Velocity analysis in the common scattering-angle/azimuth domain

Sverre Brandsberg-Dahl , Maarten V. de Hoop and Bjørn Ursin*
Center for Wave Phenomena
* *Department of Petroleum Engineering and Applied Geophysics Norwegian University of Science and Technology*

ABSTRACT

An asymptotic inversion of seismic data in a heterogeneous anisotropic elastic medium, carried out using the Generalized Radon Transform (GRT) in the Kirchhoff approximation, is naturally suited to the common scattering-angle/azimuth domain. A combination of the forward and inverse operators with a differential operator in the scattering-angle/azimuth domain gives the sensitivity transform. Applying this transform to the data after the inversion/migration, can give an estimate of how good the current background model is based on the flatness of events in the common scattering-angle/azimuth image gathers. The differential semblance operator constructed from the sensitivity transform, forms an objective function that we seek to minimize with respect to the background medium. The gradient of this operator is the key part of the minimization process.

Key words: GRT, common scattering-angle/azimuth domain, differential semblance

Introduction

One of the core problems in seismic imaging and inversion is to estimate a correct background model. To create an image based on the assumption of a linearized scattering problem, a correct smooth background model is crucial when calculating the Green's functions needed in any imaging algorithm. In general the problem is the same regardless of assumptions about the media and experiment configuration; we need to find a smooth model that prescribes a wave propagation that is similar to the propagation the real data have experienced. The problem of estimating this background model has in general been cast as an optimization problem, where the goal is to minimize some objective function with respect to the parameters in the background model. The optimization can be either global (Tarantola, 1987), trying to fit the model to the data in one complex operation, or local where only a small portion of the model is treated at a time (Deregowski, 1990). Most of the methods in use in the industry today rely heavily on human interaction, such as tideous picking of events and interpretation of curvatures in common-image gathers. In areas with complex geology, accurate velocity analysis is performed through an iterative run of pre-stack depth migrations and picking/interpretation, something that can contribute a great portion of the total processing cost of a data set. Here, we present a method that is based on local optimization and that is more or less automatic and does not require any picking.

Viewed as a kinematic problem the goal is to match the traveltimes in the model with those observed in the field data. Typically, offset has been used as the controlling parameter in this process, leading to procedures such as NMO velocity analysis and residual moveout analysis . In these techniques the traveltime for a given seismic event is matched for varying source/receiver offsets in an attempt to estimate the velocity in the subsurface. In general the velocity estimation problem appears in its simplest form in the acoustic approximation, where only one velocity needs to be estimated for the background model. This approach, however, is limited by the assumptions made when choosing the physical Earth model. First, the Earth is not a fluid, so the chosen model

might be inappropriate when we seek to describe the wave propagation and velocity with great detail. Second, in the acoustic approximation we are limited to using only one wave mode, P-waves. Even though most seismic exploration today is performed successfully using the acoustic approximation, the job of finding a good background model might be easier after all, if instead several physical parameters are incorporated into the model. A natural extension is to use an elastic medium, isotropic or anisotropic. Using an elastic medium will in addition allow for the use of several different wave modes in the estimation process. It is well known that the more wave modes that are used, say, qP, qSV and qSH, the more information can be retrieved from the data (Beylkin & Burridge, 1984). Although use of several wave modes apparently complicates the problem somewhat, the processes involved are in fact more or less the same regardless of the mode and the medium under consideration. Probably the only important factor is the increase in computational cost introduced by using a multi-mode approach. Hence, if it is possible to use additional wave modes we should incorporate them into our analysis and thereby also increase the number of parameters we can hope to estimate.

In this paper we introduce a data domain in which we can do both imaging/inversion and modeling based on a micro-local coordinate representation of the seismic data. This data domain will also be utilized in the velocity analysis that will be developed. We will use the generalized Radon transform (GRT) to invert seismic data, based on an integral over phase directions $\overset{s}{\boldsymbol{\alpha}}$ and $\overset{r}{\boldsymbol{\alpha}}$ at the image point. This naturally introduces the image-point related coordinates: migration dip $\overset{0}{\boldsymbol{\alpha}}$, scattering-angle θ, and azimuth ψ. The inversion procedure is formalized in the Kirchhoff-Born approximation and we produce an inverted data volume in the form of amplitude versus angle (AVA) sections. Besides being part of the velocity analysis, these sections can also be used in AVA analysis. Several authors have treated the GRT-based direct-inversion procedure and pointed out the benefits of the method. However, most authors transform the local geometry at the image point that is inherent in the formulation to the more conventional acquisition coordinate geometry (Beylkin & Burridge, 1984); (Miller *et al.*, 1987). In this geometry, unfortunately, these approaches suffer under the assumption of no multi-pathing; that is, these approaches exclude multiple rays between a given image point and a given source or receiver at the surface. In this paper we will be using the phase directions $(\overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})$ or alternatively the coordinates $(\overset{0}{\boldsymbol{\alpha}}, \theta, \psi)$, and parameterize the problem in a single-valued fashion at the image point (de Hoop & Brandsberg-Dahl, 1999).

The developments are carried out for general heterogeneous, anisotropic elastic media. Within such a medium we define a setting wherein a remote scattering domain is illuminated with elastic waves generated by a point source. The scattering is linearized using the familiar perturbation representation, dividing the field into a known smooth background and an unknown singular part. With the scattered field recorded by point receivers on the surface, we invert for the unknown reflectivity using the generalized Radon transform for fixed scattering-angle and azimuth. Once we have established the inversion and the modeling operators, we introduce the sensitivity transform. This is an operator made from a combination of the direct and the inverse scattering operators with a differential operator in the scattering-angle/azimuth domain. The final operator can be used to analyze the coherency between different seismic experiments in the form of the difference between neighboring traces in the common scattering-angle/azimuth sections.

After establishing the measure of coherency we proceed by defining a convex objective function $J_0[\boldsymbol{c}]$ that we seek to minimize with respect to medium parameters. This differential semblance function has previously been treated in the framework of 1D and 2D acoustic media and common-midpoint coordinates in several papers (Araya *et al.*, 1996); (Symes, 1996); (Symes, 1999).

Our formulation is general in the sense of the choice of medium and the particular choice of coordinates. Further, the approach is horizon-based, in which we allow smooth interfaces with general heterogeneous media in between. Based on the experience with the 1D and 2D version of this approach (Chauris & Noble, 1998), that this method should be a good candidate for performing velocity analysis in anisotropic elastic media for both single-mode and converted-mode seismic data.

Notation and basic equations

The domain in which the scattering takes place is $\mathcal{D} \subset \mathbb{R}^3$, and the acquisition surface, $\partial S \times \partial R$ is homeomorphic* to a part of $S^2 \times S^2$, where S^2 is the unit sphere. The acquisition coordinates $(\boldsymbol{s}, \boldsymbol{r})$ lie on surfaces ∂S and ∂R, respectively, which do not intersect the scattering domain. Measurements are taken in a time interval $[0, \mathcal{T})$. We denote Cartesian coordinates in the configuration as

$$\begin{aligned} \boldsymbol{x} &= (x_1, x_2, x_3) = \text{Cartesian position vector}, \\ \boldsymbol{s} &= (s_1, s_2, s_3) = \text{source point}, \\ \boldsymbol{r} &= (r_1, r_2, r_3) = \text{receiver point}. \end{aligned} \tag{1}$$

* homeomorphic: has the same shape as

Further, the medium is described by density ρ and the elastic stiffness tensor c_{ijkl}

$$\begin{aligned} \rho^{(0)}(\boldsymbol{x}) &= \rho + \rho^{(1)} \\ c^{(0)}_{ijkl}(\boldsymbol{x}) &= c_{ijkl} + c^{(1)}_{ijkl} \,. \end{aligned} \tag{2}$$

The wave field is described by the displacement vector

$$\boldsymbol{u}(\boldsymbol{x},t) = (u_1(\boldsymbol{x},t), u_2(\boldsymbol{x},t) u_3(\boldsymbol{x},t)) \,, \tag{3}$$

and the source is described by the body-force density distribution

$$\boldsymbol{f}(\boldsymbol{x},t) = (f_1(\boldsymbol{x},t), f_2(\boldsymbol{x},t) f_3(\boldsymbol{x},t)) \,. \tag{4}$$

The displacement field propagating in the total medium is split in two parts

$$\boldsymbol{u}^{(0)}(\boldsymbol{x}) = \boldsymbol{u} + \boldsymbol{u}^{(1)} \,, \tag{5}$$

where, within the Born approximation, the scattered wave field $\boldsymbol{u}^{(1)}$ satisfies the elastodynamic wave equation

$$\rho \partial_t^2 u_i^{(1)} - \partial_j \left(c_{ijkl} \partial_l u_k^{(1)} \right) = f_i^{(1)} \,, \tag{6}$$

with contrast source term

$$f_i^{(1)} = -\rho^{(1)} \partial_t^2 u_i + \partial_j (c^{(1)}_{ijkl} \partial_l u_k) \,.$$

Here and in the rest of the paper we use the summation notation. The causal Green's tensor

$$\boldsymbol{G}(\boldsymbol{x},\boldsymbol{x}',t) = (G_{ip}(\boldsymbol{x},\boldsymbol{x}',t)) \,, \tag{7}$$

satisfies

$$\begin{aligned} \rho \, \partial_t^2 G_{ip} - \partial_j (c_{ijkl} \partial_l G_{kp}) = \delta_{ip} \delta(\boldsymbol{x} - \boldsymbol{x}') \delta(t) \,, \\ G_{ip} = 0 \quad \text{for} \quad t < 0 \,. \end{aligned} \tag{8}$$

Geometrical ray theory

We now define the ray geometry and review the formulation of anisotropic ray theory for the evaluation of the Green's tensor (Burridge *et al.*, 1998). Since the medium is elastic a different wave mode can be associated with each of the two rays connecting the source with the scattering point back to the receiver. This is usually indicated by a superscript running over a set $\{1,2,3\}$ corresponding to $qP, qS1$ and $qS2$ wave propagation. To simplify notation we will omit this index in the further derivations.

A general ray is a path connecting two points $\boldsymbol{x}$ and $\boldsymbol{x}'$ in space. The corresponding Green's tensor is

$$G_{ip}(\boldsymbol{x},\boldsymbol{x}',t) = A(\boldsymbol{x},\boldsymbol{x}') \xi_i(\boldsymbol{x}) \xi_p(\boldsymbol{x}') \delta(t - \tau(\boldsymbol{x},\boldsymbol{x}')) \,, \tag{9}$$

where

$$\tau = \tau(\boldsymbol{x},\boldsymbol{x}') \,. \tag{10}$$

denotes the traveltime between these two points along the ray. The traveltime τ and the associated polarization vector $\boldsymbol{\xi}$ satisfies

$$(\rho \, \delta_{ik} - c_{ijkl} (\partial_l \tau)(\partial_j \tau)) \, \xi_k = 0 \quad (\text{at all } \boldsymbol{x}) \,, \tag{11}$$

which again implies the eikonal equation

$$\det(\rho \, \delta_{ik} - c_{ijkl} (\partial_l \tau) \, (\partial_j \tau)) = 0 \quad (\text{at all } \boldsymbol{x}) \,. \tag{12}$$

The polarization vectors are normalized so that $\xi_i \xi_i = 1$. We define the slowness vector at $\boldsymbol{x}$ associated with a ray emanating from a point source or a point receiver at $\boldsymbol{x}'$ to a scattering point $\boldsymbol{x} \in \mathcal{D}$, as

$$\boldsymbol{\gamma}_{\boldsymbol{x}} = \nabla_{\boldsymbol{x}} \tau(\boldsymbol{x},\boldsymbol{x}') \,. \tag{13}$$

The unit phase direction associated with the slowness vector at $\boldsymbol{x}$ along the ray is given by

$$\boldsymbol{\alpha}_{\boldsymbol{x}} = \frac{\boldsymbol{\gamma}_{\boldsymbol{x}}}{|\boldsymbol{\gamma}_{\boldsymbol{x}}|} \,, \tag{14}$$

Equation (12) constrains $\boldsymbol{\gamma}$ to lie on the sextic surface $\mathcal{A}(\boldsymbol{x})$ given by

$$\det(\rho \, \delta_{ik} - c_{ijkl} \gamma_\ell \gamma_j) = 0 \,. \tag{15}$$

The surface $\mathcal{A}(\boldsymbol{x})$ consists of three sheets corresponding to the three wave modes, each of which is a closed surface surrounding the origin. The corresponding transport equation,

$$\partial_j (c_{ijk\ell} \xi_i \xi_k \, A^2 \, \partial_\ell \tau) = 0 \,, \tag{16}$$

gives the scalar amplitudes A. Each propagation mode corresponds to a particular sheet of the slowness surface $\mathcal{A}$ with the corresponding slowness vector $\boldsymbol{\gamma}$. The group velocities $\boldsymbol{v}$ are defined to be normal to $\mathcal{A}(\boldsymbol{x})$ at $\boldsymbol{\gamma}_{\boldsymbol{x}}$ and satisfy

$$\boldsymbol{v} \cdot \boldsymbol{\gamma}_{\boldsymbol{x}} = 1 \,. \tag{17}$$

The normal or phase speeds are given by

$$V = \frac{1}{|\boldsymbol{\gamma}_{\boldsymbol{x}}|} \,, \tag{18}$$

which, with the definitions above, can be written as

$$V = |\boldsymbol{v}| \cos\chi \,, \tag{19}$$

where χ is the angle between $\boldsymbol{v}$ and $\boldsymbol{\gamma}$. According to equation (13), we can now write the slowness vectors for the source ray and the receiver ray at $\boldsymbol{x}$ as

$$\begin{aligned} \overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}} &= \nabla_{\boldsymbol{x}} \tau(\boldsymbol{x},\boldsymbol{s}) \,, \\ \overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}} &= \nabla_{\boldsymbol{x}} \tau(\boldsymbol{x},\boldsymbol{r}) \,, \end{aligned} \tag{20}$$

respectively, and the associated phase directions are

$$\begin{aligned} \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}} &= \frac{\overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}}}{|\overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}}|} = \overset{s}{V} \overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}} \,, \\ \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}} &= \frac{\overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}}}{|\overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}}|} = \overset{r}{V} \overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}} \,. \end{aligned} \tag{21}$$

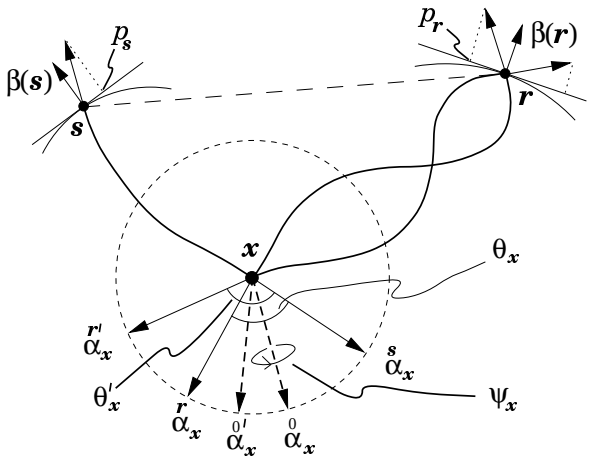

Figure 1. Ray geometry: Indicated are the phase direction at the image-point. The multi-pathing between scatterer and receiver is indicated with primed quantities. Using only offset this ambiguity would not have been resolved.

The source and receiver Green's functions are given as special cases of equation (9) by plugging in the corresponding coordinates

$$\overset{s}{\boldsymbol{G}}(\boldsymbol{x},t) \equiv \boldsymbol{G}(\boldsymbol{x},\boldsymbol{s},t)\,, \quad \overset{r}{\boldsymbol{G}}(\boldsymbol{x},t) \equiv \boldsymbol{G}(\boldsymbol{r},\boldsymbol{x},t)\,. \tag{22}$$

In *acquisition* coordinates, we define the total or two-way travel time T for the two rays connecting a source $\boldsymbol{s}$ with a receiver $\boldsymbol{r}$ via a scatterer $\boldsymbol{x}$, as

$$T(\boldsymbol{r},\boldsymbol{x},\boldsymbol{s}) = \tau(\boldsymbol{s},\boldsymbol{x}) + \tau(\boldsymbol{x},\boldsymbol{r})\,. \tag{23}$$

The mapping

$$\begin{array}{lllll} \boldsymbol{s} : (\overset{s}{S^2_{\boldsymbol{x}}}) & \rightarrow & \partial S, & \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}} & \rightarrow \quad \boldsymbol{s}(\boldsymbol{x},\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\,, \\ \boldsymbol{r} : (\overset{r}{S^2_{\boldsymbol{x}}}) & \rightarrow & \partial R, & \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}} & \rightarrow \quad \boldsymbol{r}(\boldsymbol{x},\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\,, \end{array} \tag{24}$$

transforms the the image point, $\boldsymbol{x}$, related variables $(\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}})$ into acquisition variables $(\boldsymbol{s},\boldsymbol{r})$. Here the subset $(\overset{s}{S^2_{\boldsymbol{x}}})$ of the unit sphere S^2 is the area element that contains $\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}$, defining a ray that connects $\boldsymbol{x}$ with an actual source $\boldsymbol{s}$; that is, $\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}} \in (\overset{s}{S^2_{\boldsymbol{x}}})$. For the receiver we must have $\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}} \in (\overset{r}{S^2_{\boldsymbol{x}}})$. Then we write the total travel time T as a function of the two phase directions $(\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}})$

$$T(\boldsymbol{x}) = T(\boldsymbol{x},\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}},\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) = \tau(\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}},\boldsymbol{x}) + \tau(\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}},\boldsymbol{x})\,. \tag{25}$$

We also define the gradient of the total travel time

$$\overset{0}{\boldsymbol{\gamma}}_{\boldsymbol{x}} = \overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}} + \overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}}\,, \tag{26}$$

which, in acquisition coordinates, corresponds to

$$\overset{0}{\boldsymbol{\gamma}}_{\boldsymbol{x}} = \nabla_{\boldsymbol{x}} T(\boldsymbol{r},\boldsymbol{x},\boldsymbol{s})\,. \tag{27}$$

Dip, scattering-angle and azimuth

In the framework of generalized Radon transforms (GRT) it is natural to work in local coordinates at the image-point. We will use an indexing in migration dip $\overset{0}{\boldsymbol{\alpha}}$, scattering-angle θ and azimuth ψ (third Euler angle around $\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}$). These variables will depend on the modes of the source and receiver rays, and they are uniquely defined through the relations described in the previous section. For an image-point $\boldsymbol{x} \in \mathcal{D}$ these variables are defined as

$$\begin{aligned} \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}} &= \frac{\overset{0}{\boldsymbol{\gamma}}_{\boldsymbol{x}}}{|\overset{0}{\boldsymbol{\gamma}}_{\boldsymbol{x}}|} \in S^2, \\ \cos\theta_{\boldsymbol{x}} &= \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}} \cdot \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}} \in E_\theta \subset [0,\pi), \\ \boldsymbol{\psi}_{\boldsymbol{x}} &= \frac{(\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}} \cdot \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\,\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}} - (\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}} \cdot \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\,\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}}{\sin\theta_{\boldsymbol{x}}} \in E_\psi \subset S^1\,. \end{aligned} \tag{28}$$

This gives a novel coordinate frame:

$$(\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \rightarrow (\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \theta_{\boldsymbol{x}}, \psi_{\boldsymbol{x}}) = (\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{h}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\,. \tag{29}$$

In terms of acquisition coordinates the migration dip vector replaces the midpoint and the combination of the two angles ('directivity' $\boldsymbol{\alpha}^h_{\boldsymbol{x}}$) replaces the offset vector. For the image-point dependent vectors, we use a subscript to indicate the point to which the vector is related. To perform this change in variables we need to introduce the following Jacobian (Burridge *et al.*, 1998)

$$\frac{\partial(\overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})}{\partial(\overset{0}{\boldsymbol{\alpha}}, \theta, \psi)} = \frac{\sin\theta}{1 + \left(|\overset{s}{\boldsymbol{\gamma}}||\overset{r}{\boldsymbol{\gamma}}|/|\overset{0}{\boldsymbol{\gamma}}|^2\right)\left(\tan\overset{r}{\chi} - \tan\overset{s}{\chi}\right)\sin\theta}\,, \tag{30}$$

where

$$\cos\overset{s}{\chi} = \overset{s}{\boldsymbol{n}}_{\|} \cdot \overset{s}{\boldsymbol{\alpha}} \quad \text{and} \quad \cos\overset{r}{\chi} = \overset{r}{\boldsymbol{n}}_{\|} \cdot \overset{r}{\boldsymbol{\alpha}}\,. \tag{31}$$

Here, $\overset{r}{\boldsymbol{n}}_{\|}$ and $\overset{s}{\boldsymbol{n}}_{\|}$ denote the normals to the slowness surface at the scattering point projected in the azimuth plane ψ. The source and receiver coordinates can then be written in the form

$$\boldsymbol{s} = \boldsymbol{s}(\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \theta_{\boldsymbol{x}}, \psi_{\boldsymbol{x}})\,, \qquad \boldsymbol{r} = \boldsymbol{r}(\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \theta_{\boldsymbol{x}}, \psi_{\boldsymbol{x}})\,. \tag{32}$$

It is also possible to go from the acquisition surface down to the image-point. To make this 'reverse' mapping, from a source or a receiver at the surface to an image-point along the rays unique, we introduce the projection of the slowness vectors at the source onto ∂S, and the projection of the slowness vectors at the receiver onto ∂R. These projections are given by

$$\begin{aligned} \boldsymbol{p}_{\boldsymbol{s}}(\boldsymbol{x},\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) &= \overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}} - (\beta(\boldsymbol{s})\cdot\overset{s}{\boldsymbol{\gamma}}_{\boldsymbol{x}})\beta(\boldsymbol{s})\,, \\ \boldsymbol{p}_{\boldsymbol{r}}(\boldsymbol{x},\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) &= \overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}} - (\beta(\boldsymbol{r})\cdot\overset{r}{\boldsymbol{\gamma}}_{\boldsymbol{x}})\beta(\boldsymbol{r})\,, \end{aligned} \tag{33}$$

in which

$$\begin{aligned}\boldsymbol{\beta}(\boldsymbol{s}) &= \text{unit normal to } \partial S \text{ at source,}\\ \boldsymbol{\beta}(\boldsymbol{r}) &= \text{unit normal to } \partial R \text{ at receiver.}\end{aligned} \tag{34}$$

$\boldsymbol{p_s}$ and $\boldsymbol{p_r}$ can be directly estimated from the data as a function of $\boldsymbol{s}$ and $\boldsymbol{r}$ (by slant stacking). Using these, we have the following micro-local, one-to-one mapping from a source $\boldsymbol{s}$ and a receiver $\boldsymbol{r}$ down to the image-point $\boldsymbol{x}$

$$(\boldsymbol{s}, \boldsymbol{r}, \tau, \omega \boldsymbol{p_s}, \omega \boldsymbol{p_r}, -\omega) \rightarrow (\boldsymbol{x}, \omega \overset{0}{\boldsymbol{\gamma}}_{\boldsymbol{x}}) . \tag{35}$$

For a more extended analysis of these aspects and a rigorous treatment of the various conditions that are posed upon the asymptotic framework and the mappings we use, refer to de Hoop and Brandsberg-Dahl (1999).

We summarize the different choices of variables by setting up an overview of the change from the surface/acquisition controlled variables to the subsurface/image-point controlled variables. For a given scattering point $\boldsymbol{x} \in \mathcal{D}$, with two rays connecting it with, respectively, a source $\boldsymbol{s} \in \partial S$ and a receiver $\boldsymbol{r} \in \partial R$, we have

$$\begin{aligned}&\overset{0}{\boldsymbol{x}}, \overset{h}{\boldsymbol{x}} && \text{midpoint and offset,}\\ &\quad\uparrow\downarrow\\ &\boldsymbol{r}, \boldsymbol{s} && \text{source and receiver,}\\ &\quad\uparrow\downarrow\\ &\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}} && \text{phase directions,}\\ &\quad\uparrow\downarrow\\ &\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{h}{\boldsymbol{\alpha}}_{\boldsymbol{x}} && \text{migration dip,}\\ &&& \text{scattering-angle and azimuth.}\end{aligned} \tag{36}$$

The corresponding angles and vectors are shown in Figure 1.

Medium description

We consider an anisotropic elastic medium that we represent by a smooth background medium and a perturbation along interfaces, that we describe by level surfaces, the same representation as used in de Hoop and Bleistein (1997). The relative medium perturbation is given by

$$\boldsymbol{c}^{(1)} = \left\{ \frac{\rho^{(1)}}{\rho}, \frac{c^{(1)}_{ijk\ell}}{\rho \overset{r}{V}_\circ \overset{s}{V}_\circ} \right\} ,$$

where $V_\circ$ denotes the (local) normal wave speed of the given mode in the background medium averaged over all phase directions. This normalization is performed in order to get a convenient dimensionless quantity. To describe the relative medium perturbation along the level surfaces we introduce

$$\boldsymbol{c}^{(1)}(\boldsymbol{x}) = \boldsymbol{c}^{(1)}(\boldsymbol{x}, \phi(\boldsymbol{x})) . \tag{37}$$

This is a micro-local formulation wherein we assume that the representation extends to a small region around any point under consideration. The level surfaces ϕ are smooth functions of $\boldsymbol{x}$ and describe interfaces in the subsurface. The gradient of ϕ vary rapidly normal to the level surfaces and smoothly along them, giving a formalism that can be used to represent a typical geological setting. We define the gradient by the chain rule for derivatives

$$\nabla_{\boldsymbol{x}} \boldsymbol{c}^{(1)} = (\boldsymbol{c}^{(1)})' (\nabla_{\boldsymbol{x}} \phi) + \text{smoother terms} \tag{38}$$

$$\text{where } (\boldsymbol{c}^{(1)})' \equiv \partial_\phi \boldsymbol{c}^{(1)}$$

This representation has a Dirac-distribution behavior across any geological interface and will typically represent the change in medium parameters across the level surface ϕ.

The GRT operators

The direct scattering operator

The direct scattering operators give the scattered field $\boldsymbol{u}^{(1)}(s, r, t)$ that arises from the perturbation $\boldsymbol{c}^{(1)}$ in the background medium. The operator is an integral over a hyper-surface that is the t isochrone for a fixed source-receiver pair $(\boldsymbol{s}, \boldsymbol{r})$. The conventional volume integral representation of the direct scattering operator is given in a modified form by de Hoop and Bleistein (1997). The p-component displacement at $\boldsymbol{r}$ due to a q-component point force at $\boldsymbol{s}$ is

$$\begin{aligned}u^{(1)}_{pq}(\boldsymbol{r}, \boldsymbol{s}, t) \simeq -\int_{\mathcal{D}} & \overset{s}{\xi}_q(\boldsymbol{s}) A(\boldsymbol{r}, \boldsymbol{x}, \boldsymbol{s})\\ & \cdot \overset{r}{\xi}_p(\boldsymbol{r}) \left(\boldsymbol{w}(\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\right)^T (\boldsymbol{c}^{(1)})' (\phi(\boldsymbol{x}))\\ & \cdot \left. \frac{\bar{\boldsymbol{\nu}} \cdot \nabla_{\boldsymbol{x}} \phi}{\bar{\boldsymbol{\nu}} \cdot \overset{0}{\boldsymbol{\gamma}}} \right|_{\boldsymbol{x}} \delta'[t - T(\boldsymbol{s}, \boldsymbol{x}, \boldsymbol{r})] \, \mathrm{d}\boldsymbol{x} .\end{aligned} \tag{39}$$

Here $\bar{\boldsymbol{\nu}}$ is an arbitrary unit vector $\bar{\boldsymbol{\nu}} \in S^2$, that varies slowly in space, and $\bar{\boldsymbol{\nu}} \cdot \nabla_{\boldsymbol{x}} T(\boldsymbol{x}) \neq 0$. In their paper, de Hoop and Bleistein (1997) showed that $\bar{\boldsymbol{\nu}}$ at stationarity is equal to the geological dip,

$$\boldsymbol{\alpha}_\phi = \frac{\nabla_{\boldsymbol{x}} \phi}{|\nabla_{\boldsymbol{x}} \phi|} . \tag{40}$$

Hence, at stationarity the migration dip is parallel with the geological dip. The vector $\bar{\boldsymbol{\nu}}$ can be chosen equal to the geological dip $\boldsymbol{\alpha}_\phi$, or it can be chosen equal to the migration dip $\overset{0}{\boldsymbol{\alpha}}$, which we alway control through the ray geometry. The amplitude term is

$$A(\boldsymbol{r}, \boldsymbol{x}, \boldsymbol{s}) \equiv \rho(\boldsymbol{x}) A(\boldsymbol{s}, \boldsymbol{x}) A(\boldsymbol{x}, \boldsymbol{r}) , \tag{41}$$

and $\boldsymbol{w}$ is the radiation pattern of the point contrast-source at the scattering point. The radiation patterns are symmetric dyadic products of the polarization and slowness vectors of the rays and are defined as

$$\begin{aligned} \boldsymbol{w} &= \left\{ \overset{s}{\xi}_i \overset{r}{\xi}_i, \frac{1}{2}\left[\overset{r}{a}_{ij} \overset{s}{a}_{kl} + \overset{r}{a}_{kl} \overset{s}{a}_{ij} \right] \right\} , \\ \overset{r}{a}_{ij} &= \frac{1}{2} \overset{r}{V} \left(\overset{r}{\xi}_i \overset{r}{\gamma}_j + \overset{r}{\xi}_j \overset{r}{\gamma}_i \right) , \\ \overset{s}{a}_{kl} &= \frac{1}{2} \overset{s}{V} \left(\overset{s}{\xi}_k \overset{s}{\gamma}_l + \overset{s}{\xi}_l \overset{s}{\gamma}_k \right) . \end{aligned} \tag{42}$$

In the upcoming construction of the sensitivity transform we will need a direct scattering operator to be applied to data in the common scattering-angle/azimuth domain. This will be an integral operator taken over the level surfaces in the model, something that requires that we know the dip field in the current model. From the dip field it is straight forward to find and control the scattering angles of the rays in the specular situation (see Figure 2). This procedure can be used on the true model or on any estimated model we produce. With this formulation it is possible to re-model data from a given scattering-angle section, and also utilize angle-dependent reflection coefficients along the interfaces.

The derivation starts with the volume scattering formulation given in equation (39). This integral is then recast as an integral over the level surfaces of the singular support of the medium (de Hoop & Bleistein, 1997). Let's introduce local curvi-linear coordinates on the level surface ϕ as follows: let $\boldsymbol{\sigma} = (\sigma_i, \sigma_3), i = 1, 2$ be such that σ_i are coordinates in the surface and σ_3 is the local coordinate normal to the surface, in the $\boldsymbol{\alpha}_\phi$ direction. If σ_3 represents the actual value of the level surface, we set $\sigma_3 = L$. Let $d\Sigma$ denote a surface element on the level surface, then we can recast the volume integral taken over isochrones to a surface integral taken over level surface coordinates

$$\begin{aligned} d\boldsymbol{x} &= \left. \frac{1}{|\nabla_{\boldsymbol{x}} \phi|} \right|_{\boldsymbol{x}(\boldsymbol{\sigma})} \mathrm{d}L \mathrm{d}\Sigma \, , \\ &\text{with } \mathrm{d}\Sigma = |\partial_{\sigma_1} \boldsymbol{x} \times \partial_{\sigma_2} \boldsymbol{x}| \, \mathrm{d}\sigma_1 \mathrm{d}\sigma_2 \, . \end{aligned} \tag{43}$$

Let the perturbations be represented by

$$(\boldsymbol{c}^{(1)})'(\phi(\boldsymbol{x})) = \int_{\mathbb{R}} (\boldsymbol{c}^{(1)})'(L) \delta(\phi(\boldsymbol{x}) - L) \mathrm{d}L \, . \tag{44}$$

Here the delta function will pick out the singular support of the medium and assign the appropriate medium perturbation as the integral is taken over the family of all level surfaces. To complete the reformulation of the general expression in equation (39), we make use of equation (43) and (44) to get the following identity

$$\begin{aligned} &\int_{\mathcal{D}} \ldots \delta(\phi(\boldsymbol{x}) - L) \mathrm{d}\boldsymbol{x} = \\ &\int_{\phi = L} \ldots \left. \frac{1}{|\nabla_{\boldsymbol{x}} \phi|} \right|_{\boldsymbol{x}(\boldsymbol{\sigma})} \mathrm{d}\Sigma \, . \end{aligned} \tag{45}$$

The direct scattering operator then becomes

$$\begin{aligned} &\partial_t u^{(1)}_{pq}(\boldsymbol{s}, \boldsymbol{r}, t) \simeq \\ &\quad - \int_{\mathbb{R}} \mathrm{d}L \int_{\phi = L} \overset{s}{\xi}_q (\boldsymbol{s}) A(\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \\ &\quad \cdot \overset{r}{\xi}_p (\boldsymbol{r}) \left(\boldsymbol{w}(\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \right)^T (\boldsymbol{c}^{(1)})'(L) \\ &\quad \cdot \left. \frac{(\bar{\boldsymbol{\nu}} \cdot \boldsymbol{\alpha}_\phi)}{(\bar{\boldsymbol{\nu}} \cdot \overset{0}{\boldsymbol{\gamma}})} \right|_{\boldsymbol{x}} \delta''(t - T(\boldsymbol{x})) \, \mathrm{d}\Sigma(\boldsymbol{x}) \, , \end{aligned} \tag{46}$$

where we make use of the ray-induced mapping from phase directions at the scattering point $\boldsymbol{x}$ to source and receiver at the surface $(\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \rightarrow (\boldsymbol{s}, \boldsymbol{r})$. The scattering-angle and azimuth are directly related to $(\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}})$, as given by equation (28). To retrieve the reflection coefficient we extract the velocities from the amplitude and create the quantity

$$A_u(\boldsymbol{x}) \equiv A(\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \, \overset{s}{V}(\boldsymbol{x}) (\overset{r}{V}(\boldsymbol{x}))^3)^{1/2} \, . \tag{47}$$

Then equation (46) can be written as

$$\begin{aligned} &\partial_t u^{(1)}_{pq}(\boldsymbol{s}, \boldsymbol{r}, t) \simeq (\boldsymbol{L}[\boldsymbol{c}^{(0)}] R_L)(\boldsymbol{r}, \boldsymbol{s}, t) = \\ &\quad - \int_{\mathbb{R}} \left[\int_{\phi = L} \overset{s}{\xi}_q (\boldsymbol{s}) A_u(\boldsymbol{x}) \, \overset{r}{\xi}_p (\boldsymbol{r}) \right. \\ &\quad \cdot R_L(\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \, (\bar{\boldsymbol{\nu}} \cdot \boldsymbol{\alpha}_\phi)(\bar{\boldsymbol{\nu}} \cdot \overset{0}{\boldsymbol{\gamma}}) \Big|_{\boldsymbol{x}} \\ &\quad \left. \cdot \delta''(t - T(\boldsymbol{x})) \, \frac{1}{|\nabla_{\boldsymbol{x}} \phi|} \mathrm{d}\Sigma(\boldsymbol{x}) \right] \mathrm{d}L \, , \end{aligned} \tag{48}$$

where we have defined

$$\begin{aligned} &R_L[\boldsymbol{c}^{(1)}](\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \equiv \\ &\quad \left. \frac{(\boldsymbol{w}(\boldsymbol{x}, \overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}))^T (\boldsymbol{c}^{(1)})' |\nabla_{\boldsymbol{x}} \phi|(\boldsymbol{x})}{(\overset{s}{V}(\boldsymbol{x}) (\overset{r}{V}(\boldsymbol{x}))^3)^{1/2} |\boldsymbol{\gamma}^0_{\boldsymbol{x}}|^2 \, (\bar{\boldsymbol{\nu}} \cdot \overset{0}{\boldsymbol{\alpha}})^2} \right|_{\boldsymbol{x}} . \end{aligned} \tag{49}$$

This is the scattering-matrix representation of the reflection coefficient at $\boldsymbol{x}$ such that $\phi(\boldsymbol{x}) = L$. Due to the parameterization in curvi-linear coordinates, R_L actually depends on $\boldsymbol{\sigma}_{1,2}$. The integral over L serves the purpose of picking up the singular support of the medium.

In order to get the Kirchhoff-Born approximation, we use the specular ray geometry and Snell's law and substitute into equation (51). The specular point is the point where the migration dip is parallel to the geological dip $\overset{0}{\boldsymbol{\alpha}} \, || \boldsymbol{\alpha}_\phi$. At the specular point-set $(\boldsymbol{s}(\overset{s}{\boldsymbol{\alpha}}), \boldsymbol{x}, \boldsymbol{r}(\overset{r}{\boldsymbol{\alpha}}))$, for

a given geological dip, $(\overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})$ satisfies Snell's law

$$\frac{\overset{s}{\boldsymbol{\alpha}}}{\overset{s}{V}} \cdot (I - \boldsymbol{\alpha}_\phi \cdot \boldsymbol{\alpha}_\phi) = -\frac{\overset{r}{\boldsymbol{\alpha}}}{\overset{r}{V}} \cdot (I - \boldsymbol{\alpha}_\phi \cdot \boldsymbol{\alpha}_\phi)\,, \tag{50}$$

The solution $\overset{\sim}{\overset{r}{\boldsymbol{\alpha}}}$ of this equation is substituted into the expression for the scattering matrix (49), and we get an expression for the *linearized reflection/transmission coefficients* r, a coefficient that is understood in a distributional sense. To simplify notation, considering a true curved reflector at $L = L_0$, we introduce

$$\mathsf{R}_L(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}\,(\cdot)) = \\ \mathsf{r}_{L_0}(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}\,(\cdot))\,\delta(L - L_0) = R_L(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}\,(\cdot), \overset{\sim}{\overset{r}{\boldsymbol{\alpha}}}(\cdot))\,. \tag{51}$$

Hence, for specular reflections we have $\overset{0}{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_\phi$, and the specular ray angles are given by

$$\cos\overset{s}{\theta} = \overset{s}{\boldsymbol{\alpha}} \cdot \boldsymbol{\alpha}_\phi \quad \cos\overset{r}{\theta} = \overset{\sim}{\overset{r}{\boldsymbol{\alpha}}} \cdot \boldsymbol{\alpha}_\phi \quad \theta = \overset{s}{\theta} + \overset{r}{\theta}\,. \tag{52}$$

The Kirchhoff-Born approximation of the scattered field is then given by equation (48) after setting $\overset{0}{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_\phi$ and substituting in equation (51).

To summarize this reformulation of the direct scattering operator, we have derived an operator that performs the mapping

$$\boldsymbol{L} : \mathsf{R}_L(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}\,(\cdot)) \to \boldsymbol{u}^{(1)} \tag{53}$$

in a setting of specular ray geometry. The modeling procedure in the common scattering-angle/azimuth domain can then be pictured as follows: take a fixed scattering-angle and a fixed azimuth, then for every point along an interface, the local geological dip specifies the bisection of the scattering-angle. Together with the azimuth this specifies the two phase directions needed to shoot rays to a source and a receiver at the surface. A complete common scattering-angle/azimuth dataset is created by moving the "V"of fixed opening angle across the various interfaces in the subsurface and tracing rays to the surface within the given azimuth (see Figure 2). The reflection coefficient is determined from the opening angle at the interface and the medium perturbation in the location under consideration. In the stationary-phase approximation, the direct scattering operator becomes nothing other than a generalization of the convolutional-model approximation to three dimensions. This implies that the scattered field can be viewed as a scaled and polarized version of the reflectivity. The direct scattering operator simply scales and vectorizes the reflection coefficient, and assigns a traveltime.

The inverse GRT operator

The inversion is carried out for one image-point $\boldsymbol{y}$ at a time as an integral over phase directions. Shooting the rays from $\boldsymbol{y}$, we control the initial phase directions and pick out a source and a receiver at the surface from the locations where the rays intersect the acquisition surfaces ∂S and ∂R, respectively. As we outlined in a previous section, we are then free to change variables to migration dip, scattering-angle and azimuth; see equation (28). This produces a three-fold integral over $(\theta, \psi, \overset{0}{\boldsymbol{\alpha}})$. This integral can be evaluated in separate steps. The inverse GRT operator $\boldsymbol{U}$ for a fixed pair $(\theta_{\boldsymbol{y}}, \psi_{\boldsymbol{y}})$ is given by

$$\int_{\mathbb{R}} \mathsf{r}_L(\boldsymbol{y}, \theta, \psi)\mathrm{d}L \simeq \boldsymbol{U}[\boldsymbol{c}]u^{(1)}(\boldsymbol{y}) = \\ \frac{1}{8\pi^2}\int_{E_{\overset{0}{\boldsymbol{\alpha}}}(\theta,\psi)} \frac{|\,\overset{0}{\boldsymbol{\gamma}}_{\boldsymbol{y}}\,(\boldsymbol{y}, \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}}, \theta_{\boldsymbol{y}}, \psi_{\boldsymbol{y}})|^4}{A(\boldsymbol{y}, \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}}, \theta_{\boldsymbol{y}}, \psi_{\boldsymbol{y}})} \\ \cdot[\Lambda(\boldsymbol{y}, \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}})]^{-1}\, \boldsymbol{w}(\boldsymbol{y}, \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}}, \theta_{\boldsymbol{y}}, \psi_{\boldsymbol{y}})\, \overset{r}{\xi}_p\,(\boldsymbol{r}) \\ \cdot\, \partial_t u^{(1)}_{pq}(\boldsymbol{s}, \boldsymbol{r}, T(\boldsymbol{y}, \overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}}, \theta_{\boldsymbol{y}}, \psi_{\boldsymbol{y}}))\, \overset{s}{\xi}_q\,(\boldsymbol{s}) \\ \cdot \left.\frac{\partial(\overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})}{\partial(\overset{0}{\boldsymbol{\alpha}}, \theta, \psi)}\right|_{\boldsymbol{y}} \mathrm{d}\,\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}}\,. \tag{54}$$

Where

$$\Lambda(\,\cdot\,, \overset{0}{\boldsymbol{\alpha}}) = \int_{E_\psi}\int_{E_\theta} \boldsymbol{w}(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}\,(\cdot), \overset{r}{\boldsymbol{\alpha}}\,(\cdot)) \\ \cdot\, \boldsymbol{w}(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}\,(\cdot), \overset{r}{\boldsymbol{\alpha}}\,(\cdot))^T \left.\frac{\partial(\overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})}{\partial(\overset{0}{\boldsymbol{\alpha}}, \theta, \psi)}\right|\, \mathrm{d}\theta\,\mathrm{d}\psi\,,$$

unravels the radiation pattern at the point contrast-source. The outcome of equation (54) is a reflectivity section for a fixed scattering-angle/azimuth $(\theta_{\boldsymbol{y}}, \psi_{\boldsymbol{y}})$.

In the stationary-phase approximation, the geological and migration dip coincide and we have $(\boldsymbol{\alpha}_\phi \cdot \overset{0}{\boldsymbol{\alpha}}) = 1$. The inverted data volume can then be viewed as a scaled version of the scattering matrix R_L given in equation

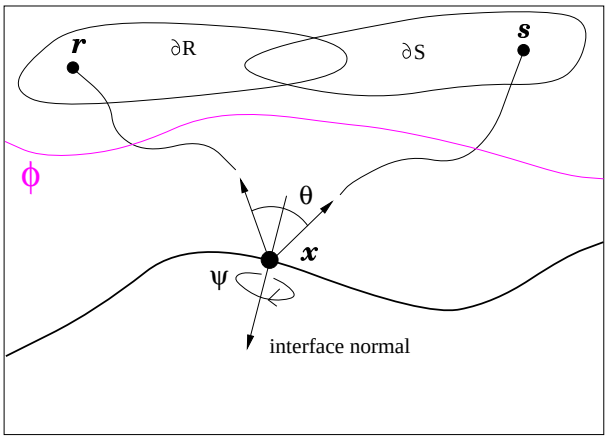

Figure 2. The Kirchhoff modeling procedure with known geological dip.

(49), which again corresponds to the linearized reflection coefficient in the specular geometry as given by equation (51). Using equation (51) for the linearized reflection coefficient it is then possible to undo this scaling and to extract the angle-independent perturbation $\boldsymbol{c}^{(1)}$. The geological dip that goes into this calculation can be estimated from any of the inverted sections as we scan over different values of the migration dip $\overset{0}{\boldsymbol{\alpha}}$.

In order to construct the angle independent quantity we need to undo the action of the radiation pattern $\boldsymbol{w}$. The radiation pattern can be calculated in the current background model based on the scattering-angle, azimuth and migration dip that we control through the raytracing. In the specular approximation, where the migration dip is parallel to the given geological dip $\overset{0}{\boldsymbol{\alpha}} \,||\boldsymbol{\alpha}_\phi$, the only degrees of freedom in the matrix $\boldsymbol{w}$ are (θ, ψ). Hence, to undo the action of $\boldsymbol{w}$ we must construct the pseudo-inverse of $\boldsymbol{w}\boldsymbol{w}^T$. This can be achieved through calculating a singular value decomposition for $\boldsymbol{w}\boldsymbol{w}^T$, say $\boldsymbol{w}\boldsymbol{w}^T = BSK^T$, where B and K are 22×22 matrices containing the eigenvectors and S is the matrix containing the eigenvalues. The pseudo-inverse is then given by

$$\lambda = KS^+B^T \,. \tag{55}$$

Multiplying R_L from the left by $\lambda\boldsymbol{w}$ leads to a quantity that will be independent of scattering-angle and azimuth (radiation pattern) if the $\boldsymbol{w}$ calculated in the current background model is equal to the true radiation pattern. An estimate of this quantity is given by

$$\begin{aligned}\widehat{\mathsf{R}}(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}}) = \lambda(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})\boldsymbol{w}(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})\, R_L(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}}) \\ \cdot \left(\overset{s}{V}(\cdot)(\overset{r}{V}(\cdot))^3\right)^{1/2} |\overset{0}{\boldsymbol{\gamma}}(\,\cdot\,, \overset{s}{\boldsymbol{\alpha}}, \overset{r}{\boldsymbol{\alpha}})|^2 (\bar{\boldsymbol{\nu}}\cdot\overset{0}{\boldsymbol{\alpha}})^2 \,,\end{aligned} \tag{56}$$

where the radiation pattern, amplitudes and traveltimes that are part of this expression are estimates computed in the current background model. Equation (56) will have important implications when we later start to compare neighboring image sections in the angle domain.

Finally, the relation between the reflectivity and the medium perturbation, that is the full inversion, is given by

$$\langle (\boldsymbol{c}^{(1)})'(\phi(\boldsymbol{y}))\,|\nabla_{\boldsymbol{x}}\phi|\boldsymbol{y}\rangle \simeq \sum_{\overset{0}{\boldsymbol{\alpha}}=\pm\boldsymbol{\alpha}_\phi} \frac{1}{2}\int_{E_\theta\times E_\psi}\int_{\mathbb{R}} \mathsf{r}_L \,\mathrm{d}L\,\mathrm{d}\theta\,\mathrm{d}\psi \,,$$

which is just a sum over scattering-angles and azimuths. The integration over $\overset{0}{\boldsymbol{\alpha}}$ in the inverse operator should produce the same image of the singular support of the perturbation $\widehat{\mathsf{R}} \sim \boldsymbol{c}^{(1)}$ for each pair of (θ, ψ). The possible redundancy between these images is what we will use to construct an update of the background medium in the next section.

Velocity analysis

Now, having both the direct and the inverse scattering operators available we can start the development of the velocity analysis. The velocity analysis is performed using experiments created with different scattering-angle/azimuth pairs. Parametrizing the problem in scattering-angle and azimuth ensures a one-to-one mapping of events between the data and the image domain and hence eliminating ambiguity due to multi-pathing in the velocity analysis. It is the degree of coherency between the different common scattering-angle/azimuth sections that will serve as the detector of the medium variations and ultimately give a criterion upon which the model can be updated. The principle for the velocity analysis is that the estimated medium perturbations are independent of scattering-angle and azimuth if and only if the velocity is correct.

With this assumption the coherency can be measured through some sort of semblance measure. Several authors have described various forms of semblance measures, either stack-power or more indirect measures such as data misfit. We will use a slightly different approach that has been suggested through several papers of The Rice Inversion Project (Araya *et al.*, 1996); (Symes, 1999), describing what they call a 'differential measure of semblance'. They have shown that the differential semblance measure has mathematical properties that make it better suited for our purpose than are other semblance measures. This measure responds smoothly to variations in the medium, and has a well-defined high-frequency limit, which allows for the use of asymptotic theory to determine the properties of the measure.

Sensitivity transform

To reveal the degree of coherency between the separate reconstructions made using equation (54), we introduce the following derivative in scattering-angle/azimuth

$$D_{(\theta,\psi)}\widehat{\mathsf{R}}_L(\boldsymbol{y},\theta,\psi) \equiv \frac{\partial^2\,\widehat{\mathsf{R}}_L(\boldsymbol{y},\theta,\psi)}{\partial\theta\,\partial\psi} \,, \qquad \boldsymbol{y}\in\mathcal{D} \,. \tag{57}$$

Here $\widehat{\mathsf{R}}$ is the quantity we constructed in the inversion and that is independent of angle unless the velocity model is incorrect. The derivative measures the similarity between inversions made with different scattering-angles and azimuths. If two "neighboring" inversions are exact alike, the result of applying equation (57) is zero. That is, an event should appear at the same spatial location and have a uniform signature in the different reconstructions, independent of the scattering-angle θ and azimuth ψ (see Figure 3). The derivative in θ, ψ then measures the extent to which this is the case. The need for calculating this derivative poses certain demands on

the data sampling. A sufficiently dense sampling in space is needed such that the derivative can be evaluated with discrete differences.

Kim and Symes (1996) show that in order to get a global, well-behaved measure on how good the current computational state relates to the seismic data, it is wise to do a comparison in the data domain. This corresponds to regenerating data from the results of the inversion after applying the differential operator (57). In computational terms this is simply applying the direct scattering operator in the common scattering-angle/azimuth domain to the output of the inversion in the current computational state $\boldsymbol{c}$. The resulting operator,the *sensitivity Transform*, is defined as

$$\begin{aligned}\boldsymbol{S} &= [I-\Delta]^{-1/2}\boldsymbol{L}\left[\frac{\partial^2 \widehat{\mathsf{R}}_L}{\partial\theta\,\partial\psi}\right] \\ &= [I-\Delta]^{-1/2}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\,, \qquad \boldsymbol{y}\in\mathcal{D}\,, \\ &\text{where}\quad \Delta \equiv \partial_\theta\partial_\psi + \boldsymbol{c}_\Delta^{-2}\partial_t^2\,. \end{aligned} \tag{58}$$

Taking the derivative enhances the high-frequency content of the data. The factor $[I-\Delta]^{-1/2}$ is applied in order to balance the derivative with respect to all the variables on which the data depend, to ensure a stable high-frequency limit. It serves as a smooth taper on the frequency content. This formulation of the sensitivity transform gives a direct measure of how good is the current background model. If all neighboring traces are equal, applying the operator should give zero as a result. To get a useful measure, however, we will construct an objective function using the L^2 norm.

To derive an explicit expression for the sensitivity transform, we need to investigate the derivative of the direct scattering operator and its inverse with respect to scattering-angle, azimuth and the medium. We will perform an analysis up to only leading-order terms, neglecting lower-order contributions as we proceed. The medium derivatives in this asymptotic setting and under the assumption that $\boldsymbol{L}[\boldsymbol{c}]\boldsymbol{U}[\boldsymbol{c}] \sim I$ can be used to make several useful approximations. For the derivative $\delta\boldsymbol{L}[\boldsymbol{c}] \simeq D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c}$ of the direct scattering operator with respect to $\boldsymbol{c}$ in the direction of $\delta\boldsymbol{c}$ the following relation holds

$$\big(D(\boldsymbol{L}[\boldsymbol{c}])\delta\boldsymbol{c}\big)\boldsymbol{U}[\boldsymbol{c}] + \boldsymbol{L}[\boldsymbol{c}]\big(D(\boldsymbol{U}[\boldsymbol{c}])\delta\boldsymbol{c}\big) = 0\,. \tag{59}$$

Then, under the assumption that $\boldsymbol{L}[\boldsymbol{c}]\boldsymbol{U}[\boldsymbol{c}]$ and $\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{L}[\boldsymbol{c}]$ are independent of the background medium $\boldsymbol{c}$, the following relation holds for the derivative of the inverse operator

$$\delta\boldsymbol{U}[\boldsymbol{c}] \simeq D(\boldsymbol{U}[\boldsymbol{c}])\delta\boldsymbol{c} \simeq -\boldsymbol{U}[\boldsymbol{c}]\big(D(\boldsymbol{L}[\boldsymbol{c}])\delta\boldsymbol{c}\big)\boldsymbol{U}[\boldsymbol{c}]\,. \tag{60}$$

Hence the derivative of the inverse operator need not be computed explicitly. The derivative with respect to the medium of the forward operator is

$$\begin{aligned} &D(\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c})\mathsf{R}_L)(\boldsymbol{s},\boldsymbol{r},t) = \\ &\quad -\int_{\mathbb{R}}\left[\int_{\phi=L} \overset{s}{\xi}_q(\boldsymbol{s})A_u(\boldsymbol{x})\overset{r}{\xi}_p(\boldsymbol{r})\right. \\ &\quad \cdot R_L(\boldsymbol{x},\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}},\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\ (\bar{\boldsymbol{\nu}}\cdot\boldsymbol{\alpha}_\phi)(\bar{\boldsymbol{\nu}}\cdot\overset{0}{\boldsymbol{\gamma}})\Big|_{\boldsymbol{x}} \\ &\quad \cdot\delta''(t-T(\boldsymbol{x}))(\delta\tau(\boldsymbol{s},\boldsymbol{x})+\delta\tau(\boldsymbol{x},\boldsymbol{r})) \\ &\quad \left.\cdot\frac{1}{|\nabla_{\boldsymbol{x}}\phi|}\mathrm{d}\Sigma(\boldsymbol{x})\right]\mathrm{d}L\,, \end{aligned} \tag{61}$$

where $\delta\tau(\,\cdot\,,\boldsymbol{x})$ are the perturbations in traveltime along the source- and receiver-ray respectively.

The derivative with respect to scattering-angle and azimuth of the direct and the inverse scattering operators follow

$$\begin{aligned} &D_{(\theta,\psi)}\boldsymbol{L}[\boldsymbol{c}] \sim \big(D_{(\theta,\psi)}\,T(\,.\,,\,.\,,\,.\,)\big)\ \partial_t + D_{(\theta,\psi)} \\ &D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}] \sim D_{(\theta,\psi)} - \big(D_{(\theta,\psi)}\,T(\,.\,,\,.\,,\,.\,)\big)\ \partial_t\,. \end{aligned} \tag{62}$$

This approximation assumes that the contribution from all amplitude terms, weight functions and radiation patterns are of lower order in frequency, and corresponds to terms neglected in the derivation of the Kirchhoff-Born approximation. For the inverse operator this gives

$$\begin{aligned} &D_{(\theta,\psi)}(\boldsymbol{U}[\boldsymbol{c}]u^{(1)})(\theta_{\boldsymbol{y}},\psi_{\boldsymbol{y}},\boldsymbol{y}) \sim \\ &\quad \int\big[\delta(t-T(\boldsymbol{y}))D_{(\theta,\psi)}\{\text{smooth terms}\} \\ &\quad -\{\text{smooth terms}\}(D_{(\theta,\psi)}T(\boldsymbol{y}))\delta'(t-T(\boldsymbol{y}))\big] \\ &\quad \cdot u^{(1)}(\theta_{\boldsymbol{y}},\psi_{\boldsymbol{y}},\boldsymbol{y})\mathrm{d}\overset{0}{\boldsymbol{\alpha}}\ . \end{aligned} \tag{63}$$

Partial integration of the second term, only keeping the most singular part, that is leaving out all terms that arise from the derivative of the smooth part of the operator, gives the expression

$$\begin{aligned} &D_{(\theta,\psi)}(\boldsymbol{U}[\boldsymbol{c}]u^{(1)})(\theta_{\boldsymbol{y}},\psi_{\boldsymbol{y}},\boldsymbol{y}) \sim \frac{1}{8\pi^2}\int_{E_{\boldsymbol{\alpha}^{\boldsymbol{0}}}}\int_{\mathbb{R}} \\ &\cdot\left[\frac{|\boldsymbol{\gamma}^0_{\boldsymbol{y}}(\boldsymbol{y},\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}},\theta_{\boldsymbol{y}},\psi_{\boldsymbol{y}})|^4}{A(\boldsymbol{y})}[\Lambda(\boldsymbol{y},\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}})]^{-1}\boldsymbol{w}(\boldsymbol{y},\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}},\theta_{\boldsymbol{y}},\psi_{\boldsymbol{y}})\right] \\ &\delta(t-T(\boldsymbol{y}))\big(D_{(\theta,\psi)} - D_{(\theta,\psi)}T(\boldsymbol{y})\partial_t\big) \\ &\cdot\overset{r}{\xi}_p(\boldsymbol{r})\partial_t u^{(1)}(\boldsymbol{s},\boldsymbol{r},t)\overset{s}{\xi}_q(\boldsymbol{s})\frac{\partial(\overset{s}{\boldsymbol{\alpha}},\overset{r}{\boldsymbol{\alpha}})}{\partial(\overset{0}{\boldsymbol{\alpha}},\theta,\psi)}\mathrm{d}t\,\mathrm{d}\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{y}} \\ &\sim \boldsymbol{U}[\boldsymbol{c}](D_{(\theta,\psi)} - P[\boldsymbol{c}]\partial_t)\partial_t u^{(1)}(\boldsymbol{s},\boldsymbol{r},t)\,. \end{aligned} \tag{64}$$

Here the acquisition coordinates $\boldsymbol{s}$ and $\boldsymbol{r}$ are functions of $\theta,\psi,\overset{0}{\boldsymbol{\alpha}}$ according to the mapping given in Section 2. The last result can be viewed as a commutator

$$\big[D_{(\theta,\psi)},\boldsymbol{U}[\boldsymbol{c}]\big] = -\boldsymbol{U}[\boldsymbol{c}]P[\boldsymbol{c}]\partial_t\,. \tag{65}$$

For the direct scattering operator, a similar calculation

gives

$$
\begin{aligned}
&D_{(\theta,\psi)}(\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L)(\boldsymbol{s},\boldsymbol{r},t) \sim \\
&-\int_{\mathbb{R}}\left[\int_{\phi=L} \overset{s}{\xi}_q\,(\boldsymbol{s})A(\boldsymbol{x})\,\overset{r}{\xi}_p\,(\boldsymbol{r})\;(\boldsymbol{\alpha}\cdot\overset{0}{\boldsymbol{\gamma}})\right|_{\boldsymbol{x}} \\
&\cdot\left(D_{(\theta,\psi)}+D_{(\theta,\psi)}T(\boldsymbol{x})\right)\delta(t-T(\boldsymbol{x}))\,\frac{1}{|\nabla_{\boldsymbol{x}}\phi|} \\
&\cdot\mathsf{R}_L(\boldsymbol{x},\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}},\overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}})\mathrm{d}\Sigma(\boldsymbol{x})\mathrm{d}L \\
&\sim\; \boldsymbol{L}[\boldsymbol{c}^{(0)}]D_{(\theta,\psi)}+P[\boldsymbol{c}^{(0)}]\partial_t\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L\,. \qquad (66)
\end{aligned}
$$

This again gives the commutator

$$
\left[D_{(\theta,\psi)},\boldsymbol{L}[\boldsymbol{c}^{(0)}]\right]=P[\boldsymbol{c}^{(0)}]\partial_t\boldsymbol{L}[\boldsymbol{c}^{(0)}]\,. \qquad (67)
$$

For notational purposes in both of these expressions we have defined

$$
P[\,\cdot\,]=D_{(\theta,\psi)}T(\,\cdot\,)=D_{(\theta,\psi)}\big(\tau(\boldsymbol{s},\,\cdot\,)+\tau(\,\cdot\,,\boldsymbol{r})\big)\,. \qquad (68)
$$

Using the fact that, for the true background $\boldsymbol{c}^{(0)}$ and some perturbation $\boldsymbol{c}^{(1)}$, the data satisfy

$$
\boldsymbol{u}^{(1)}(\boldsymbol{s},\boldsymbol{r},t)\sim\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L\,, \qquad (69)
$$

and that $D_{(\theta,\psi)}\mathsf{R}_L=0$, we calculate the following derivative

$$
\begin{aligned}
&D_{(\theta,\psi)}\big(\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L\big) \\
&\sim\big(\boldsymbol{U}[\boldsymbol{c}]D_{(\theta,\psi)}+\left[D_{(\theta,\psi)},\boldsymbol{U}[\boldsymbol{c}]\right]\big)\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L \\
&\sim\boldsymbol{U}[\boldsymbol{c}]\big(D_{(\theta,\psi)}\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L-P[\boldsymbol{c}]\partial_t\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L\big) \\
&\sim\boldsymbol{U}[\boldsymbol{c}]\big(\left[D_{(\theta,\psi)},\boldsymbol{L}[\boldsymbol{c}^{(0)}]\right]\mathsf{R}_L+\boldsymbol{L}[\boldsymbol{c}^{(0)}]D_{(\theta,\psi)}\mathsf{R}_L \\
&\quad -P[\boldsymbol{c}]\partial_t\boldsymbol{L}[\boldsymbol{c}^{(0)}]\mathsf{R}_L\big) \\
&\sim\boldsymbol{U}[\boldsymbol{c}]\left(P[\boldsymbol{c}^{(0)}]-P[\boldsymbol{c}]\right)\partial_t\boldsymbol{u}^{(1)}(\boldsymbol{s},\boldsymbol{r},t)\,. \qquad (70)
\end{aligned}
$$

Then, by applying the direct scattering operator $\boldsymbol{L}$ on the left-hand side of this expression and rearranging the terms we get

$$
\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\sim\left[P[\boldsymbol{c}^{(0)}]-P[\boldsymbol{c}]\right]\partial_t\,. \qquad (71)
$$

Using this, the sensitivity transform can be written as

$$
\boldsymbol{S}\;=\;[I-\Delta]^{-1/2}\left[P[\boldsymbol{c}^{(0)}]-P[\boldsymbol{c}]\right]\partial_t\,. \qquad (72)
$$

This can be interpreted as a weighted version of the error in ray takeoff slowness at the image-point. This quantity is directly related to scattering-angle, azimuth, local geological dip and spatial position (see Figure 4) since a perturbation in the velocity necessarily will change the traveltime along the ray-paths. This again will induce changes in the ray-slowness and hence in scattering-angle/azimuth and the corresponding estimate of the geological dip. If we have the correct background model when doing the inversion, that is $P[\boldsymbol{c}]=P[\boldsymbol{c}^{(0)}]$, the sensitivity transform of the result will vanish or be close to zero.

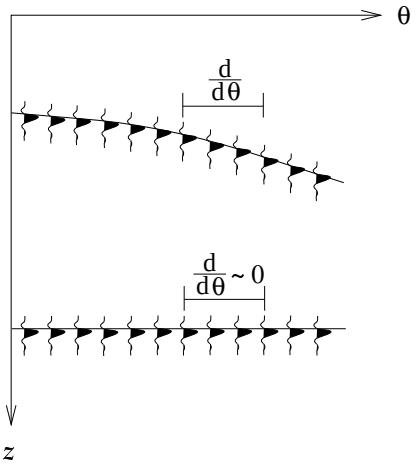

Figure 3. In 2D: the differential semblance between neighboring traces in a common scattering-angle gather ($\psi=0$).

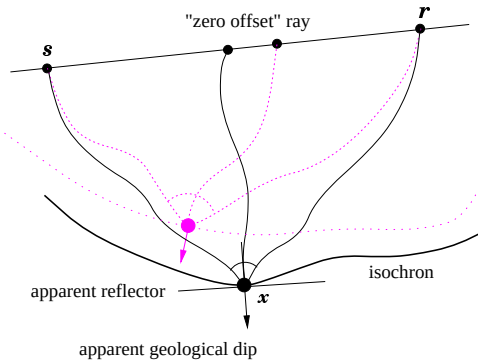

Figure 4. In 2D: slowness perturbation and the induced perturbations in position, dip and scattering-angle

The differential semblance operator

To define what Symes (1999) denotes as raw differential semblance, the L^2 norm must be applied to the sensitivity transform. Let $\langle\,.\,,\,.\,\rangle$ be the L^2 inner product. Then the objective function J_0 is the differential semblance operator (DSO), defined as

$$
\begin{aligned}
&J_0[\boldsymbol{c}]\sim\langle\boldsymbol{S},\boldsymbol{S}\rangle= \\
&\frac{1}{2}\,\|\,[I-\Delta]^{-1/2}\left[P[\boldsymbol{c}^{(0)}]-P[\boldsymbol{c}]\right]\partial_t\boldsymbol{u}^{(1)}(\boldsymbol{s},\boldsymbol{r},t)\|^2\,. \qquad (73)
\end{aligned}
$$

This is a "raw" measure since it does not take into account any noise in the data. The objective function J_0 is what needs to be minimized in order to get the correct background model. The way to achieve this is to match the slowness in the trial background model with that in the true model. Note that the amplitudes are completely missing in this expression. This is due to the application of the direct scattering operator $\boldsymbol{L}$ to the inverted sections and the fact that the expression includes only the leading-order terms. To have control over the amplitude

is, on the other hand, important when applying the derivative in the angle domain. In our approach we try to compensate for amplitude versus angle (AVA) behavior; in theory all of the AVA behavior can be removed by reconstructing the angle-independent perturbation $\widehat{\mathsf{R}}$ given in equation (56). All the images made with different scattering-angles/azimuths should have the same image pulse signature, so that only spatial variations remain as an angle-dependent feature. This ensures that the derivative only measures residual moveout and not artifacts in pulse shape, and so on.

The gradient

The gradient of the raw differential semblance is taken with respect to the background medium parameters $\boldsymbol{c}$. This gradient corresponds to a Frechét derivative of equation (73) (Tarantola, 1987). The objective function can be minimized after finding the gradient and setting $\nabla_{\boldsymbol{c}} J_0[\boldsymbol{c}] = 0$. Using the L^2 inner product, the first-order perturbation of $J_0[\boldsymbol{c}]$ due to a medium perturbation $\delta\boldsymbol{c}$ is

$$\begin{aligned}\delta J_0[\boldsymbol{c}] &= DJ_0[\boldsymbol{c}]\delta\boldsymbol{c} \\ &= \Big\langle [I-\Delta]^{-1/2}\delta\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\,\boldsymbol{u}^{(1)} + \\ &\quad \boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\delta\boldsymbol{U}[\boldsymbol{c}]\,\boldsymbol{u}^{(1)} \\ &\quad , [I-\Delta]^{-1/2}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\,\boldsymbol{u}^{(1)}\Big\rangle \\ &= \langle \delta\boldsymbol{c}, \nabla_{\boldsymbol{c}} J_0\rangle\,. \end{aligned} \tag{74}$$

This expression can be manipulated such that we can separate out the gradient $\nabla_{\boldsymbol{c}} J_0$ inside the inner product

$$\delta J_0 = \langle \delta\boldsymbol{c}, \nabla_{\boldsymbol{c}} J_0\rangle = \cdots = \int \mathrm{d}\boldsymbol{x}\,\delta\boldsymbol{c}\,(\nabla_{\boldsymbol{c}} J_0)\,. \tag{75}$$

Using the relations for the derivative with respect to the medium given in equations (65) and (67) we can write

$$\begin{aligned}&\delta\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}] + \boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\delta\boldsymbol{U}[\boldsymbol{c}] \sim \\ &\quad (D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c})D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}] - \\ &\quad \boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}](D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c})\boldsymbol{U}[\boldsymbol{c}]\,. \end{aligned} \tag{76}$$

Using this the perturbation in J_0 can be written

$$\begin{aligned}\delta J_0[\boldsymbol{c}] \sim\ & \Big\langle [I-\Delta]^{-1/2}\big((D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c})D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\big)\boldsymbol{u}^{(1)} \\ &-\big(\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}](D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c})\boldsymbol{U}[\boldsymbol{c}]\big)\boldsymbol{u}^{(1)} \\ &, [I-\Delta]^{-1/2}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\,\boldsymbol{u}^{(1)}\Big\rangle\,. \end{aligned} \tag{77}$$

Then, by use of the adjoint operators the terms can be re-arranged as

$$\begin{aligned}\delta J_0[\boldsymbol{c}] \sim\ & \Big\langle \delta\boldsymbol{c}\,,\, D\boldsymbol{L}[\boldsymbol{c}]^*\big[D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{u}^{(1)}\,, \\ &[I-\Delta]^{-1}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{u}^{(1)}\big] \\ &-D\boldsymbol{L}[\boldsymbol{c}]^*\big[\boldsymbol{U}[\boldsymbol{c}])\boldsymbol{u}^{(1)}\,,\,(\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}])^\star \\ &\cdot[I-\Delta]^{-1}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\,\boldsymbol{u}^{(1)}\big]\Big\rangle \\ =\ & \langle \delta\boldsymbol{c}\,,\, D\boldsymbol{L}[\boldsymbol{c}]^* W\rangle\,, \end{aligned} \tag{78}$$

where the smooth contributions from the commutators have been combined into the weight factor

$$\begin{aligned}W =\ & \big[D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{u}^{(1)}\,,\,[I-\Delta]^{-1}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{u}^{(1)}\big] \\ -\ & \big[\boldsymbol{U}[\boldsymbol{c}])\boldsymbol{u}^{(1)}\,,\,(\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}])^\star \\ &\cdot[I-\Delta]^{-1}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\,\boldsymbol{u}^{(1)}\big]\,. \end{aligned} \tag{79}$$

The gradient can be identified on the right-hand side of equation (78) as

$$\begin{aligned}\nabla_{\boldsymbol{c}} J_0[\boldsymbol{c}] \sim\ & D\boldsymbol{L}[\boldsymbol{c}]^*\big[D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{u}^{(1)}\,,\,[I-\Delta]^{-1}\boldsymbol{L}[\boldsymbol{c}] \\ &\cdot D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\boldsymbol{u}^{(1)}\big] - D\boldsymbol{L}[\boldsymbol{c}]^*\big[\boldsymbol{U}[\boldsymbol{c}])\boldsymbol{u}^{(1)}\,, \\ &(\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}])^\star[I-\Delta]^{-1}\boldsymbol{L}[\boldsymbol{c}]D_{(\theta,\psi)}\boldsymbol{U}[\boldsymbol{c}]\ \ \boldsymbol{u}^{(1)}\big] \\ &= D\boldsymbol{L}[\boldsymbol{c}]^* W\,. \end{aligned} \tag{80}$$

To get an explicit integral representation of the gradient we substitute the direct and inverse scattering operators given in Section 4, and the expression for $D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c}$ given in equation (61) into (80). Starting from the definition of δJ_0 with the inner product taken directly as integrals over the phase angles $\overset{s}{\boldsymbol{\alpha}}$ and $\overset{r}{\boldsymbol{\alpha}}$, the gradient can be expressed as

$$\begin{aligned}&\delta J_0[\boldsymbol{c}] = \\ &\int\int\int d\,\overset{s}{\boldsymbol{\alpha}}\,d\,\overset{r}{\boldsymbol{\alpha}}\,dt(D\boldsymbol{L}[\boldsymbol{c}]\delta\boldsymbol{c})\,W(\overset{s}{\boldsymbol{\alpha}},\overset{r}{\boldsymbol{\alpha}},t) = \\ &\int\int \mathrm{d}\,\overset{s}{\boldsymbol{\alpha}}\int \mathrm{d}\boldsymbol{x}\,(D\tau[\boldsymbol{c}]\delta\boldsymbol{c})(\overset{s}{\boldsymbol{\alpha}},\boldsymbol{x})\delta(t-\tau(\overset{s}{\boldsymbol{\alpha}},\boldsymbol{x})-\tau(\overset{r}{\boldsymbol{\alpha}},\boldsymbol{x})) \\ &\quad\cdot\left(\int \mathrm{d}\,\overset{r}{\boldsymbol{\alpha}}\,A(\overset{s}{\boldsymbol{\alpha}},\boldsymbol{x},\overset{r}{\boldsymbol{\alpha}})W(\overset{s}{\boldsymbol{\alpha}},\overset{r}{\boldsymbol{\alpha}},t)\right) + \\ &\int\int \mathrm{d}\,\overset{r}{\boldsymbol{\alpha}}\int \mathrm{d}\boldsymbol{x}\,(D\tau[\boldsymbol{c}]\delta\boldsymbol{c})(\overset{r}{\boldsymbol{\alpha}},\boldsymbol{x})\delta(t-\tau(\overset{s}{\boldsymbol{\alpha}},\boldsymbol{x})-\tau(\overset{r}{\boldsymbol{\alpha}},\boldsymbol{x})) \\ &\quad\cdot\left(\int \mathrm{d}\,\overset{s}{\boldsymbol{\alpha}}\,A(\overset{s}{\boldsymbol{\alpha}},\boldsymbol{x},\overset{r}{\boldsymbol{\alpha}})W(\overset{s}{\boldsymbol{\alpha}},\overset{r}{\boldsymbol{\alpha}},t)\right) = \\ &\int \mathrm{d}\boldsymbol{x}\,\delta\boldsymbol{c}(\boldsymbol{x}) \\ &\cdot\left(\int \mathrm{d}\,\overset{s}{\boldsymbol{\alpha}}\,(D\tau)^*[\boldsymbol{c}](\overset{s}{\boldsymbol{\alpha}},\boldsymbol{x})\int \mathrm{d}\,\overset{r}{\boldsymbol{\alpha}}\ W(\overset{s}{\boldsymbol{\alpha}},\overset{r}{\boldsymbol{\alpha}},T(\boldsymbol{x}))\right. \\ &\left.+\int \mathrm{d}\,\overset{r}{\boldsymbol{\alpha}}\,(D\tau)^*[\boldsymbol{c}](\boldsymbol{x},\overset{r}{\boldsymbol{\alpha}})\int \mathrm{d}\,\overset{s}{\boldsymbol{\alpha}}\ W(\overset{s}{\boldsymbol{\alpha}},\overset{r}{\boldsymbol{\alpha}},T(\boldsymbol{x}))\right)\,. \end{aligned} \tag{81}$$

The polarization vectors and all other smooth contributions are collected into W to simplify the expression.

Now, changing the order of integration, collecting terms with equal dependency and making use of the mapping

$$(\overset{0}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \theta_{\boldsymbol{x}}, \psi_{\boldsymbol{x}}) \rightarrow (\overset{s}{\boldsymbol{\alpha}}_{\boldsymbol{x}}, \overset{r}{\boldsymbol{\alpha}}_{\boldsymbol{x}}) \rightarrow (\boldsymbol{s}, \boldsymbol{r}),$$

that was treated in Section 2.2., the gradient can be separated inside the L^2 norm over the model space

$$\nabla_{\boldsymbol{c}} J_0[\boldsymbol{c}] = \int \mathrm{d}\boldsymbol{s} (D\tau)^{\star}[\boldsymbol{c}](\boldsymbol{s}, \cdot) \left(\int \mathrm{d}\boldsymbol{r} W(\boldsymbol{s}, \boldsymbol{r}, T(\cdot)) \right) + \int \mathrm{d}\boldsymbol{r} (D\tau)^{\star}[\boldsymbol{c}](\boldsymbol{r}, \cdot) \left(\int \mathrm{d}\boldsymbol{s} W(\boldsymbol{s}, \boldsymbol{r}, T(\cdot)) \right), \quad (82)$$

where now also the amplitudes are included in W.

Traveltime mappings, derivatives and their adjoints

To be able to compute $\nabla J_0[\boldsymbol{c}]$ we need to relate the traveltime perturbations to perturbations in the medium through the adjoint mapping $(D\tau[\boldsymbol{c}])^{\star}$. Araya et al. (1996) have performed this calculation for a 1D layered medium, and we have closely followed this derivation, but for a general velocity and with scattering-angle/azimuth as the parameters of the DSO. The scattering-angle/azimuth domain creates a setting that closely resembles the layered 1D medium case. This is due to the fact that we use the dip field in the model and the specular ray geometry. The analogy is further improved if we assume that within each Fresnel zone in the subsurface there is only one point where the migration dip and local geological coincide, giving what we can define as a *pseudo-layered* medium. This simplifies the GRT's and make them into maps that only change variables. A rigorous treatment of this concept and an extension of the properties derived for the gradient in 1D (Symes, 1999), will be the subject of another paper. Here we focus only on the proposed algorithm and some of its aspects.

The traveltime perturbation can be related to a perturbation in slowness through ray-perturbation theory (Snieder & Spencer, 1993) , (Farra & Madriaga, 1987), and the first-order perturbation is given by

$$\delta\tau = \int_{raypath} \delta s \, \mathrm{d}l, \quad (83)$$

where

δ s = 1/velocity: slowness along the raypath,
l = the distance along the ray path.

The slowness can be found directly from the medium representation $\boldsymbol{c}$, but to keep the notation simple while explaining the concept we will continue to use $\boldsymbol{c}$. The expression above is a consequence of Fermat's principle and says that the first-order perturbation in traveltime is just the integral of the slowness perturbation along the reference ray. Based on ray perturbation theory, the traveltime perturbations can be related to the model perturbations even when the reference ray is not a true ray; hence we can correct explicitly for the error in traveltime due to the fact that the curves used in the inversion are not true rays in the current model. Equation (83) is the basis for linearized traveltime tomography, which will be used to perform the adjoint mapping $(D\tau[\boldsymbol{c}])^{\star}$, in form of a back-projection of slowness perturbations into the model space.

The traveltime perturbation is related to the medium perturbation by

$$\delta\tau = D\tau[\boldsymbol{c}]\delta\boldsymbol{c}. \quad (84)$$

Hence the adjoint operator $(D\tau)^{\star}$ is nothing other than the linearized back-projection operator in tomography, which relates the traveltime perturbation to a distributed slowness perturbation along the raypaths. The adjoint is taken with respect to the model parameterization and is given by the inner product

$$\langle \delta\boldsymbol{c}, (D\tau[\boldsymbol{c}])^{\star}\delta\tau \rangle = \langle D\tau[\boldsymbol{c}]\delta\boldsymbol{c}, \delta\tau \rangle. \quad (85)$$

A short calculation gives the adjoint mapping

$$\delta\boldsymbol{c} = (D\tau[\boldsymbol{c}])^{\star}\delta\tau = \frac{\int_{E_\psi}\int_{E_\theta} l(\theta,\psi)\delta\tau(\theta,\psi)\delta N(\psi)}{\int_{E_\psi}\int_{E_\theta} l^2(\theta,\psi)\delta N(\psi)}, \quad (86)$$

where the quantity δN is related to the ray density in the model

$$\delta N(\psi) = \frac{\#\text{ rays}}{\theta_{max}} \quad (87)$$

Here the integrals are justified if there is approximately a continuum in the ray coverage in the model space. Whenever this is not the case other formulations can be used (Stork & Clayton, 1991). Equation (86) performs a spreading of the adjoint input along the rays and is computed over all rays connecting an image point in the subsurface with the acquisition surfaces. Applying the adjoint operator will produce a mesh of updated control points in the model, and the weighted sum from all rays is used to construct the update.

Discussion

We have presented a method that will allow us to perform velocity analysis in a highly automatic way. It eliminates the need for picking common-image gathers or employing any other interpretive step. The differential semblance principle will perform the "picking" automatically and provide us with a background model that minimizes the misfit in traveltimes. The sensitivity trans-

form and the differential semblance operator analyze the lateral continuity in common scattering-angle/azimuth gathers, where the reconstructed reflectivity should be independent of (θ, ψ) as long as the velocity model is correct. Before the lateral comparison we remove any angle-dependent variations in the reflectivity by undoing the effect of the point contrast-source radiation pattern. This makes the reconstructed quantity $\widehat{\mathsf{R}}$ independent of any angular effects but velocity.

We postulate that the differential semblance operator will have the same asymptotic behavior in the angle domain as in the 1D layered case; that is, all stationary points are global minimizers (Symes, 1999). The property that we describe as pseudo-layering, and various tests performed in common offset, all point in the direction of this being true. A rigorous proof of this property is the subject of some work in progress.

The general algorithm we propose looks something like

I Invert the data using equation (66), creating the angle sections and an estimate of the dip field.

II Remove AVA variations from the amplitudes assuming linearized reflection coefficients.

III Apply the derivative in the angle domain: take difference between adjacent common angle sections.

IV Model data from the result of the previous step using equation (28).

V Compute the gradient and run through a optimization scheme, i.e., Quasi-Newton method to find global minimum.

VI Update the background model using tomographic back-projection on the current ray set.

VII Repeat with the updated model until misfit is less than some specified threshold.

Implementation of this method has started, and we hope to show some numerical results in a few months.

Acknowledgments

Discussions with Sam Gray have been of great importance in this development. Sverre Brandsberg-Dahl also wants to thank Amoco Norway Oil Company w/Olav I. Barkved and The Norwegian Research Council for financial support.

References

Araya, K., Kim, S., Nolan, J., & Symes, W. 1996. Velocity inversion and high frequency asymptotics. *In: Mathematical Method in Geophysical Imaging.* Society of Photo-optical Instrumentation Engineering.

Beylkin, G, & Burridge, R. 1984. The inversion problem and applications of the generalized Radon Transform. *Comm. Pure Appl. Math.*, **37**, 579–599.

Burridge, R., de Hoop, M. V., Miller, D., & Spencer, C. 1998. Multiparameter inversion in anisotropic elastic media. *Geophys. J. Int.*, **138**, 757–777.

Chauris, H., & Noble, M. 1998. Testing the behavior of Differential Semblance for velocity estimation. *In: Expanded Abstracts.* SEG.

de Hoop, M. V., & Bleistein, N. 1997. Generalized Radon transform inversion for reflectivity in anisotropic elastic media. *Inverse Problems*, **13**, 669–690.

de Hoop, M. V., & Brandsberg-Dahl, S. 1999 (March). *Maslov extension of generalized Radon transform inversion in anisotropic elastic media: A least-squares approach.* CWP 279P.

Deregowski, S. M. 1990. Common Offset Migration and Velocity Analysis. *First Break*, **8**(6), 225–234.

Farra, V., & Madriaga, R. 1987. Seismic waveform modeling in heterogeneous media by ray perturbation theory. *Journal of Geophysical Research*, **92**(B3), 2697–2712.

Miller, D., Oristaglio, M., & Beylkin, G. 1987. A new slant on seismic imaging: migration and integral geometry. *Geophysics*, **52**(6), 943–964.

Snieder, R, & Spencer, C. 1993. A unified approach to ray bending, ray perturbation and paraxial ray theories. *Geophys. J. Int.*, **115**, 456–470.

Stork, C., & Clayton, R. W. 1991. An implementation of tomographic velocity analysis. *Geophysics*, **56**(4), 483–495.

Symes, W. 1996. *High Frequency Asymptotics and Velocity Estimation.* Rice Inversion Project, Annual Report.

Symes, W. 1999. *All stationary points of differential semblance are asymptotic global minimizers: layered acoustics.* Presented at CWP seminar February 1999.

Tarantola, A. 1987. *Inverse Problem Theory.* 3 edn. Elsevier.