Frequency domain elastic FWI for VTI media
Junxiao Li$^1$, Kristopher A. Innanen$^1$, Wenyong Pan$^1$ and Geng Yu$^1$
$^1$Dept. of Geoscience, CREWES Project, University of Calgary, Calgary, AB.

SUMMARY
In this study, a frequency-domain elastic full waveform inversion algorithm for 2-D VTI media has been developed. The forward problem used in this inversion algorithm is simulated by applying frequency domain finite difference method, which is a fast approach for multi-source and multi-receiver acquisition. For the anisotropic inversion of VTI media, five elastic constants ($c_{11}$, $c_{13}$, $c_{33}$, $c_{44}$ and density) have to be dealt with. In this paper, the gradients of four elastic constants are calculated in matrix forms. To accelerate the convergence rate of inversion, the pseudo-Hessian is also implemented in the objective function. The inversion results in the paper indicate that parameters $c_{11}$, $c_{33}$ and $c_{44}$ can be inverted properly, yet $c_{13}$ can not be inverted properly due to the severe crosstalk.

INTRODUCTION
In oil and gas exploration, seismic inversion plays a key role in delineating subsurface structures (Tarantola, 1984; Pratt et al., 1998; Shin and Min, 2006). Most of the FWI techniques are performed under the assumption that the underground formations are isotropic, yet, according to Thomsen (1986), shale, laminated thin-layers and oriented vertical fractures are transversely isotropic (TI) media. And transversely isotropic media with a vertical symmetry axis (VTI) are commonly observed in sedimentary formations. For this reason, anisotropy should be taken into account in FWI.

In the anisotropic waveform inversion, more formation parameters should be inverted than in isotropic inversion (Chang and McMechan, 2009; Kamath and Tsvankin, 2016; Pan et al., 2016). This implies multiparameter anisotropic elastic FWI is a highly non-linear problem. Plessix (2009) proposed that the inversion resolution can be greatly enhanced by taking anisotropy into consideration. Barnes et al. (2008) used the full waveform inversion in transversely isotropic media. In his paper, the Thomsen parameters (Thomsen, 1986) (including isotropic parameters of $\delta$ and $\epsilon$) are inverted. It showed that the isotropic parameters cannot be reconstructed properly the anisotropic parameters can not be well restored. According to Gholami and Siahkohi (2010), Thomsen’s parameters can not be inverted simultaneously even if source-receiver stations are well distributed at all directions. Gholami et al. (2013) also pointed out that the choice of parameterization is important for anisotropic FWI. Lee et al. (2010) applied a frequency-selection strategy, moving from lower to higher frequencies to invert elastic constants($c_{11}$, $c_{13}$, $c_{33}$, $c_{44}$) in VTI media. In his paper, he coupled elastic constants $c_{11}$ and $c_{33}$ based on Thomsen’s relationship. And the steepest-descent method based on the adjoint state of the wave equations (Lailly, 1983; Tarantola, 1984; Pratt et al., 1998) are used to updated by model parameters. In this paper, we will invert elastic constants $c_{11}$, $c_{13}$, $c_{33}$, $c_{44}$ while keeping density fixed. As an inversion problem, FWI requires intensive computation. An efficient step-length formula is extremely important to accelerate the convergence rate. In this study, we estimate the step length by a modified quadratic interpolation method. The inversion results show that satisfying reconstruction can be obtained for elastic constants $c_{11}$, $c_{33}$ and $c_{13}$, yet the inversion of $c_{44}$ needs to be enhanced.

THEORY AND PRINCIPLES OVERVIEW

Frequency domain forward modeling

Wave propagation in an elastic medium is governed by equation:

$$\rho \omega^2 u_i = \sigma_{ij},$$  (1)

where $i, j = 1, 2, 3$, $\rho$ is the density, $u_i$ is the displacement vector and $\sigma_{ij}$ is stress tensor, and where $\sigma_{(x, y)}$ represent spatial derivatives of the stress tensor. The comma between subscripts is used for spatial derivatives. The summation convention for repeated subscripts is assumed. According to Hooke’s law, the relationship between the stress and strain tensors is,

$$\sigma_{ij} = c_{ijkl} \varepsilon_{kl},$$

(2)

where $c_{ijkl}$ are the elastic stiffness coefficients. The strain tensor $\varepsilon_{ij}$ is

$$\varepsilon_{ij} = \frac{1}{2} (u_{ij} + u_{ji}).$$

(3)

In the case of a transverse isotropic medium, the second-order wave equation system in frequency domain can be written as

$$-\rho \omega^2 u_i = \frac{\partial^2 \sigma_{ij}}{\partial x_j} + f_i,$$

(4)

For formations with a vertical symmetry axis (VTI), the elastic stiffness tensor is

$$c_{ijkl} = \begin{bmatrix} c_{11} & c_{11} - 2c_{66} & c_{13} & 0 & 0 & 0 \\
0 & c_{11} & c_{11} & 0 & 0 & 0 \\
c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\
0 & 0 & 0 & c_{44} & 0 & 0 \\
0 & 0 & 0 & 0 & c_{44} & 0 \\
0 & 0 & 0 & 0 & 0 & c_{66} \end{bmatrix},$$

(5)

And the 2D elastic stiffness tensor is reduced to

$$c_{ijkl} = \begin{bmatrix} c_{11} & c_{13} & 0 \\
c_{13} & c_{33} & 0 \\
0 & 0 & c_{44} \end{bmatrix}.$$  (6)

The 2D elastic wave equations for VTI media can be written as

$$-\rho \omega^2 \ddot{u}_i = \frac{\partial}{\partial x_i} \left( c_{11} \frac{\partial u_i}{\partial x_j} + c_{13} \frac{\partial u_j}{\partial x_j} \right) + \frac{\partial}{\partial x_j} \left( c_{44} \left( \frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} \right) \right) + f_i(\omega),$$

(7)

The above equations can further be written as

$$W\ddot{u} = f.$$  (8)

Rewrite the above equations into the matrix formalism (Pratt and Worthington, 1988), we have

$$\begin{bmatrix} W_{xx}(x, \omega) & W_{x\omega}(x, \omega) & \ddot{u}_i(x, \omega) \\
W_{\omega x}(x, \omega) & W_{\omega\omega}(x, \omega) & \ddot{u}_i(x, \omega) \end{bmatrix} = \begin{bmatrix} f_x(x, \omega) \\
f_{x\omega}(x, \omega) \end{bmatrix},$$

(9)

where, $[f_x(x, \omega) \ f_{x\omega}(x, \omega)]^T$ is the source vector $f$, and the wave operator $W(x, \omega)$ is defined as

$$W(x, \omega) = \begin{bmatrix} W_{xx}(x, \omega) & W_{x\omega}(x, \omega) \\
W_{\omega x}(x, \omega) & W_{\omega\omega}(x, \omega) \end{bmatrix}.$$  (10)
Frequency domain elastic FWI

In VTI media, \( W(x, \omega) \) can be written as

\[
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right. \\
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right. \\
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right. \\
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right]
\]

(11)

In isotropic media, \( W(x, \omega) \) can be written as

\[
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right. \\
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right. \\
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right. \\
W_{\omega}(x, \omega) = -\rho(x)\omega^2 - \frac{1}{2}\rho(x)\left[ c_{11} \frac{\partial}{\partial x} \omega^2 - \frac{1}{2}\rho(x) \right]
\]

(12)

**Gradient direction**

The main objective of FWI is to find a suitable velocity model by minimizing the objective function based on the residuals between modelled and field data (Laillly, 1983; Tarantola, 1984; Virieux and Operto, 2009). The general relation between the model \( m \) and the data \( u \) can be expressed as

\[
u = g(m).
\]

The objective function can be written as

\[
E(m) = \frac{1}{2} \sum_{\omega} \sum_{s} |u - d|^T (u - d)^*.
\]

(14)

where \( u \) and \( d \) are the modelled and observed data, superscripts \( T \) and \( * \) are transpose and complex conjugate, respectively. \( \omega \) and \( s \) denote the frequency and source.

The inversion problem now has changed into finding the global minimum of this misfit function. We choose an initial model and perform iterations to reach its neighboring minimum. And then the best model calculated at one iteration can be treated as new initial model for the next iteration. For a given initial model \( m_0 \), take a second-order Taylor-Lagrange expansion, the misfit function of equation (14) can be expressed as

\[
E(m_0 + \delta m) = E(m_0) + \nabla_m E(m_0)^T \delta m + 1/2 \delta m^T \nabla_m^2 E(m_0) \delta m + O(\delta m^2),
\]

(15)

where, \( \delta m \) is parameter perturbation and \( O(\delta m^2) \) is second-order Lagrange remainder, and \( \nabla_m E \) is the gradient.

When a local minimum of \( E \) is reached with a suitable increment of \( \delta m \), the above equation can be written as

\[
\nabla_m E(m_0) = \nabla_m^2 E(m_0) \delta m = H(m_0) \delta m,
\]

(16)

The derivative of the gradient with respect to model \( m \) is Hessian \( H \). For a perturbation of the \( k \)th parameter of a model \( m \), its gradient can be obtained by taking the partial derivative of the objective function of equation (14) with respect to the \( k \)th model parameter as

\[
\nabla_{mk} E(m_k) = \sum_\omega \sum_s \Re \left[ \left( \frac{\partial u}{\partial m_k} \right)^T \left( \tilde{u} - \tilde{d} \right) \right].
\]

(17)

The first term on the RHS of the above equation can be obtained by taking the partial derivative of both sides of equation (8) as

\[
\frac{\partial u}{\partial m_k} = W^{-1} \tilde{f}_k.
\]

(18)

And \( \tilde{f}_k \) is the virtual source vector (Pratt et al., 1998) of the \( k \)th model parameter, which can be written as

\[
\tilde{f}_k = -\frac{\partial W}{\partial m_k} u.
\]

(19)

Substitute equation (18) and (19) into equation (17), we have

\[
\nabla_{mk} E(m_k) = \sum_\omega \sum_s \Re \left[ \frac{\partial u}{\partial m_k} \left( W^{-1} \frac{\partial \tilde{f}_k}{\partial m_k} (\tilde{u} - \tilde{d}) \right) \right] = \sum_\omega \sum_s \Re \left[ \left( \frac{\partial u}{\partial m_k} \frac{\partial \tilde{f}_k}{\partial m_k} \right) (\tilde{u} - \tilde{d}) \right].
\]

(20)

This equation shows the gradient can be built as a product between the incident wavefields \( \tilde{u} \) and the back-propagated wavefields \( (W^{-1} \frac{\partial \tilde{f}_k}{\partial m_k}) \), with residuals at receiver positions as a back-propagated source. The operator \( \frac{\partial \tilde{f}_k}{\partial m_k} \) with respect to elastic constants in VTI media can be calculated based on equation (11). Take \( \frac{\partial \tilde{f}_k}{\partial m_k} \) as an example, it can be expressed as

\[
\frac{\partial \tilde{f}_k}{\partial m_k} = \frac{\partial}{\partial m_k} \frac{\partial \tilde{f}_k}{\partial \lambda} \frac{\partial \lambda}{\partial m_k}.
\]

(21)

By scaling the gradients of model parameters, the inversion convergence can be accelerated. Considering the computational overburden, the pseudo-Hessian matrix instead of the full Hessian matrix is applied to scale the gradients (Shin et al., 2001). The gradient for each parameter is then calculated using

\[
\nabla_{mk} E(m_k) = \sum_\omega \sum_s \Re \left[ \frac{\partial u}{\partial m_k} \left( W^{-1} \frac{\partial \tilde{f}_k}{\partial m_k} (\tilde{u} - \tilde{d}) \right) \right].
\]

(22)

in which \( \lambda \) is the damping factor and \( I \) is the identity matrix in Marquardt-Levenberg regularization. We choose \( \lambda = 0.01 \) in this paper.

Combining all the ingredients, we can calculate the gradients with respect to different parameters. We start from an isotropic model with P- and S-wave velocity of 1500 m/s and 1200 m/s for the background.
The wavefields of a homogeneous initial model: (a) Real and (c) imaginary part of the wavefield in X component; (b) Real and (d) imaginary part of the wavefield in Z component.

Figure 3: The wavefields of an X-component back propagated source with a P-wave anomaly in the middle: (a) Real and (c) imaginary part of the wavefield in X component; (b) Real and (d) imaginary part of the wavefield in Z component.

Figure 4: The wavefields of n Z-component back propagated source in z-direction with a P-wave anomaly in the middle: (a) Real and (c) imaginary part of the wavefield in X component; (b) Real and (d) imaginary part of the wavefield in Z component.

velocities, respectively. A round anomaly (P-wave velocity in this anomaly is 1800 m/s) with a radius of 50 m is located in the middle of the model (the size of the model is 1560 x 1560m², with a PML of 50m at each side; the space-sample is 10m.). A vertical point source S is located at S = (200m, 780m), and a receiver R is located at R = (1300m, 780m). The inverted frequency is 7 Hz. The wavefields of this true model is shown in Figure 1. Take the background without the P-wave velocity anomaly as the initial model, the wavefields with respect to x- and z-component can be obtained with the same source-receiver distribution, which are shown in Figure 2. Take the residuals of receivers in both x- and z-components for the true and initial models as sources, the wavefields of horizontal and vertical sources can thus be calculated, shown in Figure 3 and ??.

Taking each part as a matrix, the gradient in equation (17) can be written as matrix form

\[ \nabla_{m} E(m) = \begin{bmatrix} \frac{\Delta W_{xx}}{\Delta m_{x}} & \frac{\Delta W_{xz}}{\Delta m_{x}} & \frac{\Delta W_{x\omega}}{\Delta m_{x}} \\ \frac{\Delta W_{xz}}{\Delta m_{x}} & \frac{\Delta W_{zz}}{\Delta m_{x}} & \frac{\Delta W_{z\omega}}{\Delta m_{x}} \end{bmatrix} \begin{bmatrix} \hat{u}_{x} \\ \hat{u}_{z} \end{bmatrix} \]

In Figure 5, the gradients of \( V_{p}, V_{z} \) and \(\delta \) are calculated for this isotropic medium. If we change this isotropic medium into VTI medium, whose stiffness tensor is

\[ c_{VTI} = \begin{bmatrix} 23.87 & 9.79 & 0 \\ 9.79 & 15.33 & 0 \\ 0 & 0 & 2.77 \end{bmatrix} \times 10^{9}N/m^{2}. \] (24)

The gradients with respect to different stiffness tensor can also be calculated with equation (21), which are shown in Figure 6.

Figure 7: Illustration that schematically outlines the principle of the modified quadratic interpolation step-length formula. (a) Original step length obtained by line search method and (b) Modified step length after interpolation.
As an inversion problem, FWI requires intensive computation. An efficient step-length formula is extremely important to accelerate the convergence rate. In this study, we estimate the step length by a modified quadratic interpolation method. A step length that minimizes the misfit function can be found by using the line search method, shown in Figure 7(a). Unlike directly using the step length in the iteration, the adjacent left and right closest step lengths are also used (red curve in Figure 7(b)) to interpolate a modified step length that minimizes the misfit function.

EXAMPLES

In this section, a two-layer VTI model with two finite-sized anomalous circles of radius 100 m is presented, shown in Figure 8(a). The sources and receivers are distributed around the model, shown as dotted points (80 sources and 200 receivers).

There is a finite-sized anomalous circle at each layer. The circle anomaly in the first layer is a $c_{11}$ perturbation. For the first model, shown in 8(b), the circle anomaly in the second layer is also a $c_{11}$ perturbation. While keeping the circle anomaly of $c_{11}$ perturbation in the first layer, we change this circle anomaly in the second layer into $c_{33}$ perturbation to form model 2 (8(c)) and $c_{44}$ perturbation to form model 3 (8(d)).

We choose a start frequency of 2 Hz to 17 Hz with a sample of 3 Hz. The initial medium is the constant background VTI model. For each iteration, the updated initial model comes from the previously inverted model with different frequencies. The maximum iteration is 10 for each frequency. The inversion results of the first model ($c_{11}$ perturbations in both layers) are shown in Figure 9. Both the shape of the anomalies and parameter amplitudes are well reconstructed. The inversion results depend on the selected frequencies. The more the inverted frequencies are to be used, the more finely the inversion results would be.

For the inversion of model 2, in which, a $c_{33}$ perturbation is located in the second layer, the inversion results are shown in Figure 10. However, for the $c_{11}$ reconstruction in the first layer, the presence of the crosstalk caused by the $c_{33}$ perturbation is obvious in $c_{11}$ inversion results. The same crosstalk can also be detected for model 3 when there are $c_{11}$ and $c_{44}$ perturbations in different layers, which is shown in Figure 11.

CONCLUSIONS

In this paper, the frequency-domain elastic inversion algorithm for elastic constants in VTI media has been discussed. According to the frequency domain forward modeling, the gradient direction can be eas-