Available online at www.sciencedirect.com
ScienceDirect
Wave Motion 43 (2006) 579–598
WAVE MOTION
www.elsevier.com/locate/wavemoti

Seismic inverse scattering in the downward continuation approach ☆

Christiaan C. Stolk [a,*], Maarten V. de Hoop [b]

[a] *University of Twente, Department of Applied Mathematics, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands*
[b] *Center for Computational and Applied Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907, USA*

Received 17 November 2005; received in revised form 4 May 2006; accepted 9 May 2006
Available online 7 July 2006

Abstract

Seismic data are commonly modeled by a linearization around a smooth background medium in combination with a high frequency approximation. The perturbation of the medium coefficient is assumed to contain the discontinuities. This leads to two inverse problems, first the linearized inverse problem for the perturbation, and second the estimation of the background, which is a priori unknown (velocity estimation). Here we give a reconstruction formula for the linearized problem using the downward continuation approach. The reconstruction is done microlocally, up to an explicitly given pseudodifferential factor that depends on the aperture. Our main result is a characterization of the wave-equation angle transform, derived from downward continuation, that generates the common image point gathers as an invertible Fourier integral operator, microlocally. We show that the common image point gathers obtained with this particular angle transform are free of so called kinematic artifacts, even in the presence of caustics. The assumption is that the rays in the background that are associated with the reflections due to the medium perturbation are nowhere horizontal. Finally, making use of the mentioned angle transform, pseudodifferential annihilators of the data are constructed. These annihilators detect whether the data are contained in the range of the modeling operator, which is the precise criterion in migration velocity analysis to determine whether a background medium is acceptable, even in the presence of caustics.

Keywords: Seismic inversion; Microlocal analysis; Double-square-root equation

1. Introduction

In reflection seismology one places point sources and point receivers on the earth's surface. Each source generates acoustic waves in the subsurface, that are reflected where the medium properties vary discontinuously. The recorded reflections that can be observed in the data are used to reconstruct these discontinuities. In this

☆ The authors thank The Mathematical Sciences Research Institute for partial support, through NSF Grant DMS-9810361. They also thank the members of the institute, and in particular Gunther Uhlmann, for providing a very stimulating environment during the Inverse Problems program in Fall 2001.
* Corresponding author. Tel.: +31 53 4893460; fax: +31 53 4894833.
E-mail address: c.c.stolk@ewi.utwente.nl (C.C. Stolk).

0165-2125/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.wavemoti.2006.05.003

List of symbols

Symbol	*Description and Equation*
$c_0(z, x)$	background wavespeed, below (1)
$\delta c(z, x)$	wavespeed contrast, below (1)
η	bicharacteristics (wave equation), (2)
τ	frequency, (2)
G	Green's function (wave equation) below, (1)
δG	scattered wavefield (Born approximation), (3)
Y	acquisition manifold
X	subsurface manifold
t_s	source-ray traveltime, (4)
t_r	receiver-ray traveltime, (4)
d	data
$C_{-,\theta}$	cone of propagating wave directions
B_-	single-square-root operator, (5)
b	principal symbol of B_-
C	damping operator, (6)
Q_-	decomposition operator (component), (5)
G_-	one-way wave propagator
γ	bicharacteristics (one-way wave equation), (7)
H	double-square-root propagator, (8)
Γ	bicharacteristics (double-square-root equation), (10)
E_1	extension operator, (13)
$R_1 = E_1^*$	restriction operator, (29)
E_2	extension operator, (14)
$R_2 = E_2^*$	restriction operator, (28)
L	(15)
$\bar{L}$	(34)
K	depth-to-time conversion operator, (17)
$\bar{K}$	(33)
A	amplitude ($\bar{K}$), (38)
S	generating function ($\bar{K}$), (38)
Σ	map associated with canonical relation of K
ψ_D	(18)
F_D	scattering operator, (16), (25)
Ω_θ	(19)
Z_{max}	below, (19)
T_{max}	(19)
ψ_Y	tapered mute, (20)
N	normal operator, (20)
p	(22)
Λ	(22)
E_3	extension operator, (23)
$R_3 = E_3^*$	restriction operator above, (56)
E_4	extension operator, (24)
A_{WE}	angle transform below, (25), (74)
Θ	frequency-to-vertical-wavenumber map
a_z	(36)
Ξ	(35)

$\bar{\Xi}$	(37)
Ψ, Ψ_0	(45)
Φ, Φ_{n-1}	(47), (46)
$\bar{\Psi}$	above, (53)
$\bar{\Phi}$	(55)
V	above, (50)
χ	cutoff function, (56)
C_0	(58)
Ψ_{WE}	above, (74), (75)
$\jmath$	amplitude correcting factor, (86)
M_j	offset multiplication operator above, (92)
W_j	annihilator, (93)

paper, together with an earlier paper, which will be referred to as Paper I [27], we develop an inverse scattering method using a downward continuation approach. The notion of downward continuation refers to the continuation of surface seismic data into the subsurface. The purpose is to provide a mathematical foundation of the downward continuation approach to seismic data processing. (In particular that part of seismic data processing that involves migration.)

Our main analytical tools will be from microlocal analysis, see e.g. the general references [9,32,12–14]. This should not come as a surprise. Indeed, the successful formation of images by migration is based on the propagation of wave fronts along the rays of geometrical acoustics. The theory of microlocal analysis precisely captures this behavior of the solutions of wave equations, and is therefore a natural tool for the analysis of seismic migration. A number of the references below also use microlocal techniques.

We study the inversion of seismic data with an acoustic model. Let $z \in \mathbb{R}$ denote the vertical (depth) coordinate, $x \in \mathbb{R}^{n-1}$ the horizontal coordinate(s) and t the time, and let $c = c(z, x)$ be the wave speed, then, assuming constant density of mass, the scalar wave equation for acoustics is given by

$$Pu = f, \quad P = c(z,x)^{-2}\partial_t^2 - \partial_z^2 - \sum_{j=1}^{n-1} \partial_{x_j}^2. \tag{1}$$

When the wavespeed c is smooth, the singularities of the solution propagate along curved trajectories, characteristics, while discontinuities in the wavespeed lead to reflections.

In practice, to model seismic reflection data (1) is not used directly, but a linearization is invoked, together with a separation of "scales". The linearization is in the coefficient c around a smooth background c_0, $c = c_0 + \delta c$. We denote the Green's function of (1) with velocity c_0 by G, with $G = 0$ for $t < 0$. The perturbation δc is assumed to contain the short wavelength variations of c, and has no long wavelength components. Note that this is only a partial linearization: The dependence of the wavefield on c_0 is still nonlinear. The high-frequency part of the perturbation in the Green's function, which we denote by δG, models the once reflected waves, whence this approximation is called the high-frequency single scattering approximation.

In the seismic inverse problem, both the smooth background c_0 and the perturbation δc are unknown and have to be reconstructed from the reflection data. This leads to two related inverse problems. The first one is the linear inverse problem of reconstructing δc given c_0. The solution of this problem is used in the estimation of c_0, the second problem, which is highly nonlinear. The purpose of this paper is to develop and analyze methods based on downward continuation for these two inverse problems; the advantages of the downward continuation approach in the presence of caustics are elucidated in Section 1.4 below.

1.1. Microlocal analysis and linearization of the scattering problem

The bicharacteristics, η with components $(\eta_z, \eta_x, \eta_\zeta, \eta_\xi)$ in phase space, that describe the propagation of singularities by the wave equation are the solutions to a Hamilton system, derived from the principal symbol of

P, and will be parameterized by initial position (z_0, x_0), take-off direction $\alpha \in S^{n-1}$, frequency τ, which together define the initial cotangent vector $(\zeta_0, \xi_0) = -\tau c(z_0, x_0)^{-1}\alpha$, and time t,

$$\eta(t,z_0,x_0,\alpha,\tau) = (\eta_z(t,z_0,x_0,\alpha,\tau), \eta_x(t,z_0,x_0,\alpha,\tau), t, \eta_\zeta(t,z_0,x_0,\alpha,\tau), \eta_\xi(t,z_0,x_0,\alpha,\tau), \tau). \tag{2}$$

Here the evolution parameter is the time t. Note that τ is invariant along the Hamilton flow (see e.g. the introduction and chapter 5 of [9]).

Hörmander's wave front set will be denoted by WF(u), for a distribution u in $\mathscr{D}'(\mathbb{R}^k)$ for some k [12, chapter 8]. For the further analysis, we also need the following notion of microlocal equivalence. Let $u, v \in \mathscr{D}'(\mathbb{R}^k)$. Then $u \equiv v$, microlocally on $\Gamma \subset T^*\mathbb{R}^k \setminus 0$, if $\mathrm{WF}(u - v) \cap \Gamma = \emptyset$.

The data are assumed to be modeled by $\delta G(z_r, r, t, z_s, s)$, for $z_r = z_s = 0$. Here $s \in \mathbb{R}^{n-1}$, $r \in \mathbb{R}^{n-1}$ denote the horizontal source and receiver position, respectively. We have (cf. Paper I and [16])

$$\delta G(0,r,t,0,s) = \int \mathrm{d}z \int \mathrm{d}x \int \mathrm{d}t' G(0,r,t-t',z,x) \frac{2\delta c(z,x)}{c_0(z,x)^3} \partial_t^2 G(z,x,t',0,s). \tag{3}$$

We assume that the data are available for (s, r, t) in a bounded open subset Y, called the acquisition manifold, of $\mathbb{R}^{2n-2} \times \mathbb{R}_+$. Except for Remark 2.5 we assume that no further restrictions apply to Y; the acquisition geometry is maximal. The modeling or scattering operator F is, for given c_0, defined as the map from δc to δG restricted to Y. Since Y is bounded and the waves propagate with finite speed we may assume that δc is supported in a bounded open subset X of $\mathbb{R}^{n-1} \times \mathbb{R}_+$. We furthermore assume that $\overline{X} \cap \{z = 0\} = \emptyset$.

Under certain conditions, the operator F is a Fourier integral operator [21] with canonical relation

$$\begin{aligned}&\{(\eta_x(t_s,z,x,\beta,\tau), \eta_x(t_r,z,x,\alpha,\tau), t_s + t_r, \eta_\xi(t_s,z,x,\beta,\tau), \eta_\xi(t_r,z,x,\alpha,\tau), \tau; z, x, \zeta, \xi) | t_s, t_r > 0,\\ &\eta_z(t_s,z,x,\beta,\tau) = \eta_z(t_r,z,x,\alpha,\tau) = 0, (\zeta,\xi) = -\tau c_0(z,x)^{-1}(\alpha + \beta), (z,x,\alpha,\beta,\tau)\\ &\in \text{ subset of } X \times (S^{n-1})^2 \times \mathbb{R} \setminus 0\} \subset T^*\mathbb{R}^{2n-1}_{(s,r,t)} \times T^*\mathbb{R}^n_{(z,x)},\end{aligned} \tag{4}$$

see also the discussion in the introduction of Paper I; for a general reference of Fourier integral operators, see [9]. We note that WF(δG) is contained in the composition of this canonical relation with WF(δc).

1.2. The downward continuation method

The main topic of this paper is the so-called wave-equation or *downward continuation* approach [5,4,19] to seismic inverse scattering. We summarize and discuss the results of Paper I. Key ingredients in this approach are an operator $H(z, z_0)$ and its adjoint $H(z, z_0)^*$, where for technical reasons $z_0 > z$. Let the data be denoted by $d = d(s, r, t)$. The result of acting with $H(0, z)^*$ on d has the geophysical interpretation of the data that would have been measured if the plane containing sources and receivers would be at depth z (this interpretation is not exact, it ignores certain pseudodifferential factors affecting the amplitude in the data, and can only be made after restriction to positive times of $H(0, z)^*d$). Thus $H(z, z_0)$ and its adjoint are operators acting on distributions of $(s,r,t) \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \times \mathbb{R}$. In the following paragraphs we will discuss this in more detail. We start with so called one-way wave equations.

1.2.1. One-way wave propagation

Let $C_{-,\theta}$ be the cone $C_{-,\theta} = \{(v_z, v_x) \in \mathbb{R} \times \mathbb{R}^{n-1} | \|v_x\| < \tan(\theta)|v_z| \text{ and } v_z < 0\}$. One-way wave equations describe only waves propagating with ray velocities, $(v_z, v_x) = \frac{\mathrm{d}(\eta_z,\eta_x)}{\mathrm{d}t}$, inside such a cone, where $\theta < \frac{\pi}{2}$. Numerical one-way wave equations have been extensively studied, see e.g. [4,22], or [8] and references therein. For the analysis of such equations, see [25]. In this section, we summarize some of the aspects needed here.

The essential property of one-way wave equations is that they give solutions to the wave equation microlocally for singularities propagating along bicharacteristics that stay in C_{-,θ_1}, for some given θ_1. To make this precise, we consider a bicharacteristic going through some point (z_0, x_0, t_0, ζ_0, ξ_0, τ), and assume that

(1) the associated velocity vector $\frac{\mathrm{d}(\eta_z,\eta_x)}{\mathrm{d}t}$ stays in C_{-,θ_1} along some open interval including (z_0, x_0, t_0, ζ_0, ξ_0, τ);
(2) u satisfies Eq. (1) with f satisfying $f \equiv 0$ microlocally on this interval of the bicharacteristic;

(3) the function f is in C^∞ on a neighborhood of (z_0, x_0, t_0) and $(z_0, x_0, t_0, -\zeta_0, \xi_0, \tau_0) \notin \mathrm{WF}(u)$ (this would generate a singularity with vertical propagation velocity $v_z > 0$, i.e. in the opposite direction as those in C_{-,θ_1}).

Then $Q_-^* u_- \equiv u$ on this bicharacteristic, where $Q_- = Q_-(z, x, D_x, D_t)$ is an elliptic pseudodifferential operator with symbol i sgn $(\tau)|\tau|^{-1/2}(c_0^{-2} - \tau^{-2}||\xi||^2)^{-1/4}$, and u_- is the solution to an initial value problem in z,

$$(\partial_z - \mathrm{i}B_-)u_- = 0, \quad z < z_0, \quad Q_-^* u_-|_{z=z_0} \equiv u|_{z=z_0}, \tag{5}$$

with B_- a pseudodifferential operator with principal symbol given by

$$-b(z, x, \xi, \tau) = \tau\sqrt{c_0(z, x)^{-2} - \tau^{-2}||\xi||^2}.$$

Here we use the notations $D_x = -\mathrm{i}\partial_x$. Such B_- and Q_- exist for any fixed $\theta_1 < \frac{\pi}{2}$.

We assume that singularities propagating with velocity outside the cone C_{-,θ_2}, $\theta_1 < \theta_2 < \frac{\pi}{2}$, are strongly damped. This damping can be obtained by introducing a new term Cu_- in the differential equation, where $C(z, x, D_x, D_t)$ is a suitably chosen pseudodifferential operator, see [25]; such an approach was followed in a finite-difference scheme for the paraxial approximation to the one-way wave equation in [29, IV.D]. We thus replace (5) by the one-way wave equation with damping,

$$(\partial_z - \mathrm{i}B_- - C)u_= 0, \quad z < z_0. \tag{6}$$

Denote by $G_-(z, z_0)$ the evolution operator associated with this equation, mapping initial values at z_0 to the solution at z. Then $\mathscr{H}Q_-^*(z, x, D_x, D_t)G_-(z, z_0)Q_-(z_0, x_0, D_{x_0}, D_{t_0})$ models microlocally the upgoing constituents of the solution operator to (1), in the sense that its distribution kernel is microlocally equal to $G(z, x, t, z_0, x_0)$. Here $\mathscr{H}$ denotes the Hilbert transform in time.

$G_-(z, z_0)$ propagates singularities along the bicharacteristics of $\partial_z - \mathrm{i}\, B_-$; they are parameterized by z, initiated at (z_0, x_0) and ξ_0, and denoted as $\gamma(z, z_0, x_0, \xi_0, \tau)$. In components, we write them as (note that they are time translation invariant)

$$\gamma(z, z_0, x_0, t_0, \xi_0, \tau) = (z, \gamma_x(z, z_0, x_0, \xi_0, \tau), \gamma_t(z, z_0, x_0, \xi_0, \tau) + t_0, -b(z, \gamma_x, \gamma_\xi, \tau), \gamma_\xi(z, z_0, x_0, \xi_0, \tau), \tau). \tag{7}$$

1.2.2. Upward/downward continuation

In Paper I, we defined an operator $H(z, z_0)$, called the upward continuation operator, by a composition of two one-way wave Green's functions. It acts on distributions of $(s, r, t) \in \mathbb{R}^{n-1} \times \mathbb{R}^{n-1} \times \mathbb{R}$, and is given by

$$H(z, z_0) = (\mathrm{Id}_s \otimes G_{-,r}(z, z_0)) \circ (G_{-,s}(z, z_0) \otimes \mathrm{Id}_r). \tag{8}$$

Here, $G_{-,s}(z, z_0)$ acts in the (s, t) variables, and Id_s is the identity operator on functions of s, and a similar notation is assumed for $G_{-,r}(z, z_0)$ and Id_r. We observed in Paper I that $G_{-,r}$, $B_{-,r}$, C_r commute with $G_{-,s}$, $B_{-,s}$, C_s. From this it follows that $H(z, z_0)$ is the solution operator of the initial value problem for the so-called *double-square-root (DSR) equation*

$$\left(\frac{\partial}{\partial z} - \mathrm{i}B_-(z, s, D_s, D_t) - \mathrm{i}B_-(z, r, D_r, D_t) - C(z, s, D_s, D_t) - C(z, r, D_r, D_t)\right)u = 0, \quad 0 \leqslant z < z_0, \tag{9}$$

i.e., the solution to (9) with initial value $u(z_0, s_0, r_0, t_0) = v(s_0, r_0, t_0)$ is given by $H(z, z_0)v$. There are bicharacteristics associated with (9) and they are given by

$$\begin{aligned}\Gamma(z, z_0; s_0, r_0, t_0, \sigma_0, \rho_0, \tau) = {} & (\gamma_x(z, z_0, s_0, \sigma_0, \tau), \gamma_x(z, z_0, r_0, \rho_0, \tau), t_0 + \gamma_t(z, z_0, s_0, \sigma_0, \tau) \\ & + \gamma_t(z, z_0, r_0, \rho_0, \tau), \gamma_\xi(z, z_0, s_0, \sigma_0, \tau), \gamma_\xi(z, z_0, r_0, \rho_0, \tau), \tau).\end{aligned} \tag{10}$$

The adjoint $H(0, z)^*$ is used to compute the downward continued data (for example, by method of generalized screen expansion [8]).

To justify the substitution for the time convolution of the two Green's functions in (3) by H, we will need

Assumption 1. (DSR assumption) If $(z, x) \in X$ and $\alpha, \beta \in S^{n-1}$, $t_s, t_r > 0$ depending on (z, x, α, β) are such that $\eta_z(t_s, z, x, \beta, \tau) = \eta_z(t_r, z, x, \alpha, \tau) = 0$ and $(\eta_x(t_s, z, x, \beta, \tau), \eta_x(t_r, z, x, \alpha, \tau), t_s + t_r) \in Y$ (cf. (4)), then

$$c(z,x)^{-1}\frac{\partial\eta_z}{\partial t}(t,z,x,\beta,\tau) < -\cos(\theta_1), t\in[0,t_{\mathrm{s}}], \tag{11}$$

$$c(z,x)^{-1}\frac{\partial\eta_z}{\partial t}(t,z,x,\alpha,\tau) < -\cos(\theta_1), t\in[0,t_{\mathrm{r}}]. \tag{12}$$

1.2.3. Factorization of the scattering operator

We will denote by F_{D} the scattering operator but with the one-way Green's functions substituted for the Green's functions (cf. (3)). In Section 5 of Paper I, we compared this scattering operator, in the downward continuation approach, with the original operator F. To write F_{D} in convenient form, we first define two mappings,

$$E_1 : \mathscr{D}'(\mathbb{R}^n) \to \mathscr{D}'(\mathbb{R}^{2n-1}) : c(z,x) \mapsto \delta(x-\bar{x})c\left(z,\frac{\bar{x}+x}{2}\right), \tag{13}$$

$$E_2 : \mathscr{D}'(\mathbb{R}^{2n-1}) \to \mathscr{D}'(\mathbb{R}^{2n}) : h(z,\bar{x},x) \mapsto \delta(t)h(z,\bar{x},x), \tag{14}$$

and a mapping $L : \mathscr{E}'(\mathbb{R}^{2n}_{(z,s,r,t)}) \to \mathscr{D}'(\mathbb{R}^{2n-1}_{(s,r,t)})$ by

$$Lg = Q^*_{-,s}(0)Q^*_{-,r}(0)\int_0^Z H(0,z)Q_{-,s}(z)Q_{-,r}(z)g(z,\cdot,\cdot,\cdot))(s,r,t)\,\mathrm{d}z. \tag{15}$$

It was shown that

$$F_{\mathrm{D}} : \frac{1}{2}c_0^{-3}\delta c \to D_t^2 L E_2 E_1(\tfrac{1}{2}c_0^{-3}\delta c). \tag{16}$$

In the above, for given z, $g = E_2E_1(\frac{1}{2}c_0^{-3}\delta c)$ has the interpretation of kernel of a reflection operator. (Note that we have absorbed a multiplication by $2c_0^3$ in F_{D} unlike the definition in Paper I.)

For the composition LE_2 we introduced the notation K,

$$K = LE_2. \tag{17}$$

The operators L and K are Fourier integral operators, see Paper I, Lemma 3.1 and Theorem 4.2. We found (Theorem 5.1) that there is a pseudodifferential operator ψ_{D} such that

$$F_{\mathrm{D}}\left(\tfrac{1}{2}c_0^{-3}\delta c\right) \equiv \psi_{\mathrm{D}}F\ \delta c \text{ or } F_{\mathrm{D}}\tfrac{1}{2}c_0^{-3}\cdot \equiv \psi_{\mathrm{D}}F. \tag{18}$$

The operator ψ_{D} can be called a pseudodifferential cutoff, it is 1 on the set Ω_{θ_1}, and in $S^{-\infty}$ outside Ω_{θ_2}, with $0<\theta_1<\theta_2<\pi/2$ and where Ω_θ is given by

$$\Omega_\theta = \{(s,r,t,\sigma,\rho,\tau)|t < T_{\max}(0,s,r,\sigma,\rho,\tau,\theta)\}. \tag{19}$$

Here the maximal time $T_{\max} = T_{\max}(0, s, r, \sigma, \rho, \tau, \theta)$ is determined by the canonical relation of K from the maximal depth, $Z_{\max} = Z_{\max}(0, s, r, \sigma, \rho, \tau, \theta)$, that is the upper boundary of the maximal interval containing $z = 0$ for which the source and receiver rays satisfy the DSR assumption at $\theta = \theta_1$.

1.3. The reconstruction of δc given c_0

There are a number of results concerning the approximate reconstruction of δc given c_0. Those differ in the assumptions on the dimension of the acquisition manifold and on the behavior of the rays (e.g. whether caustics are allowed) [1,11,17,16,23]. The approach common in these works is to first establish that the scattering operator is a Fourier integral operator and determine its canonical relation, and then to introduce a trial inverse scattering operator with the transpose canonical relation, but with a symbol that needs to be determined. This approach stems from the notion of imaging.

We introduce a function $\psi_Y \in C_0^\infty(Y)$ that goes smoothly to zero near the boundary of Y. Applying the imaging after modeling yields the composition of operators $N:= F^*\psi_Y F$, called the *normal operator*. The Bolker condition of Guillemin [10] is

Assumption 2. The projection of the canonical relation (4) on $T^*Y\backslash 0$ is an embedding.

Since (4) is a canonical relation that projects submersively on the subsurface variables (z, x, ζ, ξ), the projection of (4) on $T^*Y\backslash 0$ is immersive [14, 25.3.6]. Therefore only the injectivity in the assumption needs to be verified [16].

Theorem 1.1. *With Assumption 2 the operator N is a pseudodifferential operator.*

Applying the adjoint $F^*\psi_Y$ to the data d yields the normal equation,

$$N(z,x,D_z,D_x)\delta c = F^*\psi_Y d. \tag{20}$$

Thus, the problem becomes to invert this normal operator. Microlocally the inverse of N exists, where its principal symbol is nonzero. The subset of $T^*\mathbb{R}^n_{(z,x)}\setminus 0$ where this is the case is determined as follows. The principal symbol of the normal operator is nonzero at (z, x, ζ, ξ) whenever there is a point (s, r, t, σ, ρ, τ; z, x, ζ, ξ) in the canonical relation (4) with (s, r, t, σ, ρ, τ) in the support of ψ_Y. The mentioned subset is representative for the ''illumination'' of δc by the available data.

In the reconstruction of δc, seismologists distinguish between imaging, which produces a function that has singularities at the same position as the ones in δc, and methods that also correctly compute the size of the singularities or the discontinuities. In the latter case they speak of inversion, or of true-amplitude imaging. Thus $F^*\psi_Y$ is an imaging operator, while $\langle N\rangle^{-1}F^*\psi_Y$ is an inversion operator, where $\langle N\rangle^{-1}$ is the regularized parametrix of N, that is an inverse microlocally for a subset of $T^*\mathbb{R}^n_{(z,x)}$ where the symbol of N is nonzero. The inversion operator reconstructs microlocally δc.

The first result of this paper concerns the reconstruction of δc in the downward continuation approach. In Section 2 we derive the normal equation in this approach from the Born DSR modeling operator, $F_{\rm D}$, and derive a reconstruction equation in Theorem 2.2. This theorem is the downward continuation counterpart of Theorem 1.1. The adjoint $F^*_{\rm D}$ is factorized into a product of operators to emphasize the different steps that compose an algorithm for reconstruction. It is anticipated that, because of the cutoff $\psi_{\rm D}$ in (18), fewer singularities in δc will be reconstructed than if the original operator F^* were used for the imaging. For example, vertical reflectors that could be illuminated by turning rays are not reconstructed. In many practical cases this disadvantage is however not so important.

1.4. A processing method for identifying acceptable c_0

The results of Section 3 concern the determination of c_0 through migration velocity analysis. Here it is exploited that the reconstruction of δc is an overdetermined problem. Indeed δc is a function of n variables, and d of $2n-1$ variables. Beylkin [1] gives conditions so that δc can be reconstructed by an operator $A_{\rm ss}$, say, from a subset of dimension n of the data, given c_0. An n dimensional subset of data are obtained, for instance, by taking the source coordinate constant. In conventional migration velocity analysis the data are viewed as an $(n-1)$-parameter family of such subsets, which results in an $(n-1)$-dimensional family of reconstructions. If we denote the $(n-1)$ parameters by p, and the family of reconstructions of δc by $A_{\rm ss}[c_0]d(z, x, p)$, then this gives the following criterion for the determination of c_0:

$$(A_{\rm ss}[c_0]d)(z,x,p) \text{ is independent of } p, \tag{21}$$

at least microlocally. In migration velocity analysis some c_0 is constructed based on this criterion. Note that c_0 need not be uniquely determined by the singular part of the data, that is why we say ''acceptable'' c_0.

In the presence of caustics the conditions of [1] are generally violated, and the application of $A_{\rm ss}$ (which can still be defined in this case) results in images with so called kinematic artifacts, that correspond to *nonlocal singular contributions to* $A_{\rm ss}F$ [17,18]. As a result, (21) is no longer valid. To remedy this, the parameter p can be chosen to no longer parameterize a subset of data, but instead to parameterize the angle between in- and out-going rays at the image point. In the Kirchhoff approach to seismic inverse scattering it has been proposed to use a generalized Radon transform to generate a set of images parameterized by angle (such an operator will be referred to as angle transform) [3,33]. However, in the presence of caustics, artifacts were identified in numerical examples [3,28], also in this case. Such artifacts are attributed to contributions from nonlocal operators to the composition of the generalized Radon transform with the modeling operator F. By microlocal analysis of the Kirchhoff approach the presence of artifacts was shown in [24].

The second and main result of this paper is to show that there is a downward continuation based angle transform, A_{WE}, that is artifact free under much weaker conditions than those of Beylkin [1] (Theorem 3.1). In particular, the presence of caustics is allowed. The map A_{WE} is defined in Section 3, and resembles an operator introduced in [6] to study angle dependent reflection coefficients. In (16) we observe that

$$\begin{aligned} E_2E_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right) &= (2\pi)^{-n}\int \left(\tfrac{1}{2}c_0^{-3}\delta c\right)\left(z,\frac{\bar{x}+x}{2}\right) \times \exp[\mathrm{i}\langle(\bar{x}-x),p\rangle\tau]\exp(\mathrm{i}t\tau)|\tau|^{n-1}\mathrm{d}p\mathrm{d}\tau \\ &=: \Lambda E_3E_4\left(\tfrac{1}{2}c_0^{-3}\delta c\right), \end{aligned} \tag{22}$$

where Λ is a pseudodifferential operator with symbol $|\tau|^{n-1}$, and $E_3 : \mathscr{D}'(\mathbb{R}^{2n-1}) \to \mathscr{D}'(\mathbb{R}^{2n})$, $E_4 : \mathscr{D}'(\mathbb{R}^{n}) \to \mathscr{D}'(\mathbb{R}^{2n-1})$ are given by

$$E_3f(z,\bar{x},x,t) = (2\pi)^{-(n-1)}\int \delta(t-\langle(x-\bar{x}),p\rangle)\, f\left(z,\frac{\bar{x}+x}{2},p\right)\,\mathrm{d}p, \tag{23}$$

$$\left(E_4\left(\tfrac{1}{2}c_0^{-3}\delta c\right)\right)(z,x,p) = \left(\frac{1}{2}c_0^{-3}\delta c\right)(z,x), \tag{24}$$

respectively, so that

$$F_{\mathrm{D}} = D_t^2L\Lambda E_3E_4. \tag{25}$$

By the map E_4 the perturbation $\frac{1}{2}c_0^{-3}\delta c$ is viewed as being p dependent as if it were a reflection coefficient. The angle transform A_{WE} is derived, in the imaging process, from the adjoint $E_3^*\Lambda L^*$. In the above, p is, given the singular direction from the wavefront set of δc, obtained through trigonometric formulae related to scattering angles [7].

We show that $A_{\mathrm{WE}}F$ is a p-family of pseudodifferential operators, which implies that nonlocal singular contributions are absent and data are mapped to a set of images (Theorem 3.1). In addition we show that A_{WE} is an invertible Fourier integral operator. We introduce an appropriate modification, denoted by $\widetilde{A}_{\mathrm{WE}} = \widetilde{A}_{\mathrm{WE}}[c_0]$ such that $\widetilde{A}_{\mathrm{WE}}[c_0]F[c_0]$ is a p-family of pseudodifferential operators with symbol 1, microlocally, hence mapping data to a set of reconstructions (Proposition 3.2). In [7] it was confirmed numerically that this angle transform does not generate artifacts in the presence of caustics.

The third result concerns the construction of operators we will refer to as annihilators of the data. The notion of acceptable background model c_0 can now be made precise by requiring that the data are in the range of the modeling operator $F_{\mathrm{D}}[c_0]$. The criterion that the data $d \in \mathrm{range}(F_{\mathrm{D}}[c_0])$ becomes equivalent to

$$(\widetilde{A}_{\mathrm{WE}}[c_0]d)(z,x,p) \text{ is independent of } p, \tag{26}$$

microlocally, where p can be identified with the integration variables in E_3 (cf. (23)). Thus, the criterion for migration velocity analysis (21) is extended to allow for background media with caustics.

Eq. (26) reveals the redundancy in the data. The existence of pseudodifferential operators that annihilate the singular part of the data are associated with this redundancy [26]. The annihilators are related to the differential semblance measure in the framework of Beylkin's conditions, for estimating the background medium [30]. Annihilators of the data in the downward continuation approach can be derived from $\widetilde{A}_{\mathrm{WE}}$. In addition, we construct in Section 4 a different annihilator, W, such that $\|Wd\|_{L^2}$ measures the *focusing* at $r = s$ of the downward continued data restricted to $t = 0$.

2. Imaging and reconstruction in the downward continuation approach

Conceptually, the first step in the reconstruction of the perturbation δc given the background, c_0, is applying the adjoint of the linear modeling map to the data. This is the process of imaging. We present a normal equation similar to (20) based on the double-square-root (DSR) modeling operator (16). The operator F^* in the right hand side of (20) is replaced by F_{D}^*, so that the right hand side can be computed using the downward continuation approach summarized in the introduction. As in Theorem 1.1 and the discussion below it, this leads to reconstruction modulo a pseudodifferential operator for which an explicit expression is given.

First, we discuss the geometric or kinematic considerations pertaining to the reconstruction and the prevention of imaging artifacts. The operator F_{D} is a Fourier integral operator whose canonical relation is a subset of that of F containing the elements with (z, x, α, β) such that Assumption 1 applies, with angle given by θ_2 below (18). It follows that in this case Assumption 2 can be replaced by the following weaker assumption

Assumption 3. If several elements of (4) project to the same point in T^*Y, then none of the (z, x, α, β) that describe these elements are such that Assumption 1 is satisfied.

In effect this assumption implies that the cutoff ψ_{D} has symbol in $S^{-\infty}$ for points in T^*Y where the projection from (4) is not injective. A cartoon of two ray pairs such that Assumption 3 is violated is given in Fig. 1. As discussed in Paper I, Section 6, the DSR Assumption 1 is stronger than Assumption 2, so is also sufficient for Theorem 1.1 to hold.

Second, we discuss a factorization of F^*_{D}, and the different adjoints of operators that make up this factorization. In this way we obtain insight in the structure of the imaging operator and in the different steps in algorithms to implement it.

Consider the factorization of F_{D} given in Eq. (16). The adjoint operator $H(0, z)^*$ propagates the data downward and backward in time. The adjoint of operator L is given by

$$(L^*d)(z,s,r,t) = Q^*_{-,s}(z)Q^*_{-,r}(z)H(0,z)^*Q_{-,s}(0)Q_{-,r}(0)d, \tag{27}$$

which yields, in general, a nonvanishing outcome for $t < 0$ also. The adjoint of extension operator E_2 is given by the restriction R_2 defined by

$$g(z,s,r,t)\mapsto(R_2g)(z,s,r) = g(z,s,r,0), \tag{28}$$

while the adjoint of extension operator E_1 is given by the restriction R_1 defined by

$$h(z,s,r)\mapsto(R_1h)(z,x) = h(z,x,x). \tag{29}$$

In any algorithm derived from the inverse scattering approach developed in this paper, the operator L^* is the key component; the computation of L^*d would be implemented as a 'marching' in depth z. The restrictions R_1, R_2 represent the imaging conditions and are applied at each depth after the data are downward continued.

The adjoint of the depth-to-time conversion operator, $K = L\ E_2$ equals

$$K^* = R_2L^*. \tag{30}$$

The canonical relation of K maps points $(z, s, r, \zeta, \sigma, \rho)$ in a subset of $T^*\mathbb{R}^{2n-1}\setminus 0$ diffeomorphically to points $(s_0, r_0, t_0, \sigma_0, \rho_0, \tau)$ in a subset of the cotangent acquisition space $T^*\mathbb{R}^{2n-1}\setminus 0$ (Theorem 4.2 of Paper I). We denote this map by Σ. The adjoint K^* maps singularities according to the inverse Σ^{-1}. Furthermore, for given (z, s, r, σ, ρ), we have a mapping $\tau\mapsto\zeta = \Theta(z, s, r, \sigma, \rho, \tau)$ with inverse $\tau = \Theta^{-1}(z, s, r, \zeta, \sigma, \rho)$ (Lemma 4.1 of Paper I).

Third, we analyze the normal operator derived from F_{D}. In the downward continuation version of (20), the operator on the left hand side is given by $F^*_{\mathrm{D}}\psi_YF$. Here $\psi_Y = \psi_Y(s, r, t)$ is a smooth cutoff function on Y that is

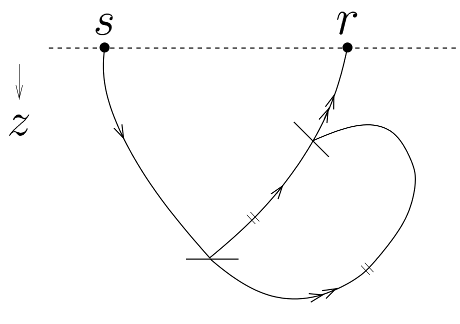

Fig. 1. Cartoon of two ray pairs such that Assumption 3 is violated. The two trajectories between the reflection points are assumed to have equal traveltime.

zero near the boundary of Y, that we introduced to account for the limited acquisition aperture. Suppose now that K' is defined as K, but with angles for the cutoff (contained in $G_{-,s}$, $G_{-,r}$) given by θ_1', θ_2' instead of θ_1, θ_2, such that $0 < \theta_1 < \theta_2 < \theta_1' < \theta_2' < \pi/2$. Then K is supported in the region where the cutoff that forms part of K' is equal to 1. Define an operator F_D' by $F_{\mathrm{D}}' = D_t^2 K' E_1$ (cf. (16)–(18)). Then F_{D} can be written as $F_D = \psi_D F_D'$, modulo a regularizing operator. Hence

$$F_{\mathrm{D}}^* \psi_Y F = F_{\mathrm{D}}' * \psi_{\mathrm{D}}^* \psi_Y F_{\mathrm{D}}' \left(\frac{1}{2} c_0^{-3} \cdot\right), \tag{31}$$

modulo a regularizing term. Thus, we can compute the symbol of $F^*{}_{\mathrm{D}} \psi_Y F$ using the downward continuation expression for F. Substituting (16), we have that

$$F_{\mathrm{D}}^* \psi_Y F_{\mathrm{D}} = R_1 \; K^* D_t^2 \psi_Y D_t^2 K \; E_1. \tag{32}$$

The first step in our computation of the symbol of this operator will be to evaluate $K^* D_t^2 \psi_Y D_t^2 K$. To simplify the evaluation, we introduce intermediate operators $\bar{L}$ and $\bar{K}$ that correspond with operators L and K, respectively, with the Q_- operators removed,

$$\bar{K} = \bar{L} E_2 \tag{33}$$

with

$$\bar{L} g = \int_0^Z H(0,z) g(z,\cdot,\cdot,\cdot) \; \mathrm{d}z. \tag{34}$$

We note, however, that $Q_{-,s}$ and $Q_{-,r}$ do not commute with E_2.

Lemma 2.1. *The composition $K^* D_t^2 \psi_Y D_t^2 K$ is a pseudodifferential operator of order 0 with principal symbol $\Sigma^*(\psi_Y |\psi_D|^2)(z, s, r, \zeta, \sigma, \rho)\; \Xi(z, s, r, \zeta, \sigma, \rho)$, where Ξ is given by*

$$\Xi(z,s,r,\zeta,\sigma,\rho) = (\Sigma^* a_0)(z,s,r,\zeta,\sigma,\rho) \times a_z(s,r,\sigma,\rho,\Theta^{-1}(z,s,r,\zeta,\sigma,\rho))\bar{\Xi}(z,s,r,\zeta,\sigma,\rho) \tag{35}$$

where

$$a_z(s,r,\sigma,\rho,\tau) = \tau^2 |b(z,s,\sigma,\tau)|^{-1} |b(z,r,\rho,\tau)|^{-1} \tag{36}$$

and

$$\begin{aligned}\bar{\Xi}(z,s,r,\zeta,\sigma,\rho)^{-1} &= \left|\frac{\partial\Theta}{\partial\tau}(z,s,r,\sigma,\rho,\Theta^{-1}(z,s,r,\zeta,\sigma,\rho))\right| \\ &= \left[c_0(z,s)^{-2}(c_0(z,s)^{-2} - \tau^{-2}||\sigma||^2)^{-1/2} + c_0(z,r)^{-2}(c_0(z,r)^{-2} - \tau^{-2}||\rho||^2)^{-1/2}\right]_{\tau=\Theta^{-1}(z,s,r,\zeta,\sigma,\rho)}.\end{aligned} \tag{37}$$

Proof. From the fact that the canonical relation of K is invertible (the graph of a diffeomorphism denoted by Σ before) between subsets of $T^*\mathbb{R}^{2n-1} \setminus 0$, it follows that $K^* \psi_Y K$ is a pseudodifferential operator. We calculate the principal symbol of $K^* D_t^2 \psi_Y D_t^2 K$.

First, we evaluate the symbol of $\bar{K}^* \bar{K}$ microlocally, ignoring the cutoffs; its principal part we denote by $\bar{\Xi}(z,s,r,\zeta,\sigma,\rho)$. We recall that the kernel of operator $H(0,z)$ has microlocally an oscillatory integral representation with amplitude $A = A(z, y_0, \eta_{0J}, s, r, t)$ and a phase function associated with generating function, $S = S(z, y_{0I}, \eta_{0J}, s, r, t)$, where $y_0 = (s_0, r_0, t_0)$ and η_0 is the corresponding cotangent vector, and $\{I, J\}$ is a partition of $\{1, \ldots, 2n-1\}$ (Lemma 3.1 of Paper I). The kernel of operator $\bar{K}$ has an oscillatory integral representation similar to the one of $H(0, z)$,

$$\bar{K}(y_0,z,s,r) = (2\pi)^{-(2n-1+|J|)/2} \int A(z,y_0,\eta_{0J},s,r,0) \mathrm{e}^{\mathrm{i}(S(z,y_{0I},\eta_{0J},s,r,0) + \langle \eta_{0J}, y_{0J}\rangle)} \; \mathrm{d}\eta_{0J} \tag{38}$$

which follows upon considering the action of $\bar{L}$ on $\delta(t) \cdot$ distribution (z, s, r) and carrying out the t integration. (The factors of 2π follow the convention used, for example, in [9].)

We evaluate, microlocally, $\bar{K}^*\bar{K}$. Using (38) and integrating out one set of η_{0J} variables, the kernel of $\bar{K}^*\bar{K}$ can be written as

$$(2\pi)^{-(2n-1)} \int \overline{A(z', y_0, \eta_{0J}, s', r', 0)} A(z, y_0, \eta_{0J}, s, r, 0) \exp(\mathrm{i}[-S(z', y_{0I}, \eta_{0J}, s', r', 0) + S(z, y_{0I}, \eta_{0J}, s, r, 0)]) \, \mathrm{d}\eta_{0J} \, \mathrm{d}y_{0I} \tag{39}$$

Next we identify the gradient

$$-\frac{\partial S}{\partial(z, s, r)}(z, y_{0I}, \eta_{0J}, s, r, 0) = (\zeta(z, y_{0I}, \eta_{0J}, s, r, 0), \sigma(z, y_{0I}, \eta_{0J}, s, r, 0), \rho(z, y_{0I}, \eta_{0J}, s, r, 0)). \tag{40}$$

Applying a change of variables of integration, $(y_{0I}, \eta_{0J}) \mapsto (\zeta, \sigma, \rho)$, the phase in the oscillatory integral representation of the kernel of $\bar{K}^*\bar{K}$ takes the form

$$\langle (\zeta, \sigma, \rho), (z' - z, s' - s, r' - r) \rangle; \tag{41}$$

a Jacobian, $|\frac{\partial(\zeta,\sigma,\rho)}{\partial(y_{0I},\eta_{0J})}|^{-1}$, appears in the amplitude, so that the amplitude of $\bar{K}^*\bar{K}$ has principal part (using the expression for A from Lemma 3.1 of Paper I)

$$\left|\frac{\partial(\zeta, \sigma, \rho)}{\partial(y_{0I}, \eta_{0J})}\right|^{-1} \left|\frac{\partial(\sigma, \rho, \tau)}{\partial(y_{0I}, \eta_{0J})}\right|_{t=0}. \tag{42}$$

For fixed (z, s, r, σ, ρ) the map $\tau \mapsto \zeta = \Theta(z, s, r, \sigma, \rho, \tau)$ is invertible on a set given by $|\tau|$ sufficiently large: $\tau = \Theta^{-1}(z, s, r, \zeta, \sigma, \rho)$ (Lemma 4.1 of Paper I). Carrying out the multiplication of determinants, it follows that the principal part of the symbol of $\bar{K}^*\bar{K}$ is microlocally given by

$$\bar{\Xi}(z, s, r, \zeta, \sigma, \rho) = \left|\frac{\partial\Theta}{\partial\tau}(z, s, r, \sigma, \rho, \Theta^{-1}(z, s, r, \zeta, \sigma, \rho))\right|^{-1}. \tag{43}$$

To obtain $K^* D_t^2 \psi_Y D_t^2 K$ involves the composition of $\bar{K}^*\bar{K}$ with pseudodifferential operators. First, the composition

$$\bar{K}^* Q_{-,s}(0) Q_{-,r}(0) D_t^2 \psi_Y D_t^2 \; Q^*_{-,s}(0) Q^*_{-,r}(0) \bar{K} \tag{44}$$

is carried out with the aid of Egorov's theorem. This leads to a factor $\Sigma^*(\tau^2 \psi_Y a_0)$ in the principal symbol (for the definition of Σ, see the text below (30)). Secondly, having obtained a pseudodifferential operator (in (44)), the inclusion of the factors $Q^*_{-,s}(z), Q^*_{-,r}(z)$ and $Q_{-,s}(z)$, $Q_{-,r}(z)$ – in between $\bar{L}$ and E_2 – is carried out with the standard calculus of pseudodifferential operators. This leads to a factor $(\Theta^{-1}(z, s, r, \zeta, \sigma, \rho))^{-2}$ $a_z(s, r, \sigma, \rho, \Theta^{-1}(z, s, r, \zeta, \sigma, \rho))$ in the principal symbol. It follows that the principal part of the symbol of $K^* D_t^2 \psi_Y D_t^2 K$ is, microlocally, given by $\Xi\Sigma^* \, \psi_Y$. With the pseudodifferential cutoff ψ_{D} taken into account we obtain the result of the lemma. □

We now determine the symbol of the normal operator in (32) and develop the downward continuation analogue of (20). To this end, we define an operator Ψ using the operator K' introduced above (31), by $\Psi = K' * D_t^2 \psi_{\mathrm{D}}^* \psi_Y D_t^2 K'$. It follows from the lemma that Ψ is pseudodifferential in $S^0_\rho(\mathbb{R}^{2n-1} \times \mathbb{R}^{2n-1})$ (for some ρ, $\frac{1}{2} < \rho < 1$, see Section 2 of Paper I) with principal symbol, that we denote by Ψ_0, given by

$$\Psi_0(z, s, r, \zeta, \sigma, \rho) = \Xi(z, s, r, \zeta, \sigma, \rho)(\Sigma^*(\overline{\psi_{\mathrm{D}}}\psi_Y))(z, s, r, \zeta, \sigma, \rho), \tag{45}$$

in which the second factor reveals the illumination. Starting from the DSR Born modeling (16), we obtain the following reconstruction result.

Theorem 2.2. *Suppose Assumption 3 is satisfied. Then there is a pseudodifferential operator* $\Phi = \Phi(z, x, D_z, D_x)$ *of order* $n - 1$, $\Phi \in S^{n-1}_\rho$ $\left(\frac{1}{2} < \rho < 1\right)$, *with principal symbol*

$$\Phi_{n-1}(z, x, \zeta, \xi) = \int_{\mathbb{R}^{n-1}} \Psi_0\left(z, x, x, \zeta, \tfrac{1}{2}\xi - \theta, \frac{1}{2}\xi + \theta\right) \, \mathrm{d}\theta, \tag{46}$$

such that

$$\Phi(z,x,D_z,D_x)\left(\tfrac{1}{2}c_0^{-3}\delta c\right) = R_1 K^* D_t^2 \psi_Y d, \tag{47}$$

where $d = F\delta c$ is the Born modeled data (16).

Proof. Because of Theorem 1.1 and the modeling formula (16), the right-hand side of (47) is equal to a pseudodifferential operator acting on $\frac{1}{2}c_0^{-3}\delta c$,

$$R_1 K^* D_t^2 \psi_Y d = R_1 \Psi E_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right).$$

It remains therefore to compute the symbol of the composition of operators $R_1\Psi E_1$, cf. (31). The kernel of this composition has an oscillatory integral representation

$$\begin{aligned}&(2\pi)^{-(2n-1)}\int_{\mathbb{R}^{2n-1}} \Psi(z,x,x,\zeta,\sigma,\rho)\ \mathrm{e}^{\mathrm{i}\langle (z,x,x)-(z',x',x'),(\zeta,\sigma,\rho)\rangle}\ \mathrm{d}\rho\mathrm{d}\sigma\mathrm{d}\zeta \\ &\quad = (2\pi)^{-(2n-1)}\int\int_{\mathbb{R}^n}\int_{\mathbb{R}^{n-1}} \Psi(z,x,x,\zeta,\tfrac{1}{2}\xi-\theta,\tfrac{1}{2}\xi+\theta)\ \mathrm{d}\theta\ \mathrm{e}^{\mathrm{i}((z-z')\zeta+\langle (x-x'),\xi\rangle)}\ \mathrm{d}\xi\mathrm{d}\zeta,\end{aligned} \tag{48}$$

upon changing variables of integration, $\sigma = \frac{1}{2}\xi - \theta$, $\rho = \frac{1}{2}\xi + \theta$. The support of the integrand in the inner integral is bounded by an inequality of the form $|\theta| \leqslant C(|\zeta| + 1)$, due to the cutoff ψ_{D}. Using this we next show that the inner integral is a symbol in $S_\rho^{n-1}(\mathbb{R}^n\times\mathbb{R}^n)$. We have the inequality

$$\begin{aligned}&\frac{\partial^\alpha}{\partial(z,x)^\alpha}\frac{\partial^\beta}{\partial(\zeta,\xi)^\beta}\int_{\mathbb{R}^{n-1}} \Psi\left(z,x,x,\zeta,\tfrac{1}{2}\xi-\theta,\tfrac{1}{2}\xi+\theta\right)\ \mathrm{d}\theta \\ &\quad \leqslant C'(|\zeta|+1)^{n-1}\max_{|\theta|<C(|\zeta|+1)}\frac{\partial^\alpha}{\partial(z,x)^\alpha}\frac{\partial^\beta}{\partial(\zeta,\xi)^\beta}\Psi\left(z,x,x,\zeta,\tfrac{1}{2}\xi-\theta,\tfrac{1}{2}\xi+\theta\right),\end{aligned} \tag{49}$$

so that the symbol estimates follow from the symbol estimates for Ψ.

From this symbol property it follows that $R_1\Psi E_1$ is a pseudodifferential operator of order $n-1$, with principal symbol (46).

Applying the above results to the expression (16) for the Born modeled data leads to the statement of the theorem. □

The symbol of $\frac{1}{2}c_0^{-3}F_{\mathrm{D}}^*\psi_Y F = \underline{F^*\psi_{\mathrm{D}}^*}\psi_Y F$ also follows from the computation of F^*F of Ten Kroode et al. [16, theorem 4.1], upon inserting $\overline{\psi}_{\mathrm{D}}(s,r,t,\sigma,\rho,\tau)\psi_Y(s,r,t)$ in their formula (64). Our purpose was to show that this computation is independent of the results of [16]. Also, our method is easily modified for the result (55) below.

The right-hand side of (47) produces an image of (the singularities of) δc; to obtain an estimate of δc requires the construction and application of the parametrix of Φ microlocally.

Remark 2.3. Note that in (47) the operator $\Phi(z, x, D_z, D_x)$ depends on the Hamiltonian flow associated with the background medium in the depth interval $[0, z]$. To account for the principal part of this operator, cf. (45) and (46), one requires a ray computation to determine Σ^* separate from the downward continuation of the data with K^*.

Remark 2.4. Theorem 2.2 follows a least-squares data fitting approach to linearized inverse scattering. However, departing from this approach, we can derive a reconstruction equation the pseudodifferential operator in which only accounts for illumination effects. By the calculus of Fourier integral operators, the operator K can be written as $K = Q_{-,s}(0)Q_{-,r}(0)\bar{K}V$ modulo a smoothing operator, where V is an operator with principal symbol

$$V(z,s,r,\zeta,\sigma,\rho) = |\tau|^{-1}(c_0(z,s)^{-2}-\tau^{-2}||\sigma||^2)^{-1/4}(c_0(z,r)^{-2}-\tau^{-2}||\rho||^2)^{-1/4}\Big|_{\tau=\Theta^{-1}(z,s,r,\zeta,\sigma,\rho)}$$

We replace the normal equation in the downward continuation approach by the following equation,

$$\bar{K}^*\bar{K}VE_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right) = \bar{K}^*Q_{-,s}^*(0)^{-1}Q_{-,r}^*(0)^{-1}D_t^{-2}d. \tag{50}$$

Applying $V^{-1}\bar{\Xi}^{-1}$, we obtain the equation

$$V^{-1}\bar{\Xi}^{-1}\bar{K}^*\bar{K}V\ E_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right) = V^{-1}\bar{\Xi}^{-1}\bar{K}^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}D_t^{-2}d, \tag{51}$$

microlocally, which reduces by (38)–(43) in the proof of Lemma 2.1 to

$$\bar{\Psi}\ E_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right) = V^{-1}\bar{\Xi}^{-1}\bar{K}^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}D_t^{-2}\psi_Y d \tag{52}$$

in which $\bar{\Psi}$ is a pseudodifferential operator with principal symbol $\bar{\Psi}_0 = \Sigma^*(\overline{\psi_{\mathrm{D}}}\psi_Y)$ (compare (45)). Therefore

$$R_1\ \bar{\Psi}\ E_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right) = R_1V^{-1}\bar{\Xi}^{-1}\bar{K}^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}D_t^{-2}\psi_Y d. \tag{53}$$

Substituting $\bar{\Psi}$ for Ψ in (46) then leads to the reconstruction equation

$$\bar{\Phi}(z,x,D_z,D_x)\left(\frac{1}{2}c_0^{-3}\delta c\right) = R_1V^{-1}\bar{\Xi}^{-1}\bar{K}^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}D_t^{-2}\psi_Y d \tag{54}$$

with

$$\bar{\Phi}(z,x,\zeta,\xi) = \int_{\mathbb{R}^{n-1}} \bar{\Psi}(z,x,x,\zeta,\tfrac{1}{2}\xi-\theta,\tfrac{1}{2}\xi+\theta)\ \mathrm{d}\theta. \tag{55}$$

The principal symbol of $\bar{\Phi}(z,x,D_z,D_x)$ can in general be zero for some parts of phase space, due to aperture effects (see e.g. chapter 4 of [2]) so that a regularized inverse is required for the final reconstruction of $\frac{1}{2}c_0^{-3}\delta c$ (microlocally where $\bar{\Phi}(z,x,D_z,D_x)$ is invertible). The symbol $\bar{\Psi}$ contains a factor $\bar{\Xi}$, which is simpler than the factor Ξ contained in Ψ, because of the absence of the pullback Σ^*a_0, compare (35) and (37). So the expression for $\bar{\Phi}$ is simpler than the expression for Φ, and therefore we prefer this equation for the development of a practical algorithm.

Remark 2.5. Depending on the background medium, the reconstruction can also be done using data on a submanifold Y' of Y. Let R' be the restriction of a function on Y to Y', so that the forward map for this case is given by $R'F$. In suitable local coordinates (y', y'') on Y such that $y''=0$ defines Y', the adjoint E' of R' is given by the map $(E'f)(y', y'')=f(y')\delta(y'')$. Conditions such that $F^*E'\psi'_YR'F$ is pseudodifferential are given in [17]. Reconstruction modulo a pseudodifferential operator is done in this case by first applying the map E' to the data, and then applying the previous procedure. Applying E' to the data simply means adding zeroes where there is no data in Y.

3. The wave-equation angle transform, common-image-point gathers

A method to generate multiple images of the medium contrast from the data follows from beamforming, here denoted by R_3, applied to the downward continued data, g say,

$$R_3 : g(z,s,r,t)\mapsto(R_3g)(z,x,p) = \int_{\mathbb{R}^{n-1}} g\left(z,x-\tfrac{h}{2},x+\tfrac{h}{2},\langle p,h\rangle\right)\ \mathrm{d}h.$$

Indeed, with the aid of (22) and (23), identifying R_3 with the adjoint of E_3, the modeling operator can be written in the form $F_{\mathrm{D}} = D_t^2L\Lambda E_3E_4$ (cf. (25)). Following (partly) the process of imaging, that is, taking adjoints, we define the wave-equation angle transform, A_{WE}, containing a cutoff function χ, as

$$A_{\mathrm{WE}} = R_3\chi L^*D_t^2, \tag{56}$$

where the composition of χ and R_3 is

$$R_3\chi : g(z,s,r,t)\mapsto(R_3\chi g)(z,x,p) = \int_{\mathbb{R}^{n-1}} g(z,x-\tfrac{h}{2},x+\tfrac{h}{2},\langle p,h\rangle)\chi(z,x,h)\ \mathrm{d}h, \tag{57}$$

where $\chi(z, x, h)$ is a compactly supported cutoff function the support of which contains $h=0$. For each x, the function $(z,p)\mapsto(A_{\mathrm{WE}}\psi_Y d)(z,x,p)$ is a so-called common-image-point gather.

The work in this section was motivated by the papers [6] and [20], where a map similar to A_{WE} was introduced. In the paper [6] the purpose was imaging of reflectors with angle-dependent reflection coefficients. The paper [20] suggested that this map could be suitable to obtain common-image point gathers in the presence of multipathing, which is indeed what we show here.

First, we analyze the properties of A_{WE} mostly from a geometrical perspective. Then we derive a p-dependent reconstruction equation that replaces (47) or (54).

Theorem 3.1. *Suppose Assumption 1 holds. Let C_0 be an upper bound for c_0. Assume that*

$$\|p\| < p_{\max} < C_0^{-1}. \tag{58}$$

Then A_{WE} is a Fourier integral operator such that $A_{\mathrm{WE}}\,F$ is a smooth p-family of pseudodifferential operators in (z, x). Let C_1 be an upper bound for $\frac{\partial c_0^{-2}}{\partial x}$, C_2 an upper bound for c_0^{-1}. If in addition the function $h \mapsto \chi(z, x, h)$ of (56), contained in A_{WE}, is supported in $B(0, R)$, where R depends on θ_2, C_0, C_1, C_2, then the canonical relation of A_{WE} corresponds to an invertible map from a subset of $T^*\mathbb{R}^{2n-1}_{(s,r,t)}$ to a subset of $T^*\mathbb{R}^{2n-1}_{(z,x,p)}$ that has nonempty intersection with the set $\vartheta = 0$ (where ϑ denotes the p-covector).

Proof. By Assumption 1 we have that $F = F_{\mathrm{D}}(\frac{1}{2}c_0^{-3}\cdot)$, modulo a regularizing term. So for the first statement it is sufficient to show that $A_{\mathrm{WE}}F_{\mathrm{D}}$ is a p-family of pseudodifferential operators (pseudodifferential operators depending on a parameter p).

The Schwartz kernel of the map $R_3\chi$ equals

$$\begin{aligned}&\delta\Big(x - \frac{s+r}{2}\Big)\ \delta(\langle p, r-s\rangle - t)\ \delta(z-z')\ \chi(z,x,r-s)\\ &\quad = (2\pi)^{-n-1}\int \chi(z,x,r-s)\ \mathrm{e}^{\mathrm{i}(\langle\xi, x-\frac{s+r}{2}\rangle + \tau(\langle p, r-s\rangle - t) + \zeta(z-z'))}\ \mathrm{d}\xi\ \mathrm{d}\tau\ \mathrm{d}\zeta.\end{aligned} \tag{59}$$

It defines a Fourier integral operator with canonical relation

$$\begin{aligned}&\Big\{\Big(z, \frac{s+r}{2}, p, \zeta, \xi, (r-s)\tau; z, s, r, \langle p, r-s\rangle, \zeta, \frac{\xi}{2} + p\tau, \frac{\xi}{2} - p\tau, \tau\Big)\ |(z,s,r,p,\zeta,\xi,\tau) \in \text{subset of } \mathbb{R}^{4n-1}\Big\}\\ &\quad \subset T^*\mathbb{R}^{2n-1}_{(z,x,p)} \setminus 0 \times T^*\mathbb{R}^{2n}_{(z,s,r,t)} \setminus 0.\end{aligned} \tag{60}$$

Using the change of variables, $\xi = \sigma + \rho$, $p = \frac{\sigma-\rho}{2\tau}$, this canonical relation can be parameterized by the coordinates of $T^*\mathbb{R}^{2n}_{(z,s,r,t)} \setminus 0$ except t, that is $(z, s, r, \zeta, \sigma, \rho, \tau)$

$$\begin{aligned}&\Big\{\Big(z, \frac{s+r}{2}, \frac{\sigma-\rho}{2\tau}, \zeta, \sigma+\rho, (r-s)\tau; z, s, r, \Big\langle\frac{\sigma-\rho}{2\tau}, r-s\Big\rangle, \zeta, \sigma, \rho, \tau\Big)\ |(z,s,r,\zeta,\sigma,\rho,\tau) \in \text{ subset of } \mathbb{R}^{4n-1}\Big\}\\ &\quad \subset T^*\mathbb{R}^{2n-1}_{(z,x,p)} \setminus 0 \times T^*\mathbb{R}^{2n}_{(z,s,r,t)} \setminus 0.\end{aligned} \tag{61}$$

The projection of this canonical relation on $T^*\mathbb{R}^{2n}_{(z,s,r,t)} \setminus 0$ is a hypersurface defined by

$$t = \Big\langle\frac{\sigma-\rho}{2\tau}, (r-s)\Big\rangle. \tag{62}$$

The map $d \mapsto L^*\psi_Y d$ is a Fourier integral operator with canonical relation

$$\begin{aligned}&\{(z, \gamma_x(z,0,s_0,\sigma_0,\tau), \gamma_x(z,0,r_0,\rho_0,\tau), t_0 + \gamma_t(z,0,s_0,\sigma_0,\tau) + \gamma_t(z,0,r_0,\rho_0,\tau),\\ &-b(z,s_0,\sigma_0,\tau) - b(z,r_0,\rho_0,\tau), \gamma_\xi(z,0,s_0,\sigma_0,\tau), \gamma_\xi(z,0,r_0,\rho_0,\tau), \tau; s_0, r_0, t_0, \sigma_0, \rho_0, \tau)\ |\\ &(s_0,r_0,t_0,\sigma_0,\rho_0,\tau) \in T^*\mathbb{R}^{2n-1} \setminus 0, z \in \mathbb{R}_+\} \subset T^*\mathbb{R}^{2n}_{(z,s,r,t)} \times T^*\mathbb{R}^{2n-1}_{(s_0,r_0,t_0)}\end{aligned} \tag{63}$$

derived from the canonical relation of $H(0,z)^*$, cf. (10). This canonical relation is parameterized by $(z, s_0, r_0, t_0, \sigma_0, \rho_0, \tau)$, and is time translation invariant (in t_0). The line in the projection of this canonical relation on $T^*\mathbb{R}^{2n}_{(z,s,r,t)} \setminus 0$ parameterized by t_0 for fixed $(s_0, r_0, \sigma_0, \rho_0, \tau)$ intersects the hypersurface (62) transversally. It follows that the composition of the canonical relations (63) and (60) is transversal. This composition is parameterized by $(z, s_0, r_0, \sigma_0, \rho_0, \tau)$. It follows that $R_3\chi L^*$, and hence $A_{\mathrm{WE}} = R_3\chi L^* D_t^2$, are Fourier integral operators. Let $p(z)$ denote the value of $\frac{\sigma-\rho}{2\tau}$ along a certain DSR bicharacteristic with initial values

$(s_0, r_0, t_0, \sigma_0, \rho_0, \tau)$, and let $t(z)$ denote the time and $h(z)$ denote the value of $r - s$. The elements of the canonical relation of A_{WE} correspond to solutions $(z, s_0, r_0, \sigma_0, \rho_0, \tau)$ of $t(z) - \langle p(z), h(z)\rangle = 0$.

To analyze the composition $A_{\mathrm{WE}}F_{\mathrm{D}}$, we first analyze the composition $L^*D_t^4\psi_Y LE_2E_1$, that maps the perturbation $\frac{1}{2}c_0^{-3}\delta c = (\frac{1}{2}c_0^{-3}\delta c)(z', x)$ to the downward continued data as a function of (z, s, r, t). This composition is a Fourier integral operator with canonical relation

$$\begin{aligned}&\{(z, \gamma_x(z,z',x,\sigma',\tau), \gamma_x(z,z',x,\rho',\tau), \gamma_t(z,z',x,\sigma',\tau) + \gamma_t(z,z',x,\rho',\tau),\\ &\zeta, \gamma_\xi(z,z',x,\sigma',\tau), \gamma_\xi(z,z',x,\rho',\tau), \tau; z', x, \zeta', \sigma' + \rho')|\\ &(z,z',x,\zeta',\sigma',\rho') \in \text{ a subset of } \mathbb{R}^{3n}, \tau \text{ such that } \zeta' = \Theta(z', x, x, \sigma', \rho', \tau), \zeta\\ &= \Theta(z, \gamma_x(z,z',x,\sigma',\tau), \gamma_x(z,z',x,\rho',\tau), \gamma_\xi(z,z',x,\sigma',\tau), \gamma_\xi(z,z',x,\rho',\tau), \tau)\}.\end{aligned} \tag{64}$$

Here, the propagation of singularities upward by L and downward by L^* is along the same DSR bicharacteristics.

We now show that the composition $A_{\mathrm{WE}}D_t^2LE_2E_1$ is a Fourier integral with canonical relation contained in

$$\{(z, x, p, \zeta, \xi, 0; z, x, \zeta, \xi) \mid (z, x, \zeta, \xi) \in T^*\mathbb{R}^n_{(z,x)} \setminus 0, \|p\| < p_{\max}\}. \tag{65}$$

From that it follows that $A_{\mathrm{WE}}F_{\mathrm{D}}$ is a p-family of pseudodifferential operators. Indeed, the projection of (64) on $T^*\mathbb{R}^{2n}_{(z,s,r,t)}$ intersects the hypersurface (62) at $z' = z$, since only then $r - s = 0$ and $t = 0$, leading to elements in (65), see Fig. 2. Since singularities propagate with speed less than C_0, if $(1, v_s, v_r, v_t, v_\zeta, v_\sigma, v_\rho, 0)$ derived from $\frac{\mathrm{d}\Gamma}{\mathrm{d}z}$ is a tangent vector to the DSR bicharacteristic Γ (cf. (10); the first component $v_z = 1$), then $\frac{\|v_r - v_s\|}{v_t} \leqslant C_0$, cf. (10). By integrating this inequality we find that $\|r - s\| \leqslant C_0|t|$ along Γ. On the other hand, for points $(z, x, p, \zeta, \xi, \theta; s, r, t, \sigma, \rho, \tau)$ in (60), one has that $|t| = |\langle p, r - s\rangle| \leqslant p_{\max}\|r - s\| < C_0^{-1}\|r - s\|$. From this, it follows that the composition of (64) with (60) contains no elements outside (65). It follows also that the composition of (64) with (60) is transversal.

Finally, we show that A_{WE} is invertible. To this end, we investigate the projections of the canonical relation of A_{WE} on the first component $T^*\mathbb{R}^{2n-1}_{(z,x,p)} \setminus 0$ and on the second component $T^*\mathbb{R}^{2n-1}_{(s,r,t)} \setminus 0$, respectively. The projection of the canonical relation of A_{WE} on the second component $T^*\mathbb{R}^{2n-1}_{(s,r,t)} \setminus 0$ is invertible if each DSR bicharacteristic with initial values $(s_0, r_0, t_0, \sigma_0, \rho_0, \tau)$, parameterized by z, intersects the hypersurface (62) at most once and transversally. We reconsider the equation, $t(z) - \langle p(z), h(z)\rangle = 0$, defining the canonical relation of A_{WE}. To estimate the derivative of the left-hand side, we will argue that

$$\frac{\partial \boldsymbol{t}}{\partial z} - \left\langle \boldsymbol{p}(z), \frac{\partial \boldsymbol{h}}{\partial z}(z)\right\rangle < -\epsilon_0 \tag{66}$$

for some $\epsilon_0 > 0$ depending on $p_{\max}$ and R as in the theorem. Indeed, note that the factor

$$1 - \left\langle \boldsymbol{p}(z), \frac{\partial \boldsymbol{h}}{\partial z}(z) \middle/ \frac{\partial \boldsymbol{t}}{\partial z}\right\rangle$$

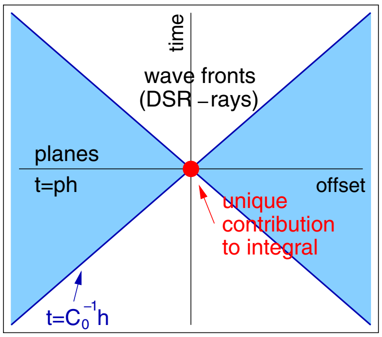

Fig. 2. Contribution to the kernel of $A_{\mathrm{WE}}D_t^2LE_2E_1$ (here, $n = 2$; $h = r - s$ is offset).

is bounded away from zero. Also, note that $\frac{\partial t}{\partial z}$ is strictly negative while $|\frac{\partial t}{\partial z}| \geqslant \frac{2}{C_0}$. Since according to the Hamilton equations,

$$\frac{\mathrm{d}\gamma_\xi}{\mathrm{d}z} = \frac{\partial b}{\partial x} = -\frac{\tau}{2\sqrt{c_0^{-2} - \tau^{-2}\|\xi\|^2}} \frac{\partial c_0^{-2}}{\partial x}, \tag{67}$$

it follows that $\|\frac{\partial p}{\partial z}(z)\| \leqslant \frac{C_2}{2}\frac{C_1}{\cos(\theta_2)}$. So, if $\|h\| < C_3\frac{\cos(\theta_2)}{C_0 C_1 C_2}$, it follows that

$$\left|\left\langle \frac{\partial \boldsymbol{p}}{\partial z}(z), \boldsymbol{h}(z)\right\rangle\right| < C_3\frac{1}{2C_0} < \epsilon_0 \tag{68}$$

by an appropriate choice of C_3. This implies that the function $z \mapsto t(z) - \langle p(z), h(z)\rangle$ is monotone. Hence the projection of the canonical relation of A_{WE} on $T^*\mathbb{R}^{2n-1}_{(s,r,t)} \setminus 0$ is invertible.

From canonical relation (60) we now establish that, given $z, \frac{s+r}{2} = x, \sigma, \rho$, the map $(h = s - r, \tau) \mapsto (\vartheta = (r - s)\tau,\ \zeta = \Theta(z, s, r, \sigma, \rho, \tau))$ is invertible for h sufficiently small. Assume that (h_1, τ_1) maps to (ϑ_1, ζ_1) and that (h_2, τ_2) maps to (ϑ_2, ζ_2) with $\vartheta_2 = \vartheta_1$. We show that $\zeta_1 > \zeta_2$ if $\tau_1 > \tau_2$. Estimating the difference

$$\Theta\left(z, x + \tfrac{1}{2}h_1, x - \tfrac{1}{2}h_1, \sigma, \rho, \tau_1\right) - \Theta\left(z, x + \tfrac{1}{2}h_2, x - \tfrac{1}{2}h_2, \sigma, \rho, \tau_2\right), \tag{69}$$

using the bounds

$$\left|\frac{\partial b}{\partial \tau}\right| = \left|\frac{c_0^{-2}}{\sqrt{c_0^{-2} - \tau^{-2}\|\xi\|^2}}\right| \geqslant \frac{C_5}{C_4}, \tag{70}$$

in which C_4 depends on the lower and upper bounds of c_0, and $C_5 = \frac{1}{C_2^2}$ is a lower bound for c_0^{-2}, and (cf. (67))

$$\left\|\frac{\partial b}{\partial x}\right\| = \left\|\frac{\tau}{2\sqrt{c_0^{-2} - \tau^{-2}\|\xi\|^2}} \frac{\partial c_0^{-2}}{\partial x}\right\| \leqslant \frac{\tau\left\|\frac{\partial c_0^{-2}}{\partial x}\right\|_{L^\infty}}{C_6}, \tag{71}$$

where C_6 depends on θ_2 and the upper and lower bounds of c_0, yields the estimate

$$\zeta_1 - \zeta_2 \geqslant \frac{C_5(\tau_1 - \tau_2)}{C_4} - \frac{\left\|\frac{\partial c_0^{-2}}{\partial x}\right\|_{L^\infty}\tau_1\|h_1 - h_2\|}{C_6}. \tag{72}$$

Since $\vartheta_2 = \vartheta_1$, $(h_1 - h_2)\tau_1 = h_2(\tau_2 - \tau_1)$, so that

$$\frac{C_5(\tau_1 - \tau_2)}{C_4} - \frac{\left\|\frac{\partial c_0^{-2}}{\partial x}\right\|_{L^\infty}\tau_1\|h_1 - h_2\|}{C_6} = (\tau_1 - \tau_2)\left[\frac{C_5}{C_4} - \frac{\left\|\frac{\partial c_0^{-2}}{\partial x}\right\|_{L^\infty}}{C_6}\|h_2\|\right]. \tag{73}$$

The right-hand side of this equality is strictly greater than zero for $\|h_2\|$ sufficiently small, since $\tau_1 - \tau_2 > 0$. It follows from this that $\zeta_1 > \zeta_2$. We conclude that the projection of the canonical relation of A_{WE} on $T^*\mathbb{R}^{2n-1}_{(z,x,p)} \setminus 0$ is invertible as well. It follows also that the linearization of this projection is invertible. This establishes the last statement of the theorem. □

To conclude this section we determine, at the principal symbol level, the modification of (56) that leads to reconstruction of singularities of δc microlocally for each p subject to $\|p\| < p_{\max}$. Like the reconstruction in (47) of Theorem 2.2, the reconstruction is microlocal. The three cutoffs, χ and ψ_Y, ψ_{D}, must be taken into account. The canonical relation of A_{WE} defines a map $(s, r, t, \sigma, \rho, \tau) \to (z, x, p, \zeta, \xi, \vartheta)$ (where ϑ is the p-covector); there is also an associated value of $h = r - s$ through (60). By pull back with the inverse of the mentioned map, one can map the symbols ψ_Y, ψ_{D} to symbols in the variables $(z, x, p, \zeta, \xi, \vartheta)$. By the mentioned evaluation of h one obtains by pull back the cutoff χ in these variables also. We define $\Psi_{\mathrm{WE}} = \Psi_{\mathrm{WE}}(z, x, p, \zeta, \xi, \vartheta)$ as the product of these symbols. With this definition we have

Proposition 3.2. *Define* $\widetilde{A}_{\mathrm{WE}}$ *by*

$$(\widetilde{A}_{\mathrm{WE}}d)(z,x,p) = \jmath^{-1}R_3\bar{\Xi}^{-1}Q_{-,s}(z)^{-1}Q_{-,r}(z)^{-1}\bar{L}^*Q^*_{-,s}(0)^{-1}Q^*_{-,r}(0)^{-1}D_t^{-2}\psi_Y d, \tag{74}$$

in which $\bar{\Xi} = \bar{\Xi}(z,s,r,\Theta(z,s,r,\sigma,\rho,\tau),\sigma,\rho)$ *is independent of* t (*the corresponding operator is convolutional in time*), *and where the symbol of* $\jmath$ *is defined in* (86) *and* (87) *below. Suppose Assumption* 1 *holds. Suppose that* χ *is* 1 *on a neighborhood of* $h = 0$ *and* $h \mapsto \chi(z, x, h)$ *is supported in* $B(0, R)$ (*cf. Theorem* 3.1), *then* $\widetilde{A}_{\mathrm{WE}}$ *is an invertible Fourier integral operator. Let the symbol* Ψ_{WE} *be as defined above. The composition* $\widetilde{A}_{\mathrm{WE}}F$ *is a p-family of pseudodifferential operators with principal symbol* $\Psi_{\mathrm{WE}}(z, x, p, \zeta, \xi, 0)$.

The microlocal reconstruction hence follows from

$$(\Psi_{\mathrm{WE}}(z,x,p,D_z,D_x,0) + \mathrm{order}(-1))\left(\tfrac{1}{2}c_0^{-3}\delta c\right) = \widetilde{A}_{\mathrm{WE}}d. \tag{75}$$

We note that Ψ_{WE} accounts purely for illumination and cannot be compensated for.

Proof. We consider the operator $R_3\bar{\Xi}^{-1}\bar{L}^*\bar{L}E_2E_1$ or the map

$$\left(\tfrac{1}{2}c_0^{-3}\delta c\right)\mapsto \int \left(\bar{\Xi}(z)^{-1}H(0,z)^* \int H(0,z')E_2E_1\left(\tfrac{1}{2}c_0^{-3}\delta c\right)\mathrm{d}z'\right)\left(x-\tfrac{h}{2},x+\tfrac{h}{2},ph\right)\mathrm{d}h, \tag{76}$$

and evaluate, microlocally, its principal symbol. This principal symbol will be the principal symbol of operator $\jmath$ in the theorem. In this proof, we will omit the cutoff functions that are part of the symbols; the calculations will be valid microlocally on the support of a cutoff.

Using an oscillatory integral representation of H similar to the one in the proof of Theorem 2.2, we find that the principal contribution to the kernel of this map, as a function of $(z, x, p;\ z', x')$, can be written as

$$\begin{aligned}(2\pi)^{-(2n-1)}\int \bar{\Xi}\left(z,x-\frac{1}{2}h,x+\frac{1}{2}h,\Theta\left(z,x-\tfrac{1}{2}h,x+\tfrac{1}{2}h,-\frac{\partial S}{\partial s},-\frac{\partial S}{\partial r},-\frac{\partial S}{\partial t}\right),-\frac{\partial S}{\partial s},-\frac{\partial S}{\partial r}\right)^{-1}\\ \times \overline{A(z,x-\tfrac{1}{2}h,x+\tfrac{1}{2}h,ph,y_0,\eta_{0J})}A(z',x',x',0,y_0,\eta_{0J}) \times \mathrm{e}^{\mathrm{i}[\Phi_R(z,x,p,z',x',y_{0I},\eta_{0J})]}\,\mathrm{d}y_{0I}\,\mathrm{d}\eta_{0J}\,\mathrm{d}h,\end{aligned} \tag{77}$$

in which $\frac{\partial S}{\partial s},\frac{\partial S}{\partial r},\frac{\partial S}{\partial t}$ are evaluated at $(z,x-\frac{1}{2}h,x+\frac{1}{2}h,ph,y_{0I},\eta_{0J})$, and where

$$\Phi_R(z,x,p,z',x',y_{0I},\eta_{0J}) = -S\left(z,x-\tfrac{1}{2}h,x+\frac{1}{2}h,ph,y_{0I},\eta_{0J}\right) + S(z',x',x',0,y_{0I},\eta_{0J}). \tag{78}$$

We expand this phase in a Taylor series about $(z', x', h) = (z, x, 0)$ and identify the gradient at $(z, x, 0)$,

$$-\frac{\partial\Phi_R}{\partial z}(z,x,x,0,y_{0I},\eta_{0J}) =\zeta(z,x,x,0,y_{0I},\eta_{0J}), \tag{79}$$

$$-\frac{\partial\Phi_R}{\partial x}(z,x,x,0,y_{0I},\eta_{0J}) =\sigma(z,x,x,0,y_{0I},\eta_{0J}) + \rho(z,x,x,0,y_{0I},\eta_{0J}), \tag{80}$$

$$-\frac{\partial\Phi_R}{\partial h}(z,x,x,0,y_{0I},\eta_{0J}) = -\tfrac{1}{2}\sigma(z,x,x,0,y_{0I},\eta_{0J}) + \tfrac{1}{2}\rho(z,x,x,0,y_{0I},\eta_{0J}) + p\tau(z,x,x,0,y_{0I},\eta_{0J}), \tag{81}$$

where

$$-\frac{\partial S}{\partial(s,r,t,z)}(z,x,x,0,y_{0I},\eta_{0J}) = (\sigma(z,x,x,0,y_{0I},\eta_{0J}),\rho(z,x,x,0,y_{0I},\eta_{0J}),\tau(z,x,x,0,y_{0I},\eta_{0J}),\zeta(z,x,x,0,y_{0I},\eta_{0J})). \tag{82}$$

Applying a change of variables, $(y_{0I}, \eta_{0J}) \mapsto (\zeta, \sigma, \rho)$, the phase takes the form

$$\zeta(z-z') + \langle\sigma+\rho,x-x'\rangle + \left\langle\tfrac{1}{2}(\rho-\sigma)+p\tau,h\right\rangle,\ \text{with}\ \tau = \Theta^{-1}(z,x,x,\zeta,\sigma,\rho). \tag{83}$$

The amplitude factor $\bar{\Xi}^{-1}\overline{A}A$ (at $h = 0$) becomes equal to one by the calculations in the proof of Theorem 2.2. Upon changing integration variables, $\sigma = \frac{1}{2}\xi - \vartheta, \rho = \frac{1}{2}\xi + \vartheta$, the oscillatory integral (77) takes the leading-order form

$$(2\pi)^{-(2n-1)}\int \mathrm{e}^{\mathrm{i}(\zeta(z-z')+\langle\xi,x-x'\rangle+\langle\vartheta+\tau p,h\rangle)}\,\mathrm{d}\zeta\,\mathrm{d}\xi\,\mathrm{d}\vartheta\,\mathrm{d}h,\ \tau=\Theta^{-1}\left(z,x,x,\zeta,\tfrac{1}{2}\xi-\vartheta,\tfrac{1}{2}\xi+\vartheta\right). \tag{84}$$

By the method of stationary phase (see e.g. [9, section 1.2]) the following integral is a symbol

$$(2\pi)^{-(n-1)}\int \mathrm{e}^{\mathrm{i}\langle\vartheta+\Theta^{-1}(z,x,x,\zeta,\frac{1}{2}\xi-\vartheta,\frac{1}{2}\xi+\vartheta)p,h\rangle}\,\mathrm{d}\vartheta\,\mathrm{d}h; \tag{85}$$

it follows from (84) that this equals to highest order the symbol of $\widetilde{A}_{\mathrm{WE}}F$. By the method of stationary phase the principal part of this symbol is given by

$$\jmath(z,x,p,\zeta,\xi):=\left|\frac{\partial(\langle\vartheta+\Theta^{-1}p,h\rangle)}{\partial(h,\vartheta)}\right|^{-\frac{1}{2}}, \tag{86}$$

evaluated where $h=0$ and ϑ is such that $\vartheta+\Theta^{-1}p=0$. By evaluating this, we find that

$$\left|\frac{\partial(\langle\vartheta+\Theta^{-1}p,h\rangle)}{\partial(h,\vartheta)}\right|^{\frac{1}{2}}=\left|\frac{\partial(\vartheta+\Theta^{-1}p)}{\partial\vartheta}\right|=\left|\det\left(I+p\otimes\frac{\partial\Theta^{-1}}{\partial\vartheta}\right)\right|=\left|\det\left(I+p\otimes\left(\frac{\partial\Theta^{-1}}{\partial\rho}-\frac{\partial\Theta^{-1}}{\partial\sigma}\right)\right)\right|, \tag{87}$$

evaluated at $s=r=x$ and $\sigma=\frac{1}{2}\xi-\vartheta,\rho=\frac{1}{2}\xi+\vartheta$. This completes the proof. □

Remark 3.3. Since $\widetilde{A}_{\mathrm{WE}}$ is invertible, microlocally, we have obtained the following diagram as suggested by Symes [31]

$$\begin{array}{ccc}\mathscr{E}'(X\times E) & \overset{\widetilde{A}_{\mathrm{WE}}^{-1}}{\longrightarrow} & \mathscr{D}'(Y)\\ \Psi_{\mathrm{WE}}\ \uparrow & & \uparrow\ \mathrm{Id}\\ \mathscr{E}'(X) & \overset{F_{\mathrm{D}}}{\rightarrow} & \mathscr{D}'(Y)\end{array} \tag{88}$$

where $E=\{p\in\mathbb{R}^{n-1}\mid\|p\|<p_{\max}\}$ as in Theorem 3.1 (cf. (58)). In this diagram – through the introduction of $\mathscr{E}'(X\times E)$ – the redundancy in the data, measured by dim $Y-$ dim X, is manifest.

4. Annihilators

As discussed in the introduction, the inverse problem of determining the background medium c_0 can be addressed by making use of the redundancy in the data. The background medium must be such that the data is in the range of the operator F. The data are in the range of F_{D}, if the angle transform generates a collection of identical images of $\frac{1}{2}c_0^{-3}\delta c$, parameterized by $p=(p_1,\ldots,p_{n-1})$, microlocally. It follows that we have the following criterion

$$(\widetilde{A}_{\mathrm{WE}}[c_0]d)(z,x,p)\text{ is independent of }p, \tag{89}$$

microlocally, or

$$\frac{\partial}{\partial p_i}(\widetilde{A}_{\mathrm{WE}}[c_0]d)(z,x,p)=0, \tag{90}$$

microlocally. Seismologists recognize this as "alignment" of the singularities in common-image-point gathers. Of course it must be taken into acount that $\widetilde{A}_{\mathrm{WE}}$ is only a microlocal inverse, so (90) is not valid globally.

The criterion (90) can be restated as that certain pseudodifferential operators annihilate the singular part of the data, see [26]. With the approach based on inversion for δc using subsets of data (discussed in the introduction) this is closely related to differential semblance [30]. A construction of such operators, annihilators, follows straightforwardly from the angle transform introduced in the previous section. On transformed data $\widetilde{A}_{\mathrm{WE}}d$, i.e. on *common-image-point gathers*, annihilators are given by $\frac{\partial}{\partial p_i}$, $i=1,\ldots,n-1$, cf. (90). Indeed, using Proposition 3.2, it follows that where there is illumination, $\frac{\partial}{\partial p_i}\widetilde{A}_{\mathrm{WE}}d=0$ (cf. (75)). Annihilators of the *data* then follow to be

$$\left(\frac{\partial}{\partial t}\right)^{-1} \langle \widetilde{A}_{\mathrm{WE}} \rangle^{-1} \frac{\partial}{\partial p_i} \widetilde{A}_{\mathrm{WE}} \tag{91}$$

as pseudodifferential operators of order zero, where $\langle \widetilde{A}_{\mathrm{WE}} \rangle^{-1}$ is a regularized inverse for $\widetilde{A}_{\mathrm{WE}}$, that is supported microlocally on a subset of $T^*\mathbb{R}^{2n-1}_{(z,x,p)}$ where $\widetilde{A}_{\mathrm{WE}}$ is invertible. Such annihilators are not uniquely defined. Indeed (91) can be composed on the left by any invertible order 0 pseudodifferential operator and we still have an annihilator, of the same order.

We now derive an alternative definition of an annihilator. This definition is motivated by the observation that

$$\frac{\partial}{\partial p_i} R_3 \left(\frac{\partial}{\partial t}\right)^{-1} g \Bigg|_{p=0} = \int_{\mathbb{R}^{n-1}} M_j (R_2 g)\left(z, x - \tfrac{h}{2}, x + \tfrac{h}{2}\right) \chi(z,x,h) \mathrm{d}h,$$

where M_j denotes the multiplication by $h_j = (r_j - s_j)$, and χ was introduced in (56). We have $M_j E_1 = 0$, since $(r_j - s_j)\delta(r - s) = 0$. We observe first that using Theorem 4.2 of Paper I,

$$M_j K^* D_t^2 d = K^* D_t^4 K M_j E_1 \left(\tfrac{1}{2} c_0^{-3} \delta c\right) + [M_j, K^* D_t^4 K] E_1 \left(\tfrac{1}{2} c_0^{-3} \delta c\right). \tag{92}$$

The first term vanishes (because $M_j E_1 = 0$), and the second term is of lower order, so M_j is to highest order an annihilator for $K^* D_t^2 d$.

To derive from (92) a pseudodifferential annihilator of the data, $M_j K^* D_t^2$ is multiplied on the left by $D_t^2 K$. We must also have the support of the operator on the left smaller than the support of the operator on the right of M_j. So suppose ψ'_Y is in $C_0^\infty(Y)$ and is 1 on $\mathrm{supp}(\psi_Y)$, and K' is defined as K but with pseudodifferential cutoff at larger angles of propagation θ'_1, θ'_2 as introduced above (31). Define

$$W_j = \psi_Y D_t^2 K M_j \langle K'^* D_t^2 \psi'_Y D_t^2 K' \rangle^{-1} K'^* D_t^2 \psi'_Y, \tag{93}$$

where we recognize $N_{K'} := K'^* D_t^2 \psi'_Y D_t^2 K'$ from Lemma 2.1. Then we have

Theorem 4.1. *With Assumption 1, the W_j and the operators (91) are annihilators of the data.*

We note that annihilators W_j are essentially multiplicative and do not require a differentiation as in (91), and hence are the ones to be used in practical implementations.

We could define $\widetilde{K} = K N_K^{-1/2}$ as a modification of K so that $D_t^2 \widetilde{K}$ is unitary – at least where ψ_{D} is 1. Then $\frac{1}{2}\sum_j \|W_j[c_0]d\|^2$ simplifies approximately to (the norm does not change by removing $D_t^2 \widetilde{K}$ from the annihilators)

$$\tfrac{1}{2} \sum_j \|W_j[c_0]d\|^2 \approx \frac{1}{2} \int (\|r - s\|^2 |\widetilde{K}^* D_t^2 d|^2 + \text{ lower order terms}) \mathrm{d}z\, \mathrm{d}s\, \mathrm{d}r, \tag{94}$$

This expression is small when the time-to-depth converted data $\widetilde{K}^* D_t^2 d$ are "focused" at $r = s$.

Migration velocity analysis is the estimation of c_0 (the background or "velocity model") based upon the alignment expressed by (90). Traditionally, the updating of the velocity model is carried out interactively. The problem of updating c_0 so that the data will be contained in the range of F_{D} is now cast in a minimum norm optimization problem. In practice, such problems can be addressed by, for example, the conjugate gradient method. Liu and Bleistein [15] developed an automated method for updating c_0 on the basis of the curvature of misalignment at $p = 0$ using ray perturbation theory. It was also done using the data subset based annihilators (differential semblance [30]). Annihilators, $W_j = W_j[c_0]$, replace the necessity to estimate this curvature from the p-family of images. They depend on the background medium. The semi-norm $(\sum_j \|W_j[c_0]d\|^2)^{\frac{1}{2}}$ detects whether c_0 was an acceptable choice or not. The functional $\frac{1}{2}\sum_j \|W_j[c_0]d\|^2$ can be viewed as the downward continuation analog of the differential semblance functional of Symes [30] with the advantage that our annihilator admits the formation of caustics. Of course, the same is true with the W_j replaced by the operators in (91).

References

[1] G. Beylkin, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, J. Math. Phys. 26 (1985) 99–108.

[2] N. Bleistein, J.K. Cohen, J.W. Stockwell Jr., Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion, Springer-Verlag, New York, 2001.
[3] S. Brandsberg-Dahl, M.V. de Hoop, B. Ursin, Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers, Geophysics 68 (2003) 232–254.
[4] J.F. Claerbout, Imaging the Earth's Interior, Blackwell Scientific Publications, Oxford, 1985.
[5] R.W. Clayton, Common midpoint migration, Technical Report SEP-14, Stanford University, 1978.
[6] C.G.M. de Bruin, C.P.A. Wapenaar, A.J. Berkhout, Angle-dependent reflectivity by means of prestack migration, Geophysics 55 (1990) 1223–1234.
[7] M.V. de Hoop, J.H. Le Rousseau, B. Biondi, Symplectic structure of wave-equation imaging: a path-integral approach based on the double-square-root equation, Geophys. J. Int. 153 (2003) 52–74.
[8] M.V. de Hoop, J.H. Le Rousseau, R.-S. Wu, Generalization of the phase-screen approximation for the scattering of acoustic waves, Wave Motion 31 (2000) 43–70.
[9] J.J. Duistermaat, Fourier Integral Operators, Birkhäuser, Boston, MA, 1996.
[10] V. Guillemin, On some results of Gel'fand in integral geometry, in: Pseudodifferential Operators and Applications (Notre Dame, Ind., 1984), Amer. Math. Soc., Providence, RI, 1985, pp. 149–155.
[11] S. Hansen, Solution of a hyperbolic inverse problem by linearization, Comm. Partial Differential Equations 16 (1991) 291–309.
[12] L. Hörmander, The Analysis of Linear Partial Differential Operators, vol. 1, Springer-Verlag, Berlin, 1983.
[13] L. Hörmander, The Analysis of Linear Partial Differential Operators, vol. 3, Springer-Verlag, Berlin, 1985.
[14] L. Hörmander, The Analysis of Linear Partial Differential Operators, vol. 4, Springer-Verlag, Berlin, 1985.
[15] Z. Liu, N. Bleistein, Velocity analysis by inversion, Geophysics 60 (1995) 142–153.
[16] A.P.E. ten Kroode, D.J. Smit, A.R. Verdel, A microlocal analysis of migration, Wave Motion 28 (1998) 149–172.
[17] C.J. Nolan, W.W. Symes, Global solution of a linearized inverse problem for the wave equation, Comm. Partial Differential Equations 22 (1997) 919–952.
[18] C.J. Nolan, Scattering near a fold caustic, SIAM J. Appl. Math. 61 (2000) 659–672.
[19] A.M. Popovici, Prestack migration by split-step DSR, Geophysics 61 (1996) 1412–1416.
[20] M. Prucha, B. Biondi, W.W. Symes, Angle-domain common image gathers by wave-equation migration, in: 69th Ann. Internat. Mtg., Soc. of Expl. Geophys., 1999, pp. 824–827.
[21] Rakesh, A linearized inverse problem for the wave equation, Comm. Partial Differential Equations 13 (1988) 573–601.
[22] D. Ristow, T. Ruhl, Fourier finite-difference migration, Geophysics 59 (12) (1994) 1882–1893.
[23] C.C. Stolk, Microlocal analysis of a seismic linearized inverse problem, Wave Motion 32 (2000) 267–290.
[24] C.C. Stolk, Microlocal analysis of the scattering angle transform, Comm. Partial Differential Equations 27 (2002) 1879–1900.
[25] C.C. Stolk, A pseudodifferential equation with damping for one-way wave propagation in inhomogeneous acoustic media, Wave Motion 40 (2) (2004) 111–121.
[26] C.C. Stolk, M.V. de Hoop, Microlocal analysis of seismic inverse scattering in anisotropic elastic media, Comm. Pure Appl. Math. 55 (2002) 261–301.
[27] C.C. Stolk, M.V. de Hoop, Modeling of seismic data in the downward continuation approach, SIAM J. Appl. Math. 65 (2005) 1388–1406.
[28] C.C. Stolk, W.W. Symes, Kinematic artifacts in prestack depth migration, Geophysics (2004) (in print).
[29] M.J.N. van Stralen, M.V. de Hoop, H. Blok, Generalized Bremmer series with rational approximation for the scattering of waves in inhomogeneous media, J. Acoust. Soc. Am. 104 (1998) 1943–1963.
[30] W.W. Symes, A differential semblance algorithm for the inverse problem of reflection seismology, Comput. Math. Appl. 22 (1991) 147–178.
[31] W.W. Symes, personal communication.
[32] F. TrevesIntroduction to Pseudodifferential and Fourier Integral Operators, vol. 2, Plenum Press, New York, NY, 1980.
[33] S. Xu, H. Chauris, G. Lambaré, M. Noble, Common angle migration: a strategy for imaging complex media, Geophysics 66 (2001) 1877–1894.