



MICROLOCAL ANALYSIS OF BOREHOLE SEISMIC DATA

RALUCA FELEA

School of Mathematical Sciences
Rochester Institute of Technology, Rochester, NY, 14623
USA

ROMINA GABURRO

Department of Mathematics and Statistics, CONFIRM, HRI
University of Limerick, Limerick, V94 T9PX
Ireland

ALLAN GREENLEAF*

Department of Mathematics
University of Rochester, Rochester, NY, 14627
USA

CLIFFORD NOLAN

Department of Mathematics and Statistics, CONFIRM, HRI
University of Limerick, Limerick, V94 T9PX
Ireland

*Dedicated to the memory of Victor Isakov,
for his contributions to the field and for his friendship.*

ABSTRACT. Borehole seismic data is obtained by receivers located in a well, with sources located on the surface or another well. Using microlocal analysis, we study possible approximate reconstruction, via linearized, filtered backprojection, of an isotropic sound speed in the subsurface for three types of data sets. The sources may form a dense array on the surface, or be located along a line on the surface (walkaway geometry) or in another borehole (crosswell). We show that for the dense array, reconstruction is feasible, with no artifacts in the absence of caustics in the background ray geometry, and mild artifacts in the presence of fold caustics in a sense that we define. In contrast, the walkaway and crosswell data sets both give rise to strong, nonremovable artifacts.

1. Introduction. In seismic acoustic imaging, borehole data refers to measurements of waves made by receivers (sensors) at various depths in a well; applications include prospecting for CO₂ sequestration sites or geothermal reservoirs, and monitoring aquifer pollution or existing hydrocarbon reservoirs. In Vertical Seismic Profiling, the waves are excited by sources located at positions on the surface [3, 31]; in crosswell (or crosshole) imaging, the sources are in another well [5, 6, 1, 33]. Compared to data resulting from traditional seismic experiments, where both the sources and receivers are located on the surface, the decreased travel distance for

2020 *Mathematics Subject Classification.* Primary: 86A22, 35R30, 35S30.

Key words and phrases. Seismic imaging, inverse problem, artifact, Fourier integral operator, microlocal analysis, caustic.

* Corresponding author: Allan Greenleaf, allan@math.rochester.edu.

the waves traveling to receivers in a borehole results in less attenuation and allows the use of higher frequency waves, potentially resulting in more sensitive and higher resolution imaging of material parameters in the subsurface [31].

The purpose of this work is to formulate a general approach to the analysis of borehole seismic data, using techniques of microlocal analysis that have previously been successful for conventional seismic data [24, 30, 28, 34, 8, 9, 11, 12] (and also for a variety of other imaging problems). As in those works, here we analyze the relation between features in the subsurface, in the form of singularities of the sound speed profile, and the resulting singularities of the data. In some cases, the latter accurately encode the former; in others, imaging artifacts arise from the data acquisition geometry, the presence of caustics (multi-pathing) in the subsurface, or the interaction of the two. For several specific borehole data geometries, we either show that the imaging is artifact free, or determine the location and strength of the artifacts.

In applying microlocal analysis to inverse problems, one is using a set of tools whose theoretical foundation is rigorously established in the high frequency limit to problems where the data is intrinsically band-limited. Past work has shown, however, that this can be a fruitful approach, since in practice frequencies do not have to be very high for the high frequency limit to be a good enough approximation that useful conclusions can be drawn regarding the structure and strength of artifacts in images produced from the data.

For simplicity, assume the Earth is $\mathbb{R}_+^3$; its surface, $\mathbb{R}^2 = \{x_3 = 0\}$, is flat; and x_3 increases with depth. Throughout, we will assume that the set Σ_R of receivers occupies a line segment located in a vertical borehole along the positive x_3 -axis,

$$\Sigma_R := \{(0, 0, r) : r_{min} < r < r_{max}\}.$$

In Vertical Seismic Profiling, waves are generated by impulses located at a set Σ_S of sources located on the surface $\mathbb{R}^2$, scatter off of features in the subsurface, and are then measured at points of Σ_R at all times $t \in T = (t_{min}, t_{max})$. By contrast, for crosswell imaging the sources are located in another borehole some distance from the borehole containing the receivers. Possible data sets $\mathbb{D} = \Sigma_S \times \Sigma_R \times T$ may be crudely classified by their dimensionality, depending on whether S is zero-, one- or two-dimensional. If $\dim \Sigma_S = 0$, then the source set is at a single offset, or at most a discrete set of points; in this case, $\dim \mathbb{D} = 2$, so that the data set is underdetermined, which is not of interest for the questions we pose. Instead, we study three basic data sets:

Overdetermined ($\dim \mathbb{D} = 4$):

For the **dense array** data acquisition geometry (also called 3D Vertical Seismic Profiling), $\Sigma_S \subset \partial \mathbb{R}_+^3 \setminus \mathbf{0} = \mathbb{R}^2 \setminus (0, 0)$ is an open subset on the surface.

Determined ($\dim \mathbb{D} = 3$):

In the **crosswell** geometry, the sources are located in another vertical borehole, along a line segment,

$$\Sigma_S := \{(s_0, s) : s_{min} < s < s_{max}\}, \quad s_0 \neq (0, 0).$$

For the **walkaway** geometry, Σ_S is contained in a line on the surface, passing over the borehole top $(0, 0, 0)$. Without loss of generality, we assume that

$$\Sigma_S := \{(s, 0, 0) : 0 < s_{min} < s < s_{max}\}.$$

Remark 1. To avoid having to deal with uninteresting degeneracies arising from sources located very close to the top of the borehole, for all of the geometries we only consider sources $S = (s_0, 0) = (s_1, s_2, 0)$ with $|(s_1, s_2)| > \epsilon > 0$.

Our goal is to study, for each of the data sets $\mathbb{D}$ considered here, the formal linearization $d\mathcal{F}$ about a known smooth background sound speed, $c_0(x)$ of the sound speed-to-data map, restricted to $\mathbb{D}$.

In Sec. 2, we describe the model and the linearization of the forward scattering operator and recall the Fourier integral operator theory needed in the paper. (The C^∞ singularity theory, describing the types of degeneracies of smooth functions - folds, blowdowns, submersions with folds and cross caps - that we will use to understand the structure of the forward and normal operators for the various data sets, is included as an appendix, Sec. 7.)

The remainder of the paper analyzes, for the data acquisition geometries described above, the linearized scattering operator $F := d\mathcal{F}$, which is a Fourier integral operator; the geometry of its canonical relation; and the implications for the associated normal operator, F^*F . We focus on the overdetermined dense array data set, for which the results are the most positive. First, in Thm. 3.1, we show that if the ray geometry of the background sound speed has no caustics, and satisfies two additional technical assumptions, then F satisfies the Traveltime Injectivity Condition, introduced in [24, 30] when both the sources and receivers are on the surface. This implies that the normal operator is a pseudodifferential operator, and filtered backprojection does not give rise to artifacts in the images. We first prove this for a constant background sound speed in Sec. 3, where the additional assumptions are unnecessary. This is followed by the analysis in Sec. 4 of variable c_0 with no caustics, where the additional assumptions are used.

Then, in Sec. 5 we study the situation for the dense array geometry when the most commonly encountered form of caustics (or multipathing) is present. We formulate a notion of *caustics of fold type* appropriate for this setting and show that, in the presence of caustics no worse than this type, the canonical relation of F , while degenerate, has a structure, that of a *folded cross cap*, introduced previously by two of the authors in the context of marine seismic imaging [11]; see Thm. 5.2 and Def. 5.3. This allows a precise description of the normal operator and characterization of imaging artifacts microlocally away from high codimension sets. We point out that the marine source-receiver manifold treated in [11] and the borehole dense array here are inequivalent: one cannot be reduced to the other by changes of variables in the source and receiver spaces, and necessarily the definition of fold caustics are distinct. Nevertheless, the microlocal geometries of the resulting canonical relations are of the same type (the folded cross caps), and although the calculations needed to show that the dense array leads to a folded cross cap are similar to those in [11], for completeness we present these in full.

In contrast, for the crosswell and walkaway data sets, the microlocal geometry is less favorable, resulting in strong, nonremovable artifacts. In Sec. 6 we analyze their normal operators when the background sound speed c_0 is constant. Calculating and analyzing their canonical relations shows that attempted inversion by filtered back projection results in strong artifacts; see Thms. 6.1 and 6.2. Since these are already badly behaved for a constant background sound speed, we do not pursue the analysis of the crosswell and walkaway geometries for variable backgrounds.

2. Scattering model, microlocal analysis and singularity theory. The idea of using techniques from microlocal analysis to study linearized seismic imaging was introduced by Beylkin [4], and expanded upon by Nolan and Symes [30] and ten Kroode, Smit and Verdel [24]. In the forward scattering problem, acoustic waves are generated at the surface of the Earth, scatter off of features in the subsurface, and some of the reflected waves return to the surface to be detected by receivers, which in these works were also at the surface. The goal of the full inverse problem is to obtain an image of the subsurface using measurements of the pressure field at various receivers. Due to the strong nonlinearity of the full problem, works such as [4, 30, 24] instead considered a formal linearization of the nonlinear sound speed-to-data map, $\mathcal{F}$. The linearization maps perturbations of a smooth background sound speed in the subsurface (assumed known), to perturbations of the resulting pressure field at receivers. The linearized problem was explored further by several authors for a variety of data acquisition geometries; see, e.g., [28, 34, 8, 9, 11, 12]. In this section, we review the scattering model and its linearization, and set down or give references to the basic microlocal analysis needed for the remainder of the paper; a summary of the requisite singularity theory is in the appendix, Sec. 7. Since this material, in forms suitable for what is needed here, is standard, we will keep the presentation as brief as possible.

2.1. Scattering model and normal operator. We now recall the scattering problem and its linearization. Represent the Earth as $Y = \mathbb{R}_+^3 = \{x \in \mathbb{R}^3, x_3 \geq 0\}$, consisting of isotropic material with sound speed $c(x)$, the recovery of which is our goal. An impulse at a source $x = s, t = 0$, assumed for simplicity to be a delta waveform, creates a pressure field $p(s; x, t)$ which solves the acoustic wave equation,

$$\begin{aligned} \frac{1}{c^2(x)} \frac{\partial^2 p}{\partial t^2}(s; x, t) - \Delta p(s; x, t) &= \delta(t) \delta(x - s) \\ p(s; x, t) &= 0, \quad t < 0, \end{aligned} \quad (1)$$

where Δ is the Laplacian on $\mathbb{R}^3$. Fixing a data acquisition set $\mathbb{D} = \Sigma_S \times \Sigma_R \times T$, where Σ_S is a set of sources, Σ_R is a set of receivers, and $T = (t_{\min}, t_{\max})$ is a time interval, the corresponding forward map $\mathcal{F} = \mathcal{F}_{\mathbb{D}}$ is the sound speed-to-data map, $c \xrightarrow{\mathcal{F}} p(s; r, t)|_{(s,r,t) \in \mathbb{D}}$; the full inverse problem is to reconstruct c from $\mathcal{F}(c)$.

Due to the nonlinearity of $\mathcal{F}$, the works cited below instead considered the formal linearization of $\mathcal{F}$, assuming c to be of the form $c = c_0 + \delta c$, with c_0 a known smooth background sound speed. Thus, associated Green's function, i.e., background pressure field p_0 satisfying (1) for c_0 , is also known (in principle). The linearization $d\mathcal{F}$ of $\mathcal{F}$ then arises from writing $p = p_0 + \delta p \bmod (\delta c)^2$, where $\delta p =: (d\mathcal{F})(\delta c)$ satisfies

$$\begin{aligned} \square_{c_0}(\delta p) &:= \frac{1}{c_0^2(x)} \frac{\partial^2 \delta p}{\partial t^2}(s; x, t) - \Delta \delta p(s; x, t) = \frac{2}{c_0^3(x)} \cdot \frac{\partial^2 p_0}{\partial t^2} \cdot \delta c(x) \\ \delta p &= 0, \quad t < 0. \end{aligned} \quad (2)$$

We denote $d\mathcal{F}$ by F , using a subscript $\mathbb{D}$ if needed for clarity. Thus,

$$F(\delta c)(s, r, t) = \square_{c_0}^{-1} \left(\frac{2}{c_0^3(x)} \cdot \frac{\partial^2 p_0}{\partial t^2} \cdot \delta c(x) \right) (s; x, t) \Big|_{x=r}, \quad (3)$$

where $\square_{c_0}^{-1}$ is the forward solution operator. One assumes that δc is supported at a positive distance from all sources s , so that its product above with the Green's

function with pole at s only involves the singularities of p_0 on the wavefront, and not at s . (The linearization can be justified in terms of Fréchet differentiability of $\mathcal{F}$ between certain pairs of function spaces [23].)

Beylkin [4] showed that, for a single source on the surface and an open set of receivers, also on the surface, if caustics do not occur for the background sound speed, then the normal operator $N := F^*F$ is a pseudodifferential operator (ΨDO). To avoid difficulties with even the basics of the operator theory, for more general ray geometries two conditions have been commonly assumed in the literature: (i) no single (unbroken) ray connects a source to a receiver; and (ii) no ray originating in the subsurface grazes Σ_S or Σ_R . Condition (ii) ensures that F is a Fourier integral operator (FIO) in the sense of Hörmander [22], while (i) ensures that the composition F^*F makes sense on distributions. Under these assumptions, in the case of a single source on the surface, and receivers forming an open subset of the surface $\{x_3 = 0\}$, Rakesh [32] showed that F is an FIO, and this was extended to other surface-surface data sets $\mathbb{D}$ in [21, 30, 24]. However, for *borehole* data sets these conditions no longer make physical sense since, for example, one fully expects unbroken rays to travel from source points on the surface to receivers in the borehole (as they already do for a constant sound speed), and so we replace (i) and (ii) by Assumption 3.2 below which can be satisfied if the data has been suitably filtered (muted). See Sec. 3 for the statement of Assumption 3.2 and the description of a filter to implement it.

The invertibility of F (modulo C^∞) was established in [4, 30, 24] under various combinations of assumptions on the data acquisition set $\mathbb{D}$ (assumed to be a smooth manifold in $\partial Y \times \partial Y$) and the ray geometry for the background sound speed c_0 . In these cases, the FIO F is associated with a canonical relation $C \subset T^*\mathbb{D} \times T^*Y$ which satisfies the so-called *traveltime injectivity condition* (TIC) described below. By the standard theory of FIO with nondegenerate canonical relations, it follows that the normal operator N is a pseudodifferential operator on Y , $N \in \Psi(Y)$; furthermore, N is elliptic (and hence invertible microlocally) under an illumination condition. If $Q \in \Psi(Y)$ is a left-parametrix for N (i.e., a left-inverse modulo C^∞), then $Q \circ F^*$ is a left-parametrix for F . This implies the injectivity of $dF \bmod C^\infty$, so that the singularities of $F(\delta c)$ determine the singularities of δc , as well as giving an approximate reconstruction formula via filtered backprojection. In [34], part of the TIC is relaxed, but the composition F^*F is still covered by the standard clean composition calculus for FIO.

On the other hand, combinations of data sets and background ray geometries for which the TIC is violated were studied in [28, 8, 9, 11, 12]. In each of these, the composition forming the normal operator lies outside the clean intersection calculus and N is not a pseudodifferential operator. The wavefront relation of N is larger than that of a ΨDO , including an additional part, which gives rise to artifacts in the image when attempting filtered back projection; the strength of that artifact depends both on the geometry of $\mathbb{D}$, the nature of the multi-pathing (if any), the background ray geometry, and their interaction.

2.2. Microlocal analysis. We now recall some basic definitions and results from the theory of FIOs [22]. Let X and Y be smooth manifolds, of (possibly different) dimensions, n_X, n_Y , resp. A *Fourier integral operator* is a continuous linear map $A : \mathcal{E}'(Y) \rightarrow \mathcal{D}'(X)$, whose Schwartz kernel is a locally finite sum of oscillatory integrals of the form

$$K_A(x, y) = \int_{\mathbb{R}^N} e^{i\varphi(x, y; \theta)} a(x, y; \theta) d\theta,$$

where φ is a nondegenerate operator phase function on $X \times Y \times (\mathbb{R}^N \setminus 0)$, for some $N \geq 1$, and a is a Hörmander class amplitude of order μ and type $(1, 0)$. The *order* of A is defined to be

$$m := \mu + \frac{2N - n_X - n_Y}{4},$$

and the *canonical relation* of A is

$$\begin{aligned} C_A &:= \{(x, d_x \varphi; y, -d_y \varphi) : (x, y; \theta) \in \text{supp}(a), d_\theta \varphi(x, y; \theta) = 0\} \\ &\subset (T^*X \setminus 0) \times (T^*Y \setminus 0). \end{aligned}$$

If $WF(\cdot)$ denotes the C^∞ wavefront set of a distribution, the *wavefront relation* of A , $WF(A) := WF(K_A)'$, is the image of the wavefront set of the Schwartz kernel of A under the map $(x, y, \xi, \eta) \rightarrow (x, \xi; y, -\eta)$; from the general theory of Fourier integral distributions, one knows that $WF(A) \subseteq C_A$. Thus, by the Hörmander-Sato Lemma, for all $u \in \mathcal{E}'(Y)$,

$$WF(Au) \subseteq WF(A) \circ WF(u) \subseteq C_A \circ WF(u), \quad (4)$$

where $WF(A)$ and C_A are considered as relations from $T^*Y \setminus 0$ to $T^*X \setminus 0$.

For any canonical relation $C \subset (T^*X \setminus 0) \times (T^*Y \setminus 0)$ and $m \in \mathbb{R}$, $I^m(X, Y; C)$ denotes the class of properly supported m -th order FIOs A with $C_A \subseteq C$. Thus, for any A in this class, $WF(Au) \subseteq C \circ WF(u)$, $\forall u \in \mathcal{E}'(Y)$.

A generalization of Fourier integral operators are the *paired Lagrangian* operators of Melrose and Uhlmann [26] and Guillemin and Uhlmann [20]. These are associated to cleanly intersecting pairs of canonical relations, $C_0, C_1 \subset T^*X \times T^*Y$, and are indexed by bi-orders $(p, l) \in \mathbb{R}^2$. We will not need the definitions and characterizations of these operators, but note two properties for later use. First, if $A \in I^{p, l}(X, Y; C_0, C_1)$, then

$$WF(A) \subseteq C_0 \cup C_1. \quad (5)$$

Secondly, microlocally away from $C_0 \cap C_1$,

$$A \in I^{p+l}(C_0 \setminus C_1) \text{ and } A \in I^p(C_1 \setminus C_0). \quad (6)$$

Now let $C_1 \subset (T^*X \setminus 0) \times (T^*Y \setminus 0)$ and $C_2 \subset (T^*Y \setminus 0) \times (T^*Z \setminus 0)$ be two canonical relations, and $A_1 \in I^{m_1}(X, Y; C_1)$ and $A_2 \in I^{m_2}(Y, Z; C_2)$. If $C_1 \times C_2$ intersects $T^*X \times \Delta_{T^*Y} \times T^*Z$ *transversely*, then Hörmander proved that $A_1 \circ A_2 \in I^{m_1+m_2}(X, Z; C_1 \circ C_2)$ where $C_1 \circ C_2$ is the composition of C_1 and C_2 as relations in $T^*X \times T^*Y$ and $T^*Y \times T^*Z$. Duistermaat and Guillemin [7] and Weinstein [36] extended this to the case of *clean intersection* and showed that if $C_1 \times C_2$ and $T^*X \times \Delta_{T^*Y} \times T^*Z$ intersect cleanly with excess e then, as in the transverse case, $C_1 \circ C_2$ is again a smooth canonical relation, and

$$A_1 \circ A_2 \in I^{m_1+m_2+e/2}(X, Z; C_1 \circ C_2).$$

We say that a canonical relation $C \subset T^*X \times T^*Y$ satisfies the *traveltime injectivity condition* (TIC) [30, 24] (equivalent to the earlier *Bolker* condition in tomography [19]) if the natural projection to the left, $\pi_L : C \rightarrow T^*X$, satisfies the following two conditions. First,

$$\pi_L \text{ is an immersion, i.e., } d\pi_L \text{ is injective everywhere.} \quad (7)$$

(By results for general canonical relations, this is equivalent with $\pi_R : C \rightarrow T^*Y$ being a submersion, i.e., $d\pi_R$ is surjective.) Secondly,

$$\pi_L \text{ is globally injective.} \quad (8)$$

(Note that (7) already implies that π_L is *locally* injective; (8) demands that the injectivity holds globally.)

If $A \in I^m(X, Y; C)$, then $A^* \in I^m(Y, X; C^t)$. If C satisfies the TIC, then it follows from (7) that the composition A^*A is covered by the clean intersection calculus, with excess $e = \dim(X) - \dim(Y)$. Furthermore, from (8) it follows that $C^t \circ C \subseteq \Delta_{T^*Y}$, the diagonal of $T^*Y \times T^*Y$. Thus, the normal operator

$$N := A^*A \in I^{2m + \frac{\dim(X) - \dim(Y)}{2}}(Y, Y; \Delta_{T^*Y}), \quad (9)$$

i.e., is a pseudodifferential operator on Y . N is elliptic if A is, which in applications corresponds to an illumination condition. In that case, N admits a left parametrix $Q \in \Psi^{-2m}(Y)$, i.e., $Q \circ N - I$ is a smoothing operator, and then QA^* is a left parametrix for A , so that, for all $u \in \mathcal{E}'(Y)$, $Au \bmod C^\infty$ determine $u \bmod C^\infty$.

However, in many inverse problems, the TIC condition fails, and to understand the possibility of imaging using filtered back projection, it is important to analyze the composition A^*A and the nature of the resulting normal operator, N . Any component of the wavefront relation of N in the complement of the diagonal Δ_{T^*Y} will produce *artifacts*, i.e., features in Nu which are not present in u . It turns out that the geometry of the canonical relation C , as expressed by degeneracies of projections π_L and π_R , if they exist, plays an important role in determining the nature, location and strength of artifacts.

It is known that if either $d\pi_L$ or $d\pi_R$ has maximal rank, so does the other one and we say that the canonical relation C is *nondegenerate*. In this case the composition $C^t \circ C$ is clean, and (9) holds.

On the other hand, if C is degenerate (the differentials of the projections fail to be of maximal rank), there is no general theory that applies to the compositions $C^t \circ C$ and A^*A . However, certain particular geometries have been analyzed, and one in particular is relevant here, for the dense array in the presence of fold caustics.

When one of the projections drops rank, then the other one does, too, and their coranks are the same. However, although $\text{corank}(d\pi_L) = \text{corank}(d\pi_R)$ at all points, the two projections might have the same type of singularity, or quite different ones. The singularities needed in this article are *blowdowns*, *folds*, *submersion with folds* and *cross caps*. We refer the reader to Sec. 7 for a concise summary of these classes; see [17, 37, 27] for more background, and to Sec. 5 for the existing composition calculus [11], originating in marine seismic imaging, that we show is relevant for the dense array with fold caustics.

3. Dense array: Constant c_0 . We start with the dense array geometry, with sources $S = (s_1, s_2, 0)$ in an open subset Σ_S of the surface, and receivers in the borehole Σ_R , $R = (0, 0, r)$, $r \in (r_{\min}, r_{\max})$. The perturbation in sound speed is a function of the three variables, $y = (y_1, y_2, y_3)$, while the resulting data is a function of four variables, $(s_1, s_2, r, t) = (s, r, t)$. We make the following assumptions, which, as discussed above in Sec. 2.1 are modifications appropriate to the borehole setting of standard assumptions in the literature (cf. [32, 24, 30]), needed for the operators to satisfy even the most basic conditions:

Assumption 3.1. *The perturbation δc in the sound speed has compact support at a positive depth below the surface.*

Assumption 3.2. *For any unbroken ray connecting a source to a receiver which intersects the support of the reflectivity function, its contribution to the data has been muted by application of the filter described below. (See Figure 1 for an illustration of this.)*

Filter Construction: Suppose a ray connects a source to a receiver located at $r_0 \in \Sigma_R$ and that this ray arrives at r_0 in a direction $\hat{\rho}_0$. Let ρ_0 be the orthogonal projection of $\hat{\rho}_0$ onto $T_{r_0}\Sigma_r$.

(i) Localize the data $d(s, r, t)$ by multiplying it by a cutoff function $\chi_1(r)$ supported near $r = r_0$, and then take the partial Fourier transform of $\chi_1 d(s, r, t)$ with respect to r to get $\hat{d}_1(s, \rho, t)$, where ρ is the Fourier variable dual to r .

(ii) Multiply $\hat{d}_1$ by a cutoff function $\chi_2(\rho)$ which is homogeneous of degree 1 and vanishes in a conic neighborhood of the direction of ρ_0 whenever $\hat{d}_1(s, \lambda \rho_0, t)$ is not rapidly decaying as $\lambda \rightarrow \infty$.

(iii) Apply the inverse Fourier transform (w.r.t. ρ) to $\chi_2 \hat{d}_1$, and use the result as the suitably modified data, referred to in Assumption 3.2 above.

In this section and the next, we show that the linearized scattering operator, filtered as above, satisfies the Traveltime Injectivity Condition:

Theorem 3.1. *Suppose, in addition to Assumptions 3.1 and 3.2, the ray geometry of a smooth background sound speed $c_0(x)$ satisfies Assumptions 4.1 and 4.2 below. Then the linearized scattering operator for the dense array data set, $F : \mathcal{E}'(\mathbb{R}_+^3) \rightarrow \mathcal{D}'(\mathbb{D})$, is a Fourier integral operator, $F \in I^{\frac{3}{4}}(\mathbb{D}, \mathbb{R}_+^3; C)$.*

If c_0 also has no caustics, then the canonical relation $C \subset T^\mathbb{D} \times T^*\mathbb{R}_+^3$ satisfies the Traveltime Injectivity Condition (7), (8), and thus the normal operator is a pseudodifferential operator of order 2, $F^*F \in \Psi^2(\mathbb{R}_+^3)$.*

In the current section, we consider first the model case of constant sound speed, normalized to $c_0 = 1$, for which Assumptions 4.1 and 4.2 hold automatically. We will show that the canonical relation of the linearized forward scattering operator F satisfies the traveltime injectivity condition (7), (8). As discussed in Sec. 2.2, this implies that the normal operator F^*F is a pseudodifferential operator, and a perturbation δc of the sound speed can be reconstructed from $F(\delta c)$ by filtered backprojection.

Proof of Thm. 3.1. In the case of constant background sound speed c_0 , we compute the canonical relations of $F = d\mathcal{F}$ for each of the data geometries (dense array here; crosswell and walkaway in Sec. 6), restricting to the various data sets the basic phase function

$$\phi(s, r, t; \omega) = (t - |y - R| - |y - S|) \omega, \quad (10)$$

where $S = S(s)$, $R = R(r)$ and y denote a general source, receiver and a point in the subsurface, respectively, and $\omega \in \mathbb{R} \setminus 0$ is a phase variable. Note that the function ϕ defined in (10) is a non-degenerate phase function as can be easily verified by noting that $\omega \neq 0$ and also using Assumption 3.2. The Schwartz kernel of F is

$$K(s, r, t, y) = \int e^{i\phi(s, r, t, y; \omega)} a(s, r, t, y; \omega) d\omega,$$

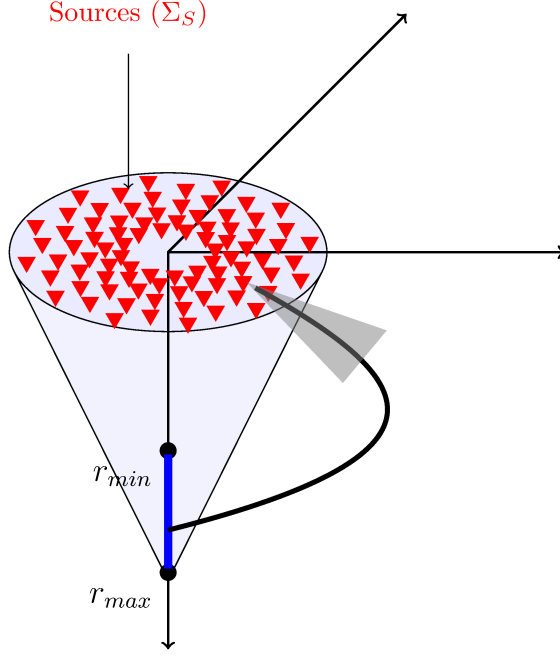


FIGURE 1. Illustration of data acquisition geometry and filtering: contributions to the data from unbroken rays such as that illustrated here are filtered out by removing data associated to those rays arriving from nearby directions, as indicated by the gray cone.

where ϕ is given by

$$\begin{aligned}\phi(s, r, t, y; \omega) &= (t - |y - S| - |y - R|) \omega \\ &= \left(t - \sqrt{(y_1 - s_1)^2 + (y_2 - s_2)^2 + y_3^2} - \sqrt{y_1^2 + y_2^2 + (y_3 - r)^2} \right) \omega\end{aligned}$$

and $a \in S_{cl}^2$ is a classical symbol of order 2. Thus, F is a Fourier integral operator of order $m = 2 + \frac{1}{2} - \frac{4+3}{4} = \frac{3}{4}$ associated with the canonical relation $C \subset T^*\mathbb{R}^4 \times T^*\mathbb{R}^3$ parametrized by ϕ . See [30] for a discussion of F as an FIO for general data sets.

Coordinates on the 7-dimensional C can be taken to be $(s, r, y, \omega) \in \mathbb{R}^6 \times (\mathbb{R} \setminus 0)$:

$$\begin{aligned}C &= \left\{ (s_1, s_2, r, A + B, \frac{y_1 - s_1}{A} \omega, \frac{y_2 - s_2}{A} \omega, \frac{y_3 - r}{B} \omega, \omega; \right. \\ &\quad \left. y_1, y_2, y_3, \left(\frac{y_1 - s_1}{A} + \frac{y_1}{B} \right) \omega, \left(\frac{y_2 - s_2}{A} + \frac{y_2}{B} \right) \omega, \left(\frac{y_3}{A} + \frac{y_3 - r}{B} \right) \omega \right\}, \quad (11)\end{aligned}$$

where $A = \sqrt{(y_1 - s_1)^2 + (y_2 - s_2)^2 + y_3^2}$ and $B = \sqrt{y_1^2 + y_2^2 + (y_3 - r)^2}$. Thus, $C \subset (T^*\mathbb{R}^4 \setminus 0) \times (T^*\mathbb{R}^3 \setminus 0)$, and we see that the projections to left and right are

$$\pi_L(s, r, y, \omega) = \left(s_1, s_2, r, A + B, \frac{y_1 - s_1}{A} \omega, \frac{y_2 - s_2}{A} \omega, \frac{y_3 - r}{B} \omega, \omega \right) \quad (12)$$

and

$$\begin{aligned} \pi_R(s, r, y, \omega) \\ = \left(y_1, y_2, y_3, \left(\frac{y_1 - s_1}{A} + \frac{y_1}{B} \right) \omega, \left(\frac{y_2 - s_2}{A} + \frac{y_2}{B} \right) \omega, \left(\frac{y_3}{A} + \frac{y_3 - r}{B} \right) \omega \right). \end{aligned} \quad (13)$$

Since π_R is the identity in y variables, $\text{rank}(d\pi_R) = 3 + \text{rank}\left(\frac{D\eta}{D(s, r, \omega)}\right)$, where η is dual to y , and the minor $\frac{D\eta}{D(s, \omega)}$ is

$$\frac{D\eta}{D(s, \omega)} = \begin{bmatrix} -\frac{(y_2 - s_2)^2 + y_3^2}{A^3} \omega & \frac{(y_1 - s_1)(y_2 - s_2)}{A^3} \omega & \frac{y_1 - s_1}{A} + \frac{y_1}{B} \\ \frac{(y_1 - s_1)(y_2 - s_2)}{A^3} \omega & -\frac{(y_1 - s_1)^2 + y_3^2}{A^3} \omega & \frac{y_2 - s_2}{A} + \frac{y_2}{B} \\ \frac{y_3(y_1 - s_1)}{A^3} \omega & \frac{(y_2 - s_2)y_3}{A^3} \omega & \frac{y_3}{A} + \frac{y_3 - r}{B} \end{bmatrix},$$

which has determinant

$$\omega^2 y_3 A^{-3} \left(1 + \frac{(y - S) \cdot (y - R)}{AB} \right).$$

Using $\omega \neq 0$, $y_3 > 0$, the Cauchy-Schwarz inequality and Assumption 3.2, one sees that $\det\left[\frac{D\eta}{D(s, \omega)}\right] \neq 0$, so that $\text{rank}(d\pi_R) = 6$. It follows that $d\pi_R$ has maximal rank and π_R is a submersion; hence π_L is an immersion, and C is a nondegenerate canonical relation.

To verify the Traveltime Injectivity Condition, it remains to show the injectivity of π_L ; we do this using Assumption 3.2. The unit vectors $(y - (0, 0, r))/B$ and $(y - (s_1, s_2, 0))/A$ point to y from the source $S = (s_1, s_2, 0)$ and from the receiver $R = (0, 0, r)$, resp. In terms of these, the condition in Assumption 3.2 is that

$$\frac{1}{A} (y - (s_1, s_2, 0)) + \frac{1}{B} (y - (0, 0, r)) \neq 0.$$

To prove that π_L is injective, let us consider

$$S = (s_1, s_2, 0); \quad R = (0, 0, r); \quad y = (y_1, y_2, y_3); \quad \omega;$$

and

$$\tilde{S} = (\tilde{s}_1, \tilde{s}_2, 0); \quad \tilde{R} = (0, 0, \tilde{r}); \quad \tilde{y} = (\tilde{y}_1, \tilde{y}_2, \tilde{y}_3); \quad \tilde{\omega};$$

such that $\pi_L(s, r, y, \omega) = \pi_L(\tilde{s}, \tilde{r}, \tilde{y}, \tilde{\omega})$. Then

$$s_1 = \tilde{s}_1, \quad s_2 = \tilde{s}_2, \quad r = \tilde{r}, \quad \omega = \tilde{\omega}, \quad A + B = \tilde{A} + \tilde{B}$$

$$\frac{y_1 - s_1}{A} \omega = \frac{\tilde{y}_1 - \tilde{s}_1}{\tilde{A}} \tilde{\omega}, \quad \frac{y_2 - s_2}{A} \omega = \frac{\tilde{y}_2 - \tilde{s}_2}{\tilde{A}} \tilde{\omega}, \quad \frac{y_3 - r}{B} \omega = \frac{\tilde{y}_3 - \tilde{r}}{\tilde{B}} \tilde{\omega}, \quad (14)$$

where

$$A = |y - S|; \quad B = |y - R|; \quad \tilde{A} = |\tilde{y} - \tilde{S}|; \quad \tilde{B} = |\tilde{y} - \tilde{R}|. \quad (15)$$

The first four equalities in (14) imply that $\tilde{S} = S$, $\tilde{R} = R$, $\tilde{\omega} = \omega$, so to prove injectivity of π_L we only need to verify that $\tilde{y} = y$. Defining

$$\sigma := \frac{y - S}{A}; \quad \tilde{\sigma} := \frac{\tilde{y} - \tilde{S}}{\tilde{A}}; \quad \rho := \frac{y - R}{B}; \quad \tilde{\rho} := \frac{\tilde{y} - \tilde{R}}{\tilde{B}}, \quad (16)$$

the last three equalities in (14) (together with the fact that $\omega = \tilde{\omega}$) imply that

$$\sigma_i = \tilde{\sigma}_i, \quad i = 1, 2; \quad (17)$$

$$\rho_3 = \tilde{\rho}_3. \quad (18)$$

Recalling that $\sigma = (\sigma_1, \sigma_2, \sigma_3)$ and $\tilde{\sigma} = (\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3)$ are unit vectors, we obtain

$$\begin{aligned}\sigma_3^2 &= 1 - \sigma_1^2 - \sigma_2^2 = 1 - \tilde{\sigma}_1^2 - \tilde{\sigma}_2^2 = \tilde{\sigma}_3^2 \\ &\Rightarrow \tilde{\sigma}_3 = \pm \sigma_3.\end{aligned}$$

However, $y_3, \tilde{y}_3 > 0 \Rightarrow \sigma_3, \tilde{\sigma}_3 > 0$, hence $\sigma = \tilde{\sigma}$, that is

$$\frac{\tilde{y} - S}{|\tilde{y} - S|} = \frac{y - S}{|y - S|}, \quad (19)$$

which shows that y and $\tilde{y}$ lie on the same ray emanating from S . The last equality in (14) can be expressed as

$$T(s, r, y) = T(s, r, \tilde{y}), \quad \text{where } T(s, r, y) = A(y, s) + B(y, r). \quad (20)$$

In the argument below, if $z \in \mathbb{R}^3$ and $z \neq 0$, we denote the corresponding unit vector as $\hat{z} := z/|z|$. We also denote a point on the line segment between y and $\tilde{y}$ as $y_t := y + t(\tilde{y} - y)$, where $t \in [0, 1]$.

To prove that π_L is an injection, we argue by contradiction. Assume that $\tilde{y} \neq y$; the Mean Value Theorem then implies $\exists c \in (0, 1)$ such that

$$\begin{aligned}\frac{d}{dt} \{T(s, r, y_t)\} \Big|_{t=c} &= 0 \\ \Rightarrow \nabla_y T(s, r, y_c) \cdot (\tilde{y} - y) &= 0 \\ \Rightarrow (\widehat{y_c - S} + \widehat{y_c - R}) \cdot (\tilde{y} - y) &= 0 \\ \Rightarrow (\widehat{y_c - S} + \widehat{y_c - R}) \cdot (\widehat{\tilde{y} - y}) &= 0 \\ \Rightarrow (\widehat{y_c - S} + \widehat{y_c - R}) \cdot (\widehat{y_c - S}) &= 0 \\ \Rightarrow (\widehat{y_c - R}) \cdot (\widehat{y_c - S}) &= -1,\end{aligned}$$

where we have used the fact that $(\widehat{\tilde{y} - y}) = (\widehat{y_c - S})$, which in turn is true because $s, y, \tilde{y}, y_c$ all lie on a single ray emanating from S . However, the last equality contradicts Assumption (3.2), and therefore $\tilde{y} = y$.

Hence, C satisfies the Traveltime Injectivity Condition. Taking the clean intersection calculus (9) into account, since $\dim(\mathbb{D}) - \dim(\mathbb{R}_+^3) = 1$ we should write the order of F as $\frac{3}{4} = 1 - \frac{1}{4}$; since the ‘effective’ order of F is 1, the composition F^*F is a pseudodifferential operator of order 2 on $\mathbb{R}_+^3$, which will be elliptic at those points in $T^*\mathbb{R}_+^3$ where F is. Thus, under an illumination assumption, a perturbation δc of the sound speed can be reconstructed without artifacts by filtered backprojection: $F(\delta c) \bmod C^\infty$ determines $\delta c \bmod C^\infty$. $\square$

4. Dense array: No caustics. In this section, we continue the proof of Thm. 3.1, modifying the argument of the previous section to a variable background speed, $c_0 = c_0(x)$, satisfying assumptions which we now describe.

Parametrize the maximally defined characteristic curve (i.e., a ray) departing z in the direction

$$\nu(\theta, \varphi) = (\sin(\varphi) \cos(\theta), \sin(\varphi) \sin(\theta), \cos(\varphi)) \quad (21)$$

by a smooth function

$$\mathbb{R} \supseteq I \ni p \mapsto x(p; \theta, \varphi, z) \in \mathbb{R}_+^3, \quad (22)$$

where $x(0; \theta, \varphi, z) = z$. The angle φ is the polar angle with respect to the x_3 -axis. If the take-off angle corresponds to $\varphi = 0, \pi$, then we change coordinates so φ is the polar angle with respect to another axis and adjust (21) correspondingly. In the following discussion, we proceed as though polar angles φ are the polar angles with respect to the x_3 -axis but none of the arguments depend on this and will work just as well if φ is another polar angle. Following [35], we make the following assumption:

Assumption 4.1 (No Caustics). *Assume $\text{sing supp}(V)$ is contained in a region $\Omega \subset \mathbb{R}_+^3$ which is completely illuminated by each source and receiver with a unique minimal traveltime ray connecting each point $y \in \Omega$ with each $z \in \Sigma_S \cup \Sigma_R$. Also assume that there are no caustic points in Ω on rays issuing from any $z \in \Sigma_S \cup \Sigma_R$.*

Under this assumption, we have a well-defined and smooth traveltime function, $t_{c_0}(z, y) = t_{c_0}(y, z)$, which is the minimal travel time between any $z \in \Sigma_S \cup \Sigma_R$ and $y \in \Omega$. We will also make the standard no-grazing ray assumption:

Assumption 4.2 (No grazing rays). *Whenever $z \in \text{sing supp}(V) \subset \mathbb{R}_+^3$, we assume*

$$\frac{\partial x_3}{\partial p}(p; \theta, \varphi, z) \neq 0, \text{ when } x(p, \theta, \varphi, z) \in \Sigma_S, \quad (23)$$

which means that there are no rays emanating from the subsurface to graze Σ_S .

Remark 2. With this setup, the phase function ϕ from the previous section is replaced, as in [30], by

$$\begin{aligned} \phi(s, r, t, y, \omega) &:= \omega(t - T(s, y, r)), \\ \text{where } T(s, y, r) &:= t_{c_0}((s_1, s_2, 0), y) + t_{c_0}((0, 0, r), y), \end{aligned}$$

the sum of the travel times of the incident and reflected rays, is the total travel time. One easily verifies, using Assumption 3.2, that ϕ is a non-degenerate phase function, so that F is an FIO.

Writing $z = (z_1, z_2, z_3) = (z', z_3)$, the wavefront relation of F is contained in the canonical relation

$$\begin{aligned} C = \Big\{ (s, r, T(s, r, y), \omega\sigma(s, y), \omega\rho(r, y), \omega; y, -\omega[\eta_s(s, y) + \eta_r(r, y)]) \mid \\ (0, 0, r) \in \Sigma_R, (s_1, s_2, 0) \in \Sigma_S, \omega \in \mathbb{R} \setminus 0 \Big\}, \end{aligned} \quad (24)$$

where

$$\sigma(s, y) := \nabla_{z'} t_{c_0}((s_1, s_2, 0), y); \quad \rho(r, y) := \frac{\partial t_{c_0}}{\partial z_3}((0, 0, r), y); \quad (25)$$

$$\eta_s(s, y) := \nabla_y t_{c_0}((s_1, s_2, 0), y); \quad \eta_r(r, y) := \nabla_y t_{c_0}((0, 0, r), y). \quad (26)$$

Since F is an FIO, C' is a 7-dimensional conic Lagrangian submanifold of $T^*\mathbb{R}^7$. We note that the above canonical relation avoids the zero section due to ϕ being a Hormander-type non-degenerate phase function. One can check that Assumptions 4.1, 4.2 imply that we may parametrize C using coordinates $(r, t_{ref}, \check{\varphi}, \hat{\theta}, \hat{\varphi}, \check{\theta}, \omega)$, defined as follows, describing each broken ray backwards: a ray, departing $(0, 0, r)$ in direction $(\sin(\hat{\varphi}) \cos(\hat{\theta}), \sin(\hat{\varphi}) \sin(\hat{\theta}), \cos(\hat{\varphi}))$ and traveling for time $t_{ref} > 0$, arrives at location $y := x(t_{ref}; \hat{\varphi}, \hat{\theta}, r) \in \mathbb{R}_+^3$ (see (22) for a reminder of the definition of the function x here). Then a ray leaving y in the direction

$$(\sin(\check{\varphi}) \cos(\check{\theta}), \sin(\check{\varphi}) \sin(\check{\theta}), \cos(\check{\varphi}))$$

arrives at $(s, 0) \in \Sigma_S$, where $(s, 0) = x(t_{inc}; \check{\varphi}, \check{\vartheta}, y)$; here, the travel time function $t_{inc} := t_{inc}(y, \varphi, \vartheta)$ is the travel time needed for this ray to reach Σ_S . Note that $t_{inc}(y, \varphi, \vartheta)$ is smooth and guaranteed to exist by the implicit function theorem and the non-grazing ray assumption. Finally $t := t_{inc} + t_{ref}$ is the two-way traveltime from $(s, 0)$ to y and from y to $(0, 0, r)$.

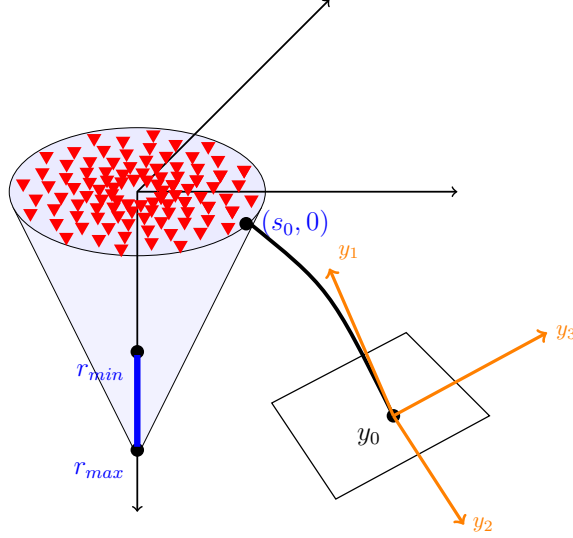


FIGURE 2. Construction of the y -coordinates, with the y_1 -direction being tangent to the ray connecting y to a source $(s, 0) \in \Sigma_S$.

We now verify that C satisfies the Traveltime Injectivity Condition, starting by showing that π_L is an immersion. We check this by showing that

$$\left| \frac{\partial(s, \sigma, T)}{\partial(t_{ref}, \hat{\varphi}, \hat{\theta}, \check{\varphi}, \check{\theta})} \right| \neq 0. \quad (27)$$

Observe that the assumption of no caustics (4.1) implies

$$\left| \frac{\partial y}{\partial(t_{ref}, \hat{\theta}, \hat{\varphi})} \right| \neq 0; \quad (28)$$

using the chain rule, it will follow that π_L is an immersion once we establish

$$\left| \frac{\partial(s, \sigma, T)}{\partial(y, \check{\theta}, \check{\varphi})} \right| \neq 0 \quad (29)$$

as follows. Make a change of y -coordinates so that the y_1 -direction is parallel with the velocity of a specific ray departing $y_0 \in \mathbb{R}_+^3$ and arriving at $(s_0, 0) \in \Sigma_S$. Let (y_2, y_3) be coordinates on the plane that contains y_0 and is also orthogonal to the y_1 -direction, as illustrated in Figure 2. With this choice of coordinates, the no-grazing ray and no caustics assumptions imply that

$$\left| \frac{\partial(s_1, s_2)}{\partial(\check{\varphi}, \check{\theta})} \right| \neq 0 \quad (30)$$

and also, the no-grazing ray assumption guarantees that

$$\left| \frac{\partial(\sigma_1, \sigma_2)}{\partial(y_2, y_3)} \right| \neq 0. \quad (31)$$

Let $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{R}_+^3, \gamma(t) = y(t)$ be a parametrization of an open interval of the ray connecting y_0 to s_0 , for a suitably small $\epsilon \in \mathbb{R}_+$, with $\gamma(0) = y_0$. By construction, the y_1 -direction is tangent to the ray connecting y_0 to $(s_0, 0)$ and since $(s(\gamma(t)), \sigma(\gamma(t))) = (s_0, \sigma_0), \forall t \in (-\epsilon, \epsilon)$, we therefore have

$$\frac{1}{\dot{\gamma}(0)} \frac{d}{dt}(s(\gamma(t)), \sigma(\gamma(t))) = \left(\frac{\partial s}{\partial y_1}(y_0), \frac{\partial \sigma}{\partial y_1}(y_0) \right) = (0, 0). \quad (32)$$

We also have

$$\begin{aligned} \frac{\partial T}{\partial y_1}(s, r, y) &= \frac{\partial t_{c_0}((0, 0, r), y)}{\partial y_1} + \frac{\partial t_{c_0}((s_1, s_2, 0), y)}{\partial y_1} \\ &= \frac{\partial t_{c_0}((0, 0, r), y)}{\partial y_1} - c^{-1}(y), \end{aligned} \quad (33)$$

with the latter equality following from the definition of the y_1 -direction. Additionally,

$$\left| \frac{\partial t_{c_0}((0, 0, r), y)}{\partial y_1} \right| \leq |\nabla t_{c_0}((0, 0, r), y)| = c_0^{-1}(y), \quad (34)$$

where we have used the eikonal equation in the last equality. Furthermore, equality is attained in the left side of (34) if and only if we have scattering over π , which is ruled out by Assumption 3.2. It now follows from (33-34) that

$$\frac{\partial T}{\partial y_1}(s_0, r_0, y_0) \neq 0, \quad (35)$$

for any $r_0 \in (r_{min}, r_{max})$. Therefore, (30-32, 35) establish (29) and so we have shown that π_L is an immersion.

We now verify that π_L is injective. To prove this, it will be convenient to use (s, r, y, ω) as coordinates on C . Suppose that $\pi_L(s, r, y, \omega) = \pi_L(\tilde{s}, \tilde{r}, \tilde{y}, \tilde{\omega})$. Then we immediately have $s = \tilde{s}, r = \tilde{r}, \omega = \tilde{\omega}$ and we deduce that

$$\nabla s t_{c_0}((s, 0), y) = \nabla s t_{c_0}((s, 0), \tilde{y}) =: (\sigma_1, \sigma_2), \quad T(s, r, y) = T(s, r, \tilde{y}). \quad (36)$$

To prove injectivity of π_L , it remains to show that $\tilde{y} = y$. Condition (36) implies y and $\tilde{y}$ lie on a common ray issuing from $(s, 0)$ in the direction

$$\begin{aligned} &(\sin(\varphi_s) \sin(\theta_s), \sin(\varphi_s) \cos(\theta_s), \cos(\varphi_s)) := \\ &\quad \left(-\sigma_1, -\sigma_2, -\sqrt{c_0^{-2}(r, 0) - \sigma_1^2 - \sigma_2^2} \right) \end{aligned} \quad (37)$$

for some angles (φ_s, θ_s) . Let p_1, p_2 be the values of p that satisfy

$$y = x(p_1; \theta_s, \varphi_s, (s, 0)); \quad \tilde{y} = x(p_2; \theta_s, \varphi_s, (s, 0)),$$

which then implies that

$$T(s, r, x(p_1; \theta_s, \varphi_s, (s, 0))) = T(s, r, x(p_2; \theta_s, \varphi_s, (s, 0))). \quad (38)$$

We can now use the same argument used at the end of Sec. 3 to show that, if we assume that $\tilde{y} \neq y$, then (38) contradicts Assumption 3.2. So, under the above assumptions, π_L is injective; combining this with π_L being an immersion, established earlier, we have shown that the Traveltime Injectivity Condition is satisfied. Thus, as in the case of constant c_0 , $F^*F \in \Psi^2(\mathbb{R}_+^3)$, concluding the proof of Thm. 3.1.

5. Dense array: Fold caustics. We start by formulating a notion of what it means for the ray geometry of c_0 to have fold caustics with respect to borehole data acquisition.

5.1. Fold caustics: Single receiver. First recall the concept for conventional seismic data, where both sources and receivers are on the surface [30, 28]. For a single receiver, r , the cotangent space $\Lambda_r := T_r^*\mathbb{R}^3$ is a Lagrangian submanifold of $T^*\mathbb{R}^3$, on which the canonical dual variables $q = (q_1, q_2, q_3)$ are coordinates. The exponential map, $\chi := \exp(H_{c_0}) : T^*\mathbb{R}^3 \rightarrow T^*\mathbb{R}^3$, of the Hamiltonian vector field of the (nonhomogeneous) symbol $\frac{1}{2}(c_0(x))^{-2} - |\xi|^2$ is a (nonhomogeneous) canonical transformation of $T^*\mathbb{R}^3$. Thus, the image $\Lambda_r^{c_0} := \chi(\Lambda_r)$ is also a Lagrangian, which is not conic since χ is not homogeneous in ξ ; on $\Lambda_r^{c_0}$, the pushforwards $\tilde{q} := \chi_*(q)$ by χ are coordinates. On $\Lambda_r^{c_0}$ there is a well-defined acoustical distance function, which is the integral of c_0^{-1} along each bicharacteristic.

A *caustic* of c_0 (with respect to r) is a point $\lambda_0 = (x^0, \xi^0) \in \Lambda_r^{c_0}$ where the spatial projection $\pi_X : \Lambda_r^{c_0} \rightarrow \mathbb{R}^3$ has a noninvertible differential, and λ_0 is a *fold* caustic if π_X has a Whitney fold singularity at λ_0 (see Def. 7.2). At such a point, $d\pi_X(T_{\lambda_0}\Lambda_r^{c_0})$ is a hyperplane $\Pi \subset T_{x^0}\mathbb{R}^3$. For the following, assume that Π is not vertical; otherwise, the discussion needs to be slightly modified. Since $\dim(\Pi) = 2$ and is nonvertical, x_1, x_2 have linearly independent gradients and thus are independent functions on $\Lambda_r^{c_0}$. By Darboux's Theorem, these may be augmented with $p_3 := \xi_3|_{\Lambda_r^{c_0}}$ to obtain a coordinate system on $\Lambda_r^{c_0}$ near λ_0 . On $\Lambda_r^{c_0}$, the restrictions of the other canonical coordinates on $T^*\mathbb{R}^3$ are functions of (x_1, x_2, p_3) : $x_3 = f(x_1, x_2, p_3)$ and $(p_1, p_2) := (\xi_1, \xi_2)|_{\Lambda_r^{c_0}} = (g_1(x_1, x_2, p_3), g_2(x_1, x_2, p_3))$. The fold caustic at λ_0 then implies that

$$\frac{\partial f}{\partial p_3} = 0, \quad \frac{\partial^2 f}{\partial p_3^2} \neq 0. \quad (39)$$

Note also that the acoustical distance function described above is a smooth function of (x_1, x_2, p_3) , since they are coordinates on $\Lambda_r^{c_0}$.

5.2. Fold caustics: Borehole data. Now let $\mathbb{D}$ be the dense array data set as in the previous two sections, for which $\Sigma_S \subset \partial\mathbb{R}_+^3 \setminus \mathbf{0} \simeq \mathbb{R}^2 \setminus (0, 0)$ is an open subset and $\Sigma_R = \{(0, 0, r) : r_{\min} < r < r_{\max}\}$. For each value of r , one can repeat the above constructions and analysis. Since each $\Lambda_r = T_r^*\mathbb{R}^3$ is Lagrangian, it follows that $\Gamma := \bigcup_r \Lambda_r$ is a 4-dimensional coisotropic (or involutive) submanifold of $T^*\mathbb{R}^3$, and Γ is foliated by the family of Λ_r . Thus, with the canonical transformation χ as in Sec. 5.1, the image $\Gamma^{c_0} := \chi(\Gamma)$ is also a (nonconic) four-dimensional coisotropic submanifold of $T^*\mathbb{R}^3$, foliated by $\{\Lambda_r^{c_0} : r_{\min} < r < r_{\max}\}$. At a regular point of the spatial projection, $\pi_X : \Gamma^{c_0} \rightarrow \mathbb{R}^3$, $\text{rank}(d\pi_X) = 3$ is maximal, while a caustic is a γ_0 where $\text{rank}(d\pi_X(\gamma_0)) \leq 2$. We will demand that π_X has at most fold singularities. Due to the difference in dimensions, this means that it is a submersion with folds (see Def. 7.3).

Definition 5.1. We say that the ray geometry of c_0 has *at most fold caustics with respect to the borehole* Σ_R if

- (i) for each $r_{\min} < r < r_{\max}$, the only singularities of the spatial projection $\pi_X : \Lambda_r^{c_0} \rightarrow \mathbb{R}_+^3$ are Whitney folds; and
- (ii) the only singularities of $\pi_X : \Gamma^{c_0} \rightarrow \mathbb{R}_+^3$ are submersions with folds.

Remark 3. One can compare conditions (i) and (ii). These are in fact independent of each other: The receiver-by-receiver Whitney fold condition (i) does not imply

(ii), since the latter requires the invariantly defined Hessian to have rank two, which cannot be derived from the rank one Hessian coming from a Whitney fold. Conversely, (ii) only implies that the $\pi_X : \Lambda_r^{c_0} \rightarrow \mathbb{R}_+^3$ are Whitney folds under a tangent space condition which, while generic, does not appear to be physically required. This is because the restriction of a submersion with folds to a submanifold passing through the critical set is not necessarily a fold; for example, the function $F(x_1, x_2) = x_1 x_2$ is a submersion with folds $\mathbb{R}^2 \rightarrow \mathbb{R}^1$, but restricted to either axis it is not a Whitney fold $\mathbb{R} \rightarrow \mathbb{R}$.

We also mention that this Def. 5.1 is related to but differs from the notion of fold caustic formulated in [11] for another overdetermined data set, the marine data acquisition geometry.

The main result of this section is the following; the terminology used and the consequences for imaging are explained after its statement.

Theorem 5.2. *Under the fold caustic assumption, and a small slope assumption on the caustic surface (see (47)), the linearized forward operator F is a Fourier integral operator, $F \in I^{\frac{3}{4}}(\mathbb{D}, \mathbb{R}_+^3; C)$, whose canonical relation C is a folded crosscap in the sense of Def. 5.3 below, away from a possible set of codimension at least four.*

We now recall the class of degenerate canonical relations and Fourier integral operators referred to in the theorem, which was originally introduced by two of the authors in the context of marine seismic imaging. Suppose that $\dim(X) = n + 1$, $\dim(Y) = n$ and $C \subset (T^*X \setminus 0) \times (T^*Y \setminus 0)$ is a canonical relation, so that

$$\dim(T^*Y) = 2n < \dim(C) = 2n + 1 < \dim(T^*X) = 2n + 2.$$

Definition 5.3. [11] The canonical relation C is a *folded cross cap* if

- (i) $\pi_R : C \rightarrow T^*Y$ is a submersion with folds (see Def. 7.3) and the image of its critical manifold, $\pi_R(\Sigma(\pi_R))$, is a nonradial hypersurface in T^*Y ;
- (ii) $\pi_L : C \rightarrow T^*X$ is a cross cap (see Def. 7.4) and $\pi_L(\Sigma(\pi_L))$, which is a codimension three, immersed submanifold in T^*X , is also nonradial.

Remark 4. Recall that *nonradial* means that the restriction of the canonical 1-form does not vanish anywhere. Also, from [11] one knows that $\pi_L(\Sigma(\pi_L))$ must be maximally noninvolutive, i.e., the restriction to it of the canonical two form has maximal possible rank everywhere, which on the $(2n - 1)$ -dimensional $\pi_L(\Sigma(\pi_L))$ is $2n - 2$.

For a folded cross cap, the composition $C^t \circ C$ lies outside the clean intersection calculus, but the following holds.

Theorem 5.4. [11] *If C is a folded cross cap and $A \in I^{m-\frac{1}{4}}(X, Y; C)$, then $A^*A \in I^{2m-\frac{1}{2}, \frac{1}{2}}(\Delta_{T^*Y}, \tilde{C})$, where $\tilde{C} \subset T^*Y \times T^*Y$ intersects Δ_{T^*Y} cleanly in codimension 1, and $\tilde{C}$ is a folding canonical relation, i.e., both π_L and π_R are Whitney folds.*

It follows that the wavefront relation of $N := A^*A$ is contained in $\Delta_{T^*Y} \cup \tilde{C}$, with $\tilde{C}$ a folding canonical relation. By (4), (5) and that Δ_{T^*Y} acts as the identity relation on T^*Y , for any $u \in \mathcal{E}'(Y)$, we have

$$WF(Nu) \subseteq WF(u) \cup (\tilde{C} \circ WF(u)). \quad (40)$$

Furthermore, by (6), microlocally away from $\Delta_{T^*Y} \cap \tilde{C}$, we have $N \in I^{2m}(\Delta_{T^*Y} \setminus \tilde{C})$, and $N \in I^{2m-\frac{1}{2}}(\tilde{C} \setminus \Delta_{T^*Y})$; thus, away from $\Delta_{T^*Y} \cap \tilde{C}$, the order of the non-pseudodifferential operator part of N , which constitutes an artifact, is $1/2$ lower order than the pseudodifferential part of N . Although the artifact's order is $1/2$ lower, the two-sided fold degeneracy of $\tilde{C}$, combined with the paired Lagrangian nature of the normal operator, produce a situation where it is not known whether the artifact is completely removable; see [13] for further analysis and discussion.

Under the assumptions of Theorem 5.2 and away from a very small microlocal set, this composition result and its implications apply to the dense array borehole data set, resulting in artifacts $1/2$ order smoother than the primary image.

5.3. Proof of Thm. 5.2. The analysis here has a great deal of similarity with that for the marine surface data set in the presence of fold caustics [11]; however, the dense array data set differs in fundamental ways from the marine data set (e.g., for each source in the dense array, the line of receivers is the same, while that is not the case for the marine geometry), and can not be reduced to that case.

Let $r_{min} < r_0 < r_{max}$ and $\gamma_0 \in \chi(\Lambda_{r_0}) \subset \Gamma^{c_0}$ be a Whitney fold point for $\pi_X : \Lambda_{r_0} \rightarrow \mathbb{R}_+^3$. Repeating the analysis from Sec. 5.1, we can assume that $x_1, x_2, p_3 := \xi_3|_{\Gamma_R^{c_0}}$ have independent gradients near γ_0 . Since ∂_r is transverse to $T\Lambda_r$, $d\chi(\partial_r)$ is transverse to $T_{\gamma_0}\chi(\Lambda_r)$. Thus, (x_1, x_2, r, p_3) form coordinates on Γ^{c_0} near γ_0 , the acoustical distance function described above is a smooth function of (x_1, x_2, r, p_3) , and we can express x_3 and $(p_1, p_2) := (\xi_1, \xi_2)$ on Γ^{c_0} in terms of them: $x_3 = f(x_1, x_2, r, p_3)$ and $(p_1, p_2) = (g_1(x_1, x_2, r, p_3), g_2(x_1, x_2, r, p_3))$ on Γ^{c_0} . With respect to these coordinates, $\pi_X(x_1, x_2, r, p_3) = (x_1, x_2, f(x_1, x_2, r, p_3))$ and

$$d\pi_X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial r} & \frac{\partial f}{\partial p_3} \end{pmatrix}.$$

From this we see that

$$\text{rank}(d\pi_X) = \begin{cases} 2, & \text{if } \frac{\partial f}{\partial r} = \frac{\partial f}{\partial p_3} = 0 \\ 3, & \text{if } \frac{\partial f}{\partial r} \neq 0 \text{ or } \frac{\partial f}{\partial p_3} \neq 0, \end{cases}$$

and $\Sigma(\pi_X) = \{\frac{\partial f}{\partial p_3} = \frac{\partial f}{\partial r} = 0\}$. At points of $\Sigma(\pi_X)$, $\ker(d\pi_X)$ is spanned by $\{(0, 0, \delta r, \delta p_3) : \delta r, \delta p_3 \in \mathbb{R}\}$, and the tangent space to $\Sigma(\pi_X)$ is

$$T\Sigma(\pi_X) = \ker\left(d_{x_1, x_2, p_3, r}\left(\frac{\partial f}{\partial p_3}\right)\right) \cap \ker\left(d_{x_1, x_2, p_3, r}\left(\frac{\partial f}{\partial r}\right)\right).$$

We have:

$$d_{x_1, x_2, p_3, r}\left(\frac{\partial f}{\partial p_3}\right) = \left(\frac{\partial^2 f}{\partial x_1 \partial p_3}, \frac{\partial^2 f}{\partial x_2 \partial p_3}, \frac{\partial^2 f}{\partial p_3^2}, \frac{\partial^2 f}{\partial r \partial p_3}\right)$$

and

$$d_{x_1, x_2, p_3, r}\left(\frac{\partial f}{\partial r}\right) = \left(\frac{\partial^2 f}{\partial x_1 \partial r}, \frac{\partial^2 f}{\partial x_2 \partial r}, \frac{\partial^2 f}{\partial p_3 \partial r}, \frac{\partial^2 f}{\partial r^2}\right).$$

The assumption Def. 5.1(ii) that $\pi_X : \Gamma^{c_0} \rightarrow \mathbb{R}_+^3$ is a submersion with folds implies that $\Sigma(\pi_X)$ is smooth (i.e., these gradients are linearly independent), and

$$T\Sigma(\pi_X) \text{ is transverse to } \ker(d\pi_X). \quad (41)$$

We can parametrize the canonical relation C in terms of $r, x_1, x_2, p_3; (\alpha_1, \alpha_2)$, where $(\alpha_1, \alpha_2, \sqrt{1 - |\alpha|^2})$ is the unit take off direction of the reflected ray; and τ , the variable dual to time. The incident ray travel time t_{inc} , the time it takes for a ray to travel from a source $S = (s, 0)$ to an incident point x , i.e., the acoustical distance from S to x , can, by the nongrazing assumption at the surface described in Sec. 2.1 and symmetry, be expressed in terms of x and α , $t_{inc} = t_{inc}(x, \alpha)$. On the other hand, the reflected ray travel time t_{ref} , the time it then takes for the reflected ray to reach the borehole at point $R = (0, 0, r)$, is, again by symmetry, the acoustical distance from R to x , and thus $t_{ref} = t_{ref}(x_1, x_2, r, p_3)$; the total time for the single-reflection event is $t = t_{inc} + t_{ref}$. Letting $(\rho, \sigma_1, \sigma_2, \tau)$ be the coordinates dual to r, s_1, s_2, t in $T^*\mathbb{D}$, we can take $(x_1, x_2, r, p_3, \alpha_1, \alpha_2, \tau)$ as coordinates on C , and

$$\begin{aligned} C = \bigg\{ & (r, s_1(x_1, x_2, f(x_1, x_2, r, p_3), \alpha), s_2(x_1, x_2, f(x_1, x_2, r, p_3), \alpha), \\ & t_{ref}(x_1, x_2, r, p_3) + t_{inc}(x_1, x_2, f(x_1, x_2, r, p_3), \alpha), \\ & \rho(x_1, x_2, f(x_1, x_2, r, p_3), p_3, \tau), \sigma_1(x_1, x_2, f(x_1, x_2, r, p_3), \alpha, \tau), \\ & \sigma_2(x_1, x_2, f(x_1, x_2, r, p_3), \alpha, \tau), \tau; \\ & x_1, x_2, f(x_1, x_2, r, p_3); -\tau (c_0^{-1}(x_1, x_2, f(x_1, x_2, r, p_3))\alpha_1 + g_1(x_1, x_2, r, p_3)), \\ & -\tau (c_0^{-1}(x_1, x_2, f(x_1, x_2, r, p_3))\alpha_2 + g_2(x_1, x_2, r, p_3)), \\ & -\tau (c_0^{-1}(x_1, x_2, f(x_1, x_2, r, p_3)) \sqrt{1 - |\alpha|^2} + p_3) \bigg\}, \end{aligned}$$

where $\rho(\cdot)$ is homogeneous of degree 1 in p_3, τ , and $\sigma_1(\cdot), \sigma_2(\cdot)$ are homogeneous of degree 1 in τ .

We now show that C is a folded cross cap in the sense of Def. 5.3, except possibly on a set of codimension four. In terms of the above coordinates on C and the standard canonical coordinates (x, ξ) on $T^*\mathbb{R}^3$, we can write π_R as a map $\pi_R : \mathbb{R}^7 \rightarrow \mathbb{R}^6$, given by

$$\begin{aligned} \pi_R(x_1, x_2, r, p_3, \alpha_1, \alpha_2, \tau) = & (x_1, x_2, f(x_1, x_2, r, p_3), \\ & -\tau (c_0^{-1}(x_1, x_2, f(x_1, x_2, r, p_3))\alpha_1 + g_1(x_1, x_2, r, p_3)), \\ & -\tau (c_0^{-1}(x_1, x_2, f(x_1, x_2, r, p_3))\alpha_2 + g_2(x_1, x_2, r, p_3)), \\ & -\tau (c_0^{-1}(x_1, x_2, f(x_1, x_2, r, p_3)) \sqrt{1 - |\alpha|^2} + p_3)). \end{aligned}$$

Thus,

$$d\pi_R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial r} & \frac{\partial f}{\partial p_3} & 0 & 0 & 0 \\ A_1 & A_2 & A_3 & A_4 & -\tau c_0^{-1} & 0 & -(c_0^{-1}\alpha_1 + g_1) \\ B_1 & B_2 & B_3 & B_4 & 0 & -\tau c_0^{-1} & -(c_0^{-1}\alpha_2 + g_2) \\ C_1 & C_2 & C_3 & C_4 & \tau c_0^{-1} \frac{\alpha_1}{\sqrt{1 - |\alpha|^2}} & \tau c_0^{-1} \frac{\alpha_2}{\sqrt{1 - |\alpha|^2}} & -(c_0^{-1} \sqrt{1 - |\alpha|^2} + p_3) \end{pmatrix} \quad (42)$$

for some A_j, B_j, C_j . The lower right 3×3 submatrix

$$\begin{pmatrix} -\tau c_0^{-1} & 0 & -(c_0^{-1}\alpha_1 + g_1) \\ 0 & -\tau c_0^{-1} & -(c_0^{-1}\alpha_2 + g_2) \\ \tau c_0^{-1} \frac{\alpha_1}{\sqrt{1-|\alpha|^2}} & \tau c_0^{-1} \frac{\alpha_2}{\sqrt{1-|\alpha|^2}} & -(c_0^{-1}\sqrt{1-|\alpha|^2} + p_3) \end{pmatrix}$$

is nonsingular, since its determinant

$$-\tau^2 c_0^{-2} (1 - |\alpha|)^{-\frac{1}{2}} \left[c_0^{-1} + \alpha_1 g_1 + \alpha_2 g_2 + p_3 (1 - |\alpha|)^{\frac{1}{2}} \right]$$

is nonzero: $(g_1, g_2, p_3) = \xi|_{\Gamma^{c_0}}$ and by Assumption 3.2, any scattering over π has been filtered out, i.e.,

$$(\alpha_1, \alpha_2, (1 - |\alpha|)^{\frac{1}{2}}) \cdot (p_1, p_2, p_3) \neq -c_0^{-1}.$$

It follows that

$$\text{rank}(d\pi_R) = \begin{cases} 5, & \text{if } \frac{\partial f}{\partial p_3} = \frac{\partial f}{\partial r} = 0, \\ 6, & \text{if } \frac{\partial f}{\partial p_3} \neq 0 \text{ or } \frac{\partial f}{\partial r} \neq 0. \end{cases} \quad (43)$$

Now, $\ker d\pi_R = \{(0, 0, \delta r, \delta p_3, \delta \alpha_1, \delta \alpha_2, \delta \tau)\}$, where $\delta \alpha_1, \delta \alpha_2, \delta \tau$ depend on $\delta p_3, \delta r$. On the other hand, the tangent space to $\Sigma(\pi_R)$ is

$$T\Sigma(\pi_R) = \ker \left(d_{x_1, x_2, r, p_3, \alpha_1, \alpha_2, \tau} \left(\frac{\partial f}{\partial p_3} \right) \right) \cap \ker \left(d_{x_1, x_2, r, p_3, \alpha_1, \alpha_2, \tau} \left(\frac{\partial f}{\partial r} \right) \right),$$

where

$$d_{x_1, x_2, r, p_3, \alpha_1, \alpha_2, \tau} \left(\frac{\partial f}{\partial p_3} \right) = \left(\frac{\partial^2 f}{\partial x_1 \partial p_3}, \frac{\partial^2 f}{\partial x_2 \partial p_3}, \frac{\partial^2 f}{\partial r \partial p_3}, \frac{\partial^2 f}{\partial p_3^2}, 0, 0, 0 \right)$$

and

$$d_{x_1, x_2, r, p_3, \alpha_1, \alpha_2, \tau} \left(\frac{\partial f}{\partial r} \right) = \left(\frac{\partial^2 f}{\partial x_1 \partial r}, \frac{\partial^2 f}{\partial x_2 \partial r}, \frac{\partial^2 f}{\partial r^2}, \frac{\partial^2 f}{\partial p_3 \partial r}, 0, 0, 0 \right).$$

Combining this with (41), one sees that $\ker(d\pi_R)$ is transverse to $T\Sigma(\pi_R)$ and thus, π_R is a submersion with folds. One can also check that, off an exceptional set, the image of the critical set is nonradial, i.e., $\xi \cdot dx \neq 0$ on $\Sigma(\pi_R)$. Since $T(\Sigma(\pi_R))$ is the span of the columns of the matrix in (42) representing $d\pi_R$, while $\tau^{-1}\xi$ consists of the last three entries in its last column, and $\frac{\partial f}{\partial p_3} = \frac{\partial f}{\partial r} = 0$ at $\Sigma(\pi_R)$, we see that $(\xi \cdot dx)(W) = 0$ for all $W \in T(\Sigma(\pi_R))$ if and only if

$$c_0^{-1}\alpha_j + g_j + c_0^{-1}f_{x_j}\sqrt{1-|\alpha|^2} + f_{x_j}p_3 = 0, \quad j = 1, 2,$$

which defines a codimension 2 submanifold in $\Sigma(\pi_R)$ and hence codimension 4 in C .

We next need to show that π_L is a cross cap. As for any canonical relation, $\Sigma(\pi_L) = \Sigma(\pi_R)$. Similar to the analysis for π_R , the projection $\pi_L : C \rightarrow T^*\mathbb{D}$ can be treated as mapping $\mathbb{R}^7 \rightarrow \mathbb{R}^8$, with (reordering the variables for convenience)

$$\begin{aligned} \pi_L(r, x_1, x_2, \alpha_1, \alpha_2, p_3, \tau) &= (r, s_1(x_1, x_2, f(x_1, x_2, r, p_3), \alpha), s_2(x_1, x_2, f(x_1, x_2, r, p_3), \alpha), \\ &\quad t_{ref}(x_1, x_2, r, p_3) + t_{inc}(x_1, x_2, f(x_1, x_2, r, p_3), \alpha), \\ &\quad \rho(x_1, x_2, f(x_1, x_2, r, p_3), p_3, \tau), \\ &\quad \sigma_1(x_1, x_2, f(x_1, x_2, r, p_3), \alpha, \tau), \\ &\quad \sigma_2(x_1, x_2, f(x_1, x_2, r, p_3), \alpha, \tau), \tau) \end{aligned}$$

and thus $d\pi_L =$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial s_1}{\partial x_3} \frac{\partial f}{\partial r} & \frac{\partial s_1}{\partial x_1} + \frac{\partial s_1}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial s_1}{\partial x_2} + \frac{\partial s_1}{\partial x_3} \frac{\partial f}{\partial x_2} & \frac{\partial s_1}{\partial \alpha_1} & \frac{\partial s_1}{\partial \alpha_2} & \frac{\partial s_1}{\partial x_3} \frac{\partial f}{\partial p_3} & 0 \\ \frac{\partial s_2}{\partial x_3} \frac{\partial f}{\partial r} & \frac{\partial s_2}{\partial x_1} + \frac{\partial s_2}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial s_2}{\partial x_2} + \frac{\partial s_2}{\partial x_3} \frac{\partial f}{\partial x_2} & \frac{\partial s_2}{\partial \alpha_1} & \frac{\partial s_2}{\partial \alpha_2} & \frac{\partial s_2}{\partial x_3} \frac{\partial f}{\partial p_3} & 0 \\ \frac{\partial t_{ref}}{\partial r} + \frac{\partial t_{inc}}{\partial x_3} \frac{\partial f}{\partial r} & \frac{\partial t_{ref}}{\partial x_1} + \frac{\partial t_{inc}}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial t_{ref}}{\partial x_2} + \frac{\partial t_{inc}}{\partial x_3} \frac{\partial f}{\partial x_2} & \frac{\partial t_{inc}}{\partial \alpha_1} & \frac{\partial t_{inc}}{\partial \alpha_2} & \frac{\partial t_{ref}}{\partial p_3} + \frac{\partial t_{inc}}{\partial x_3} \frac{\partial f}{\partial p_3} & 0 \\ \frac{\partial \rho}{\partial r} & \frac{\partial \rho}{\partial x_1} + \frac{\partial \rho}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial \rho}{\partial x_2} + \frac{\partial \rho}{\partial x_3} \frac{\partial f}{\partial x_2} & 0 & 0 & \frac{\partial \rho}{\partial x_3} \frac{\partial f}{\partial p_3} + \frac{\partial \rho}{\partial p_3} & \frac{\partial \rho}{\partial \tau} \\ \frac{\partial \sigma_1}{\partial x_3} \frac{\partial f}{\partial r} & \frac{\partial \sigma_1}{\partial x_1} + \frac{\partial \sigma_1}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial \sigma_1}{\partial x_2} + \frac{\partial \sigma_1}{\partial x_3} \frac{\partial f}{\partial x_2} & \frac{\partial \sigma_1}{\partial \alpha_1} & \frac{\partial \sigma_1}{\partial \alpha_2} & \frac{\partial \sigma_1}{\partial x_3} \frac{\partial f}{\partial p_3} & \frac{\partial \sigma_1}{\partial \tau} \\ \frac{\partial \sigma_2}{\partial x_3} \frac{\partial f}{\partial r} & \frac{\partial \sigma_2}{\partial x_1} + \frac{\partial \sigma_2}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial \sigma_2}{\partial x_2} + \frac{\partial \sigma_2}{\partial x_3} \frac{\partial f}{\partial x_2} & \frac{\partial \sigma_2}{\partial \alpha_1} & \frac{\partial \sigma_2}{\partial \alpha_2} & \frac{\partial \sigma_2}{\partial x_3} \frac{\partial f}{\partial p_3} & \frac{\partial \sigma_2}{\partial \tau} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Since $\text{corank}(d\pi_L) = \text{corank}(d\pi_R)$, it follows from (43) that $d\pi_L$ is injective except where it has a one-dimensional kernel above caustic points, where $\frac{\partial f}{\partial p_3} = \frac{\partial f}{\partial r} = 0$. We will use these two conditions, plus one more, in order to simplify this matrix. Namely, by rotation about the borehole, we can assume that, for the point of interest, the tangent plane $\Pi = \pi_X(T_{\lambda_0}\Lambda_r^{\text{c}_0})$ from Sec. 5.1 is the graph of x_3 as a linear function independent of x_2 . As a consequence, $f_{x_2} = 0$ at this point. The matrix for $d\pi_L$ then becomes

$$d\pi_L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\partial s_1}{\partial x_1} + \frac{\partial s_1}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial s_1}{\partial x_2} & \frac{\partial s_1}{\partial \alpha_1} & \frac{\partial s_1}{\partial \alpha_2} & 0 & 0 \\ 0 & \frac{\partial s_2}{\partial x_1} + \frac{\partial s_2}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial s_2}{\partial x_2} & \frac{\partial s_2}{\partial \alpha_1} & \frac{\partial s_2}{\partial \alpha_2} & 0 & 0 \\ \frac{\partial t_{ref}}{\partial r} & \frac{\partial t_{ref}}{\partial x_1} + \frac{\partial t_{inc}}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial t_{ref}}{\partial x_2} & \frac{\partial t_{inc}}{\partial \alpha_1} & \frac{\partial t_{inc}}{\partial \alpha_2} & \frac{\partial t_{ref}}{\partial p_3} & 0 \\ \frac{\partial \rho}{\partial r} & \frac{\partial \rho}{\partial x_1} + \frac{\partial \rho}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial \rho}{\partial x_2} & 0 & 0 & \frac{\partial \rho}{\partial p_3} & \frac{\partial \rho}{\partial \tau} \\ 0 & \frac{\partial \sigma_1}{\partial x_1} + \frac{\partial \sigma_1}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial \sigma_1}{\partial x_2} & \frac{\partial \sigma_1}{\partial \alpha_1} & \frac{\partial \sigma_1}{\partial \alpha_2} & 0 & \frac{\partial \sigma_1}{\partial \tau} \\ 0 & \frac{\partial \sigma_2}{\partial x_1} + \frac{\partial \sigma_2}{\partial x_3} \frac{\partial f}{\partial x_1} & \frac{\partial \sigma_2}{\partial x_2} & \frac{\partial \sigma_2}{\partial \alpha_1} & \frac{\partial \sigma_2}{\partial \alpha_2} & 0 & \frac{\partial \sigma_2}{\partial \tau} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (44)$$

Writing a spanning element $V_L \in \ker(d\pi_L)$ as $(\delta r, \delta x_1, \delta x_2, \delta \alpha_1, \delta \alpha_2, \delta p_3, \delta \tau)$, we see from (44) that $\delta r = \delta \tau = 0$, so that the $\frac{\partial \sigma_1}{\partial \tau}, \frac{\partial \sigma_2}{\partial \tau}$ terms in (44) can be ignored.

Furthermore, under the assumption that the rays from the sources are transverse to the caustic surface (see [28]), the determinant

$$\begin{vmatrix} \frac{\partial s_1}{\partial x_1} & \frac{\partial s_1}{\partial x_2} & \frac{\partial s_1}{\partial \alpha_1} & \frac{\partial s_1}{\partial \alpha_2} \\ \frac{\partial s_2}{\partial x_1} & \frac{\partial s_2}{\partial x_2} & \frac{\partial s_2}{\partial \alpha_1} & \frac{\partial s_2}{\partial \alpha_2} \\ \frac{\partial \sigma_1}{\partial x_1} & \frac{\partial \sigma_1}{\partial x_2} & \frac{\partial \sigma_1}{\partial \alpha_1} & \frac{\partial \sigma_1}{\partial \alpha_2} \\ \frac{\partial \sigma_2}{\partial x_1} & \frac{\partial \sigma_2}{\partial x_2} & \frac{\partial \sigma_2}{\partial \alpha_1} & \frac{\partial \sigma_2}{\partial \alpha_2} \end{vmatrix} \neq 0. \quad (45)$$

The matrix in (45) is almost a minor of (44): they differ only in the first column, by

$$f_{x_1} \cdot [(s_1)_{x_3}, (s_2)_{x_3}, (\sigma_1)_{x_3}, (\sigma_2)_{x_3}]^T. \quad (46)$$

Thus, if we make the small slope assumption that

$$|f_{x_1}| \quad \text{is sufficiently small,} \quad (47)$$

which, since the x_1 direction was chosen for convenience, can only be ensured if we assume $|d_{x_1, x_2} f|$ is small, i.e., the normal to the fold caustic surface is sufficiently close to vertical, then the corresponding minor of (44) is nonsingular, which implies that $\delta x_1 = \delta x_2 = \delta \alpha_1 = \delta \alpha_2 = 0$. (A “small dip” assumption such as (47) could be conceivably verified from a sufficiently accurate background velocity model, but is included here in order for the mathematical analysis to be valid.) Thus, δp_3 is the only nonzero entry in V_L ; note also that this then implies that $\frac{\partial t_{ref}}{\partial p_3} = \frac{\partial \rho}{\partial p_3} = 0$ at caustics. It then follows from (39) that (iii) below Def. 7.4 is satisfied, and (i), (ii) follow from this analysis as well. Hence, π_L is a cross cap. Again, one can

check that, away from a set of high codimension, the image of the cross cap points is nonradial in $T^*\mathbb{D}$. A point is radial if and only if $(\rho, \sigma_1, \sigma_2, \tau)^T \cdot W = 0$ for all $W \in T(\pi_L(\Sigma(\pi_L)))$, i.e., for W in the span of the columns of the upper 4×7 submatrix of $d\pi_L$, this becomes

$$\begin{aligned} \rho &= -(t_{ref})_r \tau \\ ((s_1)_{x_1} + \epsilon_1)\sigma_1 + (s_2)_{x_1}\sigma_2 &= -(t_{ref})_{x_1} \tau \\ ((s_1)_{x_2} + \epsilon_2)\sigma_1 + (s_2)_{x_2}\sigma_2 &= -(t_{ref})_{x_2} \tau \\ ((s_1)_{\alpha_1} + \epsilon_3)\sigma_1 + (s_2)_{\alpha_1}\sigma_2 &= -(t_{inc})_{\alpha_1} \tau \\ ((s_1)_{\alpha_2} + \epsilon_4)\sigma_1 + (s_2)_{\alpha_2}\sigma_2 &= -(t_{inc})_{\alpha_2} \tau, \end{aligned}$$

for some small ϵ_j , $1 \leq j \leq 4$. The first equation imposes one condition. On the other hand, the last four equations impose two more, since the coefficient matrix is the upper 2×4 submatrix of the matrix in (45) and thus has rank two, meaning that the right hand sides of these last four equations must satisfy two linear conditions in order for the equations to be solvable. Thus, the set of possibly radial points of $\pi_L(\Sigma(\pi_L))$ is of codimension at least three in the critical set, and thus codimension 5 in C . Combined with the codimension 4 set of possible radial points of π_R , we see that the nonradiality conditions of Def. 5.3 are satisfied away from a set of codimension at least 4 in C .

In summary, we have shown that if the ray geometry of the background sound speed c_0 has at most fold caustics with respect to the borehole, and the small slope assumption (47) holds, then away from a codimension 4 set the canonical relation C is a folded cross cap, finishing the proof of Thm. 5.2. $\square$

Theorem 5.2 then implies that Theorem 5.4 applies to the composition forming the normal operator F^*F (away from the possible bad set microlocally), with the consequences for artifacts as described above.

6. Crosswell and walkaway geometries. As for the dense array in Sec. 3, for the crosswell and walkaway geometries we compute $F = d\mathcal{F}$ at the constant background sound speed $c_0 = 1$ by restricting the basic phase function (10) to each data set $\mathbb{D}$.

6.1. Crosswell geometry. For the crosswell (CW) geometry, we assume that the sources and receivers are located in parallel, vertical boreholes. For simplicity, assume that the sources form an open interval along the line $y_1 = s_0$, $y_2 = 0$, for some $s_0 > 0$,

$$\Sigma_S = \{(s_0, 0, s) : s \in (s_{min}, s_{max}) =: I_S\},$$

and the receivers are similarly located, as for the other geometries, in the borehole along the y_3 axis, say

$$\Sigma_R = \{(0, 0, r) : r \in (r_{min}, r_{max}) =: I_R\},$$

and we identify $\mathbb{D} = \mathbb{D}_{CW} = (s_{min}, s_{max}) \times (r_{min}, r_{max}) \times (t_{min}, t_{max})$.

The associated linearized scattering operator F is then a Fourier integral operator with phase function obtained by restricting (10) to $\mathbb{D}_{CW}$:

$$\phi_{CW}(s, r, t, y; \omega) = \left(t - \sqrt{(y_1 - s_0)^2 + y_2^2 + (y_3 - s)^2} - \sqrt{y_1^2 + y_2^2 + (y_3 - r)^2} \right) \omega.$$

The structure of the linearized scattering operator F for the crosswell geometry is summarized by the following.

Theorem 6.1. *The linearized scattering operator F for the crosswell imaging geometry is a Fourier integral operator, $F \in I^{-\frac{1}{2}}(C_{CW})$, whose canonical relation C_{CW} is singular on the union of two hypersurfaces, $\Sigma^1 \cup \Sigma^2$, with Σ^1 and Σ^2 intersecting transversally. On $\Sigma^1 \setminus \Sigma^2$, π_L has a fold singularity and π_R is a blowdown, while on $\Sigma^2 \setminus \Sigma^1$, both of the projections π_L and π_R have fold singularities.*

Proof of Thm. 6.1. Let

$$A := \sqrt{(y_1 - s_0)^2 + y_2^2 + (y_3 - s)^2}, \quad B := \sqrt{y_1^2 + y_2^2 + (y_3 - r)^2}.$$

We calculate the canonical relation, C_{CW} , parametrized by ϕ_{CW} , and classify the singularities of the left and right projections. We have:

$$\begin{aligned} C_{CW} = \left\{ \left(s, r, A + B, \frac{y_3 - s}{A} \omega, \frac{y_3 - r}{B} \omega, \omega; \right. \right. \\ \left. y_1, y_2, y_3, \left(\frac{y_1 - s_0}{A} + \frac{y_1}{B} \right) \omega, \left(\frac{y_2}{A} + \frac{y_2}{B} \right) \omega, \left(\frac{y_3 - s}{A} + \frac{y_3 - r}{B} \right) \omega \right) \\ \left. : y \in \mathbb{R}^3, s \in I_S, r \in I_R, \omega \neq 0 \right\}. \end{aligned}$$

With respect to these coordinates, the left projection, $\pi_L : C_{CW} \rightarrow T^*\mathbb{D}_{CW}$, is

$$\pi_L(y, s, r, \omega) = \left(s, r, \omega, A + B, \frac{y_3 - s}{A} \omega, \frac{y_3 - r}{B} \omega \right)$$

and the right projection, $\pi_R : C_{CW} \rightarrow T^*\mathbb{R}^3$, is

$$\pi_R(y, s, r, \omega) = \left(y_1, y_2, y_3, \left(\frac{y_1 - s_0}{A} + \frac{y_1}{B} \right) \omega, \left(\frac{y_2}{A} + \frac{y_2}{B} \right) \omega, \left(\frac{y_3 - s}{A} + \frac{y_3 - r}{B} \right) \omega \right).$$

We first study π_L . Denote the variables dual to s, r, t by σ, ρ, τ , resp. Since π_L is the identity in the s, r, ω variables, $\det d\pi_L$ equals $\det D(\sigma, \rho, \tau)/Dy$, i.e.,

$$\begin{aligned} \det(d\pi_L) &= \begin{vmatrix} \frac{y_1 - s_0}{A} + \frac{y_1}{B} & \frac{y_2}{A} + \frac{y_2}{B} & \frac{y_3 - s}{A} + \frac{y_3 - r}{B} \\ -\frac{(y_3 - s)(y_1 - s_0)}{A^3} \omega & -\frac{(y_3 - s)y_2}{A^3} \omega & \frac{(y_1 - s_0)^2 + y_2^2}{A^3} \omega \\ -\frac{(y_3 - r)y_1}{B^3} \omega & -\frac{(y_3 - r)y_2}{B^3} \omega & \frac{y_1^2 + y_2^2}{B^3} \omega \end{vmatrix} \\ &= -\frac{\omega^2}{A^3 B^3} s_0 y_2 \left(\frac{y_3 - s}{A} + \frac{y_3 - r}{B} \right). \end{aligned}$$

Thus, $\det(d\pi_L) = 0$ on $\Sigma^1 \cup \Sigma^2$, where

$$\Sigma^1 := \{y_2 = 0\}, \quad \Sigma^2 := \left\{ \frac{y_3 - s}{A} + \frac{y_3 - r}{B} = 0 \right\}.$$

Note that points in $\Sigma^1 \cap \Sigma^2$ correspond to unbroken rays from Σ_S to Σ_R , not undergoing any scattering, and thus are first arrival events. One can thus filter the data away from $\Sigma^1 \cap \Sigma^2$ by multiplying $d(s, r, t)$ by a smooth cutoff $\chi(c_0 t - |S(s) - R(r)|)$, where $\text{supp}(\chi) \subseteq \{t \geq \epsilon\}$ for some $\epsilon > 0$. Hence, we do not need to consider the more singular structure of C_{CW} at $\Sigma^1 \cap \Sigma^2$.

Along each of $\Sigma^1 \setminus \Sigma^2$ and $\Sigma^2 \setminus \Sigma^1$, $\det(d\pi_L)$ vanishes simply, and thus $d\pi_L$ drops rank by 1. One easily sees that, along $\Sigma^1 \setminus \Sigma^2$, $\ker(d\pi_L) = \frac{\partial}{\partial y_2}$ and hence (cf. Def. 7.2) π_L has a fold singularity at points of $\Sigma^1 \setminus \Sigma^2$. Similarly, $\ker(d\pi_L) = \frac{\partial}{\partial y_2}$ at points of $\Sigma^2 \setminus \Sigma^1$, and hence π_L has a fold singularity there as well.

Next, we consider π_R . As for any canonical relation, $d\pi_R$ also drops rank by the same amount as $d\pi_L$, and hence by 1 on $(\Sigma^1 \setminus \Sigma^2) \cup (\Sigma^2 \setminus \Sigma^1)$. We find its kernel by computing

$$\frac{D\eta}{D(s, r, \omega)} = \begin{bmatrix} \frac{(y_1-s_0)(y_3-s)}{A^3} \omega & \frac{y_1(y_3-r)}{B^3} \omega & \frac{y_1-s_0}{A} + \frac{y_1}{B} \\ \frac{y_2(y_3-s)}{A^3} \omega & \frac{y_2(y_3-r)}{B^3} \omega & \frac{y_2}{A} + \frac{y_2}{B} \\ -\frac{(y_1-s_0)^2+y_2^2}{A^3} \omega & -\frac{y_1^2+y_2^2}{B^3} \omega & \frac{y_3-s}{A} + \frac{y_3-r}{B} \end{bmatrix}. \quad (48)$$

The kernel of $d\pi_R$ is contained in $\text{span} \left\{ \frac{\partial}{\partial s}, \frac{\partial}{\partial r}, \frac{\partial}{\partial \omega} \right\}$, which when applied to the defining function y_2 of Σ^1 gives 0. Hence, along $(\Sigma^1 \setminus \Sigma^2)$, $\ker(d\pi_R) \subset T\Sigma^1$ and thus (cf. Def. 7.1) π_R has a blowdown singularity along $\Sigma^1 \setminus \Sigma^2$. On the other hand, along $\Sigma^2 \setminus \Sigma^1$, the first 2 entries of the last row are nonzero while the last one is 0. Hence the kernel of $d\pi_R$ is spanned by $\frac{\partial}{\partial s}$ or $\frac{\partial}{\partial r}$, which is transverse to Σ^2 , and so π_R has a fold singularity. $\square$

6.2. Walkaway geometry. For the walkaway geometry, the set of sources is assumed to be an open subset of the y_1 axis,

$$\Sigma_S = \{(s, 0, 0) : s \in I_S = (s_{\min}, s_{\max})\}$$

and the set of receivers is as throughout an open subset of the y_3 axis,

$$\Sigma_R = \{(0, 0, r) : r \in I_R = (r_{\min}, r_{\max})\},$$

so that $\mathbb{D} = \mathbb{D}_{WA} = I_S \times I_R \times I_T$. Restricting (10) to $\mathbb{D}_{WA}$, the phase function of F is

$$\phi_{WA}(s, r, t, y; \omega) = \left(t - \sqrt{(y_1-s)^2 + y_2^2 + y_3^2} - \sqrt{y_1^2 + y_2^2 + (y_3-r)^2} \right) \omega.$$

Let

$$A := \sqrt{(y_1-s)^2 + y_2^2 + y_3^2}; \quad (49)$$

$$B := \sqrt{y_1^2 + y_2^2 + (y_3-r)^2}. \quad (50)$$

The structure of the linearized scattering operator F for the walkaway geometry is summarized by the following.

Theorem 6.2. *The linearized scattering operator F for the walkaway geometry is a Fourier integral operator, $F \in I^{-\frac{1}{2}}(C_{WA})$, whose canonical relation C_{WA} is singular at the union of two smooth hypersurfaces Σ^1 and Σ^2 , which intersect transversally. At $\Sigma^1 \setminus \Sigma^2$, π_L has a fold singularity and π_R has a blowdown singularity, while at $\Sigma^2 \setminus \Sigma^1$, π_L is a fold at all points and π_R is a fold away from a hypersurface.*

Proof of Thm. 6.2. The canonical relation C_{WA} of F is

$$\begin{aligned} C_{WA} = \left\{ \left(s, r, A+B, \frac{y_1-s}{A} \omega, \frac{y_3-r}{B} \omega, \omega; \right. \right. \\ \left. y_1, y_2, y_3, \left(\frac{y_1-s}{A} + \frac{y_1}{B} \right) \omega, \left(\frac{y_2}{A} + \frac{y_2}{B} \right) \omega, \left(\frac{y_3}{A} + \frac{y_3-r}{B} \right) \omega \right) \\ \left. : y \in \mathbb{R}^3, s \in I_S, r \in I_R, \omega \neq 0 \right\}. \end{aligned}$$

The right projection $\pi_R : C_{WA} \rightarrow T^*\mathbb{R}^3$ is

$$\pi_R(y, s, r, \omega) = \left(y_1, y_2, y_3, \left(\frac{y_1 - s}{A} + \frac{y_1}{B} \right) \omega, \left(\frac{y_2}{A} + \frac{y_2}{B} \right) \omega, \left(\frac{y_3}{A} + \frac{y_3 - r}{B} \right) \omega \right). \quad (51)$$

Since π_R is the identity in the y variables, to compute $\det(d\pi_R)$ we only need to compute the Jacobian in the remaining variables s, r, ω , which (in this order) is

$$\frac{D(\eta_1, \eta_2, \eta_3)}{D(s, r, \omega)} = \begin{bmatrix} -\frac{y_2^2 + y_3^2}{A^3} \omega & \frac{y_1(y_3 - r)}{B^3} \omega & \frac{y_1 - s}{A} + \frac{y_1}{B} \\ \frac{y_2(y_1 - s)}{A^3} \omega & \frac{y_2(y_3 - r)}{B^3} \omega & \frac{y_2}{A} + \frac{y_2}{B} \\ \frac{y_3(y_1 - s)}{A^3} \omega & -\frac{y_1 + y_2^2}{B^3} \omega & \frac{y_3}{A} + \frac{y_3 - r}{B} \end{bmatrix}. \quad (52)$$

A calculation yields that

$$\det(d\pi_R) = -\frac{\omega^2 y_2}{A^2 B^2} \left(\frac{y_1(y_1 - s) + y_2^2 + y_3^2}{A} + \frac{y_1^2 + y_2^2 + y_3(y_3 - r)}{B} \right);$$

the expression in the parentheses can be written as $y \cdot \left(\frac{y - S}{A} + \frac{y - R}{B} \right)$. If we let

$$\Sigma^1 := \{f_1 := y_2 = 0\}, \quad \Sigma^2 := \left\{ f_2 := y \cdot \left(\frac{y - S}{A} + \frac{y - R}{B} \right) = 0 \right\}, \quad (53)$$

then Σ^1, Σ^2 intersect transversally. Furthermore, on $(\Sigma^1 \setminus \Sigma^2) \cup (\Sigma^2 \setminus \Sigma^1)$, $\det(d\pi_R)$ vanishes simply, and thus $d\pi_R$ drops rank by 1 there; by general principles concerning canonical relations, the same facts hold for $\det(d\pi_L) = \det(d\pi_R)$ and $\text{rank}(d\pi_L)$, resp.

$\Sigma^1 \setminus \Sigma^2$: From (51) we see that $\ker d\pi_R \subset \text{span} \left\{ \frac{\partial}{\partial s}, \frac{\partial}{\partial r}, \frac{\partial}{\partial \omega} \right\}$, which is contained in $T\Sigma^1$, and is one-dimensional at points of $\Sigma^1 \setminus \Sigma^2$; hence, π_R has a blowdown singularity there. Next consider $\pi_L : C_{WA} \rightarrow T^*\mathbb{D}_{WA}$,

$$\pi_L(y, s, r, \omega) = \left(s, r, A + B, \frac{y_1 - s}{A} \omega, \frac{y_3 - r}{B} \omega, \omega \right). \quad (54)$$

As noted above, $d\pi_L$ drops rank by the same amount as $d\pi_R$ and so also has a one-dimensional kernel along $\Sigma^1 \setminus \Sigma^2$. Since π_L is the identity in the s, r, ω variables, we only need compute the differential in the remaining variables, y ,

$$\frac{D(t, \sigma, \rho)}{D(y_1, y_2, y_3)} = \begin{bmatrix} \frac{(y_1 - s)}{A} + \frac{y_1}{B} & \frac{y_2}{A} + \frac{y_2}{B} & \frac{y_3}{A} + \frac{y_3 - r}{B} \\ \frac{y_2 + y_3^2}{A^3} \omega & -\frac{y_2(y_1 - s)}{A^3} \omega & -\frac{y_3(y_1 - s)}{A^3} \omega \\ -\frac{y_1(y_3 - r)}{B^3} \omega & -\frac{y_2(y_3 - r)}{B^3} \omega & \frac{y_1^2 + y_2^2}{B^3} \omega \end{bmatrix}, \quad (55)$$

and $\ker d\pi_L$ is contained in $\text{span} \left\{ \frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}, \frac{\partial}{\partial y_3} \right\}$. Since the entries in the middle column are multiples of y_2 , which vanishes on Σ^1 , one sees that, on $\Sigma^1 \setminus \Sigma^2$, $\ker(d\pi_L) = \text{span} \left\{ \frac{\partial}{\partial y_2} \right\}$, which is transverse to $\Sigma^1 = \{y_2 = 0\}$. Thus, π_L has a fold singularity along $\Sigma^1 \setminus \Sigma^2$.

$\Sigma^2 \setminus \Sigma^1$: We show that all the singularities of π_L are of fold type, while π_R has fold singularities on the complement of a subset defined by a polynomial equation. At points of $\Sigma^2 \setminus \Sigma^1$, as was the case on $\Sigma^1 \setminus \Sigma^2$, $\ker(d\pi_L) \subset \left\{ \frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}, \frac{\partial}{\partial y_3} \right\}$ and is one-dimensional. If $V_L \neq 0$ spans $\ker(d\pi_L)$, then it is annihilated by all of the rows of (55), in particular the first row, and thus $\langle d_y(A + B), V_L \rangle = 0$.

There is a geometric interpretation of this last fact: for fixed s, r , the family of level surfaces of $A + B$, $E_{s,r,t} := \{y : A + B = t\}$, indexed by $t > \sqrt{s^2 + t^2}$, are ellipsoids with foci at s and r , and outward pointing (nonunit) normal $\nu :=$

$d_y(A+B)$. Then, since $\langle \nu, V_L \rangle = 0$, we see that V_L is tangent to $E_{s,r,t}$; on the other hand, by (53), y (considered as a vector) is also tangent to $E_{s,r,t}$.

Notice that, with $f_2 = y \cdot \nu$ as in (53), $d_y f_2 = \nu + y^t d_y \nu$. One has

$$\langle d_y f_2, V_L \rangle = \langle \nu, V_L \rangle + y^t (d_y \nu) V_L. \quad (56)$$

The first term on the right hand side of (56) is zero and the second one is positive since the ellipsoid $E_{s,r,t}$ has positive curvature (for every $V, V' \in T_y E_{s,r,t}$, $V^t (d_y \nu) V' > 0$) and $y, \nu \in T_y E_{s,r,t}$. Thus π_L has a fold singularity along Σ^2 .

For π_R , $\ker(d\pi_R) \subset \text{span} \left\{ \frac{\partial}{\partial s}, \frac{\partial}{\partial r}, \frac{\partial}{\partial \omega} \right\}$ is one-dimensional, and thus spanned by a $V_R = \delta s \frac{\partial}{\partial s} + \delta r \frac{\partial}{\partial r} + \delta \omega \frac{\partial}{\partial \omega}$. From the matrix (52) representing the essential part of $d\pi_R$, we use the second row to solve for $\delta \omega$ in terms of δs and δr (the value of which will be irrelevant below), and the first row to solve for δs in terms of δr , namely $\delta s = -\delta r \frac{A}{B} \frac{s}{r}$. Thus, $V_R = -\delta r \frac{A}{B} \frac{s}{r} \frac{\partial}{\partial s} + \delta r \frac{\partial}{\partial r} + \delta \omega \frac{\partial}{\partial \omega}$; applying this to f_2 (which is independent of ω), a calculation gives the critical set

$$\Sigma(\pi_R|_{\Sigma^2 \setminus \Sigma^1}) = \{s^2 B^2 (y_2^2 + y_3^2) - r^2 A^2 (y_1^2 + y_2^2) = 0\}. \quad (57)$$

We can see that the polynomial defining function in (57) is nonzero at some points, e.g., by taking s or $r \rightarrow \infty$ and considering the leading coefficient in s or r , resp. Therefore, $\Sigma(\pi_R|_{\Sigma^2 \setminus \Sigma^1})$ is a lower dimensional variety, whose complement in $\Sigma^2 \setminus \Sigma^1$ is dense; on that set, $V_R f_2 \neq 0$ so that π_R has a fold singularity at those points. This finishes the proof of Thm. 6.2. $\square$

6.3. Artifacts for crosswell and walkaway. A canonical relation similar to C_{CW} described in Thm. 6.1, with similar geometry for Σ^1 , Σ^2 , and singularities types of the projections from them, was shown to appear in the context of synthetic aperture radar and analyzed in [2]. The open dense subset of C_{WA} described in Thm. 6.2 has a similar structure. It was shown in [2] that, if A is an FIO of order m associated with such a canonical relation, then

$$A^* A \in I^{2m,0}(\Delta, C_1) + I^{2m,0}(\Delta, C_2) + I^{2m,0}(C_1, C_2), \quad (58)$$

where C_1 is the graph of a canonical involution χ , and C_2 is a two-sided fold. It follows from (6) that the order of $A^* A$ is the same (namely $2m$) on all three of Δ , C_1 and C_2 , away from their intersections, and hence the artifacts created by C_1 and C_2 when attempting imaging by backprojection are as strong as the true image, and thus are nonremovable.

Due to the presumed absence of normal forms for canonical relations with this structure, the results of [2], and thus (58), cannot be applied directly to the linearized scattering map $F \in I^{-\frac{1}{2}}(C_{CW})$, but the negative implications for artifacts are nevertheless relevant here, as can be seen by microlocalizing to $\Sigma^1 \setminus \Sigma^2$ and $\Sigma^2 \setminus \Sigma^1$, strongly indicating but not proving the presence of strong, nonremovable artifacts in reconstructions from crosswell and walkaway data. However, the presence of strong artifacts can definitely be deduced from the microlocal structure of C_{CW} and C_{WA} near points where both π_L and π_R are folds. Folding canonical relations, for which both π_L and π_R are folds, were first studied in the context of scattering by obstacles [25], and then for linearized seismics in [28, 8, 9]. It was shown in [28, 8] that, if $A \in I^m(C)$, then $A^* A \in I^{2m,0}(\Delta, C_1)$ where C_1 is another folding canonical relation. Since $A^* A \in I^{2m}(\Delta \setminus C_1)$ and $A^* A \in I^{2m}(C_1 \setminus \Delta)$ by (6), the artifact created by C_1 is as strong as the true image, again resulting in a nonremovable artifact.

7. Appendix: Singularity classes. Let V and W be smooth manifolds, initially of the same dimension, n , and let $f : V \rightarrow W$ be a smooth function. Let $\Sigma(f) := \{x \in V : \det(df(x)) = 0\}$ be the set of critical points of f . (This and all of the sets defined below are coordinate-independent.) The only singularities we will be concerned with are those which are *corank one*, by which we mean points $x_0 \in V$ such that

$$\text{rank}(df(x_0)) = n - 1 \text{ and } d(\det(df))(x_0) \neq 0. \quad (59)$$

If f only has corank one singularities, then Σ is a smooth hypersurface in V .

Definition 7.1. $f : V \rightarrow W$ is a blowdown if $\ker(df) \subseteq T\Sigma(f)$ at all points of $\Sigma(f)$.

Definition 7.2. f has singularities of (Whitney) fold type if, for every $x \in \Sigma(f)$, $\ker(df(x))$ intersects $T_x\Sigma(f)$ transversally.

Now consider the non-equidimensional situation. There is some variation in the literature in terms of how these singularities are denoted. The analogues of Whitney folds are called *submersions with folds* (if $\dim(V) > \dim(W)$) or *cross caps* (if $\dim(V) < \dim(W)$). Suppose that $\dim(V) = N$, $\dim(W) = M$, with $N \geq M$. For $N = M$, submersions with folds are Whitney folds, and are denoted by $S_{1,0}$ (in the Thom theory of C^∞ singularities [17]) and by $\Sigma_{1,0}$ (in the Boardman-Morin theory [27]) in the equidimensional case. In general,

Definition 7.3. f is a submersion with folds if the only singularities of f are of type $S_{1,0}$ (Thom) or $\Sigma_{N-M+1,0}$ (Boardman-Morin).

For our purposes, we do not need to define the classes $S_{1,0}$ or $\Sigma_{N-M+1,0}$, but simply recall that one can verify that f is a submersion with folds as follows. At points where $\text{rank } df \geq M - 1$, by [27], we can choose suitable adapted local coordinates on V and W such that f has the form: $f(x_1, x_2, \dots, x_{M-1}, x_M, \dots, x_N) = (x_1, x_2, \dots, x_{M-1}, g(x))$. The set $\Sigma(f)$ where f drops rank (by 1, by assumption) is described by $\Sigma(f) = \{x : \frac{\partial g}{\partial x_i} = 0, M \leq i \leq N\}$. Then f is a submersion with folds if, for all $x \in \Sigma(f)$,

(i) $\left\{d\left(\frac{\partial g}{\partial x_i}\right) : M \leq i \leq N\right\}$ is linearly independent (so that $\Sigma(f)$ is a smooth submanifold of V); and

(ii) the $(N - M + 1)$ -dimensional kernel of $df(x)$ is transversal to the tangent space of $\Sigma(f)$ in $T_x V$.

These conditions can be combined [27] into

$$\det \left[\frac{\partial^2 g}{\partial x_i \partial x_j} \right]_{M \leq i, j \leq N} \neq 0, \quad (60)$$

and this is independent of the choice of adapted coordinates.

For each N, M , there are a finite number of local normal forms for a submersion with folds, determined by the signature of the Hessian of f [17]:

$$f(x_1, x_2, \dots, x_N) = (x_1, x_2, \dots, x_{M-1}, x_M^2 \pm x_{M+1}^2 \pm \dots \pm x_N^2).$$

In the case relevant here, $N = 4 = M + 1$ and the last entry is a quadratic form in two variables.

We now define the final singularity class of interest, assuming that $f : V \rightarrow W$, with $\dim(V) = N < \dim(W) = M$.

Definition 7.4. f is a cross cap if the only singularities of f are of type $S_{1,0}$ (Thom) or $\Sigma_{1,0}$ (Boardman-Morin).

One can identify a cross cap as follows [27]. At a point where df has rank $\geq N-1$, we can find suitable adapted coordinates such that

$$f(x_1, x_2, \dots, x_{N-1}, x_N) = (x_1, x_2, \dots, x_{N-1}, g_1, g_2, \dots, g_q),$$

where $q = M - N + 1$. The set $\Sigma(f)$ where f drops rank by 1 from its maximal possible value, N , is given by $\Sigma(f) = \{x : \frac{\partial g_i}{\partial x_N} = 0, \quad 1 \leq i \leq q\}$. Assume that there is an i_0 , such that $\frac{\partial^2 g_{i_0}}{\partial x_N^2}(0) \neq 0$. Then, g has a cross cap singularity near 0 if the map $\chi : \mathbb{R}^N \rightarrow \mathbb{R}^q$ given by $\chi(x_1, x_2, \dots, x_N) = \left(\frac{\partial g_1}{\partial x_N}, \frac{\partial g_2}{\partial x_N}, \dots, \frac{\partial g_q}{\partial x_N}\right)$ satisfies $\text{rank}(d\chi(0)) = q$. (Notice that this forces $N \geq q$, i.e., $M \leq 2N - 1$.) These conditions can be expressed as:

- (i) $\Sigma(f)$ is smooth and of codimension q ;
- (ii) the $N \times N$ minors of df generate the ideal of $\Sigma(f)$; and
- (iii) $\ker(df) \cap T\Sigma(f) = (0)$.

As for folds, there is a local normal form for cross caps, due to Whitney [37] and Morin [27]:

$$f(x_1, x_2, \dots, x_N) = (x_1, x_2, \dots, x_{N-1}, x_1 x_N, \dots, x_{M-N} x_N, x_N^2). \quad (61)$$

Acknowledgments. This paper grew out of work supported by an American Institute of Mathematics Structured Quartet Research Experience (SQuaRE). The authors would like to thank Olga Podgornova for useful conversations. RG and CN were partly supported by Science Foundation Ireland under Grant number 16/RC/3918. AG was partly supported by US NSF Applied Mathematics Program grant DMS-1906186.

REFERENCES

- [1] J. Ajo-Franklin, J. Peterson, J. Doetsch and T. Daleya, [High-resolution characterization of a CO2 plume using crosswell seismic tomography: Cranfield, MS, USA](#), *Int. J. Greenhouse Gas Control*, **18** (2013), 497–509.
- [2] G. Ambartsoumian, R. Felea, V. Krishnan, C. Nolan and E. T. Quinto, [A class of singular Fourier integral operators in synthetic aperture radar imaging, II: Transmitter and receiver with different speeds](#), *SIAM J. Math. Analysis*, **50** (2018), 591–621.
- [3] A. Balch and M. Lee, *Vertical Seismic Profiling: Technique, Applications, and Case Histories*, International Human Resources Development Corporation (Boston), 1984.
- [4] G. Beylkin, [Imaging of discontinuities in the inverse problem by inversion of a generalized Radon transform](#), *J. Math. Phys.*, **26** (1985), 99–108.
- [5] T. Daley, E. Majer and J. Peterson, [Crosswell seismic imaging in a contaminated basalt aquifer](#), *Geophysics*, **69** (2004), 16–24.
- [6] T. Daley, L. Myer, J. Peterson, E. Majer and G. Hoversten, Time-lapse crosswell seismic and VSP monitoring of injected CO₂ in a brine aquifer, *Environ. Geology*, **54** (2008), 1657–1665.
- [7] J. J. Duistermaat and V. Guillemin, [The spectrum of positive elliptic operators and periodic bicharacteristics](#), *Inv. Math.*, **29** (1975), 39–79.
- [8] R. Felea, [Composition of Fourier integral operators with fold and blowdown singularities](#), *Comm. Partial Differential Equations*, **30** (2005), 1717–1740.
- [9] ———, Displacement of artefacts in inverse scattering, *Inverse Problems*, **23** (2007), 1519.
- [10] R. Felea, R. Gaburro, A. Greenleaf and C. Nolan, [Microlocal analysis of Doppler synthetic aperture radar](#), *Inverse Prob. Imaging*, **13** (2019), 1283–1307.
- [11] R. Felea and A. Greenleaf, [An FIO calculus for marine seismic imaging: Folds and cross-caps](#), *Comm. Partial Differential Equations*, **33** (2008), 45–77.

- [12] R. Felea and A. Greenleaf, Fourier integral operators with open umbrellas and seismic inversion for cusp caustics, *Math. Research Lett.*, **17** (2010), 867–886.
- [13] R. Felea, A. Greenleaf and M. Pramanik, An FIO calculus for marine seismic imaging, II: Sobolev estimates, *Math. Ann.*, **352** (2012), 293–337.
- [14] R. Felea, V. P. Krishnan, C. Nolan and E. T. Quinto, Common midpoint versus common offset acquisition geometry in seismic imaging, *Inverse Probl. Imaging*, **10** (2016), 87–102.
- [15] R. Felea and C. Nolan, Monostatic SAR with fold/cusp singularities, *J. Fourier Anal. Appl.*, **21** (2015), 799–821.
- [16] R. Felea and E. T. Quinto, The microlocal properties of the local 3-D SPECT operator, *SIAM J. Math. Anal.*, **43** (2011), 1145–1157.
- [17] M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities*, Graduate Texts in Mathematics, **14**, Springer-Verlag, New York-Heidelberg, 1973.
- [18] A. Greenleaf and G. Uhlmann, Estimates for singular Radon transforms and pseudodifferential operators with singular symbols, *J. Funct. Anal.*, **89** (1990), 202–232.
- [19] V. Guillemin, On some results of Gelfand in integral geometry, *Proc. Symp. Pure Math.*, **43** (1985), 149–155.
- [20] V. Guillemin and G. Uhlmann, Oscillatory integrals with singular symbols, *Duke Math. J.*, **48** (1981), 251–267.
- [21] S. Hansen, Solution of a hyperbolic inverse problem by linearization, *Comm. Partial Differential Equations*, **16** (1991), 291–309.
- [22] L. Hörmander, Fourier integral operators, I, *Acta Math.*, **127** (1971), 79–183.
- [23] A. Kirsch and A. Rieder, On the linearization of operators related to the full waveform inversion in seismology, *Math. Meth. Appl. Sci.*, **37** (2014), 2995–3007.
- [24] A. ten Kroode, D. Smit and A. Verdel, A microlocal analysis of migration, *Wave Motion*, **28** (1998), 149–172.
- [25] R. Melrose and M. Taylor, Near peak scattering and the corrected Kirchhoff approximation for a convex obstacle, *Adv. in Math.*, **55** (1985), 242–315.
- [26] R. Melrose and G. Uhlmann, Lagrangian intersection and the Cauchy problem, *Comm. Pure Appl. Math.*, **32** (1979), 483–519.
- [27] B. Morin, Formes canoniques des singularités d’une application différentiable, I, *C. R. Acad. Sc. Paris*, **260** (1965), 5662–5665; II, *C. R. Acad. Sc. Paris*, **260** (1965), 6503–6506.
- [28] C. Nolan, Scattering in the presence of fold caustics, *SIAM J. Appl. Math.*, **61** (2000), 659–672.
- [29] C. Nolan and M. Cheney, Microlocal analysis of synthetic aperture radar, *J. Fourier Anal. Appl.*, **10** (2004), 133–148.
- [30] C. J. Nolan and W. W. Symes, Global solution of a linearized inverse problem for the wave equation, *Comm. Partial Differential Equations*, **22** (1997), 919–952.
- [31] O. Podgornova, Lecture at ICERM, 2015, https://icerm.brown.edu/video_archive/?play=720.
- [32] Rakesh, A linearized inverse problem for the wave equation, *Comm. Partial Differential Equations*, **13** (1988), 573–601.
- [33] C. Schmeltz, et al., Advanced seismic processing/imaging techniques and their potential for geothermal exploration, *Interpretation*, **4** (2016).
- [34] C. Stolk, Microlocal analysis of a seismic linearized inverse problem, *Wave Motion*, **32** (2000), 267–290.
- [35] W. W. Symes, The seismic reflection inverse problem, *Inverse Problems*, **25** (2009), 123008.
- [36] A. Weinstein, On Maslov’s quantization condition, In *Fourier Integral Operators and Partial Differential Equations*, Lecture Notes in Math., Springer-Verlag, New York, **459** (1975), 341–372.
- [37] H. Whitney, The general type of singularity of a set of $2n - 1$ smooth functions of n variables, *Duke Math. J.*, **10** (1943), 161–172.

Received October 2021; revised March 2022; early access May 2022.

E-mail address: rxfsma@rit.edu

E-mail address: Romina.Gaburro@ul.ie

E-mail address: allan@math.rochester.edu

E-mail address: Clifford.Nolan@ul.ie