Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN), Vol. 1 (2008) pp. 57-91.

SENSITIVITY ANALYSIS OF WAVE-EQUATION TOMOGRAPHY: A MULTI-SCALE APPROACH *

VALERIY BRYTIK†, MAARTEN V. DE HOOP‡, AND MIKKO SALO§

Abstract. Earthquakes, viewed as passive sources, or controlled sources, like explosions, excite seismic body waves in the earth. One detects these waves at seismic stations distributed over the earth's surface. Wave-equation tomography is derived from cross correlating, at each station, data simulated in a reference model with the observed data, for a (large) set of seismic events. The times corresponding with the maxima of these cross correlations replace the notion of residual travel times used as data in traditional tomography. Using first-order perturbation, we develop an analysis of the transform, mapping a wavespeed contrast (between the 'true' and reference models) to these maxima. We develop a construction using curvelets, while maintaining an imprint of geometrical optics reminiscent of the geodesic X-ray transform. We then introduce the adjoint of the transform, which defines the imaging of wavespeed variations from 'finite-frequency travel time' residuals. The key underlying component is the construction of the Fréchet derivative of the solution to the seismic Cauchy initial value problem in wavespeed models of limited smoothness. The construction developed in this paper essentially clarifies how a wavespeed model is probed by the method of wave-equation tomography.

Key words. tomography, wave equation, Fréchet derivative, harmonic analysis

AMS subject classifications. 86A15, 35R30

DOI.

1. Introduction. Earthquakes, viewed as passive sources, or controlled sources, like explosions, excite seismic body waves in the earth. One detects these waves at seismic stations distributed over the earth's surface. It is common practice to try and estimate the travel (or arrival) times of phases of interest in the seismic records. These travel times, together with the source-receiver positions, can be compiled as "data" for reconstruction of the wavespeed in Earth's interior by methods of tomography.

Indeed, travel times of seismic body waves have played a key role in many seismological studies of Earth's interior structure. For instance, tomographic inversions of vast amounts of routinely processed travel time residuals from international data centers, such as the International Seismological Centre, have been used to delineate three-dimensional heterogeneity in rather spectacular detail [23, 29, 8, 1]. These studies use geometrical (optics) ray theory with first-order perturbation assuming a well chosen reference model and form differential travel times.

The above mentioned rays can be viewed as geodesics, travel times as their lengths, while describing the wavespeed in Earth's interior by a Riemannian metric. The 'data' then comprise knowledge of the boundary distance function; for the analysis of the associated inverse problem, see [14] and references cited therein. In this framework, the above mentioned first-order perturbation leads to the introduction of the geodesic X-ray transform. For the range characterization of, and stability estimates for the geodesic X-ray transform restricted to geodesic complex (the dimension of which equals the dimension of space), see [6]. For each source-receiver pair, the kernel of the geodesic X-ray transform is supported on the geodesic connecting them. In this paper, we, essentially, consider and analyze a generalization of this transform arising in seismic applications.

Since the advent of routine, broadband digital recording, an increasing number of studies of Earth's interior has been relying on knowledge gleaned from cross correlations of data simulated in

*This research was supported in part under NSF CMG grant DMS 0724644, and by the members of the Geo-Mathematical Imaging Group.

†Center for Computational and Applied Mathemematics, and Geo-Mathematical Imaging Group, Purdue University, West Lafayette, IN 47907, USA (vbrytik@math.purdue.edu).

‡Center for Computational and Applied Mathematics, and Geo-Mathematical Imaging Group, Purdue University, West Lafayette, IN 47907, USA (mdehoop@purdue.edu).

§Departement of Mathematics and Statistics/ RNI, University of Helsinki, Helsinki, Finland (mikko.salo@helsinki.fi)

the reference model with the observed data. The times corresponding with the maxima of these cross correlations then replace the above mentioned differential travel times, and were considered, for example, in [12, 30]. Using first-order perturbation, we develop an analysis of the transform, mapping a wavespeed contrast (between the 'true' and reference models) to these maxima; in the seismic literature one refers to such a transform as the 'sensitivity', and to the maxima as 'finite-frequency travel times'. In an earlier paper [27] we established that this transform, in the limit of infinite bandwidth and assuming smooth wavespeed models, asymptotically reduces to the geodesic X-ray transform.

The generalization encompasses the following aspects. The transform is derived from the *full wave* solutions to the seismic Cauchy initial value problems (one for each source), and no longer infers information of the singular supports (or wavefront sets) of the solutions. We develop a *multi-scale* approach − this replaces the straightforward decomposition of data into separate time-frequency bands in use in seismic applications. The wavespeed models can be of *limited smoothness*; they will be assumed to belong to $C^{1,1}$, which allows a natural connection with the integral-geometric formulation of travel time tomography to be established. We consider the *restriction to a finite set of* (isolated) *source-receiver pairs*. The kernel of the transform will have a volumetric extent, rather than being supported on the rays connecting receivers to sources. This facilitates *data fusion*, that is, assimilation of different wave types in the data in different, but overlapping, frequency bands.

In this paper, we develop a construction of the above mentioned transform using curvelets, while maintaining an imprint of geometrical optics reminiscent of the geodesic X-ray transform. We also introduce the adjoint of the transform, which defines the imaging of wavespeed variations from 'finite-frequency travel time' data. The key underlying component is the construction of the Fréchet derivative of the solution to the seismic Cauchy initial value problem in wavespeed models of limited smoothness. In principle, this construction can also be incorporated in the process of 'waveform tomography' [17] based on least-squares fitting of selected waveforms. Moreover, one can consider the reverse-time wave-equation formulation of annihilator-based reflection tomography [21, 22, 28] and develop the linearization of the map from wavespeed model to annihilated data.

1.1. Sensitivity kernels in transmission tomography. The sensitivity of the cross correlation approach to tomography has been studied in [13, 7, 31] and other publications. The distribution kernel of the associated transform has received attention, in particular, while trying to understand how it generalizes the geodesic X-ray transform and changes the scope of tomography. The construction developed in this paper essentially provides the anatomy of the transform and clarifies how a wavespeed model is probed by the method of wave-equation tomography.

1.2. The underlying initial value problem. We begin with a few definitions. We assume that $U \subset \mathbb{R}^n$ is open, and $0 < \gamma \leq 1$. With $x \in \mathbb{R}^n$, we set $D_{x_j} = \mathrm{i}^{-1}\partial_{x_j}$, while $D_t = \mathrm{i}^{-1}\partial_t$.

Definition 1.1. (i) *If $u:\ U \to \mathbb{R}$ is bounded and continuous, we write*

$$\|u\|_{C(\bar{U})} := \sup_{x\in U}|u(x)|.$$

(ii) *The γ^{th} Hölder seminorm of $u:\ U \to \mathbb{R}$ is given by*

$$[u]_{C^{0,\gamma}(\bar{U})} := \sup_{x,y\in U, x\neq y}\left\{\frac{|u(x)-u(y)|}{|x-y|^{\gamma}}\right\},$$

and the γ^{th} Hölder norm is

$$\|u\|_{C^{0,\gamma}(\bar{U})} := \|u\|_{C(\bar{U})} + [u]_{C^{0,\gamma}(\bar{U})}\,.$$

DEFINITION 1.2. *The Hölder space, $C^{k,\gamma}(\bar{U})$, consists of all $u \in C^k(\bar{U})$ for which the norm*

$$\|u\|_{C^{k,\gamma}(\bar{U})} := \sum_{|\alpha|\leq k} \|D^\alpha u\|_{C(\bar{U})} + \sum_{|\alpha|=k} [D^\alpha u]_{C^{0,\gamma}(\bar{U})}$$

is finite.

In this paper, we encounter the spaces $C^{k-1,1}$, and use the equivalent norms $\|u\|_{C^{k-1,1}} = \sum_{|\alpha|\leq k} \|D^\alpha u\|_{L^\infty}$.

DEFINITION 1.3. *The space, $L^p([0,T]\,;X)$, consists of all measurable functions $u : [0,T] \to X$ with*

$$\|u\|_{L^p([0,T];X)} := \left(\int_0^T \|u(t)\|^p \mathrm{d}t\right)^{\frac{1}{p}} < \infty,$$

for $1 \leq p < \infty$, while

$$\|u\|_{L^\infty([0,T];X)} := \operatorname{ess}\sup_{0\leq t\leq T} \|u(t)\| < \infty.$$

Throughout the paper, we will use an abbreviation for the latter space, namely $L_t^\infty X$, where X is a Sobolev space. In a similar fashion, we denote by $C_t^0 X$ the space of functions, $u(t,x)$, continuous in time, for which

$$\|u\|_{C^0([0,T];X)} := \sup_{0\leq t\leq T} \|u(t)\|_X < \infty,$$

and by $C_t^1 X$ the space of functions, $u(t,x)$, for which

$$u(t,x) \in C^0([0,T]\,;X) \text{ and } \partial_t u(t,x) \in C^0([0,T]\,;X).$$

Let $a^{ij}(x)$ be functions in $C^{1,1}(\mathbb{R}^n)$ such that the matrix $(a^{ij}(x))$ is symmetric and positive definite with a uniform bound when $x \in \mathbb{R}^n$. Then the operator $A(x, D_x) = \sum_{i,j=1}^n a^{ij}(x) D_{x_i} D_{x_j}$ is uniformly elliptic, and we may consider the Cauchy initial value problem for the wave equation,

$$\left\{\begin{aligned} [D_t^2 - A(x,D_x)]u(t,x) &= 0,\\ u|_{t=0} &= 0,\\ \partial_t u|_{t=0} &= g. \end{aligned}\right. \tag{1.1}$$

Here, $g \in H^\alpha$ and $u(t,x)$ is a function in $\mathbb{R}_t \times \mathbb{R}_x^n$. The regularity condition for the coefficient functions is natural in the context of tomography, since the Hamiltonian flow is well defined for $C^{1,1}$ metrics, but not, for example, for $C^{1,\gamma}$ metrics where $\gamma < 1$.

If M is a large positive constant such that

$$\|a^{ij}\|_{C^{1,1}} \leq M, \quad a^{ij}\xi_i\xi_j \geq M^{-2}|\xi|^2,$$

and $-1 \leq \alpha \leq 2$, it is well known that the Cauchy problem above has a unique weak solution $u \in C^0([-M,M]; H^{\alpha+1}) \cap C^1([-M,M]; H^\alpha)$. We will write this solution as

$$u(t,x) = u_A(t,x) = (S_A(t)g)(x).$$

If $a^{ij} = \delta^{ij}$, which corresponds to the Euclidean metric, it is an elementary fact that

$$S(t)g = \frac{\sin(t\sqrt{-\Delta})}{\sqrt{-\Delta}} g, \quad \text{with } \Delta = -\sum_{i=1}^n D_{x_i}^2.$$

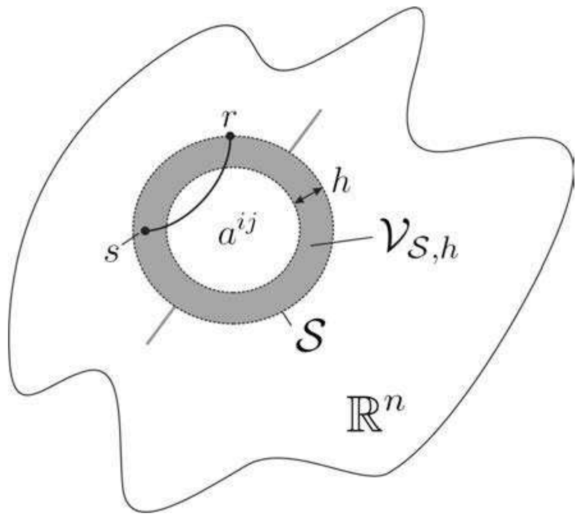

FIG. 1. *Configuration for wave-equation tomography.*

The solution operator $S_A(t)$ is the corresponding "sine propagator" in the variable coefficient case.

We are interested in studying the properties of $u_A = S_A(t)g$ when g is fixed, but the coefficient matrix of A varies. This is a common situation in tomographic imaging problems. In these problems, for the purpose of imaging and optimization, one performs a linearization about a fixed wavespeed model (metric). The main result of this paper is a multi-scale approach to constructing the linearization of operator $S_A(t)$ with respect to A. This approach reveals the imprint of traditional tomography following geodesics.

We assume the functions g represent earthquakes, parametrized by their locations, s, and write $g = g_s$. To express the source parametrization, we also write $u_A = u_A(t, x; s)$. The wave solution is observed in seismic stations located at r. We assume that $r \in \mathcal{S} \sim S^{n-1} \subset \mathbb{R}^n$; s lies in an annulus, $\mathcal{V}_{\mathcal{S},h}$ of 'thickness' h with outer boundary coinciding with $\mathcal{S}$; see Fig. 1.

2. Main results. In this section, we present our main results pertaining to the existence and mapping properties of the Fréchet derivative of the wave solution operator, and its application to wave-equation tomography.

2.1. Fréchet derivative. If g is a fixed function in $\mathbb{R}^n$, we consider the map $A \mapsto S_A(t)g$ from $C^{1,1}$ metrics to solutions of the wave equation. To begin with, the solution depends continuously on the metric. This is a consequence of energy estimates (see Stolk [20]), but our starting point is the constructive proof in Salo [15], which is based on curvelets.

THEOREM 2.1. *Let $A = (a^{ij})$ and $B = (b^{ij})$ be two $C^{1,1}$ metrics in $\mathbb{R}^n$, and let M be a large constant such that*

$$\|a^{ij}\|_{C^{1,1}} \le M, \quad \sum_{i,j=1}^{n} a^{ij}\xi_i\xi_j \ge M^{-2}|\xi|^2; \tag{2.1}$$

similar conditions are satisfied for the metric B. Let $t \in [-M, M]$; if $-1 \le \alpha \le 1$, we have

$$\|(S_A(t) - S_B(t))g\|_{H^{\alpha+1}} \le C\|A - B\|_{C^{0,1}}\|g\|_{H^{\alpha+1}}, \tag{2.2}$$

where C depends only on M and n.

We will use the notation $\lesssim$ to indicate $\le C\cdot$ where C depends only on M and n. If A is a $C^{1,1}$ metric then $S_A(t)$ is a bounded operator from $H^\alpha(\mathbb{R}^n)$ to $H^{\alpha+1}(\mathbb{R}^n)$ if $-1 \le \alpha \le 2$, by energy estimates [20] or curvelet methods [18]. Theorem 2.1 gives Lipschitz stability in the case where the initial value g is one derivative smoother than required by the mapping properties of $S_A(t)$. This "loss of derivatives" is a natural feature of the stability results and will be present throughout this paper.

One of the contributions of this paper is the constructive proof establishing that the map $A \mapsto S_A(t)g$ is not just continuous, but also Fréchet differentiable with respect to the metric. By definition, the Fréchet derivative exists at A if for small $C^{1,1}$ perturbations, $\varDelta$, one has

$$S_{A+\varDelta}(t) = S_A(t) + (DS)_{A,\varDelta}(t) + (RS)_{A,\varDelta}(t),$$

where $(DS)_{A,\varDelta}(t)$ is linear in $\varDelta$, and $(RS)_{A,\varDelta}(t) = o(\|\varDelta\|)$ in suitable norms. This is the content of the following theorem. Again, the result can be proved in an abstract way via energy estimates [20]. We present, here, a multi-scale approach and a constructive proof using curvelets.

Theorem 2.2. *Let A and $A+\Delta$ be two metrics satisfying (2.1). There exists a linear operator $(DS)_{A,\varDelta}(t)$, which acts linearly in Δ and satisfies*

$$\|(DS)_{A,\varDelta}(t)g\|_{H^{\alpha+1}} \lesssim \|\Delta\|_{C^{0,1}}\|g\|_{H^{\alpha+1}}, \quad -1 \le \alpha \le 1, \tag{2.3}$$

while

$$\|(RS)_{A,\varDelta}(t))g\|_{H^{\alpha+1}} \lesssim \|\Delta\|^2_{C^{0,1}}\|g\|_{H^{\alpha+2}}, \quad -1 \le \alpha \le 0. \tag{2.4}$$

Note that there is a "loss of one derivative" in (2.3) and a "loss of two derivatives" in (2.4). This is consistent with [20] where one has estimates in spaces which have one derivative less regularity each time.

A direct consequence of Theorems 2.1 and 2.2 is

Corollary 2.3. *If $g \in H^{\alpha+1}$ with $-1 \le \alpha < 1$, then*

$$\|(RS)_{A,\varDelta}(t))g\|_{H^{\alpha+1}} = o(\|\Delta\|_{C^{0,1}}) \tag{2.5}$$

as $\|\Delta\|_{C^{0,1}} \to 0$.

Proof. We decompose $g = g_s + g_r$, where $g_s \in H^2$ and $\|g_r\|_{H^{\alpha+1}}$ is small. Then

$$\begin{aligned}\|(RS)_{A,\varDelta}(t)g_s\|_{H^2} &\le \|(S_{A+\varDelta}(t) - S_A(t))g_s\|_{H^2} + \|(DS)_{A,\varDelta}(t)g_s\|_{H^2} \\ &\lesssim \|\varDelta\|_{C^{0,1}}\|g_s\|_{H^2}.\end{aligned} \tag{2.6}$$

Also, by Theorem 2.2

$$\|(RS)_{A,\varDelta}(t)g_s\|_{H^1} \lesssim \|\varDelta\|^2_{C^{0,1}}\|g_s\|_{H^2}.$$

Interpolating the last two estimates gives

$$\|(RS)_{A,\varDelta}(t)g_s\|_{H^{\alpha+1}} \lesssim \|\varDelta\|^{2-\alpha}_{C^{0,1}}\|g_s\|_{H^2}$$

for $0 \le \alpha \le 1$. By the same argument as in (2.6),

$$\|(RS)_{A,\varDelta}(t)g_r\|_{H^{\alpha+1}} \lesssim \|\varDelta\|_{C^{0,1}}\|g_r\|_{H^{\alpha+1}}.$$

Combining the last two estimates implies (2.5) upon choosing g_r so that $\|g_r\|_{H^{\alpha+1}}$ is sufficiently small. ∎

We need this corollary to obtain *pointwise* estimates, pertaining to the solution of the wave equation and its Fréchet derivatives in three dimensions, that apply to observations in isolated ("point") seismic stations. Indeed, if $n = 3$, Sobolev embedding [5, p.270, Theorem 6] implies that $H^s(\mathbb{R}^3)$ embeds in the bounded, continuous functions if $s > 3/2$, and then $\|u\|_{L^\infty(\mathbb{R}^3)} \le C\|u\|_{H^s(\mathbb{R}^3)}$, thus leaving admissable α-values $1/2 < \alpha < 1$.

2.2. Wave-equation tomography. In wave-equation tomography the dimension is $n = 3$, and the data, $d = d(t, r; s)$, are modelled by $u_{A_0}(t, r; s)$, with $t \in [0, T]$ (assuming $T < M$) while $r \in \mathcal{S}$ belongs to a finite set and s belongs to a finite set contained in $\mathcal{V}_{\mathcal{S},h}$. We assume that the interior of the earth can be described by the $C^{1,1}$ metric A_0. In the process of tomography, the sources, g_s, are known while the metric, A_0, is unknown. The objective is to find a good $C^{1,1}$ approximation A to A_0, by developing a mismatch criterion based on the cross-correlations

$$\mathsf{C}(A, t) = \int_{\mathbb{R}} u_A^\chi(\bar{t}, r; s) d^\chi(t + \bar{t}, r; s)\, \mathrm{d}\bar{t}. \tag{2.7}$$

Here, $u_A^\chi(t, r; s) = \chi_{[T_1,T_2]}(t)\, u_A(t, r; s)$, with u_A representing data modeled with coefficients A assuming the same source $g = g_s$ that generates d; also, $d^\chi(t, r; s) = \chi_{[T_1,T_2]}(t)\, d(t, r; s)$. The window function, $\chi_{[T_1,T_2]} \in C_0^\infty$, is chosen to select particular time intervals in the seismic records, with $0 \leq T_1 < T_2 < M$. We assume that g is compactly supported. If $g \in H^{1/2+\varepsilon}$ then u_A and d are continuous by Sobolev embedding.

For the true metric A_0, we have $\mathsf{C}(A_0, -t) = \mathsf{C}(A_0, t)$ whence $t \mapsto \mathsf{C}(A_0, t)$ has a maximum at $t = 0$. If $g \in H^{3/2+\varepsilon}$ then $\partial_t u \in C_t C_x$ by Sobolev embedding, and one may compute the time derivative

$$\mathsf{F}(A, t) := \partial_t \mathsf{C}(A, t) = \int u_A^\chi(\bar{t}) \partial_t u_{A_0}^\chi(t + \bar{t})\, \mathrm{d}\bar{t}.$$

Here, and below, we write $u_A^\chi(t) = u_A^\chi(t, r; s)$ and similarly for $u_{A_0}^\chi$. We have the property that $\mathsf{F}(A_0, 0) = 0$.

The “finite-frequency” traveltime difference $\Delta t = \Delta t(s, r)$ is defined as the real number with smallest absolute value which solves

$$\mathsf{F}(A, \Delta t) = 0.$$

If A coincides with the true model, A_0, then $\Delta t(s, r) = 0$ for all (s, r). The next result states that Δt is well defined for metrics near A_0

LEMMA 2.4. *If (A, t) is sufficiently close to $(A_0, 0)$, then there is a C^1 function Δt defined for $C^{1,1}$ metrics near A_0 such that*

$$\mathsf{F}(A, t) = 0 \textit{ for } (A, t) \textit{ near } (A_0, 0) \qquad \Leftrightarrow \qquad t = \Delta t(A).$$

Proof. In a neighborhood of $(A_0, 0)$ in $C^{1,1} \times \mathbb{R}$, the function F is C^1 because

$$\partial_t \mathsf{F}(A, t) = \partial_t \int u_A^\chi(\bar{t} - t) \partial_t u_{A_0}^\chi(\bar{t})\, \mathrm{d}\bar{t} = -\int \partial_t u_A^\chi(\bar{t} - t) \partial_t u_{A_0}^\chi(\bar{t})\, \mathrm{d}\bar{t},$$

while

$$\partial_A \mathsf{F}(A, t)\delta A = \int ((DS)_{A,\delta A}(\bar{t}) g) \chi_{[T_1,T_2]}(\bar{t})\, \partial_t u_{A_0}^\chi(t + \bar{t})\, \mathrm{d}\bar{t},$$

so that

$$\begin{aligned}
&\mathsf{F}(A + \delta A, t + \delta t) - \mathsf{F}(A, t) - \partial_t \mathsf{F}(A, t)\delta t - \partial_A \mathsf{F}(A, t + \delta t)\delta A \\
&\qquad = \int [(S_{A+\delta A}(\bar{t}) - S_A(\bar{t}) - (DS)_{A,\delta A}(\bar{t}))g] \chi_{[T_1,T_2]}(\bar{t})\, \partial_t u_{A_0}^\chi(\bar{t} + t + \delta t)\, \mathrm{d}\bar{t} \\
&\qquad\qquad + \int [u_A^\chi(\bar{t} - t - \delta t) - u_A^\chi(\bar{t} - t) - \partial_t u_A^\chi(\bar{t} - t)(-\delta t)]\, \partial_t u_{A_0}^\chi(\bar{t})\, \mathrm{d}\bar{t}.
\end{aligned}$$

We use that $g \in H^{3/2+\varepsilon}$. Moreover, we use the estimates

$$\|(S_{A+\delta A}(\bar{t}) - S_A(\bar{t}) - (DS)_{A,\delta A}(\bar{t}))g\|_{H^{3/2+\varepsilon}} = o(\|\delta A\|_{C^{1,1}}),$$

and

$$\|u_A(t' - \delta t) - u_A(t') - \partial_t u_A(t')(-\delta t)\|_{H^{3/2+\varepsilon}} = o(|\delta t|).$$

Since $\mathsf{F}(A_0, 0) = 0$ and $\partial_t \mathsf{F}(A_0, 0) = -\int |\partial_t u^{\chi}_{A_0}(\bar{t})|^2 \, \mathrm{d}\bar{t} < 0$, the existence of a C^1 function Δt with stated properties follows from the implicit function theorem. □

Differentiating the identity $\mathsf{F}(A, \Delta t(A)) = 0$ near A_0 gives

$$(\partial_A \Delta t)(A)\delta A = -\frac{\partial_A \mathsf{F}(A, \Delta t(A))\delta A}{\partial_t \mathsf{F}(A, \Delta t(A))}.$$

This can be interpreted as the sensitivity with respect to changes in A for wave-equation tomography. Using that $\Delta t(A_0) = 0$, we obtain the *sensitivity map*, $\delta A \mapsto \delta t$ with

$$\delta t := \partial_A \Delta t(A_0)\delta A = -\frac{\partial_A \mathsf{F}(A_0, 0)\delta A}{\partial_t \mathsf{F}(A_0, 0)} = \frac{\displaystyle\int ((DS)_{A_0,\delta A}(\bar{t})g)\chi_{[T_1,T_2]}(\bar{t})\, \partial_t u^{\chi}_{A_0}(\bar{t})\, \mathrm{d}\bar{t}}{\displaystyle\int \partial_t u^{\chi}_{A_0}(\bar{t}) \partial_t u^{\chi}_{A_0}(\bar{t})\, \mathrm{d}\bar{t}}. \tag{2.8}$$

(We substitute $u_{A_0}(\bar{t}) = S_{A_0}(\bar{t})g$.) This map replaces the geodesic X-ray transform in conventional tomography. Its construction – with an "imprint" of the geodesic X-ray transform – is the topic of Sections 4 and 5.

3. Multi-scale approach.

3.1. Wavepackets and curvelets. Originally, the parametrix construction based on wavepackets (or curvelets) for the wave equation with $C^{1,1}$ coefficients appeared in Smith [18]. Smith initially constructed a tight frame of curvelets, where the frame elements were compactly supported in the frequency domain. These tight frames were also considered by Candès and Donoho [2, 3, 4]. The analysis of the parametrix construction is somewhat simplified by using a continuous wavepacket representation, also called the FBI transform. This idea appeared in the works of Tataru [24, 25, 26] for wavepackets based on the Gaussian. Smith [19] gave a construction based on wavepackets compactly supported in frequency, which approach will be taken here.

We will use similar notation as in [15]. Let ϕ be a real, even Schwartz function in $\mathbb{R}^n$ with $\|\phi\|_{L^2} = (2\pi)^{-n/2}$, and assume $\hat{\phi}$ is supported in the unit ball. For $\lambda \geq 1$ and $y, x, \xi \in \mathbb{R}^n$, we define

$$\phi_\lambda(y; x, \xi) = \lambda^{n/4} \mathrm{e}^{\mathrm{i}\langle \xi, y-x\rangle} \phi(\lambda^{1/2}(y - x)). \tag{3.1}$$

This is a wavepacket at frequency level λ, centered in space at x and in frequency at ξ, see Fig. 2. Its Fourier transform is given by

$$\hat{\phi}_\lambda(\eta; x, \xi) = \lambda^{-n/4} \mathrm{e}^{-\mathrm{i}\langle \eta, x\rangle} \hat{\phi}(\lambda^{-1/2}(\eta - \xi)).$$

The FBI transform of a function $f \in \mathscr{S}(\mathbb{R}^n)$ is given by

$$T_\lambda f(x, \xi) = \int f(y) \overline{\phi_\lambda(y; x, \xi)} \, \mathrm{d}y. \tag{3.2}$$

Suppose that $\lambda \geq 2^6$. Then, if $\hat{f}$ is supported in $\frac{1}{4}\lambda < |\xi| < \lambda$, $T_\lambda f$ vanishes unless $\frac{1}{8}\lambda < |\xi| < 2\lambda$. If $F \in \mathscr{S}(\mathbb{R}^{2n}_{x,\xi})$, the adjoint T^*_λ of T_λ has the form

$$T^*_\lambda F(y) = \iint F(x, \xi) \phi_\lambda(y; x, \xi) \, \mathrm{d}x \mathrm{d}\xi. \tag{3.3}$$

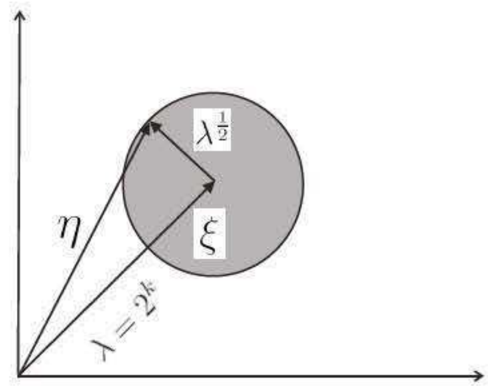

FIG. 2. *Support of a wavepacket* $\hat{g}_\lambda(\eta; x, \xi)$.

It follows that $T_\lambda^* T_\lambda = I$, and $\|T_\lambda f\|_{L^2(\mathbb{R}^{2n}_{x,\xi})} = \|f\|_{L^2(\mathbb{R}^n)}$.

The following result [19, Lemma 3.1] states the L^2 boundedness of FBI transform type operators:

LEMMA 3.1. *Suppose that $\phi_{x,\xi}$ is a (x,ξ)-family of Schwartz functions on $\mathbb{R}^n$, whose Schwartz seminorms are bounded uniformly in x and ξ. Let*

$$(\phi_{x,\xi})_\lambda(y; x, \xi) = \lambda^{n/4} \mathrm{e}^{\mathrm{i}\langle \xi, y-x\rangle} \phi_{x,\xi}(\lambda^{1/2}(y-x)).$$

The operator

$$T'_\lambda f(x,\xi) = \int f(y) \overline{(\phi_{x,\xi})_\lambda(y; x, \xi)} \, \mathrm{d}y \tag{3.4}$$

is bounded from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^{2n}_{x,\xi})$. Furthermore, its adjoint, given by

$$(T'_\lambda)^* F(y) = \int F(x,\xi) (\phi_{x,\xi})_\lambda(y; x, \xi) \, \mathrm{d}x \mathrm{d}\xi \tag{3.5}$$

is bounded from $L^2(\mathbb{R}^{2n}_{x,\xi})$ to $L^2(\mathbb{R}^n)$.

The norms of T'_λ, $(T'_\lambda)^*$ are bounded by $\sup_{(x,\xi)} C_{N+1}(\phi_{x,\xi})$, with

$$C_N(\phi_{x,\xi}) := \sum_{|\alpha| \le N} \|\langle y\rangle^N \partial_y^\alpha \phi_{x,\xi}(y)\|_{L^\infty}, \tag{3.6}$$

and $N > 2n$.

Proof. The Schwartz kernel of the composition $T'_\lambda (T'_\lambda)^*$ is given by

$$\mathcal{K}(x,\xi; x'\xi') = \mathrm{e}^{\mathrm{i}(\langle \xi, x\rangle - \langle \xi', x'\rangle)} \int \mathrm{e}^{\mathrm{i}\lambda^{-1/2}\langle \xi' - \xi, y\rangle} \, \overline{\phi_{x,\xi}(y - \lambda^{1/2}x)} \phi_{x',\xi'}(y - \lambda^{1/2}x') \, \mathrm{d}y.$$

This kernel satisfies the estimates [19]

$$|\mathcal{K}(x,\xi; x'\xi')| \le C\,(1 + \lambda^{-1/2}|\xi - \xi'| + \lambda^{1/2}\,|x - x'|)^{-N}$$

for all $N \in \mathbb{N}$. Here, the constant C depends on N and can be further estimated by finitely many Schwartz semi-norms of $\phi_{x,\xi}$, that is, by

$$\sup_{(x,\xi)} C_{N+1}(\phi_{x,\xi})^2.$$

Choosing $N > 2n$, we can use Schur's lemma to show that $T'_\lambda\,(T'_\lambda)^*$ is bounded on $L^2(\mathbb{R}^{2n}_{x,\xi})$. ∎

We will use this lemma several times in the proofs of Section 5, with different types of wavepackets, $\phi_{x,\xi}$. Furthermore, we will identify λ with dyadic scales, that is, $\lambda = 2^k$, and write T_k for T_λ.

3.2. Symbol smoothing and model decomposition. We discuss the principal multi-scale decomposition. We let $\chi(\xi)$ be a smooth cutoff, supported in the unit ball, with $\chi = 1$ for $|\xi| \le 1/2$. We subject the coefficients, $a^{ij} \in C^{1,1}$, in the wave equation to the smoothing

$$a^{ij}_k(x) = \chi(2^{-k/2} D_x) a^{ij}(x). \tag{3.7}$$

This operation directly implies smoothing of the symbol, $A(x,\xi)$, in (1.1).

In conjunction with the symbol smoothing, we use the Littlewood-Paley frequency decomposition. We take $\beta_k(D)$ to be a Littlewood-Paley partition of unity, with

$$\beta_0(\xi) + \sum_{k=1}^{\infty} \beta_k(\xi) = 1,$$

where β_0 is supported in the unit ball, β_1 is supported in $\{\frac{1}{2} < |\xi| < 2\}$, and $\beta_k(\xi) = \beta_1(2^{-k+1}\xi)$. We subject the initial data, g, to such a decomposition:

$$g_k = \beta_k(D) g. \tag{3.8}$$

4. Summary of parametrix construction. In this section, we summarize the parametrix construction for wave equations with $C^{1,1}$ metrics, cf. (1.1). We follow the construction in [15], which was based on [18] and [19].

Following the Littlewood-Paley decomposition, for each k, we consider the initial value problem,

$$\left\{ \begin{aligned} [D_t^2 - A_k(x, D_x)] u_k(t,x) &= 0, \\ u_k|_{t=0} &= 0, \\ \partial_t u_k|_{t=0} &= g_k; \end{aligned} \right. \tag{4.1}$$

here, the metric is smooth. We begin with constructing an approximate solution while distinguishing forward and backward time directions. We construct pseudodifferential operators, $P^{\pm}_{A;k}(x, D_x)$, with symbols

$$p^{\pm}_{A;k}(x,\xi) = \pm \chi(2^{-k/2} D_x) \sqrt{A_k(x,\xi)} \tag{4.2}$$

so that $p^{\pm}_{A;k}(x,\xi)\beta_k(\xi) \in S^1_{1,1/2}$. Moreover, we construct pseudodifferential operators, $Q^{\pm}_{A;k}(x, D_x)$, with symbols

$$q^{\pm}_{A;k}(x,\xi) = \chi(2^{-k/2} D_x) \frac{1}{p^{\pm}_{A;k}(x,\xi)} \tag{4.3}$$

so that $q^{\pm}_{A;k}(x,\xi)\beta_k(\xi) \in S^{-1}_{1,1/2}$.

Remark 4.1. *We discuss the decomposition into forward and backward solutions, starting from (4.1). Since $A_k(x, D_x)$ is elliptic, we can introduce its square root, $B_k(x, D_x) = \sqrt{A_k(x, D_x)}$. Then*

$$u_\pm(t,x) = \tfrac{1}{2} u_k(t,x) \pm \tfrac{1}{2} \mathrm{i} B_k(x, D_x)^{-1} \partial_t u_k(t,x)$$

so that

$$u_\pm(t=0,x) = \tfrac{1}{2} u_k(t,x)|_{t=0} \pm \tfrac{1}{2} \mathrm{i} B_k(x,D_x)^{-1} \partial_t u(t,x)|_{t=0} = \pm \tfrac{1}{2} \mathrm{i} B_k(x,D_x)^{-1} g_k,$$

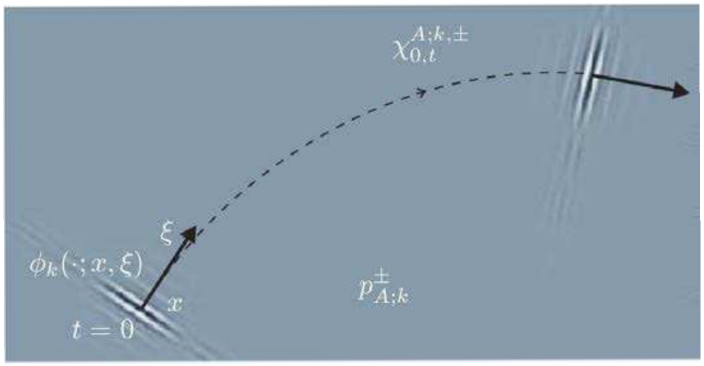

FIG. 3. *Illustration of the "flow" operator,* $U^{\pm}_{A;k}(t)$.

solve the initial value problems

$$\left\{\begin{aligned} [\partial_t \pm \mathrm{i}B_k(x,D_x)]u_\pm(t,x) &= 0, \\ u_\pm|_{t=0} &= \pm\tfrac{1}{2}\mathrm{i}B_k(x,D_x)^{-1}g_k. \end{aligned}\right. \tag{4.4}$$

In the above, $P^{\pm}_{A;k}(x,D_x)$ *essentially coincides with* $\pm B_k(x,D_x)$*, up to principal parts, and* $Q^{\pm}_{A;k}(x,D_x)$ *essentially coincides with* $\pm B_k(x,D_x)^{-1}$*, up to principal parts.*

We build "flow" operators for the metric subjected to smoothing according to:

$$E^{\pm}_{A;k}(t)g = T_k^* U^{\pm}_{A;k}(t)\, T_k(\tfrac{1}{2}\mathrm{i}Q^{\pm}_{A;k}\beta_k(D)g). \tag{4.5}$$

Here, $U^{\pm}_{A;k}(t)$ represents a rigid motion of wavepackets along the Hamilton flow associated with $p^{\pm}_{A;k}$, that is,

$$U^{\pm}_{A;k}(t)F = F\circ\chi^{A;k,\pm}_{t,0}, \tag{4.6}$$

$$\begin{aligned} T_k^* U^{\pm}_{A;k}(t)\, T_k f &= \iint (T_k f)(\chi^{A;k,\pm}_{t,0}(x',\xi'))\,\phi_k(.;x',\xi')\,\mathrm{d}x'\mathrm{d}\xi', \\ &\overset{(x',\xi')=\Phi^{\pm}_{A;k}(t)(x,\xi)}{=} \iint (T_k f)(x,\xi)\,\phi_k(.;\Phi^{\pm}_{A;k}(t)(x,\xi))\,\mathrm{d}x\mathrm{d}\xi, \end{aligned}$$

where $\Phi^{\pm}_{A;k}(t) = \chi^{A;k,\pm}_{0,t}$, with the flow $\chi^{A;k,\pm}_{t,0}:\ (x,\xi)\mapsto(x(t;x,\xi),\xi(t;x,\xi))$ being generated by the Hamilton system,

$$\begin{aligned} \dot{x}(t) &= \ \ \partial_\xi p^{\pm}_{A;k}(x(t),\xi(t)), \\ \dot{\xi}(t) &= -\partial_x p^{\pm}_{A;k}(x(t),\xi(t)), \end{aligned}$$

subject to initial conditions $(x(0),\xi(0)) = (x,\xi)$. The following property will be used throughout the paper: If $(x(t;x,\xi),\xi(t;x,\xi))$ satisfy the Hamilton system and $|\xi(0)| = |\xi| \approx 2^k$ then $|\xi(t;x,\xi)| \approx 2^k$ for $|t|\le M$ [15, pp.4-5].

We then construct the leading-order approximation to the solution operator of (4.1) by summing over scales:

$$\widetilde{S}_A(t)g = t\sum_{k<k_0} g_k + \sum_{k\ge k_0}(u_k^+ + u_k^-), \quad u_k^\pm = E^{\pm}_{A;k}(t)g, \tag{4.7}$$

with k_0 sufficiently large. We introduce the shorthand notation, $u_k^+ + u_k^- = \sum_\pm u_k = \sum_\pm E_{A;k}(.)g$. $\widetilde{S}_A(t)g$ does not satisfy the initial conditions, that is, $\widetilde{S}_A(t)g|_{t=0} = 0$ but $\partial_t \widetilde{S}_A(t)g|_{t=0} \neq g$ in general. The value for the initial velocity follows to be

$$\partial_t \widetilde{S}_A(t)g \Big|_{t=0} = \sum_{k<k_0} g_k + \mathrm{i} \sum_\pm \sum_{k \geq k_0} (D_t + P_{A;k}) u_k \Big|_{t=0} \\ - \mathrm{i} \sum_\pm \sum_{k \geq k_0} P_{A;k} u_k \Big|_{t=0} = (I + K_A) g, \tag{4.8}$$

where K_A is given by

$$K_A = \sum_\pm \sum_{k \geq k_0} \Big[\mathrm{i} \widetilde{R}_{A;k}(0) + \tfrac{1}{2} R_{A;k} \beta_k(D) \Big]. \tag{4.9}$$

Here,

$$\widetilde{R}^\pm_{A;k}(t) = (D_t + P^\pm_{A;k})\, E^\pm_{A;k}(t), \tag{4.10}$$

while $R^\pm_{A;k} \beta_k(D)$ are pseudodifferential operators with symbols in $S^0_{1,1/2}$ (in fact, also in $S^{-1}_{1,1/2}$), defined by

$$P^\pm_{A;k} Q^\pm_{A;k} \beta_k(D) = (I + R^\pm_{A;k})\, \beta_k(D). \tag{4.11}$$

In the above, k_0 is chosen sufficiently large such that $(I + K_A)^{-1}$ exists. It follows that

$$\widehat{S}_A(t) = \widetilde{S}_A(t)\, (I + K_A)^{-1} \tag{4.12}$$

has the property that

$$\widehat{S}_A(t)g|_{t=0} = 0 \quad \text{and} \quad \partial_t \widehat{S}_A(t)g|_{t=0} = g.$$

To construct the solution of (1.1) one introduces a "residual force" source, G:

$$u(t,.) = \widehat{S}_A(t)g + \int_0^t \widehat{S}_A(t,s) G(s,.)\, \mathrm{d}s, \qquad \widehat{S}_A(t,s) = \widehat{S}_A(t-s). \tag{4.13}$$

To find the residual force $G(s,x)$, we apply the wave operator operator, $D_t^2 - A(x, D_x)$, to this equality. With

$$\partial_t^2 \Big(\int_0^t \widehat{S}_A(t,s) G(s,.)\, \mathrm{d}s \Big) = G(t,.) + \int_0^t \partial_t^2 \widehat{S}_A(t,s) G(s,.)\, \mathrm{d}s,$$

we obtain

$$[D_t^2 - A(x, D_x)] u(t,.) = \mathbf{T}_A(t) g - G(t,.) + \int_0^t \mathbf{T}_A(t,s) G(s,.)\, \mathrm{d}s,$$

where

$$\mathbf{T}_A(t,s) = [D_t^2 - A(x, D_x)]\, \widehat{S}_A(t,s), \quad \text{setting } \mathbf{T}_A(t) = \mathbf{T}_A(t,0), \tag{4.14}$$

is an operator of order 0, that is, it is bounded on $H^\alpha(\mathbb{R}^n)$ for $-1 \leq \alpha \leq 2$. (In the later analysis, we will also introduce $\widetilde{\mathbf{T}}_A(t,s) = [D_t^2 - A(x, D_x)]\, \widetilde{S}_A(t,s)$.) Then u will be a solution of (1.1) provided that $G(t,.) = V_A(\mathbf{T}_A(t)g)$, where $G = V_A(h)$ solves the Volterra equation

$$G(t,.) - \int_0^t \mathbf{T}_A(t,s) G(s,.)\, \mathrm{d}s = h(t,.).$$

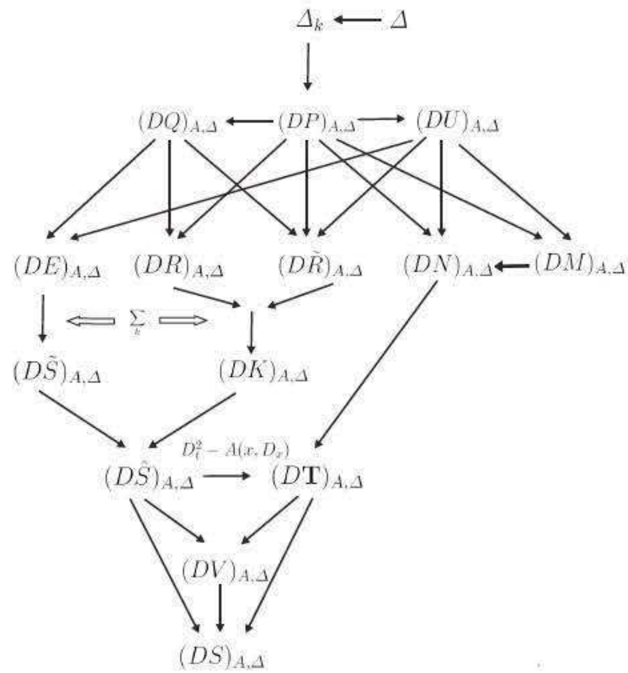

FIG. 4. *Diagram describing the proof of Fréchet differentiability of $S_A(t)$, and the interrelation of the operators involved.*

Since $\mathbf{T}_A(t,s)$ is bounded on H^α, V_A is bounded on $L_t^\infty H_x^\alpha$, for $-1 \leq \alpha \leq 2$. The full solution operator follows to be

$$S_A(t)g = \widehat{S}_A(t)g + \int_0^t \widehat{S}_A(t,s)V_A(\mathbf{T}_A g)(s,.)\,\mathrm{d}s. \tag{4.15}$$

In the process of obtaining the appropriate estimates, one introduces the operators

$$M_{A;k}^{\pm} = (D_t + P_{A;k}^{\pm})\,T_k^* U_{A;k}^{\pm} T_k \quad \text{and} \quad N_{A;k}^{\pm} = [D_t^2 - (P_{A;k}^{\pm})^2]\,T_k^* U_{A;k}^{\pm} T_k, \tag{4.16}$$

cf. (4.10) and (4.14), respectively.

5. Fréchet derivative of $S_A(t)$: Proof of Theorem 2.2. We begin with some general considerations. We let $X, Y = \mathbb{R}^n$ or $\mathbb{R}^{2n}$. A map $C^{1,1} \ni A \mapsto H_A$, $H_A:\ H^\alpha(X) \to H^{\alpha'}(Y)$, is Fréchet differentiable at A if

1. H_A is continuous at A,
2. there is a map, $(DH)_{A,\Delta}:\ H^\alpha(X) \to H^{\alpha'}(Y)$, which acts linearly in $\Delta = B - A$, such that

$$\|H_B - H_A - (DH)_{A,\Delta}\|_{L(H^\alpha(X),H^{\alpha'}(Y))} = o(\|\Delta\|_{C^{1,1}}), \tag{5.1}$$

with Δ satisfying $\|\Delta\|_{C^{1,1}} < \varepsilon$ for some $\epsilon > 0$.

Instead of (5.1), one typically establishes that

$$\|(H_B - H_A - (DH)_{A,\Delta})g\|_{H^{\alpha'}} \leq C(\Delta)\,\|g\|_{H^\alpha} \quad \text{for all } g \in H^\alpha, \tag{5.2}$$

in which

$$C(\Delta) = o(\|\Delta\|_{C^{1,1}}). \tag{5.3}$$

We adopt notation

$$(RH)_{A,\Delta} := (H_B - H_A) - (DH)_{A,\Delta}. \tag{5.4}$$

For the operators relevant to our analysis we establish, in fact, that

$$C(\Delta) \lesssim \|\Delta\|^2_{C^{0,1}}. \tag{5.5}$$

Moreover, we will refine the measure of the coefficient perturbation; that is, we will use nested spaces, $L^\infty(\mathbb{R}^n)$, and $C^{m-1,1}(\mathbb{R}^n)$ with norms $\|a^{ij}\|_{C^{m-1,1}} = \sum_{|\alpha|\le m} \|\partial^\alpha a^{ij}\|_{L^\infty}$ for $m = 1, 2, \ldots$, with $\ldots \subset C^{1,1} \subset C^{0,1} \subset L^\infty \subset \ldots$

We simplify the notation, by denoting the Fréchet derivative of operators $H_{A;k}$ by $(DH)_{A,\Delta}$ instead of $(DH)_{A,\Delta;k}$ whenever it is clear from the context. Furthermore, for simplicity of notation, we will suppress the superscripts $^\pm$. We define $\bar{\Delta}(x, D_x) = B(x, D_x) - A(x, D_x) = \sum_{i,j=1}^n \Delta^{ij}(x) D_{x_i} D_{x_j}$. In the following lemmata, we will use functions $f \in L^2$ the Fourier transform $\hat{f}$ of which is understood to be supported in $|\xi| \approx 2^k$. Such a function is reminiscent of the Littlewood-Paley decomposition of g. We will follow the diagram in Fig. 4.

Lemma 5.1. *Let $(DP)_{A,\Delta}$ and $(DQ)_{A,\Delta}$ be the pseudodifferential operators with symbols*

$$(Dp)_{A,\Delta}(x,\xi) = \chi(2^{-k/2} D_x) \left[\frac{\bar{\Delta}_k(x,\xi)}{\sqrt{A_k(x,\xi)}} \right], \tag{5.6}$$

$$(Dq)_{A,\Delta}(x,\xi) = \chi(2^{-k/2} D_x) \left[\frac{(Dp)_{A,\Delta}(x,\xi)}{p^2_{A;k}(x,\xi)} \right], \tag{5.7}$$

and $(DR)_{A,\Delta}$ be the pseudodifferential operator defined through the composition,

$$(DR)_{A,\Delta} = (DP)_{A,\Delta} Q_{A;k} + P_{A;k} (DQ)_{A,\Delta}. \tag{5.8}$$

If $f \in L^2$ and $\hat{f}$ is supported in $|\xi| \approx 2^k$, then

$$\|(DP)_{A,\Delta} f\|_{L^2} \lesssim 2^k \|\Delta\|_{L^\infty} \|f\|_{L^2}, \tag{5.9}$$
$$\|(DQ)_{A,\Delta} f\|_{L^2} \lesssim 2^{-k} \|\Delta\|_{L^\infty} \|f\|_{L^2}, \tag{5.10}$$
$$\|(DR)_{A,\Delta} f\|_{L^2} \lesssim \|\Delta\|_{L^\infty} \|f\|_{L^2}, \tag{5.11}$$

while

$$\|(RP)_{A,\Delta} f\|_{L^2} \lesssim 2^k \|\Delta\|^2_{L^\infty} \|f\|_{L^2}, \tag{5.12}$$
$$\|(RQ)_{A,\Delta} f\|_{L^2} \lesssim 2^{-k} \|\Delta\|^2_{L^\infty} \|f\|_{L^2}, \tag{5.13}$$
$$\|(RR)_{A,\Delta} f\|_{L^2} \lesssim \|\Delta\|^2_{L^\infty} \|f\|_{L^2}. \tag{5.14}$$

Proof. We consider a function

$$\begin{aligned}\tilde{h}(x,\xi) = (2^{kn/2} \chi(2^{k/2}\,.\,) \overset{(x)}{*} [F(B_k(\,.\,,\xi)) - F(A_k(\,.\,,\xi)) \\ - F'(A_k(\,.\,,\xi))(B_k - A_k)(\,.\,,\xi)])\, \tilde{\beta}_k(\xi),\end{aligned}$$

where $\tilde{\beta}_k$ is a cutoff to $|\xi| \approx 2^k$, and $F(t) = t^{1/2}$. We wish to show that

$$|\partial_x^\alpha \partial_\xi^\beta \tilde{h}(x,\xi)| \le C_{M,\alpha,\beta} \|B - A\|^2_{L^\infty} (2^k)^{1-|\beta|+\frac{1}{2}|\alpha|}, \tag{5.15}$$

which implies estimate (5.12). In $\partial_x^\alpha \partial_\xi^\beta \tilde{h}$ we let the x-derivatives, in the convolution, act on the mollifier, χ, which yields the desired growth. We now assume that $\alpha = 0$. Each ξ-derivative acting on $\tilde{\beta}_k(\xi)$ generates a factor 2^{-k}. It remains to consider the ξ-derivatives acting on the factor $F(B_k) - F(A_k) - F'(A_k)(B_k - A_k)$. We write

$$F(B_k) - F(A_k) - F'(A_k)(B_k - A_k) = \int_0^1 (1-r) F''(rB_k + (1-r)A_k)(B_k - A_k)^2 \, \mathrm{d}r. \tag{5.16}$$

The matrix $r b_k^{ij} + (1-r) a_k^{ij}$ satisfies (2.1), and it appears natural to introduce the one-parameter family of symbols,

$$C_{r;k}(x,\xi) = rB_k(x,\xi) + (1-r)A_k(x,\xi), \quad r \in [0,1].$$

Moreover, the integrand is homogeneous of degree 1 in ξ. Hence, observing that $\|B_k - A_k\|_{L^\infty} \lesssim \|B - A\|_{L^\infty}$,

$$|(\partial_\xi^\beta [F(B_k) - F(A_k) - F'(A_k)(B_k - A_k)]) \, \tilde{\beta}_k(\xi)| \le C_{M,\beta} \|B - A\|_{L^\infty}^2 (2^k)^{1-|\beta|}.$$

The constant in this inequality depends on M through the lower bounds for the symbols of A and B (cf. (2.1)). We then identify $(Dp)_{A,\varDelta}$ in (5.6) with $\chi(2^{-k/2}D) F'(A_k) \, \bar{\varDelta}_k$.

To obtain estimate (5.9), we consider the function

$$\check{h}(x,\xi) = (2^{kn/2} \chi(2^{k/2}\, .\,) \overset{(x)}{*} [F'(A_k(\,.\,,\xi))(B_k - A_k)(\,.\,,\xi)]) \, \tilde{\beta}_k(\xi).$$

It follows that

$$|\partial_x^\alpha \partial_\xi^\beta \check{h}(x,\xi)| \le C_{M,\alpha,\beta} \|B - A\|_{L^\infty} (2^k)^{1-|\beta|+\frac{1}{2}|\alpha|}, \tag{5.17}$$

from which estimate (5.9) follows.

Next, we consider the operator difference, $Q_{B;k} - Q_{A;k}$. For the symbols, we have

$$q_{B;k} - q_{A;k} = \chi(2^{-k/2}D) \left(\frac{1}{p_{B;k}} - \frac{1}{p_{A;k}} \right).$$

We note that

$$\frac{1}{p_{B;k}} - \frac{1}{p_{A;k}} = \frac{p_{B;k} - p_{A;k}}{p_{A;k}^2} - \frac{(p_{B;k} - p_{A;k})^2}{p_{A;k}^2 p_{B;k}}.$$

With (5.7) we arrive at the remainder (cf. (5.4))

$$(Rq)_{A,\varDelta} = \chi(2^{-k/2}D) \left[\frac{(Rp)_{A,\varDelta}}{p_{A;k}^2} - \frac{(p_{B;k} - p_{A;k})^2}{p_{A;k}^2 p_{B;k}} \right].$$

The symbol in (5.7) satisfies the estimate

$$|\partial_x^\alpha \partial_\xi^\beta (Dq)_{A,\varDelta}(x,\xi) \, \tilde{\beta}_k(\xi)| \le C_{M,\alpha,\beta} \|B - A\|_{L^\infty}^2 (2^k)^{-1-|\beta|+\frac{1}{2}|\alpha|},$$

from which (5.10) follows. Moreover, it can be shown that $(p_{B;k} - p_{A;k})(x,\xi) \, \tilde{\beta}_k(\xi)$ satisfies the estimates for $\check{h}(x,\xi)$ in (5.17). Using this estimate, the estimate for $(Dp)_{A,\varDelta}$, and standard symbol calculus yields the result,

$$|\partial_x^\alpha \partial_\xi^\beta (Rq)_{A,\varDelta}(x,\xi) \, \tilde{\beta}_k(\xi)| \le C_{M,\alpha,\beta} \|B - A\|_{L^\infty}^2 (2^k)^{-1-|\beta|+\frac{1}{2}|\alpha|}.$$

We then consider the operator difference, $R_{B;k} - R_{A;k}$. Suppressing the cutoffs, β_k, we have (cf. (4.11))

$$R_{B;k} - R_{A;k} = P_{B;k}Q_{B;k} - P_{A;k}Q_{A;k} = (P_{B;k} - P_{A;k})\,Q_{A;k} \\ + P_{A;k}\,(Q_{B;k} - Q_{A;k}) + (P_{B;k} - P_{A;k})\,(Q_{B;k} - Q_{A;k}).$$

It follows that the remainder (cf. (5.4)) attains the form

$$(RR)_{A,\Delta} = (RP)_{A,\Delta}\,Q_{A;k} + P_{A;k}\,(RQ)_{A,\Delta} + (P_{B;k} - P_{A;k})\,(Q_{B;k} - Q_{A;k}).$$

By composition, the estimates above imply (5.11) and (5.14). □

We now consider the Fréchet derivative of $U_{k,A}(t)\,T_k$ (cf. 4.6). To this end, we need to perturb the reverse Hamiltonian flow, $\Phi_{A;k}(t)$. To arrive at this perturbation, we introduce the one-parameter family of symbols,

$$C_r(x,\xi) = rB(x,\xi) + (1-r)A(x,\xi), \quad r \in [0,1],$$

defining, upon smoothing, the Hamiltonians

$$p_{C_r;k}(x,\xi) = \chi(2^{-k/2}D_x)\,\sqrt{C_{r;k}(x,\xi)}$$

(cf. (4.2)). These define the smooth family, $(\Phi_{C_r;k})_{r\in[0,1]}$, of symplectic diffeomorphisms on $T^*\mathbb{R}^n$, where $\Phi_{C_r;k}(t) = \chi_{0,t}^{C_r;k}$, with the flow
$\chi_{t,0}^{C_r;k} :\ (x,\xi) \mapsto (x_r(t;x,\xi), \xi_r(t;x,\xi))$ being generated by the Hamilton system,

$$\dot{x}(t) = \partial_\xi p_{C_r;k}(x(t),\xi(t)),$$
$$\dot{\xi}(t) = -\partial_x p_{C_r;k}(x(t),\xi(t)),$$

subject to initial conditions $(x(0),\xi(0)) = (x,\xi)$. We differentiate the Hamilton system, noting that

$$\partial_r p_{C_r;k} = -\chi(2^{-k/2}D_x)\,\frac{B_k(x,\xi) - A_k(x,\xi)}{2\sqrt{C_{r;k}(x,\xi)}} = -\chi(2^{-k/2}D_x)\,\frac{\bar{\Delta}_k(x,\xi)}{2\sqrt{C_{r;k}(x,\xi)}}.$$

For the Hamiltonian, and its first-order derivatives, we have the estimates

$$|\partial_x^\alpha \partial_\xi^\beta p_{C_r;k}(x,\xi)| \le C_{M,\alpha,\beta}(2^k)^{1-|\beta|+\frac{1}{2}\max(0,|\alpha|-2)}, \tag{5.18}$$

$$|\partial_x^\alpha \partial_\xi^\beta (\partial_r p_{C_r;k})(x,\xi)| \le C_{M,\alpha,\beta}\|B-A\|_{C^{0,1}}(2^k)^{1-|\beta|+\frac{1}{2}\max(0,|\alpha|-1)}, \tag{5.19}$$

or

$$|\partial_x^\alpha \partial_\xi^\beta (\partial_r p_{C_r;k})(x,\xi)| \le C_{M,\alpha,\beta}\|B-A\|_{C^{1,1}}(2^k)^{1-|\beta|+\frac{1}{2}\max(0,|\alpha|-2)}, \tag{5.20}$$

for $|\xi| \approx 2^k$ (compare, also, (5.17)). The first-order perturbation of the flow is then described by the Hamilton-Jacobi equations,

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \partial_r x_r \\ \partial_r \xi_r \end{pmatrix} = \begin{pmatrix} \partial_x\partial_\xi p_{C_r;k} & \partial_\xi^2 p_{C_r;k} \\ -\partial_x^2 p_{C_r;k} & -\partial_\xi\partial_x p_{C_r;k} \end{pmatrix}\begin{pmatrix} \partial_r x_r \\ \partial_r \xi_r \end{pmatrix} + \begin{pmatrix} \partial_\xi\partial_r p_{C_r;k} \\ -\partial_x\partial_r p_{C_r;k} \end{pmatrix}, \tag{5.21}$$

subject to substituting for (x,ξ) in the expressions on the right-hand side the solution $(x_r(t;x,\xi),\xi_r(t;x,\xi))$, and supplemented by the initial conditions $(\partial_r x_r(0), \partial_r\xi_r(0)) = (0,0)$. We use the notation $\Theta_k =$

$\Theta_k(r;t,x,\xi) = (x_r(t;x,\xi),\xi_r(t;x,\xi))$ for bicharacteristics, and denote the fundamental matrix associated with system (5.21) by $\Psi_r(t)$. Then

$$(\partial_r\Theta_k)(t) = \int_0^t \Psi_r(t)\,\Psi_r(t')^{-1} b_r(t')\,\mathrm{d}t', \quad \text{with } b_r = \begin{pmatrix} \partial_\xi\partial_r p_{C_r;k} \\ -\partial_x\partial_r p_{C_r;k} \end{pmatrix}$$

(cf. (5.21)). In particular, $(\partial_r\Theta_k)_{r=0}(-t)$ defines $(D\Phi)_{A;\Delta}(t)(x,\xi)$; we write

$$(D\Phi)_{A;\Delta,1}(t)(x,\xi) = (\partial_r x_r)_{r=0}(-t;x,\xi), \quad (D\Phi)_{A;\Delta,2}(t)(x,\xi) = (\partial_r\xi_r)_{r=0}(-t;x,\xi).$$

To obtain estimates, we compensate for the fact that if $|\xi(0)| = |\xi| \approx 2^k$ then $|\xi(t;x,\xi)| \approx 2^k$, as usual for $|t| \le M$, by redefining Θ_k as

$$\Theta'_k = \Theta'_k(r;t,x,\xi) = (x_r(t;x,\xi), 2^{-k}\xi_r(t;x,\xi))$$

with $\Theta'_k(r;0,x,\xi) = (x, 2^{-k}\xi)$. Using estimates (5.18)-(5.19), system (5.21) and the homogeneity of the Hamiltonian, we find that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\partial_r\Theta'_k(t)\right| \lesssim |\partial_r\Theta'_k(t)| + \|B-A\|_{C^{0,1}}. \tag{5.22}$$

Because $\partial_r\Theta'_k(0) = 0$, Gronwall's lemma implies that $|\partial_r\Theta'_k(t)| \lesssim \|B-A\|_{C^{0,1}}$, leading to

$$|(\partial_r x_r)(t;x,\xi)| \lesssim \quad \|B-A\|_{C^{0,1}}, \tag{5.23}$$

$$|(\partial_r\xi_r)(t;x,\xi)| \lesssim 2^k\|B-A\|_{C^{0,1}}. \tag{5.24}$$

An expression for the remainder, $(R\Phi)_{A;\Delta}$, is obtained by considering the Taylor expansions,

$$(x_B - x_A)(t;x,\xi) = (\partial_r x_r)_{r=0}(t;x,\xi) + \int_0^1 (1-r)(\partial_r^2 x_r)(t;x,\xi)\,\mathrm{d}r, \tag{5.25}$$

$$(\xi_B - \xi_A)(t;x,\xi) = (\partial_r\xi_r)_{r=0}(t;x,\xi) + \int_0^1 (1-r)(\partial_r^2\xi_r)(t;x,\xi)\,\mathrm{d}r. \tag{5.26}$$

To estimate the integrals on the right-hand sides, we need to develop the second-order perturbation of the flow, which requires taking the second-order derivative,

$$\partial_r^2 p_{C_r;k} = \chi(2^{-k/2}D_x)\,\frac{[B_k(x,\xi) - A_k(x,\xi)]^2}{4C_{r;k}(x,\xi)^{3/2}} = \chi(2^{-k/2}D_x)\,\frac{\bar{\Delta}_k(x,\xi)^2}{4C_{r;k}(x,\xi)^{3/2}},$$

satisfying the estimate

$$|\partial_x^\alpha\partial_\xi^\beta(\partial_r^2 p_{C_r;k})(x,\xi)| \le C_{M,\alpha,\beta}\|B-A\|^2_{C^{0,1}}(2^k)^{1-|\beta|+\frac{1}{2}\max(0,|\alpha|-1)}, \tag{5.27}$$

for $|\xi| \approx 2^k$ (compare, also, (5.15)). Taking the derivative of (5.21) with respect to r, using estimates (5.18), (5.19) and (5.27), and the homogeneity of the Hamiltonian, leads to

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}(\partial_r^2 x_r)(t;x,\xi)\right| \lesssim |(\partial_r^2 x_r)(t;x,\xi)| + 2^{-k}|(\partial_r^2\xi_r)(t;x,\xi)| + \|B-A\|^2_{C^{0,1}},$$
$$\left|\frac{\mathrm{d}}{\mathrm{d}t}(\partial_r^2\xi_r)(t;x,\xi)\right| \lesssim 2^k|(\partial_r^2 x_r)(t;x,\xi)| + |(\partial_r^2\xi_r)(t;x,\xi)| + 2^{3k/2}\|B-A\|^2_{C^{0,1}};$$

hence,

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\partial_r^2\Theta'_k(t)\right| \lesssim |\partial_r^2\Theta'_k(t)| + 2^{k/2}\|B-A\|^2_{C^{0,1}}. \tag{5.28}$$

Because $\partial_r^2 \Theta_k'(0) = 0$, Gronwall's lemma implies that $|\partial_r^2 \Theta_k'(t)| \lesssim 2^{k/2} \|B - A\|_{C^{0,1}}^2$, leading to

$$|(\partial_r^2 x_r)(t; x, \xi)| \lesssim 2^{k/2} \ \|B - A\|_{C^{0,1}}^2, \tag{5.29}$$

$$|(\partial_r^2 \xi_r)(t; x, \xi)| \lesssim 2^{3k/2} \|B - A\|_{C^{0,1}}^2. \tag{5.30}$$

These estimates carry over directly to $(R\Phi)_{A;\Delta}$ using (5.25)-(5.26).

The transform in (3.2), initiating the approximate solution via $U_{A;k}^{\pm}(t)$ (cf. (4.5)), has the properties

$$\partial_x^\alpha T_k f = T_k(\partial^\alpha f), \quad \partial_\xi^\beta T_k f = (2^k)^{-\frac{1}{2}|\beta|} \, \widetilde{T}_k^\beta f, \tag{5.31}$$

where

$$\widetilde{T}_k^\beta f(x, \xi) = \int f(y) \overline{(\widetilde{\phi}^\beta)_k(y; x, \xi)} \, \mathrm{d}y, \quad \widetilde{\phi}^\beta(y) = (\mathrm{i}y)^\beta \, \phi(y), \tag{5.32}$$

satisfying bounds similar to those of T_k. These properties are used in

Lemma 5.2. *Let $\widetilde{U}_{A;k}(t) = U_{A;k}(t) T_k$, and $(D\widetilde{U})_{A,\Delta} : \ H^\alpha(\mathbb{R}^n) \to H^{\alpha-1}(\mathbb{R}^{2n})$ be the mapping defined according to* [1]

$$\begin{aligned}((D\widetilde{U})_{A,\Delta} f)(x, \xi) = \sum_{j=1}^n (D\Phi)_{A;\Delta,1,j}(x, \xi) \, (T_k(\partial_{x_j} f))(\Phi_{A;k}(x, \xi)) \\ + 2^{-k/2} \sum_{j=1}^n (D\Phi)_{A;\Delta,2,j}(x, \xi) \, (\widetilde{T}_k^{\varepsilon^j} f)(\Phi_{A;k}(x, \xi)).\end{aligned} \tag{5.33}$$

If $f \in L^2$ and $\hat{f}$ is supported in $|\xi| \approx 2^k$, we have

$$\|(D\widetilde{U})_{A,\Delta} \, f\|_{L^2(\mathbb{R}^{2n}_{(x,\xi)})} \lesssim 2^k \, \|\Delta\|_{C^{0,1}} \|f\|_{L^2}, \tag{5.34}$$

while

$$\|((U_{B;k} - U_{A;k}) \, T_k - (D\widetilde{U})_{A,\Delta}) \, f\|_{L^2(\mathbb{R}^{2n}_{(x,\xi)})} \lesssim (2^k)^2 \, \|\Delta\|_{C^{0,1}}^2 \|f\|_{L^2}. \tag{5.35}$$

Proof. Using the family, $(\Phi_{C_r;k})_{r \in [0,1]}$, of symplectic diffeomorphisms on $T^*\mathbb{R}^n$, we can write $((D\widetilde{U})_{A,\Delta} f)(x, \xi)$ in the form,

$$\begin{aligned}(\partial_r T_k f(\Phi_{C_r;k}(x, \xi)))_{r=0} = \sum_{j=1}^n (\partial_r x_{r;j})_{r=0}(.; x, \xi) \, (T_k(\partial_{x_j} f))(\Phi_{A;k}(x, \xi)) \\ + 2^{-k/2} \sum_{j=1}^n (\partial_r \xi_{r;j})_{r=0}(.; x, \xi) \, (\widetilde{T}_k^{\varepsilon^j} f)(\Phi_{A;k}(x, \xi)).\end{aligned} \tag{5.36}$$

We use estimates (5.23)-(5.24) and account for the derivative, $\partial_{x_j} f$, from which it follows that the right-hand side is dominated by the first summation, and obtain (5.34).

[1] We use the multi-index notation, ε^j: $\varepsilon^j_l = 0$ if $l \neq j$, while $\varepsilon^j_j = 1$. Moreover we use the notation ε^{ij}: If $i \neq j$ then $\varepsilon^{ij}_l = 0$ if $l \neq i$ and $l \neq j$, while $\varepsilon^{ij}_i = \varepsilon^{ij}_j = 1$; otherwise, $\varepsilon^{ii}_l = 0$ if $l \neq i$ and $\varepsilon^{ii}_i = 2$.

We can now write the remainder, $((R\widetilde{U})_{A,\Delta}f)(x,\xi)$, in the form,

$$
\begin{aligned}
(5.37)\quad & \int_0^1 \int_0^{r'} \partial_r \Bigg[\sum_{j=1}^{n} (\partial_r x_{r;j})(.;x,\xi)\,(T_k(\partial_{x_j} f))(\Phi_{C_r;k}(x,\xi)) \\
&\qquad + 2^{-k/2} \sum_{j=1}^{n} (\partial_r \xi_{r;j})(.;x,\xi)\,(\widetilde{T}_k^{\varepsilon^j} f)(\Phi_{C_r;k}(x,\xi)) \Bigg]\, \mathrm{d}r\mathrm{d}r' \\
&= \int_0^1 \int_0^{r'} \Bigg[\sum_{i,j=1}^{n} (\partial_r x_{r;i})(.;x,\xi)(\partial_r x_{r;j})(.;x,\xi)\,(T_k(\partial_{x_i}\partial_{x_j} f))(\Phi_{C_r;k}(x,\xi)) \\
&\quad + 2^{-k/2} \sum_{i,j=1}^{n} (\partial_r \xi_{r;i})(.;x,\xi)(\partial_r x_{r;j})(.;x,\xi)\,(\widetilde{T}_k^{\varepsilon^i}(\partial_{x_j} f))(\Phi_{C_r;k}(x,\xi)) \\
&\qquad + \sum_{j=1}^{n} (\partial_r^2 x_{r;j})(.;x,\xi)\,(T_k(\partial_{x_j} f))(\Phi_{C_r;k}(x,\xi)) \\
&\quad + 2^{-k/2} \sum_{i,j=1}^{n} (\partial_r x_{r;i})(.;x,\xi)(\partial_r \xi_{r;j})(.;x,\xi)\,(\widetilde{T}_k^{\varepsilon^j}(\partial_{x_i} f))(\Phi_{C_r;k}(x,\xi)) \\
&\qquad + 2^{-k} \sum_{i,j=1}^{n} (\partial_r \xi_{r;i})(.;x,\xi)(\partial_r \xi_{r;j})(.;x,\xi)\,(\widetilde{T}_k^{\varepsilon^{ij}} f)(\Phi_{C_r;k}(x,\xi)) \\
&\qquad\qquad + 2^{-k/2} \sum_{j=1}^{n} (\partial_r^2 \xi_{r;j})(.;x,\xi)\,(\widetilde{T}_k^{\varepsilon^j} f)(\Phi_{C_r;k}(x,\xi)) \Bigg]\, \mathrm{d}r\mathrm{d}r'.
\end{aligned}
$$

We use estimates (5.23)-(5.24) and (5.29)-(5.30), and account for the derivatives, $\partial_{x_i} f$, $\partial_{x_j} f$ and $\partial_{x_i}\partial_{x_j} f$, from which it follows that the right-hand side is dominated by the first summation, and obtain (5.35). ∎

Making use of the fact that T_k^* is an isometry, it is immediate that

$$\|(T_k^*(D\widetilde{U})_{A,\Delta})\, f\|_{L^2} \lesssim 2^k\, \|\Delta\|_{C^{0,1}} \|f\|_{L^2},$$

and that

$$\|(T_k^*(U_{B;k} - U_{A;k})\, T_k - T_k^*(D\widetilde{U})_{A,\Delta})\, f\|_{L^2} \lesssim 2^{2k}\, \|\Delta\|_{C^{0,1}}^2 \|f\|_{L^2},$$

assuming that $f \in L^2$ and that $\hat{f}$ is supported in $|\xi| \approx 2^k$. Together with Lemma 5.1, this leads to (cf. (4.5))

Lemma 5.3. *Let $(DE)_{A,\Delta} :\ H^\alpha(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$ be the mapping defined according to*

$$(5.38)\quad (DE)_{A,\Delta}(t)g \;=\; T_k^*(D\widetilde{U})_{A,\Delta}(t)\,(\tfrac{1}{2}\mathrm{i}\, Q_{A;k}\beta_k(D)g) \;+\; (T_k^* U_{A;k}(t) T_k)(\tfrac{1}{2}\mathrm{i}\,(DQ)_{A,\Delta}\beta_k(D)g).$$

We have

$$(5.39)\qquad \|(DE)_{A,\Delta}(t)g\|_{L^2} \lesssim \|\Delta\|_{C^{0,1}} \|\beta_k(D)g\|_{L^2},$$

while

$$(5.40)\qquad \|(RE)_{A,\Delta} g\|_{L^2} \lesssim 2^k\, \|\Delta\|_{C^{0,1}}^2 \|\beta_k(D)g\|_{L^2}.$$

Following (4.7) then leads to the Fréchet derivative, $(D\widetilde{S})_{A,\Delta}(t) :\ H^\alpha(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$, of $\widetilde{S}_A(t)$,

$$(5.41)\qquad (D\widetilde{S})_{A,\Delta}(t)g = \sum_{\pm} \sum_{k \geq k_0} (DE)_{A,\Delta}(t)g.$$

Since $(DE)_{A,\Delta}(t)g$ is localized near $|\xi| \approx 2^k$, the sum over scales converges in $H^\alpha(\mathbb{R}^n)$ [2]. The Fréchet derivative satisfies the estimate

$$\|(D\widetilde{S})_{A,\Delta}(t)g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{0,1}}\|g\|_{H^\alpha}, \tag{5.42}$$

while

$$\|(R\widetilde{S})_{A,\Delta}(t)g\|_{H^{\alpha-1}} \lesssim \|\Delta\|^2_{C^{0,1}}\|g\|_{H^\alpha}. \tag{5.43}$$

To obtain the Fréchet derivative of $\widehat{S}_A(t)$ (cf. (4.8)), we need to consider the operators $M_{A;k}$ defined in (4.16). Using (4.6), we note that

$$M_{A;k}f = \iint (T_k f)(x,\xi) \\ ([P_{A;k}(y,D_y) + L^{A;k}_{\Phi_{A;k}(t)(x,\xi)}(D_x,D_\xi)]\,\phi_k)(y;\Phi_{A;k}(t)(x,\xi))\,\mathrm{d}x\mathrm{d}\xi, \tag{5.44}$$

where

$$L^{A;k}_{x,\xi}(D_x,D_\xi) = \langle(\partial_\xi p_{A;k})(x,\xi),D_x\rangle - \langle(\partial_x p_{A;k})(x,\xi),D_\xi\rangle. \tag{5.45}$$

Within the integral on the right-hand side of (5.44), we analyze

$$[P_{A;k}(y,D_y) + L^{A;k}_{x,\xi}(D_x,D_\xi)]\,\phi_k(y;x,\xi) =: (\phi^A_{x,\xi})_k(y;x,\xi).$$

Following [15, Lemma 5.3], $\phi^A_{x,\xi}$ can be generated with a pseudodifferential operator, $m^{A;k}_{x,\xi}(z,D_z)$, with symbol

$$m^{A;k}_{x,\xi}(z,\zeta) = \int_0^1 (1-s)\,\partial_s^2[p_{A;k}(x+s\,2^{-k/2}z,\xi+s\,2^{k/2}\zeta)]\,\mathrm{d}s; \tag{5.46}$$

that is,

$$\phi^A_{x,\xi}(z) = m^{A;k}_{x,\xi}(z,D_z)\phi(.), \tag{5.47}$$

from which $(\phi^A_{x,\xi})_k(y;x,\xi)$ is obtained according to (3.1). Now we can write $M_{A;k}f = (T^A_k)^* U_{A;k} T_k f$, with

$$(T^A_k)^* F(y) = \int F(x,\xi)(\phi^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi, \tag{5.48}$$

cf. Lemma 3.1 upon substituting $T'_k = T^A_k$. Thus, in the process of constructing Fréchet derivatives, we need to differentiate $(T^A_k)^*$, and, hence, $\phi^A_{x,\xi}$:

$$(D\phi_{x,\xi})_{A,\Delta}(z) = (2\pi)^{-n}\int \mathrm{e}^{\mathrm{i}\langle\zeta,z\rangle}\,(Dm_{x,\xi})_{A,\Delta}(z,\zeta)\,\hat{\phi}(\zeta)\,\mathrm{d}\zeta, \tag{5.49}$$

where

$$(Dm_{x,\xi})_{A,\Delta}(z,\zeta) = \int_0^1 (1-s)\,\partial_s^2[(Dp)_{A,\Delta}(x+s\,2^{-k/2}z,\xi+s\,2^{k/2}\zeta)]\,\mathrm{d}s. \tag{5.50}$$

[2] For the Sobolev norms, we have $\|g\|^2_{H^\alpha} \sim \sum_k (2^k)^{2\alpha}\|g_k\|^2_{L^2}$.

In a similar fashion, the remainder can be expressed as

$$(R\phi_{x,\xi})_{A,\Delta}(z) = \phi^B_{x,\xi}(z) - \phi^A_{x,\xi}(z) - (D\phi_{x,\xi})_{A,\Delta}(z) = (2\pi)^{-n} \int e^{i\langle\zeta,z\rangle}\, (Rm_{x,\xi})_{A,\Delta}(z,\zeta)\, \hat{\phi}(\zeta)\, d\zeta, \tag{5.51}$$

with

$$(Rm_{x,\xi})_{A,\Delta}(z,\zeta) = \int_0^1 (1-s)\, \partial_s^2[(Rp)_{A,\Delta}(x + s\,2^{-k/2}z, \xi + s\,2^{k/2}\zeta)]\, ds. \tag{5.52}$$

While making use of the symbol estimates in the proof of Lemma 5.1, we obtain

LEMMA 5.4. *The Schwartz seminorms of $(D\phi_{x,\xi})_{A,\Delta}$ are $\lesssim 2^k\|\Delta\|_{L^\infty}$, uniformly in x and ξ. Moreover, the Schwartz seminorms of $(R\phi_{x,\xi})_{A,\Delta}$ are $\lesssim 2^k\|\Delta\|^2_{L^\infty}$, uniformly in x and ξ.*

Proof. We use elements of the proofs in [15, Lemma 5.3, Lemma 6.2], and set

$$(\tilde{x}, \tilde{\xi}) = (x + s\,2^{-k/2}z, \xi + s\,2^{k/2}\zeta),$$

so that the integrand of (5.50) is given by

$$\partial_s^2[(Dp)_{A,\Delta}(\tilde{x},\tilde{\xi})] = \sum_{j,k} \Big[(\partial_{x_j}\partial_{x_k}(Dp)_{A,\Delta})(\tilde{x},\tilde{\xi})\; 2^{-k} z_j z_k + (\partial_{x_j}\partial_{\xi_k}(Dp)_{A,\Delta})(\tilde{x},\tilde{\xi})\; z_j\zeta_k + (\partial_{\xi_j}\partial_{\xi_k}(Dp)_{A,\Delta})(\tilde{x},\tilde{\xi})\; 2^k \zeta_j\zeta_k \Big].$$

In view of the support of $\hat{\phi}$, we only need to consider $|\zeta| < 2$. Applying (5.17), and observing that $2^{k/2}|\zeta| \ll 2^k \approx |\xi|$, we obtain

$$|\partial_z^\alpha \partial_\zeta^\beta\; \partial_s^2[(Dp)_{A,\Delta}(\tilde{x},\tilde{\xi})]| \lesssim \|B - A\|_{L^\infty}\, (2^k)^{1-\frac{1}{2}|\beta|} \langle z\rangle^2.$$

Using (5.50), these estimates carry directly over to

$$|(\partial_z^\alpha \partial_\zeta^\beta (Dm_{x,\xi})_{A,\Delta})(z,\zeta)| \lesssim (2^k)^{1-\frac{1}{2}|\beta|}\, \|B-A\|_{L^\infty}\, \langle z\rangle^2, \qquad |\zeta| < 2. \tag{5.53}$$

The Schwartz seminorms of $(D\phi_{x,\xi})_{A,\Delta}$ are determined by this estimate, upon integration by parts.
In a similar fashion, applying (5.15) to (5.52) we obtain,

$$|(\partial_z^\alpha \partial_\zeta^\beta (Rm_{x,\xi})_{A,\Delta})(z,\zeta)| \lesssim (2^k)^{1-\frac{1}{2}|\beta|}\, \|B-A\|^2_{L^\infty}\, \langle z\rangle^2, \qquad |\zeta| < 2. \tag{5.54}$$

Again, the Schwartz seminorms of $(R\phi_{x,\xi})_{A,\Delta}$ are determined by this estimate using integration by parts. ∎

We have now the tools to analyze the Fréchet derivative of the residual operator $(D\tilde{R})_{A,\Delta}$, following (4.10), which we will need at zero (initial) time (cf. (4.9) while $\Phi_{A;k}(0) = I$):

LEMMA 5.5. *Let $((D\widetilde{R})_{A,\Delta})(0) :\ H^\alpha(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$ be the operator defined by*

$$(D\widetilde{R})_{B,\Delta}(0)g = \tfrac{1}{2}i\Big[\iint (T_k(DQ)_{A,\Delta}\beta_k(D)g)(x,\xi)(\phi^A_{x,\xi})_k(y;x,\xi)\, dx d\xi + \iint (T_k Q_{A;k}\beta_k(D)g)(x,\xi)(D\phi_{x,\xi})_{A,\Delta}(y;x,\xi)\, dx d\xi \Big]; \tag{5.55}$$

we have

$$\|(D\widetilde{R})_{A,\Delta}(0)g\|_{L^2} \lesssim \|\Delta\|_{L^\infty}\|\beta_k(D)g\|_{L^2}, \tag{5.56}$$

while

$$\|(R\widetilde{R})_{A,\Delta}(0)g\|_{L^2} \lesssim \|\Delta\|^2_{L^\infty}\|\beta_k(D)g\|_{L^2}. \tag{5.57}$$

Proof. We have

$$\begin{aligned}(\widetilde{R}_{B;k}(0) - \widetilde{R}_{A;k}(0))g = \tfrac{1}{2}\mathrm{i}\Big[&\iint (T_k(Q_{B;k} - Q_{A;k})\beta_k(D)g)(x,\xi)(\phi^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ &+ \iint (T_kQ_{A;k}\beta_k(D)g)(x,\xi)(\phi^B_{x,\xi} - \phi^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ &+ \iint (T_k(Q_{B;k} - Q_{A;k})\beta_k(D)g)(x,\xi)(\phi^B_{x,\xi} - \phi^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \Big];\end{aligned}$$

this leads to the introduction of expression (5.55) for the Fréchet derivative (noting that $(D\Phi)_{A;\Delta}(0) = 0$) with remainder

$$\begin{aligned}(R\widetilde{R})_{B,\Delta}(0)g = \tfrac{1}{2}\mathrm{i}\Big[&\iint (T_k(RQ)_{A,\Delta}\beta_k(D)g)(x,\xi)(\phi^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ &+ \iint (T_kQ_{A;k}\beta_k(D)g)(x,\xi)(R\phi_{x,\xi})_{A,\Delta}(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ &+ \iint (T_k(Q_{B;k} - Q_{A;k})\beta_k(D)g)(x,\xi)(\phi^B_{x,\xi} - \phi^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \Big].\end{aligned}$$

Using the estimates in (5.13), Lemma 5.4, and [15, Lemma 6.1 and Lemma 6.2] it follows that the first term on the right-hand side is bounded by $2^{-k}\|B - A\|^2_{L^\infty}\|\beta_k(D)g\|_{L^2}$, and the second and third terms on the right-hand side are bounded by $\|B-A\|^2_{L^\infty}\|\beta_k(D)g\|_{L^2}$. We obtain (5.57). The estimate in (5.56) follows from (5.10) and Lemma 5.4. $\square$

The Fréchet derivative, $(DK)_{A,\Delta} :\ H^\alpha(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$, of K_A (cf. (4.9)) now follows to be

$$(DK)_{A,\Delta} = \sum_{\pm}\sum_{k\geq k_0}\Big[\,\mathrm{i}\,(D\widetilde{R})_{A,\Delta}(0) + \tfrac{1}{2}(DR)_{A,\Delta}\beta_k(D)\,\Big], \tag{5.58}$$

satisfying

$$\|(DK)_{A,\Delta}g\|_{H^\alpha} \lesssim \|\Delta\|_{L^\infty}\|g\|_{H^\alpha}, \tag{5.59}$$

while

$$\|(RK)_{A,\Delta}g\|_{H^{\alpha-1}} \lesssim \|\Delta\|^2_{L^\infty}\|g\|_{H^\alpha}. \tag{5.60}$$

Since

$$\begin{aligned}(I + K_B)^{-1} - (I + K_A)^{-1} = &-(I + K_A)^{-1}(K_B - K_A)(I + K_A)^{-1} \\ &- ((I + K_B)^{-1} - (I + K_A)^{-1})(K_B - K_A)(I + K_A)^{-1},\end{aligned}$$

it follows that the Fréchet derivative, $(D(I+K)^{-1})_{A,\varDelta} :\ H^\alpha(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$, of $(I+K_A)^{-1}$ is given by

$$(D(I+K)^{-1})_{A,\varDelta} = -(I+K_A)^{-1}(DK)_{A,\varDelta}(I+K_A)^{-1}, \tag{5.61}$$

and, with (5.59), satisfies the estimate

$$\|(D(I+K)^{-1})_{A,\varDelta} g\|_{H^\alpha} \lesssim \|\varDelta\|_{L^\infty} \|g\|_{H^\alpha}. \tag{5.62}$$

For the remainder,

$$\begin{aligned}(R(I+K)^{-1})_{A,\varDelta} = -(I+K_A)^{-1}(RK)_{A,\varDelta}(I+K_A)^{-1} \\ - ((I+K_B)^{-1} - (I+K_A)^{-1})(K_B - K_A)(I+K_A)^{-1},\end{aligned}$$

we obtain the estimate

$$\|(R(I+K)^{-1})_{A,\varDelta} g\|_{H^{\alpha-1}} \lesssim \|\varDelta\|_{L^\infty}^2 \|g\|_{H^\alpha}. \tag{5.63}$$

We arrive at

Lemma 5.6. *Let $(D\widehat{S})_{A,\varDelta} :\ H^\alpha(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$ be the operator defined by*

$$(D\widehat{S})_{A,\varDelta}(t) = (D\widetilde{S})_{A,\varDelta}(t)(I+K_A)^{-1} + \widetilde{S}_A(t)(D(I+K)^{-1})_{A,\varDelta}; \tag{5.64}$$

we have

$$\|(D\widehat{S})_{A,\varDelta}(t) g\|_{H^\alpha} \lesssim \|\varDelta\|_{C^{0,1}} \|g\|_{H^\alpha}, \tag{5.65}$$

while

$$\|(R\widehat{S})_{A,\varDelta}(t) g\|_{H^{\alpha-1}} \lesssim \|\varDelta\|_{C^{0,1}}^2 \|g\|_{H^\alpha}. \tag{5.66}$$

Proof. Estimates (5.65) and (5.62) imply (5.65). The remainder can be written in the form

$$\begin{aligned}(R\widehat{S})_{A,\varDelta}(t) = (R\widetilde{S})_{A,\varDelta}(I+K_A)^{-1} + \widetilde{S}_A(t)(R(I+K)^{-1})_{A,\varDelta} \\ + (\widetilde{S}_B(t) - \widetilde{S}_A(t))((I+K_B)^{-1} - (I+K_A)^{-1}).\end{aligned}$$

We apply (5.66), (5.63), and [15, (19)] and the proof of [15, Lemma 6.6] to obtain (5.66). ∎

Operator $N^{\pm}_{A;k}$ in (4.16) essentially measures, upon symbol smoothing, how accurate the solution operator $\widehat{S}_A(t)$ is. We note that D_t commutes with $P^{\pm}_{A;k}$, whence

$$D_t^2 - (P^{\pm}_{A;k})^2 = (D_t + P^{\pm}_{A;k})^2 - 2P^{\pm}_{A;k}(D_t + P^{\pm}_{A;k}).$$

Similar to (5.44), we have

$$\begin{aligned}(D_t + P_{A;k})^2\, T_k^* U_{A;k} T_k = \iint (T_k f)(x,\xi) \\ ([P_{A;k}(y, D_y) + L^{A;k}_{\varPhi_{A;k}(t)(x,\xi)}(D_x, D_\xi)]\, (\phi^A_{x,\xi})_k)(y; \varPhi_{A;k}(t)(x,\xi))\, \mathrm{d}x \mathrm{d}\xi.\end{aligned} \tag{5.67}$$

Within the integral on the right-hand side, we analyze the factor

$$[P_{A;k}(y, D_y) + L^{A;k}_{x,\xi}(D_x, D_\xi)]\, (\phi^A_{x,\xi})_k)(y; x, \xi) =: (\widetilde{\phi}^A_{x,\xi})_k(y; x, \xi).$$

Following [15, Lemma 5.5], $\widetilde{\phi}^A_{x,\xi}$ can be generated by pseudodifferential operators $m^{A;k}_{x,\xi}(z,D_z)$ (with symbol given in (5.46)) and $\widetilde{m}^{A;k}_{x,\xi}(z,D_z)$, with symbol

$$\widetilde{m}^{A;k}_{x,\xi}(z,\zeta) = \int_0^1 (1-s)\,\partial_s^2[(L^{A;k}_{x,\xi}(D_x,D_\xi)p_{A;k})(x+s\,2^{-k/2}z,\xi+s\,2^{k/2}\zeta)]\,\mathrm{d}s; \tag{5.68}$$

that is,

$$\widetilde{\phi}^A_{x,\xi}(z) = m^{A;k}_{x,\xi}(z,D_z)\phi^A_{x,\xi}(.) + \widetilde{m}^{A;k}_{x,\xi}(z,D_z)\phi(.), \tag{5.69}$$

from which $(\widetilde{\phi}^A_{x,\xi})_k(y;x,\xi)$ is obtained according to (3.1). Now we can write $(D_t+P_{A;k})^2\,T_k^* U_{A;k} T_k f = (\widetilde{T}^A_k)^* U_{A;k} T_k f$, with

$$(\widetilde{T}^A_k)^* F(y) = \int F(x,\xi)(\widetilde{\phi}^A_{x,\xi})_k(y;x,\xi)\,\mathrm{d}x\mathrm{d}\xi, \tag{5.70}$$

cf. Lemma 3.1 upon substituting $T'_k = \widetilde{T}^A_k$.

The symbol integrand in (5.68) attains the form

$$\begin{aligned}\partial_s^2[(L^{A;k}_{x,\xi}(D_x,D_\xi)p_{A;k})(\tilde{x},\tilde{\xi})] = \sum_{j,k,l}(\partial_{\xi_l}p_{A;k})(x,\xi)\Big[\,&(\partial_{x_j}\partial_{x_k}D_{x_l}p_{A;k})(\tilde{x},\tilde{\xi})\;2^{-k}z_j z_k\\ &+(\partial_{x_j}\partial_{\xi_k}D_{x_l}p_{A;k})(\tilde{x},\tilde{\xi})\;z_j\zeta_k + (\partial_{\xi_j}\partial_{\xi_k}D_{x_l}p_{A;k})(\tilde{x},\tilde{\xi})\;2^k\zeta_j\zeta_k\,\Big]\\ -\sum_{j,k,l}(\partial_{x_l}p_{A;k})(x,\xi)\Big[\,&(\partial_{x_j}\partial_{x_k}D_{\xi_l}p_{A;k})(\tilde{x},\tilde{\xi})\;2^{-k}z_j z_k + (\partial_{x_j}\partial_{\xi_k}D_{\xi_l}p_{A;k})(\tilde{x},\tilde{\xi})\;z_j\zeta_k\\ &+(\partial_{\xi_j}\partial_{\xi_k}D_{\xi_l}p_{A;k})(\tilde{x},\tilde{\xi})\;2^k\zeta_j\zeta_k\,\Big],\end{aligned}$$

by which we obtain the estimates

$$|\partial_z^\alpha\partial_\zeta^\beta\widetilde{m}^{A;k}_{x,\xi}(z,\zeta)| \lesssim (2^k)^{\frac12-\frac12|\beta|}\,\langle z\rangle^2,\qquad |\zeta|<2.$$

To develop the Fréchet derivative of $\widetilde{T}^A_k$, we need the Fréchet derivative,

$$\begin{aligned}(D\widetilde{\phi}_{x,\xi})_{A,\Delta}(z) = (Dm_{x,\xi})_{A,\Delta}(z,D_z)\phi^A_{x,\xi}(.) + m^{A;k}_{x,\xi}(z,D_z)(D\phi_{x,\xi})_{A,\Delta}(.)\\ + (D\widetilde{m}_{x,\xi})_{A,\Delta}(z,D_z)\phi(.),\end{aligned}$$

with remainder given by

$$\begin{aligned}(R\widetilde{\phi}_{x,\xi})_{A,\Delta}(z) &= \widetilde{\phi}^B_{x,\xi}(z) - \widetilde{\phi}^A_{x,\xi}(z) - (D\widetilde{\phi}_{x,\xi})_{A,\Delta}(z)\\ &= (Rm_{x,\xi})_{A,\Delta}(z,D_z)\phi^A_{x,\xi}(.) + m^{A;k}_{x,\xi}(z,D_z)(R\phi_{x,\xi})_{A,\Delta}(.)\\ &\quad + (R\widetilde{m}_{x,\xi})_{A,\Delta}(z,D_z)\phi(.) + [m^{B;k}_{x,\xi}(z,D_z) - m^{A;k}_{x,\xi}(z,D_z)]\,(\phi^B_{x,\xi}(.) - \phi^A_{x,\xi}(.)).\end{aligned} \tag{5.71}$$

The symbols of the relevant, additional pseudodifferential operators are given by

$$\begin{aligned}(D\widetilde{m}_{x,\xi})_{A,\Delta}(z,\zeta) = \int_0^1 (1-s)\,\partial_s^2[&(L^{A;k}_{x,\xi}(D_x,D_\xi)(Dp)_{A;\Delta})(\tilde{x},\tilde{\xi})\\ &+\langle(\partial_\xi(Dp)_{A;\Delta})(x,\xi),(D_x p_{A;k})(\tilde{x},\tilde{\xi})\rangle - \langle(\partial_x(Dp)_{A;\Delta})(x,\xi),(D_\xi p_{A;k})(\tilde{x},\tilde{\xi})\rangle]\,\mathrm{d}s\end{aligned}$$

and

$$\begin{aligned}(R\widetilde{m}_{x,\xi})_{A,\Delta}(z,\zeta) = \int_0^1 (1-s)\,\partial_s^2[&(L^{A;k}_{x,\xi}(D_x,D_\xi)(Rp)_{A;\Delta})(\tilde{x},\tilde{\xi})\\ &+\langle(\partial_\xi(Rp)_{A;\Delta})(x,\xi),(D_x p_{A;k})(\tilde{x},\tilde{\xi})\rangle - \langle(\partial_x(Rp)_{A;\Delta})(x,\xi),(D_\xi p_{A;k})(\tilde{x},\tilde{\xi})\rangle\\ &+\langle(\partial_\xi(p_{B;k}-p_{A;k}))(x,\xi),(D_x(p_{B;k}-p_{A;k}))(\tilde{x},\tilde{\xi})\rangle\\ &-\langle(\partial_x(p_{B;k}-p_{A;k}))(x,\xi),(D_\xi(p_{B;k}-p_{A;k}))(\tilde{x},\tilde{\xi})\rangle]\,\mathrm{d}s.\end{aligned}$$

LEMMA 5.7. *The Schwartz seminorms of $(D\widetilde{\phi}_{x,\xi})_{A,\Delta}$ are $\lesssim (2^k)^{3/2}\|\Delta\|_{L^\infty}$, uniformly in x and ξ. Moreover, the Schwartz seminorms of $(R\widetilde{\phi}_{x,\xi})_{A,\Delta}$ are $\lesssim (2^k)^2\|\Delta\|^2_{L^\infty}$, uniformly in x and ξ.*

Proof. By methods used in the proof of Lemma 5.4 we find that

$$|(\partial_z^\alpha\partial_\zeta^\beta(D\widetilde{m}_{x,\xi})_{A,\Delta})(z,\zeta)| \lesssim (2^k)^{\frac{3}{2}-\frac{1}{2}|\beta|}\,\|B-A\|_{L^\infty}\,\langle z\rangle^2, \quad |\zeta|<2. \tag{5.72}$$

The Schwartz seminorms of $(D\widetilde{\phi}_{x,\xi})_{A,\Delta}$ are dominated by this estimate for $\alpha=\beta=0$, which follows upon integration by parts.

In a similar fashion, we obtain,

$$|(\partial_z^\alpha\partial_\zeta^\beta(R\widetilde{m}_{x,\xi})_{A,\Delta})(z,\zeta)| \lesssim (2^k)^{\frac{3}{2}-\frac{1}{2}|\beta|}\,\|B-A\|^2_{L^\infty}\,\langle z\rangle^2, \quad |\zeta|<2. \tag{5.73}$$

Thus, the contribution to the Schwartz seminorms of $(R\phi_{x,\xi})_{A,\Delta}$ from the third term on the right-hand side of (5.71) is dominated by $(2^k)^{3/2}\,\|B-A\|^2_{L^\infty}$. The contributions from the first two terms are dominated by $2^k\,\|B-A\|^2_{L^\infty}$. However, the contribution from the fourth term on the right-hand side of (5.71) is dominated by $(2^k)^2\,\|B-A\|^2_{L^\infty}$ [15], from which the second statement in the lemma is a consequence. □

LEMMA 5.8. *Let $(DM)_{A,\Delta}:\ H^\alpha(\mathbb{R}^n)\to H^{\alpha-1}(\mathbb{R}^n)$ be the operator defined by*

$$\begin{aligned}(DM)_{A,\Delta}f = \iint ((D\widetilde{U})_{A,\Delta}f)(x,\xi)\,(\phi^A_{x,\xi})_k(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ + \iint (U_{A;k}(T_kf))(x,\xi)\,(D\phi_{x,\xi})_{A,\Delta}(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi,\end{aligned} \tag{5.74}$$

cf. (5.33), and let $(DN)_{A,\Delta}:\ H^\alpha(\mathbb{R}^n)\to H^{\alpha-2}(\mathbb{R}^n)$ be the operator defined by

$$\begin{aligned}(DN)_{A,\Delta}f = -2(DP)_{A;\Delta}M_{A;k}f - 2P_{A;k}(DM)_{A,\Delta}f \\ + \iint ((D\widetilde{U})_{A,\Delta}f)(x,\xi)\,(\widetilde{\phi}^A_{x,\xi})_k(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ + \iint (U_{A;k}(T_kf))(x,\xi)\,(D\widetilde{\phi}_{x,\xi})_{A,\Delta}(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi.\end{aligned} \tag{5.75}$$

If $f\in L^2$ and $\hat{f}$ is supported in $|\xi|\approx 2^k$, we have

$$\|(DM)_{A,\Delta}f\|_{L^2} \lesssim 2^k\quad \|\Delta\|_{C^{0,1}}\|f\|_{L^2}, \tag{5.76}$$

$$\|(DN)_{A,\Delta}f\|_{L^2} \lesssim (2^k)^2\,\|\Delta\|_{C^{0,1}}\|f\|_{L^2}, \tag{5.77}$$

while

$$\|(RM)_{A,\Delta}f\|_{L^2} \lesssim (2^k)^2\,\|\Delta\|^2_{C^{0,1}}\|f\|_{L^2}, \tag{5.78}$$

$$\|(RN)_{A,\Delta}f\|_{L^2} \lesssim (2^k)^3\,\|\Delta\|^2_{C^{0,1}}\|f\|_{L^2}. \tag{5.79}$$

Proof. The estimate for $(DM)_{A,\Delta}$ follows directly from (5.34) and Lemma 5.4. For the remainder,

$$\begin{aligned}(RM)_{A,\Delta}f = \iint ((R\widetilde{U})_{A,\Delta}f)(x,\xi)\,(\phi^A_{x,\xi})_k(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ + \iint (U_{A;k}(T_kf))(x,\xi)\,(R\phi_{x,\xi})_{A,\Delta}(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi \\ + \iint ((U_{B;k}-U_{A;k})(T_kf))(x,\xi)\,(\phi^B_{x,\xi}-\phi^A_{x,\xi})_k(.;x,\xi)\,\mathrm{d}x\mathrm{d}\xi,\end{aligned} \tag{5.80}$$

we use (5.35) to estimate the first term on the right-hand side, Lemma 5.4, again, to estimate the second term, and [15, Lemma 6.2 and Lemma 6.4] to estimate the third term.

We write operators $N_{A;k}$ in the form

$$N_{A;k} = (\widetilde{T}_k^A)^* U_{A;k} T_k - 2P_{A;k} M_{A;k}. \tag{5.81}$$

The estimate for $(DN)_{A,\Delta}$ follows directly from (5.9), (5.76), (5.34) and Lemma 5.7. To analyze the remainder, we consider the two terms on the right-hand side of (5.81) separately. We get

$$(R(P\,M))_{A,\Delta} = (RP)_{A,\Delta} M_{A;k} + P_{A;k}(RM)_{A,\Delta} + (P_{B;k} - P_{A;k})(M_{B;k} - M_{A;k}),$$

and using (5.12), (5.78) and [15, Lemma 6.1 and Lemma 6.5], we find that

$$\begin{aligned} \|(R(P\,M))_{A,\Delta} f\|_{L^2} &\lesssim 2^k \|B-A\|_{L^\infty}^2 \|f\|_{L^2} \\ &+ (2^k)^3 \|B-A\|_{C^{0,1}}^2 \|f\|_{L^2} + (2^k)^2 \|B-A\|_{L^\infty} \|B-A\|_{C^{0,1}} \|f\|_{L^2} \\ &\lesssim (2^k)^3 \|B-A\|_{C^{0,1}}^2 \|f\|_{L^2}. \end{aligned} \tag{5.82}$$

Furthermore,

$$\begin{aligned} (R((\widetilde{T}_k)^* U T_k))_{A,\Delta} f = &\iint ((R\widetilde{U})_{A,\Delta} f)(x,\xi)\, (\widetilde{\phi}^A_{x,\xi})_k(.;x,\xi)\, \mathrm{d}x\mathrm{d}\xi \\ &+ \iint (U_{A;k}(T_k f))(x,\xi)\, (R\widetilde{\phi}_{x,\xi})_{A,\Delta}(.;x,\xi)\, \mathrm{d}x\mathrm{d}\xi \\ &+ \iint ((U_{B;k} - U_{A;k})(T_k f))(x,\xi)\, (\widetilde{\phi}^B_{x,\xi} - \widetilde{\phi}^A_{x,\xi})_k(.;x,\xi)\, \mathrm{d}x\mathrm{d}\xi, \end{aligned}$$

and using (5.35), Lemma 5.7, and [15, Lemma 6.4 and Lemma 6.5], we find that

$$\begin{aligned} \|(R((\widetilde{T}_k)^* U T_k))_{A,\Delta} f\|_{L^2} &\lesssim (2^k)^{5/2} \|B-A\|_{C^{0,1}}^2 \|f\|_{L^2} \\ &+ (2^k)^2 \|B-A\|_{L^\infty}^2 \|f\|_{L^2} + (2^k)^3 \|B-A\|_{L^\infty} \|B-A\|_{C^{0,1}} \|f\|_{L^2} \\ &\lesssim (2^k)^3 \|B-A\|_{C^{0,1}}^2 \|f\|_{L^2}. \end{aligned} \tag{5.83}$$

Adding (5.82) and (5.83) yields (5.79). ∎

For the Fréchet derivative of $\mathbf{T}_A(t) = \widetilde{\mathbf{T}}_A(t)(I+K_A)^{-1}$ (cf. (4.14)) we have

Lemma 5.9. *Let $(D\mathbf{T})_{A,\Delta}:\ H^{\alpha+1}(\mathbb{R}^n) \to H^\alpha(\mathbb{R}^n)$, for $-1 \le \alpha \le 1$, be defined by*

$$(D\mathbf{T})_{A,\Delta} g = (D\widetilde{\mathbf{T}})_{A,\Delta}(I+K_A)^{-1} g + \widetilde{\mathbf{T}}_A\, (D(I+K)^{-1})_{A,\Delta} g, \tag{5.84}$$

in which

$$\begin{aligned} (D\widetilde{\mathbf{T}})_{A,\Delta} g = &-t \sum_{k<k_0} \bar{\Delta}\, \beta_k(D) g \\ &+ \sum_{\pm} \sum_{k \ge k_0} \Big[\, (DN)_{A,\Delta}(\tfrac{1}{2}\mathrm{i}\, Q_{A;k}\beta_k(D)g) + N_{A;k}\,(\tfrac{1}{2}\mathrm{i}\,(DQ)_{A,\Delta}\beta_k(D)g)\,\Big] \\ &+ \sum_{\pm} \sum_{k \ge k_0} \Big[\, P_{A;k}(DP)_{A,\Delta} E_{A;k} g + (DP)_{A,\Delta} P_{A;k} E_{A;k} g\,\Big] \\ &+ \sum_{\pm} \sum_{k \ge k_0} ((P_{A;k})^2 - A_k)(DE)_{A,\Delta} g \\ &+ \sum_{\pm} \sum_{k \ge k_0} (A_k - A)(DE)_{A,\Delta} g + \sum_{\pm} \sum_{k \ge k_0} (\bar{\Delta}_k - \bar{\Delta})\, E_{A;k} g. \end{aligned} \tag{5.85}$$

We have

$$\|(D\mathbf{T})_{A,\Delta} g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{0,1}} \|g\|_{H^{\alpha+1}}. \tag{5.86}$$

For $-1 \leq \alpha \leq 0$, *it holds true that*

$$\|(R\mathbf{T})_{A,\Delta} g\|_{H^\alpha} \lesssim \|\Delta\|^2_{C^{0,1}} \|g\|_{H^{\alpha+2}}. \tag{5.87}$$

REMARK 5.10. *It holds also true that* $\|(D\mathbf{T})_{A,\Delta} g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{1,1}} \|g\|_{H^{\alpha+1}}$ *and* $\|(R\mathbf{T})_{A,\Delta} g\|_{H^\alpha} \lesssim \|\Delta\|^2_{C^{1,1}} \|g\|_{H^{\alpha+2}}$ *both for* $-1 \leq \alpha \leq 2$. *However, these estimates do not apply to next lemma.*

Proof. We decompose $\widetilde{\mathbf{T}}_A(t) = [D_t^2 - A(x, D_x)]\, \widetilde{S}_A(t)$ into four contributions:

$$\widetilde{\mathbf{T}}_{1,A}(t) g = -t \sum_{k<k_0} A\, \beta_k(D) g, \tag{5.88}$$

$$\widetilde{\mathbf{T}}_{2,A}(t) g = \sum_{\pm} \sum_{k \geq k_0} N_{A;k}\, (\tfrac{1}{2}\mathrm{i}\, Q_{A;k} \beta_k(D) g), \tag{5.89}$$

$$\widetilde{\mathbf{T}}_{3,A}(t) g = \sum_{\pm} \sum_{k \geq k_0} ((P_{A;k})^2 - A_k) E_{A;k}(t) g, \tag{5.90}$$

$$\widetilde{\mathbf{T}}_{4,A}(t) g = \sum_{\pm} \sum_{k \geq k_0} (A_k - A)\, E_{A;k}(t) g. \tag{5.91}$$

The Fréchet derivative of the first term follows immediately to be

$$(D\widetilde{\mathbf{T}}_1)_{A,\Delta}(t) g = \widetilde{\mathbf{T}}_{1,B}(t) g - \widetilde{\mathbf{T}}_{1,A}(t) g = -t \sum_{k<k_0} \bar{\Delta}\, \beta_k(D) g.$$

We have the estimate

$$\|(D\widetilde{\mathbf{T}}_1)_{A,\Delta} g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{0,1}} \|g\|_{H^\alpha}, \tag{5.92}$$

while, clearly, $(R\widetilde{\mathbf{T}}_1)_{A,\Delta}(t) \equiv 0$.

From the second term (cf. (5.89)) we deduce that

$$(D\widetilde{\mathbf{T}}_2)_{A,\Delta}(t) g = \sum_{\pm} \sum_{k \geq k_0} \Big[\, (DN)_{A,\Delta}(\tfrac{1}{2}\mathrm{i}\, Q_{A;k} \beta_k(D) g) + N_{A;k}\, (\tfrac{1}{2}\mathrm{i}\, (DQ)_{A,\Delta} \beta_k(D) g)\, \Big];$$

using (5.77) and (5.10), we obtain

$$\begin{aligned}
&\|(DN)_{A,\Delta}(\tfrac{1}{2}\mathrm{i}\, Q_{A;k} \beta_k(D) g) + N_{A;k}\, (\tfrac{1}{2}\mathrm{i}\, (DQ)_{A,\Delta} \beta_k(D) g)\|_{L^2} \\
&\qquad \lesssim (2^k)^2 \|B - A\|_{C^{0,1}} 2^{-k} \|\beta_k(D) g\|_{L^2} + 2^k 2^{-k} \|B - A\|_{L^\infty} \|\beta_k(D) g\|_{L^2} \\
&\qquad\qquad \lesssim 2^k\, \|B - A\|_{C^{0,1}} \|\beta_k(D) g\|_{L^2},
\end{aligned}$$

hence, upon summation over scales,

$$\|(D\widetilde{\mathbf{T}}_2)_{A,\Delta}(t) g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{0,1}} \|g\|_{H^{\alpha+1}}. \tag{5.93}$$

Furthermore,

$$\begin{aligned}
(R\widetilde{\mathbf{T}}_2)_{A,\Delta}(t) g = \sum_{\pm} \sum_{k \geq k_0} \Big[\, &(RN)_{A,\Delta}(\tfrac{1}{2}\mathrm{i}\, Q_{A;k} \beta_k(D) g) \\
&+ N_{A;k}\, (\tfrac{1}{2}\mathrm{i}\, (RQ)_{A,\Delta} \beta_k(D) g) + (N_{B;k} - N_{A;k})\, (\tfrac{1}{2}\mathrm{i}\, (Q_{B;k} - Q_{A;k}) \beta_k(D) g)\, \Big];
\end{aligned}$$

using (5.79), (5.13), and [15, Lemma 6.1 and Lemma 6.5.] we obtain that

$$\begin{aligned}
&\|(RN)_{A,\Delta}(\tfrac{1}{2}\mathrm{i}\,Q_{A;k}\beta_k(D)g) + N_{A;k}\,(\tfrac{1}{2}\mathrm{i}\,(RQ)_{A,\Delta}\beta_k(D)g) \\
&\qquad + (N_{B;k} - N_{A;k})\,(\tfrac{1}{2}\mathrm{i}\,(Q_{B;k} - Q_{A;k})\beta_k(D)g)\|_{L^2} \\
&\lesssim (2^k)^3\|B - A\|^2_{C^{0,1}}2^{-k}\|\beta_k(D)g\|_{L^2} + 2^k2^{-k}\|B - A\|^2_{L^\infty}\|\beta_k(D)g\|_{L^2} \\
&\qquad + (2^k)^2\|B - A\|_{C^{0,1}}2^{-k}\|B - A\|_{L^\infty}\|\beta_k(D)g\|_{L^2} \\
&\qquad\qquad \lesssim (2^k)^2\|B - A\|^2_{C^{0,1}}\|\beta_k(D)g\|_{L^2},
\end{aligned}$$

whence, upon summation,

$$\|(R\widetilde{\mathbf{T}}_2)_{A,\Delta}(t)g\|_{H^\alpha} \lesssim \|\Delta\|^2_{C^{0,1}}\|g\|_{H^{\alpha+2}}. \tag{5.94}$$

From the third term (cf. (5.90)) we need to account for

$$\begin{aligned}
(D\widetilde{\mathbf{T}}_3)_{A,\Delta}(t)g = \sum_{\pm}\sum_{k\geq k_0}\Big[\, & P_{A;k}(DP)_{A,\Delta}E_{A;k}(t)g + (DP)_{A,\Delta}P_{A;k}E_{A;k}(t)g \\
& + (B_k - A_k)E_{A;k}(t)g + (P^2_{A;k} - A_k)(DE)_{A,\Delta}g\,\Big];
\end{aligned} \tag{5.95}$$

for the first two terms on the righthand side of the (5.95) we use the estimate

$$\|(DP^2)_{A,\Delta}E_{A;k}(t)g\|_{L^2} \lesssim 2^k\|\Delta\|_{C^{0,1}}\|\beta_k(D)g\|_{L^2}$$

and for the term $(P^2_{A;k} - A_k)(DE)_{A,\Delta}g$ we use the estimate [15, Lemma 5.5]

$$\|(P^2_{A;k} - A_k)(DE)_{A,\Delta}g\|_{L^2} \lesssim 2^k\|\Delta\|_{C^{0,1}}\|\beta_k(D)g\|_{L^2}.$$

The third term on the righthand side of the (5.95) is pseudodifferential operator of type $S^0_{1,\frac{1}{2}}$, [18] which will be cancelled with part of $(D\tilde{\mathbf{T}}_4)_{A,\Delta}(t)$.; see Appendix, discussion after formula (A.5).

The remainder is

$$\begin{aligned}
(R\widetilde{\mathbf{T}}_3)_{A,\Delta}(t)g = \sum_{\pm}\sum_{k\geq k_0}\Big[\, & P_{A;k}(RP)_{A,\Delta}E_{A;k}(t)g + (RP)_{A,\Delta}P_{A;k}E_{A;k}(t)g \\
& + (P_{B;k} - P_{A;k})^2E_{A;k}(t)g + (P^2_{A;k} - A_k)(RE)_{A,\Delta}g \\
& + ((P^2_{B;k} - B_k - (P^2_{A;k} - A_k))(E_{B;k} - E_{A;k})(t)g\,\Big];
\end{aligned}$$

the first three terms are covered by the estimate

$$\|(RP^2)_{A,\Delta}E_{A;k}(t)g\|_{L^2} \lesssim 2^k\|\Delta\|^2_{C^{0,1}}\|\beta_k(D)g\|_{L^2},$$

while for the fourth term, we use

$$\|(P^2_{A;k} - A_k)(RE)_{A,\Delta}g\|_{L^2} \lesssim (2^k)^2\|\Delta\|^2_{C^{0,1}}\|\beta_k(D)g\|_{L^2}.$$

The last term is written as

$$\begin{aligned}
&\sum_{\pm}\sum_{k\geq k_0}((P^2_{B;k} - B_k - (P^2_{A;k} - A_k))(E_{B;k} - E_{A;k})(t)g \\
&\quad = \sum_{\pm}\sum_{k\geq k_0}(P^2_{B;k} - P^2_{A;k})(E_{B;k} - E_{A;k})(t)g + \sum_{\pm}\sum_{k\geq k_0}(A_k - B_k)(E_{B;k} - E_{A;k})(t)g.
\end{aligned} \tag{5.96}$$

Because $P^2_{B;k} - P^2_{A;k} = \frac{1}{2}[(P_{B;k} - P_{A;k})(P_{B;k} + P_{A;k}) + (P_{B;k} + P_{A;k})(P_{B;k} - P_{A;k})]$, we have

$$\|((P^2_{B;k} - P^2_{A;k})(E_{B;k} - E_{A;k})g\|_{L^2} \lesssim (2^k)^2 \|\Delta\|^2_{C^{0,1}} \|\beta_k(D)g\|_{L^2}.$$

The term $\sum_k (A_k - B_k)(E_{B;k} - E_{A;k})(t)g$ will be cancelled with part of $(R\tilde{\mathbf{T}}_4)_{A,\Delta}(t)$.; see Appendix, discussion after formula (A.5). Hence, upon summation,

$$\|(D\widetilde{\mathbf{T}}_3)_{A,\Delta}(t)g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{0,1}} \|g\|_{H^{\alpha+1}}, \tag{5.97}$$

$$\|(R\widetilde{\mathbf{T}}_3)_{A,\Delta}(t)g\|_{H^\alpha} \lesssim \|\Delta\|^2_{C^{0,1}} \|g\|_{H^{\alpha+2}}. \tag{5.98}$$

The Fréchet derivative of the fourth term, $\widetilde{\mathbf{T}}_{4,A}(t)$, follows to be

$$(D\widetilde{\mathbf{T}}_4)_{A,\Delta}(t)g = \sum_{\pm} \sum_{k \geq k_0} [(A_k - A)(DE)_{A,\Delta}(t)g + (\bar{\Delta}_k - \bar{\Delta})E_{A;k}(t)g];$$

the associated remainder is given by

$$(R\widetilde{\mathbf{T}}_4)_{A,\Delta}(t)g = \sum_{\pm} \sum_{k \geq k_0} [(A_k - A)(RE)_{A;\Delta}(t)g + (\bar{\Delta}_k - \bar{\Delta})(E_{B;k}(t) - E_{A;k}(t))g].$$

These operators sense the 'rough' parts of the wavespeed model. We use Lemma A.1, and Lemma 5.3 to obtain

$$\|(D\widetilde{\mathbf{T}}_4)_{A,\Delta}(t)g\|_{H^\alpha} \lesssim \|\Delta\|_{C^{0,1}} \|g\|_{H^{\alpha+1}}, \quad -1 \leq \alpha \leq 1, \tag{5.99}$$

$$\|(R\widetilde{\mathbf{T}}_4)_{A,\Delta}(t)g\|_{H^\alpha} \lesssim \|\Delta\|^2_{C^{0,1}} \|g\|_{H^{\alpha+2}}, \quad -1 \leq \alpha \leq 0. \tag{5.100}$$

Combining these with (5.92), (5.93), (5.97) and (5.62) yields (5.86), and with (5.94), (5.98) and (5.63) yields (5.87). ∎

We then consider the Fréchet derivative of the Volterra solution operator. We have

LEMMA 5.11. *Let* $(DV)_{A,\Delta} : L^\infty_t H^{\alpha+1}_x \to L^\infty_t H^\alpha_x$ *be given by*

$$(DV)_{A,\Delta}F(t,x) = \sum_{j=1}^{\infty} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{j-1}} \sum_{l=1}^{j} \mathbf{T}_A(t,s_1) \ldots \mathbf{T}_A(s_{l-2},s_{l-1}) (D\mathbf{T})_{A,\Delta}(s_{l-1},s_l)\mathbf{T}_A(s_l,s_{l+1}) \cdots \mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j \ldots \mathrm{d}s_1. \tag{5.101}$$

For $-1 \leq \alpha \leq 1$, *we have*

$$\|(DV)_{A,\Delta}F\|_{L^\infty_t H^\alpha_x} \lesssim \|\Delta\|_{C^{0,1}} \|F\|_{L^\infty_t H^{\alpha+1}_x}; \tag{5.102}$$

for $-1 \leq \alpha \leq 0$ *it holds true that*

$$\|(RV)_{A,\Delta}F\|_{L^\infty_t H^\alpha_x} \lesssim \|\Delta\|^2_{C^{0,1}} \|F\|_{L^\infty_t H^{\alpha+2}_x}. \tag{5.103}$$

Proof. The formula for the Volterra solution operator V_A is

$$V_A F(t,x) = F(t,x) + \sum_{j=1}^{\infty} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{j-1}} \mathbf{T}_A(t,s_1)\mathbf{T}_A(s_1,s_2) \cdots \mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j \cdots \mathrm{d}s_1.$$

Then

$$
\begin{aligned}
V_B F(t,x) - V_A F(t,x) = \sum_{j=1}^{\infty}\sum_{l=1}^{j}\int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_B(t,s_1)\cdots\mathbf{T}_B(s_{l-2},s_{l-1})& \\
(\mathbf{T}_B - \mathbf{T}_A)(s_{l-1},s_l)\mathbf{T}_A(s_l,s_{l+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1& \\
= \sum_{j=1}^{\infty}\sum_{l=1}^{j}\int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_A(t,s_1)\cdots\mathbf{T}_A(s_{l-2},s_{l-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{l-1},s_l)& \\
\mathbf{T}_A(s_l,s_{l+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1& \\
+ \sum_{j=1}^{\infty}\sum_{l=1}^{j}\sum_{f=l+1}^{j}\int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_B(t,s_1)\cdots\mathbf{T}_B(s_{l-2},s_{l-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{l-1},s_l)& \\
\mathbf{T}_B(s_l,s_{l+1})\cdots\mathbf{T}_B(s_{f-2},s_{f-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{f-1},s_f)& \\
\mathbf{T}_A(s_f,s_{f+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1.&
\end{aligned}
$$

If $(DV)_{A,\Delta}F(s,x)$ is given by (5.101), the remainder is

$$
\begin{aligned}
(RV)_{A,\Delta}F(t,x) = \sum_{j=1}^{\infty}\int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}}\sum_{l=1}^{j} \mathbf{T}_A(t,s_1)\ \ldots\ \mathbf{T}_A(s_{l-2},s_{l-1})& \\
(R\mathbf{T})_{A,\Delta}(s_{l-1},s_l)\mathbf{T}_A(s_l,s_{l+1})\ \cdots\ \mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\ \ldots\ \mathrm{d}s_1& \\
+ \sum_{j=1}^{\infty}\sum_{l=1}^{j}\sum_{f=l+1}^{j}\int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_B(t,s_1)\cdots\mathbf{T}_B(s_{l-2},s_{l-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{l-1},s_l)& \\
\mathbf{T}_B(s_l,s_{l+1})\cdots\mathbf{T}_B(s_{f-2},s_{f-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{f-1},s_f)& \\
\mathbf{T}_A(s_f,s_{f+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1.&
\end{aligned}
$$

We need to estimate the $L_t^\infty H_x^\alpha$ norms of three types of integral expressions,

$$
\begin{aligned}
\tilde{I}(t,x) = \int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_A(t,s_1)\cdots\mathbf{T}_A(s_{l-2},s_{l-1})(D\mathbf{T})_{A,\Delta}(s_{l-1},s_l)& \\
\mathbf{T}_A(s_l,s_{l+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1,&
\end{aligned}
$$

$$
\begin{aligned}
\check{I}(t,x) = \int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_A(t,s_1)\cdots\mathbf{T}_A(s_{l-2},s_{l-1})(R\mathbf{T})_{A,\Delta}(s_{l-1},s_l)& \\
\mathbf{T}_A(s_l,s_{l+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1,&
\end{aligned}
$$

$$
\begin{aligned}
\hat{I}(t,x) = \int_0^t\int_0^{s_1}\cdots\int_0^{s_{j-1}} \mathbf{T}_B(t,s_1)\cdots\mathbf{T}_B(s_{l-2},s_{l-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{l-1},s_l)& \\
\mathbf{T}_B(s_l,s_{l+1})\cdots\mathbf{T}_B(s_{f-2},s_{f-1})(\mathbf{T}_B - \mathbf{T}_A)(s_{f-1},s_f)& \\
(\mathbf{T}_A)(s_f,s_{f+1})\cdots\mathbf{T}_A(s_{j-1},s_j)F(s_j,x)\,\mathrm{d}s_j\cdots\mathrm{d}s_1.&
\end{aligned}
$$

Choose $C = C(M)$ such that for $t, s \in [-M, M]$,

$$
\begin{aligned}
\|T_A(t,s)g\|_{H^\alpha} + \|T_B(t,s)g\|_{H^\alpha} &\le C\|g\|_{H^\alpha}, \quad -1 \le \alpha \le 2, \\
\|(T_B - T_A)(t,s)g\|_{H^\alpha} &\le C\|\Delta\|_{C^{0,1}}\|g\|_{H^{\alpha+1}}, \quad -1 \le \alpha \le 1. \\
\|(D\mathbf{T})_{A,\Delta}(t,s)g\|_{H^\alpha} &\le C\|\Delta\|_{C^{0,1}}\|g\|_{H^{\alpha+1}}, \quad -1 \le \alpha \le 1. \\
\|(R\mathbf{T})_{A,\Delta}(t,s)g\|_{H^\alpha} &\le C\|\Delta\|_{C^{0,1}}^2\|g\|_{H^{\alpha+2}}, \quad -1 \le \alpha \le 0.
\end{aligned}
$$

Then, for $-1 \le \alpha \le 1$,

$$\|\tilde{I}(t,\cdot)\|_{H^\alpha} \le \frac{C^j t^j}{j!} \|\Delta\|_{C^{0,1}} \|F\|_{L_t^\infty H_x^{\alpha+1}}.$$

In (5.101) there are j terms of the form $\tilde{I}(t,x)$ at level j. It follows that

$$\|(DV)_{A,\Delta} F\|_{L_t^\infty H_x^\alpha} \le \Big(\sum_{j=1}^\infty \frac{j(CM)^j}{j!}\Big) \|\Delta\|_{C^{0,1}} \|F\|_{L_t^\infty H_x^{\alpha+1}}.$$

For $(RV)_{A,\Delta}$, let $-1 \le \alpha \le 0$, and note that

$$\|\check{I}(t,\cdot)\|_{H^\alpha} \le \frac{C^j t^j}{j!} \|\Delta\|_{C^{0,1}}^2 \|F\|_{L_t^\infty H_x^{\alpha+2}},$$
$$\|\hat{I}(t,\cdot)\|_{H^\alpha} \le \frac{C^j t^j}{j!} \|\Delta\|_{C^{0,1}}^2 \|F\|_{L_t^\infty H_x^{\alpha+2}}.$$

There are j terms of the form $\check{I}(t,x)$ and $\frac{j(j-1)}{2}$ terms of the form $\check{I}(t,x)$ at level j. It follows that

$$\|(RV)_{A,\Delta} F\|_{L_t^\infty H_x^{\alpha-1}} \le \Big(\sum_{j=1}^\infty \frac{[j+\frac{j(j-1)}{2}](CM)^j}{j!}\Big) \|\Delta\|_{C^{0,1}}^2 \|F\|_{L_t^\infty H_x^{\alpha+2}}.$$

This proves the lemma. ∎

Finally, let $S_A(t)$ and $S_B(t)$ be given by

$$S_{A,B}(t)g = \widehat{S}_{A,B}(t)g + \int_0^t \widehat{S}_{A,B}(t,s) V_{A,B}(\mathbf{T}_{A,B}g)(s,.)\,\mathrm{d}s,$$

cf. (4.15). Then,

$$\begin{aligned}
(S_B(t) - S_A(t))g = (\widehat{S}_B - \widehat{S}_A)(t)g &+ \int_0^t (\widehat{S}_B - \widehat{S}_A)(t,s) V_B(\mathbf{T}_B g)(s,.)\,\mathrm{d}s \\
&+ \int_0^t \widehat{S}_A(t,s)(V_B(\mathbf{T}_B g) - V_A(\mathbf{T}_A g))(s,.)\,\mathrm{d}s \\
= (\widehat{S}_B - \widehat{S}_A)(t)g &+ \int_0^t (\widehat{S}_B - \widehat{S}_A)(t,s) V_A(\mathbf{T}_A g)(s,.)\,\mathrm{d}s \\
&- \int_0^t (\widehat{S}_B - \widehat{S}_A)(t,s)(V_B(\mathbf{T}_B g) - V_A(\mathbf{T}_A g))(s,.)\,\mathrm{d}s \\
&+ \int_0^t \widehat{S}_A(t,s)(V_B(\mathbf{T}_B g) - V_A(\mathbf{T}_A g))(s,.)\,\mathrm{d}s,
\end{aligned}$$

and we note that

$$\begin{aligned}
V_B(\mathbf{T}_B g)(s,.) - V_A(\mathbf{T}_A g)(s,.) &= (V_B - V_A)(\mathbf{T}_A g)(s,.) \\
&\quad + V_A((\mathbf{T}_B - \mathbf{T}_A)g)(s,.) - (V_B - V_A)((\mathbf{T}_B - \mathbf{T}_A)g)(s,.).
\end{aligned}$$

We combine the estimates in the lemmata above, and obtain the statement of Theorem 2.2, with Fréchet derivative DS of S given by

$$\begin{aligned}
(DS)_{A,\Delta}(t)g &= (D\widehat{S})_{A,\Delta}(t)g + \int_0^t (D\widehat{S})_{A,\Delta}(t,s) V_A(\mathbf{T}_A g)(s,.)\,\mathrm{d}s \\
&\quad + \int_0^t \widehat{S}_A(t,s)(DV)_{A,\Delta}(\mathbf{T}_A g)(s,.)\,\mathrm{d}s + \int_0^t \widehat{S}_A(t,s) V_A((D\mathbf{T})_{A,\Delta} g)(s,.)\,\mathrm{d}s,
\end{aligned} \tag{5.104}$$

and remainder

$$(RS)_{A,\varDelta}(t)g = (R\widehat{S})_{A,\varDelta}(t)g + \int_0^t (R\widehat{S})_{A,\varDelta}(t,s)V_A(\mathbf{T}_A g)(s,.)\,\mathrm{d}s \\ + \int_0^t \widehat{S}_A(t,s)(RV)_{A,\varDelta}(\mathbf{T}_A g)(s,.)\,\mathrm{d}s + \int_0^t \widehat{S}_A(t,s)V_A((R\mathbf{T})_{A,\varDelta}g)(s,.)\,\mathrm{d}s \\ - \int_0^t (\widehat{S}_B - \widehat{S}_A)(t,s)(V_B(\mathbf{T}_B g) - V_A(\mathbf{T}_A)g))(s,.)\,\mathrm{d}s \\ - \int_0^t (\widehat{S}_A)(t,s)(V_B - V_A)(\mathbf{T}_B - \mathbf{T}_A))(s,.)\,\mathrm{d}s. \tag{5.105}$$

6. Discussion. We address how the results presented here fit in with the literature on, and current practice of wave-equation tomograhy. Typically, one uses the Born approximation, $S^b_{A,\varDelta}(t)g$ say, in place of $(DS)_{A,\varDelta}(t)$ as the starting point in deriving the sensitivity in accordance with (2.8) [16]. In the framework of the analysis in this paper, the Born approximation is obtained as follows.

We consider the Cauchy initial value problem,

$$\begin{cases} [D_t^2 - A(x,D_x)]u(t,x) &= F(t,x), \\ u|_{t=0} &= 0, \\ \partial_t u|_{t=0} &= g \end{cases} \tag{6.1}$$

cf. (1.1). Here, $g \in H^\alpha$, $F \in L^1_t H^\alpha_x$, and $u(t,x)$ is a function in $\mathbb{R}_t \times \mathbb{R}^n_x$; $t \in [-M,M]$ as before. The unique solution can be written in the form

$$u(t,.) = S_A(t)g + \int_0^t S_A(t-t')\,F(t',.)\,\mathrm{d}t'. \tag{6.2}$$

We then consider a second Cauchy initial value problem, with the same right-hand side and initial condition, but with A replaced by $A+\varDelta$, that is, B. As in Theorem 2.2, we now assume that $g \in H^{\alpha+2}$ (accounting for the "loss of two derivatives") and $-1 \le \alpha \le 0$. The difference, $v = u_{A+\varDelta} - u_A$, satisfies the Cauchy initial value problem,

$$\begin{cases} [D_t^2 - A(x,D_x)]v(t,x) &= \bar{\varDelta}(x,D_x)u_{A+\varDelta}(t,.), \\ v|_{t=0} &= 0, \\ \partial_t v|_{t=0} &= 0. \end{cases} \tag{6.3}$$

By our choice of α, the contrast source is in $L^1_t L^2_x$. Hence, we have

$$u_{A+\varDelta}(t,.) - u_A(t,.) = \int_0^t S_A(t-t')\,\bar{\varDelta}(x,D_x)u_{A+\varDelta}(t',.)\,\mathrm{d}t'. \tag{6.4}$$

The Born approximation is standardly obtained by replacing $u_{A+\varDelta}(t',.)$ in the integrand by $u_A(t',.)$. This defines a map $S^b_{A,\varDelta}(t)$, acting linearly in $\varDelta$, given by

$$S^b_{A,\varDelta}(t)g := \int_0^t S_A(t-t')\,\bar{\varDelta}(x,D_x)\,S_A(t')g\,\mathrm{d}t'. \tag{6.5}$$

It follows that the remainder,

$$R^b_{A,\varDelta}(t)g = u_{A+\varDelta}(t,.) - u_A(t,.) - S^b_{A,\varDelta}(t)g = \int_0^t S_A(t-t')\,\bar{\varDelta}(x,D_x)\,v(t',.)\,\mathrm{d}t' \tag{6.6}$$

satisfies the estimate

$$\begin{aligned}(6.7)\quad \|R^b_{A,\varDelta}(t)g\|_{H^{\alpha+1}} &\lesssim \int_0^t \|\bar{\varDelta}(x,D_x)\,v(t',.)\|_{H^\alpha}\,\mathrm{d}t' \\ &\lesssim \|\Delta\|_{C^{0,1}} \sup_{t\in[-M,M]} \|v(t,.)\|_{H^{\alpha+2}} \lesssim \|\Delta\|^2_{C^{0,1}}\|g\|_{H^{\alpha+2}}.\end{aligned}$$

Moreover, $\|S^b_{A,\varDelta}(t)g\|_{H^{\alpha+1}} \lesssim \|\Delta\|_{C^{0,1}}\|g\|_{H^{\alpha+1}}$. These estimates confirm that the Born approximation coincides with the Fréchet derivative, $(DS)_{A,\varDelta}(t)$, as it should.

Finally, the symbol smoothing in accordance with (3.7) leads to a representation of the contrast in terms of $\varDelta_k$. In practice, it is advantageous to carry out tomographic inversion using an ℓ^1-norm regularization strategy [11] based on the possibility of obtaining sparse representations of $\varDelta$; such representations have been obtained in earth sciences applications with dual-tree complex wavelets (for $n = 2$). Indeed, the 'low-pass filtering' (with respect to cubes rather than spheres in frequency) leading to $\varDelta_k$ can be obtained through the dual-tree complex wavelet transform [9, 10]. The mentioned strategy allows fine-scale features and rapid variations across boundaries to be honored in the inversion locally in regions of 'proper' illumination, while emphasizing coarse-scale features locally in regions of 'poor' illumination.

Appendix A. Paradifferential estimates.

In this appendix, we prove paradifferential type estimates which are suitable for handling the parts of $\mathbf{T}_A(t)$ which depend on the nonsmooth metric. We follow ideas in [18].

Let a be a function in $C^{m-1,1}(\mathbb{R}^n)$ where $m \geq 2$ is an integer. Define frequency truncated versions

$$a_k = \chi(2^{-k/2}D)a,$$

where χ is a smooth function in $\mathbb{R}^n$, supported in the unit ball and satisfying $\chi(\xi) = 1$ for $|\xi| \leq \frac{1}{2}$. Considering $a_k = \chi(2^{-k/2}D)a$ as convolution, expanding in Taylor series, and using the moment conditions $\int y^\alpha \hat{\chi}(y)\,dy = 0$ for $\alpha \neq 0$, we obtain the estimate

$$\|a_k - a\|_{L^\infty} \lesssim 2^{-\frac{mk}{2}}\|a\|_{C^{m-1,1}}. \tag{A.1}$$

As in the main text, $A \lesssim B$ means $A \leq CB$ where C is a constant only depending on M, m, and n.

Futher, let $\{F_k\}_{k=1}^\infty$ be a family of operators on $L^2(\mathbb{R}^n)$ satisfying estimates

$$\|F_k g\|_{L^2} \lesssim R2^{kr}\|\beta_k(D)g\|_{L^2}, \tag{A.2}$$

where $R > 0$ and $r \in \mathbb{R}$. We also assume that F_k are frequency localized, in the sense that $F_k g = \tilde{\beta}_k(D)F_k g$ for cutoffs $\tilde{\beta}_k(\xi)$ supported in $|\xi| \approx 2^k$. More precisely, we assume that $\tilde{\beta}_k(\xi) = \tilde{\beta}(2^{-k}\xi)$ where $\tilde{\beta}$ is supported in $\{2^{-l_0} \leq |\xi| \leq 2^{l_0}\}$, and $l_0 = l_0(M)$ is an integer.

Lemma A.1. *Define the operator*

$$\Gamma g = \sum_{k=1}^\infty (a - a_k)F_k g.$$

If $0 \leq s \leq \frac{m}{2}$, *then*

$$\|\Gamma g\|_{H^{\alpha+s}} \lesssim R\|a\|_{C^{m-1,1}}\|g\|_{H^{\alpha+r}}, \qquad -m \leq \alpha < m - s. \tag{A.3}$$

Furthermore, if $s = 1$, *then*

$$\|\Gamma g\|_{H^{\alpha+1}} \lesssim R\|a\|_{C^{m-1,1}}\|g\|_{H^{\alpha+r}}, \qquad -m \leq \alpha \leq m - 1. \tag{A.4}$$

Proof. We first assume $-m \le \alpha < m - s$, and proceed to prove (A.3). Define

$$\Gamma_{jk} = \beta_j(D)(a - a_k)\tilde{\beta}_k(D).$$

By looking at supports on the Fourier side, we get

$$\Gamma_{jk} = \begin{cases} \beta_j(D)(a - a_{2k-4l_0})\tilde{\beta}_k(D), & j \le k - 4l_0, \\ \beta_j(D)(a - a_k)\tilde{\beta}_k(D), & k - 4l_0 \le j \le k + 4l_0, \\ \beta_j(D)(a - a_{2j-4l_0})\tilde{\beta}_k(D), & j \ge k + 4l_0. \end{cases}$$

By (A.1), these satisfy

$$\|\Gamma_{jk}\|_{L^2 \to L^2} \lesssim \begin{cases} \|a\|_{C^{m-1,1}} 2^{-mk}, & j \le k - 4l_0, \\ \|a\|_{C^{m-1,1}} 2^{-\frac{mk}{2}}, & k - 4l_0 \le j \le k + 4l_0, \\ \|a\|_{C^{m-1,1}} 2^{-mj}, & j \ge k + 4l_0. \end{cases}$$

By considering the sum over even and odd j separately, and using a similar argument for the sum over k, we obtain from (A.2) that

$$\begin{aligned} \|\Gamma g\|_{H^{\alpha+s}}^2 = \|\sum_{j,k} \Gamma_{jk} F_k g\|_{H^{\alpha+s}}^2 &\lesssim \sum_j \|\sum_k \Gamma_{jk} F_k g\|_{H^{\alpha+s}}^2 \lesssim \sum_j \sum_k \|\Gamma_{jk} F_k g\|_{H^{\alpha+s}}^2 \\ \lesssim R^2 \|a\|_{C^{m-1,1}}^2 \Big(\sum_j \sum_{k \le j} & 2^{2j(\alpha+s)-2mj-2k\alpha} A_k + \sum_j \sum_{j-4l_0 \le k \le j+4l_0} 2^{2j(\alpha+s)-mj-2k\alpha} A_k \\ &+ \sum_j \sum_{k \ge j} 2^{2j(\alpha+s)-2mk-2k\alpha} A_k \Big). \end{aligned}$$

We have written $A_k = 2^{2k(\alpha+r)} \|\tilde{\beta}_k(D) g\|_{L^2}^2$. Using the assumption that $\alpha < m - s$, the first sum in the parentheses is bounded by

$$\sum_k 2^{-2k\alpha} A_k \sum_{j \ge k} 2^{2j(\alpha+s-m)} \lesssim \sum_k 2^{-2k(m-s)} A_k \lesssim \|g\|_{H^{\alpha+r}}^2.$$

Since $s \le m/2$, the second sum in parentheses is bounded by

$$\sum_j 2^{-2j(m/2-s)} A_j \lesssim \|g\|_{H^{\alpha+r}}^2.$$

The third sum is bounded by

$$\sum_k 2^{-2k(m+\alpha)} A_k \sum_{j \le k} 2^{2j(\alpha+s)}.$$

By assumption, $m + \alpha \ge 0$. If $\alpha + s < 0$, the sum over j is $\lesssim 1$, and if $\alpha + s = 0$ then $m + \alpha > 0$ and $k 2^{-2k(m+\alpha)} \lesssim 1$. In both cases one gets a bound by $\|g\|_{H^{\alpha+r}}^2$. Furthermore, if $\alpha + s > 0$, then the third sum is bounded by

$$\sum_k 2^{2k(\alpha+s)-2k(m+\alpha)} A_k = \sum_k 2^{-2k(m-s)} A_k \lesssim \|g\|_{H^{\alpha+r}}^2.$$

This proves (A.3).

If $s = 1$, then (A.3) shows that

$$\|\Gamma g\|_{H^{\alpha+1}} \lesssim R \|a\|_{C^{m-1,1}} \|g\|_{H^{\alpha+r}}, \qquad -m \le \alpha < m - 1.$$

For (A.4), we need to show that the estimate holds true also for $\alpha = m-1$. Note that

$$\|\Gamma g\|_{H^m} \lesssim \|\Gamma g\|_{H^{m-1}} + \|\nabla \Gamma g\|_{H^{m-1}}.$$

The first term is $\lesssim R\|a\|_{C^{m-1,1}}\|g\|_{H^{m-2+r}} \lesssim R\|a\|_{C^{m-1,1}}\|g\|_{H^{m-1+r}}$. For the other term we compute

$$\nabla \Gamma g = (\nabla a)\sum F_k g - \sum (\nabla a_k) F_k g + \sum (a - a_k)\nabla F_k g. \tag{A.5}$$

We estimate the last three sums separately. Since ∇a is a multiplier in H^{m-1}, the bound (A.2) for frequency localized operators shows that the first sum has the bound

$$\|(\nabla a)\sum F_k g\|_{H^{m-1}} \lesssim \|a\|_{C^{m-1,1}}\|\sum_k F_k g\|_{H^{m-1}} \lesssim R\|a\|_{C^{m-1,1}}\|g\|_{H^{m-1+r}}.$$

For the second sum, note that multiplication by ∇a_k preserves localization to frequency $\sim 2^k$, since the Fourier transform of ∇a_k is supported in $\{|\xi| \leq 2^{k/2}\}$. The family of operators $\{(\nabla a_k)F_k\}$ is then frequency localized (the amount of localization being controlled by M) and satisfies $\|(\nabla a_k)F_k g\|_{L^2} \lesssim R\|a\|_{C^{m-1,1}} 2^{kr}\|\beta_k(D)g\|_{L^2}$. The second sum satisfies the estimate

$$\|\sum_k (\nabla a_k)F_k g\|_{H^{m-1}}^2 \lesssim \sum_k \|(\nabla a_k)F_k g\|_{H^{m-1}}^2 \lesssim R^2\|a\|_{C^{m-1,1}}^2\|g\|_{H^{m-1+r}}^2.$$

Finally, applying (A.3) to the family of operators $\{\nabla F_k\}$ which satisfy $\|\nabla F_k g\|_{L^2} \lesssim R 2^{k(r+1)}\|\beta_k(D)g\|_{L^2}$, the third sum has the bound,

$$\|\sum (a-a_k)\nabla F_k g\|_{H^{m-1}} \lesssim R\|a\|_{C^{m-1,1}}\|g\|_{H^{m-1+r}}.$$

This proves (A.4) also for $\alpha = m-1$. $\square$

Remark A.2. *For the corresponding estimates with a $C^{0,1}$ metric, see [15, Lemma 6.7].*

REFERENCES

[1] H. Bijwaard, W. Spakman, and E.R. Engdahl, *Closing the gap between regional and global travel time tomography*, J.geophys. Res., 103 (1998), pp. 30055–30078.

[2] E.J. Candès and D. Donoho, *New tight frames of curvelets abd optimal representations of objects with piecewise-C^2 singularities*, Comm. Pure Appl. Math., 57 (2004), pp. 219–266.

[3] ———, *Continuous curvelet transform: I. Resolution of the wavefront set*, Applied and Computational Harmonic Analysis, 19 (2005), pp. 162–197.

[4] ———, *Continuous curvelet transform: II. Discretization and frames*, Applied and Computational Harmonic Analysis, 19 (2005), pp. 198–222.

[5] L.C. Evans, *Partial Differential Equations*, vol. 19, Amer. Math. Soc., 2002.

[6] A. Greenleaf and G. Uhlmann, *Estimates for singular radon transforms and pseudodifferential operators with singular symbols*, Journal of functional analysis, 89 (1990), pp. 202–232.

[7] S.H. Hung, F.A. Dahlen, and G. Nolet, *Frchet kernels for finite-frequency traveltimes-ii.*, Geoph. J. Int., 141 (2000), pp. 175–203.

[8] B.L.N. Kennett, S. Widiyantoro, and R.D. Van der Hilst, *Joint seismic tomography for bulk-sound and shear wavespeed*, J.geophys. Res., 103 (1998), pp. 12469–12493.

[9] N. Kingsbury, *Image processing with complex wavelets*, Phil. Trans.Roy.Soc.Lond., A357 (1999), pp. 2543–2560.

[10] ———, *Complex wavelets for shift invariant analysis and filtering of signals*, Appl. Computat. Harmon. Anal., 10 (2002), pp. 234–253.

[11] I. Loris, G. Nolet, I. Daubechies, and F. A. Dahlen, *Tomographic inversion using* 1*-norm regularization of wavelet coefficients*, Geoph. J. Int., 170 (2007), pp. 359–370.

[12] Y. Luo and G.T. Schuster, *Wave-equation travel time inversion*, Geophysics, 56 (1991), pp. 645–653.

[13] H. Marquering, G. Nolet, and F.A. Dahlen, *Three-dimensional waveform sensitivity kernels.*, Geoph. J. Int., 132 (1998), pp. 521–534.

[14] L. Pestov and G. Uhlmann, *Two dimensional compact simple riemannian manifolds are boundary distance rigid*, Annals of Math., 161 (2005), pp. 1089–1106.

[15] M. Salo, *Stability for solutions of wave equations with $C^{1,1}$ coefficients*, Inverse Problems and Imaging, 1 (2007), pp. 537–556.
[16] A. Sieminski, Q. Liu, J. Trampert, and J. Tromp, *Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods.*, Geoph. J. Int., 171(1) (2007), pp. 368–389.
[17] L. Sirgue and R. Pratt, *Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies*, Geophysics, 69 (2004), pp. 231–248.
[18] H.F. Smith, *A parametrix construction for wave equations with $C^{1,1}$ coefficients*, Ann. Inst. Fourier, Grenoble, 48 (1998), pp. 797–835.
[19] ———, *Spectral cluster estimates for $C^{1,1}$ metrics*, Amer. J. Math., 128 (2006), pp. 1069–1103.
[20] C.C. Stolk, *On the modeling and inversion of seismic data*, Ph.D.thesis, Utrecht University, 2000.
[21] C.C. Stolk and M.V. de Hoop, *Microlocal analysis of seismic inverse scattering in anisotropic, elastic media*, Comm. Pure Appl. Math., 55 (2002), pp. 261–301.
[22] ———, *Seismic inverse scattering in the downward continuation approach*, Wave Motion., 43 (2006), pp. 579–598.
[23] W.-J. Su and A.M. Dziewonski, *Simultaneous inversion for 3-d variations in shear and bulk velocity in the mantle*, Phys.Earth planet. Inter., 100 (1997), pp. 135–156.
[24] D. Tataru, *Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation*, Amer. J. Math., 122 (2000), pp. 349–376.
[25] ———, *Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II*, Amer. J. Math., 123 (2001), pp. 385–423.
[26] ———, *Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III*, J. Amer. Math. Soc., 15 (2002), pp. 419–442.
[27] M. V. de Hoop and R.D. Van der Hilst, *On sensitivity kernels for "wave-equation" transmission tomography*, Geophys. J. Int., 160 (2005), pp. 621–633.
[28] M. V. de Hoop, R.D. Van der Hilst, and P. Shen, *Wave-equation reflection tomography: Annihilators and sensitivity kernels*, Geoph. J. Int., 167 (2006), pp. 1332–1352.
[29] R.D. Van der Hilst, S. Widyantoro, and E.R. Engdahl, *Evidence for deep mantle circulation from global tomography*, Nature, 386 (1997), pp. 578–584.
[30] M.J. Woodward, *Wave-equation tomography*, Geophysics, 57 (1992), pp. 12–26.
[31] L. Zhao, T. Jordan, and C. Chapman, *Three-dimensional frechet differential kernels for seismic delay times.*, Geoph. J. Int., 141 (2000), pp. 558–576.