Anisotropic full waveform inversion

Huy Le

ABSTRACT

I develop a time-domain method for anisotropic full waveform inversion based on the second-order system of pseudo-acoustic wave equations in vertical transverse isotropic media. I use a synthetic model with both reflections and diving waves to test the method's performance. Three inversion passes are carried out in a frequency continuation manner using three source wavelets with peak frequencies at 5 Hz, 10 Hz, and 20 Hz. The residual is reduced significantly after each pass, and the inverted models converge toward the true ones. I also investigate the use of the Hessian to precondition the objective function's gradient and to reduce crosstalks between different parameters. The synthetic examples show that the inverse of the Hessian spatially balances the amplitude of the gradients, focuses energy to the actual model perturbations' locations, and reduces parameter crosstalks.

INTRODUCTION

Full waveform inversion seeks to match the modeled data and the observed data in both kinematics and amplitude. Its practical applications, however, have been focusing on the kinematics, based on acoustic wave equations (Gholami et al., 2013b; Warner et al., 2013). As more field data are acquired with longer offsets and wider azimuths, it is necessary to incorporate anisotropy in order to match the kinematics accurately.

The pseudo-acoustic anisotropic wave equations were introduced by Alkhalifah (2000), and has been demonstrated to accurately capture the kinematics of seismic waves. Several of formualtions for these equations have been used in the industry. They differ in their exact details but are all equivalent to the originally proposed system by Alkhalifah (2000). Here I choose a system similar to the one used in Duneneck and Bakker (2011). Expressed in second-order form, this system of equations has been studied for imaging and inversion purposes (Fletcher et al., 2009; Zhang et al., 2011; Bube et al., 2012; Warner et al., 2013). For the simplest anisotropic model, vertical transverse isotropic (VTI), the system can be written as:

$$\begin{cases} \partial_t^2 p = c_{11}\partial_x^2 p + c_{13}\partial_z^2 q + f_x, \\ \partial_t^2 q = c_{13}\partial_x^2 p + c_{33}\partial_z^2 q + f_z. \end{cases} \tag{1}$$

Here p and q are the normal stresses in the x-direction and z-direction, f_i are the sources, and c_{ij} are the density-normalized stiffness coefficients, which are related to

vertical P-velocity, v_{pz}, horizontal P-velocity, v_{px}, and Thomsen's parameters, ϵ and δ, by:

$$c_{11} = v_{pz}^2(1+2\epsilon) = v_{px}^2, \tag{2a}$$
$$c_{13} = v_{pz}^2\sqrt{1+2\delta}, \tag{2b}$$
$$c_{33} = v_{pz}^2. \tag{2c}$$

With $\mathbf{m}$ being the vector of model parameters, the modeled pressure field is defined as the average stress, $d(\mathbf{x}, t; \mathbf{m}) = \frac{1}{2}(p+q)$.

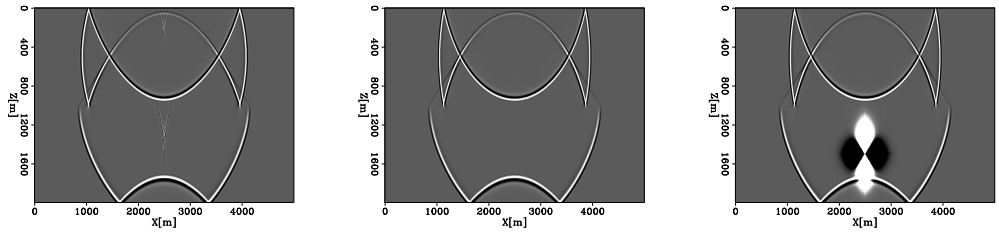
There are two issues with system 1. Firstly, it suffers from a so-called shear artifact, i.e., the diamond-shaped event shown in Figure 1a. This artifact can be reduced when the source is placed in an isotropic region, as in Figure 1b. Secondly, for the system 1 to be stable, the stiffness matrix, $\mathbf{C} = \begin{bmatrix} c_{11} & c_{13} \\ c_{13} & c_{33} \end{bmatrix}$, has to be semi-positive definite:

$$\begin{cases} c_{11} \geq 0, \\ c_{11}c_{33} - c_{13}^2 \geq 0. \end{cases} \tag{3}$$

In terms of Thomsen's parameters, these constraints mean:

$$\begin{cases} \epsilon \geq -\frac{1}{2}, \\ \epsilon \geq \delta. \end{cases} \tag{4}$$

Failure to maintain these constrants leads to solutions that grow unstably (Figure 1c).



(a) (b) (c)

Figure 1: Snapshots of solution wavefield for system 1: (a) showing the diamond-shaped shear artifact, (b) with the source in water, and (c) growing unstably when the medium properties does not satisfy constraints 4, i.e. $\epsilon < \delta$. [**ER**]

ADJOINT STATE METHOD FOR FIRST DERIVATIVES

Full waveform inversion aims to find a subsurface model that best describes the observed data by minimizing an objective function. This objective function can defined

as the l_2-norm of the difference between the modeled data, $d(\mathbf{x}_r, t; \mathbf{m})$, and the observed data, $d_0(\mathbf{x}_r, t)$:

$$\chi(\mathbf{m}) = \frac{1}{2}\|d(\mathbf{x}_r, t; \mathbf{m}) - d_0(\mathbf{x}_r, t)\|_2^2. \tag{5}$$

Following the adjoint state method outlined in Fichtner (2011), the first derivatives of the objective function with respect to model parameters are computed as cross-correlations of the adjoint wavefields, p_1 and q_1, and the forward modeling wavefields, p and q:

$$\frac{\partial \chi}{\partial c_{11}} = \int_0^T p_1 \partial_x^2 p dt, \tag{6a}$$

$$\frac{\partial \chi}{\partial c_{13}} = \int_0^T (p_1 \partial_z^2 q + q_1 \partial_x^2 p) dt, \tag{6b}$$

$$\frac{\partial \chi}{\partial c_{33}} = \int_0^T q_1 \partial_z^2 q dt. \tag{6c}$$

The adjoint wavefields are solutions to the adjoint equations, for which the sources are the data residual injected at receivers' locations:

$$\begin{cases} \partial_t^2 p_1 = \partial_x^2 (c_{11} p_1 + c_{13} q_1) + \frac{1}{2}(d - d_0)\delta(\mathbf{x} - \mathbf{x_r}), \\ \partial_t^2 q_1 = \partial_z^2 (c_{13} p_1 + c_{33} q_1) + \frac{1}{2}(d - d_0)\delta(\mathbf{x} - \mathbf{x_r}). \end{cases} \tag{7}$$

I use a simple synthetic model to demonstate the potential of using equations 1 to estimate anisotropic medium parameters. The initial model for c_{11} are shown in Figure 2. It consists of a water layer, a velocity gradient layer, and a basement. Initial models for the other two parameters, c_{13} and c_{33}, have similar structures. The true models include three perturbations, shown in Figure 3, at different locations for three parameters. I use 800 receivers and 40 shots uniformly distributed on the surface. Receiver spacing is 10 m and source spacing is 200 m. I use the nonlinear conjugate gradient algorithm (Nocedal and Wright, 2006) to minimize the objective function. Three inversion passes are performed using Ricker wavelets with peak frequencies at 5 Hz, 10 Hz, and 20 Hz. The final models of the lower-frequency inversion pass are used as starting models for the higher-frequency inversion.

Figure 4 shows the inverted model perturbations plotted at the same clip as the true perturbations (Figure 3). The perturbations' locations are identified correctly. The magnitudes of the inverted perturbations converge toward, but are still quite far from, the true ones. c_{11} seems to be the best resolved parameter and c_{13} the least. This observation is in agreement with a previous study by Gholami et al. (2013a), who used radiation patterns to analyze parameter sensitivity. They showed that the amplitude of the radiation pattern of c_{13} is three times less than those of the other two parameters, which indicates weak influence of c_{13} to the observed data.

When plotted at a lower clip (Figure 5), the inverted perturbations show crosstalks between parameters. Strong imprints of c_{11} and c_{33} on c_{13} are noticed, whereas there

is a very weak leakage from c_{13} to the other two parameters and between c_{11} and c_{33}. This is because these two parameters are sensitive to different wave paths. c_{11} is sensitive to horizontally traveling waves while c_{33} is sensitive to vertically traveling waves. Figure 6 shows the normalized objective functions for three inversion passes. The objective fucntion is reduced significantly: almost 100% after the 5 Hz inversion, 80% after the 10 Hz inversion, and 65% after the 20 Hz inversion.

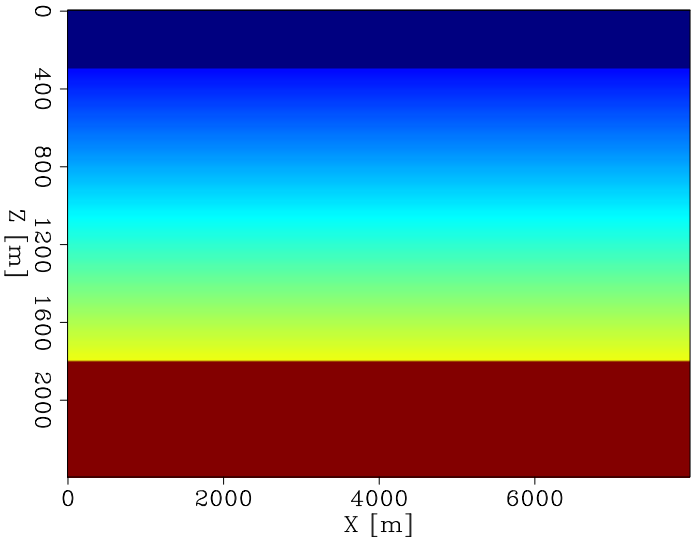
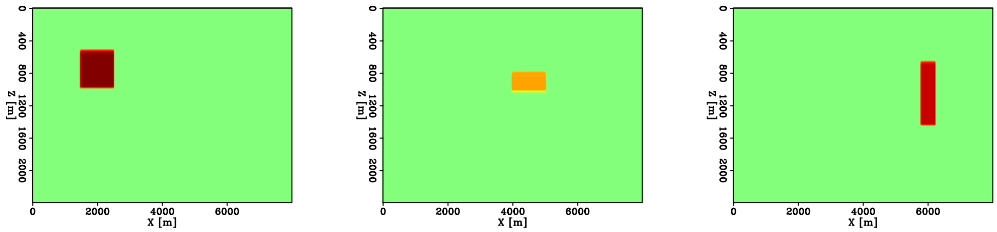


Figure 2: Initial model for c_{11} consists of a water layer, a velocity gradient layer, and a basement. Initial models for c_{13} and c_{33} are similar. [**ER**]

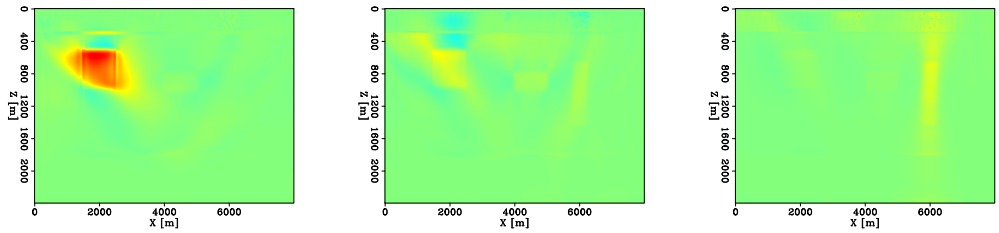


(a) (b) (c)

Figure 3: Perturbations in (a) c_{11}, (b) c_{13}, and (c) c_{33}, that are included in the true models. [**ER**]

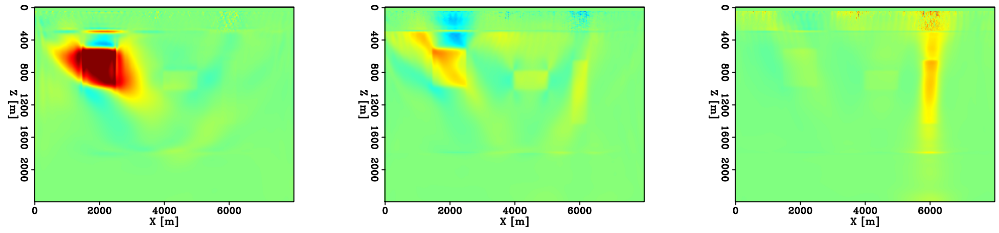
ADJOINT STATE METHOD FOR SECOND DERIVATIVES

Inversions carried out in the previous example utilize only the first-order information of the objective function (its gradients). As a result, the final inverted models exhibit crosstalks between different parameters. In this section, I investigate the use of second-order information in the Hessian to reduce the crosstalks. Ajoint state methods provide an efficient way to compute the Hessian's application to a model perturbation without explicitly calculating and storing it. This involves interactions



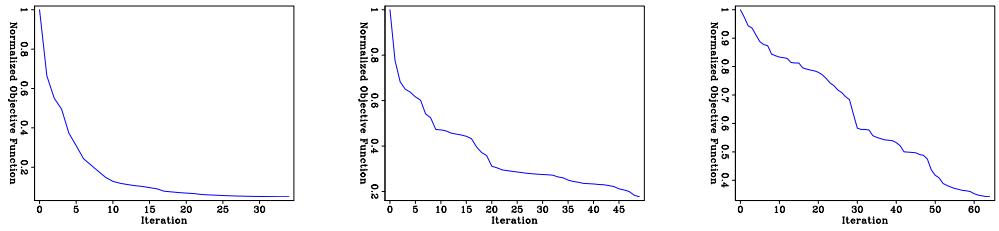
(a) (b) (c)

Figure 4: Inverted model perturbations plotted at the same clip of the true perturbations (Figure 3) for comparion: (a) c_{11}, (b) c_{13}, and (c) c_{33}. [**CR**]



(a) (b) (c)

Figure 5: Inverted model perturbations plotted at a lower clip to show the crosstalk: (a) c_{11}, (b) c_{13}, and (c) c_{33}. [**CR**]



(a) (b) (c)

Figure 6: Normalized objective functions: (a) 5 Hz inversion, (b) 10 Hz inversion, and (c) 20 Hz inversion. [**CR**]

between four wavefields, two of which are already found when computing the gradients: the forward wavefields p and q , and the primary adjoint wavefields p_1 and q_1, and two additional ones: the forward scattering wavefields δp and δq, and the secondary adjoint wavefields δp_1 and δq_1. As a result, the cost of computing one application of the Hessian is about twice that of computing one gradient.

The forward scattering wavefields come from the first-order Born approximation. They are the solutions of the linearized wave equations:

$$\begin{cases} \partial_t^2 \delta p = c_{11}\partial_x^2 \delta p + c_{13}\partial_z^2 \delta q + \delta c_{11}\partial_x^2 p + \delta c_{13}\partial_z^2 q, \\ \partial_t^2 \delta q = c_{13}\partial_x^2 \delta p + c_{33}\partial_z^2 \delta q + \delta c_{13}\partial_x^2 p + \delta c_{33}\partial_z^2 q. \end{cases} \tag{8}$$

The secondary adjoint wavefield are the resulting wavefields from scattering of the forward and adjoint wavefields off a model perturbations They are the solutions of:

$$\begin{cases} \partial_t^2 \delta p_1 = \partial_x^2(c_{11}\delta p_1 + c_{13}\delta q_1) + \frac{1}{4}(\delta p + \delta q)\delta(\mathbf{x} - \mathbf{x_r}) + \partial_x^2(\delta c_{11} p_1 + \delta c_{13} q_1), \\ \partial_t^2 \delta q_1 = \partial_z^2(c_{13}\delta p_1 + c_{33}\delta q_1) + \frac{1}{4}(\delta p + \delta q)\delta(\mathbf{x} - \mathbf{x_r}) + \partial_z^2(\delta c_{13} p_1 + \delta c_{33} q_1). \end{cases} \tag{9}$$

The Hessian applied to a model perturbation, $\delta\mathbf{m}$, is:

$$H\delta\mathbf{m} = \begin{bmatrix} \int_0^T (p_1\partial_x^2 \delta p + \delta p_1 \partial_x^2 p)dt \\ \int_0^T (p_1\partial_z^2 \delta q + q_1\partial_x^2 \delta p + \delta p_1 \partial_z^2 q + \delta q_1 \partial_x^2 p)dt \\ \int_0^T (q_1\partial_z^2 \delta q + \delta q_1 \partial_z^2 q)dt \end{bmatrix}. \tag{10}$$

The Gauss-Newton Hessian is obtained in the same fashion as the full Hessian, except only the first source terms are injected for the secondary adjoint wavefield (Equations 9). This means that the computation of Gauss-Newton Hessian requires the propagation of three wavefields: the forward wavefield, the Born wavefield, and the secondary adjoint wavefield. Consequently, the cost of this Hessian is about 1.5 times that of a gradient.

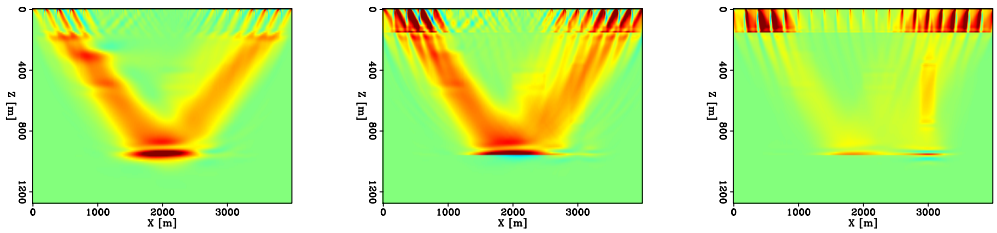
A Newton search direction can be obtained by solving the Newton equation:

$$\mathbf{Hp} = -\mathbf{g}, \tag{11}$$

where $\mathbf{H}$ can be either the full or approximate Hessian, $\mathbf{p}$ is the search direction, and $\mathbf{g}$ is the gradient. This equation can be solved iteratively with the linear conjugate gradient algorithm (Nocedal and Wright, 2006).

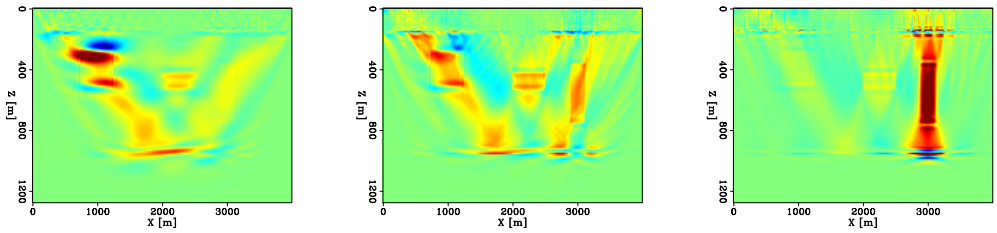
Figure 7 shows the steppest descent directions, $-\mathbf{g}$, computed for a similar model as in the previous example. The steepest descent directions, or the gradients, are contaminated with undesired, strong, long-wavelength energy from the sources at the edge of the survey, especially in the water layer. The inverse of the Hessian refocuses energy in the gradient to where the actual perturbations exists (Figure 8). There is still quite noticeable crosstalk between the parameters after 100 iteration

of linear conjugate gradient solving the Newton equation (Equation 11). A better approximation to the inverse of the Hessian (after 1000 iterations of linear conjuagte gradient) significantly reduces the crosstalk (Figure 9). Comparing the performance of the two Hessian, Figure 9 and Figure 10, the Gauss-Newton Hessian seems to do a better job than the full Hessian in focusing energy at the true perturbations. However, strong energy from the perturbations in c_{11} and c_{33} are leaked into the search direction of c_{13}, whereas much weaker energy is leaked between c_{11} and c_{33} and from c_{13} to the other two parameters. This behavior has been also observed in the inversion results in the previous example (Figure 5).



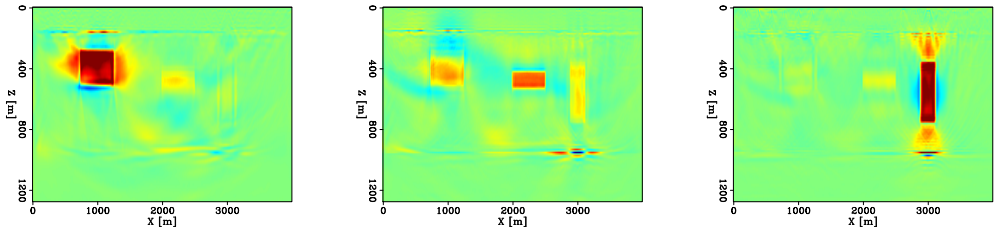
(a) (b) (c)

Figure 7: Steepest descent directions for: (a) c_{11}, (b) c_{13}, and (c) c_{33}. [**CR**]



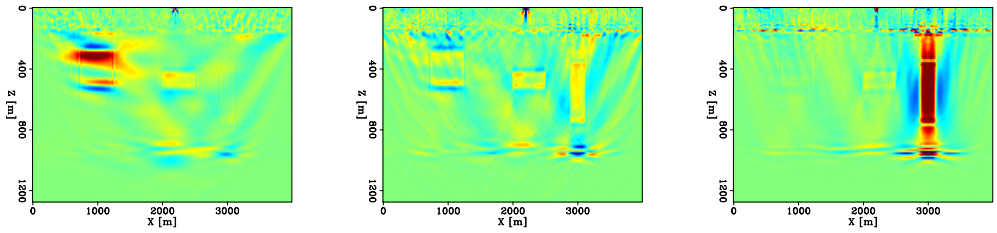
(a) (b) (c)

Figure 8: Gauss-Newton search directions for: (a) c_{11}, (b) c_{13}, and (c) c_{33}, after 100 iterations of linear conjugate gradient. [**CR**]



(a) (b) (c)

Figure 9: Gauss-Newton search directions for: (a) c_{11}, (b) c_{13}, and (c) c_{33}, after 1000 iterations of linear conjugate gradient. [**CR**]



(a) (b) (c)

Figure 10: Newton search directions for: (a) c_{11}, (b) c_{13}, and (c) c_{33}, after 1000 iterations of linear conjugate gradient. [**CR**]

CONCLUSIONS

I have shown that the second-order system of pseudo-acoustic anisotropic wave equations in vertical transverse isotropic media can be used for estimating anisotropic parameters. Tested on a simple synthetic model using a nonlinear conjugate gradient algorithm, my inversion scheme was able to reduce the difference between the modeled data and the observed data significantly. Eventhough the inverted models converge toward the true models, their magnitudes are still far from the true ones. This leaves room for future improvements, for example with regularization.

I have also presented a method to compute the Hessian for this system of wave equations and demonstated two benefits of using the Hessian. Firstly, the Hessian improves the frequency content and balances energy spatially in the gradient. Secondly, the Hessian shows the ability to reduce the crosstalk between different parameters but at a very high computational cost when its inverse is more accurately approximated. In these aspects, the Gauss-Newton Hessian performs better than the full Hessian and it is also less expensive to compute. I have shown that search directions obtained by solving the Newton equation show the same crosstalk patterns as in the inversion results. In particular, when parameterized with stiffness coefficients, c_{ij}, strong energy from c_{11} and c_{33} leaks into c_{13}, but very weakly in the other directions. Although using the Hessian in full waveform inversion might not be practical, on small models in which it is affordable, crosstalk between different parameters can be studied by observing the Newton search directions.

REFERENCES

Alkhalifah, T., 2000, An acoustic wave equation for anisotropic media: Geophysics, **65**, 1239–1250.

Bube, K. P., T. Nemeth, J. P. Stefani, R. Ergas, W. Liu, K. T. Nihei, and L. Zhang, 2012, On the instability in second-order systems for acoustic VTI and TTI media: Geophysics, **77**, No. 5, T171–T186.

Duneneck, E. and P. M. Bakker, 2011, Stable P-wave modeling for reverse-time migration in tilted TI media: Geophysics, **76**, No. 2, WA3–WA11.

Fichtner, A., 2011, Full seismic waveform modeling and inversion: Springer.

Fletcher, R. P., X. Du, and P. J. Fowler, 2009, Reverse time migration in tilted transversely isotropic (TTI) media: Geophysics, **74**, No. 6, WCA179–WCA187.

Gholami, Y., R. Brossier, S. Operto, A. Ribodetti, and J. Virieux, 2013a, Which parameterization is suitable for acoustic vertical transver isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis: Geophysics, **78**, No. 2, R81–R105.

———, 2013b, Which parameterization is suitable for acoustic vertical transver isotropic full waveform inversion? Part 2: Synthetic and real data case studies from Valhall: Geophysics, **78**, No. 2, R107–R124.

Nocedal, J. and S. J. Wright, 2006, Numerical optimization: Springer.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, No. 2, R59–R80.

Zhang, Y., H. Zhang, and G. Zhang, 2011, A stable TTI reverse time migration and its implementation: Geophysics, **76**, No. 3, WA3–WA11.