Geophys. J. Int. (2005) **161,** 722–738 doi: 10.1111/j.1365-246X.2005.02599.x

Linearized 2.5-dimensional parameter imaging inversion in anisotropic elastic media

Stig-Kyrre Foss,[1,*] Maarten V. de Hoop[2] and Bjørn Ursin[3]

[1] *Institute of Mathematics, NTNU, Trondheim, Norway. E-mail: stigfo@math.ntnu.no*
[2] *Center for Wave Phenomena, Colorado School of Mines, Golden,* CO 80401-1887, *USA. E-mail: mdehoop@mines.edu*
[3] *Institute of Petroleum Technology, NTNU, Trondheim, Norway. E-mail: bjornu@ipt.ntnu.no*

Accepted 2005 January 14. Received 2004 October 7; in original form 2003 June 24

SUMMARY

In this paper we derive 2.5-D short-period seismic modelling and imaging-inversion formulae in the Born approximation for anisotropic elastic media. The 2.5-D approach encompasses 3-D wave scattering measured in a common-azimuth acquisition geometry subject to 2-D computations under appropriate assumptions. The lowest possible symmetry of the medium in this approach, in principle, is monoclinic, while the medium must be translationally invariant in the normal direction to the associated symmetry plane. In the presence of caustics, artefacts may be generated by the imaging-inversion procedures. We show that in the 2.5-D approach the analysis of artefacts in the 2-D symmetry plane implies the corresponding analysis in 3-D in the framework of the common-azimuth acquisition geometry. An interesting aspect of our results is the occurrence of out-of-plane geometrical spreading in the least-squares removal of the contrast source radiation patterns on the data. We finally introduce the 2.5-D generalized Radon transform that generates common-image-point gathers. The reconsideration of 2.5-D scattering theory is motivated by the increase in use of ocean-bottom acquisition technology. It is not uncommon for ocean bottom cable (OBC) seismic data to be collected along a single line, and the question arises of how to make optimal use of these data. We show the effect of the factors making up the amplitude in the 2.5-D generalized Radon transform on OBC field data from the North Sea.

Key words: anisotropy, inversion, ray theory, reflection seismology, seismic modelling.

SYMBOLS AND NOTATION

Symbols	Description
t	time
ω	angular frequency
$\mathbf{x} = (x_1, x_3)$ $(= (x_1, x_2, x_3))$	position vector (before redefinition in main text)
$G_{in}(\mathbf{x}^{\mathrm{r}}, \omega, \mathbf{x}^{\mathrm{s}})$	Green's function in the frequency domain
$\mathbf{x}^{\mathrm{s}}$, $\mathbf{x}^{\mathrm{r}}$	source and receiver positions, respectively
$T(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	traveltime along a ray connecting $\mathbf{x}$ with $\mathbf{x}^{\mathrm{s}}$
$\kappa(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	KMAH index along a ray connecting $\mathbf{x}$ with $\mathbf{x}^{\mathrm{s}}$
$\mathbf{k}$	wave vector
$\mathbf{p} = (p_1, p_2, p_3)$	slowness vector
v	phase velocity
$\mathbf{h} = (h_1, h_2, h_3)$	polarization vector
$A(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	amplitude for a ray at $\mathbf{x}$ initiating at $\mathbf{x}^{\mathrm{s}}$
$\det \mathbf{Q}_2(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	relative geometrical spreading in local surface coordinates on the wave front
$Q_2^{\parallel}(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	in-plane relative geometrical spreading
$Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	out-of-plane relative geometrical spreading
(q_1, q_2)	local wave front coordinates
$\mathbf{V} = (V_1, V_2, V_3)$	group velocity vector

*Now at: Statoil Research, Arkitekt Ebbels vei 10, 7005 Trondheim, Norway. E-mail: skyf@statoil.com

$\hat{n}$	normal to slowness surface
$A^{\|\|}(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$	amplitude with in-plane geometrical spreading only
$\rho(\mathbf{x}) = \rho^{(0)}(\mathbf{x}) + \rho^{(1)}(\mathbf{x})$	density as a sum of a smoothly varying term and a perturbation
$c_{ijkl}(\mathbf{x}) = c^{(0)}_{ijkl}(\mathbf{x}) + c^{(1)}_{ijkl}(\mathbf{x})$	elastic stiffness tensor as a sum of smooth background parameters and a perturbation
$u_{mn}(\mathbf{x}^{\mathrm{r}}, t, \mathbf{x}^{\mathrm{s}})$,	scattered field in time
$U_{mn}(\mathbf{x}^{\mathrm{r}}, \omega, \mathbf{x}^{\mathrm{s}})$	scattered field in frequency
$\mathbf{L}$	modelling or scattering operator
$X \subset \mathrm{IR}^2$	open subset in configuration plane
$\Sigma \subset \partial X \times \partial X$	set of acquisition lines
$T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})$	two-way traveltime
$\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})$	two-way out-of-plane geometrical spreading
$\mathbf{w}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})$	radiation patterns 'vector'
$\mathbf{c}^{(1)}(\mathbf{x})$	medium parameter perturbation 'vector'
$\mathbf{c}^{(0)}(\mathbf{x})$	background medium parameters 'vector'
$v^{\mathrm{s}}_{\circ}$, $v^{\mathrm{r}}_{\circ}$	angular averaged phase velocities
Λ^{S}	2.5-D canonical relation
$\mathbf{p}^{\mathrm{s}}$, $\mathbf{p}^{\mathrm{r}}$	slowness vector at scattering point of the ray connecting to the source and the receiver, respectively
$\boldsymbol{\alpha}^{\mathrm{s}}$, $\boldsymbol{\alpha}^{\mathrm{r}}$	take-off ray directions at scattering point
$\mathbf{k}^{\mathrm{s}}$, $\mathbf{k}^{\mathrm{r}}$	wave vector for source and receiver rays, respectively
Λ^{CA}	common-azimuth (CA) canonical relation
$\mathbf{L}^{\star}$	imaging operator
ψ	tapered mute
$\mathbf{N}$	normal operator
$\mathcal{N}$	kernel of normal operator
$\mathcal{J}$	extended Jacobian
$\boldsymbol{\nu}^{\mathrm{m}}$	migration dip
μ_{LS}	least-squares weight
θ	scattering angle
Γ	matrix used in parameter inversion for a given dip
δ_{BL}	band-limited delta function
$\mathbf{K}$	angle transform
$\mathcal{K}_{mn}$	kernel of angle transform
Φ	phase function
$\underline{\mathbf{m}} = \{\mathbf{m}_1, \ldots\}$	collection of parametrized slices of background medium
J	misfit function for reflection tomography
Ξ	reflectivity matrix used in reflection tomography
λ	statistical regularization parameter

1 INTRODUCTION

In this paper, we derive general single-scattering 2.5-D modelling, imaging, and parameter inversion formulae in anisotropic elastic media under precise assumptions. In principle, the lowest possible symmetry of the medium in this approach is monoclinic, while the medium must be translationally invariant in the direction normal to the associated symmetry plane. It is natural to formulate our results in the context of common-azimuth (CA) migration. If the Earth's properties vary predominantly in two directions, a representative plane containing these directions will suffice to describe the subsurface; the 2.5-D framework applies approximately (it is exact if there are precisely two directions of change). For applicability of the 2.5-D scattering theory in reflection seismology, it typically suffices to assume that the properties (and the geology) vary in two directions only in a slab of sufficient width compared with the dominant wavelength in the data. The application and development of the 2-D approximation in global seismology can be found in Bostock *et al.* (2001). Fig. 1 shows a representative slice in the direction of dominant change (left) and a family of parallel such slices in a medium with a smooth out-of-plane behaviour (right).

The motivation to reconsider 2.5-D scattering, imaging and parameter inversion in elastic media comes from the increasing use of multicomponent ocean bottom acquisition technology. It is not uncommon for ocean bottom cable (OBC) seismic data to be collected along a single line, and the question arises of how to make optimal use of these data. Also, 2.5-D scattering theory can be applied in reflection tomography or migration velocity analysis (Foss *et al.* 2004). Such an application leads to fast, though approximate, algorithms that carry out the reflection tomography slicewise (Fig. 1 (right)). The models resulting from this slicewise approach can play the role of initial models in comprehensive 3-D reflection tomography.

Goldin (1986) and Červený (1981) considered the notion of 3-D wave propagation in 2-D media, the 2-D medium being contained in the mentioned plane. Bleistein (1986) introduced the notion of 2.5-D in seismic applications when restricting attention to waves that travel and scatter in this plane alone, consistent with the CA acquisition geometry if such geometry is aligned with the plane, but which exhibit 3-D geometrical spreading. He considered the acoustic wave equation and derived modelling and Kirchhoff migration formulae for this case. Several authors have since considered 2.5-D Kirchhoff migration in isotropic elastic media (Tygel *et al.* 1998; Dellinger *et al.* 2000). Geoltrain

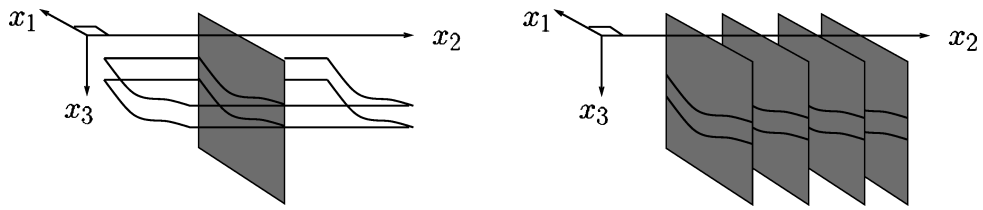

Figure 1. Cylindrical reflectors and the (x_1, x_3)-plane of consideration (left); parallel slices of the medium in the direction of the most dominant change.

(1989) extended the approach to Kirchhoff migration in transversely isotropic media with a vertical symmetry axis. Sollid & Ursin (2003) derived a 2.5-D migration formula using the generalized Radon transform (GRT) in transversely isotropic media. However, no results in 2.5-D have been published admitting the formation of caustics and exploring the weakest possible conditions subject to which the 2.5-D scattering applies.

The imaging inversion results in this paper are derived using the inverse GRT (Miller *et al.* 1987) and natural coordinates at each subsurface point to be imaged, namely scattering angle and migration dip (Burridge *et al.* 1998). Using this choice of coordinates removes the need for a Beylkin determinant (de Hoop *et al.* 1999). In that paper, it was also explained how parameter resolution is coupled to spatial resolution. In the presence of caustics, strictly speaking the GRT should be developed with Maslov Green's functions. Through a stationary phase argument, however, de Hoop & Brandsberg-Dahl (2000) carried out an analysis that showed that as long as there are no caustics occurring at the source or receiver positions, the Maslov formulation reduces to a GRT based upon the geometrical ray approximation (GRA) for the Green's functions. Following this observation, we employ the GRA Green's function in our development. The 2.5-D theory developed in this paper is applicable in the presence of multipathing and caustics under assumptions that will be clarified. On the one hand we extend the results of Burridge *et al.* (1998) to admit the presence of caustics and reformulate them as a genuine least-squares inversion leading to a slightly different amplitude in the GRT. On the other hand, we restrict and approximate the results of Burridge *et al.* (1998) from 3-D to 2.5-D. The emphasis of our analysis is on the geometry and symmetry considerations underlying the GRT application.

The high-frequency linearized (Born) inversion, given a smooth background medium, yields the most singular part of the unknown medium contrast and has been developed in the framework of pseudodifferential and Fourier integral operators (FIOs); see, for example, Duistermaat (1996), or for a short introduction de Hoop *et al.* (2003). This was done in the acoustic case by Rakesh (1988), Hansen (1991) and others, and in the anisotropic elastic case by Stolk & de Hoop (2002). We derive a new 2.5-D linearized inversion formula, using the GRT, for anisotropic media under appropriate assumptions. These assumptions are extensions of those given by Bleistein (1986) necessitated by the introduction of anisotropy, and give rise to certain restrictions on the types of media possible in the 2.5-D framework. The 2.5-D Born modelling integral is obtained by recognizing the out-of-plane variable as a phase variable. Following Bleistein (1986), the out-of-plane integral in the 3-D Born modelling integral is approximated using the method of stationary phase. This is not defined in the presence of out-of-plane caustics (although the 3-D integral is still defined (de Hoop & Brandsberg-Dahl 2000)). Using the inversion result, we derive a 2.5-D least-squares amplitude versus angle (AVA) compensated formula for the generation of common-image-point gathers used in reflection tomography (Brandsberg-Dahl *et al.* 2003a). Implicitly our formulae yield FIOs in 3-D subject to a 2-D computation. We base our numerical computations upon ray tracing; quasi-Monte Carlo sampling can be exploited to minimize the number of ray directions to be traced (de Hoop & Spencer 1996).

In the presence of caustics, imaging artefacts may occur in the CA geometry (for the acoustic case see Nolan & Symes 1997). An artefact is defined as a false event in the image that is not contained in the medium contrast, i.e. an image of a reflector that is not there. The inversion is free of artefacts under the so-called *Bolker condition* (Guillemin 1985; Stolk & de Hoop 2002). When this condition is violated, the image resulting from the inversion procedure will contain artefacts, which we call *image artefacts*. The transformation of seismic data into common-image-point gathers based upon the GRT can be viewed as introducing a restriction to a fixed scattering angle in the inversion formula (Brandsberg-Dahl *et al.* 2003b). This restriction, in the presence of caustics, may give rise to *image-gather artefacts* (for an exhaustive analysis of these artefacts see Stolk 2002). Brandsberg-Dahl *et al.* (2003b) suppressed such artefacts using a procedure called focusing in dip, which selects contributions to the imaging-inversion integral from a subset of the set of all isochrones.

The outline of the paper is as follows. In Section 2, we introduce the notation and the fundamental assumptions pertaining to 2.5-D. We also show by an example that, due to the anisotropy, an additional assumption to those of Bleistein (1986) is required to restrict the rays to travel in-plane. A detailed description of all aspects of the 2.5-D Born single-scattering modelling formula is given. The derivation of this formula can be found in Appendix A. We give a description of which elastic parameters can be determined in the 2.5-D framework of CA data. Additionally, we give the precise assumptions subject to which our 2.5-D modelling is well-defined and show how it pertains to the CA acquisition geometry. Section 3 contains the detailed guide through our inversion procedure, which is an analogue to least-squares inversion. The inversion follows closely that of de Hoop & Brandsberg-Dahl (2000). We show that the 2-D Bolker condition pertaining to the symmetry plane implies the same condition in 3-D, subject to the restriction to CA. First, we construct the imaging operator of the 2.5-D modelling operator. Second, we evaluate the normal operator (Appendix B) and identify its 'inverse' up to leading order. Third, we compose this 'inverse' with the mentioned imaging operator (Appendix C) to find the 2.5-D inversion operator. The actual inversion result, for the most singular part of the medium contrast, is given in Section 3.2 as an inverse by GRT. In Section 4 we present the 2.5-D transformation of the

data to common-image-point gathers. In Section 5 we illustrate the relative significance of the different factors of the amplitude in 2.5-D GRT inversion, using ocean-bottom-cable (OBC) seismic field data from the North Sea. We conclude with a discussion on future applications.

2 MODELLING

2.1 Green's functions in a smoothly varying medium

The geometrical ray approximation (GRA) to the Green's functions is a causal, short-period approximate solution to the elastic wave equation in the frequency domain given by

$$\rho(\mathbf{x})\omega^2 G_{in} + \partial_j(c_{ijkl}(\mathbf{x})\partial_l G_{kn}) = -\delta_{in}\delta(\mathbf{x}-\mathbf{x}^{\mathrm{s}}) \quad i,j,k,l,n = 1,2,3, \tag{1}$$

where ω is angular frequency and the position vector is denoted by $\mathbf{x} = (x_1, x_2, x_3)$; $\rho(\mathbf{x})$ and $c_{ijkl}(\mathbf{x})$ are density and the stiffness tensor, respectively. The Kronecker delta, δ_{in}, represents the body-force source on the right-hand side in the canonical directions, operative at the source point, $\mathbf{x}^{\mathrm{s}}$, through the delta function δ. The summation convention applies here and in the following. The Green's function is a sum over the different wave modes, where each term is of the form

$$G_{ip}(\mathbf{x},\omega,\mathbf{x}^{\mathrm{s}}) = A(\mathbf{x},\mathbf{x}^{\mathrm{s}})h_i^{\mathrm{s}}(\mathbf{x})h_p(\mathbf{x}^{\mathrm{s}})\exp[\mathrm{i}\omega T(\mathbf{x},\mathbf{x}^{\mathrm{s}})], \tag{2}$$

in which $T(\mathbf{x},\mathbf{x}^{\mathrm{s}})$ is the traveltime along the ray connecting $\mathbf{x}$ with $\mathbf{x}^{\mathrm{s}}$. (We do not explicitly indicate the mode of propagation; we treat the modes of propagation separately.) h_i^{s} and h_p are components of the unit polarization vectors at the endpoints of the ray, where the superscript s indicates that this polarization vector is associated with the ray originating at $\mathbf{x}^{\mathrm{s}}$. This convention will be used in the following: superscripts s and r indicate the association of a variable with either a source or a receiver point, $\mathbf{x}^{\mathrm{s}}$ and $\mathbf{x}^{\mathrm{r}}$, respectively. The amplitude, $A(\mathbf{x},\mathbf{x}^{\mathrm{s}})$, becomes complex in the presence of caustics, and can be written in the form (Červený 2001)

$$A(\mathbf{x},\mathbf{x}^{\mathrm{s}}) = \frac{\exp\left[\mathrm{i}\frac{\pi}{2}\kappa(\mathbf{x},\mathbf{x}^{\mathrm{s}})\right]}{4\pi[\rho(\mathbf{x})v^{\mathrm{s}}(\mathbf{x})\rho(\mathbf{x}^{\mathrm{s}})v(\mathbf{x}^{\mathrm{s}})]^{1/2}|\det\mathbf{Q}_2(\mathbf{x},\mathbf{x}^{\mathrm{s}})|^{1/2}}. \tag{3}$$

Here, $v^{\mathrm{s}}(\mathbf{x})$ and $v(\mathbf{x}^{\mathrm{s}})$ are the phase velocities at $\mathbf{x}$ and $\mathbf{x}^{\mathrm{s}}$, respectively, in the direction of the ray connecting these points; $|\det\mathbf{Q}_2(\mathbf{x},\mathbf{x}^{\mathrm{s}})|^{1/2}$ is the relative geometrical spreading, where

$$[\mathbf{Q}_2]^{-1}_{ij}(\mathbf{x},\mathbf{x}^{\mathrm{s}}) = -\frac{\partial^2 T(\mathbf{x},\mathbf{x}^{\mathrm{s}})}{\partial q_i^{\mathrm{s}}\partial q_j}, \quad i,j = 1,2, \tag{4}$$

in which q_j and q_i^{s} are local phase-related coordinates in the plane normal to the slowness vectors at $\mathbf{x}^{\mathrm{s}}$ and $\mathbf{x}$, respectively. We also refer to these coordinates as the *wave front coordinates*. In (3), $\kappa(\mathbf{x},\mathbf{x}^{\mathrm{s}})$ denotes the KMAH index (named after Keller, Maslov, Arnold and Hörmander), which counts the caustics that the ray encounters between $\mathbf{x}^{\mathrm{s}}$ and $\mathbf{x}$; it determines the resulting phase shift. The subtleties concerning the computation of this index in the presence of anisotropy for quasi-shear waves are elucidated in (Klimeš 1997).

We are concerned with 2.5-D modelling, and imaging inversion of elastic waves in anisotropic media in the Born approximation. To justify the 2.5-D approach we need to invoke appropriate assumptions. Two of the assumptions pertain to restricting the ray geometry to the (x_1, x_3)-plane in the modelling, the first being:

Assumption 1: symmetries
(a) The medium is translationally invariant in the out-of-plane direction (x_2).
(b) The (x_1, x_3)*-plane is a plane of mirror symmetry.*

The translational invariance of Assumption 1(a) guarantees that the out-of-plane wave slowness is constant (is conserved). Assumption 1(b) ensures that the wave front is symmetrical about the symmetry plane. Then the relative geometrical spreading can be factorized into an in-plane and out-of-plane component. The lowest possible anisotropic symmetry of the medium is hence monoclinic, having only one plane of mirror symmetry. Assumption 1 is essential for the 2.5-D theory to be applicable.

Because of the symmetry, we may choose one of the wave front coordinates to coincide with the out-of-plane direction, say $q_2 = x_2$ and $q_2^{\mathrm{s}} = x_2^{\mathrm{s}}$. The first wave front coordinate, say q_1, is chosen in-plane. Thus, in the 2.5-D approach (due to Assumption 1(b)), $\partial T(\mathbf{x},\mathbf{x}^{\mathrm{s}})/\partial q_1^{\mathrm{s}}$ and $\partial T(\mathbf{x},\mathbf{x}^{\mathrm{s}})/\partial q_1$ are even functions in x_2, so that $\partial^2 T(\mathbf{x},\mathbf{x}^{\mathrm{s}})/\partial q_1^{\mathrm{s}}\partial x_2 = \partial^2 T(\mathbf{x},\mathbf{x}^{\mathrm{s}})/\partial x_2^{\mathrm{s}}\partial q_1 = 0$ at $x_2 = 0$. In those wave front coordinates, the relative geometrical spreading matrix in eq. (4) in the (x_1, x_3)-plane becomes a diagonal matrix:

$$\mathbf{Q}_2(\mathbf{x},\mathbf{x}^{\mathrm{s}}) = -\begin{bmatrix} \dfrac{1}{\dfrac{\partial^2 T(\mathbf{x},\mathbf{x}^{\mathrm{s}})}{\partial q_1^{\mathrm{s}}\partial q_1}} & 0 \\ 0 & \dfrac{1}{\dfrac{\partial^2 T(\mathbf{x},\mathbf{x}^{\mathrm{s}})}{\partial x_2^{\mathrm{s}}\partial x_2}} \end{bmatrix} = \begin{bmatrix} Q_2^{\parallel}(\mathbf{x},\mathbf{x}^{\mathrm{s}}) & 0 \\ 0 & Q_2^{\perp}(\mathbf{x},\mathbf{x}^{\mathrm{s}}) \end{bmatrix}, \tag{5}$$

where $Q_2^{\parallel}(\mathbf{x},\mathbf{x}^{\mathrm{s}})$ and $Q_2^{\perp}(\mathbf{x},\mathbf{x}^{\mathrm{s}})$ are the in-plane and out-of-plane relative geometrical spreading factors, respectively. This structure implies that the relative geometrical spreading in eq. (3) factors into an in-plane and an out-of-plane component,

$$|\det \mathbf{Q}_2(\mathbf{x}, \mathbf{x}^s)|^{1/2} = |Q_2^{\|}(\mathbf{x}, \mathbf{x}^s) Q_2^{\perp}(\mathbf{x}, \mathbf{x}^s)|^{1/2}. \tag{6}$$

We observe that $p_2^s = \partial T(\mathbf{x}, \mathbf{x}^s)/\partial x_2^s$ is an odd function of x_2. Hence, in the limit $x_2 \to 0$, $p_2^s \to 0$ and we can replace the differential $\partial p_2^s/\partial x_2$ by the fraction p_2^s/x_2, while $p_2^s = p_2$. It follows that the out-of-plane geometrical spreading can be expressed as

$$Q_2^{\perp}(\mathbf{x}, \mathbf{x}^s) = \frac{x_2}{p_2} = \int_{\text{ray}} \frac{V_2}{p_2}\, dt, \tag{7}$$

where $\mathbf{V}$ denotes the group velocity, and the integral is along the ray parametrized by traveltime t connecting $\mathbf{x}^s$ with $\mathbf{x}$. The integrand V_2/p_2 tends to a finite value as both V_2 and p_2 tend to zero (in accordance with the 2.5-D approach, where $p_2 = 0$ and $V_2 = 0$ at $x_2 = 0$). Closed-form expressions for the integrand for isotropic, transversely isotropic and orthorhombic media can be found in Ettrich *et al.* (2002), eqs (15), (19) and (22), respectively. Note that the argument in eq. (5) for the splitting of the in-plane and out-of-plane spreading is not dependent on the particular constant value of $p_2 = \partial T(\mathbf{x}, \mathbf{x}^s)/\partial x_2 = p_2^s$.

For the later analysis, we introduce the amplitude $A^{\|}(\mathbf{x}, \mathbf{x}^s)$ (*cf.* eq. 3) containing the in-plane geometrical spreading factor $Q_2^{\|}$ only:

$$A^{\|}(\mathbf{x}, \mathbf{x}^s) = \frac{\exp\left[\mathrm{i}\frac{\pi}{2}\kappa(\mathbf{x}, \mathbf{x}^s)\mathrm{sgn}\omega\right]}{4\pi[\rho(\mathbf{x})v^s(\mathbf{x})\rho(\mathbf{x}^s)v(\mathbf{x}^s)]^{1/2}\left|Q_2^{\|}(\mathbf{x}, \mathbf{x}^s)\right|^{1/2}}. \tag{8}$$

2.2 Scattering in the high-frequency Born approximation

The Born approximation for scattered waves is of the form of a time convolution of an incoming and outgoing Green's function in a medium background that is spatially coupled to a medium contrast. We denote a source position (associated with the incoming Green's function) by $\mathbf{x}^s$ as before, and a receiver position (associated with the outgoing Green's function) by $\mathbf{x}^r$; $\mathbf{x}$ will denote any scattering point at which the medium contrast differs from zero. Substituting for the Green's functions asymptotic expressions of the form (2) leads to the high-frequency or short-period Born approximation; the associated ray geometry consists of a source ray connecting $\mathbf{x}$ with $\mathbf{x}^s$ and a receiver ray connecting $\mathbf{x}^r$ with $\mathbf{x}$. The total traveltime of the scattered wave in this approximation is the sum of traveltimes associated with the incoming and outgoing Green's functions:

$$T(\mathbf{x}^r, \mathbf{x}, \mathbf{x}^s) = T(\mathbf{x}^s, \mathbf{x}) + T(\mathbf{x}, \mathbf{x}^r). \tag{9}$$

The expression for the 3-D high-frequency Born approximation is given by eq. (A1). The symmetry plane is given by $x_2 = 0$ as before.

2.2.1 Ray geometry

The 2.5-D Born modelling formula is derived from the 3-D Born modelling formula by a stationary phase argument integrating out the out-of-plane coordinate (here x_2). The stationary phase argument involves the stationarity condition (Appendix A, eq. A3):

$$\partial_{x_2}[T(\mathbf{x}^s, \mathbf{x}) + T(\mathbf{x}, \mathbf{x}^r)] = p_2^s + p_2^r = 0, \tag{10}$$

where $\mathbf{p}^s = (p_1^s, p_2^s, p_3^s)$ and $\mathbf{p}^r = (p_1^r, p_2^r, p_3^r)$ are the slowness vectors associated with the source ray and the receiver ray, respectively. The simplification of the 2.5-D theory is dependent on whether source and receiver rays are contained in the plane reflecting the symmetries, $x_2 = 0$. However, the stationary phase argument in eq. (10) admits out-of-plane contributions. We give an example of this next.

We consider *qSV–qSV* scattering ('q' stands for 'quasi' in the following) and use that $p_2^s = -p_2^r$ equals a constant (see Fig. 2). The medium here is transversely isotropic with a vertical symmetry axis (x_3-axis). By controlling the Thomsen parameters, ϵ and δ (Thomsen 1986), we may change the shape of the slowness surface smoothly as shown on the right of the figure for three different depths in the model used in the ray tracing on the left. The group velocity vector, $\mathbf{V}$, is perpendicular to the slowness surface and points the direction of the ray. This is indicated with the surface normal to the slowness surface, $\hat{n}$. As suggested in the figure, we may smoothly change the direction of the group velocity, sending the rays out of the plane and back again while keeping p_2 fixed, as implied by Assumption 1(a). Two rays are shot at a small positive and negative angle with the x_3-axis in the out-of-plane direction ($p_1^s = p_1^r = 0$) from $(x_1, x_2, x_3) = (0, 0, 200)$. They travel in the (x_2, x_3)-plane and intersect at the point (0, 0, 2800) with incoming angles following Snell's law and satisfying eq. (10). We note that the anisotropy in this example allows triplications on the symmetry axis, which is uncommon, for example, in a sedimentary sequence setting.

We admit in-plane scattering events only by imposing:

Assumption 2: seismic phase restriction
Only seismic events with at least one ray, or leg, associated with a wave type that pertains to a convex slowness surface are considered.

If one of the ray legs is associated with a wave type pertaining to a convex slowness surface, the only solution to eq. (10) is $p_2^r = p_2^s = 0$. Because of the symmetry in Assumption 1(b) this also implies that $V_2^s = V_2^r = 0$. This follows because $p_2^r = -p_2^s \neq 0$ on a convex slowness surface would induce group velocity vectors pointing out of the plane (the middle slowness surface on the right of Fig. 2 illustrates this). Such group velocity vectors would send seismic energy away from the plane never to return. In particular, *P* waves always have a convex slowness surface (see, e.g. Musgrave 1970); hence, *qP–qP* and *qP–qSV* scattering events always satisfy Assumption 2.

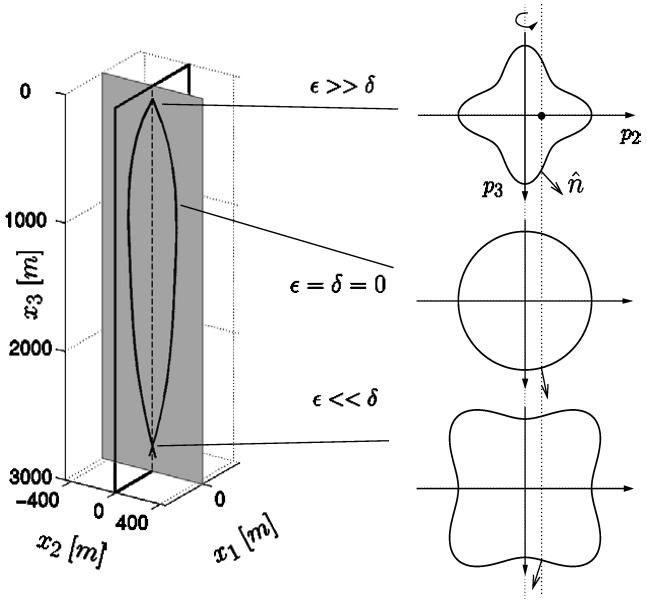

Figure 2. An example of rays scattering out-of-plane that are in the stationary point set defined by eq. (10); ϵ and δ are Thomsen's parameters in transversely isotropic media.

2.2.2 *Scattered field*

From now on, let $\mathbf{x} = (x_1, x_3)$ denote a position in the plane reflecting the symmetries. We have dropped the x_2 coordinate. Indeed, the medium parameters are independent of x_2, and we restrict the wavefield to $x_2 = 0$, i.e. $x_2^{\rm s} = x_2^{\rm r} = 0$. We assume that the medium contrast is supported in $X \subset \mathrm{IR}^2 \ni \mathbf{x}$. The medium parameters are represented by a sum of a smooth part, $\rho^{(0)}$ and $c^{(0)}_{ijkl}$, and a singular perturbation, $\rho^{(1)}$ and $c^{(1)}_{ijkl}$:

$$\rho(\mathbf{x}) = \rho^{(0)}(\mathbf{x}) + \rho^{(1)}(\mathbf{x}), \quad c_{ijkl}(\mathbf{x}) = c^{(0)}_{ijkl}(\mathbf{x}) + c^{(1)}_{ijkl}(\mathbf{x}). \tag{11}$$

Note that both the smooth part and the perturbation are restricted in accordance with Assumptions 1(a) and 1(b), making the reflectors parts of cylindrical surfaces (see Fig. 1(left)). In the Born approximation, the waves travel in the smooth part of the medium and are scattered off the perturbation once. In the imaging-inversion problem the smooth medium is assumed to be known. It is the medium perturbation we will invert for. The medium perturbation contains the volume scatterers and reflectors.

In the high-frequency Born approximation, the multicomponent data collected in a seismic experiment, u_{mn}, under the condition of a smooth background, will be the asymptotic part of the modelling formula. The subscripts indicate that u_{mn} is the m-component of the recorded displacement due to a body force in the n-direction. The data (scattered wavefield) can be modelled by an operator $\mathbf{L}$ acting on the medium perturbation (see Appendix A for a derivation, where u_{mn} is the time-Fourier transform of the data U_{mn})

$$\begin{aligned} u_{mn}(\mathbf{x}^{\rm s}, \mathbf{x}^{\rm r}, t) &= \left(\mathbf{L}\mathbf{c}^{(1)}(\mathbf{x}^{\rm s}, \mathbf{x}^{\rm r}, t)\right)_{mn} \\ &\approx \sum_{i\in I} \int_X \left(\sqrt{\frac{\mathrm{i}}{2\pi}} \int_{\mathbb{R}} h_m(\mathbf{x}^{\rm r}) \omega^{3/2} \rho^{(0)}(\mathbf{x}) \frac{A^{\|}(\mathbf{x}^{\rm s}, \mathbf{x}) A^{\|}(\mathbf{x}, \mathbf{x}^{\rm r})}{\mathcal{L}^{\perp}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s})} \right. \\ &\quad \left. \times \mathbf{w}^{\rm T}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s}) \mathbf{c}^{(1)}(\mathbf{x}) \exp\{\mathrm{i}\omega[T^{(i)}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s}) - t]\}\, h_n(\mathbf{x}^{\rm s})\, d\omega \right) d\mathbf{x}, \end{aligned} \tag{12}$$

where

$$\mathcal{L}^{\perp}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s}) = \left| Q_2^{\perp}(\mathbf{x}^{\rm s}, \mathbf{x}) + Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\rm r}) \right|^{1/2} \tag{13}$$

denotes the two-way out-of-plane geometrical spreading, and $T^{(i)}$ is the two-way traveltime given by

$$T^{(i)}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s}) = T(\mathbf{x}^{\rm s}, \mathbf{x}) + T(\mathbf{x}, \mathbf{x}^{\rm r}). \tag{14}$$

The significance of the superscript (i) on $T^{(i)}$ (that we suppress in all the factors making up the amplitude in 12) arises in the case of multivaluedness of two-way traveltime in the presence of caustics. In that case $i \in I$ labels the branches of the two-way traveltime and I is the collection of them. The set $\{T^{(i)}\}_{i\in I}$ describes the two-way traveltimes for all branches. We will omit this indexing of the branches for clarity of notation until the distinction becomes important again (see the discussion on artefacts). The sum over the different traveltime branches in eq. (12) will thus be suppressed and assumed implicit in the following.

In (12), the medium perturbations are collected in the matrix $\mathbf{c}^{(1)}(\mathbf{x})$ and the radiation patterns equivalently in $\mathbf{w}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s})$. Since by Assumption 1 the medium is restricted to monoclinic or higher symmetry, there are at most 13 stiffness parameters plus density as

independent unknown distributions. If the (x_1, x_3)-plane is the plane reflecting the symmetries, as it should, the stiffness tensor is given by:

$$[c_{ijkl}] = \begin{bmatrix} c_{1111} & c_{1122} & c_{1133} & 0 & c_{1113} & 0 \\ c_{1122} & c_{2222} & c_{2233} & 0 & c_{2213} & 0 \\ c_{1133} & c_{2233} & c_{3333} & 0 & c_{3313} & 0 \\ 0 & 0 & 0 & c_{2323} & 0 & c_{2312} \\ c_{1113} & c_{2213} & c_{3313} & 0 & c_{1313} & 0 \\ 0 & 0 & 0 & c_{2312} & 0 & c_{1212} \end{bmatrix}. \tag{15}$$

Therefore, we restrict the indices of $c^{(1)}_{ijkl}$ to the 13 independent components of the stiffness tensor (*cf.* eq. 11). The medium perturbations are collected in the 14×1 matrix for monoclinic anisotropy (Burridge *et al.* 1998)

$$\mathbf{c}^{(1)}(\mathbf{x}) = \left\{ \frac{\rho^{(1)}(\mathbf{x})}{\rho^{(0)}(\mathbf{x})}, \frac{c^{(1)}_{ijkl}(\mathbf{x})}{\rho^{(0)}(\mathbf{x}) v^{\mathrm{s}}_{\circ}(\mathbf{x}) v^{\mathrm{r}}_{\circ}(\mathbf{x})} \right\}, \tag{16}$$

where $v^{\mathrm{s}}_{\circ}$ and $v^{\mathrm{r}}_{\circ}$ are local phase velocities averaged over all phase angles. These are introduced for computational purposes so that the matrix has components of similar magnitude. With higher symmetry, such as isotropy, the matrix reduces accordingly (Beylkin & Burridge 1990). The radiation pattern matrix is defined similarly as the 14×1 matrix (Burridge *et al.* 1998):

$$\mathbf{w}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) = \left\{ h^{\mathrm{s}}_m(\mathbf{x}) h^{\mathrm{r}}_m(\mathbf{x}), [h^{\mathrm{s}}_i(\mathbf{x}) p^{\mathrm{s}}_j(\mathbf{x}) h^{\mathrm{r}}_k(\mathbf{x}) p^{\mathrm{r}}_l(\mathbf{x})] v^{\mathrm{s}}_{\circ}(\mathbf{x}) v^{\mathrm{r}}_{\circ}(\mathbf{x}) \right\}, \tag{17}$$

where the indices follow those of the stiffness matrix in the ordering defined by the matrix multiplication or inner product $\mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})\, \mathbf{c}^{(1)}(\mathbf{x})$ in the modelling eq. (12).

Since the out-of-plane slownesses will be zero, $p^{\mathrm{s}}_2 = p^{\mathrm{r}}_2 = 0$, we see from eq. (17) that the contributions to $\mathbf{w}^{\mathrm{T}}\mathbf{c}^{(1)}$ vanish for $j, l = 2$ in $c^{(1)}_{ijkl}$. In view of Assumptions 1 and 2, the polarizations of both *qP* and *qSV* waves satisfy $h_2 = 0$, and hence the contributions for $i, k = 2$ in $c^{(1)}_{ijkl}$ are also zero. Thus the parameters to which the modelling operator is sensitive for *qP*–*qP* and *qP*–*qSV* scattering are the seven parameters out of the 14 independent ones in a monoclinic medium, $c^{(1)}_{1111}, c^{(1)}_{1133}, c^{(1)}_{3333}, c^{(1)}_{1113}, c^{(1)}_{3313}, c^{(1)}_{1313}$, and density $\rho^{(1)}$. The kinematic aspects for a monoclinic background medium are governed by the same partitioning of parameters (Chapman & Pratt 1992; Foss & Ursin 2003).

Remark 1
In modelling and in imaging, with the smooth background medium given, the relevant amplitudes can be computed in the lowest possible symmetry admitted by the 2.5-D framework: monoclinic. For inversion and reflection tomography with observations restricted to the plane under consideration, however, only parameters associated with this plane can be estimated. Hence the lowest possible symmetry is transversely isotropic with a symmetry axis in the plane. Then, because of the rotational symmetry of the medium, parameters needed in out-of-plane amplitude calculations can be found from in-plane propagation (Ettrich et al. 2002; eq. 19).

Remark 2
For the stationary phase in x_2 to be applicable, we mute events associated with rays forming out-of-plane caustics (see Remark 4 in Appendix A).

We note that the modelling eq. (12) is valid up to leading order since we inserted the GRA amplitudes in the representations for the Green's functions. We restrict $\mathbf{x}^{\mathrm{s}}$, $\mathbf{x}^{\mathrm{r}}$ to smooth source and receiver lines forming the manifold Σ; we note that in ocean bottom acquisition these lines do not coincide but are still contained in the plane. Each traveltime branch $T^{(i)}$ is defined on a subset $D^{(i)}$ of $\Sigma \times X$, i.e. a particular subset of acquisition and scattering points $(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, \mathbf{x})$.

2.2.3 Ocean-bottom boundary conditions

The displacements in (12) are the ones associated with the scattered field. In the case of four-component ocean bottom acquisition, the particle velocity (and hence, displacement) and the pressure are measured at just above the fluid–solid interface. Decomposition into up- and downgoing waves of such data is then necessary to identify the scattered field (the scattered field is the upgoing constituent). We use the fluid–solid boundary conditions in this process, in particular, so that the components of the traction parallel to the fluid–solid interface vanish (in the solid). The decomposition requires the estimation of the elastic parameters of the seafloor, and is discussed, for example, by (Amundsen & Reitan (1995, eqs 23, 25); although the procedure in that paper applies to a laterally homogeneous and isotropic ocean bottom, it can be straightforwardly extended to anisotropy and heterogeneity by invoking techniques developed by Taylor (1975). We note that in our theory we only capture body wave scattering; Scholte waves, for example, are unaccounted for.

2.3 The modelling operator in common azimuth

The 3-D Born modelling operator is a Fourier integral operator (FIO) under the mild conditions that there are no direct rays between the source and the receiver reaching the medium perturbation (i.e. rays that have scattered off a subsurface point over an angle π) and no grazing rays (i.e. rays that reach the acquisition surface tangentially to the surface) (Rakesh 1988; Hansen 1991). The 3-D modelling operator with

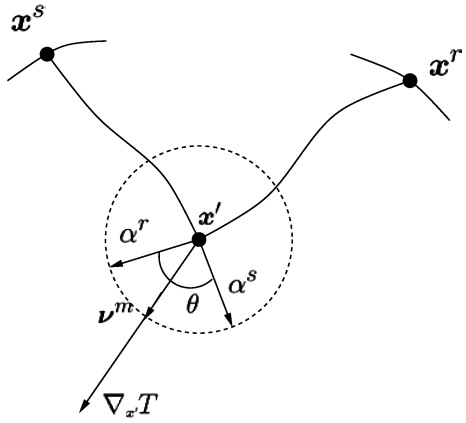

Figure 3. Phase directions, and dip and scattering angle: change from surface to subsurface coordinates.

common-azimuth (CA) acquisition geometry (Biondi & Palacharla 1996) is also an FIO under similar conditions (Nolan & Symes 1997; de Hoop *et al.* 2003). By repetition, the notion of 2.5-D implies CA (but not vice versa) by aligning our symmetry plane with the acquisition geometry, i.e. $x_2^{\rm s} = x_2^{\rm r}$. It follows that the 2.5-D modelling operator, $\mathbf{L}$: $\mathbf{c}^{(1)} \rightarrow u_{mn}$, defined by (12), has the properties of an FIO the kernel of which has a phase function $\Phi = \Phi(\mathbf{x}^{\rm s}, \mathbf{x}^{\rm r}, t, \mathbf{x}, \omega) = \omega\,[T^{(i)}(\mathbf{x}^{\rm r}, \mathbf{x}, \mathbf{x}^{\rm s}) - t]$ in an integral representation over phase variable ω. The propagation of singularities of such an operator is governed by its canonical relation, $\Lambda^{\rm S} = \{[(\mathbf{x}^{\rm s}, \mathbf{x}^{\rm r}, t), \nabla_{(\mathbf{x}^{\rm s},\mathbf{x}^{\rm r},t)}\Phi; \mathbf{x}, \nabla_{\mathbf{x}}\Phi] \,|\, \partial_\omega \Phi = 0\}$ (superscript S indicates that this is a canonical relation in two space dimensions coinciding with the symmetry plane). The canonical relation is a table that describes how reflectors map to reflections geometrically, also known as map demigration. The canonical relation is intimately connected to ray tracing in phase space (Billette & Lambaré 1998).

Using ray tracing in phase space, $\Lambda^{\rm S}$ can be written in the form

$$\Lambda^{\rm S} = \{[\mathbf{x}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}), \mathbf{x}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r}), T(\mathbf{x}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}), \mathbf{x}, \mathbf{x}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r})), \mathbf{k}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}, \omega), \mathbf{k}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r}, \omega), \omega;$$

$$\mathbf{x}, -\mathbf{k}(\mathbf{x}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}), \mathbf{x}, \mathbf{x}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r}), \omega)]|(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}, \boldsymbol{\alpha}^{\rm r}) \in K, \omega \in \mathbb{R}\backslash 0\}\,. \tag{18}$$

Here, $\boldsymbol{\alpha}^{\rm s}(\mathbf{x})$ and $\boldsymbol{\alpha}^{\rm r}(\mathbf{x})$ are the take-off directions of the rays originating at the scattering point $\mathbf{x} \in X$ and are given by the unit phase vectors $\boldsymbol{\alpha}^{\rm s}(\mathbf{x}) = \mathbf{p}^{\rm s}(\mathbf{x})/|\mathbf{p}^{\rm s}(\mathbf{x})|$ and $\boldsymbol{\alpha}^{\rm r}(\mathbf{x}) = \mathbf{p}^{\rm r}(\mathbf{x})/|\mathbf{p}^{\rm r}(\mathbf{x})|$ (see Fig. 3). On the canonical relation (18)

$$\mathbf{k}(\mathbf{x}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}), \mathbf{x}, \mathbf{x}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r}), \omega) = \omega \nabla_{\mathbf{x}} T(\mathbf{x}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}), \mathbf{x}, \mathbf{x}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r}))$$

has the interpretation of wave vector (the Fourier dual of $\mathbf{x}$). Given a take-off direction ($\boldsymbol{\alpha}^{\rm s}$ or $\boldsymbol{\alpha}^{\rm r}$) at a scattering point $\mathbf{x}$, rays are traced to an acquisition line (in Σ) giving unique intersection points ($\mathbf{x}^{\rm s} = \mathbf{x}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s})$ or $\mathbf{x}^{\rm r} = \mathbf{x}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r})$) as well as slowness vectors. These slowness vectors are projected onto the aforementioned acquisition line resulting in $\mathbf{k}^{\rm s}/\omega$, $\mathbf{k}^{\rm r}/\omega$, from which $\mathbf{k}^{\rm s}(\mathbf{x}, \boldsymbol{\alpha}^{\rm s}, \omega)$, $\mathbf{k}^{\rm r}(\mathbf{x}, \boldsymbol{\alpha}^{\rm r}, \omega)$ are derived. The set K is an appropriately chosen subset of $X \times S \times S$ (S is the unit circle) so that the 'no grazing rays' and 'no scattering over π' conditions are satisfied.

Remark 3
The 3-D CA Born scattering operator yields the following canonical relation (compare with the 2.5-D Born modelling operator above in eq. 18)

$$\Lambda^{\rm CA} = \left\{\left(x_1^{\rm s}, x_2^{\rm s}, x_1^{\rm r}, t, k_1^{\rm s}, k_2^{\rm s} + k_2^{\rm r}, k_1^{\rm r}, \omega; x_1, x_2, x_3, -k_1, -k_2, -k_3\right)\right\}, \tag{19}$$

where $(x_1^{\rm s}, x_2^{\rm s}, x_1^{\rm r}, x_2^{\rm r}, t, k_1^{\rm s}, k_2^{\rm s}, k_1^{\rm r}, k_2^{\rm r}, \omega; x_1, x_2, x_3, -k_1, -k_2, -k_3)$ *are contained in the 3-D Born modelling canonical relation, which is a straightforward extension of* $\Lambda^{\rm S}$ *to three space dimensions (e.g. Stolk & de Hoop 2002). Note that t is here a general two-way traveltime function in order to separate it from the one in eq. (14), which is limited to a plane. Due to the translational invariance of the medium, Assumption 1(a), the individual* $k_2^{\rm s}$ *and* $k_2^{\rm r}$ *are preserved. Hence at the scattering point* $k_2 = k_2^{\rm s} + k_2^{\rm r}$, *in which* $k_2^{\rm s}$ *and* $k_2^{\rm r}$ *attain their values at the source and the receiver, respectively. This equality can also be written as* $k_2 = \omega\,(p_2^{\rm s} + p_2^{\rm r})$. *In view, again, of the invariance in the out-of-plane direction the 2-component of any reflector dip must vanish. Hence, given a source and a receiver ray such that* $k_2^{\rm s} + k_2^{\rm r} \neq 0$, *no specular reflection will return to the acquisition manifold. This is also confirmed by the stationary phase argument underlying the 2.5-D analysis; see eq. (10).*

3 2.5-D IMAGING INVERSION

In this section, we develop the inversion for the medium parameters collected in matrix $\mathbf{c}^{(1)}$ in eq. (16). The inversion can be viewed as an analogue to matrix least-squares inversion. In this context, we introduce the adjoint $\mathbf{L}^\star$ of the modelling operator, the so-called *imaging operator*, defined in the standard manner through ($\langle .,. \rangle_{(\mathbf{y})}$ denotes the L^2 inner product over $\mathbf{y}$-space)

$$\left\langle u, \mathbf{L}\mathbf{c}^{(1)} \right\rangle_{(\mathbf{x}^{\rm s},\mathbf{x}^{\rm r},t)} = \left\langle \mathbf{L}^\star u, \mathbf{c}^{(1)} \right\rangle_{(\mathbf{x})}.$$

We then apply the imaging operator to eq. (12),

$$\mathbf{L}^{\star}\left(\psi\mathbf{L}\mathbf{c}^{(1)}\right)=\mathbf{L}^{\star}u, \tag{20}$$

where u is shorthand for the data. The composition of the imaging operator with the modelling operator gives us the *normal operator*, $\mathbf{N}=\mathbf{L}^{\star}\psi\,\mathbf{L}$. Here ψ is a tapered mute to ensure the conditions of Subsection 2.3, which make the modelling operator an FIO, are satisfied.

Assumption 3.
The projection of the canonical relation Λ^{S} of the 2.5-D modelling operator on the acquisition variables $(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, t, \mathbf{k}^{\mathrm{s}}, \mathbf{k}^{\mathrm{r}}, \omega)$ is one-to-one.

Assumption 3 is consistent with the 2-D Bolker condition (Guillemin 1985); for the acoustic case see also Ten Kroode *et al.* (1998). It means that $(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, t, \mathbf{k}^{\mathrm{s}}, \mathbf{k}^{\mathrm{r}}, \omega)$, i.e. the reflection times and slopes in the data, determine a scattering point $\mathbf{x}$ and associated wave vector $\mathbf{k}$, i.e. the reflector position and dip, uniquely and smoothly. This condition ensures that the normal operator is an elliptic pseudodifferential operator. Then the procedure of imaging the modelled data, eq. (20), does not generate artificial reflectors that do not really exist, i.e. that are not present in $\mathbf{c}^{(1)}$ (Stolk & de Hoop 2002).

Subject to Assumption 3, we can construct the parametrix of $\mathbf{N}$, denoted by $\langle\mathbf{N}^{-1}\rangle$. The brackets indicate that this is a generalized inverse. Following the analogue of least-squares inversion, an estimate $\hat{\mathbf{c}}^{(1)}$ of the medium perturbation in eq. (16) can be recovered from the operator composition:

$$\hat{\mathbf{c}}^{(1)}\approx\langle\mathbf{N}^{-1}\rangle\mathbf{L}^{\star}u, \tag{21}$$

where u is the data. If Assumption 3 is violated this process would generate image artefacts (discussed in the introduction). A less restrictive condition and a discussion of the implied artefacts can be found in de Hoop & Brandsberg-Dahl (2000) and Stolk (2000).

In view of Remark 3, Assumption 3 (or the 2-D Bolker condition) in the symmetry plane implies the CA Bolker condition in 3-D. This condition encompasses that for any point in Λ^{CA}, i.e. for an intersecting pair of source and receiver rays, given $(x_1^{\mathrm{s}}, x_2^{\mathrm{s}}, x_1^{\mathrm{r}}, x_2^{\mathrm{s}}, t, k_1^{\mathrm{s}}, k_2^{\mathrm{s}}+k_2^{\mathrm{r}}, k_1^{\mathrm{r}}, \omega)$, there is only one (t, k_2^{s}) that explains the reflection.

We note that the Bolker condition is ray-geometrical and implies the 'invertibility' of the modelling operator for waves.

3.1 The imaging and normal operators

In this subsection, we evaluate the adjoint $\mathbf{L}^{\star}$ of the modelling operator $\mathbf{L}$ and deduce the leading order contribution to the normal operator $\mathbf{N}$ for which we can derive a parametrix (i.e. inverse up to smoothing contributions). We will write the action of the normal operator as a pseudodifferential operator, i.e. as integrals over $\mathbf{x}$ and its Fourier dual $\mathbf{k}$, the wave vector. The integrand yields the so-called *symbol* of the normal operator.

The action of the adjoint or imaging operator $\mathbf{L}^{\star}$ can be written as

$$\begin{aligned}\mathbf{L}^{\star}u(\mathbf{x}')\approx\frac{1}{\sqrt{2\pi\mathrm{i}}}\int_{\Sigma}\int_{\mathbb{R}}\int_{\mathbb{R}\geq 0}(\omega')^{3/2}\rho^{(0)}(\mathbf{x}')\frac{(A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{\|}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})^{\star}}\\ \times\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')u_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},t)h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')\exp\{\mathrm{i}\omega'[T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})-t]\}\,d\omega'\,d\mathbf{x}^{\mathrm{s}}\,d\mathbf{x}^{\mathrm{r}}\,dt,\end{aligned} \tag{22}$$

where ω' is the frequency and * indicates the adjoint as well as complex conjugation. The polarization vectors associated with scattering off point $\mathbf{x}'$ are denoted by $h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')$ and $h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')$ to distinguish them from the polarization vectors due to scattering off $\mathbf{x}$, as in eq. (12). Note the summation over the different indices of the data u_{mn} in the integrand of the adjoint, following the summation convention; hence the notation of $\mathbf{L}^{\star}$ acting on all the data denoted by u. Note also that the radiation patterns, $\mathbf{w}=(\mathbf{w}^{\mathrm{T}})^{\star}$.

Composing the imaging operator with the modelling operator (having used a pseudodifferential cut-off for when it fails to satisfy Assumption 3) gives us the normal operator. We carry out this composition in Appendix B. We choose to work with coordinates directly at the image point, $\mathbf{x}'$ say, and introduce the change of variables

$$(\mathbf{x}',\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}})\leftarrow(\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}},\boldsymbol{\alpha}^{\mathrm{r}}), \tag{23}$$

where $\boldsymbol{\alpha}^{\mathrm{s}}$ and $\boldsymbol{\alpha}^{\mathrm{r}}$ are the phase directions associated with the rays connecting the image point and the source at $\mathbf{x}^{\mathrm{s}}$ and receiver at $\mathbf{x}^{\mathrm{r}}$, respectively (see Fig. 3). The mapping is invertible, but in the case of multipathing per branch only (see Subsection 2.2). Transformation of (23) removes the use of the Beylkin determinant (de Hoop *et al.* 1999). All functions of $(\mathbf{x}',\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}})$ become functions of $(\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}},\boldsymbol{\alpha}^{\mathrm{r}})$. The domain of $(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}})$-integration becomes

$$\Sigma\leftarrow S^{\mathrm{s}}\times S^{\mathrm{r}}\text{ for given }\mathbf{x}', \tag{24}$$

where S^{s} and S^{r} are two subsets of the unit circle in the symmetry plane for the sources and receivers, respectively.

We introduce the function containing the Jacobian,

$$\mathcal{J}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\equiv\left|\rho^{(0)}(\mathbf{x}')A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{\|}(\mathbf{x}',\mathbf{x}^{\mathrm{r}})\right|^2\frac{\partial(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}})}{\partial(\boldsymbol{\alpha}^{\mathrm{s}},\boldsymbol{\alpha}^{\mathrm{r}})}, \tag{25}$$

reflecting the change of coordinates (23). Furthermore, we introduce the weight,

$$\mu_{\mathrm{LS}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\equiv\frac{\mathcal{J}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})}{|\nabla_{\mathbf{x}}T(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})|^4}. \tag{26}$$

We change variables further, from the phase directions to scattering angle θ, which is the angle between $\boldsymbol{\alpha}^{\mathrm{s}}$ and $\boldsymbol{\alpha}^{\mathrm{r}}$, and the migration dip $\boldsymbol{\nu}^{\mathrm{m}}$, $\boldsymbol{\nu}^{\mathrm{m}} = \nabla_{\mathbf{x}'} T/|\nabla_{\mathbf{x}'} T|$

(see Fig. 3). The scattering angle satisfies

$$\theta = \theta(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}}) = \arccos(\boldsymbol{\alpha}^{\mathrm{s}} \cdot \boldsymbol{\alpha}^{\mathrm{r}}). \tag{27}$$

To be able to integrate over the migration dip, there is assumed to be no scattering over $\theta = \pi$ so that $\nabla_{x'} T \neq 0$, and $\boldsymbol{\nu}^{\mathrm{m}}$ is defined. We have

$$(\mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}}, \boldsymbol{\alpha}^{\mathrm{r}}) \leftrightarrows (\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta). \tag{28}$$

In 2.5-D this mapping exists both ways; clearly, there is no integration over the azimuth (de Hoop & Brandsberg-Dahl 2000). All functions of $(\mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}}, \boldsymbol{\alpha}^{\mathrm{r}})$ become functions of $(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)$. The domain of $(\boldsymbol{\nu}^{\mathrm{m}}, \theta)$-integration is written as

$$S^{\mathrm{s}} \times S^{\mathrm{r}} \rightarrow E_{\boldsymbol{\nu}^{\mathrm{m}}} \times E_\theta \text{ for given } \mathbf{x}'. \tag{29}$$

We introduce the square matrix (of the dimensionality of $\mathbf{w}$)

$$\Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}) = \frac{1}{2} \int_{E_\theta} \frac{\mu_{\mathrm{LS}}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)}{|\mathcal{L}^{\perp}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)|^2} \mathbf{w}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta) \mathbf{w}^{\mathrm{T}}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta) \frac{\partial(\boldsymbol{\alpha}^{\mathrm{s}}, \boldsymbol{\alpha}^{\mathrm{r}})}{\partial(\boldsymbol{\nu}^{\mathrm{m}}, \theta)} \, d\theta + (\ldots)(\mathbf{x}', -\boldsymbol{\nu}^{\mathrm{m}}), \tag{30}$$

where the second term is the same as the first but with $\boldsymbol{\nu}^{\mathrm{m}}$ replaced by $-\boldsymbol{\nu}^{\mathrm{m}}$. Then

$$\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \frac{1}{(2\pi)^2} \int_{E_{\nu^{\mathrm{m}}}} \int_{\mathbb{R}\geq 0} \int_X k^2 8\pi^2 \Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}) \mathbf{c}^{(1)}(\mathbf{x}) \exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\mathbf{x} - \mathbf{x}')] \, d\mathbf{x} \, k \, dk \, d\boldsymbol{\nu}^{\mathrm{m}}, \tag{31}$$

cf. eq. (B8), where we have rewritten the k integral over postive values only. The right-hand side of (31) has the explicit form of a pseudodifferential operator with integration over $\mathbf{x}$ and its dual $\mathbf{k}$ through the identification $\mathbf{k} = k\,\boldsymbol{\nu}^{\mathrm{m}}$. We recognize here the leading order symbol, $k^2 8\pi^2 \Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}})$, of $\mathbf{N}$. Note the differences when compared with the 3-D case found in de Hoop & Brandsberg-Dahl (2000): the out-of-plane geometrical spreading $\mathcal{L}^{\perp}$ naturally appears in the leading order symbol of the normal operator in eq. (30) combined with the radiation patterns. Also the power of 2π has been modified in accordance with the stationary phase calculation in 2.5-D. Furthermore, in the formulation of Burridge *et al.* (1998), the weight μ_{LS} is equal to 1.

3.2 Least-squares inversion

To leading order, the parametrix of the normal operator composed with the normal operator from the last subsection should yield the identity. Any departure from the identitity is due to smoother contributions to the normal operator, as well as a consequence of having to take a generalized inverse of its symbol. We denote the generalized inverse of $\mathbf{N}$ by $\langle \mathbf{N}^{-1} \rangle$. The resolution is controlled by $\langle \mathbf{N}^{-1} \rangle \, \mathbf{N}$; general analysis of the resolution is given by de Hoop *et al.* (1999).

We note that

$$\frac{1}{(2\pi)^2} \int_{\mathbb{R}\geq 0} \int_S \exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\mathbf{x} - \mathbf{x}')] k \, dk \, d\boldsymbol{\nu}^{\mathrm{m}} = \delta(\mathbf{x} - \mathbf{x}'). \tag{32}$$

In view of limited illuminations, the support of k will be bounded while $\boldsymbol{\nu}^{\mathrm{m}} \in E_{\nu^{\mathrm{m}}} \subset S$. Eq. (32) becomes a 'band-limited' delta function, $\delta_{\mathrm{BL}}(\mathbf{x})$, reflecting the spatial resolution. The kernel of the parametrix of $\mathbf{N}$ (we denote the kernel of $\mathbf{N}$ as $\mathcal{N}$) is found to be

$$\langle \mathcal{N}^{-1} \rangle(\mathbf{x}', \mathbf{x}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}\geq 0} \int_{E_{\nu^{\mathrm{m}}}} k^{-2} \langle 8\pi^2 \Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}) \rangle^{-1} \exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\mathbf{x} - \mathbf{x}')] k \, dk \, d\boldsymbol{\nu}^{\mathrm{m}}. \tag{33}$$

Note that one of the k variables is placed with the integration variables since $(k, \boldsymbol{\nu}^{\mathrm{m}})$ yield polar coordinates in $\mathbf{k}$-space.

In Appendix C, we carry out the composition $\langle \mathbf{N}^{-1} \rangle \, \mathbf{L}^{\star} u$, to yield the estimate $\hat{\mathbf{c}}^{(1)}$ for $\mathbf{c}^{(1)}$. The result is

$$\hat{\mathbf{c}}^{(1)}(\mathbf{x}') \approx \frac{1}{(2\pi)^{5/2}} \mathrm{Re} \Bigg(\int_{E_{\nu^{\mathrm{m}}} \times E_\theta} d\theta \, d\boldsymbol{\nu}^{\mathrm{m}} \frac{\partial(\boldsymbol{\alpha}^{\mathrm{s}}, \boldsymbol{\alpha}^{\mathrm{r}})}{\partial(\boldsymbol{\nu}^{\mathrm{m}}, \theta)} \int_{\mathbb{R}\geq 0} \frac{d\omega'}{(\omega' \mathrm{i})^{1/2}} \mu_{\mathrm{LS}}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta) \langle \Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}) \rangle^{-1}$$

$$\times \frac{|\nabla_{\mathbf{x}'} T(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)|^2 \mathbf{w}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta) h_m(\mathbf{x}^{\mathrm{r}}; \mathbf{x}') U_{mn}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, \omega') h_n(\mathbf{x}^{\mathrm{s}}; \mathbf{x}')}{\rho^{(0)}(\mathbf{x}') A^{||}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}') A^{||}(\mathbf{x}', \mathbf{x}^{\mathrm{r}}) \mathcal{L}^{\perp}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)^{\star}} \exp[\mathrm{i}\omega' T(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)] \Bigg). \tag{34}$$

(We changed variables of integration in (C7) from phase directions to scattering angle θ and migration dip $\boldsymbol{\nu}^{\mathrm{m}}$ as in (28).) In this expression, $\mathbf{x}^{\mathrm{s}}$ and $\mathbf{x}^{\mathrm{r}}$ are functions of $(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta)$. The integral over migration dip makes this an inverse by the generalized Radon transform (GRT). In the GRT, the acquisition footprint is contained in $E_{\nu^{\mathrm{m}}} \times E_\theta$; the domain of integration E_θ is compensated for by θ-integration in Γ.

4 TRANSFORMATION INTO ANGLE COMMON-IMAGE-POINT GATHERS

From the derivation in Appendix C it seems that it is possible to generate an image from reflection data by integration over migration dip only. Indeed, in principle, we can generate an image trace for a given horizontal position for multiple scattering angles and collect them in a gather. When constructing a trace in these so-called angle common-image-point gathers we fix the scattering angle, θ, and integrate over the

migration dip, $\boldsymbol{\nu}^{\mathrm{m}}$. The restriction can introduce artefacts in the angle gathers when the medium is heterogeneous; namely, in the presence of caustics. These artefacts, which differ from the image artefacts caused by the failure of Assumption 3, are not present in the inversion by eq. (34) since they stack destructively (Stolk & de Hoop 2002).

Since the restriction to a fixed scattering angle means that we no longer stack over all the data but rather over subsets of the data that change with image point and scattering angle, we reintroduce $\mathbf{x}^{\mathrm{s}}$ and $\mathbf{x}^{\mathrm{r}}$ as the variables of integration. This requires the notion of branches of the two-way traveltime (see eq. 14). We define $\theta^{(i)}$ as $\theta(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}})$ (*cf.* eqs 27–28) composed with the inverse of map (23), associating the scattering angle with the acquisition coordinates $\mathbf{x}^{\mathrm{s}}$, $\mathbf{x}^{\mathrm{r}}$. We define the 'angle' transform $\mathbf{K}$ via a restriction of the imaging operator $\mathbf{L}^{\star}$ in eq. (22) to fixed angle $\theta^{(i)} = \theta'$, where i indexes the traveltime branch. We then reintroduce the sum over the different traveltime branches suppressed since eq. (12).

We obtain the kernel of the angle transform upon multiplying the kernel of $\mathbf{L}^{\star}$ in eq. (22) by a delta function, $\delta(\theta^{(i)} - \theta')$. The kernel of $\mathbf{K}$, denoted by $\mathcal{K}_{mn}$, can be written as an oscillatory integral (after a change of $\mathbf{x}'$ to $\mathbf{x}$)

$$\mathcal{K}_{mn}(\mathbf{x}, \theta', \mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, t) = \sum_{i\in I}(2\pi\sqrt{\mathrm{i}})^{-1}\int\int(\omega')^{3/2}\rho^{(0)}(\mathbf{x}) \times\frac{(A^{||}(\mathbf{x}^{\mathrm{s}}, \mathbf{x})A^{||}(\mathbf{x}, \mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})^{\star}}\mathbf{w}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})h_m(\mathbf{x}^{\mathrm{r}}; \mathbf{x})h_n(\mathbf{x}^{\mathrm{s}}; \mathbf{x})\exp\big[-\mathrm{i}\Phi^{(i)}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, t, \theta', \mathbf{x}, \omega', \varepsilon)\big]\,d\omega'\,d\varepsilon. \tag{35}$$

Here, ε is the Fourier dual of the scattering angle θ (Stolk & de Hoop 2002), and

$$\Phi^{(i)}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, t, \theta', \mathbf{x}, \omega', \varepsilon) = \omega'\big[T^{(i)}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) - t\big] + \varepsilon\big[\theta' - \theta^{(i)}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})\big]. \tag{36}$$

The artefacts of the restriction can be evaluated by considering the composition $\mathbf{K}\,\mathbf{L}$, which is a θ'-family of operators each member of which resembles the normal operator. The artefacts in the angle gathers can be recognized by their 'move-out' in angle. A multidimensional filter in the Fourier domain (see eq. 35) can be applied to remove the artefacts associated with $|\varepsilon| \geq \varepsilon_0 > 0$. Brandsberg-Dahl *et al.* (2003b) suppressed the artefacts by so-called *focusing in dip*.

By proceeding as suggested above using the appropriate changes of variables leading up to eq. (34), we obtain the angle-dependent parameter estimates (used to construct common-image-point gathers):

$$\hat{\mathbf{c}}^{(1)}(\mathbf{x}'; \theta') \approx \frac{1}{(2\pi)^{5/2}}\mathrm{Re}\Bigg(\int_{E_{\boldsymbol{\nu}^{\mathrm{m}}}} d\boldsymbol{\nu}^{\mathrm{m}}\frac{\partial(\boldsymbol{\alpha}^{\mathrm{s}}, \boldsymbol{\alpha}^{\mathrm{r}})}{\partial(\boldsymbol{\nu}^{\mathrm{m}}, \theta')}\int_{\mathbb{R}\geq 0}\frac{d\omega'}{(\omega'\mathrm{i})^{1/2}}\mu_{\mathrm{LS}}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')\langle\Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1} \times\frac{|\nabla_{\mathbf{x}'}T(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')|^2\mathbf{w}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')h_m(\mathbf{x}^{\mathrm{r}}; \mathbf{x}')U_{mn}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, \omega')h_n(\mathbf{x}^{\mathrm{s}}; \mathbf{x}')}{\rho^{(0)}(\mathbf{x}')A^{||}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}')A^{||}(\mathbf{x}', \mathbf{x}^{\mathrm{r}})\mathcal{L}^{\perp}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')^{\star}}\exp[-\mathrm{i}\omega'T(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')]\Bigg). \tag{37}$$

Note that we again have suppressed the summation of the traveltime branches. $\langle\Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}$ acts as a least-squares (LS) removal of the radiation patterns. Hence, $\hat{\mathbf{c}}^{(1)}$ in (37) still depends on scattering angle. Since, upon filtering, $\hat{\mathbf{c}}^{(1)}(\mathbf{x}'; \theta')$ should only depend smoothly on θ' if the correct background medium is used, the detection of smoothness can be used as a criterion for velocity analysis. Brandsberg-Dahl *et al.* (2003a) followed such a tomographic approach to determine the background medium.

5 EXAMPLE

In this section, we illustrate the relative significance (magnitude) of the factors making up the amplitude in 2.5-D imaging inversion according to eq. (37). We use ocean bottom cable (OBC) data (a single cable) from the North Sea, the same data we used in another paper, in which we developed and carried out reflection tomography in 2.5-D (Foss *et al.* 2004). We apply the transform in (37) both for *PP* and *PSV* reflection data and assume a transversely isotropic background medium with a vertical axis of symmetry.

Figs 4 and 5 show common-image-point gathers for *PP* and *PS* scattering events, respectively. The quantity that is displayed is obtained by (37) subject to replacing $\langle\Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}\,\mathbf{w}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')$ (the radiation pattern compensation) inside the transform by a normalization by $|\mathbf{w}(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}, \theta')|$. If $\mathbf{e} \equiv \mathbf{w}/|\,\mathbf{w}|$, the displayed quantity is approximately the parameter combination $\mathbf{e}^{\mathrm{T}}\mathbf{c}^{(1)}$, where $\mathbf{e}$ is evaluated at specular reflection. The vertical scale is in metres. Following convention, the common-image-point gathers are plotted as functions of incident *P*-wave angle, ranging from 0 to 45° (horizontal scale). In the mode-converted *PS* gathers, we plot negative angles to the left and positive angles to the right; note that the zero-angle contribution of the *PS* gathers is set to zero as this is not defined.

The geology in this example varies mostly in the imaging plane making this an accurate 2.5-D problem under our Assumptions 1 and 2. However, at depths greater than 3000 m, the 2.5-D framework seems to deteriorate. There the geology also varies significantly in the out-of-plane direction, scattering waves out of the computational plane. Hence, sufficient information on the reflectors here is simply not contained in our data.

We plot four cases, namely, from left to right: 2-D GRT, 2.5-D GRT with $\mathcal{L}^{\perp}$ set to 1, 2.5-D GRT, and 2.5-D GRT with $\langle\Gamma\rangle^{-1}\mathbf{w}$ removed. We focus on reflectors around 3000 m depth. The most pronounced effect observed going from the first to the second gather is the phase rotation (the factor $(\mathrm{i}\omega')^{1/2}$); comparing the second and third gathers, the amplitudes with angle clearly differ; comparing the third with the fourth gather, a dimming of amplitude with scattering angle is noticable, which can be explained by the absence of radiation pattern corrections in the fourth gather. Furthermore, we notice how well the reflectors correlate between the *PP* and *PS* gathers, simply a confirmation of the validity of the background medium.

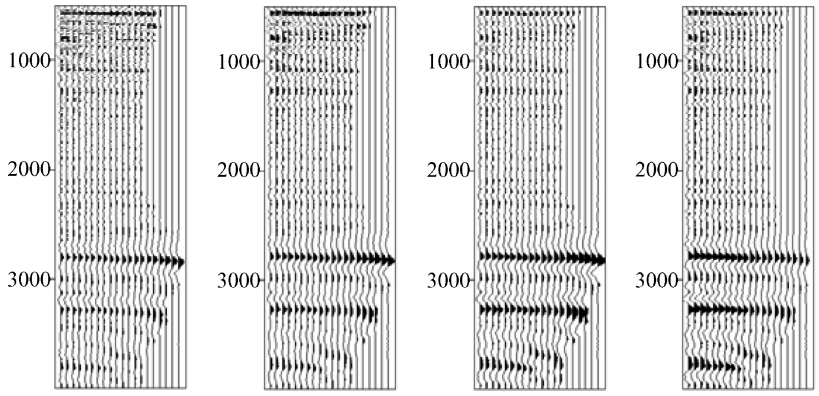

Figure 4. Common-image-point gathers in angle from *PP* OBC data. From left to right: 2-D GRT (in particular, the absence of the phase correction), 2.5-D GRT with $\mathcal{L}^{\perp}$ set to 1 (absence of out-of-plane geometrical spreading correction), 2.5-D GRT and 2.5-D GRT with $\langle \Gamma \rangle^{-1}$ **w** removed (absence of radiation pattern correction).

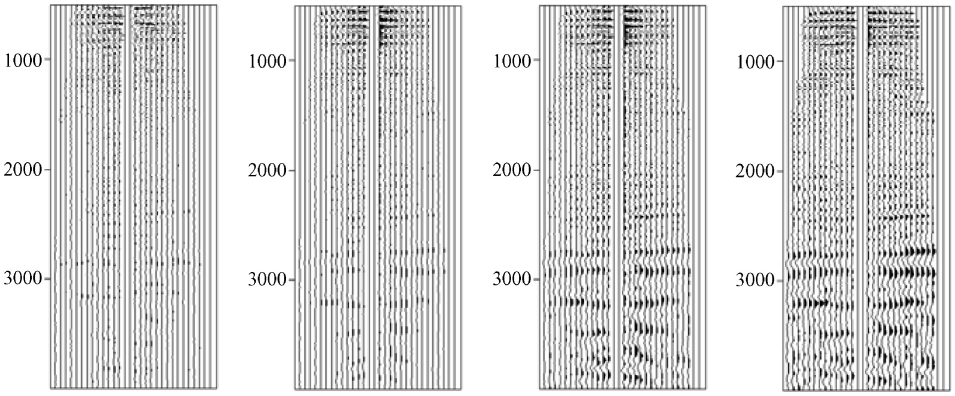

Figure 5. Common-image-point gathers in angle from *PS* OBC data. From left to right: 2-D GRT (in particular, the absence of the phase correction), 2.5-D GRT with $\mathcal{L}^{\perp}$ set to 1 (absence of out-of-plane geometrical spreading correction), 2.5-D GRT and 2.5-D GRT with $\langle \Gamma \rangle^{-1}$ **w** removed (absence of radiation pattern correction).

6 DISCUSSION

This paper presents the general form of 2.5-D modelling, imaging, parameter inversion and generalized Radon transform formulae in anisotropic elastic media under necessary and sufficient assumptions. The results extend those of Bleistein (1986) to the cases of multipathing and anisotropic elasticity. Assumption 1 restricts us to a monoclinic medium as the lowest possible symmetry, and the addition of Assumption 2 restricts wave propagation to the plane. Any symmetries lower than a transversely isotropic medium with the symmetry axis contained in the plane of consideration have to be excluded for the purpose of parameter inversion and reflection tomography. Our analysis is largely based on detailed symmetry considerations.

In the presence of caustics, artefacts may be generated by the imaging inversion procedures. Subject to the so-called Bolker condition, such artefacts are avoided. We have made the observation that the 2-D Bolker condition in the symmetry plane implies the Bolker condition in three dimensions in the framework of the common-azimuth acquisition geometry. Artefacts in the common-image-point gathers are also controlled by the ray geometry in the symmetry plane only and can thus be filtered out by 2-D considerations. An interesting aspect of our results is the appearance of out-of-plane contributions, due to the geometrical spreading contained naturally in the least-squares removal of the radiation patterns. Intuitively this can be understood as the removal of the stationary contribution to the radiation patterns from the out-of-plane geometrical spreading.

The primary application of our 2.5-D theory is the optimal use of single ocean bottom cable seismic reflection data. However, a second application is a slicewise approach to reflection tomography. We elaborate on this application. The transform of seismic data into common-image-point gathers in anisotropic media enables a slicewise approach to migration velocity analysis. Given a medium that has its dominant change in two directions we may invoke the 2.5-D assumptions approximately to slices in the smoothly varying direction. We parametrize the medium in a finite-dimensional subspace of possible smooth background models by a collection of slices $\underline{\mathbf{m}} = \{\mathbf{m}(x_2)\}$, for all values of x_2 under consideration, where each $\mathbf{m}(x_2)$ is the 2-D parametrization of a slice at x_2. Reflection tomography (Brandsberg-Dahl *et al.* 2003a) can be done per slice by differential semblance (Symes & Carazzone 1991) in angle on the common-image-point gathers generated by eq. (37).

We extend this by applying the appropriate out-of-plane 'annihilator' as a Tikhonov regularizer to the misfit function

$$J[\underline{\mathbf{m}}] = \frac{1}{2}\int\left(\int\int\left|\left[\partial_\theta - (\partial_\theta\,\Xi(\mathbf{x},\theta,x_2))\,\Xi(\mathbf{x},\theta,x_2)^{-1}\right]\hat{\mathbf{c}}^{(1)}[\underline{\mathbf{m}}](\mathbf{x};\theta,x_2)\right|^2 d\mathbf{x}\,d\theta + \lambda\int\left|\partial_{x_2}^2\mathbf{c}^{(0)}[\underline{\mathbf{m}}](\mathbf{x},x_2)\right|^2 d\mathbf{x}\right)dx_2, \tag{38}$$

where (*cf.* eq. 30)

$$\Xi(\mathbf{x},\theta,x_2) = \frac{\mu_{\mathrm{LS}}(\mathbf{x},\boldsymbol{\nu}^{\mathrm{m}},\theta)}{|\mathcal{L}^{\perp}(\mathbf{x},\boldsymbol{\nu}^{\mathrm{m}},\theta)|^2}\,\frac{\partial(\boldsymbol{\alpha}^{\mathrm{s}},\boldsymbol{\alpha}^{\mathrm{r}})}{\partial(\boldsymbol{\nu}^{\mathrm{m}},\theta)}\langle\Gamma(\mathbf{x},\boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}\mathbf{w}(\mathbf{x},\boldsymbol{\nu}^{\mathrm{m}},\theta)\mathbf{w}^{\mathrm{T}}(\mathbf{x},\boldsymbol{\nu}^{\mathrm{m}},\theta). \tag{39}$$

All relevant parameters of eq. (39) are calculated for $(\mathbf{x},\boldsymbol{\nu}^{\mathrm{m}})$ in the wavefront set of $\hat{\mathbf{c}}^{(1)}$. (Here, the wavefront set consists of the reflector positions and associated dips.) Furthermore, $\mathbf{c}^{(0)}[\underline{\mathbf{m}}](\mathbf{x},x_2)$ is the vector of parameters describing the background medium given the current parametrization $\underline{\mathbf{m}}$; λ in eq. (38) is a statistical quantity controlling the trade-off between in-plane and out-of-plane fit (Tenorio 2001) and $\mathbf{x} = (x_1, x_3)$; ∂_θ and ∂_{x_2} are the partial derivatives with respect to the scattering angle and out-of-plane coordinate, respectively. The outcome of the transform to common-image-point gathers (*cf.* 37) is denoted by $\hat{\mathbf{c}}^{(1)}[\underline{\mathbf{m}}](\mathbf{x};\theta,x_2)$ for the slice at the out-of-plane coordinate x_2. The minimum of this function indicates a smooth background model such that the data are in the range of the 2.5-D modelling operator.

ACKNOWLEDGMENTS

SKF would like to thank the URE-project, NTNU, Trondheim, Norway for financial support. The authors thank Statoil for providing the North Sea data, Børge Arntsen for the data handling, Kees Wapenaar for his valuable comments, and Barbara McLenon for her help preparing the manuscript.

REFERENCES

Amundsen, L. & Reitan, A., 1995. Decomposition of multicomponent sea-floor data into upgoing and downgoing P- and S-waves, *Geophysics,* **60,** 563–572.

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

Beylkin, G. & Burridge, R., 1990. Linearized inverse scattering problems in acoustics and elasticity, *Wave Motion,* **12,** 15–52.

Billette, F. & Lambaré, G., 1998. Velocity macro model estimation from seismic reflection sata by stereotomography, *Geophys. J. Int.,* **135,** 671–680.

Biondi, B. & Palacharla, G., 1996. 3-D prestack migration of common-azimuth data, *Geophysics,* **61,** 1822–1832.

Bleistein, N., 1986. Two-and-one-half dimensional in-plane wave propagation, *Geophys. Prospect.,* **34,** 686–703.

Bostock, M.G., Rondenay, S. & Shragge, J., 2001. Multiparameter two-dimensional inversion of scattered teleseismic body waves: 1. Theory for oblique incidence, *J. geophys. Res.,* **106,** 30 771–30 782.

Brandsberg-Dahl, S., Ursin, B. & de Hoop, M.V., 2003a. Seismic velocity analysis in the scattering-angle/azimuth domain, *Geophys. Prospect.,* **51,** 295–314.

Brandsberg-Dahl, S., de Hoop, M.V. & Ursin, B., 2003b. Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers, *Geophysics,* **68,** 232–254.

Burridge, R., de Hoop, M.V., Miller, D. & Spencer, C., 1998. Multiparameter inversion in anisotropic elastic media, *Geophys. J. Int.,* **134,** 757–777.

Červený, V., 1981. *Computation of Geometrical Spreading by Dynamic Ray Tracing,* Stanford Exploration Project (SEP) Report 28, pp. 49–59, SEP, Stanford, CA.

Červený, V., 2001. *Seismic Ray Theory,* Cambridge University Press, Cambridge.

Chapman, C.H. & Pratt, R.G., 1992, Traveltime tomography in anisotropic media-I. Theory, *Geophys. J. Int.,* **109,** 1–19.

de Hoop, M.V. & Brandsberg-Dahl, S., 2000. Maslov asymptotic extension of generalized Radon transform inversion in anisotropic elastic media: a least-squares approach, *Inverse Problems,* **16,** 519–562.

de Hoop, M.V. & Spencer, C., 1996. Quasi Monte-Carlo integration over $S^2 \times S^2$ for migration × inversion, 1996. *Inverse Problems,* **12,** 219–239.

de Hoop, M.V., Spencer, C. & Burridge, R., 1999. The resolving power of seismic amplitude data: an anisotropic inversion/migration approach, *Geophysics,* **64,** 852–873.

de Hoop, M.V., Le Rousseau, J.H. & Biondi, B., 2003. Symplectic structure of wave-equation imaging: a path-integral approach based on the double-square-root equation, *Geophys. J. Int.,* **153,** 52–74.

Dellinger, J.A., Gray, S.H., Murphy, G.E. & Etgen, J.T., 2000. Efficient 2.5-D true-amplitude migration, *Geophysics,* **65,** 943–950.

Duistermaat, J.J., 1996. *Fourier Integral Operators,* Birkhäuser, Boston.

Ettrich, N., Sollid, A. & Ursin, B., 2002. Out-of-plane geometrical spreading in anisotropic media, *Geophys. Prospect.,* **50,** 383–392.

Foss, S.K. & Ursin, B., 2004. 2.5-D modeling, inversion and angle migration in anisotropic elastic media, *Geophys. Prospect.,* **52,** 65–84.

Foss, S.K., Ursin, B. & de Hoop, M.V., 2005. Depth-consistent reflection tomography using PP and PS seismic data, *Geophysics,* in print.

Geoltrain, S., 1989. Asymptotic solutions to direct and inverse scattering in anisotropic media, *PhD thesis*, Colorado School of Mines, Golden, CO.

Goldin, S.V., 1986. *Seismic Traveltime Inversion,* SEG Investigations in Geophysics No 1, Society of Exploration Geophysicists, Tulsa, OK.

Guillemin, V., 1985. On some results of Gelfand in integral geometry, *Proc. Symp. Pure Math., Am. Math. Soc.,* **43,** 149–155.

Hansen, S., 1991. Solution of a hyberbolic inverse problem by linearization, *Comm. Partial Differential Equations,* **16,** 291–309.

Hörmander, L., 1985. *The Analysis of Linear Partial Differential Operators,* Vol. 3, Springer, Berlin.

Klimeš, L., 1997. Phase shift of the Green function due to caustics in anisotropic media, *Expanded Abstracts, 68th Annual International Meeting, SEG,* pp. 1834–1837, Society of Exploration Geophysicists, Tulsa, OK.

Miller, D., Oristaglio, M. & Beylkin, G., 1987. A new slant on seismic imaging: migration and integral geometry, *Geophysics,* **52,** 943–964.

Musgrave, M.J.P., 1970. *Crystal Acoustics,* Holden-Day, San Francisco, CA.

Nolan, C.J. & Symes, W.W., 1997. Global solution of a linearized inverse problem for the wave equation. *Comm. Partial Differential Equations,* **22,** 919–952.

Rakesh, 1988. A linearised inverse problem for the wave equation, *Comm. Partial Differential Equations,* **13,** 573–601.

Sollid, A. & Ursin, B., 2003. Scattering-angle migration of {OBS} data in weakly anisotropic media, *Geophysics,* **68,** 641–655.

Stolk, C.C., 2000. Microlocal analysis of a seismic linearized inverse problem, *Wave Motion,* **32,** 267–290.

Stolk, C.C., 2002. Microlocal analysis of the scattering angle transform, *Comm. Partial Differential Equations,* **27,** 1879–1900.

Stolk, C.C. & de Hoop, M.V., 2002. Microlocal analysis of seismic inverse scattering in anisotropic elastic media, *Comm. Pure Appl. Math.,* **55,** 261–301.

Symes, W. & Carazzone, J., 1991. Velocity inversion by differential semblance optimization, *Geophysics,* **56,** 654–663.

Taylor, M.E., 1975. Reflection of singularities of solutions to systems of differential equations, *Comm. Pure Appl. Math.,* **28,** 457–478.

Ten Kroode, A.P.E., Smit, D.J. & Verdel, A.R., 1998. A microlocal analysis of migration, *Wave Motion,* **28,** 149–172.

Tenorio, L., 2001. Statistical regularization of inverse problems, *SIAM Rev.,* **43,** 347–366.

Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics,* **51,** 1954–1966.

Treves, F., 1980. *Introduction to Pseudodifferential and Fourier Integral Operators,* Vol. 1, Plenum Press, New York.

Tygel, M., Schleicher, J., Hubral, P. & Santos, L.T., 1998. 2.5-D true-amplitude Kirchhoff migration to zero offset in laterally inhomogeneous media, *Geophysics,* **63,** 557–573.

APPENDIX A: THE 2.5-D MODELLING

The m-component of the scattered field displacement at the receiver position $\mathbf{x}^{\mathrm{r}}$ due to a n-component source at $\mathbf{x}^{\mathrm{s}}$ in the short-period or high-frequency Born approximation in 3-D is given by

$$U_{mn}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}^{\mathrm{s}}, \omega) \approx \int_X \left(\int_{\mathbb{R}} \omega^2 h_m(\mathbf{x}^{\mathrm{r}}) \rho^{(0)}(\mathbf{x}) A(\mathbf{x}^{\mathrm{s}}, \mathbf{x}) A(\mathbf{x}, \mathbf{x}^{\mathrm{r}}) \mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) \exp[\mathrm{i}\omega T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})] h_n(\mathbf{x}^{\mathrm{s}})\, dx_2 \right) \times \mathbf{c}^{(1)}(x_1, x_3)\, dx_1\, dx_3, \tag{A1}$$

where the amplitude A is given in (3) and the two-way traveltime is given in (14). We have accounted for the fact that $\mathbf{c}^{(1)}$ is independent of x_2 under Assumption 1. The medium perturbation $\mathbf{c}^{(1)}$ and the radiation patterns $\mathbf{w}$ are, for the lowest possible symmetry (triclinic), 22×1 matrices of the form as in eqs (16) and (17), respectively. The domain of integration of the (x_1, x_3)-coordinates, X, is defined in Section 2.2.

Upon scaling, $x_2 = \tilde{x}_2/|\omega|$, we recognize the phase variables $(\tilde{x}_2, \omega)$. We proceed as in Bleistein (1986) and use the method of stationary phase to integrate out the out-of-plane variable x_2 in (A1). The 1-D stationary phase formula approximates integrals of the type

$$\int f(\sigma) \exp(\mathrm{i}\omega T(\sigma))\, d\sigma \approx \sqrt{\frac{2\pi}{|\omega||\partial_\sigma^2 T(\sigma_0)|}}\, f(\sigma_0) \exp[\mathrm{i}\omega T(\sigma_0) + \mathrm{i}(\pi/4)\mathrm{sgn}(\omega)\mathrm{sgn}(\partial_\sigma^2 T(\sigma_0))], \tag{A2}$$

for sufficiently large $|\omega|$ and $\partial_\sigma^2 T(\sigma_0) \neq 0$, where σ_0 is the stationary point satisfying $\partial_\sigma T(\sigma)|_{\sigma=\sigma_0} = 0$. In the integral in (A1), σ is identified with x_2, and the stationary point is given by

$$\partial_{x_2} T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) = p_2^{\mathrm{s}} + p_2^{\mathrm{r}} = 0. \tag{A3}$$

By Assumptions 1 and 2, the only solution to (A3) is $p_2^{\mathrm{s}} = p_2^{\mathrm{r}} = 0$, which implies that the stationary point is $x_2 = 0$. We observe that the sum of the slownesses in eq. (A3) also occurs in the common-azimuth case, eq. (19).

The second derivative of the phase function at the stationary point is

$$\partial_{x_2}^2 T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})\big|_{x_2=0,\, p_2=0} = \left(\partial_{x_2} p_2^{\mathrm{s}} + \partial_{x_2} p_2^{\mathrm{r}}\right)\big|_{x_2=0,\, p_2=0} = \frac{1}{Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{s}})} + \frac{1}{Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{r}})}, \tag{A4}$$

where $Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$ and $Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{r}})$ are the out-of-plane geometrical spreading factors defined in eq. (5) for the rays connecting the image point $\mathbf{x}$ with the source $\mathbf{x}^{\mathrm{s}}$ and receiver $\mathbf{x}^{\mathrm{r}}$, respectively.

Remark 4

At points in phase space where either $Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{s}})$ or $Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{r}})$ tends to zero, the traveltime function is not smooth. We observe this in eq. (A4) where the Hessian of the traveltime function will tend towards infinity and will not be defined. This essentially means that the stationary phase argument does not hold. The integral over x_2 in eq. (A1) remains. Out-of-plane caustics are thus not allowed in order for the stationary phase formula to be applicable. We restrict the analysis in the following to rays with no out-of-plane caustics.

From this it follows that $\mathrm{sgn}(\partial_{x_2}^2 T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})|_{x_2=0}) = 1$ because the out-of-plane geometrical spreading is positive. The stationary phase formula (A2) for the x_2 integral (in view of Remark 4) then yields

$$\begin{aligned} &\int_X \left(\omega^2 \int_{\mathbb{R}} h_m(\mathbf{x}^{\mathrm{r}}) \rho^{(0)}(\mathbf{x}) A(\mathbf{x}^{\mathrm{s}}, \mathbf{x}) A(\mathbf{x}, \mathbf{x}^{\mathrm{r}}) \mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) \mathbf{c}^{(1)}(\mathbf{x}) \exp[\mathrm{i}\omega T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})] h_n(\mathbf{x}^{\mathrm{s}})\, dx_2 \right) dx_1\, dx_3 \\ &\approx \sqrt{2\pi \mathrm{i}}\, \omega^{3/2} \int_X h_m(\mathbf{x}^{\mathrm{r}}) \frac{(Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{s}}) Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{r}}))^{1/2}}{(Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{s}}) + Q_2^{\perp}(\mathbf{x}, \mathbf{x}^{\mathrm{r}}))^{1/2}} \rho^{(0)}(\mathbf{x}) A(\mathbf{x}^{\mathrm{s}}, \mathbf{x}) A(\mathbf{x}, \mathbf{x}^{\mathrm{r}}) \\ &\times \mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) \mathbf{c}^{(1)}(\mathbf{x}) \exp[\mathrm{i}\omega T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})] h_n(\mathbf{x}^{\mathrm{s}})\big|_{x_2=0,\, p_2=0}\, dx_1\, dx_3. \end{aligned} \tag{A5}$$

Using eqs (3) and (6), this equation reduces to the 2.5-D modelling formula in the main text (here in the frequency domain),

$$U_{mn}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, \omega) = \sqrt{2\pi \mathrm{i}}\, \omega^{3/2} \int_X h_m(\mathbf{x}^{\mathrm{r}}) \rho^{(0)}(\mathbf{x}) \frac{A^{||}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}) A^{||}(\mathbf{x}, \mathbf{x}^{\mathrm{r}})}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})} \mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}}) \mathbf{c}^{(1)}(\mathbf{x}) \exp[\mathrm{i}\omega T(\mathbf{x}^{\mathrm{r}}, \mathbf{x}, \mathbf{x}^{\mathrm{s}})] h_n(\mathbf{x}^{\mathrm{s}})\, d\mathbf{x}, \tag{A6}$$

where $\mathcal{L}^{\perp}$ is defined in (13).

APPENDIX B: THE NORMAL OPERATOR

In this appendix we derive an oscillatory integral representation for the kernel of the normal operator. To this end, the u_{mn} from the modelling eq. (12) is inserted into (22) to yield

$$\begin{aligned}\mathbf{L}^{\star}\psi\mathbf{L}\mathbf{c}^{(1)}(\mathbf{x}') \approx \frac{1}{2\pi}\int_{\Sigma}\int_{X}\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}_{\geq 0}}(\omega')^{3/2}\omega^{3/2}\rho^{(0)}(\mathbf{x}')\rho^{(0)}(\mathbf{x})\frac{(A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{\|}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})^{\star}}\frac{A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x})A^{\|}(\mathbf{x},\mathbf{x}^{\mathrm{r}})}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})}\\ \times\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})\mathbf{c}^{(1)}(\mathbf{x})h_m(\mathbf{x}^{\mathrm{r}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')h_n(\mathbf{x}^{\mathrm{s}})h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')\\ \times\exp\{\mathrm{i}\omega[T(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})-t]-\mathrm{i}\omega'[T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})-t]\}\,dt\,d\omega\,d\omega'\,d\mathbf{x}\,d\mathbf{x}^{\mathrm{s}}\,d\mathbf{x}^{\mathrm{r}}.\end{aligned}\tag{B1}$$

If Assumption 3 is satisfied, this composition is an elliptic pseudodifferential operator. This justifies the expansions we make below.

The integral over t yields a delta function $\delta(\omega-\omega')$ so we can collapse the integral over ω' and set $\omega=\omega'$. Invoking a Taylor expansion about $\mathbf{x}'$ for the two-way traveltime $T(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})$ yields,

$$\omega[T(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})-T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})]=\omega[(\nabla_{\mathbf{x}'}T)(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\cdot(\mathbf{x}-\mathbf{x}')]+\ldots,\tag{B2}$$

due to the fact that higher-order derivatives give smoother contributions to the amplitude (Hörmander 1985) through expansion of the exponential, Inserting (B2) in the phase function of (B1) yields

$$\begin{aligned}\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \int_{\Sigma}\int_{X}\int_{\mathbb{R}}\omega^3\rho^{(0)}(\mathbf{x}')\rho^{(0)}(\mathbf{x})\frac{(A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{\|}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})^{\star}}\frac{A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x})A^{\|}(\mathbf{x},\mathbf{x}^{\mathrm{r}})}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})}\\ \times\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})\mathbf{c}^{(1)}(\mathbf{x})h_m(\mathbf{x}^{\mathrm{r}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')h_n(\mathbf{x}^{\mathrm{s}})h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')\\ \times\exp[\mathrm{i}\omega\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\cdot(\mathbf{x}-\mathbf{x}')]\,d\omega\,d\mathbf{x}\,d\mathbf{x}^{\mathrm{s}}d\mathbf{x}^{\mathrm{r}}.\end{aligned}\tag{B3}$$

The most dominant contribution to the oscillatory integral (B3) occurs when $\mathbf{x}$ and $\mathbf{x}'$ are near one another making up a small neighbourhood of support (Beylkin 1985). Additionally, the amplitudes, polarization vectors and the radiation patterns vary slowly compared with the exponential contribution. To leading order, we may therefore assume that they are constant in a small neighbourhood of $\mathbf{x}'$. Since the polarization vectors are normalized, we obtain

$$\begin{aligned}\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \int_{\Sigma}\int_{X}\int_{\mathbb{R}}\omega^3\rho^{(0)}(\mathbf{x}')^2\frac{|A^{\|}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{\|}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))|^2}{|\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})|^2}\\ \times\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\mathbf{w}^{\mathrm{T}}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\mathbf{c}^{(1)}(\mathbf{x})\exp[\mathrm{i}\omega\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\cdot(\mathbf{x}-\mathbf{x}')]\,d\omega\,d\mathbf{x}\,d\mathbf{x}^{\mathrm{s}}\,d\mathbf{x}^{\mathrm{r}}.\end{aligned}\tag{B4}$$

We change coordinates of integration through the mapping in equation (23) in all relevant places. The in-plane geometrical spreading of $A^{\|}$ and the Jacobian cancel each other up to projection factors. Using the expression for the Jacobian and (25) in eq. (B4) yields

$$\begin{aligned}\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \int_{S^{\mathrm{s}}\times S^{\mathrm{r}}}\int_{X}\int_{\mathbb{R}}\frac{\omega^3}{|\mathcal{L}^{\perp}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})|^2}\mathbf{w}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\mathbf{w}^{T}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\\ \times\mathbf{c}^{(1)}(\mathbf{x})\exp[\mathrm{i}\omega\nabla_{\mathbf{x}'}T(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\cdot(\mathbf{x}-\mathbf{x}')]\mathcal{J}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\,d\omega\,d\mathbf{x}\,d\boldsymbol{\alpha}^{\mathrm{s}}\,d\boldsymbol{\alpha}^{\mathrm{r}}.\end{aligned}\tag{B5}$$

In accordance with the reasoning in Section 3.1, we write eq. (B5) as an integral over $\mathbf{x}$ and its Fourier dual $\mathbf{k}=\omega\,\nabla_{\mathbf{x}}T$; i.e. in the form of a pseudodifferential operator. We introduce a new 'frequency' $k=\omega|\nabla_{\mathbf{x}'}T|$. Then $\omega\,\nabla_{\mathbf{x}'}T=k\,\boldsymbol{\nu}^{\mathrm{m}}$, where $\boldsymbol{\nu}^{\mathrm{m}}=\nabla_{\mathbf{x}'}T/|\nabla_{\mathbf{x}'}T|$ is the migration dip and is a function of $\boldsymbol{\alpha}^{\mathrm{r}}$, $\mathbf{x}'$, $\boldsymbol{\alpha}^{\mathrm{s}}$. With the appropriate substitution this yields

$$\begin{aligned}\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \int_{S^{\mathrm{s}}\times S^{\mathrm{r}}}\int_{X}\int_{\mathbb{R}}\frac{|\nabla_{\mathbf{x}'}T(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})|^{-4}k^3}{|\mathcal{L}^{\perp}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})|^2}\mathbf{w}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\mathbf{w}^{\mathrm{T}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\\ \times\mathbf{c}^{(1)}(\mathbf{x})\exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\cdot(\mathbf{x}-\mathbf{x}')]\mathcal{J}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\,dk\,d\mathbf{x}\,d\boldsymbol{\alpha}^{\mathrm{s}}\,d\boldsymbol{\alpha}^{\mathrm{r}}.\end{aligned}\tag{B6}$$

Making use of the quantity defined in (26) in this equation, results in the simplified expression

$$\begin{aligned}\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \int_{S^{\mathrm{s}}\times S^{\mathrm{r}}}\int_{X}\int_{\mathbb{R}}k^3\frac{\mu_{\mathrm{LS}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})}{|\mathcal{L}^{\perp}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})|^2}\mathbf{w}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\mathbf{w}^{\mathrm{T}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\\ \times\mathbf{c}^{(1)}(\mathbf{x})\exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})\cdot(\mathbf{x}-\mathbf{x}')]\,dk\,d\mathbf{x}\,d\boldsymbol{\alpha}^{\mathrm{s}}\,d\boldsymbol{\alpha}^{\mathrm{r}}.\end{aligned}\tag{B7}$$

We change variables of integration again, from the phase directions to scattering angle θ and the migration dip $\boldsymbol{\nu}^{\mathrm{m}}$, according to mapping (28). We finally obtain

$$\begin{aligned}\mathbf{N}\mathbf{c}^{(1)}(\mathbf{x}') \approx \int_{E_{\boldsymbol{\nu}^{\mathrm{m}}}}\int_{\mathbb{R}}\int_{X}k^3\left(\int_{E_{\theta}}\frac{\mu_{\mathrm{LS}}(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}},\theta)}{|\mathcal{L}^{\perp}(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}},\theta)|^2}\mathbf{w}(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}},\theta)\mathbf{w}^{\mathrm{T}}(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}},\theta)\frac{\partial(\boldsymbol{\alpha}^{\mathrm{s}},\boldsymbol{\alpha}^{\mathrm{r}})}{\partial(\boldsymbol{\nu}^{\mathrm{m}},\theta)}\,d\theta\right)\\ \times\mathbf{c}^{(1)}(\mathbf{x})\exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\mathbf{x}-\mathbf{x}')]\,d\mathbf{x}\,dk\,d\boldsymbol{\nu}^{\mathrm{m}}.\end{aligned}\tag{B8}$$

APPENDIX C: ESTIMATING THE MEDIUM CONTRAST

In this appendix we derive the least-squares estimate of the medium perturbation. We use eqs (21) and (22).

We replace $\mathbf{x}'$ with $\mathbf{x}$ in the imaging process (22). The real part of the composition is symmetric in ω' while the imaginary part is odd and vanishes. The integration over ω' then becomes one-sided by taking the real part and multiplying by 2 (de Hoop & Brandsberg-Dahl 2000). Using eq. (21), we find that

$$\begin{aligned}\hat{\mathbf{c}}^{(1)}(\mathbf{x}') &= \langle \mathbf{N}^{-1}\rangle \mathbf{L}^{\star} u(\mathbf{x}')\\ &\approx \frac{-1}{(2\pi)^2}\frac{1}{\pi}\mathrm{Re}\Bigg(\int_X d\mathbf{x}\int_{\mathbb{R}\geq 0} k^{-1}\,dk\int_{E_{\nu^{\mathrm{m}}}} d\boldsymbol{\nu}^{\mathrm{m}}\int_{\Sigma} d\mathbf{x}^{\mathrm{s}}\,d\mathbf{x}^{\mathrm{r}}\int_{\mathbb{R}\geq 0} d\omega'\langle 8\pi^2\Gamma(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}\\ &\quad\times\sqrt{2\pi}(\mathrm{i}\omega')^{3/2}\rho^{(0)}(\mathbf{x})\frac{(A^{||}(\mathbf{x}^{\mathrm{s}},\mathbf{x})A^{||}(\mathbf{x},\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})^{\star}}\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x})U_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},\omega')h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x})\\ &\quad\times\exp[\mathrm{i}k\boldsymbol{\nu}^{\mathrm{m}}(\mathbf{x}-\mathbf{x}')-\mathrm{i}\omega' T(\mathbf{x}^{\mathrm{r}},\mathbf{x},\mathbf{x}^{\mathrm{s}})]\Bigg).\end{aligned}\tag{C1}$$

We have introduced the time-Fourier transform $U_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},\omega')$ of the data, $u_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},t')$. We proceed by Taylor expanding $T(\mathbf{x}^{\mathrm{s}},\mathbf{x},\mathbf{x}^{\mathrm{r}})$ about $\mathbf{x}'$ as in eq. (B2). By the same argument as around eq. (B4), $\mathbf{x}'$ is assumed to be close to $\mathbf{x}$ as this yields the largest contribution to the oscillatory integral. The amplitude factors and the radiation patterns are slowly varying and can be considered constant in a small neighbourhood of $\mathbf{x}'$ to leading order. This means that we may substitute $\mathbf{x}'$ for $\mathbf{x}$ for slowly varying components,

$$\begin{aligned}\hat{\mathbf{c}}^{(1)}(\mathbf{x}') &\approx \frac{-1}{(2\pi)^{3/2}\pi}\mathrm{Re}\Bigg(\int_X d\mathbf{x}\int_{\mathbb{R}\geq 0} k^{-1}\,dk\int_{E_{\nu^{\mathrm{m}}}} d\boldsymbol{\nu}^{\mathrm{m}}\int_{\Sigma} d\mathbf{x}^{\mathrm{s}}\,d\mathbf{x}^{\mathrm{r}}\int_{\mathbb{R}\geq 0} d\omega'\langle 8\pi^2\Gamma(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}\\ &\quad\times(\mathrm{i}\omega')^{3/2}\rho^{(0)}(\mathbf{x}')\frac{(A^{||}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{||}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})^{\star}}\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')U_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},\omega')h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')\\ &\quad\times\exp[-\mathrm{i}\omega' T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})]\exp\{\mathrm{i}[k\boldsymbol{\nu}^{\mathrm{m}}-\omega'\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})]\cdot(\mathbf{x}-\mathbf{x}')\}\Bigg).\end{aligned}\tag{C2}$$

We observe that the integration over $\mathbf{x}$ now becomes a delta function,

$$\begin{aligned}\int_X d\mathbf{x}\exp\{\mathrm{i}[k\boldsymbol{\nu}^{\mathrm{m}}-\omega'\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})]\cdot(\mathbf{x}-\mathbf{x}')\} &= 4\pi^2\delta[k\boldsymbol{\nu}^{\mathrm{m}}-\omega'\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})]\\ &= 4\pi^2\delta[k\boldsymbol{\nu}^{\mathrm{m}}-\omega'\boldsymbol{\nu}^{\mathrm{m}}|\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})|],\end{aligned}\tag{C3}$$

since $\boldsymbol{\nu}^{\mathrm{m}}=\nabla_{\mathbf{x}'}T/|\nabla_{\mathbf{x}'}T|$. Using that $k=\omega\,|\nabla_{\mathbf{x}'}T|$ and the result in eq. (C2), the integration over k and $\boldsymbol{\nu}^{\mathrm{m}}$ collapses to

$$\begin{aligned}\hat{\mathbf{c}}^{(1)}(\mathbf{x}') &\approx \frac{-\sqrt{2\pi}}{\pi}\mathrm{Re}\Bigg(\int_{\Sigma} d\mathbf{x}^{\mathrm{s}}d\mathbf{x}^{\mathrm{r}}\int_{\mathbb{R}\geq 0}\frac{d\omega'}{(\omega'|\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})|)^2}\langle 8\pi^2\Gamma(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}(\mathrm{i}\omega')^{3/2}\\ &\quad\times\rho^{(0)}(\mathbf{x}')\frac{(A^{||}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{||}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})^{\star}}\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')U_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},\omega')h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')\\ &\quad\times\exp[-\mathrm{i}\omega' T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})\Bigg).\end{aligned}\tag{C4}$$

Taking 8π out of the generalized inverse and making the appropriate changes yields

$$\begin{aligned}\hat{\mathbf{c}}^{(1)}(\mathbf{x}') &\approx \frac{1}{(2\pi)^{5/2}}\mathrm{Re}\Bigg(\int_{\Sigma} d\mathbf{x}^{\mathrm{s}}\,d\mathbf{x}^{\mathrm{r}}\int_{\mathbb{R}\geq 0}\frac{d\omega'}{(\omega'\mathrm{i})^{1/2}|\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})|^2}\langle\Gamma(\mathbf{x}',\boldsymbol{\nu}^{\mathrm{m}})\rangle^{-1}\\ &\quad\times\rho^{(0)}(\mathbf{x}')\frac{(A^{||}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{||}(\mathbf{x}',\mathbf{x}^{\mathrm{r}}))^{\star}}{\mathcal{L}^{\perp}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})^{\star}}\mathbf{w}(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})h_m(\mathbf{x}^{\mathrm{r}};\mathbf{x}')U_{mn}(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}},\omega')h_n(\mathbf{x}^{\mathrm{s}};\mathbf{x}')\\ &\quad\times\exp[\mathrm{i}\omega' T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})]\Bigg).\end{aligned}\tag{C5}$$

(Eq. (C5) is a direct manifestation of the composition of a pseudodifferential operator with a FIO (Treves 1980, Section 6.1, Chapter VIII)). The surface coordinates $\mathbf{x}^{\mathrm{r}}$ and $\mathbf{x}^{\mathrm{s}}$ are changed to the phase directions at the scattering point as in relation (23). Using eqs (25) and (26), the appropriate Jacobian can be expressed as

$$\frac{\partial(\mathbf{x}^{\mathrm{s}},\mathbf{x}^{\mathrm{r}})}{\partial(\boldsymbol{\alpha}^{\mathrm{s}},\boldsymbol{\alpha}^{\mathrm{r}})}=\frac{\mu_{\mathrm{LS}}(\boldsymbol{\alpha}^{\mathrm{r}},\mathbf{x}',\boldsymbol{\alpha}^{\mathrm{s}})|\nabla_{\mathbf{x}'}T(\mathbf{x}^{\mathrm{r}},\mathbf{x}',\mathbf{x}^{\mathrm{s}})|^4}{|\rho^{(0)}(\mathbf{x}')A^{||}(\mathbf{x}^{\mathrm{s}},\mathbf{x}')A^{||}(\mathbf{x}',\mathbf{x}^{\mathrm{r}})|^2}.\tag{C6}$$

Inserting the result in eq. (C5) yields the estimate of the medium perturbation

$$\hat{\mathbf{c}}^{(1)}(\mathbf{x}') \approx \frac{1}{(2\pi)^{5/2}} \mathrm{Re}\Bigg(\int_{S^{\mathrm{s}} \times S^{\mathrm{r}}} d\boldsymbol{\alpha}^{\mathrm{s}}\, d\boldsymbol{\alpha}^{\mathrm{r}} \int_{\mathbb{R}_{\geq 0}} \frac{d\omega'}{(\omega' \mathrm{i})^{1/2}} \mu_{\mathrm{LS}}(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}}) \langle \Gamma(\mathbf{x}', \boldsymbol{\nu}^{\mathrm{m}}) \rangle^{-1}$$

$$\times \frac{|\nabla_{\mathbf{x}'} T(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}})|^2 \mathbf{w}(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}}) h_m(\mathbf{x}^{\mathrm{r}}; \mathbf{x}') U_{mn}(\mathbf{x}^{\mathrm{s}}, \mathbf{x}^{\mathrm{r}}, \omega') h_n(\mathbf{x}^{\mathrm{s}}; \mathbf{x}')}{\rho^{(0)}(\mathbf{x}') A^{\|}(\boldsymbol{\alpha}^{\mathrm{s}}, \mathbf{x}') A^{\|}(\mathbf{x}', \boldsymbol{\alpha}^{\mathrm{r}}) \mathcal{L}^{\perp}(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}})^{\star}} \exp[\mathrm{i}\omega' T(\boldsymbol{\alpha}^{\mathrm{r}}, \mathbf{x}', \boldsymbol{\alpha}^{\mathrm{s}})] \Bigg). \tag{C7}$$

In the denominator, we observe the out-of-plane geometrical spreading, which also exists in the expression for Γ (eq. 30).