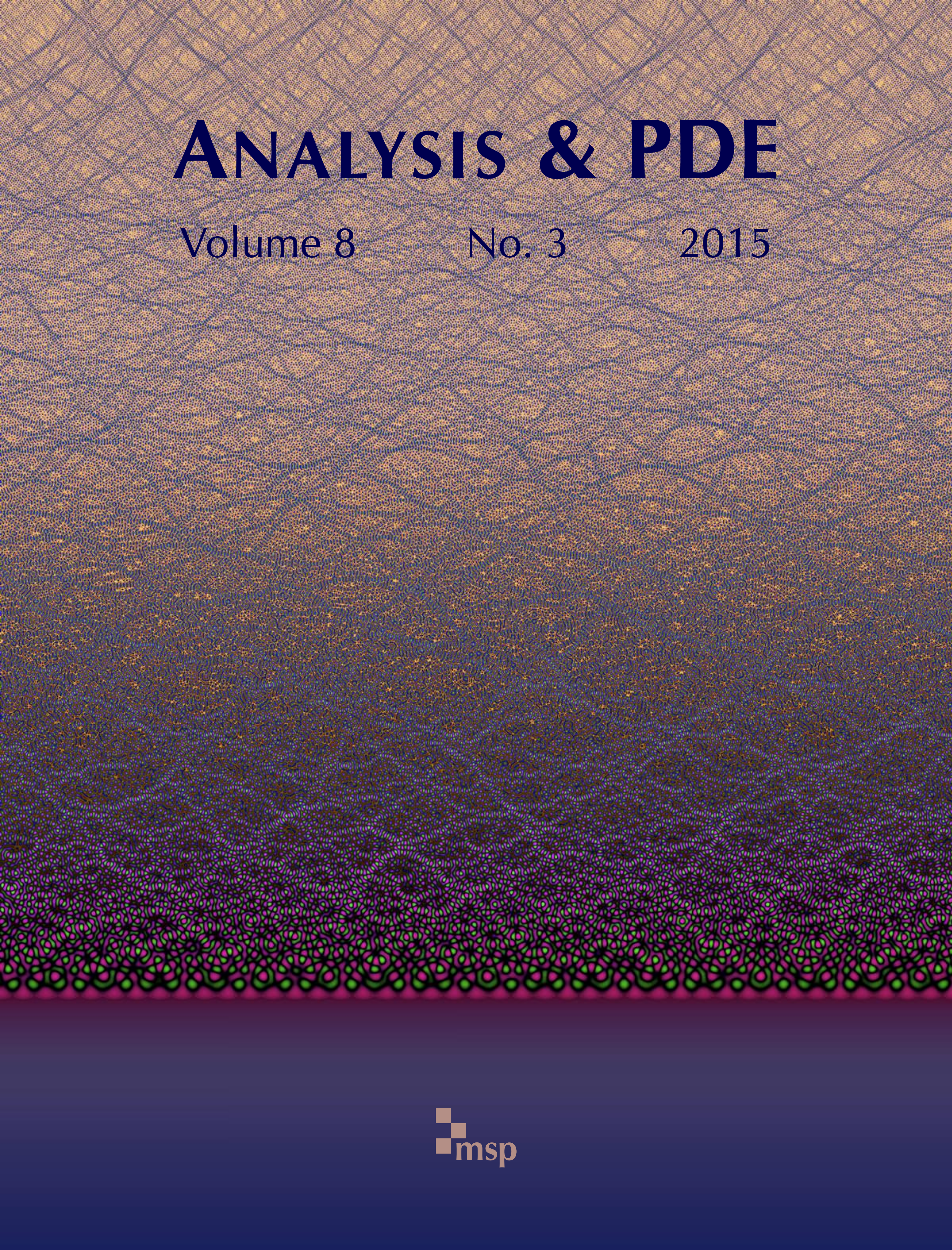

Analysis & PDE

Volume 8 No. 3 2015

msp

Analysis & PDE

msp.org/apde

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2015 is US $205/year for the electronic version, and $390/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 8, No. 3, 2015
dx.doi.org/10.2140/apde.2015.8.513

INVERSE SCATTERING WITH PARTIAL DATA ON ASYMPTOTICALLY HYPERBOLIC MANIFOLDS

RAPHAEL HORA AND ANTÔNIO SÁ BARRETO

We prove a local support theorem for the radiation fields on asymptotically hyperbolic manifolds and use it to show that the scattering operator restricted to an open subset of the boundary of the manifold determines the manifold and the metric modulo isometries that are equal to the identity on the open subset where the scattering operator is known.

1. Introduction

We recall that the ball model of the hyperbolic space $\mathbb{H}^{n+1}$ is given by

$$\mathbb{B}^{n+1} = \{z \in \mathbb{R}^{n+1} : |z| < 1\} \quad \text{equipped with the metric } g = \frac{4\,dz^2}{(1-|z|^2)^2}.$$

It is well known that $(\mathbb{B}^{n+1}, g)$ is a complete manifold with constant curvature -1. On the other hand, $(\mathbb{B}^{n+1}, (1-|z|^2)^2 g)$ is the interior of a compact Riemannian manifold with boundary. This structure can be generalized by replacing $\mathbb{B}^{n+1}$ with the interior of a C^∞ compact manifold X, with boundary ∂X, of dimension $n+1$ and replacing $1-|z|^2$ with a function $\rho \in C^\infty(X)$ which defines ∂X; that is, $\rho > 0$ in the interior of X, $\{\rho = 0\} = \partial X$, and $d\rho \neq 0$ at ∂X. Such a function ρ will be called a boundary-defining function. We will denote the interior of X by $\mathring{X}$. If g is a Riemannian metric on $\mathring{X}$ such that

$$\rho^2 g = H \tag{1-1}$$

is C^∞ and nondegenerate up to ∂X then, according to [Mazzeo and Melrose 1987], g is complete and its sectional curvatures approach $-|d\rho|_H^2$ as $\rho \downarrow 0$. In particular, when

$$|d\rho|_{H^2} = 1 \quad \text{at } \partial X, \tag{1-2}$$

the sectional curvatures converge to -1 at the boundary. A Riemannian manifold $(\mathring{X}, g)$, where X is a compact C^∞ manifold with boundary and where (1-1) and (1-2) hold, is said to be an asymptotically hyperbolic manifold (AHM). Any compact C^∞ Riemannian manifold with boundary X can be equipped with such a metric.

We will study certain properties of the asymptotic behavior of solutions to the Cauchy problem for the wave equation on $(\mathring{X}, g)$. In particular, we will study the Friedlander radiation fields on AHM, and show that the support of the radiation fields restricted to an open subset of ∂X controls the support of the initial

MSC2010: 35P25, 58J50.
Keywords: inverse scattering, asymptotically hyperbolic manifolds.

data of the Cauchy problem for the wave equation. Such theorems are usually called support theorems; see, for example, [Helgason 1999]. When $\mathring{X} = \mathbb{H}^{n+1}$, the radiation fields are given by the Lax–Phillips transform which involves the horocyclic Radon transform, and our support theorem generalizes the results of [Lax and Phillips 1982] to this setting.

We will use this result and adapt the boundary control theory of [Belishev 1987; Belishev and Kurylev 1992; Tataru 1995; 1999], and a refinement of the results of [Belishev and Kurylev 1992] due to Kurylev and Lassas [2002] and Katchalov, Kurylev and Lassas [Katchalov et al. 2001], to prove that the scattering operator restricted to a nonempty open set $\Gamma \subset \partial X$ determines (X, g) modulo isometries that are equal to the identity on Γ. There is a very large body of work on scattering and inverse scattering for Schrödinger operators, obstacle problems, etc., however much less is known about inverse scattering on manifolds. It was proved in [Sá Barreto 2005] that the scattering operator on the entire boundary of an AHM (X, g) determines the manifold and the metric modulo isometries that are the identity at ∂X. Guillarmou and Sá Barreto [2008] extended the result of [Sá Barreto 2005] to asymptotically complex hyperbolic manifolds. Isozaki, Kurylev and Lassas [Isozaki et al. 2010; 2013] studied the case of manifolds of cylindrical ends and asymptotically hyperbolic orbifolds; see also their survey paper [Isozaki et al. 2014]. One should also mention the book by Isozaki and Kurylev [2014], where they discuss spectral theory and inverse problems on AHM. If an AHM manifold is also Einstein, Guillarmou and Sá Barreto [2009] showed that the scattering matrix at one energy determines the manifold.

2. Preliminaries and statements of the results

We begin by recalling the definition of the radiation fields and the scattering operator. Let $u(t, z)$ satisfy the wave equation

$$\begin{gathered} \big(D_t^2 - \Delta_g - \tfrac{1}{4}n^2\big)u = 0 \quad \text{on } \mathbb{R}_\pm \times \mathring{X}, \\ u(0, z) = f_1, \quad D_t u(0, z) = f_2, \quad f_1,\ f_2 \in C_0^\infty(\mathring{X}). \end{gathered} \tag{2-1}$$

The spectrum of the Laplacian Δ_g, denoted by $\sigma(\Delta_g)$, was studied by [Mazzeo 1988; 1991; Mazzeo and Melrose 1987] and more recently by Bouclet [2013]. They showed that $\sigma(\Delta_g) = \sigma_{\mathrm{pp}}(\Delta_g) \cup \sigma_{\mathrm{ac}}(\Delta_g)$, where $\sigma_{\mathrm{pp}}(\Delta_g)$ is the finite point spectrum, $\sigma_{\mathrm{ac}}(\Delta_g)$ is the absolutely continuous spectrum and

$$\sigma_{\mathrm{ac}}(\Delta_g) = \big[\tfrac{1}{4}n^2, \infty\big), \quad \sigma_{\mathrm{pp}}(\Delta_g) \subset \big(0, \tfrac{1}{4}n^2\big). \tag{2-2}$$

The role of the factor $n^2/4$ in (2-1) is to shift the continuous spectrum of Δ_g to $[0, \infty)$.

Equation (2-1) has a conserved energy given by

$$\begin{aligned} E(u, \partial_t u)(t) &= \int_X \big(|du(t)|^2 - \tfrac{1}{4}n^2|u(t)|^2 + |\partial_t u(t)|^2\big)\, d\,\mathrm{vol}_g, \\ E(u, \partial_t u)(0) = E(f_1, f_2) &= \int_X \big(|df_1|^2 - \tfrac{1}{4}n^2[t]\,|f_1|^2 + |f_2|^2\big)\, d\,\mathrm{vol}_g\,. \end{aligned} \tag{2-3}$$

However, $E(f_1, f_2)$ is a nonnegative quadratic form only when projected onto $L^2_{\mathrm{ac}}(X)$. As in [Sá Barreto 2005], we define the energy space

$$H_E(X) = \{(f_1, f_2) : f_1,\ f_2 \in L^2(X),\ df_1 \in L^2(X) \text{ and } E(f_1, f_2) < \infty\}$$

and, if $\{\phi_j : 1 \leq j \leq N\}$ are the eigenfunctions of Δ_g, we define the projector

$$\mathcal{P}_{\mathrm{ac}} : L^2(X) \to L^2_{\mathrm{ac}}(X), \quad f \mapsto f - \sum_{j=1}^{N} \langle f, \phi_j \rangle \phi_j,$$

and the space $E_{\mathrm{ac}}(X) = \mathcal{P}_{\mathrm{ac}}(H_E(X))$.

The wave group induces a strongly continuous group of unitary operators,

$$U(t) : E_{\mathrm{ac}}(X) \to E_{\mathrm{ac}}(X), \quad (f_1, f_2) \mapsto (u(t), \partial_t u(t)).$$

Next we recall the definition of the forward and backward radiation fields from [Sá Barreto 2005]. We will work with a specific boundary defining function and, since our definition will depend on this choice, we will recall the construction from [Graham 2000]. Since any two defining functions of ∂X, ρ and $\tilde{\rho}$, satisfy $\rho = e^{\omega}\tilde{\rho}$ with $\omega \in C^\infty(X)$, if $H = \rho^2 g$ and $\tilde{H} = \tilde{\rho}^2 g$ then $H|_{\partial X} = e^{2\omega(0,y)}\tilde{H}|_{\partial X}$. Hence, $\rho^2 g|_{\partial X}$ determines a conformal class of metrics on ∂X. We have $H = \rho^2 g = e^{2\omega}\tilde{\rho}^2 g$, and so $H = e^{2\omega}\tilde{H}$. Since $d\rho = e^{\omega}(\tilde{\rho}\, d\omega + d\tilde{\rho})$, we have

$$|d\rho|^2_H = |d\tilde{\rho} + \tilde{\rho}\, d\omega|^2_{\tilde{H}} = |d\tilde{\rho}|^2_{\tilde{H}} + \tilde{\rho}^2 |d\omega|^2_{\tilde{H}} + 2\tilde{\rho}(\nabla_{\tilde{H}}\tilde{\rho})\omega.$$

Hence,

$$|dx|_H = 1 \qquad \text{if and only if} \qquad 2(\nabla_{\tilde{H}}\tilde{\rho})\omega + \tilde{\rho}|d\omega|^2_{\tilde{H}} = \frac{1}{\tilde{\rho}}(1 - |d\tilde{\rho}|^2_{\tilde{H}}), \quad \omega|_{\partial X} = 0.$$

Since, by assumption, $|d\tilde{\rho}|_{\tilde{H}} = 1$ at ∂X, this is a noncharacteristic ODE, and hence it has a solution in a neighborhood of ∂X. Notice that the function ρ is in principle defined only on a collar neighborhood of ∂X, but it can be extended to the whole manifold as a boundary-defining function.

The boundary-defining function ρ gives an identification between $[0, \varepsilon) \times \partial X$ and a collar neighborhood U of ∂X,

$$\Psi : [0, \varepsilon) \times \partial X \to U \subset X, \quad (x, y) \mapsto \exp(x\nabla_H \rho)(y),$$

where $\exp(x\nabla_H \rho)(y)$ just means that one follows the integral curve of $\nabla_H \rho$ starting at y for x units of time. In this case,

$$\begin{gathered} \Psi^* g = \frac{dx^2}{x^2} + \frac{h(x)}{x^2} \quad \text{on } (0, \varepsilon) \times \partial X, \quad h(0) = H|_{\partial X}, \\ \Psi = \mathrm{Id} \quad \text{on } \partial X, \end{gathered} \tag{2-4}$$

where $h(x)$ is a C^∞ family of metrics ∂X for $x \in [0, \varepsilon)$. From now on we will use this identification $U \sim [0, \varepsilon)_x \times \partial X$.

In the coordinates (2-4), for fixed $y \in \partial X$ the curve $\gamma(s) = (s, y)$ is a geodesic for the metric g, the distance between (x, y) and (x', y), $x < x'$, is $\log(x'/x)$, and if time t is the arc-length parameter then $t = \log x' - \log x$. So, to analyze global properties of $u(t, z)$ in space and time, it is convenient to work with an exponential compactification of $\mathbb{R} \ni t$, and we choose a function T such that $\{T = 0\} = \{t = 0\}$, $T = 1 - e^{-t}$ if $t > 1$, and $T = -1 + e^t$ if $t < -1$. Let $Y = [-1, 1] \times X$ be the compactified space; see Figure 1. The light cones will converge to the corners of the manifold Y and to separate them one blows up the intersection of ∂X with $\{T = -1\}$ and $\{T = 1\}$. This gives a manifold with corners $\tilde{Y}$, pictured in

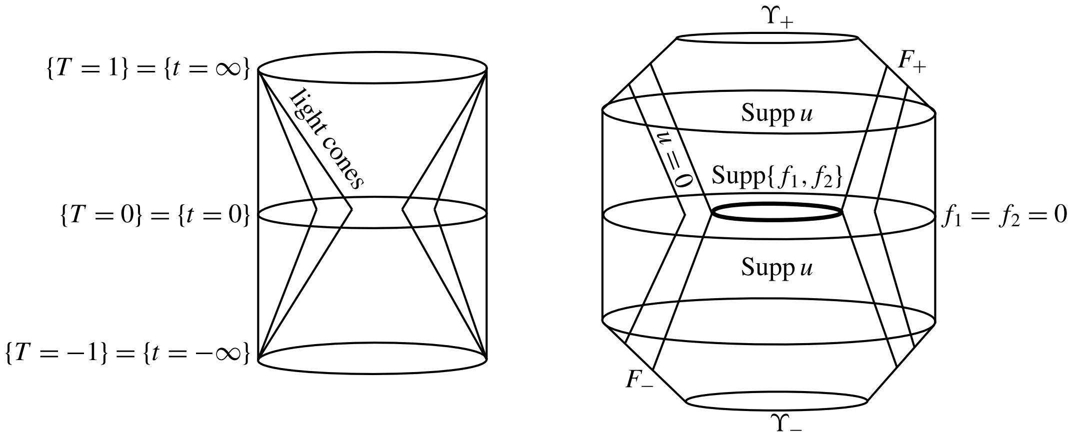

Figure 1. The manifold $Y=[-1,1]\times X$, and $\tilde{Y}$, its blow-up along $\partial X\times\{T=\pm 1\}$. All the light cones intersect at $\{x=0, T=\pm 1\}$ in Y, but in $\tilde{Y}$ they are separated at the faces F_+ and F_-.

Figure 1. In local coordinates, the blow-up is the equivalent of introducing polar coordinates $x=r\cos\theta$, $T\pm 1=r\sin\theta$.

It was proved in [Sá Barreto 2005] that, if $(f_1, f_2)\in C_0^\infty(X)$, the solution u to the wave equation (2-1) is in $C^\infty(\tilde{Y}\setminus(\overline{\Upsilon_+\cup\Upsilon_-}))$ (see Figure 1 for the definition of $\Upsilon_\pm$). The analysis of the behavior of $u(t,z)$ on the faces $\Upsilon_\pm$ give, among other things, information about the local energy decay, and will not be studied here. A similar discussion about the asymptotic solutions of the wave equation on de Sitter–Schwarzschild space, including the pictures, can be found in [Melrose et al. 2014a; 2014b]; see also [Vasy 2013].

Following Friedlander [1980; 2001], one defines the forward and backward radiation fields of u as

$$\mathcal{R}_+(f_1, f_2)=x^{-n/2}\partial_t u|_{F_+\setminus\overline{\Upsilon}_+},$$
$$\mathcal{R}_-(f_1, f_2)=x^{-n/2}\partial_t u|_{F_-\setminus\overline{\Upsilon}_-}.$$

If we use projective coordinates x and $\tau_+=x/(1-T)$, valid near $F_+\setminus\overline{\Upsilon}_+$, and $\tau_-=x/(1+T)$, valid near $F_-\setminus\overline{\Upsilon}_-$, and set $s_+=\log\tau_+$ and $s_-=-\log\tau_-$, then, for $(f_1, f_2)\in C_0^\infty(\mathring{X})\times C_0^\infty(\mathring{X})$, the solution $u(t,z)$ to (2-1), with $z=(x,y)$, satisfies

$$\begin{aligned} V_+(x,s_+,y)&=x^{-n/2}u(s_+-\log x, x, y)\in C^\infty([0,\varepsilon)_x\times\mathbb{R}_{s_+}\times\partial X)\\ V_-(x,s,y)&=x^{-n/2}u(s_-+\log x, x, y)\in C^\infty([0,\varepsilon)_x\times\mathbb{R}_{s_-}\times\partial X).\end{aligned}\tag{2-5}$$

In these coordinates, the forward and backward radiation fields can be expressed as

$$\begin{aligned} &\mathcal{R}_+ : C_0^\infty(\mathring{X})\times C_0^\infty(\mathring{X})\to C^\infty(\mathbb{R}\times\partial X), \quad \mathcal{R}_+(f_1, f_2)(s_+, y)=D_{s_+}V_+(0, s_+, y),\\ &\mathcal{R}_- : C_0^\infty(\mathring{X})\times C_0^\infty(\mathring{X})\to C^\infty(\mathbb{R}\times\partial X), \quad \mathcal{R}_-(f_1, f_2)(s_-, y)=D_{s_-}V_-(0, s_-, y).\end{aligned}\tag{2-6}$$

It was shown in [Sá Barreto 2005] that $\mathcal{R}_\pm$ extend to unitary operators

$$\mathcal{R}_\pm : E_{\mathrm{ac}}(X) \to L^2(\mathbb{R} \times \partial X), \quad (f_1, f_2) \mapsto \mathcal{R}_\pm(f_1, f_2), \tag{2-7}$$

where the measure on ∂X is the one induced by the metric h_0 defined in (2-4).

It follows from the definitions that $\mathcal{R}_\pm$ are translation representations of the wave group as in the Lax–Phillips theory [1989], i.e.,

$$\mathcal{R}_\pm(U(T)(f_1, f_2))(s, y) = \mathcal{R}_\pm(f_1, f_2)(s + T, y). \tag{2-8}$$

One can define the scattering operator

$$\mathcal{S} : L^2(\mathbb{R} \times \partial X) \to L^2(\mathbb{R} \times \partial X), \quad \mathcal{S} = \mathcal{R}_+ \circ \mathcal{R}_-^{-1}, \tag{2-9}$$

which is unitary in $L^2(\partial X \times \mathbb{R})$ and, in view of (2-8), commutes with translations in the s variable.

The scattering matrix $\mathcal{A}(\lambda)$ is defined by conjugating $\mathcal{S}$ with the Fourier transform in the s variable:

$$\mathcal{A}(\lambda) = \mathcal{F} \circ \mathcal{S} \circ \mathcal{F}^{-1}, \quad \mathcal{F} f(\lambda) = \int_{\mathbb{R}} e^{-i\lambda s} f(s)\, ds. \tag{2-10}$$

In particular, $\mathcal{S}$ determines $\mathcal{A}(\lambda)$, $\lambda \in \mathbb{R}$ and vice versa. It was proved in [Joshi and Sá Barreto 2000] that $\mathcal{A}(\lambda)$ continues meromorphically to $\mathbb{C} \setminus D$, where D is a discrete subset of $\mathbb{C}$.

As discussed above, the distance between (x, y) and (x', y), $x < x' < \varepsilon$, is $\log(x'/x)$. The finite speed of propagation for the wave equation implies that the solution $u(t, z)$ of (2-1) satisfies $u(t, z) = 0$ if $t < d_g(z, \operatorname{Supp}(f_1, f_2))$. In particular, if $f_1(x', y) = f_2(x', y) = 0$ for all $x' < \rho$, then $u(t, x) = 0$ for $x < x' < \rho$ and $t < \log(x'/x)$. This implies that $V_+(s, x, y) = x^{-n/2}\partial_t u(s - \log x, x, y) = 0$ provided $x < x' < \rho$ and $s = t + \log x < \log x' < \log \rho$. This shows that, if $f_1(x', y) = f_2(x', y) = 0$ in $x' \le \rho$, then $\mathcal{R}_+(f_1, f_2)(s, y) = 0$ for $s \le \log \rho$. The converse of this statement for initial data of the type $(0, f)$ was proved in [Sá Barreto 2005]: if $f \in L^2_{\mathrm{ac}}(X)$ and $\mathcal{R}_+(0, f)(s, y) = 0$ for $s \le \log \rho \ll 0$ and $y \in \partial X$, then $f(x, y) = 0$ in $x \le \rho$. Due to possible cancelations, one cannot expect the converse to be true for an arbitrary pair (f_1, f_2). In this paper we prove the following refinement of this result:

Theorem 2.1. *Let $\Gamma \subset \partial X$ be a nonempty open subset, let $f \in L^2_{\mathrm{ac}}(X)$ and let $s_0 \in \mathbb{R}$. Let $\varepsilon > 0$ be such that* (2-4) *holds in $(0, \varepsilon) \times \partial X$, and let $\bar\varepsilon = \min\{\varepsilon, e^{s_0}\}$. Then $\mathcal{R}_+(0, f)(s, y) = 0$ in $\{s < s_0,\ y \in \Gamma\}$ if and only if, for every $z = (x, y) \in (0, \bar\varepsilon) = U_{\bar\varepsilon}$,*

$$d_g(z, \operatorname{Supp} f) > \log \frac{e^{s_0}}{x}, \tag{2-11}$$

where d_g denotes the distance function with respect to the metric g and Supp f *denotes the support of f. Another way of stating* (2-11) *is to say that $f = 0$ on the set*

$$\mathcal{D}_{s_0}(\Gamma) = \left\{ z \in X : \exists q = (x, y) \in U_{\bar\varepsilon},\ d_g(z, q) < \log \frac{e^{s_0}}{x} \right\} = \bigcup_{(x,y) \in U_{\bar\varepsilon}} B\left((x, y), \log \frac{e^{s_0}}{x}\right), \tag{2-12}$$

where $B(p, r)$ denotes the open ball of radius r centered at p with respect to the metric g.

If $\Gamma = \partial X$ and $\bar{\varepsilon} = e^{s_0}$ then, for any $z = (\alpha, y)$ with $\alpha < e^{s_0}$, pick $q = (x, y)$ with $x < \alpha < e^{s_0}$. Then $d_g((\alpha, y), (x, y)) = \log(\alpha/x) < \log(e^{s_0}/x)$. Therefore, $\{(\alpha, y) : \alpha < e^{s_0},\ y \in \partial X\} \subset \mathcal{D}_{s_0}(\partial X)$, and hence Theorem 2.1 shows that, if $f \in L^2_{\mathrm{ac}}(X)$ and $\mathcal{R}_+(0, f)(s, y) = 0$ for $s \le s_0$ and $y \in \partial X$, then $f(x, y) = 0$ for $x < e^{s_0}$. This particular case of Theorem 2.1, when $\Gamma = \partial X$ and $\bar{\varepsilon} = e^{s_0}$ was proved in [Sá Barreto 2005].

Lax and Phillips [1982] proved Theorem 2.1 for the case when (X, g) is the hyperbolic space. In that case the radiation field is given in terms of the horocyclic Radon transform, and their result says that, if the integral of f over all horospheres tangent to points $(0, y)$ with $y \in \Gamma$ and radii less than or equal to $\frac{1}{2}e^{s_0}$ is equal to zero, then $f = 0$ in the region given by the union of these horocycles. It is useful to explain what the set $\mathcal{D}_{s_0}(\Gamma)$ is when (X, g) is the hyperbolic space, and verify that Theorem 2.1 implies the result of Lax and Phillips. It is easier to do the computations for the half-space model of hyperbolic space, which is given by

$$\mathbb{H}^{n+1} = \{(x, y) : x > 0,\ y \in \mathbb{R}^n\} \quad \text{with the metric } g = \frac{dx^2 + dy^2}{x^2}.$$

The distance function between $z = (x, y)$ and $w = (\alpha, y')$ satisfies

$$\cosh d_g(z, w) = \frac{x^2 + \alpha^2 + |y - y'|^2}{2x\alpha}.$$

Since $d_g(z, z') \le \log(e^{s_0}/\alpha)$, we obtain

$$\big(x - \tfrac{1}{2}e^{s_0}(1 + \alpha^2 e^{-2s_0})\big)^2 + |y - y'|^2 \le \tfrac{1}{4}e^{2s_0}(1 + \alpha^2 e^{-2s_0})^2 - \alpha^2 = \tfrac{1}{4}e^{2s_0}(1 - \alpha^2 e^{-2s_0})^2,$$

which corresponds to a ball $D(\alpha)$ centered at $\big(\frac{1}{2}e^{s_0}(1 + \alpha^2 e^{-2s_0}), y'\big)$ and radius $\frac{1}{2}e^{s_0}(1 - \alpha^2 e^{-2s_0})$. Since $\alpha < e^{s_0}$, we have $D(\alpha) \subset D(0)$, as shown in Figure 2. This ball is tangent to the plane $x = e^{s_0}$ at the point (e^{s_0}, y'). When $\alpha = 0$, the ball $D(0)$ has center $\big(\frac{1}{2}e^{s_0}, y'\big)$ and radius $\frac{1}{2}e^{s_0}$ and is also tangent to the plane $\{x = 0\}$. The boundary of $D(0)$ is called a horosphere since it is orthogonal to the geodesics emanating from the point $(0, y')$. When $\alpha = e^{s_0}$, $D(e^{s_0}) = (e^{s_0}, y')$. The set $\mathcal{D}_{s_0}(\Gamma)$ consists of the union

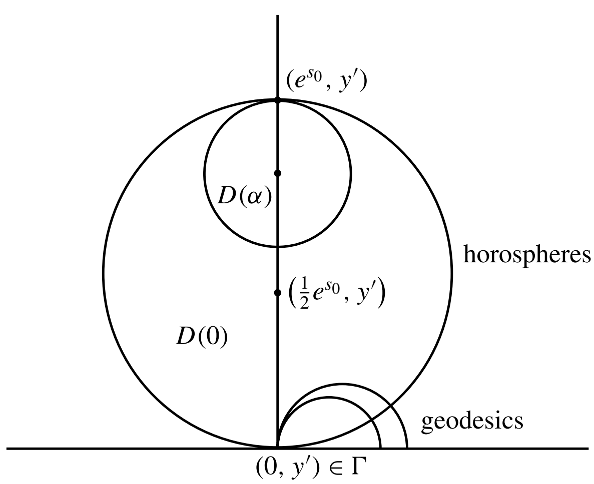

Figure 2. The horospheres tangent at $(0, y')$ and the balls $D(\alpha)$.

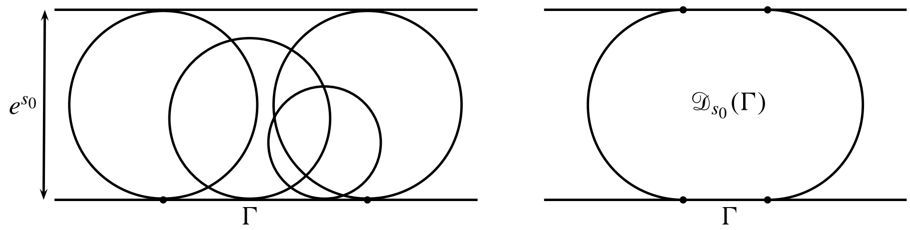

Figure 3. The set $\mathcal{D}_{s_0}(\Gamma)$ when (X, g) is the hyperbolic space is given by the union of horospheres tangent to points on Γ and radii less than or equal to $\frac{1}{2}e^{s_0}$.

of horospheres with radii less than or equal to $\frac{1}{2}e^{s_0}$ tangent to points $(0, y')$ with $y' \in \Gamma$; see Figure 3.

Theorem 2.1 can be explained in terms of the sojourn time along a geodesic. In this setting, the sojourn time plays the role of the distance function to the boundary of X and is closely related to the Busemann function used in differential geometry. Let $\gamma(t)$ be a geodesic, parametrized by the arc length, passing through $z = \gamma(0)$ and such that $\gamma(t) \to y \in \partial X$ as $t \to \infty$. We define

$$s(z, \gamma) = \lim_{t\to\infty} \big(t + \log x(\gamma(t))\big).$$

The relationship between the sojourn times and the radiation fields for nontrapping asymptotically hyperbolic manifolds was studied in [Sá Barreto and Wunsch 2005]. We have the following consequence of Theorem 2.1:

Corollary 2.2. *Let f and $\Gamma \subset \partial X$ satisfy the hypotheses of Theorem 2.1; then $f = 0$ on the set of points $z \in \mathring{X}$ such that there exists a geodesic $\gamma(t)$, parametrized by the arc length, with $\gamma(0) = z$, $\gamma(t) \to y \in \Gamma$ as $t \to \infty$, and $s(z, \gamma) < s_0$.*

Proof. Suppose there exists a geodesic $\gamma(t)$, parametrized by the arc length t, such that $\gamma(0) = z$, $\lim_{t\to\infty} \gamma(t) = y$ and

$$\lim_{t\to\infty} \big(t + \log x(\gamma(t))\big) = s < s_0.$$

Since t is the arc-length parameter, $d\big(z, \big(x(\gamma(t)), y\big)\big) \le t$ and $s < s_0$, there exists $T > 0$ such that, for $t > T$, $\gamma(t) \in U \sim [0, \varepsilon) \times \partial X$ where the coordinates (2-4) are valid and $t + \log x(\gamma(t)) < s_0$. Therefore, if $t > T$,

$$d\big(z, (x(t), y)\big) \le t < s_0 - \log x(\gamma(t)) = \log \frac{e^{s_0}}{x(\gamma(t))}.$$

Hence $z \in \mathcal{D}_{s_0}(\Gamma)$. □

Theorem 2.1 says that the support of the radiation field $\mathcal{R}_+(0, f)$ controls the support of the initial data $(0, f)$. We will use this result to adapt the boundary control method of [Belishev 1987; Belishev and Kurylev 1992; Kurylev and Lassas 2002; Katchalov et al. 2001] to study the inverse scattering problem with partial data.

Let $\Gamma \subset \partial X$ be an open subset and let $\mathcal{S}$ denote the scattering operator as in (2-9). We define the restriction of $\mathcal{S}$ to $\mathbb{R} \times \Gamma$ as

$$\mathcal{S}_\Gamma : L^2(\mathbb{R} \times \Gamma) \to L^2(\mathbb{R} \times \Gamma), \quad F \mapsto (\mathcal{S}F)|_\Gamma. \tag{2-13}$$

In other words, one starts with an $F \in L^2(\mathbb{R} \times \Gamma)$, finds the solution of the wave equation that has backward radiation field equal to F, then finds the corresponding forward radiation field, and restricts it to the subset $\mathbb{R} \times \Gamma$. We study the problem of determining (X, g) from $\mathcal{S}_\Gamma$. Recall that our definition of $\mathcal{S}$ depends on the choice of the product structure (2-4). In fact, the method used in [Graham 2000] and discussed above to construct the diffeomorphism (2-4) can also be used to show that, given two AHM (X_j, g_j), $j = 1, 2$, there exists $\varepsilon > 0$ such that (2-14) holds for both metrics. Recall that x is just the time through which one flows along the integral curves of $\nabla_H \rho$. One can take ε to the smallest one that works for both metrics, and one finds that there exist collar neighborhoods $U_j \subset X_j$ of ∂X_j and C^∞ diffeomorphisms

$$\Psi_j : (0, \varepsilon) \times \partial X_j \to U_j$$

such that

$$\Psi_j^* g_j = \frac{dx^2}{x^2} + \frac{h_j(x)}{x^2} \quad \text{in } (0, \varepsilon) \times \partial X_j, \quad h_j(0) = h_{j0}, \quad j = 1, 2, \tag{2-14}$$

where $h_j(x)$ is a C^∞ family of metrics on ∂X_j for $x \in [0, \varepsilon)$, and $\Psi_j = \mathrm{Id}$ on ∂X_j. In particular, if there exists an open set $\Gamma \subset \partial X_1 \cap \partial X_2$, as manifolds, then (2-14) holds on $(0, \varepsilon) \times \Gamma$, and $h_j(x)$ are C^∞ families of metrics on Γ. We prove the following:

Theorem 2.3. *Let (X_1, g_1) and (X_2, g_2) be connected, asymptotically hyperbolic manifolds and suppose there exists an nonempty open set $\Gamma \subset \partial X_1 \cap \partial X_2$ (as manifolds). Let x be such that* (2-14) *holds on a collar neighborhood of ∂X_j for $j = 1, 2$. Suppose that $h_1(0) = h_2(0)$ on Γ. Let $\mathcal{S}_{j,\Gamma}$, $j = 1, 2$, be the corresponding scattering operators restricted to Γ, and suppose that $\mathcal{S}_{1,\Gamma} = \mathcal{S}_{2,\Gamma}$. Then there exists a C^∞ diffeomorphism*

$$\Psi : X_1 \to X_2 \quad \textit{such that} \quad \Psi = \mathrm{Id} \ \textit{on } \Gamma \quad \textit{and} \quad \Psi^* g_2 = g_1. \tag{2-15}$$

Since we only know $\mathcal{S}$ on part of the boundary, we can only expect to recover information on the connected components of (X, g) that contain Γ, so we assume that X is connected. This result guarantees that the scattering operator restricted to Γ determines (X, g), including its topology and C^∞ structure, modulo isometries that are equal to the identity on Γ.

Theorem 2.3, and the method we use to prove it, are related to the question of reconstructing a compact Riemannian manifold with boundary from the Dirichlet-to-Neumann map (DTNM) for the wave equation. One may think of the scattering operator as the DTNM on the boundary at infinity. Belishev and Kurylev [1992] showed that the DTNM for the wave equation determines a compact manifold and its Riemannian metric using the boundary control method and a unique continuation result later proved by Tataru [1995; 1999]. Different proofs, which also rely on the result of Tataru, were given in [Katchalov et al. 2001]. This result of Tataru will be important in the proof of Theorem 2.1. The reconstruction of a compact manifold in the case where the Dirichlet-to-Neumann map is only known on part of the boundary was carried out by Kurylev and Lassas [2000] using a modification of the boundary control method; see also Section 4.4 of [Katchalov et al. 2001]. We will adapt the boundary control methods to this setting by using the radiation fields.

3. The proof of Theorem 2.1

The sufficiency of condition (2-11) in Theorem 2.1 is just a consequence of the finite speed of propagation for the wave equation.

Lemma 3.1. *Let* $f \in L^2_{\text{ac}}(X)$ *be such that* $d_g(z, \operatorname{Supp} f) > \log(e^{s_0}/x)$ *for all* $z = (x, y) \in (0, \bar{\varepsilon}) \times \Gamma$. *Then* $\mathcal{R}_+(0, f)(s, y) = 0$ *if* $s \leq s_0$ *and* $y \in \Gamma$.

Proof. Let $u(t, z)$ satisfy the wave equation (2-1) with initial data $(0, f)$. The finite speed of propagation for solutions of the wave equation guarantees that $u(t, z) = 0$ if $0 \leq t < d_g(z, \operatorname{Supp} f)$. In particular, if $z = (x, y)$ with $x < \bar{\varepsilon}$, $y \in \Gamma$, then $u(t, x, y) = 0$ if $0 \leq t \leq s_0 - \log x < d_g(z, \operatorname{Supp} f)$. Since $s = t + \log x$, we have that $V_+(x, s, y) = x^{-n/2} u(s - \log x, x, y) = 0$ provided $\log x \leq s \leq s_0$, $x < \bar{\varepsilon}$, $y \in \Gamma$. This implies that $\mathcal{R}_+(0, f)(s, y) = 0$ if $s \leq s_0$ and $y \in \Gamma$. □

We will first outline the proof of the converse, which is based on unique continuation arguments. We state three propositions, and indicate how to use them to prove the converse of Theorem 2.1. We will finish the proof of Theorem 2.1 at the end of the section, after we have proved the three propositions.

In the region where (2-4) holds, the Cauchy problem (2-1), with initial data $(0, f)$ translates into the following initial value problem for $V_+(x, s, y) = x^{-n/2} u(s + \log x, x, y)$:

$$\begin{gathered} P V_+(x, s, y) = 0 \quad \text{in } \log x < s,\ x < \varepsilon,\ y \in \partial X, \\ V_+(x, \log x, y) = 0, \quad D_s V_+(x, \log x, y) = x^{-n/2} f(x, y), \quad x < \varepsilon,\ y \in \partial X, \end{gathered} \tag{3-1}$$

where

$$P = -x^{-n/2-1}\big(D_t^2 - \Delta - \tfrac{1}{4}n^2\big)x^{n/2} = \partial_x(2\partial_s + x\partial_x) - x\Delta_h + A\partial_s + Ax\partial_x + \tfrac{1}{2}nA. \tag{3-2}$$

Here, Δ_h is the (positive) Laplace operator on ∂X corresponding to the metric $h(x)$, in local y coordinates,

$$\begin{gathered} \Delta_h = -\frac{1}{\sqrt{\theta}} \partial_{y_i}(\sqrt{\theta}\, h^{ij} \partial_{y_j}), \\ \text{where} \quad h = (h_{ij}(x, y)), \quad h^{-1} = (h^{ij}(x, y)), \quad \theta = \det(h_{ij}) \quad \text{and} \quad A = \frac{1}{\sqrt{\theta}} \partial_x \sqrt{\theta}. \end{gathered} \tag{3-3}$$

In the first proposition, we are interested in the behavior of $V_+(x, s, y)$ for x near $\{x = 0\}$ and $\{s = -\infty\}$. As in [Sá Barreto 2005], we work in the compactified space $\tilde{Y}$ — see Figure 1 — and set

$$\mu = e^{-s_-/2} \quad \text{and} \quad \nu = e^{s_+/2}. \tag{3-4}$$

This implies that $s = 2\log \nu$ and $x = \mu\nu$. Notice that $\mu = \sqrt{\tau_+}$ and $\nu = \sqrt{\tau_-}$ and that, in these coordinates, the lateral face Σ of $\tilde{Y}$ is given by $\Sigma = \{\tau_+ = \tau_- = 0\} = \{\mu = \nu = 0\}$, and one may think of this as collapsing the lateral face Σ, as shown in Figure 4.

In coordinates (μ, ν, y), the operator P defined in (3-2) has the form

$$\tilde{P} = \partial_\mu \partial_\nu - \mu\nu\Delta_h + \tfrac{1}{2}A(\mu\partial_\mu + \nu\partial_\nu) + \tfrac{1}{2}nA, \tag{3-5}$$

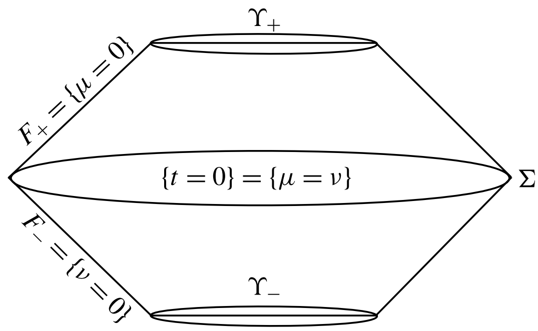

Figure 4. A compactification of $\mathbb{R}_t \times X$ with the face Σ collapsed.

where $h = h(\mu\nu)$, $A = A(\mu\nu, y)$. If

$$W(\mu, \nu, y) = V_+(\mu\nu, 2\log\nu, y) = (\mu\nu)^{-n/2} u\Big(\log\frac{\nu}{\mu}, \mu\nu, y\Big), \tag{3-6}$$

the Cauchy problem (3-1) becomes

$$\begin{gathered} \tilde{P}W = 0, \quad \mu,\ \nu \in (0, \varepsilon),\ y \in \partial X, \\ W(\mu, \mu, y) = 0, \quad \partial_\mu W(\mu, \mu, y) = -\mu^{-1-n} f(\mu^2, y). \end{gathered} \tag{3-7}$$

The fact that the initial data is of the form $(0, f)$ implies that the solution $u(t, z)$ to (2-1) satisfies $u(t, z) = -u(-t, z)$, and this implies that $W(\mu, \nu, y) = -W(\nu, \mu, y)$.

Proposition 3.2. *Let $f \in L^2_{\mathrm{ac}}(X)$ be such that $\mathcal{R}_+(0, f)(s, y) = 0$ in $\{s < s_0\} \times \Gamma$. Let u satisfy the initial value problem for the wave equation* (2-1) *with initial data* $(0, f)$, *and let $W(\mu, \nu, y)$ be defined as in* (3-6). *Then, in the sense of distributions $\partial_\mu^k W(\mu, \nu, y)|_{\{\mu=0\}} = 0$ in $[0, e^{s_0/2}) \times \Gamma$ and $\partial_\nu^k W(\mu, \nu, y)|_{\{\nu=0\}} = 0$ in $[0, e^{s_0/2}) \times \Gamma$ for $k = 0, 1, \ldots$. Moreover, for every $p \in \Gamma$ there exists $\delta > 0$ such that $W(\mu, \nu, y) = 0$ if $0 < \mu < \delta$, $0 < \nu < \delta$ and $|y - p| < \delta$.* (See Figure 5.)

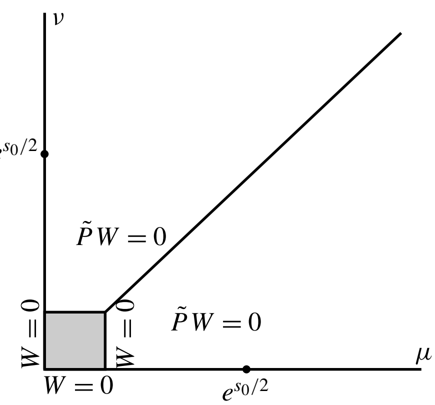

Figure 5. Unique continuation from infinity: if $\mathcal{R}_+(0, f)(s, y) = 0$ for $s \le s_0$ and a.e. $y \in \Gamma$ then, for every $p \in \Gamma$, there exists $\delta > 0$ such that $W(\mu, \nu, y) = 0$ in the region shown provided that $|y - p| < \delta$.

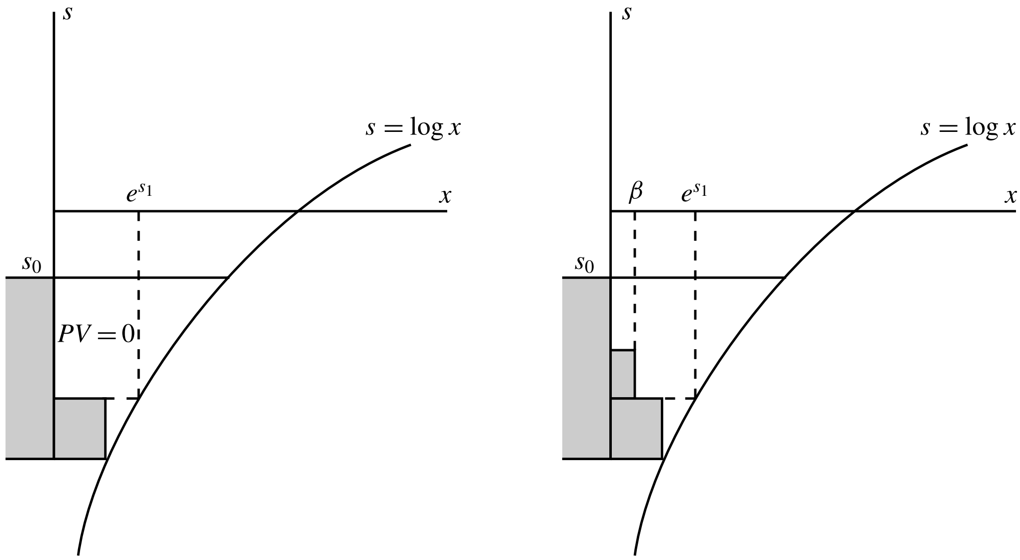

Figure 6. If $PV = 0$, and $V = 0$ in the dark region on the left, then $V = 0$ in the dark region on the right. This establishes unique continuation across the wedge $\{\log x < s < s_1,\ x < \delta,\ |y - p| < \delta\} \cup \{x \le 0,\ s < s_0,\ |y - p| < \delta\}$.

Next we need to show that we can increase the size of the neighborhood where $V_+ = 0$, and to do this we will use an iteration scheme involving the next two propositions. We will again use variables (x, s, y), and this time we will apply Hörmander's unique continuation theorem [1994b, Theorem 28.2.3], to prove:

Proposition 3.3. *Let $V(x, s, y) \in H^1_{\mathrm{loc}}$ in the region $|x| < \varepsilon$, $y \in \Gamma$ and $s \in \mathbb{R}$, satisfy $PV = 0$, where P is given by* (3-2)*. Let $s_1 < s_0$, $\delta > 0$ and $p \in \Gamma$ and suppose that*

$$V(x, s, y) = 0 \quad \text{on } \{x \in (-\varepsilon, 0],\ s < s_0,\ y \in \Gamma\} \cup \{\log x < s < s_1,\ x < \delta,\ |y - p| < \delta\}.$$

Then there exists $\beta \in (0, \delta)$ such that $V(x, s, y) = 0$ if $x < \beta$, $|y - p| < \beta$ and $\log x < s < s_1 + \frac{1}{4}(s_0 - s_1)$. (Figure 6 illustrates the result.)

We know from Proposition 3.2 that $V_+(x, s, y) = 0$ for $x < \delta$, $|y - y_0| < \delta$ and $\log x \le s \le \log \delta$. We set $s_1 = \log \delta$. Proposition 3.3 shows that $V_+(s, x, y) = 0$ in $x < \beta < \delta$, $|y - y_0| < \beta < \delta$ and $s < s_1 + \frac{1}{4}(s_0 - s_1)$. In other words, $V_+(x, s, y) = 0$ in a larger interval in the s variable at the expense of shrinking the neighborhood of $\{x = 0,\ y = p\}$.

The second piece of the scheme is a consequence of a result of Tataru [1995; 1999], and it shows that, while the neighborhood of p might shrink, the neighborhood of $x = 0$ in fact does not. Figure 7 illustrates the result.

Proposition 3.4. *Let $u(t, z)$ satisfy* (2-1) *with initial data $f_1 = 0$, $f_2 = f \in L^2(X)$. Let $V_+(x, s, y) = x^{-n/2} u(s - \log x, x, y)$. Let $p \in \Gamma$, and suppose that there exist $s_2 \in \mathbb{R}$, $\gamma > 0$ and $\delta > 0$ such that $V_+(x, s, y) = 0$ if $0 < x < \gamma$, $\log x < s < s_2$ and $|y - p| < \delta$. Then $u(t, z) = 0$ if there is (x, y) with $x < \gamma$ and $|y - p| < \delta$ such that $|t| + d_g(z, (x, y)) < \log(e^{s_2}/x)$, where d_g is the distance with respect to the*

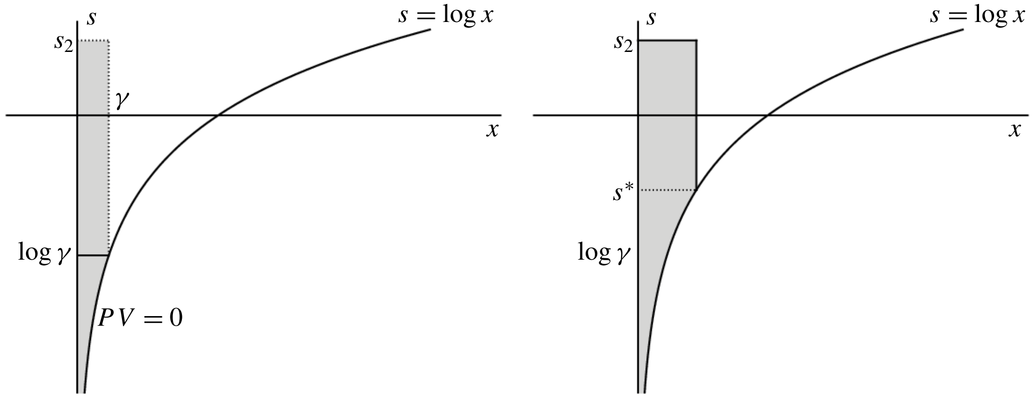

Figure 7. If $PV = 0$ and $V = 0$ in the dark region on the left, then $V = 0$ in the dark region on the right.

metric g. In particular, if $s^* < s_2$ *is such that coordinates* (2-4) *holds for* $x < e^{s^*}$, *then*

$$V_+(x, s, y) = 0 \quad \textit{if } |y - p| < \delta,\ 0 < x < e^{s^*} \textit{ and } \log x < s < s_2. \tag{3-8}$$

The idea is to iterate Propositions 3.3 and 3.4 to prove Theorem 2.1. We know from Proposition 3.2 that for any $p \in \Gamma$ there exists $\delta > 0$ such that

$$V_+(x, s, y) = 0 \quad \text{if } x < \delta,\ \log x < s < \log \delta,\ |y - p| < \delta.$$

Moreover, $V_+(x, s, y) = 0$ if $x < 0$, $s < s_0$ and $y \in \Gamma$. Applying Proposition 3.3 with $s_1 = \log \delta$, we find that there exists $\beta_1 < \delta$ such that

$$V_+(x, s, y) = 0 \quad \text{provided } x < \beta_1,\ |y - p| < \beta_1 \text{ and } \log x < s < \log \delta + \tfrac{1}{4}(s_0 - \log \delta).$$

Then Proposition 3.4 guarantees that there exists $s^* \ll 0$ independent of p such that

$$V_+(x, s, y) = 0 \quad \text{if } x < e^{s^*},\ |y - p| < \beta_1,\ s < s_2 = \log \delta + \tfrac{1}{4}(s_0 - \log \delta).$$

The main point is that, while the neighborhood of p shrinks from one step to the next, the neighborhood of $x = 0$ stays the same. Since $p \in \Gamma$ is arbitrary, it follows that in fact

$$V_+(x, s, y) = 0 \quad \text{if } x < e^{s^*},\ y \in \Gamma,\ s < s_2 = \log \delta + \tfrac{1}{4}(s_0 - \log \delta). \tag{3-9}$$

After using this argument n times, we find that

$$V_+(x, s, y) = 0 \quad \text{if } x < e^{s^*},\ y \in \Gamma,\ s < s_n = s_{n-1} + \tfrac{1}{4}(s_0 - s_{n-1}).$$

The sequence $\{s_n = s_{n-1} + \frac{1}{4}(s_0 - s_{n-1})\}$ is monotone and bounded by s_0. So it has a limit which is obviously equal to s_0. This implies that

$$V_+(x, s, y) = 0 \quad \text{if } x < e^{s^*},\ y \in \Gamma,\ s < s_0. \tag{3-10}$$

This does not quite yet prove Theorem 2.1, and the proof will be completed after the proof of Proposition 3.4. Now we will prove the three propositions above.

Proof of Proposition 3.2. First we claim that, without loss of generality, we may assume that $f \in L^2_{\mathrm{ac}}(X) \cap C^\infty(\mathring{X})$. To do this we need to characterize the range $\mathcal{R}_+(0, f)$, $f \in L^2_{\mathrm{ac}}(X)$. Notice that the solution $u(t, z)$ of (2-1) with data $(0, f)$ satisfies $u(-t, z) = -u(t, z)$, and hence $V_+(s, x, y) = x^{-n/2} u(s - \log x, x, y)$ and $V_-(s, x, y) = x^{-n/2} u(s + \log x, x, y)$ satisfy

$$V_+(x, -s, y) = x^{-n/2} u(-s - \log x, x, y) = -V_-(x, s, y). \tag{3-11}$$

In particular, we have

$$\mathcal{R}_+(0, f)(-s, y) = -(\partial_s V_+)(0, -s, y) = \partial_s V_-(0, s, y) = \mathcal{R}_-(0, f)(s, y).$$

Similarly,

$$\mathcal{R}_+(h, 0)(-s, y) = -\mathcal{R}_-(h, 0)(s, y).$$

So, if $F = \mathcal{R}_+(h, f)$ satisfies $F^*(s, y) = F(-s, y)$, then

$$F^*(s, y) = -\mathcal{R}_-(h, 0)(s, y) + \mathcal{R}_-(0, f)(s, y).$$

We apply $\mathcal{S} = \mathcal{R}_+ \mathcal{R}_-^{-1}$ to this identity and obtain

$$\mathcal{S} F^* = -\mathcal{R}_+(h, 0) + \mathcal{R}_+(0, f),$$

and we conclude that

$$\begin{aligned} \tfrac{1}{2}(\mathcal{S} F^* + F) &= \mathcal{R}_+(0, f), \\ \tfrac{1}{2}(\mathcal{S} F^* - F) &= \mathcal{R}_+(h, 0). \end{aligned} \tag{3-12}$$

Hence, $\mathcal{S} F^* = F^*$ if and only if $\mathcal{R}_+(h, 0) = 0$, and thus $h = 0$. Similarly, $\mathcal{S} F^* = -F$ if and only if $\mathcal{R}_+(0, f) = 0$ and hence $f = 0$. Therefore, we conclude that

$$\begin{aligned} \{F \in L^2(\mathbb{R} \times \partial X) : \mathcal{S} F^* = F\} &= \{\mathcal{R}_+(0, f) : f \in L^2_{\mathrm{ac}}(X)\}, \\ \{F \in L^2(\mathbb{R} \times \partial X) : \mathcal{S} F^* = -F\} &= \{\mathcal{R}_+(h, 0) : (h, 0) \in E_{\mathrm{ac}}(X)\}. \end{aligned} \tag{3-13}$$

The same argument applied to the backward radiation field shows that

$$\begin{aligned} \{F \in L^2(\mathbb{R} \times \partial X) : F^* = \mathcal{S} F\} &= \{\mathcal{R}_-(0, f) : f \in L^2_{\mathrm{ac}}(X)\}, \\ \{F \in L^2(\mathbb{R} \times \partial X) : F^* = -\mathcal{S} F\} &= \{\mathcal{R}_-(h, 0) : (h, 0) \in E_{\mathrm{ac}}(X)\}. \end{aligned} \tag{3-14}$$

Since $\mathcal{R}_+(0, f)(s, y) = 0$ in $\{s < s_0\} \times \Gamma$, we may take the convolution of $\mathcal{R}_+(0, f)$ with $\psi_\delta(s) \in C_0^\infty(\mathbb{R})$ even and supported in $(-\delta, \delta)$, with $\int \psi_\delta(s)\, ds = 1$. If $F(s, y) = \mathcal{R}_+(0, f)(s, y)$ and $F(s, y) = 0$ for $s \le s_0$, and

$$H_\delta(s, y) = \psi_\delta * F(s, y) = \int_{\mathbb{R}} \psi_\delta(s - s') F(s', y)\, ds',$$

then $H_\delta(s, y) = 0$ if $s \le s_0 - \delta$ and, since ψ_δ is even,

$$H_\delta^*(s, y) = H_\delta(-s, y) = \int_{\mathbb{R}} \psi_\delta(-s-s')F(s', y)\, ds' = \int_{\mathbb{R}} \psi_\delta(s+s')F(s', y)\, ds' \\ = \int_{\mathbb{R}} \psi_\delta(s-s')F(-s', y)\, ds' = \psi_\delta * F^*.$$

But the scattering operator commutes with translations in s, and hence it commutes with convolutions in the variable s. Therefore, in view of (3-13),

$$\mathcal{S} H_\delta^* = \psi_\delta * \mathcal{S} F^* = \psi_\delta * F = H_\delta.$$

We then use (3-13) to show that there exists $f_\delta \in L^2_{\mathrm{ac}}(X)$ such that $H_\delta = \mathcal{R}_+(0, f_\delta)$. Since $\mathcal{R}_+$ is unitary, $\|F - H_\delta\|_{L^2(\mathbb{R}\times\partial X)} = \|f - f_\delta\|_{L^2(X)}$, and hence $\|f_\delta - f\|_{L^2(X)} \to 0$ as $\delta \to 0$. Moreover, since $\partial_s^2 \mathcal{R}_+(0, f) = \mathcal{R}_+(0, (\Delta - n^2/4)f)$, it follows that, for every $k \ge 0$,

$$\partial_s^{2k} H_\delta(s, y) = \mathcal{R}_+\big(0, \big(\Delta - \tfrac{1}{4}n^2\big)^k f_\delta\big) \in L^2(\mathbb{R}\times\partial X),$$

and thus $(\Delta - n^2/4)^k f_\delta \in L^2(X)$ for all $k \ge 0$, using that $\mathcal{R}_+$ is unitary. Therefore, by elliptic regularity, $f_\delta \in C^\infty(\mathring{X})$. If one proves Theorem 2.1 for $f \in C^\infty(\mathring{X}) \cap L^2_{\mathrm{ac}}(X)$, then we conclude that $f_\delta(z) = 0$ for $z \in \mathcal{D}_{s_0-\delta}(\Gamma)$. But, since $f_\delta \to f$ as $\delta \to 0$, it follows that $f(z) = 0$ in $\mathcal{D}_{s_0}(\Gamma)$.

Next we will show that, if $\mathcal{R}(0, f)(s, y) = 0$ in $\{s < s_0\} \times \Gamma$, then in the sense of distributions W vanishes to infinite order at $\{\mu = 0,\ \nu < e^{s_0/2}\} \times \Gamma \cup \{\nu = 0,\ \mu < e^{s_0/2}\} \times \Gamma$. Recall that we are assuming that $f \in C^\infty(\mathring{X})$, so the solution W to (3-7) is C^∞ in the region $\{\mu > 0,\ \nu > 0\}$. The issue here is the behavior of W at $\{\mu = 0\} \cup \{\nu = 0\}$.

Notice that, if $F(\mu, y) = \mu^{-1-n} f(\mu^2, y)$, then

$$\int_0^\varepsilon \int_{\partial X} \mu |F(\mu, y)|^2 \theta^{\frac12}(\mu^2, y)\, dy\, d\mu = \frac12 \int_0^{\varepsilon^2} \int_{\partial X} |f(x, y)|^2 x^{-n-1} \theta^{\frac12}(x, y)\, dy\, dx \le \tfrac12 \|f\|^2_{L^2(X)}. \tag{3-15}$$

We know from Theorem 2.1 of [Sá Barreto 2005] that, if $f \in C_0^\infty(\mathring{X}) \cap L^2_{\mathrm{ac}}(X)$, then W has a C^∞ extension up to $\{\mu = 0\} \cup \{\nu = 0\}$ and, since $\partial_s = \frac12(\nu\partial_\nu - \mu\partial_\mu)$, then, provided $f \in C_0^\infty(\mathring{X}) \cap L^2_{\mathrm{ac}}(X)$,

$$\mathcal{R}_+(0, f)(2\log\nu, y) = \tfrac12[(\nu\partial_\nu - \mu\partial_\mu)W(\mu, \nu, y)]\big|_{\mu=0} = \tfrac12 \nu\partial_\nu W(0, \nu, y), \tag{3-16}$$

and we want to show that this restriction makes sense for $f \in L^2_{\mathrm{ac}}(X)$. We will work in the region $\{\nu \ge \mu\}$, but since the solution to (3-7) is odd under the change $(\mu, \nu) \mapsto (\nu, \mu)$, the same holds for the backward radiation field in the region $\{\nu \le \mu\}$.

Again, we assume that $f \in C_0^\infty(\mathring{X}) \cap L^2_{\mathrm{ac}}(X)$, and W satisfies (3-7). If one multiplies the equation $\tilde{P}W = 0$ by $\nu\partial_\nu W - \mu\partial_\mu W$, one obtains the identity

$$\frac{1}{2\sqrt{h(\mu\nu, y)}} \partial_\mu\big[\big(\nu|\partial_\nu W|^2 + \mu^2\nu|d_h(\mu\nu)W|^2\big)\sqrt{h}\big] - \frac{1}{2\sqrt{h(\mu\nu, y)}} \partial_\nu\big[\big(\mu|\partial_\mu W|^2 + \nu^2\mu|d_h(\mu\nu)W|^2\big)\sqrt{h}\big] \\ + \mu\nu\delta_{h(\mu\nu)}((\nu\partial_\nu W - \mu\partial_\mu W)d_{h(\mu\nu)}V) + Q(W, \mu\partial_\mu W, \nu\partial_\nu W, \mu\nu\partial_{y_j} W) = 0,$$

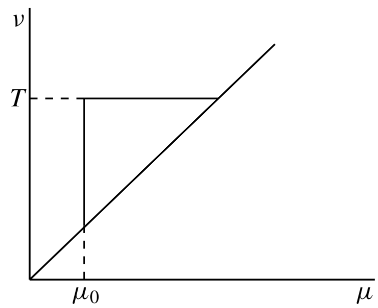

Figure 8. The region of integration in (3-17).

where $\delta_{h(\mu\nu)}$ is the divergence operator on the section ∂X dual to $d_{h(\mu\nu)}$ with respect to the metric $h(\mu\nu)$, and Q is a quadratic form. One then integrates this identity in the region $\Omega_{\mu_0,T} \times \partial X$, where $\Omega_{\mu_0,T} = \{\mu_0 \le \mu \le \nu,\ \mu \le \nu \le T\}$ is pictured in Figure 8, uses the divergence theorem and then the analogue of Gronwall's inequality, to arrive at the following inequality: for $0 \le \mu_0 \le T$, $T \in (0, e^{s_0/2})$, with T small enough that coordinates (2-4) hold for $x = \mu\nu$, there exists $C > 0$ which does not depend on f or W such that

$$\int_{\mu_0}^{T}\int_{\partial X}\big[(|W|^2+\mu|\partial_\mu W|^2+\mu\nu^2|d_{h(\mu\nu)}W|^2)\sqrt{\theta}(\mu\nu)\big]\big|_{\nu=T}\,dy\,d\mu \\ +\int_{\mu_0}^{T}\int_{\partial X}\big[(|W|^2+\nu|\partial_\nu W|^2+\mu^2\nu|d_{h(\mu\nu)}W|^2)\,\sqrt{\theta}(\mu\nu)\big]\big|_{\mu=\mu_0}\,dy\,d\nu \le C\|f\|^2_{L^2(X)}. \quad (3\text{-}17)$$

We refer the reader to the proof of Lemma 4.1 of [Sá Barreto 2005] for the details. In fact, this follows from equations (4.11), (4.14) and (4.15) of [Sá Barreto 2005], and (3-15) above.

We let

$$I(W,\mu_0,T) = \int_{\mu_0}^{T}\int_{\partial X}\big[(|W|^2+\nu|\partial_\nu W|^2+\mu^2\nu|d_{h(\mu\nu)}W|^2)\,\sqrt{\theta}(\mu\nu)\big]\big|_{\mu=\mu_0}\,dy\,d\nu.$$

If $f \in L^2_{\mathrm{ac}}(X)$ and if we take a sequence $f_j \in C_0^\infty(\mathring{X}) \cap L^2_{\mathrm{ac}}(X)$ with $\|f - f_j\|_{L^2(X)} \to 0$, (3-17) shows that, for fixed $\mu_0 \in [0, T]$,

$$I(W_j - W_k, \mu_0, T) \le C\|f_j - f_k\|^2_{L^2(X)},$$

and in particular, if $\mu_0 \in [0, T]$ and W is a solution of (3-7) with $f \in L^2_{\mathrm{ac}}(X)$, then, for $\mu_0 \in [0, T]$, the integral

$$\int_{\mu_0}^{T}\int_{\partial X}\nu|\partial_\nu W(\mu_0,\nu,y)|^2\sqrt{\theta}(\mu_0\nu,y)\,d\nu\,dy \le C\|f\|^2_{L^2(X)} \quad (3\text{-}18)$$

is well defined uniformly up to $\mu_0 = 0$. Since the radiation field is unitary, then in the sense of (3-18) the restriction $\nu\partial_\nu W(\mu_0,\nu,y)|_{\{\mu_0=0\}}$ is well defined, and hence (3-16) holds for $f \in L^2_{\mathrm{ac}}(X)$.

As was done in [Sá Barreto 2005], it is convenient to get rid of the term $A(\mu\partial_\mu + \nu\partial_\nu)$ in (3-5), by conjugating the operator by $\theta^{-1/4}$. Since Δ_h is the positive Laplacian, we find that, in local coordinates near a point $p \in \Gamma$,

$$\tilde{Q} = \theta^{1/4} \tilde{P} \theta^{-1/4} = \partial_\mu \partial_\nu + \mu\nu \sum_{i,j} h^{ij}(\mu\nu, y) \partial_{y_i} \partial_{y_j} + \mu\nu \sum_j \mathcal{B}_j(\mu\nu, y) \partial_{y_j} + C(\mu\nu, y), \tag{3-19}$$

where $C(\mu\nu, y)$ and $\mathcal{B}_j(\mu\nu, y)$ are C^∞, and $h^{-1} = (h^{ij})$ is the matrix associated with the metric h. Let $\tilde{W} = \theta^{1/4} W$; then $\tilde{Q}\tilde{W} = 0$. For $\phi(y) \in C_0^\infty(U)$, where $U \Subset \Gamma$ is such that (3-19) holds in $[0, \varepsilon] \times [0, \varepsilon] \times U$, let

$$G(\mu, \nu) = \int_{\partial X} \tilde{W}(\mu, \nu, y) \overline{\phi(y)} \, dy. \tag{3-20}$$

Notice that this is consistent with the conjugation of $\tilde{P}$ by $\theta^{1/4}$, and the factor $\theta^{1/2}$ is no longer present in the L^2 product. Let

$$Z(\mu\nu, y, D_y) = \tilde{Q} - \partial_\mu \partial_\nu = \mu\nu \sum_{i,j} h^{ij}(\mu\nu, y) \partial_{y_i} \partial_{y_j} + \mu\nu \sum_j \mathcal{B}_j(\mu\nu, y) \partial_{y_j} + C(\mu\nu, y),$$

and let $Z^*(\mu\nu, y, D_y)$ denote its adjoint with respect to the $L^2(\partial X)$ product defined by (3-20); then

$$\partial_\mu \partial_\nu G(\mu, \nu) = \int_{\partial X} \tilde{W}(\mu, \nu, y) \overline{Z^*(\mu\nu, y, D_y)\phi(y)} \, dy \tag{3-21}$$

It follows from (3-17) that there exists $C > 0$ such that

$$\begin{aligned} &\int_0^T |\partial_\mu \partial_\nu G(\mu, T)|^2 \, d\mu \le C \Big(\sum_{|\alpha| \le 2} \sup |\partial_y^\alpha \phi| \Big)^2 \|f\|^2_{L^2(X)}, \\ &\int_{\mu_0}^T |\partial_\mu \partial_\nu G(\mu_0, \nu)|^2 \, d\nu \le C \Big(\sum_{|\alpha| \le 2} \sup |\partial_y^\alpha \phi| \Big)^2 \|f\|^2_{L^2(X)} \quad \text{for } \mu_0 \in (0, T]. \end{aligned} \tag{3-22}$$

Let us write $K = \big(\sum_{|\alpha| \le 2} \sup |\partial_y^\alpha \phi| \big) \|f\|_{L^2(X)}$. Therefore, if $\delta < \mu < \varepsilon$,

$$|\partial_\nu G(\mu, \nu) - \partial_\nu G(\delta, \nu)| = \left| \int_\delta^\mu \partial_s \partial_\nu G(s, \nu) \, ds \right| \le CK(\mu - \delta)^{1/2}.$$

Hence, for $\nu > 0$,

$$\limsup_{\delta \to 0} |\partial_\nu G(\delta, \nu)| \le \liminf_{\mu \to 0} |\partial_\nu G(\mu, \nu)|,$$

so $\lim_{\mu \to 0} |\partial_\nu G(\mu, \nu)|$ exists. On the other hand, $\mathcal{R}_+(0, f)(s, y) = 0$ for $y \in \Gamma$ and $s \le s_0$, so according to (3-16) it follows that

$$\partial_\nu G(0, \nu) = 0, \quad \nu \in (0, T).$$

Now we use (3-22) to show that, if $0 \le \mu \le \nu \le T$, then there exists $C > 0$ such that

$$\begin{aligned} |\partial_\nu G(\mu, \nu)| = \left| \int_0^\mu \partial_s \partial_\nu G(s, \nu) \, ds \right| &\le \mu^{1/2} \Big(\int_0^\mu |\partial_s \partial_\nu G(s, \nu)|^2 \, ds \Big)^{\frac{1}{2}} \\ &\le \mu^{1/2} \Big(\int_0^\nu |\partial_s \partial_\nu G(s, \nu)|^2 \, ds \Big)^{\frac{1}{2}} \le CK\mu^{1/2}. \end{aligned} \tag{3-23}$$

Since $W(\mu, \mu, y) = 0$, we have, for $\mu \le \nu \le T$,

$$|G(\mu, \nu)| = \left| \int_\mu^\nu \partial_s G(\mu, s)\, ds \right| \le CK\mu^{1/2}(\nu - \mu). \tag{3-24}$$

This shows that, for every $\phi \in C_0^\infty(U)$,

$$\left| \int_{\partial X} \tilde{W}(\mu, \nu, y)\overline{\phi(y)}\, dy \right| \le CK\mu^{1/2},$$
$$\left| \int_{\partial X} \partial_\nu \tilde{W}(\mu, \nu, y)\overline{\phi(y)}\, dy \right| \le CK\mu^{1/2}.$$

Since $C_0^\infty(\mathbb{R}^2) \times C_0^\infty(U)$ spans $C_0^\infty(\mathbb{R}^2 \times U)$, it follows that for any $\psi(\mu, \nu, y)$, with $\mu, \nu \in [0, T]$,

$$\begin{aligned} \left| \int_{\partial X} \tilde{W}(\mu, \nu, y)\overline{\psi(\mu, \nu, y)}\, dy \right| &\le C\Big(\sum_{|\alpha| \le 2} \sup |\partial_y^\alpha \psi| \Big) \|f\|_{L^2(X)} \mu^{1/2}, \\ \left| \int_{\partial X} \partial_\nu \tilde{W}(\mu, \nu, y)\overline{\psi(\mu, \nu, y)}\, dy \right| &\le C\Big(\sum_{|\alpha| \le 2} \sup |\partial_y^\alpha \psi| \Big) \|f\|_{L^2(X)} \mu^{1/2}. \end{aligned} \tag{3-25}$$

Now we differentiate (3-21) with respect to ∂_ν. We have, for $\mu, \nu \in [0, T]$,

$$\partial_\nu \partial_\mu \partial_\nu G(\mu, \nu) = \int_{\partial X} \big[\partial_\nu \tilde{W}(\mu, \nu, y) \overline{Z^*(\mu\nu, y, D_y)\phi(y)} + \tilde{W}(\mu, \nu, y) \partial_\nu \overline{Z^*(\mu\nu, y, D_y))\phi(y)} \big]\, dy,$$

we apply (3-25) to $\psi(\mu, \nu, y) = Z^*(\mu\nu, y, D_y)\phi(y)$ and $\psi(\mu, \nu, y) = \partial_\nu Z^*(\mu\nu, y, D_y)\phi(y)$, and we conclude that

$$|\partial_\mu \partial_\nu^2 G(\mu, \nu, y)| \le C\Big(\sum_{|\alpha| \le 4} |\sup \partial_y^\alpha \phi| \Big) \|f\|_{L^2(X)} \mu^{1/2}$$

Let us denote $K_N(\phi) = \big(\sum_{|\alpha| \le N} |\sup \partial_y^\alpha \phi|\big) \|f\|_{L^2(X)}$. Since $\tilde{W}(\mu, \mu, y) = 0$, we have $\partial_\mu \partial_\nu G(\mu, \mu) = 0$, and so

$$|\partial_\mu \partial_\nu G(\mu, \nu)| = \left| \int_\mu^\nu \partial_\mu \partial_s^2 G(\mu, s)\, ds \right| \le K_4(\phi)\mu^{1/2}. \tag{3-26}$$

On the other hand, since $W(\mu, \mu, y) = 0$, it follows that $(\partial_\mu W)(\mu, \mu, y) = -(\partial_\nu W)(\mu, \mu, y)$. In particular, when $\nu = \mu$, we have

$$|\partial_\mu G(\mu, \mu)| \le CK_2(\phi)\mu^{1/2}$$

and, since

$$\partial_\mu G(\mu, \nu) = (\partial_\mu G)(\mu, \mu) + \int_\mu^\nu \partial_s \partial_\mu G(\mu, s)\, ds,$$

we have

$$|\partial_\mu G(\mu, \nu)| \le C(K_2(\phi) + K_4(\phi))\mu^{1/2}. \tag{3-27}$$

Proceeding as above, since $\partial_\nu G(0, \nu) = 0$, it follows from (3-26) that $|\partial_\nu G(\mu, \nu)| \le CK_4(\phi)\mu^{3/2}$ and, since $G(\mu, \mu) = 0$, we have $|G(\mu, \nu)| \le CK_4(\phi)\mu^{3/2}$ and $|\partial_\mu \partial_\nu^2 G(\mu, \nu)| \le CK_6(\phi)\mu^{3/2}$. Iterating this argument, and using the symmetry of W, we get that, for $k \geq 0$,

$$\partial_\mu^k G(0,\nu) = 0, \quad \partial_\nu^k G(\mu,0) = 0, \quad |(\partial_\mu G)(\mu,\mu)| = |(\partial_\nu G)(\mu,\mu)| \leq C\mu^k. \tag{3-28}$$

This shows that, in the sense of distributions, $\tilde{W}(\mu,\nu,y)$ vanishes to infinite order at

$$\{\mu = 0,\ \nu < T\} \times \Gamma \cup \{\nu = 0,\ \mu < T\} \times \Gamma,$$

where T has been chosen to be small enough that (2-4) holds for $x = \mu\nu < \varepsilon$. But this argument can be used finitely many times to show this holds for any $T \in (0, e^{s_0/2})$. In particular this shows that in the sense of distributions $\tilde{W}$ can be extended across the wedge $\{\mu = 0\} \cup \{\nu = 0\}$ so that

$$\begin{aligned} \tilde{Q}\tilde{W} &= 0 \quad \text{in } (-e^{s_0/2}, e^{s_0/2}) \times (-e^{s_0/2}, e^{s_0/2}) \times \Gamma = \mathcal{O}, \\ \tilde{W} &= 0 \quad \text{in } \{\mu < 0,\ 0 \leq \nu < e^{s_0/2}\} \times \Gamma \cup \{\nu < 0,\ 0 \leq \mu < e^{s_0/2}\} \times \Gamma. \end{aligned} \tag{3-29}$$

From (3-17) we know more about the regularity of $\tilde{W}$. We also know that

$$\tilde{W} \in C^\infty\big(\mathcal{O} \setminus (\{\mu = 0,\ \nu \geq 0\} \cup \{\nu = 0,\ \mu \geq 0\})\big),$$

and in fact Hörmander's propagation of singularities theorem implies that

$$WF(\tilde{W}) \subset \{\mu = 0,\ \nu \geq 0,\ \xi_1 = \xi_2 = 0\} \cup \{\nu = 0,\ \mu \geq 0,\ \xi_1 = \xi_2 = 0\}, \tag{3-30}$$

where ξ_1 and ξ_2 are dual to μ and ν respectively. If this were not true, singularities would propagate into the region where we know $\tilde{W}$ is C^∞. Indeed, the principal symbol of $\tilde{Q}$ is

$$q = -\xi_1\xi_2 - \mu\nu h(\mu\nu, y, \eta),$$

and hence its bicharacteristics satisfy

$$\begin{aligned} &\dot{\mu} = -\xi_2, && \mu(0) = \mu_0, && \dot{\nu} = -\xi_1, && \nu(0) = \nu_0, \\ &\dot{\xi}_1 = \nu(h + \mu\nu(\partial_x h)), && \xi_1(0) = \xi_{10}, && \dot{\xi}_2 = \mu(h + \mu\nu(\partial_x h)), && \xi_2(0) = \xi_{20}, \\ &\dot{y}_j = -\mu\nu\partial_{\eta_j} h, && y_j(0) = y_{j0}, && \dot{\eta}_j = \mu\nu\partial_{y_j} h, && \eta_j(0) = \eta_{j0}. \end{aligned}$$

Therefore, the bicharacteristics over $\mu = 0$ satisfy $\mu = 0$, $\xi_2 = 0$, $y = y_0$ and $\eta = \eta_0$ and

$$\dot{\nu} = -\xi_1, \quad \nu(0) = \nu_0, \quad \nu_0 \geq 0, \qquad \dot{\xi}_1 = \nu h(0, y_0, \eta_0), \quad \xi_1(0) = \xi_{10},$$

and hence, if we denote $h_0 = h(0, y_0, \eta_0)$,

$$\nu(t) = \nu_0 \cos(t\sqrt{h_0}) - \frac{\xi_{10}}{\sqrt{h_0}} \sin(t\sqrt{h_0}), \quad \xi_1(t) = \xi_{10}\cos(t\sqrt{h_0}) + \nu_0\sqrt{h_0}\sin(t\sqrt{h_0}).$$

If $(0, \nu_0, y_0, \xi_{10}, 0, \eta_0) \in WF(\tilde{W})$ with $\nu_0 \geq 0$ and $\xi_{10} > 0$, then $\nu(T) = -(\nu_0 + \xi_1)/\sqrt{2} < 0$ for $T = 3\pi/(4\sqrt{h_0})$, and so the point

$$\Big(0, -\frac{1}{\sqrt{2}}(\nu_0 + \xi_{10}), y_0, \frac{1}{\sqrt{2}}(-\xi_{10} + h_0\nu_0), 0, \eta_0\Big)$$

lies in $WF(\tilde{W})$. On the other hand, if $\xi_{10} < 0$, take $T = 5\pi/(4\sqrt{h_0})$ and so

$$\Big(0, \frac{1}{\sqrt{2}}(-\nu_0+\xi_{10}), y_0, -\frac{1}{\sqrt{2}}(\xi_{10}+h_0\nu_0), 0, \eta_0\Big) \in WF(\tilde{W}).$$

But this is not possible, since $\tilde{W} \in C^\infty$ in $\{\nu < 0\}$. The same analysis applies to $\{\nu = 0,\ \mu \geq 0\}$.

The next step is to prove the following unique continuation result:

Lemma 3.5. *Let $\Gamma \subset \partial X$ be open and not empty. Let $W(\mu, \nu, y)$ satisfy* (3-17), *and let $\tilde{W} = \theta^{1/4} W$ satisfy* (3-29). *Then for any $p \in \Gamma$ there exists $\delta > 0$ such that $\tilde{W}(\mu, \nu, y) = 0$ provided $|\mu| < \delta$, $|\nu| < \delta$ and $|y - p| < \delta$.*

Proof. It is not clear that this result is a consequence of Theorem 1.1.2 of [Alinhac 1984], but (3-31) below is similar to the estimates in Section 4.1 of [Alinhac 1984]. As usual, the proof of this result is based on a Carleman estimate. However, we need to be quite careful when applying the Carleman estimate, which is proved for C_0^∞ functions, to $\tilde{W}$. In general, one would have to cut off and mollify $\tilde{W}$ and then apply Friedrich's lemma; see for example the proof of [Hörmander 1994b, Theorem 28.3.4]. This usually requires the solution to be in H^1_{loc}. However, here the regularity for $\tilde{W}$ is given by (3-17), which is not quite H^1_{loc} near $\{\mu = 0\}$ or $\{\nu = 0\}$. We will avoid cutting $\tilde{W}$ in the variables (μ, ν), as the commutator of $\tilde{Q}$ with the cut-off function would produce terms in $\partial_\mu W$ and $\partial_\nu W$, which we cannot yet control. However cut offs in the y do not offer any problem, since the commutator of $\tilde{Q}$ with a cut-off function in y only would produce terms like $\mu\nu\partial_{y_j}\tilde{W}$, which can be controlled by (3-17). We will prove the following Carleman inequality, which will be used to prove the stated unique continuation from infinity, and will also be used to improve the regularity of $\tilde{W}$.

Lemma 3.6. *Let $p \in \Gamma$, and let $\tilde{Q}$ be the operator defined in* (3-19). *For $0 < \nu_0 \leq e^{s_0/2}$, let*

$$\Omega_\varepsilon = \{(\mu, \nu, y) : |\mu| < \varepsilon,\ |\nu| \leq \nu_0,\ |y - p| < 2\varepsilon\}, \quad \Sigma_{1,\varepsilon} = \big\{\nu = \nu_0,\ 0 \leq \mu \leq \tfrac{1}{2}\varepsilon,\ |y - p| < 2\varepsilon\big\},$$
$$\Omega_\varepsilon^+ = \{(\mu, \nu, y) \in \Omega_\varepsilon : \mu \geq 0,\ \nu \geq 0\}, \quad \Sigma_{2,\varepsilon} = \big\{\mu = \tfrac{1}{2}\varepsilon,\ 0 \leq \nu \leq \nu_0,\ |y - p| < 2\varepsilon\big\}.$$

Let $C_0 = \sup_{\Omega_\varepsilon} |C|$, where C is the zeroth order term of $\tilde{Q}$. Let $\gamma > 0$ be such that $\gamma C_0^2 \nu_0^3$ is small enough, and let $\varphi_a(\mu, \nu, y) = \mu + \gamma\nu + \frac{1}{2}a\gamma|y - p|^2$, where $a = 0$ or $a = 1$. Then there exist $\varepsilon_0 > 0$, $M > 0$ such that if $0 < \varepsilon < \varepsilon_0$ and $k \geq \frac{1}{4}$, then the following estimate holds for all $v(\mu, \nu, y) \in C^\infty(\Omega_\varepsilon)$ supported in $\{(\mu, \nu, y) :\ \mu \geq 0,\ \nu \geq 0,\ |y - p| \leq \varepsilon\}$:

$$\begin{aligned} M\|\varphi^{-k}\tilde{Q}v\| + Mk\int_{\Sigma_{1,\varepsilon}} [\mu\nu\varphi^{-1}|\nabla_y\varphi^{-k}v|^2 + k^2\varphi^{-3-2k}|v|^2]\,d\mu\,dy \\ + Mk\int_{\Sigma_{2,\varepsilon}} [\mu\nu\varphi^{-1}|\nabla_y\varphi^{-k}v|^2 + k^2\varphi^{-3-2k}|v|^2]\,d\nu\,dy \\ \geq k^3\|\varphi^{-k-2}v\|^2 + k^2\|\varphi^{-1}\partial_\mu\varphi^{-k}v\|^2 + k^2\|\varphi^{-1}\partial_\nu\varphi^{-k}v\|^2 + k\|(\mu+\gamma\nu)^{1/2}\varphi^{-1/2}\nabla_y\varphi^{-k}v\|^2, \end{aligned} \tag{3-31}$$

where $\|v\|^2 = \int_{\Omega_\varepsilon^+} |v|^2\,d\mu\,d\nu\,dy$.

Proof. The estimate with $a = 0$ was proved in [Sá Barreto 2005]. We are doing it again here for the convenience of the reader, and we will use it to improve the regularity of $\tilde{W}$. But this estimate with $a = 0$ is not strong enough to prove the unique continuation result, for which we need the estimate with $a = 1$. We will use $\varphi = \varphi_a$ in the proof to simplify the already heavy notation.

Without loss of generality, we assume that $p=0$ and that v is real-valued. We know from (3-19) that

$$\tilde{Q}(\mu,\nu,y,\partial_\mu,\partial_\nu,\partial_y)=\partial_\mu\partial_\nu+\mu\nu\sum_{i,j=1}^{n}h^{ij}(\mu\nu,y)\partial_{y_i}\partial_{y_j}+\mu\nu\sum_{j=1}^{n}\mathcal{B}_j(\mu\nu,y)\partial_{y_j}+C(\mu\nu,y).$$

As usual, we define $\tilde{Q}_k=\varphi^{-k}\tilde{Q}\varphi^k$ and, since $\partial_\mu\varphi=1$, $\partial_\nu\varphi=\gamma$ and $\partial_{y_j}\varphi=a\gamma y_j$, we have

$$\tilde{Q}_k=\varphi^{-k}\tilde{Q}\varphi^k=\tilde{Q}(\mu,\nu,y,\partial_\mu+k\varphi^{-1},\partial_\nu+k\gamma\varphi^{-1},\partial_y+ka\gamma y\varphi^{-1}),$$

and we write

$$\tilde{Q}_k=\mathcal{Q}_k+k\mathcal{L},$$

with

$$\mathcal{L}=\varphi^{-1}(\partial_\nu+\gamma\partial_\mu),$$

$$\mathcal{Q}_k=\partial_\mu\partial_\nu+\gamma(k^2-k)\varphi^{-2}+\mu\nu h^{ij}(\mu\nu,y)(\partial_{y_i}+ka\gamma y_i\varphi^{-1})(\partial_{y_j}+ka\gamma y_j\varphi^{-1}) \\ +\mu\nu\mathcal{B}_j(\partial_{y_j}+ka\gamma y_j\varphi^{-1})+C,$$

where we used the notation $\sum_{ij=1}^{n}A_{ij}B_{ij}=A_{ij}B_{ij}$ and $D_jE_j=\sum_{j=1}^{n}D_jE_j$ to indicate sums over repeated indices. Therefore,

$$\|\tilde{Q}_kv\|^2=\|\mathcal{Q}_kv\|^2+k^2\|\mathcal{L}v\|^2+2k\langle\mathcal{Q}_kv,\mathcal{L}v\rangle, \tag{3-32}$$

where

$$\langle u,v\rangle=\int_{\Omega_\varepsilon^+}uv\,dy\,d\mu\,d\nu\quad\text{and}\quad\|v\|^2=\langle v,v\rangle.$$

The first term of (3-32) is positive and we will compute $k^2\|\mathcal{L}v\|^2+2k\langle\mathcal{Q}_kv,\mathcal{L}v\rangle$. Since v is supported in $\{\mu\geq 0,\ \nu\geq 0\}$, we will assume that $\mu\geq 0$ and $\nu\geq 0$ in the computations below. We will also use M for a generic constant. The first term of $\langle\mathcal{Q}_kv,\mathcal{L}v\rangle$ is

$$\begin{aligned}&\langle\partial_\mu\partial_\nu v,\varphi^{-1}(\partial_\nu+\gamma\partial_\mu)v\rangle\\ &\quad=\frac{1}{2}\int_{\Omega_\varepsilon^+}\varphi^{-1}(\partial_\mu(\partial_\nu v)^2+\gamma\partial_\nu(\partial_\mu v)^2)\,dy\,d\mu\,d\nu\\ &\quad=\frac{1}{2}\int_{\Omega_\varepsilon^+}(\partial_\mu(\varphi^{-1}(\partial_\nu v)^2)+\partial_\nu(\gamma\varphi^{-1}\partial_\mu v)^2)\,dy\,d\mu\,d\nu+\frac{1}{2}\int_{\Omega_\varepsilon^+}\varphi^{-2}(\gamma^2(\partial_\mu v)^2+(\partial_\nu v)^2)\,dy\,d\mu\,d\nu\\ &\quad\geq\tfrac{1}{2}(\gamma^2\|\varphi^{-1}\partial_\mu v\|^2+\|\varphi^{-1}\partial_\nu v\|^2).\end{aligned} \tag{3-33}$$

Here we used that v and all its derivatives vanish at $\{\mu=0\}\cup\{\nu=0\}$, and the boundary terms in $\Sigma_{j,\varepsilon}$, $j=1,2$ are nonnegative. The next term is

$$\begin{aligned}&\gamma(k^2-k)\langle\varphi^{-2}v,\varphi^{-1}(\gamma\partial_\mu+\partial_\nu)v\rangle\\ &\quad=\tfrac{1}{2}\gamma(k^2-k)\int_{\Omega_\varepsilon^+}\varphi^{-3}(\gamma\partial_\mu+\partial_\nu)v^2\,dy\,d\mu\,d\nu\\ &\quad=\tfrac{1}{2}\gamma(k^2-k)\int_{\Omega_\varepsilon^+}(\gamma\partial_\mu+\partial_\nu)(\varphi^{-3}v^2)\,dy\,d\mu\,d\nu+3\gamma^2(k^2-k)\int_{\Omega_\varepsilon^+}\varphi^{-4}|v|^2\,d\mu\,dy\\ &\quad=\tfrac{1}{2}\gamma(k^2-k)\int_{\Sigma_{1,\varepsilon}}\varphi^{-3}v^2\,d\mu\,dy+\tfrac{1}{2}\gamma^2(k^2-k)\int_{\Sigma_{2,\varepsilon}}\varphi^{-3}v^2\,d\nu\,dy+3\gamma^2(k^2-k)\|\varphi^{-2}v\|^2.\end{aligned} \tag{3-34}$$

Since we want to prove (3-31) for all $k \geq \frac{1}{4}$, we need to get rid of the negative term $-3k\gamma^2\|\varphi^{-2}v\|^2$ in (3-34). To do this we use the term $\|\varphi^{-1}\partial_\nu v\|^2$ from (3-33). Notice that $\varphi^{-1}\partial_\nu v = \partial_\nu(\varphi^{-1}v) + \gamma\varphi^{-2}v$, and hence

$$(\varphi^{-1}\partial_\nu v)^2 \geq \gamma^2\varphi^{-4}v^2 + 2\gamma\varphi^{-2}v\partial_\nu(\varphi^{-1}v) = \gamma^2\varphi^{-4}v^2 + \gamma\varphi^{-1}\partial_\nu(\varphi^{-1}v)^2.$$

Therefore,

$$\|\varphi^{-1}\partial_\nu v\|^2 \geq 2\gamma^2\|\varphi^{-2}v^2\|^2,$$

and so

$$3\gamma^2(k^2-k)\|\varphi^{-2}v\|^2 + \tfrac{7}{16}\|\varphi^{-1}\partial_\nu v\|^2 \geq 3\gamma^2\big(k^2-k+\tfrac{7}{24}\big)\|\varphi^{-2}v\|^2 \geq \tfrac{3}{8}k^2\gamma^2\|\varphi^{-2}v\|^2.$$

Hence, the first two terms satisfy

$$\begin{aligned}\langle\partial_\mu\partial_\nu v, \varphi^{-1}(\partial_\nu+\gamma\partial_\nu)v\rangle + (k^2-k)\langle\varphi^{-2}v, \varphi^{-1}(\gamma\partial_\mu+\partial_\nu)v\rangle \\ \geq \tfrac{1}{2}\gamma^2\|\varphi^{-1}\partial_\mu v\|^2 + \tfrac{1}{16}\|\varphi^{-1}\partial_\nu v\|^2 + \tfrac{3}{8}k^2\gamma^2\|\varphi^{-2}v\|^2 \\ + \tfrac{1}{2}(k^2-k)\int_{\Sigma_{1,\varepsilon}}\varphi^{-3}v^2\,d\mu\,dy + \tfrac{1}{2}\gamma^2(k^2-k)\int_{\Sigma_{2,\varepsilon}}\varphi^{-3}v^2\,d\nu\,dy.\end{aligned} \tag{3-35}$$

To estimate the third term, we integrate by parts in y_j, recalling that v is compactly supported in the y variable in the interior of Ω_ε^+. We use that h^{ij} is symmetric to write it as

$$\begin{aligned}&\langle\mu\nu h^{ij}(\partial_{y_i}+ka\gamma y_i\varphi^{-1})(\partial_{y_j}+ka\gamma y_j\varphi^{-1})v, \mathscr{L}v\rangle \\ &\quad= \frac{1}{2}\int_{\Omega_\varepsilon^+}\mu\nu h^{ij}\,[(\partial_{y_i}+ka\gamma y_i\varphi^{-1})(\partial_{y_j}+ka\gamma y_j\varphi^{-1})v]\mathscr{L}v\,dy\,d\mu\,d\nu \\ &\qquad+ \frac{1}{2}\int_{\Omega_\varepsilon^+}\mu\nu h^{ij}\,[(\partial_{y_j}+kay_j\varphi^{-1})(\partial_{y_i}+ka\gamma y_i\varphi^{-1})v]\mathscr{L}v\,dy\,d\mu\,d\nu = I + II,\end{aligned}$$

where

$$\begin{aligned}I = &-\frac{1}{2}\int_{\Omega_\varepsilon^+}\mu\nu h^{ij}(\partial_{y_j}v+ka\gamma y_j\varphi^{-1}v)[(\partial_{y_i}-ka\gamma y_i\varphi^{-1})\mathscr{L}v]\,dy\,d\mu\,d\nu \\ &-\frac{1}{2}\int_{\Omega_\varepsilon^+}\mu\nu h^{ij}(\partial_{y_i}v+ka\gamma y_i\varphi^{-1}v)[(\partial_{y_j}-ka\gamma y_j\varphi^{-1})\mathscr{L}v]\,dy\,d\mu\,d\nu,\end{aligned}$$

$$II = -\int_{\Omega_\varepsilon^+}[\partial_{y_i}(\mu\nu h^{ij})](\partial_{y_j}v+ka\gamma y_j\varphi^{-1}v)\mathscr{L}v\,dy\,d\mu\,d\nu.$$

We can bound II from below by using that

$$\begin{aligned}\partial_{y_i}(\mu\nu h^{ij})(\partial_{y_j}v+ka\gamma y_j\varphi^{-1}v)\mathscr{L}v &\geq -M(\mu\nu)^{3/4}|\partial_{y_j}v+ka\gamma y_j\varphi^{-1}v|\,(\mu\nu)^{1/4}|\mathscr{L}v| \\ &\geq -M\big((\mu\nu)^{3/2}|\nabla_y v|^2 + k^2a^2\gamma^2(\mu\nu)^{3/2}|y|^2\varphi^{-2}v^2 + (\mu\nu)^{1/2}|\mathscr{L}v|^2\big).\end{aligned}$$

Hence,

$$II \geq -M\big(\|(\mu\nu)^{3/4}\nabla_y v\|^2 + \gamma^2k^2a^2\|(\mu\nu)^{3/4}|y|\varphi^{-1}v\|^2 + \|(\mu\nu)^{1/4}\mathscr{L}v\|^2\big). \tag{3-36}$$

Using that

$$(\partial_{y_i}-ka\gamma y_i\varphi^{-1})\mathscr{L}v = \mathscr{L}(\partial_{y_i}-ka\gamma y_i\varphi^{-1})v - a\gamma y_j\varphi^{-1}\mathscr{L}v - 2ka\gamma^2 y_i\varphi^{-3}v,$$

we write $I = I_1 + I_2$, where

$$\begin{aligned} I_1 &= -\frac{1}{2}\int_{\Omega_\varepsilon^+} \mu\nu h^{ij}(\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)[\mathscr{L}(\partial_{y_i} v - ka\gamma y_i\varphi^{-1}v)]\,dy\,d\mu\,d\nu \\ &\quad -\frac{1}{2}\int_{\Omega_\varepsilon^+} \mu\nu h^{ij}(\partial_{y_i} v + ka\gamma y_i\varphi^{-1}v)[\mathscr{L}(\partial_{y_j} v - ka\gamma y_j\varphi^{-1}v)]\,dy\,d\mu\,d\nu \qquad (3\text{-}37) \\ I_2 &= a\int_{\Omega_\varepsilon^+} \mu\nu h^{ij}(\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)(\gamma y_i\varphi^{-1}\mathscr{L}v + 2k\gamma^2 y_i\varphi^{-3}v)\,dy\,d\mu\,d\nu. \end{aligned}$$

To bound the term I_2 from below, we write

$$\begin{aligned} &\mu\nu h^{ij}(\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)(\gamma y_i\varphi^{-1}\mathscr{L}v + 2k\gamma^2 y_i\varphi^{-3}v) \\ &\qquad \geq -M|y|^{1/2}\mu\nu\varphi^{-1}(|\partial_{y_i} v| + ka\gamma|y|\varphi^{-1}|v|)|y|^{1/2}(\gamma|\mathscr{L}v| + ka\gamma^2\varphi^{-2}|v|) \\ &\qquad \geq -M\big(|y|(\mu\nu)^2\varphi^{-2}|\nabla_y v|^2 + \gamma^2|y|(\mathscr{L}v)^2 + k^2a^2\gamma^2|y|^3(\mu\nu)^2\varphi^{-4}|v|^2 + k^2a^2\gamma^4|y|\varphi^{-4}|v|^2\big). \end{aligned}$$

Therefore,

$$I_2 \geq -Ma\big(\||y|^{1/2}\mu\nu\varphi^{-1}\nabla_y v\|^2 + \gamma^2\||y|^{1/2}\mathscr{L}v\|^2 + k^2a^2\gamma^2\||y|^{3/2}\mu\nu\varphi^{-2}v\|^2 + k^2a^2\gamma^4\||y|^{1/2}\varphi^{-2}v\|^2\big) \qquad (3\text{-}38)$$

Next we consider the term I_1. Since $\mathscr{L} = \varphi^{-1}(\partial_\mu + \partial_\nu)$, integrating by parts in μ and ν we conclude that the term I_1 satisfies

$$\begin{aligned} I_1 &= -\frac{1}{2}\int_{\Omega_\varepsilon^+} \mu\nu h^{ij}\mathscr{L}\big[(\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)(\partial_{y_i} v - ka\gamma y_i\varphi^{-1}v)\big]\,dy\,d\mu\,d\nu \\ &= -\frac{1}{2}\int_{\Omega_\varepsilon^+} (\gamma\partial_\mu + \partial_\nu)\big[(\mu\nu\varphi^{-1}h^{ij})(\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)(\partial_{y_i} v - ka\gamma y_i\varphi^{-1}v)\big]\,dy\,d\mu\,d\nu \\ &\qquad + \frac{1}{2}\int_{\Omega_\varepsilon^+} \big[(\gamma\partial_\mu + \partial_\nu)(\mu\nu\varphi^{-1}h^{ij})\big](\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)(\partial_{y_i} v - ka\gamma y_i\varphi^{-1}v)\,dy\,d\mu\,d\nu \\ &= -\frac{1}{2}\int_{\Sigma_{1,\varepsilon}} \mu\nu\varphi^{-1}h^{ij}((\partial_{y_i} + ka\gamma y_j\varphi^{-1})v)((\partial_{y_j} - ka\gamma y_j\varphi^{-1})v)\,d\mu\,dy \\ &\qquad - \frac{\gamma}{2}\int_{\Sigma_{2,\varepsilon}} \mu\nu\varphi^{-1}h^{ij}((\partial_{y_i} + ka\gamma y_j\varphi^{-1})v)((\partial_{y_j} - ka\gamma y_j\varphi^{-1})v)\,d\nu\,dy \\ &\qquad\quad + \frac{1}{2}\int_{\Omega_\varepsilon^+} \big[(\gamma\partial_\mu + \partial_\nu)(\mu\nu\varphi^{-1}h^{ij})\big](\partial_{y_j} v + ka\gamma y_j\varphi^{-1}v)(\partial_{y_i} v - ka\gamma y_i\varphi^{-1}v)\,dy\,d\mu\,d\nu. \end{aligned}$$

Notice that

$$\begin{aligned} (\gamma\partial_\mu + \partial_\nu)(\mu\nu h^{ij}(\mu\nu, y)\varphi^{-1}) &= \big[(\gamma\nu + \mu)\varphi^{-1} - 2\gamma\mu\nu\varphi^{-2}\big]h^{ij} + (\mu + \gamma\nu)\mu\nu\varphi^{-1}(\partial_x h^{ij}) \\ &= \varphi^{-2}\big[\big((\mu + \gamma\nu)\big(\mu + \gamma\nu + \tfrac{1}{2}a\gamma|y|^2\big) - 2\gamma\mu\nu\big)h^{ij}(\mu\nu, y) \\ &\qquad + \mu\nu(\mu + \gamma\nu)\big(\mu + \gamma\nu + \tfrac{1}{2}a\gamma|y|^2\big)(\partial_x h^{ij})(\mu\nu, y)\big] \\ &= \varphi^{-2}\big[\big(\mu^2 + \gamma^2\nu^2 + \tfrac{1}{2}a\gamma(\mu + \gamma\nu)|y|^2\big)h^{ij}(\mu\nu, y) \\ &\qquad + \mu\nu(\mu + \gamma\nu)\big(\mu + \gamma\nu + \tfrac{1}{2}a\gamma|y|^2\big)(\partial_x h^{ij})(\mu\nu, y)\big]. \qquad (3\text{-}39) \end{aligned}$$

Hence,

$$|(\gamma\partial_\mu + \partial_\nu)(\mu\nu h^{ij}(\mu\nu, y)\varphi^{-1})| \le M\varphi^{-1}(\mu + \gamma\nu). \tag{3-40}$$

On the other hand, since h^{ij} is positive definite, we know that there exists $M > 0$ such that

$$h^{ij} W_i W_j \ge M|W|^2, \quad W \in \mathbb{R}^n, \tag{3-41}$$

We conclude from (3-39), (3-40), (3-41) and the symmetry of h^{ij} that, for ε small enough, there exists M such that

$$\begin{aligned}
&[(\partial_\mu + \partial_\nu)(\mu\nu h^{ij}\varphi^{-1})](\partial_{y_j} v + ka\gamma y_j \varphi^{-1} v)(\partial_{y_j} v - ka\gamma y_j\varphi^{-1} v) \\
&\qquad = [(\partial_\mu + \partial_\nu)(\mu\nu h^{ij}\varphi^{-1})](\partial_{y_i} v \partial_{y_j} v - k^2a^2\gamma^2 y_i y_j \varphi^{-2} v^2) \\
&\qquad \ge M(\mu + \gamma\nu)\varphi^{-1}|\nabla_y v|^2 - Mk^2a^2(\mu+\gamma\nu)\gamma^2|y|^2\varphi^{-3}|v|^2.
\end{aligned} \tag{3-42}$$

Hence, for ε small enough,

$$\begin{aligned}
I_1 \ge M\|\varphi^{-1/2}(\mu+\gamma\nu)^{1/2}\nabla_y v\|^2 &- Mk^2a^2\gamma^2\| |y|(\mu+\gamma\nu)^{1/2}\varphi^{-3/2} v\|^2 \\
&- M\int_{\Sigma_{1,\varepsilon}} \mu\nu(\varphi^{-1}|\nabla_y v|^2 + k^2a^2\varphi^{-3}|y|^2v^2)\,d\mu\,dy \\
&- M\int_{\Sigma_{2,\varepsilon}} \mu\nu(\varphi^{-1}|\nabla_y v|^2 + k^2a^2\varphi^{-3}|y|^2v^2)\,d\nu\,dy.
\end{aligned} \tag{3-43}$$

We write the last term of $\langle \mathcal{Q}_k v, \mathcal{L}v\rangle$ as

$$\begin{aligned}
&\langle \mu\nu\mathcal{B}_j(\partial_{y_j} + ka\gamma y_j\varphi^{-1})v + Cv, \mathcal{L}v\rangle \\
&\quad = \langle \mu\nu\varphi^{-1/2}\mathcal{B}_j(\partial_{y_j} + ka\gamma y_j\varphi^{-1})v + \varphi^{-1/2}Cv, \varphi^{1/2}\mathcal{L}v\rangle \\
&\quad \ge -\|\varphi^{1/2}\mathcal{L}v\|^2 - \|C\varphi^{-1/2}v\|^2 - Mk^2a^2\gamma^2\| |y|\mu\nu\varphi^{-3/2}v\|^2 - M\|(\mu\nu)\varphi^{-1/2}\nabla_y v\|^2,
\end{aligned} \tag{3-44}$$

Therefore, provided ε_0 is small enough, we deduce from equations (3-35), (3-36), (3-38), (3-43) and (3-44) that

$$\begin{aligned}
k^2\|\mathcal{L}v\|^2 + 2k\langle \mathcal{Q}_k v, \mathcal{L}v\rangle &+ Mk\int_{\Sigma_{1,\varepsilon}} (\mu\nu\varphi^{-1}|\nabla_y v|^2 + k^2\varphi^{-3}v^2)\,d\mu\,dy \\
&+ Mk\int_{\Sigma_{2,\varepsilon}} (\mu\nu\varphi^{-1}|\nabla_y v|^2 + k^2\varphi^{-3}v^2)\,d\nu\,dy \\
\ge \tfrac{1}{2}k\gamma^2\|\varphi^{-1}\partial_\mu v\|^2 &+ \tfrac{1}{16}k\|\varphi^{-1}\partial_\nu v\|^2 + \int_{\Omega_\varepsilon^+} (k^2 - kMF_1(\mu,\nu,y))|\mathcal{L}v|^2\,d\mu\,d\nu\,dy \\
&+ k\int_{\Omega_\varepsilon^+} |\nabla_y v|^2(M_1(\mu+\gamma\nu)\varphi^{-1} - MF_2(\mu,\nu,y))\,d\mu\,d\nu\,dy \\
&+ k\int_{\Omega_\varepsilon^+} k^2\gamma^4\varphi^{-4}v^2\big(\tfrac{3}{8} - MF_3(\mu,\nu,y)\big)\,d\mu\,d\nu\,dy - k\int_{\Omega_\varepsilon^+} |C|\varphi^{-1}v^2\,d\mu\,d\nu\,dy,
\end{aligned} \tag{3-45}$$

where

$$F_1(\mu,\nu,y)=(\mu\nu)^{1/2}+\gamma^2|y|+\varphi,$$
$$F_2(\mu,\nu,y)=(\mu\nu)^{3/2}+|y|(\mu\nu)^2\varphi^{-2}+(\mu\nu)^2\varphi^{-1},$$
$$F_3(\mu,\nu,y)=(\mu\nu)^{3/2}|y|^2\varphi^2+|y|^3(\mu\nu)^2+\gamma^2|y|+|y|^2(\mu+\gamma\nu)\varphi+|y|^2(\mu\nu)^2\varphi.$$

The term involving C is the most problematic. Recall that $\varphi=\mu+\gamma\nu+\frac{1}{2}a\gamma|y|^2$ and, since $|\mu|\le\varepsilon$, $|y|\le\varepsilon$ and $\nu\le\nu_0$, it follows that $\varphi\le\varepsilon+\gamma\nu_0+\frac{1}{2}a\gamma\varepsilon^2$. Therefore, if $C_0=\sup_{\Omega_\varepsilon}|C|$,

$$\tfrac{3}{8}k^2\gamma^2\varphi^{-4}-|C|^2\varphi^{-1}\ge\varphi^{-4}\big(\tfrac{3}{8}k^2\gamma^2-C_0^2\varphi^3\big)\ge\varphi^{-4}\big(\tfrac{3}{8}k^2\gamma^2-9C_0^2\big(\varepsilon^3+\gamma^3\nu_0^3+\tfrac{1}{8}a\gamma^3\varepsilon^6\big)\big).$$

If one picks γ such that $9\gamma C_0^2\nu_0^3<\frac{3}{256}$, then

$$\tfrac{3}{8}k^2-9\gamma C_0^2\nu_0^3\ge\tfrac{3}{16}k^2\quad\text{for all }\ k\ge\tfrac{1}{4},$$

and therefore

$$\tfrac{3}{8}k^2\gamma^2\varphi^{-4}-|C|^2\varphi^{-1}\ge\varphi^{-4}\big(\tfrac{3}{16}k^2\gamma^2-9C_0^2\big(\varepsilon^3+\tfrac{1}{8}a\gamma^3\varepsilon^6\big)\big)\quad\text{for all }\ k\ge\tfrac{1}{4}.$$

Notice also that $\mu\le\varphi$, and hence the coefficient of $|\nabla v|^2$ in (3-45) satisfies

$$\begin{aligned}M_1(\mu+\gamma\nu)\varphi^{-1}-M((\mu\nu)^{3/2}+|y|(\mu\nu)^2\varphi^{-2}+(\mu\nu)^2\varphi^{-1})&\\ \ge\varphi^{-1}\big(M_1(\mu+\gamma\nu)-M|((\mu\nu)^{3/2}\varphi+|y|\mu\nu^2+(\mu\nu)^2)\big)&\\ \ge\tfrac{1}{2}M_1(\mu+\gamma\nu)\varphi^{-1}\quad\text{for }\ \varepsilon_0\ \text{ small enough.}&\end{aligned}$$

One can then pick ε_0, such that for every $\varepsilon\in(0,\varepsilon_0)$,

$$\begin{aligned}k^2\|\mathscr{L}v\|^2+2k\langle\mathfrak{Q}_kv,\mathscr{L}v\rangle\ &+Mk\int_{\Sigma_{1,\varepsilon}}(\mu\nu\varphi^{-1}|\nabla_yv|^2+k^2\varphi^{-3}v^2)\,d\mu\,dy\\ &+Mk\int_{\Sigma_{2,\varepsilon}}(\mu\nu\varphi^{-1}|\nabla_yv|^2+k^2\varphi^{-3}v^2)\,d\nu\,dy\\ \ge M\big(k\|(\mu+\gamma\nu)^{1/2}\varphi^{-1/2}\nabla_yv\|^2&+k^2\|\mathscr{L}v\|^2+k\|\varphi^{-1}\partial_\mu v\|^2+k\|\varphi^{-1}\partial_\nu v\|^2+k^3\gamma^2\|\varphi^{-2}v\|^2\big),\end{aligned}$$

This ends the proof of Lemma 3.6. □

Next we want to use (3-31) to prove Lemma 3.5. Let $\chi\in C_0^\infty\big(\{|y|<\varepsilon/4\}\big)$, $\chi=1$ on $\{|y|\le\varepsilon/8\}$. Let $V(\mu,\nu,y)=\chi(y)\tilde{W}(\mu,\nu,y)$. We choose $\psi(y)$ to be a C_0^∞ function supported in $\{|y|<\varepsilon/4\}$ with $\int\psi(y)\,dy=1$, and define $\psi_\delta(y)=(\delta)^{-n}\psi(y/\delta)$, $\delta>0$. Then, for δ small enough,

$$V_\delta=\psi_\delta*'V\in C_0^\infty(\Omega_{2\varepsilon})\quad\text{is supported in}\quad\big\{\mu\ge0,\ \nu\ge0,\ |y|\le\tfrac{1}{2}\varepsilon\big\}.$$

where $*'$ denotes convolution in the y variable. To see that, let $\zeta(\mu,\nu)\in C_0^\infty$; then the Fourier transform $\widehat{\zeta V}_\delta$ satisfies

$$\widehat{\zeta V}_\delta(\xi_1,\xi_2,\eta)=\widehat{\psi}(\delta\eta)(\widehat{\zeta V})(\xi_1,\xi_2,\eta),$$

which in view of (3-30) is rapidly decaying in any conic neighborhood of a point $(\xi_{10},\xi_{20},\eta_0)\ne0$. Hence $V_\delta\in C^\infty$, and (3-31) holds for V_δ. Now we would like to take the limit of (3-31) for V_δ as $\delta\to0$.

Notice that $\varphi \geq \varepsilon$ on $\Sigma_{2,\varepsilon}$ and $\varphi \geq \gamma\nu_0$ on $\Sigma_{1\varepsilon}$ and, in view of (3-17),

$$\begin{aligned}
&\int_{\Sigma_{1,\varepsilon}} [\mu\nu|\nabla_y \varphi^{-k}\tilde{W}|^2 + k^2|\varphi^{-2-k}\tilde{W}|^2]\, d\mu\, dy < M(\gamma\nu_0)^{-k}, \\
&\int_{\Sigma_{2,\varepsilon}} [\mu\nu|\nabla_y \varphi^{-k}\tilde{W}|^2 + k^2|\varphi^{-2-k}\tilde{W}|^2]\, d\nu\, dy < M\varepsilon^{-k},
\end{aligned} \tag{3-46}$$

and these terms in (3-31) do not offer any problem when passing to the limit.

One cannot use (3-31) with $a = 0$ to prove Lemma 3.5, however we will use it here to show that

$$\begin{gathered}
(\mu+\gamma\nu)^{-k}\nabla V, \quad (\mu+\gamma\nu)^{-k-1}\partial_\nu V, \quad (\mu+\gamma\nu)^{-k-1}\partial_\mu V, \\
(\mu+\gamma\nu)^{-k-2} V \in L^2(\Omega_\varepsilon) \quad \text{with } k \geq \tfrac{1}{4}.
\end{gathered} \tag{3-47}$$

For now, we take $a = 0$ and $\varphi = \mu+\gamma\nu$. We know from (3-17) that $\tilde{W}$, $[\mu\nu(\mu+\gamma\nu)]^{1/2}\nabla_y\tilde{W} \in L^2(\Omega_\varepsilon)$. Since $\mu\gamma\nu \leq \frac{1}{2}(\mu+\gamma\nu)^2$, it follows that $\gamma(\mu+\gamma\nu)^{-1}(\mu\nu)^2 \leq (\mu+\gamma\nu)\mu\nu$, and hence one can apply Friedrich's lemma — see, for example, Lemma 17.1.5 of [Hörmander 1994a] — to show that

$$\lim_{\delta\to 0} \big\| (\mu+\gamma\nu)^{-1/4}\mu\nu[(h^{ij}\partial_{y_i}\partial_{y_j} + \mathcal{B}_j\partial_{y_j})\psi_\delta *' V - \psi_\delta *' (h^{ij}\partial_{y_i}\partial_{y_j} + \mathcal{B}_j\partial_{y_j} V)] \big\| = 0 \tag{3-48}$$

We also know from (3-17) that, for fixed $T > 0$, $\mu^{1/2}\partial_\mu \tilde{W}(\mu, T, y) \in L^2([0, T] \times \partial X)$. Hence the same holds for V and for V_δ for all $\delta > 0$. One can easily show that

$$\mu(\partial_\mu V_\delta)^2 \geq \tfrac{1}{4}\mu^{-1}(\log\mu)^{-2}V_\delta^2 - \partial_\mu((-\log\mu)^{-1}V_\delta^2).$$

Since V_δ vanishes to infinite order at $\mu = 0$, if we integrate the above on $\big[0, \frac{1}{2}\varepsilon\big] \times \partial X$ we obtain

$$\int_{\partial X} \Big(\log\frac{2}{\varepsilon}\Big)^{-1} V_\delta\big(\tfrac{1}{2}\varepsilon, T, y\big)\, dy + \int_0^T\int_{\partial X} \mu(\partial_\mu V_\delta)^2\, dy\, d\mu \geq \int_0^T\int_{\partial X} \mu^{-1}(\log\mu)^{-2}V_\delta^2\, dy\, d\mu. \tag{3-49}$$

Since, in view of (3-17), the left-hand side is finite for V, if one applies (3-49) to $V_\delta - V_{\delta'}$ it follows that V_δ is a Cauchy sequence in the norm given by the right-hand side of (3-49). So it converges and, since V_δ converges weakly to V, we conclude that $\mu^{-1/2}|\log\mu|^{-1}V \in L^2(\Omega_\varepsilon)$, and in particular

$$(\mu+\nu)^{-1/4} V \in L^2(\Omega_\varepsilon). \tag{3-50}$$

Since $\tilde{Q}$ is given by (3-19), it follows from (3-48) and (3-50) that

$$\lim_{\delta\to 0} \big\| (\mu+\nu)^{-1/4}(\tilde{Q}(\psi_\delta *' V) - \psi_\delta *' (\tilde{Q}V)) \big\| = 0. \tag{3-51}$$

Since $\tilde{Q}\tilde{W} = 0$, it follows that

$$\tilde{Q}V = \tilde{Q}(\chi(y)\tilde{W}) = \mu\nu h^{ij}(\tilde{W}\partial_{y_i}\partial_{y_j}\chi + 2\partial_{y_i}\chi\partial_{y_j}\tilde{W}) + \mu\nu(\mathcal{B}_j\partial_{y_j}\chi)\tilde{W}.$$

So we conclude that, in view of (3-17), $(\mu+\nu)^{-1/4}\tilde{Q}V \in L^2(\Omega_\varepsilon)$ and hence

$$\lim_{\delta\to 0} \|(\mu+\gamma\nu)^{-k}\tilde{Q}V_\delta\|_{L^2(\Omega_\varepsilon)} = \|(\mu+\gamma\nu)^{-k}\tilde{Q}V\| < \infty, \quad k = \tfrac{1}{4}. \tag{3-52}$$

Therefore (3-31), with $a = 0$ and $k = \frac{1}{4}$, holds for V in place of V_δ, and in particular we conclude that (3-47) holds for $k = \frac{1}{4}$ (notice that in this case $(\mu + \gamma\nu)\varphi^{-1} = 1$). We then apply the argument used above to show that (3-31) holds for $k = \frac{1}{4} + 1$, and hence (3-47) holds for $k = \frac{1}{4} + 1$, and by induction and interpolation, this shows that (3-47) holds for all $k \geq \frac{1}{4}$.

Now we use the same argument with $\varphi = \varphi_1 = \mu + \gamma\nu + \frac{1}{2}\gamma|y|^2$. Notice that in this case $\varphi \geq \mu + \gamma\nu$ and we have from (3-47) that

$$\varphi^{-k}\nabla_y V, \quad \varphi^{-k-1}\partial_\nu V, \quad \varphi^{-k-1}\partial_\mu V, \quad \varphi^{-k-2}V \in L^2(\Omega_\varepsilon), \quad k \geq \tfrac{1}{4}. \tag{3-53}$$

Since φ depends on y, it is not clear how to apply Friedrich's lemma in the bootstrapping argument above to prove (3-53), as one would have to analyze the commutator of the convolution and the weight, which is of course singular. But, given (3-53), Friedrich's lemma can be easily applied and we conclude that (3-31) holds for V and $\varphi = \varphi_1$. In particular we conclude from (3-46) that, for ε small enough,

$$Mk^3\varepsilon^{-k} + C\|\varphi^{-k}\tilde{Q}\chi(y)\tilde{W}\|^2 \geq k^3\|\varphi^{-2-k}\chi(y)\tilde{W}\|^2. \tag{3-54}$$

Now we really use the power of (3-31) with $a = 1$: since $\tilde{Q}\tilde{W} = 0$, and $\chi = 1$ for $|y| \leq \frac{1}{8}\varepsilon$, $\tilde{Q}(\chi(y)\tilde{W}) = [\tilde{Q}, \chi(y)]\tilde{W}$ is supported in $|y| \geq \frac{1}{8}\varepsilon$, and hence $\varphi \geq \lambda\varepsilon^2$ on the support of $\tilde{Q}V$, where $\lambda = \frac{1}{128}\gamma$. We deduce from (3-54) that, for ε small enough, there exists $C = C(\tilde{W}) > 0$ such that

$$C(\lambda\varepsilon^2)^{-2k} \geq \|\varphi^{-2-k}\chi(y)\tilde{W}\|^2.$$

Hence,

$$\left\|\left(\frac{\varphi}{\lambda\varepsilon^2}\right)^{-k}\chi(y)\tilde{W}\right\| \leq C, \quad k > 1,$$

and therefore $\tilde{W}(\mu, \nu, y) = 0$ if $\varphi \leq \lambda\varepsilon^2$, and in particular $\tilde{W} = 0$ if $0 \leq \mu \leq \frac{1}{3}\lambda\varepsilon^2$, $0 \leq \gamma\nu \leq \frac{1}{3}\lambda\varepsilon^2$ and $\gamma|y|^2 \leq \frac{1}{3}\lambda\varepsilon^2$. This ends the proof of Lemma 3.5, and consequently the proof of Proposition 3.2. □

Notice that since $\nu_0 \in (0, e^{s_0/2})$ is arbitrary, this result also establishes regularity for $\tilde{W}$, and in particular it shows that $\tilde{W} \in H^1_{\text{loc}}$.

Proof of Proposition 3.3. We will use Hörmander's unique continuation theorem, and we will find a function whose level surfaces are strictly pseudoconvex. The key point here is that the coefficients of the operator P defined in (3-2) do not depend on s, and hence P is invariant under translations in the variable s. Let

$$\varphi(x, s, y) = -x - \kappa(s - s_1) - |y - p|^2, \quad \text{where } \kappa > 0 \text{ small will be chosen later.}$$

Since, for $|y - p| < \delta$, $V = 0$ if $x \in (-\varepsilon, 0]$ and $s < s_0$, or if $x < \delta$ and $\log x < s < s_1$, we have — see Figure 6 —

$$V(x, s, y) = 0 \quad \text{if } \varphi > 0, \ -\varepsilon < x < \delta, \ \text{and } |y - p| < \delta. \tag{3-55}$$

The principal symbol of the operator P is

$$p = -2\sigma\xi - x\xi^2 - xh(x, y, \eta), \tag{3-56}$$

where (ξ, σ, η) are the dual variables to (x, s, y). Since $\nabla\varphi(x, s, y) = (-1, -\kappa, -2(y-p))$, we have

$$p(x, s, y, \nabla\varphi(x, s, y)) = -2\kappa - x\big(1 + h(x, y, 2(y-p))\big). \tag{3-57}$$

If $|y-p| < \beta$ is small enough and $x > -\kappa$, then $x\big(1 + h(x, y, 2(y-p))\big) > -\frac{3}{2}\kappa$, and hence $p(x, s, y, \nabla\varphi) < -\frac{1}{2}\kappa$. Therefore φ is not characteristic at (x, s, y) if $x > -\kappa$ and $|y-p| < \beta$, for small enough β.

The Hamilton vector field of p is

$$H_p = -2\xi\partial_s - 2(\sigma + x\xi)\partial_x - xH_h + (\xi^2 + h + x\partial_x h)\partial_\xi, \tag{3-58}$$

where H_h denotes the Hamilton vector field of $h(x, y, \eta)$ in the variables (y, η). Hence,

$$(H_p\varphi)(x, s, y, \xi, \sigma, \eta) = 2(\sigma + x\xi) + 2\kappa\xi + xH_h|y-p|^2 \quad \text{and}$$

$$\begin{aligned}(H_p^2\varphi)&(x, s, y, \xi, \sigma, \eta)\\ &= -2(\sigma + x\xi)(2\xi + H_h|y-p|^2 + x\partial_x H_h|y-p|^2) - (xH_h)^2|y-p|^2 + 2(\kappa + x)(\xi^2 + h + x\partial_x h).\end{aligned}$$

If $H_p\varphi = 0$, it follows that

$$\begin{aligned}H_p^2\varphi(x, s, y, \xi, \sigma, \eta) = 2(x + 3\kappa)\xi^2 + 2\xi((x+\kappa)H_h|y-p|^2 + \kappa x\partial_x H_h|y-p|^2) + 2(\kappa + x)(h + x\partial_x h)\\ + x((H_h|y-p|^2)^2 + xH_h|y-p|^2\partial_x H_h|y-p|^2 - xH_h^2|y-p|^2).\end{aligned}$$

If $|y-p| < \beta$ is small enough, there exists $C > 0$ depending on h only such that

$$|H_p|y-p|^2| \le C\beta|\eta| \quad \text{and} \quad |\partial_x H_p|y-p|^2| \le C\beta|\eta|.$$

If we impose that $-\frac{1}{2}\kappa < x < \beta$, it follows that there exists $\varepsilon_0 > 0$ depending on h such that, if $\beta, \kappa \in (0, \varepsilon_0)$ are small, then there exists $C > 0$ such that

$$h + x\partial_x h \ge C|\eta|^2,$$

and hence

$$\begin{aligned}H_p^2\varphi(x, s, p, \xi, \sigma, \eta) &\ge \kappa C(\xi^2 - \beta|\xi||\eta| + |\eta|^2)\\ &\ge C\kappa(\xi^2 + |\eta|^2) \qquad \text{if } -\tfrac{1}{2}\kappa < x < \beta,\ |y-p| < \beta \text{ and } \kappa, \delta \in (0, \varepsilon_0).\end{aligned}$$

So we conclude that there exists $\varepsilon_0 > 0$ depending on h such that

$$\begin{aligned}p(x, s, y, \xi, \sigma, \eta) = H_p\varphi(x, s, y, \xi, \sigma, \eta) = 0 \quad \Longrightarrow \quad H_p^2\varphi(x, s, y, \xi, \sigma, \eta) > 0\\ \text{if } (\xi, \sigma, \eta) \ne 0,\ -\tfrac{1}{2}\kappa < x < \beta,\ |y-p| < \beta,\ \kappa, \beta \in (0, \varepsilon_0).\end{aligned} \tag{3-59}$$

Since P is of second order, we deduce from (3-57) and (3-59) that the level surfaces of φ are strictly pseudoconvex in the region

$$-\tfrac{1}{2}\kappa < x < \beta,\ |y-p| < \beta \quad \text{provided } \kappa,\ \beta \in (0, \varepsilon_0); \tag{3-60}$$

see for example the first paragraph of Section 28.4 of [Hörmander 1994b]. As mentioned above, the fact that the coefficients of P do not depend on s imply that the conditions in (3-60) do not depend on s. Now we appeal to Theorem 28.2.3 and Proposition 28.3.3 of [Hörmander 1994b] and conclude that, if

$$Y = \left\{-\tfrac{1}{4}\kappa < x < \tfrac{1}{2}\beta,\ |y-p| < \tfrac{1}{\sqrt{2}}\beta,\ |s-s_1| < s_0 - s_1\right\},$$

then there exist $C > 0$, $\tau_0 > 0$ and $\lambda > 0$ large such that, if $\psi = e^{\lambda\varphi}$,

$$C\|e^{\tau\psi}Pv\|^2 \geq \tau^2\|e^{\tau\psi}v\|^2 + \tau\|e^{\tau\psi}v\|^2_{H^1} \quad \text{for all } v \in C_0^\infty(Y) \text{ and } \tau \geq \tau_0 > 0. \tag{3-61}$$

Let $\theta \in C_0^\infty(Y)$ with $\theta = 1$ if $-\frac{1}{8}\kappa < x < \frac{1}{4}\beta$, $|y-p| < \frac{1}{2}\beta$ and $|s-s_1| < \frac{3}{4}(s_0-s_1)$. Since $PV = 0$, it follows that

$$P(\theta V) = [P, \theta]V.$$

But, for $(x, s, y) \in Y$, $V(x, s, y)$ is supported in the region $x > 0$, $s > s_1$, so we conclude that

$$P(\theta(x, s, y)V) \text{ is supported in } (x, s, y) \in Y,\ x \geq \tfrac{1}{4}\beta, \text{ or } s - s_1 \geq \tfrac{3}{4}(s_0 - s_1), \text{ or } |y-p| \geq \tfrac{1}{2}\beta.$$

Therefore, by the definition of φ we have

$$\varphi(x, s, y) \leq -\min\left\{\tfrac{1}{4}\beta,\ \tfrac{1}{4}3\kappa(s_0-s_1),\ \tfrac{1}{4}\beta^2\right\} \quad \text{on the support of } P(\theta V). \tag{3-62}$$

Pick κ small so that $\min\left\{\frac{1}{4}\beta, \frac{3}{4}\kappa(s_0-s_1), \frac{1}{4}\beta^2\right\} = \frac{3}{4}\kappa(s_0-s_1) = \gamma$. We deduce from (3-61) and (3-62) that

$$\tau^2\|e^{\tau(e^{\lambda\varphi}-e^{-\lambda\gamma})}\theta V\|^2 \leq C, \quad \tau > \tau_0.$$

We remark that, due to Friedrich's lemma, one can apply (3-61) to θV even though V is not C^∞; see [Hörmander 1994b]. Therefore, $\theta V = 0$ if $e^{\lambda\varphi} - e^{-\lambda\gamma} > 0$, so $\theta V = 0$ if $\varphi > -\gamma$. So we deduce that

$$\theta V(x, s, y) = 0 \quad \text{provided } \kappa(s-s_1) < \tfrac{1}{3}\gamma,\ 0 < x < \tfrac{1}{3}\gamma\ |y-p|^2 < \tfrac{1}{3}\gamma.$$

In particular,

$$V(x, s, y) = 0 \quad \text{provided } s < s_1 + \tfrac{1}{4}(s_0-s_1),\ 0 < x < \tfrac{1}{3}\gamma,\ |y-p|^2 < \tfrac{1}{3}\gamma. \tag{3-63}$$

This concludes the proof of Proposition 3.3. □

Proof of Proposition 3.4. The key point in the proof is the following consequence of Tataru's theorem [1995; 1999]; see also [Hörmander 1997; Robbiano and Zuily 1998]. Let Ω be a C^∞ manifold equipped with a C^∞ Riemannian metric g. Let L be a first-order C^∞ operator that does not depend on t. If $u(t, z)$ is a C^∞ function that satisfies

$$(D_t^2 - \Delta_g + L(z, D_z))u = 0 \quad \text{in } (-\tilde{T}, \tilde{T}) \times \Omega,$$
$$u(t, z) = 0 \text{ in a neighborhood of } \{z_0\} \times (-T, T),\ T < \tilde{T},$$

then

$$u(t, z) = 0 \quad \text{if } |t| + d_g(z, z_0) < T, \tag{3-64}$$

where d_g is the distance measured with respect to the metric g.

Since the initial data of (2-1) is $(0, f)$, $u(t, z) = -u(-t, z)$. If $0 < x < \gamma$, $\log x < s < s_1$, and $|y - p| < \delta$, it follows from the definition of V_+ that

$$u(t, x, y) = 0 \quad \text{if } 0 < x < \gamma,\ |y - p| < \delta \text{ and } |t| \le s_2 - \log x = \log \frac{e^{s_2}}{x}.$$

Applying (3-64) with $z_0 = (x, y)$, we obtain

$$u(t, z) = 0 \quad \text{provided } |t| + d_g(z; (x, y)) < \log \frac{e^{s_2}}{x} \text{ with } 0 < x < \delta,\ |y - p| < \delta.$$

If $z = (\alpha, y)$ with $e^{s^*} > \alpha > x$, $d_g((x, y); (\alpha, y)) = \log(\alpha/x)$, it follows from (3-64) that

$$u(t, (\alpha, y)) = 0 \quad \text{if } t + \log \frac{\alpha}{x} < \log \frac{e^{s_2}}{x}.$$

In particular this guarantees that $u(t, \alpha, y) = 0$ if $0 < t < \log(e^{s_2}/\alpha)$ and, since $s = t + \log \alpha$, we have $V_+(\alpha, s, y) = 0$ if $\alpha < e^{s^*}$, $\log x < s < s_2$ and $|y - p| < \delta$. This ends the proof of Proposition 3.4. □

Proof of Theorem 2.1. As promised at the beginning of the section, we shall now finish the proof of Theorem 2.1. We start with (3-10), which says that $V_+(x, s, y) = x^{-n/2} u(s - \log x, x, y)$ satisfies $V_+(x, s, y) = 0$ if $y \in \Gamma$, $x < e^{s^*}$ and $\log x < s < s < s_0$.

Now we recall that $V_+(x, s, y) = x^{-n/2} u(s - \log x, x, y)$ and so, if $w = (\alpha, p)$ with $0 < \alpha < e^{s^*}$ and $p \in \Gamma$, then the solution $u(t, z)$ vanishes in a neighborhood of $\{w\} \times (0, \log(e^{s_0}/\alpha))$. Again we use that the data is of the form $(0, f)$, and hence $u(-t, z) = -u(t, z)$. So in fact $u(t, z)$ vanishes in a neighborhood of $\{w\} \times (-\log(e^{s_0}/\alpha), \log(e^{s_0}/\alpha))$. Therefore, by (3-64),

$$u(t, z) = \partial_t u(t, z) = 0 \quad \text{if } |t| + d_g(z, w) < \log \frac{e^{s_0}}{\alpha}.$$

In particular, when $t = 0$ we find that $\partial_t u(0, z) = f(z) = 0$ provided $d_g(z, w) < \log(e^{s_0}/\alpha)$, and this concludes the proof of Theorem 2.1. □

Final remarks. The following result will be useful in the next section:

Corollary 3.7. *Let (X, g) be a connected AHM and let $\Gamma \subset \partial X$ be open, $\Gamma \neq \varnothing$. If $f \in L^2_{\mathrm{ac}}(X)$ and $\mathcal{R}_+(0, f)(s, y) = 0$ in $\mathbb{R} \times \Gamma$, then $f = 0$. Similarly, if $(h, 0) \in E_{\mathrm{ac}}(X)$ and $\mathcal{R}_+(h, 0)(s, y) = 0$ in $\mathbb{R} \times \Gamma$, then $h = 0$.*

Proof. If $\mathcal{R}_+(0, f)(s, y) = 0$ in $\mathbb{R} \times \Gamma$, then $f(z) = 0$ if $z \in \mathcal{D}_{s_0}(\Gamma)$ for every s_0. Since the distance between any two points in the interior of X is finite, it follows that $f = 0$.

Suppose $F = \mathcal{R}_+(h, 0)(s, y) = 0$ in $\mathbb{R} \times \Gamma$. As in the proof of Proposition 3.2, by taking the convolution of F with $\phi \in C_0^\infty(\mathbb{R})$, even, we may assume that $(\Delta_g - n^2/4)^k h \in L^2_{\mathrm{ac}}(X)$ for every $k \ge 0$. Let $u(t, z)$ satisfy (2-1) with initial data $(h, 0)$ and let $V = \partial_t u$. Then V satisfies (2-1) with initial data $(0, (\Delta_g - n^2/4)h)$ and $\mathcal{R}_+(0, (\Delta_g - n^2/4)h)(s, y) = 0$ in $\mathbb{R} \times \Gamma$. But, as we have shown, this implies that $(\Delta_g - n^2/4)h = 0$. Since $(h, 0) \in E_{\mathrm{ac}}(X)$, this implies that $h = 0$. □

One should remark that this result can be proved by applying a result of Mazzeo [1991]; see also [Vasy and Wunsch 2005]. The solution to (2-1) with initial data $(0, f)$ is odd and, since $\mathcal{R}_+(0, f)(s, y) = 0$ for $s \in \mathbb{R}$, $y \in \Gamma$, it follows that $\mathcal{R}_-(0, f)(s, y) = 0$ for $s \in \mathbb{R}$, $y \in \Gamma$. Taking the Fourier transform in t, we find that

$$\big(\Delta_g - \lambda^2 - \tfrac{1}{4}n^2\big)\hat{u}(\lambda, z) = 0$$

and, using that $\mathcal{R}_+(0, f)(s, y) = \mathcal{R}_-(0, f)(s, y) = 0$, one deduces that $\hat{u}(\lambda, z)$ vanishes to infinite order at Γ, using a formal power series argument as in the proof of Proposition 3.4 of [Graham and Zworski 2003]. Theorem 14 of [Mazzeo 1991] implies that $\hat{u} = 0$ and hence $u = 0$. In particular, $f = 0$.

4. The control space

As we saw in (3-13) and (3-14), the ranges of the forward and backward radiation fields

$$\mathcal{R}_\pm(0, L^2_{\text{ac}}(X)) = \{\mathcal{R}_\pm(0, f) : f \in L^2_{\text{ac}}(X)\}$$

are closed subspaces of $L^2(\mathbb{R} \times \partial X)$ and are characterized by the scattering operator. Moreover, since $\mathcal{R}_\pm$ are unitary, $\|\mathcal{R}_\pm(0, f)\|_{L^2(\mathbb{R}\times\partial X)} = \|f\|_{L^2(X)}$. The main goal of this section is to show that the ranges $\{\mathcal{R}_\pm(0, f)|_{\mathbb{R}\times\Gamma}\}$ are determined by the restriction of the scattering operator to Γ, as defined in (2-13). Throughout the remainder of the paper we shall write

$$L^2(\mathbb{R} \times \Gamma) = \{F|_{\mathbb{R}\times\Gamma} : F \in L^2(\mathbb{R} \times \partial X)\}.$$

The key observation is:

Lemma 4.1. *If $F = \mathcal{R}_+(h, f) \in L^2(\mathbb{R} \times \Gamma)$, then*

$$\|f\|_{L^2(X)} = \big\langle F, \tfrac{1}{2}(F + S_\Gamma F^*)\big\rangle,$$

and in particular $\|f\|_{L^2(X)}$ is determined by $\mathcal{S}_\Gamma F$.

Proof. If $F(s, y) = \mathcal{R}_+(h, f) \in L^2(\mathbb{R} \times \Gamma)$, so in particular F is supported in $\mathbb{R} \times \Gamma$ then, according to (3-12) and the fact that $\mathcal{R}_+$ is unitary,

$$\begin{aligned}\big\langle F, \tfrac{1}{2}(F + \mathcal{S}_\Gamma F^*)\big\rangle = \big\langle F, \tfrac{1}{2}(F + (\mathcal{S}F^*)|_{\mathbb{R}\times\Gamma})\big\rangle = \big\langle F, \tfrac{1}{2}(F + \mathcal{S}F^*)\big\rangle \\ = \langle \mathcal{R}_+(h, f), \mathcal{R}_+(0, f)\rangle = \|f\|^2_{L^2(X)}.\end{aligned}$$ □

This suggests that

$$\mathbb{C}^n_+\big(\mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}\big) = \|f\|_{L^2(X)}$$

defines a norm on the space $\{\mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma} : f \in L^2_{\text{ac}}(X)\}$. We shall prove that it does and, moreover, the norm is determined by $\mathcal{S}_\Gamma$.

Theorem 4.2. *Let $\Gamma \subset \partial X$ be a nonempty open subset such that $\partial X \setminus \Gamma$ does not have empty interior. The space*

$$\mathcal{M}(\Gamma)^\pm = \{\mathcal{R}_\pm(0, f)|_{\mathbb{R}\times\Gamma} : f \in L^2_{\text{ac}}(X)\},$$

equipped with norm $\mathbb{C}^n_\pm$ defined by

$$\mathbb{C}^n_\pm(\mathcal{R}_\pm(0, f)|_{\mathbb{R}\times\Gamma}) = \|f\|_{L^2(X)}, \tag{4-1}$$

is a Hilbert space determined by $\mathcal{S}_\Gamma$.

Proof. We shall work with the forward radiation field. The proof of the result for $\mathcal{R}_-$ is identical. Since $\mathcal{R}_+$ is linear, the triangle inequality for the $L^2(X)$-norm implies that $\mathbb{C}^n_+$ is a norm, and that

$$\big\langle \mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}, \mathcal{R}_+(0, h)|_{\mathbb{R}\times\Gamma}\big\rangle_{\mathbb{C}^n_+} = \langle f, h\rangle_{L^2(X)}$$

is an inner product. Since $\mathcal{R}_+$ is continuous and $L^2_{\mathrm{ac}}(X)$ is complete, it follows that $(\mathcal{M}(\Gamma)^+, \mathbb{C}^n_+)$ is a Hilbert space. We need to show that it is determined by $\mathcal{S}_\Gamma$. We recall from (3-12) that if $F = \mathcal{R}_+(f, h)$ then

$$\tfrac{1}{2}(F + \mathcal{S}F^*)|_{\mathbb{R}\times\Gamma} = \mathcal{R}_+(0, h)|_{\mathbb{R}\times\Gamma}. \tag{4-2}$$

So, if $F \in L^2(\mathbb{R}\times\Gamma)$, then $F^* \in L^2(\mathbb{R}\times\Gamma)$ and hence $(F + \mathcal{S}F^*)|_{\mathbb{R}\times\Gamma} = F + \mathcal{S}_\Gamma F^*$. We shall let

$$\begin{aligned} \mathcal{L} : L^2(\mathbb{R}\times\Gamma) &\to L^2(\mathbb{R}\times\Gamma) \\ F &\mapsto \tfrac{1}{2}(F + \mathcal{S}_\Gamma F^*). \end{aligned} \tag{4-3}$$

Since $\mathcal{S}$ is unitary, it follows that $\|\mathcal{L}\| \le 1$. Since $\mathcal{R}_+$ is unitary, given $F \in L^2(\mathbb{R}\times\Gamma)$ there exists $(f, h) \in E_{\mathrm{ac}}(X)$ such that $\mathcal{R}_+(f, h) = F$. We can say the following about such initial data:

Lemma 4.3. *Let $\Gamma \subset \partial X$ be a nonempty open subset such that $\partial X \setminus \Gamma$ contains an open set $\mathcal{O}$, and let $h \in L^2_{\mathrm{ac}}(X)$. Then there exists at most one f such that $(f, 0) \in E_{\mathrm{ac}}(X)$ and $\mathcal{R}_+(f, h)$ is supported in $\mathbb{R}\times\Gamma$. Moreover, the set*

$$\mathcal{C}(\Gamma) = \{h \in L^2_{\mathrm{ac}}(X) : \textit{there exists } (f, 0) \in E_{\mathrm{ac}}(X) \textit{ such that } \mathcal{R}_+(f, h)(s, y) = 0,\ y \in \partial X \setminus \Gamma\}$$

is dense in $L^2_{\mathrm{ac}}(X)$.

Proof. First, if $\mathcal{R}_+(f_1, h)$ and $\mathcal{R}_+(f_2, h)$ are supported in $\mathbb{R}\times\overline{\Gamma}$, then $\mathcal{R}_+(f_1 - f_2, 0)$ is supported in $\mathbb{R}\times\overline{\Gamma}$, but this implies that $\mathcal{R}_+(f_1 - f_2, 0) = 0$ in $\mathbb{R}\times\mathcal{O}$, and so Corollary 3.7 implies that $f_1 = f_2$.

If $v \in L^2_{\mathrm{ac}}(X)$ is such that $\langle v, h\rangle_{L^2(X)} = 0$ for all $h \in \mathcal{C}(\Gamma)$ then, since $\mathcal{R}_+$ is unitary, for all $(f, 0) \in E_{\mathrm{ac}}(X)$,

$$\langle v, h\rangle_{L^2(X)} = \langle \mathcal{R}_+(0, v), \mathcal{R}_+(f, h)\rangle_{L^2(\mathbb{R}\times\partial X)}$$

Since $h \in \mathcal{C}(\Gamma)$ is arbitrary, it follows that

$$\langle \mathcal{R}_+(0, v), F\rangle_{L^2(\mathbb{R}\times\partial X)} = 0 \quad \text{for all } F \in L^2(\mathbb{R}\times\Gamma).$$

Hence $\mathcal{R}_+(0, v) = 0$ on $\mathbb{R}\times\Gamma$ and, by Corollary 3.7, $v = 0$. □

Lemma 4.4. *If $\Gamma \subset \partial X$ is open, nonempty and $\partial X \setminus \Gamma$ contains an open subset, then the map $\mathcal{L}$ is injective and has dense range.*

Proof. If $F = \mathcal{R}_+(f, h) \in L^2(\mathbb{R}\times\Gamma)$, then $\mathcal{L}F = \mathcal{R}_+(0, h)|_{\mathbb{R}\times\Gamma}$. If $\mathcal{L}F = 0$ then $\mathcal{R}_+(0, h) = 0$ on $\mathbb{R}\times\Gamma$. It follows from Corollary 3.7 that $h = 0$, and hence $F = \mathcal{R}(f, 0)$. Since there exists an open subset $\mathcal{O} \subset (\partial X \setminus \Gamma)$, and F is supported in $\mathbb{R}\times\overline{\Gamma}$, it follows that $F = \mathcal{R}_+(f, 0) = 0$ in $\mathbb{R}\times\mathcal{O}$, and again by Corollary 3.7, $f = 0$ and so $F = 0$.

Now we prove that its range is dense. Let $H \in L^2(\mathbb{R}\times\Gamma)$ be orthogonal to the range of $\mathcal{L}$. Suppose that $H = \mathcal{R}_+(h_1, h_2)$, with $(h_1, h_2) \in E_{\mathrm{ac}}(X)$. Then for every $F = \mathcal{R}_+(f, h) \in L^2(\mathbb{R}\times\Gamma)$, $h \in \mathscr{C}(\Gamma)$,

$$\begin{aligned}0 = \langle H, (F+\mathcal{S}F^*)|_{\mathbb{R}\times\Gamma}\rangle_{L^2(\mathbb{R}\times\Gamma)} = \langle H, F+\mathcal{S}F^*\rangle_{L^2(\mathbb{R}\times\Gamma)} &= \langle H, \mathcal{R}_+(0,h)\rangle_{L^2(\mathbb{R}\times\partial X)}\\ &= \langle \mathcal{R}_+(h_1,h_2), \mathcal{R}_+(0,h)\rangle_{L^2(\mathbb{R}\times\partial X)}\\ &= \langle h_2, h\rangle_{L^2(X)}.\end{aligned}$$

Since $\mathscr{C}(\Gamma)$ is dense in $L^2_{\mathrm{ac}}(X)$, $h_2 = 0$. Hence $H = \mathcal{R}_+(h_1, 0) = 0$ on $\mathbb{R}\times\mathbb{O}$, and so $H = 0$. □

We shall let

$$\mathcal{F}^+(\Gamma) = \mathcal{L}(L^2(\mathbb{R}\times\Gamma)) = \{\mathcal{R}_+(0,f)|_{\mathbb{R}\times\Gamma} : f \in \mathscr{C}(\Gamma)\}, \tag{4-4}$$

and equip $\mathcal{F}^+(\Gamma)$ with the norm given by Lemma 4.1,

$$\mathbb{C}^n_+(\mathcal{R}_+(0,f)) = \|f\|_{L^2(X)}.$$

Thus $(\mathcal{F}^+(\Gamma), \mathbb{C}^n_+)$ is a normed vector space and, since $\mathscr{C}(\Gamma)$ is dense in $L^2(X)$, $\mathcal{F}^+(\Gamma)$ is dense in $(\mathcal{M}^+(\Gamma), \mathbb{C}^n_+)$. Hence $(\mathcal{M}^+(\Gamma), \mathbb{C}^n_+)$ is the completion of $(\mathcal{F}^+(\Gamma), \mathbb{C}^n_+)$ into a Hilbert space, and therefore it is determined by $\mathcal{S}_\Gamma$. Notice that the completion of $\mathcal{F}^+(\Gamma)$ with the $L^2(\mathbb{R}\times\Gamma)$-norm is $L^2(\mathbb{R}\times\Gamma)$. But

$$\|\mathcal{R}_+(0,h)|_{(\mathbb{R}\times\Gamma)}\|_{L^2(\mathbb{R}\times\Gamma)} \le \|h\|_{L^2(X)},$$

so $\mathbb{C}^n_+$ is a stronger norm and $(\mathcal{M}^+(\Gamma), \mathbb{C}^n_+)$ is a smaller space. This ends the proof of Theorem 4.2. □

5. Proof of Theorem 2.3

The operators $\mathcal{S}_{j,\Gamma}$, $j = 1, 2$ were defined in terms of boundary-defining functions for which (2-14) holds for both g_1 and g_2 in $U_j \sim [0,\varepsilon)\times\partial X_j$. In particular,

$$\Psi_j^* g_j = \frac{dx^2}{x^2} + \frac{h_j(x)}{x^2} \quad \text{on } (0,\varepsilon)\times\Gamma, \quad h_1(0) = h_2(0) = h_0 \text{ on } \Gamma. \tag{5-1}$$

Our first step will be to prove that there exists $\varepsilon > 0$ such that the tensors $h_1(x)$ and $h_2(x)$ are equal on $[0,\varepsilon)\times\Gamma$. Once this is done, if $\Psi_j : [0,\varepsilon)\times\partial X_j \to U_j$, $j = 1, 2$, are the maps that satisfy (2-14), and if $W_{1,\varepsilon} = \Psi_1([0,\varepsilon)\times\Gamma)$, $W_{2,\varepsilon} = \Psi_{2,\varepsilon}([0,\varepsilon)\times\Gamma)$, then

$$\Psi_1^*(g_1|_{W_{1,\varepsilon}}) = \Psi_2^*(g_2|_{W_{2,\varepsilon}}) \quad \text{on } [0,\varepsilon)\times\Gamma. \tag{5-2}$$

Since $\Psi_j = \mathrm{Id}$ on Γ, $j = 1, 2$, this implies that

$$\Psi_\varepsilon = \Psi_2\circ\Psi_1^{-1} : W_{1,\varepsilon} \mapsto W_{2,\varepsilon}, \quad (\Psi_2\circ\Psi_1)^{-1} g_2 = g_1, \quad \Psi_\varepsilon = \mathrm{Id} \text{ on } \Gamma \tag{5-3}$$

gives an isometry between neighborhoods of Γ.

The local diffeomorphism. We will prove that, if $h_j(x)$ are such that (5-1) holds, then $h_1(x) = h_2(x)$ on $[0, \varepsilon) \times \Gamma$, and hence this gives the map Ψ_ε defined in (5-3). Our first step in this construction will be:

Proposition 5.1. *Let* (X_1, g_1), (X_2, g_2) *and* Γ *satisfy the hypotheses of Theorem 2.3, and denote by* $\mathcal{R}_{j,\pm}(s, y, x', y')$ *the Schwartz kernels of* $\mathcal{R}_{j,\pm}$ *acting on* $(0, f)$. *Then there exists* $\epsilon > 0$ *such that* (2-14) *holds on* $[0, \varepsilon) \times \partial X_j$, $j = 1, 2$, *and*

$$\begin{aligned} h_1(x, y, dy) &= h_2(x, y, dy) && \text{if } x \in [0, \varepsilon),\ y \in \Gamma, \\ \mathcal{R}_{1,\pm}(s, y, x', y') &= \mathcal{R}_{2,\pm}(s, y, x', y') && \text{if } y, y' \in \Gamma,\ x' < \varepsilon. \end{aligned} \tag{5-4}$$

Proof. The proof of Proposition 5.1 is an adaptation of the boundary control method of [Belishev 1987; Belishev and Kurylev 1992] to this setting. By working on an open subset of Γ if necessary, we may assume that $\partial X \setminus \Gamma$ does not have empty interior. As in [Sá Barreto 2005], pick $x_1 < \varepsilon$ and consider the spaces

$$\begin{aligned} \mathcal{M}^+_{x_1}(\Gamma) &= \{F \in \mathcal{M}^+(\Gamma) : F(s, y) = 0,\ s \le \log x_1\}, \\ \mathcal{M}^-_{x_1}(\Gamma) &= \{F \in \mathcal{M}^-(\Gamma) : F(s, y) = 0,\ s \ge -\log x_1\}, \end{aligned}$$

and let

$$\mathcal{P}^+_{x_1} : \mathcal{M}^+(\Gamma) \to \mathcal{M}^+_{x_1}(\Gamma) \quad \text{and} \quad \mathcal{P}^-_{x_1} : \mathcal{M}^-(\Gamma) \to \mathcal{M}^-_{x_1}(\Gamma) \tag{5-5}$$

denote the respective orthogonal projections with respect to the norms $\mathbb{C}^n_\pm$ defined in (4-1). Since $\mathcal{M}^\pm(\Gamma)$ and $\mathcal{M}^\pm_{x_1}(\Gamma)$ are determined by $\mathcal{S}_\Gamma$, the projections $\mathcal{P}^\pm_{x_1}$ are also determined by $\mathcal{S}_\Gamma$. Notice that $(\mathcal{P}^+_{x_1} F)(s, y)$ is not necessarily equal to $H(s - \log x_1)F(s, y)$, where H is the Heaviside function, as $H(s - \log x_1)F(s, y)$ may not be in $\mathcal{M}^+(\Gamma)$.

In view of finite speed of propagation and Theorem 2.1,

$$\begin{aligned} \mathcal{M}^+_{x_1}(\Gamma) &= \{\mathcal{R}_+(0, h)|_{\mathbb{R}\times\Gamma} : h \in L^2_{\mathrm{ac}}(X),\ h(z) = 0,\ z \in \mathcal{D}_{\log x_1}(\Gamma)\}, \\ \mathcal{M}^-_{x_1}(\Gamma) &= \{\mathcal{R}_-(0, h)|_{\mathbb{R}\times\Gamma} : h \in L^2_{\mathrm{ac}}(X),\ h(z) = 0,\ z \in \mathcal{D}_{\log x_1}(\Gamma)\}. \end{aligned}$$

As in [Sá Barreto 2005], the key to proving Proposition 5.1 is to understand the effect of the projectors $\mathcal{P}^\pm_{x_1}$ on the initial data. First we deal with the case where Δ_{g_j}, $j = 1, 2$, have no eigenvalues. In this case, $L^2(X_j) = L^2_{\mathrm{ac}}(X_j)$.

Lemma 5.2. *Let* (X, g) *be an asymptotic hyperbolic manifold such that* Δ_g *has no eigenvalues. Let* x *be such that* (2-4) *holds in* $(0, \epsilon) \times \partial X$. *For* $x_1 \in (0, \epsilon)$, *let* $\mathcal{P}^+_{x_1}$ *denote the orthogonal projector defined in* (5-5). *Let* χ_{x_1} *be the characteristic function of the set* $X_{x_1} = X \setminus \mathcal{D}_{\log x_1}(\Gamma)$. *Then, for every* $f \in L^2_{\mathrm{ac}}(X) = L^2(X)$,

$$\mathcal{P}^+_{x_1}(\mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}) = \mathcal{R}_+(0, \chi_{x_1} f)|_{\mathbb{R}\times\Gamma}.$$

Proof. Since $\mathcal{P}^+_{x_1}$ is a projector, there exists $f_{x_1} \in L^2(X)$ such that

$$\mathcal{P}^+_{x_1}(\mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}) = \mathcal{R}_+(0, f_{x_1})|_{\mathbb{R}\times\Gamma}$$

and, for every $h \in L^2(X)$ supported in X_{x_1},

$$\big\langle \mathcal{R}_+(0, f_{x_1})|_{\mathbb{R}\times\Gamma}, \mathcal{R}_+(0, h)|_{\mathbb{R}\times\Gamma} \big\rangle_{\mathbb{C}^n_+} = \langle f_{x_1}, h\rangle_{L^2(X)} = \langle f, h\rangle_{L^2(X)}.$$

Hence $f_{x_1} = \chi_{x_1} f$. □

Next we will analyze the singularities $\mathcal{R}_+(0, \chi_{x_1} f)$ at $\{s = \log x_1\}$ and, as in the proof of Proposition 3.2, we may assume that f is C^∞. In the case where $\Gamma = \partial X$, χ_{x_1} is the characteristic function of the set $\{x \geq x_1\}$ and the singularities of $\mathcal{R}_+(0, \chi_{x_1} f)$ can be computed using the plane wave expansion of the solution to the Cauchy problem

$$PV = 0, \quad V|_{s=\log x} = 0 \quad \text{and} \quad \partial_s V|_{s=\log x} = f(x, y)\chi_{x_1}, \tag{5-6}$$

where P is the operator defined in (3-2). In this case, one just writes

$$V(x, s, y) = V^+(x, s, y) + V^-(x, s, y), \quad \text{where}$$
$$V^+(x, s, y) \sim \sum_{j=1}^{\infty} v_j^+(x, y)(s - \log x_1)_+^j \quad \text{and} \quad V^-(x, s, y) \sim \sum_{j=1}^{\infty} v_j^-(x, y)(2\log x - x_1 - s)_+^j,$$

where $s = \log x_1$ and $s = 2\log x + \log x_1$ correspond to the forward and backward waves emanating from $\{x = x_1,\ s = \log x\}$. One then computes the coefficients of the expansion by using a series of transport equations. The wave $V^-(x, s, y)$ goes towards the interior and will hit $\{x = 0\}$ for $s > \log x_1$, but the wave $V^+(x, s, y)$ will intersect $\{x = 0\}$ at $s = \log x_1$. The first coefficient in the expansion of $V^+(x, s, y)$ is given by $v_1^+(x, y) = \frac{1}{2}(|h|^{1/4}(x_1, y)/|h|^{1/4}(x, y))x_1^{-n/2-1} f(x_1, y)$. Since (3-16) is well defined for L^2_{ac} initial data, $\mathcal{R}_+(0, \chi_{x_1} f) = \partial_s V(x, s, y)|_{\{x=0\}}$, and hence near $\{s = \log x_1\}$ one has an expansion

$$\mathcal{R}_+(0, \chi_{x_1} f) \sim \frac{1}{2}\frac{|h|^{1/4}(x_1, y)}{|h|^{1/4}(0, y)} x_1^{-n/2-1} f(x_1, y)(s - \log x_1)_+^0 + \sum_{j=1}^{\infty} v_j(0, y)^+(s - \log x_1)_+^j. \tag{5-7}$$

We refer the reader to the proof of Lemma 8.9 of [Sá Barreto 2005] for details.

In the case studied here, when $\Gamma \neq \partial X$, this is not so clear since χ_{x_1} is the characteristic function of $X_{x_1} = X \backslash \mathcal{D}_{\log x_1}(\Gamma)$, which is a more complicated set. However, if x_1 is small enough, the boundary of X_{x_1} contains $\Gamma_{x_1} = \{(x_1, y) : y \in \Gamma\}$. We will show that the singularities of $\mathcal{R}_+(0, \chi_{x_1} f)$ at $\{s = \log x_1,\ y \in \Gamma\}$ can be computed as in the previous case. The singularities of $\chi_{x_1} f$ lie on the set

$$\partial \mathcal{D}_{\log x_1} = \{z \in \mathring{X} : \text{there exists } (\bar{x}, \bar{y}) \in U_{\bar{\varepsilon}} \text{ such that } d_g(z, (\bar{x}, \bar{y})) = \log x_1 - \log \bar{x}\}$$

Since $\mathring{X}$ is complete, there exists a geodesic γ joining $z \in \partial \mathcal{D}_{\log x_1}$ and $(\bar{x}, \bar{y})$ such that

$$\gamma(0) = z, \quad \gamma(\bar{t}) = (\bar{x}, \bar{y}) \quad \text{and} \quad \bar{t} = d_g(z, (\bar{x}, \bar{y})).$$

One can think of this in terms of the wave equation with γ being the projection of a null bicharacteristic of $p = \frac{1}{2}(\tau^2 - x^2\xi^2 - x^2 h(x, y, \eta))$ in $\{p = 0,\ \tau = 1\}$ starting at z and going to $(\bar{x}, \bar{y})$. If one then sets $s = t + \log x$ it follows that, along this bicharacteristic, $s = t + \log x(\gamma(t))$. Hence, at $\hat{t}$, $s(\bar{t}) = \log x_1$. In these coordinates (we are using ξ by abuse of notation but we should use $\tilde{\xi}$, where $\tilde{\xi} = \xi - \tau/x$),

$$\{p = 0,\ \tau = 1\} = \left\{p = \sigma\xi + \tfrac{1}{2}x\xi^2 + \tfrac{1}{2}xh(x, y, \eta) = 0,\ \sigma = 1\right\}$$

and we have that, for $1+x\xi\neq 0$,

$$\frac{ds}{dx}=\frac{\xi}{1+x\xi},\quad \frac{d\xi}{dx}=-\frac{\xi^2+h+x\partial_x h}{2(1+x\xi)},\quad \frac{d\eta}{dx}=-\frac{x\partial_y h}{2(1+x\xi)},\quad \frac{dy}{dx}=-\frac{x\partial_\eta h}{2(1+x\xi)}.$$

So, unless $\xi=\eta=0$, $ds/dx\neq 0$. But, if $\xi=\eta=0$ at a point then, by uniqueness, $\xi=\eta=0$ along the curve. In the latter case $s=\log x_1$, $y=\bar{y}\in\Gamma$ along the curve. If $\xi\neq 0$, the geodesic will reach $\{x=0\}$ for $s\neq\log x_1$. So we conclude that (5-7) holds for $y\in\Gamma_1$, where $\overline{\Gamma}_1$ is a compact subset of Γ. The precise propagation of singularities is given by:

Lemma 5.3. *Let x be a defining function of ∂X such that* (2-4) *holds. Let $\mathcal{M}^+(\Gamma)\ni F=\mathcal{R}_+(0,f)|_{\mathbb{R}\times\Gamma}$ with f smooth. Let $\Theta(x_1,s,y)=\frac{1}{2}x_1^{-n/2-1}f(x_1,y)(|h|^{1/4}(x_1,y)/|h|^{1/4}(0,y))(s-\log x_1)_+^0$. There exists $\varepsilon>0$ such that, for any $x_1\in(0,\varepsilon)$,*

$$\mathcal{P}^+_{x_1}F(s,y)-\Theta(x_1,s,y)\in H^1_{\mathrm{loc}}(\mathbb{R}\times\Gamma). \tag{5-8}$$

Since $\mathcal{P}^+_{x_1}$ and $\mathcal{M}^+(\Gamma)$ are determined by $\mathcal{S}_\Gamma$ in view of (5-8), $\Theta(x_1,s,y)$ is determined by $\mathcal{S}_\Gamma$ provided $x_1\in(0,\varepsilon)$ and $y\in\Gamma$. By assumption in Theorem 2.3, $h_{0,1}=h_{0,2}$ on Γ. Therefore, $|h_1|(0,y)=|h_2|(0,y)$, $y\in\Gamma$ and, since $F=\mathcal{R}_+(0,f)|_{\mathbb{R}\times\Gamma}$ in Lemma 5.3, we obtain the following result:

Corollary 5.4. *Let (X_1,g_1) and (X_2,g_2) be asymptotically hyperbolic manifolds satisfying the hypothesis of Theorem 2.3. Moreover, assume that Δ_{g_j}, $j=1,2$, have no eigenvalues. Let $\mathcal{R}_{j,\pm}$, $j=1,2$, denote the corresponding forward or backward radiation fields defined in coordinates in which* (2-4) *holds. Then there exists an $\epsilon>0$ such that, for $(x,y)\in[0,\epsilon)\times\Gamma$,*

$$\begin{aligned}|h_1|^{1/4}(x,y)\mathcal{R}_{1,-}^{-1}F(x,y)&=|h_2|^{1/4}(x,y)\mathcal{R}_{2,-}^{-1}F(x,y)\quad \textit{for all } F\in\mathcal{M}^-(\Gamma),\\ |h_1|^{1/4}(x,y)\mathcal{R}_{1,+}^{-1}F(x,y)&=|h_2|^{1/4}(x,y)\mathcal{R}_{2,+}^{-1}F(x,y)\quad \textit{for all } F\in\mathcal{M}^+(\Gamma).\end{aligned} \tag{5-9}$$

Proposition 5.1 easily follows from this result. Indeed, since

$$\mathcal{R}_{j,-}^{-1}\left(\frac{\partial^2}{\partial s^2}F\right)=\left(\Delta_{g_j}-\tfrac{1}{4}n^2\right)\mathcal{R}_{j,-}^{-1}F, \tag{5-10}$$

if we apply Corollary 5.4 to ∂_s^2F we obtain

$$|h_1|^{1/4}(x,y)\left(\Delta_{g_1}-\tfrac{1}{4}n^2\right)\mathcal{R}_{1,-}^{-1}F(x,y)=|h_2|^{1/4}(x,y)\left(\Delta_{g_2}-\tfrac{1}{4}n^2\right)\mathcal{R}_{2,-}^{-1}F(x,y). \tag{5-11}$$

If $\mathcal{R}_{1,-}^{-1}F=(0,f)$, where $F\in\mathcal{M}(\Gamma)^-$ is arbitrary and the metrics have no eigenvalues, equations (5-9) and (5-11) give

$$|h_1|^{1/4}(x,y)\left(\Delta_{g_1}-\tfrac{1}{4}n^2\right)f(x,y)=|h_2|^{1/4}(x,y)\left(\Delta_{g_2}-\tfrac{1}{4}n^2\right)\frac{|h_1|^{1/4}(x,y)}{|h_2|^{1/4}(x,y)}f(x,y) \tag{5-12}$$

for all $f\in C_0^\infty((0,\epsilon)\times\Gamma)\cap L^2_{\mathrm{ac}}(X)$. Therefore the operators on both sides of (5-12) are equal. In particular, the coefficients of the principal parts of Δ_{g_1} are equal to those of Δ_{g_2}, and hence the tensors h_1 and h_2 from (2-4) are equal. This proves that

$$\begin{aligned}\mathcal{R}_{1,-}^{-1}(s,y,x',y')&=\mathcal{R}_{2,-}^{-1}(s,y,x',y'),\quad && y,y'\in\Gamma,\ x'\in[0,\varepsilon),\\ h_1(x,y,dy)&=h_2(x,y,dy), && y\in\Gamma,\ x\in[0,\varepsilon),\end{aligned} \tag{5-13}$$

and of course the same holds for the forward radiation field. Since $\mathcal{R}_\pm$ are unitary, $\mathcal{R}_\pm^{-1} = \mathcal{R}_\pm^*$, and hence this determines the kernel of $\mathcal{R}_\pm$. This proves Proposition 5.1 in the case of no eigenvalues.

Now we remove the assumption that there are no eigenvalues. We need to show that, if $\mathcal{S}_{1,\Gamma} = \mathcal{S}_{2,\Gamma}$, then the eigenvalues of Δ_{g_1} and Δ_{g_2} are equal, and the eigenfunctions can be reordered in such a way that their traces are equal on Γ. In fact they agree to infinite order at Γ. To show that, we need to appeal to the stationary version of scattering theory, and we have to recall the relationship between the scattering operator, the scattering matrix and the resolvent from [Sá Barreto 2005]. It was shown in [Joshi and Sá Barreto 2000] that $\mathcal{A}(\lambda)$, defined in (2-10), continues meromorphically to $\mathbb{C} \setminus D$, where D is a discrete set. The eigenvalues of Δ_g correspond to poles of $\mathcal{A}(\lambda)$ on the negative imaginary axis. Proposition 3.6 of [Graham and Zworski 2003] states that, if $\lambda_0 \in i\mathbb{R}_-$ is such that $\frac{1}{4}n^2 + \lambda_0^2$ is an eigenvalue of Δ_g, then the scattering matrix $\mathcal{A}(\lambda)$ has a pole at λ_0 and its residue is given by

$$\operatorname{Res}_{\lambda_0} A(\lambda) = \begin{cases} \Pi_{\lambda_0} & \text{if } -i\lambda_0 \notin \frac{1}{2}\mathbb{N}, \\ \Pi_{\lambda_0} - P_l & \text{if } -i\lambda_0 = \frac{1}{2}l,\ l \in \mathbb{N}, \end{cases} \tag{5-14}$$

where P_l is a differential operator whose coefficients depend on derivatives of the tensor h at ∂X, and the Schwartz kernel of Π_{λ_0} is

$$K(\Pi_{\lambda_0})(y, y') = -2i\lambda_0 \sum_{j=1}^{N_0} \phi_j^0 \otimes \phi_j^0(y, y'), \tag{5-15}$$

where N_0 is the multiplicity of the eigenvalue $\frac{1}{4}n^2 + \lambda_0^2$, the ϕ_j, $1 \le j \le N_0$, are the corresponding orthonormalized eigenfunctions and $\phi_j^0(y)$ is defined by

$$\phi_j^0(y) = x^{-n/2-\lambda_0}\phi_j(x, y)|_{x=0}. \tag{5-16}$$

Since $A_{1,\Gamma} = A_{2,\Gamma}$, $\lambda \in \mathbb{R} \setminus 0$, it follows from Theorem 1.2 of [Joshi and Sá Barreto 2000] that, in coordinates where (2-14) is satisfied, all derivatives of h_1 and h_2 agree at $x = 0$ on Γ. Therefore the operators $P_{l,j}$ in (5-14) corresponding to (X_j, g_j) are the same in Γ. Then (5-14), (5-15), and the meromorphic continuation of the scattering matrix show that Δ_{g_1} and Δ_{g_2} have the same eigenvalues with the same multiplicity. Moreover, (5-15) implies that if ϕ_j and ψ_j, $1 \le j \le N_0$, are orthonormal sets of eigenfunctions of Δ_{g_1} and Δ_{g_2}, respectively, corresponding to the eigenvalue $\frac{1}{4}n^2 + \lambda_0^2$, then there exists a constant orthogonal $(N_0 \times N_0)$-matrix A such that $\Phi^0\big|_\Gamma = A\Psi^0\big|_\Gamma$, where $(\Phi^0)^T = (\phi_1^0, \phi_2^0, \dots, \phi_{N_0}^0)$ and $(\Psi^0)^T = (\psi_1^0, \psi_2^0, \dots, \psi_{N_0}^0)$. So, by redefining one set of eigenfunctions from, let us say, Ψ to $A\Psi$, where $\Psi^T = (\psi_1, \psi_2, \dots, \psi_{N_0})$, we may assume that

$$\phi_j^0(y) = \psi_j^0(y), \quad y \in \Gamma, \quad j = 1, 2, \dots, N_0. \tag{5-17}$$

Note that this does not change the orthonormality of the eigenfunctions in X_2 because A is orthogonal. Denote the eigenvalues of Δ_{g_1} and Δ_{g_2}, which we know are equal, by

$$\mu_j = \tfrac{1}{4}n^2 + \lambda_j^2, \quad \lambda_j \in i\mathbb{R}_-, \quad 1 \le j \le N. \tag{5-18}$$

They are also ordered so that $\mu_1 \le \mu_2 \le \cdots \le \mu_N$.

Again, we use that the singularities of $\chi_{x_1} f$ at Γ_{x_1} produce the singularities of $\mathcal{R}_+(0, \chi_{x_1} f)$ at $\{s = \log x_1, y \in \Gamma\}$ and expand the solution to (2-1) with initial data, $(0, \chi_{x_1} f)$. However, in this case $L^2(X) \neq L^2_{\text{ac}}(X)$ and hence Lemma 5.2 is not valid, and we have to replace it by the following:

Lemma 5.5. *Let (X, g) be an asymptotic hyperbolic manifold and let ϕ_j, $1 \le j \le N$, denote the orthonormal set of eigenfunctions of Δ_g. Let x be such that* (2-4) *holds in $(0, \epsilon) \times \partial X$. For $x_1 \in (0, \epsilon)$, let $\mathcal{P}^+_{x_1}$ denote the orthogonal projector defined in* (5-5). *Let χ_{x_1} be the characteristic function of the set $X_{x_1} = X \setminus \mathcal{D}_{\log x_1}(\Gamma)$. There exists ε_0 such that, if $\varepsilon < \varepsilon_0$, then for every $f \in L^2_{\text{ac}}(X)$ there exists $\alpha(x_1, f)$, which is a linear function of f, such that*

$$\mathcal{P}^+_{x_1}(\mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}) = \mathcal{R}_+\biggl(0, \chi_{x_1}\biggl(f - \sum_{j=1}^{N} \alpha_j(x_1, f)\phi_j\biggr)\biggr)\bigg|_{\mathbb{R}\times\Gamma}.$$

Proof. Let $h \in L^2_{\text{ac}}(X)$ be supported in X_{x_1}. This means that $\langle h, \chi_{x_1}\phi_j\rangle = 0$ for $1 \le j \le N$. Then, since $\mathcal{P}^+_{x_1}$ is a projector, there exists $f_{x_1} \in L^2_{\text{ac}}(X)$, supported in X_{x_1}, such that $\mathcal{P}^+_{x_1}(\mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}) = \mathcal{R}_+(0, f_{x_1})|_{\mathbb{R}\times\Gamma}$ and, for every $h \in L^2_{\text{ac}}(X)$ supported in X_{x_1},

$$\bigl\langle \mathcal{R}_+(0, f_{x_1})|_{\mathbb{R}\times\Gamma}, \mathcal{R}_+(0, h)|_{\mathbb{R}\times\Gamma}\bigr\rangle_{\mathbb{C}^n_+} = \langle f_{x_1}, h\rangle_{L^2(X)} = \langle f, h\rangle_{L^2(X)}.$$

Hence $\langle (f_{x_1} - f), h\rangle = 0$ for all $h \in C_0^\infty(X) \cap L^2_{\text{ac}}(X)$ supported in X_{x_1}. We claim that there exist $\alpha_j = \alpha_j(x_1, f) \in \mathbb{C}$ such that

$$f_{x_1} - \chi_{x_1} f - \chi_{x_1} \sum_{j=1}^{N} \alpha_j \phi_j = 0 \quad \text{for } x_1 \text{ small enough.}$$

If such a formula were to hold, since $\langle f_{x_1}, \chi_{x_1}\phi_j\rangle = 0$ one would have to have

$$\langle f, \chi_{x_1}\phi_k\rangle_{L^2(X)} = \sum_{j=1}^{N} \alpha_j \langle \chi_{x_1}\phi_j, \chi_{x_1}\phi_k\rangle_{L^2(X)}.$$

This gives a linear system of equations

$$M\alpha = F, \quad \alpha^T = (\alpha_1, \dots, \alpha_N), \quad F^T = (F_1(x_1), \dots, F_N(x_1)),$$
$$M_{jk}(x_1) = \langle \chi_{x_1}\phi_j, \chi_{x_1}\phi_k\rangle_{L^2(X)}, \quad F_k(x_1) = \langle f, \chi_{x_1}\phi_k\rangle_{L^2(X)}.$$

Since the eigenfunctions are orthonormal, for $x_1 = 0$ we have $M_{jk}(0) = \delta_{jk}$. Therefore, there exists $\varepsilon_0 > 0$, which depends on the matrix M, and hence only on the eigenfunctions and not on f, such that the system has a solution if $x_1 < \varepsilon_0$. Notice that, since $f \in L^2_{\text{ac}}(X)$, for $x_1 = 0$ we have $F_k(0) = 0$, and hence $\alpha(0, f) = 0$.

With this choice of α_j, the function

$$G = f_{x_1} - \chi_{x_1} f - \chi_{x_1} \sum_{j=1}^{N} \alpha_j \phi_j$$

is supported in X_{x_1} and $\langle G, \phi_j\rangle_{L^2(X)} = 0$, so $G \in L^2_{\text{ac}}(X)$. But at the same time $\langle F, h\rangle_{L^2(X)} = 0$ for all $h \in L^2_{\text{ac}}(X)$ supported in X_{x_1}. Therefore $\langle G, G\rangle_{L^2(X)} = 0$, and so $G = 0$. ☐

As in [Sá Barreto 2005], we shall denote

$$T(x_1)f = \sum_j \alpha_j(x_1, f)\phi_j.$$

Since $\alpha(0, f) = 0$, $T(0) = 0$. Therefore one can pick ε small so that

$$\|T(x_1)\| < \tfrac{1}{2} \quad \text{for } x_1 < \varepsilon. \tag{5-19}$$

In this case, Lemma 5.3 and Corollary 5.4 have to be substituted by:

Lemma 5.6. *Let (X, g) be an asymptotically hyperbolic manifold, and let x be a defining function of ∂X such that (2-4) holds. Let ϕ_j, $1 \le j \le N$, denote the eigenfunctions of Δ_g and let $T(x_1)$ be defined as above. Let $F \in \mathcal{M}^+(\Gamma)$, $F = \mathcal{R}_+(0, f)|_{\mathbb{R}\times\Gamma}$ with f smooth and let*

$$\Xi(x_1, s, y) = \tfrac{1}{2} x_1^{-n/2-1} \frac{|h|^{1/4}(x_1, y)}{|h|^{1/4}(0, y)} \big[(\mathrm{Id} - T(x_1)) f\big](x_1, y)(s - \log x_1)_+^0.$$

There exists $\varepsilon > 0$ such that, for any $x_1 \in (0, \varepsilon)$,

$$\mathcal{P}^+_{x_1} F(s, y) - \Xi(x_1, s, y) \in H^1_{\mathrm{loc}}(\mathbb{R} \times \Gamma). \tag{5-20}$$

Corollary 5.7. *Let (X_1, g_1) and (X_2, g_2) be asymptotically hyperbolic manifolds satisfying the hypothesis of Theorem 2.3. Let $\mathcal{R}_{j,\pm}$, $j = 1, 2$, denote the corresponding forward or backward radiation fields defined in coordinates in which (2-4) holds. Then there exists an $\varepsilon > 0$ such that, for $(x, y) \in (0, \varepsilon) \times \Gamma$,*

$$\begin{aligned} |h_1|^{1/4}(x, y)(\mathrm{Id} - T_1(x))\mathcal{R}^{-1}_{1,-} F(x, y) &= |h_2|^{1/4}(x, y)(\mathrm{Id} - T_2(x))\mathcal{R}^{-1}_{2,-} F(x, y) \quad \textit{for all } F \in \mathcal{M}^-(\Gamma), \\ |h_1|^{1/4}(x, y)(\mathrm{Id} - T_1(x))\mathcal{R}^{-1}_{1,+} F(x, y) &= |h_2|^{1/4}(x, y)(\mathrm{Id} - T_2(x))\mathcal{R}^{-1}_{2,+} F(x, y) \quad \textit{for all } F \in \mathcal{M}^+(\Gamma). \end{aligned} \tag{5-21}$$

We write $\mathcal{R}^{-1}_{j,-} F(x, y) = f_j(x, y)$, and pick ε small so that (5-19) holds. We apply (5-21) to f_1 and f_2 and to $\big(\Delta_{g_1} - \frac{1}{4}n^2\big) f_1$ and $\big(\Delta_{g_2} - \frac{1}{4}n^2\big) f_2$ for $(x, y) \in [0, \varepsilon) \times \Gamma$ and find that

$$\begin{aligned} |h_1(x)|^{1/4}(\mathrm{Id} - T_1(x)) f_1 &= |h_2(x)|^{1/4}(\mathrm{Id} - T_2(x)) f_2, \\ |h_1(x)|^{1/4}(\mathrm{Id} - T_1(x))\big(\Delta_{g_1} - \tfrac{1}{4}n^2\big) f_1(x, y) &= |h_2(x)|^{1/4}(\mathrm{Id} - T_2(x))\big(\Delta_{g_2} - \tfrac{1}{4}n^2\big) f_2(x, y). \end{aligned} \tag{5-22}$$

Therefore,

$$f_2(x, y) = (\mathrm{Id} - T_2(x))^{-1} \frac{|h_1|^{1/4}}{|h_2|^{1/4}} (\mathrm{Id} - T_1(x)) f_1(x, y) = \frac{|h_1|^{1/4}}{|h_2|^{1/4}} f_1(x, y) + K(x) f_1(x, y),$$

where K is a compact operator. If one substitutes this into the second equation in (5-22), one obtains

$$|h_1|^{1/4}(\mathrm{Id} - T_1)\big(\Delta_{g_1} - \tfrac{1}{4}n^2\big) f_1 = |h_2|^{1/4}(\mathrm{Id} - T_2)\big(\Delta_{g_2} - \tfrac{1}{4}n^2\big)\bigg(\frac{|h_1|^{1/4}}{|h_2|^{1/4}} f_1 + K f_1\bigg)$$

Hence,

$$\big(\Delta_{g_1} - \tfrac{1}{4}n^2\big) f_1(x, y) - \frac{|h_2|^{1/4}}{|h_1|^{1/4}} \big(\Delta_{g_2} - \tfrac{1}{4}n^2\big)\bigg(\frac{|h_1|^{1/4}}{|h_2|^{1/4}} f_1\bigg)(x, y) = (\mathcal{K} f_1)(x, y),$$

where $\mathcal{K}$ is a compact operator. Since the operator on the left-hand side is a differential operator, and the operator on the right-hand side is compact, they both must be equal to zero. As above, we conclude that in coordinates (x, y), the coefficients of the operators Δ_{g_1} are equal to those of Δ_{g_2}. Hence, we must have $h_1(x, y, dy) = h_2(x, y, dy)$.

We still have to show that (5-4) holds in the case where eigenvalues exist. Let $F \in \mathcal{M}^+(\Gamma)$, and let $f_j = \mathcal{R}_{j,+}^{-1} F$. Let v_j satisfy (2-1) with initial data $(0, f_j)$. Let $V_j(x, s, y) = x^{-n/2} v_j(s - \log x, x, y)$. Since $\mathcal{R}_+(0, f_j) = F$, we have $\partial_s V_j(0, s, y) = F$. Since $\Delta_{g_1} = \Delta_{g_2}$ in $(0, \varepsilon) \times \Gamma$, for P as defined in (3-2),

$$\begin{gathered} P(V_1 - V_2) = 0 \quad \text{in } \log x < s,\ x < \varepsilon,\ y \in \Gamma \\ (V_1 - V_2)(x, \log x, y) = 0, \quad \partial_s (V_1 - V_2)(x, \log x, y) = f_1(x, y) - f_2(x, y) \text{ on } x < \varepsilon,\ y \in \Gamma, \\ \partial_s (V_1 - V_2)(0, s, y) = 0, \quad y \in \Gamma,\ s \in \mathbb{R}. \end{gathered} \tag{5-23}$$

Now we apply Propositions 3.2, 3.3 and 3.4 as in the proof of Theorem 2.1, to conclude that there exists s^* such that

$$V_1(x, s, y) = V_2(x, s, y) \quad \text{provided } x < e^{s^*},\ y \in \Gamma,\ s \in \mathbb{R}.$$

We then apply Tataru's theorem, as in the argument used in the final step of the proof of Theorem 2.1, to conclude that $f_1(z) - f_2(z) = 0$ for every $z \in (0, \varepsilon) \times \Gamma$ such that there exists $(x, y) \in (0, e^{s^*}) \times \Gamma$ with $d(z, (x, y)) < e^s / x$. In particular this shows that $f_1 = f_2$ in $(0, \varepsilon) \times \Gamma$. One cannot say that $f_1 = f_2$ on X since (5-23) only holds on $(0, \varepsilon) \times \Gamma$. Since F is arbitrary, (5-4) follows. □

Since $h_1(x) = h_2(x)$ on $[0, \varepsilon) \times \Gamma$, this finishes the construction of the map Ψ_ε defined in (5-3). We will use both equalities in (5-4) to extend Ψ_ε to a global diffeomorphism $\Psi : X_1 \to X_2$ satisfying (2-15).

The construction of the global diffeomorphism. First we need to show that if the eigenfunctions are reordered such that (5-17) holds, then in fact $\phi_{j,1}(x, y) = \phi_{j,2}(x, y)$ on $(0, \varepsilon) \times \Gamma$. To prove this we have to appeal again to the stationary scattering theory. We know from [Joshi and Sá Barreto 2000] that the operator

$$E_+(\lambda)\psi(\lambda, y) = \widehat{\mathcal{R}_+(0, \psi)}(\lambda, y) = \int_{\mathbb{R}} e^{-i\lambda s} \mathcal{R}_+(0, f)(s, y)\, ds,$$

continues meromorphically to $\mathbb{C} \setminus D$, where D is a discrete subset. Since their Schwartz kernels satisfy $E_1(\lambda, y', x, y) = E_2(\lambda, y', x, y)$ for $x \in [0, \varepsilon)$ and $y, y' \in \Gamma$, $\lambda \in \mathbb{R}$, this equality must remain for $\mathbb{C} \setminus D$.

We also know from equation (3.15) of [Graham and Zworski 2003] that $\frac{1}{4}n^2 + \lambda_0^2$ is an eigenvalue of Δ_g if and only if $\lambda_0 \in i\mathbb{R}_-$ is a pole of $E(\lambda, y, z)$, with the same multiplicity, and its residue is given by

$$\frac{1}{2i\lambda_0} \sum_{k=1}^{K} \phi_k^0(y) \phi_k(z), \quad y \in \partial X,\ z \in X, \tag{5-24}$$

where $\phi_k^0(y)$ is defined in (5-16) and K is the multiplicity of the eigenvalue. We know from (5-17) and (5-18) that the eigenvalues and the traces of the eigenfunctions are equal. So if $\phi_k^{(j)}(x', y')$ $j = 1, 2$,

$1 \le k \le K$, denote the eigenfunctions, we must have

$$\sum_{k=1}^{K} (\phi_k^{(1)}(x', y') - \phi_k^{(2)}(x', y'))\phi_k^0(y) = 0, \quad x' \in [0, \varepsilon),\ y, y' \in \Gamma.$$

Since the points (x', y'), $x' \in [0, \varepsilon)$ and $y, y' \in \Gamma$ are arbitrary and can be independently chosen, we must have

$$\phi_k^{(1)}(x', y') = \phi_k^{(2)}(x', y') \quad \text{for all } x' \in [0, \varepsilon),\ y' \in \Gamma. \tag{5-25}$$

We know that the Schwartz kernels of the radiation fields $\mathcal{R}_{j,\pm}$, $j = 1, 2$, acting on data $(0, f)$, and the metric tensors $h_j(x, y, dy)$, $j = 1, 2$, satisfy (5-4). However, if $\phi \in C_0^\infty((0, \varepsilon) \times \Gamma)$ and $(\phi, 0) \in E_{\mathrm{ac}}(X_j)$, then

$$\partial_s \mathcal{R}_{j,\pm}(\phi, 0)(s, y) = \mathcal{R}_{j,\pm}\big(0, \big(\Delta_{g_j} - \tfrac{1}{4}n^2\big)\phi\big)(s, y).$$

Since ϕ is compactly supported, $\mathcal{R}_+\big(0, \big(\Delta_{g_j} - \tfrac{1}{4}n^2\big)\phi\big)(s, y) = 0$ for $s \ll 0$. So,

$$\mathcal{R}_{j,+}(\phi, \psi) = \mathcal{R}_{j,+}(0, \psi) + \int_{-\infty}^{s} \mathcal{R}_{j,+}\big(0, \big(\Delta_{g_j} - \tfrac{1}{4}n^2\big)\phi\big)(\tau, y)\, d\tau,$$
$$\mathcal{R}_{j,-}(\phi, \psi) = \mathcal{R}_{j,-}(0, \psi) + \int_{s}^{\infty} \mathcal{R}_{j,-}\big(0, \big(\Delta_{g_j} - \tfrac{1}{4}n^2\big)\phi\big)(\tau, y)\, d\tau,$$

provided $(\phi, \psi) \in (C_0^\infty((0, \varepsilon) \times \Gamma) \times C_0^\infty((0, \varepsilon) \times \Gamma)) \cap E_{\mathrm{ac}}(X_j)$. Since we know from (5-13) that $\Delta_{g_1} = \Delta_{g_2}$ on $[0, \varepsilon) \times \Gamma$, and we also know from (5-25) that

$$\begin{aligned}\mathcal{A}((0, \varepsilon) \times \Gamma) &\doteq \big(C_0^\infty((0, \varepsilon) \times \Gamma) \times C_0^\infty((0, \varepsilon) \times \Gamma)\big) \cap E_{\mathrm{ac}}(X_1) \\ &= \big(C_0^\infty((0, \varepsilon) \times \Gamma) \times C_0^\infty((0, \varepsilon) \times \Gamma)\big) \cap E_{\mathrm{ac}}(X_2),\end{aligned}$$

we deduce that

$$\mathcal{R}_{1,\pm}(\phi, \psi)(s, y) = \mathcal{R}_{2,\pm}(\phi, \psi)(s, y), \quad (s, y) \in \mathbb{R} \times \Gamma,\ (\phi, \psi) \in \mathcal{A}((0, \varepsilon) \times \Gamma). \tag{5-26}$$

But $\mathcal{R}_\pm$ are unitary operators, and so their inverses are equal to their adjoints, and we deduce from (5-26) that the Schwartz kernels of the full operators $\mathcal{R}_{j,\pm}$ acting on $\mathcal{A}((0, \varepsilon) \times \Gamma)$ are determined by the scattering operator $\mathcal{S}_\Gamma$. We conclude that if $F \in L^2(\mathbb{R} \times \Gamma)$, and if $\mathcal{R}_{j,\pm}^{-1}\big|_\Gamma : L^2(\mathbb{R} \times \Gamma) \to E_{\mathrm{ac}}(X_1)|_{(0,\varepsilon)\times\Gamma}$, $j = 1, 2$, is given by

$$F(s, y) \mapsto (\phi_j, \psi_j) = (u_j(0), \partial_t u_j(0))\big|_{(0,\varepsilon)\times\Gamma},$$

then $(\phi_1, \psi_1) = (\phi_2, \psi_2)$. Here $u_j(t, z)$ denotes the solution to the Cauchy problems for the wave equation (2-1) for the metric g_j. But, on the other hand, $\mathcal{R}_{j,\pm}$ are translation representations of the wave group, and therefore

$$\mathcal{R}_{j,+}^{-1}\big|_\Gamma F(s + t) = (u_j(t), \partial_t u(t)),$$

where $u_j(t)$ satisfies (2-1) with initial data $(\phi, \psi) = \mathcal{R}_{j,\pm}^{-1}\big|_\Gamma \in \mathcal{A}((0, \varepsilon) \times \Gamma)$. We conclude that, if $u_j(t, z)$ solves (2-1) for the metric g_j, with initial data supported in $(0, \varepsilon) \times \Gamma$, then $u_1(t, z) = u_2(t, z)$, provided $z \in (0, \varepsilon) \times \Gamma$. This implies that, if $U_j(t, z, z')$ is the forward fundamental solution of the Cauchy problem for the wave equation in (X_j, g_j), then

$$U_1(t, z, z') = U_2(t, z, z'), \quad z,\ z' \in (0, \varepsilon) \times \Gamma,\ t > 0. \tag{5-27}$$

By Duhamel's principle, if

$$\begin{gathered}\big(D_t^2 - \Delta_{g_j} - \tfrac{1}{4}n^2\big)\tilde{U}_j(t, t', z, z') = \delta(x, y)\delta(t - t') \quad \text{in } X_j \times \mathbb{R}, \\ \tilde{U}_j(0) = \partial_t \tilde{U}_j(0) = 0,\end{gathered} \tag{5-28}$$

then

$$\tilde{U}_1(t, t'z, z') = \tilde{U}_2(t, t', z, z'), \quad t,\ t' \in \mathbb{R}_+,\ z,\ z' \in (0, \varepsilon) \times \Gamma. \tag{5-29}$$

So we have reduced the extension of the diffeomorphism to the following:

Proposition 5.8. *Let (X_1, g_1) and (X_2, g_2) be AHM such that*:

(A) *There exists a nonempty open subset $\Gamma \subset \partial X_1 \cap \partial X_2$ as manifolds and an open subset $\mathbb{O} \sim \Gamma \times (0, \varepsilon)$ such that $\mathbb{O} \subset \mathring{X}_1 \cap \mathring{X}_2$ as manifolds.*

(B) *The metric tensors g_j, $j = 1, 2$, satisfy $g_1 = g_2$ on $\mathbb{O}$.*

(C) *If $\tilde{U}_j(t, t', z, z')$, $j = 1, 2$ is the forward fundamental solution of the wave equation in (X_j, g_j), $j = 1, 2$, defined in* (5-28), *then $U_1(t, t', z, z') = U_2(t, t', z, z')$ for $t, t' \in \mathbb{R}_+$ and $z, z' \in \mathbb{O}$.*

Then there exists

$$\Psi : X_1 \to X_2 \quad \textit{such that} \quad \Psi^* g_2 = g_1 \ \textit{and} \ \Psi = \mathrm{Id} \ \textit{in} \ \mathbb{O}. \tag{5-30}$$

This is similar to the inverse boundary value problem with data on part of the boundary, studied for example in [Katchalov et al. 2001; Kurylev and Lassas 2000], except that we are not dealing with boundary control but control from an open set in the interior. A somewhat similar problem for closed manifolds was studied in [Krupchyk et al. 2008]. Lassas and Oksanen [2014] also dealt with a problem of this nature. This is also related to the problem studied by Lassas, Taylor and Uhlmann on complete real analytic manifolds without boundary M_j, $j = 1, 2$, where the Green functions for the Laplace operator agree on $U \times U$, with $U \subset M_1 \cap M_2$; see Theorem 4.1 of [Lassas et al. 2003]. The difference here is that we do not have real analyticity of the manifolds, but we are dealing with the wave equation instead of the Laplace equation.

Proof. We adapt the proof of Theorem 4.33 in [Katchalov et al. 2001]. Instead of working with X_1 and X_2, we will fix $X = X_1$ and reconstruct $(X, g) = (X_1, g_1)$ from (A), (B) and (C). Of course, we are reconstructing (X_2, g_2) as well. First of all, we observe that an AHM has a uniform radius of injectivity for the geodesic flow. In other words, there exists a $\rho_0 > 0$ such that, if $S_p X = \{v \in T_p X : \|v\|_g = 1\}$, the map

$$\exp_p : [0, \rho_0) \times S_p X \to X, \quad (t, v) \mapsto \exp_p(tv),$$

is well defined for all $p \in X$. We pick a point $p \in \mathbb{O}$ and let $\rho \in (0, \rho_0)$ be such that the geodesic ball $B(p, \rho) \subset \mathbb{O}$. Let $f(t, z) \in C_0^\infty(\mathbb{R} \times B(p, \rho))$, $f(t, z) = 0$ for $t < 0$, and let $u^f(t, z)$ be the solution to

$$\begin{gathered} \big(D_t^2 - \Delta_g - \tfrac{1}{4}n^2\big)u^f(t, z) = f(t, z) \quad \text{in } \mathbb{R} \times X, \\ u^f(0) = \partial_t u^f(0) = 0. \end{gathered} \tag{5-31}$$

From the hypothesis (C) above, we know $u^f(t, z)$ for $z \in B(p, \rho)$, $t > 0$. We then define the map

$$\mathcal{B}(T) : C_0^\infty((0, T) \times B(p, \rho)) \to C^\infty((0, T) \times B(p, \rho)), \quad f \mapsto u^f\big|_{(0,T)\times B(p,\rho)}. \tag{5-32}$$

For $T > 0$ we will work with the space of functions

$$\mathscr{C}_0 = \mathscr{C}_0(p, \rho, T) \doteq \{\phi \in C_0^\infty((0, T] \times B(p, \rho)) : \phi(T) = 0\},$$

and the quotient space

$$\mathscr{C} = \mathscr{C}(p, \rho, T) \doteq \mathscr{C}_0 \big/ \big(D_t^2 - \Delta_g - \tfrac{1}{4}n^2\big)\mathscr{C}_0.$$

In other words,

$$\mathscr{C} = \{[\psi] : \psi \in \mathscr{C}_0\}, \quad \text{where } [\psi] = \big\{\phi \in \mathscr{C}_0 : \text{there is } \zeta \in \mathscr{C}_0 \text{ such that } \phi = \psi + \big(D_t^2 - \Delta_g - \tfrac{1}{4}n^2\big)\zeta\big\}.$$

Since we know g in $\mathbb{O}$, the space $\mathscr{C}$ is determined by hypotheses (A), (B) and (C).

For $\phi \in \mathscr{C}$, let u^ϕ be the solution to (5-31) in $\mathbb{R} \times X$. We define the map

$$C_T : \mathscr{C} \to C_0^\infty(X), \quad \phi \mapsto u^\phi(T, z).$$

The formal adjoint of this map is given by

$$C_T^* : \{w \in C_0^\infty(\{z \in X : d_g(z, B(p, \rho)) < T\})\} \to \mathscr{C}, \quad w \mapsto v|_{(0,T)\times B(p,\rho)},$$

where v is the solution to the Cauchy problem

$$\begin{gathered} \big(D_t^2 - \Delta_g - \tfrac{1}{4}n^2\big)v(t, z) = 0 \quad \text{in } \{t < T\} \times X, \\ v(T, z) = 0, \quad \partial_t v(T, z) = w. \end{gathered} \tag{5-33}$$

As in the boundary control method, we define

$$S_T = C_T^* C_T : \mathscr{C} \to \mathscr{C}.$$

The next step is to prove a Blagovestchenskii-type identity to show that S_T is determined by the map $\mathcal{B}(2T)$, which the map defined in (5-32) but in the time interval $(0, 2T)$, and hence is determined from (A), (B) and (C). Let $\phi(t, z)$, $\psi(t, z) \in \mathscr{C}$ and let $u^\phi(t, z)$, $u^\psi(t, z)$ be the solutions to (5-31), with left-hand side ϕ and ψ respectively. Let

$$W(s, t) = \int_X u^\phi(t, z)u^\psi(s, z)\, d\,\mathrm{vol}_g(z).$$

Notice that this integration is defined over the entire manifold. But, after integrating by parts, we obtain

$$\begin{aligned}(\partial_t^2-\partial_s^2)W(s,t)&=\int_X(\phi(t,z)u^{\psi}(s,z)-u^{\phi}(t,z)\psi(s,z))\,d\,\mathrm{vol}_g(z)\\&=\int_X[\phi(t,z)\mathcal{B}(T)\psi(s,z)-\psi(s,z)\mathcal{B}(T)\phi(t,z)]\,d\,\mathrm{vol}_g(z),\\W(0,t)=\partial_sW(0,t)&=0,\qquad W(s,0)=\partial_tW(s,0)=0,\end{aligned}$$

and, since ϕ and ψ are supported in $(0,T)\times B(p,\rho)$, the last integration is restricted to $B(p,\rho)$. We can find $W(T,T)$ explicitly in terms of d'Alembert's formula, but we need to extend ϕ and ψ to the interval $(0,2T)$. As in [Belishev and Kurylev 1992], we define $\tilde{\phi}$ and $\tilde{\psi}$ to be the odd extensions of ϕ and ψ across $t=T$, in other words

$$\tilde{\phi}(t)=\begin{cases}\phi(t) & \text{if } t\in(0,T),\\ -\phi(2T-t) & \text{if } t\in(T,2T),\end{cases}$$

and similarly for $\tilde{\psi}$. This gives

$$W(T,T)=\int_0^T\int_t^{2T-t}\biggl(\int_X\bigl(\tilde{\phi}(t,z)\mathcal{B}(2T)(\tilde{\psi})(s,z)-\mathcal{B}(2T)(\tilde{\phi})(t,z)\tilde{\psi}(s,z)\bigr)\,d\,\mathrm{vol}_g(z)\biggr)\,ds\,dt$$

Since $\tilde{\psi}(s,z)$ is odd with respect to $s=T$, it follows that

$$\begin{aligned}W_j(T,T)&=\int_0^T\int_t^{2T-t}\int\tilde{\phi}(t,z)\mathcal{B}(2T)(\tilde{\psi})(s,z)\,d\,\mathrm{vol}_g(z)\,ds\,dt\\&=\int_0^T\int_X\phi(t,z)\biggl(\int_t^{2T-t}\mathcal{B}(2T)\tilde{\psi}(s,z)\,ds\biggr)\,d\,\mathrm{vol}_g(z)\,dt.\end{aligned}$$

On the other hand, since

$$W(T,T)=\langle C_T\phi,C_T\psi\rangle=\langle\phi,C_T^*C_T\psi\rangle,$$

it follows that

$$C_T^*C_T\psi(t,z)=\int_t^{2T-t}\mathcal{B}(2T)\tilde{\psi}(s,z)\,ds.$$

Now we define the following inner product in the space $\mathcal{C}$:

$$\langle\phi,\psi\rangle_{\mathcal{C}}=\langle u^{\phi}(T,z),u^{\psi}(T,z)\rangle_{L^2(X)}.$$

As shown above, this is determined by the map $\mathcal{B}$. We need to show that this is a nondegenerate inner product. First we show that the range $\{u^{\phi}(T):\phi\in\mathcal{C}\}$ is dense in the space

$$L^2\bigl(\{z\in X_j:d(z,B(p,\rho))\le T\}\bigr)=\bigl\{u\in L^2(X_j):\mathrm{Supp}(u)\subset\{z:d(z,B(p,\rho))\le T\}\bigr\}.$$

Suppose that $w\in L^2(\{z\in X_j:d(z,B(p,\rho))\le T\})$ is such that

$$\langle w,u^{\phi}(T)\rangle=0\quad\text{for all }\phi\in\mathcal{C}.$$

Let v satisfy (5-33) and let u^ϕ satisfy (5-31) with right-hand side equal to ϕ. Integrating the identity

$$v\big(D_t^2-\Delta_g-\tfrac{1}{4}n^2\big)u^\phi-u^\phi\big(D_t^2-\Delta_g-\tfrac{1}{4}n^2\big)v=v(t,z)\phi(t,z)$$

in the domain of influence of ϕ and w, we find that

$$\int_{B(p,\rho)\times(0,T)} v(t,z)\phi(t,z)\,dt\,d\,\mathrm{vol}_g(z)=0 \quad \text{for all } \phi\in\mathscr{C}. \tag{5-34}$$

But, again using the fact that v satisfies (5-33), we see that

$$\int_{B(p,\rho)\times(0,T)} v(t,z)\big(D_t^2-\Delta_g-\tfrac{1}{4}n^2\big)\phi(t,z)\,dt\,d\,\mathrm{vol}_g(z)=0 \quad \text{for all } \phi\in\mathscr{C}_0.$$

This means that (5-34) is satisfied for every $\phi\in\mathscr{C}_0$, and hence $v(t,z)=0$ in $(0,T)\times B(p,\rho)$. Now the odd extension $\tilde{v}(t,z)$ of $v(t,z)$ across $t=T$ satisfies (5-33) in $(0,2T)\times\{z:d(z,B(p,\rho))<T+\rho\}$ and $\tilde{v}(t,z)=0$ in $(0,2T)\times B(p,\rho)$. An application of Tataru's theorem implies that $\tilde{v}(t,z)=0$ if $|t|+d(z,B(p,\rho))\le T$ for any $q\in B(p,\rho)$. In particular, this implies that $w(z)=\partial_t v(T,z)=0$ provided $d(z,B(p,\rho))\le T$, and hence $w=0$.

Now suppose that $\phi\in\mathscr{C}$ is such that $\langle\phi,\psi\rangle_{\mathscr{C}}=0$ for every $\psi\in\mathscr{C}$. From the previous discussion, it follows that $u^\phi(T)=0$. Then

$$\tilde{u}(t,z)=\begin{cases} u^\phi(t,z) & \text{if } t<T,\\ -u^\phi(2T-t,z) & \text{if } t>T\end{cases}$$

satisfies

$$\begin{aligned}\big(D_t^2-\Delta_g-\tfrac{1}{4}n^2\big)\tilde{u}&=\tilde{\phi} \quad \text{in } \mathbb{R}\times X_j\\ \tilde{u}&=0 \quad \text{in } \mathbb{R}\times\{z:d(z,B(p,\rho))>T\}.\end{aligned}$$

Again, Tataru's theorem and finite speed of propagation implies that $u^\phi\in C_0^\infty((0,T]\times B(p,\rho))$ and $u^\phi(T)=0$. This of course means that $u^\phi\in\mathscr{C}_0$, and hence $[\phi]=0$.

Next we define $\overline{\mathscr{C}}$ as the Hilbert space given by the closure of $\mathscr{C}$ with the norm given by the inner product $\langle\phi,\psi\rangle_{\mathscr{C}}$, and set up a scheme which is very similar to the one used in the proof of Lemma 5.3, which is of course similar to the arguments used in [Belishev and Kurylev 1992; Katchalov et al. 2001]. For $\tau\in(0,T)$ define

$$\overline{\mathscr{C}}_\tau=\{\phi\in\overline{\mathscr{C}}:\phi(t,z)=0,\ t<\tau\},$$

and let

$$\mathscr{P}_\tau:\overline{\mathscr{C}}\to\overline{\mathscr{C}}_\tau$$

be the orthogonal projection to $\overline{\mathscr{C}}_\tau$. Then, using propagation of singularities (and here we do not have to project onto the continuous spectrum), and that the choices for $t=0$ and $t=T$ are arbitrary, we recover the metric tensor g and the fundamental solution of wave equation in $B(p,r)$, where $r=r(p)$ is the radius of injectivity of $\exp_p$. In other words, we recover

$$g(z),\quad z\in B(p,r)\quad\text{and}\quad \tilde{U}(t,t',z,z'),\quad t,\ t'\in\mathbb{R},\ z,\ z'\in B(p,r),\ r=r(p).$$

We repeat the process for every $p \in \mathbb{O}$, and we would like to define $\mathcal{M} = \bigcup_{p\in\mathbb{O}} B(p, r(p))$. However, we have to make sure the inclusion map $\iota : \mathcal{M} \hookrightarrow X$ is injective, which would guarantee that $\iota(\mathcal{M})$ is an open embedded submanifold of X. Therefore we need to identify the points that are in $B(p, r(p))$ and $B(q, r(q))$. In Section 4.4.9 of [Katchalov et al. 2001], since they are working on a compact manifold, they use the family of eigenfunctions to do that. Here the precise analogue is to use $\tilde{U}(t, t', z, z')$, and we shall say that $z \in B(p, r(p))$, and $w \in B(q, r(q))$ are equivalent, and we denote $z \equiv w$ if $\tilde{U}(t, t', z, z') = \tilde{U}(t, t', w, z')$ for all $t, t' > 0$ and $z' \in \mathbb{O}$. In this case, the points z and w correspond to the same point in X. This is the equivalent of saying that $u^\phi(t, z) = u^\phi(t, w)$ for all $t \in \mathbb{R}$ and for all $\phi \in C_0^\infty(\mathbb{R}\times\mathbb{O})$. We also use the same identification for points in $\mathbb{O}$ and $B(p, r(p))$, $p \in \mathbb{O}$. With this identification, we set $\mathbb{O}_1 = \big(\bigcup_{p\in\mathbb{O}} B(p, r(p))\big) \cup \mathbb{O}$.

We have constructed an open C^∞ submanifold $\mathbb{O}_1 \subset X$ such that $\mathbb{O} = \mathbb{O}_0 \subset \mathbb{O}_1$ and such that hypotheses (A), (B) and (C) are satisfied for $\mathbb{O}_1$. Now we repeat the process for $\mathbb{O}_1$. Thus we obtain a sequence of C^∞ open submanifolds $\mathbb{O}_j \subset X$ satisfying $\mathbb{O}_j \subset \mathbb{O}_{j+1} \subset X$, $j = 0, 1, \dots$, and satisfying the hypotheses (A), (B) and (C) above. As in Section 4.4.9 of [Katchalov et al. 2001], we claim that for any compact subset $K \subset X$ there exists $J \in \mathbb{N}$ such that $K \subset \mathbb{O}_J$. To see that, we observe that, since (X, g) is complete, there exists $M > 0$ such that, for any $p \in K$, $\delta < \varepsilon$ and $\Gamma' \Subset \Gamma$, $d_g(p, \Gamma' \times \{\delta\}) \le M$. We also assume that $\delta < \delta_0$, where δ_0 is the radius of injectivity of X. Since X is complete, given a point $p \in K$ there is a geodesic $\mu(s)$, parametrized by the arc length $0 \le s \le L \le M$, joining p to a point $z \in \Gamma' \times \delta$. Let $x_0 = z$ and $x_k = \mu(k\delta)$, with $k = 0, 1, \dots, [L/\delta] = J$. By definition $x_0 = z \in \Gamma \times \{\delta\} \subset \mathbb{O} = \widetilde{\mathbb{O}}_0$. Suppose that $x_k \in \widetilde{\mathbb{O}}_k$; then there exists $\rho > 0$ such that $B(x_k, \rho) \subset \mathbb{O}_k$ but, since δ is less than the radius of injectivity, $\overline{B(x_k, \delta)} \subset \mathbb{O}_{k+1}$ and, since s is the arc length, in particular $x_{k+1} \in \mathbb{O}_{k+1}$. By induction it follows that $p \in \mathbb{O}_{J+1} \subset \mathbb{O}_{[M/\delta]}$.

This shows that we can reconstruct $(\mathring{X}, g)$ from (A), (B) and (C). But we know a priori that (X, g) is an AHM, and so $\mathring{X}$ can be compactified into a C^∞ with boundary, and there exists a defining function x of ∂X for which (2-4) holds. The construction of the function x shows that the compactification is uniquely defined modulo diffeomorphisms that are equal to the identity in $\mathbb{O}$. □

Acknowledgements

Hora and Sá Barreto are grateful to two anonymous referees for carefully reviewing the paper and making numerous very useful suggestions. Both authors were supported by the NSF grant DMS-0901334.

References

[Alinhac 1984] S. Alinhac, "Unicité du problème de Cauchy pour des opérateurs du second ordre à symboles réels", *Ann. Inst. Fourier* (*Grenoble*) **34**:2 (1984), 89–109. MR 86a:35003 Zbl 0507.35001

[Belishev 1987] M. I. Belishev, "An approach to multidimensional inverse problems for the wave equation", *Dokl. Akad. Nauk SSSR* **297**:3 (1987), 524–527. In Russian; translated in *Soviet Math. Dokl.* **36**:3 (1988), 172–182. MR 89c:35152 Zbl 0661.35084

[Belishev and Kurylev 1992] M. I. Belishev and Y. V. Kurylev, "To the reconstruction of a Riemannian manifold via its spectral data (BC-method)", *Comm. Partial Differential Equations* **17**:5-6 (1992), 767–804. MR 94a:58199 Zbl 0812.58094

[Bouclet 2013] J.-M. Bouclet, "Absence of eigenvalue at the bottom of the continuous spectrum on asymptotically hyperbolic manifolds", *Ann. Global Anal. Geom.* **44**:2 (2013), 115–136. MR 3073583 Zbl 1286.58019

[Friedlander 1980] F. G. Friedlander, "Radiation fields and hyperbolic scattering theory", *Math. Proc. Cambridge Philos. Soc.* **88**:3 (1980), 483–515. MR 82f:35112 Zbl 0465.35068

[Friedlander 2001] F. G. Friedlander, "Notes on the wave equation on asymptotically Euclidean manifolds", *J. Funct. Anal.* **184**:1 (2001), 1–18. MR 2002k:58059 Zbl 0997.58013

[Graham 2000] C. R. Graham, "Volume and area renormalizations for conformally compact Einstein metrics", pp. 31–42 in *The Proceedings of the 19th Winter School "Geometry and Physics"* (Srní, 1999), vol. 63, 2000. MR 2002c:53073 Zbl 0984.53020

[Graham and Zworski 2003] C. R. Graham and M. Zworski, "Scattering matrix in conformal geometry", *Invent. Math.* **152**:1 (2003), 89–118. MR 2004c:58064 Zbl 1030.58022

[Guillarmou and Sá Barreto 2008] C. Guillarmou and A. Sá Barreto, "Scattering and inverse scattering on ACH manifolds", *J. Reine Angew. Math.* **622** (2008), 1–55. MR 2010i:58028 Zbl 1159.58016

[Guillarmou and Sá Barreto 2009] C. Guillarmou and A. Sá Barreto, "Inverse problems for Einstein manifolds", *Inverse Probl. Imaging* **3**:1 (2009), 1–15. MR 2010m:58040 Zbl 1229.58025

[Helgason 1999] S. Helgason, *The Radon transform*, 2nd ed., Progress in Mathematics **5**, Birkhäuser, Boston, 1999. MR 2000m:44003 Zbl 0932.43011

[Hörmander 1994a] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudo-differential operators*, Grundlehren der Mathematischen Wissenschaften **274**, Springer, Berlin, 1994. MR 95h:35255 Zbl 1115.35005

[Hörmander 1994b] L. Hörmander, *The analysis of linear partial differential operators, IV: Fourier integral operators*, Grundlehren der Mathematischen Wissenschaften **275**, Springer, Berlin, 1994. MR 98f:35002 Zbl 1178.35003

[Hörmander 1997] L. Hörmander, "On the uniqueness of the Cauchy problem under partial analyticity assumptions", pp. 179–219 in *Geometrical optics and related topics* (Cortona, 1996), edited by F. Colombini and N. Lerner, Progr. Nonlinear Differential Equations Appl. **32**, Birkhäuser, Boston, 1997. MR 2033496 Zbl 0907.35002

[Isozaki and Kurylev 2014] H. Isozaki and Y. Kurylev, *Introduction to spectral theory and inverse problem on asymptotically hyperbolic manifolds*, MSJ Memoirs **32**, Mathematical Society of Japan, Tokyo, 2014. MR 3222614 Zbl 06317422

[Isozaki et al. 2010] H. Isozaki, Y. Kurylev, and M. Lassas, "Forward and inverse scattering on manifolds with asymptotically cylindrical ends", *J. Funct. Anal.* **258**:6 (2010), 2060–2118. MR 2010k:58047 Zbl 1204.58023

[Isozaki et al. 2013] H. Isozaki, Y. Kurylev, and M. Lassas, "Spectral theory and inverse problem on asymptotically hyperbolic orbifolds", preprint, 2013. arXiv 1312.0421

[Isozaki et al. 2014] H. Isozaki, Y. Kurylev, and M. Lassas, "Recent progress of inverse scattering theory on non-compact manifolds", pp. 143–163 in *Inverse problems and applications*, edited by P. Stefanov et al., Contemp. Math. **615**, Amer. Math. Soc., Providence, RI, 2014. MR 3221603

[Joshi and Sá Barreto 2000] M. S. Joshi and A. Sá Barreto, "Inverse scattering on asymptotically hyperbolic manifolds", *Acta Math.* **184**:1 (2000), 41–86. MR 2002g:58052 Zbl 1142.58309

[Katchalov et al. 2001] A. Katchalov, Y. Kurylev, and M. Lassas, *Inverse boundary spectral problems*, Monographs and Surveys in Pure and Applied Mathematics **123**, Chapman & Hall/CRC, Boca Raton, FL, 2001. MR 2003e:58045 Zbl 1037.35098

[Krupchyk et al. 2008] K. Krupchyk, Y. Kurylev, and M. Lassas, "Inverse spectral problems on a closed manifold", *J. Math. Pures Appl.* (9) **90**:1 (2008), 42–59. MR 2010h:58052 Zbl 1146.58020

[Kurylev and Lassas 2000] Y. V. Kurylev and M. Lassas, "Hyperbolic inverse problem with data on a part of the boundary", pp. 259–272 in *Differential equations and mathematical physics* (Birmingham, AL, 1999), edited by R. Weikard and G. Weinstein, AMS/IP Stud. Adv. Math. **16**, Amer. Math. Soc., Providence, RI, 2000. MR 2001f:58061 Zbl 1161.58315

[Kurylev and Lassas 2002] Y. Kurylev and M. Lassas, "Hyperbolic inverse boundary-value problem and time-continuation of the non-stationary Dirichlet-to-Neumann map", *Proc. Roy. Soc. Edinburgh Sect. A* **132**:4 (2002), 931–949. MR 2003h:35279 Zbl 1034.35152

[Lassas and Oksanen 2014] M. Lassas and L. Oksanen, "Inverse problem for the Riemannian wave equation with Dirichlet data and Neumann data on disjoint sets", *Duke Math. J.* **163**:6 (2014), 1071–1103. MR 3192525 Zbl 06298159

[Lassas et al. 2003] M. Lassas, M. Taylor, and G. Uhlmann, "The Dirichlet-to-Neumann map for complete Riemannian manifolds with boundary", *Comm. Anal. Geom.* **11**:2 (2003), 207–221. MR 2004h:58033 Zbl 1077.58012

[Lax and Phillips 1982] P. D. Lax and R. S. Phillips, "A local Paley–Wiener theorem for the Radon transform of L_2 functions in a non-Euclidean setting", *Comm. Pure Appl. Math.* **35**:4 (1982), 531–554. MR 83i:43016 Zbl 0511.58037

[Lax and Phillips 1989] P. D. Lax and R. S. Phillips, *Scattering theory*, 2nd ed., Pure and Applied Mathematics **26**, Academic Press, Boston, 1989. MR 90k:35005 Zbl 0697.35004

[Mazzeo 1988] R. Mazzeo, "The Hodge cohomology of a conformally compact metric", *J. Differential Geom.* **28**:2 (1988), 309–339. MR 89i:58005 Zbl 0656.53042

[Mazzeo 1991] R. Mazzeo, "Unique continuation at infinity and embedded eigenvalues for asymptotically hyperbolic manifolds", *Amer. J. Math.* **113**:1 (1991), 25–45. MR 92f:58187 Zbl 0725.58044

[Mazzeo and Melrose 1987] R. R. Mazzeo and R. B. Melrose, "Meromorphic extension of the resolvent on complete spaces with asymptotically constant negative curvature", *J. Funct. Anal.* **75**:2 (1987), 260–310. MR 89c:58133 Zbl 0636.58034

[Melrose et al. 2014a] R. Melrose, A. Sá Barreto, and A. Vasy, "Asymptotics of solutions of the wave equation on de Sitter–Schwarzschild space", *Comm. Partial Differential Equations* **39**:3 (2014), 512–529. MR 3169793 Zbl 1286.35145

[Melrose et al. 2014b] R. Melrose, A. Sá Barreto, and A. Vasy, "Analytic continuation and semiclassical resolvent estimates on asymptotically hyperbolic spaces", *Comm. Partial Differential Equations* **39**:3 (2014), 452–511. MR 3169792 Zbl 06291160

[Robbiano and Zuily 1998] L. Robbiano and C. Zuily, "Uniqueness in the Cauchy problem for operators with partially holomorphic coefficients", *Invent. Math.* **131**:3 (1998), 493–539. MR 99e:35004 Zbl 0909.35004

[Sá Barreto 2005] A. Sá Barreto, "Radiation fields, scattering, and inverse scattering on asymptotically hyperbolic manifolds", *Duke Math. J.* **129**:3 (2005), 407–480. MR 2006g:58058 Zbl 1154.58310

[Sá Barreto and Wunsch 2005] A. Sá Barreto and J. Wunsch, "The radiation field is a Fourier integral operator", *Ann. Inst. Fourier* (*Grenoble*) **55**:1 (2005), 213–227. MR 2006a:58033 Zbl 1091.58018

[Tataru 1995] D. Tataru, "Unique continuation for solutions to PDE's; between Hörmander's theorem and Holmgren's theorem", *Comm. Partial Differential Equations* **20**:5-6 (1995), 855–884. MR 96e:35019 Zbl 0846.35021

[Tataru 1999] D. Tataru, "Unique continuation for operators with partially analytic coefficients", *J. Math. Pures Appl.* (9) **78**:5 (1999), 505–521. MR 2000e:35005 Zbl 0936.35038

[Vasy 2013] A. Vasy, "Microlocal analysis of asymptotically hyperbolic spaces and high-energy resolvent estimates", pp. 487–528 in *Inverse problems and applications: inside out, II*, edited by G. Uhlmann, Math. Sci. Res. Inst. Publ. **60**, Cambridge Univ. Press, Cambridge, 2013. MR 3135765

[Vasy and Wunsch 2005] A. Vasy and J. Wunsch, "Absence of super-exponentially decaying eigenfunctions of Riemannian manifolds with pinched negative curvature", *Math. Res. Lett.* **12**:5-6 (2005), 673–684. MR 2006m:58048 Zbl 1121.58025

Received 3 Sep 2013. Revised 23 Dec 2014. Accepted 22 Jan 2015.

RAPHAEL HORA: rhora@mtm.ufsc.br
Departamento de Matemática, Universidade Federal de Santa Catarina, Campus Universitário Trindade, 88040-900 Florianópolis, SC, Brazil

ANTÔNIO SÁ BARRETO: sabarre@math.purdue.edu
Department of Mathematics, Purdue University, 150 North University Street, West Lafayette, IN 47907, United States

ANALYSIS AND PDE
Vol. 8, No. 3, 2015
dx.doi.org/10.2140/apde.2015.8.561

LOW TEMPERATURE ASYMPTOTICS FOR QUASISTATIONARY DISTRIBUTIONS IN A BOUNDED DOMAIN

TONY LELIÈVRE AND FRANCIS NIER

We analyze the low temperature asymptotics of the quasistationary distribution associated with the overdamped Langevin dynamics (also known as the Einstein–Smoluchowski diffusion equation) in a bounded domain. This analysis is useful to rigorously prove the consistency of an algorithm used in molecular dynamics (the hyperdynamics) in the small temperature regime. More precisely, we show that the algorithm is exact in terms of state-to-state dynamics up to exponentially small factors in the limit of small temperature. The proof is based on the asymptotic spectral analysis of associated Dirichlet and Neumann realizations of Witten Laplacians. In order to widen the range of applicability, the usual assumption that the energy landscape is a Morse function has been relaxed as much as possible.

1.	Introduction	561
2.	Assumptions and statements of the main results	565
3.	A priori exponential decay and first consequences	571
4.	Quasimodes for $\Delta^{D,(0)}_{f,h}(\Omega_+)$ and $\Delta^{D,(1)}_{f,h}(\Omega_+)$	579
5.	Analysis of the restricted differential β	585
6.	Proof of Theorem 2.4 and two corollaries	594
7.	About Hypotheses 3 and 4	604
Appendix: Riemannian geometry formulas		622
Acknowledgements		627
References		627

1. Introduction

The motivation of this work comes from the mathematical analysis of an algorithm used in molecular dynamics, called the hyperdynamics [Voter 1997]. The aim of this algorithm is to generate very efficiently the discrete state-to-state dynamics associated with a continuous state space, metastable, Markovian dynamics, by modifying the potential function. In Section 1A, we explain the principle of the algorithm and state the mathematical problem. In Section 1B, the main result of this article is given in a simple setting.

MSC2010: 58J10, 58J32, 60J65, 60J70, 81Q20.
Keywords: quasistationary distributions, Witten Laplacian, low temperature asymptotics and semiclassical asymptotics.

1A. ***Molecular dynamics, hyperdynamics and the quasistationary distribution.*** Molecular dynamics calculations consist in simulating very long trajectories of a particle model of matter, in order to infer macroscopic properties from an atomic description. Examples include the study of the change of conformation of large molecules (such as proteins), with applications in biology, or the description of the motion of defects in materials.

In a constant-temperature environment, the dynamics used in practice contains stochastic terms which model thermostatting. The prototypical example, which is the focus of this work, is the overdamped Langevin dynamics,

$$dX_t = -\nabla f(X_t)\,dt + \sqrt{2\beta^{-1}}\,dB_t, \tag{1-1}$$

where $X_t \in \mathbb{R}^{3N}$ is the position vector of N particles, $f : \mathbb{R}^{3N} \to \mathbb{R}$ is the potential function (assumed to be smooth here), and $\beta^{-1} = k_B T$ with k_B the Boltzmann constant and T the temperature. The stochastic process B_t is a standard $3N$-dimensional Brownian motion. The dynamics (1-1) admits the canonical ensemble $\mu(dx) = Z^{-1}\exp(-\beta f(x))\,dx$ as an invariant probability measure.

To relate the macroscopic properties of matter to the microscopic phenomenon, one simulates the process $(X_t)_{t\geq 0}$ (or processes following related dynamics, like the Langevin dynamics) over very long times. The difficulty associated with such simulations is *metastability*, namely the fact that the stochastic process remains trapped for very long times in some regions of the configurational space, called the metastable states. The time step used to obtain stable discretization is typically 10^{-15} s, while the macroscopic timescales of interest range from a few microseconds to a few seconds. At the macroscopic level, the details of the dynamics $(X_t)_{t\geq 0}$ do not matter. The important information is the history of the visited metastable states, the so-called *state-to-state dynamics*.

The principle of the hyperdynamics algorithm [Voter 1997] is to modify the potential f in order to accelerate the exit from metastable states, while keeping a correct state-to-state dynamics. Here, we focus on one elementary brick of this dynamics, namely the exit event from a given metastable state.

In mathematical terms, the problem is as follows (we refer to [Le Bris et al. 2012] for the mathematical proofs of the statements below). Assuming that the process remains trapped for a very long time in a domain $\Omega_+ \subset \mathbb{R}^{3N}$ (Ω_+ is a metastable state,[1] as mentioned above), it is known that the process reaches a local equilibrium called the quasistationary distribution (QSD) ν attached to the domain Ω_+, before leaving it. We assume that Ω_+ is a smooth bounded domain in $\mathbb{R}^{3N}$. The probability distribution ν has support Ω_+ and is such that, for all smooth test function $\varphi : \mathbb{R}^{3N} \to \mathbb{R}$,

$$\lim_{t\to\infty} \mathbb{E}(\varphi(X_t)|\tau > t) = \int_{\Omega_+} \varphi\,d\nu, \tag{1-2}$$

where

$$\tau = \inf\{t > 0 : X_t \notin \Omega_+\}$$

is the first exit time from Ω_+ for X_t. The metastability of the well Ω_+ can be quantified through the rate of convergence of the limit in (1-2); in the following, it is assumed that this convergence is infinitely fast. From a PDEs viewpoint, ν has a density v with respect to the Boltzmann–Gibbs measure

[1] We use the notation Ω_+ since, in the following, we will need a subdomain Ω_- such that $\overline{\Omega}_- \subset \Omega_+$.

$\mu(dx) = e^{-\beta f(x)}\,dx$, v being the first eigenvector of the infinitesimal generator of the dynamics (1-1) with Dirichlet boundary conditions on $\partial\Omega_+$:

$$\begin{cases} -\nabla f \cdot \nabla v + \beta^{-1}\Delta v = -\lambda v & \text{in } \Omega_+, \\ v = 0 & \text{on } \partial\Omega_+, \end{cases} \tag{1-3}$$

where $-\lambda < 0$ is the first eigenvalue. In other words,

$$d\nu = \frac{1_{\Omega_+}(x) v(x) \exp(-\beta f(x))\,dx}{\int_{\Omega_+} v(x) \exp(-\beta f(x))\,dx}.$$

Starting from the QSD ν (namely if $X_0 \sim \nu$), the way the stochastic process X_t, solution to (1-1), leaves the well Ω_+ is known: the law of the pair of random variables (τ, X_τ) (exit time, exit point) is characterized by the following three properties, the first two of which are the building blocks of a Markovian transition starting from Ω_+:

(i) τ and X_τ are independent.

(ii) τ is exponentially distributed with parameter λ:

$$\tau \sim \mathcal{E}(\lambda), \tag{1-4}$$

where the notation $\sim$ is used to indicate the law of a random variable.

(iii) The exit point distribution has an analytic expression in terms of v: for all smooth test functions $\varphi : \partial\Omega_+ \to \mathbb{R}$,

$$\mathbb{E}^\nu(\varphi(X_\tau)) = -\frac{\int_{\partial\Omega_+} \varphi \partial_n (v \exp(-\beta f))\,d\sigma}{\beta\lambda \int_{\Omega_+} v(x) \exp(-\beta f(x))\,dx}, \tag{1-5}$$

where, for any smooth function $w : \Omega_+ \to \mathbb{R}$, $\partial_n w = \nabla w \cdot n$ denotes the outward normal derivative, σ is the Lebesgue measure on $\partial\Omega_+$ and $\mathbb{E}^\nu$ indicates the expectation for the stochastic process X_t following (1-1) and starting under the QSD, $X_0 \sim \nu$.

In practical cases of interest, the typical exit time is very large ($\mathbb{E}(\tau) = 1/\lambda$ is very large). The principle of the hyperdynamics is to modify the potential f in the state Ω_+ to lead to smaller exit times, while keeping a correct statistics on the exit points. Let us make this more precise, and let us consider the process $X_t^{\delta f}$ which evolves on a new potential $f + \delta f$:

$$dX_t^{\delta f} = -\nabla(f + \delta f)(X_t^{\delta f})\,dt + \sqrt{2\beta^{-1}}\,dB_t. \tag{1-6}$$

Instead of simulating $(X_t)_{t\geq 0}$ following the dynamics (1-1) and considering the associated random variables (τ, X_τ), the hyperdynamics algorithm consists in simulating $(X_t^{\delta f})_{t\geq 0}$ and considering the associated random variables $(\tau^{\delta f}, X^{\delta f}_{\tau^{\delta f}})$, where $\tau^{\delta f}$ is the first exit time from Ω_+ for $X_t^{\delta f}$.

The assertion underlying the hyperdynamics algorithm is the following: under appropriate assumptions on the perturbation δf, (i) the exit point distribution of $X_t^{\delta f}$ from Ω_+ is (almost) the same as the exit point distribution of X_t from Ω_+, and (ii) the exit time distribution for X_t can be inferred from the exit time distribution for $X_t^{\delta f}$ by a simple multiplicative factor (see (1-7)–(1-8) below).

More precisely, the assumptions on δf in [Voter 1997] can be stated as follows: (i) δf is sufficiently small that Ω_+ is still a metastable state for $X_t^{\delta f}$, and (ii) δf is zero on the boundary of Ω_+. The first hypothesis implies that we can assume that $X_0^{\delta f}$ is distributed according to the QSD $\nu^{\delta f}$ associated with (1-6) and Ω_+. The aim of this paper is to prove that, in the small temperature regime (namely $\beta \to \infty$) and under appropriate assumptions on δf, we indeed have the equality in law

$$(\tau, X_\tau) \overset{\mathcal{L}}{\simeq} (B\tau^{\delta f}, X^{\delta f}_{\tau^{\delta f}}), \tag{1-7}$$

where, in the left-hand side, $X_0 \sim \nu$ and, in the right-hand side, $X_0^{\delta f} \sim \nu^{\delta f}$. The so-called boost factor B has the expression

$$B = \frac{\int_{\Omega_+} \exp(-\beta f)}{\int_{\Omega_+} \exp(-\beta(f+\delta f))} = \int_{\Omega_+} \exp(\beta \delta f) \frac{\exp(-\beta(f+\delta f))}{\int_{\Omega_+} \exp(-\beta(f+\delta f))}. \tag{1-8}$$

The second formula is interesting because it shows that B can be approximated through ergodic averages on the process $(X_t^{\delta f})_{t \geq 0}$ (and this is actually exactly what is done in practice).

In view of the formulas (1-4)–(1-5) for the laws of the distributions of the two random variables exit time and exit point, a crucial point for the mathematical analysis of the hyperdynamics algorithm is to study how the first eigenvalue λ and the normal derivative $\partial_n v$ (v being the first eigenvector; see (1-3)) are modified when changing the potential f to $f + \delta f$. More precisely, we would like to check that, in the limit $\beta \to \infty$, $\lambda^{\delta f} = B\lambda$ and, up to a multiplicative constant, $\partial_n v^{\delta f} \propto \partial_n v$, where, with obvious notation, $(-\lambda^{\delta f}, v^{\delta f})$ denotes the first eigenvalue–eigenfunction pair solution to (1-3) when f is replaced by $f + \delta f$.

1B. ***The main results in a simple setting.*** Let us state the main results obtained in this paper in a simple and restricted setting. For the potential f, we assume that there exists a subdomain Ω_- such that $\overline{\Omega}_- \subset \Omega_+$ and:

(i) f and $f\big|_{\partial\Omega_+}$ are Morse functions, namely $\mathcal{C}^\infty$ functions with nondegenerate critical points;

(ii) $|\nabla f| \neq 0$ in $\overline{\Omega}_+ \setminus \Omega_-$, $\partial_n f > 0$ on $\partial\Omega_-$ and $\min_{\partial\Omega_+} f \geq \min_{\partial\Omega_-} f$;

(iii) the critical values of f in Ω_- are all distinct and the differences $f(U^{(1)}) - f(U^{(0)})$, where $U^{(0)}$ ranges over the local minima of $f\big|_{\Omega_-}$ and $U^{(1)}$ ranges over the critical points of $f\big|_{\Omega_-}$ with index 1, are all distinct;

(iv) the maximal value of f at critical points, denoted by $\mathrm{cvmax} = \max\{f(x) : x \in \Omega_+, |\nabla f(x)| = 0\} = \max\{f(x) : x \in \Omega_-, |\nabla f(x)| = 0\}$, satisfies

$$\min_{\partial\Omega_-} f - \mathrm{cvmax} > \mathrm{cvmax} - \min_{\Omega_-} f. \tag{1-9}$$

Concerning the perturbation δf, let us assume that $f + \delta f$ satisfies the same four above hypotheses as f, and that, in addition,

$$\delta f = 0 \quad \text{on } \Omega_+ \setminus \Omega_-.$$

Under these assumptions on f and δf, it can be shown that the first eigenvalue–eigenfunction pairs $(-\lambda, v)$ and $(-\lambda^{\delta f}, v^{\delta f})$, the respective solutions to (1-3) with the potential f and $f+\delta f$, satisfy the following estimate: for some positive constant c, in the limit $\beta \to \infty$,

$$\frac{\lambda^{\delta f}}{\lambda} = B(1+\mathcal{O}(e^{-\beta c})),$$

where, we recall, B is defined by (1-8) and

$$\frac{\partial_n v\big|_{\partial\Omega_+}}{\|\partial_n v\|_{L^1(\partial\Omega_+)}} = \frac{\partial_n v^{\delta f}\big|_{\partial\Omega_+}}{\|\partial_n v^{\delta f}\|_{L^1(\partial\Omega_+)}} + \mathcal{O}(e^{-\beta c}) \quad \text{in } L^1(\partial\Omega_+).$$

These results are simple consequences of the general Theorem 2.4 below (see Corollary 2.9) together with Proposition 7.1 and Remark 7.2.

For readers who are familiar with the Agmon distance, let us note that condition (1-9) can actually be replaced by Hypothesis 2 (stated in Section 2) and condition (7-1). Condition (7-1) explicitly states that the potential function f on $\partial\Omega_-$ should be larger than the largest barrier (difference of potential between index-one critical points and local minima) within Ω_-.

1C. ***Outline of the article.*** The main result of this article, Theorem 2.4, gives general asymptotic formulas for the first eigenvalue λ and the normal derivative $\partial_n v$ in the limit of small temperature. This theorem will be proven under assumptions involving the low-lying spectra of Witten Laplacians on Ω_- and on $\Omega_+ \setminus \overline{\Omega}_-$. These assumptions hold for potentials satisfying the four conditions (i)–(iv) stated above, but they are also valid in much more general cases. In particular, we have in mind assumptions stated only in terms of Ω_+ (see Remark 7.4), or potentials not fulfilling the Morse assumption (see Section 7B).

The outline of the article is as follows: In Section 2, we specify our general assumptions and state the two main theorems, Theorem 2.4 and Theorem 2.10. In Section 3, exponential decay estimates for the eigenvectors in terms of Agmon distances are reviewed. In Section 4, approximate eigenvectors for the Dirichlet Witten Laplacians on Ω_+ are constructed in terms of eigenvectors for the Neumann Witten Laplacians on Ω_- and eigenvectors for the Dirichlet Witten Laplacians on the shell $\Omega_+ \setminus \overline{\Omega}_-$. Following the strategy of [Helffer et al. 2004; Helffer and Nier 2006; Le Peutrec 2009; 2010b; 2011; Le Peutrec et al. 2013], accurate approximations of singular values of the Witten differential $d_{f,h}$ are computed using matrix arguments in Section 5. Theorem 2.4 and Theorem 2.10 are finally proved in Section 6. The general assumptions used to prove the theorems are then thoroughly discussed and illustrated with various examples in Section 7. Our approach relies on the introduction of boundary Witten Laplacians (namely Witten Laplacians with Dirichlet or Neumann boundary conditions) and requires notions and notation of Riemannian differential geometry. A short presentation of these notions is given in the Appendix.

2. Assumptions and statements of the main results

In order to prove the main result, we first need to restate the eigenvalue problem (1-3) with the standard notation used in the framework of Witten Laplacians, which will be our central tool. It is easy to check that (λ, v) satisfies (1-3) if and only if (λ_1, u_1) satisfies

$$\Delta_{f,h}^{D,(0)}(\Omega_+)u_1 = \lambda_1 u_1$$

with

$$h = \frac{2}{\beta}, \quad \lambda_1 = \frac{4}{\beta}\lambda = 2h\lambda, \quad u_1 = \exp\left(-\tfrac{1}{2}\beta f\right)v = \exp\left(-\frac{f}{h}\right)v$$

and where $\Delta_{f,h}^{D,(0)}(\Omega_+)$ is the Witten Laplacian on zero-forms on $\Omega_+ \subset \mathbb{R}^d$, $d = 3N$, with homogeneous Dirichlet boundary conditions on $\partial\Omega_+$ (see (2-3) below for more general formulas on p-forms),

$$\Delta_{f,h}^{D,(0)}(\Omega_+)u_1 = (-h\nabla + \nabla f)\cdot((h\nabla + \nabla f)u_1) = -h^2\Delta u_1 + (|\nabla f|^2 - h\Delta f)u_1. \tag{2-1}$$

Notice that the operator $\Delta_{f,h}^{D,(0)}(\Omega_+)$ is a *positive* symmetric operator. We recall that Ω_+ is the metastable domain of interest, and Ω_- is a subdomain of Ω_+, where the potential f is modified in the hyperdynamics algorithm. We will thus study how the first eigenvalue λ_1 and eigenfunction u_1 of the Witten Laplacian $\Delta_{f,h}^{D,(0)}(\Omega_+)$ depend on $f\big|_{\Omega_-}$. We will state the results in a very general setting, namely for open, regular, bounded, connected subsets Ω_- and Ω_+ of a d-dimensional Riemannian manifold (M, g) such that $\overline{\Omega}_- \subset \Omega_+$.

The first assumption we make on f is the following:

Hypothesis 1. *The function $f : M \to \mathbb{R}$ is a $\mathcal{C}^\infty$ function satisfying*

$$|\nabla f| > 0 \text{ on } \overline{\Omega}_+ \setminus \Omega_-, \quad \partial_n f > 0 \text{ on } \partial\Omega_- \quad \text{and} \quad \min_{\partial\Omega_+} f \geq \min_{\partial\Omega_-} f. \tag{2-2}$$

In (2-2), n denotes the unit normal vector on $\partial\Omega_-$ that points outward from Ω_-. This first assumption has simple consequences that will be used repeatedly.

Lemma 2.1. *Under Hypothesis 1, for all $x \in \overline{\Omega}_+ \setminus \Omega_-$,*

$$f(x) \geq \min_{\partial\Omega_-} f > \min_{\Omega_-} f = \min_{\Omega_+} f.$$

Proof. The last equality is a simple consequence of the fact that the critical points are in Ω_- and of the inequality $\min_{\partial\Omega_+} f \geq \min_{\partial\Omega_-} f$. Let us now consider the first inequality. Let us denote by $\gamma_x(t)$ the gradient trajectory $\dot{\gamma}_x = -\nabla f(\gamma_x)$ starting from $x \in \Omega_+$ ($\gamma_x(0) = x$). Let us consider $x \in \overline{\Omega}_+ \setminus \Omega_-$ such that $f(x) < \min_{\partial\Omega_+} f$. Since $t \mapsto f(\gamma_x(t))$ is nonincreasing, $(\gamma_x(t))_{t\geq 0}$ remains in the bounded domain Ω_+ and is thus well defined for all positive times. Moreover, necessarily, the distance of $\gamma_x(t)$ to the set of critical points of f tends to 0 as $t \to \infty$. This implies that there exists $t_0 > 0$ such that $\gamma_x(t_0) \in \Omega_-$ and, thus, $f(x) = f(\gamma_x(0)) \geq f(\gamma_x(t_0)) \geq \min_{\partial\Omega_-} f$. This concludes the proof of the first inequality. The second inequality is a consequence of the assumption $\partial_n f > 0$ on $\partial\Omega_-$, and is proven by considering the trajectory $(\gamma_x(t))_{t\geq 0}$ with $x \in \arg\min_{\partial\Omega_-} f$. □

Remark 2.2. One can easily check, using the same arguments, that the condition $\partial_n f > 0$ on $\partial\Omega_+$, together with the two first conditions of Hypothesis 1, implies $\min_{\partial\Omega_+} f > \min_{\partial\Omega_-} f$.

The second assumption on f is:

Hypothesis 2. *There exists $c_0 > 0$ such that the set of critical points of f in Ω_+ is included in $\{f < \min_{\partial\Omega_+} f - c_0\}$:*

$$\{x \in \Omega_+ : \nabla f(x) = 0\} \subset \Big\{x \in \Omega_+ : f(x) < \min_{\partial\Omega_+} f - c_0\Big\}.$$

In addition to Hypotheses 1 and 2, our main results are stated under assumptions on the spectrum of the Witten Laplacians associated with f on Ω_- and $\Omega_+ \setminus \overline{\Omega}_-$ (see Hypotheses 3 and 4 below). We will discuss more explicit assumptions on f for which those additional hypotheses are satisfied in Section 7. Let us first define the Witten Laplacians. We refer the reader to [Witten 1982; Helffer and Sjöstrand 1985b; Cycon et al. 1987; Burghelea 1997; Zhang 2001] for introductory texts on the semiclassical analysis of Witten Laplacians and its famous application to Morse inequalities, and related results.

The Witten Laplacians are defined on $\bigwedge \mathcal{C}^\infty(M) = \bigoplus_{p=0}^d \bigwedge^p \mathcal{C}^\infty(M)$ as

$$\Delta_{f,h} = (d_{f,h}^* + d_{f,h})^2 = d_{f,h}^* d_{f,h} + d_{f,h} d_{f,h}^*, \qquad \text{where } d_{f,h} = e^{-f/h}(hd)e^{f/h} \text{ and } d_{f,h}^* = e^{f/h}(hd^*)e^{-f/h}. \tag{2-3}$$

On a domain $\Omega \subset M$ and for $m \in \mathbb{N}$, the Sobolev space $\bigwedge W^{m,2}(\Omega)$ is defined as the set of $u \in \Lambda L^2(\Omega)$ such that, locally, $\partial_x^\alpha u \in \bigwedge L^2(\Omega)$ for all $\alpha \in \mathbb{N}^d$ with $|\alpha| \le m$ (this property does not depend on the local coordinate system $(x^1, \dots, x^d)$). When Ω is a regular bounded domain, $\bigwedge W^{m,2}(\Omega)$ coincides with the set of $u \in \bigwedge L^2$ such that there exists $\tilde{u} \in \bigwedge W^{m,2}(M)$ such that $\tilde{u}\big|_\Omega = u$. The spaces $\bigwedge W^{s,2}(\Omega)$ for $s \in \mathbb{R}$ are then defined by duality and interpolation. For $m = 1$, the quantity $\sqrt{\|u\|_{L^2(\Omega)}^2 + \|du\|_{L^2(\Omega)}^2 + \|d^*u\|_{L^2(\Omega)}^2}$ is equivalent to the $W^{1,2}(\Omega)$-norm. This is a well-known result when $\Omega = \mathbb{R}^d$. The extension to a regular, bounded domain is proved by using local charts and the reflexion principle; see [Taylor 1997; Chazarain and Piriou 1982].

In a regular, bounded domain Ω of M, various self-adjoint realizations of $\Delta_{f,h}$ can be considered:

- The Dirichlet realization $\Delta_{f,h}^D(\Omega)$ with domain

$$D(\Delta_{f,h}^D(\Omega)) = \Big\{\omega \in \bigwedge W^{2,2}(\Omega) : \boldsymbol{t}\omega\big|_{\partial\Omega} = 0,\ \boldsymbol{t} d_{f,h}^*\omega\big|_{\partial\Omega} = 0\Big\}.$$

This is the Friedrichs extension of the closed quadratic form

$$\mathcal{D}(\omega, \omega') = \langle d_{f,h}\omega, d_{f,h}\omega'\rangle_{L^2} + \langle d_{f,h}^*\omega, d_{f,h}^*\omega'\rangle_{L^2} \tag{2-4}$$

defined on the domain

$$\bigwedge W_D^{1,2}(\Omega) = \Big\{\omega \in \bigwedge W^{1,2}(\Omega) : \boldsymbol{t}\omega\big|_{\partial\Omega} = 0\Big\}.$$

Its restriction to zero-forms (functions) is simply the operator (2-1) on Ω with homogeneous Dirichlet boundary conditions. It is associated with the stochastic process (1-1) killed at the boundary.

- The Neumann realization $\Delta_{f,h}^N(\Omega)$ with domain

$$D(\Delta_{f,h}^N)(\Omega) = \Big\{\omega \in \bigwedge W^{2,2}(\Omega) : \boldsymbol{n}\omega\big|_{\partial\Omega} = 0,\ \boldsymbol{n} d_{f,h}\omega\big|_{\partial\Omega} = 0\Big\}.$$

This is the Friedrichs extension of the closed quadratic form (2-4) defined on the domain

$$\bigwedge W_N^{1,2}(\Omega) = \Big\{\omega \in \bigwedge W^{1,2}(\Omega) : \boldsymbol{n}\omega\big|_{\partial\Omega} = 0\Big\}.$$

Its restriction to zero-forms (functions) is simply the operator (2-1) on Ω with homogeneous Neumann boundary conditions. It is associated with the stochastic process (1-1) reflected at the boundary.

We will handle exponentially small quantities and we shall use the following notation, which is convenient when comparing them.

Definition 2.3. Let $(E, \|\ \|)$ be a normed space. For two functions $a : \mathbb{R}_+ \to E$ and $b : \mathbb{R}_+ \to \mathbb{R}_+$, we write:

- $a(h) = \mathcal{O}(b(h))$ if there exist $h_0 > 0$ and $C > 0$ such that $\|a(h)\| \le Cb(h)$ for all $h \in (0, h_0)$;
- $a(h) = \tilde{\mathcal{O}}(b(h))$ if, for every $\varepsilon > 0$, $a(h) = \mathcal{O}(b(h)e^{\varepsilon/h})$, or, equivalently,

$$\forall\varepsilon > 0\ \ \exists h_0 > 0\ \ \exists C > 0\ \ \forall h \in (0, h_0) \quad \|a(h)\| \le Cb(h)e^{\varepsilon/h}.$$

Notice that $a(h) = \tilde{\mathcal{O}}(b(h))$ is equivalent to $\limsup_{h\to 0} h \log(\|a(h)\|/b(h)) \le 0$. Note in particular the identity $\mathcal{O}(e^{-c_1/h})\tilde{\mathcal{O}}(e^{-c_2/h}) = \tilde{\mathcal{O}}(e^{-(c_1+c_2)/h}) = \mathcal{O}(e^{-c'/h})$ for any fixed $c' < c_1 + c_2$, independently of $h \in (0, h_0)$.

We are now in position to state the two additional hypotheses on f, which are stated as assumptions on the eigenvalues of Witten Laplacians on Ω_- and $\Omega_+ \setminus \overline{\Omega}_-$. We assume that there exist a constant $c_0 > 0$ and a function $\nu : (0, h_0) \to (0, +\infty)$ with

$$\forall\varepsilon > 0\ \ \exists C_\varepsilon > 1 \quad \frac{1}{C_\varepsilon}e^{-\varepsilon/h} \le \nu(h) \le h, \tag{2-5}$$

or, equivalently,

$$\log\left(\frac{\nu(h)}{h}\right) \le 0 \quad \text{and} \quad \lim_{h\to 0} h\log(\nu(h)) = 0,$$

and such that the following hypotheses are fulfilled:

Hypothesis 3. *The Neumann Witten Laplacian defined on Ω_- and restricted to forms of degree* 0 *and* 1*, $\Delta_{f,h}^{N,(p)}(\Omega_-)$, $p = 0, 1$, satisfies*

$$\#\big[\sigma(\Delta_{f,h}^{N,(p)}(\Omega_-)) \cap [0, \nu(h)]\big] =: m_p^N(\Omega_-)\,, \tag{2-6}$$

$$\sigma(\Delta_{f,h}^{N,(p)}(\Omega_-)) \cap [0, \nu(h)] \subset [0, e^{-c_0/h}] \tag{2-7}$$

with $m_p^N(\Omega_-)$ independent of $h \in (0, h_0)$. Throughout, eigenvalues are counted with multiplicity, and the symbol # denotes the cardinal of a finite ensemble.

In addition, there exists in $\overline{\Omega}_-$ an open neighborhood $\mathcal{V}_-$ of $\partial\Omega_-$ such that any eigenfunction $\psi(h)$ of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ associated with a small nonzero eigenvalue $\mu(h)$ (namely $0 < \mu(h) \le \nu(h)$) satisfies

$$\|\psi(h)\|_{L^2(\mathcal{V}_-)} = \tilde{\mathcal{O}}(\sqrt{\mu(h)}). \tag{2-8}$$

Hypothesis 4. *The Dirichlet Witten Laplacian on* $\Omega_+ \setminus \bar{\Omega}_-$ *restricted to one-forms satisfies*

$$\#\big[\sigma(\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \bar{\Omega}_-)) \cap [0, \nu(h)]\big] =: m_1^D(\Omega_+ \setminus \bar{\Omega}_-)\,, \tag{2-9}$$

$$\sigma(\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \bar{\Omega}_-)) \cap [0, \nu(h)] \subset [0, e^{-c_0/h}] \tag{2-10}$$

with $m_1^D(\Omega_+ \setminus \bar{\Omega}_-)$ *independent of* $h \in (0, h_0)$.

Our main results concern the smallest eigenvalue as well as properties of the associated eigenfunction of $\Delta_{f,h}^{D,(0)}(\Omega_+)$.

Theorem 2.4. *Assume Hypotheses 1, 2, 3, 4 and that* $h \in (0, h_0)$ *with* $h_0 > 0$ *small enough. The eigenvalues contained in* $[0, \nu(h)]$ *of the Dirichlet Witten Laplacians* $\Delta_{f,h}^{D,(p)}(\Omega_+)$ *for* $p = 0, 1$, *satisfy*:

$$m_0^D(\Omega_+) := \#\big[\sigma(\Delta_{f,h}^{D,(0)}(\Omega_+)) \cap [0, \nu(h)]\big] = m_0^N(\Omega_-),$$
$$m_1^D(\Omega_+) := \#\big[\sigma(\Delta_{f,h}^{D,(1)}(\Omega_+)) \cap [0, \nu(h)]\big] = m_1^N(\Omega_-) + m_1^D(\Omega_+ \setminus \bar{\Omega}_-),$$
$$\sigma(\Delta_{f,h}^{D,(p)}(\Omega_+)) \cap [0, \nu(h)] \subset [0, e^{-c/h}].$$

Let $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar{\Omega}_-)}$ *be an orthonormal basis of the spectral subspace* $\operatorname{Ran} 1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \bar{\Omega}_-))$ *and set*

$$\kappa_f = \min_{\partial\Omega_+} f - \min_{\Omega_+} f.$$

The smallest eigenvalue of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ *satisfies, in the limit* $h \to 0$,

$$\lim_{h\to 0} h \log \lambda_1^{(0)}(\Omega_+) = -2\kappa_f, \tag{2-11}$$

$$\lambda_1^{(0)}(\Omega_+) = \frac{h^2 \sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \big|\int_{\partial\Omega_+} e^{-f/h} u_k^{(1)}(n)(\sigma)\, d\sigma\big|^2}{\int_{\Omega_+} e^{-2f(x)/h}\, dx}(1 + \mathcal{O}(e^{-c/h})) \tag{2-12}$$

for some constant $c > 0$ *and* $u_k^{(1)}(n)(\sigma) = \boldsymbol{i}_n u_k^{(1)}(\sigma)$ *with the interior product notation* (A-1). *Moreover, the nonnegative* $L^2(\Omega_+)$*-normalized eigenfunction* $u_1^{(0)}$ *satisfies*

$$\bigg\| u_1^{(0)} - \frac{e^{-f/h}}{\big(\int_{\Omega_+} e^{-2f(x)/h}\, dx\big)^{1/2}} \bigg\|_{W^{2,2}(\Omega_+)} = \mathcal{O}(e^{-c/h}), \tag{2-13}$$

$$\bigg\| d_{f,h} u_1^{(0)} + \sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \frac{h \int_{\partial\Omega_+} e^{-f(\sigma)/h} u_k^{(1)}(n)(\sigma)\, d\sigma}{\big(\int_{\Omega_+} e^{-2f(x)/h}\, dx\big)^{1/2}} u_k^{(1)} \bigg\|_{W^{p,2}(\mathcal{V})} = \mathcal{O}(e^{-(\kappa_f + c_{\mathcal{V}})/h}) \tag{2-14}$$

for all $p \in \mathbb{N}$, *where* $\mathcal{V}$ *is any neighborhood of* $\partial\Omega_+$ *lying in* $\Omega_+ \setminus \bar{\Omega}_-$ *and* $c_{\mathcal{V}} > 0$ *is a constant independent of* p *and* h. *The symbols* $d\sigma$ *and* $n(\sigma)$, *respectively, denote the infinitesimal volume on* $\partial\Omega_+$ *and the outward normal vector at* $\sigma \in \partial\Omega_+$.

We would like to stress again that Theorem 2.4 does not require f to be a Morse function on Ω_+, nor on $\partial\Omega_+$.

Remark 2.5. It would be interesting for practical applications to relax the assumption $|\nabla f| > 0$ on $\bar{\Omega}_+ \setminus \Omega_-$ in Hypothesis 1 in order to be able to consider saddle points on $\partial\Omega_+$.

Remark 2.6. While proving these results, we will actually show that, necessarily, $m_1^D(\Omega_+ \setminus \bar{\Omega}_-) \neq 0$; see Remark 5.6 below.

Remark 2.7. All the terms in the sum in (2-14) are exponentially small, but at least one is larger than the remainder $\mathcal{O}(e^{-(\kappa_f + c_{\mathcal{V}})/h})$ (see (5-10) and Proposition 6.4). The number of terms which are indeed larger than the remainder depends on the precise value of $c_{\mathcal{V}}$, which depends on the geometry, the global topology of the domain and the function f (the possibility of several terms is discussed in Remarks 7.7 and 7.9 after Proposition 7.5). In particular, if f is a Morse function, the heights of the generalized critical points of index 2 along $\partial\Omega_+$ play a role.

Remark 2.8. In spectral theory, it is natural to work with complex-valued functions or complex-valued forms. In view of the probabilistic interpretation of our results, the above result is stated — and, actually, most of the analysis of this text is carried out — with real-valued functions or forms. One exception is Section 4A, which requires functional calculus and resolvents for complex spectral parameters. Notice that it is straightforward to write a complex-valued version of the previous results, by replacing the real scalar product by the hermitian scalar product. For example, in (2-14), this simply consists in changing $\int_{\partial\Omega_+} e^{-f(\sigma)/h} u_k^{(1)}(n)(\sigma)\, d\sigma$ to $\int_{\partial\Omega_+} e^{-f(\sigma)/h} \overline{u_k^{(1)}}(n)(\sigma)\, d\sigma$.

Note that the numerators in the estimates (2-12) and (2-14) of the eigenvalue $\lambda_1^{(0)}(\Omega_+)$ and of $d_{f,h} u_1^{(0)}$ depend only on the values of f and the geometry of Ω_+ around $\partial\Omega_+$. More precisely, they do not change when f is modified inside Ω_-. This allows us to understand the variations of $\lambda_1^{(0)}(\Omega_+)$ and $\partial_n u_1^{(0)}\big|_{\partial\Omega_+}$ with respect to f, which is needed in the hyperdynamics algorithm (see Section 1A).

Corollary 2.9. *Let f_1 and f_2 be two functions which fulfill Hypotheses 1, 2, 3 and 4. Let $\lambda_1^{(0)}(f_1)$ be the first eigenvalue of $\Delta_{f_1,h}^{D,(0)}(\Omega_+)$ associated with the nonnegative normalized eigenvector $u_1^{(0)}(f_1)$, and $\lambda_1^{(0)}(f_2)$ the first eigenvalue of $\Delta_{f_2,h}^{D,(0)}(\Omega_+)$ associated with the eigenvector $u_1^{(0)}(f_2)$. Assume additionally $f_1 = f_2$ in $\Omega_+ \setminus \bar{\Omega}_-$. The quantities $\lambda_1^{(0)}(f_{1,2})$ and $\partial_n[e^{-f_{1,2}/h} u_1^{(0)}(f_{1,2})]\big|_{\partial\Omega_+} = e^{-f_{1,2}/h}[\partial_n u_1^{(0)}(f_{1,2})]\big|_{\partial\Omega_+}$ satisfy*

$$\frac{\lambda_1^{(0)}(f_2)}{\lambda_1^{(0)}(f_1)} = \frac{\int_{\Omega_+} e^{-2f_1(x)/h}\, dx}{\int_{\Omega_+} e^{-2f_2(x)/h}\, dx}(1 + \mathcal{O}(e^{-c/h})), \tag{2-15}$$

$$\frac{\partial_n[e^{-f_2/h} u_1^{(0)}(f_2)]\big|_{\partial\Omega_+}}{\|\partial_n[e^{-f_2/h} u_1^{(0)}(f_2)]\|_{L^1(\partial\Omega_+)}} = \frac{\partial_n[e^{-f_1/h} u_1^{(0)}(f_1)]\big|_{\partial\Omega_+}}{\|\partial_n[e^{-f_1/h} u_1^{(0)}(f_1)]\|_{L^1(\partial\Omega_+)}} + \mathcal{O}(e^{-c/h}) \quad \textit{in } L^1(\partial\Omega_+). \tag{2-16}$$

Other corollaries and variations of Theorem 2.4 are given in Section 6. Among the consequences, one can prove the following result when, additionally, $f\big|_{\partial\Omega_+}$ is a Morse function and $\partial_n f > 0$ on $\partial\Omega_+$.

Theorem 2.10. *Assume Hypotheses 1, 2, 3 and 4 and $h \in (0, h_0)$ with $h_0 > 0$ small enough. Assume moreover that $f\big|_{\partial\Omega_+}$ is a Morse function and $\partial_n f > 0$ on $\partial\Omega_+$. Then the first eigenvalue $\lambda_1^{(0)}(\Omega_+)$ of*

$\Delta^{D,(0)}_{f,h}(\Omega_+)$ *and the corresponding* $L^2(\Omega_+)$*-normalized nonnegative eigenfunction* $u_1^{(0)}$ *satisfy*

$$\lambda_1^{(0)}(\Omega_+) = \frac{\int_{\partial\Omega_+} 2\partial_n f(\sigma) e^{-2f(\sigma)/h}\, d\sigma}{\int_{\Omega_+} e^{-2f(x)/h}\, dx}(1+\mathcal{O}(h)), \tag{2-17}$$

$$-\frac{\partial_n[e^{-f/h}u_1^{(0)}]\big|_{\partial\Omega_+}}{\|\partial_n[e^{-f/h}u_1^{(0)}]\|_{L^1(\partial\Omega_+)}} = \frac{(2\partial_n f)e^{-2f/h}\big|_{\partial\Omega_+}}{\|(2\partial_n f)e^{-2f/h}\|_{L^1(\partial\Omega_+)}} + \mathcal{O}(h) \quad \text{in } L^1(\partial\Omega_+). \tag{2-18}$$

The proof of Theorem 2.4 is given in Proposition 3.12, Lemma 5.9, Proposition 6.1 and Proposition 6.8. The proof of Corollary 2.9 is given in Section 6D. The proof of Theorem 2.10 is given in Section 7A2.

3. A priori exponential decay and first consequences

By applying Agmon's type estimate (see, for example, [Helffer 1988; Dimassi and Sjöstrand 1999] for a general introduction) for boundary Witten Laplacians, we give here exponential decay estimates for the eigenvectors of $\Delta^N_{f,h}(\Omega_-)$, $\Delta^D_{f,h}(\Omega_+\setminus\overline{\Omega}_-)$ and $\Delta^D_{f,h}(\Omega_+)$.

3A. ***Agmon identity.*** We shall use an identity for boundary Witten Laplacians, proved in [Helffer and Nier 2006] in the Dirichlet case and in [Le Peutrec 2010b] in the Neumann case.

Lemma 3.1. *Let* Ω *be a regular bounded domain of* (M,g) *and let* $\Delta^D_{f,h}(\Omega)$ (*resp.* $\Delta^N_{f,h}(\Omega)$) *be the Dirichlet* (*resp. Neumann*) *realization of* $\Delta_{f,h}(\Omega)$. *Let* φ *be a real-valued Lipschitz function on* $\overline{\Omega}$. *Then, for any real-valued* $\omega\in D(\Delta^D_{f,h}(\Omega))$ (*resp.* $\omega\in D(\Delta^N_{f,h}(\Omega))$),

$$\begin{aligned}\langle\omega, e^{2\varphi/h}\Delta^D_{f,h}(\Omega)\omega\rangle_{L^2(\Omega)} &= h^2\|de^{\varphi/h}\omega\|^2_{L^2(\Omega)} + h^2\|d^*e^{\varphi/h}\omega\|^2_{L^2(\Omega)} + \langle(|\nabla f|^2-|\nabla\varphi|^2+h\mathcal{L}_{\nabla f}+h\mathcal{L}^*_{\nabla f})e^{\varphi/h}\omega, e^{\varphi/h}\omega\rangle_{L^2(\Omega)}\\ &\quad - h\int_{\partial\Omega}\langle\omega,\omega\rangle_{T^*_\sigma\Omega}e^{2\varphi(\sigma)/h}\frac{\partial f}{\partial n}(\sigma)\,d\sigma.\end{aligned}$$

$$\begin{aligned}\langle\omega, e^{2\varphi/h}\Delta^N_{f,h}(\Omega)\omega\rangle_{L^2(\Omega)} &= h^2\|de^{\varphi/h}\omega\|^2_{L^2(\Omega)} + h^2\|d^*e^{\varphi/h}\omega\|^2_{L^2(\Omega)} + \langle(|\nabla f|^2-|\nabla\varphi|^2+h\mathcal{L}_{\nabla f}+h\mathcal{L}^*_{\nabla f})e^{\varphi/h}\omega, e^{\varphi/h}\omega\rangle_{L^2(\Omega)}\\ &\quad + h\int_{\partial\Omega}\langle\omega,\omega\rangle_{T^*_\sigma\Omega}e^{2\varphi(\sigma)/h}\frac{\partial f}{\partial n}(\sigma)\,d\sigma.\end{aligned}$$

In the previous formulas, the notation $\mathcal{L}_X$ refers to the Lie derivative; see (A-2). We shall use this lemma with specific functions φ associated with the metric $|\nabla f|^2 g$.

Lemma 3.2. *Let* Ω *be an open subset of* M, $f\in\mathcal{C}^\infty(\overline{\Omega})$, *and let* d_{Ag} *be the geodesic pseudodistance on* $\overline{\Omega}$ *associated with the possibly degenerate metric* $|\nabla f|^2 g$. *The function* $(x,y)\mapsto d_{Ag}(x,y)$ *is Lipschitz* (*and thus almost everywhere differentiable*) *and satisfies*

$$\begin{aligned}&|\nabla_x d_{Ag}(x,y_0)|\le|\nabla f(x)| \quad \text{for all } y_0\in\overline{\Omega} \text{ and for a.e. } x\in\Omega,\\ &|f(x)-f(y)|\le d_{Ag}(x,y) \quad \text{for all } x,y\in\overline{\Omega}.\end{aligned} \tag{3-1}$$

The equality $d_{Ag}(x,y)=|f(x)-f(y)|$ occurs if there is an integral curve of ∇f joining x to y. Moreover, for any $A\subset\overline{\Omega}$, the function $x\mapsto d_{Ag}(x,A)$ (where $d_{Ag}(x,A)=\inf_{a\in A}d_{Ag}(x,a)$) is Lipschitz and satisfies

$$|\nabla_x d_{Ag}(x,A)|\le|\nabla f(x)|\quad \textit{for a.e. } x\in\Omega.$$

Proof. The Lipschitz property comes from the triangular inequality for $d_{\mathrm{Ag}}(x,y)$. It carries over to $d_{\mathrm{Ag}}(x,A)$. The comparison between $|f(x)-f(y)|$ and $d_{\mathrm{Ag}}(x,y)$ comes from

$$|f(x)-f(y)|=\left|\int_0^1\nabla f(\gamma(t))\cdot\dot{\gamma}(t)\,dt\right|\le\int_0^1|\nabla f(\gamma(t))|\,|\dot{\gamma}(t)|\,dt=|\gamma|_{\mathrm{Ag}}$$

for any $\mathcal{C}^1$-path γ joining x to y and denoting by $|\gamma|_{\mathrm{Ag}}$ its length according to d_{Ag}. □

Remark 3.3. A detailed discussion about the equality $d_{\mathrm{Ag}}(x,y)=|f(x)-f(y)|$ when f is a Morse function, which involves the notion of generalized integral curves of ∇f, can be found in [Helffer and Sjöstrand 1985b].

3B. ***Exponential decay for the eigenvectors of $\Delta^{N,(p)}_{f,h}(\Omega_-)$ ($p=0,1$).*** Notice that, from Hypothesis 1, there exists an open set U such that

$$\overline{U}\subset\Omega_-\qquad\text{and}\qquad|\nabla f|\neq0\quad\text{in }\overline{\Omega}_-\setminus U.\tag{3-2}$$

The following proposition will be useful to prove that all the eigenvectors of $\Delta^{N,(p)}_{f,h}$ are exponentially small in the neighborhood of $\partial\Omega_-$ (see Proposition 3.5). It actually holds for any open set $U\subset\Omega_-$ which contains all the critical points, without the additional requirement $\overline{U}\subset\Omega_-$.

Proposition 3.4. *Let U be an open subset of Ω_- such that $|\nabla f|\neq0$ in $\overline{\Omega}_-\setminus U$ and let $d_{Ag}(x,U)$ be the Agmon distance to U defined for $x\in\Omega_-$. There exists a constant $C>0$ independent of $h\in[0,h_0]$ such that every normalized eigenvector ω_{λ_h} of $\Delta^N_{f,h}(\Omega_-)$ associated with an eigenvalue $\lambda_h\in[0,\nu(h)]$ satisfies*

$$\|e^{d_{Ag}(\cdot,U)/h}\omega_{\lambda_h}\|_{L^2(\Omega_-\setminus U)}\le\|e^{d_{Ag}(\cdot,U)/h}\omega_{\lambda_h}\|_{L^2(\Omega_-)}\le C,$$
$$\|e^{d_{Ag}(\cdot,U)/h}\omega_{\lambda_h}\|_{W^{1,2}(\Omega_-\setminus U)}\le\|e^{d_{Ag}(\cdot,U)/h}\omega_{\lambda_h}\|_{W^{1,2}(\Omega_-)}\le\frac{C}{h^{1/2}}.$$

Proof. The function $d_{\mathrm{Ag}}(\cdot,U)$ vanishes in $\overline{U}$ and satisfies the properties of Lemma 3.2 with $(\Omega,A)=(\Omega_-,\overline{U})$. Let us now apply Lemma 3.1 on $\Delta^N_{f,h}(\Omega_-)$ with the function $\varphi=(1-\alpha h)d_{\mathrm{Ag}}(\cdot,U)$ (where α is a positive constant to be fixed later on) and a normalized eigenvector ω: $\Delta^N_{f,h}(\Omega_-)\omega=\lambda\omega$, where $\lambda\in[0,\nu(h)]$. With $\partial f/\partial n>0$ on $\partial\Omega_-$, $\nu(h)\le h$ and $|\nabla\varphi|^2\le(1-\alpha h)|\nabla f|^2$ (for $h<1/\alpha$), we obtain

$$0\ge h^2\|de^{\varphi/h}\omega\|^2_{L^2(\Omega_-)}+h^2\|d^*e^{\varphi/h}\omega\|^2_{L^2(\Omega_-)}+h[\alpha\langle e^{\varphi/h}\omega,|\nabla f|^2e^{\varphi/h}\omega\rangle_{L^2(\Omega_-)}-C_f\|e^{\varphi/h}\omega\|^2_{L^2(\Omega_-)}].\tag{3-3}$$

Here, we have used the fact that, for any vector field X, $\mathcal{L}_X+\mathcal{L}_X^*$ is a differential operator of order 0 involving derivatives of X and g that are uniformly bounded in $\overline{\Omega}_-$.

Using (3-2), choose α such that $\alpha\min_{x\in\overline{\Omega}_-\setminus U}|\nabla f(x)|^2\ge2C_f$ and add $2C_fh\|e^{\varphi/h}\omega\|^2_{L^2(U)}$ on both sides of the inequality (3-3). Using the fact that

$$2C_fh\ge2C_fh\|\omega\|^2_{L^2(U)}=2C_fh\|e^{\varphi/h}\omega\|^2_{L^2(U)},$$

one obtains

$$2C_f h \geq h^2 \|de^{\varphi/h}\omega\|^2_{L^2(\Omega_-)} + h^2 \|d^* e^{\varphi/h}\omega\|^2_{L^2(\Omega_-)} + C_f h \|e^{\varphi/h}\omega\|^2_{L^2(\Omega_-)}.$$

This implies $\|e^{(1-\alpha h) d_{\mathrm{Ag}}(\cdot, U)/h}\omega\|^2_{L^2(\Omega_-)} \leq 2$ and

$$\|(hd) e^{(1-\alpha h) d_{\mathrm{Ag}}(\cdot, U)/h}\omega\|^2_{L^2(\Omega_-)} + \|(hd)^* e^{(1-\alpha h) d_{\mathrm{Ag}}(\cdot, U)/h}\omega\|^2_{L^2(\Omega_-)} \leq 2C_f h.$$

Since $d_{\mathrm{Ag}}(\cdot, U)$ is a Lipschitz (and thus also bounded) function on $\overline{\Omega}_-$, this ends the proof. □

Here is a useful consequence of Proposition 3.4:

Proposition 3.5. *Let $(\psi_j^{(0)})_{1\leq j\leq m_0^N(\Omega_-)}$ (resp. $(\psi_k^{(1)})_{1\leq k\leq m_1^N(\Omega_-)}$) be an orthonormal basis of eigenvectors of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ (resp. $\Delta_{f,h}^{N,(1)}(\Omega_-)$) associated with the eigenvalues lying in $[0, \nu(h)]$ (or, owing to Hypothesis 3, in $[0, e^{-c_0/h}]$). Let $U \subset \Omega_-$ be an open set satisfying (3-2). Let $\chi_- \in \mathcal{C}_0^\infty(\Omega_-)$ be a cut-off function such that $0 \leq \chi_- \leq 1$ and $\chi_- \equiv 1$ on a neighborhood of U. The functions $v_j^{(0)} = \chi_- \psi_j^{(0)}$, $1 \leq j \leq m_0^N(\Omega_-)$ (resp. one-forms $v_k^{(1)} = \chi_- \psi_k^{(1)}$, $1 \leq k \leq m_1^N(\Omega_-)$) belong to the domain $D(\Delta_{f,h}^{D,(0)}(\Omega_+))$ (resp. $D(\Delta_{f,h}^{D,(1)}(\Omega_+))$) of the Dirichlet realization of $\Delta_{f,h}$ in Ω_+ and they satisfy: for $h \in [0, h_0]$,*

$$\sum_{j=1}^{m_0^N(\Omega_-)} \|\psi_j^{(0)} - v_j^{(0)}\|_{W^{1,2}(\Omega_-)} + \sum_{k=1}^{m_1^N(\Omega_-)} \|\psi_k^{(1)} - v_k^{(1)}\|_{W^{1,2}(\Omega_-)} = \mathcal{O}(e^{-c_{\chi_-}/h}),$$

$$(\langle v_j^{(0)}, v_{j'}^{(0)}\rangle_{L^2(\Omega_+)})_{j,j'} = \mathrm{Id}_{m_0^N(\Omega_-)} + \mathcal{O}(e^{-c_{\chi_-}/h}), \quad (\langle v_k^{(1)}, v_{k'}^{(1)}\rangle_{L^2(\Omega_+)})_{k,k'} = \mathrm{Id}_{m_1^N(\Omega_-)} + \mathcal{O}(e^{-c_{\chi_-}/h}),$$

$$\langle v_j^{(0)}, \Delta_{f,h}^{D,(0)}(\Omega_+) v_j^{(0)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c_{\chi_-}/h}), \quad \langle v_k^{(1)}, \Delta_{f,h}^{D,(1)}(\Omega_+) v_k^{(1)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c_{\chi_-}/h}),$$

where the $\mathcal{O}(e^{-c_{\chi_-}/h})$ remainders can be bounded from above by $C_{\chi_-} e^{-c_{\chi_-}/h}$ for some constants C_{χ_-}, $c_{\chi_-} > 0$ independent of $h \in [0, h_0]$. Throughout, Id_m denotes the identity matrix of size $m \times m$.

Proof. Let ψ be a $L^2(\Omega_-)$-normalized eigenvector of $\Delta_{f,h}^{N,(p)}(\Omega_-)$, $p = 0, 1$, associated with the eigenvalue $\lambda = \mathcal{O}(e^{-c_0/h})$, and set $v = \chi_- \psi$. Since χ_- belongs to $\mathcal{C}_0^\infty(\Omega_-)$ the form $v = \chi_- \psi$ belongs to $D(\Delta_{f,h}^{D,(p)}(\Omega_+))$.

The $W^{1,2}(\Omega_-)$ estimates as well as the result on the Gram matrices are consequences of

$$\|\psi - v\|_{W^{1,2}(\Omega_-)} = \|(1-\chi_-)\psi\|_{W^{1,2}(\Omega_-)} \leq \|\psi\|_{W^{1,2}(\Omega_- \setminus \{\chi_- = 1\})} \leq C'_{\chi_-} e^{-c'_{\chi_-}/h} \tag{3-4}$$

for some constants $c'_{\chi_-} > 0$ and $C'_{\chi_-} > 0$. The estimate (3-4) is derived from Proposition 3.4 by using the fact that there exists $c > 0$ such that $d_{\mathrm{Ag}}(x, U) \geq c$ for all $x \in \Omega_- \setminus \{\chi_- = 1\}$ (this is a consequence of (3-2)).

For the last estimate of Proposition 3.5, we use Lemma 3.1 with $\varphi = 0$. Considering first the estimate on $\Delta_{f,h}^D$ with $\Omega = \Omega_+$, $\omega = v = \chi_- \psi$ and then the estimate on $\Delta_{f,h}^N$ with $\Omega = \Omega_-$, $\omega = \psi$, one obtains

$$\langle \chi_- \psi, \Delta_{f,h}^D(\Omega_+) \chi_- \psi\rangle_{L^2(\Omega_+)}$$
$$= h^2 \|d\chi_- \psi\|^2_{L^2(\Omega_+)} + h^2 \|d^* \chi_- \psi\|^2_{L^2(\Omega_+)} + \langle (|\nabla f|^2 + h\mathcal{L}_{\nabla f} + h\mathcal{L}^*_{\nabla f})\chi_- \psi, \chi_- \psi\rangle_{L^2(\Omega_+)} + 0 \tag{3-5}$$

and (since $\partial f/\partial n > 0$ on $\partial\Omega_-$)

$$e^{-c_0/h} \geq \lambda \geq h^2\|d\psi\|^2_{L^2(\Omega_-)} + h^2\|d^*\psi\|^2_{L^2(\Omega_-)} + \langle(|\nabla f|^2 + h\mathcal{L}_{\nabla f} + h\mathcal{L}^*_{\nabla f})\psi, \psi\rangle_{L^2(\Omega_-)}. \tag{3-6}$$

By considering the difference between (3-5) and (3-6), we thus have

$$\begin{aligned}\langle \chi_-\psi, \Delta^D_{f,h}(\Omega_+)\chi_-\psi\rangle_{L^2(\Omega_+)} &\\ \leq e^{-c_0/h} + h^2(\|d\chi_-\psi\|^2_{L^2(\Omega_+)} - \|d\psi\|^2_{L^2(\Omega_-)}) &+ h^2(\|d^*\chi_-\psi\|^2_{L^2(\Omega_+)} - \|d^*\psi\|^2_{L^2(\Omega_-)})\\ + \big(\langle(|\nabla f|^2 + h\mathcal{L}_{\nabla f} + h\mathcal{L}^*_{\nabla f})\chi_-\psi, \chi_-\psi\rangle_{L^2(\Omega_+)} &- \langle(|\nabla f|^2 + h\mathcal{L}_{\nabla f} + h\mathcal{L}^*_{\nabla f})\psi, \psi\rangle_{L^2(\Omega_-)}\big).\end{aligned}$$

The last three terms in the right-hand side are all of order $\mathcal{O}(e^{-c_{\chi_-}/h})$. Indeed, for the first term (the two other terms are estimated in the same way),

$$\begin{aligned}\big|\|d\chi_-\psi\|^2_{L^2(\Omega_+)} - \|d\psi\|^2_{L^2(\Omega_-)}\big| &= |\langle d(1-\chi_-)\psi, d(1+\chi_-)\psi\rangle_{L^2(\Omega_-)}|\\ &\leq C''_{\chi_-}\|\psi\|^2_{W^{1,2}(\Omega_-\setminus\{\chi_-=1\})} \leq C^{(3)}_{\chi_-}e^{-2c'_{\chi_-}/h},\end{aligned}$$

using again (3-4). This proves the last estimate. □

According to the terminology of [Le Peutrec 2009], the property on the Gram matrices in Proposition 3.5 is equivalent to the almost orthonormality of the family $(v_j^{(p)})_{1\leq j\leq m_p^N(\Omega_-)}$, $p = 0, 1$, in $L^2(\Omega_+)$.

Definition 3.6. A finite family of h-dependent vectors $(u_k^h)_{1\leq k\leq N}$ in a Hilbert space $\mathcal{H}$ is almost orthonormal if the Gram matrix satisfies

$$(\langle u_j^h, u_k^h\rangle)_{1\leq j,k\leq N} = \mathrm{Id}_N + \mathcal{O}(e^{-c/h})$$

for some $c > 0$ independent of h.

We end this subsection with some remarks on the spectrum of $\Delta^{N,(0)}_{f,h}(\Omega_-)$, which we denote (as usual, in increasing order and with multiplicity) by $(\mu_k^{(0)}(\Omega_-))_{k\geq 1}$. The first eigenvalue of $\Delta^{N,(0)}_{f,h}(\Omega_-)$ is $\mu_1^{(0)}(\Omega_-) = 0$ associated with the eigenvector

$$\psi_1^{(0)} = \frac{e^{-f/h}}{\big(\int_{\Omega_-} e^{-2f(x)/h}\,dx\big)^{1/2}}.$$

One can prove that the second eigenvalue $\mu_2^{(0)}(\Omega_-)$ of $\Delta^{N,(0)}_{f,h}(\Omega_-)$ is exponentially large compared to $e^{-2\kappa_f/h}$, where we recall $\kappa_f = \min_{\partial\Omega_+} f - \min_{\Omega_+} f = \min_{\partial\Omega_+} f - \min_{\Omega_-} f$.

Proposition 3.7. *Let* cvmax *be the maximum critical value of f in Ω_-:*

$$\mathrm{cvmax} = \max\{f(x) : x \in \Omega_-, \nabla f(x) = 0\}.$$

Then the second eigenvalue $\mu_2^{(0)}(\Omega_-)$ of $\Delta^{N,(0)}_{f,h}(\Omega_-)$ satisfies

$$\liminf_{h\to 0} h\log(\mu_2^{(0)}(\Omega_-)) \geq -2(\mathrm{cvmax} - \min_{\Omega_-} f) \geq -2\kappa_f + 2c_0,$$

where c_0 denotes the positive constant used in Hypothesis 2.

Proof. The second inequality $-2(\mathrm{cvmax}-\min_{\Omega_-} f) \geq -2\kappa_f + 2c_0$ is of course a consequence of Hypothesis 2. To prove the first inequality, let us reason by contradiction and assume that there exists $\varepsilon_0 > 0$ and a sequence h_n such that $\lim_{n\to\infty} h_n = 0$ and

$$\min\{\sigma(\Delta^{N,(0)}_{f,h_n}(\Omega_-)) \setminus \{0\}\} \leq C e^{-2(\mathrm{cvmax}-\min_{\Omega_-} f+\varepsilon_0)/h_n}.$$

To simplify the notation, let us drop the subscript n in h_n. Let $\psi_2^{(0)}$ be a normalized eigenfunction of $\Delta^{N,(0)}_{f,h}(\Omega_-)$ associated with $\mu_2^{(0)}(\Omega_-) > 0$. It is orthogonal to $\psi_1^{(0)}$ in $L^2(\Omega_-)$ and it satisfies: for any $\Omega \subset \Omega_-$,

$$\|d_{f,h}\psi_2^{(0)}\|^2_{L^2(\Omega)} \leq \|d_{f,h}\psi_2^{(0)}\|^2_{L^2(\Omega_-)} = \langle \psi_2^{(0)}, \Delta^{N,(0)}_{f,h}(\Omega_-)\psi_2^{(0)}\rangle_{L^2(\Omega_-)} = \mu_2^{(0)} \leq C e^{-2(\mathrm{cvmax}-\min_{\Omega_-} f+\varepsilon_0)/h}.$$

In particular, for $\Omega = \{x \in \Omega_- : f(x) < \mathrm{cvmax} + \frac{1}{2}\varepsilon_0\}$, this gives

$$\begin{aligned}\|d(e^{(f-\min_{\Omega_-} f)/h}\psi_2^{(0)})\|^2_{L^2(\Omega)} &\leq h^{-2}\max_{x\in\Omega}|e^{(f(x)-\min_{\Omega_-} f)/h}|^2 \|d_{f,h}\psi_2^{(0)}\|^2_{L^2(\Omega)} \\ &\leq Ch^{-2}e^{-2(\mathrm{cvmax}-\min_{\Omega_-} f+\varepsilon_0)/h}\max_{x\in\Omega}|e^{(f(x)-\min_{\Omega_-} f)/h}|^2 \leq C' e^{-\varepsilon_0/h}.\end{aligned}$$

Using the spectral gap estimate for the Neumann Laplacian in Ω (or equivalently the Poincaré–Wirtinger inequality on Ω), there is a constant C_h (depending on $\psi_2^{(0)}$) such that

$$\|\psi_2^{(0)} - C_h e^{-(f-\min_{\Omega_-} f)/h}\|_{L^2(\Omega)} = \mathcal{O}(e^{-\varepsilon_0/(2h)}).$$

Equivalently, there is a constant C_h such that

$$\|\psi_2^{(0)} - C_h\psi_1^{(0)}\|_{L^2(\Omega)} = \mathcal{O}(e^{-\varepsilon_0/(2h)}). \tag{3-7}$$

Further, using Proposition 3.4 with $U = \{x \in \Omega_- : f(x) < \mathrm{cvmax} + \frac{1}{4}\varepsilon_0\} \subset \Omega$, and a lower bound on $d_{\mathrm{Ag}}(x, U)$ (see (3-4) for a similar argument), one obtains

$$\|\psi_1^{(0)}\|_{L^2(\Omega_-\setminus\Omega)} + \|\psi_2^{(0)}\|_{L^2(\Omega_-\setminus\Omega)} \leq C_{\varepsilon_0} e^{-c_{\varepsilon_0}/h}. \tag{3-8}$$

The two estimates (3-7) and (3-8) contradict the orthogonality of $\psi_2^{(0)}$ and $\psi_1^{(0)}$ in $L^2(\Omega_-)$ in the limit $h \to 0$ (actually $n \to \infty$). □

3C. ***Exponential decay for the eigenvectors of $\Delta^{D,(p)}_{f,h}(\Omega_+ \setminus \bar{\Omega}_-)$.*** In this section, we will check that $\sigma(\Delta^{D,(0)}_{f,h}(\Omega_+ \setminus \bar{\Omega}_-)) \cap [0, \nu(h)] = \varnothing$ and provide the same results as in the previous section for the eigenvectors of $\Delta^{D,(1)}_{f,h}(\Omega_+ \setminus \bar{\Omega}_-)$. Let us start with an equivalent of Proposition 3.4.

Proposition 3.8. *Let $\mathcal{V}$ be a subset of $\bar{\Omega}_+ \setminus \Omega_-$ such that $\partial\Omega_+ \subset \mathcal{V}$ and let $d_{Ag}(x, \mathcal{V})$ be the Agmon distance to $\mathcal{V}$ defined for $x \in \Omega_+ \setminus \bar{\Omega}_-$. There exists a constant $C > 0$ independent of $h \in [0, h_0]$ such that every normalized eigenvector ψ of $\Delta^{D,(1)}_{f,h}(\Omega_+ \setminus \bar{\Omega}_-)$ associated with an eigenvalue $\lambda \in [0, \nu(h)]$ satisfies*

$$\|e^{d_{Ag}(\cdot,\mathcal{V})/h}\psi\|_{W^{1,2}(\Omega_+\setminus\bar{\Omega}_-)} \leq \frac{C}{h}.$$

Proof. The proof follows ideas from [Dimassi and Sjöstrand 1999]. Using Lemma 3.1, the fact that $\lambda \le h$ and the assumption on the sign of the normal derivative of f on $\partial\Omega_-$ stated in Hypothesis 1, we have

$$0 \ge h^2\|de^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)} + h^2\|d^*e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)} + \langle(|\nabla f|^2 - |\nabla\varphi|^2)e^{\varphi/h}\psi,\, e^{\varphi/h}\psi\rangle_{L^2(\Omega_+\setminus\bar\Omega_-)} - hC_f\|e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)} - h\int_{\partial\Omega_+}\langle\psi,\psi\rangle_{\bigwedge T^*_\sigma\Omega_+} e^{2\varphi(\sigma)/h}\frac{\partial f}{\partial n}(\sigma)\,d\sigma. \tag{3-9}$$

Using the trace theorem, there exists a constant $C_{\mathcal{V}}$ such that, for any $\omega \in \bigwedge W^{1,2}(\mathcal{V})$,

$$\int_{\partial\Omega_+}\langle\omega,\omega\rangle_{\bigwedge T^*_\sigma\Omega_+}\,d\sigma \le C_{\mathcal{V}}[\|\omega\|^2_{L^2(\mathcal{V})} + \|\omega\|_{W^{1,2}(\mathcal{V})}\|\omega\|_{L^2(\mathcal{V})}].$$

By applying this inequality to $\omega = e^{\varphi/h}\psi$ and using

$$\|\omega\|^2_{W^{1,2}(\mathcal{V})} \le C_{\mathcal{V}}[\|\omega\|^2_{L^2(\mathcal{V})} + \|d\omega\|^2_{L^2(\mathcal{V})} + \|d^*\omega\|^2_{L^2(\mathcal{V})}],$$

the last term of (3-9) is estimated by

$$\begin{aligned}\left|h\int_{\partial\Omega_+}\langle\psi,\psi\rangle_{\bigwedge T^*_\sigma\Omega_+}e^{2\varphi(\sigma)/h}\frac{\partial f}{\partial n}(\sigma)\,d\sigma\right| &\le \tfrac12 h^2[\|de^{\varphi/h}\psi\|^2_{L^2(\mathcal{V})} + \|d^*e^{\varphi/h}\psi\|^2_{L^2(\mathcal{V})}] + C_{f,\mathcal{V}}\|e^{\varphi/h}\psi\|^2_{L^2(\mathcal{V})}\\ &\le \tfrac12 h^2[\|de^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)} + \|d^*e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)}] + C_{f,\mathcal{V}}\end{aligned}$$

since $\varphi \equiv 0$ on $\mathcal{V}$. Taking $\varphi = (1-\alpha h)d_{\mathrm{Ag}}(x,\mathcal{V})$ in (3-9) gives (using $|\nabla\varphi|^2 \le (1-\alpha h)|\nabla f|^2$ and the inequality $\|e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)} = \|e^{\varphi/h}\psi\|^2_{L^2(\mathcal{V})} + \|e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\overline{\Omega_-\cup\mathcal{V}})} \le C'_{\mathcal{V}} + \|e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\overline{\Omega_-\cup\mathcal{V}})}$)

$$C'_{f,\mathcal{V}} \ge \tfrac12 h^2[\|de^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)} + \|d^*e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\bar\Omega_-)}] + h\big(\alpha\min_{x\in\Omega_+\setminus\overline{\Omega_-\cup\mathcal{V}}}|\nabla f(x)|^2 - C_f\big)\|e^{\varphi/h}\psi\|^2_{L^2(\Omega_+\setminus\overline{\Omega_-\cup\mathcal{V}})}.$$

By taking α large enough, this yields the exponential decay estimate

$$\|e^{d_{\mathrm{Ag}}(\cdot,\mathcal{V})/h}\psi\|_{W^{1,2}(\Omega_+\setminus\bar\Omega_-)} \le \frac{C''_{f,\mathcal{V}}}{h}.$$

□

We are now in position to state the main result of this section, which can be seen as an equivalent of Proposition 3.5 for $\Delta^{D,(p)}_{f,h}(\Omega_+\setminus\bar\Omega_-)$.

Proposition 3.9. (1) *There is a constant $c > 0$ such that*

$$\sigma(\Delta^{D,(0)}_{f,h}(\Omega_+\setminus\bar\Omega_-)) \cap [0,c] = \varnothing \quad \textit{for all } h \in (0,h_0). \tag{3-10}$$

(2) *Let $(\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^N(\Omega_-)+m_1^D(\Omega_+\setminus\bar\Omega_-)}$ be an orthonormal basis of eigenvectors of $\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\bar\Omega_-)$ associated with the eigenvalues in $[0,\nu(h)]$, and let $\chi_+ \in \mathcal{C}^\infty(\bar\Omega_+)$ be such that $\chi_+ \equiv 1$ in a neighborhood of $\partial\Omega_+$ and $\chi_+ \equiv 0$ in a neighborhood of $\bar\Omega_-$. For all $k \in \{m_1^N(\Omega_-)+1,\dots,m_1^N(\Omega_-)+m_1^D(\Omega_+\setminus\bar\Omega_-)\}$, set $v_k^{(1)} = \chi_+\psi_k^{(1)}$. Then*

$$\sum_{k=m_1^N(\Omega_-)+1}^{m_1^N(\Omega_-)+m_1^D(\Omega_+\setminus\bar\Omega_-)}\|\psi_k^{(1)} - v_k^{(1)}\|_{W^{1,2}(\Omega_+\setminus\bar\Omega_-)} = \mathcal{O}(e^{-c_{\chi_+}/h}), \tag{3-11}$$

that is, the one-forms $v_k^{(1)}$ *are close to* $\psi_k^{(1)}$ *for* $k \in \{m_1^N(\Omega_-)+1, \dots, m_1^N(\Omega_-)+m_1^D(\Omega_+ \setminus \bar{\Omega}_-)\}$. *They are almost orthonormal in* $L^2(\Omega_+)$:

$$(\langle v_k^{(1)}, v_{k'}^{(1)}\rangle_{L^2(\Omega_+)})_{k,k'} = \mathrm{Id}_{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} + \mathcal{O}(e^{-c_{\chi_+}/h}).$$

Moreover, they belong to $D(\Delta_{f,h}^{D,(1)}(\Omega_+))$ *and they satisfy*

$$\langle v_k^{(1)}, \Delta_{f,h}^{D,(1)}(\Omega_+)v_k^{(1)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c_{\chi_+}/h}) \qquad \text{and} \qquad d_{f,h}^* v_k^{(1)} \equiv 0 \quad \text{in } \{\chi_+ = 1\}.$$

All the $\mathcal{O}(e^{-c_{\chi_+}/h})$ *remainders can be bounded from above by* $C_{\chi_+}e^{-c_{\chi_+}/h}$ *for some constants* $C_{\chi_+}, c_{\chi_+} > 0$ *independent of* $h \in [0, h_0]$.

Proof. (1) The lower bound on the spectrum of $\Delta_{f,h}^{D,(0)}(\Omega_+ \setminus \bar{\Omega}_-)$ comes from Lemma 3.1, used with $\varphi = 0$, and Hypothesis 1: for any function $\omega \in D(\Delta_{f,h}^{D,(0)}(\Omega_+ \setminus \bar{\Omega}_-))$,

$$\begin{aligned}&\langle \omega, \Delta_{f,h}^{D,(0)}(\Omega_+ \setminus \bar{\Omega}_-)\omega\rangle_{L^2(\Omega_+\setminus\bar{\Omega}_-)}\\ &= h^2\|d\omega\|^2_{L^2(\Omega_+\setminus\bar{\Omega}_-)} + h^2\|d^*\omega\|^2_{L^2(\Omega_+\setminus\bar{\Omega}_-)} + \langle(|\nabla f|^2 + h\mathcal{L}_{\nabla f} + h\mathcal{L}^*_{\nabla f})\omega, \omega\rangle_{L^2(\Omega_+\setminus\bar{\Omega}_-)} \geq C_f\|\omega\|^2_{L^2(\Omega_+\setminus\bar{\Omega}_-)}.\end{aligned}$$

(2) Let us start by proving that $d_{f,h}^* v_k^{(1)} \equiv 0$ in $\{\chi_+ = 1\}$. Let ψ be an eigenvector of $\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \bar{\Omega}_-)$ associated with an eigenvalue $\lambda \in [0, \nu(h)]$. Then, $d_{f,h}^*\psi$ belongs to $D(\Delta_{f,h}^{D,(0)}(\Omega_+ \setminus \bar{\Omega}_-))$ and

$$\Delta_{f,h}^{D,(0)}(d_{f,h}^*\psi) = \lambda d_{f,h}^*\psi,$$

according to [Helffer and Nier 2006] (see also (4-3) below). Using now (3-10) and $\lambda \leq \nu(h) \leq h$, this implies

$$d_{f,h}^*\psi \equiv 0, \tag{3-12}$$

and thus $d_{f,h}^* v \equiv 0$ in $\{\chi_+ \equiv 1\}$.

All the other estimates are proved like in Proposition 3.5 as consequences of the exponential decay estimate for the eigenvector ψ, stated in Proposition 3.9, using a neighborhood $\mathcal{V} \subset \bar{\Omega}_+ \setminus \Omega_-$ of $\partial\Omega_+$ such that $\chi_+ \equiv 1$ in a neighborhood of $\bar{\mathcal{V}}$.

For example, for (3-11), using $d_{\mathrm{Ag}}(x, \mathcal{V}) \geq 2c'_{\chi_+} > 0$ for $x \in \mathrm{supp}(1-\chi_+)$, Proposition 3.9 provides

$$\|(1-\chi_+)\psi\|_{W^{1,2}(\Omega_+\setminus\bar{\Omega}_-)} \leq C'_{\chi_+}e^{-c'_{\chi_+}/h}. \tag{3-13}$$

The proofs of the two other estimates on $\langle v_k^{(1)}, v_{k'}^{(1)}\rangle_{L^2(\Omega_+)}$ and $\langle v_k^{(1)}, \Delta_{f,h}^{D,(1)}(\Omega_+)v_k^{(1)}\rangle_{L^2(\Omega_+)}$ follow the same lines as in the proof of Proposition 3.5. □

3D. ***Exponential decay for the eigenvectors of*** $\boldsymbol{\Delta_{f,h}^{D,(p)}(\Omega_+)}$***,*** $(\boldsymbol{p = 0, 1})$***.*** We will use the two operators $\Delta_{f,h}^N(\Omega_-)$ and $\Delta_{f,h}^D(\Omega_+ \setminus \bar{\Omega}_-)$ to analyze the spectrum of $\Delta_{f,h}^D(\Omega_+)$.

Definition 3.10. On $\bigwedge L^2(\Omega_+) = \bigwedge L^2(\Omega_-) \oplus \bigwedge L^2(\Omega_+ \setminus \bar{\Omega}_-)$, let $\Delta_{f,h}^{\oplus}(\Omega_+)$ be the self-adjoint operator $\Delta_{f,h}^N(\Omega_-) \oplus \Delta_{f,h}^D(\Omega_+ \setminus \bar{\Omega}_-)$.

In other words, for any form u such that $u1_{\Omega_-} \in D(\Delta_{f,h}^N(\Omega_-))$ and $u1_{\Omega_+\setminus\bar{\Omega}_-} \in D(\Delta_{f,h}^D(\Omega_+ \setminus \bar{\Omega}_-))$ (namely if $u \in D(\Delta_{f,h}^{\oplus}(\Omega_+))$),

$$\Delta_{f,h}^{\oplus}(\Omega_+)u = \Delta_{f,h}^N(\Omega_-)(u1_{\Omega_-}) + \Delta_{f,h}^D(\Omega_+ \setminus \bar{\Omega}_-)(u1_{\Omega_+\setminus\bar{\Omega}_-}).$$

It is easy to check that the spectrum of $\Delta^{\oplus,(p)}_{f,h}(\Omega_+)$ is the union of the two spectra $\sigma(\Delta^{N,(p)}_{f,h}(\Omega_-))$ and $\sigma(\Delta^{D,(p)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-))$. Bases of eigenvectors are given by the direct sum structure. In particular, we have

$$m^{\oplus}_p(\Omega_+) = m^N_p(\Omega_-) + m^D_p(\Omega_+ \setminus \overline{\Omega}_-),$$

where $m^{\oplus}_p(\Omega_+) = \#\big[\sigma(\Delta^{\oplus}_{f,h}(\Omega_+)) \cap [0, \nu(h)]\big]$ denotes the number of small eigenvalues of $\Delta^{\oplus}_{f,h}(\Omega_+)$.

Proposition 3.11. *Let U be an open set satisfying* (3-2). *Let $(\psi^{(p)}_k)_{1\le k\le m^D_p(\Omega_+)}$, $p = 0$ or 1, be an orthonormal basis of eigenvectors of $\Delta^{D,(p)}_{f,h}(\Omega_+)$ associated with the eigenvalues in* $[0, \nu(h)]$, *and let $\chi \in \mathcal{C}^\infty(\overline{\Omega}_+)$ be such that $\chi \equiv 1$ in a neighborhood of $\partial\Omega_+ \cup U$ and $\chi \equiv 0$ in a neighborhood of $\partial\Omega_-$. For all $k \in \{1, \dots, m^D_p(\Omega_+)\}$, set $v^{(p)}_k = \chi\psi^{(p)}_k$. The forms $v^{(p)}_k$ are close to $\psi^{(p)}_k$ for $k \in \{1, \dots, m^D_p(\Omega_+)\}$:*

$$\sum_{k=1}^{m^D_p(\Omega_+)} \|\psi^{(p)}_k - v^{(p)}_k\|_{W^{1,2}(\Omega_+)} = \mathcal{O}(e^{-c_\chi/h}).$$

They are almost orthonormal in $L^2(\Omega_+)$:

$$(\langle v^{(p)}_k, v^{(p)}_{k'}\rangle_{L^2(\Omega_+)})_{k,k'} = \mathrm{Id}_{m^D_p(\Omega_+)} + \mathcal{O}(e^{-c_\chi/h}).$$

Moreover, they belong to the domain $D(\Delta^{\oplus,(p)}_{f,h}(\Omega_+))$ and they satisfy

$$\langle v^{(p)}_k, \Delta^{\oplus,(p)}_{f,h}(\Omega_+) v^{(p)}_k\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c_\chi/h}).$$

All the $\mathcal{O}(e^{-c_\chi/h})$ remainders can be bounded from above by $C_\chi e^{-c_\chi/h}$ for some constants $C_\chi, c_\chi > 0$ independent of $h \in [0, h_0]$.

Proof. The proof for $p = 0$ follows the same lines as the proofs of Proposition 3.4 and Proposition 3.5, because the boundary term in Lemma 3.1 disappears for functions vanishing along $\partial\Omega_+$.

For $p = 1$, the boundary term has to be taken into account as we did in the proofs of Proposition 3.8 and Proposition 3.9. A neighborhood $\mathcal{V}$ of $\partial\Omega_+$ has to be introduced and the function φ used in Lemma 3.1 is $\varphi(x) = (1-\alpha h)d_{\mathrm{Ag}}(x, U \cup \mathcal{V})$ with $\alpha > 0$ large enough. □

Notice that the number $m^D_p(\Omega_+)$ of small eigenvalues for $\Delta^{D,(p)}_{f,h}(\Omega_+)$ is a priori dependent on h. We did not explicitly indicate this dependency since the result of the next section is that $m^D_p(\Omega_+)$ is actually independent of h.

3E. *On the number of small eigenvalues of $\Delta^{D,(p)}_{f,h}(\Omega_+)$.* Using the results of the three previous sections, one can show that the number $m^D_p(\Omega_+)$ of eigenvalues of $\Delta^{D,(p)}_{f,h}(\Omega_+)$ in $[0, \nu(h)]$, is actually independent of $h \in (0, h_0)$.

Proposition 3.12. *For $p \in \{0, 1\}$, the number of eigenvalues of $\Delta^{D,(p)}_{f,h}(\Omega_+)$ lying in $[0, \nu(h)]$ is given by*

$$m^D_p(\Omega_+) = m^N_p(\Omega_-) + m^D_p(\Omega_+ \setminus \overline{\Omega}_-)\,,$$

where we recall (*see* (3-10)) *that $m^D_0(\Omega_+ \setminus \overline{\Omega}_-) = 0$. Moreover all these eigenvalues are exponentially small, i.e., there exists $c'_0 > 0$ such that*

$$\sigma(\Delta^{D,(p)}_{f,h}(\Omega_+)) \cap [0, \nu(h)] \subset [0, e^{-c'_0/h}] \quad \textit{for all } h \in (0, h_0),\ p = 0, 1.$$

Proof. This is obtained as an application of the min–max principle. Indeed, we know that the spectrum of $\Delta_{f,h}^{D,(p)}(\Omega_+)$ is given by the formula

$$\lambda_k^{(p)}(\Omega_+) = \sup_{\{\omega_1,\dots,\omega_{k-1}\}} Q(\omega_1,\dots,\omega_{k-1}) \quad \text{for } k \ge 1,$$

where

$$Q(\omega_1,\dots,\omega_{k-1}) = \inf_v \left\{ \frac{\langle v, \Delta_{f,h}^{D,(p)}(\Omega_+) v\rangle_{L^2(\Omega_+)}}{\|v\|^2_{L^2(\Omega_+)}} : v \in D(\Delta_{f,h}^{D,(p)}(\Omega_+)),\, v \in \operatorname{Span}(\omega_1,\dots,\omega_{k-1})^\perp \right\}.$$

By convention, for $k = 1$, the supremum is taken over an empty set (and can thus be neglected). Using Proposition 3.5 and Proposition 3.9, one can build $m_p := m_p^N(\Omega_-) + m_p^D(\Omega_+ \setminus \bar{\Omega}_-)$, almost orthonormal vectors for which the Rayleigh quotients associated with $\Delta_{f,h}^{D,(p)}(\Omega_+)$ are exponentially small. Let us fix $\varepsilon > 0$ and consider $\{\omega_1,\dots,\omega_{m_p-1}\}$ such that $\lambda_{m_p}^{(p)}(\Omega_+) \le Q(\omega_1,\dots,\omega_{m_p-1}) + \varepsilon$. Since, in the limit $h \to 0$, the m_p vectors built in Proposition 3.5 and Proposition 3.9 are linearly independent, there exists a linear combination $v \in D(\Delta_{f,h}^{D,(p)}(\Omega_+))$ of these vectors which is in $\operatorname{Span}(\omega_1,\dots,\omega_{m_p-1})^\perp$. Using the estimates on the Rayleigh quotients and the almost orthonormality of these vectors, one obtains that $\langle v, \Delta_{f,h}^{D,(p)}(\Omega_+) v\rangle_{L^2(\Omega_+)}/\|v\|^2_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$ for some positive constant c. This implies that $Q(\omega_1,\dots,\omega_{k-1}) = \mathcal{O}(e^{-c/h})$ and thus $\lambda_{m_p}^{(p)}(\Omega_+) = \mathcal{O}(e^{-c/h})$. Therefore, one gets $m_p^D(\Omega_+) \ge m_p = m_p^N(\Omega_-) + m_p^D(\Omega_+ \setminus \bar{\Omega}_-)$.

Similar reasoning on $\Delta_{f,h}^{\oplus,(p)}(\Omega_+)$ using Proposition 3.11 gives the opposite inequality $m_p^{\oplus}(\Omega_+) = m_p^N(\Omega_-) + m_p^D(\Omega_+ \setminus \bar{\Omega}_-) \ge m_p^D(\Omega_+)$. This ends the proof. ☐

4. Quasimodes for $\Delta_{f,h}^{D,(0)}(\Omega_+)$ and $\Delta_{f,h}^{D,(1)}(\Omega_+)$

In this section, we specify the quasimodes which will be useful for the analysis of the spectrum of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ lying in $[0, \nu(h)]$. In our context, for $p = 0, 1$, a quasimode for $\Delta_{f,h}^{D,(p)}(\Omega_+)$ is simply a function v in the domain $D(\Delta_{f,h}^{D,(p)}(\Omega_+))$ such that $\langle v, \Delta_{f,h}^{D,(p)}(\Omega_+) v\rangle_{L^2(\Omega_+)}/\|v\|^2_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$. Quasimodes for $\Delta_{f,h}^{D,(0)}(\Omega_+)$ (resp. $\Delta_{f,h}^{D,(1)}(\Omega_+)$) will be built from the eigenvectors of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ (resp. of $\Delta_{f,h}^{N,(1)}(\Omega_-)$ and $\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \bar{\Omega}_-)$).

4A. ***The restricted differential β.*** We recall here basic properties of boundary Witten Laplacians.

Proposition 4.1. *Let Ω be a regular bounded domain of (M, g) and consider the Dirichlet (resp. Neumann) realization $A = \Delta_{f,h}^D(\Omega)$ (resp. $A = \Delta_{f,h}^N(\Omega)$) of the Witten Laplacian with form domain $Q(A) = W_D^{1,2}(\Omega)$ (resp. $Q(A) = W_N^{1,2}(\Omega)$). The differential $d_{f,h}$ and codifferential $d_{f,h}^*$ satisfy the commutation property: for all $z \in \mathbb{C} \setminus \sigma(A)$ and $u \in Q(A)$,*

$$d_{f,h}(z - A)^{-1}u = (z - A)^{-1} d_{f,h} u \quad \textit{and} \quad d_{f,h}^*(z - A)^{-1}u = (z - A)^{-1} d_{f,h}^* u.$$

Consequently, for any $\ell \in \mathbb{R}_+$,

$$d_{f,h} \circ 1_{[0,\ell]}(A^{(p)}) = 1_{[0,\ell]}(A^{(p+1)}) \circ d_{f,h}, \quad \textit{and} \quad d_{f,h}^* \circ 1_{[0,\ell]}(A^{(p)}) = 1_{[0,\ell]}(A^{(p-1)}) \circ d_{f,h}^*, \tag{4-1}$$

where $A^{(p)}$ denotes the restriction of A to p-forms. Moreover, if $F_\ell^{(p)}$ denotes the spectral subspace $\operatorname{Ran} 1_{[0,\ell]}(A^{(p)})$, *the chain complex*

$$0 \longrightarrow F_\ell^{(0)} \longrightarrow \cdots \longrightarrow F_\ell^{(p-1)} \xrightarrow{-d_{f,h}} F_\ell^{(p)} \xrightarrow{-d_{f,h}} F_\ell^{(p+1)} \longrightarrow \cdots \longrightarrow F_\ell^{(d)} \longrightarrow 0 \tag{4-2}$$

is quasi-isomorphic to the relative (*resp. absolute*) *Hodge–de Rham chain complex. The Witten codifferential $d^*_{f,h}$ implements the dual chain complex.*

Relative and absolute homologies are standard notions in algebraic topology and Morse theory (see, for example, [Hatcher 2002; Milnor 1963]). Their translations to cohomology and boundary value Hodge theory is presented, for example, in [Taylor 1997; Schwarz 1995]. A quasi-isomorphism is a morphism of complexes which induces an isomorphism of homology groups.

We refer to [Chang and Liu 1995; Helffer and Nier 2006; Le Peutrec 2010b] for the adaptation to boundary cases of these well-known properties of Witten Laplacians [Cycon et al. 1987, Chapter 11].

Let us give two consequences of that result that are useful in our context. First, the following property, which was already used in the proof of Proposition 3.9, holds (using the notation of Proposition 4.1):

$$A^{(p)}\psi = \lambda\psi \implies \begin{cases} A^{(p+1)} d_{f,h}\psi = \lambda d_{f,h}\psi \\ A^{(p-1)} d^*_{f,h}\psi = \lambda d^*_{f,h}\psi \end{cases} \tag{4-3}$$

with the convention $A^{(-1)} = A^{(d+1)} = 0$. Secondly, we have the orthogonal decompositions

$$F_\ell = \operatorname{Ker}\big[A\big|_{F_\ell}\big] \overset{\perp}{\oplus} \operatorname{Ran}\big[d_{f,h}\big|_{F_\ell}\big] \overset{\perp}{\oplus} \operatorname{Ran}\big[d^*_{f,h}\big|_{F_\ell}\big], \tag{4-4}$$

$$\operatorname{Ran}\big[d^*_{f,h}\big|_{F_\ell}\big]^\perp = \operatorname{Ker}\big[d_{f,h}\big|_{F_\ell}\big] = \operatorname{Ker}\big[A\big|_{F_\ell}\big] \overset{\perp}{\oplus} \operatorname{Ran}\big[d_{f,h}\big|_{F_\ell}\big],$$

$$\operatorname{Ran}\big[d_{f,h}\big|_{F_\ell}\big]^\perp = \operatorname{Ker}\big[d^*_{f,h}\big|_{F_\ell}\big] = \operatorname{Ker}\big[A\big|_{F_\ell}\big] \overset{\perp}{\oplus} \operatorname{Ran}\big[d^*_{f,h}\big|_{F_\ell}\big],$$

where $F_\ell = \bigoplus_{p=0}^d F_\ell^{(p)}$. In our problem, we shall use the following notation:

Definition 4.2. Consider the Dirichlet realization $\Delta^D_{f,h}(\Omega_+)$ of $\Delta_{f,h}$ on Ω_+. For $p = 0, 1$, the operators $\Pi^{(p)}$ are the spectral projections

$$\Pi^{(p)} = 1_{[0,\nu(h)]}(\Delta^{D,(p)}_{f,h}(\Omega_+)), \quad p = 0, 1,$$

and their range is denoted by $F^{(p)}$. Moreover, the Witten differential $d_{f,h}$ restricted to $F^{(0)}$ is written as $\beta = d_{f,h}\big|_{F^{(0)}} : F^{(0)} \to F^{(1)}$, so that $\Delta^{D,(0)}_{f,h}(\Omega_+)\big|_{F^{(0)}} = \beta^*\beta$, where $\beta^* = d^*_{f,h}\big|_{F^{(1)}} : F^{(1)} \to F^{(0)}$.

A consequence of the commutation properties (4-1) is the identity

$$\beta = \Pi^{(1)} d_{f,h} = d_{f,h}\Pi^{(0)} = \Pi^{(1)}\beta\Pi^{(0)}. \tag{4-5}$$

Moreover, (4-4) becomes

$$F^{(0)} = \operatorname{Ran}[\beta^*], \quad \text{since } \operatorname{Ker}(\beta) = \{0\}$$

because $\beta u = d_{f,h} u = 0$ and $u = 0$ on $\partial\Omega$ imply $u = 0$, and

$$F^{(1)} = \operatorname{Ker}[\beta^*] \overset{\perp}{\oplus} \operatorname{Ran}[\beta] = \operatorname{Ker}[\Delta^{D,(1)}_{f,h}(\Omega_+)] \overset{\perp}{\oplus} \operatorname{Ran}(\beta) \overset{\perp}{\oplus} \operatorname{Ran}\big[d^*_{f,h}\big|_{F^{(2)}}\big]. \tag{4-6}$$

4B. ***Truncated eigenvectors.*** Let us recall the eigenvectors that have been introduced in Propositions 3.5 and 3.9:

- $(\psi_j^{(0)})_{1\le j\le m_0^N(\Omega_-)}$ are eigenvectors for the operator $\Delta_{f,h}^{N,(0)}(\Omega_-)$ associated with the eigenvalues $0=\mu_1^{(0)}(\Omega_-)\le C_0 e^{-2(\kappa_f-c_0)/h}\le \mu_2^{(0)}(\Omega_-)\le\cdots\le \mu_{m_0^N(\Omega_-)}^{(0)}(\Omega_-)\le e^{-c_0/h}\le \nu(h)$. The first eigenvector $\psi_1^{(0)}$ associated with the eigenvalue $\mu_1^{(0)}(\Omega_-)=0$ is $\psi_1^{(0)}=e^{-f/h}1_{\Omega_-}/\bigl(\int_{\Omega_-}e^{-2f(x)/h}\,dx\bigr)^{1/2}$. The lower bound on $\mu_2^{(0)}(\Omega_-)$ stated above is valid for sufficiently small h and was proven in Proposition 3.7.
- $(\psi_k^{(1)})_{1\le k\le m_1^N(\Omega_-)}$ are eigenvectors for the operator $\Delta_{f,h}^{N,(1)}(\Omega_-)$ associated with the $m_1^N(\Omega_-)$ eigenvalues smaller than $\nu(h)$. Using (4-3), those eigenvectors can be labeled so that
$$\psi_k^{(1)}=(\mu_{k+1}^{(0)}(\Omega_-))^{-1/2}d_{f,h}\psi_{k+1}^{(0)}=(\mu_{k+1}^{(0)}(\Omega_-))^{-1/2}\beta\psi_{k+1}^{(0)}\quad\text{for } k\in\{1,\dots,m_0^N(\Omega_-)-1\}.$$
Notice that we may have $m_1^N(\Omega_-)=m_0^N(\Omega_-)-1$. If not, using (4-6), $\beta^*\psi_k^{(1)}=d_{f,h}^*\psi_k^{(1)}=0$ for $k\ge m_0^N(\Omega_-)$.
- $(\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^N(\Omega_-)+m_1^D(\Omega_+\setminus\bar\Omega_-)}$ are eigenvectors for the operator $\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-)$ associated with the $m_1^D(\Omega_+\setminus\bar\Omega_-)$ eigenvalues smaller than $\nu(h)$. From (3-12) in the proof of Proposition 3.9, we know that $d_{f,h}^*\psi_k^{(1)}=\beta^*\psi_k^{(1)}=0$.

In Proposition 3.12 we proved that $m_0^D(\Omega_+)=m_0^N(\Omega_-)$ and $m_1^D(\Omega_+)=m_1^N(\Omega_-)+m_1^D(\Omega_+\setminus\bar\Omega_-)$. The families $(\psi_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$ and $(\psi_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$ are orthonormal bases of eigenvectors for $\Delta_{f,h}^{\oplus,(0)}(\Omega_+)$ and $\Delta_{f,h}^{\oplus,(1)}(\Omega_+)$, respectively, restricted to the spectral range $[0,\nu(h)]$. These two families will be used to construct quasimodes for the operator $\Delta_{f,h}^{D,(p)}(\Omega_+)$ restricted to the spectral range $[0,\nu(h)]$. This will require some appropriate truncations or extrapolations, detailed below.

Let us start with $\psi_1^{(0)}$ and let us introduce
$$\tilde\psi_1^{(0)}=\frac{e^{-f/h}1_{\Omega_+}(x)}{\bigl(\int_{\Omega_+}e^{-2f(x)/h}\,dx\bigr)^{1/2}}.\tag{4-7}$$

These two functions are exponentially close in $L^2(\Omega_+)$, that is,
$$\|\psi_1^{(0)}-\tilde\psi_1^{(0)}\|_{L^2(\Omega_+)}\le Ce^{-c/h},$$
owing to $f(x)\ge\min_{\partial\Omega_-}f>\min_{\Omega_+}f$ for all $x\in\bar\Omega_+\setminus\Omega_-$ and the following upper and lower bounds of the integral factor:

Lemma 4.3. *Let Ω be a regular bounded domain of (M,g) and let f belong to $\mathcal{C}^\infty(\bar\Omega)$ such that $\min_{\bar\Omega}f$ is achieved in Ω. Then there exists a constant $C_f>0$ such that*
$$\frac{1}{C_f}h^{d/2}e^{-2(\min_\Omega f)/h}\le\int_\Omega e^{-2f(x)/h}\,dx\le \mathrm{Vol}_g(\Omega)e^{-2(\min_\Omega f)/h},$$
where $\mathrm{Vol}_g(\Omega)$ denotes the volume of Ω for the metric g.

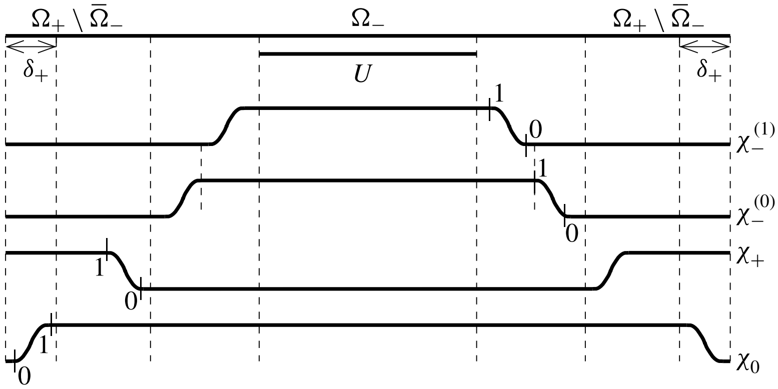

Figure 1. Positions of the domains Ω_+, Ω_- and U, and of the supports of the cut-off functions $\chi_-^{(0)}$, $\chi_-^{(1)}$, χ_+, χ_0.

Proof. The upper bound is obvious since $e^{-2f(x)/h} \le e^{-2(\min_\Omega f)/h}$ for all $x \in \Omega$. For the lower bound, write

$$\int_\Omega e^{-2f(x)/h}\,dx = \int_\Omega \int_{2f(x)}^{+\infty} e^{-t/h}\,\frac{dt}{h}\,dx = \int_{2\min_\Omega f}^{+\infty} \mathrm{Vol}_g(2f < t) e^{-t/h}\,\frac{dt}{h}$$
$$= e^{-2(\min_\Omega f)/h} \int_0^{+\infty} \mathrm{Vol}_g\big(2f < 2\min_\Omega f + hs\big) e^{-s}\,ds.$$

We assumed the existence of $x_0 \in \Omega$ such that $f(x_0) = \min_\Omega f$. Using the Taylor expansion of f around x_0, there exist $r > 0$, $h_0 > 0$ and $s_0 > 0$ such that the ball $B(x_0, (hs)^{1/2}/r)$ is included in $\{f < \min_\Omega f + \frac{1}{2}hs\}$ for all $s < s_0$ and $h < h_0$. Since $\mathrm{Vol}_g[B(x_0, (hs)^{1/2}/r)] \ge (hs)^{d/2}/C_r$, we get

$$\int_\Omega e^{-2f(x)/h}\,dx \ge \frac{1}{C_r} e^{-2(\min_\Omega f)/h} \int_0^{s_0} e^{-s}(hs)^{d/2}\,ds \ge \frac{h^{d/2} e^{-2(\min_\Omega f)/h}}{C_f}. \qquad \square$$

Compared to the standard Laplace estimate, the interest of Lemma 4.3 is that it holds even if the minimum of f is degenerate.

In all of what follows, U denotes a fixed subset of Ω_- satisfying (3-2). Let us introduce various cut-off functions, which all satisfy $0 \le \chi \le 1$. We refer to Figure 1 for an illustration of these cut-off functions with respect to the three sets $U \subset \Omega_- \subset \Omega_+$.

- $\chi_-^{(0)}$ and $\chi_-^{(1)}$ are two cut-off functions like χ_- in Proposition 3.5, that is, $\chi_-^{(p)} \in \mathcal{C}_0^\infty(\Omega_-)$ and $\chi_-^{(p)} \equiv 1$ in a neighborhood of U with the additional condition that $\chi_-^{(0)} \equiv 1$ in a neighborhood of $\operatorname{supp} \chi_-^{(1)}$.
- χ_+ is chosen as in Proposition 3.9, that is, $\chi_+ \in \mathcal{C}^\infty(\bar{\Omega}_+)$, $\chi_+ \equiv 1$ in a neighborhood of $\partial\Omega_+$ and $\chi_+ \equiv 0$ in a neighborhood of $\bar{\Omega}_-$. Let us introduce $c_+ > 0$ such that $\chi_+ \equiv 1$ on $\{x \in \bar{\Omega}_+ : d(x, \partial\Omega_+) \le c_+\}$.
- χ_0 belongs to $\mathcal{C}_0^\infty(\Omega_+)$, $\chi_0 \equiv 1$ in a neighborhood of $\bar{\Omega}_-$ and is chosen in such a way that its gradient is supported in $\{x \in \Omega_+ : d(x, \partial\Omega_+) \le \delta_+\}$, where $\delta_+ \in (0, c_+)$ will be fixed later.

We are now in position to introduce a family of quasimodes for the operator $\Delta_{f,h}^{D,(p)}(\Omega_+)$.

Definition 4.4. Let $\chi_-^{(0)}$, $\chi_-^{(1)}$, χ_+ and χ_0 be the cut-off functions defined above. Let $(\psi_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$ and $(\psi_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$ be the previously gathered families of eigenvectors of $\Delta_{f,h}^{N,(p)}(\Omega_-)$ and $\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-)$, and finally let $\tilde{\psi}_1^{(0)}$ be given by (4-7). The families of vectors $(v_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$ and $(v_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$ are defined by:

- $v_1^{(0)} = \chi_0 \tilde{\psi}_1^{(0)}$;
- $v_j^{(0)} = \chi_-^{(0)} \psi_j^{(0)}$ for $j \in \{2, \dots, m_0^D(\Omega_+)\}$;
- $v_k^{(1)} = \chi_-^{(1)} \psi_k^{(1)}$ for $k \in \{1, \dots, m_1^N(\Omega_-)\}$;
- $v_k^{(1)} = \chi_+ \psi_k^{(1)}$ for $k \in \{m_1^N(\Omega_-)+1, \dots, m_1^D(\Omega_+)\}$.

Proposition 4.5. *The families* $(v_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$ *and* $(v_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$ *of Definition 4.4 satisfy*:

(1) *They are almost orthonormal in* $L^2(\Omega_+)$:

$$(\langle v_j^{(0)}, v_{j'}^{(0)}\rangle_{L^2(\Omega_+)})_{1\le j,j'\le m_0^D(\Omega_+)} = \mathrm{Id}_{m_0^D(\Omega_+)} + \mathcal{O}(e^{-c/h}),$$
$$(\langle v_k^{(1)}, v_{k'}^{(1)}\rangle_{L^2(\Omega_+)})_{1\le k,k'\le m_1^D(\Omega_+)} = \mathrm{Id}_{m_1^D(\Omega_+)} + \mathcal{O}(e^{-c/h})$$

for some constant $c>0$ *independent of* δ_+.

(2) *The elements* $v_j^{(0)}$, $1\le j\le m_0^D(\Omega_+)$ (*resp.* $v_k^{(1)}$, $1\le k\le m_1^D(\Omega_+)$) *belong to* $D(\Delta_{f,h}^{D,(0)}(\Omega_+))$ (*resp.* $\Delta_{f,h}^{D,(1)}(\Omega_+)$) *and satisfy*

$$\langle v_j^{(0)}, \Delta_{f,h}^{D,(0)}(\Omega_+) v_j^{(0)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h}) \quad \textit{and} \quad \langle v_k^{(1)}, \Delta_{f,h}^{D,(1)}(\Omega_+) v_k^{(1)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h}),$$

respectively, for some constant $c>0$ *independent of* δ_+.

(3) *Let us consider the spectral projections* $\Pi^{(0)}$ *and* $\Pi^{(1)}$ *associated with* $\Delta_{f,h}^D(\Omega_+)$ *introduced in Definition 4.2. The elements* $v_j^{(0)}$, $1\le j\le m_0^D(\Omega_+)$ (*resp.* $v_k^{(1)}$, $1\le k\le m_1^D(\Omega_+)$) *satisfy*:

$$\|v_j^{(0)} - \Pi^{(0)} v_j^{(0)}\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h}) \quad \textit{and} \quad \|v_k^{(1)} - \Pi^{(1)} v_k^{(1)}\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h}),$$

respectively, for some constant $c>0$ *independent of* δ_+.

Proof. (1) The families $(\psi_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$ and $(\psi_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$ are orthonormal bases of eigenvectors of $\Delta_{f,h}^{\oplus,(0)}$ and $\Delta_{f,h}^{\oplus,(1)}$, respectively. Proposition 3.5 implies that the family $(\chi_-^{(0)}\psi_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$ is almost orthonormal. The estimate $\|\chi_0\tilde{\psi}_1^{(0)} - \chi_-^{(0)}\psi_1^{(0)}\|_{L^2(\Omega_+)} \le Ce^{-c/h}$ (which is a consequence of Lemma 4.3 and $f(x) \ge \min_{\partial\Omega_-} f > \min_{\Omega_+} f$ for all $x \in \overline{\Omega}_+ \setminus \Omega_-$) ends the proof of the almost orthonormality of $(v_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}$. For $p=1$, the two families $(v_k^{(1)} = \chi_-^{(1)}\psi_k^{(1)})_{1\le k\le m_1(\Omega_-)}$ and $(v_k^{(1)} = \chi_+\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)}$ have disjoint supports and therefore lie in orthogonal subspaces of $L^2(\Omega_+)$. Also, the almost orthonormality of both families is again a consequence of the exponential decay of the $\psi_k^{(1)}$; see Proposition 3.5 and Proposition 3.11.

(2) With the chosen truncations, all the vectors $v_j^{(0)}$ (resp. $v_k^{(1)}$) belong to the domain $D(\Delta_{f,h}^{D,(0)}(\Omega_+))$ (resp. $D(\Delta_{f,h}^{D,(1)}(\Omega_+))$). In all cases except $p=0$ and $k=1$, we obtain, for $v=\chi\psi$ (we omit the index k and the superscript (p)) and $A\psi = \lambda\psi$, where $A = \Delta_{f,h}^N(\Omega_-)$ or $A = \Delta_{f,h}^D(\Omega_+ \setminus \overline{\Omega}_-)$,

$$\langle v, \Delta_{f,h}^D(\Omega_+)v\rangle_{L^2(\Omega_+)} = \|d_{f,h}v\|_{L^2}^2 + \|d_{f,h}^* v\|_{L^2(\Omega_+)}^2 \le \langle \psi, A\psi\rangle + C\|\psi\|_{W^{1,2}(\{\chi\ne 1\})}^2 \le Ce^{-c/h},$$

owing to $\langle\psi, A\psi\rangle = \lambda = \mathcal{O}(e^{-c_0/h})$ and to the estimates on $\psi - v$ given in Proposition 3.5 and Proposition 3.11. For $p = 0$ and $k = 1$, it is even simpler because $d_{f,h}\tilde{\psi}_1^{(0)} = 0$ implies

$$\langle v_1^{(0)}, \Delta_{f,h}^{D,(0)}(\Omega_+)v_1^{(0)}\rangle_{L^2(\Omega_+)} = \|d_{f,h}(\chi_0\tilde{\psi}_1^{(0)})\|^2_{L^2(\Omega_+)} = \|(hd\chi_0)\tilde{\psi}_1^{(0)}\|^2_{L^2(\Omega_+)} \le Ce^{-c/h}$$

as a consequence of Lemma 4.3 (see (5-8) below for a more precise estimate).

(3) All the $v_j^{(0)}$ and $v_k^{(1)}$ satisfy $\langle v, Av\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$ with $A = \Delta_{f,h}^{D,(0)}(\Omega_+)$ or $A = \Delta_{f,h}^{D,(1)}(\Omega_+)$, and recall that $\Pi^{(0)}$ and $\Pi^{(1)}$ are the spectral projectors $1_{[0,\nu(h)]}(A)$. The last estimates are consequences of

$$\nu(h)\|1_{(\nu(h),+\infty)}(A)v\|^2_{L^2(\Omega_+)} \le \langle v, Av\rangle_{L^2(\Omega_+)} \le Ce^{-c/h}$$

together with the fact that $\lim_{h\to 0} h\log\nu(h) = 0$; see (2-5). □

In the next section, we will need these calculations:

Proposition 4.6. *The coefficients* $\langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)}$, $j \in \{1, \dots, m_0^D(\Omega_+)\}$, $k \in \{1, \dots, m_1^D(\Omega_+)\}$, *satisfy*:

(1) *For* $j = 1$ *and* $k \in \{1, \dots, m_1^N(\Omega_-)\}$, $\langle v_k^{(1)}, d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)} = 0$.

(2) *For* $j = 1$ *and* $k \in \{m_1^N(\Omega_-) + 1, \dots, m_1^D(\Omega_+)\}$,

$$\langle v_k^{(1)}, d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)} = -\frac{h\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}},$$

where $d\sigma$ *is the infinitesimal volume on* $\partial\Omega_+$ *and* $n(\sigma)$ *the outward normal vector at* $\sigma \in \partial\Omega_+$.

(3) *For* $j \in \{2, \dots, m_0^D(\Omega_+)\}$ *and* $k \in \{1, \dots, m_1^N(\Omega_-)\}$,

$$\langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)} = \sqrt{\mu_j^{(0)}(\Omega_-)}(\delta_{k,j-1} + \mathcal{O}(e^{-c/h})).$$

(4) *For* $j \in \{2, \dots, m_0^D(\Omega_+)\}$ *and* $k \in \{m_1^N(\Omega_-) + 1, \dots, m_1^D(\Omega_+)\}$, $\langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)} = 0$.

Proof. Cases (1) and (4) are due to the disjoint supports of $d_{f,h}v_j^{(0)}$ and $v_k^{(1)}$ (see Figure 1).

Case (3) comes from the computation

$$\begin{aligned} d_{f,h}v_j^{(0)} &= d_{f,h}(\chi_-^{(0)}\psi_j^{(0)}) = \chi_-^{(0)}d_{f,h}\psi_j^{(0)} + (hd\chi_-^{(0)}) \wedge \psi_j^{(0)} \\ &= \sqrt{\mu_j^{(0)}(\Omega_-)}\chi_-^{(0)}\psi_{j-1}^{(1)} + \psi_j^{(0)}hd\chi_-^{(0)}. \end{aligned}$$

The condition $\chi_-^{(0)} \equiv 1$ in a neighborhood of $\operatorname{supp}\chi_-^{(1)}$ then leads to

$$\begin{aligned} \langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)} &= \langle \chi_-^{(1)}\psi_k^{(1)}, \sqrt{\mu_j^{(0)}(\Omega_-)}\psi_{j-1}^{(1)}\rangle_{L^2(\Omega_-)} \\ &= \sqrt{\mu_j^{(0)}(\Omega_-)}\delta_{k,j-1} + \sqrt{\mu_j^{(0)}(\Omega_-)}\|(1-\chi_-^{(1)})\psi_k^{(1)}\|_{L^2(\Omega_-)}, \end{aligned}$$

and we conclude with the exponential decay of $\psi_k^{(1)}$ given by (3-4) in the proof Proposition 3.5.

For case (2), we first use

$$d_{f,h}v_1^{(0)} = d_{f,h}(\chi_0\tilde{\psi}_1^{(0)}) = \frac{e^{-f/h}}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}} h\,d\chi_0.$$

The assumption on the supports of χ_0 and χ_+ (see Figure 1) implies that $d\chi_0$ is supported in the interior of $\{x \in \overline{\Omega}_+ : \chi_+(x) = 1\}$, so that

$$\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{\frac{1}{2}} \langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle = \langle \chi_+\psi_k^{(1)}, e^{-f/h}h\,d\chi_0\rangle = \langle \psi_k^{(1)}, e^{-f/h}h\,d\chi_0\rangle.$$

The definition of the Hodge $\star$ operation gives

$$\langle \psi_k^{(1)}, e^{-f/h}h\,d\chi_0\rangle = h\int_{\Omega_+} d\chi_0 \wedge [\star(e^{-f/h}\psi_k^{(1)})] = -h\int_{\Omega_+\setminus\overline{\Omega}_-} d(1-\chi_0)\wedge[\star(e^{-f/h}\psi_k^{(1)})].$$

We recall (see (3-12) in the proof of Proposition 3.9) that $d^*_{f,h}\psi_k^{(1)} = 0$ in $\Omega_+ \setminus \overline{\Omega}_-$, which means

$$d[\star(e^{-f/h}\psi_k^{(1)})] = (-1)^{1+1}\star\left[\frac{e^{-f/h}}{h}d^*_{f,h}\psi_k^{(1)}\right] = 0 \quad \text{in } \Omega_+ \setminus \overline{\Omega}_-\,.$$

Hence, we get

$$d(1-\chi_0)\wedge[\star(e^{-f/h}\psi_k^{(1)})] = d\big[(1-\chi_0)\wedge[\star(e^{-f/h}\psi_k^{(1)})]\big],$$

and Stokes' formula yields

$$\langle \psi_k^{(1)}, e^{-f/h}h\,d\chi_0\rangle = -h\int_{\partial\Omega_+} e^{-f/h}\star\psi_k^{(1)} = -h\int_{\partial\Omega_+} e^{-f/h}\boldsymbol{t}(\star\psi_k^{(1)}).$$

Using the relations (A-8), $\boldsymbol{t}\star = \star\boldsymbol{n}$, and (A-10) $\omega_1 \wedge (\star\boldsymbol{n}\omega_2) = \langle \omega_1, \boldsymbol{i}_n\omega_2\rangle_{\bigwedge^{p-1}T^*_\sigma\Omega_+}\,d\sigma$ along $\partial\Omega_+$ (where $d\sigma$ is the infinitesimal volume on $\partial\Omega_+$ and $n(\sigma)$ the outward normal vector at $\sigma \in \partial\Omega_+$) with $p = 1$, $\omega_1 = 1$ and $\omega_2 = \psi_k^{(1)}$, we get

$$\langle \psi_k^{(1)}, e^{-f/h}h\,d\chi_0\rangle = -h\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma.$$

This concludes the proof of case (2), and of Proposition 4.6. □

5. Analysis of the restricted differential β

It is in this section that the assumption (2-8) is used. We assume that the open subset U of Ω_- that has been used to build the cut-off functions in the previous section satisfies (in addition to (3-2))

$$U \cup \mathcal{V}_- = \overline{\Omega}_-, \tag{5-1}$$

where $\mathcal{V}_-$ is the neighborhood of $\partial\Omega_-$ introduced in the assumption (2-8).

The main result of this section is the following:

Proposition 5.1. *The singular values of* $\beta = d_{f,h}\big|_{F^{(0)}} : F^{(0)} \to F^{(1)}$, *labeled in decreasing order*, *are given by*

$$s_j(\beta) = \sqrt{\mu^{(0)}_{m_0^D(\Omega_+)+1-j}(\Omega_-)(1+\mathcal{O}(e^{-c/h}))} \quad \text{for } j \in \{1, \dots, m_0^D(\Omega_+) - 1\},$$

$$s_{m_0^D(\Omega_+)}(\beta) = \frac{h\sqrt{\sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} \left|\int_{\partial\Omega_+} e^{-f(\sigma)/h} \boldsymbol{i}_n \psi_k^{(1)}(\sigma)\, d\sigma\right|^2}}{\sqrt{\int_{\Omega_+} e^{-2f(x)/h}\, dx}}(1+\mathcal{O}(e^{-c/h}))$$

for some $c > 0$.

According to the notation of Section 4B, $(\mu_j^{(0)}(\Omega_-))_{1\le j\le m_0^D(\Omega_+)}$ are the eigenvalues of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ and $(\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)}$ are the eigenvectors of $\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \overline{\Omega}_-)$. Notice that, contrary to the eigenvalues of the operators considered in the previous sections which were labeled in increasing order, the singular values are naturally labeled in decreasing order. Of course, the singular values of β are related to the small eigenvalues of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ through the relation

$$\sigma(\Delta_{f,h}^{D,(0)}(\Omega_+)) \cap [0, \nu(h)] = \{s_k(\beta)^2 : 1 \le k \le m_0^D(\Omega_+)\}, \tag{5-2}$$

since $\Delta_{f,h}^{D,(0)}\big|_{F^{(0)}} = \beta^*\beta$. Proposition 5.1 will thus be instrumental in proving Theorem 2.4.

The idea of the proof of Proposition 5.1 follows the linear algebra argument used in [Helffer et al. 2004; Helffer and Nier 2006; Le Peutrec 2010b; Le Peutrec et al. 2013] and well summarized in [Le Peutrec 2009]. Notice that $\beta = d_{f,h}\big|_{F^{(0)}}$ is a *finite-dimensional* linear operator. The proof then relies on the following fundamental property for singular values of matrices. Let us denote by $s_k(B)$, $k \in \{1, \dots, \max(n_0, n_1)\}$, the singular values of a matrix $B \in \mathcal{M}_{n_1,n_0}(\mathbb{C})$. Then, for any matrices $C_0 \in \mathcal{M}_{n_0}(\mathbb{C})$ and $C_1 \in \mathcal{M}_{n_1}(\mathbb{C})$,

$$s_k(BC_0) \le s_k(B)\|C_0\|, \quad s_k(C_1B) \le \|C_1\| s_k(B), \tag{5-3}$$

and, for any matrices $C_0 \in GL_{n_0}(\mathbb{C})$ and $C_1 \in GL_{n_1}(\mathbb{C})$,

$$\frac{1}{\|C_0^{-1}\|\,\|C_1^{-1}\|} s_k(B) \le s_k(C_1BC_0) \le \|C_0\|\,\|C_1\| s_k(B), \tag{5-4}$$

where $\|A\| = (\max \sigma(AA^T))^{1/2}$ denotes the spectral radius of a matrix A. The inequalities (5-3) are specific and simple cases of the Ky Fan inequalities (see, for example, [Simon 1979] for a generalization). In particular, when $C_p^*C_p = \mathrm{Id}_{n_p} + \mathcal{O}(\varepsilon)$ ($p = 0, 1$), the k-th singular value of B is close to the k-th singular value of C_1BC_0, that is, $s_k(C_1BC_0) = s_k(B)(1+\mathcal{O}(\varepsilon))$. In particular, computing the singular values of β in almost orthonormal bases (according to Definition 3.6) changes every $s_k(\beta)$ into $s_k(\beta)(1+\mathcal{O}(e^{-c/h}))$. To analyze the singular values of β, we will use the almost orthonormal bases built in the previous section.

Remark 5.2. Our approach, which emphasizes the differential $d_{f,h}$ and allows almost orthonormal changes of bases, is very close to [Bismut and Zhang 1994] (see in particular their Section 6), where an isomorphism between the Thom–Smale complex and the Witten complex is constructed.[2] The interest of our technique, following [Helffer and Nier 2006; Le Peutrec 2010b; Le Peutrec et al. 2013], is that the

[2]F. Nier thanks J. M. Bismut for mentioning this point.

hierarchy of long range tunnel effects can be analyzed accurately using a Gauss elimination algorithm (see [Le Peutrec 2009]). This makes more explicit the inductive process which was used by Bovier, Eckhoff, Gayrard and Klein [Bovier et al. 2004; 2005]. Actually, the present analysis shows that the Thom–Smale transversality condition and the Morse condition are not necessary: introducing the suitable block structure associated with the assumed geometry of the tunnel effect (see in particular Hypothesis 2) suffices.

5A. ***Structure of*** $\boldsymbol{\beta}$**.** The estimates $\|v_j^{(0)}-\Pi^{(0)}v_j^{(0)}\|_{L^2(\Omega_+)}=\mathcal{O}(e^{-c/h})$ and $\|v_k^{(1)}-\Pi^{(1)}v_k^{(1)}\|_{L^2(\Omega_+)}=\mathcal{O}(e^{-c/h})$ of Proposition 4.5 together with the results stated in Proposition 4.5(1) ensure that

$$\mathcal{B}^{(0)}=(\Pi^{(0)}v_j^{(0)})_{1\le j\le m_0^D(\Omega_+)}\quad\text{and}\quad \mathcal{B}^{(1)}=(\Pi^{(1)}v_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$$

are almost orthonormal bases of $F^{(0)}$ and $F^{(1)}$. The same holds for their dual bases (in $L^2(\Omega_+)$), denoted by $\mathcal{B}^{(0),*}$ and $\mathcal{B}^{(1),*}$. The matrix of $\beta=d_{f,h}\big|_{F^{(0)}}:F^{(0)}\to F^{(1)}$ in the bases $\mathcal{B}^{(0)}$, $\mathcal{B}^{(1),*}$ is given by

$$M(\beta,\mathcal{B}^{(0)},\mathcal{B}^{(1),*})=B=(b_{k,j})_{1\le k\le m_1^D(\Omega_+),\,1\le j\le m_0^D(\Omega_+)}\quad\text{with}\quad b_{k,j}=\langle\Pi^{(1)}v_k^{(1)},\beta\Pi^{(0)}v_j^{(0)}\rangle_{L^2(\Omega_+)}.$$

Remember that the coefficients are equivalently written, by using (4-5), as

$$b_{k,j}=\langle\Pi^{(1)}v_k^{(1)},\beta\Pi^{(0)}v_j^{(0)}\rangle_{L^2(\Omega_+)}=\langle\Pi^{(1)}v_k^{(1)},d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)}=\langle v_k^{(1)},d_{f,h}\Pi^{(0)}v_j^{(0)}\rangle_{L^2(\Omega_+)}.\tag{5-5}$$

Following the various cases discussed in Proposition 4.6, where the scalar products $\langle v_k^{(1)},d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)}$ were studied, we shall write the matrix B in block form:

$$B=\begin{pmatrix}B_{1,1}&B_{1,2}\\B_{2,1}&B_{2,2}\end{pmatrix},\quad\text{where}\quad\begin{cases}B_{1,1}=(\langle\Pi^{(1)}v_k^{(1)},d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)})_{1\le k\le m_1^N(\Omega_-)},\\B_{1,2}=(\langle\Pi^{(1)}v_k^{(1)},d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)})_{2\le j\le m_0^D(\Omega_+),\,1\le k\le m_1^N(\Omega_-)},\\B_{2,1}=(\langle\Pi^{(1)}v_k^{(1)},d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)})_{m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)},\\B_{2,2}=(\langle\Pi^{(1)}v_k^{(1)},d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)})_{2\le j\le m_0^D(\Omega_+),\,m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)}\end{cases}$$

In the following, we will give some estimates of each of these blocks in the asymptotic regime $h\to 0$. We let

$$C_0=2\|\nabla f\|_{L^\infty(\operatorname{supp}(\nabla\chi_0))}.\tag{5-6}$$

Notice that $C_0>0$. We assume that $\delta_+>0$ is chosen so that

$$\delta_+<\frac{\kappa_f}{C_0}.\tag{5-7}$$

The assumption (2-8) will be useful to study the blocks $B_{1,2}$ and $B_{2,2}$ and the parameter $\delta_+>0$ (see Figure 1) will be further adjusted when considering the blocks $B_{1,1}$ and $B_{2,1}$.

5B. ***The blocks*** $\boldsymbol{B_{1,2}}$ ***and*** $\boldsymbol{B_{2,2}}$**.** Estimates for both blocks rely on assumption (2-8). Let us start with $B_{1,2}$.

Lemma 5.3. *The coefficients of* $B_{1,2}$ *satisfy*

$$b_{k,j}=\langle\Pi^{(1)}v_k^{(1)},d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)}=\sqrt{\mu_j^{(0)}(\Omega_-)}\,(\delta_{k,j-1}+\mathcal{O}(e^{-c/h}))$$

for $j\in\{2,\dots,m_0^D(\Omega_+)\}$ *and* $k\in\{1,\dots,m_1^N(\Omega_-)\}$.

Proof. Let us first estimate $\|d_{f,h}v_j^{(0)}\|_{L^2(\Omega_+)}$ by writing

$$d_{f,h}v_j^{(0)} = d_{f,h}(\chi_-^{(0)}\psi_j^{(0)}) = \chi_-^{(0)}d_{f,h}\psi_j^{(0)} + h\psi_j^{(0)}d\chi_-^{(0)} = \chi_-^{(0)}\sqrt{\mu_j^{(0)}(\Omega_-)}\psi_{j-1}^{(1)} + h\psi_j^{(0)}d\chi_-^{(0)}.$$

Since $\operatorname{supp} d\chi_-^{(0)} \subset \Omega_- \setminus U \subset \mathcal{V}_-$ (see (5-1)), (2-8) implies $\|d_{f,h}v_j^{(0)}\|_{L^2(\Omega_+)} = \tilde{\mathcal{O}}\big(\sqrt{\mu_j^{(0)}(\Omega_-)}\big)$. The difference

$$|\langle \Pi^{(1)}v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)} - \langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)}|$$

is thus bounded from above by

$$\|\Pi^{(1)}v_k^{(1)} - v_k^{(1)}\|_{L^2(\Omega_+)}\tilde{\mathcal{O}}\big(\sqrt{\mu_j^{(0)}(\Omega_-)}\big) \le Ce^{-c'/(2h)}\sqrt{\mu_j^{(0)}(\Omega_-)},$$

owing to the estimate $\|\Pi^{(1)}v_k^{(1)} - v_k^{(1)}\| = \mathcal{O}(e^{-c'/h})$ obtained in Proposition 4.5(3). The result then comes from the expression of $\langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)}$ given in Proposition 4.6(3). □

The estimate of the block $B_{2,2}$ follows the same lines:

Lemma 5.4. *The coefficients of $B_{2,2}$ satisfy*

$$b_{k,j} = \langle \Pi^{(1)}v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle_{L^2(\Omega_+)} = \mathcal{O}\big(\sqrt{\mu_j^{(0)}(\Omega_-)}e^{-c/h}\big)$$

for $j \in \{2, \dots, m_0^D(\Omega_+)\}$ and $k \in \{m_1^N(\Omega_-)+1, \dots, m_1^D(\Omega_+)\}$.

Proof. Using $\|d_{f,h}v_j^{(0)}\| = \tilde{\mathcal{O}}\big(\sqrt{\mu_j^{(0)}(\Omega_-)}\big)$ again, $\|\Pi^{(1)}v_k^{(1)} - v_k^{(1)}\| = \mathcal{O}(e^{-c'/h})$ and, according to Proposition 4.6(4), $\langle v_k^{(1)}, d_{f,h}v_j^{(0)}\rangle = 0$ we get $|b_{k,j}| \le Ce^{-c'/(2h)}\sqrt{\mu_j^{(0)}(\Omega_-)}$. □

5C. *The block $B_{1,1}$.* In this section, the value of the parameter δ_+ is adjusted. This value will possibly be changed twice more: for the estimate of the block $B_{2,1}$ and in the final proof of Theorem 2.4; see Sections 6A and 6B. Remember that the constant c occurring in the remainders $\mathcal{O}(e^{-c/h})$ introduced in Proposition 4.5 does not depend on $\delta_+ > 0$.

Lemma 5.5. *For any $k \in \{1, \dots, m_1^N(\Omega_-)\}$, the matrix element $b_{k,1}$ satisfies*

$$b_{k,1} = \langle \Pi^{(1)}v_k^1, d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-(\kappa_f + c - C_0\delta_+)/h}),$$

where $\kappa_f = \min_{\partial\Omega_+} f - \min_{\Omega_+} f$, and the constants $c > 0$ and $C_0 > 0$ (defined by (5-6)) are independent of $\delta_+ > 0$. In particular, when $\delta_+ > 0$ is chosen smaller than c/C_0, one gets

$$b_{k,1} = \mathcal{O}(e^{-(\kappa_f + c)/h})$$

for a positive constant c, which depends on δ_+.

Proof. Remember that $v_1^{(0)} = \chi_0\tilde{\psi}_1^{(0)} = \chi_0 e^{-f/h}/\big(\int_{\Omega_+} e^{-2f(x)/h}\,dx\big)^{1/2}$, where $\nabla\chi_0$ is supported in $\{x \in \Omega_+ : d(x, \partial\Omega_+) < \delta_+\}$ (see Figure 1). The Witten differential of $v_1^{(0)}$ satisfies

$$d_{f,h}v_1^{(0)} = \frac{e^{-f/h}}{\big(\int_{\Omega_+} e^{-2f(x)/h}\,dx\big)^{1/2}}(h\,d\chi_0)$$

and its L^2-norm can be estimated by

$$\|d_{f,h}v_1^{(0)}\|^2_{L^2(\Omega_+)} \le C_{\chi_0}\frac{\int_{\mathrm{supp}(\nabla\chi_0)} e^{-2f(x)/h}\,dx}{\int_{\Omega_+} e^{-2f(x)/h}\,dx}.$$

With $f(x) \ge \min_{\partial\Omega_+} f - \frac{1}{2}C_0\delta_+$ for $x \in \mathrm{supp}(\nabla\chi_0)$ (where C_0 is defined by (5-6) and does not depend on δ_+) and the lower bound $\int_{\Omega_+} e^{-2f(x)/h}\,dx \ge h^{d/2}e^{-2(\min_{\Omega_+} f)/h}/C_1$ of Lemma 4.3, we get

$$\|d_{f,h}v_1^{(0)}\|^2_{L^2(\Omega_+)} \le C_1h^{-d/2}e^{-2(\kappa_f - C_0\delta_+/2)/h} \le C_2e^{-2(\kappa_f - C_0\delta_+)/h} \tag{5-8}$$

provided that h is small enough. Then, like in Lemma 5.3, using

$$|b_{k,1} - \langle v_k^{(1)}, d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)}| \le \|\Pi^{(1)}v_k^{(1)} - v_k^{(1)}\|_{L^2(\Omega_+)}\|d_{f,h}v_1^{(0)}\|_{L^2(\Omega_+)} \le C_3e^{-c'/h}e^{-(\kappa_f - C_0\delta_+)/h},$$

the equality $\langle v_k^{(1)}, d_{f,h}v_1^{(0)}\rangle = 0$ (see Proposition 4.6(1)) yields the result. □

Remark 5.6. If $m_1^D(\Omega_+ \setminus \Omega_-) = 0$ (and thus $m_1^N(\Omega_-) = m_1^D(\Omega_+)$), the previous lemma shows that

$$\langle \Pi^{(0)}v_1^{(0)}, \beta^*\beta\Pi^{(0)}v_1^{(0)}\rangle_{F^{(0)}} = \|\beta\Pi^{(0)}v_1^{(0)}\|^2_{F^{(0)}} = \sum_{k=1}^{m_1^N(\Omega_-)} |b_{k,1}|^2(1 + \mathcal{O}(e^{-c/h})) = \mathcal{O}(e^{-(\kappa_f + c)/h}).$$

This implies that $\beta^*\beta$ (and therefore $\Delta_{f,h}^{D,(0)}(\Omega_+)$) has an eigenvalue of the order $\mathcal{O}(e^{-(\kappa_f+c)/h})$, which contradicts Lemma 5.9 below. Therefore, $m_1^D(\Omega_+ \setminus \Omega_-)$ is not zero.

5D. *The block $B_{2,1}$.* We shall first give an approximate expression for the coefficients of the column $B_{2,1}$.

Proposition 5.7. *For any $k \in \{m_1^N(\Omega_-) + 1, \dots, m_1^D(\Omega_+)\}$, the matrix element*

$$b_{k,1} = \langle \Pi^{(1)}v_k^1, d_{f,h}v_1^{(0)}\rangle_{L^2(\Omega_+)}$$

satisfies

$$b_{k,1} = -\frac{h\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}} + \mathcal{O}(e^{-(\kappa_f+c)/h}), \tag{5-9}$$

where c is a positive constant which depends on $\delta_+ > 0$ chosen to be sufficiently small, and $\kappa_f = \min_{\partial\Omega_+} f - \min_{\Omega_+} f$. Moreover, these coefficients $b_{k,1}$ satisfy

$$\lim_{h\to 0} h\log\left[\sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} |b_{k,1}|^2\right] = -2\kappa_f. \tag{5-10}$$

The estimate (5-10) shows that the approximation (5-9) is meaningful, in the sense that some of the coefficients $b_{k,1}$ are indeed larger than the error term $\mathcal{O}(e^{-(\kappa_f+c)/h})$. In particular, we have

$$\sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} |b_{k,1}|^2 = \frac{h^2\sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)}\left(\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma\right)^2}{\int_{\Omega_+} e^{-2f(x)/h}\,dx}(1 + \mathcal{O}(e^{-c/h})). \tag{5-11}$$

Proof. The first statement is proved like in Lemma 5.5, after recalling

$$\langle v_k^{(1)}, d_{f,h} v_1^{(0)}\rangle = -\frac{h\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}},$$

according to Proposition 4.6(2).

For the equality (5-10), the upper bound

$$\limsup_{h\to 0} h\log\left[\sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} |b_{k,1}|^2\right] \le -2\kappa_f$$

is a consequence of

$$\left|\frac{\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}}\right| \le C\frac{\left(\int_{\partial\Omega_+} |\boldsymbol{i}_n\psi_k^{(1)}(\sigma)|^2\,d\sigma\right)^{1/2}}{\left(\int_{\Omega_+} e^{-2(f(x)-\min_{\Omega_+} f)/h}\,dx\right)^{1/2}} e^{-\kappa_f/h},$$

where the denominator is bounded from below by Lemma 4.3. The numerator is estimated by

$$\|\psi_k^{(1)}|_{\partial\Omega_+}\|_{L^2(\partial\Omega_+)} \le C\|\psi_k^{(1)}\|_{W^{1,2}(\mathcal{V})} = \mathcal{O}(h^{-1}) = \tilde{\mathcal{O}}(1)$$

owing to Proposition 3.8, since $d_{\mathrm{Ag}}(x,\mathcal{V}) = 0$ for $x\in\mathcal{V}$. Using Lemma 5.5, the lower bound for (5-10) is equivalent to

$$\liminf_{h\to 0} h\log\left[\sum_{k=1}^{m_1^D(\Omega_+)} |b_{k,1}|^2\right] \ge -2\kappa_f. \tag{5-12}$$

Since $b_{k,1} = \langle \Pi^{(1)} v_k^{(1)}, d_{f,h}\Pi^{(0)} v_1^{(0)}\rangle_{L^2(\Omega_+)}$ is the k-th component of $d_{f,h}\Pi^{(0)} v_1^{(0)} \in F^{(1)}$ in the almost orthonormal basis $\mathcal{B}^{(1),*}$ of $F^{(1)}$, the inequality (5-12) is equivalent to

$$\liminf_{h\to 0} h\log(\|d_{f,h}\Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)}) = \liminf_{h\to 0} h\log\bigl(\langle \Pi^{(0)} v_1^{(0)}, \Delta_{f,h}^{D,(0)}(\Omega_+)\Pi^{(0)} v_1^{(0)}\rangle_{L^2(\Omega_+)}\bigr) \ge -2\kappa_f.$$

With $\|\Pi^{(0)} v_1^{(0)}\|_{L^2(\Omega_+)} = 1 + \mathcal{O}(e^{-c/h})$, the last inequality is a consequence of

$$\liminf_{h\to 0} h\log\bigl[\min\sigma(\Delta_{f,h}^{D,(0)}(\Omega_+))\bigr] \ge -2\kappa_f,$$

which is proved in the next lemma. □

Remark 5.8. Using Lemma 5.5, the asymptotic result (5-10) is actually equivalent to

$$\lim_{h\to 0} h\log\left[\sum_{k=1}^{m_1^D(\Omega_+)} |b_{k,1}|^2\right] = -2\kappa_f.$$

We end this section with an estimate on the bottom of the spectrum of $\Delta_{f,h}^{D,(0)}(\Omega_+)$, which was used to conclude the proof of Proposition 5.7 above.

Lemma 5.9. *The bottom of the spectrum of* $\Delta^{D,(0)}_{f,h}(\Omega_+)$ *satisfies*

$$\lim_{h\to 0} h\log\bigl[\min\sigma(\Delta^{D,(0)}_{f,h}(\Omega_+))\bigr] = -2\kappa_f.$$

In particular, we have

$$\forall\varepsilon>0\ \exists C_\varepsilon>1\ \exists h_\varepsilon>0\ \forall h\in(0,h_\varepsilon]\quad \min\sigma(\Delta^{D,(0)}_{f,h}(\Omega_+))\geq \frac{1}{C_\varepsilon}e^{-2(\kappa_f+\varepsilon)/h}.$$

Proof. Let us introduce a function $w_1^{(0)}$ defined similarly to $v_1^{(0)}$ by $w_1^{(0)}=\tilde\chi_0\tilde\psi_1^{(0)}$, where $\tilde\chi_0$ is a $\mathcal{C}_0^\infty(\Omega_+)$ function, equal to 1 in a neighborhood of Ω_- and such that $d\tilde\chi_0$ is supported in $\{x\in\Omega_+ : d(x,\partial\Omega_+)\leq\delta\}$. The estimate $\limsup_{h\to 0} h\log\bigl[\min\sigma(\Delta^{D,(0)}_{f,h}(\Omega_+))\bigr]\leq -2\kappa_f$ is then a consequence of the computation

$$\langle w_1^{(0)},\Delta_{f,h}w_1^{(0)}\rangle_{L^2(\Omega_+)} = \|d_{f,h}w_1^{(0)}\|^2_{L^2(\Omega_+)} = \tilde{\mathcal{O}}(e^{-2(\kappa_f-C_0\delta)/h}) \tag{5-13}$$

by considering δ arbitrarily small. The last equality is proved like (5-8) above.

It remains to prove that $\liminf_{h\to 0} h\log\bigl[\min\sigma(\Delta^{D,(0)}_{f,h}(\Omega_+))\bigr]\geq -2\kappa_f$. The proof is very similar to that of Proposition 3.7. Assume on the contrary that there exists $\varepsilon_0>0$ and a sequence h_n such that $\lim_{n\to\infty}h_n=0$ and

$$\min\sigma(\Delta^{D,(0)}_{f,h_n}(\Omega_+))\leq Ce^{-2(\kappa_f+\varepsilon_0)/h_n}.$$

To simplify the notation, let us drop the subscript n in h_n. The previous inequality means that there exists $v_h\in L^2(\Omega_+)$ and $\lambda_h\geq 0$ such that

$$\Delta^{D,(0)}_{f,h}v_h=\lambda_h v_h \quad\text{in }\ \Omega_+,\qquad v_h\big|_{\partial\Omega_+}=0,\qquad \|v_h\|_{L^2(\Omega_+)}=1, \tag{5-14}$$

$$\lambda_h=\langle v_h,\Delta^{D,(0)}_{f,h}(\Omega_+)v_h\rangle_{L^2(\Omega_+)}=\|d_{f,h}v_h\|^2_{L^2(\Omega_+)}\leq Ce^{-2(\kappa_f+\varepsilon_0)/h}. \tag{5-15}$$

For a small $t>0$, let us consider the domain

$$\Omega_t=\bigl\{x\in\Omega_+ : f(x)<\min_{\partial\Omega_+}f+t\bigr\}.$$

With $d_{f,h}=e^{-(f-\min_{\Omega_+}f)/h}(hd)e^{(f-\min_{\Omega_+}f)/h}$, the estimate (5-15) implies

$$\begin{aligned}\|d(e^{(f-\min_{\Omega_+}f)/h}v_h)\|_{L^2(\Omega_t)} &\leq h^{-1}\max_{x\in\Omega_t}e^{(f(x)-\min_{\Omega_+}f)/h}\|d_{f,h}v_h\|_{L^2(\Omega_t)}\\ &\leq Ch^{-1}e^{-(\varepsilon_0-t)/h}=\mathcal{O}(e^{-\varepsilon_0/(2h)})\end{aligned} \tag{5-16}$$

as soon as $t<\frac12\varepsilon_0$.

For a given $t\in\bigl(0,\frac12\varepsilon_0\bigr)$, let us now prove that $\|v_h\|_{L^2(\Omega_t)}$ is close to 1, using the same reasoning as in the proof of Proposition 3.4. There exists is an open neighborhood $\mathcal{V}$ of $\{x\in\Omega_- : \nabla f(x)=0\}$ such that $\mathcal{V}\subset\Omega_t$ and

$$d_{\mathrm{Ag}}(\Omega_+\setminus\Omega_t,\mathcal{V})\geq c>0, \tag{5-17}$$

where c can be chosen independently of t, and ε_0 and is positive according to Hypothesis 2. Applying Lemma 3.1 with $\Omega=\Omega_+$ and $\varphi=(1-\alpha h)d_{\mathrm{Ag}}(\cdot,\mathcal{V})$, one gets, for $h<1/\alpha$ (similarly to (3-3)),

$$0\geq h^2\|d(e^{\varphi/h}v_h)\|^2_{L^2(\Omega_+)}+h\bigl[\alpha\langle e^{\varphi/h}v_h,|\nabla f|^2e^{\varphi/h}v_h\rangle_{L^2(\Omega_+)}-C_f\|e^{\varphi/h}v_h\|^2_{L^2(\Omega_+)}\bigr].$$

By choosing α sufficiently large that $\alpha \min_{\Omega_+\setminus\mathcal{V}} |\nabla f|^2 \geq 2C_f$, we get

$$0 \geq h^2\|d(e^{\varphi/h}v_h)\|^2_{L^2(\Omega_+)} + h\big[C_f\|e^{\varphi/h}v_h\|^2_{L^2(\Omega_+\setminus\mathcal{V})} - C_f\|e^{\varphi/h}v_h\|^2_{L^2(\mathcal{V})}\big].$$

Using the fact that $\|e^{\varphi/h}v_h\|^2_{L^2(\mathcal{V})} = \|v_h\|^2_{L^2(\mathcal{V})} \leq 1$, we obtain, by adding $2C_f h\|v_h\|^2_{L^2(\mathcal{V})}$ on both sides of the previous inequality,

$$2C_f h \geq 2C_f h\|v_h\|^2_{L^2(\mathcal{V})} \geq h^2\|d(e^{\varphi/h}v_h)\|^2_{L^2(\Omega_+)} + hC_f\|e^{\varphi/h}v_h\|^2_{L^2(\Omega_+)}.$$

This implies, in particular,

$$\|e^{d_{\mathrm{Ag}}(\cdot,\mathcal{V})/h}v_h\|^2_{L^2(\Omega_+)} \leq 2,$$

and thus, using (5-17),

$$\|v_h\|^2_{L^2(\Omega_+\setminus\Omega_t)} \leq Ce^{-c/h}.$$

This implies

$$\|e^{(f-\min_{\Omega_+} f)/h}v_h\|_{L^2(\Omega_t)} \geq \|v_h\|_{L^2(\Omega_t)} \geq 1 - Ce^{-c/h}, \tag{5-18}$$

where, we recall, c is independent of t and ε_0, supposed to be small enough.

The two estimates (5-16) and (5-18) lead to a contradiction. Indeed, let us now set $t = \frac{1}{4}\varepsilon_0$. The Poincaré–Wirtinger inequality or, equivalently, the spectral gap estimate for the Neumann Laplacian in $\Omega_{\varepsilon_0/4}$, implies that there exists a constant C_h such that

$$\|(e^{(f-\min_{\Omega_+} f)/h}v_h) - C_h\|_{L^2(\Omega_{\varepsilon_0/4})} = \mathcal{O}(e^{-\varepsilon_0/(2h)}),$$

and therefore

$$\|(e^{(f-\min_{\Omega_+} f)/h}v_h) - C_h\|_{W^{1,2}(\Omega_{\varepsilon_0/4})} = \mathcal{O}(e^{-\varepsilon_0/(2h)}).$$

Since $\Omega_{\varepsilon_0/4}\cap\partial\Omega_+$ has a nonempty interior U_{ε_0}, the trace theorem implies

$$\|(e^{(f-\min_{\Omega_+} f)/h}v_h) - C_h\|_{L^2(U_{\varepsilon_0})} = \mathcal{O}(e^{-\varepsilon_0/(2h)}).$$

Since $v_h\big|_{\partial\Omega_+} \equiv 0$ and since U_{ε_0} is fixed by ε_0 and independent of h, this implies $C_h = \mathcal{O}(e^{-\varepsilon_0/(2h)})$. We are led to

$$1 - Ce^{-c/h} \leq \|v_h\|_{L^2(\Omega_{\varepsilon_0/4})} \leq \|e^{(f-\min_{\Omega_+} f)/h}v_h\|_{L^2(\Omega_{\varepsilon_0/4})} \leq \|C_h\|_{L^2(\Omega_{\varepsilon_0/4})} + Ce^{-\varepsilon_0/(2h)} \leq C'e^{-\varepsilon_0/(2h)},$$

which is impossible when h is small enough. □

This lemma shows the equality (2-11) stated in Theorem 2.4.

5E. ***Singular values of β.*** We are now in position to complete the proof of Proposition 5.1.

Proof of Proposition 5.1. Let $e^{(0)} = (e^{(0)}_1, \dots, e^{(0)}_{m^D_0(\Omega_+)})$ (resp. $e^{(1)} = (e^{(1)}_1, \dots, e^{(1)}_{m^D_1(\Omega_+)})$) denote an orthonormal basis of $F^{(0)}$ (resp. of $F^{(1)}$) and let C_0 (resp. C_1) be the matrix of the change of basis from $e^{(0)}$ (resp. from $\mathcal{B}^{(1),*}$) to $\mathcal{B}^{(0)}$ (resp. to $e^{(1)}$). Let $A = M(\beta, e^{(0)}, e^{(1)})$ denote the matrix of β in the bases $e^{(0)}$ and $e^{(1)}$, so that

$$A = C_1BC_0,$$

where, we recall, $B = M(\beta, \mathcal{B}^{(0)}, \mathcal{B}^{(1),*})$. Using the fact that $\mathcal{B}^{(0)}$ and $\mathcal{B}^{(1)}$ are almost orthonormal bases, the matrices C_0 and C_1 satisfy $C_p^* C_p = \mathrm{Id} + \mathcal{O}(\varepsilon)$, so that (according to (5-4))

$$s_j(\beta) = s_j(A) = s_j(C_1 B C_0) = s_j(B)(1 + \mathcal{O}(e^{-c/h})).$$

The singular values of β can be understood from those of B, up to exponentially small relative errors.

Now Lemmas 5.3, 5.4, 5.5 and Proposition 5.7 can be gathered (using the block structure of B introduced in Section 5A), in the asymptotic regime $h \to 0$, as

$$B = \left(\begin{array}{c|cccc}
\mathcal{O}(b_{k_0,1}e^{-c/h}) & b_{1,2} & \mathcal{O}(b_{2,3}e^{-c/h}) & \cdots & \mathcal{O}(b_{m_0^D-1,m_0^D}e^{-c/h}) \\
\vdots & \mathcal{O}(b_{1,2}e^{-c/h}) & b_{2,3} & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \mathcal{O}(b_{m_0^D-1,m_0^D}e^{-c/h}) \\
\mathcal{O}(b_{k_0,1}e^{-c/h}) & \mathcal{O}(b_{1,2}e^{-c/h}) & \mathcal{O}(b_{2,3}e^{-c/h}) & \cdots & b_{m_0^D-1,m_0^D} \\
\mathcal{O}(b_{k_0,1}e^{-c/h}) & \mathcal{O}(b_{1,2}e^{-c/h}) & \mathcal{O}(b_{2,3}e^{-c/h}) & \cdots & \mathcal{O}(b_{m_0^D-1,m_0^D}e^{-c/h}) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mathcal{O}(b_{k_0,1}e^{-c/h}) & \mathcal{O}(b_{1,2}e^{-c/h}) & \mathcal{O}(b_{2,3}e^{-c/h}) & \cdots & \mathcal{O}(b_{m_0^D-1,m_0^D}e^{-c/h}) \\
\hline
b_{m_1^N+1,1} & \mathcal{O}(b_{1,2}e^{-c/h}) & \mathcal{O}(b_{2,3}e^{-c/h}) & \cdots & \mathcal{O}(b_{m_0^D-1,m_0^D}e^{-c/h}) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
b_{m_1^D,1} & \mathcal{O}(b_{1,2}e^{-c/h}) & \mathcal{O}(b_{2,3}e^{-c/h}) & \cdots & \mathcal{O}(b_{m_0^D-1,m_0^D}e^{-c/h})
\end{array}\right),$$

where we used m_0^D (resp. m_1^N, m_1^D) instead of $m_0^D(\Omega_+)$ (resp. $m_1^N(\Omega_-)$, $m_1^D(\Omega_+)$) and where k_0 is a (possibly h-dependent) index such that $|b_{k_0,1}| = \max_{m_1^N+1 \le k \le m_1^D} |b_{k,1}|$. By Gaussian elimination (see [Le Peutrec 2009] for more details), one can find a matrix $R \in \mathcal{M}_{m_1^D}(\mathbb{R})$ with $\|R\| = \mathcal{O}(e^{-c/h})$ such that

$$(\mathrm{Id}_{m_1^D} + R)B = \tilde{B} = \left(\begin{array}{c|c}
0(m_0^D-1,1) & \tilde{B}_{1,2} \\
0(m_1^N - m_0^D + 1, 1) & 0(m_1^N - m_0^D + 1, m_0^D - 1) \\
\hline
\tilde{B}_{3,1} & 0(m_1^D - m_1^N, m_0^D - 1)
\end{array}\right)$$

with

$$\tilde{B}_{3,1} = \begin{pmatrix} b_{m_1^N+1,1} \\ \vdots \\ b_{m_1^D,1} \end{pmatrix} \quad \text{and} \quad \tilde{B}_{1,2} = \begin{pmatrix}
b_{1,2}(1+\mathcal{O}(e^{-c/h})) & 0 & \cdots & 0 \\
0 & \ddots & & \vdots \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & b_{m_0^D-1,m_0^D}(1+\mathcal{O}(e^{-c/h}))
\end{pmatrix},$$

where $0(i,j)$ is the null matrix in $\mathcal{M}_{i,j}(\mathbb{R})$. We deduce that the singular values of B are approximated (up to exponentially small relative error terms) by the ones of $\tilde{B}$, which are given by its block structure. We find (recall that the singular values are labeled in decreasing order):

$$s_j(B) = |b_{m_0^D-j,m_0^D-j+1}|(1+\mathcal{O}(e^{-c/h})) \quad \text{for } j \in \{1, \dots, m_0^D - 1\}$$

and

$$s_{m_0^D}(B)^2 = \Bigg[\sum_{k=m_1^N+1}^{m_1^D} |b_{k,1}|^2 \Bigg] (1+\mathcal{O}(e^{-c/h})).$$

We conclude the proof of Proposition 5.1 using the approximate values of $b_{k,k+1}$ ($k \in \{1,\dots,m_0^D-1\}$) and $b_{k,1}$ ($k \in \{m_1^N+1,\dots,m_1^D\}$) given in Lemma 5.3 and Proposition 5.7:

$$|b_{m_0^D-j,m_0^D-j+1}| = \sqrt{\mu^{(0)}_{m_0^D-j+1}(\Omega_-)}(1+\mathcal{O}(e^{-c/h})) \quad \text{for } j \in \{1,\dots,m_0^D-1\},$$

$$\sum_{k=m_1^N+1}^{m_1^D} |b_{k,1}|^2 = \frac{h^2 \sum_{k=m_1^N+1}^{m_1^D} \big(\int_{\partial\Omega_+} e^{-f(\sigma)/h} \boldsymbol{i}_n \psi_k^{(1)}(\sigma)\, d\sigma\big)^2}{\int_{\Omega_+} e^{-2f(x)/h}\, dx} + \mathcal{O}(e^{-(2\kappa_f+c)/h}).$$

In particular, for h small enough, we indeed have

$$|b_{m_0^D-1,m_0^D}|^2 \ge \cdots \ge |b_{1,2}|^2 \ge \sum_{k=m_1^N+1}^{m_1^D} |b_{k,1}|^2,$$

the last inequality being a consequence of (5-10) and $|b_{1,2}|^2 = \mu_2^{(0)}(\Omega_-)(1+\mathcal{O}(e^{-c/h})) \ge C_\varepsilon e^{-2(\kappa_f-c_0)/h}$ using Proposition 3.7. □

6. Proof of Theorem 2.4 and two corollaries

Proposition 5.1 already provides a precise asymptotic result on the exponentially small eigenvalues of $\Delta^{D,(0)}_{f,h}(\Omega_+)$, using (5-2):

$$\lambda_j^{(0)}(\Omega_+) = s_{m_0^D(\Omega_+)+1-j}(\beta)^2 = \mu_j^{(0)}(\Omega_-)(1+\mathcal{O}(e^{-c/h})) \quad \text{for } j \in \{2,\dots,m_0^D(\Omega_+)\}, \tag{6-1}$$

$$\lambda_1^{(0)}(\Omega_+) = s_{m_0^D(\Omega_+)}(\beta)^2 = \frac{h^2 \sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} \big(\int_{\partial\Omega_+} e^{-f(\sigma)/h} \boldsymbol{i}_n \psi_k^{(1)}(\sigma)\, d\sigma\big)^2}{\int_{\Omega_+} e^{-2f(x)/h}\, dx}(1+\mathcal{O}(e^{-c/h})), \tag{6-2}$$

the second estimate being a consequence of Proposition 5.7 (see (5-11)). This is essentially the result of Theorem 2.4 about $\lambda_1^{(0)}(\Omega_+)$ (see (2-12)); it remains to show that the basis $(\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)}$ in (6-2) (which was introduced in Section 4B) can be replaced by *any* orthonormal basis $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\Omega_-)}$ of $\operatorname{Ran} 1_{[0,\nu(h)]}(\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-))$. This will be done in Section 6C.

In addition, it also remains to prove the estimates (2-13) and (2-14) on the eigenvector $u_1^{(0)}$ associated with the smallest eigenvalue $\lambda_1^{(0)}(\Omega_+)$. This will be the subject of Sections 6A and 6B. We recall that the spectral subspace associated with $\lambda_1^{(0)}(\Omega_+)$ is one-dimensional (since $\lambda_2^{(0)}(\Omega_+) \ge \lambda_1^{(0)}(\Omega_+)e^{c/h}$). We thus have

$$u_1^{(0)} = \frac{\Pi_0 v_1^{(0)}}{\|\Pi_0 v_1^{(0)}\|_{L^2(\Omega_+)}}, \tag{6-3}$$

where Π_0 denotes the spectral projection associated with $\lambda_1^{(0)}(\Omega_+)$:

$$\Pi_0 = 1_{\{\lambda_1^{(0)}(\Omega_+)\}}(\Delta_{f,h}^{D,(0)}(\Omega_+)). \tag{6-4}$$

The fact that $\Pi_0 v_1^{(0)} \neq 0$ follows from the fact that $\Pi_0\Pi^{(0)} = \Pi_0$ and the estimate, for small h,

$$\begin{aligned}\frac{\langle \Pi^{(0)} v_1^{(0)}, \Delta_{f,h}^{D,(0)} \Pi^{(0)} v_1^{(0)}\rangle_{L^2(\Omega_+)}}{\|\Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)}} &= \frac{\|d_{f,h}\Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)}}{\|\Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)}} = \|\beta \Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)}(1+\mathcal{O}(e^{-c/h}))\\ &= \sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} |b_{k,1}|^2(1+\mathcal{O}(e^{-c/h}))\\ &= \lambda_1^{(0)}(\Omega_+)(1+\mathcal{O}(e^{-c/h})) \le \lambda_2^{(0)}(\Omega_+)e^{-c/h}\end{aligned} \tag{6-5}$$

for some positive constant c. The second and third equalities are consequences of the almost orthonormality of the bases $\mathcal{B}^{(0)}$ and $\mathcal{B}^{(1),*}$ (see Proposition 4.5). The third one comes from (6-2) and (5-11). The last inequality is a consequence of (6-1) and Proposition 3.7.

Finally, Section 6D is devoted to two corollaries of Theorem 2.4.

6A. *Approximation of* $u_1^{(0)}$.

Let us first prove the estimate (2-13) on $u_1^{(0)}$.

Proposition 6.1. *There exists $c > 0$ such that*

$$\left\| u_1^{(0)} - \frac{e^{-f/h}}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}} \right\|_{W^{2,2}(\Omega_+)} = \mathcal{O}(e^{-c/h}).$$

Proof. Since $\left\| v_1^{(0)} - e^{-f/h}/\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}\right\|_{W^{2,2}(\Omega_+)} = \mathcal{O}(e^{-c/h})$ (which is a simple consequence of Lemma 4.3), it suffices to prove $\|u_1^{(0)} - v_1^{(0)}\|_{W^{2,2}(\Omega_+)} = \mathcal{O}(e^{-c/h})$.

Let us first prove the result in the $L^2(\Omega_+)$-norm. From (6-5), we have $\|d_{f,h}\Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)} \le \lambda_2^{(0)}(\Omega_+)e^{-c/h}$, and thus

$$\begin{aligned}\lambda_2^{(0)}(\Omega_+)\left\| 1_{[\lambda_2^{(0)}(\Omega_+),+\infty)}(\Delta_{f,h}^{D,(0)}(\Omega_+))\Pi^{(0)} v_1^{(0)}\right\|^2_{L^2(\Omega_+)} &\le \langle \Pi^{(0)} v_1^{(0)}, \Delta_{f,h}^{D,(0)}(\Omega_+)\Pi^{(0)} v_1^{(0)}\rangle_{L^2(\Omega_+)}\\ &\le \lambda_2^{(0)}(\Omega_+)e^{-c/h}.\end{aligned}$$

Since $\Pi_0 = \Pi_0\Pi^{(0)}$, we deduce

$$\left\|\Pi_0 v_1^{(0)} - \Pi^{(0)} v_1^{(0)}\right\|_{L^2(\Omega_+)} = \left\| 1_{[\lambda_2^{(0)}(\Omega_+),+\infty)}(\Delta_{f,h}^{D,(0)}(\Omega_+))\Pi^{(0)} v_1^{(0)}\right\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h}).$$

Using in addition the facts that $\|\Pi^{(0)} v_1^{(0)} - v_1^{(0)}\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$ and $\|v_1^{(0)}\|_{L^2(\Omega_+)} = 1+\mathcal{O}(e^{-c/h})$ (see Proposition 4.5), this proves

$$\|u_1^{(0)} - v_1^{(0)}\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c'/h}). \tag{6-6}$$

The estimate in the $W^{2,2}(\Omega_+)$-norm is then obtained by a bootstrap argument that will be used many times again below. The following equations hold:

$$\begin{cases} \Delta^{(0)}_{f,h} u^{(0)}_1 = \lambda^{(0)}_1(\Omega_+) u^{(0)}_1, \\ u^{(0)}_1\big|_{\partial\Omega_+} = 0, \end{cases} \quad \text{and} \quad \begin{cases} \Delta^{(0)}_{f,h} v^{(0)}_1 = g_h, \\ v^{(0)}_1\big|_{\partial\Omega_+} = 0, \end{cases}$$

where g_h is defined by the equation $g_h = \Delta^{(0)}_{f,h} v^{(0)}_1$ and, using the same arguments as in the proof of (5-8), $\|g_h\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-(\kappa_f - C_0\delta_+)/h})$. Recall that, by the assumption (5-7), δ_+ is small enough that $C_0\delta_+ < \kappa_f$, and thus $\|g_h\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$. We then deduce that, with Δ_{H} denoting the Hodge Laplacian (A-3), $u^{(0)}_1 - v^{(0)}_1$ solves

$$\begin{cases} \Delta^{(0)}_{\mathrm{H}} (u^{(0)}_1 - v^{(0)}_1) = \tilde{g}_h, \\ (u^{(0)}_1 - v^{(0)}_1)\big|_{\partial\Omega_+} = 0. \end{cases}$$

Again, $\tilde{g}_h$ is defined by the first equation. Using the formula (A-6), which relates the Hodge and the Witten Laplacians and the estimate (6-6), $\|\tilde{g}_h\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c'/h})$. The elliptic regularity of the Dirichlet Hodge Laplacian then implies $\|u^{(0)}_1 - v^{(0)}_1\|_{W^{2,2}(\Omega_+)} = \mathcal{O}(e^{-c'/h})$. □

6B. ***Approximation of*** $\boldsymbol{d_{f,h} u^{(0)}_1}$**.** We now consider $d_{f,h} u^{(0)}_1$. In this section, we will first prove (2-14) using for the $u^{(1)}_k$ the special basis considered in Section 5. This will be generalized to any orthonormal basis of $\operatorname{Ran} 1_{[0,\nu(h)]}(\Delta^{D,(1)}_{f,h}(\Omega_+ \setminus \overline{\Omega}_-))$ in the next section.

Let us start with an estimate in the $L^2(\Omega_+)$-norm.

Proposition 6.2. *Let $\mathcal{B}^*_1 = (w_k)_{1 \le k \le m^D_1(\Omega_+)}$ be the basis of $F^{(1)} = \operatorname{Ran} 1_{[0,\nu(h)]}(\Delta^{D,(1)}_{f,h}(\Omega_+))$ dual (in $L^2(\Omega_+)$) to $\mathcal{B}_1 = (\Pi^{(1)} v^{(1)}_k)_{1 \le k \le m^D_1(\Omega_+)}$. Then the eigenvector $u^{(0)}_1$ of $\Delta^{D,(0)}_{f,h}(\Omega_+)$ given by* (6-3) *satisfies*

$$\Bigg\| d_{f,h} u^{(0)}_1 - \sum_{k=m^N_1(\Omega_-)+1}^{m^D_1(\Omega_+)} b_{k,1} w_k \Bigg\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-(\kappa_f + c)/h}) \tag{6-7}$$

for some $c > 0$ and where the coefficients $b_{k,1}$ are defined by (5-5).

Proof. By definition of the matrix $B = M(\beta, \mathcal{B}^{(0)}, \mathcal{B}^{(1)*})$,

$$d_{f,h}(\Pi^{(0)} v^{(0)}_1) = \beta(\Pi^{(0)} v^{(0)}_1) = \sum_{k=1}^{m^D_1(\Omega_+)} b_{k,1} w_k = \sum_{k=m^N_1(\Omega_-)+1}^{m^D_1(\Omega_+)} b_{k,1} w_k + r_h$$

with $\|r_h\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-(\kappa_f + c)/h})$, this estimate being a consequence of the almost orthonormality of the one-forms w_k, and of Lemma 5.5. Equation (6-7) is thus equivalent to:

$$\big\| d_{f,h}(u^{(0)}_1 - \Pi^{(0)} v^{(0)}_1) \big\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-(\kappa_f + c)/h}).$$

Notice that

$$u^{(0)}_1 - \Pi^{(0)} v^{(0)}_1 = \|\Pi_0 v^{(0)}_1\|^{-1}_{L^2(\Omega_+)} (\Pi_0 - \Pi^{(0)}) v^{(0)}_1 + (\|\Pi_0 v^{(0)}_1\|^{-1}_{L^2(\Omega_+)} - 1) \Pi^{(0)} v^{(0)}_1.$$

We recall that $\|\Pi_0 v_1^{(0)}\|_{L^2(\Omega_+)} = 1+\mathcal{O}(e^{-c/h})$ and $\|d_{f,h}\Pi^{(0)} v_1^{(0)}\|_{L^2(\Omega_+)} = \|\beta\Pi^{(0)} v_1^{(0)}\|_{L^2(\Omega_+)} = \tilde{\mathcal{O}}(e^{-\kappa_f/h})$ (see (6-5)). This implies that

$$\left\|d_{f,h}(u_1^{(0)} - \Pi^{(0)} v_1^{(0)})\right\|_{L^2(\Omega_+)} = \left\|d_{f,h}((\Pi_0 - \Pi^{(0)}) v_1^{(0)})\right\|_{L^2(\Omega_+)}(1+\mathcal{O}(e^{-c/h})) + \mathcal{O}(e^{-(\kappa_f+c)/h}).$$

Moreover, using the fact that $\Pi_0\Pi^{(0)} = \Pi_0$ and $\Pi^{(0)} - \Pi_0 = 1_{[\lambda_2^{(0)}(\Omega_+),+\infty)}(\Delta_{f,h}^{D,(0)}(\Omega_+))$ commutes with $\Delta_{f,h}^{D,(0)}(\Omega_+)$,

$$\begin{aligned}
\|d_{f,h}((\Pi_0 - \Pi^{(0)}) v_1^{(0)})\|^2_{L^2(\Omega_+)} &= \langle(\Pi^{(0)} - \Pi_0)v_1^{(0)}, \Delta_{f,h}^{D,(0)}(\Omega_+)(\Pi^{(0)} - \Pi_0)v_1^{(0)}\rangle_{L^2(\Omega_+)}\\
&= \|\beta\Pi^{(0)} v_1^{(0)}\|^2_{L^2(\Omega_+)} - \lambda_1^{(0)}(\Omega_+)\|\Pi_0 v_1^{(0)}\|^2_{L^2(\Omega_+)}\\
&= \lambda_1^{(0)}(\Omega_+)(1+\mathcal{O}(e^{-c/h})) - \lambda_1^{(0)}(\Omega_+)(1+\mathcal{O}(e^{-c/h}))\\
&= \mathcal{O}(e^{-2(\kappa_f+c')/h}).
\end{aligned}$$

The third equality is obtained from (6-5) and the last one from the estimate on the bottom of the spectrum in Lemma 5.9. This concludes the proof of (6-7). □

To perform a bootstrap argument to extend the previous result to stronger norms, we need an intermediate lemma:

Lemma 6.3. *For any $p \in \mathbb{N}$, there exists $C_p > 0$ and $N_p \in \mathbb{N}$ such that*

$$\|u\|_{W^{p,2}(\Omega_+)} \le C_p h^{-N_p}\|u\|_{L^2(\Omega_+)} \quad \text{for all } u \in F^{(1)} = \operatorname{Ran} 1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+)).$$

Proof. Let us introduce an orthonormal basis $(e_k)_{1\le k\le m_1^D(\Omega_+)}$ of eigenvectors of $\Delta_{f,h}^{D,(1)}(\Omega_+)$ associated with the small eigenvalues $\lambda_k^{(1)}(\Omega_+) \le \nu(h)$, so $\Delta_{f,h}^{D,(1)} e_k = \lambda_k^{(1)} e_k$. We have

$$\|d_{f,h}e_k\|^2_{L^2(\Omega_+)} + \|d_{f,h}^* e_k\|^2_{L^2(\Omega_+)} = \lambda_k^{(1)} \le \nu(h).$$

For any $u \in F^{(1)}$, there exist some reals $(u_k)_{1\le k\le m_1^D(\Omega_+)}$ such that

$$u = \sum_{k=1}^{m_1^D(\Omega_+)} u_k e_k \quad \text{with} \quad \sum_{k=1}^{m_1^D(\Omega_+)} |u_k|^2 = \|u\|^2_{L^2(\Omega_+)}.$$

Lemma 6.3 will be proven if one can show that, for all $p \in \mathbb{N}$, there exist $C_p > 0$ and $N_p \in \mathbb{N}$ such that $\|e_k\|_{W^{p,2}(\Omega_+)} \le C_p h^{-N_p}$ for all $k \in \{1, \dots, m_1^D(\Omega_+)\}$. From

$$4\||\nabla f|e_k\|^2_{L^2(\Omega_+)} + 2\|d_{f,h}e_k\|^2_{L^2(\Omega_+)} + 2\|d_{f,h}^* e_k\|^2_{L^2(\Omega_+)} \ge h^2[\|de_k\|^2_{L^2(\Omega_+)} + \|d^* e_k\|^2_{L^2(\Omega_+)}]$$

(which is obtained from the formulas (A-4) and (A-5) that relate $d_{f,h}$ to d and $d_{f,h}^*$ to d^*), we deduce $\|e_k\|_{W^{1,2}(\Omega_+)} \le Ch^{-1}$. Then the equation $\Delta_{f,h}^{D,(1)}(\Omega_+)e_k = \lambda_k^{(1)} e_k$ can be written

$$\begin{cases}\Delta_H^{(1)} e_k = r_k(h)\\ \mathbf{t}e_k\big|_{\partial\Omega_+} = 0, \quad \mathbf{t}d^* e_k\big|_{\partial\Omega_+} = \rho_k(h)\end{cases}$$

with $\|r_k(h)\|_{L^2(\Omega_+)} + \|\rho_k(h)\|_{W^{1/2,2}(\partial\Omega_+)} = \mathcal{O}(h^{-2})$. The estimate on $\rho_k(h)$ follows from $0 = \mathbf{t}d_{f,h}^* e_k = h\mathbf{t}d^* e_k + \boldsymbol{i}_{\nabla f} e_k$, so that $\|\rho_k(h)\|_{W^{1/2,2}(\partial\Omega_+)} = h^{-1}\|\boldsymbol{i}_{\nabla f} e_k\|_{W^{1/2,2}(\partial\Omega_+)} \le Ch^{-1}\|e_k\|_{W^{1,2}(\Omega_+)} \le C' h^{-2}$. The estimate on $r_k(h)$ comes from the relation (A-6) between the Hodge and the Witten Laplacians. The elliptic regularity of the above system (see, for example, [Schwarz 1995, Theorem 2.2.6]) implies $\|e_k\|_{W^{2,2}(\Omega_+)} = \mathcal{O}(h^{-2})$. Finally, the result for a general $p \in \mathbb{N}$ is obtained by a bootstrap argument. □

We are now in position to restate the result of Proposition 6.2 in terms of the $W^{p,2}(\mathcal{V})$-norm.

Proposition 6.4. *Let $(\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)}$ be the orthonormal basis of eigenvectors chosen in Section 4B and let χ_+ be the cut-off function of Definition 4.4. For any $p \in \mathbb{N}$, there exists a constant $C_p > 0$ such that*

$$\Big\| d_{f,h}u_1^{(0)} - \sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} b_{k,1}\psi_k^{(1)} \Big\|_{W^{p,2}(\mathcal{V})} \le C_p e^{-(\kappa_f+c)/h},$$

where $\mathcal{V}$ is any neighborhood of $\partial\Omega_+$ contained in $\{\chi_+ = 1\}$, c is a positive constant and, we recall (see Proposition 5.7), the coefficients $b_{k,1}$ defined by (5-5) *satisfy*

$$b_{k,1} = -\frac{h\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n\psi_k^{(1)}(\sigma)\,d\sigma}{\big(\int_{\Omega_+} e^{-2f(x)/h}\,dx\big)^{1/2}} + \mathcal{O}(e^{-(\kappa_f+c)/h}).$$

Proof. From Proposition 6.2 and Lemma 6.3, we deduce

$$\Big\| d_{f,h}u_1^{(0)} - \sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} b_{k,1}w_k \Big\|_{W^{p,2}(\Omega_+)} \le C_p h^{-N_p} e^{-(\kappa_f+c)/h} \le C'_p e^{-(\kappa_f+c/2)/h}.$$

Since, by the almost orthonormality of the family $(\Pi^{(1)}v_k^{(1)})_{1\le k\le m_1^D(\Omega_+)}$, $\|w_k - \Pi^{(1)}v_k^{(1)}\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$ and $\max\{|b_{k,1}|, m_1^N(\Omega_-)+1 \le k \le m_1^D(\Omega_+)\} = \tilde{\mathcal{O}}(e^{-\kappa_f/h})$ (see Proposition 5.7), Lemma 6.3 also leads to

$$\Big\| d_{f,h}u_1^{(0)} - \sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} b_{k,1}\Pi^{(1)}v_k^{(1)} \Big\|_{W^{p,2}(\Omega_+)} \le C''_p e^{-(\kappa_f+c/2)/h}.$$

By recalling the definition of $v_k^{(1)} = \chi_+\psi_k^{(1)}$, it suffices now to check that $\|v_k^{(1)} - \Pi^{(1)}v_k^{(1)}\|_{W^{p,2}(\Omega_+)}$ is of order $\mathcal{O}(e^{-c'/h})$ for some $c' > 0$. We already know

$$\|v_k^{(1)} - \Pi^{(1)}v_k^{(1)}\|_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$$

from Proposition 4.5.

For the $W^{1,2}(\Omega_+)$ estimates, notice that

$$\|d_{f,h}v_k^{(1)}\|^2_{L^2(\Omega_+)} + \|d^*_{f,h}v_k^{(1)}\|^2_{L^2(\Omega_+)} = \langle v_k^1, \Delta_{f,h}^{D,(1)}(\Omega_+)v_k^{(1)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h})$$

(again from Proposition 4.5), while $\Pi^{(1)}v_k^{(1)} \in F^{(1)} = \mathrm{Ran}\,1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+))$ implies

$$\|d_{f,h}\Pi^{(1)}v_k^{(1)}\|^2_{L^2(\Omega_+)} + \|d^*_{f,h}\Pi^{(1)}v_k^{(1)}\|^2_{L^2(\Omega_+)} = \langle \Pi^{(1)}v_k^1, \Delta_{f,h}^{D,(1)}(\Omega_+)\Pi^{(1)}v_k^{(1)}\rangle_{L^2(\Omega_+)} = \mathcal{O}(e^{-c/h}).$$

We deduce

$$\begin{aligned}
&\|d(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|^2_{L^2(\Omega_+)}+\|d^*(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|^2_{L^2(\Omega_+)}\\
&\quad\le\frac{2}{h^2}\big[\|d_{f,h}(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|^2_{L^2(\Omega_+)}+\|d^*_{f,h}(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|^2_{L^2(\Omega_+)}+2\||\nabla f|(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|^2_{L^2(\Omega_+)}\big]\\
&\quad\le\frac{Ce^{-2c/h}}{h^2}.
\end{aligned}$$

This gives the $W^{1,2}$ estimate $\|v_k^{(1)}-\Pi^{(1)}v_k^{(1)}\|_{W^{1,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})$.

The $W^{p,2}$ estimates ($p\ge 2$) are then obtained by an argument based on the elliptic regularity of the (nonhomogeneous) Dirichlet Hodge Laplacian. On the one hand, $\|\Pi^{(1)}v_k^{(1)}\|_{L^2(\Omega_+)}=1+\mathcal{O}(e^{-c/h})$, $\Pi^{(1)}v_k^{(1)}\in F^{(1)}$ and $\big\|\Delta_{f,h}^{D,(1)}\big|_{F^{(1)}}\big\|=\mathcal{O}(e^{-c/h})$ (see Proposition 3.12) imply that $\|\Delta_{f,h}^{D,(1)}\Pi^{(1)}v_k^{(1)}\|_{L^2(\Omega_+)}=\mathcal{O}(e^{-c/h})$. Lemma 6.3 can then be used to obtain $\|\Delta_{f,h}^{D,(1)}\Pi^{(1)}v_k^{(1)}\|_{W^{p,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})$ for any integer p. Here, $\big\|\Delta_{f,h}^{D,(1)}\big|_{F^{(1)}}\big\|=\sup_{u\in F^{(1)}}(\|\Delta_{f,h}^{D,(1)}u\|_{L^2(\Omega_+)}/\|u\|_{L^2(\Omega_+)})$ is simply the spectral radius of the finite-dimensional operator $\Delta_{f,h}^{D,(1)}:F^{(1)}\to F^{(1)}$. On the other hand, Lemma 6.5 below implies $\|\Delta_{f,h}^{D,(1)}v_k^{(1)}\|_{W^{p,2}(\Omega_+)}=\|\Delta_{f,h}^{D,(1)}(\chi_+\psi_k^{(1)})\|_{W^{p,2}(\Omega_+)}=\mathcal{O}(e^{-c/h})$ for any integer p, using the arguments of the proofs of Proposition 3.5 or 3.9 to get the estimate on the truncated eigenvector from the exponential decay of the eigenvector. Thus, for $p\ge 1$, if $\|(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|_{W^{p,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})$ then the difference $v_k^{(1)}-\Pi^{(1)}v_k^{(1)}$ satisfies

$$\begin{cases}\Delta_H^{(1)}(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})=r_k(h),\\ \mathbf{t}(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})=0,\quad \mathbf{t}d^*(v_k^{(1)}-\Pi_k^{(1)}v_k^{(1)})=\varrho_k(h)\end{cases}$$

with $\|r_k(h)\|_{W^{p,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})$ and $\|\varrho_k(h)\|_{W^{p-1/2,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})$.

This implies $\|(v_k^{(1)}-\Pi^{(1)}v_k^{(1)})\|_{W^{p+2,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})$. A bootstrap argument (induction on p) thus shows that, for any p, $\|v_k^{(1)}-\Pi^{(1)}v_k^{(1)}\|_{W^{p,2}(\Omega_+)}=\tilde{\mathcal{O}}(e^{-c/h})\le\mathcal{O}(e^{-c'/h})$ for any $c'<c$. □

We end this section with an estimate on the exponential decay (in a neighborhood of supp χ_+) of the eigenvectors of $\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-)$ in $\mathcal{C}^\infty$ norm. This is a refinement of Proposition 3.8, which was needed in the previous proof.

Lemma 6.5. *For every $\varepsilon\in(0,1)$, there exists a function $\varphi_\varepsilon\in\mathcal{C}_0^\infty(\Omega_+\setminus\bar{\Omega}_-)$ such that, for all $x\in\bar{\Omega}_+\setminus\Omega_-$,*

$$\begin{gathered}|\nabla\varphi_\varepsilon(x)|\le(1-\varepsilon)|\nabla f(x)|,\\ d(x,\partial\Omega_+\cup\partial\Omega_-)\le\tfrac12\varepsilon\implies\varphi_\varepsilon(x)=0,\\ \varphi_\varepsilon(x)\ge 0\quad\text{and}\quad d_{Ag}(x,\partial\Omega_+\cup\partial\Omega_-)-C\varepsilon\le\varphi_\varepsilon(x),\end{gathered}$$

where $C>0$ is a constant independent of ε. For every $p\in\mathbb{N}$, and once φ_ε is fixed, there exist $C_{\varepsilon,p}>0$ and $N_p>0$ independent of $h\in[0,h_0]$ such that every normalized eigenvector ψ of $\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-)$ associated with an eigenvalue $\lambda\in[0,\nu(h)]$ satisfies

$$\|e^{\varphi_\varepsilon/h}\psi\|_{W^{p,2}(\Omega_+\setminus\bar{\Omega}_-)}\le C_{\varepsilon,p}h^{-N_p}.$$

As explained in the proof, we cannot state this result with φ_ε equal to the Agmon distance to a neighborhood of $\partial\Omega_+$ as in Proposition 3.8 because the Agmon distance is not a sufficiently regular function.

Proof. The function $\varphi_\varepsilon \in \mathcal{C}_0^\infty(\Omega_+ \setminus \overline{\Omega}_-)$ is built as an accurate enough mollified version of $\theta_\varepsilon(x) = (1-2\varepsilon)d_{\mathrm{Ag}}(x, \mathcal{V}_+^\varepsilon \cup \mathcal{V}_-^\varepsilon)$, where

$$\mathcal{V}_\pm^\varepsilon = \{x \in \overline{\Omega}_+ \setminus \Omega_- : d(x, \partial\Omega_\pm) \le \varepsilon\}.$$

Indeed, the function θ_ε is a Lipschitz function such that

$$|\nabla\theta_\varepsilon(x)| \le (1-2\varepsilon)|\nabla f(x)| \quad \text{a.e.},$$
$$d(x, \partial\Omega_+ \cup \partial\Omega_-) \le \varepsilon \implies \theta_\varepsilon(x) = 0,$$
$$d(x, \partial\Omega_+ \cup \partial\Omega_-) - C_1\varepsilon \le \theta_\varepsilon(x) \le d(x, \partial\Omega_+ \cup \partial\Omega_-)$$

hold in $\overline{\Omega}_+ \setminus \Omega_-$, with $C_1 \ge 0$ independent of ε. Since θ_ε fulfills uniform Lipschitz estimates and $|\nabla f(x)| \ge c > 0$ on $\overline{\Omega}_+ \setminus \Omega_-$, all the properties of φ_ε are obtained by considering the convolution of θ_ε with a mollifier with a sufficiently small compact support. We cannot simply take $\varphi_\varepsilon = d_{\mathrm{Ag}}(x, \partial\Omega_+ \cup \partial\Omega_-)$, or even $\varphi_\varepsilon = d_{\mathrm{Ag}}(\,\cdot\,, \mathcal{V}_+^\varepsilon \cup \mathcal{V}_-^\varepsilon)$, because the argument requires us to consider high-order derivatives of φ_ε.

Let ψ be a normalized eigenvector of $\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \overline{\Omega}_-)$ associated with an eigenvalue $\lambda \in [0, \nu(h)]$. We already know from Proposition 3.8 that

$$\|e^{\varphi_\varepsilon/h}\psi\|_{W^{1,2}(\Omega_+\setminus\overline{\Omega}_-)} \le C_\varepsilon h^{-1}. \tag{6-8}$$

The argument to obtain the estimates in $W^{p,2}(\Omega_+ \setminus \overline{\Omega}_-)$-norms is based on a bootstrap argument, using the elliptic regularity of nonhomogeneous Dirichlet boundary problems for the Hodge Laplacian.

Indeed, we have

$$e^{-\varphi_\varepsilon/h}\Delta_{f,h}e^{\varphi_\varepsilon/h} = \Delta_{f,h} - h\mathcal{L}_{\nabla\varphi_\varepsilon} + h\mathcal{L}^*_{\nabla\varphi_\varepsilon} - |\nabla\varphi_\varepsilon|^2,$$

and thus

$$\Delta_{f,h}(e^{\varphi_\varepsilon/h}\psi) = \lambda e^{\varphi_\varepsilon/h}\psi - he^{\varphi_\varepsilon/h}\mathcal{L}_{\nabla\varphi_\varepsilon}\psi + he^{\varphi_\varepsilon/h}\mathcal{L}^*_{\nabla\varphi_\varepsilon}\psi - |\nabla\varphi_\varepsilon|^2 e^{\varphi_\varepsilon/h}\psi.$$

Using the fact that $\Delta_{f,h} = h^2(dd^* + d^*d) + h(\mathcal{L}_{\nabla f} + \mathcal{L}^*_{\nabla f}) + |\nabla f|^2$, we obtain

$$\Delta_H v = h^{-2}(\lambda v - he^{\varphi_\varepsilon/h}\mathcal{L}_{\nabla\varphi_\varepsilon}e^{-\varphi_\varepsilon/h}v + he^{\varphi_\varepsilon/h}\mathcal{L}^*_{\nabla\varphi_\varepsilon}e^{-\varphi_\varepsilon/h}v - |\nabla\varphi_\varepsilon|^2 v - h\mathcal{L}_{\nabla f}v - h\mathcal{L}^*_{\nabla f}v - |\nabla f|^2 v), \tag{6-9}$$

where

$$v = e^{\varphi_\varepsilon/h}\psi.$$

For the boundary conditions, we have, of course,

$$\boldsymbol{t}v = 0, \tag{6-10}$$

and

$$0 = \boldsymbol{t}d^*_{f,h}\psi = e^{\varphi_\varepsilon/h}\boldsymbol{t}d^*_{f,h}\psi = \boldsymbol{t}d^*_{f,h}e^{\varphi_\varepsilon/h}\psi + e^{\varphi_\varepsilon/h}\boldsymbol{t}\boldsymbol{i}_{\nabla\varphi_\varepsilon}\psi.$$

The condition $\varphi_\varepsilon = 0$ in a neighborhood of $\partial\Omega_+ \cup \partial\Omega_-$ implies $\nabla\varphi_\varepsilon = 0$ on $\partial\Omega_+ \cup \partial\Omega_-$, and thus $\boldsymbol{t}\boldsymbol{i}_{\nabla\varphi_\varepsilon}\psi = 0$. Since $d^*_{f,h} = hd^* + \boldsymbol{i}_{\nabla f}$, we thus obtain

$$\boldsymbol{t}d^*v = -\frac{1}{h}\boldsymbol{i}_{\nabla f}v. \tag{6-11}$$

By considering the boundary value problem (6-9)–(6-11) and using the $W^{1,2}(\Omega_+\setminus\bar\Omega_-)$ estimate (6-8), we thus obtain, by the elliptic regularity of the Dirichlet Hodge Laplacian,

$$\|e^{\varphi_\varepsilon/h}\psi\|_{W^{2,2}(\Omega_+\setminus\bar\Omega_-)}\le C_{2,\varepsilon}h^{-3}.$$

This is due to the fact that the right-hand side in (6-9) (resp. (6-11)) is a differential operator of order 1 (resp. 0). The $W^{p,2}(\Omega_+\setminus\bar\Omega_-)$ estimates for $p\ge 3$ are then obtained by induction on p. □

6C. *Change of basis in $F^{(1)}$.* In the previous sections, the estimates (2-12) and (2-14) of the eigenvalue $\lambda_1^{(0)}$ and $d_{f,h}u_1^{(0)}$ in a neighborhood of $\partial\Omega_+$ have been proven with the basis $(\psi_k^{(1)})_{m_1^N(\Omega_-)+1\le k\le m_1^D(\Omega_+)}$ of $\operatorname{Ran}1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-))$. The aim of this section is to show that the estimates (2-12) and (2-14) are valid for any almost orthonormal basis (according to Definition 3.6)

$$(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar\Omega_-)}\quad\text{of}\quad \operatorname{Ran}1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-)).$$

The next proposition thus concludes the proof of Theorem 2.4.

Remark 6.6. We thus prove a slightly more general result than the one stated in Theorem 2.4, since it is only required that $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar\Omega_-)}$ is an *almost* orthonormal basis of $\operatorname{Ran}1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-))$.

Remark 6.7. All the results below extend to complex-valued eigenbases, by simply replacing the real scalar product by the hermitian scalar product.

Proposition 6.8. *Let $\lambda_1^{(0)}$ be the first eigenvalue of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ and $u_1^{(0)}$ the associated $L^2(\Omega_+)$-normalized nonnegative eigenfunction. For any almost orthonormal basis*

$$(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar\Omega_-)}\quad\textit{of}\quad \operatorname{Ran}1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-)),$$

the approximate expressions (2-12) *and* (2-14) *for $\lambda_1^{(0)}$ and $d_{f,h}u_1^{(0)}$ hold true.*

Proof. Let $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar\Omega_-)}$ be an almost orthonormal basis of $\operatorname{Ran}1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-))$. Then there exists a matrix $C(h)=(c_{k,k'})_{1\le k,k'\le m_1^D(\Omega_+\setminus\bar\Omega_-)}$ such that

$$C(h)C(h)^*=\mathrm{Id}_{m_1^D(\Omega_+\setminus\bar\Omega_-)}+\mathcal{O}(e^{-c/h}),\quad C(h)^*C(h)=\mathrm{Id}_{m_1^D(\Omega_+\setminus\bar\Omega_-)}+\mathcal{O}(e^{-c/h}),$$
$$\text{and}\quad \psi^{(1)}_{k+m_1^N(\Omega_-)}=\sum_{k'=1}^{m_1^D(\Omega_+\setminus\bar\Omega_-)}c_{k,k'}u_{k'}^{(1)}\quad\text{for all } k\in\{1,\dots,m_1^D(\Omega_+\setminus\bar\Omega_-)\}.\tag{6-12}$$

Here, $C(h)^*$ denotes the transpose of the matrix $C(h)$.

Let L_1 (resp. L_2) be a continuous linear mapping from $\operatorname{Ran}1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar\Omega_-))$, the finite-dimensional space endowed with the scalar product of $L^2(\Omega_+\setminus\bar\Omega_-)$, to $\mathbb{R}$ (resp. to some vector space E).

Then, using (6-12),

$$\sum_{k=m_1^N(\Omega_-)+1}^{m_1^D(\Omega_+)} L_1(\psi_k^{(1)})L_2(\psi_k^{(1)}) = \sum_{k,k_1,k_2=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} c_{k,k_1}c_{k,k_2}L_1(u_{k_1}^{(1)})L_2(u_{k_2}^{(1)})$$
$$= \sum_{k'=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} L_1(u_{k'}^{(1)})L_2(u_{k'}^{(1)}) + \mathcal{O}(\|L_1\|\|L_2\|e^{-c/h}), \tag{6-13}$$

where $\|L_1\|$ and $\|L_2\|$ denote the operator norms of the linear mappings L_1 and L_2.

The estimate (2-12) is then a consequence of (6-2) and (6-13) with

$$L_1 = L_2 : \operatorname{Ran} 1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-)) \to \mathbb{R}, \quad u \mapsto -\frac{\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n u(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}}$$

with $\|L_1\| = \|L_2\| = \tilde{\mathcal{O}}(e^{-\kappa_f/h})$ due to $\lambda_1^{(0)}(\Omega_+) = \tilde{O}(e^{-2\kappa_f/h})$ (see (6-2)) and the orthonormality of the basis $(\psi^{(1)}_{k+m_1^N(\Omega_-)})_{1\le k\le m_1^D(\Omega_+\setminus\Omega_-)}$. The estimate (2-14) is a consequence of Proposition 6.4 and of (6-13) with L_1 like before and

$$L_2 : \operatorname{Ran} 1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-)) \to \bigwedge^1 W^{p,2}(\mathcal{V}), \quad u \mapsto u\big|_{\mathcal{V}}$$

with $\|L_2\| = \tilde{\mathcal{O}}(1)$ according to Lemma 6.3 applied with $\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-)$ instead of $\Delta_{f,h}^{D,(1)}(\Omega_+)$. □

6D. ***Corollaries.*** The estimate (2-14) contains accurate information about the trace $\partial_n u_1^{(0)}\big|_{\partial\Omega_+}$:

Corollary 6.9. *Let $n : \sigma \mapsto n(\sigma)$ be the outward normal vector field on $\partial\Omega_+$ and let $\partial_n = \boldsymbol{i}_n d$ be the outward normal derivative for functions. For any almost orthonormal basis $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar{\Omega}_-)}$ of* $\operatorname{Ran} 1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+\setminus\bar{\Omega}_-))$, *the normal derivative of the nonnegative and normalized first eigenfunction $u_1^{(0)}$ of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ satisfies*

$$\partial_n u_1^{(0)}(\sigma) \le 0 \quad \textit{for all } \sigma\in\partial\Omega_+$$

and

$$\left\| \partial_n u_1^{(0)}\big|_{\partial\Omega_+} + \sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \frac{\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}}\boldsymbol{i}_n u_k^{(1)} \right\|_{W^{p,2}(\partial\Omega_+)} = \mathcal{O}(e^{-(\kappa_f+c)/h}) \quad \textit{for all } p\in\mathbb{N}$$

for some $c>0$ independent of p.

Proof. The sign condition for $\partial_n u_1^{(0)}(\sigma)$ is a consequence of $u_1^{(0)} \ge 0$ in Ω_+ and $u_1^{(0)}\big|_{\partial\Omega_+} = 0$.

The trace theorem with (2-14) implies

$$d_{f,h}u_1^{(0)}\big|_{\partial\Omega_+} = -h \sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \frac{\int_{\partial\Omega_+} e^{-f(\sigma)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}} u_k^{(1)} + \mathcal{O}(e^{-(\kappa_f+c)/h})$$

in any Sobolev space $W^{p,2}(\partial\Omega_+)$. Recalling

$$d_{f,h}u_1^{(0)} = hdu_1^{(0)} + u_1^{(0)}df \quad \text{and} \quad u_1^{(0)}\big|_{\partial\Omega_+} = 0$$

yields the result. □

Proof of Corollary 2.9. First, note that the equality

$$\partial_n[e^{-f_{1,2}/h}u_1^{(0)}(f_{1,2})]\big|_{\partial\Omega_+} = e^{-f_{1,2}/h}[\partial_n u_1^{(0)}(f_{1,2})]\big|_{\partial\Omega_+}$$

is simply due to the Dirichlet boundary condition $u_1^{(0)}\big|_{\partial\Omega_+} = 0$. The identity (2-15) is then a direct consequence of (2-12), since the same basis $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\overline{\Omega}_-)}$ can be picked for f_1 and f_2 because these two functions coincide on $\Omega_+\setminus\overline{\Omega}_-$.

Second, for (2-16), it is more convenient to write (2-14) with f_j, $j=1,2$, in the form

$$\begin{aligned}\left(\int_{\Omega_+} e^{-2f_j(x)/h}\,dx\right)^{\frac12} & d_{f_j,h}u_1^{(0)}(f_j) \\ &= -h\sum_{k=1}^{m_1^D(\Omega_+\setminus\overline{\Omega}_-)}\left(\int_{\partial\Omega_+} e^{-f_j(\sigma)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma\right)u_k^{(1)} + \mathcal{O}(e^{-(\min_{\partial\Omega_+} f_j+c)/h}),\end{aligned}$$

the estimate being true in any Sobolev space $\bigwedge^1 W^{p,2}(\mathcal{V})$. Using the fact that $f_1\equiv f_2\equiv f$ in $\Omega_+\setminus\overline{\Omega}_-$, taking the trace along $\partial\Omega_+$ and multiplying by $e^{-(f-\min_{\partial\Omega_+} f)/h}$, which is less than 1 on $\partial\Omega_+$, and then by $e^{(\min_{\partial\Omega_+} f)/h}$, lead to

$$\begin{aligned}&\left(\int_{\Omega_+} e^{-2f_j(x)/h}\,dx\right)^{\frac12} e^{-(f-2\min_{\partial\Omega_+} f)/h}\partial_n u_1^{(0)}(f_j)\big|_{\partial\Omega_+} \\ &\quad= -\sum_{k=1}^{m_1^D(\Omega_+\setminus\overline{\Omega}_-)}\left(\int_{\partial\Omega_+} e^{-(f(\sigma)-\min_{\partial\Omega_+} f)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma\right)e^{-(f-\min_{\partial\Omega_+} f)/h}\boldsymbol{i}_n u_k^{(1)} + \mathcal{O}(e^{-c/h}),\end{aligned}$$

the estimate being true in $L^1(\partial\Omega_+)$. The left-hand side is negative and its L^1-norm is thus given by the absolute value of its integral. Let us estimate this norm, using Lemma 4.3 and Lemma 5.9: for any positive ε,

$$\begin{aligned}-\left(\int_{\Omega_+} e^{-2f_j(x)/h}\,dx\right)^{\frac12}&\int_{\partial\Omega_+} e^{-(f-2\min_{\partial\Omega_+} f)/h}\partial_n u_1^{(0)}(f_j)(\sigma)\,d\sigma \\ &= \sum_{k=1}^{m_1^D(\Omega_+\setminus\overline{\Omega}_-)}\left(\int_{\partial\Omega_+} e^{-(f(\sigma)-\min_{\partial\Omega_+} f)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma\right)^2 + \mathcal{O}(e^{-c/h}) \\ &= e^{(2\min_{\partial\Omega_+} f)/h}\lambda_1^{(0)}(f_1)h^{-2}\int_{\Omega_+} e^{-2f_1(x)/h}\,dx + \mathcal{O}(e^{-c/h}) \\ &\ge C_\varepsilon e^{(2\min_{\partial\Omega_+} f)/h}e^{-2(\kappa_f+\varepsilon)/h}h^{-2}\frac{1}{C_{f_1}}h^{d/2}e^{-(2\min_{\Omega_+} f_1)/h} + \mathcal{O}(e^{-c/h}) \\ &= C_\varepsilon e^{-2\varepsilon/h}\frac{h^{-2+d/2}}{C_{f_1}} + \mathcal{O}(e^{-c/h}) \ge Ce^{-c/(2h)}.\end{aligned}$$

Thus,

$$-\frac{e^{-f_j/h}\partial_n u_1^{(0)}(f_j)\big|_{\partial\Omega_+}}{\big\|e^{-f_j/h}\partial_n u_1^{(0)}(f_j)\big|_{\partial\Omega_+}\big\|_{L^1(\partial\Omega_+)}} = \frac{\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\big(\int_{\partial\Omega_+} e^{-(f(\sigma)-\min_{\partial\Omega_+}f)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma\big)e^{-(f-\min_{\partial\Omega_+}f)/h}\boldsymbol{i}_n u_k^{(1)}}{\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\big(\int_{\partial\Omega_+} e^{-(f(\sigma)-\min_{\partial\Omega_+}f)/h}\boldsymbol{i}_n u_k^{(1)}(\sigma)\,d\sigma\big)^2} + \mathcal{O}(e^{-c/(2h)}).$$

This concludes the proof, since the right-hand side does not depend on f_j. □

7. About Hypotheses 3 and 4

We have chosen to set the Hypotheses 3 and 4 in terms of some spectral properties of the Witten Laplacians $\Delta^N_{f,h}(\Omega_-)$ and $\Delta^D_{f,h}(\Omega_-\setminus\bar{\Omega}_-)$ in order to be general enough and to cover possible further advances about the low spectrum of Witten Laplacians. These hypotheses can actually be translated into very explicit and simple geometric conditions on the function f when f is a Morse function such that $f\big|_{\partial\Omega_+}$ is a Morse function. We recall that a Morse function is a $\mathcal{C}^\infty$ function whose critical points are all nondegenerate. Section 7A is devoted to a verification of Hypotheses 3 and 4 when f and $f\big|_{\partial\Omega_+}$ are Morse functions, using the results of [Helffer and Nier 2006; Le Peutrec 2010b]. Theorem 2.10 is then obtained as a consequence of the accurate results under the Morse conditions and the estimates stated in Corollary 2.9.

Finally, Section 7B is devoted to a discussion about potentials that are not Morse functions. In particular, examples of functions f which are not Morse functions and for which Hypotheses 3 and 4 hold are presented.

7A. *The case of a Morse function f.*

7A1. *Verifying Hypotheses 3 and 4.* Let us first specify the assumptions which allow us to use the results of [Helffer and Nier 2006; Le Peutrec 2010b], in addition to Hypotheses 1 and 2, which were already explicitly formulated in terms of the function f:

Hypothesis 5. *The functions f and $f\big|_{\partial\Omega_+}$ are Morse functions.*

Hypothesis 6. *The critical values of f are all distinct and the differences $f(U^{(1)})-f(U^{(0)})$, where $U^{(0)}$ ranges over the local minima of f and $U^{(1)}$ ranges over the critical points of f with index* 1, *are all distinct.*

Although $f\big|_{\partial\Omega_-}$ is not assumed to be a Morse function (see the discussion below), Hypotheses 1, 5 and 6 ensure that the results of [Helffer and Nier 2006; Le Peutrec 2010b] on small eigenvalues of $\Delta^D_{f,h}(\Omega_+)$, $\Delta^N_{f,h}(\Omega_-)$ and $\Delta^D_{f,h}(\Omega_+\setminus\bar{\Omega}_-)$ apply. Following [Le Peutrec 2010b], Hypothesis 6 is useful to get accurate scaling rates for the small eigenvalues of $\Delta^{N,(0)}_{f,h}(\Omega_-)$. In particular, the information on the size of the second eigenvalue $\mu_2^{(0)}(\Omega_-)>\mu_1^{(0)}(\Omega_-)=0$ of $\Delta^{N,(0)}_{f,h}(\Omega_-)$ is important to prove (2-8) in Hypothesis 3. Hypothesis 6 also implies that f has a unique global minimum. Hypothesis 6 could certainly be relaxed.

Let us recall the general results of [Helffer and Nier 2006; Le Peutrec 2010b] on the number and the scaling of small eigenvalues for boundary Witten Laplacians in a regular domain Ω (see also [Chang and Liu 1995; Laudenbach 2011] for related results). The potential f is assumed to be a Morse function f on Ω such that $|\nabla f| \neq 0$ on $\partial\Omega$ and $f\big|_{\partial\Omega}$ is also a Morse function. The notion of critical points with index p for f has to be extended as follows, in order to take into account points on the boundary $\partial\Omega$.

• *In the interior* Ω*:* A generalized critical point with index p is, as usual, a critical point at which the Hessian of f has p negative eigenvalues. It is a local minimum for $p = 0$, a saddle point for $p = 1$ and a local maximum for $p = \dim M = d$.

• *Along the boundary* $\partial\Omega$ *in the Dirichlet case:* A generalized critical point with index $p \geq 1$ is a critical point σ of $f\big|_{\partial\Omega}$ with index $p-1$ such that the outward normal derivative is positive ($\partial_n f(\sigma) > 0$). Therefore, along the boundary, there is no generalized critical point with index 0, and critical points with index 1 coincide with the local minima σ of $f\big|_{\partial\Omega}$ such that $\partial_n f(\sigma) > 0$. Intuitively, this definition can be understood by interpreting the homogeneous Dirichlet boundary conditions as an extension of the potential by $-\infty$ outside Ω.

• *Along the boundary* $\partial\Omega$ *in the Neumann case:* A generalized critical point with index p is a critical point σ of $f\big|_{\partial\Omega}$ with index p such that the outward normal derivative is negative ($\partial_n f(\sigma) < 0$). Therefore, along the boundary, a generalized critical point with index 0 is a local minimum of $f\big|_{\overline{\Omega}}$ and a critical point with index 1 is a saddle point σ of $f\big|_{\partial\Omega}$ such that $\partial_n f(\sigma) < 0$. Intuitively, this definition can be understood by interpreting the homogeneous Neumann boundary conditions as an extension of the potential by $+\infty$ outside Ω.

The number of generalized critical points in Ω with index p is denoted by $\tilde{m}_p^D(\Omega)$ or $\tilde{m}_p^N(\Omega)$, depending on whether the boundary Witten Laplacian on Ω with Dirichlet or Neumann boundary conditions is considered.

One result of [Helffer and Nier 2006; Le Peutrec 2010b] says that, for $\nu(h) = h^{6/5}$, one has, for the Dirichlet Witten Laplacian,

$$\#\big[\sigma(\Delta_{f,h}^{D,(p)}(\Omega)) \cap [0, \nu(h)]\big] = \tilde{m}_p^D(\Omega), \quad \sigma(\Delta_{f,h}^{D,(p)}(\Omega)) \cap [0, \nu(h)] \subset [0, e^{-c_0/h}],$$

and, for the Neumann boundary Witten Laplacian,

$$\#\big[\sigma(\Delta_{f,h}^{N,(p)}(\Omega)) \cap [0, \nu(h)]\big] = \tilde{m}_p^N(\Omega), \quad \sigma(\Delta_{f,h}^{N,(p)}(\Omega)) \cap [0, \nu(h)] \subset [0, e^{-c_0/h}]$$

for some positive constant c_0. These results rely, like in [Cycon et al. 1987] for the boundaryless case, on the introduction of an h-dependent partition of unity and a rough analysis of boundary local models.

Let us now apply these general results in our context. Under Hypotheses 1 and 5, we have:

- $\tilde{m}_p^N(\Omega_-)$ is the number of critical points with index p in the interior of Ω_-.
- $\tilde{m}_p^D(\Omega_+ \setminus \overline{\Omega}_-)$ is the number of critical points σ with index $p-1$ of $f\big|_{\partial\Omega_+}$ such that $\partial_n f(\sigma) > 0$. In particular, $\tilde{m}_0^D(\Omega_+ \setminus \overline{\Omega}_-) = 0$, and $\tilde{m}_1^D(\Omega_+ \setminus \overline{\Omega}_-)$ is the number of local minima of $f\big|_{\partial\Omega_+}$ with positive normal derivatives.

- $\tilde{m}_p^D(\Omega_+)$ is the number of critical points with index p in the interior of Ω_- plus the number of critical points σ of $f\big|_{\partial\Omega_+}$ with index $p-1$ such that $\partial_n f(\sigma) > 0$. For $p=0$, $\tilde{m}_0^D(\Omega_+)$ equals $\tilde{m}_0^N(\Omega_-)$ while $\tilde{m}_1^D(\Omega_+)$ is $m_1^N(\Omega_-)$ augmented by the number of local minima of $f\big|_{\partial\Omega_+}$ with positive normal derivatives.

As already mentioned above, we can use the results of [Helffer and Nier 2006; Le Peutrec 2010b] without assuming that $f\big|_{\partial\Omega_-}$ is a Morse function. The reason is that $\partial_n f > 0$ on $\partial\Omega_-$ and, thus, there is no generalized critical point on $\partial\Omega_-$ associated with $\Delta_{f,h}^{N,(p)}(\Omega_-)$ and $\Delta_{f,h}^{D,(p)}(\Omega_+ \setminus \overline{\Omega}_-)$.

In summary, using these results, conditions (2-6), (2-7), (2-9) and (2-10) are fulfilled with $\nu(h) = h^{6/5}$, some $c_0 > 0$ and $m_p^{N,D}(\Omega) = \tilde{m}_p^{N,D}(\Omega)$, $p \in \{0, 1\}$ and $\Omega = \Omega_-$ or $\Omega = \Omega_+ \setminus \overline{\Omega}_-$. Hence, all the conditions of Hypotheses 3 and 4 are satisfied except (2-8). Note in particular that two of the results in Theorem 2.4,

$$m_0^D(\Omega_+) = m_0^N(\Omega_-) \quad \text{and} \quad m_1^D(\Omega_+) = m_1^N(\Omega_-) + m_1^D(\Omega_+ \setminus \overline{\Omega}_-),$$

are consistent with the relations on the numbers of generalized critical points:

$$\tilde{m}_0^D(\Omega_+) = \tilde{m}_0^N(\Omega_-) \quad \text{and} \quad \tilde{m}_1^D(\Omega_+) = \tilde{m}_1^N(\Omega_-) + \tilde{m}_1^D(\Omega_+ \setminus \overline{\Omega}_-).$$

As explained in the proof below, Hypothesis 6 is particularly useful to verify condition (2-8) in Hypothesis 3.

The following proposition thus yields a simple set of assumptions on f such that Theorem 2.4 holds:

Proposition 7.1. *Assume Hypotheses 1, 5 and 6 and let $\mathcal{U}^{(0)}$ (resp. $\mathcal{U}^{(1)}$) denote the set of critical points with index* 0 (*resp.* 1) *of $f\big|_{\Omega_-}$. Let us consider the Agmon distance d_{Ag} introduced in Lemma 3.2. Then the inequality*

$$d_{Ag}(\partial\Omega_-, \mathcal{U}^{(0)}) > \max_{U^{(1)} \in \mathcal{U}^{(1)},\, U^{(0)} \in \mathcal{U}^{(0)}} f(U^{(1)}) - f(U^{(0)}) \tag{7-1}$$

implies (2-8). *As a consequence, the inequality* (7-1) *together with Hypotheses 1, 2, 5 and 6 are sufficient conditions for the results of Theorem 2.4 and its corollaries to hold.*

Figures 2 and 3 give examples of functions f for which the inequality (7-1) together with Hypotheses 1, 2, 5 and 6 are fulfilled. Figure 4 is an example of a function f which satisfies Hypotheses 1, 2, 5 and 6, but not the inequality (7-1).

Remark 7.2. Since $d_{\mathrm{Ag}}(x, y) \geq |f(x) - f(y)|$ (see (3-1)), the condition (1-9) given in the introduction is a sufficient condition for (7-1). Condition (1-9) also implies Hypothesis 2. Thus, a set of sufficient conditions for Theorem 2.4 to hold is Hypotheses 1, 5 and 6 together with (1-9). This is indeed the simple setting presented in the introduction (see the four assumptions stated in Section 1B).

Remark 7.3. It may happen that $\mathcal{U}^{(1)} = \varnothing$. In this case, the inequality (7-1) is automatically satisfied, and there are no exponentially small nonzero eigenvalue for $\Delta_{f,h}^{N,(0)}(\Omega_-)$. Consistently, (2-8) is a void condition in this case.

Proof of Proposition 7.1. By the previous discussion, it only remains to prove that Hypotheses 1, 5 and 6 together with (7-1) imply (2-8) for the proposition to hold. According to [Le Peutrec 2010b], the smallest

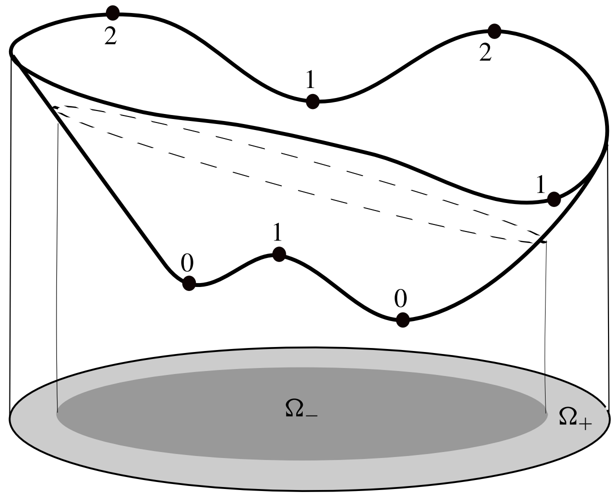

Figure 2. A two-dimensional example where the inequality (7-1) together with Hypotheses 1, 2, 5 and 6 are fulfilled. The generalized critical points are labeled by their indices.

nonzero eigenvalue of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ (namely $\mu_2^{(0)}(\Omega_-)$), satisfies, under Hypotheses 5 and 6, the inequality

$$\lim_{h\to 0} h\log(\mu_2^{(0)}(\Omega_-)) = -2(f(U_{j_1}^{(1)}) - f(U_{j_0}^{(0)})) \geq -2 \max_{U^{(1)}\in\mathcal{U}^{(1)},\, U^{(0)}\in\mathcal{U}^{(0)}} f(U^{(1)}) - f(U^{(0)}),$$

where $U_{j_0}^{(0)}$ and $U_{j_1}^{(1)}$ are two critical points of index 0 and 1, respectively.

Let us now consider the exponential decay near $\partial\Omega_-$ of an eigenfunction of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ associated with a nonzero, exponentially small eigenvalue. A stronger version of Proposition 3.4 can be given because under Hypotheses 1, 5 and 6 the critical points of $f\big|_{\Omega_-}$ which are not local minima are not associated with small eigenvalues of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ (they are so-called *nonresonant* wells; see [Helffer and Sjöstrand 1985a]). Indeed, when U is a critical point of $f\big|_{\Omega_-}$ with $U \notin \mathcal{U}^{(0)}$, the local model of $\Delta_{f,h}^{D,(0)}(B(U,r))$ has his spectrum included in $[h/C(U,r), +\infty)$ for $r>0$ small enough (see, for example, [Cycon et al. 1987]). Then, Corollary 2.2.7 of [Helffer and Sjöstrand 1985a] implies that any normalized eigenfunction $\psi(h)$ of $\Delta_{f,h}^{N,(0)}(\Omega_-)$ associated with an eigenvalue $\mu(h)\in[0, e^{-c_0/h}]$ satisfies

$$\forall\varepsilon>0\;\; \exists C_\varepsilon>0\;\; \forall x\in\Omega_- \quad |\psi_h(x)| \leq C_\varepsilon(e^{-(d_{\mathrm{Ag}}(x,\mathcal{U}_0)+\varepsilon)/h})$$

(compare with the result of Proposition 3.4). Hence, condition (7-1) implies that, in a small neighborhood $\mathcal{V}_-$ of $\partial\Omega_-$, the eigenfunction $\psi(h)$ is estimated by

$$\|\psi(h)\|_{L^2(\mathcal{V}_-)} = \tilde{\mathcal{O}}(e^{-d_{\mathrm{Ag}}(\mathcal{V}_-,\mathcal{U}^{(0)})/h}) \leq C\exp\biggl(-\frac{\max_{U^{(1)}\in\mathcal{U}^{(1)},\, U^{(0)}\in\mathcal{U}^{(0)}} f(U^{(1)}) - f(U^{(0)}) + c}{h}\biggr)$$
$$\leq \tilde{\mathcal{O}}\Bigl(\sqrt{\mu_2^{(0)}(\Omega_-)}\Bigr) \leq \tilde{\mathcal{O}}(\sqrt{\mu(h)})$$

provided that $\mu(h)\neq\mu_1^{(0)}(\Omega_-)=0$. This is exactly (2-8). □

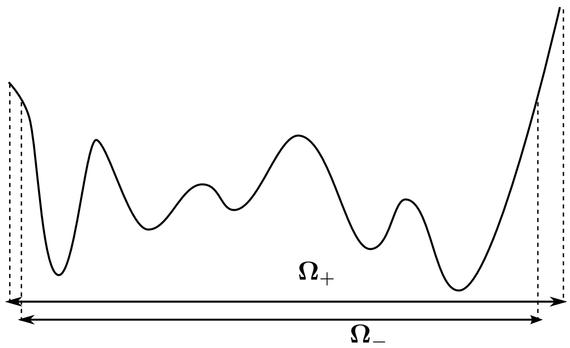

Figure 3. A one-dimensional example where the inequality (7-1) together with Hypotheses 1, 2, 5 and 6 are fulfilled.

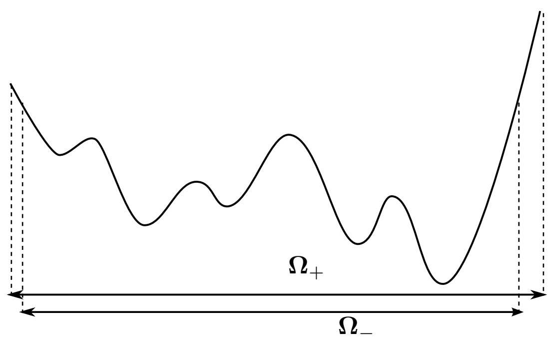

Figure 4. A one-dimensional example where Hypotheses 1, 2, 5 and 6 are fulfilled, but the inequality (7-1) is not satisfied. The condition (7-1) would be fulfilled with a lower local minimum on the left-hand side, for example (see Figure 3).

Remark 7.4 (assumptions in terms of Ω_+ only). Let us assume that Hypotheses 2, 5 and 6 hold. Then, it is easy to check that, if

$$\partial_n f\big|_{\partial\Omega_+} > 0 \tag{7-2}$$

and

$$d_{\mathrm{Ag}}(\partial\Omega_+, \mathcal{U}^{(0)}) > \max_{U^{(1)}\in\mathcal{U}^{(1)},\, U^{(0)}\in\mathcal{U}^{(0)}} f(U^{(1)}) - f(U^{(0)}), \tag{7-3}$$

then there exists a regular open domain Ω_- such that $\overline{\Omega}_- \subset \Omega_+$ and Hypothesis 1 and condition (7-1) hold. Indeed, conditions (7-2) and (7-3) are open and allow small deformation from Ω_+ to some subset Ω_-. Note that condition (7-2) implies that this small deformation can be chosen so that all the critical points of f are indeed in Ω_-; this is exactly Hypothesis 1. As a consequence, under Hypotheses 2, 5 and 6 and assumptions (7-2) and (7-3), the results of Theorem 2.4 hold for a well-chosen domain Ω_- such that $\overline{\Omega}_- \subset \Omega_+$.

In addition, following Remark 7.2 above, it is easy to check that the inequality

$$\min_{\partial\Omega_+} f - \text{cvmax} > \text{cvmax} - \min_{\Omega_+} f \tag{7-4}$$

is a sufficient condition for (7-3). It also implies Hypothesis 2. Thus, under Hypotheses 5 and 6 and the two assumptions (7-2) and (7-4), the results of Theorem 2.4 hold for a well-chosen domain Ω_- such that $\overline{\Omega}_- \subset \Omega_+$.

7A2. *Proof of Theorem 2.10.* In this section, more explicit formulas for $\lambda_1^{(0)}(\Omega_+)$ and $\partial_n(e^{-f/h}u_1^{(0)})$ are given under the Morse assumption on f and $f\big|_{\partial\Omega_+}$. We shall prove:

Proposition 7.5. *Assume Hypotheses 1, 2, 5, 6, the condition* (7-1) *and, moreover,*

$$\partial_n f > 0 \quad \text{on } \partial\Omega_+ . \tag{7-5}$$

Then the first eigenvalue $\lambda_1^{(0)}(\Omega_+)$ of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ satisfies

$$\begin{aligned}\lambda_1^{(0)}(\Omega_+) &= \sum_{k=1}^{m_1^D(\Omega_+\setminus\overline{\Omega}_-)} \left(\frac{h\det(\operatorname{Hess} f)(U_0)}{\pi\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})}\right)^{\frac12} 2\partial_n f(U_k^{(1)})e^{-2(f(U_k^{(1)})-f(U_0))/h}(1+\mathcal{O}(h)) &(7\text{-}6)\\ &= \frac{\int_{\partial\Omega_+} 2\partial_n f(\sigma)e^{-2f(\sigma)/h}\,d\sigma}{\int_{\Omega_+} e^{-2f(x)/h}\,dx}(1+\mathcal{O}(h)), &(7\text{-}7)\end{aligned}$$

where U_0 is the (*unique*) *global minimum of f in Ω_+ and the $U_k^{(1)}$ are the local minima of $f\big|_{\partial\Omega_+}$. Moreover, the normalized nonnegative eigenfunction $u_1^{(0)}$ of $\Delta_{f,h}^{D,(0)}(\Omega_+)$ associated with $\lambda_1^{(0)}(\Omega_+)$ satisfies*

$$-\frac{\partial_n[e^{-f/h}u_1^{(0)}]\big|_{\partial\Omega_+}}{\big\|\partial_n[e^{-f/h}u_1^{(0)}]\big\|_{L^1(\partial\Omega_+)}} = \frac{(2\partial_n f)e^{-2f/h}\big|_{\partial\Omega_+}}{\|(2\partial_n f)e^{-2f/h}\|_{L^1(\partial\Omega_+)}} + \mathcal{O}(h) \quad \text{in } L^1(\partial\Omega_+). \tag{7-8}$$

Remark 7.6. The hypothesis $\partial_n f > 0$ on $\partial\Omega_+$ ensures that the set of all the local minima $U_k^{(1)}$ of $f\big|_{\partial\Omega_+}$ coincides with the set of generalized critical points with index 1 for $\Delta_{f,h}^D(\Omega_+\setminus\overline{\Omega}_-)$. The results of Proposition 7.5 also hold under the more general assumption that $\partial_n f(\sigma) > 0$ when $\sigma\in\partial\Omega_+$ is such that $f(\sigma) \le \min_{\partial\Omega_+} f + \varepsilon_0$ for some $\varepsilon_0 > 0$, by adapting the arguments below.

Remark 7.7. It is possible to write explicitly a first-order approximation for the probability density $-\partial_n(e^{-f/h}u_1^{(0)})\big|_{\partial\Omega_+}/\|\partial_n(e^{-f/h}u_1^{(0)})\|_{L^1(\partial\Omega_+)}$, in the spirit of the approximation (7-6) for $\lambda_1^{(0)}(\Omega_+)$. This approximation uses second-order Taylor expansions of f around the local minima $U_k^{(1)}$; see (7-20) below. More precisely, this approximation becomes

$$-\frac{\partial_n[e^{-f/h}u_1^{(0)}]\big|_{\partial\Omega_+}}{\|\partial_n[e^{-f/h}u_1^{(0)}]\|_{L^1(\partial\Omega_+)}} = \frac{\sum_{k=1}^{m_1^D(\Omega_+\setminus\overline{\Omega}_-)} t_k(h)G_k(h)}{\sum_{k=1}^{m_1^D(\Omega_+\setminus\overline{\Omega}_-)} t_k(h)} + \mathcal{O}(h), \tag{7-9}$$

where the $G_k(h)$ are Gaussian densities centered at the $U_k^{(1)}$ and the weights $t_k(h)$ are such that $\lim_{h\to 0} h\log t_k(h) = -f(U_k^{(1)})$. When $f\big|_{\partial\Omega_+}$ has a unique global minimum, the sums in (7-6) and (7-9) reduce to a single term.

Remark 7.8. As explained in Remark 7.4 above, it is again possible to write a set of assumptions in terms of Ω_+ only. In particular, the results of Proposition 7.5 hold under Hypotheses 2, 5 and 6 and assumptions (7-2) and (7-3).

Remark 7.9. It is possible to extend our analysis to the case of an h-dependent function $f = f_h$ such that our assumptions are verified with uniform constants. For example, the results hold if the values $f(U_k^{(1)})$ of f at the local minima $U_k^{(1)}$ are moved in an $\mathcal{O}(h)$ range without changing $f - f(U_k^{(1)})$ locally. This would change the coefficients $t_k(h)$ in (7-9) accordingly by $\mathcal{O}(1)$ factors.

Most of our effort will be devoted to the proof of Proposition 7.5. Let us first conclude the proof of Theorem 2.10 using the result of Proposition 7.5.

Proof of Theorem 2.10. Let f be a function such that Hypotheses 1, 2, 3 and 4 are satisfied. Let us assume moreover that $f\big|_{\partial\Omega_+}$ is a Morse function and $\partial_n f > 0$ on $\partial\Omega_+$. It is possible to build a $\mathcal{C}^\infty$ function $\tilde{f}$ such that $\tilde{f} = f$ on $\Omega_+ \setminus \Omega_-$ and Hypotheses 1, 2, 5 and 6 and condition (7-1) are satisfied by $\tilde{f}$. This relies in particular on the fact that Morse functions are dense in $\mathcal{C}^\infty$ functions. The condition (7-1) may require us to slightly change the local minimal values of the Morse function $\tilde{f}$.

The function $\tilde{f}$ now fulfills all the requirements of Proposition 7.5 and thus, with obvious notation,

$$\tilde{\lambda}_1^{(0)}(\Omega_+) = \frac{\int_{\partial\Omega_+} 2\partial_n f(\sigma) e^{-2f(\sigma)/h}\, d\sigma}{\int_{\Omega_+} e^{-2\tilde{f}(x)/h}\, dx}(1 + \mathcal{O}(h))$$

and

$$-\frac{\partial_n[e^{-\tilde{f}/h}\tilde{u}_1^{(0)}]\big|_{\partial\Omega_+}}{\|\partial_n[e^{-\tilde{f}/h}\tilde{u}_1^{(0)}]\|_{L^1(\partial\Omega_+)}} = \frac{(2\partial_n f)e^{-2f/h}\big|_{\partial\Omega_+}}{\|(2\partial_n f)e^{-2f/h}\|_{L^1(\partial\Omega_+)}} + \mathcal{O}(h) \quad \text{in } L^1(\partial\Omega_+).$$

Here, we have used the fact that $\tilde{f} = f$ on $\Omega_+ \setminus \Omega_-$. Notice that the function $\tilde{f}$ satisfies Hypotheses 1, 2, 3 and 4 by the results of the previous section. We thus conclude the proof by referring to Corollary 2.9. □

The proof of Proposition 7.5 is done in two steps: We first apply Theorem 2.4 using a very specific basis of Ran $1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \overline{\Omega}_-))$ to get estimates of

$$\lambda_1^{(0)}(\Omega_+) \quad \text{and} \quad -\frac{\partial_n(e^{-f/h}u_1^{(0)})\big|_{\partial\Omega_+}}{\|\partial_n(e^{-f/h}u_1^{(0)})\|_{L^1(\partial\Omega_+)}}$$

in terms of second-order Taylor expansions of f around the local minima $U_k^{(1)}$ (see (7-6) and (7-20)). We then show that these expansions coincide with (7-7) and (7-8).

Before this, we explain how to build the almost orthonormal basis of Ran $1_{[0,\nu(h)]}(\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \overline{\Omega}_-))$ that is needed to prove our results. This construction relies heavily on the Morse assumption on f and $f\big|_{\partial\Omega_+}$ (see Hypothesis 5). We need the results of [Helffer and Nier 2006, Chapter 4] on approximate formulas for a basis of the eigenspace of $\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \overline{\Omega}_-)$ associated with $\mathcal{O}(e^{-c_0/h})$ eigenvalues (see also [Le Peutrec 2010a] for a more general analysis). In what follows, it is assumed that Hypotheses 1, 5 and 6 and condition (7-5) hold. The one-forms of that basis are constructed via a WKB expansion around each local minimum $U_k^{(1)}$ of $f\big|_{\partial\Omega_+}$ $(1 \le k \le m_1^D(\Omega_+ \setminus \overline{\Omega}_-))$. In a neighborhood $\mathcal{V}_k$ of $U_k^{(1)}$, consider the function φ_k defined in a neighborhood of $U_k^{(1)}$ as follows: We assume that all the $\mathcal{V}_k$ are disjoint subsets of $\overline{\Omega}_+ \setminus \overline{\Omega}_-$. The function φ_k satisfies the eikonal equation

$$|\nabla\varphi_k|^2 = |\nabla f|^2, \quad \varphi_k\big|_{\partial\Omega_+} = (f - f(U_k^{(1)}))\big|_{\partial\Omega_+}, \quad \partial_n\varphi_k\big|_{\partial\Omega_+} = -\partial_n f\big|_{\partial\Omega_+}.$$

In the neighborhood $\mathcal{V}_k$, one can build coordinates $(x', x^d) = (x^1, \dots, x^{d-1}, x^d)$ such that:

- The open set Ω_+ looks like a half-space:

$$\Omega_+ \cap \mathcal{V}_k = \{(x', x^d) : |x'| \le r,\ x^d < 0\},$$
$$\partial\Omega_+ \cap \mathcal{V}_k = \{(x', x^d) : |x'| \le r,\ x^d = 0\}.$$

- The metric has the form $g_{d,d}(x)(dx^d)^2 + \sum_{i,j=1}^{d-1} g_{i,j}(x)\,dx^i\,dx^j$ with $g_{i,j}(0) = \delta_{i,j}$ (notice that a different normalization of $g_{d,d}(0)$ was used in [Helffer and Nier 2006]).
- The coordinates (x', x^d) are Morse coordinates both for f and φ_k:

$$f(x) - f(U_k^{(1)}) = \partial_n f(U_k^{(1)})x^d + \frac{1}{2}\sum_{j=1}^{d-1}\lambda_j(x^j)^2, \quad \varphi_k(x) = -\partial_n f(U_k^{(1)})x^d + \frac{1}{2}\sum_{j=1}^{d-1}\lambda_j(x^j)^2, \quad (7\text{-}10)$$

  where the λ_j are the eigenvalues of $\operatorname{Hess}\big(f\big|_{\partial\Omega_+}\big)(U_k^{(1)})$.

In [Helffer and Nier 2006] a local self-adjoint realization of $\Delta_{f,h}^{(1)}$ around $U_k^{(1)}$ is introduced with the same boundary conditions along $\partial\Omega_+$ as for $\Delta_{f,h}^{D,(1)}(\Omega_+)$, with a unique exponentially small eigenvalue $\zeta_k(h) = \mathcal{O}(e^{-c_k/h})$. A corresponding approximate eigenvector is given by the WKB expansion (in the limit of small h)

$$z_k^{\mathrm{wkb},(1)}(x,h) = a_k(x,h)e^{-\varphi_k(x)/h}, \quad \text{where } a_k(x,h) \sim a_{k,0}(x)\,dx^d + \sum_{\ell=1}^{\infty} b_{k,\ell}h^\ell \qquad (7\text{-}11)$$

with $b_{k,\ell} = \sum_{j=1}^d a_{k,\ell,j}(x)\,dx^j$ and $a_{k,0}(0) = 1$. The symbol $\sim$ stands for the equality of asymptotic expansions. Let $z_k^{(1)}$ be the eigenvector of the self-adjoint realization of $\Delta_{f,h}^{(1)}$ around $U_k^{(1)}$ introduced above, associated with $\zeta_k(h)$ and normalized by $\boldsymbol{i}_{\partial_{x^d}} z_k^{(1)}(0) = \boldsymbol{i}_{\partial_{x^d}} z_k^{\mathrm{wkb},(1)}(0)$. It is shown in [Helffer and Nier 2006, Proposition 4.3.2(b,d)] that the estimates

$$\forall\alpha \in \mathbb{N}^d\ \ \exists C_\alpha > 0\ \ \exists N_\alpha \in \mathbb{N} \quad |\partial_x^\alpha z_k^{(1)}(x)| \le C_\alpha h^{-N_\alpha} e^{-\varphi_k(x)/h}, \qquad (7\text{-}12)$$

$$\forall N \in \mathbb{N}\ \ \forall\alpha \in \mathbb{N}^d\ \ \exists C_{\alpha,N} > 0 \quad |\partial_x^\alpha(z_k^{\mathrm{wkb},(1)} - z_k^{(1)})(x)| \le C_{N,\alpha} h^N e^{-\varphi_k(x)/h} \qquad (7\text{-}13)$$

hold for all x in a neighborhood $\mathcal{V}_k' \subset \mathcal{V}_k$ of $U_k^{(1)}$. Notice that the one-forms $z_k^{\mathrm{wkb},(1)}$ and $z_k^{(1)}$ are real-valued. By taking a cut-off function $\chi_k \in \mathcal{C}_0^\infty(\mathcal{V}_k')$ with $\chi_k \equiv 1$ in a neighborhood of $U_k^{(1)}$, a normalized quasimode for $\Delta_{f,h}^{D,(1)}(\Omega_+ \setminus \Omega_-)$ is given by

$$w_k^{(1)} = \frac{\chi_k z_k^{(1)}}{\|\chi_k z_k^{(1)}\|_{L^2(\mathcal{V}_k')}}.$$

The set of functions $(w_k^{(1)})_{k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}}$ is orthonormal, owing to the disjoint supports of the functions $(\chi_k)_{k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}}$. According to [Helffer and Nier 2006, Proposition 6.6], those quasimodes belong to the form domain of $\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-)$, and there exist two constants $C, c>0$ such that

$$\|d_{f,h}w_k^{(1)}\|^2_{L^2(\Omega_+)}+\|d^*_{f,h}w_k^{(1)}\|^2_{L^2(\Omega_+)}\le Ce^{-c/h} \tag{7-14}$$

holds for all $k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}$. In addition, the estimates (7-12) and (7-13) with $\zeta_k(h)=\mathcal{O}(e^{-c_k/h})$ imply that the $w_k^{(1)}$ solve

$$\begin{cases}\Delta^{(1)}_{f,h}w_k^{(1)}=r_k \quad \text{on } \Omega_+\setminus\overline{\Omega}_-,\\ \boldsymbol{t}w_k^{(1)}\big|_{\partial\Omega_+\cup\partial\Omega_-}=0,\quad \boldsymbol{t}d^*_{f,h}w_k^{(1)}\big|_{\partial\Omega_-}=0,\quad \boldsymbol{t}d^*_{f,h}w_k^{(1)}\big|_{\partial\Omega_+}=\rho_k,\end{cases} \tag{7-15}$$

where r_k and ρ_k satisfy

$$\forall p\in\mathbb{N}\ \exists C_p>0\ \forall k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}\quad \|r_k\|_{W^{p,2}(\overline{\Omega}_+\setminus\Omega_-)}+\|\rho_k\|_{W^{p+1/2,2}(\partial\Omega_+)}\le C_pe^{-c'/h} \tag{7-16}$$

for some $c'>0$. The construction of the almost orthonormal basis of $\operatorname{Ran} 1_{[0,\nu(h)]}(\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-))$ is completed with the next lemma.

Lemma 7.10. *Assume Hypotheses 1, 5 and 6 and condition* (7-5), *and set*

$$u_k^{(1)}=1_{[0,\nu(h)]}(\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-))w_k^{(1)}$$

for any $k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}$. *Then* $(u_k^{(1)})_{k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}}$ *is an almost orthonormal basis of* $\operatorname{Ran} 1_{[0,\nu(h)]}(\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-))$.

Moreover,

$$\exists c>0\ \forall p\in\mathbb{N}\ \exists C_p>0\ \forall k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}\quad \|u_k^{(1)}-w_k^{(1)}\|_{W^{p,2}(\Omega_+\setminus\overline{\Omega}_-)}\le C_pe^{-c/h} \tag{7-17}$$

for all sufficiently small h.

Proof. Let us introduce $v_k^{(1)}=u_k^{(1)}-w_k^{(1)}$ for $k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}$. The one-form $v_k^{(1)}$ belongs to the form domain of $\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-)$ and the spectral theorem leads to

$$\nu(h)\|v_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}\le\|d_{f,h}w_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}+\|d^*_{f,h}w_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}\le Ce^{-c/h}\le Ce^{-c_1/h}$$

owing to (7-14) and $\sigma(\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\overline{\Omega}_-))\cap[0,\nu(h)]\subset[0,e^{-c_0/h}]$. With (2-5), this implies that $\|v_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}=\mathcal{O}(e^{-c_2/h})$. By using

$$\begin{aligned}h^2(\|dv_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}&+\|d^*v_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)})\\ &\le 2\|d_{f,h}v_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}+2\|d^*_{f,h}v_k^{(1)}\|^2_{L^2(\Omega_+\setminus\overline{\Omega}_-)}+C\|v_k^{(1)}\|^{(2)}_{L^2(\Omega_+\setminus\overline{\Omega}_-)},\end{aligned}$$

we obtain

$$\|v_k^{(1)}\|^2_{W^{1,2}(\Omega_+\setminus\overline{\Omega}_-)}=\mathcal{O}(h^{-2}e^{-c_2/h})=\mathcal{O}(e^{-c_2/(2h)}).$$

Thus, the almost orthonormality property of $(u_k^{(1)})_{k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}}$ is due to the orthonormality of $(w_k^{(1)})_{k\in\{1,\dots,m_1^D(\Omega_+\setminus\overline{\Omega}_-)\}}$.

The $W^{p,2}$ estimates (7-17) are then obtained by a bootstrap argument (induction on p) using the elliptic regularity of the Hodge Laplacian. With $\Delta^{D,(1)}_{f,h}(\Omega_+\setminus\Omega_-)u_k^{(1)}=\tilde{\mathcal{O}}(e^{-c_0/h})$ in any $W^{p,2}$ (see Lemma 6.3), (7-15) leads to

$$\begin{cases}\Delta_H v_k^{(1)} = r_k'(h) - h^{-2}(\Delta_{f,h}-h^2\Delta_H)v_k^{(1)},\\ \boldsymbol{t}v_k^{(1)}\big|_{\partial\Omega_+\cup\partial\Omega_-}=0, \quad \boldsymbol{t}d^*v_k^{(1)}\big|_{\partial\Omega_-}=0, \quad \boldsymbol{t}d^*v_k^{(1)}\big|_{\partial\Omega_+}=-h^{-1}\rho_k - h^{-1}\boldsymbol{i}_{\nabla f}v_k^{(1)},\end{cases}$$

where $\|r_k'(h)\|_{W^{p,2}(\Omega_+\setminus\bar{\Omega}_-)}$ satisfies the same estimate (7-16) as $\|r_k(h)\|_{W^{p,2}(\Omega_+\setminus\bar{\Omega}_-)}$. Using the fact that the zeroth-order differential operator $\Delta_{f,h}-h^2\Delta_H=|\nabla f|^2+h(\mathcal{L}_{\nabla f}+\mathcal{L}^*_{\nabla f})$ is bounded in L^∞-norm, we thus obtain the $W^{p,2}$ estimates (7-17) by induction on p. □

Proof of Proposition 7.5. Let us apply Theorem 2.4 and Corollary 6.9 to the almost orthonormal basis $(u_k^{(1)})_{1\le k\le m_1^D(\Omega_+\setminus\bar{\Omega}_-)}$ introduced in Lemma 7.10 (see Remark 6.6). From the estimate (7-17) and the fact that $\lim_{h\to 0} h\log\lambda_1^{(0)}(\Omega_+)=-2\kappa_f$, we deduce

$$\lambda_1^{(0)}(\Omega_+)=\frac{h^2\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\big(\int_{\partial\Omega_+}e^{-f/h}\boldsymbol{i}_n w_k^{(1)}(\sigma)\,d\sigma\big)^2}{\int_{\Omega_+}e^{-2f(x)/h}\,dx}(1+\mathcal{O}(e^{-c/h})),$$

$$\partial_n u_1^{(0)}\big|_{\partial\Omega_+}=-\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\frac{\int_{\partial\Omega_+}e^{-f(\sigma)/h}\boldsymbol{i}_n w_k^{(1)}(\sigma)\,d\sigma}{\big(\int_{\Omega_+}e^{-2f(x)/h}\,dx\big)^{1/2}}\boldsymbol{i}_n w_k^{(1)}+\mathcal{O}(e^{-(\kappa_f+c)/h}),$$

where the last remainder term is measured in $W^{p,2}(\partial\Omega_+)$-norm for any $p\in\mathbb{N}$. In particular, we deduce

$$\frac{e^{-f/h}}{\big(\int_{\Omega_+}e^{-2f(x)/h}\,dx\big)^{1/2}}\partial_n u_1^{(0)}\big|_{\partial\Omega_+}=-\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\bigg(\int_{\partial\Omega_+}\theta_k(\sigma)\,d\sigma\bigg)\theta_k+\mathcal{O}(e^{-(2\kappa_f+c)/h}) \quad \text{in } L^1(\partial\Omega_+)$$

and

$$\lambda_1^{(0)}(\Omega_+)=h^2\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\bigg(\int_{\partial\Omega_+}\theta_k\,d\sigma\bigg)^2(1+\mathcal{O}(e^{-c/h})), \tag{7-18}$$

where $\theta_k=\big(e^{-f/h}/\big(\int_{\Omega_+}e^{-2f(x)/h}\,dx\big)^{1/2}\big)\boldsymbol{i}_n w_k^{(1)}\big|_{\partial\Omega_+}$.

Using $\partial_n u_1^{(0)}\big|_{\partial\Omega_+}\le 0$ and the fact that the θ_k have disjoint supports, the following estimates hold:

$$\begin{aligned}\bigg(\int_{\Omega_+}e^{-2f(x)/h}\,dx\bigg)^{-\frac12}\big\|e^{-f/h}\partial_n u_1^{(0)}\big|_{\partial\Omega_+}\big\|_{L^1(\partial\Omega_+)} &= \sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)}\bigg(\int_{\partial\Omega_+}\theta_k(\sigma)\,d\sigma\bigg)^2+\mathcal{O}(e^{-(2\kappa_f+c)/h})\\ &= h^{-2}\lambda_1^{(0)}(\Omega_+)(1+\tilde{\mathcal{O}}(e^{-c/h})).\end{aligned}$$

In the last equality, we used (2-11) to get a lower bound on $\lambda_1^{(0)}(\Omega_+)$. By recalling that the Dirichlet boundary condition $u_1^{(0)}\big|_{\partial\Omega_+}=0$ implies

$$\partial_n[e^{-f/h}u_1^{(0)}]\big|_{\partial\Omega_+}=e^{-f/h}\partial_n u_1^{(0)}\big|_{\partial\Omega_+},$$

we thus get

$$-\frac{\partial_n[e^{-f/h}u_1^{(0)}]\big|_{\partial\Omega_+}}{\|\partial_n[e^{-f/h}u_1^{(0)}]\|_{L^1(\partial\Omega_+)}} = \frac{\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar\Omega_-)}\left(\int_{\partial\Omega_+}\theta_k\,d\sigma\right)\theta_k}{\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar\Omega_-)}\left(\int_{\partial\Omega_+}\theta_k\,d\sigma\right)^2} + \tilde{\mathcal{O}}(e^{-c/h}) \quad \text{in } L^1(\partial\Omega_+). \tag{7-19}$$

In order to get estimates from (7-18) and (7-19) in terms of f, it remains to approximate the quantities θ_k and $\int_{\partial\Omega_+}\theta_k\,d\sigma$ in the limit $h\to 0$. Recall that

$$\theta_k = \frac{e^{-f/h}}{\left(\int_{\Omega_+} e^{-2f(x)/h}\,dx\right)^{1/2}}\boldsymbol{i}_n w_k^{(1)}\big|_{\partial\Omega_+} \quad\text{and}\quad w_k^{(1)} = \frac{\chi_k z_k^{(1)}}{\|\chi_k z_k^{(1)}\|_{L^2(\mathcal{V}_k')}}.$$

The estimates are obtained using the Laplace method and the WKB expansion (7-11) together with (7-13) to approximate $z_k^{(1)}$.

- $\int_{\Omega_+} e^{-2f(x)/h}\,dx$: A direct application of the Laplace method gives

$$\int_{\Omega_+} e^{-2f(x)/h}\,dx = e^{-2f(U_0)/h}(\pi h)^{d/2}\big(\det(\operatorname{Hess} f)(U_0)\big)^{-1/2}(1+\mathcal{O}(h)),$$

  where U_0 is the unique global minimum of f.

- $\|\chi_k z_k^{(1)}\|_{L^2(\mathcal{V}_k')}$: Recall the coordinates around $U_k^{(1)}$ used in (7-10) and (7-11). Using these coordinates and (7-13), there is a $\mathcal{C}_0^\infty(\{x^d\le 0\})$ function $\alpha(x,h)\sim\sum_{k=0}^\infty \alpha_k(x)h^k$ with $\alpha_0(0)=1$ such that

$$\begin{aligned}
\|\chi_k z_k^{(1)}\|^2_{L^2(\mathcal{V}_k')} &= \int_{\{x^d\le 0\}} e^{-2\varphi_k(x)/h}\alpha(x,h)\,dx^1\cdots dx^d \\
&= \int_{\{x^d\le 0\}} e^{2\partial_n f(U_k^{(1)})x^d/h} e^{-\sum_{j=1}^{d-1}\lambda_j(x^j)^2/h}\alpha(x,h)\,dx^1\cdots dx^d \\
&= \frac{h}{2\partial_n f(U_k^{(1)})}\frac{(\pi h)^{(d-1)/2}}{\sqrt{\lambda_1\cdots\lambda_{d-1}}}(1+\mathcal{O}(h)) \\
&= \frac{(\pi h)^{(d+1)/2}}{2\pi\,\partial_n f(U_k^{(1)})\big(\det(\operatorname{Hess} f\big|_{\partial\Omega_+})(U_k^{(1)})\big)^{1/2}}(1+\mathcal{O}(h)).
\end{aligned}$$

  We applied the Laplace method to get the estimate of the integral (using the fact that $\partial_n f(U_k^{(1)})>0$ by (7-5)).

- θ_k: On the one hand, using $f(x)=f(U_k^{(1)})+\partial_n f(U_k^{(1)})x^d+\frac{1}{2}\sum_{j=1}^{d-1}\lambda_j(x^j)^2$ in a neighborhood of $U_k^{(1)}$ (see (7-10)), we have, on $\partial\Omega_+$ (so that $x^d=0$),

$$\begin{aligned}
&\chi_k\frac{e^{-f/h}}{\left(\int_{\Omega_+}e^{-f(x)/h}\,dx\right)^{1/2}} \\
&\qquad = \chi_k e^{-(f(U_k^{(1)})-f(U_0))/h}(\pi h)^{-d/4}\big(\det(\operatorname{Hess} f)(U_0)\big)^{1/4}e^{-\sum_{j=1}^{d-1}\lambda_j(x^j)^2/(2h)}(1+\mathcal{O}(h)).
\end{aligned}$$

On the other hand, the function $\boldsymbol{i}_n w_k^{(1)}\big|_{\partial\Omega_+} = \chi_k \boldsymbol{i}_n z_k^{(1)}\big|_{\partial\Omega_+} / \|\chi_k z_k^{(1)}\|_{L^2(\mathcal{V}_k')}$ satisfies

$$\boldsymbol{i}_n w_k^{(1)}\big|_{\partial\Omega_+} = \chi_k \frac{\sqrt{2\pi\,\partial_n f(U_k^{(1)})}\big(\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})\big)^{1/4}}{(\pi h)^{(d+1)/4}} e^{-\sum_{j=1}^{d-1}\lambda_j (x^j)^2/(2h)}(1+\mathcal{O}(h)).$$

From these two estimates, θ_k satisfies

$$\theta_k = A_k \chi_k e^{-\sum_{j=1}^{d-1}\lambda_j (x^j)^2/h}(1+\mathcal{O}(h)),$$

where

$$A_k = \frac{\sqrt{2\pi\,\partial_n f(U_k^{(1)})}}{(\pi h)^{(2d+1)/4}}\big(\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})\big)^{1/4}\big(\det(\operatorname{Hess} f)(U_0)\big)^{1/4} e^{-(f(U_k^{(1)})-f(U_0))/h}.$$

- $\int_{\partial\Omega_+}\theta_k$: The Laplace method implies that

$$\int e^{-\sum_{j=1}^{d-1}\lambda_j (x^j)^2/h}\,dx^1\cdots dx^{d-1} \\ = \frac{(\pi h)^{(d-1)/2}}{\sqrt{\lambda_1\cdots\lambda_{d-1}}}(1+\mathcal{O}(h)) = (\pi h)^{(d-1)/2}\big(\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})\big)^{-1/2}(1+\mathcal{O}(h)).$$

We thus obtain

$$\int_{\partial\Omega_+}\theta_k = \frac{\sqrt{2\pi\,\partial_n f(U_k^{(1)})}\big(\det(\operatorname{Hess} f)(U_0)\big)^{1/4}}{(\pi h)^{3/4}\big(\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})\big)^{1/4}} e^{-(f(U_k^{(1)})-f(U_0))/h}(1+\mathcal{O}(h)).$$

Putting together the above information and using (7-18) and (7-19) finally implies

$$\lambda_1^{(0)}(\Omega_+) = \sqrt{\frac{h\det(\operatorname{Hess} f)(U_0)}{\pi}}\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar\Omega_-)}\frac{2\partial_n f(U_k^{(1)})}{\sqrt{\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})}} e^{-2(f(U_k^{(1)})-f(U_0))/h}(1+\mathcal{O}(h)),$$

which is exactly (7-6), and

$$-\frac{\partial_n[e^{-f/h}u_1^{(0)}]\big|_{\partial\Omega_+}}{\big\|\partial_n[e^{-f/h}u_1^{(0)}]\big\|_{L^1(\partial\Omega_+)}} \\ = \frac{\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar\Omega_-)}\partial_n f(U_k^{(1)})e^{-2(f(U_k^{(1)})-f(U_0))/h}\chi_k e^{-\sum_{j=1}^{d-1}\lambda_j(x^j)^2/h}}{(\pi h)^{(d-1)/2}\sum_{k'=1}^{m_1^D(\Omega_+\setminus\bar\Omega_-)}\big(\partial_n f(U_{k'}^{(1)})/\sqrt{\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_{k'}^{(1)})}\big)e^{-2(f(U_{k'}^{(1)})-f(U_0))/h}}(1+\mathcal{O}(h)). \tag{7-20}$$

We thus obtain estimates of $\lambda_1^{(0)}(\Omega_+)$ and $-\partial_n(e^{-f/h}u_1^{(0)})\big|_{\partial\Omega_+}/\|\partial_n(e^{-f/h}u_1^{(0)})\|_{L^1(\partial\Omega_+)}$ in terms of second-order Taylor expansions of f around the local minima $U_k^{(1)}$. This ends the first step of the proof.

Actually, the two estimates (7-6) and (7-20) can be rewritten in a simpler form using the Laplace method again. By recalling the equality $f(x) = f(U_k^{(1)}) + \partial_n f(U_k^{(1)})x^d + \frac{1}{2}\sum_{j=1}^{d-1}\lambda_j(x^j)^2$ in a neighborhood of $U_k^{(1)}$, the Laplace method gives, by similar computations to those performed above,

$$\frac{\int_{\partial\Omega_+} 2\partial_n f(\sigma) e^{-2f(\sigma)/h}\, d\sigma}{\int_{\Omega_+} e^{-2f(x)/h}\, dx} = \sqrt{\frac{h \det(\operatorname{Hess} f)(U_0)}{\pi}} \sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \frac{2\partial_n f(U_k^{(1)})}{\sqrt{\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_k^{(1)})}} e^{-2(f(U_k^{(1)})-f(U_0))/h}(1+\mathcal{O}(h)),$$

$$\frac{(2\partial_n f) e^{-2f/h}\big|_{\partial\Omega_+}}{\|(2\partial_n f)e^{-2f/h}\|_{L^1(\partial\Omega_+)}} = \frac{\sum_{k=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \partial_n f(U_k^{(1)}) e^{-2(f(U_k^{(1)})-f(U_0))/h} \chi_k e^{-\sum_{j=1}^{d-1}\lambda_j (x^j)^2/h}}{(\pi h)^{(d-1)/2} \sum_{k'=1}^{m_1^D(\Omega_+\setminus\bar{\Omega}_-)} \big(\partial_n f(U_{k'}^{(1)})/\sqrt{\det\big(\operatorname{Hess} f\big|_{\partial\Omega_+}\big)(U_{k'}^{(1)})}\big) e^{-2(f(U_{k'}^{(1)})-f(U_0))/h}} + \mathcal{O}(h),$$

where the last remainder term is measured in $L^1(\partial\Omega_+)$-norm. Comparing with the two estimates (7-6) and (7-20) above, we thus obtain (7-7) and (7-8). This concludes the proof. □

7B. ***Beyond Morse assumptions.*** In this section, we discuss Hypotheses 3 and 4 for functions f which do not fulfill the Morse assumptions of Hypothesis 5 above. In Sections 7B2 and 7B3, we present two examples (respectively in dimension 1 and 2) of functions f which do not fulfill Hypothesis 5 but for which Hypotheses 3 and 4 still hold true. Section 7B1 is first devoted to a few remarks that will be useful in the examples we will discuss below.

7B1. *General remarks.* First, we will use the duality between the chain complexes associated with $d_{f,h}$ and $d_{f,h}^*$. More precisely, conjugating with the Hodge $\star$-operator exchanges p- and $(\dim M - p)$-forms, d and d^*, f and $-f$, Neumann and Dirichlet boundary conditions. This was used extensively in [Le Peutrec 2011; Le Peutrec et al. 2013].

Second, the following lemma will also be useful. It is a variant of Proposition 3.7.

Lemma 7.11. *Let Ω be a regular bounded domain of the Riemannian manifold (M, g) and let $f \in \mathcal{C}^\infty(\bar{\Omega})$ be such that $(\nabla f)^{-1}(\{0\})$ has a unique nonempty connected component in Ω.*

- *If $\partial_n f\big|_{\partial\Omega} > 0$ then the two first eigenvalues of $\Delta_{f,h}^{N,(0)}(\Omega)$ satisfy*

$$\mu_1^{(0)}(\Omega) = 0 \quad \textit{and} \quad \lim_{h\to 0} h \log \mu_2^{(0)}(\Omega) = 0.$$

- *If $\partial_n f\big|_{\partial\Omega} < 0$ and $|\nabla f|^2 - h\Delta f \geq 0$ in Ω for all $h \in (0, h_0)$, then the first eigenvalue of $\Delta_{f,h}^{D,(0)}(\Omega)$ satisfies*

$$\lim_{h\to 0} h \log \lambda_1^{(0)}(\Omega) = 0.$$

Proof. Up to the addition of a constant to the function f (which only affects the normalization of $e^{-f/h}$), one may assume without loss of generality that $f \equiv 0$ on $(\nabla f)^{-1}(\{0\})$ (using the connectedness assumption on $(\nabla f)^{-1}(\{0\})$). Then, $f \geq 0$ in Ω when $\partial_n f\big|_{\partial\Omega} > 0$, and $f \leq 0$ when $\partial_n f\big|_{\partial\Omega} < 0$.

The fact that $\mu_1^{(0)}(\Omega)=0$ is obvious, by considering the associated eigenvector $e^{-f/h}$. The Witten Laplacian acting on functions is the Schrödinger-type operator

$$\Delta^{(0)}_{f,h}=-h^2\Delta+|\nabla f|^2-h(\Delta f).$$

Since the function $|\nabla f|^2-h\Delta f$ is uniformly bounded in $\overline{\Omega}$, the two inequalities

$$\limsup_{h\to 0} h\log \mu_2^{(0)}(\Omega)\le 0 \quad\text{and}\quad \limsup_{h\to 0} h\log\lambda_1^{(0)}(\Omega)\le 0$$

are consequences of the min–max principle. For the Dirichlet case, any fixed nonzero function in $\mathcal{C}_0^\infty(\Omega)$ will provide an $\mathcal{O}(1)$ Rayleigh quotient. For the Neumann case, consider two regular functions $\chi_1,\ \chi_2\in\mathcal{C}_0^\infty(\Omega)$ such that $\operatorname{supp}\chi_1\cap\operatorname{supp}\chi_2=\varnothing$ and $\|\chi_1\|_{L^2(\Omega)}=\|\chi_2\|_{L^2(\Omega)}=1$, and take $\psi_h=\alpha_1(h)\chi_1+\alpha_2(h)\chi_2$ such that $\|\psi_h\|^2_{L^2}=|\alpha_1(h)|^2+|\alpha_2(h)|^2=1$ and $\langle\psi_h, e^{-f/h}\rangle_{L^2(\Omega)}=0$. We get $\langle\psi_h,\Delta^{N,(0)}_{f,h}\psi_h\rangle_{L^2(\Omega)}=\mathcal{O}(1)$ and the min–max principle applied to $\Delta^{N,(0)}_{f,h}(\Omega)$ on the orthogonal of $e^{-f/h}$ yields $\mu_2^{(0)}(\Omega)=\mathcal{O}(1)$ as $h\to 0$.

Let us first consider the case where $\partial_n f\big|_{\partial\Omega}<0$ and $|\nabla f|^2-h\Delta f\ge 0$. It remains to prove that $\liminf_{h\to 0} h\log\lambda_1^{(0)}(\Omega)\ge 0$. Let ω be a normalized eigenfunction associated with $\lambda_1^{(0)}(\Omega)$, so $\Delta^{D,(0)}_{f,h}(\Omega)\omega=\lambda_1^{(0)}(\Omega)\omega$ and $\|\omega\|_{L^2(\Omega)}=1$. Using Lemma 3.1 with $\varphi=0$ and the Poincaré inequality, we get

$$\lambda_1^{(0)}(\Omega)\ge h^2\|\nabla\omega\|^2_{L^2(\Omega)}\ge C_\Omega h^2.$$

This concludes the proof in the case $\partial_n f\big|_{\partial\Omega}<0$ and $|\nabla f|^2-h\Delta f\ge 0$.

Let us now consider the case $\partial_n f\big|_{\partial\Omega}>0$. It remains to prove that $\liminf_{h\to 0} h\log\mu_2^{(0)}(\Omega)\ge 0$. Let us reason by contradiction, by assuming that there exists $c>0$ and a sequence $(h_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to\infty} h_n=0 \quad\text{and}\quad \mu_2^{(0)}(\Omega)\le e^{-c/h_n} \quad\text{with } c>0.$$

Notice that $\mu_2^{(0)}(\Omega)$ depends on n. Let us introduce ω_n, a normalized eigenfunction associated with $\mu_2^{(0)}(\Omega)$, so $\Delta^{N,(0)}_{f,h_n}\omega_n=\mu_2^{(0)}(\Omega)\omega_n$ and $\|\omega_n\|_{L^2(\Omega)}=1$. Notice that $\int_\Omega\omega_n e^{-f/h_n}=0$. For $\varepsilon>0$, consider the open set

$$K_\varepsilon=\big\{x\in\overline{\Omega}: d\big(x,(\nabla f)^{-1}(\{0\})\big)<\varepsilon\big\},$$

so that $\overline{K}_\varepsilon$ is contained in Ω for $\varepsilon\in(0,\varepsilon_0)$ and ε_0 sufficiently small. Take a partition of unity $\chi_1^2+\chi_2^2\equiv 1$ in $\overline{\Omega}$ such that $\chi_i\in\mathcal{C}^\infty(\overline{\Omega})$, $\chi_1\equiv 1$ in a neighborhood of $K_{\varepsilon/2}$ and $\operatorname{supp}\chi_1\subset K_\varepsilon$. The IMS localization formula (see, for example, [Cycon et al. 1987]) gives

$$\begin{aligned}e^{-c/h_n}&\ge\langle\omega_n,\Delta^{N,(0)}_{f,h_n}(\Omega)\omega_n\rangle_{L^2(\Omega)}\\&=\langle\chi_1\omega_n,\Delta^{N,(0)}_{f,h_n}(\Omega)\chi_1\omega_n\rangle_{L^2(\Omega)}+\langle\chi_2\omega_n,\Delta^{N,(0)}_{f,h_n}(\Omega)\chi_2\omega_n\rangle_{L^2(\Omega)}-h_n^2\sum_{j=1}^2\|\omega_n\nabla\chi_j\|^2_{L^2(\Omega)}.\end{aligned}\tag{7-21}$$

The lower bound (which is a consequence of $|\nabla f|^2>0$ on $\operatorname{supp}\chi_2$ and $\partial_n\chi_2=0$ on $\partial\Omega$)

$$\langle\chi_2\omega_n,\Delta^{N,(0)}_{f,h_n}(\Omega)\chi_2\omega_n\rangle_{L^2(\Omega)}\ge\langle\chi_2\omega_n,|\nabla f|^2\chi_2\omega_n\rangle_{L^2(\Omega)}-Ch_n\|\chi_2\omega_n\|^2_{L^2(\Omega)}\ge\frac{1}{C_\varepsilon}\|\chi_2\omega_n\|^2_{L^2(\Omega)}$$

for n sufficiently large together with (7-21) implies

$$\forall\delta > 0 \ \forall\varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N \quad \|\omega_n\|^2_{L^2(K_\varepsilon)} \geq 1 - \delta.$$

Since $(\nabla f)^{-1}(\{0\})$ is assumed to be connected and, for every point of the open set K_ε, the gradient flow associated with f defines a path to $(\nabla f)^{-1}(\{0\})$, K_ε is a connected open set. The function $v_n = \omega_n\big|_{K_\varepsilon}$ belongs to $W^{1,2}(K_\varepsilon)$ with

$$h_n^2 e^{-2C\varepsilon^2/h_n} \|de^{f/h_n} v_n\|^2_{L^2(K_\varepsilon)} \leq \|d_{f,h} v_n\|^2_{L^2(K_\varepsilon)} \leq e^{-c/h_n},$$

thanks to the fact that

$$\exists C > 0 \ \forall x \in K_\varepsilon \quad 0 \leq f(x) \leq C\varepsilon^2.$$

By choosing $\varepsilon > 0$ so that $c - 2C\varepsilon^2 > 0$, the spectral gap estimate for the Neumann Laplacian in Ω (or equivalently the Poincaré–Wirtinger inequality in Ω) provides a constant C_n such that

$$\lim_{n\to\infty} \|e^{f/h_n} v_n - C_n\|_{L^2(K_\varepsilon)} = 0.$$

We thus deduce

$$\lim_{n\to\infty} \|\omega_n - C_n e^{-f/h_n}\|_{L^2(K_\varepsilon)} = 0 \quad \text{with} \quad \|\omega_n\|^2_{L^2(K_\varepsilon)} \geq 1 - \delta, \ \|\omega_n\|_{L^2(\Omega)} = 1.$$

For $\delta < 1$, this is in contradiction with $\int_\Omega \omega_n e^{-f/h_n} = 0$. □

7B2. *A one-dimensional example.* In this section, we exhibit a simple one-dimensional example of a function f satisfying Hypotheses 3 and 4 though not being a Morse function. An extension is then briefly discussed.

Proposition 7.12. *Consider a function $f \in \mathcal{C}^\infty(\overline{\Omega}_+)$, $\Omega_+ = (a_+, b_+)$ with $a_+ < b_+$ two real numbers, such that*

$$f^{-1}(0) = (f')^{-1}(0) = [a_1, b_1], \quad -\infty < a_+ < a_1 \leq b_1 < b_+ < +\infty,$$
$$f'(a_+) < 0 \quad \text{and} \quad f'(b_+) > 0.$$

Then, for any $\Omega_- = (a_-, b_-)$ such that $a_+ < a_- < a_1 \leq b_1 < b_- < b_+$, Hypotheses 3 and 4 are valid with $m_0^N(\Omega_-) = 1$, $m_1^N(\Omega_-) = 0$ and $m_1^D(\Omega_+ \setminus \overline{\Omega}_-) = 2$.

Notice that, for this example, Hypotheses 1 and 2 are also satisfied, which means that the results of Theorem 2.4 are valid.

Proof. On an interval I with the Euclidean metric, the one-forms can be written as $u^{(1)} = u_1(x)\,dx$. The Witten Laplacians $\Delta^{(p)}_{f,h}(I)$ with $p = 0, 1$ are then given by

$$\Delta^{(0)}_{f,h}(I) u^{(0)} = (-h^2 \partial_{x,x} + |\partial_x f|^2 - h(\partial_{x,x} f)) u^{(0)},$$
$$\Delta^{(1)}_{f,h}(I)(u_1 dx) = \big[(-h^2 \partial_{x,x} + |\partial_x f|^2 + h(\partial_{x,x} f)) u_1\big]\,dx.$$

The Dirichlet boundary conditions are given by

$$u^{(0)} = 0 \quad \text{on } \partial I \qquad \text{and} \qquad -h\partial_x u_1 + (\partial_x f) u_1 = 0 \quad \text{on } \partial I,$$

while the Neumann boundary conditions are given by

$$h\partial_x u^{(0)} + (\partial_x f)u^{(0)} = 0 \quad \text{on } \partial I \qquad \text{and} \qquad u_1 = 0 \quad \text{on } \partial I.$$

This is a particular case of the general duality recalled at the beginning of Section 7B1. Let us now check Hypotheses 3 and 4.

First, $e^{-f/h}$ belongs to the kernel of $\Delta^{N,(0)}_{f,h}(\Omega_-)$. A direct application of Lemma 7.11 shows that (2-6) holds for $p=0$ with $m_0^N(\Omega_-)=1$. Second, by the duality argument, proving that (2-6) holds for $p=1$ with $m_1^N(\Omega_-)=0$ is equivalent to proving that there are no exponentially small eigenvalues for $\Delta^{D,(0)}_{-f,h}(\Omega_-)$ (notice that f has been changed to $-f$). But this is a consequence of the second part of Lemma 7.11, since f is convex. Finally, note that the condition (2-8) is empty, since the only exponentially small eigenvalue of $\Delta^{N,(0)}_{f,h}(\Omega_-)$ is 0. This shows that Hypothesis 3 holds.

The open set $\Omega_+ \setminus \overline{\Omega}_-$ is the disjoint union of the two open intervals (a_+, a_-) and (b_-, b_+). On each of them, $\partial_x f$ does not vanish and the Morse assumptions of Hypothesis 5 are satisfied. On (a_+, a_-) (resp. (b_-, b_+)), f has one generalized critical point of index 1 at a_+ (resp. at b_+). Therefore, using the results of [Helffer and Nier 2006] (see Section 7A1), (2-9) holds with $m_1^D(\Omega_+ \setminus \overline{\Omega}_-)=2$. This shows that Hypothesis 4 holds. □

It is not difficult to treat the case when $f \in \mathcal{C}^\infty([a_+, b_+])$ has a finite number of critical intervals,

$$(f')^{-1}(\{0\}) = \bigcup_{n=1}^{2N+1} [a_n, b_n], \qquad a_+ < a_1 \le b_1 < \cdots < a_{2N+1} \le b_{2N+1} < b_+,$$

with $f'(a_+) < 0$ and $f'(b_+) > 0$. Again, $\Omega_- = (a_-, b_-)$, with $a_+ < a_- < a_1 < b_{2N+1} < b_- < b_+$. The local problems around every $[a_n, b_n]$ can be studied with the help of the duality argument and Lemma 7.11. Using an argument based on a partition of unity, one can check that (2-6) and (2-9) hold with $m_0^N(\Omega_-)=2N+1$, $m_1^N(\Omega_-)=2N$ and $m_1^D(\Omega_+ \setminus \overline{\Omega}_-)=2N+2$. Hypothesis 1 is of course satisfied. Ensuring that Hypothesis 2 and condition (2-8) hold then requires us to correctly choose the heights of the critical values. They hold, for example, when $\max_{1\le n\le 2N+1} f(a_i) < \min\{f(a_+),\ f(b_+)\}$ and when $f(a_1)$ and $f(b_{2N+1})$ are the two smallest critical values.

7B3. *A two-dimensional example.* This example is inspired by the work of [Bismut 1986; Helffer and Sjöstrand 1987; 1988] on Bott inequalities. We consider the following $\mathcal{C}^\infty$ radial functions in $\mathbb{R}^2$:

$$\begin{aligned} \varphi_{\text{in}}(x) &= e^{-1/(|x|^2-1)^2} 1_{[0,1]}(|x|), \\ \varphi_{\text{ext}} &\equiv 0 \quad \text{for } |x| \le 1, \quad \varphi_{\text{ext}} \text{ strictly convex in } \{|x|>1\}. \end{aligned}$$

The domain Ω_+ is the disc $D((-R,0),2R)$ and Ω_- the disc $D((-R,0),2R-1)$ with $R>3$. The function f is defined by $f(x)=\varphi_{\text{in}}(x)+\varphi_{\text{ext}}\big(\tfrac{1}{2}x\big)$. The level sets of the function f are represented in Figure 5.

Proposition 7.13. *When $R>3$ is chosen large enough, the above triple (Ω_+, Ω_-, f) fulfills Hypotheses 1, 2, 3 and 4 with $m_0^N(\Omega_-)=1$, $m_1^N(\Omega_-)=1$ and $m_1^D(\Omega_+ \setminus \overline{\Omega}_-)=1$.*

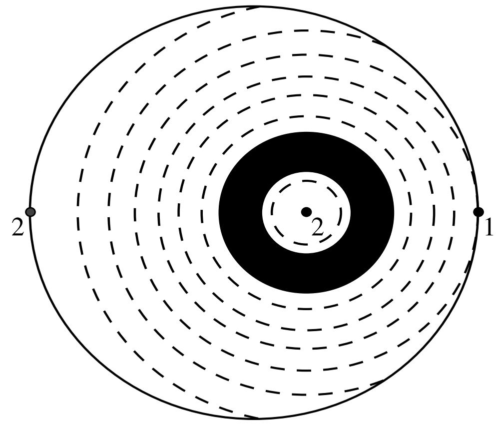

Figure 5. Only Ω_+ is represented. The level sets of f are represented by dashed lines. The black area is the 0 level set. The dots indicate the generalized critical points, together with their indices (for Dirichlet boundary conditions).

Proof. Thanks to the convexity assumption on $x \mapsto \varphi_{\text{ext}}\big(\frac{1}{2}x\big)$ and its local behavior around $\{|x| = 2\}$, Hypotheses 1 and 2 hold for $R > 3$ large enough.

The choice of non-0-centered disks for Ω_+ and Ω_- while f is a radial function implies that $f\big|_{\partial\Omega_+}$ has a unique local minimum and therefore, using the results recalled in Section 7A1, (2-9) is satisfied with $m_1^D(\Omega_+ \setminus \overline{\Omega}_-) = 1$. This shows that Hypothesis 4 holds.

The fact that (2-6) holds for $p = 0$ with $m_0^N(\Omega_-) = 1$ is a direct application of Lemma 7.11. This also implies that the condition (2-8) is void. It only remains to prove that (2-6) holds for $p = 1$ with $m_1^N(\Omega_-) = 1$. We will actually prove that (2-6) holds for $p = 2$ with $m_2^N(\Omega_-) = 1$. Then the quasi-isomorphism with the absolute cohomology of the disc (see Section 4A) gives $m_2^N(\Omega_-) - m_1^N(\Omega_-) + m_0^N(\Omega_-) = 1$, which indeed implies $m_1^N(\Omega_-) = 1$. Moreover, by the duality argument, (2-6) holds for $p = 2$ with $m_2^N(\Omega_-) = 1$ if (2-9) holds for $p = 0$ with $m_0^D(\Omega_-) = 1$, f being changed into $-f$. The proof of this claim will conclude the demonstration.

In the rest of this proof, $m_0^D(\Omega_-)$ denotes the number of small eigenvalues for $\Delta_{-f,h}^{D,(0)}(\Omega_-)$. The function $-f$ has a local minimum at $x = (0, 0)$. Applying the min–max principle with a quasimode $\chi(x)e^{f(x)/h}$, where χ is a smooth nonnegative function such that $\chi \equiv 1$ on $\big\{|x| \le \frac{1}{4}\big\}$ and $\chi \equiv 0$ on $\big\{|x| \ge \frac{1}{2}\big\}$, implies that $m_0^D(\Omega_-) \ge 1$.

Let us now consider $\omega \in D(\Delta_{-f,h}^{D,(0)}(\Omega_-))$, a normalized eigenvector associated with an exponentially small eigenvalue, so $\langle \omega, \Delta_{-f,h}^{D,(0)}(\Omega_-)\omega\rangle_{L^2(\Omega_-)} \le e^{-c/h}$ for some $c > 0$. Let $\chi_1^2 + \chi_2^2 = 1$ be a partition of unity on Ω_- with $\chi_1^2 \equiv 1$ on $\{|x| \le \varepsilon\}$ and $\chi_1^2 \equiv 0$ on $\{|x| \ge 2\varepsilon\}$ (for $\varepsilon < \frac{1}{4}$). The IMS localization formula gives

$$\langle \omega, \Delta_{-f,h}^{D,(0)}(\Omega_-)\omega\rangle_{L^2(\Omega_-)} = \langle \chi_1\omega, \Delta_{-f,h}^{D,(0)}(\Omega_-)\chi_1\omega\rangle_{L^2(\Omega_-)} + \langle \chi_2\omega, \Delta_{-f,h}^{D,(0)}(\Omega_-)\chi_2\omega\rangle_{L^2(\Omega_-)} - h^2 \sum_{j=1}^{2} \|\omega\nabla\chi_j\|_{L^2(\Omega_-)}^2. \quad (7\text{-}22)$$

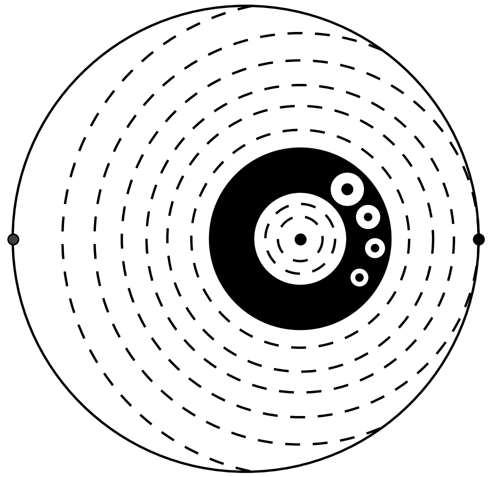

Figure 6. A variant of Figure 5 with $N = 4$. The supports of the additional terms in f (compared with Figure 5) are represented by the white disks.

The second term of the right-hand side equals $\langle \chi_2\omega, \Delta^{D,(0)}_{-f,h}(\Omega)\chi_2\omega\rangle_{L^2(\Omega_-\setminus\Omega_\varepsilon)}$ with $\Omega_\varepsilon = \{x \in \Omega_- : |x| \le \varepsilon\}$. Our choice of the function $f(x) = \varphi_{\text{in}}(x) + \varphi_{\text{ext}}\big(\frac{1}{2}x\big)$ ensures that, for $h \in (0, h_0)$ with h_0 small enough, $|\nabla f|^2 + h\Delta f$ is nonnegative on $\Omega_- \setminus \Omega_\varepsilon$. The second part of Lemma 7.11 thus implies that there exists a function ν of h such that

$$\langle \chi_2\omega, \Delta^{D,(0)}_{-f,h}(\Omega)\chi_2\omega\rangle_{L^2(\Omega_-\setminus\Omega_\varepsilon)} \ge \nu(h)\|\chi_2\omega\|^2_{L^2(\Omega_-\setminus\Omega_\varepsilon)}$$

with $\liminf_{h\to 0} h \log \nu(h) = 0$. In addition, exponential decay estimates based on the Agmon identity imply that $\sum_{j=1}^2 \|\omega\nabla\chi_j\|^2_{L^2(\Omega_-)} = \mathcal{O}(e^{-c/h})$, since $|\nabla f| > 0$ on $\operatorname{supp}(\chi_1) \cup \operatorname{supp}(\chi_2)$ (this is obtained by adapting the arguments of Proposition 3.4, for example). By using the IMS localization formula (7-22), we thus obtain that $\|\chi_2\omega\|_{L^2(\Omega\setminus\Omega_\varepsilon)}$ goes to zero when h goes to zero, and thus that $\lim_{h\to 0} \|\chi_1\omega\|_{L^2(\Omega_-)} = \lim_{h\to 0} \|\omega\|_{L^2(\Omega_\varepsilon)} = 1$. Using then the same argument as in the end of the proof of the first part of Lemma 7.11, we obtain that, for sufficiently small ε, $\lim_{h\to 0} \|\omega - C_h e^{f/h}\|_{L^2(\Omega_\varepsilon)} = 0$ for some constant $C_h \in \mathbb{R}$. The two limits $\lim_{h\to 0} \|\omega\|_{L^2(\Omega_\varepsilon)} = 1$ and $\lim_{h\to 0} \|\omega - C_h e^{f/h}\|_{L^2(\Omega_\varepsilon)} = 0$ imply that, in the asymptotic $h \to 0$, ω cannot be orthogonal to $\chi e^{f/h}$ (recall that $\chi \equiv 1$ on Ω_ε), which is in the spectral subspace associated with exponentially small eigenvalues. This concludes the proof. □

It is not difficult to adapt the previous argument to the case when the function f has several local maxima. Set $(x_0, r_0) = (0, 1)$ and consider a finite number of points and radii $(x_k, r_k)_{1\le k\le N}$ such that the open discs $D(x_k, r_k)$, $k = 0, \dots, N$, are all disjoint and included in $D(0, 2)$. Let us consider the function $f(x) = \varphi_{\text{ext}}\big(\frac{1}{2}x\big) + \sum_{k=0}^N \varphi_{\text{in}}((x - x_k)/r_k)$ (see Figure 6). Then Hypotheses 1, 2, 3 and 4 hold with $m_0^N(\Omega_-) = 1$, $m_1^N(\Omega_-) = N + 1$ and $m_1^D(\Omega_+ \setminus \overline{\Omega}_-) = 1$.

Remark 7.14. Interestingly, one can extend the last example to build a function f for which Hypothesis 3 is *not* satisfied. Consider an infinite sequence $(x_k, r_k)_{k\in\mathbb{N}}$ with $x_0 = 0$ and $r_0 = 1$ such that the open discs $D(x_k, r_k)$, $k \ge 0$, are all disjoint and included in $D(0, 2)$. Take the function $f(x) =$

$\varphi_{\text{ext}}\left(\frac{1}{2}x\right)+\sum_{k=0}^{\infty}(r_k^k/(1+k^2))\varphi_{\text{in}}((x-x_k)/r_k)$ in the domain $\Omega_-=D((-R,0),2R-1)$ with $R>3$ large enough. By Lemma 7.11, we know $m_0^N(\Omega_-)=1$, while quasimodes associated with every x_k show that the number of eigenvalues of $\Delta_{f,h}^{N,(2)}(\Omega_-)$ (or equivalently $\Delta_{-f,h}^{D,(0)}(\Omega_-)$) lying in $[0,e^{-\delta/h}]$ is larger than any fixed $n\in\mathbb{N}$ for h sufficiently small. Using, as in the proof of Proposition 7.13, the identity $m_2^N(\Omega_-)-m_1^N(\Omega_-)+m_0^N(\Omega_-)=1$, the number of eigenvalues of $\Delta_{f,h}^{N,(1)}(\Omega_-)$ lying in $[0,e^{-\delta/h}]$ is thus also larger than any $n\in\mathbb{N}$ for h sufficiently small. Thus Hypothesis 3 is not satisfied.

Actually, there are up to now no satisfactory necessary and sufficient conditions which guarantee that Witten Laplacians with general $\mathcal{C}^\infty$ potentials have a finite number of exponentially small eigenvalues.

Appendix: Riemannian geometry formulas

For the sake of completeness and in order to help the reader not so familiar with those tools, here is a list of formulas of Riemannian geometry which were used in this text. We refer the reader, for example, to [Abraham and Marsden 1978; Cycon et al. 1987; Gallot et al. 2004; Sternberg 1964; Goldberg 1970] for introductory texts in differential and Riemannian geometry. We also consider here only real-valued differential forms (the extension to complex-valued differential forms is easy).

Let (M,g) be a d-dimensional Riemannian manifold. The tangent (resp. cotangent) bundle is denoted by TM (resp. T^*M) and its fiber over $x\in M$ by T_xM (resp. T_x^*M). The exterior algebra over T_x^*M is $\bigwedge T_x^*M=\bigoplus_{p=0}^d\bigwedge^p T_x^*M$ endowed with the exterior product $\wedge$, and the associated fiber bundle is denoted by $\bigwedge T^*M=\bigoplus\bigwedge^p T^*M$. The exterior product of p elements $(\varphi_i)_{1\le i\le p}$ of T_x^*M is defined by

$$\varphi_1\wedge\cdots\wedge\varphi_p=\sum_{\sigma\in\mathfrak{S}_{\{1,\dots,p\}}}\epsilon_{\{1,\dots,p\}}(\sigma)\varphi_{\sigma(1)}\otimes\cdots\otimes\varphi_{\sigma(p)},$$

where $\epsilon_E(\pi)$ is the signature of the permutation $\pi\in\mathfrak{S}_E$. Differential forms are sections of this fiber bundle and their regularity is encoded by the notation: $\bigwedge\mathcal{C}^\infty(M)$ is the set of $\mathcal{C}^\infty$-differential forms, $\bigwedge L^2(M)$ is the set of L^2-differential forms, and so on. This notation was used in the present text for the sake of conciseness. A more standard and general notation would be $\mathcal{C}^\infty(M;\bigwedge T^*M)$, where $\mathcal{C}^\infty(M;E)$ more generally stands for the set of $\mathcal{C}^\infty$ sections of the differential fiber bundle (E,Π) on M with $\Pi:E\to M$ (a section $x\mapsto s(x)$ satisfies $\Pi(s(x))=x$).

In a local coordinate system $(x^1,\dots,x^d)$, a basis of $\bigwedge^p T_x^*M$ is formed by the elements

$$dx^I=dx^{i_1}\wedge\cdots\wedge dx^{i_p},\quad I=\{i_1,\dots,i_p\},\quad i_1<\cdots<i_p.$$

Here and in the following, $I=\{i_1,\dots,i_p\}$ denotes a subset of $\{1,\dots,d\}$ with $\#I=p$ elements, which can be described equivalently as an ordered p-tuple $(i_1,\dots,i_p)$ with $i_1<\cdots<i_p$.

A differential form $\omega\in\bigwedge^p T^*M$ is written

$$\omega=\sum_{\#I=p}\omega_I(x)\,dx^I,$$

and its differential is given by

$$d\omega=\sum_{\#I=p}\partial_{x^i}\omega_I(x)\,dx^i\wedge dx^I.$$

Remember that the exterior product is bilinear associative and antisymmetric:

$$\omega_1 \wedge \omega_2 = (-1)^{p_1 p_2} \omega_2 \wedge \omega_1, \quad \omega_i \in \bigwedge^{p_i} T_x^* M.$$

The differential and the $\wedge$ product satisfy $d \circ d = 0$ and

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^{p_1} \omega_1 \wedge d\omega_2, \quad \omega_i \in \bigwedge^{p_i} \mathcal{C}^\infty(M).$$

A $\mathcal{C}^\infty$ vector field X on M is a $\mathcal{C}^\infty$ section of TM, that is, $X \in \mathcal{C}^\infty(M; TM)$. The interior product $\boldsymbol{i}_X$ is the local operation defined for $X_x \in T_x M$ and $\omega_x \in \bigwedge^p T_x^* M$ by

$$\boldsymbol{i}_{X_x} \omega_x(T_2, \dots, T_p) = \omega_x(X_x, T_2, \dots, T_p) \quad \text{for all } T_2, \dots, T_p \in T_x M. \tag{A-1}$$

For $X \in \mathcal{C}^\infty(M; TM)$ and $\omega_i \in \bigwedge^{p_i} \mathcal{C}^\infty(M)$, one has

$$\boldsymbol{i}_X(\omega_1 \wedge \omega_2) = (\boldsymbol{i}_X \omega_1) \wedge \omega_2 + (-1)^{p_1} \omega_1 \wedge (\boldsymbol{i}_X \omega_2).$$

When $\Phi : M \to N$ is a $\mathcal{C}^\infty$ map, Φ_* denotes the functorial push-forward and Φ^* the functorial pull-back. For a $\mathcal{C}^\infty$ map Φ and two forms ω_1, ω_2, one has

$$\Phi^*(d\omega_1) = d(\Phi^* \omega_1), \quad \Phi^*(\omega_1 \wedge \omega_2) = (\Phi^* \omega_1) \wedge (\Phi^* \omega_2).$$

When Φ is a diffeomorphism, ω a p-form and X a vector field,

$$\Phi^* \boldsymbol{i}_X \omega = \boldsymbol{i}_{\Phi^* X} \Phi^* \omega.$$

When Φ is a diffeomorphism given by the exponential map of a vector field X, we can define the Lie derivative

$$\mathcal{L}_X \omega = \frac{d}{dt} (e^{tX})^* \omega \Big|_{t=0} \quad \text{for } \omega \in \bigwedge \mathcal{C}^\infty(M). \tag{A-2}$$

The Lie derivative satisfies

$$\mathcal{L}_X(\omega_1 \wedge \omega_2) = (\mathcal{L}_X \omega_1) \wedge \omega_2 + \omega_1 \wedge (\mathcal{L}_X \omega_2),$$

and Cartan's magic formula says

$$\mathcal{L}_X = \boldsymbol{i}_X \circ d + d \circ \boldsymbol{i}_X.$$

Differential forms $d\omega$ with degree $p+1$ can be integrated along a $(p+1)$-chain, or more specifically a $(p+1)$-dimensional submanifold with boundary; let us write it as C with boundary ∂C. Stokes' formula is written

$$\int_C d\omega = \int_{\partial C} \omega,$$

and it is the ground for de Rham's cohomology.

The Riemannian structure adds the pointwise dependent scalar product $g(x)$ given by

$$\langle S, T \rangle_{T_x M} = \sum_{1 \le i, j \le d} g_{i,j}(x) S^i T^j$$

with a dual metric $(g^{i,j}(x))_{1\le i,j\le d} := g(x)^{-1}$ defined on T_x^*M. This is also written with Einstein's conventions as

$$g = g_{i,j}\, dx^i\, dx^j, \quad g_{i,j} g^{j,k} = \delta_i^k.$$

Both $g(x)$ and $g(x)^{-1}$ are extended by tensor product to $\bigwedge T_x M$ and $\bigwedge T_x^* M$: for $\omega, \omega' \in \bigwedge^p \mathcal{C}^\infty(M)$,

$$\langle \omega', \omega\rangle_{\bigwedge^p T_x^* M} = \sum_{\#I=p}\sum_{\#J=p}\left(\prod_{k=1}^{p} g^{i_k, j_k}\right)\omega'_I \omega_J,$$

where $I = \{i_1, \dots, i_p\}$ $(i_1 < \dots < i_p)$ and $J = \{j_1, \dots, j_p\}$ $(j_1 < \dots < j_p)$. The Riemannian infinitesimal volume (denoted simply by dx in the text) is in an oriented local coordinate system:

$$d\,\mathrm{Vol}_g(x) = (\det g)^{1/2}\, dx^1 \wedge \dots \wedge dx^d = (\det g)^{1/2}\, dx^1 \cdots dx^d.$$

Those scalar products as nondegenerate bilinear forms allow identifications between forms and vectors. Here are examples: when $\omega = \omega_i(x) dx^i$ is a one-form, the vector $\omega^\#$ is given by $(\omega^\#)^i = g^{i,j}\omega_j$; when $X = X^i \partial_{x^i}$ is a vector field, $X^\flat$ is the one-form defined by $(X^\flat)_i = g_{i,j} X^j$. As an application, the gradient for a function is nothing but $\nabla f = (df)^\#$. Similarly, the Hessian of a function f at a point x, initially defined as a bilinear form, can be viewed a linear map of $T_x M$.

Another duality between forms of complementary degrees $p + p' = d = \dim M$ is provided by the Hodge $\star$ operator. When the Riemannian manifold (M, g) is orientable (locally this is always the case), the operator $\star : \bigwedge^p \mathcal{C}^\infty(M) \to \bigwedge^{d-p} \mathcal{C}^\infty(M)$ is defined by

$$\int \langle \omega', \omega\rangle_{\bigwedge^p T_x^* M}\, d\,\mathrm{Vol}_g(x) = \int \omega' \wedge (\star\omega), \quad \omega, \omega' \in \bigwedge^p \mathcal{C}^\infty(M).$$

In a coordinate system it is given by

$$(\star\omega)_J = \sum_I \delta_{I\cup J}^{\{1,\dots,d\}} \epsilon_{\{1,\dots,d\}}(I, J)(\det g)^{1/2}(\omega^\#)^I, \quad \begin{cases} I = \{i_1, \dots, i_p\},\ i_1 < \dots < i_p, \\ J = \{j_1, \dots, j_{d-p}\},\ j_1 < \dots < j_{d-p}, \\ (I, J) = (i_1, \dots, i_p,\ j_1, \dots, j_{d-p}), \end{cases}$$

where $\delta_A^B = 1$ when $A = B$ and $\delta_A^B = 0$ otherwise. We have the additional properties, for $\omega, \omega' \in \bigwedge^p \mathcal{C}^\infty(M)$,

$$\begin{aligned} \star(\lambda\omega + \omega') &= \lambda \star \omega + \star\omega', \quad \lambda \in \mathcal{C}^\infty(M), \\ \star\star\,\omega &= (-1)^{p(d+1)}\omega, \\ \omega \wedge (\star\omega') &= \omega' \wedge (\star\omega), \\ \star 1 &= d\,\mathrm{Vol}_g(x) \quad \text{(assuming } M \text{ is oriented).} \end{aligned}$$

The codifferential d^* is defined as the formal adjoint of the differential $d : \bigwedge \mathcal{C}^\infty(M) \to \bigwedge \mathcal{C}^\infty(M)$,

$$\langle d\omega, \omega'\rangle = \langle \omega, d^*\omega'\rangle.$$

With the Hodge $\star$ operator (do the identification on a compact oriented manifold without boundary with $\int_M d\eta = 0$),

$$\begin{cases} \star d^*\omega = (-1)^p d \star \omega, \\ \star d\omega = (-1)^{p+1} d^* \star \omega, \\ d^*\omega = (-1)^{pd+d+1} \star d \star \omega \end{cases} \qquad \text{for all } \omega \in \bigwedge^p \mathcal{C}^\infty(M).$$

The Hodge Laplacian is then given by

$$\Delta_{\mathrm{H}} = (d + d^*)^2 = dd^* + d^*d. \tag{A-3}$$

It is possible to write d^* and Δ_{H} in a coordinate system. For example,

$$\begin{aligned} (d^*\omega)_I &= -g^{i,j} \delta^J_{i\cup I} \epsilon_J(i, I) \nabla_j \omega_J, \quad (i, I) = (i, i_1, \dots, i_{p-1}), \\ \nabla_j \omega_J &= \partial_{x^j} \omega_J - \sum_{\ell=1}^{p} \omega_{I\cup\{k\}\setminus i_\ell} \epsilon_{I\cup\{k\}\setminus i_\ell}(i_1, \dots, i_{\ell-1}, k, i_{\ell+1}, \dots i_p) \Gamma^k_{i_\ell\, j}, \\ \Gamma^k_{i_\ell, j} &= \tfrac{1}{2} g^{k,m} (\partial_{x^{i_\ell}} g_{j,m} + \partial_{x^j} g_{m,i_\ell} - \partial_{x^m} g_{i_\ell, j}), \end{aligned}$$

where one recognizes the covariant derivative ∇_j associated with the metric g (the Levi–Civita connection) and the Christoffel symbols $\Gamma^j_{k\,\ell}$. The writing of Δ_{H} involves the Riemann curvature tensor and is known as Weitzenbock's formula. We wrote the above example to convince the unfamiliar reader that the explicit writing in a coordinate system is not always more informative than the intrinsic formula.

Here is the example of the Witten Laplacian, $\Delta_{f,h} = (d_{f,h} + d^*_{f,h})^2 = d^*_{f,h} d_{f,h} + d_{f,h} d^*_{f,h}$:

$$d_{f,h} = e^{-f/h}(hd)e^{f/h} = hd + df\wedge, \tag{A-4}$$

$$d^*_{f,h} = e^{f/h}(hd^*)e^{-f/h} = hd^* + \boldsymbol{i}_{\nabla f}, \tag{A-5}$$

$$\begin{aligned} \Delta_{f,h} &= d_{f,h} d^*_{f,h} + d^*_{f,h} d_{f,h} = (hd + df\wedge)(hd^* + \boldsymbol{i}_{\nabla f}) + (hd^* + \boldsymbol{i}_{\nabla f})(hd + df\wedge) \\ &= h^2(dd^* + d^*d) + [(df\wedge)\circ \boldsymbol{i}_{\nabla f} + \boldsymbol{i}_{\nabla f} \circ (df\wedge)] + h[d\boldsymbol{i}_{\nabla f} + \boldsymbol{i}_{\nabla f} d] + h[(df\wedge)\circ d^* + d^* \circ (df\wedge)] \\ &= h^2 \Delta_{\mathrm{H}} + |\nabla f|^2 + h(\mathcal{L}_{\nabla f} + \mathcal{L}^*_{\nabla f}), \end{aligned} \tag{A-6}$$

where we used $\boldsymbol{i}_X(df \wedge \omega) = df(X)\omega - df \wedge (\boldsymbol{i}_X \omega)$ with $X = \nabla f$, Cartan's magic formula and an easy identification of $\mathcal{L}^*_{\nabla f}$. No explicit computation of d^* or the Hodge Laplacian is necessary to understand the structure of the Witten Laplacian. In particular, $\mathcal{L}_X + \mathcal{L}^*_X$ is clearly a zeroth-order differential operator because in a coordinate system the formal adjoint of $a^j(x)\partial_{x^j}$ in $L^2(\mathbb{R}^d, \varrho(x)\, dx)$ equals $-a^j(x)\partial_{x^j} + b[a, \varrho](x)$, where $b[a, \varrho]$ is the multiplication by a function of x. The operator $\mathcal{L}_{\nabla f} + \mathcal{L}^*_{\nabla f}$ is not the local action of a tensor field on M because it does not follow the change of coordinates rule for tensors. Actually, one can give a meaning to the general expression

$$\Delta^{(p)}_{f,h} = h^2 \Delta^{(p)}_{\mathrm{H}} + |\nabla f|^2 - h(\Delta f) + 2h(\mathrm{Hess}\, f)_p,$$

where $(\mathrm{Hess}\, f)_p$ is an element of the curvature tensor algebra (see [Jammes 2012] and references therein).

Let us conclude this appendix with integration by parts formulas in the case of a manifold with a boundary. All these formulas rely first on Stokes' formula $\int_\Omega d\omega = \int_{\partial\Omega} \omega$ when $\omega \in \bigwedge^{d-1} \mathcal{C}^\infty(\overline{\Omega})$.

Note that the right-hand side of Stokes' formula may equivalently (and more explicitly) be written $\int_{\partial\Omega}\omega = \int_{\partial\Omega} j^*\omega$, where $j : \partial\Omega \to \overline{\Omega}$ is the natural embedding map (a trace along $\partial\Omega$ is taken and, pointwise, $j^*\omega_x$ is evaluated only on $(d-1)$-vectors tangent to $\partial\Omega$). Another expression taken initially from [Schwarz 1995] is also convenient. For $\sigma \in \partial\Omega$ let $n(\sigma)$ be the outward normal vector and write, for any element $X \in T_\sigma M$, $X = X_T + X_n n$.

For $\omega \in \bigwedge^p \mathcal{C}^\infty(\overline{\Omega})$, define $\boldsymbol{t}\omega$ and $\boldsymbol{n}\omega = \omega - \boldsymbol{t}\omega$ by

$$\forall \sigma \in \partial\Omega \ \ \forall X_1, \dots, X_p \in T_\sigma\Omega \quad \boldsymbol{t}\omega(X_1, \dots, X_p) = \omega(X_{1,T}, \dots, X_{p,T}).$$

If $(x^1, \dots, x^d) = (x', x^d)$ is a coordinate system in a neighborhood $\mathcal{V}$ of $\sigma_0 \in \partial\Omega$ such that $\Omega \cap \mathcal{V}$ is given locally by $\{x^d < 0\}$, $\partial\Omega \cap \mathcal{V}$ by $\{x^d = 0\}$ and $n = \partial_{x^d}$, then a p-form can be written

$$\omega = \sum_{\substack{\#I=p\\ d\notin I}} \omega_I dx^I + \sum_{\substack{\#I'=p-1\\ d\notin I'}} \omega_{I'} dx^{I'} \wedge dx^d,$$

and the operators $\boldsymbol{t}$ and $\boldsymbol{n}$ act as

$$\boldsymbol{t}\omega = \sum_{\substack{\#I=p\\ d\notin I}} \omega_I dx^I, \quad \boldsymbol{n}\omega = \sum_{\substack{\#I'=p-1\\ d\notin I'}} \omega_{I'} dx^{I'} \wedge dx^d.$$

Stokes' formula can be written now as $\int_\Omega d\omega = \int_{\partial\Omega} \boldsymbol{t}\omega$ for $\omega \in \bigwedge^{d-1} \mathcal{C}^\infty(\overline{\Omega})$, but contrary to the operator j^* the operator $\boldsymbol{t}$ makes sense in a collar neighborhood of $\partial\Omega$; locally $\boldsymbol{t}\omega_{(x',x^d)} = \boldsymbol{t}\omega_{(x',0)}$ by definition. In particular, the formula

$$\boldsymbol{t}\, d\omega = d\boldsymbol{t}\omega$$

makes sense for any $\omega \in \bigwedge \mathcal{C}^\infty(\overline{\Omega})$ and it is rather easy to check with the above coordinates description. One also gets, in the same way,

$$\boldsymbol{t}\omega = \boldsymbol{i}_n(n^\flat \wedge \omega) \quad \text{for } \omega \in \bigwedge \mathcal{C}^\infty(\overline{\Omega}), \tag{A-7}$$

$$\star\boldsymbol{n} = \boldsymbol{t}\star, \quad \star\boldsymbol{t} = \boldsymbol{n}\star, \tag{A-8}$$

$$\boldsymbol{t}d = d\boldsymbol{t}, \quad \boldsymbol{n}d^* = d^*\boldsymbol{n}, \tag{A-9}$$

$$\boldsymbol{t}\omega_1 \wedge \star\boldsymbol{n}\omega_2 = \langle \omega_1, \boldsymbol{i}_n\omega_2\rangle_{\bigwedge^p T^*_\sigma\Omega} \times d\,\mathrm{Vol}_{g,\partial\Omega} \quad \text{for } \omega_i \in \bigwedge^p \mathcal{C}^\infty(\overline{\Omega}), \tag{A-10}$$

where we recall that $d\,\mathrm{Vol}_{g,\partial\Omega}(X_1, \dots, X_{d-1}) = d\,\mathrm{Vol}_g(n, X_1, \dots, X_{d-1})$.

The above formulas, for example lead to the following integration by parts for $\omega_1, \omega_2 \in \bigwedge^p \mathcal{C}^\infty(\overline{\Omega})$:

$$\langle d_{f,h}\omega_1, d_{f,h}\omega_2\rangle_{L^2(\Omega)} + \langle d^*_{f,h}\omega_1, d^*_{f,h}\omega_2\rangle_{L^2(\Omega)} \\ = \langle \omega_1, \Delta_{f,h}\omega_2\rangle_{L^2(\Omega)} + h\int_{\partial\Omega} (\boldsymbol{t}\omega_2) \wedge \star\boldsymbol{n}d_{f,h}\omega_1 - h\int_{\partial\Omega} (\boldsymbol{t}d^*_{f,h}\omega_1) \wedge (\star\boldsymbol{n}\omega_2).$$

This shows, for example, that $\Delta^D_{f,h}$ (resp. $\Delta^N_{f,h}$) with its form domain $W^{1,2}_D = \{\omega \in \bigwedge W^{1,2} : \boldsymbol{t}\omega = 0\}$ (resp. $W^{1,2}_N = \{\omega \in \bigwedge W^{1,2} : \boldsymbol{n}\omega = 0\}$) is associated with the Dirichlet form $\|d_{f,h}\omega\|^2 + \|d^*_{f,h}\omega\|^2$. Interpreting the weak formulation of $\Delta_{f,h}\omega = f$ leads to the operator domains $D(\Delta^D_{f,h})$ and $D(\Delta^N_{f,h})$ (we refer the reader to [Helffer and Nier 2006] for details). The boundary terms of Lemma 3.1 are obtained in a very similar way.

Acknowledgements

T. Lelièvre would like to thank C. Le Bris, D. Perez and A. Voter for useful discussions. His work is supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement number 614492. This work was essentially completed while F. Nier had a "Délégation INRIA" in CERMICS at Ecole des Ponts. He acknowledges the support of INRIA and thanks the people of CERMICS for their hospitality.

References

[Abraham and Marsden 1978] R. Abraham and J. E. Marsden, *Foundations of mechanics*, 2nd ed., Benjamin/Cummings Publishing Co., Advanced Book Program, Reading, MA, 1978. MR 81e:58025 Zbl 0393.70001

[Bismut 1986] J.-M. Bismut, "The Witten complex and the degenerate Morse inequalities", *J. Differential Geom.* **23**:3 (1986), 207–240. MR 87m:58169 Zbl 0608.58038

[Bismut and Zhang 1994] J.-M. Bismut and W. Zhang, "Milnor and Ray–Singer metrics on the equivariant determinant of a flat vector bundle", *Geom. Funct. Anal.* **4**:2 (1994), 136–212. MR 96f:58179 Zbl 0830.58030

[Bovier et al. 2004] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, "Metastability in reversible diffusion processes, I: Sharp asymptotics for capacities and exit times", *J. Eur. Math. Soc.* (*JEMS*) **6**:4 (2004), 399–424. MR 2006b:82112

[Bovier et al. 2005] A. Bovier, V. Gayrard, and M. Klein, "Metastability in reversible diffusion processes, II: Precise asymptotics for small eigenvalues", *J. Eur. Math. Soc.* (*JEMS*) **7**:1 (2005), 69–99. MR 2006b:82113

[Burghelea 1997] D. Burghelea, "Lectures on Witten–Helffer–Sjöstrand theory", pp. 85–99 in *Proceedings of the Third International Workshop on Differential Geometry and its Applications and the First German–Romanian Seminar on Geometry* (Sibiu, 1997), vol. 5, edited by S. Balea et al., 1997. MR 2000h:58040 Zbl 0936.58008

[Chang and Liu 1995] K. C. Chang and J. Liu, "A cohomology complex for manifolds with boundary", *Topol. Methods Nonlinear Anal.* **5**:2 (1995), 325–340. MR 96k:58206 Zbl 0848.58001

[Chazarain and Piriou 1982] J. Chazarain and A. Piriou, *Introduction to the theory of linear partial differential equations*, Studies in Mathematics and its Applications **14**, North-Holland Publishing Co., Amsterdam–New York, 1982. MR 83j:35001 Zbl 0487.35002

[Cycon et al. 1987] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger operators with application to quantum mechanics and global geometry*, Springer, Berlin, 1987. MR 88g:35003 Zbl 0619.47005

[Dimassi and Sjöstrand 1999] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit*, London Mathematical Society Lecture Note Series **268**, Cambridge University Press, 1999. MR 2001b:35237 Zbl 0926.35002

[Gallot et al. 2004] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian geometry*, 3rd ed., Springer, Berlin, 2004. MR 2005e:53001 Zbl 1068.53001

[Goldberg 1970] S. I. Goldberg, *Curvature and homology*, 3rd ed., Academic Press, New York–London, 1970. MR 25 #2537 Zbl 0962.53001

[Hatcher 2002] A. Hatcher, *Algebraic topology*, Cambridge University Press, 2002. MR 2002k:55001 Zbl 1044.55001

[Helffer 1988] B. Helffer, *Semi-classical analysis for the Schrödinger operator and applications*, Lecture Notes in Mathematics **1336**, Springer, Berlin, 1988. MR 90c:81043 Zbl 0647.35002

[Helffer and Nier 2006] B. Helffer and F. Nier, *Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach: the case with boundary*, Mém. Soc. Math. Fr. (N.S.) **105**, Société Mathématique de France, 2006. MR 2007k:58044

[Helffer and Sjöstrand 1985a] B. Helffer and J. Sjöstrand, "Puits multiples en limite semi-classique, II: Interaction moléculaire, symétries, perturbation", *Ann. Inst. H. Poincaré Phys. Théor.* **42**:2 (1985), 127–212. MR 87a:35142

[Helffer and Sjöstrand 1985b] B. Helffer and J. Sjöstrand, "Puits multiples en mécanique semi-classique, IV: Étude du complexe de Witten", *Comm. Partial Differential Equations* **10**:3 (1985), 245–340. MR 87i:35162

[Helffer and Sjöstrand 1987] B. Helffer and J. Sjöstrand, "Puits multiples en mécanique semi-classique, VI: Cas des puits sous-variétés", *Ann. Inst. H. Poincaré Phys. Théor.* **46**:4 (1987), 353–372. MR 91g:58285

[Helffer and Sjöstrand 1988] B. Helffer and J. Sjöstrand, "A proof of the Bott inequalities", pp. 171–183 in *Algebraic analysis, Vol. I*, edited by M. Kashiwara and T. Kawai, Academic Press, Boston, 1988. MR 90g:58018 Zbl 0693.58019

[Helffer et al. 2004] B. Helffer, M. Klein, and F. Nier, "Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach", *Mat. Contemp.* **26** (2004), 41–85. MR 2005i:58025 Zbl 1079.58025

[Jammes 2012] P. Jammes, "Sur la multiplicité des valeurs propres du laplacien de Witten", *Trans. Amer. Math. Soc.* **364**:6 (2012), 2825–2845. MR 2888230 Zbl 1242.58015

[Laudenbach 2011] F. Laudenbach, "A Morse complex on manifolds with boundary", *Geom. Dedicata* **153** (2011), 47–57. MR 2819662 Zbl 1223.57020

[Le Bris et al. 2012] C. Le Bris, T. Lelièvre, M. Luskin, and D. Perez, "A mathematical formalization of the parallel replica dynamics", *Monte Carlo Methods Appl.* **18**:2 (2012), 119–146. MR 2926765 Zbl 1243.82045

[Le Peutrec 2009] D. Le Peutrec, "Small singular values of an extracted matrix of a Witten complex", *Cubo* **11**:4 (2009), 49–57. MR 2010j:58070 Zbl 1181.81050

[Le Peutrec 2010a] D. Le Peutrec, "Local WKB construction for Witten Laplacians on manifolds with boundary", *Anal. PDE* **3**:3 (2010), 227–260. MR 2011j:58035 Zbl 1225.58012

[Le Peutrec 2010b] D. Le Peutrec, "Small eigenvalues of the Neumann realization of the semiclassical Witten Laplacian", *Ann. Fac. Sci. Toulouse Math.* (6) **19**:3-4 (2010), 735–809. MR 2012c:58042 Zbl 1213.58023

[Le Peutrec 2011] D. Le Peutrec, "Small eigenvalues of the Witten Laplacian acting on p-forms on a surface", *Asymptot. Anal.* **73**:4 (2011), 187–201. MR 2012i:58025 Zbl 1232.58024

[Le Peutrec et al. 2013] D. Le Peutrec, F. Nier, and C. Viterbo, "Precise Arrhenius law for p-forms: the Witten Laplacian and Morse–Barannikov complex", *Ann. Henri Poincaré* **14**:3 (2013), 567–610. MR 3035640 Zbl 1275.58018

[Milnor 1963] J. Milnor, *Morse theory*, Annals of Mathematics Studies **51**, Princeton University Press, Princeton, N.J., 1963. MR 29 #634 Zbl 0108.10401

[Schwarz 1995] G. Schwarz, *Hodge decomposition—a method for solving boundary value problems*, Lecture Notes in Mathematics **1607**, Springer, Berlin, 1995. MR 96k:58222 Zbl 0828.58002

[Simon 1979] B. Simon, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series **35**, Cambridge University Press, 1979. MR 80k:47048 Zbl 0423.47001

[Sternberg 1964] S. Sternberg, *Lectures on differential geometry*, Prentice-Hall, Englewood Cliffs, N.J., 1964. MR 33 #1797 Zbl 0129.13102

[Taylor 1997] M. E. Taylor, *Partial differential equations, I: Basic theory*, Applied Mathematical Sciences **115**, Springer, New York, 1997. MR 98b:35002b Zbl 1206.35002

[Voter 1997] A. F. Voter, "A method for accelerating the molecular dynamics simulation of infrequent events", *J. Chem. Phys.* **106**:11 (1997), 4665–4677.

[Witten 1982] E. Witten, "Supersymmetry and Morse theory", *J. Differential Geom.* **17**:4 (1982), 661–692. MR 84b:58111 Zbl 0499.53056

[Zhang 2001] W. Zhang, *Lectures on Chern–Weil theory and Witten deformations*, Nankai Tracts in Mathematics **4**, World Scientific Publishing Co., River Edge, NJ, 2001. MR 2002m:58032 Zbl 0993.58014

Received 18 Sep 2013. Revised 22 Dec 2014. Accepted 18 Feb 2015.

TONY LELIÈVRE: lelievre@cermics.enpc.fr
CERMICS, Université Paris-Est, Joint Project-team INRIA Matherials, 6 & 8 Avenue Blaise Pascal, F-77455 Marne-la-Vallée, France

FRANCIS NIER: francis.nier@math.univ-paris13.fr
LAGA, Université de Paris 13, UMR-CNRS 7539, Avenue Jean-Baptiste Clément, F-93430 Villetaneuse, France

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 8, No. 3, 2015
dx.doi.org/10.2140/apde.2015.8.629

DYNAMICS OF COMPLEX-VALUED MODIFIED KDV SOLITONS WITH APPLICATIONS TO THE STABILITY OF BREATHERS

MIGUEL A. ALEJO AND CLAUDIO MUÑOZ

We study the long-time dynamics of complex-valued modified Korteweg–de Vries (mKdV) solitons, which are distinguished because they blow up in finite time. We establish stability properties at the H^1 level of regularity, uniformly away from each blow-up point. These new properties are used to prove that mKdV breathers are H^1-stable, improving our previous result [*Comm. Math. Phys.* **324**:1 (2013) 233–262], where we only proved H^2-stability. The main new ingredient of the proof is the use of a Bäcklund transformation which relates the behavior of breathers, complex-valued solitons and small real-valued solutions of the mKdV equation. We also prove that negative energy breathers are asymptotically stable. Since we do not use any method relying on the inverse scattering transform, our proof works even under $L^2(\mathbb{R})$ perturbations, provided a corresponding local well-posedness theory is available.

1. Introduction 629
2. Complex-valued mKdV soliton profiles 635
3. Bäcklund transformation for mKdV 638
4. Dynamics of complex-valued mKdV solitons 646
5. Complex solitons versus breathers 652
6. Double Bäcklund transformation for mKdV 657
7. Stability of breathers 665
8. Asymptotic stability 669
Appendix A: Proof of Lemma 5.1 671
Acknowledgments 672
References 672

1. Introduction

Consider the modified Korteweg–de Vries (mKdV) equation on the real line

$$u_t + (u_{xx} + u^3)_x = 0, \tag{1-1}$$

where $u = u(t, x)$ is a complex-valued function and $(t, x) \in \mathbb{R}^2$. Note that (1-1) is not $U(1)$-invariant. In the case of real-valued initial data, the associated Cauchy problem for (1-1) is globally well posed for initial data in $H^s(\mathbb{R})$ for any $s > \frac{1}{4}$; see Kenig, Ponce and Vega [Kenig et al. 1993], and Colliander, Keel, Staffilani, Takaoka and Tao [Colliander et al. 2003]. Additionally, the (real-valued) flow map is not

MSC2010: primary 35Q51, 35Q53; secondary 37K10, 37K40.
Keywords: mKdV equation, Bäcklund transformation, solitons, breather, stability.

uniformly continuous if $s < \frac{1}{4}$ [Kenig et al. 2001].[1] In order to prove this last result, Kenig, Ponce and Vega considered a very particular class of solutions of (1-1) called *breathers*, discovered by Wadati [1973].

Definition 1.1 (see, e.g., [Wadati 1973; Lamb 1980]). Let $\alpha, \beta > 0$ and $x_1, x_2 \in \mathbb{R}$ be fixed parameters. The mKdV breather is a smooth solution of (1-1) given explicitly by the formula

$$\begin{aligned} B = B(t, x; \alpha, \beta, x_1, x_2) &:= 2\sqrt{2}\partial_x \left[\arctan \frac{\beta \sin(\alpha y_1)}{\alpha \cosh(\beta y_2)} \right], \\ &= \frac{2\sqrt{2}\alpha\beta(\alpha \cos(\alpha y_1)\cosh(\beta y_2) - \beta \sin(\alpha y_1)\sinh(\beta y_2))}{\alpha^2 \cosh^2(\beta y_2) + \beta^2 \sin^2(\alpha y_1)}, \end{aligned} \tag{1-2}$$

where

$$y_1 := x + \delta t + x_1, \quad y_2 := x + \gamma t + x_2, \tag{1-3}$$

and

$$\delta := \alpha^2 - 3\beta^2, \quad \gamma := 3\alpha^2 - \beta^2. \tag{1-4}$$

Breathers are *oscillatory bound states*. They are periodic in time (after a suitable space shift) and localized in space. The parameters α and β are scaling parameters, x_1, x_2 are shifts, and $-\gamma$ represents the *velocity* of a breather. As we will see later, the main difference between *solitons*[2] and breathers is given at the level of the *oscillatory* scaling α, which is not present in the case of solitons. For a detailed account of the physics of breathers, see, e.g., [Lamb 1980; Ablowitz and Clarkson 1991; Aubry 1997; Alejo 2012; Alejo and Muñoz 2013] and references therein.

Numerical computations (see Gorria, Alejo and Vega [Gorria et al. 2013]) showed that breathers are *numerically* stable. Next, in [Alejo and Muñoz 2013] we constructed a Lyapunov functional that controls the dynamics of H^2 perturbations of (1-2). The purpose of this paper is to improve this previous result and show that mKdV breathers are indeed H^1-stable, i.e., stable in the energy space.

Theorem 1.2. *Let $\alpha, \beta > 0$ be fixed scalings. There exist parameters η_0, A_0, depending on α and β only, such that the following holds: Consider $u_0 \in H^1(\mathbb{R})$, and assume that there exists $\eta \in (0, \eta_0)$ such that*

$$\|u_0 - B(0, \cdot\,; \alpha, \beta, 0, 0)\|_{H^1(\mathbb{R})} \le \eta. \tag{1-5}$$

Then there exist functions $x_1(t), x_2(t) \in \mathbb{R}$ such that the solution $u(t)$ of the Cauchy problem for the mKdV equation (1-1) *with initial data u_0 satisfies*

$$\sup_{t\in\mathbb{R}} \|u(t) - B(t, \cdot\,; \alpha, \beta, x_1(t), x_2(t))\|_{H^1(\mathbb{R})} \le A_0\eta, \tag{1-6}$$

$$\sup_{t\in\mathbb{R}} |x_1'(t)| + |x_2'(t)| \le C A_0 \eta, \tag{1-7}$$

for some constant $C > 0$.

[1] However, one can construct a solution in L^2; see [Christ et al. 2012].

[2] See (1-8).

The initial condition (1-5) can be replaced by any initial breather profile of the form $B(t_0; \alpha, \beta, x_1^0, x_2^0)$ with $t_0, x_1^0, x_2^0 \in \mathbb{R}$, thanks to the invariance of the equation under translations in time and space.[3] Moreover, using the Miura transform [Miura et al. 1968], one can prove a natural stability property in $L^2(\mathbb{R}; \mathbb{C})$ for an associated *complex-valued* KdV breather.

One can also use the scaling invariance of the equation, $u(t,x) \mapsto \lambda u(\lambda^3 t, \lambda x)$, to reduce the problem to the case where α equals 1 and $\beta > 0$ is arbitrary, but for symmetry reasons we shall not follow this approach.[4]

Additionally, from the proof, the shifts $x_1(t)$ and $x_2(t)$ in Theorem 1.2 can be described almost explicitly[5], which is a substantial improvement with respect to [Alejo and Muñoz 2013], where no exact control on the shift parameters was given. We obtain such a control with no additional decay assumptions on the initial data other than being in $H^1(\mathbb{R})$.

Theorem 1.2 places breathers as stable objects at the same level of regularity as mKdV solitons, even if they are very different in nature. To be more precise, a (real-valued) soliton is a solution of (1-1) of the form

$$u(t,x) = Q_c(x - ct), \quad Q_c(s) := \sqrt{c}\, Q(\sqrt{c}\, s), \quad c > 0, \tag{1-8}$$

with

$$Q(s) := \frac{\sqrt{2}}{\cosh(s)} = 2\sqrt{2}\,\partial_s[\arctan(e^s)],$$

and where $Q_c > 0$ satisfies the nonlinear ODE

$$Q_c'' - cQ_c + Q_c^3 = 0, \quad Q_c \in H^1(\mathbb{R}). \tag{1-9}$$

We recall that solitons are H^1-stable (Benjamin [1972], Bona, Souganidis and Strauss [Bona et al. 1987]). See also the works by Grillakis, Shatah and Strauss [Grillakis et al. 1987] and Weinstein [1986] for the nonlinear Schrödinger case.

Even more surprising is the fact that Theorem 1.2 will arise as a consequence of a suitable stability property of the zero solution and of *complex-valued* mKdV solitons, which are singular solutions.

A complex-valued soliton is a solution of the form (1-8) of (1-1) with a complex-valued scaling and velocity, i.e.,

$$u(t,x) := Q_c(x - ct), \quad \sqrt{c} := \beta + i\alpha, \quad \alpha,\ \beta > 0; \tag{1-10}$$

see Definition 2.1 for a precise interpretation. In Lemma 2.2 we give a detailed description of the singular nature of (1-10). On the other hand, very little is known about mKdV (1-1) when the initial data is complex-valued. For instance, it is known that it has finite-time blow-up solutions, the most important

[3]Indeed, if $u(t,x)$ solves (1-1), then, for any $t_0, x_0 \in \mathbb{R}$ and $c > 0$, $u(t - t_0, x - x_0)$, $c^{1/2}u(c^{3/2}t, c^{1/2}x)$, $u(-t,-x)$ and $-u(t,x)$ are solutions of (1-1).

[4]For example, if (1-6) holds, then $v_0(y) := u_0(y/\alpha)/\alpha$ satisfies

$$\alpha \int_{\mathbb{R}} \Big(v_0 - B\Big(0, \cdot\,; 1, \frac{\beta}{\alpha}, 0, 0\Big)\Big)^2 = \int_{\mathbb{R}} (u_0 - B(0, \cdot\,; \alpha, \beta, 0, 0))^2 \le \eta^2.$$

[5]See (7-9).

examples being the complex solitons themselves; see, e.g., Bona, Vento and Weissler [Bona et al. 2013] and references therein for more details. According to [Bona et al. 2013], blow-up in the complex-valued case can be understood as the intersection with the real line $x \in \mathbb{R}$ of a curve of poles of the solution after being extended to the complex plane (i.e., now x is replaced by $z \in \mathbb{C}$). Blow-up in this case seems to have better properties than the corresponding critical blow-up described by Martel and Merle [2002].

Let $H^1(\mathbb{R}; \mathbb{C})$ denote the standard Sobolev space of complex-valued functions $f(x) \in \mathbb{C}$, $x \in \mathbb{R}$. In this paper we prove the following *stability property* for solitons, far away from each blow-up time:

Theorem 1.3. *There exists an open set of initial data in $H^1(\mathbb{R}; \mathbb{C})$ for which the mKdV complex solitons are well-defined and stable in $H^1(\mathbb{R}; \mathbb{C})$ for all times uniformly separated from a countable sequence of finite blow-up times with no limit points. Moreover, one can define a mass and an energy, both invariant* for all time.

We cannot prove an all-time stability result using the $H^1(\mathbb{R}; \mathbb{C})$-norm because even complex solitons leave that space at each blow-up time, and several computations in this paper break down. However, the previous result states that the Cauchy problem is almost globally well-posed around a soliton, and the solution can be continued after (or before) every blow-up time. The novelty with respect to the local Cauchy theory [Kenig et al. 1993] is that now it is possible to define an almost global solution instead of defining a local solution on each subinterval of time defined by two blow-up points, because from the proof we will recognize that the behavior before and after the blow-up time are deeply linked. From this property, the existence and invariance of uniquely well-defined mass and energy will be quite natural. For this particular problem, we answer positively the questions about existence, uniqueness and regularity after blow-up posed by Merle [1992]. See Theorem 4.5 and its corollaries for a more detailed statement.

Lastly, we prove that breathers behaving as standard solitons are asymptotically stable in the energy space. For previous results for the soliton and multisoliton case, see Pego and Weinstein [1994] and Martel and Merle [2005].

Theorem 1.4. *Under the hypotheses of Theorem 1.2, there exists $c_0 > 0$ depending on η, with $c_0(\eta) \to 0$ as $\eta \to 0$, such that the following holds: There exist β^* and α^* (depending on η) close enough to β and α, respectively, for which*

$$\lim_{t \to +\infty} \|u(t) - B(t; \cdot, \alpha^*, \beta^*, x_1(t), x_2(t))\|_{H^1(x \geq c_0 t)} = 0. \tag{1-11}$$

In particular, the asymptotic of $u(t)$ has new and explicit velocity parameters $\delta^* = (\alpha^*)^2 - 3(\beta^*)^2$ and $\gamma^* = 3(\alpha^*)^2 - (\beta^*)^2$ at the leading order.

The previous result is more interesting when $\gamma < 0$; see (1-4). In this case, the breather has negative energy (see [Alejo and Muñoz 2013, p. 9]) and it moves rightwards in space (the so-called physically relevant region). We recall that working in the energy space implies that small solitons moving to the right in a very slow fashion are allowed (the condition $c_0 > 0$ is essential; see, e.g., [Martel and Merle 2005]). Indeed, there are explicit solutions of (1-1) composed of one breather and one very small soliton moving rightwards, which contradicts any sort of *global* asymptotic stability result in the energy space [Lamb 1980]. Additionally, we cannot ensure that the left portion of the real line $\{x < 0\}$ corresponds to radiation only. Following [Lamb 1980], it is possible to construct a solution to (1-1) composed of two breathers, one very small with respect to the other one, the latter with positive velocity and the former with small but still negative velocity (just take the corresponding scaling parameters α and β both small so that $-\gamma < 0$). Such a solution has no radiation at infinity. Of course, working in a neighborhood of the breather using weighted spaces rules out such small perturbations.

The mechanism under which α^* and β^* are chosen is very natural and reflects the power and simplicity of the arguments of the proof: under different scaling parameters, it was impossible to describe the dynamics as in Theorem 1.2. We do indeed have two linked results: in some sense Theorem 1.2 is a consequence of Theorem 1.4 and vice versa.

It is also important to emphasize that (1-1) is a well-known *completely integrable* model [Miura et al. 1968; Ablowitz and Clarkson 1991; Lamb 1980; Lax 1968; Schuur 1986], with infinitely many conserved quantities and a suitable Lax pair formulation. The inverse scattering theory has been applied in [Schuur 1986] to describe the evolution of *rapidly decaying* initial data, by purely algebraic methods. Solutions are shown to decompose into a very particular set of solutions: solitons, breathers and radiation. Moreover, as a consequence of the integrability property, these nonlinear modes interact elastically during the dynamics, and no dispersive effects are present at infinity. In particular, even more complex solutions are present, such as *multisolitons* (explicit solutions describing the interaction of several solitons [Hirota 1972]). Multisolitons for mKdV and several integrable models of Korteweg–de Vries-type are stable in H^1; see Maddocks and Sachs [1993] for the KdV case and in a more general setting see Martel, Merle and Tsai [Martel et al. 2002].

However, the proof of Theorem 1.2 does not involve any method relying on the inverse scattering transform [Miura et al. 1968; Schuur 1986], nor the steepest descent machinery [Deift and Zhou 1993],[6] which allows us to work in the *very large* energy space $H^1(\mathbb{R})$. Note that if the inverse scattering methods are allowed, one could describe the dynamics of very general initial data with more detail. But if this is the case, additional decay and/or spectral assumptions are always needed, and, except with well-prepared initial data, such conditions are difficult to verify. We claim that our proof works even if the initial data is in $L^2(\mathbb{R})$ provided mKdV is locally well-posed at that level of regularity, which remains a very difficult open problem.

Comparing with [Alejo and Muñoz 2013], where we have proved that mKdV breathers are H^2-stable, now we are not allowed to use the *third* conservation law associated to mKdV,[7]

$$F[u](t) = \frac{1}{2}\int_{\mathbb{R}} u_{xx}^2(t,x)\,dx - \frac{5}{2}\int_{\mathbb{R}} u^2u_x^2(t,x)\,dx + \frac{1}{4}\int_{\mathbb{R}} u^6(t,x)\,dx,$$

nor the elliptic equation satisfied by *any* breather profile,

$$B_{(4x)} - 2(\beta^2-\alpha^2)(B_{xx}+B^3) + (\alpha^2+\beta^2)^2 B + 5BB_x^2 + 5B^2B_{xx} + \tfrac{3}{2}B^5 = 0$$

[6]Note that Deift and Zhou [1993] consider the *defocusing* mKdV equation, which has no smooth solitons or breathers.

[7]See (4-13) and (4-14) for the other two low-regularity conserved quantities.

since the dynamics is no longer in H^2. Moreover, since breathers are bound states, there is no associated decoupling in the dynamics as time evolves as in [Martel et al. 2002], which makes the proof of the H^1 case even more difficult. We need a different method of proof.

We follow a method of proof that is in the spirit of the seminal work by Merle and Vega [2003] (see also Alejo, Muñoz and Vega [Alejo et al. 2013]), where the L^2-stability of KdV solitons has been proved. In those cases, the use of the Miura and Gardner transformations were the new ingredients to prove stability where the standard energy is missing. Recently, the Miura transformation has been studied at very low regularity by Buckmaster and Koch [2014]; using this information, they showed that KdV solitons are orbitally stable under H^{-1} perturbations leading to a $H^n \cap H^{-3/4}$ solution, where $n \geq -1$ is an integer.

More precisely, we will make use of the Bäcklund transformation [Lamb 1980, p. 257] associated to mKdV to obtain new conserved quantities, additional to the mass and energy. Mizumachi and Pelinovsky [2012] and Hoffmann and Wayne [2013] described a similar approach for the NLS and sine-Gordon equations and their corresponding one-solitons. However, unlike those previous works, and in order to control any breather, we use the Bäcklund transformation twice: one to control an associated complex-valued mKdV soliton, and a second one to get almost complete control of the breather.

Indeed, solving the Bäcklund transformation in the vicinity of a breather leads (formally) to the emergence of complex-valued mKdV solitons, which blow up in finite time. A difficult problem arises at the level of the Cauchy theory, and any attempt to prove stability must face the ill-posedness behavior of the complex-valued mKdV equation (1-1). However, after a new use of the Bäcklund transformation around the complex soliton we end up with a small, real-valued $H^1(\mathbb{R})$ solution of mKdV which is stable for all time. The fact that a second application of the Bäcklund transformation leads to a real-valued solution is not trivial and is a consequence of a deep property called the *permutability theorem* [Lamb 1980]. Roughly speaking, that result states that the order under which we perform two inversions of the Bäcklund transformation does not matter. After some work we are able to give a rigorous proof of the following fact: we can invert a breather using Bäcklund towards two particularly well-chosen complex solitons first, and then invert once again to obtain two small solutions — say a and b — and the final result *must be the same*. Even better, one can show that a has to be the conjugate of b, which gives the real character of the solution. Now, the dynamics is real-valued and simple. We use the Kenig–Ponce–Vega theory [Kenig et al. 1993] to evolve the system to any given time. Using this trick we avoid dealing with the blow-up times of the complex soliton — for a while — and at the same time we prove a new stability result for them.

However, unlike [Mizumachi and Pelinovsky 2012; Hoffman and Wayne 2013], we cannot invert the Bäcklund transformation at any given time, and in fact each blow-up time of the complex-valued mKdV soliton is a dangerous obstacle for the breather stability. In order to extend the stability property up to the blow-up times we discard the method involving the Bäcklund transformation. Instead we run a bootstrap argument starting from a fixed time very close to each singular point, using the fact that the real-valued mKdV dynamics is continuous in time. Finally, using energy methods related to the stability of single solitons we are able to extend the uniform bounds in time to any singularity point, with a universal constant A_0 as in Theorem 1.2.

From the proof it will be evident that, even if there is no global well-posedness theory (with uniform bounds in time) below H^s, $s < \frac{1}{4}$, one can prove stability of breathers in spaces of the form $H^1 \cap H^s$, $s < \frac{1}{4}$, following the ideas of Buckmaster and Koch [2014]. We thank Professor Herbert Koch for mentioning to us this interesting property.

Our results apply without significant modifications to the case of the sine-Gordon (SG) equation in $\mathbb{R}_t \times \mathbb{R}_x$,

$$u_{tt} - u_{xx} + \sin u = 0, \quad (u, u_t)(t, x) \in \mathbb{R}^2, \tag{1-12}$$

and its corresponding breather [Lamb 1980, p. 149]. See [Birnir et al. 1994; Denzler 1993; Soffer and Weinstein 1999] for related results. Note that SG is globally well-posed in $L^2 \times H^{-1}$; then we have that breathers are stable under small perturbations in that space. Since the proofs are very similar, and in order to avoid repetition, we skip the details.

Moreover, following our proof it is possible to give a new proof of the global H^1-stability of two-solitons, first proved in [Martel et al. 2002].

We also claim that k-breathers ($k \geq 2$), namely solutions composed of k different breathers, are H^1-stable. Following the proof of Theorem 1.2, one can show by induction that a k-breather can be obtained from a $(k-1)$-breather after two Bäcklund transformations using a fixed set of complex conjugate parameters, as in Lemmas 2.4 and 5.1. After proving this identity, the rest of the proof adapts with no deep modifications.

This paper is organized as follows: In Section 2 we introduce the complex-valued soliton profiles. Section 3 is devoted to the study of the mKdV Bäcklund transformation in the vicinity of a given complex-valued mKdV solution. In Section 4 we apply the previous results to prove Theorem 1.3 (see Theorem 4.5). Section 5 deals with the relation between complex soliton profiles and breathers. In Section 6 we apply the results from Section 3 to the case of a perturbation of a breather solution. Finally, in Sections 7 and 8 we prove Theorems 1.2 and 1.4.

2. Complex-valued mKdV soliton profiles

Definition 2.1. Consider parameters α, $\beta > 0$, $x_1, x_2 \in \mathbb{R}$. We introduce the localized profile

$$\tilde{Q} = \tilde{Q}(x; \alpha, \beta, x_1, x_2),$$

defined as

$$\tilde{Q} := 2\sqrt{2} \arctan\left(e^{\beta y_2 + i\alpha y_1}\right), \tag{2-1}$$

where y_1 and y_2 are (re)defined as

$$y_1 := x + x_1, \quad y_2 := x + x_2. \tag{2-2}$$

Note that

$$\lim_{x \to -\infty} \tilde{Q}(x) = 0. \tag{2-3}$$

We define the complex-valued soliton profile as follows:

$$
\begin{aligned}
Q &:= \partial_x \tilde{Q} \\
&= \frac{2\sqrt{2}(\beta + i\alpha)e^{\beta y_2 + i\alpha y_1}}{1 + e^{2(\beta y_2 + i\alpha y_1)}} \qquad (2\text{-}4)\\
&= \sqrt{2}\,\frac{\beta\cosh(\beta y_2)\cos(\alpha y_1) + \alpha\sinh(\beta y_2)\sin(\alpha y_1)}{\cosh^2(\beta y_2) - \sin^2(\alpha y_1)} \\
&\qquad + i\sqrt{2}\,\frac{\alpha\cosh(\beta y_2)\cos(\alpha y_1) - \beta\sinh(\beta y_2)\sin(\alpha y_1)}{\cosh^2(\beta y_2) - \sin^2(\alpha y_1)}. \qquad (2\text{-}5)
\end{aligned}
$$

Finally, we write

$$\tilde{Q}_t := -(\beta + i\alpha)^2 Q, \tag{2-6}$$

and

$$\tilde{Q}_1 := \partial_{x_1}\tilde{Q}, \quad \tilde{Q}_2 := \partial_{x_2}\tilde{Q}. \tag{2-7}$$

Note that Q is complex-valued and is pointwise convergent to the soliton Q_{β^2} as $\alpha \to 0$. A second condition satisfied by $\tilde{Q}$ and Q is the following periodicity property: for all $k \in \mathbb{Z}$,

$$
\begin{cases}
\tilde{Q}(x;\alpha,\beta,x_1 + k\pi/\alpha, x_2) = (-1)^k \tilde{Q}(x;\alpha,\beta,x_1,x_2), \\
Q(x;\alpha,\beta,x_1 + k\pi/\alpha, x_2) = (-1)^k Q(x;\alpha,\beta,x_1,x_2).
\end{cases} \tag{2-8}
$$

We remark that, in what follows, $\tilde{Q}$ and Q may blow up in finite time.

Lemma 2.2. *Consider the complex-valued soliton profile defined in* (2-1)–(2-5). *Assume that, for x_2 fixed and some $k \in \mathbb{Z}$,*

$$x_1 = x_2 + \frac{\pi}{\alpha}\left(k + \tfrac{1}{2}\right). \tag{2-9}$$

Then $\tilde{Q}$ and Q cannot be defined at $x = -x_2$. Moreover, if $x_1 = x_2 = 0$, then $Q(\,\cdot\,;\alpha,\beta,0,0) \in H^1(\mathbb{R};\mathbb{C})$.

Remark. We emphasize that, given x_2 fixed, the set of points x_1 of the form (2-9) for some $k \in \mathbb{Z}$ is a countable set of real numbers with no limit points.

Remark. The complex-valued function $\arctan z$ (leading to the definition of $\tilde{Q}$) has two branches of discontinuities of the form im with $m \in \mathbb{R}$, $|m| \geq 1$, appearing from the standard branch of the complex logarithm function $\operatorname{Re} z < 0$, $\operatorname{Im} z = 0$. Such discontinuities may induce singularities on the function Q. Fortunately, both Q and functions of the type sine and cosine of arguments of the form $\tilde{Q}$ are smooth except on the points determined by Lemma 2.2. Throughout this paper we shall work with functions of the latest form instead of the original $\tilde{Q}$.

Proof. Fix $x_2 \in \mathbb{R}$. If (2-9) is satisfied for some $k \in \mathbb{Z}$, we have that, at $x = -x_2$,

$$y_1 = x + x_1 = \frac{\pi}{\alpha}\left(k + \tfrac{1}{2}\right), \quad y_2 = x + x_2 = 0,$$

and

$$\sinh(\beta y_2) = 0, \quad \cos(\alpha y_1) = 0. \tag{2-10}$$

Therefore, under (2-9), we have from (2-1) and (2-5) that $\tilde{Q}$ and Q cannot be defined at $x = -x_2$. Finally, if $x_1 = x_2 = 0$, we have

$$k + \tfrac{1}{2} = 0, \quad k \in \mathbb{Z},$$

which is impossible. □

Lemma 2.3. *Fix α, $\beta > 0$ and $x_1, x_2 \in \mathbb{R}$ such that* (2-9) *is not satisfied. Then we have*

$$Q_{xx} - (\beta + i\alpha)^2 Q + Q^3 = 0 \quad \textit{for all } x \in \mathbb{R}, \tag{2-11}$$

and

$$Q_x^2 - (\beta + i\alpha)^2 Q^2 + \tfrac{1}{2} Q^4 = 0 \quad \textit{for all } x \in \mathbb{R}. \tag{2-12}$$

Moreover, the previous identities can be extended to any $x_1, x_2 \in \mathbb{R}$ by continuity.

Proof. This is direct from the definition. □

Assume that (2-9) does not hold. Consider the sine and cosine functions applied to complex numbers. We have, from (2-1) and (2-4),

$$\begin{aligned} \sin \frac{\tilde{Q}}{\sqrt{2}} &= \sin(2 \arctan e^{\beta y_2 + i\alpha y_1}) \\ &= 2e^{\beta y_2 + i\alpha y_1} \cos^2(\arctan e^{\beta y_2 + i\alpha y_1}) \\ &= \frac{2e^{\beta y_2 + i\alpha y_1}}{1 + e^{2(\beta y_2 + i\alpha y_1)}} = \frac{1}{\beta + i\alpha} \frac{Q}{\sqrt{2}}. \end{aligned} \tag{2-13}$$

Similarly, from this identity we have

$$Q_x - (\beta + i\alpha) \cos\left(\frac{\tilde{Q}}{\sqrt{2}}\right) Q = 0, \tag{2-14}$$

so that, from (2-6) and (2-12),

$$\tilde{Q}_t + (\beta + i\alpha)\left[Q_x \cos \frac{\tilde{Q}}{\sqrt{2}} + \frac{Q^2}{\sqrt{2}} \sin \frac{\tilde{Q}}{\sqrt{2}}\right] = -(\beta + i\alpha)^2 Q + Q_x^2 Q^{-1} + \tfrac{1}{2} Q^3 = 0.$$

So far, we have proved the following result:

Lemma 2.4. *Let Q be a complex-valued soliton profile with scaling parameters α, $\beta > 0$ and shifts $x_1, x_2 \in \mathbb{R}$ such that* (2-9) *is not satisfied. Then we have*

$$\frac{Q}{\sqrt{2}} - (\beta + i\alpha) \sin \frac{\tilde{Q}}{\sqrt{2}} \equiv 0, \tag{2-15}$$

and

$$\tilde{Q}_t + (\beta + i\alpha)\left[Q_x \cos \frac{\tilde{Q}}{\sqrt{2}} + \frac{Q^2}{\sqrt{2}} \sin \frac{\tilde{Q}}{\sqrt{2}}\right] \equiv 0, \tag{2-16}$$

where $\sin z$ *and* $\cos z$ *are defined on the complex plane in the usual sense.*

We finish this section with a simple computational lemma.

Lemma 2.5. *Fix x_1, x_2 such that* (2-9) *is not satisfied. Then, for all $\alpha, \beta > 0$ we have*

$$\mathcal{N} := \frac{1}{2}\int_{-\infty}^{x} Q^2 = \frac{2(\beta+i\alpha)e^{2(\beta y_2+i\alpha y_1)}}{1+e^{2(\beta y_2+i\alpha y_1)}}, \tag{2-17}$$

and

$$\frac{1}{2}\int_{\mathbb{R}} Q^2 = 2(\beta+i\alpha), \tag{2-18}$$

no matter what x_1, x_2 are. Finally, if we let $L_1 := \log(1+e^{2(\beta x_2+i\alpha x_1)})$,

$$\int_0^x \mathcal{N} = \log(1+e^{2(\beta y_2+i\alpha y_1)}) - L_1. \tag{2-19}$$

Note that the previous formula is well-defined, since x_1 and x_2 do not satisfy (2-9).

Proof. It is not difficult to check that (2-17) is satisfied. Note that

$$\lim_{x\to-\infty}\left|\frac{2(\beta+i\alpha)e^{2(\beta y_2+i\alpha y_1)}}{1+e^{2(\beta y_2+i\alpha y_1)}}\right| = 0.$$

Identity (2-18) is a consequence of the fact that

$$\lim_{x\to+\infty}\frac{2(\beta+i\alpha)e^{2(\beta y_2+i\alpha y_1)}}{1+e^{2(\beta y_2+i\alpha y_1)}} = 2(\beta+i\alpha).$$

Finally, (2-19) is easy to check. □

3. Bäcklund transformation for mKdV

Lemma 2.4 is a consequence of a deeper result. In what follows, we fix a primitive $\tilde f$ of f, i.e.,

$$\tilde f_x := f, \tag{3-1}$$

where f is assumed only to be in $L^2(\mathbb{R})$. Notice that, even if $f = f(t,x)$ is a solution of mKdV, a corresponding term $\tilde f(t,x)$ may be unbounded in space.

Definition 3.1 (see, e.g., [Lamb 1980]). Let

$$(u_a, u_b, v_a, v_b, m) \in H^1(\mathbb{R};\mathbb{C})^2 \times H^{-1}(\mathbb{R};\mathbb{C})^2 \times \mathbb{C}.$$

We set

$$G := (G_1, G_2), \qquad G = G(u_a, u_b, v_a, v_b, m),$$

where

$$G_1(u_a, u_b, v_a, v_b, m) := \frac{u_a - u_b}{\sqrt{2}} - m\sin\frac{\tilde u_a + \tilde u_b}{\sqrt{2}}, \tag{3-2}$$

and

$$G_2(u_a, u_b, v_a, v_b, m) := v_a - v_b + m\left[((u_a)_x + (u_b)_x)\cos\frac{\tilde u_a + \tilde u_b}{\sqrt{2}} + \frac{u_a^2 + u_b^2}{\sqrt{2}}\sin\frac{\tilde u_a + \tilde u_b}{\sqrt{2}}\right]. \tag{3-3}$$

For the moment we do not specify the range of $G(u_a, u_b, v_a, v_b, m)$ for data (u_a, u_b, v_a, v_b, m) in $H^1(\mathbb{R};\mathbb{C})^2 \times H^{-1}(\mathbb{R};\mathbb{C})^2 \times \mathbb{C}$. However, thanks to Lemma 2.4, we have the following result:

Lemma 3.2. *Assume that x_1 and x_2 do not satisfy* (2-9). *Then*

$$G(Q, 0, \tilde{Q}_t, 0, \beta + i\alpha) \equiv (0, 0).$$

The previous identity can be extended by zero to the case where x_1 and x_2 satisfy (2-9), in such a form that $G(Q, 0, \tilde{Q}_t, 0, \beta+i\alpha)$, as a function of $(x_1, x_2) \in \mathbb{R}^2$, is now well-defined and continuous everywhere.

In what follows we consider the invertibility of the Bäcklund transformation on complex-valued functions. See [Hoffman and Wayne 2013] for the statement involving the real-valued solitons in the sine-Gordon case and [Mizumachi and Pelinovsky 2012] for the case of nonlinear Schrödinger solitons.

Proposition 3.3. *Let $X^0 := (u_a^0, u_b^0, v_a^0, v_b^0, m^0) \in H^1(\mathbb{R}; \mathbb{C})^2 \times H^{-1}(\mathbb{R}; \mathbb{C})^2 \times \mathbb{C}$ be such that*

$$\operatorname{Re} m^0 > 0, \tag{3-4}$$

$$G(X^0) = (0, 0), \tag{3-5}$$

$$\sin \frac{\tilde{u}_a^0 + \tilde{u}_b^0}{\sqrt{2}} \in H^1(\mathbb{R}; \mathbb{C}), \tag{3-6}$$

$$\text{and} \quad \lim_{-\infty} (\tilde{u}_a^0 + \tilde{u}_b^0) = 0, \quad \lim_{+\infty} (\tilde{u}_a^0 + \tilde{u}_b^0) = \sqrt{2}\pi. \tag{3-7}$$

Assume additionally that the ODE

$$\mu_x^0 - m^0 \cos\left(\frac{\tilde{u}_a^0 + \tilde{u}_b^0}{\sqrt{2}}\right)\mu^0 = 0, \tag{3-8}$$

has a smooth *solution $\mu^0 = \mu^0(x) \in \mathbb{C}$ satisfying*

$$\mu^0 \in H^1(\mathbb{R}; \mathbb{C}), \quad |\mu^0(x)| > 0, \quad \left|\frac{\mu_x^0(x)}{\mu^0(x)}\right| \le C, \tag{3-9}$$

$$\text{and} \quad \int_{\mathbb{R}} \sin\left(\frac{\tilde{u}_a^0 + \tilde{u}_b^0}{\sqrt{2}}\right)\mu^0 \neq 0. \tag{3-10}$$

Then there exist $\nu_0 > 0$ and $C > 0$ such that the following is satisfied: *For any $0 < \nu < \nu_0$ and any $(u_a, v_a) \in H^1(\mathbb{R}; \mathbb{C}) \times H^{-1}(\mathbb{R}; \mathbb{C})$ satisfying*

$$\|u_a - u_a^0\|_{H^1(\mathbb{R};\mathbb{C})} < \nu, \tag{3-11}$$

G is well-defined in a neighborhood of X^0 and there exists an unique (u_b, v_b, m) defined in an open *subset of $H^1(\mathbb{R}, \mathbb{C}) \times H^{-1}(\mathbb{R}; \mathbb{C}) \times \mathbb{C}$ such that*

$$G(u_a, u_b, v_a, v_b, m) \equiv (0, 0), \tag{3-12}$$

$$\|\tilde{u}_a + \tilde{u}_b - \tilde{u}_a^0 - \tilde{u}_b^0\|_{H^2(\mathbb{R};\mathbb{C})} \le C\nu, \tag{3-13}$$

$$\|u_b - u_b^0\|_{H^1(\mathbb{R};\mathbb{C})} + |m - m^0| < C\nu, \tag{3-14}$$

$$\sin \frac{\tilde{u}_a + \tilde{u}_b}{\sqrt{2}} \in H^1(\mathbb{R}; \mathbb{C}), \tag{3-15}$$

$$\text{and} \quad \lim_{-\infty}(\tilde{u}_a + \tilde{u}_b) = 0, \quad \lim_{+\infty}(\tilde{u}_a + \tilde{u}_b) = \sqrt{2}\pi. \tag{3-16}$$

Proof. Given u_a, u_b, m and v_a well-defined, v_b is uniquely defined from (3-3). We solve for u_b and m now. We will use the implicit function theorem.

We make a change of variables in order to specify a suitable range for G and to be able to prove (3-16). Define

$$u_c := u_a + u_b - u_c^0, \quad u_c^0 := u_a^0 + u_b^0 \in H^1(\mathbb{R}; \mathbb{C}), \tag{3-17}$$

and similarly for $\tilde{u}_c$ and $\tilde{u}_c^0$:

$$(\tilde{u}_c)_x = u_c, \quad (\tilde{u}_c^0)_x = u_c^0.$$

In what follows, we will look for a suitable $\tilde{u}_c$ with decay, and then we find u_b. Indeed, note that given u_c and u_a, u_b can be easily obtained. Then, with a slight abuse of notation, we consider G defined as follows:

$$G = (G_1, G_2), \quad G = G(u_a, \tilde{u}_c, v_a, v_b, m),$$

and

$$\begin{aligned} G : H^1(\mathbb{R}; \mathbb{C}) \times H^2(\mathbb{R}; \mathbb{C}) \times H^{-1}(\mathbb{R}; \mathbb{C})^2 \times \mathbb{C} &\longrightarrow H^1(\mathbb{R}; \mathbb{C}) \times H^{-1}(\mathbb{R}; \mathbb{C}) \\ (u_a, \tilde{u}_c, v_a, v_b, m) &\longmapsto G(u_a, \tilde{u}_c, v_a, v_b, m), \end{aligned}$$

where, from (3-2),

$$G_1(u_a, \tilde{u}_c, v_a, v_b, m) := \frac{2u_a - u_c^0 - u_c}{\sqrt{2}} - m \sin \frac{\tilde{u}_c^0 + \tilde{u}_c}{\sqrt{2}}, \tag{3-18}$$

and, from (3-3),

$$\begin{aligned} &G_2(u_a, \tilde{u}_c, v_a, v_b, m) \\ &\quad := v_a - v_b + m\left[(u_c^0 + u_c)_x \cos \frac{\tilde{u}_c^0 + \tilde{u}_c}{\sqrt{2}} + \frac{u_a^2 + (u_c^0 + u_c - u_a)^2}{\sqrt{2}} \sin \frac{\tilde{u}_c^0 + \tilde{u}_c}{\sqrt{2}}\right]. \end{aligned} \tag{3-19}$$

Clearly G as in (3-18)–(3-19) defines a C^1 functional in a small neighborhood of X^1 given by

$$X^1 := (u_a^0, 0, v_a^0, v_b^0, m^0) \in H^1(\mathbb{R}; \mathbb{C}) \times H^2(\mathbb{R}; \mathbb{C}) \times H^{-1}(\mathbb{R}; \mathbb{C})^2 \times \mathbb{C}, \tag{3-20}$$

where G is well-defined according to (3-6). Let us apply the implicit function theorem at this point. By (3-18) we have to show that

$$u_c + m^0 \cos\left(\frac{\tilde{u}_c^0}{\sqrt{2}}\right)\tilde{u}_c = f - m \sin \frac{\tilde{u}_c^0}{\sqrt{2}}$$

has a unique solution $(\tilde{u}_c, m)$ such that $\tilde{u}_c \in H^2(\mathbb{R}; \mathbb{C})$ for any $f \in H^1(\mathbb{R}; \mathbb{C})$ with linear bounds. From (3-7), we have

$$\lim_{x \to \pm\infty} \cos \frac{\tilde{u}_c^0}{\sqrt{2}} = \mp 1, \tag{3-21}$$

so that we can assume

$$\mu^0(x) = \exp\left(m^0 \int_0^x \cos \frac{\tilde{u}_c^0}{\sqrt{2}}\right).$$

Note that μ^0 decays exponentially in space as $x \to \pm\infty$. We have

$$\mu^0 u_c + (\mu^0)_x \tilde{u}_c = \mu^0\left[f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right].$$

Using (3-10), we choose $m \in \mathbb{C}$ such that

$$\int_{\mathbb{R}} \mu^0\left[f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right] = 0, \tag{3-22}$$

so that

$$|m| \le C\|f\|_{L^2(\mathbb{R};\mathbb{C})}$$

with $C > 0$ depending on the quantity $\left|\int_{\mathbb{R}} \mu^0 \sin(\tilde{u}_c^0/\sqrt{2})\right| \neq 0$ and $\|\mu^0\|_{L^2(\mathbb{R};\mathbb{C})}$.[8] We get

$$\tilde{u}_c = \frac{1}{\mu^0}\int_{-\infty}^{x} \mu^0\left[f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right]. \tag{3-23}$$

Finally, note that we have $u_c \in H^1(\mathbb{R};\mathbb{C})$. Indeed, first of all, thanks to (3-22), (3-8) and (3-21),

$$\lim_{x\to\pm\infty} \tilde{u}_c = \lim_{x\to\pm\infty} \frac{\mu^0}{\mu_x^0}\left[f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right] = 0.$$

If $s \le x \ll -1$, from (3-21) we get

$$\left|\frac{\mu^0(s)}{\mu^0(x)}\right| = \left|\exp\left(-m^0\int_s^x \cos\frac{\tilde{u}_c^0}{\sqrt{2}}\right)\right| \le Ce^{-\operatorname{Re} m^0(x-s)},$$

so that we have, for $x < 0$ and large,[9]

$$\begin{aligned}|\tilde{u}_c(x)| &\le C\int_{-\infty}^{x} e^{-(\operatorname{Re} m^0)(x-s)}\left|f(s) - m\sin\frac{\tilde{u}_c^0(s)}{\sqrt{2}}\right| ds \\ &\le C\mathbf{1}_{(-\infty,x]}e^{-(\operatorname{Re} m^0)(\cdot)} \star \left|f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right|, \quad \operatorname{Re} m^0 > 0.\end{aligned}$$

A similar result holds for $x > 0$ large, after using (3-22). Therefore, from Young's inequality,

$$\|\tilde{u}_c\|_{L^2(\mathbb{R};\mathbb{C})} \le C\left\|f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right\|_{L^2(\mathbb{R};\mathbb{C})} \le C\|f\|_{L^2(\mathbb{R};\mathbb{C})}, \tag{3-24}$$

as desired. On the other hand,

$$(\tilde{u}_c)_x = \left[f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right] - \frac{\mu_x^0}{(\mu^0)^2}\int_{-\infty}^{x} \mu^0\left[f - m\sin\frac{\tilde{u}_c^0}{\sqrt{2}}\right].$$

Since μ_x^0/μ^0 is bounded (see (3-9)), we have $\tilde{u}_c \in H^1(\mathbb{R};\mathbb{C})$. Finally, it is easy to see that $\tilde{u}_c \in H^2(\mathbb{R};\mathbb{C})$. Note that the constant involving the boundedness of the linear operator $f \mapsto \tilde{u}_c$ depends on the H^1-norm of μ^0, which blows up if (2-9) is satisfied.

[8]Note that $\|\mu^0\|_{L^2(\mathbb{R};\mathbb{C})}$ blows up as (2-9) is attained.

[9]Here the symbol $\star$ denotes *convolution*.

It turns out that we can apply the implicit function theorem to the operator G described in (3-18)–(3-19), so that (3-12) is satisfied, provided (3-11) holds.

First of all, note that (3-15) and (3-16) follow from $\tilde{u}_c \in H^2(\mathbb{R};\mathbb{C})$.

On the other hand, the estimate (3-13) is equivalent to

$$\|\tilde{u}_c\|_{H^2(\mathbb{R};\mathbb{C})} \le C\nu.$$

We will obtain this estimate using the almost linear character of the operator G around the point X^1. Since $\tilde{u}_c$ satisfies (3-18), we have

$$\frac{2u_a-(\tilde{u}_c)_x}{\sqrt{2}} - m\sin\frac{\tilde{u}_c^0+\tilde{u}_c}{\sqrt{2}} = 0.$$

Recall that $\tilde{u}_c$ depends on u_a. Near u_a^0, one has

$$\partial_t \tilde{u}_c[u_a^0+th]\big|_{t=0} = w[h] + O(h^2),$$

where $w = w[h]$ solves the *derivative* equation

$$w_x + m^0\cos\Big(\frac{\tilde{u}_c^0}{\sqrt{2}}\Big)w = -2h - m[h]\sin\frac{\tilde{u}_c^0}{\sqrt{2}}.$$

Here $m[h]$ is a constant that makes the right-hand side integrable, just as in (3-23). From (3-11) we know that $\|u_a-u_a^0\|_{H^1(\mathbb{R};\mathbb{C})} < \nu$. We shall use $h := u_a - u_a^0$. Following the computations after (3-23), we obtain the desired conclusion (see, e.g., (3-24)). We conclude that the L^2 norm of $\tilde{u}_c$ is bounded by $C\nu$. For the derivatives of $\tilde{u}_c$, the proof is very similar. □

Later we will need a second invertibility theorem. This time we assume that m is fixed, $u_b \sim u_b^0$ is known and we look for $u_a \sim u_a^0$. Note that the positive sign in front of (3-2) will be essential for the proof, otherwise we cannot take m fixed.

Proposition 3.4. *Let* $X^0 = (u_a^0, u_b^0, v_a^0, v_b^0, m^0) \in H^1(\mathbb{R};\mathbb{C})^2 \times H^{-1}(\mathbb{R},\mathbb{C}) \times \mathbb{C}$ *be such that* (3-4), (3-5), (3-6) *and* (3-7) *are satisfied. Assume additionally that the ODE*

$$(\mu^1)_x + m\cos\Big(\frac{\tilde{u}_a^0+\tilde{u}_b^0}{\sqrt{2}}\Big)\mu^1 = 0 \tag{3-25}$$

has a smooth *solution* $\mu^1 = \mu^1(x) \in \mathbb{C}$ *satisfying*

$$|\mu^1(x)| > 0, \qquad \left|\frac{\mu_x^1(x)}{\mu^1(x)}\right| \le C, \qquad \frac{1}{\mu^1} \in H^1(\mathbb{R};\mathbb{C}), \tag{3-26}$$

and G is smooth in a small neighborhood of X^0. Then there exists $\nu_1 > 0$ and a fixed constant $C > 0$ such that for all $0 < \nu < \nu_1$ the following is satisfied: for any $(u_b, v_b, m) \in H^1(\mathbb{R};\mathbb{C}) \times H^{-1}(\mathbb{R};\mathbb{C}) \times \mathbb{C}$ such that

$$\|u_b - u_b^0\|_{H^1(\mathbb{R};\mathbb{C})} + |m - m^0| < \nu, \tag{3-27}$$

G is well-defined and there exist unique $(u_a, v_a) \in H^1(\mathbb{R}, \mathbb{C}) \times H^{-1}(\mathbb{R}; \mathbb{C})$ such that

$$G(u_a, u_b, v_a, v_b, m) \equiv (0,0),$$

$$\int_{\mathbb{R}} (u_a - u_b)\left(\frac{1}{\mu^1}\right)_x = 0. \tag{3-28}$$

$$\|\tilde u_a + \tilde u_b - \tilde u_a^0 - \tilde u_b^0\|_{H^2(\mathbb{R};\mathbb{C})} \le C\nu, \tag{3-29}$$

$$\lim_{-\infty}(\tilde u_a + \tilde u_b) = 0, \quad \lim_{+\infty}(\tilde u_a + \tilde u_b) = \sqrt{2}\pi, \tag{3-30}$$

$$\textit{and} \quad \|u_a - u_a^0\|_{H^1(\mathbb{R};\mathbb{C})} < C\nu. \tag{3-31}$$

Proof. Given u_a, u_b and v_b well-defined, v_a is uniquely defined from (3-3). We solve for u_a now.

We follow the ideas of the proof of Proposition 3.3. However, this time we consider G defined in the opposite sense: using (3-17),

$$G = (G_3, G_4), \quad G = G(\tilde u_c, u_b, v_a, v_b, m),$$
$$G : H^2(\mathbb{R};\mathbb{C}) \times H^1(\mathbb{R};\mathbb{C}) \times H^{-1}(\mathbb{R};\mathbb{C})^2 \times \mathbb{C} \longrightarrow H^1(\mathbb{R};\mathbb{C}) \times H^{-1}(\mathbb{R};\mathbb{C})$$
$$(\tilde u_c, u_b, v_a, v_b, m) \longmapsto G(\tilde u_c, u_b, v_a, v_b, m)$$

with

$$\int_{\mathbb{R}} (\tilde u_c)_x \left(\frac{1}{\mu^1}\right)_x = 0, \tag{3-32}$$

where, from (3-2),

$$G_3(\tilde u_c, u_b, v_a, v_b, m) := \frac{u_c^0 + u_c - 2u_b}{\sqrt{2}} - m \sin\frac{\tilde u_c^0 + \tilde u_c}{\sqrt{2}}, \tag{3-33}$$

and, from (3-3),

$$G_4(\tilde u_c, u_b, v_a, v_b, m)$$
$$:= v_a - v_b + m\left[(u_c^0 + u_c)_x \cos\frac{\tilde u_c^0 + \tilde u_c}{\sqrt{2}} + \frac{(u_c^0 + u_c - u_b)^2 + u_b^2}{\sqrt{2}} \sin\frac{\tilde u_c^0 + \tilde u_c}{\sqrt{2}}\right]. \tag{3-34}$$

Clearly G as in (3-33)–(3-34) defines a C^1 functional in a small neighborhood of X^2 given by

$$X^2 := (0, u_b^0, v_a^0, v_b^0, m^0) \in H^2(\mathbb{R};\mathbb{C}) \times H^1(\mathbb{R};\mathbb{C}) \times H^{-1}(\mathbb{R};\mathbb{C})^2 \times \mathbb{C}, \tag{3-35}$$

where G is well-defined according to (3-6) and $G(X^2) = (0,0)$.

Fix m close enough to m^0. Now we have to show that

$$u_c - m\cos\left(\frac{\tilde u_c^0}{\sqrt{2}}\right)\tilde u_c = f \tag{3-36}$$

has a unique solution $\tilde u_c$ such that $u_c \in H^2(\mathbb{R};\mathbb{C})$ for any $f \in H^1(\mathbb{R};\mathbb{C})$. Indeed, consider μ^1 given by (3-25). It is not difficult to check that (see conditions (3-4), (3-27) and (3-7))

$$\operatorname{Re} m > 0, \quad \lim_{\pm\infty} \cos\frac{\tilde u_c^0}{\sqrt{2}} = \mp 1, \quad \text{and} \quad \mu^1 = \exp\left(-m\int_0^x \cos\frac{\tilde u_c^0}{\sqrt{2}}\right). \tag{3-37}$$

Note that, by (3-37) and (3-4), $|\mu^1(x)|$ is exponentially growing in space as $x \to \pm\infty$. From (3-36),

$$(\mu^1 \tilde{u}_c)_x = \mu^1 f,$$

so that, thanks to (3-26),

$$\tilde{u}_c = \frac{1}{\mu^1}\mu^1(0)\tilde{u}_c(0) + \frac{1}{\mu^1}\int_0^x \mu^1 f.$$

Clearly $\lim_{\pm\infty} \tilde{u}_c = 0$ for $f \in H^1(\mathbb{R};\mathbb{C})$. In order to ensure uniqueness, we seek $\tilde{u}_c$ satisfying

$$\int_{\mathbb{R}} u_c \left(\frac{1}{\mu^1}\right)_x = 0,$$

which is nothing but (3-32) and (3-28), which is justified by (3-26). Let us show that $\tilde{u}_c \in L^2(\mathbb{R};\mathbb{C})$. We have, for $x > 0$ large,

$$|\tilde{u}_c(x)| \le C\int_0^x e^{-(\operatorname{Re} m)(x-s)}|f(s)|\,ds = Ce^{-(\operatorname{Re} m)(\cdot)} \star |f|, \quad \operatorname{Re} m > 0.$$

A similar estimate can be established if $x < 0$. Therefore, using Young's inequality,

$$\|\tilde{u}_c\|_{L^2(\mathbb{R};\mathbb{C})} \le C\|f\|_{L^2(\mathbb{R};\mathbb{C})},$$

as desired. Now we check that $u_c \in H^1(\mathbb{R};\mathbb{C})$. Indeed, we have

$$u_c = f - \frac{\mu^1_x}{(\mu^1)^2}\int_0^x \mu^1 f.$$

Since μ^1_x/μ^1 is bounded, we have proven that $u_c \in L^2(\mathbb{R};\mathbb{C})$. A new iteration proves that $u_c \in H^1(\mathbb{R};\mathbb{C})$. Estimates (3-29)–(3-31) are consequences of the implicit function theorem and can be proved as in the previous proposition. The proof is complete. □

We finish this section by pointing out the role played by the Bäcklund transformation in the mKdV dynamics. We recall the following standard result:

Theorem 3.5. *Let $m \in \mathbb{C}$ be a* fixed *parameter, and $I \subset \mathbb{R}$ an open time interval. Assume that $u_b \in C(I; H^1(\mathbb{R};\mathbb{C}))$ solves* (1-1), *i.e.*,

$$(u_b)_t + ((u_b)_{xx} + u_b^3)_x = 0, \tag{3-38}$$

in the H^1-sense. Assume, moreover that u_b is close to u_b^0 and that (3-25) *and* (3-26) *hold. Define $v_b := -((u_b)_{xx} + u_b^3)$ as a distribution in $H^{-1}(\mathbb{R};\mathbb{C})$. Then, for each $t \in I$, the corresponding solution $(u_a(t), v_a(t))$ of $G_1 = G_2 = 0$ for m fixed, obtained in the space $H^1(\mathbb{R};\mathbb{C}) \times H^{-1}(\mathbb{R};\mathbb{C})$, satisfies the following*:

(1) $u_a \in C(I; H^1(\mathbb{R};\mathbb{C}))$;

(2) $(u_a)_t := (v_a)_x$ *is well-defined in* $H^{-2}(\mathbb{R};\mathbb{C})$; *and*

(3) u_a *solves* (1-1) *in the* H^1*-sense.*

Proof. The first step is an easy consequence of the continuous character of the solution map given by the implicit function theorem. By density we can assume $u_b(t) \in H^3(\mathbb{R};\mathbb{C})$. From (3-2) we have

$$(u_a)_x - (u_b)_x = m\cos\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)(u_a+u_b), \tag{3-39}$$

and

$$(u_a)_{xx} - (u_b)_{xx} = m\cos\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)((u_a)_x+(u_b)_x) - \frac{m}{\sqrt{2}}\sin\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)(u_a+u_b)^2.$$

Therefore, from (3-3) and (3-2),

$$\begin{aligned} v_a - v_b &= -((u_a)_{xx}-(u_b)_{xx}) - \frac{m}{\sqrt{2}}\sin\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)[(u_a+u_b)^2+(u_a^2+u_b^2)] \\ &= -((u_a)_{xx}-(u_b)_{xx}) - (u_a-u_b)(u_a^2+u_au_b+u_b^2) \\ &= -((u_a)_{xx}+u_a^3) - ((u_b)_{xx}+u_b^3). \end{aligned}$$

We have from (3-38) that $(v_b)_x + ((u_b)_{xx}+u_b^3)_x = 0$. Therefore,

$$(v_a)_x + ((u_a)_{xx}+u_a^3)_x = 0. \tag{3-40}$$

Finally, if $(u_a)_t = (v_a)_x$, we have that u_a solves (1-1). In order to prove this result, we compute the time derivative in (3-2): we get

$$(u_a)_t - (u_b)_t = m\cos\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)((\tilde{u}_a)_t+(\tilde{u}_b)_t). \tag{3-41}$$

Note that, given u_b, the solution u_a is uniquely defined, thanks to the implicit function theorem. Additionally, from (3-3),

$$\begin{aligned} (v_a)_x - (v_b)_x + m\Bigg[&((u_a)_{xx}+(u_b)_{xx})\cos\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}} - \frac{1}{\sqrt{2}}((u_a)_x+(u_b)_x)(u_a+u_b)\sin\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}} \\ &+\sqrt{2}(u_a(u_a)_x+u_b(u_b)_x)\sin\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}} + \frac{(u_a^2+u_b^2)}{2}(u_a+u_b)\cos\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\Bigg] = 0. \end{aligned}$$

We use (3-2) and (3-3) in the previous identity, and get

$$(v_a)_x - (v_b)_x + \left[m((u_a)_{xx}+(u_b)_{xx})\cos\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}} + (u_a^2-u_au_b+u_b^2)((u_a)_x-(u_b)_x)\right] = 0.$$

Finally, we use (3-39) to obtain

$$(v_a)_x - (v_b)_x + m\cos\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)((u_a)_{xx}+u_a^3+(u_b)_{xx}+u_b^3) = 0,$$

so (3-38) and (3-40) imply

$$(v_a)_x - (v_b)_x = m\cos\left(\frac{\tilde{u}_a+\tilde{u}_b}{\sqrt{2}}\right)(v_a+v_b),$$

so that from (3-41) and the uniqueness we are done. □

4. Dynamics of complex-valued mKdV solitons

In what follows we will apply the results from the previous section in a neighborhood of the complex soliton at time zero. Define (cf. (2-1)),

$$\tilde{Q}^0 := \tilde{Q}(x;\alpha,\beta,0,0), \tag{4-1}$$

and similarly for Q^0 and $\tilde{Q}^0_t$. Recall that, by Lemma 2.2, the complex soliton Q^0 is well-defined everywhere if (2-9) is not satisfied. Finally, given any

$$\tilde{z}^0_b \in \dot{H}^1(\mathbb{R};\mathbb{C}),$$

we define z^0_b by the identity (see (3-1), for instance)

$$z^0_b := (\tilde{z}^0_b)_x,$$

and, in term of distributions,

$$w^0_b := -((z^0_b)_{xx} + (z^0_b)^3) \in H^{-1}(\mathbb{R};\mathbb{C}).$$

Lemma 4.1. *There exists $\nu_0 > 0$ and $C > 0$ such that, for all $0 < \nu < \nu_0$, the following holds. For all $z^0_b \in H^1(\mathbb{R};\mathbb{C})$ satisfying*

$$\|z^0_b\|_{H^1(\mathbb{R};\mathbb{C})} < \nu, \tag{4-2}$$

there exist unique $y^0_a \in H^1(\mathbb{R},\mathbb{C})$, $y^1_a \in H^{-1}(\mathbb{R},\mathbb{C})$ and $m \in \mathbb{C}$ of the form

$$y^0_a(x) = y^0_a[z^0_b](x), \quad y^1_a(x) = y^1_a[z^0_b, w^0_b](x), \quad m := \beta + i\alpha + q^0 \tag{4-3}$$

such that

$$\begin{aligned} &\|y^0_a\|_{H^1(\mathbb{R};\mathbb{C})} + |q^0| \le C\nu, \quad \tilde{z}^0_a + \tilde{y}^0_a \in H^2(\mathbb{R};\mathbb{C}), \\ &\text{and} \quad G(Q^0 + z^0_b, y^0_a, \tilde{Q}^0_t + w^0_b, y^1_a, m) \equiv (0,0). \end{aligned} \tag{4-4}$$

Note that both $\tilde{z}^0_a$ and $\tilde{y}^0_a$ may be unbounded functions, but the sum is bounded on $\mathbb{R}$.

Proof. Let Q^0 be the soliton profile with parameters β, α and $x_1 = x_2 = 0$ (cf. (4-1)). We apply Proposition 3.3 with

$$u^0_a := Q^0, \quad u^0_b := 0, \quad v^0_a := \tilde{Q}^0_t, \quad v^0_b := 0 \quad \text{and} \quad m^0 := \beta + i\alpha.$$

Clearly $\tilde{u}^0_a + \tilde{u}^0_b = \tilde{Q}_0$ satisfies (3-6)–(3-7). From (2-15) we have

$$(Q^0)_x - (\beta + i\alpha)\cos\left(\frac{\tilde{Q}^0}{\sqrt{2}}\right) Q^0 = 0, \quad Q^0(-\infty) = 0, \tag{4-5}$$

so that we have (cf. (3-8)–(3-9))

$$\mu^0 = Q^0.$$

Clearly Q^0 is never zero. Moreover, $|(Q^0)^{-1}Q^0_x|$ is bounded on $\mathbb{R}$. Now we prove that

$$\int_{\mathbb{R}} \sin\left(\frac{\tilde{Q}^0}{\sqrt{2}}\right) Q^0 \neq 0.$$

From (2-15) and (2-18),

$$\int_{\mathbb{R}} \sin\left(\frac{\tilde{Q}^0}{\sqrt{2}}\right) Q^0 = \frac{1}{\sqrt{2}(\beta+i\alpha)} \int_{\mathbb{R}} (Q^0)^2 = \frac{4(\beta+i\alpha)}{\sqrt{2}(\beta+i\alpha)} = 2\sqrt{2}. \qquad \square$$

Before continuing, we need some definitions. We write

$$\alpha^* := \alpha + \operatorname{Im} q^0, \quad \beta^* := \beta + \operatorname{Re} q^0, \tag{4-6}$$

so that m in (4-3) satisfies

$$m = \beta + i\alpha + q^0 = \beta^* + i\alpha^*.$$

Since q^0 is small, we have that β^* and α^* are positive quantities. Similarly, define

$$\delta^* := (\alpha^*)^2 - 3(\beta^*)^2, \quad \gamma^* := 3(\alpha^*)^2 - (\beta^*)^2, \tag{4-7}$$

and compare with (1-4).

Consider the kink profile $\tilde{Q}$ introduced in (2-1). We consider, for all $t \in \mathbb{R}$, the complex (kink) profile

$$\tilde{Q}^*(t,x) := \tilde{Q}(x; \alpha^*, \beta^*, \delta^* t + x_1, \gamma^* t + x_2), \tag{4-8}$$

with δ^* and γ^* defined in (4-7), x_1 and x_2 *possibly* depending on time, and

$$Q^*(t,x) := \partial_x \tilde{Q}^*(t,x). \tag{4-9}$$

It is not difficult to see that (see, e.g., (1-10))

$$Q^*(t,x) = Q_c(x - ct - \hat{x}), \quad \sqrt{c} = \beta^* + i\alpha^*, \quad \hat{x} \in \mathbb{C},$$

which is a *complex-valued* solution of mKdV (1-1). Technically, the complex soliton $Q^*(t)$ has velocity $-\gamma^* = (\beta^*)^2 - 3(\alpha^*)^2$, a quantity that is always smaller than the corresponding speed $(\beta^*)^2$ of the associated real-valued soliton $Q_{(\beta^*)^2}$ obtained by sending α^* to zero. Finally, as in (2-6) we define

$$\tilde{Q}^*_t(t,x) := -(\beta^* + i\alpha^*)^2 Q^*(t,x).$$

Lemma 4.2. *Fix $\alpha, \beta > 0$. Assume that x_1, x_2 are time-dependent functions such that*

$$|x_1'(t)| + |x_2'(t)| \ll |\delta^* - \gamma^*| = 2((\alpha^*)^2 + (\beta^*)^2). \tag{4-10}$$

Then there exists a sequence of times $t_k \in \mathbb{R}$, $k \in \mathbb{Z}$ such that (2-9) *is satisfied. In particular*, (t_k) *is a sequence with no limit points.*

Proof. Note that (2-9) now reads

$$(\delta^* - \gamma^*)t_k + (x_1 - x_2)(t_k) = \frac{\pi}{\alpha^*}\left(k + \tfrac{1}{2}\right).$$

By (4-7), $\delta^* - \gamma^* = -2((\alpha^*)^2 + (\beta^*)^2) \neq 0$, and using (4-10) and the mean and intermediate value theorems applied to the smooth function

$$f(t) := (\delta^* - \gamma^*)t + (x_1 - x_2)(t),$$

at each value $\frac{\pi}{\alpha^*}\big(k + \frac{1}{2}\big)$, $k \in \mathbb{Z}$, we see that f satisfies

$$f'(t) = -2((\alpha^*)^2 + (\beta^*)^2) + (x_1' - x_2')(t) \sim -2((\alpha^*)^2 + (\beta^*)^2). \qquad \square$$

We conclude that $\tilde{Q}^*$ and Q^* defined in (4-8) and (4-9) are well-defined except for an isolated sequence of times t_k. We impose now the condition

$$t \in \mathbb{R} \quad \text{satisfies} \quad t \neq t_k \ \text{ for all } \ k \in \mathbb{Z}. \tag{4-11}$$

In what follows we will solve the Cauchy problem associated to mKdV with suitable initial data. Indeed, we will assume that

$$y_a^0 \ \text{ is a real-valued function and } \ y_a^0 \in H^1(\mathbb{R}). \tag{4-12}$$

We will need the following:[10]

Theorem 4.3 ([Kenig et al. 1993]). *For any $y_a^0 \in H^1(\mathbb{R})$, there exists a unique*[11] *solution $y_a \in C(\mathbb{R}, H^1(\mathbb{R}))$ with initial data $y_a(0) = y_a^0 \in H^1(\mathbb{R})$ to mKdV, and*

$$\sup_{t\in\mathbb{R}} \|y_a(t)\|_{H^1(\mathbb{R})} \leq C\|y_a^0\|_{H^1(\mathbb{R})}$$

with $C > 0$ independent of time. Moreover, the mass

$$M[y_a](t) := \frac{1}{2}\int_{\mathbb{R}} y_a^2(t,x)\,dx = M[y_a^0] \tag{4-13}$$

and energy

$$E[y_a](t) := \frac{1}{2}\int_{\mathbb{R}} (y_a)_x^2(t,x)\,dx - \frac{1}{4}\int_{\mathbb{R}} (y_a)^4(t,x)\,dx = E[y_a^0] \tag{4-14}$$

are conserved quantities.

Let $y_a \in C(\mathbb{R}, H^1(\mathbb{R}))$ denote the corresponding solution for mKdV with initial data y_a^0. Since $\|y_a^0\|_{H^1} \leq C\eta$, we have, for a (possibly different) constant $C > 0$,

$$\sup_{t\in\mathbb{R}} \|y_a(t)\|_{H^1(\mathbb{R})} \leq C\eta. \tag{4-15}$$

In particular, we can define, for all $t \in \mathbb{R}$,

$$\tilde{y}_a(t) := \int_0^x y_a(t,s)\,ds,$$

and

$$(\tilde{y}_a)_t(t) := -((y_a)_{xx}(t) + y_a^3(t)) \in H^{-1}(\mathbb{R}) \tag{4-16}$$

[10] We recall that this result is consequence of the local Cauchy theory and the conservation of mass and energy (4-13)–(4-14).
[11] In a certain sense; see [Kenig et al. 1993].

because $y_a(t) \in L^p(\mathbb{R})$ for all $p \geq 2$.

Lemma 4.4. *Assume that a time $t \in \mathbb{R}$ and y_a^0 are such that* (4-11) *and* (4-12) *hold. Then there are unique $z_b = z_b(t) \in H^1(\mathbb{R}; \mathbb{C})$ and $w_b = w_b(t) \in H^{-1}(\mathbb{R}; \mathbb{C})$ such that, for all $t \neq t_k$,*

$$\tilde{z}_b + \tilde{y}_a \in H^2(\mathbb{R}; \mathbb{C}), \tag{4-17}$$

$$\frac{1}{\sqrt{2}}(Q^* + z_b - y_a) = (\beta + i\alpha + q^0) \sin \frac{\tilde{Q}^* + \tilde{z}_b + \tilde{y}_a}{\sqrt{2}}, \tag{4-18}$$

where $\tilde{Q}$ and Q are defined in (4-8) *and* (4-9). *Moreover, we have*

$$0 = \tilde{Q}_t^* + w_b - (\tilde{y}_a)_t + (\beta + i\alpha + q^0)\left[(Q_x^* + (z_b)_x + (y_a)_x) \cos \frac{\tilde{Q}^* + \tilde{z}_b + \tilde{y}_a}{\sqrt{2}} + \frac{(Q^* + z_b)^2 + y_a^2}{\sqrt{2}} \sin \frac{\tilde{Q}^* + \tilde{z}_b + \tilde{y}_a}{\sqrt{2}}\right], \tag{4-19}$$

and, for all $t \neq t_k$,

$$\|z_b(t)\|_{H^1(\mathbb{R};\mathbb{C})} < C\nu \tag{4-20}$$

with C uniformly bounded provided t is uniformly far from each t_k.

Proof. We will use Proposition 3.4. For that it is enough to recall that, from (2-15) and (2-16), and for all $t \neq t_k$,[12]

$$\frac{1}{\sqrt{2}} Q^* = (\beta + i\alpha + q^0) \sin \frac{\tilde{Q}^*}{\sqrt{2}} \tag{4-21}$$

and

$$\tilde{Q}_t^* + (\beta + i\alpha + q^0)\left[Q_x^* \cos \frac{\tilde{Q}^*}{\sqrt{2}} + \frac{(Q^*)^2}{\sqrt{2}} \sin \frac{\tilde{Q}^*}{\sqrt{2}}\right] = 0,$$

so that we can apply Proposition 3.4 at $X^0 = (Q^*, 0, \tilde{Q}_t^*, 0, m)$, where, by, (4-21) we have $m = (\beta + i\alpha + q^0)$. It is not difficult to see that the function μ^1 in (3-25) is given by

$$\mu^1 = (Q^*)^{-1},$$

and (3-26) is satisfied. Note that we require the estimate (4-15) in order to obtain (4-18)–(4-19). Finally, (4-20) is a direct consequence of (3-31). □

Remark. Since, from (4-4), we get

$$\frac{1}{\sqrt{2}}(Q^0 + z_b^0 - y_a^0) = (\beta + i\alpha + q^0) \sin \frac{\tilde{Q}^0 + \tilde{z}_b^0 + \tilde{y}_a^0}{\sqrt{2}},$$

we have that (4-18) implies by uniqueness that

$$(Q^* + z_b - y_a)(t = 0) = Q^0 + z_b^0 - y_a^0,$$

[12] It is interesting to note that the shifts x_1, x_2 on $Q^*(t, x)$ cannot be modified, otherwise there is no continuity at $t = 0$.

i.e.,

$$(Q^* + z_b)(t = 0) = Q^0 + z_b^0.$$

We are ready to prove a detailed version of Theorem 1.3, a result on complex-valued solitons.

Theorem 4.5. *There exists $\nu_0 > 0$ such that for all $0 < \nu < \nu_0$ the following holds: Consider the initial data $u_b^0 := Q^0 + z_b^0 \in H^1(\mathbb{R}; \mathbb{C})$, where*

$$\|z_b^0\|_{H^1(\mathbb{R};\mathbb{C})} < \nu.$$

Assume in addition that the corresponding function y_a^0 given by Lemma 4.1 is real-valued and belongs to $H^1(\mathbb{R})$. Fix $\varepsilon_0 > 0$. Then, for all t such that $|t - t_k| \geq \varepsilon_0$, with t_k defined in Lemma 4.2, the function $u_b := Q^* + z_b$, with z_b introduced in Lemma 4.4, is an H^1 complex-valued solution of mKdV, it satisfies $(u_b)_t = (Q^* + z_b)_t = (\tilde{Q}_t^* + w_b)_x$ and

$$\sup_{|t-t_k|\geq\varepsilon_0} \|u_b(t) - Q^*(t)\|_{H^1(\mathbb{R};\mathbb{C})} \leq C_{\varepsilon_0}\nu. \tag{4-22}$$

Remark. The quantity $\varepsilon_0 > 0$ is just an auxiliary parameter and it can be made as small as required; however, the constant C_{ε_0} in (4-22) becomes singular as ε_0 approaches zero.

Remark. In Corollary 6.5 we will prove that there is an open set in $H^1(\mathbb{R}; \mathbb{C})$ leading to y_a^0 being real-valued. The openness of this set will be a consequence of the implicit function theorem.

Proof. We apply Lemma 4.1. Assuming (4-12) we have y_a^0 real-valued, so that there is an mKdV dynamics $y_a(t)$ constructed in Theorem 4.3. Lastly, we apply Lemma 4.4 to obtain the dynamical function $Q^*(t) + z_b(t)$. Theorem 3.5 gives the conclusion. □

Now we will prove that the mass and energy,

$$\frac{1}{2}\int_{\mathbb{R}} u_b^2(t) \quad \text{and} \quad \frac{1}{2}\int_{\mathbb{R}} (u_b)_x^2(t) - \frac{1}{4}\int_{\mathbb{R}} u_b^4(t), \tag{4-23}$$

remain conserved for all time without using the mKdV equation (1-1), only the Bäcklund transformation (4-18). The fact that $\tilde{z}_b + \tilde{y}_a$ is in $H^1(\mathbb{R}; \mathbb{C})$ will be essential for the proof.

Corollary 4.6. *Assume that $t \neq t_k$ for all $k \in \mathbb{Z}$. Then the quantity*

$$\frac{1}{2}\int_{\mathbb{R}} (Q^* + z_b)^2(t) \tag{4-24}$$

is well-defined and independent of time, and

$$\frac{1}{2}\int_{\mathbb{R}} (Q^* + z_b)^2(t) = \frac{1}{2}\int_{\mathbb{R}} (y_a^0)^2 + 2(\beta + i\alpha + q^0). \tag{4-25}$$

Moreover, (4-25) can be extended in a continuous form to every $t \in \mathbb{R}$.

Proof. Using (4-18) and multiplying each side by $(1/\sqrt{2})(Q^* + z_b + y_a)$, we obtain

$$\tfrac{1}{2}(Q^* + z_b - y_a)(Q^* + z_b + y_a) = -(\beta + i\alpha + q^0)\left[\cos\frac{\tilde{Q}^* + \tilde{z}_b + \tilde{y}_a}{\sqrt{2}}\right]_x.$$

Using (2-15) and (2-14),

$$\cos\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}} = \cos\frac{\tilde{Q}^*}{\sqrt{2}}\cos\frac{\tilde{z}_b+\tilde{y}_a}{\sqrt{2}} - \sin\frac{\tilde{Q}^*}{\sqrt{2}}\sin\frac{\tilde{z}_b+\tilde{y}_a}{\sqrt{2}} = \frac{1}{(\beta^*+i\alpha^*)}\left[\frac{Q_x^*}{Q^*}\cos\frac{\tilde{z}_b+\tilde{y}_a}{\sqrt{2}} - \frac{Q^*}{\sqrt{2}}\sin\frac{\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\right]. \tag{4-26}$$

We integrate on $\mathbb{R}$ to obtain

$$\frac{1}{2}\int_{\mathbb{R}}(Q^*+z_b-y_a)(Q^*+z_b+y_a) = -(\beta+i\alpha+q^0)\cos\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\bigg|_{-\infty}^{\infty}.$$

Since $\lim_{\pm\infty} Q^*=0$, $\lim_{\pm\infty} Q_x^*/Q^* = \mp(\beta^*+i\alpha^*)$ (see (2-4)) and $\lim_{\pm\infty}(\tilde{z}_b+\tilde{y}_a)=0$, we get (4-24)–(4-25), because the mass of $y_a(t)$ is conserved. □

Corollary 4.7. *Assume that $t \neq t_k$ for all $k\in\mathbb{Z}$. Then the quantity*

$$E[Q^*+z_b](t) := \frac{1}{2}\int_{\mathbb{R}}(Q^*+z_b)_x^2(t) - \frac{1}{4}\int_{\mathbb{R}}(Q^*+z_b)^4(t) \tag{4-27}$$

is well-defined and independent of time. Moreover, it satisfies

$$\frac{1}{2}\int_{\mathbb{R}}(Q^*+z_b)_x^2(t) - \frac{1}{4}\int_{\mathbb{R}}(Q^*+z_b)^4(t) = E[y_a^0] - \tfrac{2}{3}(\beta^*+i\alpha^*)^3.$$

Finally, this quantity can be extended in a continuous way to every $t\in\mathbb{R}$.

Proof. Let $m=(\beta+i\alpha+q^0)$. From (4-18) we have

$$(Q^*+z_b)_x-(y_a)_x = m\cos\left(\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\right)(Q^*+z_b+y_a). \tag{4-28}$$

Multiplying by $(Q^*+z_b)_x+(y_a)_x$, we get

$$(Q^*+z_b)_x^2-(y_a)_x^2 = m\cos\left(\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\right)(Q^*+z_b+y_a)(Q^*+z_b+y_a)_x = m\cos\left(\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\right)\big[(Q^*+z_b)(Q^*+z_b)_x+y_a(y_a)_x + y_a(Q^*+z_b)_x+(Q^*+z_b)(y_a)_x\big]. \tag{4-29}$$

On the other hand, we multiply (4-28) by y_a and (Q^*+z_b) to obtain

$$y_a(Q^*+z_b)_x - y_a(y_a)_x = m\cos\left(\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\right)(Q^*+z_b+y_a)y_a,$$

and

$$(Q^*+z_b)(Q^*+z_b)_x-(Q^*+z_b)(y_a)_x = m\cos\left(\frac{\tilde{Q}^*+\tilde{z}_b+\tilde{y}_a}{\sqrt{2}}\right)(Q^*+z_b+y_a)(Q^*+z_b).$$

If we subtract the latter from the former we get

$$y_a(Q^*+z_b)_x+(Q^*+z_b)(y_a)_x = (Q^*+z_b)(Q^*+z_b)_x+y_a(y_a)_x+m\cos\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)[y_a^2-(Q^*+z_b)^2]. \quad (4\text{-}30)$$

Substituting (4-30) into (4-29),

$$(Q^*+z_b)_x^2-(y_a)_x^2 = m\cos\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)[(Q^*+z_b)^2+y_a^2]_x+m^2\cos^2\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)[y_a^2-(Q^*+z_b)^2]. \quad (4\text{-}31)$$

Finally, we use (4-18) once again. We multiply by $(Q^*+z_b+y_a)$:

$$\frac{1}{\sqrt2}[(Q^*+z_b)^2-y_a^2]=m\sin\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)(Q^*+z_b+y_a).$$

Substituting in (4-31) we finally arrive to the identity

$$(Q^*+z_b)_x^2-(y_a)_x^2=m\cos\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)[(Q^*+z_b)^2+y_a^2]_x - m^3\sqrt2\cos^2\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)\sin\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)(Q^*+z_b+y_a).$$

The last term on the right-hand side above can be recognized as a total derivative. After integration and using (4-26), we obtain

$$\begin{aligned}\int_{\mathbb R}[(Q^*+z_b)_x^2-(y_a)_x^2]&=m\int_{\mathbb R}\cos\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)[(Q^*+z_b)^2+y_a^2]_x+\frac23 m^3\cos^3\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)\Big|_{-\infty}^{+\infty}\\ &=\frac{m}{\sqrt2}\int_{\mathbb R}\sin\Big(\frac{\tilde Q^*+\tilde z_b+\tilde y_a}{\sqrt2}\Big)(Q^*+z_b+y_a)[(Q^*+z_b)^2+y_a^2]-\frac43 m^3\\ &=\frac12\int_{\mathbb R}[(Q^*+z_b)^2-y_a^2][(Q^*+z_b)^2+y_a^2]-\frac43 m^3\\ &=\frac12\int_{\mathbb R}[(Q^*+z_b)^4-y_a^4]-\frac43 m^3.\end{aligned}$$

Finally,

$$\frac12\int_{\mathbb R}(Q^*+z_b)_x^2-\frac14\int_{\mathbb R}(Q^*+z_b)^4=\frac12\int_{\mathbb R}(y_a)_x^2-\frac14\int_{\mathbb R}y_a^4-\frac23(\beta+i\alpha+q^0)^3.$$

Since the right-hand side above is conserved for all time, we have proved (4-27). □

5. Complex solitons versus breathers

We introduce now the notion of breather profile. Given parameters $x_1, x_2\in\mathbb R$ and $\alpha,\beta>0$, we consider y_1 and y_2 defined in (2-2). Let $\tilde B$ be the localized profile

$$\tilde B=\tilde B(x;\alpha,\beta,x_1,x_2):=2\sqrt2\arctan\frac{\beta\sin(\alpha y_1)}{\alpha\cosh(\beta y_2)}, \quad (5\text{-}1)$$

and, with a slight abuse of notation, we redefine

$$B := \tilde{B}_x. \tag{5-2}$$

Note that

$$\tilde{B}(-\infty) = \tilde{B}(+\infty) = 0 \tag{5-3}$$

and, for $k \in \mathbb{Z}$,

$$\begin{cases} \tilde{B}(x; \alpha, \beta, x_1 + k\pi/\alpha, x_2) = (-1)^k \tilde{B}(x; \alpha, \beta, x_1, x_2), \\ B(x; \alpha, \beta, x_1 + k\pi/\alpha, x_2) = (-1)^k B(x; \alpha, \beta, x_1, x_2). \end{cases} \tag{5-4}$$

Now we introduce the directions associated to the shifts x_1 and x_2. Given a breather profile of parameters α, β, x_1 and x_2, we define

$$B_1 = B_1(x; \alpha, \beta, x_1, x_2) := \partial_{x_1} B, \tag{5-5}$$

$$B_2 = B_2(x; \alpha, \beta, x_1, x_2) := \partial_{x_2} B \tag{5-6}$$

and, for δ and γ defined in (1-4),

$$\tilde{B}_t := \delta B_1 + \gamma B_2. \tag{5-7}$$

We also have

$$\tilde{B}_t + B_{xx} + B^3 = 0; \tag{5-8}$$

see [Alejo and Muñoz 2013] for a proof of this identity.

If x_1 or x_2 are time-dependent variables, we assume that the associated B_j corresponds to the partial derivative with respect to the time-independent variable x_j, evaluated at $x_j(t)$.

In this section we will prove that there is a deep interplay between complex solitons and breather profiles. We start with the following identities:

Lemma 5.1. *Let (B, Q) be a pair breather-soliton profiles with scaling parameters $\alpha, \beta > 0$ and shifts $x_1, x_2 \in \mathbb{R}$. Assume that* (2-9) *is not satisfied. Then we have*

$$\frac{B - Q}{\sqrt{2}} - (\beta - i\alpha) \sin \frac{\tilde{B} + \tilde{Q}}{\sqrt{2}} \equiv 0, \tag{5-9}$$

$$\text{and} \quad \tilde{B}_t - \tilde{Q}_t + (\beta - i\alpha)\left[(B_x + Q_x) \cos \frac{\tilde{B} + \tilde{Q}}{\sqrt{2}} + \frac{B^2 + Q^2}{\sqrt{2}} \sin \frac{\tilde{B} + \tilde{Q}}{\sqrt{2}}\right] \equiv 0. \tag{5-10}$$

Proof. Let us assume (5-9) and prove (5-10). We have from (2-6) and (2-11) that

$$\tilde{Q}_t = -(\beta + i\alpha)^2 Q = -(Q_{xx} + Q^3).$$

Using (5-9), we have

$$B_x - Q_x - (\beta - i\alpha)(B + Q) \cos \frac{\tilde{B} + \tilde{Q}}{\sqrt{2}} = 0,$$

and

$$B_{xx} - Q_{xx} - (\beta - i\alpha)(B_x + Q_x) \cos \frac{\tilde{B} + \tilde{Q}}{\sqrt{2}} + (\beta - i\alpha) \frac{(B + Q)^2}{\sqrt{2}} \sin \frac{\tilde{B} + \tilde{Q}}{\sqrt{2}} = 0,$$

so that, using (5-9) and (5-8) once again,

$$\begin{aligned}
\tilde{B}_t - \tilde{Q}_t + (\beta - i\alpha)\Big[(B_x + Q_x)\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} + \frac{B^2+Q^2}{\sqrt{2}}\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\Big]& \\
= -(B_{xx}+B^3) + Q_{xx} + Q^3 + \Big[B_{xx} - Q_{xx} + (\beta - i\alpha)\frac{(B+Q)^2}{\sqrt{2}}\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\Big]& \\
+ (\beta - i\alpha)\frac{B^2+Q^2}{\sqrt{2}}\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}& \\
= Q^3 - B^3 + \sqrt{2}(\beta - i\alpha)(B^2+Q^2+BQ)\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}& \\
= Q^3 - B^3 + (B^2+Q^2+BQ)(B-Q) = 0.&
\end{aligned}$$

The proof of (5-9) is a tedious but straightforward computation which deeply exploits the nature of the breather and soliton profiles. For the proof of this result, see the Appendix. □

Corollary 5.2. *Under the assumptions of Lemma 5.1, for any $x \in \mathbb{R}$ one has*

$$\frac{B - \overline{Q}}{\sqrt{2}} - (\beta + i\alpha)\sin\frac{\tilde{B} + \tilde{\overline{Q}}}{\sqrt{2}} \equiv 0 \quad \textit{in } \mathbb{R},$$

where $\overline{Q}$ is the complex-valued soliton with parameters β and $-\alpha$.

In order to prove some results in the next section, we need several additional identities.

Corollary 5.3. *Under the assumptions of Lemma 5.1, for any $x \in \mathbb{R}$ one has*

$$\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = 1 - \frac{1}{2(\beta - i\alpha)}\int_{-\infty}^{x}(B^2 - Q^2) \quad \textit{and} \quad \lim_{x\to\pm\infty}\cos\Big(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\Big)(x) = \mp 1.$$

Remark. Note that both limits above make sense since, from (2-15) and (2-14), we have, for all x,

$$\begin{aligned}
\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} &= \cos\frac{\tilde{B}}{\sqrt{2}}\cos\frac{\tilde{Q}}{\sqrt{2}} - \sin\frac{\tilde{B}}{\sqrt{2}}\sin\frac{\tilde{Q}}{\sqrt{2}} \\
&= \frac{1}{\beta + i\alpha}\Big[\frac{Q_x}{Q}\cos\frac{\tilde{B}}{\sqrt{2}} - \frac{Q}{\sqrt{2}}\sin\frac{\tilde{B}}{\sqrt{2}}\Big].
\end{aligned}$$

In particular,

$$\lim_{\pm\infty}\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = \frac{1}{\beta + i\alpha}\times\mp(\beta + i\alpha) = \mp 1.$$

Proof. We multiply by $(1/\sqrt{2})(B+Q)$ in (5-9). We get

$$\frac{1}{2}(B^2 - Q^2) - (\beta - i\alpha)\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\times\frac{1}{\sqrt{2}}(B+Q) = 0,$$

i.e.,

$$\frac{1}{2}(B^2 - Q^2) + (\beta - i\alpha)\partial_x\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = 0.$$

From (2-1) and (5-1), one has

$$\lim_{x\to-\infty} \cos \frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = 1.$$

Therefore, after integration,

$$\frac{1}{2}\int_{-\infty}^{x} (B^2 - Q^2) + (\beta - i\alpha)\left[\cos \frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} - 1\right] = 0,$$

as desired. □

Lemma 5.4. $$\mathcal{M} := 2\beta\left[1 + \frac{\alpha(\beta \sin(2\alpha y_1) + \alpha \sinh(2\beta y_2))}{\alpha^2 + \beta^2 - \beta^2 \cos(2\alpha y_1) + \alpha^2 \cosh(2\beta y_2)}\right].$$

Proof. See, e.g., [Alejo and Muñoz 2013]. □

The following result is not difficult to prove:

Corollary 5.5. *We have*

$$\int_0^x \mathcal{M} = 2\beta x + \log(\alpha^2 + \beta^2 - \beta^2 \cos(2\alpha y_1) + \alpha^2 \cosh(2\beta y_2)) - L_0, \tag{5-11}$$

where

$$L_0 := \log(\alpha^2 + \beta^2 - \beta^2 \cos(2\alpha x_1) + \alpha^2 \cosh(2\beta x_2)).$$

Corollary 5.6. *Under the assumptions of Lemma 5.1, we have*

$$\begin{aligned} -(\beta - i\alpha)\int_0^x \cos \frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} \\ = (\beta + i\alpha)x + \log(\alpha^2 + \beta^2 - \beta^2 \cos(2\alpha y_1) + \alpha^2 \cosh(2\beta y_2)) - \log(1 + e^{2\beta y_2 + 2i\alpha y_1}) - L_0 + L_1, \end{aligned}$$

with L_0 and L_1 as defined in (5-11) *and* (2-19).

Proof. We have, from Corollaries 5.3 and 5.5 and (2-19),

$$\begin{aligned} &\int_0^x \cos \frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} \\ &= x - \frac{1}{\beta - i\alpha}\int_0^x (\mathcal{M} - \mathcal{N}) \\ &= x - \frac{1}{\beta - i\alpha}[2\beta x + \log(\alpha^2 + \beta^2 - \beta^2 \cos(2\alpha y_1) + \alpha^2 \cosh(2\beta y_2)) - \log(1 + e^{2(\beta y_2 + i\alpha y_1)}) - L_0 + L_1] \\ &= -\frac{1}{\beta - i\alpha}\big[(\beta + i\alpha)x + \log(\alpha^2 + \beta^2 - \beta^2 \cos(2\alpha y_1) + \alpha^2 \cosh(2\beta y_2)) \\ &\qquad\qquad - \log(1 + e^{2(\beta y_2 + i\alpha y_1)}) - L_0 + L_1\big], \end{aligned}$$

as desired. □

Corollary 5.7. *Assume that $x_1, x_2 \in \mathbb{R}$ do not satisfy* (2-9)*. Consider the function*

$$\mu(x; \alpha, \beta, x_1, x_2) := 2\sqrt{2}\alpha^2\beta^2 \frac{\cosh(\beta y_2)\cos(\alpha y_1) + i \sinh(\beta y_2)\sin(\alpha y_1)}{\alpha^2 \cosh^2(\beta y_2) + \beta^2 \sin^2(\alpha y_1)} = \beta \tilde{B}_1 - i\alpha \tilde{B}_2. \tag{5-12}$$

Then we have

$$\lim_{x\to\pm\infty} \mu(x) = 0 \tag{5-13}$$

$$\textit{and} \quad \mu_x = (\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)\mu. \tag{5-14}$$

Proof. Identity (5-13) is trivial. Let us prove (5-14). First of all, note that (cf. (2-7))

$$\beta\tilde{Q}_1 - i\alpha\tilde{Q}_2 \equiv 0. \tag{5-15}$$

On the other hand, from (5-9) we have

$$(\tilde{B}_1 - \tilde{Q}_1)_x - (\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)(\tilde{B}_1 + \tilde{Q}_1) = 0.$$

Similarly,

$$(\tilde{B}_2 - \tilde{Q}_2)_x - (\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)(\tilde{B}_2 + \tilde{Q}_2) = 0.$$

We then have

$$\begin{aligned}\mu_x &= (\beta\tilde{B}_1 - i\alpha\tilde{B}_2)_x \\ &= (\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)\mu + (\beta\tilde{Q}_1 - i\alpha\tilde{Q}_2)_x + (\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)(\beta\tilde{Q}_1 - i\alpha\tilde{Q}_2) \\ &= (\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)\mu.\end{aligned}$$

The proof is complete. □

Lemma 5.8. *Assume that* (2-9) *does not hold. Then* μ *defined in* (5-12) *has no zeroes*, *i.e.*, $|\mu(x)| > 0$ *for all* $x \in \mathbb{R}$.

Proof. From (5-12) we have $\mu(x) = 0$ if and only if $\cos(\beta y_1) = 0$ and $\sinh(\alpha y_2) = 0$, i.e., from (2-10) we have that (2-9) is satisfied. □

Now we consider the opposite case, where the sign in front of (5-14) is negative. We finish this section with the following result:

Lemma 5.9. *Assume that* (2-9) *does not hold. Then*

$$\mu^1(x;\alpha,\beta,x_1,x_2) := \frac{1}{\mu}(x;\alpha,\beta,x_1,x_2),$$

with μ *as defined in* (5-12), *is well-defined*, *has no zeroes and satisfies*

$$\lim_{x\to\pm\infty} |\mu^1(x)| = +\infty \quad \textit{and} \quad \mu^1_x = -(\beta - i\alpha)\cos\left(\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right)\mu^1.$$

Proof. This is a direct consequence of Corollary 5.7 and Lemma 5.8. □

6. Double Bäcklund transformation for mKdV

Assume that x_1 and x_2 do not satisfy (2-9). Consider the breather and soliton profiles B and Q defined in (5-2) and (2-5), which are well-defined by Lemma 2.2. From Lemma 5.1, we have the following result:

Lemma 6.1. *We have, for all* $x \in \mathbb{R}$,

$$G(B, Q, \tilde{B}_t, \tilde{Q}_t, \beta - i\alpha) = (0, 0).$$

Note that the previous identity can be extended by zero to the case where x_1 and x_2 satisfy (2-9), in such a form that now $G(B, Q, \tilde{B}_t, \tilde{Q}_t, \beta - i\alpha)$ as a function of x_1 and x_2 is well-defined and continuous everywhere in $\mathbb{R}^2$ (and identically zero).

Define (cf. (5-1)–(5-7)),

$$\begin{cases} \tilde{B}^0(x; \alpha, \beta) := \tilde{B}(x; \alpha, \beta, 0, 0), \\ \tilde{B}_t^0(x; \alpha, \beta) := \delta \tilde{B}_1(x; \alpha, \beta, 0, 0) + \gamma \tilde{B}_2(x; \alpha, \beta, 0, 0), \\ B^0(x; \alpha, \beta) := \partial_x \tilde{B}(x; \alpha, \beta, 0, 0). \end{cases} \tag{6-1}$$

Finally, for $z_a^0 \in H^1(\mathbb{R})$ we define

$$\omega_a^0 := -((z_a^0)_{xx} + (z_a^0)^3) \in H^{-1}(\mathbb{R}). \tag{6-2}$$

We will use Lemma 6.1 and apply Propositions 3.3 and 3.4 in a neighborhood of the complex soliton and the breather at time zero. Recall that, by Lemma 2.2, the complex soliton Q^0 is everywhere well-defined since (2-9) is not satisfied.

Lemma 6.2. *There exists* $\eta_0 > 0$ *and a constant* $C > 0$ *such that, for all* $0 < \eta < \eta_0$, *the following holds*: *Assume that* $z_a^0 \in H^1(\mathbb{R})$ *satisfies*

$$\|z_a^0\|_{H^1(\mathbb{R})} < \eta, \quad \omega_a^0 \text{ defined by } (6\text{-}2).$$

Then there exist unique $z_b^0 \in H^1(\mathbb{R}, \mathbb{C})$, $\omega_b^0 \in H^{-1}(\mathbb{R}; \mathbb{C})$ *and* $m_1 \in \mathbb{C}$ *of the form*

$$z_b^0(x) = z_b^0[z_a^0](x), \quad \omega_b^0(x) = \omega_b^0[z_a^0, \omega_a^0](x), \quad m_1 = m_1[z_a^0] := \beta - i\alpha + p^0$$

such that

$$\|z_b^0\|_{H^1(\mathbb{R};\mathbb{C})} + |p^0| \le C\eta,$$
$$\tilde{z}_a + \tilde{z}_b \in H^2(\mathbb{R}; \mathbb{C}),$$
$$\text{and} \quad G(B^0 + z_a^0, Q^0 + z_b^0, \tilde{B}_t^0 + \omega_a^0, \tilde{Q}_t^0 + \omega_b^0, m_1) \equiv (0, 0).$$

Proof. Let Q^0 and B^0 be the soliton and breather profiles defined in (4-1) and (6-1). We will apply Proposition 3.3 with

$$u_a^0 := B^0, \quad u_b^0 := Q^0, \quad v_a^0 := \tilde{B}_t^0, \quad v_b^0 := \tilde{Q}_t^0, \quad m^0 := \beta + i\alpha.$$

Clearly $\operatorname{Re} m^0 = \beta > 0$, so that (3-4) is satisfied. On the other hand, (3-5) is a consequence of Lemma 6.1. From (5-9), condition (3-6) reads

$$\sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} = \frac{(B^0-Q^0)}{\sqrt{2}(\beta-i\alpha)} \in H^1(\mathbb{R};\mathbb{C}).$$

Condition (3-7) is clearly satisfied (see (2-3) and (5-3)). From Corollary 5.7 we have

$$\mu^0 = \beta(\tilde{B}_1)^0 - i\alpha(\tilde{B}_2)^0.$$

Note that, from Lemmas 2.2 and 5.8, μ^0 has no zeroes in the complex plane and it is exponentially decreasing in space. Finally, let us show that

$$\int_{\mathbb{R}} \mu^0 \sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} = \frac{4i\alpha\beta}{\beta-i\alpha}.$$

First of all, we have from (5-15) that

$$\begin{aligned}[\beta(\tilde{B}_1)^0 - i\alpha(\tilde{B}_2)^0]\sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} \\ = [\beta(\tilde{B}_1+\tilde{Q}_1)^0 - i\alpha(\tilde{B}_2+\tilde{Q}_2)^0]\sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} + [-\beta(\tilde{Q}_1)^0 + i\alpha(\tilde{Q}_2)^0]\sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}}.\end{aligned}$$

Consequently,

$$[\beta(\tilde{B}_1)^0 - i\alpha(\tilde{B}_2)^0]\sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} = -\sqrt{2}\beta\partial_{x_1}\left[\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right]\Big|^0 + i\alpha\sqrt{2}\partial_{x_2}\left[\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right]\Big|^0.$$

Therefore, if R_1, $R_2 > 0$ are independent of x_1 and x_2,

$$\begin{aligned}\int_{-R_2}^{R_1} \mu^0 \sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} &= \sqrt{2}\int_{-R_2}^{R_1}\left\{-\beta\partial_{x_1}\left[\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right]\Big|^0 + i\alpha\partial_{x_2}\left[\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right]\Big|^0\right\} \\ &= \sqrt{2}\left\{-\beta\partial_{x_1}\int_{-R_2}^{R_1}\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} + i\alpha\partial_{x_2}\int_{-R_2}^{R_1}\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}}\right\}\Big|^0.\end{aligned}$$

Now we use Corollary 5.6: we have

$$\partial_{x_1}\int_{-R_2}^{R_1}\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = -\frac{1}{\beta-i\alpha}\left[\frac{2i\alpha e^{2\beta y_2+2i\alpha y_1}}{1+e^{2\beta y_2+2i\alpha y_1}} - \frac{2\alpha\beta^2\sin(2\alpha y_1)}{\alpha^2+\beta^2-\beta^2\cos(2\alpha y_1)+\alpha^2\cosh(2\beta y_2)}\right]\Big|_{-R_1}^{R_2}.$$

We have that

$$\lim_{R_1,R_2\to\infty}\partial_{x_1}\int_{-R_2}^{R_1}\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = -\frac{2i\alpha}{\beta-i\alpha}.$$

Similarly,

$$\partial_{x_2}\int_{-R_2}^{R_1}\cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = -\frac{1}{\beta-i\alpha}\left[\frac{2\beta e^{2\beta y_2+2i\alpha y_1}}{1+e^{2\beta y_2+2i\alpha y_1}} - \frac{2\alpha^2\beta\sinh(2\beta y_2)}{\alpha^2+\beta^2-\beta^2\cos(2\alpha y_1)+\alpha^2\cosh(2\beta y_2)}\right]\Big|_{-R_1}^{R_2}$$

and

$$\lim_{R_1,R_2\to\infty} \partial_{x_2} \int_{-R_2}^{R_1} \cos\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = -\frac{2\beta-4\beta}{\beta-i\alpha} = \frac{2\beta}{\beta-i\alpha}.$$

Adding the previous identities, we finally obtain

$$\int_{\mathbb{R}} \mu^0 \sin\frac{\tilde{B}^0+\tilde{Q}^0}{\sqrt{2}} = \left[\frac{2i\alpha\beta}{\beta-i\alpha}+\frac{2i\alpha\beta}{\beta-i\alpha}\right] = \frac{4i\alpha\beta}{\beta-i\alpha} \neq 0.$$

After applying Proposition 3.3, we are done. □

Now we address the following very important question: is the y_a^0 given in Lemma 4.1 real-valued for all $x \in \mathbb{R}$? In general, it seems that the answer is negative; however, if z_a^0 in Lemma 6.2 is real-valued, and z_b^0 from Lemma 6.2 satisfies (4-2), then the corresponding function y_a^0 given in Lemma 4.1 is also real-valued. This property is a consequence of a deep result called the *permutability theorem*, which we explain below.

First of all, from Lemma 6.2 we have

$$\frac{1}{\sqrt{2}}(B^0+z_a^0-Q^0-z_b^0) = (\beta-i\alpha+p^0)\sin\frac{\tilde{B}^0+\tilde{z}_a^0+\tilde{Q}^0+\tilde{z}_b^0}{\sqrt{2}} \tag{6-3}$$

for some small $p^0 \in \mathbb{C}$, and

$$\sin\frac{\tilde{B}^0+\tilde{z}_a^0+\tilde{Q}^0+\tilde{z}_b^0}{\sqrt{2}} \in H^1(\mathbb{R};\mathbb{C}). \tag{6-4}$$

Now, by taking η_0 smaller if necessary, such that $C\eta < \nu_0$ for all $0<\eta<\eta_0$, Lemma 4.1 also applies. We get

$$\frac{1}{\sqrt{2}}(Q^0+z_b^0-y_a^0) = (\beta+i\alpha+q^0)\sin\frac{\tilde{Q}^0+\tilde{z}_b^0+\tilde{y}_a^0}{\sqrt{2}}, \tag{6-5}$$

for some small q^0.

We need some auxiliary notation. Define

$$\beta_* := \beta+\operatorname{Re} p^0, \quad \alpha_* := \alpha-\operatorname{Im} p^0,$$

so that (compare with (4-6))

$$\beta-i\alpha+p^0 = \beta_*-i\alpha_*.$$

We also consider

$$\tilde{Q}_*^0 := \tilde{Q}(\,\cdot\,;-\alpha_*,\beta_*,0,0), \quad Q_*^0 := Q(\,\cdot\,;-\alpha_*,\beta_*,0,0).$$

Note that, since p^0 is small, we have that Q_*^0 and $\overline{Q}^0$ share the same properties, i.e., they are close enough. Indeed,

$$\|Q_*^0-\overline{Q}^0\|_{H^1(\mathbb{R};\mathbb{C})} \le C\eta. \tag{6-6}$$

Moreover, thanks to Lemma 2.4 applied to Q_*^0,

$$\frac{1}{\sqrt{2}}Q_*^0 = (\beta-i\alpha+p^0)\sin\frac{\tilde{Q}_*^0}{\sqrt{2}}.$$

Consequently, applying Proposition 3.4 starting at y_a^0 and using (6-6), we can define z_d^0 via the identity

$$\frac{1}{\sqrt{2}}(\bar{Q}^0 + z_d^0 - y_a^0) = (\beta - i\alpha + p^0)\sin\frac{\tilde{\bar{Q}}^0 + \tilde{z}_d^0 + \tilde{y}_a^0}{\sqrt{2}}. \tag{6-7}$$

Similarly, using (4-6) and (6-1) we define

$$(\tilde{B}^0)^* := \tilde{B}^0(\cdot\,;\alpha^*,\beta^*), \quad (B^0)^* := B(\cdot\,;\alpha^*,\beta^*), \tag{6-8}$$

so that from Lemma 5.1 we have

$$\frac{1}{\sqrt{2}}((B^0)^* - (Q^0)^*) = (\beta^* - i\alpha^*)\sin\frac{(\tilde{B}^0)^* + (\tilde{Q}^0)^*}{\sqrt{2}},$$

and applying Corollary 5.2 we get

$$\frac{1}{\sqrt{2}}((B^0)^* - \overline{(Q^0)^*}) = (\beta + i\alpha + q^0)\sin\frac{(\tilde{B}^0)^* + \overline{(\tilde{Q}^0)^*}}{\sqrt{2}}.$$

Using that

$$\|(B^0)^* - B^0\|_{H^1(\mathbb{R})} \le C\eta, \quad \|\overline{(Q^0)^*} - \bar{Q}^0\|_{H^1(\mathbb{R};\mathbb{C})} \le C\eta,$$

we can use Proposition 3.4 to obtain

$$\frac{1}{\sqrt{2}}(B^0 + z_c^0 - \bar{Q}^0 - z_d^0) = (\beta + i\alpha + q^0)\sin\frac{\tilde{B}^0 + \tilde{z}_c^0 + \tilde{\bar{Q}}^0 + \tilde{z}_d^0}{\sqrt{2}} \tag{6-9}$$

for some z_c^0 small. Note that the coefficients $(\beta - i\alpha + p^0)$ and $(\beta + i\alpha + q^0)$ were left fixed this time. Note additionally that z_d^0 and z_c^0 are bounded functions. Now we can state a permutability theorem [Lamb 1980, p. 246]. This is part of a more general result, standard in the mathematical physics literature; see [Wahlquist and Estabrook 1973] for a formal proof in the Korteweg–de Vries (KdV) case.

Theorem 6.3 (permutability theorem). *We have*

$$\tilde{z}_c^0 \equiv \tilde{z}_a^0. \tag{6-10}$$

In particular, z_c^0 is an H^1 real-valued function.

Proof. Define

$$u_0 := y_a^0, \quad u_1 := Q^0 + z_b^0, \quad u_2 := \bar{Q}^0 + z_d^0, \tag{6-11}$$

$$u_{12} := B^0 + z_a^0, \quad u_{21} := B^0 + z_c^0, \tag{6-12}$$

$$\text{and} \quad \kappa_1 := \beta + i\alpha + q^0, \quad \kappa_2 := \beta - i\alpha + p^0. \tag{6-13}$$

Since p^0 and q^0 are small quantities, we have $\kappa_1 \neq \kappa_2$, and both are nonzero complex numbers. Equations (6-5), (6-3), (6-7) and (6-9) now read

$$\frac{u_1 - u_0}{\sqrt{2}} = \kappa_1 \sin \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}}, \tag{6-14}$$

$$\frac{u_{12} - u_1}{\sqrt{2}} = \kappa_2 \sin \frac{\tilde{u}_{12} + \tilde{u}_1}{\sqrt{2}}, \tag{6-15}$$

$$\frac{u_2 - u_0}{\sqrt{2}} = \kappa_2 \sin \frac{\tilde{u}_2 + \tilde{u}_0}{\sqrt{2}},$$

$$\frac{u_{21} - u_2}{\sqrt{2}} = \kappa_1 \sin \frac{\tilde{u}_{21} + \tilde{u}_2}{\sqrt{2}}.$$

Note that u_1 and u_2 are obtained via the implicit function theorem and therefore there is an associated uniqueness property for solutions obtained in a small neighborhood of the breather. The idea is to prove that $\tilde{u}_{21} \equiv \tilde{u}_{12}$. Define $\tilde{u}_3$ via the identity

$$\frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}} = -\arctan\left[\frac{\kappa_1 - \kappa_2}{\kappa_1 + \kappa_2} \tan \frac{\tilde{u}_{12} - \tilde{u}_0}{2\sqrt{2}}\right]. \tag{6-16}$$

Whenever $u_1 = Q^0$, $u_{12} = B^0$, $u_0 = 0$, $\kappa_1 = \beta + i\alpha$ and $\kappa_2 = \beta - i\alpha$, we get from (1-2) that

$$\frac{\tilde{u}_3 - \tilde{Q}^0}{2\sqrt{2}} = -\arctan\left[i\frac{\alpha}{\beta} \tan \frac{\tilde{B}^0}{2\sqrt{2}}\right] = -\arctan\left(i\frac{\sin(\alpha x)}{\cosh(\beta x)}\right) = -\arctan \frac{e^{i\alpha x} - e^{-i\alpha x}}{e^{\beta x} + e^{-\beta x}}.$$

Therefore, using (2-1),

$$\begin{aligned}
\tilde{u}_3 &= 2\sqrt{2}\arctan(e^{(\beta+i\alpha)x}) - 2\sqrt{2}\arctan \frac{e^{i\alpha x} - e^{-i\alpha x}}{e^{\beta x} + e^{-\beta x}} \\
&= 2\sqrt{2}\arctan \frac{e^{(\beta+i\alpha)x} - (e^{i\alpha x} - e^{-i\alpha x})/(e^{\beta x} + e^{-\beta x})}{1 + e^{(\beta+i\alpha)x}(e^{i\alpha x} - e^{-i\alpha x})/(e^{\beta x} + e^{-\beta x})} \\
&= 2\sqrt{2}\arctan(e^{(\beta-i\alpha)x}) \\
&= \bar{\tilde{Q}}^0.
\end{aligned}$$

Consequently, under the smallness assumptions in (6-11)–(6-13) (the open character of these sets is essential) we have that $\tilde{u}_3$ is still well-defined on the real line with values in the complex plane, and it is close to $\bar{\tilde{Q}}^0$, as well as to $\tilde{u}_2$.

Let us find an equation for $\tilde{u}_3$. As usual, define $u_3 := (\tilde{u}_3)_x$. We claim that

$$\frac{u_3 - u_0}{\sqrt{2}} = \kappa_2 \sin \frac{\tilde{u}_3 + \tilde{u}_0}{\sqrt{2}}; \tag{6-17}$$

in other words, $\tilde{u}_3 \equiv \tilde{u}_2$. Similarly, if $\tilde{u}_4$ solves

$$\frac{\tilde{u}_2 - \tilde{u}_4}{2\sqrt{2}} = -\arctan\left[\frac{\kappa_1 - \kappa_2}{\kappa_1 + \kappa_2} \tan \frac{\tilde{u}_{21} - \tilde{u}_0}{2\sqrt{2}}\right], \tag{6-18}$$

then

$$\frac{u_4 - u_0}{\sqrt{2}} = \kappa_1 \sin \frac{\tilde{u}_4 + \tilde{u}_0}{\sqrt{2}},$$

which implies $\tilde{u}_4 \equiv \tilde{u}_1$. Finally, from (6-16) and (6-18) we have $\tilde{u}_{12} \equiv \tilde{u}_{21}$, which proves (6-10). Even better, we have[13]

$$\tan \frac{\tilde{u}_{12} - \tilde{u}_0}{2\sqrt{2}} = -\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2} \tan \frac{\tilde{u}_2 - \tilde{u}_1}{2\sqrt{2}}. \tag{6-19}$$

Now let us prove (6-17). First of all, denote

$$\ell := \frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}. \tag{6-20}$$

We have, from (6-16),

$$\frac{\tilde{u}_{12} - \tilde{u}_0}{\sqrt{2}} = -2 \arctan\left[\ell \tan \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}}\right],$$

so that

$$u_{12} - u_0 = \frac{-\ell(u_3 - u_1) \sec^2((\tilde{u}_3 - \tilde{u}_1)/(2\sqrt{2}))}{1 + \ell^2 \tan^2((\tilde{u}_3 - \tilde{u}_1)/(2\sqrt{2}))}.$$

We also check that

$$\sin \frac{\tilde{u}_{12} - \tilde{u}_0}{\sqrt{2}} = \frac{-2\ell \tan((\tilde{u}_3 - \tilde{u}_1)/(2\sqrt{2}))}{1 + \ell^2 \tan^2((\tilde{u}_3 - \tilde{u}_1)/(2\sqrt{2}))},$$

and

$$\cos \frac{\tilde{u}_{12} - \tilde{u}_0}{\sqrt{2}} = \frac{1 - \ell^2 \tan^2((\tilde{u}_3 - \tilde{u}_1)/(2\sqrt{2}))}{1 + \ell^2 \tan^2((\tilde{u}_3 - \tilde{u}_1)/(2\sqrt{2}))}.$$

Substituting in (6-15) and using (6-14) we obtain

$$\begin{aligned} &-\ell \frac{u_3 - u_1}{\sqrt{2}} \sec^2 \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}} \\ &= \kappa_1 \sin \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \left[1 + \ell^2 \tan^2 \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}}\right] + \kappa_2 \sin \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \left[1 - \ell^2 \tan^2 \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}}\right] - 2\ell\kappa_2 \cos \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \tan \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}}. \end{aligned}$$

Using (6-20) and (6-14), we have

$$\begin{aligned} &u_3 - u_0 - \sqrt{2}\kappa_1 \sin \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \\ &\quad = -\sqrt{2} \cos^2 \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}} \left[(\kappa_1 - \kappa_2) \sin \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \left(1 + \ell \tan^2 \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}}\right) - 2\kappa_2 \cos \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \tan \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}}\right], \end{aligned}$$

i.e., after some standard trigonometric simplifications,

$$u_3 - u_0 = \sqrt{2}\kappa_2 \sin \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \cos \frac{\tilde{u}_3 - \tilde{u}_1}{\sqrt{2}} + \sqrt{2}\kappa_2 \cos \frac{\tilde{u}_1 + \tilde{u}_0}{\sqrt{2}} \sin \frac{\tilde{u}_3 - \tilde{u}_1}{2\sqrt{2}} = \sqrt{2}\kappa_2 \sin \frac{\tilde{u}_3 + \tilde{u}_0}{\sqrt{2}},$$

as desired. □

Another consequence of the previous result is the following equivalent result:

[13]Note that this identity is well-defined at one particular set of functions, then extended by continuity.

Corollary 6.4. *We have*

$$z_d^0 \equiv \overline{z_b^0} \quad \textit{and} \quad p^0 = \overline{q^0}.$$

In other words, $\alpha^* = \alpha_*$ *and* $\beta^* = \beta_*$.

Proof. Note that $z_a^0 \equiv z_c^0$. From (6-9) we have

$$\frac{1}{\sqrt{2}}(B^0 + z_a^0 - Q^0 - \overline{z_d^0}) = (\beta - i\alpha + \overline{q^0}) \sin \frac{\tilde{B}^0 + \tilde{z}_a^0 + \tilde{Q}^0 + \overline{\tilde{z}_d^0}}{\sqrt{2}}.$$

The result follows from (6-3) and the uniqueness of z_b^0 and p^0 as implicit functions of z_a^0. □

The key result of this paper is the following surprising property:

Corollary 6.5. *The function* y_a^0 *is real-valued. Moreover, there is a small ball of data* z_a^0 *in* $H^1(\mathbb{R})$ *for which the corresponding data* z_b^0 *lies in an open set of* $H^1(\mathbb{R}; \mathbb{C})$.

Proof. The second statement is a consequence of the implicit function theorem. On the other hand, the first one is consequence of the permutability theorem. First of all, note that

$$\overline{\beta + i\alpha + q^0} = \beta - i\alpha + p^0 = \beta^* - i\alpha^*. \tag{6-21}$$

Now, from (6-19) we get

$$\tan \frac{B^0 + z_a^0 - y_a^0}{2\sqrt{2}} = -\frac{\beta + \operatorname{Re} p^0}{i(\alpha - \operatorname{Im} p^0)} \tan \frac{\tilde{Q}^0 + \tilde{z}_b^0 - \overline{\tilde{Q}}^0 - \overline{\tilde{z}_b}^0}{2\sqrt{2}},$$

so

$$\tan \frac{B^0 + z_a^0 - y_a^0}{2\sqrt{2}} = -\frac{\beta + \operatorname{Re} p^0}{(\alpha - \operatorname{Im} p^0)} \tanh \frac{\operatorname{Im}(\tilde{Q}^0 + \tilde{z}_b^0)}{\sqrt{2}},$$

from which we have that $y_a^0(x)$ is real-valued for all $x \in \mathbb{R}$. □

The main advantage of the double Bäcklund transformation is that now the dynamics of y_a^0 is real-valued. We apply Theorem 4.5 with the initial data z_b^0 to get a complex solution of mKdV, $u_b(t) = Q^*(t) + z_b(t)$ defined for all $t \neq t_k$ and satisfying (4-22).

Now we reconstruct $z_a(t)$. As in (6-8), let us define, using (5-1), (4-6) and (4-7),

$$\tilde{B}^*(t, x) := \tilde{B}(x; \alpha^*, \beta^*, \delta^* t + x_1, \gamma^* t + x_2) \tag{6-22}$$

and

$$B^*(t, x) = \partial_x \tilde{B}^*(t, x), \quad \tilde{B}_j^*(t, x) := \tilde{B}_j(x; \alpha^*, \beta^*, x_1, x_2)\big|_{x_1 = \delta^* t + x_1,\, x_2 = \gamma^* t + x_2}. \tag{6-23}$$

In other words, we recover the original breather in (1-2) with scaling parameters α^* and β^* and shifts x_1, x_2, provided they do not depend on time. Finally, as in (5-7) we define

$$\tilde{B}_t^*(t, x) := \delta \tilde{B}_1^*(t, x) + \gamma \tilde{B}_2^*(t, x).$$

Lemma 6.6. *Assume that $t \in \mathbb{R}$ is such that* (4-11) *holds. Then there are unique $z_a = z_a(t) \in H^1(\mathbb{R};\mathbb{C})$ and $w_a = w_a(t) \in H^{-1}(\mathbb{R};\mathbb{C})$ such that*

$$\tilde{z}_a + \tilde{z}_b \in H^2(\mathbb{R};\mathbb{C}), \tag{6-24}$$

$$\frac{1}{\sqrt{2}}(B^* + z_a - Q^* - z_b) = (\beta - i\alpha + p^0)\sin\frac{\tilde{B}^* + \tilde{z}_a + \tilde{Q}^* + \tilde{z}_b}{\sqrt{2}}, \tag{6-25}$$

where $\tilde{B}^*$ and B^* are defined in (6-22) *and* (6-23). *Moreover, we have*

$$\begin{aligned} 0 = \tilde{B}^*_t + w_a - \tilde{Q}^*_t - w_b \\ + (\beta - i\alpha + p^0)\Bigg[(B^*_x + (z_a)_x + Q^*_x + (z_b)_x)\cos\frac{\tilde{B}^* + \tilde{z}_a + \tilde{Q}^* + \tilde{z}_b}{\sqrt{2}} \\ + \frac{(B^* + z_a)^2 + (Q^* + z_b)^2}{\sqrt{2}}\sin\frac{\tilde{B}^* + \tilde{z}_a + \tilde{Q}^* + \tilde{z}_b}{\sqrt{2}}\Bigg] \end{aligned} \tag{6-26}$$

and, for all $t \neq t_k$,

$$\|z_a(t)\|_{H^1(\mathbb{R};\mathbb{C})} \le C\eta.$$

Proof. We apply Proposition 3.4 at the point

$$X^0 := (B^*, Q^*, \tilde{B}^*_t, \tilde{Q}^*_t, \beta - i\alpha + p^0),$$

because a slight variation of Lemma 6.1 shows that (compare with (6-21))

$$G(B^*, Q^*, \tilde{B}^*_t, \tilde{Q}^*_t, \beta - i\alpha + p^0) = (0, 0).$$

Since p^0 is small,

$$\operatorname{Re}(\beta - i\alpha + p^0) > 0.$$

On the other hand, (3-6) is a consequence of (6-4). Similarly, from (2-3) we get that (3-7) is satisfied. Finally, in order to ensure that (3-26) is clearly satisfied, we apply Lemma 5.9: we get

$$\mu^1 = \frac{1}{\mu^*}, \quad \text{where} \quad \mu^* := \beta^*\tilde{B}^*_1 - i\alpha^*\tilde{B}^*_2;$$

see Corollary 5.7 and (6-23). Then we conclude thanks to Proposition 3.4. □

Corollary 6.7. *The function $z_a(t)$ as defined in* (6-25) *is real-valued.*

Proof. The same proof as in Corollary 6.5 works mutatis mutandis, since now $y_a(t)$ is real-valued. □

Proposition 6.8. *For all $t \neq t_k$, $u_a = B^* + z_a$ is an H^1 real-valued solution to mKdV with initial data u_0. Therefore, by uniqueness,*[14] *$B^* + z_a \equiv u$.*

Proof. Since $u_b = Q^* + z_b$ solves mKdV, we use (6-25)–(6-26) and Theorem 3.5 to conclude. □

[14] Technically, what we need is a result about unconditional uniqueness, however, from [Kwon and Oh 2012] one can conclude that such a result is valid for mKdV on the line if we consider data with H^1 regularity.

7. Stability of breathers

We now prove Theorem 1.2. We assume that $u_0 \in H^1(\mathbb{R})$ satisfies (1-5) for some η small. Let $u \in C(\mathbb{R}; H^1(\mathbb{R}))$ be the — unique in a certain sense — associated solution of the Cauchy problem (1-1) with initial data $u(0) = u_0$. Finally, we recall the conserved quantities of mass (4-13) and energy (4-14).

Proof of Theorem 1.2. Consider $\varepsilon_0 > 0$ small but fixed, $A_0 > 1$ and $0 < \eta < \eta_0$ small. From Lemmas 6.2 and 6.6 the proof is not difficult, and we follow standard methods; see [Martel et al. 2002] for instance. Indeed, define the tubular neighborhood

$$\mathcal{V}(A_0, \eta) := \bigl\{U \in H^1(\mathbb{R}) \bigm| \inf_{\tilde{x}_1, \tilde{x}_2 \in \mathbb{R}} \|U - B(\,\cdot\,; \alpha, \beta, \tilde{x}_1, \tilde{x}_2)\| \le A_0 \eta\bigr\}. \tag{7-1}$$

Note that B represents here the breather profile defined in (5-2). The original breather $B(t)$ from (1-2) can be recovered using (6-22) as follows (there is a slight abuse of notation here, but it is easily understood):

$$B(t, x; \alpha, \beta, x_1, x_2) = B(x; \alpha, \beta, \delta t + x_1, \gamma t + x_2).$$

Clearly $u(t) \in \mathcal{V}(A_0, \eta)$ for small $t > 0$. Define the set

$$J_{\varepsilon_0} := \{t > 0 \mid |t - t_k| > \varepsilon_0 \ \text{ for all } \ k \in \mathbb{Z}\}.$$

We will prove that $u(t)$ is in $\mathcal{V}(A_0, \eta)$ for all $t \in J_{\varepsilon_0}$ provided A_0 is chosen large enough.

We argue by reductio ad absurdum. Assume that, for some $T_0 \in J_{\varepsilon_0}$, we have

$$\inf_{\tilde{x}_1, \tilde{x}_2 \in \mathbb{R}} \|u(T_0) - B(\,\cdot\,; \alpha, \beta, \tilde{x}_1, \tilde{x}_2)\|_{H^1(\mathbb{R})} = A_0 \eta, \tag{7-2}$$

and, for any $\delta > 0$ small, $\delta < \frac{1}{100}\varepsilon_0$, if $T_1 := T_0 + \delta$ then

$$\inf_{\tilde{x}_1, \tilde{x}_2 \in \mathbb{R}} \|u(T_1) - B(\,\cdot\,; \alpha, \beta, \tilde{x}_1, \tilde{x}_2)\|_{H^1(\mathbb{R})} > A_0 \eta. \tag{7-3}$$

We also assume that T_0 is the first positive time in J_{ε_0} with this property. We will show that, under this last assumption, after fixing $A_0 > 1$ large enough we will have

$$u(T_0) \in \mathcal{V}\bigl(\tfrac{1}{2} A_0, \eta\bigr), \tag{7-4}$$

which contradicts (7-2)–(7-3) and therefore proves the result for all positive times far from the points t_k.

First of all, by taking $\eta_0 > 0$ smaller if necessary, and $\eta \in (0, \eta_0)$, we can ensure that there are unique $x_1(t), x_2(t) \in \mathbb{R}$ defined on $[0, T_0]$ such that

$$z(t, x) := u(t, x) - B(x; \alpha, \beta, \delta t + x_1(t), \gamma t + x_2(t)) \tag{7-5}$$

satisfies

$$\int_{\mathbb{R}} z(t, x) B_1(x; \alpha, \beta, \delta t + x_1(t), \gamma t + x_2(t))\, dx = 0, \tag{7-6}$$

and

$$\int_{\mathbb{R}} z(t, x) B_2(x; \alpha, \beta, \delta t + x_1(t), \gamma t + x_2(t))\, dx = 0. \tag{7-7}$$

The directions B_1 and B_2 are defined in (5-5)–(5-6) (see [Alejo and Muñoz 2013] for a similar statement and its proof). Moreover, we have

$$\|z(0)\|_{H^1(\mathbb{R})} \lesssim \eta,$$

and similar estimates for $x_1(0)$ and $x_2(0)$, with constants not depending on large A_0. Therefore, (2-9) is not satisfied for $x_1(0)$ and $x_2(0)$. For the sake of simplicity, we can assume $x_1(0) = x_2(0) = 0$, otherwise we perform a shift in space and time on the solution to set them equal to zero.

Define $z_a^0 := z(0)$ and apply Lemma 6.2, and then Lemma 4.1 to the corresponding z_b^0 obtained from Lemma 6.2. We will obtain a *real-valued* seed y_a^0, small in $H^1(\mathbb{R})$. Note that the constants involved in each inversion do not depend on A_0. In particular, the differences between α and α^*, and β and β^*, are not dependent on A_0:

$$|\alpha-\alpha^*| + |\beta-\beta^*| \lesssim \eta. \tag{7-8}$$

Next, we let the mKdV equation evolve with initial data y_a^0. From Theorem 4.3 we have the bound (4-15) for the dynamics $y_a(t)$. On the other hand, the decomposition (7-6)–(7-7) implies that

$$|x_1'(t)| + |x_2'(t)| \lesssim A_0\eta, \tag{7-9}$$

from which the set of points where condition (4-11) is not satisfied is still a countable set of isolated points (see Lemma 4.2).

Now we are ready to apply Lemmas 4.4 and 6.6 with parameters α^*, β^* and shifts $x_1(t)$ and $x_2(t)$ in (4-8), (4-9) and (6-22)–(6-23). In that sense, we have chosen a unique set of parameters for each fixed time t, and the mKdV solution that we choose is the same as the original $u(t)$. Indeed, just notice that, at $t=0$, we have, from (4-18) at $t=0$ and (6-5),

$$\frac{1}{\sqrt{2}}(Q^*(0) + z_b(0) - y_a^0) = (\beta + i\alpha + q^0)\sin\frac{\tilde{Q}^*(0) + \tilde{z}_b(0) + \tilde{y}_a^0}{\sqrt{2}},$$
$$\frac{1}{\sqrt{2}}(Q^0 + z_b^0 - y_a^0) = (\beta + i\alpha + q^0)\sin\frac{\tilde{Q}^0 + \tilde{z}_b^0 + \tilde{y}_a^0}{\sqrt{2}}.$$

Using the uniqueness of the solution obtained by the implicit function theorem in a neighborhood of the base point, we have

$$z_b(0) = Q^0 - Q^*(0) + z_b^0 \sim z_b^0. \tag{7-10}$$

Now we use (6-25) at $t=0$ and (6-3):

$$\frac{1}{\sqrt{2}}(B^*(0) + z_a(0) - Q^*(0) - z_b(0)) = (\beta - i\alpha + p^0)\sin\frac{\tilde{B}^*(0) + \tilde{z}_a(0) + \tilde{Q}^*(0) + \tilde{z}_b(0)}{\sqrt{2}},$$

and

$$\frac{1}{\sqrt{2}}(B^0 + z_a^0 - Q^0 - z_b^0) = (\beta - i\alpha + p^0)\sin\frac{\tilde{B}^0 + \tilde{z}_a^0 + \tilde{Q}^0 + \tilde{z}_b^0}{\sqrt{2}}.$$

From (7-10), we have

$$\frac{1}{\sqrt{2}}(B^*(0)+z_a(0)-Q^0-z_b^0)=(\beta-i\alpha+p^0)\sin\frac{\tilde{B}^*(0)+\tilde{z}_a(0)+\tilde{Q}^0+\tilde{z}_b^0}{\sqrt{2}}.$$

Once again, since B^0 and $B^*(0)$ are close, using the uniqueness of the solution obtained via the implicit function Theorem, we conclude that

$$B^*(0)+z_a(0)=B^0+z_a^0.$$

Since both initial data are the same, we conclude that the solution obtained via the Bäcklund transformation is $u(t)$.

Note that the constants involved in the inversions are not dependent on A_0. We finally get

$$\sup_{|t-t_k|\geq\varepsilon_0}\|u(t)-B^*(t)\|_{H^1(\mathbb{R})}\leq C_0\eta, \tag{7-11}$$

where

$$B^*(t,x):=B(x;\alpha^*,\beta^*,\delta^*t+x_1(t),\gamma^*t+x_2(t)).$$

Finally, from (7-8) and after redefining the shift parameters and choosing $t=T_0$, we get the desired conclusion since, for A_0 large enough, we have $C_0\leq\frac{1}{2}A_0$ and (7-4) is proved.

Now we deal with the remaining case, $t\sim t_k$. Fix $k\in\mathbb{Z}$. Note that $z_a=u-B^*$ satisfies the equation

$$(z_a)_t+[(z_a)_{xx}+3(B^*)^2z_a+3B^*z_a^2+z_a^3]_x+x_1'(t)B_1^*+x_2'(t)B_2^*=0 \tag{7-12}$$

in the H^1-sense. In what follows, we will prove that, maybe taking ε_0 smaller but independent of k, we have

$$\sup_{|t-t_k|\leq\varepsilon_0}\|u(t)-B^*(t)\|_{H^1(\mathbb{R})}\leq 4A_0\eta. \tag{7-13}$$

Since A_0 grows with ε_0 small, this implies that, after choosing η_0 smaller if necessary, such an operation can be performed without any risk.

In what follows, we assume that there is $T^*\in(t_k-\varepsilon_0,t_k+\varepsilon_0]$ such that, for all $t\in[t_k-\varepsilon_0,T^*]$,

$$\|z_a(t)\|_{H^1(\mathbb{R})}\leq 4A_0\eta, \tag{7-14}$$

and T^* is maximal in the sense of the above definition (i.e., there is no $T^{**}>T^*$ satisfying the previous property). If $T^*=t_k+\varepsilon_0$, there is nothing to prove and (7-13) holds.

Assume $T^*<t_k+\varepsilon_0$. Now we consider the quantity

$$\frac{1}{2}\int_{\mathbb{R}}z_a^2(t),\qquad t\in[t_0-\varepsilon_0,T^*].$$

We have, from (7-12),

$$\partial_t\frac{1}{2}\int_{\mathbb{R}}z_a^2(t)=\int_{\mathbb{R}}(z_a)_x[3(B^*)^2z_a+3B^*z_a^2+z_a^3](t)+x_1'(t)\int_{\mathbb{R}}z_a(t)B_1^*+x_2'(t)\int_{\mathbb{R}}z_a(t)B_2^*.$$

Using (7-14) and (7-9), we have, for some — explicit — fixed constant $C > 0$ depending only on α, β, and η_0 even smaller if necessary,

$$\left|\partial_t \frac{1}{2}\int_{\mathbb{R}} z_a^2(t)\right| \le C A_0^2\eta^2.$$

After integration in time and using (7-11), we have

$$\int_{\mathbb{R}} z_a^2(T^*) \le \int_{\mathbb{R}} z_a^2(t_0-\varepsilon_0) + C\varepsilon_0 A_0^2\eta^2 \le 1.9A_0^2\eta^2,$$

if ε_0 is small but fixed. A similar estimate can be obtained for $\|(z_a)_x(t)\|_{H^1(\mathbb{R})}$ by proving an estimate of the form

$$\left|\partial_t \frac{1}{2}\int_{\mathbb{R}} (z_a)_x^2(t)\right| \le C A_0^2\eta^2.$$

Therefore, estimate (7-14) has been bootstrapped, which implies that $T^* = t_0+\varepsilon_0$. Note that the estimates do not depend on k, but only on the length of the intervals, which is about ε_0.[15]

We conclude that there is $\tilde{A}_0 > 0$ fixed such that

$$\sup_{t\in\mathbb{R}} \|u(t) - B^*(t)\|_{H^1(\mathbb{R})} \le \tilde{A}_0\eta.$$

Finally, estimates (1-6) and (1-7) are obtained from (7-9), using the fact that α^* and β^* are close to α and β in terms of $C\eta$. The proof is complete. □

Remark. From the proof and the results in [Colliander et al. 2003], it is easy to show that the evolution of breathers can be estimated in a polynomial form in time for any $s > \frac{1}{4}$, however, in order to make things simpler, we will not address this issue.

Corollary 7.1. *We have, for all $t \neq t_k$,*

$$\frac{1}{2}\int_{\mathbb{R}} (B^*+z_a)^2(t) = \frac{1}{2}\int_{\mathbb{R}} (Q^*+z_b)^2(t) + 2(\beta^* - i\alpha^*) = M[y_a^0] + 4\beta^*.$$

Moreover, this identity can be extended to any $t \in \mathbb{R}$.

Proof. In the same way as Corollary 4.6. □

Finally, we recall that $\gamma^* = 3(\alpha^*)^2 - (\beta^*)^2$ and $E[u] = \frac{1}{2}\int_{\mathbb{R}} u_x^2 - \frac{1}{4}\int_{\mathbb{R}} u^4$.

Corollary 7.2. *Assume that $t \neq t_k$ for all $k \in \mathbb{Z}$. Then we have*

$$E[B^*+z_a](t) = E[Q^*+z_b](t) - \tfrac{4}{3}(\beta^* - i\alpha^*)^3 = E[y_a^0] + \tfrac{4}{3}\beta^*\gamma^*.$$

Finally, this quantity can be extended in a continuous form to every $t \in \mathbb{R}$.

Proof. In the same way as Corollary 4.7. □

[15]Note that an argument involving the uniform continuity of the mKdV flow will not work in this particular case since the sequence of times (t_k) is unbounded.

8. Asymptotic stability

We finally prove Theorem 1.4. Note that, for some $c_0 > 0$ depending on $\eta > 0$,

$$\lim_{t\to+\infty} \|y_a(t)\|_{H^1(x\geq c_0 t)} = 0. \tag{8-1}$$

This result can be obtained by adapting the proof for the soliton case in [Martel and Merle 2005]. Indeed, consider

$$\phi(x) := \frac{K}{\pi}\arctan(e^{x/K}), \quad K > 0,$$

so that

$$\lim_{-\infty}\phi = 0, \quad \lim_{+\infty}\phi = 1, \quad \phi''' \leq \frac{1}{K^2}\phi', \quad \phi' > 0 \text{ on } \mathbb{R}. \tag{8-2}$$

Fix c_0, $t_0 > 0$. Consider the quantities

$$\begin{aligned} I(t) &:= \frac{1}{2}\int_{\mathbb{R}} y_a^2(t)\phi\big(x - c_0t_0 + \tfrac{1}{2}c_0(t_0 - t)\big),\\ J(t) &:= \int_{\mathbb{R}} \big[\tfrac{1}{2}(y_a)_x^2(t) - \tfrac{1}{4}y_a^4(t) + \tfrac{1}{2}y_a^2(t)\big]\phi\big(x - c_0t_0 + \tfrac{1}{2}c_0(t_0 - t)\big). \end{aligned}$$

It is not difficult to see that

$$I'(t) = -\frac{1}{4}c_0\int_{\mathbb{R}} y_a^2\phi'(t) + \frac{1}{2}\int_{\mathbb{R}} y_a^2\phi'''(t) - \frac{3}{2}\int_{\mathbb{R}}(y_a)_x^2\phi'(t) + \frac{3}{4}\int_{\mathbb{R}} y_a^4\phi'(t),$$

so that, using (8-2), and if $c_0 > 0$ is small (and, depending on η, even smaller if necessary),

$$I'(t) \leq 0.$$

We then have

$$I(t_0) \leq I(0) = \frac{1}{2}\int_{\mathbb{R}} y_a^2(0)\phi(x - c_0t_0)$$

and

$$\lim_{t\to+\infty} I(t) = 0.$$

A similar result holds for $J(t)$, which proves (8-1).

Note that $\tilde{z}_b + \tilde{y}_a \in H^2(\mathbb{R};\mathbb{C})$ (see (4-17)). In what follows, we will prove that this function satisfies better estimates than y_a and z_b if x is large.

Fix $t \neq t_k$ large with $|t - t_k| \geq \varepsilon_0$. We use the notation

$$\tilde{z}_c := \tilde{y}_a + \tilde{z}_b. \tag{8-3}$$

From (3-29) we have

$$\|\tilde{z}_c(t)\|_{H^2(\mathbb{R};\mathbb{C})} \leq C\nu$$

with $C = C(\varepsilon_0)$ independent of time. From the Bäcklund transformation (4-18) we obtain

$$\begin{aligned}(\tilde z_c)_x - 2y_a &= \sqrt{2}(\beta + i\alpha + q^0)\left[\sin\frac{\tilde Q^* + \tilde z_c}{\sqrt2} - \sin\frac{\tilde Q^*}{\sqrt2}\right]\\ &= \sqrt{2}(\beta + i\alpha + q^0)\left[\sin\frac{\tilde Q^*}{\sqrt2}\left\{\cos\frac{\tilde z_c}{\sqrt2} - 1\right\} + \sin\frac{\tilde z_c}{\sqrt2}\,\cos\frac{\tilde Q^*}{\sqrt2}\right]\\ &= Q^*\left\{\cos\frac{\tilde z_c}{\sqrt2} - 1\right\} + \sqrt2\sin\left(\frac{\tilde z_c}{\sqrt2}\right)\frac{Q^*_x}{Q^*}.\end{aligned}$$

Assume now that $x > c_0t/2$. Then we have, for some fixed constant $c > 0$,

$$\left|\frac{Q^*_x}{Q^*} + m\right| \le e^{-cx}, \quad m = \beta + i\alpha + q^0 = \beta^* + i\alpha^*,$$

and

$$(\tilde z_c)_x + m\tilde z_c = g,$$

where

$$g := Q^*\left\{\cos\frac{\tilde z_c}{\sqrt2} - 1\right\} + \sqrt2\left\{\sin\frac{\tilde z_c}{\sqrt2} - \frac{\tilde z_c}{\sqrt2}\right\}\frac{Q^*_x}{Q^*} + \tilde z_c\left\{\frac{Q^*_x}{Q^*} + m\right\} + 2y_a.$$

Solving the previous ODE, we get

$$\tilde z_c(t,x) = \tilde z_c\big(t, \tfrac12 c_0t\big)e^{-m(x-c_0t/2)} + \int_{c_0t/2}^x g(t,s)e^{-m(x-s)}\,ds,$$

so that

$$|\tilde z_c(t,x)| \lesssim \big|\tilde z_c\big(t, \tfrac12 c_0t\big)\big|e^{-\beta^*(x-c_0t/2)} + \int_{c_0t/2}^x |g(t,s)|e^{-\beta^*(x-s)}\,ds.$$

From Young's inequality we get

$$\|\tilde z_c(t)\|_{L^2(x\ge c_0t)} \lesssim \big|\tilde z_c\big(t, \tfrac12 c_0t\big)\big|e^{-\beta^*c_0t/2} + \|g(t)\|_{L^2(x\ge c_0t)}e^{-\beta^*c_0t}.$$

Clearly,

$$\big|\tilde z_c\big(t, \tfrac12 c_0t\big)\big| \lesssim \|\tilde z_c(t)\|_{H^1(\mathbb{R};\mathbb{C})} \le C\nu, \quad \|g(t)\|_{L^2(x\ge c_0t)} \le C\nu^2 + C\nu e^{-ct} + o(1).$$

Passing to the limit, we obtain that, for all $T_n \to +\infty$ with $|T_n - t_k| \ge \varepsilon_0$ for all n and k,

$$\lim_{n\to+\infty}\|\tilde z_c(T_n)\|_{L^2(x\ge c_0T_n)} = 0.$$

A similar result can be obtained for z_c and $(z_c)_x$. From (8-3), we get

$$\lim_{n\to+\infty}\|z_b(T_n)\|_{H^1(x\ge c_0T_n)} = 0. \tag{8-4}$$

Finally, we repeat the same strategy with (6-25) and (6-24) to obtain

$$\lim_{t\to+\infty}\|z_a(T_n)\|_{H^1(x\ge c_0T_n)} = 0.$$

Note that, since the flow map is continuous in time with values in H^1, we can extend the result to any sequence $T_n \to +\infty$ by choosing an $\varepsilon_0 > 0$ smaller but still independent of k.

Appendix A: Proof of Lemma 5.1

We will use the specific character of the breather and soliton profiles. Since (2-9) does not hold, both $\tilde{Q}$ and Q are well-defined everywhere. We have

$$\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = \sin(2(\arctan\Theta_1+\arctan\Theta_2)),$$

where, from (2-1) and (5-1), $\Theta_2 := e^{\beta y_2+i\alpha y_1}$ and $\Theta_1 := \dfrac{\beta\sin(\alpha y_1)}{\alpha\cosh(\beta y_2)}$. The expression in the previous display equals

$$\begin{aligned}
&2\big[\sin(\arctan\Theta_1)\cos(\arctan\Theta_2)+\sin(\arctan\Theta_2)\cos(\arctan\Theta_1)\big]\\
&\qquad\qquad\times\big[\cos(\arctan\Theta_1)\cos(\arctan\Theta_2)-\sin(\arctan\Theta_1)\sin(\arctan\Theta_2)\big]\\
&=2\big[\tan(\arctan\Theta_1)\cos^2(\arctan\Theta_1)\cos^2(\arctan\Theta_2)-\sin^2(\arctan\Theta_1)\tan(\arctan\Theta_2)\cos^2(\arctan\Theta_2)\\
&\quad+\cos^2(\arctan\Theta_1)\tan(\arctan\Theta_2)\cos^2(\arctan\Theta_2)-\sin^2(\arctan\Theta_2)\tan(\arctan\Theta_1)\cos^2(\arctan\Theta_1)\big].
\end{aligned}$$

Since $\sin^2(\arctan z)=\dfrac{z^2}{1+z^2}$ and $\cos^2(\arctan z)=\dfrac{1}{1+z^2}$, we have

$$\sin\frac{\tilde{B}+\tilde{Q}}{\sqrt{2}} = \frac{2(\Theta_1-\Theta_1^2\Theta_2+\Theta_2-\Theta_2^2\Theta_1)}{(1+\Theta_1^2)(1+\Theta_2^2)}. \tag{A-1}$$

On the other hand,

$$\frac{1}{\sqrt{2}}(B-Q)=2\partial_x(\arctan\Theta_1-\arctan\Theta_2)=2\left(\frac{\Theta_{1,x}}{1+\Theta_1^2}-\frac{\Theta_{2,x}}{1+\Theta_2^2}\right)=2\frac{(1+\Theta_2^2)\Theta_{1,x}-(1+\Theta_1^2)\Theta_{2,x}}{(1+\Theta_1^2)(1+\Theta_2^2)}.$$

Hence, collecting terms and factoring, from (5-9) we are led to prove that

$$(1+\Theta_2^2)\Theta_{1,x}-(1+\Theta_1^2)\Theta_{2,x}-(\beta-i\alpha)(\Theta_1-\Theta_1^2\Theta_2+\Theta_2-\Theta_2^2\Theta_1)=0. \tag{A-2}$$

Now we perform some computations. We have, from (2-1),

$$\Theta_{2,x}=(\beta+i\alpha)\Theta_2, \tag{A-3}$$

$$\alpha(\beta+i\alpha\Theta_1^2)\cosh^2(\beta y_2)=\beta(\alpha\cosh^2(\beta y_2)+i\beta\sin^2(\alpha y_1)) \tag{A-4}$$

and

$$\Theta_{1,x}=\left(\frac{\beta\sin(\alpha y_1)}{\alpha\cosh(\beta y_2)}\right)_x=\frac{\alpha\beta\cos(\alpha y_1)\cosh(\beta y_2)-\beta^2\sin(\alpha y_1)\sinh(\beta y_2)}{\alpha\cosh^2(\beta y_2)},$$

so that

$$\Theta_{1,x}-(\beta-i\alpha)\Theta_1=\beta\left[\frac{\alpha e^{i\alpha y_1}\cosh(\beta y_2)-\beta e^{\beta y_2}\sin(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right] \tag{A-5}$$

and

$$\begin{aligned}
[\Theta_{1,x}+(\beta-i\alpha)\Theta_1]\Theta_2^2&=\beta\left[\frac{\alpha e^{-i\alpha y_1}\cosh(\beta y_2)+\beta e^{-\beta y_2}\sin(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right]e^{2(\beta y_2+i\alpha y_1)}\\
&=\beta\Theta_2\left[\frac{\alpha e^{\beta y_2}\cosh(\beta y_2)+\beta e^{i\alpha y_1}\sin(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right].
\end{aligned} \tag{A-6}$$

Using (A-3), (A-4), (A-5) and (A-6) we have that the left-hand side of (A-2) is

$$
\begin{aligned}
&(1+\Theta_2^2)\Theta_{1,x}-2(\beta+i\alpha\Theta_1^2)\Theta_2-(\beta-i\alpha)(1-\Theta_2^2)\Theta_1\\
&\qquad=[\Theta_{1,x}-(\beta-i\alpha)\Theta_1]+[\Theta_{1,x}+(\beta-i\alpha)\Theta_1]\Theta_2^2-2(\beta+i\alpha\Theta_1^2)\Theta_2\\
&\qquad=\beta\left[\frac{\alpha e^{i\alpha y_1}\cosh(\beta y_2)-\beta e^{\beta y_2}\sin(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right]\\
&\qquad\qquad+\beta\Theta_2\left[\frac{\alpha e^{\beta y_2}\cosh(\beta y_2)+\beta e^{i\alpha y_1}\sin(\alpha y_1)-2\alpha\cosh^2(\beta y_2)-2i\beta\sin^2(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right]\\
&\qquad=\beta\left[\frac{\alpha e^{i\alpha y_1}\cosh(\beta y_2)-\beta e^{\beta y_2}\sin(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right]+\beta\Theta_2\left[\frac{-\alpha e^{-\beta y_2}\cosh(\beta y_2)+\beta e^{-i\alpha y_1}\sin(\alpha y_1)}{\alpha\cosh^2(\beta y_2)}\right]\\
&\qquad=0,
\end{aligned}
$$

which proves (A-2).

Acknowledgments

Alejo acknowledges the financial support from the University of the Basque Country-EHU. Part of this work was done while Muñoz was L. E. Dickson Instructor at the University of Chicago. He would like to express his gratitude with the members of the Department of Mathematics, in particular Professors Carlos Kenig and Wilhelm Schlag. We also thank Herbert Koch and Yvan Martel for several enlightening discussions which improved the quality of this paper, and a gap in the first version of this one. Finally, we deeply thank both referees for several comments, corrections and criticisms that led to an important improvement of our first manuscript.

References

[Ablowitz and Clarkson 1991] M. J. Ablowitz and P. A. Clarkson, *Solitons, nonlinear evolution equations and inverse scattering*, London Mathematical Society Lecture Note Series **149**, Cambridge University Press, 1991. MR 93g:35108 Zbl 0762.35001

[Alejo 2012] M. A. Alejo, "Geometric breathers of the mKdV equation", *Acta Appl. Math.* **121** (2012), 137–155. MR 2966969 Zbl 1254.35201

[Alejo and Muñoz 2013] M. A. Alejo and C. Muñoz, "Nonlinear stability of mKdV breathers", *Comm. Math. Phys.* **324**:1 (2013), 233–262. MR 3116324 Zbl 1280.35123

[Alejo et al. 2013] M. A. Alejo, C. Muñoz, and L. Vega, "The Gardner equation and the L^2-stability of the N-soliton solution of the Korteweg–de Vries equation", *Trans. Amer. Math. Soc.* **365**:1 (2013), 195–212. MR 2984057 Zbl 1278.35209

[Aubry 1997] S. Aubry, "Breathers in nonlinear lattices: existence, linear stability and quantization", *Phys. D* **103**:1-4 (1997), 201–250. MR 99b:34067 Zbl 1194.34059

[Benjamin 1972] T. B. Benjamin, "The stability of solitary waves", *Proc. Roy. Soc. London Ser. A* **328** (1972), 153–183. MR 49 #3348

[Birnir et al. 1994] B. Birnir, H. P. McKean, and A. Weinstein, "The rigidity of sine-Gordon breathers", *Comm. Pure Appl. Math.* **47**:8 (1994), 1043–1051. MR 95h:35195 Zbl 0814.35112

[Bona et al. 1987] J. L. Bona, P. E. Souganidis, and W. A. Strauss, "Stability and instability of solitary waves of Korteweg–de Vries type", *Proc. Roy. Soc. London Ser. A* **411**:1841 (1987), 395–412. MR 88m:35128 Zbl 0648.76005

[Bona et al. 2013] J. L. Bona, S. Vento, and F. B. Weissler, "Singularity formation and blowup of complex-valued solutions of the modified KdV equation", *Discrete Contin. Dyn. Syst.* **33**:11-12 (2013), 4811–4840. MR 3060816 Zbl 1290.35221

[Buckmaster and Koch 2014] T. Buckmaster and H. Koch, "The Korteweg–de Vries equation at H^{-1} regularity", *Ann. Inst. H. Poincaré Anal. Non Linéaire* (online publication June 2014).

[Christ et al. 2012] M. Christ, J. Holmer, and D. Tataru, "Low regularity bounds for mKdV", preprint, 2012. arXiv 1207.6738

[Colliander et al. 2003] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Sharp global well-posedness for KdV and modified KdV on $\mathbb{R}$ and $\mathbb{T}$", *J. Amer. Math. Soc.* **16**:3 (2003), 705–749. MR 2004c:35352 Zbl 1025.35025

[Deift and Zhou 1993] P. Deift and X. Zhou, "A steepest descent method for oscillatory Riemann–Hilbert problems: asymptotics for the mKdV equation", *Ann. of Math.* (2) **137**:2 (1993), 295–368. MR 94d:35143 Zbl 0771.35042

[Denzler 1993] J. Denzler, "Nonpersistence of breather families for the perturbed sine Gordon equation", *Comm. Math. Phys.* **158**:2 (1993), 397–430. MR 95c:35210 Zbl 0802.35132

[Gorria et al. 2013] C. Gorria, M. A. Alejo, and L. Vega, "Discrete conservation laws and the convergence of long time simulations of the mkdv equation", *J. Comput. Phys.* **235** (2013), 274–285. MR 3017597 Zbl 1291.65279

[Grillakis et al. 1987] M. Grillakis, J. Shatah, and W. Strauss, "Stability theory of solitary waves in the presence of symmetry, I", *J. Funct. Anal.* **74**:1 (1987), 160–197. MR 88g:35169 Zbl 0656.35122

[Hirota 1972] R. Hirota, "Exact solution of the modified Korteweg–de Vries equation for multiple collisions of solitons", *J. Phys. Soc. Japan* **33**:5 (1972), 1456–1458.

[Hoffman and Wayne 2013] A. Hoffman and C. E. Wayne, "Orbital stability of localized structures via Bäcklund transformations", *Differential Integral Equations* **26**:3-4 (2013), 303–320. MR 3059166 Zbl 1289.35225

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness and scattering results for the generalized Korteweg–de Vries equation via the contraction principle", *Comm. Pure Appl. Math.* **46**:4 (1993), 527–620. MR 94h:35229 Zbl 0808.35128

[Kenig et al. 2001] C. E. Kenig, G. Ponce, and L. Vega, "On the ill-posedness of some canonical dispersive equations", *Duke Math. J.* **106**:3 (2001), 617–633. MR 2002c:35265 Zbl 1034.35145

[Kwon and Oh 2012] S. Kwon and T. Oh, "On unconditional well-posedness of modified KdV", *Int. Math. Res. Not.* **2012**:15 (2012), 3509–3534. MR 2959040 Zbl 1248.35186

[Lamb 1980] G. L. Lamb, Jr., *Elements of soliton theory*, John Wiley & Sons, New York, 1980. MR 82f:35165 Zbl 0445.35001

[Lax 1968] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves", *Comm. Pure Appl. Math.* **21** (1968), 467–490. MR 38 #3620 Zbl 0162.41103

[Maddocks and Sachs 1993] J. H. Maddocks and R. L. Sachs, "On the stability of KdV multi-solitons", *Comm. Pure Appl. Math.* **46**:6 (1993), 867–901. MR 94h:35232 Zbl 0795.35107

[Martel and Merle 2002] Y. Martel and F. Merle, "Blow up in finite time and dynamics of blow up solutions for the L^2-critical generalized KdV equation", *J. Amer. Math. Soc.* **15**:3 (2002), 617–664. MR 2003c:35142 Zbl 0996.35064

[Martel and Merle 2005] Y. Martel and F. Merle, "Asymptotic stability of solitons of the subcritical gKdV equations revisited", *Nonlinearity* **18**:1 (2005), 55–80. MR 2006i:35319 Zbl 1064.35171

[Martel et al. 2002] Y. Martel, F. Merle, and T.-P. Tsai, "Stability and asymptotic stability in the energy space of the sum of N solitons for subcritical gKdV equations", *Comm. Math. Phys.* **231**:2 (2002), 347–373. MR 2003j:35280

[Merle 1992] F. Merle, "On uniqueness and continuation properties after blow-up time of self-similar solutions of nonlinear Schrödinger equation with critical exponent and critical mass", *Comm. Pure Appl. Math.* **45**:2 (1992), 203–254. MR 93e:35104 Zbl 0767.35084

[Merle and Vega 2003] F. Merle and L. Vega, "L^2 stability of solitons for KdV equation", *Int. Math. Res. Not.* **2003**:13 (2003), 735–753. MR 2004k:35330 Zbl 1022.35061

[Miura et al. 1968] R. M. Miura, C. S. Gardner, and M. D. Kruskal, "Korteweg–de Vries equation and generalizations, II: Existence of conservation laws and constants of motion", *J. Mathematical Phys.* **9** (1968), 1204–1209. MR 40 #6042b Zbl 0283.35019

[Mizumachi and Pelinovsky 2012] T. Mizumachi and D. Pelinovsky, "Bäcklund transformation and L^2-stability of NLS solitons", *Int. Math. Res. Not.* **2012**:9 (2012), 2034–2067. MR 2920823 Zbl 1239.35148

[Pego and Weinstein 1994] R. L. Pego and M. I. Weinstein, "Asymptotic stability of solitary waves", *Comm. Math. Phys.* **164**:2 (1994), 305–349. MR 95h:35209 Zbl 0805.35117

[Schuur 1986] P. C. Schuur, *Asymptotic analysis of soliton problems: an inverse scattering approach*, Lecture Notes in Mathematics **1232**, Springer, Berlin, 1986. MR 88e:35003 Zbl 0643.35003

[Soffer and Weinstein 1999] A. Soffer and M. I. Weinstein, "Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations", *Invent. Math.* **136**:1 (1999), 9–74. MR 2000k:37119 Zbl 0910.35107

[Wadati 1973] M. Wadati, "The modified Korteweg–de Vries equation", *J. Phys. Soc. Japan* **34** (1973), 1289–1296. MR 51 #7472

[Wahlquist and Estabrook 1973] H. D. Wahlquist and F. B. Estabrook, "Bäcklund transformation for solutions of the Korteweg–de Vries equation", *Phys. Rev. Lett.* **31** (1973), 1386–1390. MR 54 #3191

[Weinstein 1986] M. I. Weinstein, "Lyapunov stability of ground states of nonlinear dispersive evolution equations", *Comm. Pure Appl. Math.* **39**:1 (1986), 51–67. MR 87f:35023 Zbl 0594.35005

Received 13 Feb 2014. Revised 4 Dec 2014. Accepted 9 Feb 2015.

MIGUEL A. ALEJO: miguel.alejo@ufsc.br
Instituto Nacional de Matemática Pura e Aplicada, 22081-010 Rio de Janeiro, RJ, Brazil

CLAUDIO MUÑOZ: claudio.munoz@math.u-psud.fr
CNRS and Laboratoire de Mathématiques d'Orsay UMR 8628, Université Paris-Sud, Bât. 425 Faculté des Sciences d'Orsay, F-91405 Orsay, France

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 8, No. 3, 2015
dx.doi.org/10.2140/apde.2015.8.675

L^p ESTIMATES FOR BILINEAR AND MULTIPARAMETER HILBERT TRANSFORMS

WEI DAI AND GUOZHEN LU

Muscalu, Pipher, Tao and Thiele proved that the standard bilinear and biparameter Hilbert transform does not satisfy any L^p estimates. They also raised a question asking if a bilinear and biparameter multiplier operator defined by

$$T_m(f_1, f_2)(x) := \int_{\mathbb{R}^4} m(\xi, \eta) \hat{f}_1(\xi_1, \eta_1) \hat{f}_2(\xi_2, \eta_2) e^{2\pi i x \cdot ((\xi_1, \eta_1)+(\xi_2, \eta_2))}\, d\xi\, d\eta$$

satisfies any L^p estimates, where the symbol m satisfies

$$|\partial_\xi^\alpha \partial_\eta^\beta m(\xi, \eta)| \lesssim \frac{1}{\operatorname{dist}(\xi, \Gamma_1)^{|\alpha|}} \cdot \frac{1}{\operatorname{dist}(\eta, \Gamma_2)^{|\beta|}}$$

for sufficiently many multi-indices $\alpha = (\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$, Γ_i $(i = 1, 2)$ are subspaces in $\mathbb{R}^2$ and $\dim \Gamma_1 = 0$, $\dim \Gamma_2 = 1$. Silva partially answered this question and proved that T_m maps $L^{p_1} \times L^{p_2} \to L^p$ boundedly when $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p}$ with $p_1, p_2 > 1$, $\frac{1}{p_1} + \frac{2}{p_2} < 2$ and $\frac{1}{p_2} + \frac{2}{p_1} < 2$. One notes that the admissible range here for these tuples (p_1, p_2, p) is a proper subset of the admissible range of the bilinear Hilbert transform (BHT) derived by Lacey and Thiele.

We establish the same L^p estimates as BHT in the full range for the bilinear and d-parameter $(d \geq 2)$ Hilbert transforms with arbitrary symbols satisfying appropriate decay assumptions and having singularity sets $\Gamma_1, \ldots, \Gamma_d$ with $\dim \Gamma_i = 0$ for $i = 1, \ldots, d-1$ and $\dim \Gamma_d = 1$. Moreover, we establish the same L^p estimates as BHT for bilinear and biparameter Fourier multipliers of symbols with $\dim \Gamma_1 = \dim \Gamma_2 = 1$ and satisfying some appropriate decay estimates. In particular, our results include the L^p estimates as BHT in the full range for certain modified bilinear and biparameter Hilbert transforms of tensor-product type with $\dim \Gamma_1 = \dim \Gamma_2 = 1$ but with a slightly better logarithmic decay than that of the bilinear and biparameter Hilbert transform $\mathrm{BHT} \otimes \mathrm{BHT}$.

1. Introduction 676
2. Reduction to restricted weak type estimates of discrete bilinear model operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}_{\vec{\mathbb{P}}}^{\varepsilon}$ 681
3. Trees, L^2 sizes and L^2 energies 697
4. Proof of Theorem 1.3 699
5. Proof of Theorem 1.5 706
References 711

Dai was partly supported by grants from the NNSF of China (No.11371056), the China Postdoctoral Science Foundation and Lu was partly supported by a US NSF grant DMS #1301595.
MSC2010: 42B15, 42B20.
Keywords: bilinear and multiparameter Hilbert transforms, L^p estimates, Hölder estimates, polydiscs, multiparameter paraproducts, wave packets.

1. Introduction

The bilinear Hilbert transform is defined by

$$\mathrm{BHT}(f_1, f_2)(x) := \text{p.v.} \int_{\mathbb{R}} f_1(x-t) f_2(x+t) \, \frac{dt}{t}; \tag{1-1}$$

or, equivalently, it can be written as the bilinear multiplier operator

$$\mathrm{BHT} : (f_1, f_2) \mapsto \int_{\xi<\eta} \hat{f}_1(\xi) \hat{f}_2(\eta) e^{2\pi i x(\xi+\eta)} \, d\xi \, d\eta, \tag{1-2}$$

where f_1 and f_2 are Schwartz functions on $\mathbb{R}$. M. Lacey and C. Thiele proved the following celebrated L^p estimates for the bilinear Hilbert transform:

Theorem 1.1 [Lacey and Thiele 1997; 1999]. *The bilinear operator* BHT *maps* $L^p(\mathbb{R}) \times L^q(\mathbb{R})$ *into* $L^r(\mathbb{R})$ *boundedly for any* $1 < p,\ q \le \infty$ *with* $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ *and* $\frac{2}{3} < r < \infty$.

There are lots of works related to bilinear operators of BHT type. J. Gilbert and A. Nahmod [2001] and F. Bernicot [2008] proved that the same L^p estimates as BHT are valid for bilinear operators with more general symbols. Uniform estimates were obtained by Thiele [2002], L. Grafakos and X. Li [2004] and Li [2006]. A maximal variant of Theorem 1.1 was proved by Lacey [2000]. C. Muscalu, Thiele and T. Tao [Muscalu et al. 2004b] and J. Jung [$\ge$ 2015] investigated various trilinear variants of the bilinear Hilbert transform. For more related results involving estimates for multilinear singular multiplier operators, we refer to, for example, [Christ and Journé 1987; Coifman and Meyer 1978; 1997; Fefferman and Stein 1982; Grafakos and Torres 2002a; 2002b; Journé 1985; Kenig and Stein 1999; Muscalu and Schlag 2013; Muscalu et al. 2002; Thiele 2006] and the references therein.

Since Lacey and Thiele [1997; 1999] established the L^p estimates for $\frac{2}{3} < p < \infty$, whether the bilinear operators of BHT type satisfy L^p estimates all the way down to $\frac{1}{2}$ has remained an open problem. Though we do not have a counterexample yet for the L^p estimates for the bilinear Hilbert transform in the range of $\frac{1}{2} < p < \frac{2}{3}$, we have established in [Dai and Lu $\ge$ 2015b] a counterexample for a modified version of bilinear operators of BHT type. To describe this result, we denote by $\mathscr{F}L^p(\mathbb{R})$ the space consisting of all functions f whose Fourier transform $\hat{f}$ satisfies $\hat{f} \in L^p(\mathbb{R})$. The Hausdorff–Young inequality indicates that $\|\hat{f}\|_{L^{p'}(\mathbb{R})} \lesssim_p \|f\|_{L^p(\mathbb{R})}$ for $1 \le p \le 2$. Then, by Theorem 1.1, it implies that the bilinear Hilbert transform maps $\mathscr{F}L^{p_1'} \times L^{p_2} \to L^p$ for $p_1 \ge 2$ and maps $L^{p_1} \times \mathscr{F}L^{p_2'} \to L^p$ for $p_2 \ge 2$ with $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p}$. Thus it will be interesting to know whether the bilinear operators of BHT type map $\mathscr{F}L^{p_1'} \times L^{p_2} \to L^p$ for $p_1 < 2$ or $L^{p_1} \times \mathscr{F}L^{p_2'} \to L^p$ for $p_2 < 2$ boundedly with $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p}$. Our work in [Dai and Lu $\ge$ 2015b] gives a negative answer to the boundedness of $\mathscr{F}L^{p_1'} \times L^{p_2} \to L^p$ for $p_1 < 2$ and $L^{p_1} \times \mathscr{F}L^{p_2'} \to L^p$ for $p_2 < 2$.

To date, we are still not aware of any uniform L^p estimates for bilinear Fourier multiplier operator of BHT type in the range $p \in \left(\frac{1}{2}, \frac{2}{3}\right)$. By decomposing the bilinear multiplier operator T_m into a summation of infinitely many bilinear paraproducts without modulation invariance, we have proved in [Dai and Lu $\ge$ 2015b] that there exists a class of symbols m (with one-dimensional singularity sets), which also satisfy the symbol estimates of BHT type operators investigated in [Gilbert and Nahmod 2001] and are arbitrarily close to the symbols of BHT type operators, such that the corresponding bilinear multiplier operators T_m associated with symbols m satisfy L^p estimates all the way down to $\frac{1}{2}$.

In multiparameter cases, there are also large amounts of literature devoted to studying the estimates of multiparameter and multilinear operators (see [Chen and Lu 2014; Dai and Lu $\geq$ 2015a; Demeter and Thiele 2010; Hong and Lu 2014; Kesler $\geq$ 2015; Luthy 2013; Muscalu and Schlag 2013; Muscalu et al. 2004a; 2006; Silva 2014] and the references therein). In the bilinear and biparameter cases, let Γ_i ($i = 1, 2$) be subspaces in $\mathbb{R}^2$, we consider operators T_m defined by

$$T_m(f_1, f_2)(x) := \int_{\mathbb{R}^4} m(\xi, \eta)\hat{f}_1(\xi_1, \eta_1)\hat{f}_2(\xi_2, \eta_2)e^{2\pi i x\cdot((\xi_1,\eta_1)+(\xi_2,\eta_2))}\,d\xi\,d\eta, \tag{1-3}$$

where the symbol m satisfies[1]

$$|\partial_\xi^\alpha \partial_\eta^\beta m(\xi, \eta)| \lesssim \frac{1}{\operatorname{dist}(\xi, \Gamma_1)^{|\alpha|}} \cdot \frac{1}{\operatorname{dist}(\eta, \Gamma_2)^{|\beta|}} \tag{1-4}$$

for sufficiently many multi-indices $\alpha = (\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$. If $\dim \Gamma_1 = \dim \Gamma_2 = 0$, Muscalu, J. Pipher, Tao and Thiele proved in [Muscalu et al. 2004a; 2006] that Hölder-type L^p estimates are available for T_m; however, if $\dim \Gamma_1 = \dim \Gamma_2 = 1$, let T_m be the double bilinear Hilbert transform on polydisks $\mathrm{BHT} \otimes \mathrm{BHT}$ defined by

$$\mathrm{BHT} \otimes \mathrm{BHT}(f_1, f_2)(x, y) := \text{p.v.} \int_{\mathbb{R}^2} f_1(x - s, y - t) f_2(x + s, y + t)\,\frac{ds}{s}\,\frac{dt}{t}; \tag{1-5}$$

they also proved in [Muscalu et al. 2004a] that the operator $\mathrm{BHT} \otimes \mathrm{BHT}$ does not satisfy any L^p estimates of Hölder type by constructing a counterexample. In fact, consider bounded functions $f_1(x, y) = f_2(x, y) = e^{ixy}$; one has formally

$$\mathrm{BHT} \otimes \mathrm{BHT}(f_1, f_2)(x, y) = (f_1 \cdot f_2)(x, y) \int_{\mathbb{R}^2} \frac{e^{2ist}}{st}\,ds\,dt = i\pi (f_1 \cdot f_2)(x, y) \int_{\mathbb{R}} \frac{\operatorname{sgn}(s)}{s}\,ds,$$

then localize functions f_1, f_2 and let $f_1^N(x, y) = f_2^N(x, y) = e^{ixy}\chi_{[-N,N]}(x)\chi_{[-N,N]}(y)$. One can verify the pointwise estimate

$$|\mathrm{BHT} \otimes \mathrm{BHT}(f_1^N, f_2^N)(x, y)| \geq \left| \int_{-\frac{N}{10}}^{\frac{N}{10}} \int_{-\frac{N}{10}}^{\frac{N}{10}} \frac{e^{2ist}}{st}\,ds\,dt \right| + O(1) \geq C \log N + O(1) \tag{1-6}$$

for every $x, y \in \left[-\frac{1}{100}N, \frac{1}{100}N\right]$ and sufficiently large $N \in \mathbb{Z}^+$, which indicates that no Hölder-type L^p estimates are available for the bilinear operator $\mathrm{BHT} \otimes \mathrm{BHT}$. When $\dim \Gamma_1 = 0$ and $\dim \Gamma_2 = 1$, there is the following problem:

Question 1.2 [Muscalu et al. 2004a, Question 8.2]. Let $\dim \Gamma_1 = 0$ and $\dim \Gamma_2 = 1$ with Γ_2 nondegenerate in the sense of [Muscalu et al. 2002]. If m is a multiplier satisfying (1-4), does the corresponding operator T_m defined by (1-3) satisfy any L^p estimates?

---

[1] Throughout this paper, $A \lesssim B$ means that there exists a universal constant $C > 0$ such that $A \leq CB$. If necessary, we use explicitly $A \lesssim_{\star,\dots,\star} B$ to indicate that there exists a positive constant $C_{\star,\dots,\star}$, continuously depending only on the quantities appearing in the subscript, such that $A \leq C_{\star,\dots,\star}B$.

P. Silva [2014] answered this question partially and proved that T_m defined by (1-3), (1-4) with $\dim\Gamma_1 = 0$ and $\dim\Gamma_2 = 1$ maps $L^p \times L^q \to L^r$ boundedly when $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$ with $p, q > 1$, $\frac{1}{p}+\frac{2}{q} < 2$ and $\frac{1}{q}+\frac{2}{p} < 2$. One should observe that the admissible range for these tuples (p, q, r) is a proper subset of the region $p, q > 1$ and $\frac{3}{4} < r < \infty$, which is also properly contained in the admissible range of BHT (see Theorem 1.1).

Naturally, we may wonder whether the biparameter bilinear operator T_m given by (1-3), (1-4) (with appropriate decay assumptions on the symbol m and singularity sets Γ_1, Γ_2 satisfying $\dim\Gamma_1 = 0$ or 1, $\dim\Gamma_2 = 1$) satisfies the same L^p estimates as BHT.

To study this problem, we must find the implicit decay assumptions on symbol m to preclude the existence of those kinds of counterexamples constructed in (1-6) for BHT $\otimes$ BHT. To this end, let us consider first the bilinear operator $T_m \otimes \mathrm{BHT}$ of tensor product type that is defined by

$$T_m \otimes \mathrm{BHT}(f_1, f_2)(x, y) := \mathrm{p.v.}\int_{\mathbb{R}^2} f_1(x-s, y-t) f_2(x+s, y+t)\frac{K(s)}{t}\,ds\,dt, \tag{1-7}$$

where the symbol $m(\xi_1^1, \xi_2^1) = m(\zeta) := \hat{K}(\zeta)$ with $\zeta := \xi_1^1 - \xi_2^1$ has one-dimensional nondegenerate singularity set Γ_1. Let $f_1(x, y) = f_2(x, y) = e^{ixy}$; one can easily derive that

$$T_m \otimes \mathrm{BHT}(f_1, f_2)(x, y) = (f_1 \cdot f_2)(x, y)\int_{\mathbb{R}^2} K(s)\frac{e^{2ist}}{t}\,ds\,dt. \tag{1-8}$$

From (1-8) and the above counterexample constructed in (1-6) for the operator BHT $\otimes$ BHT, we observe that one sufficient condition for precluding the existence of these kinds of counterexamples is $K \in L^1$ or, equivalently, $m = \hat{K} \in \mathcal{F}(L^1)$. From the Riemann–Lebesgue theorem, we know that a necessary condition for $m \in \mathcal{F}(L^1)$ is $m(\zeta) \to 0$ as $|\zeta| \to \infty$. Moreover, if $K \in L^1(\mathbb{R})$ is odd, one can even derive that $\left|\int_{\mathbb{R}} m(\zeta)/\zeta\,d\zeta\right| \lesssim \|K\|_{L^1}$ (this indicates that many uniformly continuous functions with logarithmic decay rate do not belong to $\mathcal{F}(L^1)$). Therefore, in order to guarantee that the same L^p estimates as the bilinear Hilbert transform are available for the bilinear operators $T_m \otimes$ BHT and BHT $\otimes$ BHT, we need some appropriate decay assumptions on the symbol.

The purpose of this paper is to prove the same L^p estimates as BHT for modified bilinear operators $T_m \otimes$ BHT with arbitrary nonsmooth symbols which decay faster than the logarithmic rate.

For $d \geq 2$, any two generic vectors $\xi_1 = (\xi_1^i)_{i=1}^d$, $\xi_2 = (\xi_2^i)_{i=1}^d$ in $\mathbb{R}^d$ generate naturally the following collection of d vectors in $\mathbb{R}^2$:

$$\bar{\xi}_1 = (\xi_1^1, \xi_2^1), \quad \bar{\xi}_2 = (\xi_1^2, \xi_2^2), \quad \ldots, \quad \bar{\xi}_d = (\xi_1^d, \xi_2^d). \tag{1-9}$$

Let $m = m(\xi) = m(\bar{\xi})$ be a bounded symbol in $L^\infty(\mathbb{R}^{2d})$ that is smooth away from the subspaces $\Gamma_1 \cup \cdots \cup \Gamma_{d-1} \cup \Gamma_d$ and satisfies

$$\mathrm{dist}(\bar{\xi}_d, \Gamma_d)^{|\alpha_d|} \cdot \int_{\mathbb{R}^{2(d-1)}} \frac{\left|\partial_{\bar{\xi}_1}^{\alpha_1} \cdots \partial_{\bar{\xi}_d}^{\alpha_d} m(\bar{\xi})\right|}{\prod_{i=1}^{d-1} \mathrm{dist}(\bar{\xi}_i, \Gamma_i)^{2-|\alpha_i|}}\,d\bar{\xi}_1 \cdots d\bar{\xi}_{d-1} \leq B < +\infty \tag{1-10}$$

for sufficiently many multi-indices $\alpha_1, \dots, \alpha_d$, where $\dim \Gamma_i = 0$ for $i = 1, \dots, d-1$ and $\Gamma_d := \{(\xi_1^d, \xi_2^d) \in \mathbb{R}^2 : \xi_1^d = \xi_2^d\}$. Denote by $T_m^{(d)}$ the bilinear multiplier operator defined by

$$T_m^{(d)}(f_1, f_2)(x) := \int_{\mathbb{R}^{2d}} m(\xi) \hat{f}_1(\xi_1) \hat{f}_2(\xi_2) e^{2\pi i x \cdot (\xi_1 + \xi_2)} \, d\xi. \tag{1-11}$$

Our result for bilinear operators $T_m^{(d)}$ satisfying (1-10) and (1-11) is the following:

Theorem 1.3. *For any $d \ge 2$, the bilinear, d-parameter multiplier operator $T_m^{(d)}$ maps $L^{p_1}(\mathbb{R}^d) \times L^{p_2}(\mathbb{R}^d)$ into $L^p(\mathbb{R}^d)$ boundedly for any $1 < p_1, p_2 \le \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{2}{3} < p < \infty$. The implicit constants in the bounds depend only on p_1, p_2, p, d and B.*

Remark 1.4. For arbitrarily small $\varepsilon > 0$, let $m^\varepsilon = m^\varepsilon(\xi) = m^\varepsilon(\bar{\xi})$ be a bounded symbol in $L^\infty(\mathbb{R}^{2d})$ that is smooth away from the subspaces $\Gamma_1 \cup \dots \cup \Gamma_{d-1} \cup \Gamma_d$ defined as in Theorem 1.3 and satisfying differential estimates

$$|\partial_{\bar{\xi}_1}^{\alpha_1} \cdots \partial_{\bar{\xi}_d}^{\alpha_d} m^\varepsilon(\bar{\xi})| \lesssim \prod_{i=1}^{d-1} \left(\frac{1}{\operatorname{dist}(\bar{\xi}_i, \Gamma_i)^{|\alpha_i|}} \cdot \langle \log_2 \operatorname{dist}(\bar{\xi}_i, \Gamma_i) \rangle^{-(1+\varepsilon)} \right) \cdot \frac{1}{\operatorname{dist}(\bar{\xi}_d, \Gamma_d)^{|\alpha_d|}} \tag{1-12}$$

for sufficiently many multi-indices $\alpha_1, \dots, \alpha_d$; then m^ε satisfies conditions (1-10).

As shown in [Muscalu et al. 2004a], the bilinear and biparameter Hilbert transform does not satisfy any L^p estimates. This is the case when the singularity sets Γ_1 and Γ_2 satisfy $\dim \Gamma_1 = \dim \Gamma_2 = 1$. Thus, it is natural to ask if the L^p estimates will break down for any bilinear and biparameter Fourier multiplier operator with $\dim \Gamma_1 = \dim \Gamma_2 = 1$. In other words, will a nonsmooth symbol with the same dimensional singularity sets but with a slightly better decay than that for the bilinear and biparameter Hilbert transform assure the L^p estimates? Our next two theorems will address this issue.

For $d = 2$ and arbitrarily small $\varepsilon > 0$, let $\tilde{m}^\varepsilon = \tilde{m}^\varepsilon(\xi) = \tilde{m}^\varepsilon(\bar{\xi})$ be a bounded symbol in $L^\infty(\mathbb{R}^4)$ that is smooth away from the subspaces $\Gamma_1 \cup \Gamma_2$ and satisfies

$$|\partial_{\bar{\xi}_1}^{\alpha_1} \partial_{\bar{\xi}_2}^{\alpha_2} \tilde{m}^\varepsilon(\bar{\xi})| \lesssim \prod_{i=1}^{2} \frac{1}{\operatorname{dist}(\bar{\xi}_i, \Gamma_i)^{|\alpha_i|}} \cdot \langle \log_2 \operatorname{dist}(\bar{\xi}_1, \Gamma_1) \rangle^{-(1+\varepsilon)} \tag{1-13}$$

for sufficiently many multi-indices α_1, α_2, where $\langle x \rangle := \sqrt{1 + x^2}$ and $\Gamma_i := \{(\xi_1^i, \xi_2^i) \in \mathbb{R}^2 : \xi_1^i = \xi_2^i\}$ for $i = 1, 2$. Denote by $T_{\tilde{m}^\varepsilon}^{(2)}$ the bilinear multiplier operator defined by

$$T_{\tilde{m}^\varepsilon}^{(2)}(f_1, f_2)(x) := \int_{\mathbb{R}^4} \tilde{m}^\varepsilon(\xi) \hat{f}_1(\xi_1) \hat{f}_2(\xi_2) e^{2\pi i x \cdot (\xi_1 + \xi_2)} \, d\xi. \tag{1-14}$$

Our result for bilinear operators $T_{\tilde{m}^\varepsilon}^{(2)}$ satisfying (1-13) and (1-14) is the following:

Theorem 1.5. *For $d = 2$ and any $\varepsilon > 0$, the bilinear and biparameter multiplier operator $T_{\tilde{m}^\varepsilon}^{(2)}$ maps $L^{p_1}(\mathbb{R}^2) \times L^{p_2}(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$ boundedly for any $1 < p_1, p_2 \le \infty$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{2}{3} < p < \infty$. The implicit constants in the bounds depend only on p_1, p_2, p, ε and tend to infinity as $\varepsilon \to 0$.*

Our result for modified bilinear and biparameter Hilbert transform of tensor product type with a slightly better decay than that of $\mathrm{BHT} \otimes \mathrm{BHT}$ is the following:

Theorem 1.6. *For any $\varepsilon > 0$, let the bilinear and biparameter operator* $\mathrm{BHT}^{\varepsilon} \otimes \mathrm{BHT}$ *be defined by*

$$\mathrm{BHT}^{\varepsilon} \otimes \mathrm{BHT}(f_1, f_2)(x_1, x_2) = p.v. \int_{\mathbb{R}^2} f_1(x-s) f_2(x+s) \frac{\Psi^{\varepsilon}(s_1)}{s_2} \, ds_1 \, ds_2$$

with the function Ψ^{ε} satisfying

$$|\partial_{\bar{\xi}_1}^{\alpha_1} \hat{\Psi}^{\varepsilon}(\xi_1^1 - \xi_2^1)| \lesssim |\xi_1^1 - \xi_2^1|^{-|\alpha_1|} \cdot \langle \log_2 |\xi_1^1 - \xi_2^1| \rangle^{-(1+\varepsilon)} \tag{1-15}$$

for sufficiently many multi-indices α_1; then it satisfies the same L^p estimates as $T^{(2)}_{\tilde{m}^{\varepsilon}}$.

Remark 1.7. For simplicity, we will only consider the biparameter case $d = 2$ and $\Gamma_i = \{(0, 0)\}$ ($i = 1, \dots, d-1$) in the proof of Theorem 1.3. It will be clear from the proof (see Section 4) that we can extend the argument to the general d-parameter and $\dim \Gamma_i = 0$ ($i = 1, \dots, d-1$) cases straightforwardly. We will only prove Theorem 1.5 in Section 5 and omit the proof of Theorem 1.6, since one can observe from the discretization procedure in Section 2 that the bilinear and biparameter operator $\mathrm{BHT}^{\varepsilon} \otimes \mathrm{BHT}$ can be reduced to the same bilinear model operators $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ as $T^{(2)}_{\tilde{m}^{\varepsilon}}$.

It's well known that a standard approach to prove L^p estimates for one-parameter n-linear operators with singular symbols (e.g., Coifman–Meyer multiplier, BHT and one-parameter paraproducts) is by the generic estimates of the corresponding $(n+1)$-linear forms consisting of estimates for different sizes and energies (see [Jung $\geq$ 2015; Muscalu and Schlag 2013; Muscalu et al. 2002; 2004b]), which relies on the one-dimensional BMO theory, or, more precisely, the John–Nirenberg-type inequalities to get good control over the relevant sizes. Unfortunately, there is no routine generalization of such approach to multiparameter settings, for instance, we don't have analogues of the John–Nirenberg inequalities for dyadic rectangular BMO spaces in the two-parameter case (see [Muscalu and Schlag 2013]). To overcome these difficulties, Muscalu et al. [2004a] developed a completely new approach to prove L^p estimates for biparameter paraproducts; their essential idea is to apply the stopping-time decompositions based on hybrid square and maximal operators MM, MS, SM and SS, the one-dimensional BMO theory and Journé's lemma, and hence could not be extended to solve the general d-parameter ($d \geq 3$) cases. As to the general d-parameter ($d \geq 3$) cases, by proving a generic decomposition (see Lemma 4.1), Muscalu et al. [2006] simplified the arguments they introduced in [Muscalu et al. 2004a], and this simplification works equally well in all d-parameter settings. Recently, a pseudodifferential variant of the theorems in [Muscalu et al. 2004a; 2006] has been established in [Dai and Lu $\geq$ 2015a]. Moreover, J. Chen and G. Lu [Chen and Lu 2014] offer a different proof than those in [Muscalu et al. 2004a; 2006] to establish a Hörmander-type theorem of L^p estimates (and weighted estimates as well) for multilinear and multiparameter Fourier multiplier operators with limited smoothness in multiparameter Sobolev spaces.

However, in this paper, in order to prove our main results, Theorems 1.3 and 1.5 in biparameter settings, we have at least two different difficulties from [Muscalu et al. 2004a; 2006]. First, observe that if one restricts the sum of tritiles $P'' \in \mathbb{P}''$ in the definitions of discrete model operators (see Section 2) to a tree then one essentially gets a tensor product of two discrete paraproducts on x_1 and x_2, respectively, which can be estimated by the MM, MS, SM and SS functions, but, due to the extra degree of freedom in frequency in the x_2 direction, there are infinitely many such tensor products of paraproducts in the summation, so it's difficult for us to carry out the stopping-time decompositions by using the hybrid square and maximal operators as in [Muscalu et al. 2004a; 2006]. Second, in the proof of Theorem 1.5, note that there are infinitely many tritiles $P' \in \mathbb{P}'$ with the property that $I_{P'} = I_0$ for a certain fixed dyadic interval I_0 of the same length as $I_{P'}$, so we can't estimate $\sum_{P'} |I_{P'}| \lesssim |\tilde{I}|$ for all dyadic intervals $I_{P'} \subseteq \tilde{I}$ with comparable lengths, and hence we can't apply Journé's lemma as in [Muscalu et al. 2004a] either. By making use of the L^2 sizes and L^2 energies estimates of the trilinear forms, the almost orthogonality of wave packets associated with different tiles of distinct trees and the decay assumptions on the symbols, we are able to overcome these difficulties in the proof of Theorems 1.3 and 1.5 in biparameter settings.

Nevertheless, in the proof of Theorem 1.5 in general d-parameter settings ($d \geq 3$), one easily observes that the generic decomposition will destroy the perfect orthogonality of wave packets associated with distinct tiles which have disjoint frequency intervals in both the x_1 and x_2 directions, thus we can't apply the generic decomposition to extend the results of Theorem 1.5 to higher parameters $d \geq 3$ as in [Muscalu et al. 2006]. For the proof of Theorem 1.3, we are able to apply the generic decomposition lemma (Lemma 4.1) to the $d-1$ variables $x_1, \dots, x_{d-1}$. Although one can't obtain that supp $\Phi^{3,\ell}_{\tilde{P}'_3} \otimes \Phi^3_{P''_3}$ is entirely contained in the exceptional set U as in [Muscalu et al. 2006], one can observe that the support set is contained in U in all the variables $x_1, \dots, x_{d-1}$, but not the last, x_d. Therefore, we only need to consider the distance from the support set to the set E'_3 in the x_d direction and obtain enough decay factors for summation; the extension of the proof to the general d-parameter ($d \geq 3$) cases is straightforward.

The rest of this paper is organized as follows. In Section 2 we reduce the proof of Theorem 1.3 and Theorem 1.5 to proving restricted weak type estimates of discrete bilinear model operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ (Proposition 2.17). Section 3 is devoted to giving a review of the definitions and useful properties about trees, L^2 sizes and L^2 energies introduced in [Muscalu et al. 2004b]. In Sections 4 and 5 we carry out the proof of Proposition 2.17, which completes the proof of our main theorems, Theorem 1.3 and Theorem 1.5, respectively.

2. Reduction to restricted weak type estimates of discrete bilinear model operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$

2A. ***Discretization.*** As we can see from the study of multiparameter and multilinear Coifman–Meyer multiplier operators (see, e.g., [Muscalu et al. 2002; 2004a; 2004b; 2006]), a standard approach to obtain L^p estimates of bilinear operators $T^{(d)}_m$ and $T^{(2)}_{\tilde{m}^{\varepsilon}}$ is to reduce them into discrete sums of inner products with wave packets (see [Thiele 2006]).

2A1. *Discretization for bilinear, biparameter operators $T^{(2)}_m$ with $\Gamma_1 = \{(0,0)\}$.* We will use the following discretization procedure. First, we need to decompose the symbol $m(\xi)$ in a natural way. To this end, for the first spatial variable x_1, we decompose the region $\{\bar{\xi}_1 = (\xi^1_1, \xi^1_2) \in \mathbb{R}^2 \setminus \{(0,0)\}\}$ by using *Whitney squares* with respect to the singularity point $\{\xi^1_1 = \xi^1_2 = 0\}$, while, for the last spatial variable x_2, we decompose the region $\{\bar{\xi}_2 = (\xi^2_1, \xi^2_2) \in \mathbb{R}^2 : \xi^2_1 \neq \xi^2_2\}$ by using Whitney squares with respect to the singularity line $\Gamma_2 = \{\xi^2_1 = \xi^2_2\}$. In order to describe our discretization procedure clearly, let us first recall some standard notation and definitions in [Muscalu et al. 2004b].

An interval I on the real line $\mathbb{R}$ is called dyadic if it is of the form $I = 2^{-k}[n,\ n+1]$ for some $k, n \in \mathbb{Z}$. An interval is said to be a *shifted dyadic interval* if it is of the form $2^{-k}[j+\alpha,\ j+1+\alpha]$ for some $k,\ j \in \mathbb{Z}$ and $\alpha \in \{0, \frac{1}{3}, -\frac{1}{3}\}$. A *shifted dyadic cube* is a set of the form $Q = Q_1 \times Q_2 \times Q_3$, where each Q_j is a shifted dyadic interval and they all have the same length. A *shifted dyadic quasicube* is a set $Q = Q_1 \times Q_2 \times Q_3$, where Q_j $(j = 1, 2, 3)$ are shifted dyadic intervals satisfying the less restrictive condition $|Q_1| \simeq |Q_2| \simeq |Q_3|$. One easily observes that, for every cube $Q \subseteq \mathbb{R}^3$, there exists a shifted dyadic cube $\tilde{Q}$ such that $Q \subset \frac{7}{10}\tilde{Q}$ (the cube having the same center as $\tilde{Q}$ but with side length $\frac{7}{10}$ that of $\tilde{Q}$) and $\operatorname{diam}(Q) \simeq \operatorname{diam}(\tilde{Q})$.

The same terminology will also be used in the plane $\mathbb{R}^2$. The only difference is that the previous cubes become squares. For any cube or square Q, we will denote the side length of Q by $\ell(Q)$ and denote the reflection of Q with respect to the origin by $-Q$ hereafter.

Definition 2.1 [Muscalu and Schlag 2013; Muscalu et al. 2006]. For $J \subseteq \mathbb{R}$ an arbitrary interval, we say that a smooth function Φ_J is a bump adapted to J if and only if the following inequalities hold:

$$|\Phi_J^{(l)}(x)| \lesssim_{l,\alpha} \frac{1}{|J|^l} \cdot \frac{1}{(1+\operatorname{dist}(x, J)/|J|)^{\alpha}} \tag{2-1}$$

for every integer $\alpha \in \mathbb{N}$ and for sufficiently many derivatives $l \in \mathbb{N}$. If Φ_J is a bump adapted to J, we say that $|J|^{-\frac{1}{2}}\Phi_J$ is an L^2-normalized bump adapted to J.

Now let $\varphi \in \mathcal{S}(\mathbb{R})$ be an even Schwartz function such that $\operatorname{supp}\hat{\varphi} \subseteq \left[-\frac{3}{16}, \frac{3}{16}\right]$ and $\hat{\varphi}(\xi) = 1$ on $\left[-\frac{1}{6}, \frac{1}{6}\right]$, and define $\psi \in \mathcal{S}(\mathbb{R})$ to be the Schwartz function whose Fourier transform satisfies $\hat{\psi}(\xi) := \hat{\varphi}(\xi/4) - \hat{\varphi}(\xi/2)$ and $\operatorname{supp}\hat{\psi} \subseteq \left[-\frac{3}{4}, -\frac{1}{3}\right] \cup \left[\frac{1}{3}, \frac{3}{4}\right]$, such that $0 \le \hat{\varphi}(\xi),\ \hat{\psi}(\xi) \le 1$. Then, for every integer $k \in \mathbb{Z}$, we define $\hat{\varphi}_k,\ \hat{\psi}_k \in \mathcal{S}(\mathbb{R})$ by

$$\hat{\varphi}_k(\xi) := \hat{\varphi}\left(\frac{\xi}{2^k}\right), \quad \hat{\psi}_k(\xi) := \hat{\psi}\left(\frac{\xi}{2^k}\right) = \hat{\varphi}_{k+2}(\xi) - \hat{\varphi}_{k+1}(\xi) \tag{2-2}$$

and observe that

$$\operatorname{supp}\hat{\varphi}_k \subseteq \left[-\tfrac{3}{16}\cdot 2^k, \tfrac{3}{16}\cdot 2^k\right], \quad \operatorname{supp}\hat{\psi}_k \subseteq \left[-\tfrac{3}{4}\cdot 2^k, -\tfrac{1}{3}\cdot 2^k\right] \cup \left[\tfrac{1}{3}\cdot 2^k, \tfrac{3}{4}\cdot 2^k\right],$$

and $\operatorname{supp}\hat{\psi}_k \cap \operatorname{supp}\hat{\psi}_{k'} = \varnothing$ for any integers $k, k' \in \mathbb{Z}$ such that $|k-k'| \ge 2$, and $\operatorname{supp}\hat{\varphi} \cap \operatorname{supp}\hat{\psi}_k = \varnothing$ for any integer $k \ge 0$. One easily obtains the homogeneous Littlewood–Paley dyadic decomposition

$$1 = \sum_{k\in\mathbb{Z}} \hat{\psi}_k(\xi) \quad \text{for all } \xi \in \mathbb{R}\setminus\{0\} \tag{2-3}$$

and inhomogeneous Littlewood–Paley dyadic decomposition

$$1 = \hat{\varphi}(\xi) + \sum_{k\ge -1} \hat{\psi}_k(\xi) \quad \text{for all } \xi \in \mathbb{R}. \tag{2-4}$$

As a consequence, we get a decomposition for the product $1(\xi_1^1, \xi_2^1) = 1(\xi_1^1)\cdot 1(\xi_2^1)$ as follows:

$$1(\xi_1^1, \xi_2^1) = \sum_{k'\in\mathbb{Z}} \hat{\varphi}_{k'}(\xi_1^1)\hat{\psi}_{k'}(\xi_2^1) + \sum_{k'\in\mathbb{Z}} \hat{\psi}_{k'}(\xi_1^1)\hat{\tilde{\psi}}_{k'}(\xi_2^1) + \sum_{k'\in\mathbb{Z}} \hat{\psi}_{k'}(\xi_1^1)\hat{\varphi}_{k'}(\xi_2^1) \tag{2-5}$$

for every $(\xi_1^1, \xi_2^1) \neq (0,0)$, where

$$\hat{\tilde{\psi}}_{k'} := \sum_{|k-k'|\leq 1,\, k\in\mathbb{Z}} \hat{\psi}_k \quad \text{for all } k' \in \mathbb{Z}.$$

By breaking the characteristic function of the plane (ξ_1^1, ξ_2^1) into finite sums of smoothed versions of characteristic functions of cones as in (2-5), we can decompose the operator $T_m^{(2)}$ into a finite sum of several parts in the x_1 direction. Since all the operators obtained in this decomposition can be treated in the same way, we will only discuss one of them in detail. More precisely, let

$$\tilde{\mathbb{Q}}' := \left\{ \tilde{Q}' = \tilde{Q}_1' \times \tilde{Q}_2' \subseteq \mathbb{R}^2 : \tilde{Q}_1' := 2^{k'}\left[-\tfrac{1}{2}, \tfrac{1}{2}\right],\ \tilde{Q}_2' := 2^{k'}\left[\tfrac{1}{24}, \tfrac{25}{24}\right] \text{ for all } k' \in \mathbb{Z} \right\}. \tag{2-6}$$

For each square $\tilde{Q}' \in \tilde{\mathbb{Q}}'$, we define bump functions $\phi_{\tilde{Q}_i', i}$ $(i = 1, 2)$ adapted to intervals $\tilde{Q}_i'$ and satisfying $\operatorname{supp} \phi_{\tilde{Q}_i', i} \subseteq \frac{9}{10}\tilde{Q}_i'$ by

$$\phi_{\tilde{Q}_1', 1}(\xi) := \hat{\varphi}\left(\frac{\xi}{\ell(\tilde{Q}')}\right) = \hat{\varphi}_{k'}(\xi) \tag{2-7}$$

and

$$\phi_{\tilde{Q}_2', 2}(\xi) := \hat{\tilde{\psi}}\left(\frac{\xi}{\ell(\tilde{Q}')}\right) \cdot \chi_{\{\xi>0\}} = \hat{\tilde{\psi}}_{k'}(\xi) \cdot \chi_{\{\xi>0\}}, \tag{2-8}$$

respectively, and finally define smooth bump functions $\phi_{\tilde{Q}'}$ adapted to $\tilde{Q}'$ and satisfying $\operatorname{supp} \phi_{\tilde{Q}'} \subseteq \frac{9}{10}\tilde{Q}'$ by

$$\phi_{\tilde{Q}'}(\xi_1^1, \xi_2^1) := \phi_{\tilde{Q}_1', 1}(\xi_1^1) \cdot \phi_{\tilde{Q}_2', 2}(\xi_2^1). \tag{2-9}$$

Without loss of generality, we will only consider the smoothed characteristic function of the cone $\{(\xi_1^1, \xi_2^1) \in \mathbb{R}^2 : |\xi_1^1| \lesssim |\xi_2^1|,\ \xi_2^1 > 0\}$ in the decomposition (2-5) from now on, which is defined by

$$\sum_{\tilde{Q}' \in \tilde{\mathbb{Q}}'} \phi_{\tilde{Q}'}(\xi_1^1, \xi_2^1). \tag{2-10}$$

As to the x_2 direction, we consider the collection $\mathbb{Q}''$ of all shifted dyadic squares $Q'' = Q_1'' \times Q_2''$ satisfying

$$Q'' \subseteq \{(\xi_1^2, \xi_2^2) \in \mathbb{R}^2 : \xi_1^2 \neq \xi_2^2\}, \quad \operatorname{dist}(Q'', \Gamma_2) \simeq 10^4 \operatorname{diam}(Q''). \tag{2-11}$$

We can split the collection $\mathbb{Q}''$ into two disjoint subcollections, that is, define

$$\mathbb{Q}_{\mathrm{I}}'' := \{Q'' \in \mathbb{Q}'' : Q'' \subseteq \{\xi_1^2 < \xi_2^2\}\}, \quad \mathbb{Q}_{\mathrm{II}}'' := \{Q'' \in \mathbb{Q}'' : Q'' \subseteq \{\xi_1^2 > \xi_2^2\}\}. \tag{2-12}$$

Since the set of squares $\left\{\frac{7}{10}Q'' : Q'' \in \mathbb{Q}''\right\}$ also forms a finitely overlapping cover of the region $\{\xi_1^2 \neq \xi_2^2\}$, we can apply a standard partition of unity and write the symbol $\chi_{\{\xi_1^2 \neq \xi_2^2\}}$ as

$$\chi_{\{\xi_1^2 \neq \xi_2^2\}} = \sum_{Q'' \in \mathbb{Q}''} \phi_{Q''}(\xi_1^2, \xi_2^2) = \left(\sum_{Q'' \in \mathbb{Q}_{\mathrm{I}}''} + \sum_{Q'' \in \mathbb{Q}_{\mathrm{II}}''} \right) \phi_{Q''}(\xi_1^2, \xi_2^2) = \chi_{\{\xi_1^2 < \xi_2^2\}} + \chi_{\{\xi_1^2 > \xi_2^2\}}, \tag{2-13}$$

where each $\phi_{Q''}$ is a smooth bump function adapted to Q'' and supported in $\frac{8}{10}Q''$.

One can easily observe that we only need to discuss in detail one term in the decomposition (2-13), since the other term can be treated in the same way. Without loss of generality, we will only consider the first term in (2-13), that is, the characteristic function $\chi_{\{\xi_1^2<\xi_2^2\}}$ of the upper half plane with respect to the singularity line Γ_2, which can be written as

$$\chi_{\{\xi_1^2<\xi_2^2\}} = \sum_{Q''\in\mathbb{Q}''_{\mathbb{I}}} \phi_{Q''}(\xi_1^2, \xi_2^2). \tag{2-14}$$

In a word, we only need to consider the bilinear operator $T^{(2)}_{m,(lh,\mathbb{I})}$ given by

$$T^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2)(x) := \sum_{\substack{\tilde{Q}'\in\tilde{\mathbb{Q}}'\\ Q''\in\mathbb{Q}''_{\mathbb{I}}}} \int_{\mathbb{R}^4} m(\xi)\phi_{\tilde{Q}'}(\bar{\xi}_1)\phi_{Q''}(\bar{\xi}_2)\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)e^{2\pi i x\cdot(\xi_1+\xi_2)}\,d\xi \tag{2-15}$$

from now on, and the proof of Theorem 1.3 can be reduced to proving the L^p estimates

$$\|T^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2)\|_{L^p(\mathbb{R}^2)} \lesssim_{p,p_1,p_2,B} \|f_1\|_{L^{p_1}(\mathbb{R}^2)}\cdot\|f_2\|_{L^{p_2}(\mathbb{R}^2)} \tag{2-16}$$

as long as $1 < p_1, p_2 \le \infty$ and $0 < \frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} < \frac{3}{2}$.

On one hand, since $\xi_1^1 \in \operatorname{supp}\phi_{\tilde{Q}'_1,1} \subseteq \ell(\tilde{Q}')\big[-\frac{3}{16}, \frac{3}{16}\big]$ and $\xi_2^1 \in \operatorname{supp}\phi_{\tilde{Q}'_2,2} \subseteq \ell(\tilde{Q}')\big[\frac{1}{3}, \frac{3}{4}\big]$, it follows that $-\xi_1^1 - \xi_2^1 \in \ell(\tilde{Q}')\big[-\frac{15}{16}, -\frac{7}{48}\big]$, and, as a consequence, there exists an interval $\tilde{Q}'_3 := \ell(\tilde{Q}')\big[-\frac{25}{24}, -\frac{1}{24}\big]$ and a bump function $\phi_{\tilde{Q}'_3,3}$ adapted to $\tilde{Q}'_3$ such that $\operatorname{supp}\phi_{\tilde{Q}'_3,3} \subseteq \ell(\tilde{Q}')\big[-\frac{23}{24}, -\frac{1}{8}\big] \subseteq \frac{9}{10}\tilde{Q}'_3$ and $\phi_{\tilde{Q}'_3,3} \equiv 1$ on $\ell(\tilde{Q}')\big[-\frac{15}{16}, -\frac{7}{48}\big]$.

On the other hand, observe that there exist bump functions $\phi_{Q''_i,i}$ $(i = 1, 2)$ adapted to the shifted dyadic interval Q''_i such that $\operatorname{supp}\phi_{Q''_i,i} \subseteq \frac{9}{10}Q''_i$ and $\phi_{Q''_i,i} \equiv 1$ on $\frac{8}{10}Q''_i$ $(i = 1, 2)$, respectively, and $\operatorname{supp}\phi_{Q''} \subseteq \frac{8}{10}Q''$, thus one has $\phi_{Q''_1,1}\cdot\phi_{Q''_2,2} \equiv 1$ on $\operatorname{supp}\phi_{Q''}$. Since $\xi_1^2 \in \operatorname{supp}\phi_{Q''_1,1} \subseteq \frac{9}{10}Q''_1$ and $\xi_2^2 \in \operatorname{supp}\phi_{Q''_2,2} \subseteq \frac{9}{10}Q''_2$, it follows that $-\xi_1^2 - \xi_2^2 \in -\frac{9}{10}Q''_1 - \frac{9}{10}Q''_2$, and, as a consequence, one can find a shifted dyadic interval Q''_3 with the property that $-\frac{9}{10}Q''_1 - \frac{9}{10}Q''_2 \subseteq \frac{7}{10}Q''_3$ and also satisfying $|Q''_1| = |Q''_2| \simeq |Q''_3|$. In particular, there exists a bump function $\phi_{Q''_3,3}$ adapted to Q''_3 and supported in $\frac{9}{10}Q''_3$ such that $\phi_{Q''_3,3} \equiv 1$ on $-\frac{9}{10}Q''_1 - \frac{9}{10}Q''_2$.

We denote by $\tilde{\boldsymbol{Q}}'$ the collection of all cubes $\tilde{Q}' := \tilde{Q}'_1 \times \tilde{Q}'_2 \times \tilde{Q}'_3$ with $\tilde{Q}'_1 \times \tilde{Q}'_2 \in \tilde{\mathbb{Q}}'$ and $\tilde{Q}'_3$ defined as above, and denote by $\boldsymbol{Q}''$ the collection of all shifted dyadic quasicubes $Q'' := Q''_1 \times Q''_2 \times Q''_3$ with $Q''_1 \times Q''_2 \in \mathbb{Q}''_{\mathbb{I}}$ and Q''_3 defined as above.

Definition 2.2 [Muscalu et al. 2004b]. We say that a collection of shifted dyadic quasicubes (cubes) is *sparse* if and only if, for every $j = 1, 2, 3$:

(i) If Q and $\tilde{Q}$ belong to this collection and $|Q_j| < |\tilde{Q}_j|$, then $10^8|Q_j| \le |\tilde{Q}_j|$.

(ii) If Q and $\tilde{Q}$ belong to this collection and $|Q_j| = |\tilde{Q}_j|$, then $10^8 Q_j \cap 10^8\tilde{Q}_j = \varnothing$.

In fact, it is not difficult to see that the collection $\boldsymbol{Q}''$ can be split into a sum of finitely many sparse collection of shifted dyadic quasicubes. Therefore, we can assume from now on that the collection $\boldsymbol{Q}''$ is sparse.

Assuming this we then observe that, for any Q'' in such a sparse collection $\boldsymbol{Q}''$, there exists a unique shifted dyadic cube $\tilde{Q}''$ in $\mathbb{R}^3$ such that $Q'' \subseteq \frac{7}{10}\tilde{Q}''$ and with the property that $\operatorname{diam}(Q'') \simeq \operatorname{diam}(\tilde{Q}'')$.

This allows us in particular to assume further that $\boldsymbol{Q}''$ is a sparse collection of shifted dyadic cubes (that is, $|Q_1''| = |Q_2''| = |Q_3''| = \ell(Q'')$).

Now consider the trilinear form $\Lambda^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2, f_3)$ associated to $T^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2)$, which can be written as

$$\begin{aligned}\Lambda^{(2)}_{m,(lh,\mathbb{I})}&(f_1, f_2, f_3)\\ &:= \int_{\mathbb{R}^2} T^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2)(x) f_3(x)\,dx\\ &= \sum_{\substack{\tilde{Q}'\in\tilde{\boldsymbol{Q}}'\\ Q''\in\boldsymbol{Q}''}} \int_{\xi_1+\xi_2+\xi_3=0} m_{\tilde{Q}',Q''}(\xi_1,\xi_2,\xi_3) \prod_{i=1}^{3} (f_i * (\check{\phi}_{\tilde{Q}'_i,i} \otimes \check{\phi}_{Q''_i,i}))^\wedge(\xi_i)\,d\xi_1\,d\xi_2\,d\xi_3, \end{aligned}\qquad (2\text{-}17)$$

where $\xi_i = (\xi_i^1, \xi_i^2)$ for $i = 1, 2, 3$, while

$$m_{\tilde{Q}',Q''}(\xi_1,\xi_2,\xi_3) := m(\xi_1,\xi_2)\cdot(\tilde{\phi}_{\tilde{Q}'} \otimes (\phi_{Q_1''\times Q_2''}\cdot \tilde{\phi}_{Q_3'',3}))(\xi_1,\xi_2,\xi_3), \qquad (2\text{-}18)$$

where $\tilde{\phi}_{\tilde{Q}'}$ is an appropriate smooth function of $(\xi_1^1, \xi_2^1, \xi_3^1)$ which is supported on a slightly larger cube (with a constant magnification independent of $\ell(\tilde{Q}')$) than $\operatorname{supp}(\phi_{\tilde{Q}'_1,1}(\xi_1^1)\phi_{\tilde{Q}'_2,2}(\xi_2^1)\phi_{\tilde{Q}'_3,3}(\xi_3^1))$ and equals 1 on $\operatorname{supp}(\phi_{\tilde{Q}'_1,1}(\xi_1^1)\phi_{\tilde{Q}'_2,2}(\xi_2^1)\phi_{\tilde{Q}'_3,3}(\xi_3^1))$, the function $\phi_{Q_1''\times Q_2''}(\xi_1^2, \xi_2^2)$ is one term of the partition of unity defined in (2-14), and $\tilde{\phi}_{Q_3'',3}$ is an appropriate smooth function of ξ_3^2 supported on a slightly larger interval (with a constant magnification independent of $\ell(Q'')$) than $\operatorname{supp}\phi_{Q_3'',3}$ which equals 1 on $\operatorname{supp}\phi_{Q_3'',3}$. We can decompose $m_{\tilde{Q}',Q''}(\xi_1,\xi_2,\xi_3)$ as a Fourier series,

$$m_{\tilde{Q}',Q''}(\xi_1,\xi_2,\xi_3) = \sum_{\vec{n}_1,\vec{n}_2,\vec{n}_3\in\mathbb{Z}^2} C^{\tilde{Q}',Q''}_{\vec{n}_1,\vec{n}_2,\vec{n}_3} e^{2\pi i(n_1',n_2',n_3')\cdot(\xi_1^1,\xi_2^1,\xi_3^1)/\ell(\tilde{Q}')} e^{2\pi i(n_1'',n_2'',n_3'')\cdot(\xi_1^2,\xi_2^2,\xi_3^2)/\ell(Q'')}, \qquad (2\text{-}19)$$

where the Fourier coefficients $C^{\tilde{Q}',Q''}_{\vec{n}_1,\vec{n}_2,\vec{n}_3}$ are given by

$$\begin{aligned}C^{\tilde{Q}',Q''}_{\vec{n}_1,\vec{n}_2,\vec{n}_3} = \int_{\mathbb{R}^6} m_{\tilde{Q}',Q''}\big((\ell(\tilde{Q}')\xi_1^1, \ell(Q'')\xi_1^2), (\ell(\tilde{Q}')\xi_2^1, \ell(Q'')\xi_2^2), (\ell(\tilde{Q}')\xi_3^1, \ell(Q'')\xi_3^2)\big)&\\ \times e^{-2\pi i(\vec{n}_1\cdot\xi_1+\vec{n}_2\cdot\xi_2+\vec{n}_3\cdot\xi_3)}\,d\xi_1\,d\xi_2\,d\xi_3.&\end{aligned}\qquad (2\text{-}20)$$

Then, by a straightforward calculation, we can rewrite (2-17) as

$$\begin{aligned}\Lambda^{(2)}_{m,(lh,\mathbb{I})}&(f_1, f_2, f_3)\\ &= \sum_{\substack{\tilde{Q}'\in\tilde{\boldsymbol{Q}}'\\ Q''\in\boldsymbol{Q}''}} \sum_{\vec{n}_1,\vec{n}_2,\vec{n}_3\in\mathbb{Z}^2} C^{\tilde{Q}',Q''}_{\vec{n}_1,\vec{n}_2,\vec{n}_3} \int_{\mathbb{R}^2} \prod_{j=1}^{3} (f_j * (\check{\phi}_{\tilde{Q}'_j,j} \otimes \check{\phi}_{Q''_j,j}))\left(x - \left(\frac{n_j'}{\ell(\tilde{Q}')}, \frac{n_j''}{\ell(Q'')}\right)\right)dx.\end{aligned}\qquad (2\text{-}21)$$

Definition 2.3 [Muscalu et al. 2004b; Thiele 2006]. An arbitrary dyadic rectangle of area 1 in the phase-space plane is called a *Heisenberg box* or *tile*. Let $P := I_P \times \omega_P$ be a tile. An L^2-normalized wave packet on P is a function Φ_P which has Fourier support $\operatorname{supp}\hat{\Psi}_P \subseteq \frac{9}{10}\omega_P$ and obeys the estimates

$$|\Phi_P(x)| \lesssim |I_P|^{-\frac{1}{2}}\left(1 + \frac{\operatorname{dist}(x, I_P)}{|I_P|}\right)^{-M}$$

for all $M > 0$, where the implicit constant depends on M.

Now we define $\phi^{n'_i}_{\tilde{Q}'_i,i} := e^{2\pi i n'_i \xi^1_i/\ell(\tilde{Q}')} \cdot \phi_{\tilde{Q}'_i,i}$ and $\phi^{n''_i}_{Q''_i,i} := e^{2\pi i n''_i \xi^2_i/\ell(Q'')} \cdot \phi_{Q''_i,i}$ for $i = 1, 2, 3$. Since any $\tilde{Q}' \in \tilde{\boldsymbol{Q}}'$ and $Q'' \in \boldsymbol{Q}''$ are both shifted dyadic cubes, there exist integers k', $k'' \in \mathbb{Z}$ such that $\ell(\tilde{Q}') = |\tilde{Q}'_1| = |\tilde{Q}'_2| = |\tilde{Q}'_3| = 2^{k'}$ and $\ell(Q'') = |Q''_1| = |Q''_2| = |Q''_3| = 2^{k''}$, respectively. By splitting the integral region $\mathbb{R}^2$ into the union of unit squares, using the L^2-normalization procedure and simple calculations, we can rewrite (2-21) as

$$\Lambda^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2, f_3)$$

$$= \sum_{\vec{n}_1,\vec{n}_2,\vec{n}_3\in\mathbb{Z}^2} \sum_{\substack{\tilde{Q}'\in\tilde{\boldsymbol{Q}}' \\ Q''\in\boldsymbol{Q}''}} \int_0^1\int_0^1 \sum_{\substack{\tilde{I}' \text{ dyadic,} \\ |\tilde{I}'|=2^{-k'}}} \sum_{\substack{I'' \text{ dyadic,} \\ |I''|=2^{-k''}}} \frac{C^{\tilde{Q}',Q''}_{\vec{n}_1,\vec{n}_2,\vec{n}_3}}{|\tilde{I}'|^{\frac{1}{2}} \times |I''|^{\frac{1}{2}}} \prod_{j=1}^{3} \langle f_j, \check{\phi}^{n'_j,\nu'}_{\tilde{I}',\tilde{Q}'_j,j} \otimes \check{\phi}^{n''_j,\nu''}_{I'',Q''_j,j}\rangle \, d\nu'\, d\nu''$$

$$=: \sum_{\vec{n}_1,\vec{n}_2,\vec{n}_3\in\mathbb{Z}^2} \int_0^1\int_0^1 \sum_{\vec{P}:=\tilde{P}'\otimes P''\in\vec{\mathbb{P}}} \frac{C_{Q_{\vec{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3}}{|I_{\vec{P}}|^{\frac{1}{2}}} \prod_{j=1}^{3} \langle f_j, \Phi^{j,\vec{n}_j,\nu}_{\vec{P}_j}\rangle \, d\nu, \tag{2-22}$$

where $\langle\,\cdot\,,\cdot\,\rangle$ denotes the complex scalar L^2 inner product, and we have:

- Fourier coefficients $C_{Q_{\vec{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3} := C^{\tilde{Q}',Q''}_{\vec{n}_1,\vec{n}_2,\vec{n}_3}$;
- tritiles $\tilde{P}' := (\tilde{P}'_1, \tilde{P}'_2, \tilde{P}'_3)$ and $P'' := (P''_1, P''_2, P''_3)$;
- tiles $\tilde{P}'_i := I_{\tilde{P}'_i} \times \omega_{\tilde{P}'_i}$, where $I_{\tilde{P}'_i} := \tilde{I}' = 2^{-k'}[l', l'+1] =: I_{\tilde{P}'}$ and the frequency intervals are $\omega_{\tilde{P}'_i} := \tilde{Q}'_i$ for $i = 1, 2, 3$;
- tiles $P''_j := I_{P''_j} \times \omega_{P''_j}$, where $I_{P''_j} := I'' = 2^{-k''}[l'', l''+1] =: I_{P''}$ and the frequency intervals are $\omega_{P''_j} := Q''_j$ for $j = 1, 2, 3$;
- frequency cubes $Q_{\tilde{P}'} := \omega_{\tilde{P}'_1} \times \omega_{\tilde{P}'_2} \times \omega_{\tilde{P}'_3}$ and $Q_{P''} := \omega_{P''_1} \times \omega_{P''_2} \times \omega_{P''_3}$;
- $\tilde{\mathbb{P}}'$ denotes a collection of such tritiles $\tilde{P}'$ and $\mathbb{P}''$ denotes a collection of such tritiles P'';
- bitiles $\vec{P}_1$, $\vec{P}_2$ and $\vec{P}_3$ defined by

$$\vec{P}_1 := (\tilde{P}'_1, P''_1) = \big(2^{-k'}[l', l'+1] \times 2^{k'}\big[-\tfrac{1}{2}, \tfrac{1}{2}\big], 2^{-k''}[l'', l''+1] \times Q''_1\big),$$
$$\vec{P}_2 := (\tilde{P}'_2, P''_2) = \big(2^{-k'}[l', l'+1] \times 2^{k'}\big[\tfrac{1}{24}, \tfrac{25}{24}\big], 2^{-k''}[l'', l''+1] \times Q''_2\big),$$
$$\vec{P}_3 := (\tilde{P}'_3, P''_3) = \big(2^{-k'}[l', l'+1] \times 2^{k'}\big[-\tfrac{25}{24}, -\tfrac{1}{24}\big], 2^{-k''}[l'', l''+1] \times Q''_3\big);$$

- the biparameter tritile $\vec{P} := \tilde{P}' \otimes P'' = (\vec{P}_1, \vec{P}_2, \vec{P}_3)$;
- rectangles $I_{\vec{P}_i} := I_{\tilde{P}'_i} \times I_{P''_i} = I_{\tilde{P}'} \times I_{P''} =: I_{\vec{P}}$ for $i = 1, 2, 3$, and hence $|I_{\vec{P}}| = |I_{\tilde{P}'} \times I_{P''}| = |I_{\vec{P}_1}| = |I_{\vec{P}_2}| = |I_{\vec{P}_3}| = 2^{-k'} \cdot 2^{-k''}$;
- the double frequency cube $Q_{\vec{P}} := (Q_{\tilde{P}'}, Q_{P''}) = (\omega_{\tilde{P}'_1} \times \omega_{\tilde{P}'_2} \times \omega_{\tilde{P}'_3}, \omega_{P''_1} \times \omega_{P''_2} \times \omega_{P''_3})$;
- $\vec{\mathbb{P}} := \tilde{\mathbb{P}}' \times \mathbb{P}''$ denotes a collection of such biparameter tritiles $\vec{P}$;
- L^2-normalized wave packets $\Phi^{i,n'_i,\nu'}_{\tilde{P}'_i}$ associated with the Heisenberg boxes $\tilde{P}'_i$ defined by

$$\Phi^{i,n'_i,\nu'}_{\tilde{P}'_i}(x_1) := \check{\phi}^{n'_i,\nu'}_{\tilde{I}',\tilde{Q}'_i,i}(x_1) := 2^{-k'/2}\overline{\check{\phi}^{n'_i}_{\tilde{Q}'_i,i}(2^{-k'}(l'+\nu') - x_1)} \quad \text{for } i = 1, 2, 3;$$

- L^2-normalized wave packets $\Phi_{P_i''}^{i,n_i'',\nu''}$ associated with the Heisenberg boxes P_i'' defined by

$$\Phi_{P_i''}^{i,n_i'',\nu''}(x_2) := \check{\phi}_{I'',Q_i'',i}^{n_i'',\nu''}(x_2) := 2^{-k''/2}\overline{\check{\phi}_{Q_i'',i}^{n_i''}(2^{-k''}(l''+\nu'')-x_2)} \quad \text{for } i=1,2,3;$$

- smooth bump functions $\Phi_{\vec{P}_i}^{i,\vec{n}_i,\nu} := \Phi_{\tilde{P}_i'}^{i,n_i',\nu'} \otimes \Phi_{P_i''}^{i,n_i'',\nu''}$ for $i=1,2,3$.

We have the following rapid decay estimates of the Fourier coefficients $C_{Q_{\vec{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3}$ with respect to the parameters $\vec{n}_1,\vec{n}_2,\vec{n}_3 \in \mathbb{Z}^2$:

Lemma 2.4. *The Fourier coefficients $C_{Q_{\vec{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3}$ satisfy estimates*

$$|C_{Q_{\vec{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3}| \lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{n}_j|)^M} \cdot C_{|I_{\tilde{P}'}|} \tag{2-23}$$

for any biparameter tritile $\vec{P} \in \vec{\mathbb{P}}$, where M is sufficiently large and the sequence $C_{k'} := C_{|I_{\tilde{P}'}|}$ for $|I_{\tilde{P}'}| = 2^{-k'}$ ($k' \in \mathbb{Z}$) satisfies

$$\sum_{k'\in\mathbb{Z}} C_{k'} \le B < +\infty. \tag{2-24}$$

Proof. Let $\ell(Q_{\tilde{P}'}) = 2^{k'}$ and $\ell(Q_{P''}) = 2^{k''}$ for $k', k'' \in \mathbb{Z}$. For any $\vec{n}_1,\vec{n}_2,\vec{n}_3 \in \mathbb{Z}^2$ and $\vec{P} \in \vec{\mathbb{P}}$, we deduce from (2-18) and (2-20) that

$$C_{Q_{\vec{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3} = \int_{\mathbb{R}^6} m_{Q_{\tilde{P}'},Q_{P''}}((2^{k'}\xi_1^1, 2^{k''}\xi_1^2), (2^{k'}\xi_2^1, 2^{k''}\xi_2^2), (2^{k'}\xi_3^1, 2^{k''}\xi_3^2)) e^{-2\pi i(\vec{n}_1\cdot\xi_1+\vec{n}_2\cdot\xi_2+\vec{n}_3\cdot\xi_3)}\, d\xi_1\, d\xi_2\, d\xi_3, \tag{2-25}$$

where

$$m_{Q_{\tilde{P}'},Q_{P''}}((2^{k'}\xi_1^1, 2^{k''}\xi_1^2), (2^{k'}\xi_2^1, 2^{k''}\xi_2^2), (2^{k'}\xi_3^1, 2^{k''}\xi_3^2)) := m(2^{k'}\bar{\xi}_1, 2^{k''}\bar{\xi}_2)\tilde{\phi}_{Q_{\tilde{P}'}}(2^{k'}\xi_1^1, 2^{k'}\xi_2^1, 2^{k'}\xi_3^1)\phi_{\omega_{P_1''}\times\omega_{P_2''}}(2^{k''}\bar{\xi}_2)\tilde{\phi}_{\omega_{P_3''},3}(2^{k''}\xi_3^2). \tag{2-26}$$

Since $\operatorname{supp}(\tilde{\phi}_{Q_{\tilde{P}'}}(\xi_1^1,\xi_2^1,\xi_3^1)\phi_{\omega_{P_1''}\times\omega_{P_2''}}(\bar{\xi}_2)\tilde{\phi}_{\omega_{P_3''},3}(\xi_3^2)) \subseteq Q_{\tilde{P}'} \times Q_{P''}$, we have that

$$\operatorname{supp}(\tilde{\phi}_{Q_{\tilde{P}'}}(2^{k'}\xi_1^1, 2^{k'}\xi_2^1, 2^{k'}\xi_3^1)\phi_{\omega_{P_1''}\times\omega_{P_2''}}(2^{k''}\bar{\xi}_2)\tilde{\phi}_{\omega_{P_3''},3}(2^{k''}\xi_3^2)) \subseteq Q_{\tilde{P}'}^0 \times Q_{P''}^0,$$

where the cubes $Q_{\tilde{P}'}^0$ and $Q_{P''}^0$ are defined by

$$Q_{\tilde{P}'}^0 = \omega_{\tilde{P}_1'}^0 \times \omega_{\tilde{P}_2'}^0 \times \omega_{\tilde{P}_3'}^0 := \{(\xi_1^1,\xi_2^1,\xi_3^1) \in \mathbb{R}^3 : (2^{k'}\xi_1^1, 2^{k'}\xi_2^1, 2^{k'}\xi_3^1) \in Q_{\tilde{P}'}\}, \tag{2-27}$$

$$Q_{P''}^0 = \omega_{P_1''}^0 \times \omega_{P_2''}^0 \times \omega_{P_3''}^0 := \{(\xi_1^2,\xi_2^2,\xi_3^2) \in \mathbb{R}^3 : (2^{k''}\xi_1^2, 2^{k''}\xi_2^2, 2^{k''}\xi_3^2) \in Q_{P''}\} \tag{2-28}$$

and satisfy $|Q_{\tilde{P}'}^0| \simeq |Q_{P''}^0| \simeq 1$. From the properties of the Whitney squares we constructed above, one obtains that $\operatorname{dist}(2^{k'}\bar{\xi}_1, \Gamma_1) \simeq 2^{k'}$ for any $\bar{\xi}_1 \in \omega_{\tilde{P}_1'}^0 \times \omega_{\tilde{P}_2'}^0$ and $\operatorname{dist}(2^{k''}\bar{\xi}_2, \Gamma_2) \simeq 2^{k''}$ for any $\bar{\xi}_2 \in \omega_{P_1''}^0 \times \omega_{P_2''}^0$.

One can deduce from (2-25), (2-26) and integrating by parts sufficiently many times that

$$\begin{aligned}
&|C_{Q_{\tilde{P}},\vec{n}_1,\vec{n}_2,\vec{n}_3}| \\
&\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{n}_j|)^M} \\
&\qquad \times \int_{Q^0_{\tilde{P}'}\times Q^0_{P''}} \big|\partial^{\alpha_1}_{\xi_1}\partial^{\alpha_2}_{\xi_2}\partial^{\alpha_3}_{\xi_3}\big[m_{Q_{\tilde{P}'},Q_{P''}}((2^{k'}\xi_1^1,2^{k''}\xi_1^2),(2^{k'}\xi_2^1,2^{k''}\xi_2^2),(2^{k'}\xi_3^1,2^{k''}\xi_3^2))\big]\big|\,d\xi_1\,d\xi_2\,d\xi_3 \\
&\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{n}_j|)^M} \int_{\omega^0_{P''_1}\times\omega^0_{P''_2}} \operatorname{dist}(2^{k''}\bar{\xi}_2,\Gamma_2)^{|\alpha''|} \int_{\omega^0_{\tilde{P}'_1}\times\omega^0_{\tilde{P}'_2}} \operatorname{dist}(2^{k'}\bar{\xi}_1,\Gamma_1)^{|\alpha'|}|\partial^{\alpha'}_{\bar{\xi}_1}\partial^{\alpha''}_{\bar{\xi}_2} m(2^{k'}\bar{\xi}_1,2^{k''}\bar{\xi}_2)|\,d\bar{\xi}_1\,d\bar{\xi}_2 \\
&\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{n}_j|)^M}\cdot\frac{1}{\ell(Q_{P''})^2} \int_{\omega_{P''_1}\times\omega_{P''_2}} \operatorname{dist}(\bar{\xi}_2,\Gamma_2)^{|\alpha''|} \int_{\omega_{\tilde{P}'_1}\times\omega_{\tilde{P}'_2}} \operatorname{dist}(\bar{\xi}_1,\Gamma_1)^{|\alpha'|-2}|\partial^{\alpha'}_{\bar{\xi}_1}\partial^{\alpha''}_{\bar{\xi}_2} m(\bar{\xi}_1,\bar{\xi}_2)|\,d\bar{\xi}_1\,d\bar{\xi}_2 \\
&=: \prod_{j=1}^{3} \frac{1}{(1+|\vec{n}_j|)^M}\cdot C_{|I_{\tilde{P}'}|},
\end{aligned}$$

where the multi-indices $\alpha_i := (\alpha_i^1,\alpha_i^2)$ for $i=1,2,3$ and $|\alpha_1|=|\alpha_2|=|\alpha_3|=M$ are sufficiently large, the multi-indices $\alpha' := (\alpha'_1,\alpha'_2,\alpha'_3)$, $\alpha'' := (\alpha''_1,\alpha''_2,\alpha''_3)$ with $\alpha'_i \le \alpha_i^1$ and $\alpha''_j \le \alpha_j^2$ for $i, j = 1, 2, 3$. This proves the estimates (2-23).

Moreover, for $|I_{\tilde{P}'}| = 2^{-k'}$, we define the sequence $C_{k'} := C_{|I_{\tilde{P}'}|}$ ($k'\in\mathbb{Z}$). From the estimates (1-10) for symbol $m(\bar{\xi}_1,\bar{\xi}_2)$, we get that

$$\operatorname{dist}(\bar{\xi}_2,\Gamma_2)^{|\alpha''|}\cdot\int_{\mathbb{R}^2}\operatorname{dist}(\bar{\xi}_1,\Gamma_1)^{|\alpha'|-2}|\partial^{\alpha'}_{\bar{\xi}_1}\partial^{\alpha''}_{\bar{\xi}_2} m(\bar{\xi})|\,d\bar{\xi}_1 \le B < +\infty, \tag{2-29}$$

and hence we can deduce the following summable property for the sequence $\{C_{k'}\}_{k'\in\mathbb{Z}}$:

$$\begin{aligned}
\sum_{k'\in\mathbb{Z}} C_{k'} &\le \frac{1}{\ell(Q_{P''})^2}\int_{\omega_{P''_1}\times\omega_{P''_2}}\operatorname{dist}(\bar{\xi}_2,\Gamma_2)^{|\alpha''|}\int_{\bigcup_{\tilde{P}'\in\tilde{\mathbb{P}}'}(\omega_{\tilde{P}'_1}\times\omega_{\tilde{P}'_2})_{\tilde{P}'}}\operatorname{dist}(\bar{\xi}_1,\Gamma_1)^{|\alpha'|-2}|\partial^{\alpha'}_{\bar{\xi}_1}\partial^{\alpha''}_{\bar{\xi}_2} m(\bar{\xi}_1,\bar{\xi}_2)|\,d\bar{\xi}_1\,d\bar{\xi}_2 \\
&\le \frac{1}{\ell(Q_{P''})^2}\int_{\omega_{P''_1}\times\omega_{P''_2}} B\,d\bar{\xi}_2 \le B < +\infty.
\end{aligned}\tag{2-30}$$

This ends the proof of the summable estimate (2-24). □

Observe that the rapid decay with respect to the parameters $\vec{n}_1,\vec{n}_2,\vec{n}_3\in\mathbb{Z}^2$ in (2-23) is acceptable for summation, all the functions $\Phi^{i,n'_i,\nu'}_{\tilde{P}'_i}$ ($i=1,2,3$) are L^2-normalized and are wave packets associated with the Heisenberg boxes $\tilde{P}'_i$ uniformly with respect to the parameters n'_i, and all the functions $\Phi^{j,n''_j,\nu''}_{P''_j}$ ($j=1,2,3$) are L^2-normalized and are wave packets associated with the Heisenberg boxes P''_j uniformly with respect to the parameters n''_j; therefore we only need to consider from now on the part of the trilinear form $\Lambda^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2, f_3)$ defined in (2-22) corresponding to $\vec{n}_1 = \vec{n}_2 = \vec{n}_3 = \vec{0}$,

$$\dot{\Lambda}^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2, f_3) := \int_0^1 \int_0^1 \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{C_{Q_{\vec{P}}}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi^{1,\nu}_{\vec{P}_1}\rangle \langle f_2, \Phi^{2,\nu}_{\vec{P}_2}\rangle \langle f_3, \Phi^{3,\nu}_{\vec{P}_3}\rangle \, d\nu, \tag{2-31}$$

where $C_{Q_{\vec{P}}} := C_{Q_{\vec{P}},\vec{0},\vec{0},\vec{0}}$, we have parameters $\nu = (\nu', \nu'')$ and $\Phi^{i,\nu}_{\vec{P}_i} := \Phi^{i,\vec{0},\nu}_{\vec{P}_i}$ for $i = 1, 2, 3$.

Remark 2.5. We should point out two important properties of the tritiles in $\mathbb{P}''$ (see [Muscalu and Schlag 2013; Muscalu et al. 2004b]). First, if one knows the position of P''_1, P''_2 or P''_3, then one knows precisely the positions of the other two as well. Second, if one assumes for instance that all the frequency intervals $\omega_{P''_1}$ of the P''_1 tiles intersect each other (say, they are nonlacunary about a fixed frequency ξ_0), then the frequency intervals $\omega_{P''_2}$ of the corresponding P''_2 tiles are disjoint and lacunary around ξ_0 (that is, $\operatorname{dist}(\xi_0, \omega_{P''_2}) \simeq |\omega_{P''_2}|$ for all $P'' \in \mathbb{P}''$). A similar conclusion can also be drawn for the P''_3 tiles modulo certain translations. This observation motivates the introduction of trees in Definition 3.1.

We review the following definitions from [Muscalu et al. 2004b].

Definition 2.6. A collection $\mathbb{P}$ of tritiles is called sparse if all tritiles in $\mathbb{P}$ have the same shift and the sets $\{Q_P : P \in \mathbb{P}\}$ and $\{I_P : P \in \mathbb{P}\}$ are sparse.

Definition 2.7. Let P and P' be tiles. Then we write:

(i) $P' < P$ if $I_{P'} \subsetneq I_P$ and $\omega_P \subseteq 3\omega_{P'}$;

(ii) $P' \le P$ if $P' < P$ or $P' = P$;

(iii) $P' \lesssim P$ if $I_{P'} \subseteq I_P$ and $\omega_P \subseteq 10^6\omega_{P'}$;

(iv) $P' \lesssim' P$ if $P' \lesssim P$ but $P' \nleq P$.

Definition 2.8. A collection $\mathbb{P}$ of tritiles is said to have *rank* 1 if the following properties are satisfied for all $P, P' \in \mathbb{P}$:

(i) If $P \neq P'$, then $P_j \neq P'_j$ for $1 \le j \le 3$.

(ii) If $\omega_{P_j} = \omega_{P'_j}$ for some j, then $\omega_{P_j} = \omega_{P'_j}$ for all $1 \le j \le 3$.

(iii) If $P'_j \le P_j$ for some j, then $P'_j \lesssim P_j$ for all $1 \le j \le 3$.

(iv) If in addition to $P'_j \le P_j$ one also assumes that $10^8|I_{P'}| \le |I_P|$, then one has $P'_i \lesssim' P_i$ for every $i \neq j$.

It is not difficult to see that the collection of tritiles $\mathbb{P}''$ can be written as a finite union of sparse collections of rank 1; thus we may assume further that $\mathbb{P}''$ is a sparse collection of rank 1 from now on.

The bilinear operator corresponding to the trilinear form $\dot{\Lambda}^{(2)}_{m,(lh,\mathbb{I})}(f_1, f_2, f_3)$ can be written as

$$\dot{\Pi}_{\vec{\mathbb{P}}}(f_1, f_2)(x) = \int_0^1 \int_0^1 \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{C_{Q_{\vec{P}}}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi^{1,\nu}_{\vec{P}_1}\rangle \langle f_2, \Phi^{2,\nu}_{\vec{P}_2}\rangle \Phi^{3,\nu}_{\vec{P}_3}(x) \, d\nu. \tag{2-32}$$

Since $\dot{\Pi}_{\vec{\mathbb{P}}}(f_1, f_2)$ is an average of some discrete bilinear model operators depending on the parameters $\nu = (\nu_1, \nu_2) \in [0, 1]^2$, it is enough to prove the Hölder-type L^p estimates for each of them, uniformly with respect to parameters $\nu = (\nu_1, \nu_2)$. From now on, we will do this in the particular case when the parameters $\nu = (\nu_1, \nu_2) = (0, 0)$, but the same argument works in general. By Fatou's lemma, we can also replace the summation in the definition (2-32) of $\dot{\Pi}_{\vec{\mathbb{P}}}(f_1, f_2)$ on the collection $\vec{\mathbb{P}} = \tilde{\mathbb{P}}' \times \mathbb{P}''$ by arbitrary finite collections $\tilde{\mathbb{P}}'$ and $\mathbb{P}''$ of tritiles, and prove the estimates are uniform with respect to different choices of the set $\vec{\mathbb{P}}$.

Therefore, one can reduce the bilinear operator $\dot{\Pi}_{\vec{\mathbb{P}}}$ further to the discrete bilinear model operator $\Pi_{\vec{\mathbb{P}}}$ defined by

$$\Pi_{\vec{\mathbb{P}}}(f_1, f_2)(x) := \sum_{\vec{P} \in \vec{\mathbb{P}}} \frac{C_{Q_{\vec{P}}}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi^1_{\vec{P}_1} \rangle \langle f_2, \Phi^2_{\vec{P}_2} \rangle \Phi^3_{\vec{P}_3}(x), \tag{2-33}$$

where $\Phi^j_{\vec{P}_j} := \Phi^{j,(0,0)}_{\vec{P}_j}$ for $j = 1, 2, 3$, respectively, $\vec{\mathbb{P}} = \tilde{\mathbb{P}}' \times \mathbb{P}''$ with an arbitrary finite collection $\tilde{\mathbb{P}}'$ of tritiles and an arbitrary finite sparse collection $\mathbb{P}''$ of rank 1. As discussed above, we now reach a conclusion that the proof of Theorem 1.3 can be reduced to proving the following L^p estimates for discrete bilinear model operators $\Pi_{\vec{\mathbb{P}}}$:

Proposition 2.9. *If the finite set $\vec{\mathbb{P}}$ is chosen arbitrarily, as above, then the operator $\Pi_{\vec{\mathbb{P}}}$ given by* (2-33) *maps $L^{p_1}(\mathbb{R}^2) \times L^{p_2}(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$ boundedly for any $1 < p_1, p_2 \leq \infty$ satisfying $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{2}{3} < p < \infty$. Moreover, the implicit constants in the bounds depend only on p_1, p_2, p, B and are independent of the particular choice of the finite collection $\vec{\mathbb{P}}$.*

2A2. *Discretization for bilinear, biparameter operators $T^{(2)}_{\tilde{m}^\varepsilon}$.* We will use the discretization procedure as follows. First, we need to decompose the symbol $\tilde{m}^\varepsilon(\xi)$ in a natural way. To this end, for both the spatial variables x_i $(i = 1, 2)$, we decompose the regions $\{\bar{\xi}_i = (\xi^i_1, \xi^i_2) \in \mathbb{R}^2 : \xi^i_1 \neq \xi^i_2\}$ by using Whitney squares with respect to the singularity lines $\Gamma_i = \{\xi^i_1 = \xi^i_2\}$ $(i = 1, 2)$ respectively. Since the Whitney dyadic square decomposition for the x_2 direction has already been described in (2-11), (2-12), (2-13) and (2-14) in Section 2A1, we only need to discuss the Whitney decomposition with respect to the singularity line Γ_1 in the x_1 direction.

To be specific, we consider the collection $\mathbb{Q}'$ of all shifted dyadic squares $Q' = Q'_1 \times Q'_2$ satisfying

$$Q' \subseteq \{(\xi^1_1, \xi^1_2) \in \mathbb{R}^2 : \xi^1_1 \neq \xi^1_2\}, \quad \operatorname{dist}(Q', \Gamma_1) \simeq 10^4 \operatorname{diam}(Q'). \tag{2-34}$$

We can split the collection $\mathbb{Q}'$ into two disjoint subcollections, that is, define

$$\mathbb{Q}'_{\mathrm{I}} := \{Q' \in \mathbb{Q}' : Q' \subseteq \{\xi^1_1 < \xi^1_2\}\}, \quad \mathbb{Q}'_{\mathrm{II}} := \{Q' \in \mathbb{Q}' : Q' \subseteq \{\xi^1_1 > \xi^1_2\}\}. \tag{2-35}$$

Since the set of squares $\{\frac{7}{10} Q' : Q' \in \mathbb{Q}'\}$ also forms a finitely overlapping cover of the region $\{\xi^1_1 \neq \xi^1_2\}$, we can apply a standard partition of unity and write the symbol $\chi_{\{\xi^1_1 \neq \xi^1_2\}}$ as

$$\chi_{\{\xi^1_1 \neq \xi^1_2\}} = \sum_{Q' \in \mathbb{Q}'} \phi_{Q'}(\xi^1_1, \xi^1_2) = \left(\sum_{Q' \in \mathbb{Q}'_{\mathrm{I}}} + \sum_{Q' \in \mathbb{Q}'_{\mathrm{II}}} \right) \phi_{Q'}(\xi^1_1, \xi^1_2) = \chi_{\{\xi^1_1 < \xi^1_2\}} + \chi_{\{\xi^1_1 > \xi^1_2\}}, \tag{2-36}$$

where each $\phi_{Q'}$ is a smooth bump function adapted to Q' and supported in $\frac{8}{10} Q'$.

Notice that, by splitting the symbol $\tilde{m}^\varepsilon(\xi)$, we can decompose the operator $T^{(2)}_{\tilde{m}^\varepsilon}$ correspondingly into a finite sum of several parts, and we only need to discuss one of them in detail. From the decompositions

(2-13) and (2-36), we obtain that

$$\begin{aligned}\tilde{m}^{\varepsilon}(\bar{\xi}_1,\bar{\xi}_2) &= \Bigg(\sum_{\substack{Q'\in\mathbb{Q}'_{\mathrm{I}}\\ Q''\in\mathbb{Q}''_{\mathrm{I}}}}+\sum_{\substack{Q'\in\mathbb{Q}'_{\mathrm{I}}\\ Q''\in\mathbb{Q}''_{\mathrm{II}}}}+\sum_{\substack{Q'\in\mathbb{Q}'_{\mathrm{II}}\\ Q''\in\mathbb{Q}''_{\mathrm{I}}}}+\sum_{\substack{Q'\in\mathbb{Q}'_{\mathrm{II}}\\ Q''\in\mathbb{Q}''_{\mathrm{II}}}}\Bigg)\phi_{Q'}(\xi_1^1,\xi_2^1)\phi_{Q''}(\xi_1^2,\xi_2^2)\cdot\tilde{m}^{\varepsilon}(\bar{\xi}_1,\bar{\xi}_2)\\ &=: \tilde{m}^{\varepsilon}_{\mathrm{I},\mathrm{I}}(\bar{\xi}_1,\bar{\xi}_2)+\tilde{m}^{\varepsilon}_{\mathrm{I},\mathrm{II}}(\bar{\xi}_1,\bar{\xi}_2)+\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{I}}(\bar{\xi}_1,\bar{\xi}_2)+\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{II}}(\bar{\xi}_1,\bar{\xi}_2).\end{aligned}\tag{2-37}$$

One can easily see that we only need to discuss in detail one term in the decomposition (2-37), since the other terms can be treated in the same way. Without loss of generality, we will only consider the third term in (2-37), which can be written as

$$\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{I}}(\bar{\xi}_1,\bar{\xi}_2) := \sum_{\substack{Q'\in\mathbb{Q}'_{\mathrm{II}}\\ Q''\in\mathbb{Q}''_{\mathrm{I}}}} \tilde{m}^{\varepsilon}(\bar{\xi}_1,\bar{\xi}_2)\phi_{Q'}(\xi_1^1,\xi_2^1)\phi_{Q''}(\xi_1^2,\xi_2^2).\tag{2-38}$$

In other words, we only need to consider the bilinear operator $T^{(2)}_{\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{I}}}$ given by

$$T^{(2)}_{\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{I}}}(f_1,f_2)(x) := \sum_{\substack{Q'\in\mathbb{Q}'_{\mathrm{II}}\\ Q''\in\mathbb{Q}''_{\mathrm{I}}}} \int_{\mathbb{R}^4} \tilde{m}^{\varepsilon}(\xi)\phi_{Q'}(\bar{\xi}_1)\phi_{Q''}(\bar{\xi}_2)\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)e^{2\pi i x\cdot(\xi_1+\xi_2)}\,d\xi\tag{2-39}$$

from now on, and the proof of Theorem 1.5 can be reduced to proving the following L^p estimates for $T^{(2)}_{\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{I}}}$:

$$\|T^{(2)}_{\tilde{m}^{\varepsilon}_{\mathrm{II},\mathrm{I}}}(f_1,f_2)\|_{L^p(\mathbb{R}^2)} \lesssim_{\varepsilon,p,p_1,p_2} \|f_1\|_{L^{p_1}(\mathbb{R}^2)}\cdot\|f_2\|_{L^{p_2}(\mathbb{R}^2)}\tag{2-40}$$

as long as $1<p_1,\,p_2\le\infty$ and $0<\frac{1}{p}=\frac{1}{p_1}+\frac{1}{p_2}<\frac{3}{2}$.

Observe that there exist bump functions $\phi_{Q'_i,i}$ ($i=1,2$) adapted to the shifted dyadic interval Q'_i such that $\operatorname{supp}\phi_{Q'_i,i}\subseteq\frac{9}{10}Q'_i$ and $\phi_{Q'_i,i}\equiv 1$ on $\frac{8}{10}Q'_i$ ($i=1,2$) respectively, and $\operatorname{supp}\phi_{Q'}\subseteq\frac{8}{10}Q'$, so one has $\phi_{Q'_1,1}\cdot\phi_{Q'_2,2}\equiv 1$ on $\operatorname{supp}\phi_{Q'}$. Since $\xi_1^1\in\operatorname{supp}\phi_{Q'_1,1}\subseteq\frac{9}{10}Q'_1$ and $\xi_2^1\in\operatorname{supp}\phi_{Q'_2,2}\subseteq\frac{9}{10}Q'_2$, it follows that $-\xi_1^1-\xi_2^1\in-\frac{9}{10}Q'_1-\frac{9}{10}Q'_2$, and, as a consequence, one can find a shifted dyadic interval Q'_3 with the property that $-\frac{9}{10}Q'_1-\frac{9}{10}Q'_2\subseteq\frac{7}{10}Q'_3$ and also satisfying $|Q'_1|=|Q'_2|\simeq|Q'_3|$. In particular, there exists a bump function $\phi_{Q'_3,3}$ adapted to Q'_3 and supported in $\frac{9}{10}Q'_3$ such that $\phi_{Q'_3,3}\equiv 1$ on $-\frac{9}{10}Q'_1-\frac{9}{10}Q'_2$. Recall that the smooth functions $\phi_{Q''_j,j}$ ($j=1,2,3$) and shifted dyadic intervals Q''_3 have already been defined in Section 2A1.

We denote by $\boldsymbol{Q}'$ the collection of all shifted dyadic quasicubes $Q':=Q'_1\times Q'_2\times Q'_3$ with $Q'_1\times Q'_2\in\mathbb{Q}'_{\mathrm{II}}$ and Q'_3 defined as above, and denote by $\boldsymbol{Q}''$ the collection of all shifted dyadic quasicubes $Q'':=Q''_1\times Q''_2\times Q''_3$ with $Q''_1\times Q''_2\in\mathbb{Q}''_{\mathrm{I}}$ and Q''_3 defined in Section 2A1.

In fact, it is not difficult to see that the collections $\boldsymbol{Q}'$ and $\boldsymbol{Q}''$ can be split into a sum of finitely many sparse collection of shifted dyadic quasicubes. Therefore, we can assume from now on that the collections $\boldsymbol{Q}'$ and $\boldsymbol{Q}''$ are sparse.

Assuming this, we then observe that, for any Q' in such a sparse collection $\boldsymbol{Q}'$, there exists a unique shifted dyadic cube $\tilde{Q}'$ in $\mathbb{R}^3$ such that $Q'\subseteq\frac{7}{10}\tilde{Q}'$ and with the property that $\operatorname{diam}(Q')\simeq\operatorname{diam}(\tilde{Q}')$. This allows us in particular to assume further that $\boldsymbol{Q}'$ is a sparse collection of shifted dyadic cubes (that is,

$|Q'_1| = |Q'_2| = |Q'_3| = \ell(Q')$). Similarly, we can also assume that $\boldsymbol{Q}''$ is a sparse collection of shifted dyadic cubes.

Now consider the trilinear form $\Lambda^{(2)}_{\tilde{m}^{\varepsilon}_{\mathbb{0},\mathbb{0}}}(f_1, f_2, f_3)$ associated to $T^{(2)}_{\tilde{m}^{\varepsilon}_{\mathbb{0},\mathbb{0}}}(f_1, f_2)$, which can be written as

$$\begin{aligned}\Lambda^{(2)}_{\tilde{m}^{\varepsilon}_{\mathbb{0},\mathbb{0}}}&(f_1, f_2, f_3)\\ &:= \int_{\mathbb{R}^2} T^{(2)}_{\tilde{m}^{\varepsilon}_{\mathbb{0},\mathbb{0}}}(f_1, f_2)(x) f_3(x)\,dx\\ &= \sum_{\substack{Q'\in\boldsymbol{Q}'\\ Q''\in\boldsymbol{Q}''}} \int_{\xi_1+\xi_2+\xi_3=0} \tilde{m}^{\varepsilon}_{Q',Q''}(\xi_1,\xi_2,\xi_3) \prod_{j=1}^{3} (f_j * (\check{\phi}_{Q'_j,j}\otimes\check{\phi}_{Q''_j,j}))^{\wedge}(\xi_j)\,d\xi_1\,d\xi_2\,d\xi_3,\end{aligned} \tag{2-41}$$

where $\xi_i = (\xi_i^1, \xi_i^2)$ for $i = 1, 2, 3$, while

$$\tilde{m}^{\varepsilon}_{Q',Q''}(\xi_1,\xi_2,\xi_3) := \tilde{m}^{\varepsilon}(\xi_1,\xi_2)\cdot((\phi_{Q'_1\times Q'_2}\cdot\tilde{\phi}_{Q'_3,3})\otimes(\phi_{Q''_1\times Q''_2}\cdot\tilde{\phi}_{Q''_3,3}))(\xi_1,\xi_2,\xi_3), \tag{2-42}$$

where $\tilde{\phi}_{Q'_3,3}$ is an appropriate smooth function of ξ_3^1 which equals 1 on $\operatorname{supp}\phi_{Q'_3,3}$ and is supported on a slightly larger interval (with a constant magnification independent of $\ell(Q')$) than $\operatorname{supp}\phi_{Q'_3,3}$, and $\tilde{\phi}_{Q''_3,3}$ is an appropriate smooth function of ξ_3^2 which equals 1 on $\operatorname{supp}\phi_{Q''_3,3}$ and is supported on a slightly larger interval (with a constant magnification independent of $\ell(Q'')$) than $\operatorname{supp}\phi_{Q''_3,3}$. We can decompose $\tilde{m}^{\varepsilon}_{Q',Q''}(\xi_1,\xi_2,\xi_3)$ as a Fourier series,

$$\tilde{m}^{\varepsilon}_{Q',Q''}(\xi_1,\xi_2,\xi_3) = \sum_{\vec{l}_1,\vec{l}_2,\vec{l}_3\in\mathbb{Z}^2} \tilde{C}^{\varepsilon,Q',Q''}_{\vec{l}_1,\vec{l}_2,\vec{l}_3} e^{2\pi i(l'_1,l'_2,l'_3)\cdot(\xi_1^1,\xi_2^1,\xi_3^1)/\ell(Q')} e^{2\pi i(l''_1,l''_2,l''_3)\cdot(\xi_1^2,\xi_2^2,\xi_3^2)/\ell(Q'')}, \tag{2-43}$$

where the Fourier coefficients $C^{\varepsilon,Q',Q''}_{\vec{l}_1,\vec{l}_2,\vec{l}_3}$ are given by

$$\begin{aligned}\tilde{C}^{\varepsilon,Q',Q''}_{\vec{l}_1,\vec{l}_2,\vec{l}_3} = \int_{\mathbb{R}^6} \tilde{m}^{\varepsilon}_{Q',Q''}\big((\ell(Q')\xi_1^1, \ell(Q'')\xi_1^2), (\ell(Q')\xi_2^1, \ell(Q'')\xi_2^2), (\ell(Q')\xi_3^1, \ell(Q'')\xi_3^2)\big)&\\ \times e^{-2\pi i(\vec{l}_1\cdot\xi_1+\vec{l}_2\cdot\xi_2+\vec{l}_3\cdot\xi_3)}\,d\xi_1\,d\xi_2\,d\xi_3.&\end{aligned} \tag{2-44}$$

Then, by a straightforward calculation, we can rewrite (2-41) as

$$\begin{aligned}\Lambda^{(2)}_{\tilde{m}^{\varepsilon}_{\mathbb{0},\mathbb{0}}}&(f_1, f_2, f_3)\\ &= \sum_{\substack{Q'\in\boldsymbol{Q}'\\ Q''\in\boldsymbol{Q}''}} \sum_{\vec{l}_1,\vec{l}_2,\vec{l}_3\in\mathbb{Z}^2} \tilde{C}^{\varepsilon,Q',Q''}_{\vec{l}_1,\vec{l}_2,\vec{l}_3} \int_{\mathbb{R}^2} \prod_{i=1}^{3} (f_i * (\check{\phi}_{Q'_i,i}\otimes\check{\phi}_{Q''_i,i}))\left(x - \left(\frac{l'_i}{\ell(Q')}, \frac{l''_i}{\ell(Q'')}\right)\right) dx.\end{aligned} \tag{2-45}$$

Now we define $\phi^{l'_i}_{Q'_i,i} := e^{2\pi i l'_i \xi_i^1/\ell(Q')}\cdot\phi_{Q'_i,i}$ and $\phi^{l''_i}_{Q''_i,i} := e^{2\pi i l''_i \xi_i^2/\ell(Q'')}\cdot\phi_{Q''_i,i}$ for $i = 1, 2, 3$. Since any $Q' \in \boldsymbol{Q}'$ and $Q'' \in \boldsymbol{Q}''$ are shifted dyadic cubes, there exist integers $k', k'' \in \mathbb{Z}$ such that $\ell(Q') = |Q'_1| = |Q'_2| = |Q'_3| = 2^{k'}$ and $\ell(Q'') = |Q''_1| = |Q''_2| = |Q''_3| = 2^{k''}$, respectively. By splitting the integral region $\mathbb{R}^2$ into the union of unit squares, the L^2-normalization procedure and simple calculations, we can rewrite (2-45) as

$$\begin{aligned}\Lambda^{(2)}_{\tilde{m}^{\varepsilon}_{\mathbb{I},\mathbb{I}}}&(f_1,f_2,f_3)\\ &=\sum_{\vec{l}_1,\vec{l}_2,\vec{l}_3\in\mathbb{Z}^2}\sum_{\substack{Q'\in\boldsymbol{Q}'\\ Q''\in\boldsymbol{Q}''}}\int_0^1\int_0^1\sum_{\substack{I'\text{ dyadic}\\ |I'|=2^{-k'}}}\sum_{\substack{I''\text{ dyadic}\\ |I''|=2^{-k''}}}\frac{\tilde{C}^{\varepsilon,Q',Q''}_{\vec{l}_1,\vec{l}_2,\vec{l}_3}}{|I'|^{\frac12}\times|I''|^{\frac12}}\prod_{i=1}^{3}\langle f_i,\check{\phi}^{l'_i,\lambda'}_{I',Q'_i,i}\otimes\check{\phi}^{l''_i,\lambda''}_{I'',Q''_i,i}\rangle\,d\lambda'\,d\lambda''\\ &=:\sum_{\vec{l}_1,\vec{l}_2,\vec{l}_3\in\mathbb{Z}^2}\int_0^1\int_0^1\sum_{\vec{P}:=P'\otimes P''\in\vec{\mathbb{P}}}\frac{\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3}}{|I_{\vec{P}}|^{\frac12}}\prod_{i=1}^{3}\langle f_i,\Phi^{i,\vec{l}_i,\lambda}_{\vec{P}_i}\rangle\,d\lambda,\end{aligned}\tag{2-46}$$

where we have:

- Fourier coefficients $\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3}:=\tilde{C}^{\varepsilon,Q',Q''}_{\vec{l}_1,\vec{l}_2,\vec{l}_3}$;
- tritiles $P':=(P'_1,P'_2,P'_3)$ and $P'':=(P''_1,P''_2,P''_3)$;
- tiles $P'_i:=I_{P'_i}\times\omega_{P'_i}$, where $I_{P'_i}:=I'=2^{-k'}[n',n'+1]=:I_{P'}$ and the frequency intervals are $\omega_{P'_i}:=Q'_i$ for $i=1,2,3$;
- tiles $P''_j:=I_{P''_j}\times\omega_{P''_j}$, where $I_{P''_j}:=I''=2^{-k''}[n'',n''+1]=:I_{P''}$ and the frequency intervals are $\omega_{P''_j}:=Q''_j$ for $j=1,2,3$;
- frequency cubes $Q_{P'}:=\omega_{P'_1}\times\omega_{P'_2}\times\omega_{P'_3}$ and $Q_{P''}:=\omega_{P''_1}\times\omega_{P''_2}\times\omega_{P''_3}$;
- $\mathbb{P}'$ denotes a collection of such tritiles P' and $\mathbb{P}''$ denotes a collection of such tritiles P'';
- bitiles $\vec{P}_1$, $\vec{P}_2$ and $\vec{P}_3$ defined by
$$\vec{P}_i:=(P'_i,P''_i)=(2^{-k'}[n',n'+1]\times Q'_i,\,2^{-k''}[n'',n''+1]\times Q''_i)\quad\text{for } i=1,2,3;$$
- the biparameter tritile $\vec{P}:=P'\otimes P''=(\vec{P}_1,\vec{P}_2,\vec{P}_3)$;
- rectangles $I_{\vec{P}_i}:=I_{P'_i}\times I_{P''_i}=I_{P'}\times I_{P''}=:I_{\vec{P}}$ for $i=1,2,3$, and hence $|I_{\vec{P}}|=|I_{P'}\times I_{P''}|=|I_{\vec{P}_1}|=|I_{\vec{P}_2}|=|I_{\vec{P}_3}|=2^{-k'}\cdot2^{-k''}$;
- the double frequency cube $Q_{\vec{P}}:=(Q_{P'},Q_{P''})=(\omega_{P'_1}\times\omega_{P'_2}\times\omega_{P'_3},\,\omega_{P''_1}\times\omega_{P''_2}\times\omega_{P''_3})$;
- $\vec{\mathbb{P}}:=\mathbb{P}'\times\mathbb{P}''$ denotes a collection of such biparameter tritiles $\vec{P}$;
- L^2-normalized wave packets $\Phi^{i,l'_i,\lambda'}_{P'_i}$ associated with the Heisenberg boxes P'_i defined by
$$\Phi^{i,l'_i,\lambda'}_{P'_i}(x_1):=\check{\phi}^{l'_i,\lambda'}_{I',Q'_i,i}(x_1):=2^{-k'/2}\overline{\check{\phi}^{l'_i}_{Q'_i,i}(2^{-k'}(n'+\lambda')-x_1)}\quad\text{for } i=1,2,3,$$
- L^2-normalized wave packets $\Phi^{i,l''_i,\lambda''}_{P''_i}$ associated with the Heisenberg boxes P''_i defined by
$$\Phi^{i,l''_i,\lambda''}_{P''_i}(x_2):=\check{\phi}^{l''_i,\lambda''}_{I'',Q''_i,i}(x_2):=2^{-k''/2}\overline{\check{\phi}^{l''_i}_{Q''_i,i}(2^{-k''}(n''+\lambda'')-x_2)}\quad\text{for } i=1,2,3,$$
- smooth bump functions $\Phi^{i,\vec{l}_i,\lambda}_{\vec{P}_i}:=\Phi^{i,l'_i,\lambda'}_{P'_i}\otimes\Phi^{i,l''_i,\lambda''}_{P''_i}$ for $i=1,2,3$.

We have the following rapid decay estimates of the Fourier coefficients $\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3}$ with respect to the parameters $\vec{l}_1,\vec{l}_2,\vec{l}_3\in\mathbb{Z}^2$:

Lemma 2.10. *The Fourier coefficients $\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3}$ satisfy estimates*

$$|\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3}| \lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{l}_j|)^M} \cdot \langle \log_2 \ell(Q_{P'}) \rangle^{-(1+\varepsilon)} \tag{2-47}$$

for any biparameter tritile $\vec{P} \in \vec{\mathbb{P}}$, where M is sufficiently large.

Proof. Let $\ell(Q_{P'}) = 2^{k'}$ and $\ell(Q_{P''}) = 2^{k''}$ for $k', k'' \in \mathbb{Z}$. For any $\varepsilon > 0$, $\vec{l}_1, \vec{l}_2, \vec{l}_3 \in \mathbb{Z}^2$ and $\vec{P} \in \vec{\mathbb{P}}$, we deduce from (2-42) and (2-44) that

$$\begin{aligned}\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3} &= \int_{\mathbb{R}^6} \tilde{m}^{\varepsilon}_{Q_{P'},Q_{P''}}((2^{k'}\xi_1^1, 2^{k''}\xi_1^2), (2^{k'}\xi_2^1, 2^{k''}\xi_2^2), (2^{k'}\xi_3^1, 2^{k''}\xi_3^2)) e^{-2\pi i(\vec{l}_1\cdot\xi_1+\vec{l}_2\cdot\xi_2+\vec{l}_3\cdot\xi_3)}\, d\xi_1\, d\xi_2\, d\xi_3,\end{aligned} \tag{2-48}$$

where

$$\begin{aligned}&\tilde{m}^{\varepsilon}_{Q_{P'},Q_{P''}}((2^{k'}\xi_1^1, 2^{k''}\xi_1^2), (2^{k'}\xi_2^1, 2^{k''}\xi_2^2), (2^{k'}\xi_3^1, 2^{k''}\xi_3^2)) \\ &\quad := \tilde{m}^{\varepsilon}(2^{k'}\bar{\xi}_1, 2^{k''}\bar{\xi}_2)\phi_{\omega_{P_1'}\times\omega_{P_2'}}(2^{k'}\bar{\xi}_1)\tilde{\phi}_{\omega_{P_3'},3}(2^{k'}\xi_3^1)\phi_{\omega_{P_1''}\times\omega_{P_2''}}(2^{k''}\bar{\xi}_2)\tilde{\phi}_{\omega_{P_3''},3}(2^{k''}\xi_3^2).\end{aligned} \tag{2-49}$$

Since $\operatorname{supp}(\phi_{\omega_{P_1'}\times\omega_{P_2'}}(\bar{\xi}_1)\tilde{\phi}_{\omega_{P_3'},3}(\xi_3^1)\phi_{\omega_{P_1''}\times\omega_{P_2''}}(\bar{\xi}_2)\tilde{\phi}_{\omega_{P_3''},3}(\xi_3^2)) \subseteq Q_{P'} \times Q_{P''}$, we have that

$$\operatorname{supp}(\phi_{\omega_{P_1'}\times\omega_{P_2'}}(2^{k'}\bar{\xi}_1)\tilde{\phi}_{\omega_{P_3'},3}(2^{k'}\xi_3^1)\phi_{\omega_{P_1''}\times\omega_{P_2''}}(2^{k''}\bar{\xi}_2)\tilde{\phi}_{\omega_{P_3''},3}(2^{k''}\xi_3^2)) \subseteq Q^0_{P'} \times Q^0_{P''},$$

where the cubes $Q^0_{P'}$ and $Q^0_{P''}$ are defined by

$$Q^0_{P'} = \omega^0_{P_1'} \times \omega^0_{P_2'} \times \omega^0_{P_3'} := \{(\xi_1^1, \xi_2^1, \xi_3^1) \in \mathbb{R}^3 : (2^{k'}\xi_1^1, 2^{k'}\xi_2^1, 2^{k'}\xi_3^1) \in Q_{P'}\}, \tag{2-50}$$

$$Q^0_{P''} = \omega^0_{P_1''} \times \omega^0_{P_2''} \times \omega^0_{P_3''} := \{(\xi_1^2, \xi_2^2, \xi_3^2) \in \mathbb{R}^3 : (2^{k''}\xi_1^2, 2^{k''}\xi_2^2, 2^{k''}\xi_3^2) \in Q_{P''}\} \tag{2-51}$$

and satisfy $|Q^0_{P'}| \simeq |Q^0_{P''}| \simeq 1$. From the properties of the Whitney squares we constructed above, one obtains that $\operatorname{dist}(2^{k'}\bar{\xi}_1, \Gamma_1) \simeq 2^{k'}$ for any $\bar{\xi}_1 \in \omega^0_{P_1'} \times \omega^0_{P_2'}$ and $\operatorname{dist}(2^{k''}\bar{\xi}_2, \Gamma_2) \simeq 2^{k''}$ for any $\bar{\xi}_2 \in \omega^0_{P_1''} \times \omega^0_{P_2''}$.

By taking advantage of the estimates (1-13) for symbol $\tilde{m}^{\varepsilon}(\bar{\xi})$, one can deduce from (2-48), (2-49) and integrating by parts sufficiently many times that

$$\begin{aligned}&|\tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{l}_1,\vec{l}_2,\vec{l}_3}| \\ &\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{l}_j|)^M} \\ &\qquad \times \int_{Q^0_{P'}\times Q^0_{P''}} \big|\partial^{\alpha_1}_{\xi_1}\partial^{\alpha_2}_{\xi_2}\partial^{\alpha_3}_{\xi_3}\big[\tilde{m}^{\varepsilon}_{Q_{P'},Q_{P''}}((2^{k'}\xi_1^1, 2^{k''}\xi_1^2), (2^{k'}\xi_2^1, 2^{k''}\xi_2^2), (2^{k'}\xi_3^1, 2^{k''}\xi_3^2))\big]\big|\, d\xi_1\, d\xi_2\, d\xi_3 \\ &\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{l}_j|)^M} \int_{\omega^0_{P_1''}\times\omega^0_{P_2''}} \operatorname{dist}(2^{k''}\bar{\xi}_2, \Gamma_2)^{|\alpha''|} \int_{\omega^0_{P_1'}\times\omega^0_{P_2'}} \operatorname{dist}(2^{k'}\bar{\xi}_1, \Gamma_1)^{|\alpha'|} |\partial^{\alpha'}_{\bar{\xi}_1}\partial^{\alpha''}_{\bar{\xi}_2}\tilde{m}^{\varepsilon}(2^{k'}\bar{\xi}_1, 2^{k''}\bar{\xi}_2)|\, d\bar{\xi}_1\, d\bar{\xi}_2\end{aligned}$$

$$\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{l}_j|)^M} \cdot 2^{-2k'} 2^{-2k''} \int_{\omega_{P_1''}\times\omega_{P_2''}} \int_{\omega_{P_1'}\times\omega_{P_2'}} \operatorname{dist}(\bar{\xi}_2, \Gamma_2)^{|\alpha''|} \cdot \operatorname{dist}(\bar{\xi}_1, \Gamma_1)^{|\alpha'|} |\partial_{\bar{\xi}_1}^{\alpha'} \partial_{\bar{\xi}_2}^{\alpha''} \tilde{m}^{\varepsilon}(\bar{\xi}_1, \bar{\xi}_2)| \, d\bar{\xi}_1 \, d\bar{\xi}_2$$

$$\lesssim \prod_{j=1}^{3} \frac{1}{(1+|\vec{l}_j|)^M} \cdot \langle \log_2 \ell(Q_{P'}) \rangle^{-(1+\varepsilon)},$$

where the multi-indices $\alpha_i := (\alpha_i^1, \alpha_i^2)$ for $i = 1, 2, 3$ and $|\alpha_1| = |\alpha_2| = |\alpha_3| = M$ are sufficiently large, the multi-indices $\alpha' := (\alpha_1', \alpha_2', \alpha_3')$, $\alpha'' := (\alpha_1'', \alpha_2'', \alpha_3'')$ with $\alpha_i' \leq \alpha_i^1$ and $\alpha_j'' \leq \alpha_j^2$ for $i, j = 1, 2, 3$. This ends our proof of the estimates (2-47). □

Note that the rapid decay with respect to the parameters $\vec{l}_1, \vec{l}_2, \vec{l}_3 \in \mathbb{Z}^2$ in (2-47) is acceptable for summation, all the functions $\Phi_{P_i'}^{i,l_i',\lambda'}$ $(i = 1, 2, 3)$ are L^2-normalized and are wave packets associated with the Heisenberg boxes P_i' uniformly with respect to the parameters l_i', and all the functions $\Phi_{P_j''}^{j,l_j'',\lambda''}$ $(j = 1, 2, 3)$ are L^2-normalized and are wave packets associated with the Heisenberg boxes P_j'' uniformly with respect to the parameters l_j'', therefore we only need to consider from now on the part of the trilinear form $\Lambda_{\tilde{m}_{\mathbb{I}\mathbb{I},\mathbb{I}}^{\varepsilon}}^{(2)}(f_1, f_2, f_3)$ defined in (2-46) corresponding to $\vec{l}_1 = \vec{l}_2 = \vec{l}_3 = \vec{0}$,

$$\dot{\Lambda}_{\tilde{m}_{\mathbb{I}\mathbb{I},\mathbb{I}}^{\varepsilon}}^{(2)}(f_1, f_2, f_3) := \int_0^1 \int_0^1 \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{\tilde{C}_{Q_{\vec{P}}}^{\varepsilon}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi_{\vec{P}_1}^{1,\lambda} \rangle \langle f_2, \Phi_{\vec{P}_2}^{2,\lambda} \rangle \langle f_3, \Phi_{\vec{P}_3}^{3,\lambda} \rangle \, d\lambda, \tag{2-52}$$

where $\tilde{C}_{Q_{\vec{P}}}^{\varepsilon} := \tilde{C}_{Q_{\vec{P}},\vec{0},\vec{0},\vec{0}}^{\varepsilon}$, we have parameters $\lambda = (\lambda', \lambda'')$, and $\Phi_{\vec{P}_i}^{i,\lambda} := \Phi_{\vec{P}_i}^{i,\vec{0},\lambda}$ for $i = 1, 2, 3$.

The tritiles $P' = (P_1', P_2', P_3')$ in the collection $\mathbb{P}'$ also satisfy the same properties (as $P'' \in \mathbb{P}''$) described in Remark 2.5. It is not difficult to see that both the collections of tritiles $\mathbb{P}'$ and $\mathbb{P}''$ can be written as a finite union of sparse collections of rank 1; thus we may assume further that $\mathbb{P}'$ and $\mathbb{P}''$ are sparse collections of rank 1 from now on.

The bilinear operator corresponding to the trilinear form $\dot{\Lambda}_{\tilde{m}_{\mathbb{I}\mathbb{I},\mathbb{I}}^{\varepsilon}}^{(2)}(f_1, f_2, f_3)$ can be written as

$$\dot{\tilde{\Pi}}_{\vec{\mathbb{P}}}^{\varepsilon}(f_1, f_2)(x) = \int_0^1 \int_0^1 \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{\tilde{C}_{Q_{\vec{P}}}^{\varepsilon}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi_{\vec{P}_1}^{1,\lambda} \rangle \langle f_2, \Phi_{\vec{P}_2}^{2,\lambda} \rangle \Phi_{\vec{P}_3}^{3,\lambda}(x) \, d\lambda. \tag{2-53}$$

Since $\dot{\tilde{\Pi}}_{\vec{\mathbb{P}}}^{\varepsilon}(f_1, f_2)$ is an average of some discrete bilinear model operators depending on the parameters $\lambda = (\lambda_1, \lambda_2) \in [0, 1]^2$, it is enough to prove the Hölder-type L^p estimates for each of them, uniformly with respect to parameters $\lambda = (\lambda_1, \lambda_2)$. From now on, we will do this in the particular case when the parameters $\lambda = (\lambda_1, \lambda_2) = (0, 0)$, but the same argument works in general. By Fatou's lemma, we can also replace the summation in the definition (2-53) of $\dot{\tilde{\Pi}}_{\vec{\mathbb{P}}}^{\varepsilon}(f_1, f_2)$ on the collection $\vec{\mathbb{P}} = \mathbb{P}' \times \mathbb{P}''$ by arbitrary finite collections $\mathbb{P}'$ and $\mathbb{P}''$ of tritiles, and prove the estimates are uniform with respect to different choices of the set $\vec{\mathbb{P}}$.

Definition 2.11. A finite collection $\vec{\mathbb{P}} = \mathbb{P}' \times \mathbb{P}''$ of biparameter tritiles is said to be sparse and of rank 1 if both the finite collections $\mathbb{P}'$ and $\mathbb{P}''$ are sparse and of rank 1.

Therefore, one can reduce the bilinear operator $\dot{\tilde{\Pi}}^{\varepsilon}_{\vec{\mathbb{P}}}$ further to the discrete bilinear model operator $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ defined by

$$\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2)(x) := \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi^1_{\vec{P}_1}\rangle\langle f_2, \Phi^2_{\vec{P}_2}\rangle \Phi^3_{\vec{P}_3}(x), \tag{2-54}$$

where $\Phi^j_{\vec{P}_j} := \Phi^{j,(0,0)}_{\vec{P}_j}$ for $j = 1, 2, 3$, and the finite set $\vec{\mathbb{P}} = \mathbb{P}' \times \mathbb{P}''$ is an arbitrary sparse collection (of biparameter tritiles) of rank 1. As discussed above, we now reach a conclusion that the proof of Theorem 1.5 can be reduced to proving the following L^p estimates for discrete bilinear model operators $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$:

Proposition 2.12. *If the finite set $\vec{\mathbb{P}}$ is an arbitrary sparse collection of rank* 1, *then the operator $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ given by* (2-54) *maps $L^{p_1}(\mathbb{R}^2) \times L^{p_2}(\mathbb{R}^2) \to L^p(\mathbb{R}^2)$ boundedly for any $1 < p_1, p_2 \le \infty$ satisfying $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{2}{3} < p < \infty$. Moreover, the implicit constants in the bounds depend only on ε, p_1, p_2, p and are independent of the particular finite sparse collection $\vec{\mathbb{P}}$ of rank* 1.

2B. ***Multilinear interpolations.*** First, let's review the following terminologies and definitions of multilinear interpolation arguments:

Definition 2.13 [Muscalu and Schlag 2013; Muscalu et al. 2002]. An n-tuple $\beta = (\beta_1, \dots, \beta_n)$ is said to be *admissible* if and only if $\beta_j < 1$ for every $1 \le j \le n$, $\sum_{j=1}^n \beta_j = 1$ and there is at most one index j for which $\beta_j < 0$. An index j is called *good* if $\beta_j \ge 0$ and *bad* if $\beta_j < 0$. A *good tuple* is an admissible tuple that contains only good indices; a *bad tuple* is an admissible tuple that contains precisely one bad index.

Definition 2.14 [Muscalu et al. 2002]. Let E, E' be sets of finite measure. We say that E' is a *major subset* of E if $E' \subseteq E$ and $|E'| \ge \frac{1}{2}|E|$.

Definition 2.15 [Muscalu and Schlag 2013; Muscalu et al. 2002]. If $\beta = (\beta_1, \dots, \beta_n)$ is an admissible tuple, we say that an n-linear form Λ is of *restricted weak type* β if and only if, for every sequence $E_1, \dots, E_n$ of measurable sets with positive and finite measure, there exists a major subset E'_j of E_j for the bad index j (if there is one) such that

$$|\Lambda(f_1, \dots, f_n)| \lesssim |E_1|^{\beta_1} \cdots |E_n|^{\beta_n} \tag{2-55}$$

for all measurable functions $|f_i| \le \chi_{E'_i}$ $(i = 1, \dots, n)$, where we adopt the convention $E'_i = E_i$ for good indices i. If β is bad with bad index j_0, and it happens that one can choose the major subset $E'_{j_0} \subseteq E_{j_0}$ in a way that depends only on the measurable sets $E_1, \dots, E_n$ and not on β, we say that Λ is of *uniformly restricted weak type*.

Definition 2.16 [Muscalu and Schlag 2013]. Let $1 < p_1, p_2 \le \infty$ and $0 < p < \infty$ be such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. An arbitrary bilinear operator T is said to be of the restricted weak type (p_1, p_2, p) if and only if, for all measurable sets E_1, E_2, E of finite measure, there exists $E' \subseteq E$ with $|E'| \simeq |E|$ such that

$$\left| \int_{\mathbb{R}^d} T(f_1, f_2)(x) f(x)\, dx \right| \lesssim |E_1|^{1/p_1} |E_2|^{1/p_2} |E'|^{1/p'} \tag{2-56}$$

for every $|f_1| \le \chi_{E_1}$, $|f_2| \le \chi_{E_2}$ and $|f| \le \chi_{E'}$.

By using multilinear interpolation (see [Grafakos and Tao 2003; Janson 1988; Muscalu and Schlag 2013; Muscalu et al. 2002]) and the symmetry of the operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$, we can reduce further the proof of Proposition 2.9 and Proposition 2.12 to proving the following restricted weak type estimates for the model operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$:

Proposition 2.17. *Let p_1 and p_2 be such that p_1 is strictly larger than 1 and arbitrarily close to 1 and p_2 is strictly smaller than 2 and arbitrarily close to 2 and such that, for $\frac{1}{p} := \frac{1}{p_1} + \frac{1}{p_2}$, one has $\frac{2}{3} < p < 1$. Then both the model operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ defined in* (2-33) *and* (2-54) *are of the restricted weak type (p_1, p_2, p). Moreover, the implicit constants in the bounds depend only on p_1, p_2, p, ε and B, and are independent of the particular choice of the finite collection $\vec{\mathbb{P}}$.*

Indeed, first we should note that, if p_1, p_2, p are as in Propositions 2.9 and 2.12, then the 3-tuple $\left(\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{p'}\right)$ lies in the interior of the convex hull of the following six extremal points: $\beta^1 := \left(-\frac{1}{2}, \frac{1}{2}, 1\right)$, $\beta^2 := \left(-\frac{1}{2}, 1, \frac{1}{2}\right)$, $\beta^3 := \left(\frac{1}{2}, -\frac{1}{2}, 1\right)$, $\beta^4 := \left(1, -\frac{1}{2}, \frac{1}{2}\right)$, $\beta^5 := \left(\frac{1}{2}, 1, -\frac{1}{2}\right)$ and $\beta^6 := \left(1, \frac{1}{2}, -\frac{1}{2}\right)$. Then, if we assume that Proposition 2.17 has been proved, from the symmetry of operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ and their adjoints we deduce that both the trilinear forms associated to bilinear operators $\Pi_{\vec{\mathbb{P}}}$ and $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ are of uniformly restricted weak type β for 3-tuples $\beta = (\beta_1, \beta_2, \beta_3)$ arbitrarily close to the six extremal points $\beta^1, \dots, \beta^6$ inside their convex hull and satisfying that, if β_j is close to $\frac{1}{2}$ for some $j = 1, 2, 3$, then β_j is strictly larger than $\frac{1}{2}$. By using the multilinear interpolation lemma, [Muscalu and Schlag 2013, Lemmas 9.4 and 9.6] or [Muscalu et al. 2002, Lemma 3.8], we first obtain restricted weak type estimates of Λ for good tuples inside the smaller convex hull of the three coordinate points $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. After that, we use the interpolation lemma [Muscalu and Schlag 2013, Lemma 9.5] or [Muscalu et al. 2002, Lemma 3.10] to obtain restricted weak type estimates of Λ for bad tuples and finally conclude that restricted weak type estimates of Λ hold for all tuples β inside the convex hull of the six extremal points $\beta^1, \dots, \beta^6$.

It only remains to convert these restricted weak type estimates into strong type estimates. To do this, one just has to apply (exactly as in [Muscalu et al. 2002]) the multilinear Marcinkiewicz interpolation theorem in [Janson 1988] in the case of good tuples and the interpolation lemma [Muscalu et al. 2002, Lemma 3.11] in the case of bad tuples. This ends the proof of Propositions 2.9 and 2.12, and, as a consequence, completes the proof of our main results, Theorems 1.3 and 1.5. Therefore, we only have the task of proving Proposition 2.17 from now on.

3. Trees, L^2 sizes and L^2 energies

3A. ***Trees.*** We should recall that, for discrete bilinear paraproducts, the frequency intervals have already been organized with the lacunary properties (see [Muscalu and Schlag 2013; Muscalu et al. 2004a; 2006]), so we could use square function and maximal function estimates to handle the corresponding terms easily, at least in the Banach case. By the properties of the collection $\mathbb{P}''$ of tritiles we have explained in Remark 2.5, we can organize our collections of tritiles $\mathbb{P}'$, $\mathbb{P}''$ into trees as in [Grafakos and Li 2004], which satisfy lacunary properties about a certain frequency. We review the following standard definitions and properties for trees from [Muscalu et al. 2004b]:

Definition 3.1. Let $\mathbb{P}$ be a sparse rank-1 collection of tritiles and $j \in \{1, 2, 3\}$. A subcollection $T \subseteq \mathbb{P}$ is called a *j-tree* if and only if there exists a tritile P_T (called the top of the tree) such that

$$P_j \leq P_{T,j} \tag{3-1}$$

for every $P \in T$.

Remark 3.2. A tree does not necessarily have to contain the corresponding top P_T. From now on, we will write I_T and $\omega_{T,j}$ for I_{P_T} and $\omega_{P_T,j}$ for $j = 1, 2, 3$. Then, we simply say that T is a tree if it is a j-tree for some $j = 1, 2, 3$.

For every given dyadic interval I_0, there are potentially many tritiles P in $\mathbb{P}'$ and $\mathbb{P}''$ with the property that $I_P = I_0$. Due to this extra degree of freedom in frequency, we have infinitely many trees in our collections $\mathbb{P}'$ and $\mathbb{P}''$. We need to estimate each of these trees separately, and then add all these estimates together, by using the almost orthogonality conditions for distinct trees. This motivates the following definition:

Definition 3.3. Let $1 \leq i \leq 3$. A finite sequence of trees $T_1, \dots, T_M$ is said to be a *chain of strongly i-disjoint trees* if and only if:

(i) $P_i \neq P'_i$ for every $P \in T_{l_1}$ and $P' \in T_{l_2}$ with $l_1 \neq l_2$.

(ii) Whenever $P \in T_{l_1}$ and $P' \in T_{l_2}$ with $l_1 \neq l_2$ are such that $2\omega_{P_i} \cap 2\omega_{P'_i} \neq \varnothing$, then if $|\omega_{P_i}| < |\omega_{P'_i}|$ one has $I_{P'} \cap I_{T_{l_1}} = \varnothing$ and if $|\omega_{P'_i}| < |\omega_{P_i}|$ one has $I_P \cap I_{T_{l_2}} = \varnothing$.

(iii) Whenever $P \in T_{l_1}$ and $P' \in T_{l_2}$ with $l_1 < l_2$ are such that $2\omega_{P_i} \cap 2\omega_{P'_i} \neq \varnothing$, then if $|\omega_{P_i}| = |\omega_{P'_i}|$ one has $I_{P'} \cap I_{T_{l_1}} = \varnothing$.

3B. *L^2 sizes and L^2 energies.* Following [Muscalu et al. 2004b], we give the definitions of standard norms on sequences of tiles:

Definition 3.4. Let $\mathbb{P}$ be a finite collection of tritiles, $j \in \{1, 2, 3\}$, and let f be an arbitrary function. We define the *size* of the sequence $(\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}$ by

$$\operatorname{size}_j\big((\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}\big) := \sup_{T\subseteq\mathbb{P}} \left(\frac{1}{|I_T|}\sum_{P\in T}|\langle f, \Phi^j_{P_j}\rangle|^2\right)^{\frac{1}{2}}, \tag{3-2}$$

where T ranges over all trees in $\mathbb{P}$ that are i-trees for some $i \neq j$. For $j = 1, 2, 3$, we define the *energy* of the sequence $(\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}$ by

$$\operatorname{energy}_j\big((\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}\big) := \sup_{n\in\mathbb{Z}} \sup_{\mathbb{T}} 2^n \left(\sum_{T\in\mathbb{T}}|I_T|\right)^{\frac{1}{2}}, \tag{3-3}$$

where now $\mathbb{T}$ ranges over all chains of strongly j-disjoint trees in $\mathbb{P}$ (which are i-trees for some $i \neq j$) having the property that

$$\left(\sum_{P\in T}|\langle f, \Phi^j_{P_j}\rangle|^2\right)^{\frac{1}{2}} \geq 2^n|I_T|^{\frac{1}{2}} \tag{3-4}$$

for all $T \in \mathbb{T}$ and such that

$$\Big(\sum_{P\in T'} |\langle f, \Phi^j_{P_j}\rangle|^2\Big)^{\frac{1}{2}} \le 2^{n+1}|I_{T'}|^{\frac{1}{2}} \tag{3-5}$$

for all subtrees $T' \subseteq T \in \mathbb{T}$.

The size measures the extent to which the sequences $(\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}$ $(j = 1, 2, 3)$ can concentrate on a single tree and should be thought of as a phase-space variant of the BMO norm. The energy is a phase-space variant of the L^2 norm. As the notation suggests, the number $\langle f, \Phi^j_{P_j}\rangle$ should be thought of as being associated with the tile P_j $(j = 1, 2, 3)$ rather than the full tritile P.

Let $\mathbb{P}$ be a finite collection of tritiles. Denote by $\Pi_{\mathbb{P}}$ the discrete bilinear operator given by

$$\Pi_{\mathbb{P}}(f_1, f_2)(x) = \sum_{P\in\mathbb{P}} \frac{1}{|I_P|^{\frac{1}{2}}} \langle f_1, \Phi^1_{P_1}\rangle\langle f_2, \Phi^2_{P_2}\rangle \Phi^3_{P_3}(x).$$

The following proposition provides a way of estimating the trilinear form associated with the bilinear operator $\Pi_{\mathbb{P}}(f_1, f_2)$. We define

$$\Lambda_{\mathbb{P}}(f_1, f_2, f_3) := \int_{\mathbb{R}} \Pi_{\mathbb{P}}(f_1, f_2)(x) f_3(x)\, dx.$$

Proposition 3.5 [Muscalu et al. 2004b]. *Let $\mathbb{P}$ be a finite collection of tritiles. Then*

$$|\Lambda_{\mathbb{P}}(f_1, f_2, f_3)| \lesssim \prod_{j=1}^{3} \big(\mathrm{size}_j\big((\langle f_j, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}\big)\big)^{\theta_j} \big(\mathrm{energy}_j\big((\langle f_j, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}\big)\big)^{1-\theta_j} \tag{3-6}$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1+\theta_2+\theta_3 = 1$; the implicit constants depend on the θ_j but are independent of the other parameters.

3C. ***Estimates for sizes and energies.*** In order to apply Proposition 3.5, we need to estimate further the sizes and energies appearing on the right-hand side of (3-6).

Lemma 3.6 [Muscalu and Schlag 2013; Muscalu et al. 2004b]. *Let $j \in \{1, 2, 3\}$ and $f \in L^2(\mathbb{R})$. Then one has*

$$\mathrm{size}_j\big((\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}\big) \lesssim \sup_{P\in\mathbb{P}} \frac{1}{|I_P|} \int_{\mathbb{R}} |f| \tilde{\chi}^M_{I_P}\, dx \tag{3-7}$$

for every $M > 0$, where the approximate cutoff function $\tilde{\chi}^M_{I_P}(x)$ equals $(1 + \mathrm{dist}(x, I_P)/|I_P|)^{-M}$ and the implicit constants depend on M.

Lemma 3.7 (Bessel-type estimates [Muscalu et al. 2004b]). *Let $j \in \{1, 2, 3\}$ and $f \in L^2(\mathbb{R})$. Then*

$$\mathrm{energy}_j\big((\langle f, \Phi^j_{P_j}\rangle)_{P\in\mathbb{P}}\big) \lesssim \|f\|_{L^2}. \tag{3-8}$$

4. Proof of Theorem 1.3

In this section, we prove Theorem 1.3 by carrying out the proof of Proposition 2.17 for model operators $\Pi_{\vec{\mathbb{P}}}$ defined in (2-33) with $\vec{\mathbb{P}} = \tilde{\mathbb{P}}' \times \mathbb{P}''$.

Fix indices p_1, p_2, p as in the hypothesis of Proposition 2.17. Fix arbitrary measurable sets E_1, E_2, E_3 of finite measure (by using the scaling invariance of $\Pi_{\vec{\mathbb{P}}}$, we can assume further that $|E_3| = 1$). Our goal is to find $E_3' \subseteq E_3$ with $|E_3'| \simeq |E_3| = 1$ such that, when $|f_1| \leq \chi_{E_1}$, $|f_2| \leq \chi_{E_2}$ and $|f_3| \leq \chi_{E_3'}$, the trilinear form $\Lambda_{\vec{\mathbb{P}}}(f_1, f_2, f_3)$ defined by

$$\Lambda_{\vec{\mathbb{P}}}(f_1, f_2, f_3) := \int_{\mathbb{R}^2} \Pi_{\vec{\mathbb{P}}}(f_1, f_2)(x) f_3(x)\, dx \tag{4-1}$$

satisfies the estimate

$$|\Lambda_{\vec{\mathbb{P}}}(f_1, f_2, f_3)| = \left| \sum_{\vec{P} \in \vec{\mathbb{P}}} \frac{C_{Q_{\vec{P}}}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi^1_{\vec{P}_1} \rangle \langle f_2, \Phi^2_{\vec{P}_2} \rangle \langle f_3, \Phi^3_{\vec{P}_3} \rangle \right| \lesssim_{p, p_1, p_2, B} |E_1|^{1/p_1} |E_2|^{1/p_2}, \tag{4-2}$$

where p_1 is larger than but close to 1, while p_2 is smaller than but close to 2.

In order to prove our Theorem 1.3 in biparameter settings, one can easily observe that the main difficulty from [Muscalu et al. 2004a; 2006] is that, if we restrict the sum of tritiles $P'' \in \mathbb{P}''$ in the definition of discrete model operators $\Pi_{\vec{\mathbb{P}}}$ to a tree, then we essentially get a tensor product of two discrete paraproducts on x_1 and x_2 respectively, which can be estimated by the MM, MS, SM and SS functions, but, due to the extra degree of freedom in frequency in the x_2 direction, there are infinitely many such tensor products of paraproducts in the summation, so it's difficult for us to carry out the stopping-time decompositions by using the hybrid square and maximal operators as in [Muscalu et al. 2004a; 2006]. Instead, we will make use of the L^2 size and L^2 energy estimates of the trilinear forms, the almost orthogonality of wave packets associated with different tiles and the decay assumptions on the symbols. Furthermore, we can extend our proof of Theorem 1.3 to general d-parameter settings ($d \geq 3$) by applying the generic decomposition lemma (Lemma 4.1) to the $d-1$ variables $x_1, \ldots, x_{d-1}$. Although one can't obtain that supp $\Phi^{3,\ell}_{\vec{P}_3'} \otimes \Phi^3_{P_3''}$ is entirely contained in the exceptional set U as in [Muscalu et al. 2006], one can show that this support set is contained in U in all the variables $x_1, \ldots, x_{d-1}$, but not x_d. Therefore, we only need to consider the distance from this support set to the set E_3' in the x_d direction and obtain enough decay factors for summation; the extension of the proof from biparameter case to the general d-parameter ($d \geq 3$) cases is straightforward.

From [Muscalu et al. 2006], we can find the following generic decomposition lemma:

Lemma 4.1. *Let $J \subseteq \mathbb{R}$ be a fixed interval. Then every smooth bump function ϕ_J adapted to J can be naturally decomposed as*

$$\phi_J = \sum_{\ell \in \mathbb{N}} 2^{-100\ell} \phi_J^\ell,$$

where, for every $\ell \in \mathbb{N}$, ϕ_J^ℓ is also a bump function adapted to J but having the additional property that $\operatorname{supp}(\phi_J^\ell) \subseteq 2^\ell J$. *If in addition we assume that $\int_{\mathbb{R}} \phi_J(x)\, dx = 0$, then the functions ϕ_J^ℓ can be chosen so that $\int_{\mathbb{R}} \phi_J^\ell(x)\, dx = 0$ for every $\ell \in \mathbb{N}$.*

We use $2^\ell J$ to denote the interval having the same center as J but with length 2^ℓ times that of J.

By using Lemma 4.1, we can estimate the left-hand side of (4-2) by

$$|\Lambda_{\vec{\mathbb{P}}}(f_1, f_2, f_3)| \lesssim \sum_{\ell\in\mathbb{N}} 2^{-100\ell}\Lambda^{\ell}_{\vec{\mathbb{P}}}(f_1, f_2, f_3). \tag{4-3}$$

The trilinear forms $\Lambda^{\ell}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)$ ($\ell \in \mathbb{N}$) are defined by

$$\Lambda^{\ell}_{\vec{\mathbb{P}}}(f_1, f_2, f_3) := \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{|C_{Q_{\vec{P}}}|}{|I_{\vec{P}}|^{\frac{1}{2}}} |\langle f_1, \Phi^1_{\vec{P}_1}\rangle||\langle f_2, \Phi^2_{\vec{P}_2}\rangle||\langle f_3, \Phi^{3,\ell}_{\vec{P}_3}\rangle|, \tag{4-4}$$

where the new biparameter wave packets are $\Phi^{3,\ell}_{\vec{P}_3} := \Phi^{3,\ell}_{\tilde{P}'_3} \otimes \Phi^3_{P''_3}$ with the additional property that $\operatorname{supp}(\Phi^{3,\ell}_{\tilde{P}'_3}) \subseteq 2^{\ell} I_{\tilde{P}'_3} = 2^{\ell} I_{\tilde{P}'}$.

For every $\ell \in \mathbb{N}$, we define the sets

$$\Omega_{-10\ell} := \bigcup_{j=1}^{2} \left\{ x \in \mathbb{R}^2 : \mathrm{MM}\left(\frac{\chi_{E_j}}{|E_j|}\right)(x) > C2^{10\ell} \right\} \tag{4-5}$$

and

$$\tilde{\Omega}_{-10\ell} := \{x \in \mathbb{R}^2 : \mathrm{MM}(\chi_{\Omega_{-10\ell}})(x) > 2^{-\ell}\}, \tag{4-6}$$

where the double maximal operator MM is given by

$$\mathrm{MM}(h)(x, y) := \sup_{\substack{\text{dyadic rectangle } R\\ (x,y)\in R}} \frac{1}{|R|} \int_R |h(u, v)|\, du\, dv. \tag{4-7}$$

Finally, we define the exceptional set

$$U := \bigcup_{\ell\in\mathbb{N}} \tilde{\Omega}_{-10\ell}. \tag{4-8}$$

It is clear that $|U| < \frac{1}{10}$ if C is a large enough constant, which we fix from now on. Then, we define $E'_3 := E_3 \setminus U$ and note that $|E'_3| \simeq 1$.

Now fix $\ell \in \mathbb{N}$, and split the trilinear form $\Lambda^{\ell}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)$ defined in (4-4) into two parts as follows:

$$\begin{aligned}
&\Lambda^{\ell}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)\\
&= \sum_{\substack{\vec{P}\in\vec{\mathbb{P}}:\\ I_{\vec{P}}\cap\Omega^c_{-10\ell}\neq\varnothing}} \frac{|C_{Q_{\vec{P}}}|}{|I_{\vec{P}}|^{\frac{1}{2}}} |\langle f_1, \Phi^1_{\vec{P}_1}\rangle||\langle f_2, \Phi^2_{\vec{P}_2}\rangle||\langle f_3, \Phi^{3,\ell}_{\vec{P}_3}\rangle| + \sum_{\substack{\vec{P}\in\vec{\mathbb{P}}:\\ I_{\vec{P}}\cap\Omega^c_{-10\ell}=\varnothing}} \frac{|C_{Q_{\vec{P}}}|}{|I_{\vec{P}}|^{\frac{1}{2}}} |\langle f_1, \Phi^1_{\vec{P}_1}\rangle||\langle f_2, \Phi^2_{\vec{P}_2}\rangle||\langle f_3, \Phi^{3,\ell}_{\vec{P}_3}\rangle|\\
&=: \Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) + \Lambda^{\ell}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3),
\end{aligned} \tag{4-9}$$

where A^c denotes the complement of a set A.

4A. ***Estimates for trilinear form $\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3)$.*** We can decompose the collection $\tilde{\mathbb{P}}'$ of tritiles into

$$\tilde{\mathbb{P}}' = \bigcup_{k'\in\mathbb{Z}} \tilde{\mathbb{P}}'_{k'}, \tag{4-10}$$

where

$$\tilde{\mathbb{P}}'_{k'} := \{\tilde{P}' \in \tilde{\mathbb{P}}' : |I_{\tilde{P}'}| = 2^{-k'}\}. \tag{4-11}$$

As a consequence, we can split the trilinear form $\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3)$ into

$$\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) = \sum_{k'\in\mathbb{Z}} \sum_{\substack{\vec{P}\in\tilde{\mathbb{P}}'_{k'}\times\mathbb{P}'':\\ I_{\vec{P}}\cap\Omega^c_{-10\ell}\neq\varnothing}} |C_{Q_{\vec{P}}}| \frac{|I_{\tilde{P}'}|}{|I_{P''}|^{\frac{1}{2}}} \prod_{j=1}^{2} \left|\left\langle \frac{\langle f_j, \Phi^j_{\tilde{P}'_j}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\right\rangle\right| \times \left|\left\langle \frac{\langle f_3, \Phi^{3,l}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\right\rangle\right|. \tag{4-12}$$

By Lemma 2.4, we can estimate the Fourier coefficients $C_{Q_{\vec{P}}} := C_{Q_{\vec{P}},\vec{0},\vec{0},\vec{0}}$ for each $\vec{P} \in \tilde{\mathbb{P}}'_{k'} \times \mathbb{P}''$ ($k' \in \mathbb{Z}$) by

$$|C_{Q_{\vec{P}}}| \lesssim C_{k'} \quad \text{with} \quad \sum_{k'\in\mathbb{Z}} C_{k'} \le B < +\infty. \tag{4-13}$$

For each fixed $\tilde{P}' \in \tilde{\mathbb{P}}'$, we define the subcollection

$$\mathbb{P}''_{\tilde{P}'} := \{P'' \in \mathbb{P}'' : I_{\vec{P}} \cap \Omega^c_{-10\ell} \neq \varnothing\}.$$

Therefore, by using Proposition 3.5, we derive the estimates

$$\begin{aligned}
&\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3)\\
&\lesssim \sum_{k'\in\mathbb{Z}} C_{k'} \sum_{\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}} |I_{\tilde{P}'}|\\
&\qquad \times \left[\prod_{j=1}^{2}\left(\text{energy}_j\left(\left(\left\langle \frac{\langle f_j, \Phi^j_{\tilde{P}'_j}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right)\right)^{1-\theta_j} \left(\text{size}_j\left(\left(\left\langle \frac{\langle f_j, \Phi^j_{\tilde{P}'_j}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right)\right)^{\theta_j}\right]\\
&\qquad \times \left(\text{size}_3\left(\left(\left\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right)\right)^{\theta_3} \left(\text{energy}_3\left(\left(\left\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right)\right)^{1-\theta_3}
\end{aligned} \tag{4-14}$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$.

To estimate the right-hand side of (4-14), note that $I_{\vec{P}} \cap \Omega^c_{-10\ell} \neq \varnothing$ and $\operatorname{supp} f_3 \subseteq E'_3 \subseteq \mathbb{R}^2 \setminus U$; we apply the size estimates in Lemma 3.6 and get, for each $\tilde{P}' \in \tilde{\mathbb{P}}'_{k'}$,

$$\text{size}_1\left(\left(\left\langle \frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^1_{P''_1}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{\tilde{P}'}} \frac{1}{|I_{P''}|} \int_{\mathbb{R}} \left|\frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\right| \tilde{\chi}^M_{I_{P''}}\, dx \lesssim 2^{10\ell}|E_1|, \tag{4-15}$$

$$\text{size}_2\left(\left(\left\langle \frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^2_{P''_2}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{\tilde{P}'}} \frac{1}{|I_{P''}|} \int_{\mathbb{R}} \left|\frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\right| \tilde{\chi}^M_{I_{P''}}\, dx \lesssim 2^{10\ell}|E_2|, \tag{4-16}$$

$$\text{size}_3\left(\left(\left\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}'}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{\tilde{P}'}} \frac{1}{|I_{P''}|} \int_{\mathbb{R}} \left|\frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\right| \tilde{\chi}^M_{I_{P''}}\, dx \lesssim 1, \tag{4-17}$$

where $M > 0$ is sufficiently large. By applying the energy estimates in Lemma 3.7 and Hölder estimates, we have, for each $\tilde{P}' \in \tilde{\mathbb{P}}'_{k'}$,

$$\text{energy}_1\Big(\Big(\Big\langle \frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^1_{P''_1}\Big\rangle\Big)_{P''\in\mathbb{P}''_{\tilde{P}'}}\Big) \lesssim \Big\| \frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\Big\|_{L^2(\mathbb{R})} \lesssim \Big(\int_{E_1} \frac{\tilde{\chi}^{100}_{I_{\tilde{P}'}}(x_1)}{|I_{\tilde{P}'}|}\,dx_1\,dx_2\Big)^{\frac{1}{2}}, \tag{4-18}$$

$$\text{energy}_2\Big(\Big(\Big\langle \frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^2_{P''_2}\Big\rangle\Big)_{P''\in\mathbb{P}''_{\tilde{P}'}}\Big) \lesssim \Big\| \frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\Big\|_{L^2(\mathbb{R})} \lesssim \Big(\int_{E_2} \frac{\tilde{\chi}^{100}_{I_{\tilde{P}'}}(x_1)}{|I_{\tilde{P}'}|}\,dx_1\,dx_2\Big)^{\frac{1}{2}}, \tag{4-19}$$

$$\text{energy}_3\Big(\Big(\Big\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\Big\rangle\Big)_{P''\in\mathbb{P}''_{\tilde{P}'}}\Big) \lesssim \Big\| \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\Big\|_{L^2(\mathbb{R})} \lesssim \Big(\int_{E'_3} \frac{\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}(x_1)}{|I_{\tilde{P}'}|}\,dx_1\,dx_2\Big)^{\frac{1}{2}}, \tag{4-20}$$

where the approximate cutoff function $\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}(x_1)$ decays rapidly (of order 100) away from the interval $I_{\tilde{P}'}$ at scale $|I_{\tilde{P}'}|$ and satisfies the additional property that $\operatorname{supp}\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}} \subseteq 2^\ell I_{\tilde{P}'}$.

Now we insert the size and energy estimates (4-15)–(4-20) into (4-14) and get

$$\begin{aligned}&\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3)\\ &\lesssim 2^{10\ell}|E_1|^{\theta_1}|E_2|^{\theta_2}\sum_{k'\in\mathbb{Z}} C_{k'} \sum_{\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}} \Big(\int_{E_1}\tilde{\chi}^{100}_{I_{\tilde{P}'}}\,dx\Big)^{\frac{1-\theta_1}{2}}\Big(\int_{E_2}\tilde{\chi}^{100}_{I_{\tilde{P}'}}\,dx\Big)^{\frac{1-\theta_2}{2}}\Big(\int_{E'_3}\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}\,dx\Big)^{\frac{1-\theta_3}{2}}.\end{aligned} \tag{4-21}$$

Since $|I_{\tilde{P}'}| = 2^{-k'}$ for every $\tilde{P}' \in \tilde{\mathbb{P}}'_{k'}$, all the dyadic intervals $I_{\tilde{P}'}$ are disjoint, thus, by using Hölder's inequality, we can estimate the inner sum in the right-hand side of (4-21) by

$$\prod_{j=1}^{2}\Big(\sum_{\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}}\int_{E_j}\tilde{\chi}^{100}_{I_{\tilde{P}'}}\,dx\Big)^{\frac{1-\theta_j}{2}}\Big(\sum_{\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}}\int_{E'_3}\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}\,dx\Big)^{\frac{1-\theta_3}{2}} \lesssim |E_1|^{(1-\theta_1)/2}|E_2|^{(1-\theta_2)/2}. \tag{4-22}$$

Combining the estimates (4-13), (4-21) and (4-22), we arrive at

$$\begin{aligned}\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) &\lesssim 2^{10\ell}|E_1|^{\theta_1}|E_2|^{\theta_2}|E_1|^{(1-\theta_1)/2}|E_2|^{(1-\theta_2)/2}\sum_{k'\in\mathbb{Z}} C_{k'}\\ &\lesssim_{\theta_1,\theta_2,\theta_3,B} 2^{10\ell}|E_1|^{(1+\theta_1)/2}|E_2|^{(1+\theta_2)/2}\end{aligned} \tag{4-23}$$

for every $\ell \in \mathbb{N}$ and $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$.

By taking θ_1 sufficiently close to 1 and θ_2 sufficiently close to 0, one can make the exponent $2/(1+\theta_1) = p_1$ strictly larger than 1 and close to 1, and $2/(1+\theta_2) = p_2$ strictly smaller than 2 and close to 2. We finally get the estimate

$$\Lambda^{\ell}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) \lesssim_{p,p_1,p_2,B} 2^{10\ell}|E_1|^{1/p_1}|E_2|^{1/p_2} \tag{4-24}$$

for every $\ell \in \mathbb{N}$ and p, p_1, p_2 satisfying the hypothesis of Proposition 2.17.

4B. ***Estimates for the trilinear form*** $\boldsymbol{\Lambda^{\ell}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3)}$**.** If $I_{\vec{P}} \subseteq \Omega_{-10\ell}$, then $2^{\ell} I_{\tilde{P}'} \times I_{P''} \subseteq \tilde{\Omega}_{-10\ell}$. Therefore, for each fixed $\tilde{P}' \in \tilde{\mathbb{P}}'$, we define the corresponding subcollection of $\mathbb{P}''$ by

$$\mathbb{P}''_{\tilde{P}'} := \{P'' \in \mathbb{P}'' : I_{\vec{P}} \subseteq \Omega_{-10\ell}\},$$

then we can decompose the collection $\mathbb{P}''_{\tilde{P}'}$ further, as follows:

$$\mathbb{P}''_{\tilde{P}'} = \bigcup_{d'' \in \mathbb{N}} \mathbb{P}''_{\tilde{P}',d''}, \tag{4-25}$$

where

$$\mathbb{P}''_{\tilde{P}',d''} := \{P'' \in \mathbb{P}''_{\tilde{P}'} : 2^{\ell} I_{\tilde{P}'} \times 2^{d''} I_{P''} \subseteq \tilde{\Omega}_{-10\ell}\} \tag{4-26}$$

and d'' is maximal with this property.

Now we apply both the decompositions of $\tilde{\mathbb{P}}'$ and $\mathbb{P}''_{\tilde{P}'}$ defined in (4-10) and (4-25) at the same time, and split the trilinear form $\Lambda^{\ell}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3)$ into

$$\Lambda^{\ell}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3) = \sum_{k' \in \mathbb{Z}} \sum_{\tilde{P}' \in \tilde{\mathbb{P}}'_{k'}} |C_{Q_{\vec{P}}}| |I_{\tilde{P}'}| \sum_{d'' \in \mathbb{N}} \sum_{P'' \in \mathbb{P}''_{\tilde{P}',d''}} \frac{1}{|I_{P''}|^{\frac{1}{2}}} \prod_{j=1}^{2} \left| \left\langle \frac{\langle f_j, \Phi^j_{\tilde{P}'_j} \rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^j_{P''_j} \right\rangle \right| \times \left| \left\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3} \rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3} \right\rangle \right|. \tag{4-27}$$

In the inner sum of (4-27), since $2^{\ell} I_{\tilde{P}'} \times 2^{d''} I_{P''} \subseteq \tilde{\Omega}_{-10\ell}$,

$$\operatorname{supp}(\Phi^{3,\ell}_{\tilde{P}'_3}) \subseteq 2^{\ell} I_{\tilde{P}'} \quad \text{and} \quad \operatorname{supp} f_3 \subseteq E'_3 \subseteq \mathbb{R}^2 \setminus U,$$

we can assume hereafter in this subsection that

$$|f_3| \leq \chi_{E'_3} \chi_{2^{\ell} I_{\tilde{P}'}} \chi_{(2^{d''} I_{P''})^c}. \tag{4-28}$$

By using Proposition 3.5 and (4-13), we derive from (4-27) the estimates

$$\begin{aligned}
&\Lambda^{\ell}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3) \\
&\lesssim \sum_{k' \in \mathbb{Z}} C_{k'} \sum_{\tilde{P}' \in \tilde{\mathbb{P}}'_{k'}} |I_{\tilde{P}'}| \sum_{d'' \in \mathbb{N}} \Bigg[\prod_{j=1}^{2} \Bigg(\operatorname{energy}_j \Bigg(\Bigg(\Bigg\langle \frac{\langle f_j, \Phi^j_{\tilde{P}'_j} \rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^j_{P''_j} \Bigg\rangle \Bigg)_{P'' \in \mathbb{P}''_{\tilde{P}',d''}} \Bigg) \Bigg)^{1-\theta_j} \\
&\qquad \times \Bigg(\operatorname{size}_j \Bigg(\Bigg(\Bigg\langle \frac{\langle f_j, \Phi^j_{\tilde{P}'_j} \rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^j_{P''_j} \Bigg\rangle \Bigg)_{P'' \in \mathbb{P}''_{\tilde{P}',d''}} \Bigg) \Bigg)^{\theta_j} \Bigg] \\
&\times \Bigg(\operatorname{size}_3 \Bigg(\Bigg(\Bigg\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3} \rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3} \Bigg\rangle \Bigg)_{P'' \in \mathbb{P}''_{\tilde{P}',d''}} \Bigg) \Bigg)^{\theta_3} \Bigg(\operatorname{energy}_3 \Bigg(\Bigg(\Bigg\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3} \rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3} \Bigg\rangle \Bigg)_{P'' \in \mathbb{P}''_{\tilde{P}',d''}} \Bigg) \Bigg)^{1-\theta_3}
\end{aligned} \tag{4-29}$$

for any $0 \leq \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$.

To estimate the inner sum in the right-hand side of (4-29), note that $I_{\vec{P}} \subseteq \Omega_{-10\ell}$, $P'' \in \mathbb{P}''_{\tilde{P}',d''}$ and f_3 satisfies (4-28), so we apply the size estimates in Lemma 3.6 and get, for each $\tilde{P}' \in \tilde{\mathbb{P}}'_{k'}$ and $d'' \in \mathbb{N}$,

$$\text{size}_1\left(\left(\left\langle \frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^1_{P''_1}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}',d''}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{\tilde{P}',d''}} \frac{1}{|I_{P''}|}\int_{\mathbb{R}}\left|\frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\right|\tilde{\chi}^M_{I_{P''}}\,dx \lesssim 2^{11\ell+d''}|E_1|, \quad (4\text{-}30)$$

$$\text{size}_2\left(\left(\left\langle \frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^2_{P''_2}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}',d''}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{\tilde{P}',d''}} \frac{1}{|I_{P''}|}\int_{\mathbb{R}}\left|\frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\right|\tilde{\chi}^M_{I_{P''}}\,dx \lesssim 2^{11\ell+d''}|E_2|, \quad (4\text{-}31)$$

$$\text{size}_3\left(\left(\left\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}',d''}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{\tilde{P}',d''}} \frac{1}{|I_{P''}|}\int_{\mathbb{R}}\left|\frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}\right|\tilde{\chi}^M_{I_{P''}}\,dx \lesssim 2^{-(M-100)d''}, \quad (4\text{-}32)$$

where $M>0$ is arbitrarily large. Similar to the energy estimates obtained in (4-18), (4-19) and (4-20), by applying the energy estimates in Lemma 3.7 and Hölder estimates we have, for each $\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}$ and $d''\in\mathbb{N}$,

$$\text{energy}_1\left(\left(\left\langle \frac{\langle f_1, \Phi^1_{\tilde{P}'_1}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^1_{P''_1}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}',d''}}\right) \lesssim \left(\int_{E_1}\frac{\tilde{\chi}^{100}_{I_{\tilde{P}'}}(x_1)}{|I_{\tilde{P}'}|}\,dx_1\,dx_2\right)^{\frac{1}{2}}, \quad (4\text{-}33)$$

$$\text{energy}_2\left(\left(\left\langle \frac{\langle f_2, \Phi^2_{\tilde{P}'_2}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^2_{P''_2}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}',d''}}\right) \lesssim \left(\int_{E_2}\frac{\tilde{\chi}^{100}_{I_{\tilde{P}'}}(x_1)}{|I_{\tilde{P}'}|}\,dx_1\,dx_2\right)^{\frac{1}{2}}, \quad (4\text{-}34)$$

$$\text{energy}_3\left(\left(\left\langle \frac{\langle f_3, \Phi^{3,\ell}_{\tilde{P}'_3}\rangle}{|I_{\tilde{P}'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\right\rangle\right)_{P''\in\mathbb{P}''_{\tilde{P}',d''}}\right) \lesssim \left(\int_{E'_3}\frac{\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}(x_1)}{|I_{\tilde{P}'}|}\,dx_1\,dx_2\right)^{\frac{1}{2}}, \quad (4\text{-}35)$$

where the approximate cutoff function $\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}(x_1)$ decays rapidly (of order 100) away from the interval $I_{\tilde{P}'}$ at scale $|I_{\tilde{P}'}|$ and satisfies the additional property that supp $\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}\subseteq 2^\ell I_{\tilde{P}'}$.

Now we insert the size and energy estimates (4-30)–(4-35) into (4-29); by using the estimates (4-13), (4-22) and Hölder's inequality, we then get

$$\begin{aligned}
&\Lambda^\ell_{\tilde{\mathbb{P}},II}(f_1, f_2, f_3)\\
&\lesssim 2^{11\ell}|E_1|^{\theta_1}|E_2|^{\theta_2}\sum_{k'\in\mathbb{Z}}C_{k'}\sum_{d''\in\mathbb{N}}2^{-(M\theta_3-100)d''}\prod_{j=1}^{2}\left(\sum_{\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}}\int_{E_j}\tilde{\chi}^{100}_{I_{\tilde{P}'}}\,dx\right)^{\frac{1-\theta_j}{2}}\times\left(\sum_{\tilde{P}'\in\tilde{\mathbb{P}}'_{k'}}\int_{E'_3}\tilde{\chi}^{100,\ell}_{I_{\tilde{P}'}}\,dx\right)^{\frac{1-\theta_3}{2}}\\
&\lesssim_{\theta_1,\theta_2,\theta_3,B,M} 2^{11\ell}|E_1|^{(1+\theta_1)/2}|E_2|^{(1+\theta_2)/2}\sum_{d''\in\mathbb{N}}2^{-(M\theta_3-100)d''}. \qquad (4\text{-}36)
\end{aligned}$$

for every $\ell\in\mathbb{N}$ and $0\le\theta_1,\theta_2,\theta_3<1$ with $\theta_1+\theta_2+\theta_3=1$.

By taking θ_1 sufficiently close to 1 and θ_2 sufficiently close to 0, one can make the exponent $2/(1+\theta_1)=p_1$ strictly larger than 1 and close to 1, and $2/(1+\theta_2)=p_2$ strictly smaller than 2 and close to 2. The series over $d''\in\mathbb{N}$ in (4-36) is summable if we choose M large enough (say, $M\simeq 200\theta_3^{-1}$). We finally get the estimate

$$\Lambda^\ell_{\tilde{\mathbb{P}},II}(f_1, f_2, f_3)\lesssim_{p,p_1,p_2,B} 2^{11\ell}|E_1|^{1/p_1}|E_2|^{1/p_2} \quad (4\text{-}37)$$

for every $\ell\in\mathbb{N}$ and p, p_1, p_2 satisfying the hypothesis of Proposition 2.17.

4C. ***Conclusions.*** By inserting the estimates (4-9), (4-24) and (4-37) into (4-3), we finally get

$$|\Lambda_{\vec{\mathbb{P}}}(f_1, f_2, f_3)| \lesssim_{p,p_1,p_2,B} \sum_{\ell\in\mathbb{N}} 2^{-100\ell}2^{12\ell}|E_1|^{1/p_1}|E_2|^{1/p_2} \lesssim_{p,p_1,p_2,B} |E_1|^{1/p_1}|E_2|^{1/p_2}, \tag{4-38}$$

which completes the proof of Proposition 2.17 for the model operators $\Pi_{\vec{\mathbb{P}}}$.

This concludes the proof of Theorem 1.3.

5. Proof of Theorem 1.5

In this section, we prove Theorem 1.5 by carrying out the proof of Proposition 2.17 for the model operators $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$ defined in (2-54) with $\vec{\mathbb{P}} = \mathbb{P}' \times \mathbb{P}''$.

Fix indices p_1, p_2, p as in the hypothesis of Proposition 2.17. Fix arbitrary measurable sets E_1, E_2, E_3 of finite measure (by using the scaling invariance of $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$, we can assume further that $|E_3| = 1$). Our goal is to find $E_3' \subseteq E_3$ with $|E_3'| \simeq |E_3| = 1$ such that, for any functions $|f_1| \le \chi_{E_1}$, $|f_2| \le \chi_{E_2}$ and $|f_3| \le \chi_{E_3'}$, one has the corresponding trilinear forms $\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)$ defined by

$$\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2, f_3) := \int_{\mathbb{R}^2} \tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2)(x) f_3(x)\,dx \tag{5-1}$$

satisfy estimates

$$|\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)| = \left| \sum_{\vec{P}\in\vec{\mathbb{P}}} \frac{\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}}{|I_{\vec{P}}|^{\frac{1}{2}}} \langle f_1, \Phi^1_{\vec{P}_1}\rangle \langle f_2, \Phi^2_{\vec{P}_2}\rangle \langle f_3, \Phi^3_{\vec{P}_3}\rangle \right| \lesssim_{\varepsilon,p,p_1,p_2} |E_1|^{1/p_1}|E_2|^{1/p_2}, \tag{5-2}$$

where p_1 is larger than but close to 1, while p_2 is smaller than but close to 2.

In the proof of Theorem 1.5 in biparameter settings, besides the difficulty that one can't carry out the stopping-time decompositions by using the hybrid square and maximal operators as in [Muscalu et al. 2004a; 2006], we can't apply Journé's lemma as in [Muscalu et al. 2004a] either, since we can't get the estimate $\sum_{P'} |I_{P'}| \lesssim |\tilde{I}|$ for all dyadic intervals $I_{P'} \subseteq \tilde{I}$ with comparable lengths. Therefore, in order to prove Theorem 1.5, we will take advantage of the almost orthogonality of wave packets associated with different tiles of distinct trees and the decay assumptions on the symbols to overcome these difficulties.

We define the exceptional set

$$\Omega := \bigcup_{j=1}^{2} \left\{ x \in \mathbb{R}^2 : \mathrm{MM}\left(\frac{\chi_{E_j}}{|E_j|}\right)(x) > C \right\} \tag{5-3}$$

It is clear that $|\Omega| < \frac{1}{10}$ if C is a large enough constant, which we fix from now on. Then, we define $E_3' := E_3 \setminus \Omega$ and observe that $|E_3'| \simeq 1$.

Now we estimate the trilinear form $\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)$ defined in (5-1) by two terms as follows:

$$
\begin{aligned}
&|\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)| \\
&\lesssim \sum_{\substack{\vec{P}\in\vec{\mathbb{P}}:\\ I_{\vec{P}}\cap\Omega^c\neq\varnothing}} \frac{|\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}|}{|I_{\vec{P}}|^{\frac{1}{2}}}|\langle f_1, \Phi^1_{\vec{P}_1}\rangle||\langle f_2, \Phi^2_{\vec{P}_2}\rangle||\langle f_3, \Phi^3_{\vec{P}_3}\rangle| + \sum_{\substack{\vec{P}\in\vec{\mathbb{P}}:\\ I_{\vec{P}}\cap\Omega^c=\varnothing}} \frac{|\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}|}{|I_{\vec{P}}|^{\frac{1}{2}}}|\langle f_1, \Phi^1_{\vec{P}_1}\rangle||\langle f_2, \Phi^2_{\vec{P}_2}\rangle||\langle f_3, \Phi^3_{\vec{P}_3}\rangle| \\
&=: \tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) + \tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3).
\end{aligned}
\tag{5-4}
$$

5A. ***Estimates for trilinear form*** $\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3)$***.*** We can decompose the collection $\tilde{\mathbb{P}}'$ of tritiles into

$$\mathbb{P}' = \bigcup_{k'\in\mathbb{Z}} \mathbb{P}'_{k'}, \tag{5-5}$$

where

$$\mathbb{P}'_{k'} := \{P' \in \mathbb{P}' : \ell(Q_{P'}) = 2^{k'}\}. \tag{5-6}$$

As a consequence, we can split the trilinear form $\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3)$ into

$$\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) = \sum_{k'\in\mathbb{Z}} \sum_{\substack{\vec{P}\in\mathbb{P}'_{k'}\times\mathbb{P}'':\\ I_{\vec{P}}\cap\Omega^c\neq\varnothing}} |\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}| \frac{|I_{P'}|}{|I_{P''}|^{\frac{1}{2}}} \prod_{j=1}^{3} \left|\left\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\right\rangle\right|. \tag{5-7}$$

By Lemma 2.10, we can estimate the Fourier coefficients $\tilde{C}^{\varepsilon}_{Q_{\vec{P}}} := \tilde{C}^{\varepsilon}_{Q_{\vec{P}},\vec{0},\vec{0},\vec{0}}$ for each $\vec{P} \in \mathbb{P}'_{k'} \times \mathbb{P}''$ ($k' \in \mathbb{Z}$) by

$$|\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}| \lesssim \tilde{C}^{\varepsilon}_{k'} := \langle k'\rangle^{-(1+\varepsilon)} = (1 + |k'|^2)^{-(1+\varepsilon)/2}. \tag{5-8}$$

For each fixed $P' \in \mathbb{P}'$, we define the subcollection $\mathbb{P}''_{P'}$ of $\mathbb{P}''$ by

$$\mathbb{P}''_{P'} := \{P'' \in \mathbb{P}'' : I_{\vec{P}} \cap \Omega^c \neq \varnothing\}.$$

Therefore, by using Proposition 3.5, we derive the estimates

$$
\begin{aligned}
&\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1, f_2, f_3) \\
&\lesssim \sum_{k'\in\mathbb{Z}} \tilde{C}^{\varepsilon}_{k'} \sum_{P'\in\mathbb{P}'_{k'}} |I_{P'}| \prod_{j=1}^{3} \Bigg[\Bigg(\text{energy}_j \Bigg(\Bigg(\Bigg\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j} \Bigg\rangle \Bigg)_{P''\in\mathbb{P}''_{P'}} \Bigg) \Bigg)^{1-\theta_j} \\
&\qquad\qquad \times \Bigg(\text{size}_j \Bigg(\Bigg(\Bigg\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j} \Bigg\rangle \Bigg)_{P''\in\mathbb{P}''_{P'}} \Bigg) \Bigg)^{\theta_j} \Bigg]
\end{aligned}
\tag{5-9}
$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$.

To estimate the right-hand side of (5-9), note that $I_{\vec{P}} \cap \Omega^c \neq \varnothing$ and $\operatorname{supp} f_3 \subseteq E_3'$, so we apply the size estimates in Lemma 3.6 and get, for each $P' \in \mathbb{P}'_{k'}$ and $j = 1, 2, 3$,

$$\operatorname{size}_j\left(\left(\left\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\right\rangle\right)_{P''\in\mathbb{P}''_{P'}}\right) \lesssim \sup_{P''\in\mathbb{P}''_{P'}} \frac{1}{|I_{P''}|}\int_{\mathbb{R}}\left|\frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}\right| \tilde{\chi}^M_{I_{P''}}\,dx \lesssim |E_j|, \tag{5-10}$$

where $M > 0$ is sufficiently large. By applying the energy estimates in Lemma 3.7, we have, for each $P' \in \mathbb{P}'_{k'}$ and $j = 1, 2, 3$,

$$\operatorname{energy}_j\left(\left(\left\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\right\rangle\right)_{P''\in\mathbb{P}''_{P'}}\right) \lesssim \frac{1}{|I_{P'}|^{\frac{1}{2}}}\left(\int_{\mathbb{R}} |\langle f_j, \Phi^j_{P'_j}\rangle|^2\,dx_2\right)^{\frac{1}{2}}. \tag{5-11}$$

Now we insert the size and energy estimates (5-10) and (5-11) into (5-9) and get

$$\tilde{\Lambda}^{\varepsilon}_{\bar{\mathbb{P}},I}(f_1, f_2, f_3) \lesssim |E_1|^{\theta_1}|E_2|^{\theta_2}\sum_{k'\in\mathbb{Z}}\tilde{C}^{\varepsilon}_{k'}\sum_{P'\in\mathbb{P}'_{k'}}\prod_{j=1}^{3}\left(\int_{\mathbb{R}}|\langle f_j, \Phi^j_{P'_j}\rangle|^2\,dx_2\right)^{\frac{1-\theta_j}{2}}. \tag{5-12}$$

Observe that, for any different tritiles $\bar{P}' \in \mathbb{P}'_{k'}$ and $\bar{\bar{P}}' \in \mathbb{P}'_{k'}$, one has $I_{\bar{P}'} \cap I_{\bar{\bar{P}}'} = \varnothing$, or otherwise one has $I_{\bar{P}'} = I_{\bar{\bar{P}}'}$ but $\omega_{\bar{P}'_j} \cap \omega_{\bar{\bar{P}}'_j} = \varnothing$ for every $j = 1, 2, 3$. By taking advantage of such orthogonality in L^2 of the wave packets $\Phi^j_{P'_j}$ corresponding to the tiles P'_j ($j = 1, 2, 3$), one has that, for any function $F \in L^2(\mathbb{R})$ and $k' \in \mathbb{Z}$,

$$\begin{aligned}
\Big\|\sum_{P'\in\mathbb{P}'_{k'}}\langle F, \Phi^j_{P'_j}\rangle\Phi^j_{P'_j}\Big\|^2_{L^2} &\le \sum_{\substack{\bar{P}',\,\bar{\bar{P}}'\in\mathbb{P}'_{k'}:\\ \omega_{\bar{P}'_j}=\omega_{\bar{\bar{P}}'_j}\\ I_{\bar{P}'}\cap I_{\bar{\bar{P}}'}=\varnothing}} |\langle F, \Phi^j_{\bar{P}'_j}\rangle||\langle F, \Phi^j_{\bar{\bar{P}}'_j}\rangle||\langle \Phi^j_{\bar{P}'_j}, \Phi^j_{\bar{\bar{P}}'_j}\rangle| \\
&\lesssim 2^{k'}\sum_{\bar{P}'\in\mathbb{P}'_{k'}}|\langle F, \Phi^j_{\bar{P}'_j}\rangle|^2\sum_{\substack{\bar{\bar{P}}'\in\mathbb{P}'_{k'}:\\ \omega_{\bar{P}'_j}=\omega_{\bar{\bar{P}}'_j}\\ I_{\bar{P}'}\cap I_{\bar{\bar{P}}'}=\varnothing}}|\langle\tilde{\chi}^{1000}_{I_{\bar{P}'}}, \tilde{\chi}^{1000}_{I_{\bar{\bar{P}}'}}\rangle| \\
&\lesssim \sum_{\bar{P}'\in\mathbb{P}'_{k'}}|\langle F, \Phi^j_{\bar{P}'_j}\rangle|^2\sum_{\substack{\bar{\bar{P}}'\in\mathbb{P}'_{k'}:\\ \omega_{\bar{P}'_j}=\omega_{\bar{\bar{P}}'_j}\\ I_{\bar{P}'}\cap I_{\bar{\bar{P}}'}=\varnothing}}\left(1+\frac{\operatorname{dist}(I_{\bar{P}'}, I_{\bar{\bar{P}}'})}{|I_{\bar{P}'}|}\right)^{-100} \\
&\lesssim \sum_{P'\in\mathbb{P}'_{k'}}|\langle F, \Phi^j_{P'_j}\rangle|^2,
\end{aligned} \tag{5-13}$$

from which we deduce the Bessel-type inequality

$$\sum_{P'\in\mathbb{P}'_{k'}}|\langle F, \Phi^j_{P'_j}\rangle|^2 = \left|\left\langle\sum_{P'\in\mathbb{P}'_{k'}}\langle F, \Phi^j_{P'_j}\rangle\Phi^j_{P'_j}, F\right\rangle\right| \le \Big\|\sum_{P'\in\mathbb{P}'_{k'}}\langle F, \Phi^j_{P'_j}\rangle\Phi^j_{P'_j}\Big\|_{L^2}\cdot\|F\|_{L^2} \lesssim \|F\|^2_{L^2}, \tag{5-14}$$

where the implicit constants in the bounds are independent of $k' \in \mathbb{Z}$. Then, we can use the Bessel-type inequality (5-14) and Hölder's inequality to estimate the inner sum in the right-hand side of (5-12) by

$$\sum_{P'\in\mathbb{P}'_{k'}} \prod_{j=1}^{3}\Bigl(\int_{\mathbb{R}} |\langle f_j, \Phi^j_{P'_j}\rangle|^2\,dx_2\Bigr)^{\frac{1-\theta_j}{2}} \lesssim \prod_{j=1}^{3}\Bigl(\int_{\mathbb{R}} \sum_{P'\in\mathbb{P}'_{k'}} |\langle f_j, \Phi^j_{P'_j}\rangle|^2\,dx_2\Bigr)^{\frac{1-\theta_j}{2}}$$

$$\lesssim \prod_{j=1}^{3} \|f_j\|_{L^2(\mathbb{R}^2)}^{1-\theta_j} \lesssim |E_1|^{(1-\theta_1)/2}|E_2|^{(1-\theta_2)/2}. \tag{5-15}$$

Combining the estimates (5-8), (5-12) and (5-15), we arrive at

$$\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1,f_2,f_3) \lesssim |E_1|^{\theta_1}|E_2|^{\theta_2}|E_1|^{(1-\theta_1)/2}|E_2|^{(1-\theta_2)/2}\sum_{k'\in\mathbb{Z}}\tilde{C}^{\varepsilon}_{k'} \lesssim_{\varepsilon,\theta_1,\theta_2,\theta_3} |E_1|^{(1+\theta_1)/2}|E_2|^{(1+\theta_2)/2} \tag{5-16}$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1+\theta_2+\theta_3 = 1$.

By taking θ_1 sufficiently close to 1 and θ_2 sufficiently close to 0, one can make the exponent $2/(1+\theta_1) = p_1$ strictly larger than 1 and close to 1, and $2/(1+\theta_2) = p_2$ strictly smaller than 2 and close to 2. We finally get the estimate

$$\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},I}(f_1,f_2,f_3) \lesssim_{\varepsilon,p,p_1,p_2} |E_1|^{1/p_1}|E_2|^{1/p_2} \tag{5-17}$$

for every $\varepsilon > 0$, and p, p_1, p_2 satisfy the hypothesis of Proposition 2.17.

5B. ***Estimates for the trilinear form*** $\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}(f_1,f_2,f_3)$. For each fixed $P' \in \mathbb{P}'$, we define the corresponding subcollection of $\mathbb{P}''$ by

$$\mathbb{P}''_{P'} := \{P'' \in \mathbb{P}'' : I_{\vec{P}} \subseteq \Omega\},$$

then we can decompose the collection $\mathbb{P}''_{P'}$ further, as follows:

$$\mathbb{P}''_{P'} = \bigcup_{\mu\in\mathbb{N}} \mathbb{P}''_{P',\mu}, \tag{5-18}$$

where

$$\mathbb{P}''_{P',\mu} := \{P'' \in \mathbb{P}''_{P'} : \mathrm{Dil}_{2^\mu}(I_{P'}\times I_{P''}) \subseteq \Omega\} \tag{5-19}$$

and μ is maximal with this property. By $\mathrm{Dil}_{2^\mu}(I_{\vec{P}})$ we mean the rectangle having the same center as the original $I_{\vec{P}}$ but whose side lengths are 2^μ times larger.

Now we apply both the decompositions of $\tilde{\mathbb{P}}'$ and $\mathbb{P}''_{P'}$ defined in (5-5) and (5-18) at the same time, and split the trilinear form $\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}(f_1,f_2,f_3)$ into

$$\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}(f_1,f_2,f_3) = \sum_{k'\in\mathbb{Z}}\sum_{P'\in\mathbb{P}'_{k'}} |\tilde{C}^{\varepsilon}_{Q_{\vec{P}}}||I_{P'}| \sum_{\mu\in\mathbb{N}}\sum_{P''\in\mathbb{P}''_{P',\mu}} \frac{1}{|I_{P''}|^{\frac{1}{2}}} \prod_{j=1}^{3}\biggl|\biggl\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\biggr\rangle\biggr| \tag{5-20}$$

In the inner sum of (5-20), since $\mathrm{Dil}_{2^\mu}(I_{P'}\times I_{P''}) \subseteq \Omega$ and $\mathrm{supp}\, f_3 \subseteq E'_3 \subseteq \mathbb{R}^2\setminus\Omega$, we get that

$$|f_3| \le \chi_{E'_3}\chi_{(\mathrm{Dil}_{2^\mu}(I_{P'}\times I_{P''}))^c} = \chi_{E'_3}\{\chi_{(2^\mu I_{P'})^c} + \chi_{(2^\mu I_{P''})^c} - \chi_{(2^\mu I_{P'})^c}\chi_{(2^\mu I_{P''})^c}\}, \tag{5-21}$$

and hence we can assume hereafter in this subsection that

$$|f_3| \le \chi_{E_3'} \chi_{(2^\mu I_{P'})^c}, \tag{5-22}$$

and the other two terms can be handled similarly.

By using Proposition 3.5 and (5-8), we derive from (5-20) the estimates

$$\begin{aligned}\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}&(f_1, f_2, f_3)\\ &\lesssim \sum_{k'\in\mathbb{Z}} \tilde{C}^{\varepsilon}_{k'} \sum_{P'\in\mathbb{P}'_{k'}} |I_{P'}| \sum_{\mu\in\mathbb{N}} \prod_{j=1}^{3} \Biggl[\biggl(\text{energy}_j\biggl(\biggl(\biggl\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\biggr\rangle\biggr)_{P''\in\mathbb{P}''_{P',\mu}}\biggr)\biggr)^{1-\theta_j}\\ &\qquad\qquad \times \biggl(\text{size}_j\biggl(\biggl(\biggl\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\biggr\rangle\biggr)_{P''\in\mathbb{P}''_{P',\mu}}\biggr)\biggr)^{\theta_j}\Biggr]\end{aligned} \tag{5-23}$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$.

To estimate the inner sum in the right-hand side of (5-23), note that $I_{\vec{P}} \subseteq \Omega$, $P'' \in \mathbb{P}''_{P',\mu}$ and f_3 satisfies (5-22), so we apply the size estimates in Lemma 3.6 and get, for each $P' \in \mathbb{P}'_{k'}$ and $\mu \in \mathbb{N}$,

$$\text{size}_1\biggl(\biggl(\biggl\langle \frac{\langle f_1, \Phi^1_{P'_1}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^1_{P''_1}\biggr\rangle\biggr)_{P''\in\mathbb{P}''_{P',\mu}}\biggr) \lesssim \sup_{P''\in\mathbb{P}''_{P',\mu}} \frac{1}{|I_{P''}|} \int_{\mathbb{R}} \biggl|\frac{\langle f_1, \Phi^1_{P'_1}\rangle}{|I_{P'}|^{\frac{1}{2}}}\biggr| \tilde{\chi}^M_{I_{P''}}\, dx \lesssim 2^{2\mu}|E_1|, \tag{5-24}$$

$$\text{size}_2\biggl(\biggl(\biggl\langle \frac{\langle f_2, \Phi^2_{P'_2}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^2_{P''_2}\biggr\rangle\biggr)_{P''\in\mathbb{P}''_{P',\mu}}\biggr) \lesssim \sup_{P''\in\mathbb{P}''_{P',\mu}} \frac{1}{|I_{P''}|} \int_{\mathbb{R}} \biggl|\frac{\langle f_2, \Phi^2_{P'_2}\rangle}{|I_{P'}|^{\frac{1}{2}}}\biggr| \tilde{\chi}^M_{I_{P''}}\, dx \lesssim 2^{2\mu}|E_2|, \tag{5-25}$$

$$\text{size}_3\biggl(\biggl(\biggl\langle \frac{\langle f_3, \Phi^3_{P'_3}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^3_{P''_3}\biggr\rangle\biggr)_{P''\in\mathbb{P}''_{P',\mu}}\biggr) \lesssim \sup_{P''\in\mathbb{P}''_{P',\mu}} \frac{1}{|I_{P''}|} \int_{\mathbb{R}} \biggl|\frac{\langle f_3, \Phi^3_{P'_3}\rangle}{|I_{P'}|^{\frac{1}{2}}}\biggr| \tilde{\chi}^M_{I_{P''}}\, dx \lesssim 2^{-N\mu}, \tag{5-26}$$

where $M > 0$ and $N > 0$ are arbitrarily large. By applying the energy estimates in Lemma 3.7, we have, for each $P' \in \mathbb{P}'_{k'}$, $\mu \in \mathbb{N}$ and $j = 1, 2, 3$,

$$\text{energy}_j\biggl(\biggl(\biggl\langle \frac{\langle f_j, \Phi^j_{P'_j}\rangle}{|I_{P'}|^{\frac{1}{2}}}, \Phi^j_{P''_j}\biggr\rangle\biggr)_{P''\in\mathbb{P}''_{P',\mu}}\biggr) \lesssim \frac{1}{|I_{P'}|^{\frac{1}{2}}} \biggl(\int_{\mathbb{R}} |\langle f_j, \Phi^j_{P'_j}\rangle|^2\, dx_2\biggr)^{\frac{1}{2}}. \tag{5-27}$$

Now we insert the size and energy estimates (5-24)–(5-27) into (5-23); by using the estimates (5-8) and (5-15), we derive that

$$\begin{aligned}\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3) &\lesssim |E_1|^{\theta_1}|E_2|^{\theta_2} \sum_{k'\in\mathbb{Z}} \tilde{C}^{\varepsilon}_{k'} \sum_{\mu\in\mathbb{N}} 2^{-(N\theta_3-2)\mu} \sum_{P'\in\mathbb{P}'_{k'}} \prod_{j=1}^{3} \biggl(\int_{\mathbb{R}} |\langle f_j, \Phi^j_{P'_j}\rangle|^2\, dx_2\biggr)^{\frac{1-\theta_j}{2}}\\ &\lesssim_{\varepsilon,\theta_1,\theta_2,\theta_3,N} |E_1|^{(1+\theta_1)/2}|E_2|^{(1+\theta_2)/2} \sum_{\mu\in\mathbb{N}} 2^{-(N\theta_3-2)\mu}.\end{aligned} \tag{5-28}$$

for every $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$.

By taking θ_1 sufficiently close to 1 and θ_2 sufficiently close to 0, one can make the exponent $2/(1+\theta_1) = p_1$ strictly larger than 1 and close to 1, and $2/(1+\theta_2) = p_2$ strictly smaller than 2 and close to 2. The series over $\mu \in \mathbb{N}$ in (5-28) is summable if we choose N large enough (say, $N \simeq 4\theta_3^{-1}$). We finally get the estimate

$$\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}},II}(f_1, f_2, f_3) \lesssim_{\varepsilon,p,p_1,p_2} |E_1|^{1/p_1}|E_2|^{1/p_2} \tag{5-29}$$

for any $\varepsilon > 0$, and p, p_1, p_2 satisfy the hypothesis of Proposition 2.17.

5C. ***Conclusions.*** By inserting the estimates (5-17) and (5-29) into (5-4), we finally get

$$|\tilde{\Lambda}^{\varepsilon}_{\vec{\mathbb{P}}}(f_1, f_2, f_3)| \lesssim_{\varepsilon,p,p_1,p_2} |E_1|^{1/p_1}|E_2|^{1/p_2} \tag{5-30}$$

for any $\varepsilon > 0$, which completes the proof of Proposition 2.17 for the model operators $\tilde{\Pi}^{\varepsilon}_{\vec{\mathbb{P}}}$.

This concludes the proof of Theorem 1.5.

References

[Bernicot 2008] F. Bernicot, "Local estimates and global continuities in Lebesgue spaces for bilinear operators", *Anal. PDE* **1**:1 (2008), 1–27. MR 2009e:42019 Zbl 1151.42005

[Chen and Lu 2014] J. Chen and G. Lu, "Hörmander type theorems for multi-linear and multi-parameter Fourier multiplier operators with limited smoothness", *Nonlinear Anal.* **101** (2014), 98–112. MR 3178381 Zbl 1286.42010

[Christ and Journé 1987] M. Christ and J.-L. Journé, "Polynomial growth estimates for multilinear singular integral operators", *Acta Math.* **159**:1-2 (1987), 51–80. MR 89a:42024 Zbl 0645.42017

[Coifman and Meyer 1978] R. R. Coifman and Y. Meyer, *Au delà des opérateurs pseudo-différentiels*, Astérisque **57**, Société Mathématique de France, Paris, 1978. MR 81b:47061 Zbl 0483.35082

[Coifman and Meyer 1997] R. Coifman and Y. Meyer, *Wavelets: Calderón–Zygmund and multilinear operators*, Cambridge Studies in Advanced Mathematics **48**, Cambridge University Press, 1997. MR 98e:42001 Zbl 0916.42023

[Dai and Lu ≥ 2015a] W. Dai and G. Lu, "L^p estimates for multi-linear and multi-parameter pseudo-differential operators", preprint. To appear in *Bull. Soc. Math. France.* arXiv 1308.4062

[Dai and Lu ≥ 2015b] W. Dai and G. Lu, "L^p estimates for the bilinear Hilbert transform for $\frac{1}{2} < p \le \frac{2}{3}$: a counterexample and generalizations to non-smooth symbols", preprint. arXiv 1409.3875

[Demeter and Thiele 2010] C. Demeter and C. Thiele, "On the two-dimensional bilinear Hilbert transform", *Amer. J. Math.* **132**:1 (2010), 201–256. MR 2011b:42040 Zbl 1206.42010

[Fefferman and Stein 1982] R. Fefferman and E. M. Stein, "Singular integrals on product spaces", *Adv. in Math.* **45**:2 (1982), 117–143. MR 84d:42023 Zbl 0517.42024

[Gilbert and Nahmod 2001] J. E. Gilbert and A. R. Nahmod, "Bilinear operators with non-smooth symbol, I", *J. Fourier Anal. Appl.* **7**:5 (2001), 435–467. MR 2002m:42009 Zbl 0994.42014

[Grafakos and Li 2004] L. Grafakos and X. Li, "Uniform bounds for the bilinear Hilbert transforms, I", *Ann. of Math.* (2) **159**:3 (2004), 889–933. MR 2006e:42011 Zbl 1071.44004

[Grafakos and Tao 2003] L. Grafakos and T. Tao, "Multilinear interpolation between adjoint operators", *J. Funct. Anal.* **199**:2 (2003), 379–385. MR 2004b:42022 Zbl 1031.46026

[Grafakos and Torres 2002a] L. Grafakos and R. H. Torres, "Multilinear Calderón–Zygmund theory", *Adv. Math.* **165**:1 (2002), 124–164. MR 2002j:42029 Zbl 1032.42020

[Grafakos and Torres 2002b] L. Grafakos and R. H. Torres, "On multilinear singular integrals of Calderón–Zygmund type", pp. 57–91 in *Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations* (El Escorial, 2000), edited by P. Cifuentes et al., 2002. MR 2004c:42031 Zbl 1016.42009

[Hong and Lu 2014] Q. Hong and G. Lu, "Symbolic calculus and boundedness of multi-parameter and multi-linear pseudo-differential operators", *Adv. Nonlinear Stud.* **14**:4 (2014), 1055–1082. MR 3269385 Zbl 06382833

[Janson 1988] S. Janson, "On interpolation of multilinear operators", pp. 290–302 in *Function spaces and applications* (Lund, 1986), edited by M. Cwikel et al., Lecture Notes in Math. **1302**, Springer, Berlin, 1988. MR 89g:46114 Zbl 0827.46062

[Journé 1985] J.-L. Journé, "Calderón–Zygmund operators on product spaces", *Rev. Mat. Iberoamericana* **1**:3 (1985), 55–91. MR 88d:42028 Zbl 0634.42015

[Jung ≥ 2015] J. Jung, "Iterated trilinear Fourier integrals with arbitrary symbols", preprint. arXiv 1311.1574

[Kenig and Stein 1999] C. E. Kenig and E. M. Stein, "Multilinear estimates and fractional integration", *Math. Res. Lett.* **6**:1 (1999), 1–15. MR 2000k:42023a Zbl 0952.42005

[Kesler ≥ 2015] R. Kesler, "Mixed estimates for degenerate multilinear oscillatory integrals and their tensor product generalizations", preprint. arXiv 1311.2322

[Lacey 2000] M. T. Lacey, "The bilinear maximal functions map into L^p for $\frac{2}{3} < p \leq 1$", *Ann. of Math.* (2) **151**:1 (2000), 35–57. MR 2001b:42015 Zbl 0967.47031

[Lacey and Thiele 1997] M. Lacey and C. Thiele, "L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$", *Ann. of Math.* (2) **146**:3 (1997), 693–724. MR 99b:42014 Zbl 0914.46034

[Lacey and Thiele 1999] M. Lacey and C. Thiele, "On Calderón's conjecture", *Ann. of Math.* (2) **149**:2 (1999), 475–496. MR 2000d:42003 Zbl 0934.42012

[Li 2006] X. Li, "Uniform bounds for the bilinear Hilbert transforms, II", *Rev. Mat. Iberoam.* **22**:3 (2006), 1069–1126. MR 2008c:42014 Zbl 1133.42022

[Luthy 2013] P. M. Luthy, *Bi-parameter maximal multilinear operators*, Ph.D. thesis, Cornell University, 2013. MR 3152552

[Muscalu and Schlag 2013] C. Muscalu and W. Schlag, *Classical and multilinear harmonic analysis, II*, Cambridge Studies in Advanced Mathematics **138**, Cambridge University Press, Cambridge, 2013. MR 3052499 Zbl 1282.42001

[Muscalu et al. 2002] C. Muscalu, T. Tao, and C. Thiele, "Multi-linear operators given by singular multipliers", *J. Amer. Math. Soc.* **15**:2 (2002), 469–496. MR 2003b:42017 Zbl 0994.42015

[Muscalu et al. 2004a] C. Muscalu, J. Pipher, T. Tao, and C. Thiele, "Bi-parameter paraproducts", *Acta Math.* **193**:2 (2004), 269–296. MR 2005m:42028 Zbl 1087.42016

[Muscalu et al. 2004b] C. Muscalu, T. Tao, and C. Thiele, "L^p estimates for the biest, II: The Fourier case", *Math. Ann.* **329**:3 (2004), 427–461. MR 2005k:42054 Zbl 1073.42010

[Muscalu et al. 2006] C. Muscalu, J. Pipher, T. Tao, and C. Thiele, "Multi-parameter paraproducts", *Rev. Mat. Iberoam.* **22**:3 (2006), 963–976. MR 2008b:42037 Zbl 1114.42005

[Silva 2014] P. Silva, "Vector-valued inequalities for families of bilinear Hilbert transforms and applications to bi-parameter problems", *J. Lond. Math. Soc.* (2) **90**:3 (2014), 695–724. MR 3291796 Zbl 06382012

[Thiele 2002] C. Thiele, "A uniform estimate", *Ann. of Math.* (2) **156**:2 (2002), 519–563. MR 2003i:47036 Zbl 1038.42019

[Thiele 2006] C. Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics **105**, American Mathematical Society, Providence, RI, 2006. MR 2006m:42073 Zbl 1101.42001

Received 3 May 2014. Accepted 22 Jan 2015.

WEI DAI: weidai@buaa.edu.cn
School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, China

and

School of Mathematics and Systems Science, Beihang University (BUAA), Beijing 100191, China

GUOZHEN LU: gzlu@math.wayne.edu
Department of Mathematics, Wayne State University, 656 W. Kirby Street, Detroit, MI 48202, United States

ANALYSIS AND PDE
Vol. 8, No. 3, 2015
dx.doi.org/10.2140/apde.2015.8.713

LARGE BMO SPACES VS INTERPOLATION

JOSE M. CONDE-ALONSO, TAO MEI AND JAVIER PARCET

We introduce a class of BMO spaces which interpolate with L_p and are sufficiently large to serve as endpoints for new singular integral operators. More precisely, let (Ω, Σ, μ) be a σ-finite measure space. Consider two filtrations of Σ by successive refinement of two atomic σ-algebras Σ_{a} and Σ_{b} having trivial intersection. Construct the corresponding truncated martingale BMO spaces. Then, the intersection seminorm only leaves out constants and we provide a quite flexible condition on $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ so that the resulting space interpolates with L_p in the expected way. In the presence of a metric d, we obtain endpoint estimates for Calderón–Zygmund operators on (Ω, μ, d) under additional conditions on $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$. These are weak forms of the "isoperimetric" and the "locally doubling" properties of Carbonaro, Mauceri and Meda which admit less concentration at the boundary. Examples of particular interest include densities of the form $e^{\pm|x|^\alpha}$ for any $\alpha > 0$ or $(1+|x|^\beta)^{-1}$ for any $\beta \gtrsim n^{3/2}$. A (limited) comparison with Tolsa's RBMO is also possible. On the other hand, a more intrinsic formulation yields a Calderón–Zygmund theory adapted to regular filtrations over $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ without using a metric. This generalizes well-known estimates for perfect dyadic and Haar shift operators. In contrast to previous approaches, ours extends to matrix-valued functions (via recent results from noncommutative martingale theory) for which only limited results are known and no satisfactory nondoubling theory exists so far.

Introduction

A BMO space is a set of functions that enjoy bounded mean oscillation in a certain sense. Both "mean" and "oscillation" can be measured in many different ways. Most frequently, we find BMO spaces refer to averages over balls in a metric measure space. In other notable scenarios, we may replace these averages by conditional expectations with respect to a martingale filtration, or even by the action of a nicely behaved semigroup of operators. These more abstract formulations are known to be very useful given the lack of appropriate metrics. The relation between metric and martingale BMO spaces is well understood for doubling spaces, that is, when the measure of a ball in the given metric is comparable with the measure of its concentric dilations up to constants depending on the dilation factor but not on the chosen ball. Indeed, in this case the metric BMO is equivalent to a finite intersection of martingale BMO spaces constructed out of dyadic two-sided filtrations of atomic σ-algebras whose atoms look like balls; see [Conde 2013; Garnett and Jones 1982; Hytönen and Kairema 2012; Mei 2003]. What is more relevant, however, is that any of these martingale BMO spaces satisfies the following fundamental properties:

(i) Interpolation endpoint for the L_p scale.

MSC2010: 42B20, 42B35, 46L52, 60G46.
Keywords: nondoubling measures, BMO spaces, interpolation, martingales, noncommutative harmonic analysis, classical harmonic analysis, Calderón–Zygmund theory.

(ii) John–Nirenberg inequalities and H_1–BMO duality.

(iii) CZ extrapolation: L_2-boundedness $\Rightarrow L_\infty \to$ BMO boundedness.

Hence, these spaces yield at least as many endpoint estimates as the metric BMO.

The main goal of this paper is to construct BMO spaces satisfying the properties stated above for a larger class of measures, and to explore the implications of this construction to provide new endpoint estimates. The first attempts in this direction [Mateu et al. 2000; Nazarov et al. 2002] culminated in the work of Tolsa [2001] on so-called RBMO spaces. These spaces enjoy the above-mentioned properties for measures of polynomial growth. There are, however, a couple of open questions concerning Tolsa's construction. In the first place, Calderón–Zygmund extrapolation holds under a Lipschitz kernel condition instead of the more flexible Hörmander condition. Second, only interpolation of operators, has been studied but it seems to be unknown whether these spaces interpolate with the L_p scale. These two problems were solved by Carbonaro, Mauceri and Meda [Carbonaro et al. 2009; 2010] for a different class of measures, based on similar results for the Gaussian measure on Euclidean spaces [Mauceri and Meda 2007]. The properties they imposed lead to locally doubling measures with certain concentration behavior at the boundary. In both cases — up to equivalence in the norm and additional conditions — only doubling balls are used to measure the mean oscillation of the function.

We present an alternative approach to these questions. Martingale BMO spaces always satisfy conditions (i) and (ii) above, with independence of the existence of a metric in the underlying measure space. The third property however requires additional structure on our BMO spaces. Indeed, assume for a moment that we work with a two-sided filtration $(\Sigma_k)_{k\in\mathbb{Z}}$ of atomic σ-subalgebras of Σ with corresponding conditional expectations E_{Σ_k}. If $\mathbf{\Pi}$ denotes the union of atoms in our filtration, the corresponding martingale BMO norm is given by

$$\|f\|_{\mathrm{BMO}} = \sup_{k\in\mathbb{Z}} \big\| \mathsf{E}_{\Sigma_k} |f - \mathsf{E}_{\Sigma_{k-1}} f|^2 \big\|_\infty^{1/2},$$

which is larger than the function BMO norm

$$\sup_{\mathrm{A}\in\mathbf{\Pi}} \bigg(\frac{1}{\mu(\mathrm{A})} \int_{\mathrm{A}} \bigg| f(w) - \frac{1}{\mu(\mathrm{A})} \int_{\mathrm{A}} f\, d\mu \bigg|^2 d\mu(w) \bigg)^{\frac{1}{2}}.$$

Thus, if we admit from [Carbonaro et al. 2009; Tolsa 2001] that extrapolation for (nonlocal) Calderón–Zygmund operators imposes that our atoms be doubling — i.e., contained in a doubling ball of comparable measure or a union of at most C_0 sets of this kind; see below — we immediately find obstructions to constructing filtrations satisfying this assumption for nondoubling spaces. We propose to consider a sort of intersection of two large BMO spaces as follows. Consider a σ-finite measure space (Ω, Σ, μ) and two atomic σ-algebras $\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}}$ of measurable sets in Σ satisfying $\Sigma_{\mathrm{a}} \cap \Sigma_{\mathrm{b}} = \{\Omega, \varnothing\}$. Write BMO_j for any martingale BMO space over a filtration $(\Sigma_{jk})_{k\geq 1}$ with $\Sigma_{j1} = \Sigma_j$; then the seminorm

$$\|f\|_{\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)} = \max\{ \|f - \mathsf{E}_{\Sigma_{\mathrm{a}}} f\|_{\mathrm{BMO}_{\mathrm{a}}},\ \|f - \mathsf{E}_{\Sigma_{\mathrm{b}}} f\|_{\mathrm{BMO}_{\mathrm{b}}} \}$$

vanishes on constant functions precisely when $\Sigma_a \cap \Sigma_b$ is trivial. Let

$$\mathrm{BMO}_{\Sigma_{ab}}(\Omega) = \{f \in L^1_{\mathrm{loc}}(\Omega) \mid \|f\|_{\mathrm{BMO}_{\Sigma_{ab}}(\Omega)} < \infty\}/\mathbb{C}.$$

This settles a model of "large BMO spaces" which easily satisfy property (ii) and leave some room for property (iii). The problem reduces then to identify conditions on the pair (Σ_a, Σ_b) so that $\mathrm{BMO}_{\Sigma_{ab}}(\Omega)$ interpolates with the L_p scale. A standard argument shows that this is the case when

$$\|f\|_{L_p^\circ(\Omega)} := \inf_{\mathrm{k}\in\mathbb{C}} \|f - \mathrm{k}\|_p \sim \max\{\|f - \mathsf{E}_{\Sigma_a} f\|_p, \|f - \mathsf{E}_{\Sigma_b} f\|_p\} =: \|f\|_{L^p_{\Sigma_{ab}}(\Omega)}$$

for $2 \leq p < \infty$, where

$$\begin{aligned} L_p^\circ(\Omega) &= L_p(\Omega, \Sigma, \mu)/\mathbb{C}, \\ L^p_{\Sigma_{ab}}(\Omega) &= \{f \in L^1_{\mathrm{loc}}(\Omega) \mid \|f\|_{L^p_{\Sigma_{ab}}(\Omega)} < \infty\}/\mathbb{C} \\ &= L_p(\Omega, \Sigma, \mu)/\Sigma_a \wedge L_p(\Omega, \Sigma, \mu)/\Sigma_b. \end{aligned}$$

Here, $L_p(\Omega, \Sigma, \mu)/\Sigma_i$ denotes the quotient space of $L_p(\Omega, \Sigma, \mu)$ by the subspace of Σ_i-measurable functions. More precisely, we have an isomorphism $L_p^\circ(\Omega) \simeq L^p_{\Sigma_{ab}}(\Omega)$. It should be mentioned that this isomorphism fails in general, even for the Lebesgue measure in $\mathbb{R}^n$ and many "natural" choices of pairs (Σ_a, Σ_b). Recall that $L_p^\circ(\Omega) = L_p(\Omega)$ for infinite measures. Note also that we use $\wedge$ and not $\cap$ since this space is not really an intersection; we shall also write $\mathrm{BMO}_{\Sigma_{ab}}(\Omega) = \mathrm{BMO_a}(\Omega)/\Sigma_a \wedge \mathrm{BMO_b}(\Omega)/\Sigma_b$. To formulate a sufficient condition on (Σ_a, Σ_b) for $L_p^\circ(\Omega) \simeq L^p_{\Sigma_{ab}}(\Omega)$, let Π_j be the set of atoms in Σ_j. When $\mu(\Omega) < \infty$ we shall consider two distinguished atoms $(A_0, B_0) \in \Pi_a \times \Pi_b$, while for μ not finite we take $A_0 = B_0 = \varnothing$ for notation consistency. Given $(A, B) \in \Pi_a \times \Pi_b$, set

$$R_A = \{B' \in \Pi_b \mid \mu(A \cap B') > 0\} \quad \text{and} \quad R_B = \{A' \in \Pi_a \mid \mu(A' \cap B) > 0\}.$$

We will write $|R_A|$ and $|R_B|$ for the cardinality of these sets. The following is the main result of this paper, where we establish a condition on (Σ_a, Σ_b) which suffices to make intersections and quotients commute in L_p as described above. We will say that (Σ_a, Σ_b) is an *admissible covering* of (Ω, Σ, μ) when $\Sigma_a \cap \Sigma_b = \{\Omega, \varnothing\}$ and

$$\min\left\{ \sup_{A\in\Pi_a\setminus\{A_0\}} \sum_{B\in R_A} |R_B| \frac{\mu(A\cap B)^2}{\mu(A)\mu(B)}, \sup_{B\in\Pi_b\setminus\{B_0\}} \sum_{A\in R_B} |R_A| \frac{\mu(A\cap B)^2}{\mu(A)\mu(B)} \right\} < 1.$$

Theorem A. *Let* (Ω, Σ, μ) *be a* σ*-finite measure space equipped with an admissible covering* (Σ_a, Σ_b). *Then, for each* $2 \leq p < \infty$, *there exists a constant* c_p, *depending only on* p *and the admissible covering, such that*

$$L_p^\circ(\Omega) \simeq_{c_p} L^p_{\Sigma_{ab}}(\Omega).$$

Moreover, we have the desired complex interpolation result,

$$[\mathrm{BMO}_{\Sigma_{ab}}(\Omega), L_1^\circ(\Omega)]_{1/q} \simeq_{c_q} L_q^\circ(\Omega) \quad (1 < q < \infty),$$

with $\mathrm{BMO}_{\Sigma_{ab}}(\Omega)$ *defined as above for any two martingale* BMO *spaces over* (Σ_a, Σ_b).

The first assertion fails for $p = 1, \infty$. On the other hand, both the John–Nirenberg inequalities and H_1–BMO duality are easily formulated for these spaces. Therefore, we shall focus in what follows on condition (iii). Calderón–Zygmund extrapolation means that under a certain mild smoothness condition on the kernel, L_2-boundedness yields L_p-boundedness for $1 < p < \infty$. As usual, we handle it by providing an endpoint estimate for interpolation. Let d be a metric on Ω and denote by $\alpha \mathrm{B}$ the α-dilation of a ball B. We impose the standard Hörmander kernel condition

$$\sup_{\mathrm{B}\, d\text{-ball}} \sup_{z_1, z_2 \in \mathrm{B}} \int_{\Omega \setminus \alpha \mathrm{B}} |k(z_1, x) - k(z_2, x)| + |k(x, z_1) - k(x, z_2)| \, d\mu(x) < \infty.$$

Define a CZO on (Ω, μ, d) as any linear map T satisfying the following properties:

- T is well-defined and bounded on $L_2(\Omega)$.
- The kernel representation for any $f \in \mathcal{C}_c(\Omega)$,

$$Tf(x) = \int_\Omega k(x, y) f(y) \, d\mu(y) \quad \text{holds for } x \notin \operatorname{supp} f$$

  and some kernel $k : \Omega \times \Omega \setminus \Delta \to \mathbb{C}$ satisfying the Hörmander condition.

Given $C_0 > 0$, a Σ-measurable set A will be called (C_0, α, β)-doubling when it is the union of at most C_0 sets which are contained in (α, β)-doubling balls — balls B such that $\mu(\alpha B) \le \beta \mu(B)$ — of comparable measure up to the constant C_0. Recall that a filtration $(\Sigma_k)_{k \ge 1}$ is called *regular* if $\mathsf{E}_k f \lesssim \mathsf{E}_{k-1} f$ for all $k > 1$ and all $f \ge 0$.

Theorem B1. *Let* $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ *be an admissible covering of* (Ω, Σ, μ). *Assume that* (Ω, Σ, μ) *admits regular filtrations* $(\Sigma_{jk})_{k \ge 1}$ *by successive refinement of* $\Sigma_{j1} = \Sigma_j$ *for* $j = \mathrm{a}, \mathrm{b}$ *and that each atom in* Σ_{jk} *is* (C_0, α, β)*-doubling for certain absolute constants* $C_0, \alpha, \beta > 0$. *Construct the spaces* $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$ *which are defined over these filtrations. Then, every Calderón–Zygmund operator extends to a bounded map* $L_\infty(\Omega) \to \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$, *and* $L_p(\Omega) \to L_p(\Omega)$ *for* $1 < p < \infty$.

A few illustrations of Theorem B1 are the following:

• *Doubling case*: Theorem B1 recovers Calderón–Zygmund extrapolation on homogeneous spaces (Ω, μ, d). We shall construct explicit pairs $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ and martingale filtrations satisfying our assumptions.

• *Polynomial growth*: Given any (Ω, μ, d) with polynomial growth, it is not difficult to construct atomic σ-algebras composed uniquely of doubling atoms, even giving admissible coverings. Under the existence of filtrations based on $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ and composed of doubling atoms — regular or not — we may prove that Tolsa's RBMO sits inside our $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$. This condition seems, unfortunately, a restrictive limit in Theorem B1. However, it can be checked in some concrete scenarios, like for

$$d\mu(x) = \frac{dx}{1 + |x|^\beta} \quad \text{with } \beta \gtrsim n^{3/2}$$

in $\mathbb{R}^n$ equipped with the Euclidean metric. Note that μ is doubling for $\beta < n$. The key advantage over Tolsa's approach is that we only need to impose Hörmander kernel smoothness, instead of stronger Lipschitz conditions. This was also achieved by [Carbonaro et al. 2009; 2010] for another family of measures (see below) but not for the measures considered above, since they are drastically less concentrated at the boundary for any β.

• *Concentration at the boundary*: Carbonaro et al. [2009; 2010] proved that when (Ω, μ, d) is locally doubling and the measure concentrates at the boundary of open sets in a certain sense — together with a purely metric condition that does not play any role here — a BMO space satisfying (i), (ii) and (iii) is possible. Their main examples in $\mathbb{R}^n$ with a weighted Euclidean metric were $d\mu(x) = e^{\pm|x|^\alpha}\,dx$ and $\alpha > 1$. The exponentially decreasing ones behave in some sense like the Gaussian measure, which was studied a few years before by Mauceri and Meda. It is of polynomial growth, so that the kernel smoothness condition was the main advantage with respect to Tolsa's approach. The exponentially increasing ones are not of polynomial growth. In this paper we shall remove their condition $\alpha > 1$.

In the literature, we find other families of operators — with no need of a metric in the underlying space — which are close to CZOs in spirit. Martingale transforms are the simplest ones, but are local and much easier to bound. Nonlocal models include the so-called perfect dyadic CZOs and, most notably, Haar shift operators, which include prominent examples like the discrete Hilbert transform and dyadic paraproducts. In these cases, the Hörmander kernel condition can be replaced by

$$\sup_{Q \text{ dyadic cube}} \sup_{z_1, z_2 \in Q} \int_{\Omega \setminus \hat{Q}} |k(z_1, x) - k(z_2, x)| + |k(x, z_1) - k(x, z_2)|\, d\mu(x) < \infty,$$

where $\hat{Q}$ denotes the dyadic father of Q. Our BMO spaces allow us to further replace dyadic cubes in dyadically doubling measure spaces — see [López-Sánchez et al. 2014] for recent progress on more general measures in this direction — by more general atoms. Namely, assume $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ gives an admissible covering of (Ω, Σ, μ). Consider regular filtrations of atomic σ-algebras $(\Sigma_{jk})_{k\geq 1}$ with $\Sigma_{j1} = \Sigma_j$ for $j = \mathrm{a}, \mathrm{b}$. Let us write Π_{jk} for the family of atoms in the atomic σ-algebra Σ_{jk} and set $\mathbf{\Pi}_j = \bigcup_{k\geq 1} \Pi_{jk}$. Then, consider the following Hörmander-type kernel condition, where the former role of the metric d is replaced by the shape of our atoms in $\mathbf{\Pi} = \mathbf{\Pi}_{\mathrm{a}} \cup \mathbf{\Pi}_{\mathrm{b}}$:

$$\sup_{A \in \mathbf{\Pi}} \sup_{z_1, z_2 \in A} \int_{\Omega \setminus \hat{A}} |k(z_1, x) - k(z_2, x)| + |k(x, z_1) - k(x, z_2)|\, d\mu(x) < \infty.$$

Again, $\hat{A}$ denotes the minimal atom in the filtration of A which contains A properly, unless there is no such atom, in which case we pick $\hat{A} = A$. If we replace the Hörmander condition by this one, we obtain another class of "atomic" CZOs, which will be denoted in what follows by ACZO.

Theorem B2. *Let* $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ *be an admissible covering of* (Ω, Σ, μ). *Assume in addition that* (Ω, Σ, μ) *admits regular filtrations* $(\Sigma_{jk})_{k\geq 1}$ *by successive refinement of* $\Sigma_{j1} = \Sigma_j$ *for* $j = \mathrm{a}, \mathrm{b}$. *Construct the spaces* $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$ *which are defined over these filtrations. Then, every* ACZO *extends to a bounded map* $L_\infty(\Omega) \to \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$, *and* $L_p(\Omega) \to L_p(\Omega)$ *for* $1 < p < \infty$.

An advantage of Theorem B2 is that our kernel conditions are flexible, since we may carefully choose (Σ_a, Σ_b) and the regular filtrations according to the concrete singular integral operator. It is worth mentioning that every σ-finite (atomless if μ is finite) measure space (Ω, Σ, μ) has nontrivial admissible coverings. Of course, the regularity of the filtration is a light form of "doublingness" needed to emulate the classical argument in this setting. We will also provide weaker estimates for pseudolocal operators when the filtrations are not regular.

In contrast to [Carbonaro et al. 2009; 2010; Mauceri and Meda 2007; Tolsa 2001], our approach extends to matrix-valued functions, for which only limited results are known and no satisfactory nondoubling theory exists so far. In fact, this was our original motivation and the necessity of alternative arguments led to the results presented so far. We will postpone the discussion of the matrix-valued setting to the last section of this paper, which will allow those readers not familiar with noncommutative L_p theory to isolate these results.

Our results above give some insight on the relation between nondoubling and martingale BMO theories; see [Conde-Alonso and Parcet 2014; Junge et al. $\geq$ 2015] for other results along this line. In [Conde-Alonso and Parcet 2014], we adapt Tolsa's ideas to give an atomic block description of martingale H_1. Semigroup BMO spaces are used in [Junge et al. $\geq$ 2015] to construct a Calderón–Zygmund theory that incorporates noncommutative measure spaces (von Neumann algebras) to the picture.

1. Admissible coverings and BMO spaces

In this section we recall some basic background around martingale BMO spaces and introduce our new class of BMO spaces. We will study standard properties of this class, like the existence of admissible coverings, John–Nirenberg inequalities and H_1–BMO duality. The proof of Theorem A is more technical and will be postponed to Section 2.

Martingale BMO spaces. Let (Ω, Σ, μ) be a σ-finite measure space and consider a filtration $(\Sigma_k)_{k\geq 1}$ of Σ. In other words, we have $\Sigma_k \subset \Sigma_{k+1}$ and the union of the spaces $L_\infty(\Omega, \Sigma_k, \mu)$ is weak-$*$ dense in $L_\infty(\Omega, \Sigma, \mu)$. Let E_{Σ_k} denote the conditional expectation onto Σ_k-measurable functions. Then, define the martingale BMO space associated to this filtration as the space of locally integrable functions $f : \Omega \to \mathbb{C}$ whose BMO norm,

$$\|f\|_{\mathrm{BMO}} = \sup_{k\geq 1} \big\| (\mathsf{E}_{\Sigma_k} |f - \mathsf{E}_{\Sigma_{k-1}} f|^2)^1/2 \big\|_\infty,$$

is finite, where we use the convention $\mathsf{E}_{\Sigma_0} f = 0$; see [Garsia 1973]. Another expression for the norm is

$$\begin{aligned}\|f\|_{\mathrm{BMO}} &= \sup_{k\geq 1} \Big\| |df_k|^2 + \sum_{n>k} \mathsf{E}_{\Sigma_k} |df_n|^2 \Big\|_\infty^{\frac{1}{2}} \\ &\sim \Big[\sup_{k\geq 1} \big\| (\mathsf{E}_{\Sigma_k} |f - \mathsf{E}_{\Sigma_k} f|^2)^1/2 \big\|_\infty + \|\mathsf{E}_{\Sigma_1} f\|_\infty\Big] + \sup_{k>1} \|df_k\|_\infty,\end{aligned}$$

where $df_k = \Delta_k f = \mathsf{E}_{\Sigma_k} f - \mathsf{E}_{\Sigma_{k-1}} f$. According to [Janson and Jones 1982], $[\mathrm{BMO}, L_1(\Omega)]_{1/p} \simeq L_p(\Omega)$ for any filtration we pick. The bracketed term in the right-hand side above is called the martingale bmo norm of f, and it is closer to the standard expressions to measure the mean oscillation of a function. Namely, if the σ-algebras Σ_k are atomic and if Π_k denotes the atoms in Σ_k and $\mathbf{\Pi} = \bigcup_{k\geq 1} \Pi_k$, we deduce that

$$\begin{aligned}\|f\|_{\mathrm{bmo}} &= \sup_{k\geq 1}\big\|(\mathsf{E}_{\Sigma_k}|f-\mathsf{E}_{\Sigma_k}f|^2)^1/2\big\|_\infty + \|\mathsf{E}_{\Sigma_1}f\|_\infty \\ &= \sup_{A\in\mathbf{\Pi}}\biggl(\frac{1}{\mu(A)}\int_A\biggl|f(w)-\frac{1}{\mu(A)}\int_A f\,d\mu\biggr|^2 d\mu(w)\biggr)^{\frac{1}{2}} + \sup_{A\in\Pi_1}\biggl|\frac{1}{\mu(A)}\int_A f\,d\mu\biggr|.\end{aligned}$$

Of course, using a selected family of atoms makes L_p-interpolation fail in general for bmo. The extra term in BMO corrects this. This should be compared with the extra condition in the definition of Tolsa's RBMO. On the other hand, bmo spaces have good interpolation properties with little Hardy spaces h_p. Namely, according to [Bekjan et al. 2010] we have $[\mathrm{bmo}, \mathrm{h}_1]_{1/p} \simeq \mathrm{h}_p$ for any filtration, where h_p is the closure of the space of finite martingales in L_p with respect to the norm

$$\|f\|_{\mathrm{h}_p} = \biggl\|\biggl(\sum_{k\geq 1}\mathsf{E}_{\Sigma_{k-1}}|df_k|^2\biggr)^{\frac{1}{2}}\biggr\|_p;$$

this time the convention is $\mathsf{E}_{\Sigma_0}|df_1|^2 = |\mathsf{E}_{\Sigma_1}f|^2$. In contrast to other BMO spaces seminorms, paradoxically, we will need to quotient out certain spaces. Note that, for Σ_1-measurable functions, the norms above coincide with the L_∞ norm

$$\|\mathsf{E}_{\Sigma_1}f\|_{\mathrm{BMO}} = \|\mathsf{E}_{\Sigma_1}f\|_{\mathrm{bmo}} = \|\mathsf{E}_{\Sigma_1}f\|_{L_\infty(\Omega)}.$$

If we define the seminorms

$$\begin{aligned}\|f\|^\circ_{\mathrm{bmo}} &= \|f-\mathsf{E}_{\Sigma_1}f\|_{\mathrm{bmo}},\\ \|f\|^\circ_{\mathrm{BMO}} &= \|f-\mathsf{E}_{\Sigma_1}f\|_{\mathrm{BMO}},\end{aligned}$$

we obtain complemented subspaces $\mathrm{BMO}_{\Sigma_1} = J_{\Sigma_1}(\mathrm{BMO})$ using the projection $J_{\Sigma_1} = \mathrm{id} - \mathsf{E}_{\Sigma_1}$. Indeed, it is a simple exercise using Jensen's conditional inequality $|\mathsf{E}_{\Sigma_1}f|^2 \leq \mathsf{E}_{\Sigma_1}|f|^2$; details are left to the reader. Since J_{Σ_1} is also bounded on h_p and L_p, the previous interpolation results imply the following isomorphisms for $1 < p < \infty$:

$$\begin{aligned}\bigl[J_{\Sigma_1}(\mathrm{bmo}), J_{\Sigma_1}(\mathrm{h}_1(\Omega))\bigr]_{1/p} &\simeq J_{\Sigma_1}(\mathrm{h}_p(\Omega)),\\ \bigl[J_{\Sigma_1}(\mathrm{BMO}), J_{\Sigma_1}(L_1(\Omega))\bigr]_{1/p} &\simeq J_{\Sigma_1}(L_p(\Omega)).\end{aligned}$$

Note that $J_{\Sigma_1}(L_p(\Omega)) \simeq L_p(\Omega, \Sigma, \mu)/\Sigma_1$ in the terminology of the introduction.

Remark 1.1. It is worth mentioning that the Janson–Jones interpolation theorem [1982] holds for arbitrary filtrations. In particular, we could replace $(\Sigma_k)_{k\geq 1}$ by $(\Sigma_k)_{k\geq \mathrm{N}}$ for some large N, and the latter BMO comes equipped with the norm

$$\sup_{k\geq \mathrm{N}}\big\|(\mathsf{E}_{\Sigma_k}|f-\mathsf{E}_{\Sigma_k}f|^2)^1/2\big\|_\infty + \|\mathsf{E}_{\Sigma_\mathrm{N}}f\|_\infty + \sup_{k>\mathrm{N}}\|df_k\|_\infty.$$

When N is large enough, the middle term dominates the others and we get spaces which are closer and closer to $L_\infty(\Omega)$. In contrast, when we quotient out the first σ-algebra by using the J-projections, it follows from the interpolation identities above that the starting σ-algebra significantly affects the interpolated space. This justifies, in part, our need to intersect two such spaces in this paper.

BMO spaces for admissible coverings. Let (Ω, Σ, μ) be a σ-finite measure space and consider two atomic σ-algebras $\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}}$ of measurable sets in Σ. Let Π_j be the set of atoms in Σ_j for $j = \mathrm{a}, \mathrm{b}$. When $\mu(\Omega) < \infty$, we shall consider two distinguished atoms $(A_0, B_0) \in \Pi_{\mathrm{a}} \times \Pi_{\mathrm{b}}$. If μ is not finite, take $A_0 = B_0 = \varnothing$. Given $A \in \Pi_{\mathrm{a}}$, set

$$R_A = \{B' \in \Pi_{\mathrm{b}} \mid \mu(A \cap B') > 0\}.$$

Define R_B for $B \in \Pi_{\mathrm{b}}$ similarly. The pair $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ is called an *admissible covering* of (Ω, Σ, μ) when $\Sigma_{\mathrm{a}} \cap \Sigma_{\mathrm{b}} = \{\Omega, \varnothing\}$ and

$$\min\left\{ \sup_{A \in \Pi_{\mathrm{a}} \setminus \{A_0\}} \sum_{B \in R_A} |R_B| \frac{\mu(A \cap B)^2}{\mu(A)\mu(B)}, \sup_{B \in \Pi_{\mathrm{b}} \setminus \{B_0\}} \sum_{A \in R_B} |R_A| \frac{\mu(A \cap B)^2}{\mu(A)\mu(B)} \right\} < 1.$$

One can view the condition above as a weak version of the concentration of measure near the boundary that appeared in [Carbonaro et al. 2009]. In particular, it is not a geometric notion, but only a measure-theoretic one (see Remark 3.3 for more details). Now, consider any pair of filtrations $(\Sigma_{jk})_{k \geq 1}$ with $\Sigma_{j1} = \Sigma_j$ for $j = \mathrm{a}, \mathrm{b}$, and construct the corresponding martingale BMO spaces $\mathrm{BMO}_{\mathrm{a}}$ and $\mathrm{BMO}_{\mathrm{b}}$. As in the previous subsection, we quotient out the Σ_j-measurable functions and set, as we did in the introduction,

$$\mathrm{BMO}_{\Sigma_j}(\Omega) = J_{\Sigma_j}(\mathrm{BMO}_j),$$
$$\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega) = \mathrm{BMO}_{\Sigma_{\mathrm{a}}}(\Omega) \wedge \mathrm{BMO}_{\Sigma_{\mathrm{b}}}(\Omega) = \{f \in L^1_{\mathrm{loc}}(\Omega) \mid \|f\|_{\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)} < \infty\}/\mathbb{C}$$

In the following, we construct admissible coverings for σ-finite measure spaces. The procedure we employ is quite general. In concrete scenarios, other admissible coverings can be constructed enjoying additional properties as required in Theorems B1 and B2; these examples will be given later in this paper.

Remark 1.2. The classical BMO on Euclidean spaces can be decomposed as an intersection of finitely many martingale BMO spaces, the number of which depends on the dimension [Conde 2013; Garnett and Jones 1982; Mei 2003]. In contrast, we just consider "intersections" of two martingale BMOs. Note this makes our spaces larger and still amenable for interpolation, which gives some extra room to obtain endpoint estimates for singular integral operators. The main reason why this is possible is that our approach just relies on measure-theoretic properties and does not rely on the geometry of the underlying space, as will become clear in the sequel.

Lemma 1.3. *Let (Ω, Σ, μ) be a σ-finite measure space. Then*:

(i) *If $\mu(\Omega) = \infty$, it admits an admissible covering.*

(ii) *If $\mu(\Omega) < \infty$ and μ is atomless, it admits an admissible covering.*

Proof. If $\mu(\Omega) = \infty$, pick $A_0 = \tilde{A}_0 = B_0 = \tilde{B}_0 = \varnothing$,

$$A_j = \tilde{A}_j \setminus \tilde{A}_{j-1} \quad \text{and} \quad B_j = \tilde{B}_j \setminus \tilde{B}_{j-1},$$

where $\varnothing \neq \tilde{A}_1 \subsetneq \tilde{B}_1 \subsetneq \tilde{A}_2 \subsetneq \tilde{B}_2 \subsetneq \tilde{A}_3 \subsetneq \cdots$ are Σ-measurable sets chosen so that

$$\min\left\{\frac{\mu(\tilde{B}_j \setminus \tilde{B}_{j-1})}{\mu(\tilde{A}_j)}, \frac{\mu(\tilde{A}_{j+1} \setminus \tilde{A}_j)}{\mu(\tilde{B}_j)}\right\} > \lambda > 4 \quad \text{for all } \ j \geq 1.$$

It is at this point that we have used that $\mu(\Omega) = \infty$. Let Σ_a be the atomic σ-algebra generated by the atoms $(A_j)_{j\geq 1}$. Similarly, define $\Sigma_b = \sigma\langle B_j : j \geq 1\rangle$. It is clear by construction that

$$\Sigma_a \cap \Sigma_b = \{\Omega, \varnothing\}.$$

On the other hand, $|R_B| = 2$ for every atom B in Σ_b. Therefore, it remains to show that

$$\sup_{j\geq 1}\left[\frac{\mu(A_j \cap B_{j-1})^2}{\mu(A_j)\mu(B_{j-1})} + \frac{\mu(A_j \cap B_j)^2}{\mu(A_j)\mu(B_j)}\right] < \frac{1}{2}.$$

Note that the first summand above vanishes for $j = 1$. The rest of terms are smaller than $1/\lambda$, according to our conditions, so that $\lambda > 4$ suffices. When $\mu(\Omega) < \infty$ we may assume that $\mu(\Omega) = 1$, since renormalization does not affect our definition of admissible covering. We use again a "corona-type partition"

$$\varnothing \neq \tilde{A}_0 \subsetneq \tilde{B}_0 \subsetneq \tilde{A}_1 \subsetneq \tilde{B}_1 \subsetneq \tilde{A}_2 \subsetneq \cdots$$

satisfying $\mu(\tilde{A}_0) = 1 - \zeta$, $\mu(\tilde{B}_0 \setminus \tilde{A}_0) = \zeta(1-\zeta)$ and the relations

$$\mu(\tilde{A}_{j+1} \setminus \tilde{B}_j) = \zeta\mu(\tilde{B}_j \setminus \tilde{A}_j) \quad \text{and} \quad \mu(\tilde{B}_{j+1} \setminus \tilde{A}_{j+1}) = \zeta\mu(\tilde{A}_{j+1} \setminus \tilde{B}_j) \quad \text{for } \ j \geq 0.$$

This is where we use the fact that μ has no atoms. Define $A_0 = \tilde{A}_0$, $B_0 = \tilde{B}_0$, $A_j = \tilde{A}_j \setminus \tilde{A}_{j-1}$ and $B_j = \tilde{B}_j \setminus \tilde{B}_{j-1}$ for $j \geq 1$. The σ-algebras Σ_a and Σ_b are the ones generated by $(A_j)_{j\geq 0}$ and $(B_j)_{j\geq 0}$, respectively. In order to show that $\Omega = \bigcup_{j\geq 0} A_j = \bigcup_{j\geq 0} B_j$, let us prove that we have

$$\sum_{j\geq 0} \mu(A_j) = \sum_{j\geq 0} \mu(B_j) = 1.$$

Indeed, if $j \geq 2$ we have

$$\begin{aligned}
\mu(A_j) &= \mu(\tilde{A}_j \setminus \tilde{A}_{j-1}) \\
&= (1+\zeta)\mu(\tilde{B}_{j-1} \setminus \tilde{A}_{j-1}) = \zeta(1+\zeta)\mu(\tilde{A}_{j-1} \setminus \tilde{B}_{j-2}) \\
&= \zeta(1+\zeta)\Big[\mu(\tilde{A}_{j-1} \setminus \tilde{A}_{j-2}) - \frac{1}{\zeta}\mu(\tilde{A}_{j-1} \setminus \tilde{B}_{j-2})\Big] = \zeta^2\mu(A_{j-1}).
\end{aligned}$$

Therefore, since $\mu(A_0) = 1 - \zeta$ and $\mu(A_1) = \zeta(1-\zeta^2)$, we deduce immediately that $\sum_{j\geq 0} \mu(A_j) = 1$. The sum $\sum_j \mu(B_j)$ also equals 1 since the two families are nested. The condition $\Sigma_a \cap \Sigma_b = \varnothing$ follows again by construction. Finally, since $|R_B| = 2$ for all atoms $B = B_j$, it suffices one more time to prove that

$$\sup_{j\geq 1}\left[\frac{\mu(A_j \cap B_{j-1})^2}{\mu(A_j)\mu(B_{j-1})} + \frac{\mu(A_j \cap B_j)^2}{\mu(A_j)\mu(B_j)}\right] < \frac{1}{2}.$$

According to our construction, the left-hand side can be majorized by

$$\frac{\mu(A_j \cap B_{j-1})^2}{\mu(A_j)\mu(B_{j-1})} + \frac{\mu(A_j \cap B_j)^2}{\mu(A_j)\mu(B_j)} \le \frac{\mu(\tilde{B}_{j-1} \setminus \tilde{A}_{j-1})}{\mu(B_{j-1})} + \frac{\mu(\tilde{A}_j \setminus \tilde{B}_{j-1})}{\mu(A_j)}.$$

On the other hand, arguing as before, we may obtain the identities

$$\mu(A_j) = \zeta^{2(j-1)}(\zeta - \zeta^3), \qquad \mu(\tilde{A}_j \setminus \tilde{B}_{j-1}) = \zeta^{2j}(1-\zeta),$$
$$\mu(B_{j-1}) = \zeta^{2(j-1)}(1-\zeta^2), \quad \mu(\tilde{B}_{j-1} \setminus \tilde{A}_{j-1}) = \zeta^{2(j-1)}(\zeta - \zeta^2).$$

This gives a bound $2\zeta/(1+\zeta)$. It suffices for $\zeta < \frac{1}{3}$. The proof is complete. □

Remark 1.4. All fully supported probability measures on $\mathbb{R}^n$ are nondoubling. In fact, this also holds for probability measures supported on unbounded sets. In particular, we hope Lemma 1.3 together with Theorems B1 and B2 might open a door to further insight into Calderón–Zygmund theory for these measures.

John–Nirenberg inequalities, atomic* H_1 *and duality. We now transfer some well-known properties of martingale BMO spaces to our new class of spaces. The analogue of John–Nirenberg inequalities [1961] for martingale BMO spaces can be stated as follows:

$$\sup_{k\ge 1} \sup_{A\in\Sigma_k} \frac{1}{\mu(A)} \mu\big(A \cap \{|f - \mathsf{E}_{\Sigma_{k-1}} f| > \lambda\}\big) \lesssim \exp\left(-\frac{c\lambda}{\|f\|_{\mathrm{BMO}}}\right) \quad \text{for all } \lambda > 0,$$

where the martingale BMO is constructed over the filtration $(\Sigma_k)_{k\ge1}$ and we use the convention $\mathsf{E}_{\Sigma_0} f = 0$. The proof can be found in [Garsia 1973]. An important consequence of this inequality is the p-invariance of the BMO norm. To be more precise, the martingale BMO norm admits the equivalent expressions, for any $0 < p < \infty$,

$$\|f\|_{\mathrm{BMO}} \sim \sup_{k\ge1} \big\| (\mathsf{E}_{\Sigma_k} |f - \mathsf{E}_{\Sigma_{k-1}} f|^p)^{1/p} \big\|_\infty.$$

If we replace f by $J_{\Sigma_1} f = f - \mathsf{E}_{\Sigma_1} f$ in both inequalities, we immediately obtain the corresponding analogues for the BMO spaces which quotient out Σ_1-measurable functions, introduced above. Namely, the only difference is that we should read John–Nirenberg inequalities under the convention that $\mathsf{E}_{\Sigma_0} f = \mathsf{E}_{\Sigma_1} f$, and the BMO norm is given by $\|\cdot\|^\circ_{\mathrm{BMO}}$ instead. If we intersect two of these BMO spaces, we get John–Nirenberg-type inequalities for our spaces $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$ associated to an admissible covering $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ by taking again $\mathsf{E}_{\Sigma_0} f = \mathsf{E}_{\Sigma_1} f$:

$$\|f\|_{\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)} \sim \max_{j=\mathrm{a,b}} \sup_{k\ge1} \big\| (\mathsf{E}_{\Sigma_{jk}} |f - \mathsf{E}_{\Sigma_{j(k-1)}} f|^p)^{1/p} \big\|_\infty,$$
$$\sup_{j=\mathrm{a,b}} \sup_{\substack{k\ge1\\ A\in\Sigma_{jk}}} \frac{1}{\mu(A)} \mu\big(A \cap \{|f - \mathsf{E}_{\Sigma_{j(k-1)}} f| > \lambda\}\big) \lesssim \exp\left(-\frac{c\lambda}{\|f\|_{\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)}}\right).$$

Let us now consider H_1–BMO duality in our context. In the literature we find several equivalent descriptions of martingale H_1 spaces, via Doob's maximal function, martingale square function or conditional square function. Namely, H_1 can be defined as the closure of the space of finite L_1 martingales with respect to any of the norms

$$\Big\|\sup_{k\geq 1}|\mathsf{E}_{\Sigma_k}f|\Big\|_1 \sim \bigg\|\Big(\sum_{k\geq 1}|df_k|^2\Big)^{\frac{1}{2}}\bigg\|_1 \sim \sum_{k\geq 1}\|df_k\|_1 + \bigg\|\Big(\sum_{k\geq 1}\mathsf{E}_{\Sigma_{k-1}}|df_k|^2\Big)^{\frac{1}{2}}\bigg\|_1.$$

We refer to [Davis 1970] for the equivalences above and to [Garsia 1973] for the duality theorem, which claims that $\mathrm{H}_1^* \simeq \mathrm{BMO}$, a martingale analogue of the Fefferman–Stein duality theorem. Let us now consider atomic descriptions of these spaces. The term "atom" unfortunately appears here in several settings — σ-algebras, measures and Hardy spaces — with different meanings, but it will be clear which one is used from the context. Atomic descriptions are not possible for arbitrary H_1 — see [Conde-Alonso and Parcet 2014] for an "atomic block" description both in the commutative and noncommutative settings — but there are such results for h_1 (defined above). A Σ-measurable function $a \in L_2(\Omega)$ is called an atom when there exists $k \geq 1$ and $A \in \Sigma_k$ with

$$\mathrm{supp}(a) \subset A, \quad \mathsf{E}_{\Sigma_k}(a) = 0, \quad \|a\|_2 \leq \mu(A)^{-1/2}.$$

The atomic h_1 is defined as the space of functions of the form $f = \sum_j \lambda_j a_j$ with the a_j atoms. The norm is the infimum of $\sum_j |\lambda_j|$ over all such possible expressions for the function f. This space is isomorphic to h_1; see [Garsia 1973]. In particular, it is also isomorphic to H_1 when the filtration is regular. This will be enough for our purposes, since we will only use H_1–BMO duality for regular filtrations. Now, given two filtrations $(\Sigma_{jk})_{k\geq 1}$ with $\Sigma_{j1} = \Sigma_j$ for $j = \mathrm{a}, \mathrm{b}$, let H_{1j} be the corresponding H_1 spaces. Define

$$\mathrm{H}^1_{\Sigma_{\mathrm{ab}}}(\Omega) = \Big\{ f \in L_1(\Omega) \;\Big|\; \|f\|_{\mathrm{H}_1} = \inf_{\substack{f = f_1 + f_2 \\ \mathsf{E}_{\Sigma_\mathrm{a}} f_1 = \mathsf{E}_{\Sigma_\mathrm{b}} f_2 = 0}} \|f_1\|_{\mathrm{H}_{1\mathrm{a}}} + \|f_2\|_{\mathrm{H}_{1\mathrm{b}}} < \infty \Big\}.$$

Then, all the results above apply. In particular, we have

$$\mathrm{H}^1_{\Sigma_{\mathrm{ab}}}(\Omega)^* \simeq \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega).$$

2. Interpolation: proof of Theorem A

Proof of Theorem A. The argument is a bit lengthy, so we have divided it into several steps. We will assume that μ is a finite measure on Ω — normalized so that $\mu(\Omega) = 1$ — since this case is more technical. The slight modifications needed for the nonfinite case will be explained in the last step of the proof.

Step 1: Intersection of quotients. Let us first show that the interpolation result follows from the first assertion of Theorem A. Namely, given an admissible covering $(\Sigma_\mathrm{a}, \Sigma_\mathrm{b})$ of (Ω, Σ, μ) and filtrations $(\Sigma_{jk})_{k\geq 1}$ with $\Sigma_{j1} = \Sigma_j$ for $j = \mathrm{a}, \mathrm{b}$, let BMO_j be the corresponding martingale BMO spaces. It is clear that

$$\begin{aligned} \|f\|_{L_q^\circ(\Omega)} = \|f\|_{[L_\infty^\circ(\Omega), L_1^\circ(\Omega)]_{1/q}} &\gtrsim \|f\|_{[\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega), L_1^\circ(\Omega)]_{1/q}} \\ &\geq \max_{j=\mathrm{a},\mathrm{b}} \|f - \mathsf{E}_{\Sigma_j} f\|_{[J_{\Sigma_j}(\mathrm{BMO}_j), J_{\Sigma_j}(L_1(\Omega))]_{1/q}} \\ &\simeq \max_{j=\mathrm{a},\mathrm{b}} \|f - \mathsf{E}_{\Sigma_j} f\|_{L^q_{\Sigma_\mathrm{j}}(\Omega)} = \|f\|_{L^q_{\Sigma_{\mathrm{ab}}}(\Omega)}. \end{aligned}$$

For $q \geq 2$, this implies

$$L_q^\circ(\Omega) \subset [\mathrm{BMO}_{\Sigma_{ab}}(\Omega), L_1^\circ(\Omega)]_{1/q} \subset L^q_{\Sigma_{ab}}(\Omega).$$

Thus, the result follows from the isomorphism $L_q^\circ(\Omega) \simeq L^q_{\Sigma_{ab}}(\Omega)$. The interpolation result for $1 < q < 2$ follows from this and the well-known reiteration theorem [Bergh and Löfström 1976].

Step 2: Reduction to strict contractions. The rest of the proof will be devoted to justify the first assertion of Theorem A. We claim that such an isomorphism holds whenever we can find a constant $0 < c_p(\Sigma_{ab}) < 1$ such that, for every mean-zero function $f \in L_p(\Omega)$,

$$\min\{\|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p,\ \|\mathsf{E}_{\Sigma_b}\mathsf{E}_{\Sigma_a} f\|_p\} \leq c_p(\Sigma_{ab})\|f\|_p. \tag{2-1}$$

Indeed, if $\mathsf{E}\phi = \int_\Omega \phi\, d\mu$, we first observe that

$$\|\phi\|_{L_p^\circ(\Omega)} \sim \|\phi - \mathsf{E}\phi\|_p \sim \inf_{\mathrm{k}\in\mathbb{C}} \|\phi - \mathrm{k}\|_p,$$

$$\|\phi\|_{J_{\Sigma_j}(L_p(\Omega))} \sim \|\phi - \mathsf{E}_{\Sigma_j}\phi\|_p \sim \inf_{\varphi\ \Sigma_j\text{-measurable}} \|\phi - \varphi\|_p.$$

Therefore, our goal in what follows is to show that

$$\|\phi - \mathsf{E}\phi\|_p \sim \|\phi - \mathsf{E}_{\Sigma_a}\phi\|_p + \|\phi - \mathsf{E}_{\Sigma_b}\phi\|_p \quad \text{for every } \phi \in L_p(\Omega).$$

The lower estimate is trivial. For the upper estimate, we shall use (2-1). Assume that the minimum above is attained at the first term (say) and let $f = \phi - \mathsf{E}\phi$ be a mean-zero function. We then find

$$\|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p \leq c_p(\Sigma_{ab})\|f\|_p \leq c_p(\Sigma_{ab})\big[\|f - \mathsf{E}_{\Sigma_a} f\|_p + \|\mathsf{E}_{\Sigma_a}(f - \mathsf{E}_{\Sigma_b} f)\|_p + \|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p\big],$$

which implies

$$\|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p \leq \frac{c_p(\Sigma_{ab})}{1 - c_p(\Sigma_{ab})}[\|\phi - \mathsf{E}_{\Sigma_a}\phi\|_p + \|\phi - \mathsf{E}_{\Sigma_b}\phi\|_p].$$

This inequality is all we need, since the upper estimate follows from it:

$$\begin{aligned}\|\phi - \mathsf{E}\phi\|_p \leq \|\mathsf{E}_{\Sigma_a}\phi - \mathsf{E}\phi\|_p + \|\phi - \mathsf{E}_{\Sigma_a}\phi\|_p &\leq \|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p + \|\mathsf{E}_{\Sigma_a}(\phi - \mathsf{E}_{\Sigma_b}\phi)\|_p + \|\phi - \mathsf{E}_{\Sigma_a}\phi\|_p \\ &\leq \frac{1}{1 - c_p(\Sigma_{ab})}[\|\phi - \mathsf{E}_{\Sigma_a}\phi\|_p + \|\phi - \mathsf{E}_{\Sigma_b}\phi\|_p].\end{aligned}$$

Step 3: The case $p = 2$. Recall that we are assuming for the moment that $\mu(\Omega) = 1$, and in that case we may consider two distinguished atoms $(A_0, B_0) \in \Sigma_a \times \Sigma_b$. In accordance with the previous point, it suffices to show that

$$\min\{\|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_2,\ \|\mathsf{E}_{\Sigma_b}\mathsf{E}_{\Sigma_a} f\|_2\} \leq c_2(\Sigma_{ab})\|f\|_2$$

for some $0 < c_2(\Sigma_{ab}) < 1$ and every mean-zero $f \in L_2(\Omega)$. We claim that this estimate follows if the same inequality holds for Σ_j-measurable functions which vanish on the corresponding distinguished atom. More precisely, it suffices to prove that one of the following conditions holds:

- $\|\mathsf{E}_{\Sigma_a}\phi_b\|_2 \leq c_2(\Sigma_{ab})\|\phi_b\|_2$ for ϕ_b Σ_b-measurable with $\phi_b(B_0) = 0$;
- $\|\mathsf{E}_{\Sigma_b}\phi_a\|_2 \leq c_2(\Sigma_{ab})\|\phi_a\|_2$ for ϕ_a Σ_a-measurable with $\phi_a(A_0) = 0$.

Indeed, assume the first condition holds and let $\phi_{\mathrm{b}} \in L_2(\Omega, \Sigma_{\mathrm{b}}, \mu)$ be mean-zero. Then

$$\|\phi_{\mathrm{b}}\|_2^2 = \|\phi_{\mathrm{b}} - \phi_{\mathrm{b}}(B_0)\|_2^2 - |\phi_{\mathrm{b}}(B_0)|^2,$$
$$\|\mathsf{E}_{\Sigma_{\mathrm{a}}}\phi_{\mathrm{b}}\|_2^2 = \|\mathsf{E}_{\Sigma_{\mathrm{a}}}(\phi_{\mathrm{b}} - \phi_{\mathrm{b}}(B_0))\|_2^2 - |\phi_{\mathrm{b}}(B_0)|^2.$$

Subtracting and using the first condition, we get

$$\|\phi_{\mathrm{b}}\|_2^2 - \|\mathsf{E}_{\Sigma_{\mathrm{a}}}\phi_{\mathrm{b}}\|_2^2 \geq (1 - c_2(\Sigma_{\mathrm{ab}}))\|\phi_{\mathrm{b}} - \phi_{\mathrm{b}}(B_0)\|_2^2 \geq (1 - c_2(\Sigma_{\mathrm{ab}}))\|\phi_{\mathrm{b}}\|_2^2.$$

Here, $\phi_{\mathrm{b}}(B_0)$ denotes the constant value of ϕ_{b} on B_0. Rearranging, we get $\|\mathsf{E}_{\Sigma_{\mathrm{a}}}\phi_{\mathrm{b}}\|_2 \leq c_2(\Sigma_{\mathrm{ab}})\|\phi_{\mathrm{b}}\|_2$. Therefore, given any mean-zero $f \in L_2(\Omega)$, we may define $\phi_{\mathrm{b}} = \mathsf{E}_{\Sigma_{\mathrm{b}}} f$ and deduce that $\|\mathsf{E}_{\Sigma_{\mathrm{a}}}\mathsf{E}_{\Sigma_{\mathrm{b}}} f\|_2 \leq c_2(\Sigma_{\mathrm{ab}})\|f\|_2$, as desired. Alternatively, if we use the second condition above, the roles of Σ_{a} and Σ_{b} are switched and we obtain the other sufficient inequality which is implicit in the minimum above. Thus we have reduced the proof to justify one of the two conditions above. It is at this point where our definition of admissible pair comes into play. Namely, we know that

$$\min\left\{ \sup_{A \in \Pi_{\mathrm{a}} \setminus \{A_0\}} \sum_{B \in R_A} |R_B| \frac{\mu(A \cap B)^2}{\mu(A)\mu(B)}, \sup_{B \in \Pi_{\mathrm{b}} \setminus \{B_0\}} \sum_{A \in R_B} |R_A| \frac{\mu(A \cap B)^2}{\mu(A)\mu(B)} \right\} = c(\Sigma_{\mathrm{ab}})$$

for some $0 < c(\Sigma_{\mathrm{ab}}) < 1$. Let us assume (say) that the minimum above is attained by the first term and let ϕ_{a} be a Σ_{a}-measurable function in $L_2(\Omega)$ that vanishes on A_0. Then, if we write $\phi_{\mathrm{a}} = \sum_{A \neq A_0} \alpha_A \chi_A$, we have the estimate

$$\begin{aligned}
\|\mathsf{E}_{\Sigma_{\mathrm{b}}}\phi_{\mathrm{a}}\|_2^2 &= \sum_{A, A' \neq A_0} \overline{\alpha}_A \alpha_{A'} \sum_{B \in R_A \cap R_{A'}} \frac{\mu(A \cap B)}{\mu(B)^{1/2}} \frac{\mu(A' \cap B)}{\mu(B)^{1/2}} \\
&\leq \sum_{A, A' \neq A_0} \frac{1}{2} \sum_{B \in R_A \cap R_{A'}} \left(|\alpha_A|^2 \frac{\mu(A \cap B)^2}{\mu(B)} + |\alpha_{A'}|^2 \frac{\mu(A' \cap B)^2}{\mu(B)} \right) \\
&= \sum_{A \neq A_0} |\alpha_A|^2 \mu(A) \sum_{\substack{A' \neq A_0 \\ R_A \cap R_{A'} \neq \varnothing}} \sum_{B \in R_A \cap R_{A'}} \frac{\mu(A \cap B)^2}{\mu(A)\mu(B)} \\
&= \sum_{A \neq A_0} |\alpha_A|^2 \mu(A) \sum_{B \in R_A} |R_B| \frac{\mu(A \cap B)^2}{\mu(A)\mu(B)} \leq c(\Sigma_{\mathrm{ab}}) \sum_{A \neq A_0} |\alpha_A|^2 \mu(A).
\end{aligned}$$

The right-hand side equals $c(\Sigma_{\mathrm{ab}})\|\phi_{\mathrm{a}}\|_2^2$, so we obtain the second condition. The first one follows when the minimum in our definition of admissible covering is attained by the second term. This proves that the first assertion of Theorem A holds for finite measures and $p = 2$. The case $p > 2$ requires some preliminaries.

Step 4: A mass absorption principle. Let us consider a particular ordering of the atoms in Σ_{a} and Σ_{b}. According to our assumption $\Sigma_{\mathrm{a}} \cap \Sigma_{\mathrm{b}} = \{\Omega, \varnothing\}$, we may order Π_{a} so that $\Pi_{\mathrm{a}} = \{A_1, A_2, \dots\}$ and, for each $m \geq 0$, there exists $B \in \Pi_{\mathrm{b}}$ such that $\mu(A_{m+1} \cap B)$ and $\mu\big(\bigcup_{s \leq m} A_s \cap B\big)$ are both strictly positive.

Similarly, we may order Π_b satisfying the symmetric condition. Define the atomic σ-algebras

$$\Sigma_{\mathrm{a}}(m)=\sigma\Big\langle \bigcup_{s=0}^{m} A_s,\, \{A_s\}_{s\geq m+1}\Big\rangle,$$
$$\Sigma_{\mathrm{b}}(m)=\sigma\Big\langle \bigcup_{s=0}^{m} B_s,\, \{B_s\}_{s\geq m+1}\Big\rangle.$$

In this step we will prove that

$$\|f\|_{L^p_{\Sigma_{\mathrm{ab}}}(\Omega)} \simeq \|f-\mathsf{E}_{\Sigma_{\mathrm{a}}(m)}f\|_{L_p(\Omega)}+\|f-\mathsf{E}_{\Sigma_{\mathrm{b}}(m)}f\|_{L_p(\Omega)} \tag{2-2}$$

for any $m\geq 1$ and $2<p<\infty$. The constants may depend on m, p and the covering $(\Sigma_{\mathrm{a}},\Sigma_{\mathrm{b}})$. Indeed, since the result is trivial for $m=0$, we will proceed by induction and assume that the result holds for $m-1$. Moreover, the upper estimate is straightforward and by symmetry it suffices to show that

$$\|f-\mathsf{E}_{\Sigma_{\mathrm{a}}(m)}f\|_p \lesssim \|f-\mathsf{E}_{\Sigma_{\mathrm{a}}(m-1)}f\|_p+\|f-\mathsf{E}_{\Sigma_{\mathrm{b}}}f\|_p.$$

Taking $A_0(m)=\bigcup_{s\leq m}A_s$, let $f=f\chi_{A_0(m)}+f\chi_{\Omega\setminus A_0(m)}=f_1+f_2$. Since it is clear that $\mathsf{E}_{\Sigma_{\mathrm{a}}(m)}f_2=\mathsf{E}_{\Sigma_{\mathrm{a}}(m-1)}f_2$, we may concentrate only on f_1. The left-hand side for f_1 can be written as

$$\|f_1-\mathsf{E}_{\Sigma_{\mathrm{a}}(m)}f_1\|_p=\|\chi_{A_0(m)}(f-\mathsf{E}_{\Sigma_{\mathrm{a}}(m)}f)\|_p=\|f\|_{L_p^\circ(A_0(m))}\sim \sup_{\substack{\|g\|_{L_{p'}(A_0(m))}\leq 1\\ g \text{ mean-zero}}}\Big|\int_{A_0(m)} fg\,d\mu\Big|.$$

Approximating the right-hand side up to $\varepsilon>0$ by some mean-zero g_0 in the unit ball of $L_{p'}(A_0(m))$, let B be an atom in Σ_{b} satisfying that $\mu(A_0(m-1)\cap B)$ and $\mu(A_m\cap B)$ are strictly positive. Recall that this can be done by the specific enumeration of atoms we picked. Then, define

$$\begin{aligned}
g_1&=\chi_{A_m}g_0-\frac{\chi_{A_m\cap B}}{\mu(A_m\cap B)}\int_{A_m} g_0\,d\mu,\\
g_2&=\chi_{A_0(m-1)}g_0-\frac{\chi_{A_0(m-1)\cap B}}{\mu(A_0(m-1)\cap B)}\int_{A_0(m-1)} g_0\,d\mu,\\
g_3&=\frac{\chi_{A_0(m-1)\cap B}}{\mu(A_0(m-1)\cap B)}\int_{A_0(m-1)} g_0\,d\mu+\frac{\chi_{A_m\cap B}}{\mu(A_m\cap B)}\int_{A_m} g_0\,d\mu.
\end{aligned}$$

Obviously, $g_0=g_1+g_2+g_3$ and each g_j is mean-zero. Moreover, we have

$$\begin{aligned}
\|g_1\|_{L_{p'}(A_0(m))}&\leq \|\chi_{A_m}g_0\|_{L_{p'}(A_0(m))}+\Big\|\frac{\chi_{A_m\cap B}}{\mu(A_m\cap B)}\int_{A_m} g_0\,d\mu\Big\|_{L_{p'}(A_0(m))}\\
&\leq \Big(1+\frac{\mu(A_m)^{1/p}}{\mu(A_m\cap B)^{1/p}}\Big)\|g_0\|_{L_{p'}(A_0(m))}\lesssim \|g_0\|_{L_{p'}(A_0(m))}\leq 1.
\end{aligned}$$

Similar computations apply to g_2 and g_3. In summary, we obtain the estimate below, where we write f_Q to denote the average of f over a given measurable set Q:

$$\begin{aligned}\|f_1 - \mathsf{E}_{\Sigma_a(m)} f_1\|_p &\sim \Big|\int_{A_0(m)} f g_0 \, d\mu\Big| \\ &\leq \Big|\int_{A_m} (f - f_{A_m}) g_1 \, d\mu\Big| + \Big|\int_{A_0(m-1)} (f - f_{A_0(m-1)}) g_2 \, d\mu\Big| + \Big|\int_B (f - f_B) g_3 \, d\mu\Big| \\ &\lesssim \|\chi_{A_m}(f - f_{A_m})\|_p + \|\chi_{A_0(m-1)}(f - f_{A_0(m-1)})\|_p + \|\chi_B(f - f_B)\|_p \\ &\lesssim \|f - \mathsf{E}_{\Sigma_a(m-1)} f\|_p + \|f - \mathsf{E}_{\Sigma_b} f\|_p.\end{aligned}$$

This completes the proof of the norm equivalence (2-2).

Step 5: The case $p > 2$. We now complete the proof of Theorem A for probability measures. According to (2-2), it suffices to show that there exists $0 < c_p(\Sigma_{ab}) < 1$ and $m = m(p) \geq 1$ such that, for any mean-zero function $f \in L_p(\Omega)$,

$$\min\{\|\mathsf{E}_{\Sigma_a(m)} \mathsf{E}_{\Sigma_b(m)} f\|_p, \ \|\mathsf{E}_{\Sigma_b(m)} \mathsf{E}_{\Sigma_a(m)} f\|_p\} \leq c_p(\Sigma_{ab}) \|f\|_p.$$

Pick $m = m(p)$ as the smallest possible value of m satisfying

$$\min\big\{\mu(A_0(m)), \ \mu(B_0(m))\big\} > \max\bigg\{\Big(\frac{2 \cdot 4^{-p}}{1 - 2 \cdot 4^{-p}}\Big)^{\frac{1}{p-1}}, (1 - 4^{-p})^{1/p}\bigg\}$$

and $\varepsilon = \varepsilon(p) > 0$ small enough so that

$$(1 - 2 \cdot 4^{-p})^{1/2} \leq (1 - 4^{-p})^{1/(2p)} (1 - \varepsilon^3)^{1/2} - \varepsilon^{3/2}.$$

Since $L_p(\Omega) \subset L_2(\Omega)$, we know from Step 3 that f always satisfies the above inequality for $p = 2$. Assume that the minimum for $p = 2$ is attained (say) at the first term, so that $\|\mathsf{E}_{\Sigma_a(m)} \mathsf{E}_{\Sigma_b(m)} f\|_2 \leq c_2(\Sigma_{ab}) \|f\|_2$. Recall that $\mathsf{E} f = \int_\Omega f \, d\mu$. When

$$\mathsf{E}(|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2 < (1 - \varepsilon^3) \|\mathsf{E}_{\Sigma_b(m)} f\|_p^p,$$

we proceed as follows:

$$\begin{aligned}\|\mathsf{E}_{\Sigma_a(m)} \mathsf{E}_{\Sigma_b(m)} f\|_p^p &\leq \|\mathsf{E}_{\Sigma_a(m)} |\mathsf{E}_{\Sigma_b(m)} f|^{p/2}\|_2^2 - (\mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2 + (\mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2 \\ &= \|\mathsf{E}_{\Sigma_a(m)} (|\mathsf{E}_{\Sigma_b(m)} f|^{p/2} - \mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})\|_2^2 + (\mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2 \\ &\leq c_2^2(\Sigma_{ab}) \| |\mathsf{E}_{\Sigma_b(m)} f|^{p/2} - \mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2}\|_2^2 + (\mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2 \\ &\leq \big[c_2^2(\Sigma_{ab}) + (1 - c_2^2(\Sigma_{ab}))(1 - \varepsilon^3)\big] \|f\|_p^p = c_p^p(\Sigma_{ab}) \|f\|_p^p.\end{aligned}$$

If $\mathsf{E}(|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2 \geq (1 - \varepsilon^3) \|\mathsf{E}_{\Sigma_b(m)} f\|_p^p$, then one can easily show that

$$\big\| |\mathsf{E}_{\Sigma_b(m)} f|^{p/2} - \mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2} \big\|_2^2 \leq \varepsilon^3 \|\mathsf{E}_{\Sigma_b(m)} f\|_p^p.$$

Now, decomposing $\mathsf{E}_{\Sigma_b(m)} f = \mathsf{E}_{\Sigma_b(m)} f(B_0(m))\chi_{B_0(m)} + \mathsf{E}_{\Sigma_b(m)} f \chi_{\Omega\setminus B_0(m)}$, we get

$$\begin{aligned}\sqrt{\mu(B_0(m))|\mathsf{E}_{\Sigma_b(m)} f(B_0(m))|^p} &= \||\mathsf{E}_{\Sigma_b(m)} f|^{p/2}\chi_{B_0(m)}\|_2 \\ &\geq \|\mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2}\chi_{B_0(m)}\|_2 - \||\mathsf{E}_{\Sigma_b(m)} f|^{p/2} - \mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2}\|_2 \\ &\geq \mu(B_0(m))^{1/2}\mathsf{E}|\mathsf{E}_{\Sigma_b(m)} f|^{p/2} - \varepsilon^{3/2}\|\mathsf{E}_{\Sigma_b(m)} f\|_p^{p/2} \\ &\geq [(1-4^{-p})^{1/(2p)}(1-\varepsilon^3)^{1/2} - \varepsilon^{3/2}]\|\mathsf{E}_{\Sigma_b(m)} f\|_p^{p/2} \\ &\geq (1-2\cdot 4^{-p})^{1/2}\|\mathsf{E}_{\Sigma_b(m)} f\|_p^{p/2}.\end{aligned}$$

This also implies

$$\|\mathsf{E}_{\Sigma_b(m)} f\chi_{\Omega\setminus B_0(m)}\|_p^p \leq 2\cdot 4^{-p}\|\mathsf{E}_{\Sigma_b(m)} f\|_p^p.$$

On the other hand, since f is mean-zero we have

$$\mathsf{E}_{\Sigma_b(m)} f(B_0(m))\mu(B_0(m)) + \mathsf{E}(\mathsf{E}_{\Sigma_b(m)} f\chi_{\Omega\setminus B_0(m)}) = 0.$$

Rearranging and raising to the power p then gives

$$\mu(B_0(m))^p|\mathsf{E}_{\Sigma_b(m)} f(B_0(m))|^p \leq \|\mathsf{E}_{\Sigma_b(m)} f\chi_{\Omega\setminus B_0(m)}\|_p^p \leq 2\cdot 4^{-p}\|\mathsf{E}_{\Sigma_b(m)} f\|_p^p.$$

Finally, combining our two estimates so far for $\mu(B_0(m))$, we obtain

$$\mu(B_0(m)) \leq \left(\frac{2\cdot 4^{-p}}{1-2\cdot 4^{-p}}\right)^{\frac{1}{p-1}},$$

which contradicts our choice of $m = m(p)$. This shows that $\mathsf{E}(|\mathsf{E}_{\Sigma_b(m)} f|^{p/2})^2$ cannot be larger than $(1-\varepsilon^3)\|\mathsf{E}_{\Sigma_b(m)} f\|_p^p$ and completes the proof in the case the minimum for $p = 2$ is attained at the first term. When the minimum is attained at the second term, a symmetric argument applies.

Step 6: The nonfinite case. When $\mu(\Omega) = \infty$ the proof of Theorem A is a bit simpler. In the first place, note that $L_p^\circ(\Omega) = L_p(\Omega)$ in this case. In particular, the goal is to show that

$$\begin{aligned}L_q(\Omega) &\simeq [\mathrm{BMO}_{\Sigma_{ab}}(\Omega), L_1(\Omega)]_{1/q}, \\ L_p(\Omega) &\simeq L_p(\Omega,\Sigma,\mu)/\Sigma_a \wedge L_p(\Omega,\Sigma,\mu)/\Sigma_b.\end{aligned}$$

Since $L_\infty(\Omega) \subset \mathrm{BMO}_{\Sigma_{ab}}(\Omega)$, our argument in Step 1 can be easily adapted and interpolation follows from the second isomorphism above. To prove it, we follow essentially the same argument as for finite measures. Indeed, arguing as in Step 2, we see that it suffices to show that

$$\min\{\|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p,\ \|\mathsf{E}_{\Sigma_b}\mathsf{E}_{\Sigma_a} f\|_p\} \leq c_p(\Sigma_{ab})\|f\|_p$$

for some constant $0 < c_p(\Sigma_{ab}) < 1$ and every function $f \in L_p(\Omega)$. The only difference is that here it must hold for every f, not just mean-zero elements as in the finite case. The case $p = 2$ is proved following Step 3. The fact that we do not assume f to be mean-zero — or ultimately to vanish at A_0 or B_0 — is compensated by our definition of admissible coverings, which does not consider distinguished atoms for infinite measures. Finally, once we know the case $p = 2$ holds — for arbitrary functions, not only mean-zero ones — we conclude that

$$\|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b} f\|_p^p \le \|\mathsf{E}_{\Sigma_a}\mathsf{E}_{\Sigma_b}|f|^{p/2}\|_2^2 \le c_2^2(\Sigma_{ab})\||f|^{p/2}\|_2^2 \le c_p^p(\Sigma_{ab})\|f\|_p^p$$

or a similar estimate for $\mathsf{E}_{\Sigma_b}\mathsf{E}_{\Sigma_a} f$. The proof of Theorem A is now complete. □

3. Calderón–Zygmund operators, I

Let (Ω, Σ, μ) be a measure space and consider a metric d on Ω. Assume that μ is σ-finite with respect to the metric topology. In this section we will be interested in Calderón–Zygmund operators on the metric measure space (Ω, μ, d), as defined in the introduction. More precisely, we prove Theorem B1 below and, after that, we shall illustrate this result with a few constructions of admissible coverings.

Proof of Theorem B1. Our definition of CZO includes a symmetric Hörmander kernel condition. This implies that the class of Calderón–Zygmund operators is closed under taking adjoints. In particular, the L_p-boundedness for $1 < p < 2$ can be deduced by duality from the case $p > 2$. On the other hand, according to Theorem A, the latter follows by interpolation if we can prove that any CZO extends to a bounded map $L_\infty(\Omega) \to \mathrm{BMO}_{\Sigma_{ab}}(\Omega)$. Indeed, since T is L_2-bounded, Theorem A yields that $T : L_p(\Omega) \to L_p^\circ(\Omega)$. This is enough when the measure μ is infinite, since in that case $L_p(\Omega) = L_p^\circ(\Omega)$. When μ is finite we use L_2-boundedness once again together with Hölder's inequality to deduce that

$$\|Tf\|_p \le \|Tf - \mathsf{E}Tf\|_p + \mu(\Omega)^{1/p}|\mathsf{E}Tf| \lesssim \|f\|_p + \mu(\Omega)^{1/p-1/2}\|f\|_2 \lesssim \|f\|_p.$$

This completes the proof of our claim. Let us then prove the $L_\infty \to \mathrm{BMO}$ estimate. Consider an auxiliary BMO space which arises by averaging over the family of doubling balls in (Ω, Σ, μ),

$$\|f\|_{\mathrm{DBMO}} = \sup_{\substack{\mathrm{B}\, d\text{-ball} \\ \text{doubling}}} \left(\frac{1}{\mu(\mathrm{B})}\int_{\mathrm{B}} \left| f(w) - \frac{1}{\mu(\mathrm{B})}\int_{\mathrm{B}} f\, d\mu\right|^2 d\mu(w)\right)^{\frac12}.$$

Following the standard argument, it is easily checked that

$$T : L_\infty(\Omega) \to \mathrm{DBMO}.$$

Indeed, in the first place we may observe as usual that we have the equivalence

$$\|f\|_{\mathrm{DBMO}} \sim \sup_{\substack{\mathrm{B}\, d\text{-ball} \\ \text{doubling}}} \inf_{\mathrm{k_B}\in\mathbb{C}} \left(\frac{1}{\mu(\mathrm{B})}\int_{\mathrm{B}} |f(w) - \mathrm{k_B}|^2\, d\mu(w)\right)^{\frac12}.$$

Second, we decompose $f = f\chi_{\alpha\mathrm{B}} + f\chi_{\Omega\setminus\alpha\mathrm{B}} = \phi_{1\mathrm{B}} + \phi_{2\mathrm{B}}$ and pick the constant $\mathrm{k_B}$ to be the average of $T\phi_{2\mathrm{B}}$ over B. Then, we may estimate the norm of Tf in DBMO by using the L_2-boundedness of T for

$T\phi_{1\mathrm{B}}$ and the Hörmander kernel condition for $T\phi_{2\mathrm{B}}$. More precisely, we get

$$\|Tf\|_{\mathrm{DBMO}} \le \sup_{\substack{\mathrm{B}\, d\text{-ball} \\ \text{doubling}}} \Big(\frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} \Big| T(f\chi_{\alpha \mathrm{B}})(w) \Big|^2 \, d\mu(w) \Big)^{\frac12} + \Big(\frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} \Big| T(f\chi_{\Omega\setminus\alpha\mathrm{B}})(w) - \frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} T(f\chi_{\Omega\setminus\alpha\mathrm{B}})\, d\mu \Big|^2 \, d\mu(w) \Big)^{\frac12}.$$

Since we just use (α,β)-doubling balls, the first term is dominated by

$$\Big(\frac{\mu(\alpha\mathrm{B})}{\mu(\mathrm{B})} \Big)^{\frac12} \|T\|_{2\to 2} \|f\|_\infty \lesssim \|f\|_\infty.$$

On the other hand, using the kernel representation of T we may write

$$T(f\chi_{\Omega\setminus\alpha\mathrm{B}})(w) - \frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} T(f\chi_{\Omega\setminus\alpha\mathrm{B}})\, d\mu = \frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} \int_{\Omega\setminus\alpha\mathrm{B}} (k(w,\zeta) - k(\xi,\zeta)) f(\zeta)\, d\mu(\zeta)\, d\mu(\xi)$$

for $w \in \mathrm{B}$. In particular, the last term above can be majorized by $\|f\|_\infty$ using the Hörmander condition for k. This proves the $L_\infty(\Omega) \to \mathrm{DBMO}$ boundedness of our CZO. Therefore, it suffices to show that $\mathrm{DBMO} \subset \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$. This follows from the chain of inclusions

$$\mathrm{DBMO} \subset \mathrm{bmo}_{\Sigma_{\mathrm{a}}} \wedge \mathrm{bmo}_{\Sigma_{\mathrm{b}}} \subset \mathrm{BMO}_{\Sigma_{\mathrm{a}}} \wedge \mathrm{BMO}_{\Sigma_{\mathrm{b}}} = \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega).$$

Let us recall in passing the terminology we are using, namely

$$\mathrm{bmo}_{\Sigma_j} = J_{\Sigma_j}(\mathrm{bmo}_j) \quad \text{and} \quad \mathrm{BMO}_{\Sigma_j} = J_{\Sigma_j}(\mathrm{BMO}_j)$$

for $j = \mathrm{a}, \mathrm{b}$. Here, bmo_j and BMO_j are the martingale bmo and BMO spaces constructed over the filtrations $(\Sigma_{jk})_{k\ge 1}$ described in the statement of Theorem B1. If $\mathbf{\Pi}_j$ denotes the atoms in such a filtration, the norm in bmo_{Σ_j} is given by

$$\|f\|_{\mathrm{bmo}_{\Sigma_j}} = \sup_{k\ge 1} \big\| \mathsf{E}_{\Sigma_{jk}} |f - \mathsf{E}_{\Sigma_{jk}} f|^2 \big\|_\infty^{1/2} = \sup_{A\in\mathbf{\Pi}_j} \Big(\frac{1}{\mu(A)} \int_A \Big| f(w) - \frac{1}{\mu(A)} \int_A f\, d\mu \Big|^2 \, d\mu(w) \Big)^{\frac12}.$$

Now, since we assume that all atoms in $\mathbf{\Pi} = \mathbf{\Pi}_{\mathrm{a}} \cup \mathbf{\Pi}_{\mathrm{b}}$ are doubling, the seminorm above is majorized (up to absolute constants) by the seminorm in DBMO. As this holds for both $j = \mathrm{a}, \mathrm{b}$, we have proved the first inclusion. Now, for the second inclusion, we recall the seminorm in BMO_{Σ_j},

$$\|f\|_{\mathrm{BMO}_{\Sigma_j}} = \sup_{k\ge 1} \big\| \mathsf{E}_{\Sigma_{jk}} |f - \mathsf{E}_{\Sigma_{jk-1}} f|^2 \big\|_\infty^{1/2},$$

where $\mathsf{E}_{\Sigma_{j0}} f = \mathsf{E}_{\Sigma_{j1}} f$ since we quotient out Σ_{j1}-measurable functions. Note also that we are requiring the filtrations $(\Sigma_{jk})_{k\ge 1}$ to be regular. In other words, there exist absolute constants $c_j > 0$ such that $\mathsf{E}_{\Sigma_{jk}}|f| \le c_j \mathsf{E}_{\Sigma_{jk-1}}|f|$ for $j = \mathrm{a}, \mathrm{b}$ and $k \ge 1$. This yields the inequality

$$\|f\|_{\mathrm{BMO}_{\Sigma_j}} \le c_j \|f\|_{\mathrm{bmo}_{\Sigma_j}}.$$

Thus, $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega) \simeq \mathrm{bmo}_{\Sigma_{\mathrm{a}}} \wedge \mathrm{bmo}_{\Sigma_{\mathrm{b}}}$ for regular filtrations, and we are done. □

Remark 3.1. Under the same assumptions, every CZO extends to a bounded map

$$\mathrm{H}^1_{\Sigma_{\mathrm{ab}}}(\Omega) \to L_1(\Omega).$$

Indeed, this follows at once by duality and Theorem B1. Alternatively, since we need to work with regular filtrations, we may use the atomic description of $\mathrm{H}^1_{\Sigma_{\mathrm{ab}}}(\Omega)$ given in Section 1, from which an easy argument arises; details are left to the reader.

In the following subsections we shall illustrate Theorem B1 with a few examples.

Doubling case. Admissible coverings fulfilling the assumptions in Theorem B1 can always be constructed on every doubling space, so that Calderón–Zygmund extrapolation for homogeneous spaces appears as a particular application of our approach. For clarity of the exposition, we shall just indicate how to construct such admissible coverings in $\mathbb{R}^2$ with the Lebesgue measure m and the Euclidean metric, although a similar construction works in the general case. Let us pick $Q_0 = \left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right]$, the unit cube, and set $Q_s = 3^s Q_0$ for $s \geq 1$. Consider the σ-algebras

$$\Sigma_{\mathrm{a}} = \sigma\langle A_s \mid s \geq 1\rangle \quad \text{and} \quad \Sigma_{\mathrm{b}} = \sigma\langle B_s \mid s \geq 1\rangle,$$

where $(A_1, B_1) = (Q_0, Q_1)$ and $(A_s, B_s) = (Q_{2s-2} \setminus Q_{2s-4}, Q_{2s-1} \setminus Q_{2s-3})$ for $s \geq 2$. Then it follows from the proof of Lemma 1.3 that $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ is an admissible covering of the Euclidean space $(\mathbb{R}^2, m)$.

Next, we define the filtrations $(\Sigma_{jk})_{k\geq 1}$ with $\Sigma_{j1} = \Sigma_j$ for $j = \mathrm{a}, \mathrm{b}$. Except for A_1 and B_1 — which are ordinary cubes — the atoms A_s and B_s $(s \geq 2)$ are punctured cubes in which we remove a concentric cube with side-length $\frac{1}{9}$ times the side-length of the larger one. To define Σ_{j2} for $j = \mathrm{a}, \mathrm{b}$, we break each A_s, B_s into a disjoint union of 80 equal cubes of side-length $\frac{1}{9}$ times the side-length of the original punctured cube — i.e., all except for the one in the center — unless $s = 1$, in which case we also pick the center and get 81 subcubes (see Figure 1). The next generations are simpler. Indeed, since all our atoms in Σ_{j2} are already cubes, we perform dyadic partitions in each of them to provide the next generations of our filtration. This procedure completely defines two filtrations respectively based on Σ_{a} and Σ_{b}. It remains to check that these filtrations are regular and the atoms are doubling. The regularity constant is dominated uniformly by 81 when $(k-1, k) = (1, 2)$ and by 4 otherwise. On the other hand, our atoms for $k = 1$ are punctured cubes which are comparable to the corresponding unpunctured ones, which in turn are doubling with constant 4. This proves that all conditions in Theorem B1 are satisfied. In the general case, we just need to use Christ dyadic cubes [1990] and adapt our choice according to the finiteness or nonfiniteness of μ as we did in Lemma 1.3.

Polynomial growth. Assume that we have (Ω, μ, d) of k-th degree polynomial growth with $\mu(\Omega) = \infty$. The associated RBMO norm can be defined as follows:

$$\|f\|_{\mathrm{RBMO}} = \max\left\{ \|f\|_{\mathrm{DBMO}}, \sup_{\substack{\mathrm{B}\subset\mathrm{B}' \\ \mathrm{B},\mathrm{B}'\, d\text{-balls} \\ \text{doubling}}} \left| \frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} f\, d\mu - \frac{1}{\mu(\mathrm{B}')} \int_{\mathrm{B}'} f\, d\mu \right| \Big/ K_{\mathrm{B},\mathrm{B}'} \right\},$$

with $1 \leq K_{\mathrm{B},\mathrm{B}'} = 1 + \sum_{2^j\mathrm{B}\subset\mathrm{B}'} \mu(2^j\mathrm{B})/r(2^j\mathrm{B})^k$. For such measures, we may easily construct an admissible covering of (Ω, μ) composed of doubling atoms. Indeed, the construction above can easily be modified

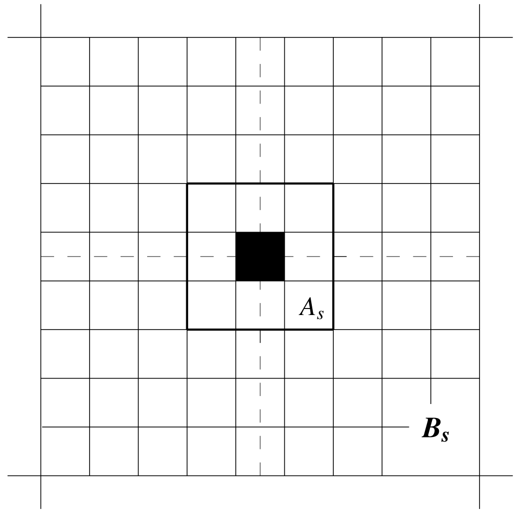

Figure 1. The admissible covering and the second generation of one of the filtrations.

using the existence of arbitrarily large doubling cubes centered at almost every point in the support of μ; see [Tolsa 2001] for details. The main difficulty relies in the construction of filtrations $(\Sigma_{jk})_{k\geq 1}$ satisfying the assumptions in Theorem B1. Note that, whenever that holds, we find

$$\mathrm{RBMO} \subset \mathrm{DBMO} \subset \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega).$$

In particular, we deduce that $[\mathrm{RBMO}, L_1(\Omega)]_{1/q} \simeq L_q(\Omega)$ when this happens. As far as we know, such interpolation identities are new since Tolsa [2001] studied interpolation of operators. Unfortunately, the construction of such filtrations seems to be a difficult task in the general case. For instance, the corona-type construction described above finds some obstructions when the measure μ is supported in Cantor-like sets. Nevertheless, we may construct these filtrations in some other cases. Let us consider the following family of measures on $\mathbb{R}^n$ equipped with the Euclidean distance

$$d\mu_\beta(x) = \frac{dx}{1+|x|^\beta}.$$

These measures are nondoubling only for $\beta > n$. We will construct an admissible covering for $\beta \gtrsim n^{3/2}$ satisfying the hypotheses of Theorem B1 when d is the Euclidean metric in $\mathbb{R}^n$. We will work with the equivalent measure

$$d\nu_\beta(x) = \min\{1, |x|^{-\beta}\}\, dx$$

for convenience. Note that this does not affect the conclusions in Theorem B1.

Pick $Q_0 = [-\lambda, \lambda]^n$ with $\lambda > 1$ to be fixed, and set $Q_s = 2^s Q_0$. Consider the σ-algebras $\Sigma_{\mathrm{a}} = \sigma\langle A_s \mid s \geq 1\rangle$ and $\Sigma_{\mathrm{b}} = \sigma\langle B_s \mid s \geq 1\rangle$, where $(A_0, B_0) = (Q_0, Q_1)$ and $(A_s, B_s) = (Q_{2s} \setminus Q_{2s-2}, Q_{2s+1} \setminus Q_{2s-1})$ for any $s \geq 1$. We clearly have $\Sigma_{\mathrm{a}} \cap \Sigma_{\mathrm{b}} = \{\mathbb{R}^n, \varnothing\}$ and $\max\{|R_A|, |R_B|\} \leq 2$ for $(A, B) \in \Pi_{\mathrm{a}} \times \Pi_{\mathrm{b}}$, by

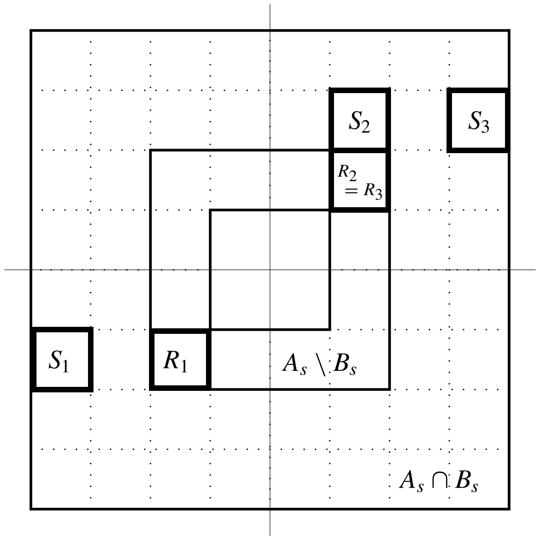

Figure 2. There is a cube R_j for each cube S_j.

construction of (Σ_a, Σ_b). Thus, it suffices to show that

$$\sup_{\substack{(A,B)\in\Pi_a\times\Pi_b\\ A\neq A_0}} \frac{\nu_\beta(A\cap B)^2}{\nu_\beta(A)\nu_\beta(B)} < \frac{1}{2}.$$

We will prove, in fact, the apparently stronger inequalities

$$\sup_{s\geq 1} \frac{\nu_\beta(A_s\cap B_{s-1})}{\nu_\beta(B_{s-1})} < \frac{1}{2} \quad \text{and} \quad \sup_{s\geq 1} \frac{\nu_\beta(A_s\cap B_s)}{\nu_\beta(A_s)} < \frac{1}{2}.$$

By symmetry of the argument, we just prove the second inequality above. Denote by L the side length of the smallest cube Q_{2s} containing A_s. Then we have that $A\cap B$ can be decomposed into $\mathrm{C}_n = 8^n - 4^n$ cubes S_j, each of which satisfies that $S_j = R_j + a_{S_j}$ for some cube $R_j = R_j(S_j) \subset A_s \setminus B_s$ of side length equal to $L/8$ and such that the angle between any point in R_j and a_{S_j} is smaller than $\pi/3$. We can also impose that $|a_{S_j}| \geq L/8$; see Figure 2. This implies that, for each x in R_j, we have

$$|x + a_{S_j}| \geq |x| + |a_{S_j}| \cos\sphericalangle(x, a_{S_j}) \geq |x| + \tfrac{1}{2}|a_{S_j}|.$$

Since $A_s \subset \mathbb{R}^n \setminus \mathrm{B}_1(0)$ for $s \geq 1$, we have

$$\begin{aligned}\nu_\beta(A_s\cap B_s) &= \int_{A_s\cap B_s} |x|^{-\beta}\,dx = \sum_{j=1}^{\mathrm{C}_n} \int_{S_j} |x|^{-\beta}\,dx \\ &= \sum_{j=1}^{\mathrm{C}_n} \int_{R_j} |x + a_{S_j}|^{-\beta}\,dx \leq \sum_{j=1}^{\mathrm{C}_n} \int_{R_j} \big(|x| + \tfrac{1}{2}|a_{S_j}|\big)^{-\beta}\,dx.\end{aligned}$$

Using that $|x| \le \sqrt{n}L/2$ for $x \in A_s$ and $|a_{S_j}| \ge L/8$,

$$\frac{1}{\nu_\beta(R_j)} \int_{R_j} \big(|x| + \tfrac{1}{2}|a_{S_j}|\big)^{-\beta} \, dx \le \sup_{x \in R_j} \frac{\big(|x| + \frac{1}{2}|a_{S_j}|\big)^{-\beta}}{|x|^{-\beta}} \le \bigg(\frac{\sqrt{n}}{\sqrt{n} + \frac{1}{8}}\bigg)^{\beta}.$$

Therefore, we obtain

$$\frac{\nu_\beta(A_s \cap B_s)}{\nu_\beta(A_s)} \le \mathrm{C}_n \bigg(\frac{\sqrt{n}}{\sqrt{n} + \frac{1}{8}}\bigg)^{\beta} \le 8^n \bigg(\frac{\sqrt{n}}{\sqrt{n} + \frac{1}{8}}\bigg)^{\beta} < \frac{1}{2} \quad \text{for } \beta \gtrsim n^{3/2}.$$

A similar argument shows that

$$\frac{\nu_\beta(A_s \cap B_{s-1})}{\nu_\beta(B_{s-1})} < \frac{1}{2}$$

for $\beta \gtrsim n^{3/2}$ and $s \ge 2$, whereas the same estimate holds for $s = 1$ as a consequence of the fact that B_0 contains $[-\lambda, \lambda]^n$ for $\lambda > 1$ large enough. This completes the construction of an admissible covering. It remains to construct filtrations $(\Sigma_{jk})_{k \ge 1}$ for $j = \mathrm{a}, \mathrm{b}$, that are regular and composed of doubling atoms. Recall that we set $\Sigma_{j1} = \Sigma_j$ and define Σ_{j2} by splitting each atom in Σ_j into a disjoint union of cubes. Namely, for $j = \mathrm{a}$ we keep A_0 and divide A_s into the cubes R_j, S_j in Figure 2. We proceed similarly for $j = \mathrm{b}$. Once we have defined Σ_{j2}, we construct Σ_{jk} by dyadic splitting of the cubes in $\Sigma_{j(k-1)}$. Note that the atoms in $\Sigma_{j1} \setminus \{A_0, B_0\}$ split at most into 8^n cubes K centered at c_K which are away from the origin. Thus

$$\nu_\beta(2K) = \int_{2K} |x|^{-\beta} \, dx \lesssim |2K| \, |c_K|^{-\beta} \lesssim |K| \, |c_K|^{-\beta} \lesssim \int_K |x|^{-\beta} \, dx = \nu_\beta(K).$$

It easily follows from this that all the atoms in Σ_{jk} are doubling up to absolute constants independent of $k \ge 1$ and that both filtrations are regular. This shows that Theorem B1 applies to $(\mathbb{R}^n, \mu_\beta)$ with the Euclidean metric.

Remark 3.2. A few comments are in order:

(i) In the light of the example above, one could wonder what happens with the positive powers $d\mu_\gamma(x) = |x|^\gamma dx$ for $\gamma > 0$, but it is straightforward to show that these measures are doubling, so that we can handle them following the construction of the previous section (see p. 731).

(ii) Our proof of Theorem B1 relies crucially on the embedding of the space DBMO in $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$ under suitable conditions. When the metric measure space (Ω, μ, d) is of polynomial growth, we know from [Tolsa 2001] that CZOs are $L_\infty \to \mathrm{RBMO}$ bounded. Since $\mathrm{RBMO} \subset \mathrm{DBMO}$, it is natural to wonder if we have

$$\mathrm{RBMO} \subset \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$$

under weaker assumptions than in Theorem B1. It turns out that this is the case when there exists filtrations composed of doubling atoms, no matter whether they are regular or not. Indeed, noticing that RBMO can be described as a subspace of DBMO with an additional condition, it is this crucial extra condition introduced by Tolsa that allows an embedding into $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$ and not into $\mathrm{bmo}_{\Sigma_{\mathrm{ab}}}(\Omega)$ for nonregular filtrations.

Concentration at the boundary. Let

$$d\mu_{\pm\alpha}(x) = e^{\pm|x|^\alpha}\,dx$$

on $\mathbb{R}^n$ equipped with the Euclidean metric. Carbonaro et al. [2009; 2010] proved that these measures satisfy their concentration condition when $\alpha > 1$. In this subsection we shall prove that our hypotheses in Theorem B1 hold for any $\alpha > 0$, hence extending their results for measures with less concentration at the boundary. Let us start with the probability measure $\mu_{-\alpha}$. Pick $K = K(n, \alpha) > 0$, a large constant of the form 2^k for some $k \geq 1$ to be fixed below. Denote by $\boldsymbol{D}(\mathbb{R}^n)$ the standard filtration of dyadic cubes in $\mathbb{R}^n$. We consider the distinguished atom $A_0 = [-K, K]^n$. The other atoms $A_s \in \Pi_{\mathrm{a}}$ for $s \geq 1$ are chosen to be the cubes in $\boldsymbol{D}(\mathbb{R}^n \setminus A_0)$ which are maximal under the following constraint on the side-length $\ell(A_s)$ in terms of the modulus of its center c_{A_s}:

$$\Pi_{\mathrm{a}} = \{A_0\} \cup \big\{A_s \text{ maximal in } \boldsymbol{D}(\mathbb{R}^n \setminus A_0) \mid \ell(A_s) \leq K|c_{A_s}|^{1-\alpha},\ s \geq 1\big\}.$$

Before defining Π_{b}, we also need another dyadic filtration $\boldsymbol{D}'(\mathbb{R}^n)$ satisfying some specific properties which we now detail. Given cubes $(A, B) \in \boldsymbol{D}(\mathbb{R}^n) \times \boldsymbol{D}'(\mathbb{R}^n)$ of comparable size — $2^{-\mathrm{k}_0} \leq \ell(A)/\ell(B) \leq 2^{\mathrm{k}_0}$ for some absolute constant k_0 — with nonempty intersection, there exists a parallelepiped $R \subset A \triangle B$ such that:

(1) R is "substantially closer" than $A \cap B$ to the origin;

(2) there exists $a = a(R) \in \mathbb{R}^n$ such that $A \cap B \subset \bigcup_{j=1}^{\mathrm{N}} R + ja$;

(3) $|a| \geq \frac{1}{\mathrm{N}} \max\{\ell(A), \ell(B)\}$ and $|x + ja| \geq |x| + \frac{1}{2}|a|$ for every $x \in R$.

Let A_1 be the cube in $\Pi_{\mathrm{a}} \setminus \{A_0\}$ whose center is the closest to the origin. Let $L = \ell(A_1)$ and pick $B_0 = A_0 + \frac{1}{3}Le_d$ with $e_d = (1, 1, \ldots, 1)$. Then, the dyadic filtration $\boldsymbol{D}'(\mathbb{R}^n)$ is defined as one of the shifted dyadic filtrations in [Conde 2013] with the initial cube being B_0. The fact that the properties above hold follows ultimately from the "good separation" between $\boldsymbol{D}(\mathbb{R}^n)$ and $\boldsymbol{D}'(\mathbb{R}^n)$. Here we pick K large enough so that the estimate $\mu_{-\alpha}\big(\frac{1}{2}A_0\big) > (1-\varepsilon)\mu_{-\alpha}(\mathbb{R}^n)$ holds. In particular, we get

$$\frac{\mu_{-\alpha}(A_0 \cap B_0)}{\mu_{-\alpha}(\mathbb{R}^n)} > 1 - \varepsilon$$

for some small $\varepsilon > 0$ to be fixed. The family Π_{b} is defined similarly by

$$\Pi_{\mathrm{b}} = \{B_0\} \cup \big\{B_s \text{ maximal in } \boldsymbol{D}'(\mathbb{R}^n \setminus B_0) \mid \ell(B_s) \leq K|c_{B_s}|^{1-\alpha},\ s \geq 1\big\}.$$

Set $\Sigma_j = \sigma(\Pi_j)$ for $j = \mathrm{a}, \mathrm{b}$ and observe that $\Sigma_{\mathrm{a}} \cap \Sigma_{\mathrm{b}} = \{\mathbb{R}^n, \varnothing\}$ by construction. Therefore, to prove that $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ yields an admissible covering we only need to check that we have

$$\sup_{A \in \Pi_{\mathrm{a}} \setminus \{A_0\}} \sum_{B \in R_A} |R_B| \frac{\mu_{-\alpha}(A \cap B)^2}{\mu_{-\alpha}(A)\mu_{-\alpha}(B)} < 1.$$

According to our definition of A_s, it is a simple exercise to check that we have $\ell(A_s) \geq \frac{1}{3}K|c_{A_s}|^{1-\alpha}$ for all $s \geq 1$ but for a finite number (independent of K) of cubes close to the origin. The same argument

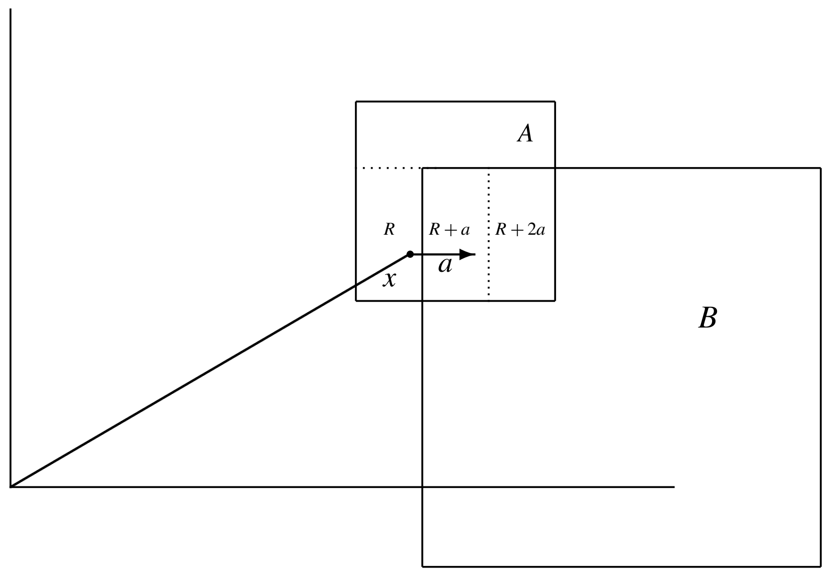

Figure 3. $A \cap B$ is covered by at most N a-translates of R; R is "substantially closer" to the origin than $A \cap B$; a is parallel to a coordinate axis for cubes A, B in a sector around that axis. In particular, x and a are "close" to being parallel, so that $|x + ja| \geq |x| + \frac{1}{2}|a|$.

holds for atoms in Π_{b}. In particular, we have $|R_A|, |R_B| \leq \mathrm{C}_n$ for all $(A, B) \in \Pi_{\mathrm{a}} \times \Pi_{\mathrm{b}}$. Therefore, when $B = B_0$ we obtain

$$|R_B| \frac{\mu_{-\alpha}(A \cap B)^2}{\mu_{-\alpha}(A)\mu_{-\alpha}(B)} \leq \mathrm{C}_n \frac{\mu_{-\alpha}(A \cap B)}{\mu_{-\alpha}(B)} < \mathrm{C}_n \frac{\varepsilon}{1-\varepsilon} < \frac{1}{2}$$

for $\varepsilon < \frac{1}{3}\mathrm{C}_n^{-1}$. Otherwise, when $B \neq B_0$, we obtain

$$\begin{aligned} \frac{\mu_{-\alpha}(A \cap B)}{\mu_{-\alpha}(R)} &= \frac{1}{\mu_{-\alpha}(R)} \int_{A \cap B} e^{-|x|^\alpha} \, dx \\ &\leq \sum_{j=1}^{\mathrm{N}} \frac{1}{\mu_{-\alpha}(R)} \int_R e^{-|x+ja|^\alpha} \, dx \\ &\leq \frac{\mathrm{N}}{\mu_{-\alpha}(R)} \int_R e^{-(|x|+|a|/2)^\alpha} \, dx \leq \mathrm{N} \sup_{x \in R} e^{-(|x|+|a|/2)^\alpha + |x|^\alpha}. \end{aligned}$$

If $\alpha = 1$, we get an estimate $\mathrm{N}e^{-|a|/2} \geq \mathrm{N}e^{-\mathrm{C}_n' K}$. For other values of $\alpha > 0$, a straightforward application of the mean value theorem gives

$$\big(|x| + \tfrac{1}{2}|a|\big)^\alpha - |x|^\alpha \geq \frac{\alpha}{18} K |c_A|^{1-\alpha} |x|^{\alpha-1} \geq \mathrm{C}_n' K$$

since $|x| \sim |c_A|$. Hence we get, for $A \neq A_0$,

$$\sum_{B \in R_A} |R_B| \frac{\mu_{-\alpha}(A \cap B)^2}{\mu{-\alpha}(A)\mu{-\alpha}(B)} < \frac{1}{2} + \mathrm{C}_n^2 \sup_{B \neq B_0} \frac{\mu_{-\alpha}(A \cap B)^2}{\mu_{-\alpha}(A)\mu_{-\alpha}(B)} \leq \frac{1}{2} + \mathrm{C}_n^2 \mathrm{N} e^{-\mathrm{C}_n' K} < 1,$$

picking $K = K(n, \alpha)$ large enough. This shows that we have an admissible covering. Note that our choice of cubes for $\alpha > 1$ is a family which becomes smaller and smaller when we get away from the origin. This is in the spirit of the Mauceri–Meda [2007] construction for the Gaussian measure. In contrast, when $\alpha < 1$ we pick larger and larger cubes as we get away from the origin. This construction seems not to be useful in [Carbonaro et al. 2009; 2010], since we may not use the locally doubling property for arbitrarily large cubes. Let us complete the proof by showing that the other hypotheses in Theorem B1 hold. Our choice of filtrations $(\Sigma_{jk})_{k\geq 1}$ for $j = \mathrm{a}, \mathrm{b}$ is by dyadic splitting of the cubes in Π_{a} and Π_{b}, respectively. The regularity of such filtrations will follow from the fact that every atom in $(\Sigma_{jk})_{k\geq 1}$ is $(3, \beta)$-doubling for some absolute constant β, and this suffices to complete the proof. If Q is any subcube of $A_0 \cup B_0$, there are dimensional constants k_n and K_n such that

$$k_n |Q| \leq \int_Q e^{-|x|^\alpha}\, dx \leq K_n |Q|,$$

and hence Q is trivially $(3, \beta)$-doubling. Otherwise, we compute

$$\begin{aligned} \frac{\mu_{-\alpha}(3Q)}{\mu_{-\alpha}(Q)} &\leq \frac{|3Q|}{|Q|} \sup_{x\in 3Q} e^{-|x|^\alpha} \sup_{x\in Q} e^{|x|^\alpha} \\ &\leq 3^n \exp\Big(\big(|x_Q| + \tfrac{1}{2}\sqrt{n}\,\ell(Q)\big)^\alpha - \big(|x_Q| - \tfrac{3}{2}\sqrt{n}\,\ell(Q)\big)^\alpha\Big) \leq \beta \end{aligned}$$

for some absolute constant $\beta > 0$, using the mean value theorem one more time.

Remark 3.3. A few comments are in order:

(i) Given $\alpha > 0$ and by minor modifications in the above arguments, we may also produce an admissible covering for $(\mathbb{R}^n, e^{|x|^\alpha} dx)$ which satisfies the hypotheses of Theorem B1 with respect to the Euclidean metric.

(ii) In this paper we $\wedge$-intersect two truncated martingale BMO spaces, but our results also hold for finite $\wedge$-intersections; details are simple and not very relevant. The Mauceri–Meda BMO space for the Gaussian measure [2007] can be described as such a finite intersection of BMO spaces using a construction similar to the one above for μ_{-2} but intersecting $n+1$ BMO spaces instead of 2. Namely, one uses as many filtrations as needed to cover all cubes in $\mathbb{R}^n$ with dyadic cubes of comparable size; see, for instance, [Conde 2013] for the optimal choice. This establishes an inclusion of their BMO space into our 2-intersection $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}$ associated to μ_{-2}, which still interpolates and is strictly larger. The latter assertion can be proved following the argument which shows that classical BMO is strictly contained in dyadic BMO.

(iii) A geometric interpretation of our definition of admissible covering could be that we still impose certain concentration at the boundary, but much less than [Carbonaro et al. 2009; 2010]. In support of this, let us consider an admissible covering $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$. Let $\mathcal{A}$ be a finite family in $\Pi_{\mathrm{a}} \setminus \{A_0\}$ and let $R_{\mathcal{A}}$ be the union $\bigcup_{A\in\mathcal{A}} R_A$. If we consider the set $R_{\mathcal{A}}$ as a measurable set and interpret $R_{\mathcal{A}} \setminus \mathcal{A}$ as the region "close to the boundary", then we can prove that

$$\mu(R_{\mathcal{A}}) \leq \frac{1}{1 - c(\Sigma_{\mathrm{ab}})} \mu(R_{\mathcal{A}} \setminus \mathcal{A})$$

or, equivalently, $\mu(\mathcal{A}) \leq c(\Sigma_{\mathrm{ab}})\mu(R_{\mathcal{A}})$. Indeed, we have

$$\begin{aligned}\mu(\mathcal{A}) &= \sum_{A\in\mathcal{A}} \sum_{B\in R_{\mathcal{A}}} \mu(A\cap B)\\ &\leq c(\Sigma_{\mathrm{ab}})^{1/2} \sum_{A\in\mathcal{A}} \mu(A)^{1/2}\biggl(\sum_{B\in R_{\mathcal{A}}} \frac{\mu(B)}{|R_B|}\biggr)^{\frac{1}{2}}\\ &\leq c(\Sigma_{\mathrm{ab}})^{1/2}\mu(\mathcal{A})^{1/2}\biggl(\sum_{A\in\mathcal{A}} \sum_{B\in R_{\mathcal{A}}} \frac{\mu(B)}{|R_B|}\biggr)^{\frac{1}{2}} \leq (c(\Sigma_{\mathrm{ab}})\mu(\mathcal{A})\mu(R_{\mathcal{A}}))^{1/2}.\end{aligned}$$

(iv) In the case of the Gaussian measure on $\mathbb{R}^n$, Mauceri, Meda and Sjögren [Mauceri et al. 2012] proved that R_i, S_i, the Riesz transforms associated with the Ornstein–Uhlenbeck semigroups, are bounded from L_∞ to Mauceri–Meda BMO spaces, but their adjoint operators R_i^*, S_i^* are not when $n \geq 2$. As explained in (ii), our BMO spaces are strictly larger than Mauceri and Meda's if the σ-algebras Σ_a, Σ_b are picked as in the beginning of this subsection. Therefore, the Riesz transforms R_i, S_i studied in [Mauceri et al. 2012] are bounded from L^∞ to our BMO spaces as well. Pierre Portal [2014] introduced a different type of Hardy spaces using truncated maximal functions and square functions. He proved that the Riesz transforms R_i, S_i and their adjoint operators R_i^*, S_i^* are all bounded from his H_1 space to L_1 [Portal 2014, Theorem 6.1]. It is interesting to determine whether R_i^*, S_i^* are L^∞-BMO bounded for our BMO spaces with carefully picked Σ_a, Σ_b.

4. Calderón–Zygmund operators, II

In this section we will study the class of atomic Calderón–Zygmund operators (ACZO) defined in the introduction over a given measure space (Ω, Σ, μ). More precisely, we shall prove Theorem B2 and illustrate it with a few constructions of dyadic operators satisfying its hypotheses.

Proof of Theorem B2. Following the same argument as in the proof of Theorem B1, we can use duality and our interpolation result in Theorem A to reduce the L_p-boundedness in the assertion to the $L_\infty(\Omega) \to \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega)$ boundedness of our ACZO. This is however standard. Indeed, since the filtration is regular we know that $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\Omega) \simeq \mathrm{bmo}_{\Sigma_{\mathrm{ab}}}(\Omega)$. Up to absolute constants, the norm in the latter space is given by

$$\|Tf\|_{\mathrm{bmo}_{\Sigma_{\mathrm{ab}}}(\Omega)} = \sup_{Q\in\mathbf{\Pi}} \inf_{\mathrm{k}_Q\in\mathbb{C}} \biggl(\frac{1}{\mu(Q)}\int_Q |f(w) - \mathrm{k}_Q|^2\, d\mu(\omega)\biggr)^{\frac{1}{2}},$$

where $\mathbf{\Pi} = \mathbf{\Pi}_{\mathrm{a}} \cup \mathbf{\Pi}_{\mathrm{b}}$ is the set of atoms in any of the two filtrations. Decompose

$$f = f\chi_{\hat{Q}} + f\chi_{\Omega\setminus\hat{Q}} = f_1 + f_2.$$

As usual, we pick $\mathrm{k}_Q = (Tf_2)_Q$. Then, we control the term for Tf_1 using the L_2-boundedness of T and the regularity of the filtrations. The term $Tf_2 - \mathrm{k}_Q$ is dominated by means of the Hörmander kernel condition given in the definition of ACZO. Namely,

$$\left(\frac{1}{\mu(Q)}\int_Q |Tf_1(\omega)|^2\,d\mu(\omega)\right)^{\frac{1}{2}} \le \|T\|_{2\to 2}\sqrt{\frac{\mu(\hat{Q})}{\mu(Q)}}\|f\|_\infty \lesssim \|f\|_\infty$$

by regularity of the filtrations. On the other hand

$$\left(\frac{1}{\mu(Q)}\int_Q |Tf_2(\omega)-(Tf_2)_Q|^2\,d\mu(\omega)\right)^{\frac{1}{2}}$$
$$\le \left(\frac{1}{\mu(Q)^2}\int_Q\int_Q\left[\int_{\Omega\setminus\hat{Q}} |K(z_1,x)-K(z_2,x)|\,d\mu(x)\right]^2 d\mu(z_1)\,d\mu(z_2)\right)^{\frac{1}{2}}\|f\|_\infty,$$

which is dominated by $\|f\|_\infty$ according to the Hörmander condition for ACZOs. □

Remark 4.1. As mentioned in the introduction, standard prototypes of atomic Calderón–Zygmund operators include martingale transforms, perfect dyadic CZOs and Haar shift operators. These are usually defined on the Euclidean space $\mathbb{R}^n$ equipped with a dyadic filtration. Nevertheless, the exact same arguments apply on any dyadically doubling measure space or even for any measure space equipped with a two-sided regular filtration. López-Sánchez et al. [2014] have studied those nondoubling measure spaces for which Haar shift operators satisfy weak type-(1, 1) estimates. Theorem B2 provides a tool to produce nondoubling measure spaces over which Haar shifts, or more general atomic CZOs, are $L_\infty \to \mathrm{BMO}$ bounded. In the case of martingale transforms, Haar shift operators and perfect dyadic CZOs in $(\mathbb{R}^n, \mu)$, all of them satisfy the Hörmander-like condition

$$\sup_{Q\in\boldsymbol{D}(\mathbb{R}^n)}\ \sup_{z_1,z_2\in Q}\int_{\mathbb{R}^n\setminus\hat{Q}} |k(z_1,x)-k(z_2,x)|+|k(x,z_1)-k(x,z_2)|\,d\mu(x)<\infty.$$

This means that these operators are ACZOs satisfying Theorem B2 as long as we can find an admissible covering $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ and regular filtrations over it such that all the atoms are cubes in $\boldsymbol{D}(\mathbb{R}^n)$ or suitable unions of those. If we review our examples in Section 3, this is not the case for our construction for $d\mu_{\pm\alpha}(x) = e^{\pm|x|^\alpha}dx$. It is however quite simple to adapt our construction for

$$d\mu_\beta(x) = \frac{dx}{1+|x|^\beta}$$

so that it satisfies the hypotheses of Theorem B2. In particular, the Haar shift operators defined on $(\mathbb{R}^n, \mu_\beta)$ are $L_\infty(\mathbb{R}^n) \to \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathbb{R}^n)$ bounded. It remains open to decide whether an admissible covering exists on the exponential measure spaces $(\mathbb{R}^n, \mu_{\pm\alpha})$ using only atoms associated to one and not two dyadic systems.

5. Matrix-valued forms of our results

In this section, we extend our main results to the context of operator-valued functions. Noncommutative forms of Calderón–Zygmund theory have been recently studied in [Junge et al. $\ge$ 2015; Mei and Parcet 2009; Parcet 2009; Mei 2007]. There are however no specific results in the context of nondoubling metric measure spaces. Unfortunately, it seems difficult to extend the approach of [Tolsa 2001] or [Mauceri and Meda 2007] to the operator-valued or even the noncommutative setting, since their interpolation results rest on good-λ inequalities which do not have a noncommutative analogue so far. On the other hand, the semicommutative approach in [Parcet 2009] is valid for doubling spaces, but again presents serious obstacles to be extended to the nondoubling setting. The crucial aspect of our approach is that it ultimately rests on martingale inequalities that have been successfully transferred to the noncommutative setting. Namely, after Pixier and Xu's [1997] seminal contribution on Burkholder–Gundy inequalities for noncommutative martingales, we find analogues of Doob's maximal inequalities, Gundy, Davis and atomic decompositions, Burkholder conditional square functions, John–Nirenberg inequalities, L_p/BMO interpolation results; see [Hong and Mei 2012; Junge 2002; Junge and Mei 2010; Junge and Musat 2007; Junge and Perrin 2014; Junge and Xu 2003; Mei 2007; Musat 2003; Parcet and Randrianantoanina 2006; Perrin 2009] and the references therein.

Let us briefly introduce the framework for our results in this section; we refer to [Parcet 2009, Section 1] for a rather complete review of the necessary background adapted to our necessities. We also refer the reader to Pisier and Xu's survey [2003] for more on noncommutative L_p theory. Let (Ω, Σ, μ) be a σ-finite measure space and consider any pair $(\mathcal{M}, \tau)$ given by a von Neumann algebra $\mathcal{M}$ equipped with a normal, semifinite, faithful trace τ. This is sometimes called a noncommutative measure space. We will write $(\mathcal{R}, \varphi)$ to denote the von Neumann algebra generated by essentially bounded functions $f : \Omega \to \mathcal{M}$ equipped with the trace

$$\varphi(f) = \int_\Omega \tau(f(\omega))\, d\mu(\omega).$$

$\mathcal{R}$ is the von Neumann algebra tensor product $\mathcal{R} = L_\infty(\Omega)\overline{\otimes}\mathcal{M}$ and we may consider the corresponding noncommutative spaces $L_p(\mathcal{R}, \varphi)$. This semicommutative model is the context where we intend to generalize our main results. Apart from its own interest as an operator-valued model, it constitutes a first step towards further results for more general von Neumann algebras. In particular, as [Junge et al. 2014] demonstrates, certain fully noncommutative questions can be reduced to the semicommutative setting. Readers not familiar with von Neumann algebra theory are encouraged to read this section restricting their attention to matrix-valued functions. In other words, replace $\mathcal{M}$ by the algebra M_m of $m \times m$ matrices and τ by the standard trace tr. The difficulties are similar in this case to in the general setting, as long as we provide results with constants independent of m. We also refer to [Parcet 2009] for a comparison between this model and the vector-valued setting, which differs substantially in the endpoint estimates.

The BMO spaces. First we review the definitions and results in Section 1 for the semicommutative setting described above. Given a filtration $(\Sigma_k)_{k\geq 1}$ of (Ω, Σ, μ), we consider the conditional expectations

$$f \mapsto \mathsf{E}_{\Sigma_k} \otimes \mathrm{id}_{\mathcal{M}}(f) \in \mathcal{R} \quad \text{for } f \in \mathcal{R},$$

still denoted by E_{Σ_k}. The martingale bmo and BMO norms are

$$\|f\|_{\mathrm{bmo}} = \max\{\|f\|_{\mathrm{bmo_c}},\ \|f^*\|_{\mathrm{bmo_c}}\},$$
$$\|f\|_{\mathrm{BMO}} = \max\{\|f\|_{\mathrm{BMO_c}},\ \|f^*\|_{\mathrm{BMO_c}}\},$$

where the column norms are defined as in the commutative case, taking into account that we use $|x|^2 = x^*x$ for any operator x on a Hilbert space. The interpolation result $[\mathrm{BMO}, L_1(\mathcal{R})]_{1/p} \simeq L_p(\mathcal{R})$ was proved by Musat [2003] for any semifinite von Neumann algebra $\mathcal{R}$. This is the noncommutative analogue of the Janson–Jones interpolation theorem. If we set $\|f\|_{\mathrm{h}_p^{\mathrm{r}}} = \|f^*\|_{\mathrm{h}_p^{\mathrm{c}}}$, where the norm in $\mathrm{h}_p^{\mathrm{c}}$ is defined as in the commutative case, then the noncommutative Hardy spaces have the form

$$\mathrm{h}_p = \begin{cases} \mathrm{h}_p^{\mathrm{r}} + \mathrm{h}_p^{\mathrm{c}} & \text{if } 1 \le p \le 2, \\ \mathrm{h}_p^{\mathrm{r}} \cap \mathrm{h}_p^{\mathrm{c}} & \text{if } 2 \le p \le \infty. \end{cases}$$

This combination of row and column square functions is known to be the right one for L_p inequalities, as was discovered for the first time with the noncommutative Khintchine inequalities [Lust-Piquard 1986; Lust-Piquard and Pisier 1991]. The interpolation result $[\mathrm{bmo}, \mathrm{h}_1]_{1/p} \simeq \mathrm{h}_p$ was proved in [Bekjan et al. 2010] for noncommutative martingales. As in the commutative case, the projections $J_{\Sigma_1} = \mathrm{id} - \mathsf{E}_{\Sigma_1}$ are bounded on bmo, BMO, L_p and h_p, so that we will be working with these complemented subspaces which enjoy the same interpolation and duality properties as the original spaces. Note that the identity

$$\begin{aligned} \|f\|_{J_{\Sigma_1}(\mathrm{bmo_c})} &= \sup_{k \ge 1} \big\| (\mathsf{E}_{\Sigma_k} |f - \mathsf{E}_{\Sigma_k} f|^2)^{1/2} \big\|_{\mathcal{M}} \\ &= \sup_{A \in \mathbf{\Pi}} \left\| \left(\frac{1}{\mu(A)} \int_A \left| f(w) - \frac{1}{\mu(A)} \int_A f \, d\mu \right|^2 d\mu(w) \right)^{\frac{1}{2}} \right\|_{\mathcal{M}} \end{aligned}$$

still holds and we have $J_{\Sigma_1}(\mathrm{bmo}) \simeq J_{\Sigma_1}(\mathrm{BMO})$ for regular filtrations. Consider an admissible covering $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ of (Ω, Σ, μ) and any pair of filtrations $(\Sigma_{jk})_{k \ge 1}$ with $\Sigma_{j1} = \Sigma_j$ for $j = \mathrm{a}, \mathrm{b}$. Denote by $\mathrm{BMO_a}$ and $\mathrm{BMO_b}$ the BMO spaces associated to these filtrations in the semicommutative algebra $\mathcal{R}$ and set

$$\mathrm{BMO}_{\Sigma_j}(\mathcal{R}) = J_{\Sigma_j}(\mathrm{BMO}_j) \quad \text{and} \quad \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R}) = \mathrm{BMO}_{\Sigma_{\mathrm{a}}}(\mathcal{R}) \wedge \mathrm{BMO}_{\Sigma_{\mathrm{b}}}(\mathcal{R}).$$

The John–Nirenberg inequalities, atomic descriptions of H_1 and duality results have also been transferred to the context of noncommutative martingales [Bekjan et al. 2010; Hong and Mei 2012; Junge and Musat 2007; Pisier and Xu 1997] and we will not review these results here, since they will not play a crucial role.

The interpolation theorem. Let us now state the analogue of Theorem A in the operator-valued setting. As usual, we will write $L_p^\circ(\mathcal{R})$ for the subspace of mean-zero elements with respect to μ. In the terminology we use for admissible coverings,

$$L_p^\circ(\mathcal{R}) \simeq J_{\Sigma_{\mathrm{a}} \cap \Sigma_{\mathrm{b}}}(L_p(\mathcal{R})).$$

Theorem 5.1. *Let $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ be an admissible covering in (Ω, Σ, μ) and consider the semicommutative space $\mathcal{R} = L_\infty(\Omega) \overline{\otimes} \mathcal{M}$. Then, for each $2 \le p < \infty$, there exists a constant $c_p \ge 1$ such that*

$$L_p^\circ(\mathcal{R}) \simeq_{c_p} J_{\Sigma_{\mathrm{a}}}(L_p(\mathcal{R})) \wedge J_{\Sigma_{\mathrm{b}}}(L_p(\mathcal{R})).$$

In particular, we have by complex interpolation that

$$[\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R}), L_1^\circ(\mathcal{R})]_{1/q} \simeq_{c_q} L_q^\circ(\mathcal{R}) \quad (1 < q < \infty),$$

with $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R})$ constructed with any two filtrations over $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$.

Sketch of the proof. Thanks to the close connection with martingales, the proof is entirely parallel to the one given in the classical case. Indeed, combining standard facts from noncommutative L_p theory with the martingale results reviewed in the previous subsection, it is a simple exercise to adapt our proof of Theorem A to the present case. The only subtle point is the inequality

$$\|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}} f\|_p^p \leq \|\mathsf{E}_{\Sigma_\mathrm{a}}|\mathsf{E}_{\Sigma_\mathrm{a}} f|^{p/2}\|_2^2,$$

which is used in the last two steps of our argument. Namely, in the classical case this is due to the conditional Jensen's inequality $\phi(\mathsf{E}_{\Sigma_k} f) \leq \mathsf{E}_{\Sigma_k}\phi(f)$ for convex functions ϕ. In contrast, its noncommutative form does not hold for all $p \geq 2$, since we need the operator-convexity of the function $\phi(x) = |x|^\beta$ for $\beta = p/2$ and x not necessarily positive. This is the case for $\beta \geq 2$, or equivalently $p \geq 4$, but it fails for $2 \leq p < 4$. Note however that the ultimate goal in Steps 5 and 6 is to show that $\|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}} f\|_p \leq c_p(\Sigma_{\mathrm{ab}})\|f\|_p$ for some $0 < c_p(\Sigma_{\mathrm{ab}}) < 1$. To prove it, we observe that

$$\mathsf{E}_{\Sigma_\mathrm{a}}(g_1^* g_2) = \xi_k(g_1)^*\xi_k(g_2),$$
$$\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}}(f_1^* f_2) = \omega_k(f_1)^*\omega_k(f_2)$$

for certain right $\mathcal{R}_k$-module maps $\xi_k, \omega_k : L_q(\mathcal{R}) \to C_q(L_q(\mathcal{R}))$ with $\mathcal{R}_k = \mathsf{E}_{\Sigma_k}(\mathcal{R})$. This follows from standard factorization properties of completely positive unital maps in terms of Hilbert modules; see, for instance, [Junge 2002]. Let us consider the polar decompositions $f = u|f|$ and $g = v|g|$ of $g = \mathsf{E}_{\Sigma_\mathrm{b}} f$. Then we can factorize $\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}} f$ in two ways:

$$\mathsf{E}_{\Sigma_\mathrm{a}} g = \mathsf{E}_{\Sigma_\mathrm{a}}(v|g|^{1/2}|g|^{1/2}) = \xi_k(|g|^{1/2}v^*)^*\xi_k(|g|^{1/2}),$$
$$\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}} f = \mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}}(u|f|^{1/2}|f|^{1/2}) = \omega_k(|f|^{1/2}u^*)^*\omega_k(|f|^{1/2}).$$

This yields the estimates

$$\begin{aligned}
\|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}} f\|_p &\leq \|\xi_k(|g|^{1/2}v^*)\|_{2p}\|\xi_k(|g|^{1/2})\|_{2p} \\
&\leq \|\xi_k(|g|^{1/2}v^*)^*\xi_k(|g|^{1/2}v^*)\|_p^{1/2}\|\xi_k(|g|^{1/2})^*\xi_k(|g|^{1/2})\|_p^{1/2} \\
&= \|\mathsf{E}_{\Sigma_\mathrm{a}}(v|g|v^*)\|_p^{1/2}\|\mathsf{E}_{\Sigma_\mathrm{a}}(|g|)\|_p^{1/2} \leq \|f\|_p^{1/2}\|\mathsf{E}_{\Sigma_\mathrm{a}}|\mathsf{E}_{\Sigma_\mathrm{b}} f|^{p/2}\|_2^{1/p},
\end{aligned}$$

and

$$\begin{aligned}
\|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}} f\|_p &\leq \|\omega_k(|f|^{1/2}u^*)\|_{2p}\|\omega_k(|f|^{1/2})\|_{2p} \\
&\leq \|\omega_k(|f|^{1/2}u^*)^*\omega_k(|f|^{1/2}u^*)\|_p^{1/2}\|\omega_k(|f|^{1/2})^*\omega_k(|f|^{1/2})\|_p^{1/2} \\
&= \|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}}(u|f|u^*)\|_p^{1/2}\|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}}(|f|)\|_p^{1/2} \leq \|f\|_p^{1/2}\|\mathsf{E}_{\Sigma_\mathrm{a}}\mathsf{E}_{\Sigma_\mathrm{b}}|f|^{p/2}\|_2^{1/p}.
\end{aligned}$$

The last inequality in both estimates follows from the Kadison–Schwarz inequality for operator-convex functions, since $\phi(x) = x^\beta$ is operator-convex on $\mathbb{R}_+$ for $\beta \geq 1$. The first estimate is the right one to use in Step 5 and the second one in Step 6. □

The Calderón–Zygmund operators. We now consider Calderón–Zygmund operators in semicommutative algebras associated to operator-valued kernels. Our construction is standard; we refer to [Duoandikoetxea 2001; Junge et al. 2014; Rubio de Francia et al. 1986] for further details. Let us write $L_0(\mathcal{M})$ for the $*$-algebra of τ-measurable operators affiliated with $\mathcal{M}$ and consider kernels $k : (\Omega \times \Omega) \setminus \Delta \to \mathcal{L}(L_0(\mathcal{M}))$

defined away from the diagonal Δ of $\Omega \times \Omega$ and which take values in linear maps on τ-measurable operators. If d is a metric in Ω, the standard Hörmander kernel condition takes the same form in this setting when we replace the absolute value by the norm in the algebra $\mathcal{B}(\mathcal{M})$ of bounded linear operators acting on $\mathcal{M}$:

$$\sup_{\substack{\mathrm{B}\,d\text{-ball} \\ z_1, z_2 \in \mathrm{B}}} \int_{\Omega \setminus \alpha \mathrm{B}} \|k(z_1, x) - k(z_2, x)\|_{\mathcal{B}(\mathcal{M})} + \|k(x, z_1) - k(x, z_2)\|_{\mathcal{B}(\mathcal{M})} \, d\mu(x) < \infty.$$

Define a CZO in $(\mathcal{R}, \varphi, d)$ as any linear map T satisfying the following properties:

- T is bounded on $L_\infty(\mathcal{M}; L_2^r(\Omega))$,
$$\Big\| \Big(\int_\Omega Tf(x) Tf(x)^* \, d\mu(x) \Big)^{\frac{1}{2}} \Big\|_{\mathcal{M}} \lesssim \Big\| \Big(\int_\Omega f(x) f(x)^* \, d\mu(x) \Big)^{\frac{1}{2}} \Big\|_{\mathcal{M}}.$$
- T is bounded on $L_\infty(\mathcal{M}; L_2^c(\Omega))$,
$$\Big\| \Big(\int_\Omega Tf(x)^* Tf(x) \, d\mu(x) \Big)^{\frac{1}{2}} \Big\|_{\mathcal{M}} \lesssim \Big\| \Big(\int_\Omega f(x)^* f(x) \, d\mu(x) \Big)^{\frac{1}{2}} \Big\|_{\mathcal{M}}.$$
- The kernel representation
$$Tf(x) = \int_\Omega k(x, y)(f(y)) \, d\mu(y) \quad \text{holds for } x \notin \operatorname{supp}_\Omega f$$
and some kernel $k : (\Omega \times \Omega) \setminus \Delta \to \mathbb{C}$ satisfying the Hörmander condition.

The first two conditions replace the usual L_2-boundedness; see [Junge et al. 2014] for explanations.

Theorem 5.2. *Let* $(\Sigma_{\mathrm{a}}, \Sigma_{\mathrm{b}})$ *be an admissible covering of* (Ω, Σ, μ). *Assume that* Σ *admits regular filtrations* $(\Sigma_{jk})_{k \ge 1}$ *by successive refinement of* $\Sigma_{j1} = \Sigma_j$ *for* $j = \mathrm{a}, \mathrm{b}$ *and that each atom in* Σ_{jk} *is a* $(\mathrm{C}_0, \alpha, \beta)$*-doubling set for certain absolute constants* $\mathrm{C}_0, \alpha, \beta > 0$. *Let* $\mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R})$ *denote the* $\wedge$*-intersection of the* BMO *spaces defined over these filtrations. Then, every* CZO *extends to a bounded map*:

(i) $\mathrm{H}^1_{\Sigma_{\mathrm{ab}}}(\mathcal{R}) \to L_1(\mathcal{R})$;

(ii) $L_\infty(\mathcal{R}) \to \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R})$;

(iii) $L_p^\circ(\mathcal{R}) \to L_p^\circ(\mathcal{R})$ *for* $1 < p < \infty$.

Moreover, if T *is* $L_2(\mathcal{R})$*-bounded then* $T : L_p(\mathcal{R}) \to L_p(\mathcal{R})$ *for all* $1 < p < \infty$.

Proof. According to Theorem 5.1 (interpolation) and the semicommutative form of Remark 3.1 (duality), it turns out that $L_\infty(\mathcal{R}) \to \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R})$ boundedness automatically implies $\mathrm{H}^1_{\Sigma_{\mathrm{ab}}}(\mathcal{R}) \to L_1(\mathcal{R})$ boundedness, as well as $L_p^\circ(\mathcal{R}) \to L_p^\circ(\mathcal{R})$ boundedness. Moreover, if T is also L_2-bounded we may reproduce the argument given in the proof of Theorem B1 to obtain L_p-boundedness for all $1 < p < \infty$. Let us then focus on the $L_\infty \to \mathrm{BMO}$ boundedness. Define

$$\mathrm{DBMO} = \mathrm{DBMO}_{\mathrm{r}} \cap \mathrm{DBMO}_{\mathrm{c}}$$

with $\|f\|_{\mathrm{DBMO_r}} = \|f^*\|_{\mathrm{DBMO_c}}$ and

$$\|f\|_{\mathrm{DBMO_c}} = \sup_{\substack{\mathrm{B\ ball}\\ d\text{-doubling}}} \left\| \left(\frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} \left| f(w) - \frac{1}{\mu(\mathrm{B})} \int_{\mathrm{B}} f \, d\mu \right|^2 d\mu(w) \right)^{\frac{1}{2}} \right\|_{\mathcal{M}}.$$

As usual, we write $|x|^2$ for x^*x. The assertion follows from

$$L_\infty(\mathcal{R}) \xrightarrow{T} \mathrm{DBMO} \xrightarrow{\mathrm{id}} \mathrm{bmo}_{\Sigma_{\mathrm{ab}}}(\mathcal{R}) \simeq \mathrm{BMO}_{\Sigma_{\mathrm{ab}}}(\mathcal{R}).$$

The boundedness of the chain above can be justified as in the proof of Theorem B1. Indeed, the analogies in the argument lead us to apply the new conditions which appear in our definition of semicommutative CZO; see [Junge et al. 2014]. □

Remark 5.3. Theorem B2 also admits a straightforward generalization to the semicommutative setting. Again, our use of martingale techniques makes the proof entirely analogous, so that we think it would be too repetitive to include it here.

Acknowledgements

Conde-Alonso and Parcet are partially supported by the ERC StG-256997-CZOSQP, the Spanish grant MTM2010-16518 and by ICMAT Severo Ochoa Grant SEV-2011-0087 (Spain). Mei is partially supported by the NSF DMS-1266042.

References

[Bekjan et al. 2010] T. N. Bekjan, Z. Chen, M. Perrin, and Z. Yin, "Atomic decomposition and interpolation for Hardy spaces of noncommutative martingales", *J. Funct. Anal.* **258**:7 (2010), 2483–2505. MR 2011d:46131 Zbl 1198.46047

[Bergh and Löfström 1976] J. Bergh and J. Löfström, *Interpolation spaces: an introduction*, Grundlehren der Math. Wissenschaften **223**, Springer, Berlin, 1976. MR 58 #2349 Zbl 0344.46071

[Carbonaro et al. 2009] A. Carbonaro, G. Mauceri, and S. Meda, "H^1 and BMO for certain locally doubling metric measure spaces", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **8**:3 (2009), 543–582. MR 2010j:42025 Zbl 1180.42008

[Carbonaro et al. 2010] A. Carbonaro, G. Mauceri, and S. Meda, "H^1 and BMO for certain locally doubling metric measure spaces of finite measure", *Colloq. Math.* **118**:1 (2010), 13–41. MR 2011b:42034 Zbl 1193.42076

[Christ 1990] M. Christ, "A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral", *Colloq. Math.* **60/61**:2 (1990), 601–628. MR 92k:42020 Zbl 0758.42009

[Conde 2013] J. M. Conde, "A note on dyadic coverings and nondoubling Calderón–Zygmund theory", *J. Math. Anal. Appl.* **397**:2 (2013), 785–790. MR 2979613 Zbl 1254.42021

[Conde-Alonso and Parcet 2014] J. M. Conde-Alonso and J. Parcet, "Atomic blocks for noncommutative martingales", preprint, 2014. arXiv 1407.5451

[Davis 1970] B. Davis, "On the integrability of the martingale square function", *Israel J. Math.* **8** (1970), 187–190. MR 42 #3863 Zbl 0211.21902

[Duoandikoetxea 2001] J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics **29**, American Mathematical Society, Providence, RI, 2001. MR 2001k:42001 Zbl 0969.42001

[Rubio de Francia et al. 1986] J. L. Rubio de Francia, F. J. Ruiz, and J. L. Torrea, "Calderón–Zygmund theory for operator-valued kernels", *Adv. in Math.* **62**:1 (1986), 7–48. MR 88f:42035 Zbl 0627.42008

[Garnett and Jones 1982] J. B. Garnett and P. W. Jones, "BMO from dyadic BMO", *Pacific J. Math.* **99**:2 (1982), 351–371. MR 85d:42021 Zbl 0516.46021

[Garsia 1973] A. M. Garsia, *Martingale inequalities: seminar notes on recent progress*, W. A. Benjamin, Reading, MA–London–Amsterdam, 1973. MR 56 #6844 Zbl 0284.60046

[Hong and Mei 2012] G. Hong and T. Mei, "John–Nirenberg inequality and atomic decomposition for noncommutative martingales", *J. Funct. Anal.* **263**:4 (2012), 1064–1097. MR 2927404 Zbl 06071594

[Hytönen and Kairema 2012] T. Hytönen and A. Kairema, "Systems of dyadic cubes in a doubling metric space", *Colloq. Math.* **126**:1 (2012), 1–33. MR 2901199 Zbl 1244.42010

[Janson and Jones 1982] S. Janson and P. W. Jones, "Interpolation between H^p spaces: the complex method", *J. Funct. Anal.* **48**:1 (1982), 58–80. MR 84c:46021 Zbl 0507.46047

[John and Nirenberg 1961] F. John and L. Nirenberg, "On functions of bounded mean oscillation", *Comm. Pure Appl. Math.* **14** (1961), 415–426. MR 24 #A1348 Zbl 0102.04302

[Junge 2002] M. Junge, "Doob's inequality for non-commutative martingales", *J. Reine Angew. Math.* **549** (2002), 149–190. MR 2003k:46097 Zbl 1004.46043

[Junge and Mei 2010] M. Junge and T. Mei, "Noncommutative Riesz transforms—a probabilistic approach", *Amer. J. Math.* **132**:3 (2010), 611–680. MR 2012b:46136 Zbl 1200.46056

[Junge and Musat 2007] M. Junge and M. Musat, "A noncommutative version of the John–Nirenberg theorem", *Trans. Amer. Math. Soc.* **359**:1 (2007), 115–142. MR 2007m:46103 Zbl 1173.46046

[Junge and Perrin 2014] M. Junge and M. Perrin, *Theory of $\mathcal{H}_p$-spaces for continuous filtrations in von Neumann algebras*, Astérisque **362**, 2014. MR 3241706

[Junge and Xu 2003] M. Junge and Q. Xu, "Noncommutative Burkholder/Rosenthal inequalities", *Ann. Probab.* **31**:2 (2003), 948–995. MR 2004f:46078 Zbl 1041.46050

[Junge et al. 2014] M. Junge, T. Mei, and J. Parcet, "Smooth Fourier multipliers on group von Neumann algebras", *Geom. Funct. Anal.* **24**:6 (2014), 1913–1980. MR 3283931

[Junge et al. $\geq$ 2015] M. Junge, T. Mei, and J. Parcet, "Algebraic Calderón–Zygmund theory". In preparation.

[López-Sánchez et al. 2014] L. D. López-Sánchez, J. M. Martell, and J. Parcet, "Dyadic harmonic analysis beyond doubling measures", *Adv. Math.* **267** (2014), 44–93. MR 3269175 Zbl 1304.42041

[Lust-Piquard 1986] F. Lust-Piquard, "Inégalités de Khintchine dans C_p $(1 < p < \infty)$", *C. R. Acad. Sci. Paris Sér. I Math.* **303**:7 (1986), 289–292. MR 87j:47032 Zbl 0592.47038

[Lust-Piquard and Pisier 1991] F. Lust-Piquard and G. Pisier, "Noncommutative Khintchine and Paley inequalities", *Ark. Mat.* **29**:2 (1991), 241–260. MR 94b:46011 Zbl 0755.47029

[Mateu et al. 2000] J. Mateu, P. Mattila, A. Nicolau, and J. Orobitg, "BMO for nondoubling measures", *Duke Math. J.* **102**:3 (2000), 533–565. MR 2001e:26019 Zbl 0964.42009

[Mauceri and Meda 2007] G. Mauceri and S. Meda, "BMO and H^1 for the Ornstein–Uhlenbeck operator", *J. Funct. Anal.* **252**:1 (2007), 278–313. MR 2008m:42024 Zbl 1136.46027

[Mauceri et al. 2012] G. Mauceri, S. Meda, and P. Sjögren, "Endpoint estimates for first-order Riesz transforms associated to the Ornstein–Uhlenbeck operator", *Rev. Mat. Iberoam.* **28**:1 (2012), 77–91. MR 2904131 Zbl 1238.42005

[Mei 2003] T. Mei, "BMO is the intersection of two translates of dyadic BMO", *C. R. Math. Acad. Sci. Paris* **336**:12 (2003), 1003–1006. MR 2004e:42037 Zbl 1032.42026

[Mei 2007] T. Mei, "Operator valued Hardy spaces", *Mem. Amer. Math. Soc.* **188**:881 (2007), vi+64. MR 2010d:46085

[Mei and Parcet 2009] T. Mei and J. Parcet, "Pseudo-localization of singular integrals and noncommutative Littlewood–Paley inequalities", *Int. Math. Res. Not.* **2009**:8 (2009), 1433–1487. MR 2010k:46063 Zbl 1175.46057

[Musat 2003] M. Musat, "Interpolation between non-commutative BMO and non-commutative L_p-spaces", *J. Funct. Anal.* **202**:1 (2003), 195–225. MR 2004g:46081 Zbl 1042.46038

[Nazarov et al. 2002] F. Nazarov, S. Treil, and A. Volberg, "Accretive system Tb-theorems on nonhomogeneous spaces", *Duke Math. J.* **113**:2 (2002), 259–312. MR 2003g:42030 Zbl 1055.47027

[Parcet 2009] J. Parcet, "Pseudo-localization of singular integrals and noncommutative Calderón–Zygmund theory", *J. Funct. Anal.* **256**:2 (2009), 509–593. MR 2010g:42035 Zbl 1179.46051

[Parcet and Randrianantoanina 2006] J. Parcet and N. Randrianantoanina, "Gundy's decomposition for non-commutative martingales and applications", *Proc. London Math. Soc.* (3) **93**:1 (2006), 227–252. MR 2007b:46115 Zbl 1119.46049

[Perrin 2009] M. Perrin, "A noncommutative Davis' decomposition for martingales", *J. Lond. Math. Soc.* (2) **80**:3 (2009), 627–648. MR 2011e:46104 Zbl 1225.46048

[Pisier and Xu 1997] G. Pisier and Q. Xu, "Non-commutative martingale inequalities", *Comm. Math. Phys.* **189**:3 (1997), 667–698. MR 98m:46079 Zbl 0898.46056

[Pisier and Xu 2003] G. Pisier and Q. Xu, "Non-commutative L^p-spaces", pp. 1459–1517 in *Handbook of the geometry of Banach spaces, II*, edited by W. B. Johnson and J. Lindenstrauss, North-Holland, Amsterdam, 2003. MR 2004i:46095 Zbl 1046.46048

[Portal 2014] P. Portal, "Maximal and quadratic Gaussian Hardy spaces", *Rev. Mat. Iberoam.* **30**:1 (2014), 79–108. MR 3186932 Zbl 1290.42048

[Tolsa 2001] X. Tolsa, "BMO, H^1, and Calderón–Zygmund operators for non doubling measures", *Math. Ann.* **319**:1 (2001), 89–149. MR 2002c:42029 Zbl 0974.42014

Received 9 Jul 2014. Revised 18 Jan 2015. Accepted 6 Mar 2015.

JOSE M. CONDE-ALONSO: jose.conde@icmat.es
, *Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM Consejo Superior de Investigaciones Científicas, C/ Nicolás Cabrera 13-15, 28049 Madrid, Spain*

TAO MEI: mei@wayne.edu
Department of Mathematics, Wayne State University, 656 W. Kirby St., Detroit, MI 48202, United States

and

Department of Mathematics, Baylor University, 1311 S 5th St., Waco, TX 77401, United States

JAVIER PARCET: javier.parcet@icmat.es
, *Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM Consejo Superior de Investigaciones Cientí{ }ficas, C/ Nicolás Cabrera 13-15, 28049 Madrid, Spain*

ANALYSIS AND PDE
Vol. 8, No. 3, 2015
dx.doi.org/10.2140/apde.2015.8.747

REFINED AND MICROLOCAL KAKEYA–NIKODYM BOUNDS FOR EIGENFUNCTIONS IN TWO DIMENSIONS

MATTHEW D. BLAIR AND CHRISTOPHER D. SOGGE

We obtain some improved essentially sharp Kakeya–Nikodym estimates for eigenfunctions in two dimensions. We obtain these by proving stronger related microlocal estimates involving a natural decomposition of phase space that is adapted to the geodesic flow.

1. Introduction and main results

Suppose that (M, g) is a two-dimensional compact Riemannian manifold and $\{e_\lambda\}$ are the associated eigenfunctions. That is, if Δ_g is the Laplace–Beltrami operator, we have

$$-\Delta_g e_\lambda(x) = \lambda^2 e_\lambda(x),$$

and we assume throughout that the eigenfunctions are normalized to have L^2-norm one, i.e.,

$$\int_M |e_\lambda|^2 \, dV_g = 1,$$

where dV_g is the volume element.

The purpose of this paper is to obtain essentially sharp estimates that link, in two dimensions, the size of L^p-norms of eigenfunctions with $2 < p < 6$ to their L^2-concentration near geodesics. Specifically, we have the following:

Theorem 1.1. *For every* $0 < \varepsilon_0 \le \frac{1}{2}$, *we have*

$$\|e_\lambda\|_{L^4(M)} \lesssim_{\varepsilon_0} \lambda^{\varepsilon_0/4} \|e_\lambda\|^{1/2}_{L^2(M)} \times |||e_\lambda|||^{1/2}_{KN(\lambda,\varepsilon_0)} \tag{1-1}$$

if

$$|||e_\lambda|||_{KN(\lambda,\varepsilon_0)} = \left(\sup_{\gamma \in \Pi} \lambda^{1/2-\varepsilon_0} \int_{\mathcal{T}_{\lambda^{-1/2+\varepsilon_0}}(\gamma)} |e_\lambda|^2 \, dV \right)^{1/2}. \tag{1-2}$$

Equivalently, if $\varepsilon_0 > 0$, *then there is a* $C = C(\varepsilon_0, M)$ *such that*

$$\|e_\lambda\|_{L^4} \le C\lambda^{1/8} \|e_\lambda\|^{1/2}_{L^2(M)} \times \left(\sup_{\gamma \in \Pi} \int_{\mathcal{T}_{\lambda^{-1/2+\varepsilon_0}}(\gamma)} |e_\lambda|^2 \, dV \right)^{1/4}, \tag{1-3}$$

Blair and Sogge were supported in part by the NSF grants DMS-1301717 and DMS-1361476, respectively.
MSC2010: primary 58J51; secondary 35L20, 42B37.
Keywords: eigenfunctions, Kakeya averages.

and therefore if $\int_M |e_\lambda|^2\,dV = 1$, *for any* $\varepsilon > 0$ *there is a* $C = C(\varepsilon, M)$ *such that*

$$\|e_\lambda\|_{L^4(M)} \le C\lambda^{1/8+\varepsilon} \sup_{\gamma\in\Pi} \|e_\lambda\|^{1/2}_{L^2(\mathcal{T}_{\lambda^{-1/2}}(\gamma))} \le C\lambda^{1/16+\varepsilon} \sup_{\gamma\in\Pi} \|e_\lambda\|^{1/2}_{L^4(\mathcal{T}_{\lambda_{j_k}^{-1/2}}(\gamma))}. \tag{1-4}$$

Here Π denotes the space of unit-length geodesics in M and the last factor in (1-2) involves averages of $|e_\lambda|^2$ over $\lambda^{-1/2+\varepsilon_0}$ tubes about $\gamma \in \Pi$. Also, for simplicity, we are only stating things here and throughout for eigenfunctions, but the results easily extend to quasimodes using results from [Sogge and Zelditch 2014].

Note that if $\varepsilon_0 = \frac{1}{2}$, then (1-1) is equivalent to the eigenfunction estimates from [Sogge 1988]

$$\|e_\lambda\|_{L^4(M)} \lesssim \lambda^{1/8}\|e_\lambda\|_{L^2(M)},$$

which are saturated by highest weight spherical harmonics on the standard two-sphere. We also remark that, up to the factor $\lambda^{\varepsilon_0/4}$, the estimate (1-1) is saturated by both the highest weight spherical harmonics and zonal functions on S^2. This is because the highest weight spherical harmonics are given by the restriction of the harmonic polynomials $\lambda^{1/4}(x_1 + ix_2)^k$, $\lambda = \sqrt{k(k+1)}$ to the unit sphere, while the L^2-normalized zonal functions centered about the north pole on S^2 behave like $\bigl(\lambda^{-1} + \operatorname{dist}(x, \pm(0,0,1))\bigr)^{-1/2}$. See, for instance, [Sogge 1986].

In [Bourgain 2009] (with a slight loss) and in [Sogge 2011], inequalities of the form (1-1) and (1-3) were proved, where the first norm on the right is raised to the $\frac{3}{4}$ power and the second to the $\frac{1}{4}$ power. The inequalities in [Sogge 2011] were not formulated in this way but easily lead to this result. The approach in [Sogge 2011] made inefficient use of the Cauchy–Schwarz inequality to handle the "easy" term (not the bilinear one), which led to the loss. The strategy for proving (1-1) will be to make an angular dyadic decomposition of a bilinear expression and pay close attention to the dependence of the bilinear estimates in terms of the angles, which we shall exploit using a multilayered microlocal decomposition of phase space.

Before turning to the details of the proof, let us record a few simple corollaries of our main estimate.

If $\{a_{\lambda_{j_k}}\}_{k=0}^\infty$ is a sequence depending on a subsequence $\{\lambda_{j_k}\}$ of the eigenvalues of Δ_g, then we say that

$$a_\lambda = o_-(\lambda^\sigma)$$

if there are some $\varepsilon > 0$ and $C < \infty$ such that

$$|a_\lambda| \le C(1+\lambda)^{\sigma-\varepsilon}.$$

Then using Theorem 1.1, we get:

Corollary 1.2. *The following are equivalent*:

$$\|e_{\lambda_{j_k}}\|_{L^4(M)} = o_-(\lambda_{j_k}^{1/8}), \tag{1-5}$$

$$\sup_{\gamma\in\Pi} \|e_{\lambda_{j_k}}\|_{L^4(\mathcal{T}_{\lambda_{j_k}^{-1/2}}(\gamma))} = o_-(\lambda_{j_k}^{1/8}), \tag{1-6}$$

$$\sup_{\gamma\in\Pi} \|e_{\lambda_{j_k}}\|_{L^2(\mathcal{T}_{\lambda_{j_k}^{-1/2}}(\gamma))} = o_-(1). \tag{1-7}$$

Also, if either

$$\sup_{\gamma\in\Pi}\int_\gamma |e_\lambda|^2\,ds = O(\lambda_{j_k}^{\varepsilon}), \quad \textit{for all } \varepsilon>0 \tag{1-8}$$

or

$$\sup_{\gamma\in\Pi}\|e_{\lambda_{j_k}}\|_{L^2(\mathcal{T}_{\lambda_{j_k}^{-1/2}}(\gamma))} = O(\lambda_{j_k}^{-1/4+\varepsilon}), \quad \textit{for all } \varepsilon>0, \tag{1-9}$$

then

$$\|e_{\lambda_{j_k}}\|_{L^4(M)} = O(\lambda_{j_k}^{\varepsilon}), \quad \textit{for all } \varepsilon>0. \tag{1-10}$$

Here, ds denotes the arc length measure on γ.

Clearly (1-5) implies (1-6). Also, (1-7) follows from (1-6) and Hölder's inequality. Since (1-1) shows that (1-7) implies (1-5), the last part of the corollary is also an easy consequence of Theorem 1.1.

Note also that (1-4) says that if $e_{\lambda_{j_k}}$ is a sequence of eigenfunctions with

$$\|e_{\lambda_{j_k}}\|_{L^4(M)} = \Omega(\lambda_{j_k}^{1/8}),$$

then for any ε, there must be a sequence of shrinking geodesic tubes $\{\mathcal{T}_{\lambda_{j_k}^{-1/2}}(\gamma_k)\}$ for which, for some $c=c_\varepsilon>0$, we have

$$\|e_{\lambda_{j_k}}\|_{L^4(\mathcal{T}_{\lambda_{j_k}^{-1/2}}(\gamma_k))} \ge c\,\lambda_{j_k}^{1/8-\varepsilon}.$$

In other words, up to a factor of $\lambda^{-\varepsilon}$ for any $\varepsilon>0$, they fit the profile of the highest weight spherical harmonics by having maximal L^4-mass on a sequence of shrinking $\lambda^{-1/2}$ tubes.

Like in Bourgain's estimate, (1-1) involves a slight loss, but this is not so important in view of the above application. In a later work we hope to show that (1-1) holds without this loss (in other words with $\varepsilon_0=0$), which should mainly involve refining the $S_{1/2,1/2}$ microlocal arguments that are to follow. Note that, because of the zonal functions on S^2, this result would be sharp.

This paper is organized as follows. In Section 2 we shall introduce a microlocal Kakeya–Nikodym norm and an inequality involving it, (2-14), which implies (1-1). This norm is associated to a decomposition of phase space which is naturally associated to the geodesic flow on the cosphere bundle. In particular, each term in the decomposition will involve bump functions which are supported in tubular neighborhoods of unit geodesics in S^*M. This decomposition and the resulting square function arguments are similar to the earlier ones in the joint paper of Mockenhaupt, Seeger and the second author [Mockenhaupt et al. 1993], but there are some differences and new technical issues that must be overcome. We do this and prove our microlocal Kakeya–Nikodym estimate in Section 3. There, after some pseudodifferential arguments, we reduce matters to an oscillatory integral estimate which is a technical variation on the classical one in Hörmander [1973], which was the main step in his proof of the Carleson–Sjölin theorem [1972]. The result which we need does not directly follow from the results in [Hörmander 1973]; however, we can prove it by adapting Hörmander's argument and using Gauss's lemma. After doing this, in Section 4 we shall see how our results are in some sense related to Zygmund's theorem [1974] saying that in two dimensions, eigenfunctions on the standard torus have bounded L^4-norms. Specifically, we shall see there that if we could obtain the endpoint version of (1-1), we would be able to recover Zygmund's theorem with no loss if we also knew a conjectured result that arcs on λS^1 of length $\lambda^{1/2}$ contain a uniformly bounded number of lattice points.

In a later paper with S. Zelditch, we hope to strengthen our results and also extend them to higher dimensions, as well as to present applications in the spirit of [Sogge and Zelditch 2012] of the microlocal bounds which we obtain. The current authors would like to thank S. Zelditch for a number of stimulating discussions.

2. Microlocal Kakeya–Nikodym norms

As in [Sogge 2011; Sogge 1993, §5.1], we use the fact that we can use a reproducing operator to write $e_\lambda = \chi_\lambda f = \rho(\lambda - \sqrt{\Delta_g})e_\lambda$, for $\rho \in \mathcal{S}$ satisfying $\rho(0) = 1$, where, if $\operatorname{supp}\hat{\rho} \subset (1, 2)$, we also have modulo $O(\lambda^{-N})$ errors (see [Sogge 1993, Lemma 5.1.3])

$$\chi_\lambda f(x) = \frac{1}{2\pi}\int \hat{\rho}(t)e^{i\lambda t}(e^{-it\sqrt{\Delta_g}}f)(x)\,dt = \lambda^{1/2}\int e^{i\lambda\psi(x,y)}a_\lambda(x, y)f(y)\,dV(y), \tag{2-1}$$

where

$$\psi(x, y) = d_g(x, y) \tag{2-2}$$

is the Riemannian distance function, and if, as we may, we assume that the injectivity radius is 10 or more, a_λ belongs to a bounded subset of C^∞ and satisfies

$$a_\lambda(x, y) = 0, \quad \text{if } d_g(x, y) \notin (1, 2). \tag{2-3}$$

Thus, in order to prove (1-1), it suffices to work in a local coordinate patch and show that if a is smooth and satisfies the support assumptions in (2-3), if $0 < \delta < \frac{1}{10}$ is small but fixed, and if

$$x_0 = (0, y_0), \quad \tfrac{1}{2} < y_0 < 4$$

is also fixed, then

$$\left\|\lambda^{1/2}\int e^{i\lambda\psi(x,y)}a(x, y)f(y)\,dy\right\|^2_{L^4(B(0,\delta))} \lesssim_{\varepsilon_0} \lambda^{\varepsilon_0/2}\|f\|_{L^2} \times |||f|||_{KN(\lambda,\varepsilon_0)}, \quad \text{if } \operatorname{supp} f \subset B(x_0, \delta). \tag{2-4}$$

Here $B(x, \delta)$ denotes the δ-ball about x in our coordinates. We may assume that in our local coordinate system the line segment $(0, y)$, $|y| < 4$ is a geodesic.

In order to prove (2-4) we also need to define a microlocal version of the above Kakeya–Nikodym norm. We first choose $0 \le \beta \in C_0^\infty(\mathbb{R}^2)$ satisfying

$$\sum_{\nu\in\mathbb{Z}^2}\beta(z+\nu) = 1 \quad \text{and} \quad \operatorname{supp}\beta \subset \{x \in \mathbb{R}^2 : |x| \le 2\}. \tag{2-5}$$

To use this bump function, let $\Phi_t(x, \xi) = (x(t), \xi(t))$ denote the geodesic flow on the cotangent bundle. Then if (x, ξ) is a unit cotangent vector with $x \in B(x_0, \delta)$ and $|\xi_1| < \delta$, with δ small enough, it follows that there is a unique $0 < t < 10$ such that $x(t) = (s, 0)$ for some $s(x, \xi)$. If $\xi(t) = (\xi_1(t), \xi_2(t))$ for this t, it follows that $\xi_2(t)$ is bounded from below. Let us then set $\varphi(x, \xi) = (s(x, \xi), \xi_1(t)/|\xi(t)|)$. Note that φ then is a smooth map from such unit cotangent vectors to $\mathbb{R}^2$. Also, φ is constant on the orbit of Φ. Therefore, $|\varphi(x,\xi)-\varphi(y,\eta)|$ can be thought of as measuring the distance from the geodesic in our coordinate patch through (x,ξ) to that of the one through (y,η).

Let $\alpha(x)$ be a nonnegative C_0^∞ function which is one in $B\big(x_0,\frac{3}{2}\delta\big)$ and zero outside of $B(x_0,2\delta)$. Given $\theta=2^{-k}$ with $\lambda^{-1/2}\le\theta\le 1$ and $\nu\in\mathbb{Z}^2$, let $\Upsilon\in C^\infty(\mathbb{R})$ satisfy

$$\Upsilon(s)=1,\ s\in[c,c^{-1}],\quad \Upsilon(s)=0,\ s\notin\left[\frac{c}{2},2c^{-1}\right], \tag{2-6}$$

for some $c>0$ to be specified later. We then put

$$Q_\theta^\nu(x,\xi)=\alpha(x)\beta\big(\theta^{-1}\varphi(x,\xi)+\nu\big)\Upsilon(|\xi|/\lambda). \tag{2-7}$$

This is a function of unit cotangent vectors, and we also denote its homogeneous of degree zero extension to the cotangent bundle with the zero section removed by $Q_\theta^\nu(x,\xi)$, $\xi\neq 0$, and the resulting pseudodifferential operator by $Q_\theta^\nu(x,D)$. Then if f is as in (2-4), we define its microlocal Kakeya–Nikodym norm corresponding to frequency λ and angle $\theta_0=\lambda^{-1/2+\varepsilon_0}$ to be

$$|||f|||_{MKN(\lambda,\varepsilon_0)}=\sup_{\theta_0\le\theta\le 1}\Big(\sup_{\nu\in\mathbb{Z}^2}\theta^{-1/2}\|Q_\theta^\nu(x,D)f\|_{L^2(\mathbb{R}^2)}\Big)+\|f\|_{L^2(\mathbb{R}^2)},\quad \theta_0=\lambda^{-1/2+\varepsilon_0}. \tag{2-8}$$

Note that

$$\sup_{\nu\in\mathbb{Z}^2}\theta^{-1/2}\|Q_\theta^\nu(x,D)f\|_{L^2(\mathbb{R}^2)}$$

measures the maximal microlocal concentration of f about all unit geodesics in the scale of θ. This is because if we consider the restriction of Q_θ^ν to unit cotangent vectors and if $Q_\theta^\nu(x,\xi)\neq 0$, then supp Q_θ^ν is contained in an $O(\theta)$ tube in the space of unit cotangent vectors about the orbit $t\to\Phi_t(x,\xi)$.

Let us collect a few facts about these pseudodifferential operators. First, the Q_θ^ν belong to a bounded subset of $S^0_{1/2+\varepsilon_0,1/2-\varepsilon_0}$ (pseudodifferential operators of order zero and type $(\frac{1}{2}+\varepsilon_0,\frac{1}{2}-\varepsilon_0)$), if $\lambda^{-1/2+\varepsilon_0}\le\theta\le 1$, with $\varepsilon_0>0$ fixed. Therefore, there is a uniform constant C_{ε_0} such that

$$\|Q_\theta^\nu(x,D)g\|_{L^2}\le C_{\varepsilon_0}\|g\|_{L^2},\quad \lambda^{-1/2+\varepsilon_0}\le\theta\le 1. \tag{2-9}$$

Similarly, if $P_\theta^\nu=(Q_\theta^\nu)^*\circ Q_\theta^\nu$ for such θ, then by (2-5), $\sum_\nu P_\theta^\nu$ belongs to a bounded subset of $S^0_{1/2+\varepsilon_0,1/2-\varepsilon_0}$, and so we also have the uniform bounds

$$\Big\|\sum_{\nu\in\mathbb{Z}^2}P_\theta^\nu(x,D)g\Big\|_{L^2}\le C_{\varepsilon_0}\|g\|_{L^2},\quad \lambda^{-1/2+\varepsilon_0}\le\theta\le 1. \tag{2-10}$$

We can relate the microlocal Kakeya–Nikodym norm to the Kakeya–Nikodym norm if we realize that if the $\delta>0$ above is small enough, then there is a unit length geodesic γ_ν such that $Q_\theta^\nu(x,\xi)=0$ for $x\notin\mathcal{T}_{C\theta_\nu}(\gamma)$, with C a uniform constant. As a result, since $Q_\theta^\nu(x,\xi)=0$ if $|\xi|$ is not comparable to λ, we can improve (2-9) and deduce that for every $N=1,2,\dots$, there is a uniform constant C' such that

$$\|Q_\theta^\nu(x,D)g\|_{L^2}\le C_{\varepsilon_0}\bigg(\int_{\mathcal{T}_{C'\theta}(\gamma_\nu)}|g|^2\,dy\bigg)^{1/2}+C_N\lambda^{-N}\|g\|_{L^2},\quad \lambda^{-1/2+\varepsilon_0}\le\theta\le 1, \tag{2-11}$$

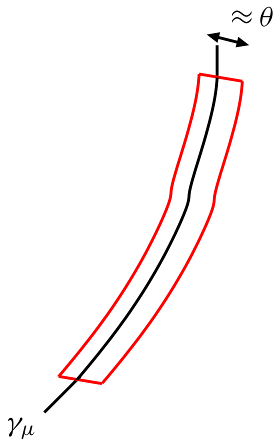

Figure 1. $\mathcal{T}_{C'\theta(\gamma_\nu)}$.

since the kernel $K_\theta^\nu(x, y)$ of $Q_\theta^\nu(x, D)$ is $O(\lambda^{-N})$ for any N if y is not in $\mathcal{T}_{C'\theta}(\gamma_\nu)$, with C' sufficiently large but fixed. (See Figure 1.) Since

$$\theta^{-1/2}\left(\int_{\mathcal{T}_{C'\theta}(\gamma_\nu)} |g|^2\,dy\right)^{1/2} \lesssim \sup_{\gamma\in\Pi}\left(\theta_0^{-1}\int_{\mathcal{T}_{\theta_0}(\gamma)} |g|^2\,dy\right)^{1/2}, \quad \lambda^{-1/2+\varepsilon_0} = \theta_0 \le \theta \le 1,$$

we have

$$\sup_{\nu\in\mathbb{Z}^2} \theta^{-1/2}\|Q_\theta^\nu(x, D)f\|_{L^2(\mathbb{R}^2)} \le C_{\varepsilon_0}|||g|||_{KN(\lambda,\varepsilon_0)}, \quad \lambda^{-1/2+\varepsilon_0} \le \theta \le 1, \tag{2-12}$$

meaning that we can dominate the microlocal Kakeya–Nikodym norm by the Kakeya–Nikodym norm.

From this, we conclude that we would have (2-4) if we could show

$$\left\|\int \lambda^{1/2} e^{i\lambda\psi(x,y)} a(x, y) f(y)\,dy\right\|^2_{L^4(B(0,\delta))} \lesssim_{\varepsilon_0} \lambda^{\varepsilon_0/2}\|f\|_{L^2} \times |||f|||_{MKN(\lambda,\varepsilon_0)}, \quad \text{if supp } f \subset B(x_0, \delta). \tag{2-13}$$

We note also that since $\chi_\lambda e_\lambda = e_\lambda$, this inequality of course yields the following microlocal strengthening of Theorem 1.1:

Theorem 2.1. *For every* $0 < \varepsilon_0 \le \frac{1}{2}$, *we have*

$$\|e_\lambda\|_{L^4(M)} \lesssim_{\varepsilon_0} \lambda^{\varepsilon_0/4}\|e_\lambda\|^{1/2}_{L^2(M)} \times |||e_\lambda|||^{1/2}_{MKN(\lambda,\varepsilon_0)}, \tag{2-14}$$

if $|||e_\lambda|||_{MKN(\lambda,\varepsilon_0)}$ *is as in* (2-8).

3. Proof of the refined two-dimensional microlocal Kakeya–Nikodym estimates

Let us now prove the estimates in (2-13). We shall follow arguments from §6 of [Mockenhaupt et al. 1993].

We first note that if $\operatorname{supp} f \subset B(x_0, \delta)$ as in (2-4), and if

$$\theta_0 = \lambda^{-1/2+\varepsilon_0} \tag{3-1}$$

with $\varepsilon_0 > 0$ fixed,

$$\chi_\lambda f = \sum_{\nu \in \mathbb{Z}^2} \chi_\lambda\big(Q^\nu_{\theta_0}(x, D) f\big) + R_\lambda f,$$

where, if $c > 0$ in (2-6) is small enough, and $N = 1, 2, 3, \dots,$

$$\|R_\lambda f\|_{L^\infty} \lesssim \lambda^{-N} \|f\|_{L^2}.$$

Therefore, in order to prove (2-4), it suffices to show that

$$\Big\| \sum_{\nu, \nu' \in \mathbb{Z}^2} \chi_\lambda Q^\nu_{\theta_0} f \, \chi_\lambda Q^{\nu'}_{\theta_0} f \Big\|_{L^2} \lesssim_{\varepsilon_0} \lambda^{\varepsilon_0/2} \|f\|_{L^2} \times |||f|||_{MKN(\lambda,\varepsilon_0)}. \tag{3-2}$$

We split the sum on the left based on the size of $|\nu - \nu'|$. Indeed, the left side of (3-2) is dominated by

$$\Big\| \sum_\nu (\chi_\lambda Q^\nu_{\theta_0} f)^2 \Big\|_{L^2} + \sum_{\ell=1}^\infty \Big\| \sum_{|\nu-\nu'| \in [2^\ell, 2^{\ell+1})} \chi_\lambda Q^\nu_{\theta_0} f \, \chi_\lambda Q^{\nu'}_{\theta_0} f \Big\|_{L^2}. \tag{3-3}$$

The square of the first term in (3-3) is

$$\sum_{\nu, \nu'} \int (\chi_\lambda Q^\nu_{\theta_0} f)^2 \overline{(\chi_\lambda Q^{\nu'}_{\theta_0} f)^2} \, dx.$$

Next we need an orthogonality result, similar to Lemma 6.7 in [Mockenhaupt et al. 1993], which says that if A is large enough we have

$$\sum_{|\nu-\nu'| \ge A} \Big| \int (\chi_\lambda Q^\nu_{\theta_0} f)^2 \overline{(\chi_\lambda Q^{\nu'}_{\theta_0} f)^2} \, dx \Big| \lesssim_{\varepsilon_0, N} \lambda^{-N} \|f\|^4_{L^2}. \tag{3-4}$$

We shall postpone the proof of this result until the end of the section, when we will have recorded the information about the kernels of $\chi_\lambda Q^\nu_\theta$ that will be needed for the proof.

Since by [Sogge 1988],

$$\|\chi_\lambda\|_{L^2 \to L^4} = O(\lambda^{1/8}),$$

if we use (3-4) we conclude that the first term in (3-3) is majorized by (2-10) and (2-12):

$$\begin{aligned} \lambda^{1/2} \sum_\nu \|Q^\nu_{\theta_0} f\|^2_{L^2} \|Q^\nu_{\theta_0} f\|^2_{L^2} + \lambda^{-N} \|f\|^4_{L^2} &\lesssim \lambda^{1/2} \|f\|^2_{L^2} \times \sup_{\nu \in \mathbb{Z}^2} \|Q^\nu_{\theta_0} f\|^2_{L^2} + \lambda^{-N} \|f\|^4_{L^2} \\ &= \lambda^{\varepsilon_0} \|f\|^2_{L^2} \times \lambda^{1/2-\varepsilon_0} \sup_{\nu \in \mathbb{Z}^2} \|Q^\nu_{\theta_0} f\|^2_{L^2} + \lambda^{-N} \|f\|^4_{L^2}. \end{aligned} \tag{3-5}$$

Therefore, the first term in (3-3) satisfies the desired bounds.

Using (2-12) again, the proof of (2-13) and hence (2-4) would be complete if we could estimate the other terms in (3-3) and show that

$$\Big\| \sum_{|\nu-\nu'|\in[2^\ell,2^{\ell+1})} \chi_\lambda Q^\nu_{\theta_0} f\, \chi_\lambda Q^{\nu'}_{\theta_0} f \Big\|^2_{L^2} \lesssim_{\varepsilon_0} \|f\|^2_{L^2} \times (2^\ell\theta_0)^{-1} \sup_{\nu\in\mathbb{Z}^2} \|Q^\nu_{2^\ell\theta_0} f\|^2_{L^2} + \lambda^{-N}\|f\|^4_{L^2}. \tag{3-6}$$

Note that if $2^\ell\theta_0 \gg 1$, the left side of (3-6) vanishes and thus, as in (2-12), we are just considering $\ell\in\mathbb{N}$ satisfying $1\le 2^\ell \le \lambda^{1/2-\varepsilon_0}$. In proving this, we may assume that ℓ is larger than a fixed constant, since the bound for small ℓ (with an extra factor of λ^{ε_0} on the right) follows from what we just did. We can handle the sum over ℓ in (3-3) due to the fact that the right side of (3-6) does not include a factor λ^{ε_0}.

We now turn to estimating the nondiagonal terms in (3-3). We first note that by (2-5),

$$\chi_\lambda Q^\nu_{\theta_0} f = \sum_{\mu\in\mathbb{Z}^2} \chi_\lambda Q^\mu_\theta Q^\nu_{\theta_0} f + O_N(\lambda^{-N}\|f\|_2), \quad \text{if } \operatorname{supp} f\subset B(x_0,\delta).$$

Furthermore, if, as we may, we assume that $\ell\in\mathbb{N}$ is sufficiently large, then given $N_0\in\mathbb{N}$, there are fixed constants $c_0>0$ and $N_1<\infty$ (with c_0 depending only on N_0 and the cutoff β in the definition of these pseudodifferential operators) such that if

$$\theta_\ell = \theta_0 2^\ell,$$

then

$$\sum_{|\nu-\nu'|\in[2^\ell,2^{\ell+1})} \chi_\lambda Q^\nu_{\theta_0} f\, \chi_\lambda Q^{\nu'}_{\theta_0} f$$
$$= \sum_{\{\mu,\mu'\in\mathbb{Z}^2: N_0\le|\mu-\mu'|\le N_1\}} \sum_{|\nu-\nu'|\in[2^\ell,2^{\ell+1})} \chi_\lambda Q^\mu_{c_0\theta_\ell} Q^\nu_{\theta_0} f\, \chi_\lambda Q^{\mu'}_{c_0\theta_\ell} Q^{\nu'}_{\theta_0} f + O_N(\lambda^{-N}\|f\|^2_{L^2}), \tag{3-7}$$

for each $N\in\mathbb{N}$. Also, given $\mu\in\mathbb{Z}^2$, there is a $\nu_0(\mu)\in\mathbb{Z}^2$ such that

$$\|Q^\mu_{c_0\theta_\ell} Q^\nu_{\theta_0} f\|_{L^2} \le C_N\lambda^{-N}\|f\|_{L^2}, \quad \text{if } |\nu-\nu_0(\mu)|\ge C2^\ell,$$

for some uniform constant C. If $|\mu-\mu'|\le N_1$, then $|\nu_0(\mu)-\nu_0(\mu')|\le C2^\ell$ for some uniform constant C. Since $\|(Q^{\nu'}_\theta)^*\circ Q^\nu_\theta\|_{L^2\to L^2} = O(\lambda^{-N})$ for every N if $|\nu-\nu'|$ is larger than a fixed constant, it follows that

$$\iint \Big| \sum_{|\nu_0(\mu)-\nu|,|\nu_0(\mu')-\nu'|\le C2^\ell} \sum_{|\nu-\nu'|\in[2^\ell,2^{\ell+1})} Q^\nu_{\theta_0} f(x) Q^{\nu'}_{\theta_0} f(y) \Big|^2 dx\,dy$$
$$\lesssim \sum_{|\nu-\nu_0(\mu)|,|\nu'-\nu_0(\mu)|\le C'2^\ell} \|Q^\nu_{\theta_0} f\|^2_{L^2} \|Q^{\nu'}_{\theta_0} f\|^2_{L^2} + O_N(\lambda^{-N}\|f\|^2_{L^2}), \quad \text{if } |\mu-\mu'|\le C_0, \tag{3-8}$$

for every N if C' is a sufficiently large but fixed constant. Also, using (2-10), we deduce that

$$\sum_{\mu\in\mathbb{Z}^2} \sum_{|\nu_0(\mu)-\nu|\le C'2^\ell} \|Q^\nu_{\theta_0} f\|^2_{L^2} \lesssim \|f\|^2_{L^2}.$$

We clearly also have

$$\sum_{|\nu(\mu)-\nu'|\le C'2^\ell} \|Q_{\theta_0}^{\nu'} f\|_{L^2}^2 \lesssim \sup_{\mu\in\mathbb{Z}^2} \|Q_{2^\ell\theta}^{\mu} f\|_{L^2}^2.$$

Using these two inequalities and (3-8), we deduce that

$$\sum_{|\mu-\mu'|\le N_1} \Bigl\| \sum_{|\nu_0(\mu)-\nu|,|\nu_0(\mu')-\nu'|<C2^\ell} \; \sum_{|\nu-\nu'|\in[2^\ell,2^{\ell+1})} Q_{\theta_0}^{\nu} f(x) Q_{\theta_0}^{\nu'} f(y) \Bigr\|_{L^2(dx\,dy)} \lesssim \|f\|_{L^2} \times \sup_{\mu\in\mathbb{Z}^2} \|Q_{2^\ell\theta}^{\mu} f\|_{L^2} + O_N(\lambda^{-N}\|f\|_{L^2}^2). \tag{3-9}$$

In addition to (3-4), we shall need another orthogonality result whose proof we postpone until the end of the section, which says that whenever θ is larger than a fixed positive multiple of θ_0 in (3-1) and N_1 is fixed,

$$\Bigl| \int \bigl(\chi_\lambda Q_\theta^{\mu} g_1\, \chi_\lambda Q_\theta^{\mu'} g_2\bigr) \overline{\bigl(\chi_\lambda Q_\theta^{\tilde\mu} g_3\, \chi_\lambda Q_\theta^{\tilde\mu'} g_4\bigr)}\, dx \Bigr| \lesssim_N \lambda^{-N} \prod_{j=1}^4 \|g_j\|_{L^2}, \quad \text{if } |\mu-\tilde\mu|+|\mu'-\tilde\mu'| \ge C \text{ and } |\mu-\mu'|,\, |\tilde\mu-\tilde\mu'| \le N_1, \tag{3-10}$$

for every $N = 1, 2, \dots$, with C being a sufficiently large uniform constant (depending on N_1 of course).

Using (3-9) and (3-10), we conclude that we would have (3-6) (and consequently (2-4)) if we could prove the following:

Proposition 3.1. *Let*

$$(T_{\lambda,\theta}^{\mu,\mu'} F)(x) = \iint (\chi_\lambda Q_\theta^{\mu})(x,y)(\chi_\lambda Q_\theta^{\mu'})(x,y') F(y,y')\, dy\, dy', \tag{3-11}$$

where

$$(\chi_\lambda Q_\theta^{\mu})(x,y)$$

denotes the kernel of $\chi_\lambda Q_\theta^{\mu}$. *Then if* $\delta > 0$ *is sufficiently small and if* θ *is larger than a fixed positive constant times* θ_0 *in* (3-1) *and if* $N_0 \in \mathbb{N}$ *is sufficiently large and if* $N_1 > N_0$ *is fixed*, *we have*

$$\|T_{\lambda,\theta}^{\mu,\mu'} F\|_{L^2(B(0,\delta))} \lesssim_{\varepsilon_0} \theta^{-1/2} \|F\|_{L^2}, \quad \text{if } N_0 \le |\mu-\mu'| \le N_1, \qquad F(y,y') = 0, \quad \text{if } (y,y') \notin B(x_0,2\delta)\times B(x_0,2\delta). \tag{3-12}$$

To prove this we shall need some information about the kernel of $\chi_\lambda Q_\theta^{\mu}$. By (2-7), the kernel is highly concentrated near the geodesic in M

$$\gamma_\mu = \bigl\{x_\mu(t) : -2 \le t \le 2,\ \Phi_t(x_\mu,\xi_\mu) = (x_\mu(t),\xi_\mu(t)),\ \theta^{-1}\varphi(x_\mu,\xi_\mu)+\mu = 0\bigr\}, \tag{3-13}$$

which corresponds to Q_θ^{μ}. We also will exploit the oscillatory behavior of the kernel near γ_μ.

Specifically, we require the following:

Lemma 3.2. *Let* $\theta \in \bigl[C_0\lambda^{-1/2+\varepsilon_0}, \frac{1}{2}\bigr]$, *where* C_0 *is a sufficiently large fixed constant*, *and*, *as above*, $\varepsilon_0 > 0$. *Then there is a uniform constant* C *such that for each* $N = 1, 2, 3, \dots$, *we have*

$$|(\chi_\lambda Q_\theta^\mu)(x, y)| \le C_N \lambda^{-N}, \quad \text{if } x \notin \mathcal{T}_{C\theta}(\gamma_\mu) \text{ or } y \notin \mathcal{T}_{C\theta}(\gamma_\mu). \tag{3-14}$$

Furthermore,

$$(\chi_\lambda Q_\theta^\mu)(x, y) = \lambda^{1/2} e^{i\lambda d_g(x,y)} a_{\mu,\theta}(x, y) + O_N(\lambda^{-N}), \tag{3-15}$$

where one has the uniform bounds

$$|\nabla_y^\alpha a_{\mu,\theta}(x, y)| \le C_\alpha \theta^{-|\alpha|}, \tag{3-16}$$

$$|\partial_t^j a_{\mu,\theta}(x, x_\mu(t))| \le C_j, \quad x \in \gamma_\mu, \tag{3-17}$$

if, *as in* (3-13), $\{x_\mu(t)\} = \gamma_\mu$.

Proof. To prove the lemma it is convenient to choose Fermi normal coordinates so that the geodesic becomes the segment $\{(0, s) : |s| \le 2\}$. Let us also write θ as

$$\theta = \lambda^{-1/2+\delta},$$

where, because of our assumptions, $c_1 \le \delta \le \frac{1}{2}$ for an appropriate $c_1 > 0$. Then in these coordinates, $Q_\theta^\mu(x, D)$ has symbol satisfying

$$q_\theta^\mu(x, \xi) = 0, \quad \text{if } \bigl|\xi_1/|\xi|\bigr| \ge C\lambda^{-1/2+\delta},\ |x_1| \ge C\lambda^{-1/2+\delta}, \text{ or } |\xi|/\lambda \notin [C^{-1}, C], \tag{3-18}$$

for some uniform constant C, and, additionally,

$$\bigl|\partial_{x_1}^j \partial_{x_2}^k \partial_{\xi_1}^l \partial_{\xi_2}^m q_\theta^\mu(x, \xi)\bigr| \le C_{j,k,l,m}(1 + |\xi|)^{j(1/2-\delta)-l(1/2+\delta)-m}. \tag{3-19}$$

Next we recall that $\chi_\lambda = \rho(\lambda - \sqrt{-\Delta_g})$, where $\rho \in \mathcal{S}(\mathbb{R})$ satisfies $\hat{\rho} \subset (1, 2)$, and that the injectivity radius of (M, g) is ten or more. Therefore, we can use Fourier integral parametrices for the wave equation to see that the kernel of χ_λ is of the form

$$\chi_\lambda(x, y) = \iint e^{iS(t,x,\xi)-iy\cdot\xi+it\lambda} \hat{\rho}(t)\alpha(t, x, y, \xi)\, d\xi\, dt,$$

where $\alpha \in S^1_{1,0}$, and S is homogeneous of degree one in ξ and is a generating function for the canonical relation for the half wave group $e^{-it\sqrt{-\Delta_g}}$. Thus,

$$\partial_t S(t, x, \xi) = -p(x, \nabla_x S(t, x, \xi)), \quad S(0, x, \xi) = x \cdot \xi. \tag{3-20}$$

Let $\tilde{\Phi}_t(x, \xi)$ denote the Hamiltonian flow generated by $p(x, \xi)$, which is homogeneous of degree one in ξ and agrees with the geodesic flow $\Phi_t(x, \xi)$ when restricted to unit cotangent vectors. The phase $S(t, x, \xi)$ also satisfies

$$\tilde{\Phi}_t(x, \nabla_x S) = (\nabla_\xi S, \xi). \tag{3-21}$$

Furthermore,

$$\det \frac{\partial S}{\partial x \partial \xi} \ne 0. \tag{3-22}$$

By (3-18), (3-19), and the proof of the Kohn–Nirenberg theorem, we have that

$$
\begin{aligned}
(\chi_\lambda Q_\theta^\mu)(x,y) &= \iint e^{iS(t,x,\xi)-iy\cdot\xi+i\lambda t}\hat\rho(t)q(t,x,y,\xi)\,d\xi\,dt + O(\lambda^{-N}),\\
&= \lambda^2 \iint e^{i\lambda(S(t,x,\xi)-y\cdot\xi+t)}\hat\rho(t)q(t,x,y,\lambda\xi)\,d\xi\,dt + O(\lambda^{-N}),
\end{aligned}
\tag{3-23}
$$

where for all t in the support of $\hat\rho$,

$$
q(t,x,y,\xi)=0 \quad \text{if } \bigl|\xi_1/|\xi|\bigr| \ge C\lambda^{-1/2+\delta},\ |x_1|\ge C\lambda^{-1/2+\delta},\ \text{ or } |\xi|/\lambda\notin[C^{-1},C], \tag{3-24}
$$

with C as in (3-19), and also

$$
\bigl|\partial_{x_1}^j\partial_{x_2}^k\partial_{\xi_1}^l\partial_{\xi_2}^m q(t,x,y,\xi)\bigr| \le C_{j,k,l,m}(1+|\xi|)^{j(1/2-\delta)-l(1/2+\delta)-m}. \tag{3-25}
$$

Let us now prove (3-14). We have the assertion if $y\notin\mathcal{T}_{C\lambda^{-1/2+\delta}}(\gamma_\mu)$ by (3-24). To prove that remaining part of (3-24) which says that this is also the case when x is not in such a tube, we note that by (3-21), if $d_g(x_0,y_0)=t_0$ and $x_0,y_0\in\gamma_\mu$, then

$$
\nabla_\xi(S(t_0,x_0,\xi)-y_0\cdot\xi)=0, \quad \text{if } \xi_1=0.
$$

By (3-22), we then have

$$
\bigl|\nabla_\xi(S(t_0,x,\xi)-y_0\cdot\xi)\bigr| \approx d_g(x,x_0), \quad \text{if } \xi_1=0.
$$

We deduce from this that if $|\xi_1|/|\xi|\le C\lambda^{-1/2+\delta}$, $|y_1|\le C\lambda^{-1/2+\delta}$, and $|\xi|\in[C^{-1},C]$, then there are a $c_0>0$ and a $C_0<\infty$ such that

$$
\bigl|\nabla_\xi(S(t_0,x,\xi)-y\cdot\xi)\bigr| \ge c_0\lambda^{-1/2+\delta}, \quad \text{if } x\notin\mathcal{T}_{C_0\lambda^{-1/2+\delta}}(\gamma_\mu).
$$

From this we obtain the remaining part of (3-14) via a simple integration by parts argument if we use the support properties (3-24) and size estimates (3-25) of $q(t,x,y,\xi)$. We note that every time we integrate by parts in ξ we gain by $\lambda^{-2\delta}$, which implies (3-14) since q vanishes unless $|\xi|\approx\lambda$ and δ is bounded below by a fixed positive constant.

To finish the proof of the lemma and obtain (3-15)–(3-17), we note that if we let

$$
\Psi(t,x,y,\xi)=S(t,x,\xi)-y\cdot\xi+t
$$

denote the phase function of the second oscillatory integral in (3-23), then at a stationary point where

$$
\nabla_{\xi,t}\Psi=0,
$$

we must have $\Psi=d_g(x,y)$, due to the fact that $S(t,x,\xi)-y\cdot\xi=0$ and $t=d_g(x,y)$ at points where the ξ-gradient vanishes. Additionally, it is not difficult to check that the mixed Hessian of the phase satisfies

$$
\det\left(\frac{\partial^2\Psi}{\partial(\xi,t)\partial(\xi,t)}\right)\neq 0
$$

on the support of the integrand. This follows from the proof of Lemma 5.1.3 of [Sogge 1993]. Moreover, since modulo $O(\lambda^{-N})$ error terms $(\chi_\lambda Q_\theta^\mu)(x,y)$ equals

$$\lambda^2 \iint e^{i\lambda\Psi}\hat{\rho}(t)q(t,x,y,\lambda\xi)\,d\xi\,dt, \tag{3-26}$$

we obtain (3-15)–(3-16) by the proof of this result if we use the stationary phase and (3-24)–(3-25). Indeed, by (3-21), (3-26) has a stationary phase expansion (see [Hörmander 2003, Theorem 7.7.5]), where the leading term is a fixed constant times

$$\lambda^{1/2}e^{i\lambda t}q(t,x,y,\lambda\xi), \quad \text{if } t=d_g(x,y) \text{ and } \tilde{\Phi}_{-t}(y,\xi)=(x,\nabla_x S(t,x,\xi)). \tag{3-27}$$

From this, we see that the leading term in the asymptotic expansion must satisfy (3-16), and subsequent terms in the expansion will satisfy better estimates, where the right-hand side involves increasing negative powers of $\lambda^{2\delta}$ (by [Hörmander 2003, (7.7.1)] and (3-25)), from which we deduce that (3-16) must be valid. Since $\xi_1=0$ and $p(y,\xi)=1$ (by (3-21)) in (3-27) when $x,y\in\gamma_\mu$, we similarly deduce from (3-25) that the leading term in the stationary phase expansion must satisfy (3-17), and since the other terms satisfy better bounds involving increasing powers of $\lambda^{-2\delta}$, we similarly obtain (3-17), which completes the proof of the lemma. □

Let us now collect some simple consequences of Lemma 3.2. First, in addition to (3-14), the kernel $(\chi_\lambda Q_\theta^\mu)(x,y)$ is also $O(\lambda^{-N})$ unless the distance between x and y is comparable to one by (2-3). From this we deduce that if $N_0\in\mathbb{N}$ is sufficiently large,

$$(\chi_\lambda Q_\theta^\mu)(x,y)(\chi_\lambda Q_\theta^{\mu'})(x,y')=O(\lambda^{-N}),$$
$$\text{unless Angle}(x;y,y')\in[\theta,C_2\theta] \text{ and } x,y,y'\in\mathcal{T}_{C_2\theta}(\gamma_\mu), \text{ if } |\mu-\mu'|\in[N_0,N_1], \tag{3-28}$$

if Angle(x,y,y') denotes the angle at x of the geodesic connecting x and y and the one connecting x and y', and where $C_2=C_2(N_1)$.

This is because in this case, if $x\in\mathcal{T}_{C\theta}(\gamma_\mu)\cap\mathcal{T}_{C\theta}(\gamma_{\mu'})$, then the tubes must be disjoint at a distance bounded below by a fixed positive multiple of θ if N_0 is large enough, and in this region their separation is bounded by a fixed constant times θ if N_1 is fixed; see Figure 2.

To exploit this key fact, as above, let us choose Fermi normal coordinates (see [Gray 2004, Chapter 2]) about γ_μ so that the geodesic becomes the segment $\{(0,s): |s|\le 2\}$. Then, as in (2-2), let

$$\psi(x;y)=d_g\big((x_1,x_2),(y_1,y_2)\big)$$

be the Riemannian distance function written in these coordinates. Then if x,y,y' are close to this segment and if the distances between x and y and x and y' are both comparable to 1 and if, as well, y is close to y', it follows from Gauss's lemma that

$$\text{Angle}\big(x;(y_1,y_2),(y_1',y_2')\big)\approx\left|\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\psi(x,y)-\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\psi(x,y')\right|. \tag{3-29}$$

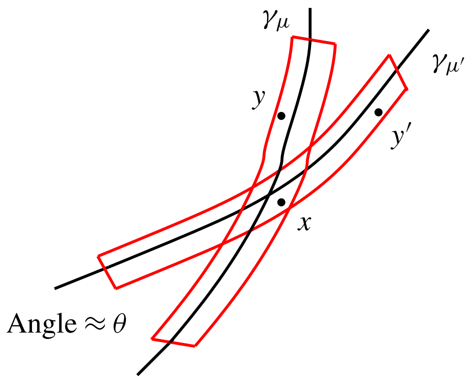

Figure 2. θ-tubes intersecting at angle $\geq N_0\theta$.

As a result, by (3-28), there must be a constant $c_0 > 0$ such that

$$(\chi_\lambda Q_\theta^\mu)(x,y)(\chi_\lambda Q_\theta^{\mu'})(x,y') = O(\lambda^{-N}),$$
$$\text{if } \left|\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\psi(x,y) - \frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\psi(x,y')\right| \leq c_0\theta \text{ and } |\mu-\mu'| \in [N_0, N_1], \tag{3-30}$$

with, as above, $N_0 \in \mathbb{N}$ sufficiently large and N_1 fixed. Another consequence of Gauss's lemma is that if x and y as in (3-29) are close to this segment and at a distance from each other which is comparable to one, then

$$\frac{\partial}{\partial x_1}\frac{\partial}{\partial y_1}\psi(x,y) \neq 0. \tag{3-31}$$

We shall also need to make use of the fact that, in these Fermi normal coordinates, we have

$$\frac{\partial}{\partial x_2}\frac{\partial}{\partial y_1}\psi\big((0,x_2),(0,y_2)\big) = \frac{\partial}{\partial x_1}\psi\big((0,x_2),(0,y_2)\big) = 0, \quad \text{if } d_g\big((0,x_2),(0,y_2)\big) \approx 1. \tag{3-32}$$

Next, by (3-15)–(3-17), modulo terms which are $O(\lambda^{-N})$ we can write

$$(\chi_\lambda Q_\theta^\mu)(x,y)(\chi_\lambda Q_\theta^{\mu'})(x,y') = \lambda e^{i\lambda(\psi(x,y)+\psi(x,y'))} b_\mu(x;y,y'),$$

where, by (3-28) and (3-30),

$$b_\mu(x;y,y') = 0, \quad \text{if } d_g(x,y) \text{ or } d_g(x,y') \notin [1,2],$$
$$\text{or } |x_1|+|y_1|+|y_1'| \geq c_0^{-1}\theta, \text{ or } \left|\frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\psi(x,y) - \frac{\partial}{\partial y_1}\frac{\partial}{\partial x_2}\psi(x,y')\right| \leq c_0\theta, \tag{3-33}$$

and, since we are working in Fermi normal coordinates,

$$\left|\frac{\partial^j}{\partial x_1^j}\frac{\partial^k}{\partial x_2^k} b_\mu(x,y,y')\right| \leq C_0\theta^{-j}, \quad 0 \leq j,k \leq 3. \tag{3-34}$$

The constants C_0 and c_0 can be chosen to be independent of $\mu \in \mathbb{Z}^2$ and $\theta \geq \lambda^{-1/2+\varepsilon_0}$ if $\varepsilon_0 > 0$. But then, by (3-33) and (3-34) if y_2 and y_2' are fixed and close to one another, and if we set

$$\Psi(x; s, t) = \psi(x, (s+t, y_2)) + \psi(x, (s-t, y_2')) \quad \text{and} \quad b(x; s, t) = b_\mu(x; s+t, y_2, s-t, y_2'),$$

there is a fixed constant C such that

$$\begin{aligned} &b(x; s, t) = 0 \quad \text{if } |x_1| + |s| + |t| \geq C\theta, \\ \text{and} \quad &\left| \frac{\partial^j}{\partial x_1^j} \frac{\partial^k}{\partial x_2^k} b(x; s, t) \right| \leq C\theta^{-j}, \quad 0 \leq j, k \leq 3, \end{aligned} \tag{3-35}$$

while, by (3-31) and (3-32),

$$\begin{aligned} \frac{\partial}{\partial x_2} \frac{\partial}{\partial s} \Psi(0, x_2; 0, 0) = \frac{\partial}{\partial x_2} \frac{\partial}{\partial t} \Psi(0, x_2; 0, 0) = \frac{\partial}{\partial x_1} \Psi(0, x_2; 0, 0) = 0, \\ \text{but} \quad \frac{\partial}{\partial x_1} \frac{\partial}{\partial s} \Psi(0, x_2; 0, 0) \neq 0 \quad \text{if } b(0, x_2; 0, 0) \neq 0, \end{aligned} \tag{3-36}$$

and, moreover, by (3-33),

$$\left| \frac{\partial}{\partial x_2} \frac{\partial}{\partial t} \Psi(x; s, t) \right| \geq c\theta, \quad \text{if } b(x; s, t) \neq 0. \tag{3-37}$$

Also, if we assume that $|y_2 - y_2'| \leq \delta$, as we may because of the support assumption in (3-12), then

$$\left| \frac{\partial}{\partial x_1} \frac{\partial}{\partial t} \Psi(x; s, 0) \right| \leq C\delta, \quad \text{if } b(x; s, t) \neq 0, \tag{3-38}$$

since the quantity on the left vanishes identically when $y_2 = y_2'$.

Another consequence of Gauss's lemma is that if y, y', x are close to the second coordinate axis and if the distances between x and each of y and y' are comparable to 1, then if θ above is bounded below, the 2×2 mixed Hessian of the function $(x; y_1, y_1') \to \psi(x, y) + \psi(x, y')$ has nonvanishing determinant. Thus, in this case (3-12) just follows from Hörmander's nondegenerate L^2-oscillatory integral lemma [1973] (see [Sogge 1993, Theorem 2.1.1]). Therefore, it suffices to prove (3-12) when θ is bounded above by a fixed positive constant, and so Proposition 3.1, and hence Theorem 1.1, is a consequence of the following:

Lemma 3.3. *Suppose that* $b \in C_0^\infty(\mathbb{R}^2 \times \mathbb{R}^2)$ *vanishes when* $|(s, t)| \geq \delta$. *Then if* $\Psi \in C^\infty(\mathbb{R}^2 \times \mathbb{R}^2)$ *is real and* (3-35)–(3-38) *are valid, there is a uniform constant C such that if* $\delta > 0$ *and* $\theta > 0$ *are smaller than a fixed positive constant and*

$$T_\lambda F(x) = \iint e^{i\lambda\Psi(x; s, t)} b(x; s, t) F(s, t) \, ds \, dt,$$

then we have

$$\|T_\lambda F\|_{L^2(\mathbb{R}^2)} \leq C\lambda^{-1}\theta^{-1/2} \|F\|_{L^2(\mathbb{R}^2)}. \tag{3-39}$$

We shall include the proof of this result for the sake of completeness even though it is a standard result. It is a slight variant of the main lemma in Hörmander's proof [1973] of the Carleson–Sjölin theorem (see [Sogge 1993, pp. 61–62]). Hörmander's proof gives this result in the special case where $y_2 = y_2'$, and, as above, Ψ is defined by two copies of the Riemannian distance function. The case where y_2 and y_2' are not equal to each other introduces some technicalities that, as we shall see, are straightforward to overcome.

Proof. Inequality (3-39) is equivalent to the statement that $\|T_\lambda^* T_\lambda\|_{L^2\to L^2} \le C\lambda^{-2}\theta^{-1}$. The kernel of $T_\lambda^* T_\lambda$ is

$$K(s,t;s',t') = \iint e^{i\lambda(\Psi(x;s,t)-\Psi(x;s',t'))} a(x;s,t,s',t')\,dx_1\,dx_2, \qquad \text{if } a(x;s,t,s',t') = b(x,s,t)\overline{b(x;s',t')}.$$

Therefore, we would have this estimate if we could show that

$$|K(s,t;s',t')| \le C\theta^{1-N}\big(1+\lambda|(s-s',t-t')|\big)^{-N} + C\theta\big(1+\lambda\theta|(s-s',t-t')|\big)^{-N}, \qquad N=0,1,2,3, \tag{3-40}$$

for then by using the $N=0$ bounds for the regions where $|(s-s',t-t')| \le (\lambda\theta)^{-1}$ and the $N=3$ bounds in the complement, we see that

$$\sup_{s,t} \iint |K|\,ds'\,dt', \ \sup_{s',t'} \iint |K|\,ds\,dt \le C\lambda^{-2}\theta^{-1},$$

which means that by Young's inequality, $\|T_\lambda^* T_\lambda\|_{L^2\to L^2} \le C\lambda^{-2}\theta^{-1}$, as desired.

The bound for $N=0$ follows from the first part of (3-35). To prove the bounds for $N=1,2,3$, we need to integrate by parts.

Let us first handle the case where

$$|s-s'| \ge A^{-1}|t-t'|, \tag{3-41}$$

where $A \ge 1$ is a possibly fairly large constant which we shall specify in the next step. By the second part of (3-36) and by (3-38), we conclude that if $\delta > 0$ is sufficiently small (depending on A), we have

$$\left|\frac{\partial}{\partial x_1}\big(\Psi(x;s,t)-\Psi(x;s',t')\big)\right| \ge c|s-s'|, \quad |s-s'| \ge A^{-1}|t-t'|, \tag{3-42}$$

for some uniform constant $c > 0$.

Since $|K|$ is trivially bounded by the second term on the right side of (3-40) when $|s-s'| \le (\lambda\theta)^{-1}$ and (3-41) is valid, we shall assume that $|s-s'| \ge (\lambda\theta)^{-1}$.

If we then write

$$e^{i\lambda(\Psi(x;s,t)-\Psi(x;s',t'))} = Le^{i\lambda(\Psi(x;s,t)-\Psi(x;s',t'))}, \qquad \text{where } L(x,D) = \frac{1}{i\lambda\big(\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big)}\frac{\partial}{\partial x_1}, \tag{3-43}$$

then we obtain

$$|K| \le \iint \big|(L^*(x,D))^N a(x;s,t,s',t')\big|\,dx.$$

Note that

$$\begin{aligned}\big|\lambda\big(\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big)\big|^N &|(L^*)^N a| \\ &\le C_N \sum_{0\le j+k\le N} \left|\frac{\partial^j}{\partial x_1^j}a\right| \times \sum_{\alpha_1+\cdots+\alpha_k\le N} \frac{\prod_{m=1}^k \left|\frac{\partial^{\alpha_m}}{\partial x_1^{\alpha_m}}\big(\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big)\right|}{\big|\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big|^k}. \end{aligned} \tag{3-44}$$

Clearly,

$$\prod_{m=1}^k \left|\frac{\partial^{\alpha_m}}{\partial x_1^{\alpha_m}}\big(\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big)\right| \le C_k|(s-s',t-t')|^k, \tag{3-45}$$

and consequently, by (3-41) and (3-42),

$$\frac{\prod_{m=1}^k \left|\frac{\partial^{\alpha_m}}{\partial x_1^{\alpha_m}}\big(\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big)\right|}{\big|\Psi'_{x_1}(x;s,t)-\Psi'_{x_1}(x;s',t')\big|^k} \le C_{A,k}. \tag{3-46}$$

Since by (3-35), we have that $|\partial^j_{x_1}a| \le C\theta^{-j}$, $j=0,1,2,3$, and (3-35) also says that a vanishes when $|x_1|$ is larger than a fixed multiple of θ, we conclude from (3-42)–(3-46) that if (3-41) holds, then $|K|$ is dominated by the first term on the right side of (3-40).

We now turn to the remaining case, which is

$$|t-t'| \ge A|s-s'|, \tag{3-47}$$

and where the parameter $A \ge 1$ will be specified. By the first part of (3-36) and by (3-37) and the fact that $|s|$, $|s'|$, $|t|$, $|t'|$ are bounded by a fixed multiple of θ in the support of a, it follows that we can fix A (independent of θ small) so that if (3-47) is valid, then

$$\left|\frac{\partial}{\partial x_2}\big(\Psi(x;s,t)-\Psi(x;s',t')\big)\right| \ge c\theta|t-t'|, \quad \text{on } \operatorname{supp} a,$$

for some uniform constant $c>0$. Then since (3-32) implies that

$$\prod_{m=1}^k \left|\frac{\partial^{\alpha_m}}{\partial x_2^{\alpha_m}}\big(\Psi'_{x_2}(x;s,t)-\Psi'_{x_2}(x;s',t')\big)\right| \le C_k\theta^k|(s-s',t-t')|^k, \quad \text{on } \operatorname{supp} a,$$

and since, by (3-35),

$$|\partial^j_{x_2}a| \le C_N, \quad 1\le j\le N,$$

we conclude that if we repeat the argument just given but now integrate by parts with respect to x_2 instead of x_1, then $|K|$ is bounded by the second term on the right side of (3-40), which completes the proof of Lemma 3.3. □

To conclude matters, we also need to prove the orthogonality estimates (3-4) and (3-10). Since (3-4) is a special case of (3-10), we just need to establish the latter.

To see this, we note that by Lemma 3.2, if $(\chi_\lambda Q_\theta^\mu)(x, y)$ denotes the kernel of $\chi_\lambda Q_\theta^\mu$, then

$$(\chi_\lambda Q_\theta^\mu)(x, y)(\chi_\lambda Q_\theta^{\mu'})(x, y')\overline{(\chi_\lambda Q_\theta^{\tilde{\mu}})(x, \tilde{y})(\chi_\lambda Q_\theta^{\tilde{\mu}'})(x, \tilde{y}')} = O_N(\lambda^{-N}),$$
$$\text{if } x \notin \mathcal{T}_{C\theta}(\gamma_\mu) \cap \mathcal{T}_{C\theta}(\gamma_{\mu'}) \cap \mathcal{T}_{C\theta}(\gamma_{\tilde{\mu}}) \cap \mathcal{T}_{C\theta}(\gamma_{\tilde{\mu}'}),$$

with C sufficiently large and the geodesics defined by (3-13). On the other hand, if x is in the above intersection of tubes, then the condition on $(\mu, \mu', \tilde{\mu}, \tilde{\mu}')$ in (3-10) ensures that if the constant C there is large enough, we have

$$\left|\nabla_x\big(d_g(x, y) + d_g(x, y') - d_g(x, \tilde{y}) - d_g(x, \tilde{y}')\big)\right| \geq c_0\theta,$$
$$\text{if } y \in \mathcal{T}_{C\theta}(\gamma_\mu),\ y' \in \mathcal{T}_{C\theta}(\gamma_{\mu'}),\ \tilde{y} \in \mathcal{T}_{C\theta}(\gamma_{\tilde{\mu}}), \text{ and } \tilde{y}' \in \mathcal{T}_{C\theta}(\gamma_{\tilde{\mu}'}),$$

for some uniform $c_0 > 0$. Thus, (3-10) follows from Lemma 3.2 and a simple integration by parts argument since we are assuming that $\theta \geq \theta_0 = \lambda^{-1/2+\varepsilon_0}$ with $\varepsilon_0 > 0$.

4. Relationships with Zygmund's L^4-toral eigenfunction bounds

Recall that for $\mathbb{T}^2$, Zygmund [1974] showed that if e_λ is an eigenfunction on $\mathbb{T}^2$, i.e.,

$$e_\lambda(x) = \sum_{\{\varepsilon\in\mathbb{Z}^2:\, |\ell|=\lambda\}} a_\ell e^{ix\cdot\ell}, \tag{4-1}$$

then

$$\|e_\lambda\|_{L^4(\mathbb{T}^2)} \leq C,$$

for some uniform constant C.

As observed in [Burq et al. 2007], using well-known pointwise estimates in two dimensions, one has

$$\sup_{\gamma\in\Pi} \int_\gamma |e_\lambda|^2\, ds = O_\varepsilon(\lambda^\varepsilon)$$

for all $\varepsilon > 0$. This of course implies that one also has

$$\sup_{\gamma\in\Pi} \int_{\mathcal{T}_{\lambda^{-1/2}}(\gamma)} |e_\lambda|^2\, dx = O_\varepsilon(\lambda^{-1/2+\varepsilon})$$

for any $\varepsilon > 0$.

Sarnak (unpublished) made an interesting observation that having $O(1)$ geodesic restriction bounds for $\mathbb{T}^2$ is equivalent to the statement that there is a uniformly bounded number of lattice points on arcs of λS^1 of aperture $\lambda^{-1/2}$. (Cilleruelo and Córdoba [1992] showed that this is the case for arcs of aperture $\lambda^{-1/2-\delta}$ for any $\delta > 0$.)

Using (1-1) we can essentially recover Zygmund's bound and obtain $\|e_\lambda\|_{L^4(\mathbb{T}^2)} = O_\varepsilon(\lambda^\varepsilon)$ for every $\varepsilon > 0$. (Of course this just follows from the pointwise estimate, but it shows how the method is natural too.)

If we could push the earlier results to include $\varepsilon_0 = 0$ and if we knew that there were uniformly bounded restriction bounds, then we would recover Zygmund's estimate.

References

[Bourgain 2009] J. Bourgain, "Geodesic restrictions and L^p-estimates for eigenfunctions of Riemannian surfaces", pp. 27–35 in *Linear and complex analysis*, edited by A. Alexandrov et al., Amer. Math. Soc. Transl. Ser. 2 **226**, Amer. Math. Soc., Providence, RI, 2009. MR 2011b:58066 Zbl 1189.58015

[Burq et al. 2007] N. Burq, P. Gérard, and N. Tzvetkov, "Restrictions of the Laplace–Beltrami eigenfunctions to submanifolds", *Duke Math. J.* **138**:3 (2007), 445–486. MR 2008f:58029 Zbl 1131.35053

[Carleson and Sjölin 1972] L. Carleson and P. Sjölin, "Oscillatory integrals and a multiplier problem for the disc", *Studia Math.* **44** (1972), 287–299. (errata insert). MR 50 #14052 Zbl 0215.18303

[Cilleruelo and Córdoba 1992] J. Cilleruelo and A. Córdoba, "Trigonometric polynomials and lattice points", *Proc. Amer. Math. Soc.* **115**:4 (1992), 899–905. MR 92j:11116 Zbl 0777.11035

[Gray 2004] A. Gray, *Tubes*, 2nd ed., Progress in Mathematics **221**, Birkhäuser, Basel, 2004. MR 2004j:53001 Zbl 1048.53040

[Hörmander 1973] L. Hörmander, "Oscillatory integrals and multipliers on FL^p", *Ark. Mat.* **11** (1973), 1–11. MR 49 #5674 Zbl 0254.42010

[Hörmander 2003] L. Hörmander, *The analysis of linear partial differential operators, I*, Springer, Berlin, 2003. MR 1996773 Zbl 1028.35001

[Mockenhaupt et al. 1993] G. Mockenhaupt, A. Seeger, and C. D. Sogge, "Local smoothing of Fourier integral operators and Carleson–Sjölin estimates", *J. Amer. Math. Soc.* **6**:1 (1993), 65–130. MR 93h:58150 Zbl 0776.58037

[Sogge 1986] C. D. Sogge, "Oscillatory integrals and spherical harmonics", *Duke Math. J.* **53**:1 (1986), 43–65. MR 87g:42026 Zbl 0636.42018

[Sogge 1988] C. D. Sogge, "Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds", *J. Funct. Anal.* **77**:1 (1988), 123–138. MR 89d:35131 Zbl 0641.46011

[Sogge 1993] C. D. Sogge, *Fourier integrals in classical analysis*, Cambridge Tracts in Mathematics **105**, Cambridge University Press, 1993. MR 94c:35178 Zbl 0783.35001

[Sogge 2011] C. D. Sogge, "Kakeya–Nikodym averages and L^p-norms of eigenfunctions", *Tohoku Math. J.* (2) **63**:4 (2011), 519–538. MR 2872954 Zbl 1234.35156

[Sogge and Zelditch 2012] C. D. Sogge and S. Zelditch, "Concerning the L^4 norms of typical eigenfunctions on compact surfaces", pp. 407–423 in *Recent developments in geometry and analysis*, edited by Y. Dong et al., Adv. Lect. Math. (ALM) **23**, Int. Press, Somerville, MA, 2012. MR 3077213

[Sogge and Zelditch 2014] C. D. Sogge and S. Zelditch, "A note on L^p-norms of quasi-modes", preprint, 2014.

[Zygmund 1974] A. Zygmund, "On Fourier coefficients and transforms of functions of two variables", *Studia Math.* **50** (1974), 189–201. MR 52 #8788 Zbl 0278.42005

Received 12 Sep 2014. Revised 31 Dec 2014. Accepted 9 Feb 2015.

MATTHEW D. BLAIR: blair@math.unm.edu
Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131, United States

CHRISTOPHER D. SOGGE: sogge@jhu.edu
Department of Mathematics, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218-2689, United States

mathematical sciences publishers msp

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 8 No. 3 2015

Inverse scattering with partial data on asymptotically hyperbolic manifolds 513
RAPHAEL HORA and ANTÔNIO SÁ BARRETO

Low temperature asymptotics for quasistationary distributions in a bounded domain 561
TONY LELIÈVRE and FRANCIS NIER

Dynamics of complex-valued modified KdV solitons with applications to the stability of breathers 629
MIGUEL A. ALEJO and CLAUDIO MUÑOZ

L^p estimates for bilinear and multiparameter Hilbert transforms 675
WEI DAI and GUOZHEN LU

Large BMO spaces vs interpolation 713
JOSE M. CONDE-ALONSO, TAO MEI and JAVIER PARCET

Refined and microlocal Kakeya–Nikodym bounds for eigenfunctions in two dimensions 747
MATTHEW D. BLAIR and CHRISTOPHER D. SOGGE

2157-5045(2015)8:3;1-B