### PP/PS anisotropic stereotomography

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Geophysical Journal International</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>draft</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Research Paper</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Nag, Steinar; SINTEF Petroleum Research, Seismic and Reservoir Department; Norwegian University of Science and Technology, Department of Petroleum Eng. and Applied Geophysics Alerini, Mathias; SINTEF Petroleum Research, Seismic and Reservoir Department Ursin, Bjørn; Norwegian University of Science and Technology, Department of Petroleum Eng. and Applied Geophysics</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Seismic tomography &lt; SEISMOLOGY, Seismic anisotropy &lt; SEISMOLOGY, Tomography &lt; GEOPHYSICAL METHODS</td>
</tr>
</tbody>
</table>
**PP/PS anisotropic stereotomography**

Steinar Nag\(^1\), Mathias Alerini\(^1\) and Bjørn Ursin\(^2\)

\(^1\) SINTEF Petroleum Research, 7465 Trondheim, Norway; email: steinar@nag.name
\(^2\) Norwegian University of Science and Technology, Department of Petroleum Eng. and Applied Geophysics, 7491 Trondheim, Norway

**SUMMARY**

Stereotomography is a slope tomographic method which gives good results for background velocity model estimation in 2D isotropic media. We develop here the extension of the method to 3D general anisotropic media for PP and PS events. As in isotropic media, the sensitivity matrix of the inversion can be computed by paraxial ray tracing. Based on ray perturbation theory, we derive all derivatives of stereotomography data parameters with respect to model parameters in a 3D general anisotropic medium. These general formulas for the derivatives can also be used in other applications that rely on anisotropic ray perturbation theory. The derivatives are expressed using the Voigt notation for the elastic model parameters. We include a Jacobian that allows us to change the model parameterization to Thomsen parameters. Explicit expressions for the derivatives of the data are given for 2D tilted transversely isotropic (TTI) media. These derivatives are tested and validated numerically for a PS event. We verify the possibilities of the method by separate estimation of each Thomsen parameter field of a 2D TTI synthetic model used for modelling data by ray tracing. For each Thomsen parameter, the estimated field fits well with the exact field.

**Key words:** seismic anisotropy; seismic tomography
1 INTRODUCTION

Seismic anisotropy is an important and challenging issue in the processing of seismic data (Helbig & Thomsen 2005). In anisotropic media, the velocity of seismic waves depends on the direction of wave propagation. Seismic anisotropy affects both the phase and amplitude of the seismic signal. Not taking anisotropy into account in seismic processing can lead to increased errors in the results, in particular in the depth of imaged subsurface reflectors. The $S$-waves are usually more affected by anisotropy than the $P$-waves. In the processing of $PP/PS$ ocean-bottom-seismic (OBS) data, anisotropy should be taken into account (Szydlik et al. 2007).

Seismic velocity estimation is a crucial step in the depth imaging process. Anisotropy not only complicates the forward modelling, but also the optimization process, by introducing more unknown parameters and non-uniqueness. Velocity model estimation from seismic reflection data is traditionally performed using a linearization of the non-linear inverse problem (Tarantola 2005). The inverse problem can then be solved by iteratively updating the model to reduce the misfit between observed and calculated data. This update of seismic velocity model parameters becomes a major challenge when the model is anisotropic and the number of velocity model parameters increases. Even in the simple case of a transversely isotropic medium (Thomsen 1986), the inverse problem suffers due to lack of data and non-uniqueness. The velocity estimation then requires regularization and additional constraints on the parameters, like well log information.

Migration-based velocity estimation methods (Al-Yahya 1989; Symes & Carazzone 1991) offer a dense data coverage, while tomography methods (Dines & Lytle 1979) are faster and therefore more appropriate for large datasets. However, reflection traveltime tomography (Bishop et al. 1985; Farra & Madariaga 1988; Chapman & Pratt 1992) faces the problem of a difficult and costly interpretative picking. To remedy this problem, stereotomography has been introduced by Billette & Lambaré (1998). Stereotomography is a slope tomographic method (Riabinkin 1957; Sword 1987) that uses the slopes of locally coherent events (Lambaré 2002) to estimate a background velocity model.

Stereotomography has shown its simplicity of use and its quality in different contexts; in particular, Alerini et al. (2007) and Alerini et al. (2008) extended stereotomography for the estimation of $P$- and $S$-wave isotropic velocity models. Although their results were very encouraging, the method should be extended to anisotropic media in order to improve processing of OBS data. Motivated by the results of Alerini et al. (2007) we give here the theoretical extension of 3D and 2D $PP/PS$-stereotomography to include anisotropy.

This anisotropic extension of stereotomography presents several difficulties. The calculation of the derivatives of the data with respect to the model parameters represents a major difficulty compared with the isotropic case. As in isotropic media, we use paraxial ray theory (Farra & Madariaga 1987)
to compute the derivatives. Paraxial ray theory can provide all the required quantities of the derivative matrix (sensitivity matrix), and several approaches are possible. We decide here to use the general 3D anisotropic ray tracing (Červený 2001) to derive our results for the most general type of anisotropic elastic media. However, as Thomsen’s (1986) parameters are often more practical for anisotropic velocity estimation, we will introduce in our equations a Jacobian allowing the change of parameters from Voigt parameters to Thomsen parameters. Notice that using other kinds of anisotropic model parameters (Mensch & Farra 1999) is therefore straightforward.

Another difficulty, which is not considered here, is the shear-wave degeneracy that may occur for 3D weak anisotropic media when the two quasi-shear waves $qS_1$ and $qS_2$ have nearly the same velocity. Using general anisotropic ray theory to compute the eigenvectors can lead to numerical instabilities when the medium is weakly anisotropic. Quasi-isotropic ray theory (Kravtsov & Orlov 1990; Pšenčík 1998) can certainly be used in such cases.

Unlike the PP/PS isotropic inversion, extra constraints are needed for PP/PS anisotropic inversion, since the inverse problem is generally ill-posed. Barbosa et al. (2008) give an anisotropic extension of stereotomography for PP events in 3D elliptic and anelliptic media. The authors study the sensitivity of stereotomography to anisotropic parameters with and without transmitted events. They solve the ill-posedness of the inverse problem by using the refracted waves recorded at wells. Foss et al. (2005) proposed, in the context of differential semblance optimization (Symes & Carazzone 1991), to add an extra constraint based on the depth mismatch between migrated PP and PS key reflectors. Their approach can be applied to stereotomography in a rather straightforward manner. Both the approach of Foss et al. (2005) and of Barbosa et al. (2008) offer additional possibilities for constraining the anisotropic velocity estimation.

The main purpose of this paper is to present the derivation of all the quantities we need in order to compute the sensitivity matrix of stereotomography in 3D general anisotropic media. In the first part we define the stereotomographic data and model vectors. Their components are given based on the choice of boundary conditions for ray pairs. The data vector depends non-linearly on the model vector through the model parameters. We use a linearization of this dependence to solve for the model vector update. The second part of this paper gives the derivation of general formulas for calculation of the stereotomographic sensitivity matrix valid in 3D general anisotropic media. Some general results from ray theory are derived in the Appendices, including explicit expressions for the derivatives of the data in the case of a 2D TTI medium. In the last part of the paper, we include numerical tests for the derivatives of stereotomographic data in a 2D TTI medium, both with respect to ray model and velocity model parameters. These tests verify that we can compute the sensitivity matrix for PP/PS stereotomography in 2D TTI media. We also give examples of anisotropic velocity estimation by
stereotomography where we estimate separately each of the five 2D TTI Thomsen parameters relevant for $P$-waves and $SV$-waves, keeping the other fields equal to the exact fields. These synthetic tests correspond to the idealized case of perfect picking of the data and dense regular coverage of the reflection/diffraction points in the subsurface.

2 METHODOLOGY

2.1 Principles of stereotomography

Slope tomographic methods (Riabinin 1957; Sword 1987) are based on picking slope information in the data in addition to the traveltimes. The picking identifies locally coherent events (Lambaré 2002) and can be automated (Billette et al. 2003). For common source and common receiver gathers, a locally coherent event can be described by the source and receiver positions $x_s, x_r$, the slopes $q_s, q_r$ picked in the gathers and the two-way traveltime $T_{sr}$. Such locally coherent events provide information on the velocity macro-model independently of each other. The idea is that a locally coherent event identified in two orthogonal gathers constrains the background velocity model.

Stereotomography was proposed by Billette & Lambaré (1998) as the most general slope tomographic method. The particularity of stereotomography is that all the data of each locally picked event are inverted. The method provides not only information on the velocity model, but also information on the parameters describing the pair of ray segments from the reflection/diffraction point in depth to the acquisition surface. For each event, the ray parameters are the reflection/diffraction point $x_0$, the polar and azimuth phase angles at $x_0$ for rays towards source and receiver, and the one-way traveltimes $T^s$ and $T^r$ for rays towards source and receiver.

In isotropic $PP$-stereotomography (Billette & Lambaré 1998), all picked events are interpreted as primary $PP$ reflections or diffractions. In isotropic $PP/PS$-stereotomography (Alerini et al. 2007), events have to be interpreted as $PP$ or $PS$ primary reflections or diffractions, where wave mode conversion occurs at the reflection or diffraction point. It should be noted that $PS$ diffractions are weak events. In our 3D anisotropic approach, similar assumptions are made: events have to be interpreted as $qPP$, $qPS_1$ or $qPS_2$ with conversion at the reflection/diffraction point.

Stereotomography provides the advantage of an easier picking compared to classical traveltime tomography methods, and it is computationally faster than migration-based velocity estimation methods. The cost function of stereotomography has been shown to be equivalent (Chauris et al. 2002) to that of differential semblance optimization (DSO) (Symes & Carazzone 1991). Stereotomography was first developed for 2D isotropic media with picking of $PP$ events in the time domain (Billette & Lambaré 1998). Later extensions of the method include 3D $PP$ streamer data (Chalard et al. 2002).
2D isotropic PP/PS events (Alerini et al. 2007) and 2D first-arrival VSP data (Gosselet et al. 2003, 2005). Although picking is easier in stereotomography than in classical tomography, it remains the bottleneck of the approach. Several solutions have been proposed: picking in the depth domain (Chauris et al. 2002; Nguyen et al. 2008), picking in the poststack time domain (Lavaud et al. 2004), and strategies for selecting the correct picked events (Lambaré 2004).

2.2 Data vector

Billette & Lambaré (1998) suggested to minimize the misfit on all data parameters for stereotomography. This allows one to consider uncertainties on all of the observables. Here, we consider that there are no uncertainties on the depths of sources and receivers (Duveneck 2004). We can therefore fix the vertical positions $x^s_3$ and $x^r_3$ at source and receiver. This leads to possible instabilities with grazing rays and some extra computations to obtain the needed derivatives (Section 4.3). However, the size of the inverse problem is significantly reduced (Section 2.7). We believe that this is an important issue, especially for anisotropic inversion where the number of unknowns may be quite high. The data vector in our approach to stereotomography is therefore

$$d = \left( [(x^s)^t, (x^r)^t, (q^s)^t, (q^r)^t, T^{sr}]_{i=1}^{N_j} \right)_{j=1}^3,$$

where from Fig. 1 we have that $x^s = (x^s_1, x^s_2)^t$ and $x^r = (x^r_1, x^r_2)^t$ are the lateral positions of the sources and receivers, $q^s = (\tilde{p}^s_1, \tilde{p}^s_2)^t$ and $q^r = (\tilde{p}^r_1, \tilde{p}^r_2)^t$ are the slopes measured relative to the acquisition surface at sources and receivers, and $T^{sr}$ is the two-way traveltime associated with each event. $N_j$ denotes the number of picked events for each wave mode ($q_{PS1}, q_{PS2}$ or $q_{PP}$, indexed by $j$) of a reflection/diffraction. The picked slopes $q^s$ and $q^r$ can be related to the horizontal slowness vector $\hat{p}$ from ray theory Červený (2001) using the rotation

$$\hat{p} = Rp,$$

where $\hat{p} = (\tilde{p}^s_1, \tilde{p}^s_2, \tilde{p}^s_3)^t$ is the slowness vector measured in the acquisition surface coordinate system $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$ (see Fig. 2) and $R$ is a rotation matrix defined from the geometry of the acquisition surface (see Appendix A).

Note that (1) is a general expression for 3D ideal acquisitions. In practice, for 3D we usually have access to only three of the four slopes present in (1). This is due to the sparse acquisition in one of the directions: cross-line source direction for conventional streamer cases and cross-line receiver direction for conventional ocean-bottom-cable (OBC) cases. Chalard et al. (2002) have shown that the missing fourth slope does not cause a problem for the 3D stereotomographic inversion for most practical applications. In particular, for the OBC case, we can therefore take $q^r = (\tilde{p}^r_1)$. In the 2D
case, we assume no velocity variation in the $x_2$-direction and the corresponding quantities can be
suppressed: $x^s = (x^s_1)$, $x^r = (x^r_1)$, $q^s = (\tilde{p}^s_1)$, $q^r = (\tilde{p}^r_1)$.

An important practical point can be noted here. As soon as we can separate the different wave
modes and an efficient multiple attenuation tool has been applied, picking the observed data for
$PP/PS$ anisotropic stereotomography can be performed exactly in the same way as for $PP/PS$
isotropic stereotomography. In particular, an automatic picking can be performed (Billette et al. 2003).

2.3 Model vector

The model vector corresponding to the choice of data vector (1) can be written as

$$m = \left( \left[ (x^s_{0i}, (\Theta^s)_i)^t, (\Theta^r)_i \right]_{i=1}^{N_j} \right)^t, m^b \right)^t. \tag{3}$$

In (3), $x_0 = (x_{01}, x_{02}, x_{03})^t$ is the reflection/diffraction point in the subsurface, while $\Theta^s = (\theta^s, \psi^s)^t$
and $\Theta^r = (\theta^r, \psi^r)^t$ are the phase direction angles at $x_0$ of rays towards source and receiver (see Fig.
3). Finally, $m_b$ denotes the collection of parameters (see Section 2.4) that describe the background
velocity model. Note here as well the necessary adaptations to the 2D case, where only $\Theta^s = (\theta^s)$
and $\Theta^r = (\theta^r)$ are needed, since $\psi^s$ and $\psi^r$ can be suppressed.

2.4 Parameterization of the background velocity model

Ray tracing can be developed both for blocky or smooth velocity models, or a combination of both
(Červený 2001). We choose a smooth velocity model description for anisotropic stereotomography to
avoid dealing with complex ray codes and boundary conditions at interfaces, that would complicate
the computation of the derivatives of the data.

The background velocity model $m_b$ can typically be described for smooth media using the notation
$m_b = \left( \left( C_{k=1}^{M_l} \right)^L \right).$ That is, each of the $L$ elastic velocity parameters denoted by $C$
(e.g. vertical velocities and anisotropic parameters) are described using $L$ smooth fields, with each field parameter-
ized by $M_l$ parameters (e.g. cubic B-spline knot values (de Boor 2001)).

There are several ways to parameterize the elastic model parameters for anisotropic velocity es-
timation. The most general way is to use the full density-normalized elastic stiffness tensor $a_{ijkl} =
c_{ijkl} / \rho$, where $\rho$ is the density and $c_{ijkl}$ is the stiffness tensor. Due to the symmetry properties of
c_{ijkl}, the Voigt notation $C_{\alpha\beta}$ is usually preferred. The Voigt notation (Voigt 1910) for elastic stiffness
components is given by a mapping between stiffness tensor $c_{ijkl}$ and the Voigt parameters $C_{\alpha\beta}$. The
mappings $(ij) \rightarrow \alpha$ and $(kl) \rightarrow \beta$ are defined by

$$11 \rightarrow 1, \quad 22 \rightarrow 2, \quad 33 \rightarrow 3, \quad 23 = 32 \rightarrow 4, \quad 13 = 31 \rightarrow 5, \quad 12 = 21 \rightarrow 6. \tag{4}$$
As for the density-normalized stiffness components $a_{ijkl} = c_{ijkl}/\rho$, we can define the density-normalized Voigt parameters $A_{\alpha\beta}$. $a_{ijkl}$ can be mapped into a symmetric matrix using the Voigt notation. For a transversely isotropic medium, the Voigt matrix in the local crystal coordinate system $\hat{x}_1\hat{x}_2\hat{x}_3$ (see Fig. 4) is given by

$$A_{\text{TI}} = \begin{pmatrix}
\hat{A}_{11} & \hat{A}_{12} & \hat{A}_{13} & 0 & 0 & 0 \\
\hat{A}_{12} & \hat{A}_{11} & \hat{A}_{13} & 0 & 0 & 0 \\
\hat{A}_{13} & \hat{A}_{13} & \hat{A}_{33} & 0 & 0 & 0 \\
0 & 0 & 0 & \hat{A}_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & \hat{A}_{44} & 0 \\
0 & 0 & 0 & 0 & 0 & \hat{A}_{66}
\end{pmatrix}, \tag{5}$$

which has 5 independent parameters $(\hat{A}_{11}, \hat{A}_{33}, \hat{A}_{44}, \hat{A}_{66}, \hat{A}_{13})$, because of the relation $\hat{A}_{12} = \hat{A}_{11} - 2\hat{A}_{66}$ (Červený 2001).

The symmetry axis can be tilted in 3D by a rotation of the local crystal coordinate system, using the angle $\varphi$ with the vertical $x_3$ axis and the angle $\xi$ with the horizontal $x_1$ axis (see Fig. 4). For the Voigt matrix this rotation can be done using a Bond transformation (Bond 1943),

$$A_{\text{TII}} = MA_{\text{TI}}M^T, \tag{6}$$

where the matrix $M$ depends only on the rotation angles $\varphi$ and $\xi$.

The resulting Voigt matrix for a tilted transversely isotropic (TTI) medium,

$$A_{\text{TII}} = \begin{pmatrix}
A_{11} & A_{12} & A_{13} & 0 & A_{15} & 0 \\
A_{12} & A_{22} & A_{23} & 0 & A_{25} & 0 \\
A_{13} & A_{23} & A_{33} & 0 & A_{35} & 0 \\
0 & 0 & 0 & A_{44} & 0 & A_{46} \\
0 & 0 & 0 & 0 & A_{55} & 0 \\
0 & 0 & 0 & 0 & 0 & A_{66}
\end{pmatrix}, \tag{7}$$

has in general 13 non-zero components, but only 7 parameters are independent (the 5 independent parameters of the crystal TI Voigt matrix (5) and the two rotation angles $\varphi$ and $\xi$).

Thomsen’s (1986) parameters for transversely isotropic (TI) media are defined from the density-
normalized Voigt parameters $\alpha_{\alpha\beta}$ by the relations

\begin{align*}
V_{P0} &= \sqrt{\hat{A}_{33}}, \\
V_{S0} &= \sqrt{\hat{A}_{44}}, \\
\varepsilon &= \frac{\hat{A}_{11} - \hat{A}_{33}}{2\hat{A}_{33}}, \\
\delta &= \frac{(\hat{A}_{44} + \hat{A}_{44})^2 - (\hat{A}_{33} - \hat{A}_{44})^2}{2\hat{A}_{33}(\hat{A}_{33} - \hat{A}_{44})}, \\
\gamma &= \frac{\hat{A}_{66} - \hat{A}_{44}}{2\hat{A}_{44}}.
\end{align*}

(8)

where $V_{P0}$ and $V_{S0}$ are the $P$-wave and $SV$-wave velocities, respectively, for propagation in the direction of the symmetry axis. $\varepsilon$, $\delta$ and $\gamma$ are dimensionless parameters that are equal to zero in isotropic media. Thomsen's parameters are therefore commonly used for anisotropic velocity estimation, after an input isotropic model has provided estimates of $V_{P0}$ and $V_{S0}$.

2.5 Stereotomographic inversion

We follow here the stereotomographic inversion strategy proposed by Billette & Lambaré (1998), which is:

- initialization (definition of an initial model, see Appendix G)
- relocalization (inversion of the ray model parameters in the initial model, stabilizing the subsequent joint inversion)
- joint inversion of both ray and velocity model parameters

The joint optimization of the full model $m$ of (3) is performed through the minimization of an $\ell_2$-norm cost function. We consider here the part of the cost function

\[ J(m) = \frac{1}{2} (\|\Delta d\|^2 + \epsilon^2 \|Lm\|^2), \]

(9)

where the data misfit

\[ \Delta d(m) = d_{\text{calc}}(m_{\text{ray}}, m_b) - d_{\text{obs}} \]

(10)

depends on both the ray model vector $m_{\text{ray}}$ and the background velocity model $m_b$. The calculated data $d_{\text{calc}}(m)$ are computed by anisotropic ray tracing in the model $m$. The second term in (9) is a regularization term, such as a weighted Laplacian operator. This operator can be defined as a constraint on the smoothness of the velocity model. By using an appropriate weighting factor $\epsilon$, it can regularize
the inverse problem and help to avoid oscillations in velocity models parameterized using cubic B
spline functions.

The non-linear operator \( m \rightarrow d \) takes a vector \( m \) from the model space into the data space. Linearizing around the current model \( m \) gives a system of equations

\[
\Delta d \approx J \Delta m,
\]

which can be solved for the update vector \( \Delta m \). The linear system (11) is typically solved by the LSQR method (Paige & Saunders 1982), which is suitable for solving systems with large sparse matrices like \( J \). The linearization of the inverse problem may not always be accurate enough to achieve convergence when updating both ray the model and the velocity model at the same time, especially for anisotropic velocity estimation. However, this issue is taken into account in our update algorithm by using a relocalization of stereotomographic rays in joint inversion when needed.

### 2.6 The sensitivity matrix

In (11), \( J \) is a Jacobian matrix or the so-called Fréchet derivative matrix (Tarantola 2005), which contains the derivatives of the data parameters with respect to the model parameters. As we will show, ray perturbation theory (Farra & Madariaga 1987) provides all the necessary quantities for computing \( J \), which we from here on refer to as the sensitivity matrix.

Consider the data and model spaces given in equations (1) and (3) for one picked event. The sensitivity matrix that is needed for anisotropic stereotomography is then

\[
J = \frac{\partial d_{\text{calc}}(m)}{\partial (m^1_{\text{ray}}, m^1_{b})} = 
\begin{pmatrix}
\frac{\partial x^s}{\partial x_0} & \frac{\partial x^s}{\partial (\Theta)^s} & 0 & \frac{\partial x^s}{\partial m^1_b} \\
\frac{\partial x^r}{\partial x_0} & 0 & \frac{\partial x^r}{\partial (\Theta)^r} & \frac{\partial x^r}{\partial m^1_b} \\
\frac{\partial q^s}{\partial x_0} & \frac{\partial q^s}{\partial (\Theta)^s} & 0 & \frac{\partial q^s}{\partial m^1_b} \\
\frac{\partial q^r}{\partial x_0} & 0 & \frac{\partial q^r}{\partial (\Theta)^r} & \frac{\partial q^r}{\partial m^1_b}
\end{pmatrix}_L
\]

The derivatives are computed relative to the depth level of source or receiver, which is indicated by the subscript \( L \) on the sensitivity matrix. Here we have used a vectorial notation to provide a general expression for 3D or 2D media. As \( x^s_0 \) and \( x^r_0 \) are fixed, we have in 3D that \( x^s = (x^s_1, x^s_2)^t \), \( x^r = (x^r_1, x^r_2)^t \) and \( \Theta = (\theta, \psi)^t \) (see Fig. 3). In the 2D case, we take \( x^s = (x^s_1)^t \), \( x^r = (x^r_1)^t \) and \( \Theta = (\theta) \), as \( x_2 \) and \( \psi \) can be suppressed.
2.7 Analysis of the size of the inverse problem

The stereotomographic inversion is typically ill-posed in some regions and over-constrained in other regions of the background velocity model. Anisotropy introduces more unknown background velocity model parameters, contributing to the non-uniqueness of the problem. In 3D, memory handling can be an issue, due to the large number of picked events. We present here an analysis of the size of the inverse problem for stereotomography.

For the 3D case, with the data defined according to (1) and the model defined according to (3), it follows that

\[
\text{dim}(\mathbf{d}) = 9 \cdot N_{\text{obs}},
\]

\[
\text{dim}(\mathbf{m}) = 7 \cdot N_{\text{obs}} + N_{\text{mb}},
\]

where \(N_{\text{obs}} = \sum_{j=1}^{3} N_j\) is the number of observed picked events and \(N_{\text{mb}}\) is the number of background velocity model parameters. It is therefore required that \(2N_{\text{obs}} \geq N_{\text{mb}}\) to sufficiently constrain all parameters of the inversion (assuming full ray coverage and independent picks). However, often only three slopes are available, so that it is required that \(N_{\text{obs}} \geq N_{\text{mb}}\). Similarly for the 2D case, we have

\[
\text{dim}(\mathbf{d}) = 5 \cdot N_{\text{obs}},
\]

\[
\text{dim}(\mathbf{m}) = 4 \cdot N_{\text{obs}} + N_{\text{mb}}.
\]

It is therefore required that \(N_{\text{obs}} \geq N_{\text{mb}}\) for the 2D case. In practice, due to limited ray coverage and errors in the data, most applications will require that \(N_{\text{obs}} \gg N_{\text{mb}}\) both for the 3D and 2D case.

Stereotomography has been formulated in a slightly different way by previous authors by also relaxing the vertical components of sources and receivers (Billette & Lambaré 1998). Then the one-way traveltimes \(T_s^r\) and \(T_r^s\) have to be included in the ray model \(\mathbf{m}_{\text{ray}}\) in order to limit the length of the rays. This approach gives an additional misfit on the vertical components of sources and receivers. The derivatives of the data would then be computed along the wavefront (indicated by subscript \(W\) on \(\mathbf{J}\)) instead of along the source/receiver depth level. The Jacobian matrix in that case would be

\[
\mathbf{J} = \frac{\partial \mathbf{d}_{\text{calc}}(\mathbf{m})}{\partial (\mathbf{m}_{\text{ray}}, \mathbf{m}_b)} = \begin{pmatrix}
\frac{\partial x_s}{\partial \mathbf{x}} & \frac{\partial x_s}{\partial \Theta_s} & 0 & \frac{\partial x_s}{\partial T_s^r} & 0 & \frac{\partial x_s}{\partial m_{\text{mb}}} \\
\frac{\partial y_s}{\partial \mathbf{x}} & \frac{\partial y_s}{\partial \Theta_s} & 0 & \frac{\partial y_s}{\partial T_s^r} & 0 & \frac{\partial y_s}{\partial m_{\text{mb}}} \\
\frac{\partial q_s}{\partial \mathbf{x}} & \frac{\partial q_s}{\partial \Theta_s} & 0 & \frac{\partial q_s}{\partial T_s^r} & 0 & \frac{\partial q_s}{\partial m_{\text{mb}}} \\
0 & 0 & 0 & \frac{\partial T_r^s}{\partial T_s^r} & \frac{\partial T_r^s}{\partial T_s^r} & 0
\end{pmatrix} \bigg|_{W}.
\]
where in this case $\mathbf{x}^a = (x_1^a, x_2^a, x_3^a)^T$ and $\mathbf{x}^i = (x_1^i, x_2^i, x_3^i)^T$, as $x_3^a$ and $x_3^i$ are not fixed.

The Jacobian matrix (12) in 3D has in general dimension $(d) \cdot \text{dim} (m) = 63 N_{\text{obs}}^2 + 9 \cdot N_{\text{obs}} N_{\text{mb}}$ elements, but the number of non-zero elements cannot be larger than $47 N_{\text{obs}}^2 + 9 \cdot N_{\text{obs}} N_{\text{mb}}$. The Jacobian matrix (15) in 3D has in general $\text{dim} (d) \cdot \text{dim} (m) = 99 N_{\text{obs}}^2 + 10 \cdot N_{\text{obs}} N_{\text{mb}}$ elements, but the number of non-zero elements cannot be larger than $62 N_{\text{obs}}^2 + 10 \cdot N_{\text{obs}} N_{\text{mb}}$.

The Jacobian matrix (12) in 2D has in general dimension $(d) \cdot \text{dim} (m) = 20 N_{\text{obs}}^2 + 5 \cdot N_{\text{obs}} N_{\text{mb}}$ elements, but the number of non-zero elements cannot be larger than $16 N_{\text{obs}}^2 + 5 \cdot N_{\text{obs}} N_{\text{mb}}$. The Jacobian matrix (15) in 2D has in general dimension $(d) \cdot \text{dim} (m) = 42 N_{\text{obs}}^2 + 7 \cdot N_{\text{obs}} N_{\text{mb}}$ elements, but the number of non-zero elements cannot be larger than $26 N_{\text{obs}}^2 + 6 \cdot N_{\text{obs}} N_{\text{mb}}$.

The number of non-zero elements of the Jacobian matrix is therefore significantly smaller when the rays are stopped at the known depth level of sources and receivers. This can be important in 3D when the memory requirements can be very high. As an example, take $N_{\text{obs}} = 10000$ and $N_{\text{mb}} = 21 \cdot 21 \cdot 11 = 4851$ for a 3D isotropic inversion of one velocity field. The size of the storage of the velocity model part of the sensitivity matrix will be less than 2 GBytes and similar using both (12) and (15). However, the larger ray model part will be reduced from $62 \cdot N_{\text{obs}}^2 \cdot 4 \text{ Bytes} \approx 25 \text{ GBytes}$ using (15) to $47 \cdot N_{\text{obs}}^2 \cdot 4 \text{ Bytes} \approx 19 \text{ GBytes}$ using (12).

3 RAY AND PARAXIAL RAY THEORY

In the high-frequency approximation of ray theory (Červený 2001) it can be shown that a ray solution of the elastodynamic equation must satisfy the Christoffel equation

$$\left( \Gamma_{ik} - \delta_{ik} \right) g_i^{(m)} = 0, \quad i, k = 1, 2, 3;$$

(16)

where $\Gamma_{ik}$ are elements of the Christoffel matrix $\Gamma$, defined by

$$\Gamma_{ik} = \alpha_{ijkl} p_j p_l, \quad \Gamma_{ik} = \Gamma_{kl}, \quad i, j, k, l = 1, 2, 3.$$

(17)

In (16), $\delta_{ik}$ is the Kronecker delta, $(m)$ represents the wave mode of the considered ray and $g_i^{(m)}$ is the polarization vector associated with ray mode $(m)$. In (17), $\alpha_{ijkl}$ is the density-normalized stiffness tensor and $p_i$ is the slowness vector. We here use the Einstein summation convention for repeated subscripts. The summation indices will not be listed in the following equations.

The $g_i^{(m)}$ are the eigenvectors of the Christoffel matrix (17) and satisfy

$$\left( \Gamma_{ik} - G^{(m)} \delta_{ik} \right) g_i^{(m)} = 0,$$

(18)

where there is no summation over $(m)$. The eigenvalues $G_i^{(m)}$ can be computed by solving the char-
acteristic equation
\[ \det \left( \Gamma_{ik} - G^{(m)} \delta_{ik} \right) = 0, \] (19)
and can also be expressed using the eigenvectors according to the relation
\[ G^{(m)} = \Gamma_{ik} g_{i}^{(m)} g_{k}^{(m)}. \] (20)

The ray Hamiltonian for general anisotropic ray tracing can be defined (Červený 2001) as
\[ \mathcal{H}(x, p) = \frac{1}{2} \left( G^{(m)}(x, p) - 1 \right). \] (21)

Each ray solution must satisfy \( \mathcal{H} = 0 \) along the ray, so the eigenvalues for the ray with mode \( (m) \) must satisfy
\[ G^{(m)} = 1. \] (22)

The unit vector \( n \) in the direction of the slowness vector \( p \) has components
\[ n_1 = \cos \psi \sin \theta, \]
\[ n_2 = \sin \psi \sin \theta, \]
\[ n_3 = \cos \theta, \] (23)
where \( \theta \) and \( \psi \) are the polar and azimuthal phase angles (see Fig. 3). Using \( n_i \) we can define the phase-normalized Christoffel matrix \( \Gamma' \) (Červený 2001) using
\[ \Gamma'_{ik} = a_{ijkl} n_j n_l, \] (24)
so that
\[ \Gamma'_{ik} = v^2 \Gamma_{ik}. \] (25)

\( G^{(m)} \) are the eigenvalues of \( \Gamma' \) and satisfy
\[ G^{(m)} = v^2, \] (26)
where \( v \) is the phase velocity of the ray with wave mode \( (m) \).

The perturbations in position and slowness of a paraxial ray along the wavefront are given from Farra & Madariaga (1987),
\[ \Delta y(t) = \Pi(t, t_0) \Delta y_0 + \int_{t_0}^{t} \Pi(t, t') B(\Delta \mathcal{H}(t')) dt', \] (27)
where \( \Delta y(t) \) is the perturbation of the phase space vector \( y = (x^t, p^t) \) at time \( t \) and \( \Delta y_0 \) is the perturbation of the phase space vector \( y_0 = (x_0^t, p_0^t) \) at time \( t_0 \). \( \Pi \) is the ray propagator matrix and
\[ B(x(t), p(t)) = \begin{pmatrix} \nabla_p \Delta \mathcal{H}(x(t), p(t)) \\ -\nabla_x \Delta \mathcal{H}(x(t), p(t)) \end{pmatrix}. \] (28)
is a source matrix for integration along the ray of the perturbations $\Delta H$ of the ray Hamiltonian due to perturbations $\Delta m_b$. As the one-way traveltimes $T^s$ and $T^r$ are not part of the ray model vector in our approach, the paraxial ray expression given in (27) does not contain the term depending on the traveltime as given by Billette & Lambaré (1998).

From (27) we can compute the derivatives $\partial y/\partial x^t$, $\partial y/\partial \Theta^t$, and $\partial y/\partial m^t_b$. A correction onto the depth level (see Section 4.3) and the rotation (2) must be performed in order to get the derivatives needed in the sensitivity matrix (12): $\partial(y_L^t, T^{sr})^t/\partial x^t$, $\partial(y_L^t, T^{sr})^t/\partial \Theta^t$, and $\partial(y_L^t, T^{sr})^t/\partial m^t_b$, where $y_L = (x^t, q^t)^t$. These derivatives are computed by paraxial ray tracing, which requires that the fields of the velocity model must be twice continuously differentiable.

### 4 CALCULATION OF THE SENSITIVITY MATRIX $J$

We can develop expressions for the perturbation of the Hamiltonian directly in terms of the perturbations of the density-normalized stiffness tensor $a_{ijkl}$ or in terms of the perturbations of the Voigt parameters $A_{ij}$. However, other medium parameters can be more convenient to use. The change of parameters can be expressed by the Jacobian

$$\frac{\partial f_V(x)}{\partial f^t(x)},$$

(29)

where the vector function $f_V(x)$ collects the fields of the relevant density-normalized elastic stiffness coefficients in the Voigt notation $A_{\alpha\beta}$, and the vector function $f(x)$ collects the relevant parameter fields of the desired notation used for velocity estimation. Notice that with this approach it is straightforward to use any type of parameters as, for example, Thomsen’s (1986) parameters or the extension of Thomsen parameters to orthorombic media as proposed by Mensch & Farra (1999). The only requirement will be to update the Jacobian $\partial f_V/\partial f^t$ with appropriate parameter fields $f$.

In the following we choose to express the perturbation of the Hamiltonian by perturbations in B-spline knot values of Thomsen parameters. We will here use the Thomsen parameters for a 3D tilted transversely isotropic (TTI) medium. The parameter fields are $f_T = (V_P^0, V_S^0, \varepsilon, \delta, \gamma, \varphi, \xi)^t$, of which the first 5 are defined in (8). The polar angle $\varphi$ and azimuth angle $\xi$ (see Fig. 4) determine the direction of the symmetry axis. A TTI medium can be described from a transverse isotropic (TI) medium by a Bond transformation (Bond 1943), using the angles $\varphi$ and $\xi$ for the rotation. The Jacobian (29) for the TTI case is given in Appendix B.

We first give the derivatives of position and slowness with respect to ray model parameters at the initial point of the central ray. Using the propagator matrix from dynamic ray tracing, we obtain the corresponding derivatives along the wavefront at the end-point of the central ray. We then give the derivatives of position and slowness due to changes in the background velocity parameters $m_b$. 

Finally, we correct all derivatives along the wavefront onto the depth level of sources or receivers by considering the change in traveltime that follows from this restriction. Some details of the derivations are given in the Appendices. Explicit expressions for the derivatives in the 2D TTI case are given in Appendix H.

4.1 Derivatives at the initial point

We compute in the following the derivatives of position and slowness at the initial point of the ray (the reflection/diffraction point $x_0$).

$J^0_{x_0}$ - derivatives with respect to reflection point position

The initial point for both rays in a stereotomography ray pair is the reflection/diffraction point $x_0 = (x_{01}, x_{02}, x_{03})$. Perturbations in the initial point $x_0$ give perturbations in the magnitude of the initial slowness due to the heterogeneity of the background velocity fields. The direction of the initial slowness is determined by the angles $\theta$ and $\psi$ that are part of the ray model vector. These angles are kept fixed for perturbations in $x_0$. The Jacobian $\partial y_0/\partial x_0$ can then be written as

$$J^0_{x_0} = \frac{\partial}{\partial x_0} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} \frac{\partial x_0}{\partial x_0} \\ \frac{\partial (x_0/\eta_0)}{\partial x_0} \end{pmatrix} = \begin{pmatrix} I \\ -\frac{1}{\eta_0} p_0 \frac{\partial \eta_0}{\partial x_0} \end{pmatrix},$$

where $I$ is the $3 \times 3$ unit matrix and $n_0$ is the fixed unit vector in the direction of the initial slowness $p_0$ (see Fig. 3). For the derivatives of the phase velocity we have in general that

$$\frac{\partial v}{\partial x_0} = \sum_{f_V} \frac{\partial v}{\partial f_V} \frac{\partial f_V}{\partial x_0} = \sum_{f_V} \sum_{f_T} \frac{\partial v}{\partial f_V} \frac{\partial f_V}{\partial f_T} \frac{\partial f_T}{\partial x_0}.$$  

In (31), there is a summation over the components of the Jacobian $\partial f_V/\partial f_T$ (given in Appendix B) from Voigt parameters to Thomsen parameters. From (E19) we have the result

$$\frac{\partial v}{\partial f_V} = \frac{v}{2} \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} \begin{pmatrix} g_i^{(m)} \\ g_k^{(m)} \end{pmatrix},$$

where

$$\frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} = \frac{\partial a_{ijkl}^{(m)}}{\partial f_V} \bigg|_{x,p} p_{0j} p_{0k},$$

from the definition of $\Gamma_{ik}$ in (17). Note that a vertical bar and subscripts have been used to indicate quantities that are kept fixed when computing derivatives. The particular expressions for $\partial \Gamma_{ik}/\partial f_V|_{x,p}$ and $\partial v/\partial f_V$ in a 2D TTI medium are derived in Appendix H.
The phase angles $\Theta = (\theta, \psi)^t$ of the ray determine the direction of the slowness vector at $x_0$. In anisotropic media, the phase angles also have an influence on the magnitude of the slowness vector. Thus

$$J_\Theta^0 = \frac{\partial}{\partial \Theta_0} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\partial}{\partial \Theta_0} \left( \frac{1}{v} n_0 \right) \end{pmatrix}$$

(34)

The derivatives of phase velocity with respect to the phase angles are given from (E16) as

$$\frac{\partial v}{\partial \Theta} = \frac{1}{2v} \frac{\partial \Gamma_{ik}'}{\partial \Theta} g^{(m)}_i g^{(m)}_k,$$

(35)

where it follows from the definition of $\Gamma_{ik}'$ in (24) that

$$\frac{\partial \Gamma_{ik}'}{\partial \Theta} \bigg|_{x_0} = (a_{ijkl} + a_{ilkj}) n_0 j \frac{\partial n_0}{\partial \Theta} L,$$

(36)

The unit phase vector $n_0$ can be expressed in terms of the angles $\Theta = (\theta, \psi)^t$ (see definition in (23)) and the derivatives $\frac{\partial n_0}{\partial \Theta}$ follow directly from the definition of $n_0$. The particular expressions for $\frac{\partial \Gamma_{ik}'}{\partial \Theta} \bigg|_{x_0}$ and $\frac{\partial v}{\partial \Theta}$ in a 2D TTI medium have are derived in Appendix H.

**$J_f^0$ - derivatives with respect to velocity fields**

Perturbations of the velocity fields $f$ around the initial point $x_0$ lead to perturbations in the magnitude of the initial slowness vector $p_0$. We therefore have the derivatives

$$J_f^0 = \frac{\partial}{\partial f} \begin{pmatrix} x_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\partial}{\partial f} \left( \frac{1}{v} n_0 \right) \end{pmatrix}$$

(37)

Here we have used that $\frac{\partial n_0}{\partial f} = 0$, since $n_0$ depends only on the initial phase angles $\Theta_0$ that are kept fixed for perturbations in the velocity fields $f$. The derivatives $\frac{\partial v}{\partial f}$ of the phase velocity with respect to the velocity fields are given above in (32) and given for a 2D TTI medium in Appendix H.

### 4.2 Derivatives along the wavefront

Given a central ray starting from an image point and ending on a source or receiver position, paraxial ray tracing can be used to compute the perturbations of the ray end-point position and slowness due to perturbations in the model parameters. We here describe how to compute the derivatives corresponding to such perturbations along the wavefront of the central ray. That is, when the paraxial ray has the same
traveltime as the central ray. When there are no perturbations along the ray, the integral term in (27) is zero, and multiplying the derivatives at the initial point of the ray by the ray propagator matrix $\Pi$ gives the derivatives
\[
J_{x_0}(y(t)) = \Pi(t,t_0)J_{x_0}^0(y(t_0)),
\]
\[
J_{\Theta}(y(t)) = \Pi(t,t_0)J_{\Theta}^0(y(t_0)),
\]
along the wavefront at the ray end point. When perturbing the velocity fields along the central ray, we also have to take into account the integral term
\[
\int_{t_0}^{t} \Pi(t,t')B(\Delta H(t')) dt'
\]
in the paraxial ray expression (27). The $B$-matrix in (39) depends on the phase-space derivatives of the perturbations of the Hamiltonian along the central ray. The computation of this matrix $B$ is certainly the most challenging part of the extension of stereotomography to anisotropic media. We present here the derivation of this quantity for 3D general anisotropic media, including the Jacobian from Voigt parameters to Thomsen parameters. We assume here that the medium is twice continuously differentiable.

The perturbation of the Hamiltonian $H$ due to perturbations of the Voigt parameters $f_V$ can be expressed in the first order approximation as
\[
\Delta H = \sum_{f_V} \frac{\partial H}{\partial f_V} \bigg|_{x,p} \Delta f_V.
\]
In (40), we can develop further for perturbations in the Voigt parameters $f_V$ in terms of perturbations in the Thomsen parameters $f_T$
\[
\Delta f_V = \sum_{f_T} \frac{\partial f_V}{\partial f_T} \bigg|_x \Delta f_T.
\]
Let us first consider the perturbations of the Thomsen parameters, $\Delta f_T$, which depend upon the choice of background velocity parameters $m_b$. We here describe each of the velocity fields of $f_T$ by a cubic B-spline representation, to make the velocity fields twice continuously differentiable. Thus, any of the Thomsen parameters of $f_T$ are defined by the cubic B-spline tensor product given by
\[
f_T(x_1, x_2, x_3) = \sum_{n=1}^{N_{x_1}} \sum_{m=1}^{N_{x_2}} \sum_{l=1}^{N_{x_3}} v^{f_T}_{nml} b_n(x_1) b_m(x_2) b_l(x_3),
\]
where $v^{f_T}_{nml}$ is the spline coefficient for the knot with indices $(n, m, l)$. The functions $b_n(x_1), b_m(x_2)$ and $b_l(x_3)$ are the 1D cubic B-spline basis functions in the $x_1$, $x_2$, and $x_3$-directions respectively, centered about the knot with indices $(n, m, l)$. Note that the number of knots $N_{x_1}, N_{x_2}$, and $N_{x_3}$ and their spacing may be different for each Thomsen field $f_T$, therefore the basis functions in one of the directions will also in general be different for each Thomsen field $f_T$. We can now express the
perturbation in the Hamiltonian (40) directly in terms of perturbations in the B-spline knot values of the Thomsen parameters as

$$\Delta H = \sum_{f_v} \left. \frac{\partial H}{\partial f_v} \right|_{x, p} \sum_{f_r} \left. \frac{\partial f_r}{\partial f_T} \right|_x \sum_{n} \sum_{m} \sum_{l} \left. \frac{\partial f_T}{\partial \nu^{fr}_{nml}} \right|_x \Delta \nu^{fr}_{nml}. \tag{43}$$

Assume that we perturbate only one value $\nu^{fr}_{nml}$, that is, only one knot value of the cubic B-spline function representing one of the Thomsen parameters. The perturbation of the Hamiltonian that we need for the integrand in (39) is then

$$\Delta H = \sum_{f_v} \left. \frac{\partial H}{\partial f_v} \right|_{x, p} \left. \frac{\partial f_r}{\partial f_T} \right|_x \left. \frac{\partial f_T}{\partial \nu^{fr}_{nml}} \right|_x \Delta \nu^{fr}_{nml}. \tag{44}$$

As $\Delta \nu^{fr}_{nml}$ is independent of time and non-zero, we can divide both sides of (27) by it. Based on the assumption that the non-linear dependence $d(m)$ is such that the data parameters are differentiable with respect to model parameters, we can let $\Delta \nu^{fr}_{nml} \to 0$, which gives

$$\frac{\partial y}{\partial \nu^{fr}_{nml}} = \Pi(t, t_0) \frac{\partial y_0}{\partial \nu^{fr}_{nml}} + \int_{t_0}^{t} \Pi(t, t') B(y(t')) \, dt', \tag{45}$$

where we have redefined the source matrix as

$$B(y(t')) = \begin{pmatrix} \nabla_p \left[ \left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} \right] & \nabla_x \left[ \left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} \right] \\ -\nabla_x \left[ \left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} \right] & \nabla_p \left[ \left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} \right] \end{pmatrix}. \tag{46}$$

In the rest of this section we will express $B(y(t'))$ in detail.

From (40) and (44) we get

$$\left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} = \sum_{f_v} \left. \frac{\partial H(x, p)}{\partial f_v} \right|_{x, p} \left. \frac{\partial f_v}{\partial \nu^{fr}_{nml}} \right|_x, \tag{47}$$

where

$$\left. \frac{\partial f_v}{\partial \nu^{fr}_{nml}} \right|_x = \left. \frac{\partial f_v}{\partial f_T} \right|_x \left. \frac{\partial f_T}{\partial \nu^{fr}_{nml}} \right|_x. \tag{48}$$

Thus for the computation of $B$ in (46) it follows that

$$\nabla_p \left[ \left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} \right] = \sum_{f_v} \left\{ \nabla_p \left[ \left. \frac{\partial H(x, p)}{\partial f_v} \right|_{x, p} \right] \left. \frac{\partial f_v}{\partial \nu^{fr}_{nml}} \right|_x \right\}, \tag{49a}$$

$$\nabla_x \left[ \left. \frac{\partial H(x, p)}{\partial \nu^{fr}_{nml}} \right|_{x, p} \right] = \sum_{f_v} \left\{ \nabla_x \left[ \left. \frac{\partial H(x, p)}{\partial f_v} \right|_{x, p} \right] \left. \frac{\partial f_v}{\partial \nu^{fr}_{nml}} \right|_x + \left. \frac{\partial H(x, p)}{\partial f_v} \right|_{x, p} \nabla_x \left[ \left. \frac{\partial f_v}{\partial \nu^{fr}_{nml}} \right|_x \right] \right\}. \tag{49b}$$
The second term in (52) is zero from (F2), so that we have the general result from (33) that
\[
\nabla_x \left[ \frac{\partial f_V(x)}{\partial v_{nml}^f} \right] = \nabla_x \left[ \frac{\partial f_V(x)}{\partial f_T} \right] \nabla_x \left[ \frac{\partial f_T(x)}{\partial v_{nml}^f} \right].
\]

We now express the quantities in (49) in more detail. Inserting \(G^{(m)}\) from (20) into the definition (21) of the Hamiltonian \(\mathcal{H}\) we get
\[
\mathcal{H}(x, p) = \frac{1}{2} \left( \Gamma_{ik} g_i^{(m)} g_k^{(m)} - 1 \right).
\]

Taking the derivative of (51) with respect to the Voigt parameter field \(f_V\) and using that \(\Gamma_{ki} = \Gamma_{ik}\), it follows that
\[
\frac{\partial \mathcal{H}(x, p)}{\partial f_V} \bigg|_{x,p} = \frac{1}{2} \left( \Gamma_{ik} \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} g_i^{(m)} g_k^{(m)} + \Gamma_{ik} \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} (g_i^{(m)} g_k^{(m)}) \right)_{x,p}.
\]

The second term in (52) is zero from (F2), so that we have the general result
\[
\frac{\partial \mathcal{H}(x, p)}{\partial f_V} \bigg|_{x,p} = \frac{1}{2} \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} g_i^{(m)} g_k^{(m)}.
\]

Since \(\partial a_{ijkl} / \partial f_V\) is obviously equal to either zero or unity using the Voigt mapping (4), it follows from (33) that
\[
\nabla_x \left[ \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} \right] = 0.
\]

The derivatives of (53) with respect to position and slowness are therefore given as
\[
\nabla_p \left[ \frac{\partial \mathcal{H}}{\partial f_V} \bigg|_{x,p} \right] \bigg|_x = \frac{1}{2} \nabla_p \left[ \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} \right] \bigg|_x g_i^{(m)} g_k^{(m)} + \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} g_i^{(m)} \nabla_p \left[ g_k^{(m)} \right]_x,
\]
\[
\nabla_x \left[ \frac{\partial \mathcal{H}}{\partial f_V} \bigg|_{x,p} \right] \bigg|_p = \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} g_i^{(m)} \nabla_x \left[ g_k^{(m)} \right]_p.
\]

\(\partial \Gamma_{ik} / \partial f_V\) has been given in (33), from which it follows that
\[
\frac{\partial}{\partial p_n} \left[ \frac{\partial \Gamma_{ik}}{\partial f_V} \bigg|_{x,p} \right] \bigg|_x = \left( \frac{\partial a_{inkl}}{\partial f_V} \bigg|_{x} + \frac{\partial a_{inkm}}{\partial f_V} \bigg|_{x} \right) p_l,
\]

to be inserted into (55a).

The derivatives of the polarization vectors \(g_i^{(m)}\) with respect to position and slowness are given in Appendix F. The full Jacobian \(\partial V / \partial f_T^T\) from Voigt parameters to Thomsen parameters is given in Appendix B. The derivative of a Thomsen parameter field \(f_T(x)\) with respect to a knot value \(v_{nml}^f\) at the spatial position \(x\) is given from the cubic B-spline representation (42) as
\[
\frac{\partial f_T(x)}{\partial v_{nml}^f} = b_n(x_1)b_m(x_2)b_l(x_3).
\]

This quantity can be found practically by interpolating at \(x\) the B-spline field for \(f_T\) that has only
\( v_{nml} = 1 \) and the rest of knot values are set to zero. The spatial derivatives

\[
\nabla_x \left[ \frac{\partial f_T(x)}{\partial v_{nml}^T} \right] = \nabla_x \left[ b_n(x_1)b_m(x_2)b_l(x_3) \right]
\]

are found by computing derivatives in the same B-spline field as used to compute (57).

### 4.3 Derivatives relative to depth level of source or receiver

The derivatives (38) and (45) at the ray end point of position, slowness and cubic B-spline knot values are computed for constant one-way ray traveltime, that is, along the wavefront. For our implementation of stereotomography, perturbations along the wavefront are transformed onto the depth level of source or receiver. The wavefront can be defined for the source and receiver rays by

\[
T(x_0, p_0, a_{ijkl}) = T_s, \quad T(x_0, p_0, a_{ijkl}) = T_r,
\]

respectively. We let subscript S correspond to perturbations along the tangent to the acquisition surface. The acquisition surface is defined at source by \( u_s(x) = 0 \) and at receiver by \( u_r(x) = 0 \). Subscript L corresponds to perturbations relative to the depth level of source, defined by \( x_3 = (x_3^{obs}) \), or of receiver, defined by \( x_3 = (x_3^{obs}) \). We use the same subscript convention for the derivatives (see Fig. 5).

#### Derivatives of traveltime

Perturbing the medium parameters along a central ray, while keeping the ray endpoints fixed, gives the traveltime perturbation (Farra et al. 1989),

\[
\Delta T_0 = -\int_{t_0}^t \Delta H(\Delta a_{ijkl}) \, dt' = -\int_{t_0}^t \Delta a_{ijkl} p_i p_j g_i g_k \, dt'.
\]

When perturbing \( a_{ijkl} \), while keeping the one-way traveltime fixed, the wavefront of the paraxial ray will be perturbed compared to the wavefront of the central ray (see Fig. 5). The ray end point of the paraxial ray will then be different from the central ray, introducing a traveltime perturbation

\[
\Delta T_1 = p^l \cdot \Delta x.
\]

When the initial point \( x_0 \) of the ray is perturbed as well, the additional traveltime perturbation (see Fig. 5) is given by

\[
\Delta T_2 = -p^l_0 \cdot \Delta x_0.
\]
At the end point \( x \) of the central ray, let \( n_S \) be the normal to the acquisition surface at source or receiver. Then it can be shown (Farra et al. 1989; Farra & Le Bégat 1995) that the first-order perturbation

\[
\Delta T_3 = -\frac{n_S^t \cdot (\Delta x)^W}{n_S^t \cdot \nabla_p H}
\]

(63)
corrects the traveltime for a paraxial ray with perturbed initial slowness angle from the wavefront of the central ray onto the tangent line of the acquisition surface. The second order correction can be found in Farra (1999). In our implementation, we stop the rays at the depth level of source and receivers, and not on the acquisition surface tangent line. We can therefore replace \( n_S \) in (63) by

\[
\begin{align*}
\begin{bmatrix} n_{L} \end{bmatrix} &= (0, 0, 1)^t. 
\end{align*}
\]

The derivatives of traveltime along the depth level \( L \) with respect to initial position and initial phase direction angles (see Fig. 5) are then given by

\[
\begin{align*}
\frac{\partial T}{\partial x_0^L} &= \frac{\partial (T_0 + T_2 + T_3)}{\partial x_0^L} = \left( p^t \left( \frac{\partial x}{\partial x_0^L} \right)^W - p_0^t \right) - \frac{n_L^t \cdot (\partial x/\partial x_0^L)^W}{n_L^t \cdot \nabla_p H}, \\
\frac{\partial T}{\partial \Theta^L} &= \frac{\partial T_3}{\partial x_0^L} = -\frac{n_L^t \cdot (\partial x/\partial x_0^L)^W}{n_L^t \cdot \nabla_p H},
\end{align*}
\]

(65a, 65b)

where the derivatives \((\partial x/\partial x_0^L)^W\) and \((\partial x/\partial \Theta^L)^W\) are found in (38).

The derivatives of the traveltime with respect to cubic B-spline knot values along the depth level \( L \) are

\[
\begin{align*}
\frac{\partial T}{\partial v_{nml}^L} &= \frac{\partial (T_0 + T_2 + T_3)}{\partial v_{nml}^L} \\
&= -\int_{t_0}^{t_1} \frac{\partial a_{ijkl}}{\partial v_{nml}^L} p_j p_g g_k \, dt' + p^t \left( \frac{\partial x}{\partial v_{nml}^L} \right)^W - \frac{n_L^t \cdot (\partial x/\partial v_{nml}^L)^W}{n_L^t \cdot \nabla_p H},
\end{align*}
\]

(66)

where the integrand in the first term can be developed as

\[
\begin{align*}
\frac{\partial a_{ijkl}}{\partial v_{nml}^L} p_j p_g g_k &= \sum_{fv} \frac{\partial a_{ijkl}}{\partial f_v^L} \frac{\partial f_v^T}{\partial v_{nml}^L} p_j p_g g_k \\
&= \sum_{fv} \left\{ \frac{\partial a_{ijkl}}{\partial f_v^L} p_j p_g g_k \left[ \frac{\partial f_v^T}{\partial f_T^L} \frac{\partial f_T}{\partial v_{nml}^L} \right] \right\}.
\end{align*}
\]

(67)

Using (32) and (33), we can rewrite (67) in the more compact form

\[
\begin{align*}
\frac{\partial a_{ijkl}}{\partial v_{nml}^L} p_j p_g g_k &= \sum_{fv} \left\{ \frac{2}{v} \frac{\partial v}{\partial f_v^L} \left[ \frac{\partial f_v^T}{\partial f_T^L} \frac{\partial f_T}{\partial v_{nml}^L} \right] \right\}.
\end{align*}
\]

(68)

The Jacobian \( \partial f_v^L / \partial f_T^L \) can be found in Appendix B and \( \partial f_T / \partial v_{nml}^L \) is given in (57). \( \partial v / \partial f_v^L \) are given in (32) and in the case of a 2D TTI medium in (H9).
Derivatives of position and slowness

The first order corrections of perturbations of the position and slowness along the wavefront onto a tangent plane are given from Farra et al. (1989) as

\[
(\Delta x)_S = (\Delta x)_W + \frac{dx}{dT} \cdot \Delta T = (\Delta x)_W + \nabla_p \mathcal{H} \cdot \Delta T, \quad (69a)
\]

\[
(\Delta p)_S = (\Delta p)_W + \frac{dp}{dT} \cdot \Delta T = (\Delta p)_W - \nabla_x \mathcal{H} \cdot \Delta T. \quad (69b)
\]

The perturbations are expressed as a sum of the perturbation along the wavefront and a term that corrects the perturbation onto the tangent to the acquisition surface. For restriction to the depth level of sources and receivers, we replace \(n_S\) in (63) by \(n_L\) defined in (64). From (63) and (69) we then get the derivatives of position and slowness components with respect to initial position \(x_0\), initial phase direction angle \(\Theta\) and B-spline knot value \(v_{nml}^{fr}\) as

\[
\begin{pmatrix}
\frac{\partial x}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_L \\
\frac{\partial x}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_W \\
\end{pmatrix}
= \begin{pmatrix}
\frac{\partial x}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_L \\
\frac{\partial x}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_W \\
\end{pmatrix}
- n_L^t \cdot \frac{\partial x}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_W.
\]

(70a)

\[
\begin{pmatrix}
\frac{\partial p}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_L \\
\frac{\partial p}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_W \\
\end{pmatrix}
= \begin{pmatrix}
\frac{\partial p}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_L \\
\frac{\partial p}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_W \\
\end{pmatrix}
- n_L^t \cdot \frac{\partial p}{\partial \left( x_0^t, \Theta^t, v_{nml}^{fr} \right)}_W
+ \nabla_x \mathcal{H} \cdot \left( n_L^t \cdot \nabla_p \mathcal{H} \right)_W.
\]

(70b)

The derivatives of the slowness vector \(p\) in (70b) can be rotated using multiplication by the fixed rotation matrix \(R\) (see Appendix A) to give the derivatives of the slowness components \(q = (\tilde{p}_1, \tilde{p}_2)^t\) tangential to the acquisition surface at the end point of the central ray. The derivatives of \(q\) are needed in the sensitivity matrix (12).

5 NUMERICAL TESTS IN A 2D TTI MEDIUM

We first check the derivatives of the stereotomography data with respect to the model parameters \(m_{ray}\) and \(m_b\) for a single PS ray pair in a 2D TTI example model. We then test the anisotropic velocity estimation on a 2D TTI example where data are modelled by ray tracing. This corresponds to perfect noise-free picking of the data. We also control the ray coverage by the choice of ray model used for...
modelling the data. Here we only invert for one 2D TTI Thomsen parameter field, keeping the other four Thomsen parameter fields of the initial model equal to the fields of the synthetic model used for modelling the data. The purpose of these tests is to show that the sensitivity matrix that we compute can be used to derive an update to reduce the cost function. Using these updates we show that for any of 2D TTI Thomsen fields, the estimated field converges to the field of the synthetic model. To invert for all the 2D TTI Thomsen parameter fields \((V_{P0}, V_{S0}, \varepsilon, \delta, \varphi)\) will require additional constraints.

5.1 Velocity model

We define smooth velocity fields for the 2D TI parameters \(V_{P0}\), \(V_{S0}\), \(\varepsilon\), \(\delta\), and \(\varphi\) as defined by Thomsen (1986), and shown in Fig. 6. In addition, we include a field shown in Fig. 7 for the tilt angle \(\varphi\) (in degrees) for the rotation of the symmetry axis.

The fields of the vertical velocities \(V_{P0}\) and \(V_{S0}\) are represented by \(21 \times 9\) cubic B-spline knot values in the lateral and vertical directions, respectively, with a knot spacing of 500 m in both directions. The \(\varepsilon\), \(\delta\) and \(\varphi\) fields are represented by \(6 \times 5\) cubic B-spline knot values in the lateral and vertical directions, respectively, with a knot spacing of 2 km laterally and 1 km vertically.

5.2 Test of the derivatives of the data

A single PS ray pair is modelled by ray tracing as shown in Fig. 8. The derivatives of the data are first computed along the wavefront, before being corrected onto the depth level of source or receiver. In this case, we set both acquisition surfaces to be horizontal, so that \(q_1 = p_1\).

We test the derivative of each data parameter with respect to each model parameter. In the case of perturbations in the polar angle \(\theta\), we can write for the horizontal component \(p_1\) of the slowness vector

\[
(p_1(\theta + \Delta \theta))_L \approx p_1(\theta) + \left(\frac{\partial p_1(\theta)}{\partial \theta}\right)_L \Delta \theta,
\]

using a Taylor expansion to the first order. This is a linearization of the inverse problem \(d = f(m)\) with \(d = p_1\) and \(m = \theta\). In (71), \(\Delta \theta\) is the perturbation in the polar slowness angle \(\theta\) at the initial point of the central ray. \((\partial p_1(\theta)/\partial \theta)_L\) is the derivative of the horizontal slowness with respect to the initial polar angle \(\theta\) of the paraxial ray. If we compute \((p_1(\theta + \Delta \theta))_L\) for a range of \(\Delta \theta\) using (71), we can plot a line in data-model-space with slope equal to the derivative \((\partial p_1(\theta)/\partial \theta)_L\). For the same range of perturbations in \(\Delta \theta\) we can also plot a curve based on computing \((p_1(\theta + \Delta \theta))_L\) by ray tracing in the perturbed model \(m = \theta + \delta \theta\) and compute a derivative numerically. This numerical derivative can be used to check the derivative computed by paraxial ray tracing. We can also check the derivative computed by paraxial tracing by visual inspection of the plots - the line with slope computed
by paraxial ray tracing should coincide with the tangent at zero perturbation to the curve of perturbed data parameter versus perturbed model parameter.

The derivatives of the data are tested along the depth level L in the synthetic 2D TTI model shown in Figs. 6-7. In Fig. 9 we show the test of the derivatives of data along the depth level L with respect to the ray model parameters. In Fig. 10 we show the derivatives of the $P$-ray data with respect to the B-spline knot values at $((x_1)_n, (x_3)_l) = (4 \text{ km}, 2 \text{ km})$ for all the 2D TTI model fields. In Fig. 11 we show the test of derivatives of the $S$-ray data with respect to the B-spline knot values at $((x_1)_n, (x_3)_l) = (6 \text{ km}, 2 \text{ km})$ for all the 2D TTI model fields. In all cases of data and model parameters in Fig. 9-11, we can check visually that the line with slope from paraxial ray tracing at zero perturbation is indeed tangent to the curve of perturbed data versus perturbed model parameter.

5.3 Stereotomographic inversion

As a test of the validity of our code we do velocity estimation for each of the fields $V_{P0}$, $V_{S0}$, $\varepsilon$, $\delta$ and $\varphi$ separately. In each case, for the four fields not estimated, we use the fields of the synthetic model used for modelling the data. Stereotomography data are modelled by kinematic ray tracing using the velocity models in Figs. 6-7. The ray coverage of 406 $PP$ ray pairs shot towards the surface from $29 \times 14$ regularly spaced reflection/diffraction points is shown in Fig. 12. The same points are used to shoot 406 $PS$ events. The initial polar phase angles $\theta$ used for shooting the rays are the same for each test and in the range $[135^\circ, 157.5^\circ]$ (see Fig. 1).

The $PP$ data are used in the separate estimation of $V_{P0}$ and $\varepsilon$, while the $PS$ data are used in the separate estimation of $V_{S0}$, $\delta$ and $\varphi$. Figs. 13-17 show the results from separate estimation of $V_{P0}$, $V_{S0}$, $\varepsilon$, $\delta$ and $\varphi$. In each case, the estimated field fits the corresponding field of the synthetic model (Figs. 6-7) with only small errors.

The inversion algorithm takes into account possible breakdown of the linear assumption between data and model by adaptively doing a relocalization of rays in the updated model before new misfit calculation. In our tests, the use of $PP$ rays to estimate $V_{P0}$ and $\varepsilon$ did not require a relocalization of rays after update of velocity model, while for the estimation of $V_{S0}$, $\delta$ and $\varphi$ using $PS$ rays this was required in nearly every update step.
6 CONCLUSION

Using ray perturbation theory, we have derived all the needed derivatives of the stereotomography data with respect to the model parameters for 3D smooth anisotropic elastic media. Our results for the derivatives of the Hamiltonian and of the phase velocity are general, and the formulas are expected to be useful in other applications that depend on ray perturbation theory.

We have shown, using numerical examples in a 2D TTI medium, the validity of our equations for the stereotomography sensitivity matrix. In addition, we have shown on a 2D TTI synthetic data example without picking that our approach allows to converge towards the correct solution when estimating only one velocity parameter field. Because of lack of data and non-uniqueness, extra constraints and regularization are needed in order to invert for all velocity parameter fields, even for the simplest cases of anisotropic media, e.g. transversely isotropic media.

Compared to the stereotomography cost function introduced by Billette & Lambaré (1998), we assume no uncertainties on the depths of sources and receivers. This allows the size of the sensitivity matrix to be reduced. By introducing a rotation of the slowness vector, we also take into account that the slopes in the data are generally measured on a non-horizontal acquisition surface.

To add extra constraints to anisotropic stereotomography, we can follow the approach of Foss et al. (2005) and include a smoothed layered part in the velocity model description. Such an approach allows us to limit the number of anisotropic parameters to invert for. The depth of interfaces associated with the layers can be updated by map migration of zero-offset data or constrained by well log information. Co-depthing of imaged PP and PS key reflectors (Foss et al. 2005) can be included as a constraint for stereotomography to get a better estimate of $\varepsilon$ and $\delta$. The regularization of the anisotropic stereotomographic inversion can also be done by use of transmission events (Barbosa et al. 2008).

ACKNOWLEDGMENTS

We thank Eric Duveneck and Gilles Lambaré for helpful discussions and suggestions. We are also grateful to Eric Duveneck for implementing an isotropic stereotomography code on which the anisotropic stereotomography code has been based. Thanks to Robert Drysdale for reading through the manuscript. The authors acknowledge The Research Council of Norway for funding this work (grant 159117/I30).
REFERENCES


For Peer Review


28  S. Nag, M. Alerini and B. Ursin

List of symbols

(·)T (superscript) transpose
(·)(m) (superscript) wavemode \((m = 1: qS_1, \ m = 2: qS_2, \ m = 3: qP)\)
(·)W (subscribe) relative to the wavefront
(·)S (subscribe) relative to the source or receiver acquisition surface
(·)L (subscribe) relative to the source or receiver depth level
\(t\) time
\(\mathbf{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)^T\) position vector in local (crystal) Cartesian coordinate system
\(\mathbf{x} = (x_1, x_2, x_3)^T\) position vector in global (model) Cartesian coordinate system
\(\mathbf{p} = (p_1, p_2, p_3)^T\) slowness vector relative to global Cartesian coordinate system
\(\mathbf{x} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^T\) position vector in local acquisition surface Cartesian coordinate system
\(\tilde{\mathbf{p}} = (\tilde{p}_1, \tilde{p}_2, \tilde{p}_3)^T\) slowness vector relative to the \(\hat{x}_1\hat{x}_2\hat{x}_3\)-coordinate system
\(\mathbf{R}\) rotation matrix from \(x_1x_2x_3\)-coordinate system to \(\hat{x}_1\hat{x}_2\hat{x}_3\)-coordinate system
\(\mathbf{n} = (n_1, n_2, n_3)^T\) unit vector in direction of the slowness vector \(\mathbf{p}\)
\(\mathbf{y} = (\mathbf{x}^T, \mathbf{p}^T)^T\) phase space vector
\(\rho(\mathbf{x})\) density
\(c_{ijkl}(\mathbf{x})\) elastic stiffness tensor
\(a_{ijkl}(\mathbf{x}) = c_{ijkl}(\mathbf{x})/\rho(\mathbf{x})\) density-normalized elastic stiffness tensor
\(\Delta a_{ijkl}\) perturbation of the density-normalized elastic stiffness tensor
\(\tilde{C}_{\alpha\beta}(\mathbf{x})\) Voigt matrix in local (crystal) coordinate system
\(C_{\alpha\beta}(\mathbf{x})\) Voigt matrix in global (model) coordinate system
\(\tilde{A}_{\alpha\beta}(\mathbf{x}) = \tilde{C}_{\alpha\beta}(\mathbf{x})/\rho(\mathbf{x})\) density-normalized Voigt matrix in local (crystal) coordinate system
\(A_{\alpha\beta}(\mathbf{x}) = C_{\alpha\beta}(\mathbf{x})/\rho(\mathbf{x})\) density-normalized Voigt matrix in global (model) coordinate system
\(V_{P\mathbf{p}}(\mathbf{x})\) P-wave velocity in the direction of the TTI symmetry axis
\(V_{S\mathbf{p}}(\mathbf{x})\) S-wave velocity in the direction of the TTI symmetry axis
\(\varepsilon(\mathbf{x}), \delta(\mathbf{x}), \gamma(\mathbf{x})\) Thomsen anisotropy parameters
\(\varphi(\mathbf{x}), \xi(\mathbf{x})\) polar and azimuth angle of the TTI symmetry axis
\(v(\mathbf{x}) = 1/\sqrt{\mathbf{p}_i \mathbf{p}_i}\) phase velocity
\(\Gamma_{ik}(\mathbf{x}, \mathbf{p}) = a_{ijkl}(\mathbf{x})p_j p_l\) Christoffel matrix \(\Gamma\)
\(\Gamma'_{ik}(\mathbf{x}, \mathbf{p}) = a_{ijkl}(\mathbf{x})n_j n_l\) phase-normalized Christoffel matrix \(\Gamma'\)
\(\mathbf{g} = (g_1, g_2, g_3)^T\) polarization vector
\(\mathcal{H}(\mathbf{x}, \mathbf{p})\) ray Hamiltonian
\(\mathcal{G}^{(m)}(\mathbf{x}, \mathbf{p})\) eigenvalues of \(\Gamma\)
\(\mathcal{G}'^{(m)}(\mathbf{x}, \mathbf{p})\) eigenvalues of \(\Gamma'\)
\[ u^s(x) = 0 \] source acquisition surface

\[ u^r(x) = 0 \] receiver acquisition surface

\[ x^s_3, x^r_3 \] depth of source and receiver

\[ x^s = (x^s_1, x^s_2)^t \] source position on acquisition surface \( u^s \)

\[ x^r = (x^r_1, x^r_2)^t \] receiver position on acquisition surface \( u^r \)

\[ p^s = (p^s_1, p^s_2, p^s_3)^t \] slowness vector at source \( x^s \)

\[ p^r = (p^r_1, p^r_2, p^r_3)^t \] slowness vector at receiver \( x^r \)

\[ q^s = (\tilde{p}^s_1, \tilde{p}^s_2)^t \] slowness vector components tangential to the acquisition surface at source \( x^s \)

\[ q^r = (\tilde{p}^r_1, \tilde{p}^r_2)^t \] slowness vector components tangential to the acquisition surface at receiver \( x^r \)

\[ x_0 = (x_{01}, x_{02}, x_{03})^t \] reflection/diffraction point of stereotomographic event

\[ p_0 = (p_{01}, p_{02}, p_{03})^t \] slowness vector at \( x_0 \)

\[ n_0 = (n_{01}, n_{02}, n_{03})^t \] unit vector in direction of the slowness vector \( p_0 \)

\[ \Theta^s = (\theta^s, \psi^s)^t \] phase direction angles at reflection point \( x_0 \) for source ray

\[ \Theta^r = (\theta^r, \psi^r)^t \] phase direction angles at reflection point \( x_0 \) for receiver ray

\( T, \Delta T \) one-way traveltime, perturbation in one-way traveltime

\( T^s, T^r \) one-way traveltime from reflection point to source and to receiver, respectively

\( T^{sr} = T^s + T^r \) two-way traveltime

\( N_j \) number of picked events of mode \( j \)

\( N_{\text{obs}} = \sum_{j=1}^3 N_j \) number of observed (picked) events

\( m_b \) background velocity model

\( m_{\text{ray}} \) ray model

\( m = (m_{\text{ray}}^t, m_b^t)^t \) model vector

\( d(m) \) data vector

\( d_{\text{obs}} \) observed data

\( d_{\text{calc}}(m) \) calculated data

\( \Delta d = d_{\text{obs}} - d_{\text{calc}}(m) \) data misfit

\( \mathcal{J}(m) \) cost function

\( \epsilon \) regularization weight

\( L \) regularization operator
$\Delta m_{\text{ray}}$ perturbation of the ray model

$\Delta m_{b}$ perturbation of the background velocity model

$\Delta m = (\Delta m_{\text{ray}}, \Delta m_{b})^T$ perturbation of the model

$\Delta y$ perturbation of phase space vector at the end point of a ray

$\Delta y_0$ perturbation of phase space vector at the initial point of a ray

$\Delta \mathcal{H}(x(t), p(t))$ perturbation of the ray Hamiltonian

$f(x)$ model fields vector

$f_{\text{T}}(x)$ Voigt model fields vector

$f_{\text{T}}(x)$ Thomsen’s model fields vector

$B(x(t), p(t))$ source matrix for integration of pertubations in $f$ along a ray

$\Pi(t, t_0)$ ray propagator matrix from time $t_0$ to time $t$

$I$ identity matrix

$J = \frac{\partial d}{\partial m^T}$ sensitivity matrix

$J_x = \frac{\partial y}{\partial x_0}$ sensitivity matrix at ray end point w.r.t. initial position $x_0$

$J_\Theta = \frac{\partial y}{\partial \Theta_0}$ sensitivity matrix at ray end point w.r.t. initial phase angles $\Theta$

$J_f = \frac{\partial y}{\partial f}$ sensitivity matrix at ray end point w.r.t. field parameters $f$

$J_x^0 = \frac{\partial y_0}{\partial x_0}$ sensitivity matrix at ray initial point w.r.t. initial position $x_0$

$J_\Theta^0 = \frac{\partial y_0}{\partial \Theta_0}$ sensitivity matrix at ray initial point w.r.t. initial phase angles $\Theta$

$J_f^0 = \frac{\partial y_0}{\partial f}$ sensitivity matrix at ray initial point w.r.t. field parameters $f$

$v_{nml}$ cubic B-spline knot values for a 3D grid indexed by $(n, m, l)$

$\Delta v_{nml}$ perturbation of cubic B-spline knot values $v_{nml}$

$b_n(x_1), b_m(x_2), b_l(x_3)$ 1D cubic B-spline basis functions in the $x_1, x_2, x_3$-directions, respectively

$N_{x_1}, N_{x_2}, N_{x_3}$ number of cubic B-spline knot values in the $x_1, x_2, x_3$-directions, respectively

$n_S$ unit normal vector to the acquisition surface at source or receiver

$n_L = (0, 0, 1)^T$ unit normal vector to depth level of source or receiver

$M$ transformation matrix from TI to TTI medium

$\zeta = \hat{A}_{33} - \hat{A}_{44}$ variable used for computing the Jacobian $\partial f_{\text{V}} / \partial \hat{f}_{\text{T}}$ for a TI medium

$\chi = \hat{A}_{13} + \hat{A}_{44}$ variable used for computing the Jacobian $\partial f_{\text{V}} / \partial \hat{f}_{\text{T}}$ for a TI medium
Figure 1. The geometry of a stereotomographic event. Rays are shot from the reflection/diffraction point $x_0$ towards source and receiver with phase angles $\mathbf{\Theta}^s = (\theta^s, \psi^s)^t$ and $\mathbf{\Theta}^r = (\theta^r, \psi^r)^t$, respectively. The calculated slowness vectors $\mathbf{p}^s = (p^s_1, p^s_2, p^s_3)^t$ and $\mathbf{p}^r = (p^r_1, p^r_2, p^r_3)^t$ are rotated according to the normal of the acquisition surfaces $u^s(x) = 0$ and $u^r(x) = 0$, respectively. The tangential components of the slowness relative to the acquisition surfaces are collected in the vectors $\mathbf{q}^s = (\tilde{p}^s_1, \tilde{p}^s_2)^t$ and $\mathbf{q}^r = (\tilde{p}^r_1, \tilde{p}^r_2)^t$.

Figure 2. The rotation from slowness vector components $p_i$ in the global $x_1x_2x_3$-coordinate system to slowness vector components $\tilde{p}_i$ in the local $\tilde{x}_1\tilde{x}_2\tilde{x}_3$-coordinate system tangential to the acquisition surface $u(x) = 0$.

Figure 3. Phase slowness vector $\mathbf{p}$ and the unit vector $\mathbf{n}$ in the direction of $\mathbf{p}$. The two angles $\theta$ and $\psi$ are the azimuthal and polar phase angles.

Figure 4. Symmetry axis for a tilted transverse isotropic medium (TTI). The two angles, $\varphi$ and $\xi$ are the polar and azimuthal angles made by the symmetry axis $\hat{x}_3$ with the $x_3$ and $x_1$ axis, respectively.

Figure 5. The perturbations in the one-way traveltime $T$ resulting from restricting the paraxial ray to the depth level $L$ when perturbing the initial polar angle $\theta$, the initial point $x_0$ and the medium parameters $a_{ijkl}$.

Figure 6. The synthetic fields for $V_{P0}$ (upper left), $V_{S0}$ (upper right), $\varepsilon$ (lower left) and $\delta$ (lower right) of the smooth 2D TTI medium.

Figure 7. The synthetic field for the tilt angle $\varphi$ of the smooth 2D TTI medium.

Figure 8. Modelling of a single PS ray pair used for testing derivatives of the data, here shown in the $V_{P0}$ field. The left ray towards the source $x^s = (x^s_1)$ is here a $P$-wave and the right ray towards the receiver $x^r = (x^r_1)$ is an $SV$-wave.
Figure 9. Test of the derivatives $\frac{\partial d}{\partial m_{ray}} = \frac{\partial (x_1, x_1', p_1', T_{sur})}{\partial (x_1, x_3, \theta_s, \theta_r)}$ along the depth level $L$ for the PS event in Fig. 8. Note that the perturbation of the ray model parameters range from -12 % to +12 % of their values.

Figure 10. Test of the derivatives $\frac{\partial (x_1, p_1', T_{sur})}{\partial (V_{P0_{nm}}, V_{S0_{nm}}, \varepsilon_{nm}, \delta_{nm}, \phi_{nm})}$ along the depth level $L$ for the left $P$-mode ray in Fig. 8. The knot location is at $(x_1)_n, (x_3)_m = (4 \text{ km}, 2 \text{ km})$. Note that the perturbation of the B-spline knot values range from -12 % to +12 % of their values.

Figure 11. Test of the derivatives $\frac{\partial (x_1', p_1', T_{sur})}{\partial (V_{P0_{nm}}, V_{S0_{nm}}, \varepsilon_{nm}, \delta_{nm}, \phi_{nm})}$ along the depth level $L$ for the right $SV$-mode ray in Fig. 8. The knot location is at $(x_1)_n, (x_3)_m = (6 \text{ km}, 2 \text{ km})$. Note that the perturbation of the B-spline knot values range from -12 % to +12 % of their values.

Figure 12. Modelling of 406 $PP$ events shown in the $V_{P0}$ model.

Figure 13. Estimation of $V_{P0}$: exact model (upper left), initial model (upper right), final output model (lower left), and difference between exact and final models (lower right). The fixed $V_{S0}, \varepsilon, \delta$ and $\phi$ models were the exact models from Figs. 6-7. Only the central part of the model, where the ray coverage is good enough, is displayed.

Figure 14. Estimation of $V_{S0}$: exact model (upper left), initial model (upper right), final output model (lower left), and difference between exact and final models (lower right). The fixed $V_{P0}, \varepsilon, \delta$ and $\phi$ models were the exact models in Figs. 6-7. Only the central part of the model, where the ray coverage is good enough, is displayed.

Figure 15. Estimation of $\varepsilon$: exact model (upper left), initial model (upper right), final output model (lower left), and difference between exact and final models (lower right). Initially $\varepsilon = 0$. The fixed $V_{P0}, V_{S0}, \delta$ and $\phi$ models were the exact models in Figs. 6-7. Only the central part of the model, where the ray coverage is good enough, is displayed.
Figure 16. Estimation of $\delta$: exact model (upper left), initial model (upper right), final output model (lower left), and difference between exact and final models (lower right). Initially $\delta = 0$. The fixed $V_{P0}$, $V_{S0}$, $\varepsilon$ and $\varphi$ models were the exact models in Figs. 6-7. Only the central part of the model, where the ray coverage is good enough, is displayed.

Figure 17. Estimation of $\varphi$: exact model (upper left), initial model (upper right), final output model (lower left), and difference between exact and final models (lower right). Initially $\varphi = 0$. The fixed $V_{P0}$, $V_{S0}$, $\varepsilon$ and $\delta$ models were the exact models in Fig. 6. Only the central part of the model, where the ray coverage is good enough, is displayed.
Figure 1
Figure 2
\[ \theta \quad \psi \quad x_2 \quad x_3 \]

Figure 3
Figure 4
\[
T(x) = t
\]

\[
T(x + \Delta x) = t
\]

\[
\Delta T = p \cdot \Delta x
\]

\[
\Delta T_0 = \int_{t_0}^{t_1} \Delta H(\Delta a_{i\mu}) \, dt
\]

\[
\Delta T_0 = n \cdot \Delta x
\]

\[
\Delta T_0 = -n \cdot \Delta x
\]

**Figure 5**
Figure 6
Figure 7
Figure 8
Figure 9
Figure 10
Figure 11
Figure 12
Figure 13
Figure 15
Figure 16
Figure 17
APPENDIX A: RELATION BETWEEN PICKED SLOPES AND THE RAY SLOWNESS VECTOR

In the high frequency approximation of ray theory (Červený 2001), the picked slopes \(q_{\text{obs}}^s = (\tilde{p}_{1s}^s, \tilde{p}_{2s}^s)^t\) and \(q_{\text{obs}}^r = (\tilde{p}_{1r}^r, \tilde{p}_{2r}^r)^t\) in the data correspond to the tangential components of the ray slowness vector along the acquisition surface (see Fig.2) at source and receiver, respectively. The slopes are picked for the same two-way traveltime \(T_{\text{sr}}^{\text{obs}}\). For picking in common-shot and common-receiver gathers, the slope components can be defined by

\[
\tilde{p}_{is}^s = \frac{\partial T_{\text{sr}}^{\text{obs}}}{\partial \tilde{x}_{is}^s}, \quad i = 1, 2; \\
\tilde{p}_{ir}^r = \frac{\partial T_{\text{sr}}^{\text{obs}}}{\partial \tilde{x}_{ir}^r}, \quad i = 1, 2; 
\]

where \(\tilde{x}_{is}^s\) and \(\tilde{x}_{ir}^r\) are coordinates of a local Cartesian coordinate system centered on the acquisition surfaces \(u^s(x) = 0\) and \(u^r(x) = 0\) (see Fig. 2).

The calculated slowness vectors \(p_s\) and \(p_r\) at source and receiver are defined Červený (2001) as

\[
p_{is}^s = \frac{\partial T^s}{\partial x_{is}}^t, \quad i = 1, 2, 3; \\
p_{ir}^r = \frac{\partial T^r}{\partial x_{ir}}^t, \quad i = 1, 2, 3. 
\]

The slowness vector \(\hat{p} = (\hat{p}_1, \hat{p}_2, \hat{p}_3)^t\) measured in the local \(\tilde{x}_1\tilde{x}_2\tilde{x}_3\)-coordinate system can be related to the same slowness vector \(p = (p_1, p_2, p_3)^t\) measured in the global \(x_1x_2x_3\)-coordinate system by a transformation. As both coordinate systems are orthonormal and the translation of origo is not needed for transformation of a vector, this transformation becomes the rotation

\[
\hat{p} = Rp, 
\]

where the rotation matrix \(R\) can be defined using the normal vector \(n_s\) to the acquisition surface \(u(x) = 0\). This normal vector is given by

\[
n_s = \frac{\nabla u(x)}{|\nabla u(x)|}. 
\]

APPENDIX B: JACOBIAN FROM TI VOIGT PARAMETERS TO TI THOMSEN PARAMETERS

Here we compute the Jacobian \(\frac{\partial f_{V}^{\text{TI}}}{\partial (f_{T}^{\text{TI}})^t}\), where \(f_{V}^{\text{TI}} = \left(\hat{A}_{11}, \hat{A}_{33}, \hat{A}_{44}, \hat{A}_{66}, \hat{A}_{13}\right)^t\) are TI Voigt parameters from (5), and \(f_{T}^{\text{TI}} = (V_P, V_S, \varepsilon, \delta, \gamma)^t\) are the Thomsen parameters defined in (8). From these definitions we can express the TI Voigt parameters \(f_{V}^{\text{TI}}\) as functions of the TI Thomsen parameters
From (B1) and (B3) it is straightforward to take derivatives

\[ \frac{\partial}{\partial f_T} \left( \begin{array}{c} \hat{A}_{11} = V_{P_0}^2 (1 + 2 \varepsilon) \\ \hat{A}_{33} = V_{P_0}^2 \\ \hat{A}_{44} = V_{S_0}^2 \\ \hat{A}_{66} = V_{S_0}^2 (1 + 2 \gamma) \\ \hat{A}_{13} = \sqrt{(V_{P_0}^2 - V_{S_0}^2)(V_{P_0}^2 (1 + 2 \delta) - V_{S_0}^2) - V_{S_0}^2} \end{array} \right), \quad (B1) \]

where we consider only the positive sign in the front of the square root in the expression of \( \hat{A}_{13} \). The negative sign of the root is needed for the rare cases (Helbig & Schoenberg 1987) that \( \hat{A}_{13} + \hat{A}_{33} < 0 \). However, as discussed by Tsvankin (1996), for seismic modeling and processing we can assume that \( \hat{A}_{13} + \hat{A}_{33} > 0 \).

To simplify the following development of the derivatives, first define

\[ \begin{align*}
\zeta &= V_{P_0}^2 - V_{S_0}^2 \\
\chi &= \sqrt{\zeta (2 \delta V_{P_0}^2 + \zeta)}
\end{align*} \]

so that

\[ \hat{A}_{13} = \chi - \hat{A}_{44}. \quad (B3) \]

From (B1) and (B3) it is straightforward to take derivatives

\[ \begin{align*}
\frac{\partial \hat{A}_{11}}{\partial f_T} &= 2 V_{P_0} \frac{\partial V_{P_0}^2}{\partial f_T} (1 + 2 \varepsilon) + 2 V_{P_0}^2 \frac{\partial \delta}{\partial f_T} \\
\frac{\partial \hat{A}_{33}}{\partial f_T} &= 2 V_{P_0} \frac{\partial V_{P_0}^2}{\partial f_T} \\
\frac{\partial \hat{A}_{44}}{\partial f_T} &= 2 V_{S_0} \frac{\partial V_{S_0}^2}{\partial f_T} \\
\frac{\partial \hat{A}_{66}}{\partial f_T} &= 2 V_{S_0} \frac{\partial V_{S_0}^2}{\partial f_T} (1 + 2 \gamma) + 2 V_{S_0}^2 \frac{\partial \gamma}{\partial f_T} \\
\frac{\partial \hat{A}_{13}}{\partial f_T} &= \frac{\partial \chi}{\partial f_T} - \frac{\partial \hat{A}_{44}}{\partial f_T}
\end{align*} \]

The only term which requires some development to compute is \( \frac{\partial \chi}{\partial f_T} \). From the definition (B2) it follows directly that

\[ \frac{\partial \chi}{\partial f_T} = \frac{1}{\chi} \left[ (\delta V_{P_0}^2 + \zeta) \frac{\partial \zeta}{\partial f_T} + \zeta V_{P_0} \frac{\partial \delta}{\partial f_T} + 2 \zeta \delta V_{P_0} \frac{\partial V_{P_0}}{\partial f_T} \right], \quad (B5) \]

where

\[ \frac{\partial \zeta}{\partial f_T} = \frac{\partial \hat{A}_{33}}{\partial f_T} - \frac{\partial \hat{A}_{44}}{\partial f_T}, \quad (B6) \]

which is non-zero only for \( f_T = V_{P_0} \) or \( f_T = V_{S_0} \). Expanding and simplifying the derivatives of (B4) gives for the non-zero components of the Jacobian \( \frac{\partial f_T^\text{III}}{\partial (f_T^\text{II})^t} \)

\[ \begin{align*}
\frac{\partial \hat{A}_{11}}{\partial f_T} &= 2 V_{P_0} \left( 1 + 2 \varepsilon \right) \\
\frac{\partial \hat{A}_{33}}{\partial f_T} &= 2 V_{P_0}^2 \\
\frac{\partial \hat{A}_{13}}{\partial f_T} &= \frac{2 V_{P_0}}{\chi} \left[ \delta V_{P_0}^2 + \zeta (1 + \delta) \right]
\end{align*} \]

(B7)
\[ \begin{align*}
\frac{\partial \hat{A}_{44}}{\partial V_{S0}} &= 2V_{S0}, \\
\frac{\partial \hat{A}_{66}}{\partial V_{S0}} &= 2V_{S0}(1 + 2\gamma), \\
\frac{\partial \hat{A}_{13}}{\partial V_{S0}} &= -2V_{S0} \left[ \frac{1}{\lambda} (\delta V_{P0} + \zeta) + 1 \right] \\
\frac{\partial \hat{A}_{11}}{\partial \epsilon} &= 2V_{P0}^2, \\
\frac{\partial \hat{A}_{66}}{\partial \gamma} &= 2V_{S0}^2, \\
\frac{\partial \hat{A}_{13}}{\partial \delta} &= \frac{\zeta}{\chi} V_{P0}^2.
\end{align*} \] (B8)

Using the above results, it is straightforward to derive the spatial derivatives \( \nabla_x \left[ \frac{\partial f_{\text{TI}}}{\partial (f_{\text{TI}})^T} \right] \) of the Jacobian from TI Voigt to TI Thomsen notation.

**APPENDIX C: BOND TRANSFORMATION FOR A 2D TTI MEDIUM**

The Bond transformation has been defined for the general 3D case in (6). For the 2D case it is more convenient to work with the explicit relations (Jech 1983) between the TTI Voigt parameters and the TI Voigt parameters. In 2D, we only need to include the polar angle \( \varphi \) for the rotation of the symmetry axis with respect to the vertical axis. In the following we will include only the Voigt parameters \( f_{\text{V}} \) and \( f_{\text{VTTI}} \) of (5) and (7) that are relevant for the propagation of \( P \) and \( SV \) waves in 2D TTI media.

The Bond transformation can then be written as a matrix-vector multiplication

\[ f_{\text{VTTI}} = M f_{\text{V}}, \] (C1)

where

\[ f_{\text{VTTI}} = (A_{11}, A_{33}, A_{44}, A_{13}, A_{15}, A_{35}), \] (C2)

and

\[ f_{\text{V}} = (\hat{A}_{11}, \hat{A}_{33}, \hat{A}_{44}, \hat{A}_{13}). \] (C3)

The \( 6 \times 4 \) matrix \( M \) can be found from the results in Jech (1983) and can be written as

\[ M = \begin{pmatrix} M_1 & M_2 \end{pmatrix}, \] (C4)

where

\[ M_1 = \begin{pmatrix}
    \cos^4 \varphi & \sin^4 \varphi \\
    \sin^4 \varphi & \cos^4 \varphi \\
    \cos^2 \varphi \sin^2 \varphi & \cos^2 \varphi \sin^2 \varphi \\
    \cos^2 \varphi \sin^2 \varphi & \cos^2 \varphi \sin^2 \varphi \\
    -\cos^3 \varphi \sin \varphi & \sin^3 \varphi \cos \varphi \\
    -\sin^3 \varphi \cos \varphi & \cos^3 \varphi \sin \varphi
\end{pmatrix}, \] (C5a)

and

\[ M_2 = \begin{pmatrix}
    \cos^2 \varphi \sin^2 \varphi & \cos^2 \varphi \sin^2 \varphi \\
    \cos^2 \varphi \sin^2 \varphi & \cos^2 \varphi \sin^2 \varphi \\
    -\cos^3 \varphi \sin \varphi & \sin^3 \varphi \cos \varphi \\
    -\sin^3 \varphi \cos \varphi & \cos^3 \varphi \sin \varphi
\end{pmatrix}. \] (C5b)
\[ M_2 = \begin{pmatrix} 4 \cos^2 \varphi \sin^2 \varphi & 2 \cos^2 \varphi \sin^2 \varphi \\ 4 \sin^2 \varphi \cos^2 \varphi & 2 \sin^2 \varphi \cos^2 \varphi \\ \frac{1}{2} (\cos 4\varphi + 1) & -2 \cos^2 \varphi \sin^2 \varphi \\ -4 \cos^2 \varphi \sin^2 \varphi & \frac{1}{4} (\cos 4\varphi + 3) \\ \frac{1}{2} \sin 4\varphi & \frac{1}{4} \sin 4\varphi \\ -\frac{1}{2} \sin 4\varphi & -\frac{1}{4} \sin 4\varphi \end{pmatrix} \]  
(C5b)

From (C5) it follows directly that

\[ \frac{\partial M_1}{\partial \varphi} = \begin{pmatrix} -4 \cos^3 \varphi \sin \varphi & 4 \sin^3 \varphi \cos \varphi \\ 4 \sin^3 \varphi \cos \varphi & -4 \cos^3 \varphi \sin \varphi \\ \frac{1}{2} \sin 4\varphi & \frac{1}{2} \sin 4\varphi \\ \frac{1}{2} \sin 4\varphi & \frac{1}{2} \sin 4\varphi \\ -\frac{1}{2} (\cos 2\varphi + \cos 4\varphi) & \frac{1}{2} (\cos 2\varphi - \cos 4\varphi) \\ -\frac{1}{2} (\cos 2\varphi - \cos 4\varphi) & \frac{1}{2} (\cos 2\varphi + \cos 4\varphi) \end{pmatrix} \]  
(C6a)

and

\[ \frac{\partial M_2}{\partial \varphi} = \begin{pmatrix} 2 \sin 4\varphi & \sin 4\varphi \\ 2 \sin 4\varphi & \sin 4\varphi \\ -2 \sin 4\varphi & -\sin 4\varphi \\ -2 \sin 4\varphi & -\sin 4\varphi \\ 2 \cos 4\varphi & \cos 4\varphi \\ -2 \cos 4\varphi & -\cos 4\varphi \end{pmatrix} \]  
(C6b)

From (C6) we also get

\[ \frac{\partial^2 M_1}{\partial \varphi^2} = \begin{pmatrix} -2 (\cos 2\varphi + \cos 4\varphi) & 2 (\cos 2\varphi - \cos 4\varphi) \\ 2 (\cos 2\varphi - \cos 4\varphi) & -2 (\cos 2\varphi + \cos 4\varphi) \\ 2 \cos 4\varphi & 2 \cos 4\varphi \\ 2 \cos 4\varphi & 2 \cos 4\varphi \\ \sin 2\varphi + 2 \sin 4\varphi & - (\sin 2\varphi - 2 \sin 4\varphi) \\ \sin 2\varphi - 2 \sin 4\varphi & - (\sin 2\varphi + 2 \sin 4\varphi) \end{pmatrix} \]  
(C7a)
and

$$\frac{\partial^2 \mathbf{M}_2}{\partial \varphi^2} = \begin{pmatrix} 8 \cos 4\varphi & 4 \cos 4\varphi \\ 8 \cos 4\varphi & 4 \cos 4\varphi \\ -8 \cos 4\varphi & -4 \cos 4\varphi \\ -8 \cos 4\varphi & -4 \cos 4\varphi \\ -8 \sin 4\varphi & -4 \sin 4\varphi \\ 8 \sin 4\varphi & 4 \sin 4\varphi \end{pmatrix}. \quad (C7b)$$

Ray tracing requires the calculation of smooth second order spatial derivatives of the medium parameters. If the model is described using Thomsen parameters, we need a Bond transformation for each evaluation of the 2D TTI Voigt matrix and its spatial derivatives. From the definition (C1) of the Bond transformation we have

$$\frac{\partial f_{TTI}^T}{\partial x_i} = \frac{\partial \mathbf{M}}{\partial \varphi} \frac{\partial \varphi}{\partial x_i} f_{TTI}^T + \mathbf{M} \frac{\partial^2 \varphi}{\partial x_i \partial x_i} f_{TTI}^T, \quad i = 1, 2, 3. \quad (C8)$$

The second order derivatives of the TTI Voigt matrix are given as

$$\frac{\partial^2 f_{TTI}^T}{\partial x_i \partial x_j} = \frac{\partial^2 \mathbf{M}}{\partial \varphi^2} \frac{\partial \varphi}{\partial x_i} \frac{\partial \varphi}{\partial x_j} f_{TTI}^T + \frac{\partial \mathbf{M}}{\partial \varphi} \frac{\partial^2 \varphi}{\partial x_i \partial x_j} f_{TTI}^T + \mathbf{M} \frac{\partial \varphi}{\partial x_i} \frac{\partial^2 \varphi}{\partial x_j \partial \varphi} f_{TTI}^T + \mathbf{M} \frac{\partial \varphi}{\partial x_j} \frac{\partial^2 \varphi}{\partial x_i \partial \varphi} f_{TTI}^T + \mathbf{M} \frac{\partial^2 \varphi}{\partial x_i \partial x_j} f_{TTI}^T, \quad i, j = 1, 2, 3. \quad (C9)$$

A way of avoiding these computations for each step in the ray tracing is to do the ray tracing in 2D TTI Voigt parameter B-spline fields instead of 2D TTI Thomsen parameter B-spline fields. The update of the model can still be made using the Thomsen parameter fields. The relations (B1), between the VTI Voigt parameters and the VTI Thomsen parameters, and the rotation (C1) can be used to update the Voigt parameter fields after each update of the Thomsen parameter fields.

**APPENDIX D: THE JACOBIAN FROM VOIGT TO THOMSEN FOR A 2D TTI MEDIUM**

Since according to our definitions $f_{TTI}^T = (f_{TTI}^T)^t$, we may write the Jacobian from TTI Voigt parameters to TTI Thomsen parameters as

$$\frac{\partial f_{TTI}^T}{\partial (f_{TTI}^T)^t} = \left( \sum_{i' \in T'} \frac{\partial f_{TTI}^T}{\partial (f_{TTI}^T)^t} \frac{\partial f_{TTI}^T}{\partial \varphi} \right). \quad (D1)$$
From the definition (C1) of the Bond transformation from a 2D TI medium to a 2D TTI medium it follows that

\[
\frac{\partial f_{\text{TTI}}^{V}}{\partial (f_{\text{TI}}^{V})^t} = \frac{\partial}{\partial \left(A_{11}, A_{33}, A_{55}, A_{13}, A_{15}, A_{35}\right)} = M
\]

(D2)

and

\[
\frac{\partial f_{\text{TTI}}^{V}}{\partial \varphi} = \frac{\partial}{\partial \varphi \left(A_{11}, A_{33}, A_{55}, A_{13}, A_{15}, A_{35}\right)} = \frac{\partial M}{\partial \varphi f_{\text{TTI}}^{V}}.
\]

(D3)

to be used in (D1). The spatial derivatives of the two parts of the Jacobian in (D1) are

\[
\nabla_x \left[ \frac{\partial f_{\text{TTI}}^{V}}{\partial (f_{\text{TI}}^{V})^t} \right] = \nabla_x [M] \frac{\partial f_{\text{TTI}}^{V}}{\partial (f_{\text{TI}}^{V})^t} + M \nabla_x \left[ \frac{\partial f_{\text{TTI}}^{V}}{\partial (f_{\text{TI}}^{V})^t} \right]
\]

(D4)

and

\[
\nabla_x \left[ \frac{\partial f_{\text{TTI}}^{V}}{\partial \varphi} \right] = \nabla_x \left[ \frac{\partial M}{\partial \varphi} \right] f_{\text{TTI}}^{V} + \frac{\partial M}{\partial \varphi} \nabla_x [f_{\text{TTI}}^{V}],
\]

(D5)

where

\[
\nabla_x [M] = \frac{\partial M}{\partial \varphi} \nabla_x \varphi, \quad \nabla_x \left[ \frac{\partial M}{\partial \varphi} \right] = \frac{\partial^2 M}{\partial \varphi^2} \nabla_x \varphi.
\]

(D6)

APPENDIX E: DERIVATIVES OF THE PHASE VELOCITY

We want to compute the derivatives of the phase velocity with respect to phase angles \( \Theta = (\theta, \psi)^t \) and the medium parameters \( f \). It is possible, in the case of special medium symmetries, like TI, to derive approximations (Thomsen 1986; Dellinger et al. 1993; Stopin 2001) of the general phase velocity equation

\[
v^2 = \Gamma_{ik} g_i g_k = a_{ijkl} n_j n_i g_j g_k.
\]

(E1)

Such approximations can be used to compute derivatives of the phase velocity. However, not only will such an approach be limited to the TI case, for example, but those computations will be very difficult for more general types of anisotropic media. It is better to develop derivatives of the phase velocity valid in the general anisotropic case. Consider that the eigenvalues \( G^{(m)} \) of the Christoffel equation are homogeneous functions of the second degree in \( p_j \) in the same way as \( \Gamma_{ik} \) (Červený 2001). This means that

\[
G^{(m)}(bp_j) = b^2 G^{(m)}(p_j),
\]

(E2)

where \( b \) is a scalar and where

\[
\Gamma_{ik}(bp_j) = b^2 \Gamma_{ik}(p_j).
\]

(E3)

For instance, if we choose \( b = v \), the phase velocity, we obtain

\[
G^{(m)}(vp_j) = G^{(m)}(n_j) = v^2 G^{(m)}(p_j),
\]

(E4)
where \( n_j \) is the \( j^{th} \)-component of the unit phase direction vector \( \mathbf{n} \). The Christoffel equation Červený (2001)

\[
\left( a_{ijkl} p_j p_l - G'(m) \delta_{ik} \right) g_i^{(m)} = 0,
\]

(E5)
can then be rewritten as

\[
\left( a_{ijkl} n_j n_l - v^2 G'(m) \delta_{ik} \right) g_i^{(m)} = 0.
\]

(E6)

Along the ray, \( G'(m) = 1 \) and we can write (E6) as

\[
\left( \Gamma'_{ik} - G'(m) \delta_{ik} \right) g_i^{(m)} = 0,
\]

(E7)

with

\[
\Gamma'_{ik} = a_{ijkl} n_j n_l \quad \text{and} \quad G'(m) = v^2.
\]

(E8)

Solving (E7) gives the phase velocity. Also, to obtain the derivative of the phase velocity, we are going to differentiate (E7) with respect to any parameter \( b \)

\[
\frac{\partial}{\partial b} \left[ \left( \Gamma'_{ik} - G'(m) \delta_{ik} \right) g_i^{(m)} \right] = 0,
\]

(E9)

\[
\left( \frac{\partial \Gamma'_{ik}}{\partial b} - \frac{\partial G'(m)}{\partial b} \delta_{ik} \right) g_i^{(m)} + \left( \Gamma'_{ik} - G'(m) \delta_{ik} \right) \frac{\partial g_i^{(m)}}{\partial b} = 0.
\]

(E10)

Multiplying each of the equations indexed by \( k \) in (E10) by \( g_k^{(m)} \) and summing for each \( k \) gives

\[
g_k^{(m)} \left( \frac{\partial \Gamma'_{ik}}{\partial b} - \frac{\partial G'(m)}{\partial b} \delta_{ik} \right) g_i^{(m)} + \frac{\partial \Gamma'_{ik}}{\partial b} g_i^{(m)} g_k^{(m)} + \frac{\partial g_i^{(m)}}{\partial b} \frac{\partial g_i^{(m)}}{\partial b} = 0.
\]

(E11)

Due to the symmetry properties of \( \Gamma'_{ik} \), it appears that the second term in (E11) is the Christoffel equation (E7) times \( \frac{\partial g_i^{(m)}}{\partial b} \). This term is therefore zero and

\[
\frac{\partial \Gamma'_{ik}}{\partial b} g_i^{(m)} g_k^{(m)} = 0,
\]

(E12)

\[
\frac{\partial G'(m)}{\partial b} g_i^{(m)} g_k^{(m)} = \frac{\partial \Gamma'_{ik}}{\partial b} g_i^{(m)} g_k^{(m)},
\]

(E13)

\[
\frac{\partial g_i^{(m)}}{\partial b} = \frac{\partial g_i^{(m)}}{\partial b} g_k^{(m)} g_k^{(m)},
\]

(E14)

as \( g_i^{(m)} g_i^{(m)} = 1 \). Replacing the eigenvalue \( G'(m) \) by the squared phase velocity gives

\[
\frac{\partial \nu^2}{\partial b} = \frac{\partial \Gamma'_{ik}}{\partial b} g_i^{(m)} g_k^{(m)},
\]

(E15)

\[
\frac{\partial \nu}{\partial b} = \frac{1}{2v} \frac{\partial \Gamma'_{ik}}{\partial b} g_i^{(m)} g_k^{(m)},
\]

(E16)

The general result (E16) is the partial derivative of the phase velocity with respect to any parameter \( b \).
With \( b = f \), a velocity model field parameter, we can rewrite (E16) as

\[
\frac{\partial v}{\partial f} = \frac{1}{2v} \frac{\partial a_{ijkl} n_j n_l}{\partial f} \bigg|_x g^{(m)}_i g^{(m)}_k
\]  
(E17)

The unit slowness vector \( n_l \) does not depend on the field parameters, so that by replacing \( n_j n_l \) by \( v^2 p_j p_l \) we have

\[
\frac{\partial v}{\partial f} = \frac{v^2}{2} \frac{\partial a_{ijkl}}{\partial f} \bigg|_x p_j p_l g^{(m)}_i g^{(m)}_k, \tag{E18}
\]
or equivalently

\[
\frac{\partial v}{\partial f} = \frac{v^2}{2} \frac{\partial \Gamma_{ik}}{\partial f} \bigg|_x p_j g^{(m)}_i g^{(m)}_k. \tag{E19}
\]

The final result (E19) is very useful, as we only need to work explicitly with the derivatives of the Christoffel matrix for fixed \( x \) and \( p \). The derivatives of the Christoffel matrix with respect to the Voigt parameters are trivial to compute as soon as the elements of the matrix have been defined by application of the Voigt mapping (4).

**APPENDIX F: DERIVATIVES OF THE POLARIZATION VECTORS**

By differentiating (E1) with respect to any parameter \( b \) and comparing the result with (E15), it also follows that

\[
\Gamma'_{ik} \frac{\partial}{\partial b} \left( g^{(m)}_i g^{(m)}_k \right) = 0, \tag{F1}
\]
and since \( \Gamma' = v^2 \Gamma_{ik} \), we have also have that

\[
\Gamma_{ik} \frac{\partial}{\partial b} \left( g^{(m)}_i g^{(m)}_k \right) = 0. \tag{F2}
\]

The derivatives with respect to position and slowness of the polarization vectors \( g^{(m)}_i \) needed by paraxial ray tracing are derived in Červený (2001). We include them here for completeness. If \( b \) is any parameter, e.g. \( x_l \) or \( p_l \), then from (Červený 2001, p.402) we have

\[
\frac{\partial g^{(1)}_j}{\partial b} = \left( \frac{1}{G^{(1)} - G^{(2)}} \frac{\partial \Gamma_{ik}}{\partial b} g^{(1)}_k g^{(1)}_i \right) g^{(2)}_j + \left( \frac{1}{G^{(1)} - G^{(3)}} \frac{\partial \Gamma_{ik}}{\partial b} g^{(1)}_k g^{(3)}_i \right) g^{(3)}_j, \tag{F3a}
\]

\[
\frac{\partial g^{(2)}_j}{\partial b} = \left( \frac{1}{G^{(2)} - G^{(1)}} \frac{\partial \Gamma_{ik}}{\partial b} g^{(2)}_k g^{(1)}_i \right) g^{(1)}_j + \left( \frac{1}{G^{(2)} - G^{(3)}} \frac{\partial \Gamma_{ik}}{\partial b} g^{(2)}_k g^{(3)}_i \right) g^{(3)}_j. \tag{F3b}
\]
\[
\frac{\partial g_j^{(3)}}{\partial b} = \left( \frac{1}{G^{(3)} - G^{(1)}} \frac{\partial \Gamma_{ik}}{\partial b} g_k^{(3)} g_i^{(1)} \right) g_j^{(1)} + \left( \frac{1}{G^{(3)} - G^{(2)}} \frac{\partial \Gamma_{ik}}{\partial b} g_k^{(3)} g_i^{(2)} \right) g_j^{(2)},
\]

where \( G^{(m)} (m = 1, 2, 3) \) are the eigenvalues of the Christoffel matrix \( \Gamma \) (17). Note that \( G^{(1)} \) and \( G^{(2)} \), the eigenvalues of the \( qS_1 \)- and \( qS_2 \)-rays, may have nearly the same value in weak anisotropic media. In such cases, the computation of the derivatives of the polarization vectors requires the use of quasi-isotropic ray theory (Bakker 1998; Pšenčík 1998).

**APPENDIX G: INITIALIZATION OF STEREOTOMOGRAPHY**

In the initialization step of stereotomography we need to define an initial velocity model and an initial ray model. One approach is to define an initial ray model by shooting rays from source and receivers into the subsurface using the picked slope information of events and an initial velocity model (Baina, personal communication (2002)). Given an initial background velocity model \( m_b \), the picked slopes \( q^s = (\tilde{p}^s_1, \tilde{p}^s_2)^T \) and \( q^r = (\tilde{p}^r_1, \tilde{p}^r_2)^T \) can be input into the Christoffel equation (16) to solve for \( \tilde{p}^s_3 \) and \( \tilde{p}^r_3 \), respectively. Before solving the Christoffel equation, the Voigt matrix has to rotated by the same rotation angles as for the vector \( \tilde{p} \) of (A3), in order to obtain the corresponding \( \tilde{p}^s_3 \) and \( \tilde{p}^r_3 \) satisfying the eikonal equation Červený (2001)

\[
\tilde{p}^2_1 + \tilde{p}^2_2 + \tilde{p}^2_3 = \frac{1}{v^2}.
\]

Note that the obtained \( \tilde{p}^s_3 \) and \( \tilde{p}^r_3 \) will not be observed data, as the initial velocity model parameters are used in solving the Christoffel equation. The slowness vector \( \tilde{p} \) in the local acquisition surface Cartesian coordinate system can be then be transformed into the slowness vector \( p \) in the glocal Cartesian coordinate system using

\[
p = R^{-1} \tilde{p}
\]

where \( R \) is the rotation matrix from Appendix A. The resulting slowness vector \( p \) from (G2) can be used to shoot rays from source and receiver into the subsurface using the initial background velocity model. The one-way traveltimes \( T^s \) and \( T^r \) of the rays are not known, only the two-ray traveltime \( T^{sr} \). In addition, the initial background velocity model is not correct. The ray pairs will therefore in general not meet in a common reflection/diffraction point. However, we can stop the rays after half of the two-way traveltime has been spent and define our initial reflection/diffraction point \( x_0 \) for the ray model using the mean value of the two ray end points. Shooting back towards the acquisition surface
from this initial \( x_0 \), with the opposite slowness direction as the downgoing rays, gives an initial misfit in the stereotomography data.

Assuming a horizontal acquisition surface and a 2D TI medium, the Christoffel equation (16) gives a second order polynomial equation in the variable \((p_3)^2\). This equation can be solved analytically to give \( p_3 \) as a function of \( p_1 \) and of the Voigt parameters \( \hat{A}_{11}, \hat{A}_{13}, \hat{A}_{33} \) and \( \hat{A}_{44} \). In detail, we first have the Christoffel equation (16) for 2D media given by

\[
\begin{pmatrix}
\Gamma_{11} - 1 & \Gamma_{13} \\
\Gamma_{13} & \Gamma_{33} - 1
\end{pmatrix}
\begin{pmatrix}
g_1 \\
g_3
\end{pmatrix}
= \begin{pmatrix}
0 \\
0
\end{pmatrix},
\]

(G3)

where \( \Gamma_{ik} \) is defined in (17) and \( g_1 \) and \( g_3 \) are the relevant polarization vector components. Equation (G3) gives two homogenous equations in the two unknowns \( g_1 \) and \( g_3 \), that can be solved to give

\[
\Gamma_{13}^2 = (\Gamma_{11} - 1) \cdot (\Gamma_{33} - 1).
\]

(G4)

For a 2D TI medium with \( \Gamma_{11}, \Gamma_{33} \) and \( \Gamma_{13} \) defined by

\[
\begin{align*}
\Gamma_{11} &= \hat{A}_{11} p_1^2 + \hat{A}_{44} p_3^2, \\
\Gamma_{33} &= \hat{A}_{44} p_1^2 + \hat{A}_{33} p_3^2, \\
\Gamma_{13} &= (\hat{A}_{13} + \hat{A}_{44}) p_1 p_3,
\end{align*}
\]

(G5)

the solution of (G4) can be written as

\[
p_3^2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},
\]

(G6)

where

\[
\begin{align*}
a &= \hat{A}_{33} \hat{A}_{44}, \\
b &= \left[ \hat{A}_{33} \cdot \hat{A}_{11} + \hat{A}_{44}^2 - (\hat{A}_{13} + \hat{A}_{44})^2 \right] p_1^2 - (\hat{A}_{33} + \hat{A}_{44}), \\
c &= \hat{A}_{11} \cdot \hat{A}_{44} \cdot p_1^2 - (\hat{A}_{11} + \hat{A}_{44}) p_1^2 + 1
\end{align*}
\]

(G7)

We take the negative sign in (G6) for \( qP \)-rays and the positive sign for \( qSV \)-rays. When solving for \( p_3 \) we take the sign that fits the orientation of the coordinate system and the ray shooting direction. For a TTI medium in the global coordinate system, the Christoffel equation gives a fourth-order polynomial equation in \( p_3 \) that needs to be solved numerically.
APPENDIX H: DERIVATIVES FOR A 2D TTI MEDIUM

For a 2D TTI medium, \( x_2 = 0 \), \( p_2 = 0 \) and \( g_2 = 0 \), and the relevant Christoffel matrix components are given from Červený (2001) as

\[
\begin{align*}
\Gamma_{11} &= A_{11}p_1^2 + 2A_{15}p_1p_3 + A_{55}p_3^2, \\
\Gamma_{33} &= A_{55}p_1^2 + 2A_{35}p_1p_3 + A_{33}p_3^2, \\
\Gamma_{13} &= A_{15}p_1^2 + (A_{13} + A_{55})p_1p_3 + A_{35}p_3^2.
\end{align*}
\]  

(H1)

The derivatives of the Christoffel matrix with respect to Voigt parameters,

\[
\frac{\partial \Gamma_{ik}}{\partial f_V}_{|_{x,p}} = \frac{\partial (\Gamma_{11}, \Gamma_{33}, \Gamma_{13})}{\partial (A_{11}, A_{33}, A_{55}, A_{13}, A_{15}, A_{35})}_{|_{x,p}},
\]  

(H2)

can be written in the table

\[
\begin{array}{|c|ccc|ccc|}
\hline
& A_{11} & A_{33} & A_{55} & A_{13} & A_{15} & A_{35} \\
\hline
\Gamma_{11} & p_1^2 & 0 & p_2^2 & 0 & 2p_1p_3 & 0 \\
\Gamma_{33} & 0 & p_3^2 & p_2^2 & 0 & 0 & 2p_1p_3 \\
\Gamma_{13} & 0 & 0 & p_1p_3 & 0 & p_1p_3 & p_2^2 \\
\hline
\end{array}
\]  

(H3)

Then we need the derivatives of \( \frac{\partial \Gamma_{ik}}{\partial f_V}_{|_{x,p}} \) with respect to \( p \), while keeping \( x \) fixed. These can be obtained directly from the table above. We get the results

\[
\begin{array}{|c|ccc|ccc|}
\hline
& A_{11} & A_{33} & A_{55} & A_{13} & A_{15} & A_{35} \\
\hline
\Gamma_{11} & 2p_1 & 0 & 0 & 0 & 2p_3 & 0 \\
\Gamma_{33} & 0 & 0 & 2p_1 & 0 & 0 & 2p_3 \\
\Gamma_{13} & 0 & 0 & p_3 & p_3 & 2p_1 & 0 \\
\hline
\end{array}
\]  

(H4)

and

\[
\begin{array}{|c|ccc|ccc|}
\hline
& A_{11} & A_{33} & A_{55} & A_{13} & A_{15} & A_{35} \\
\hline
\Gamma_{11} & 0 & 0 & 2p_3 & 0 & 2p_1 & 0 \\
\Gamma_{33} & 0 & 2p_3 & 0 & 0 & 0 & 2p_1 \\
\Gamma_{13} & 0 & 0 & p_1 & p_1 & 0 & 2p_3 \\
\hline
\end{array}
\]  

(H5)

From (E16), with \( b = \theta \) and \( g_2 = 0 \), we have

\[
\frac{\partial v}{\partial \theta} = \frac{1}{2v} \left( \frac{\partial \Gamma'_{11}}{\partial \theta} g_1^2 + \frac{\partial \Gamma'_{33}}{\partial \theta} g_3^2 + 2 \frac{\partial \Gamma'_{13}}{\partial \theta} g_1g_3 \right). 
\]  

(H6)
Using (36), with $n_2 = 0$, it then follows that

$$
\begin{pmatrix}
\frac{\partial \Gamma'}{\partial \theta} \\
\frac{\partial \Gamma'}{\partial \theta} \\
\frac{\partial \Gamma'}{\partial \theta}
\end{pmatrix}
= 2
\begin{pmatrix}
A_{11} & A_{15} & A_{15} & A_{55} \\
A_{55} & A_{35} & A_{35} & A_{33} \\
A_{15} & \frac{1}{2} (A_{13} + A_{55}) & \frac{1}{2} (A_{13} + A_{55}) & A_{35}
\end{pmatrix}
\begin{pmatrix}
n_1 \frac{\partial n_1}{\partial \theta} \\
n_1 \frac{\partial n_3}{\partial \theta} \\
n_3 \frac{\partial n_1}{\partial \theta} \\
n_3 \frac{\partial n_3}{\partial \theta}
\end{pmatrix}.
$$

(H7)

From (E16), with $b = f_V$ and $g_2 = 0$, we have

$$
\frac{\partial v}{\partial f_V} = \frac{v}{2} \left( \frac{\partial \Gamma_{11}}{\partial f_V} g_1^2 + \frac{\partial \Gamma_{33}}{\partial f_V} g_3^2 + 2 \frac{\partial \Gamma_{13}}{\partial f_V} g_1 g_3 \right),
$$

(H8)

and inserting from (H3) gives

$$
\begin{align*}
\frac{\partial v}{\partial A_{11}} &= \frac{v}{2} \left( p_1^2 g_1^2 \right), \\
\frac{\partial v}{\partial A_{33}} &= \frac{v}{2} \left( p_3^2 g_3^2 \right), \\
\frac{\partial v}{\partial A_{55}} &= \frac{v}{2} \left( p_3^2 g_1^2 + 2 p_1 p_2 g_1 g_3 + p_1^2 g_3^2 \right), \\
\frac{\partial v}{\partial A_{13}} &= \frac{v}{2} \left( 2 p_1 p_3 g_1 g_3 \right), \\
\frac{\partial v}{\partial A_{15}} &= \frac{v}{2} \left( 2 p_1 p_3 g_1^2 + 2 p_1^2 g_1 g_3 \right), \\
\frac{\partial v}{\partial A_{35}} &= \frac{v}{2} \left( 2 p_1 p_3 g_3^2 + 2 p_3^2 g_1 g_3 \right),
\end{align*}
$$

(H9a-f)

where the phase velocity $v = 1/\sqrt{p_i p_i}$ is provided from the ray tracing.