Using a nonlinear wave equation for anisotropic inversion

Huy Le, Biondo Biondi, Robert G. Clapp and Stewart A. Levin

ABSTRACT

We present some preliminary results on a specific nonlinear pseudo-acoustic wave equation in anisotropic media, including forward modeling, linearization, and adjoint method. Our objective is to find a robust and efficient method for anisotropic full-waveform inversion (FWI). The wave equation is solved discretely by the rotated staggered finite-difference scheme (RSFD) in time and space domains. The solution is more accurate than the one obtained using the centered finite-difference (CFD) scheme. The linearized equation is derived by taking only the first-order dependence of the wavefield with respect to medium parameters. The nonlinearity of the pseudo-acoustic wave equation introduces an additional term in the linearized equation. The adjoint method provides a mean to compute the gradients of the least-squares misfit objective function with respect to medium parameters through the adjoint wavefield. As a result of solving the forward equation by RSFD, the medium parameters are located on two different grids. We show that the gradients computed by the derived adjoint equation are in fact collocated consistently with the medium parameters. Applications on simple models, in both vertically transverse isotropic (VTI) and orthorhombic media, show that they also lead to the correct update directions. These results show the potential of our method for anisotropic parameter estimation.

INTRODUCTION

Though anisotropy has been recognized by the industry to play an important role in seismic imaging and inversion, its multiparameter nature remains a great challenge. In an attempt to reduce the number of parameters, Alkhalifah (1998) introduced a pseudo-acoustic approximation, under which shear-wave velocities along the symmetry axes are set to zero. This approximation reduces the number of anisotropic parameters to three for VTI (v_{pz}, ϵ, and δ) and to six for orthorhombic media (v_{pz}, ϵ_i, and δ_i). The pseudo-acoustic approximation results in an equation for P-wave in frequency-wavenumber domain as follows:

$$\left(\omega^2 - v_{pz}^2 k^2 S\right) u = 0, \tag{1}$$

where u is the pressure wavefield, ω is the angular frequency, $\mathbf{k}$ is the wave vector, and $S = S(\mathbf{n}, \epsilon_i, \delta_i)$; with $\mathbf{n}$ being the normalized wave vector. Mathematically, S

is a pseudo-differential operator, whose expressions for VTI and orthorhombic media can be found in the Appendix A (equations A-4 and A-5, respectively). Physically, it controls the degree of anisotropy along different propagation directions. For isotropic media, $S = 1$.

Equation 1 is a pseudo-differential equation, which can be computationally expensive to solve (Song and Alkhalifah, 2013; Le and Levin, 2014) because S incorporates all the anisotropic parameters. To overcome the computational intensity of solving equation 1, Xu and Zhou (2014) introduced the following approximation:

$$\mathbf{n} \approx \frac{\nabla u}{|\nabla u|}. \tag{2}$$

By regarding the wavefront normal, $\mathbf{n}$, as the direction of greatest change in the pressure wavefield, ∇u, this approximation ignores any amplitude variation with angle, but is exactly correct for plane waves. For this reason, the approximation 2 might be called plane-wave approximation. Assuming local homogeneity, the linear pseudo-differential equation 1 now becomes a nonlinear differential equation:

$$\partial_t^2 u - v_{pz}^2 \nabla \cdot (S \nabla u) = 0. \tag{3}$$

The nonlinearity comes from the dependence of the scalar term, S, on the wavefield, u, as a result of the approximation 2.

Forward modeling

Efficient solution to the forward modeling problem is crucial to any imaging and inversion process. Xu and Zhou (2014) solved equation 3 by a spectral method. However, using a spectral method to solve equation 3 is as expensive as solving equation 1. Taking advantage of the fact that equation 3 is no longer a pseudo-differential equation, we propose to use finite differences to solve this equation instead of spectral methods.

Equation 3 involves first derivatives, which can be inaccurately approximated by CFD. Figure 1 compares two snapshots of the wavefield solutions for pseudo-differential equation 1 by spectral method (panel a) and equation 3 by CFD (panel b). Similar artifacts as shown in Figure 1b were recognized by Ozdenvar and McMechan (1996) as a consequence of the nonlocality of the first-derivative CFD operator. To obtain an efficient and accurate solution, we use the RSFD scheme to solve equation 3 (Saenger et al., 2000).

Figure 2 shows a representative cell of the RSFD scheme in 2D, in which the pressure wavefield and its derivatives are located on two different grids that are staggered from each other: the main grid, denoted by circular nodes, and the staggered grid, denoted by square nodes. Consequently, the vertical P-wave velocity, which is associated with the wavefield, and the Thomsen parameters, which are associated

with its first derivatives, are also on different grids. Figure 3a shows a snapshot of the wavefield using RSFD that is more accurate than the one using CFD (Figure 1b). However, when viewed at a harder clip, Figure 3b reveals high-frequency noise generated around the source injection area. This noise can be mitigated by a smaller time step or a larger source area.

Beside the high-frequency artifacts around the source area, the RSFD solution to equation 3 (Figure 3a) shows some amplitude differences in comparison to the solution to equation 1 by a spectral method (Figure 1a). These differences are expected because the plane-wave approximation 2 neglects amplitude variation along wavefronts. Despite these differences, the RSFD solution matches the traveltimes of the spectral solution.

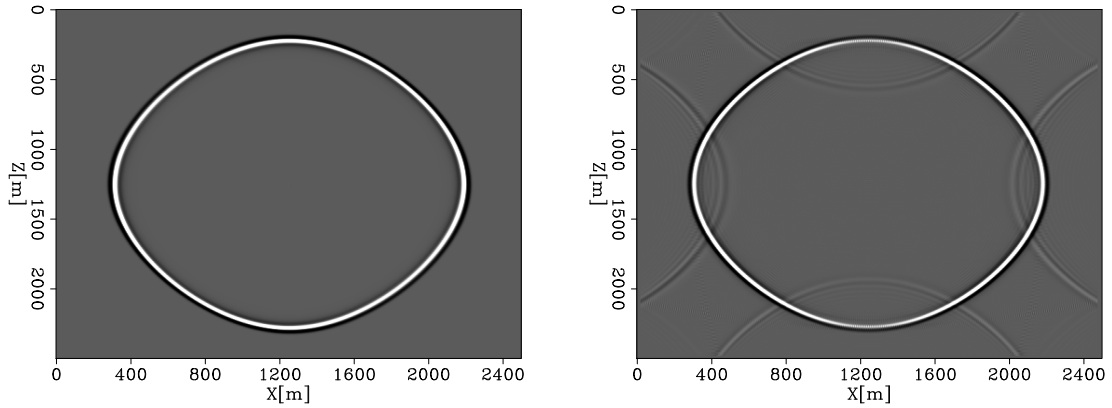


Figure 1: Wavefield snapshots of solutions to: (a) equation 1 by spectral method and (b) equation 3 by CFD in a homogeneous VTI medium. [**ER**]

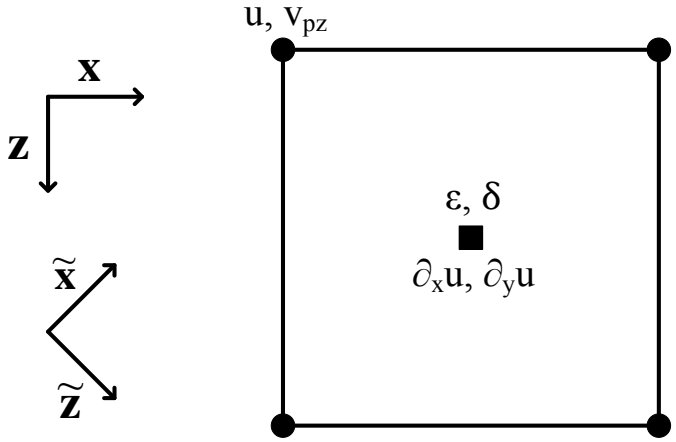


Figure 2: A representative cell of the RSFD scheme in 2D, in which the wavefield and velocity are located on the main grid while the anisotropic parameters and wavefield derivatives are located on the staggered grid. [**NR**]

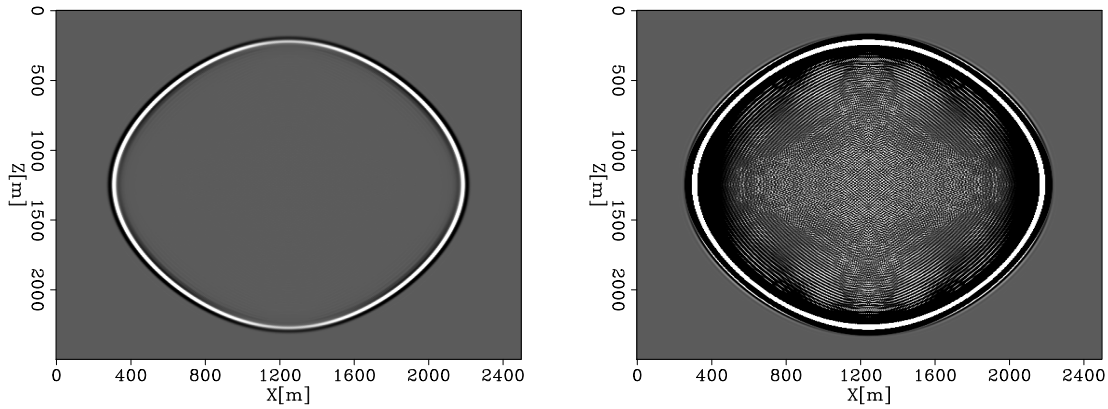


Figure 3: Modeling pseudo-acoustic wave equation in a homogeneous VTI medium using RSFD: (a) default clip and (b) hard clip revealing high-frequency noise around the source area. [**ER**]

Linearization

Define $m_1 = \frac{1}{v_{pz}^2}$ and m_i with $i > 1$ being the anisotropic parameters, ϵ_i and δ_i, or some combinations of them. In VTI for example, one possibility is $m_2 = 1 + 2\epsilon$ and $m_3 = \epsilon - \delta$, whereas in orthorhombic media, m_i with $i = 2, ..., 6$ can just be the anisotropic parameters. With the source term, f, the wave equation 3 is re-written on the domain $\mathbf{\Omega}$, together with initial and boundary conditions, as:

$$m_1 \partial_t^2 u - \nabla \cdot (S \nabla u) = f, \tag{4a}$$
$$u(\mathbf{x}, 0) = 0, \quad \partial_t u(\mathbf{x}, 0) = 0, \tag{4b}$$
$$u|_{\partial \mathbf{\Omega}} = 0. \tag{4c}$$

These equations give a nonlinear relationship between the pressure wavefield and the model parameters. For the purpose of least squares reverse time migration and inversion, the first-order Born modeling operator is needed. The Born operator is the derivative of the modeling function with respect to model parameters. It is obtained by linearizing the forward modeling equations 4.

Perturb equations 4 to obtain (see Appendix A for a detailed derivation):

$$m_1 \partial_t^2 \delta u - \nabla \cdot (\mathbf{A} \nabla \delta u) = \delta f, \tag{5a}$$
$$\delta u(\mathbf{x}, 0) = 0, \quad \partial_t \delta u(\mathbf{x}, 0) = 0, \tag{5b}$$
$$\delta u|_{\partial \mathbf{\Omega}} = 0, \tag{5c}$$

where

$$\mathbf{A} = S\mathbf{I} + \nabla u \left(\frac{\partial S}{\partial \nabla u} \right)^T, \tag{6}$$

and

$$\delta f = -\partial_t^2 u \delta m_1 + \nabla \cdot \left[\left(\sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i \right) \nabla u \right]. \tag{7}$$

Equation 5 is a linear partial differential equation for the perturbed wavefield, δu, because the differential operator now does not depend on the perturbed wavefield but on the background wavefield, u. As a result, they give a linear relationship between the perturbed wavefield and the perturbed model pamareters. Compared to the nonlinear forward modeling equation 4a, the linearized equation 5a has an additional term, $\nabla u \left(\frac{\partial S}{\partial \nabla u}\right)^T$, incorporated in $\mathbf{A}$ (equation 6). This is a result of the nonlinearity of equation 4a.

Based on the fact that when the pertubation is small enough, the perturbed wavefield is a good approximation to the difference between the full and background wavefields to the first order, we design a test for the linearized equations 5 with a point perturbation. The amount of perturbation is 5% in velocity and 50% in Thomsen parameters. Figure 4 and Figure 5 are the test results for a VTI and an orthorhombic media respectively. Although these figures display high-frequency artifacts and amplitude differences as seen earlier (Figure 3), they show a similarity between the perturbed wavefield and the wavefield difference. This similarity validates the linearized equations 5.

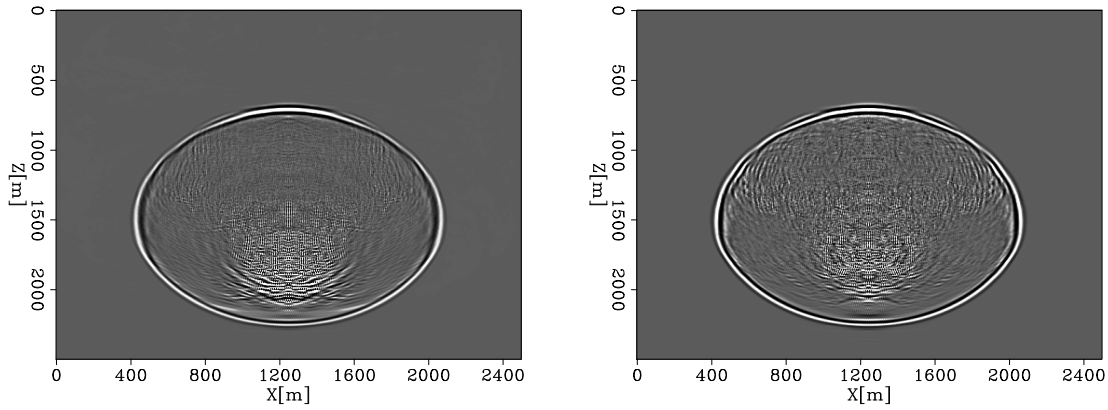
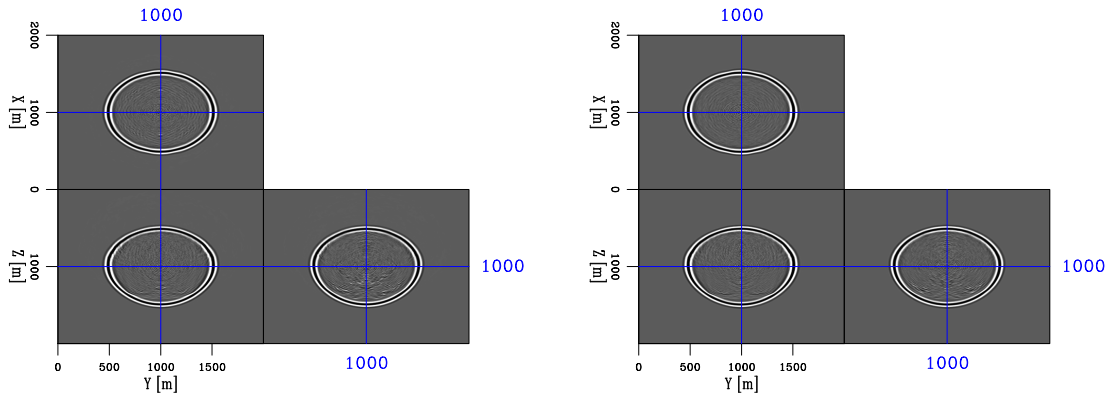


Figure 4: Linearization test in a VTI medium: (a) difference between the full and the background wavefields (b) the perturbed wavefield obtained by solving equation 5. [**ER**]

Adjoint method

Gradients of the FWI objective function were originally derived by using Green's functions (Tarantola, 1984). Because the acoustic isotropic wave equation is a linear partial differential equation, it was possible to express its solution in terms of Green's functions. The equation at hand (equation 3) is, however, not linear; and therefore



(a) (b)

Figure 5: Linearization test in an orthorhombic medium: (a) difference between the full and the background wavefields (b) the perturbed wavefield obtained by solving equation 5. [**CR**]

Green's theorem is not applicable in this case. As a result, we have to employ the adjoint method to derive the objective function's gradients.

The adjoint method provides an efficient way to obtain gradients of the FWI objective function through solution of an adjoint equation. Following Liu and Tromp (2006), the least-squares misfit function is defined as:

$$\chi = \frac{1}{2}\sum_r \int_0^T \|u(\mathbf{x}_r, t) - d(\mathbf{x}_r, t)\|^2 dt, \tag{8}$$

where $\mathbf{x}_r$ are the receivers' locations, T is the final time, and $d(\mathbf{x}_r, t)$ is the observed data.

The augmented Lagrangian is:

$$\chi = \frac{1}{2}\sum_r \int_0^T \|u(\mathbf{x}_r, t) - d(\mathbf{x}_r, t)\|^2 dt + \int_0^T \int_{\mathbf{\Omega}} \lambda \left[m_1 \partial_t^2 u - \nabla \cdot (S \nabla u) - f \right] dV dt, \tag{9}$$

where $\int_{\mathbf{\Omega}} dV$ is the volumetric integration.

The perturbed augmented Lagrangian is:

$$\begin{aligned} \delta\chi = &\int_0^T \int_{\mathbf{\Omega}} \sum_r \left[u(\mathbf{x}_r, t) - d(\mathbf{x}_r, t) \right] \delta(\mathbf{x} - \mathbf{x}_r) \delta u dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left[\delta m_1 \partial_t^2 u - \nabla \cdot (\delta S \nabla u) - \delta f \right] dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left[m_1 \partial_t^2 \delta u - \nabla \cdot (S \nabla \delta u) \right] dV dt, \end{aligned} \tag{10}$$

which, after some integration by parts and algebra manipulations (Appendix B), becomes:

$$\begin{aligned}\delta\chi = &\int_0^T \int_{\mathbf{\Omega}} \sum_r \left[u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\right] \delta(\mathbf{x}-\mathbf{x}_r)\delta u dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \left[\delta m_1 \lambda \partial_t^2 u + \nabla\lambda \cdot \left(\sum_{i>1} \frac{\partial S}{\partial m_i}\delta m_i\right)\nabla u - \lambda \delta f\right] dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \left[m_1 \partial_t^2 \lambda - \nabla\cdot\left(\mathbf{A}^T\nabla\lambda\right)\right] \delta u dV dt \\ &+ \int_{\mathbf{\Omega}} m_1 \left(\lambda\partial_t\delta u - \partial_t\lambda\delta u\right)|_T dV \\ &- \int_0^T \oint_{\partial\mathbf{\Omega}} \hat{\mathbf{n}}\cdot\lambda\left[\left(\sum_{i>1}\frac{\partial S}{\partial m_i}\delta m_i\right)\nabla u + \mathbf{A}\nabla\delta u\right] ds dt,\end{aligned} \tag{11}$$

where $\delta(\mathbf{x}-\mathbf{x}_r)$ is the delta function centered at the receivers' locations and $\oint_{\partial\mathbf{\Omega}} ds$ is the surface integration.

Define the adjoint equations, with final and boundary conditions, as:

$$m_1\partial_t^2\lambda - \nabla\cdot\left(\mathbf{A}^T\nabla\lambda\right) = -\sum_r \left[u(\mathbf{x}_r,t) - d(\mathbf{x}_r,t)\right]\delta(\mathbf{x}-\mathbf{x}_r), \tag{12a}$$

$$\lambda(\mathbf{x},T) = 0, \quad \partial_t\lambda(\mathbf{x},T) = 0, \tag{12b}$$

$$\lambda|_{\partial\mathbf{\Omega}} = 0. \tag{12c}$$

Notice that, as with the linearized equation (equation 5), the above equation is also linear with respect to the adjoint wavefield because the differential operator of equation 12 does not depend on λ but on the background wavefield, u. Now equation 11 becomes:

$$\delta\chi = \int_0^T \int_{\mathbf{\Omega}} \left[\delta m_1 \lambda \partial_t^2 u + \nabla\lambda\cdot\left(\sum_{i>1}\frac{\partial S}{\partial m_i}\delta m_i\right)\nabla u - \lambda\delta f\right] dV dt. \tag{13}$$

Partial derivatives of the misfit function with respect to the model parameters can now be read from equation 13 as:

$$\frac{\partial\chi}{\partial m_1} = \int_0^T \lambda\partial_t^2 u dt, \tag{14}$$

and for $i > 1$:

$$\frac{\partial\chi}{\partial m_i} = \int_0^T \nabla\lambda\cdot\frac{\partial S}{\partial m_i}\nabla u dt. \tag{15}$$

These derivatives can be easily casted in terms of Thomsen paramaters, for example in VTI, as:

$$\frac{\partial\chi}{\partial\epsilon} = 2\frac{\partial\chi}{\partial m_2} + \frac{\partial\chi}{\partial m_3}, \tag{16}$$

and

$$\frac{\partial \chi}{\partial \delta} = -\frac{\partial \chi}{\partial m_3}. \tag{17}$$

Equation 14 shows that the gradient of the objective function with respect to the squared vertical P-wave slowness is calculated from the adjoint wavefield and the second-order time derivative of the forward wavefield, both of which are located on the main grid. Equation 15 shows that the gradients with respect to m_i with $i > 1$ or the Thomsen parameters are computed from first derivatives of the forward and adjoint wavefields, which are located on the staggered grid. These computations are consistent with the locations where these parameters are defined (Figure 2).

We have solved the forward modeling and adjoint equations for simple cases. The first one is a model with a small perturbation point at the center of a homogeneous VTI medium. There are 800 receivers at 5 meters apart placed everywhere on the surface. Figures 6a through 6d show, respectively, the model, full data, and background data; as well as the residual for one shot. The full and background data are displayed at clip of 95% to reveal the reflection from the perturbation. These data and the residual are contaminated by the high-frequency artifacts mentioned earlier. Figures 7a through 7f show snapshots of the source wavefield and adjoint wavefield. Although the adjoint wavefield is degraded by noise, it is still able to focus at the perturbation point at 0.4 second. Cross-corelation and stacking over multiple shots help to reduce the effect of this noise. Figures 8a through 8c show, respectively, gradients of slowness squares, ϵ, and δ, using 40 shots at 100 meters spacing everywhere on the surface. These gradients focus clearly at the perturbation point.

In another model, the background consists of two homogeneous VTI layers and the perturbation is a rectangular box embedded in the top layer (Figure 9). We use positive perturbations in velocity and Thomsen parameters. There are 800 receivers and 40 shots placed everywhere on the surface with similar spacing as in the previous example. Figures 10a through 10c show to the full data, background data, and the residual for one shot. Figure 11 shows the gradients of the objective function with respect to: squared vertical P-wave slowness (panel a), ϵ (panel b), and δ (panel c). These figures show that the gradients accurately identify the perturbation, and indeed, lead to the correct update directions. Similarities between the slowness squares and ϵ gradients can be observed. These similarities might indicate cross-talks between these two parameters. The δ gradient is the most accurate in identifying the perturbation's shape but is the weakeast in terms of magnitude relative to the other two gradients. The small magnitude of the δ gradient is an indication of limited ability to invert for δ from surface reflection data. Additionally, notice the V-shaped sensitivity regions of the shots that are located on the edges of the model.

We carried out 100 inversion iterations in the time domain using a Ricker wavelet of 40-Hertz (Hz) fundamental frequency. We employ the steepest decent algorithm and calculate the step length from a parabolic interpolation in the search direction (Vigh and Starr, 2008). Figure 12 shows the model updates after the first iteration (left column) and after 100 iterations (right column) in velocity (top row), ϵ (middle

row), and δ (bottom row). Improvements in resolving the magnitudes and shape of the perturbations are noticeable after 100 inversion iterations compared to the first-iteration results. Figure 13 shows the objective function decreasing with iterations.

For orthorhombic media, we only computed the sensitivity kernels for one shot and one receiver in different directions in a homogeneous subsurface. These kernels are shown in Figures 14, 15, and 16 for sources and receivers separated in the x-, y-, and z-directions, respectively. It can be noted from these figures that the sensitivity kernels of slowness squares are always nonzero, regardless of the source-receiver's direction. On the other hands, ϵ_i and δ_i are insensitive in certain planes when the source and receiver are apart in certain directions. This difference between the sensitivity kernels of slowness squares and Thomsen paramaters might be because slowness or velocity has first-order influence on seismic signature in anisotropic media in comparison to Thomsen parameters. The insensitivity of Thomsen parameters in certain planes, however, will not affect the ability to invert for these parameters when there are more sources and receivers covering wider apertures and azimuths. It can also be observed that the planes of insensitivity interchange between ϵ_i and between δ_i when the source and receiver are separated in different directions. For example, ϵ_1 is insensitive in the xz-plane when the source and receiver are apart in the x-direction; whereas, ϵ_2 is insensitive in the yz-plane when the source and receiver are apart in the y-direction. This interchangeability is consistent with how these parameters are defined: ϵ_1 is the Thomsen parameter in the yz-plane; whereas, ϵ_2 is the equivalent parameter in the xz-plane. By these definitions, the two parameters' roles are interchangeable.

CONCLUSIONS

With numerical examples for VTI and orthorhombic media, I have shown that the nonlinear pseudo-acoustic wave equation in anisotropic media can be modeled, linearized, and adjointed. The forward modeling equation can be solved accurately using RSFD. Its nonlinearity introduces an additional term in the linearized and adjoint equations. Together, these three equations can be used in an inversion framework to estimate medium parameters.

APPENDIX A

In this appendix, I derive the linearized equation 5. The perturbed wave equation is:

$$(m_1 + \delta m_1)\partial_t^2(u + \delta u) - \nabla \cdot [(S + \delta S)\nabla(u + \delta u)] = f. \tag{A-1}$$

Neglect the second-order terms:

$$m_1\partial_t^2\delta u - \nabla \cdot (S\nabla\delta u) = -\partial_t^2 u\delta m_1 + \nabla \cdot (\delta S\nabla u). \tag{A-2}$$

After the plane-wave approximation (equation 2), $S = S(\nabla u, m_i)$ with $i > 1$.

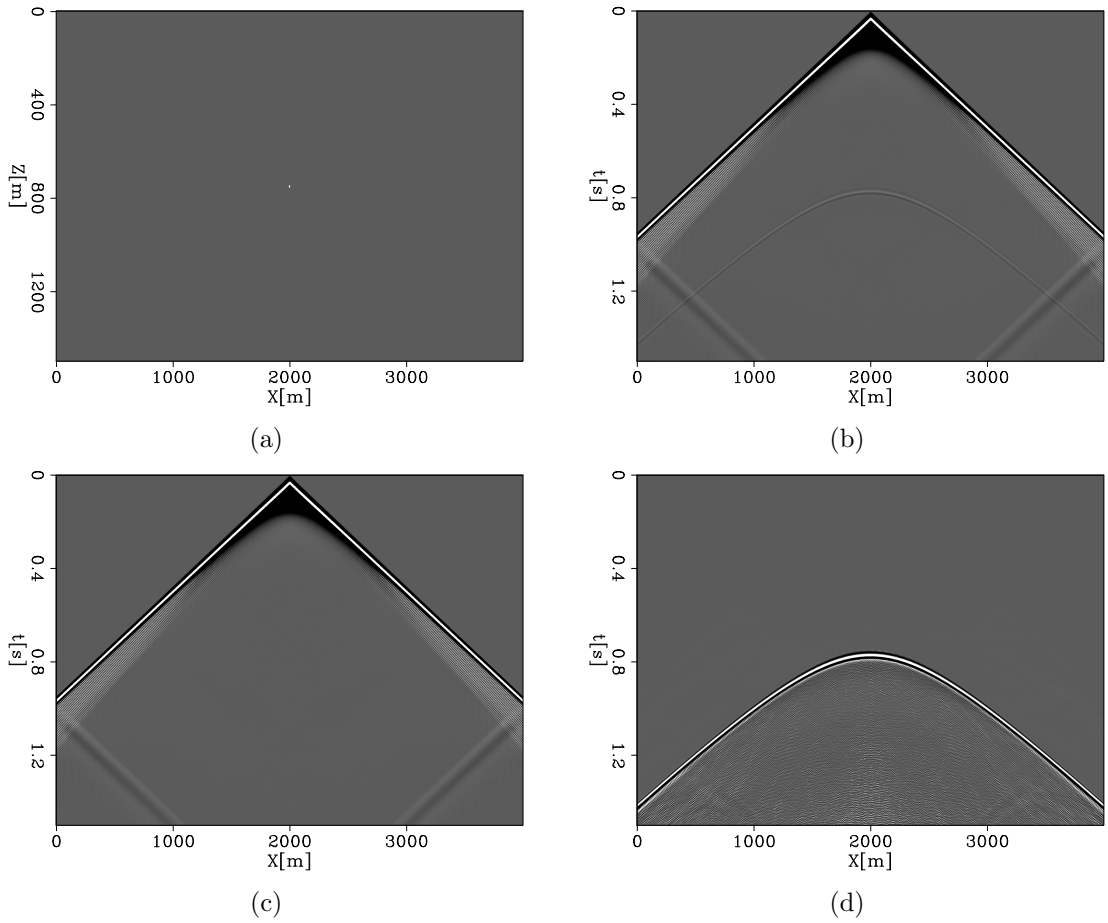


Figure 6: (a) Model with a small perturbation point in a homogeneous VTI medium, (b) full data, (c) background data, and (d) residual. The data, displayed at clip of 95%, and the residual are contaminated by high-frequency noise. [**ER**]

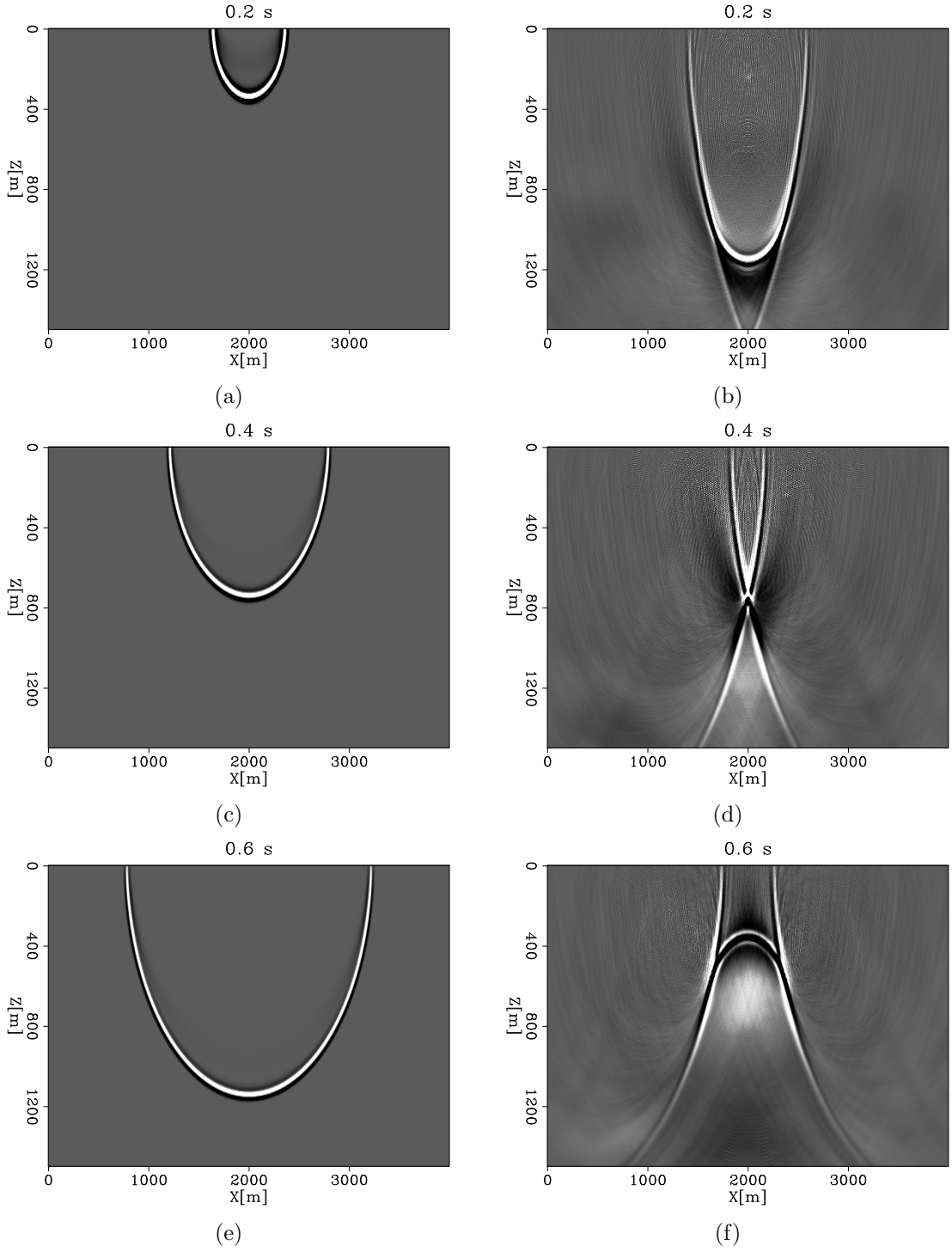


Figure 7: Snapshots of the source wavefield (left column) and adjoint wavefield (right column) at different times: (a) and (b) 0.2 second, (c) and (d) 0.4 second, and (e) and (f) 0.6 second. Although the adjoint wavefield is severely degraded by high-frequency noise, it is still able to focus at the perturbation point at 0.4 second (panel d). [**CR**]

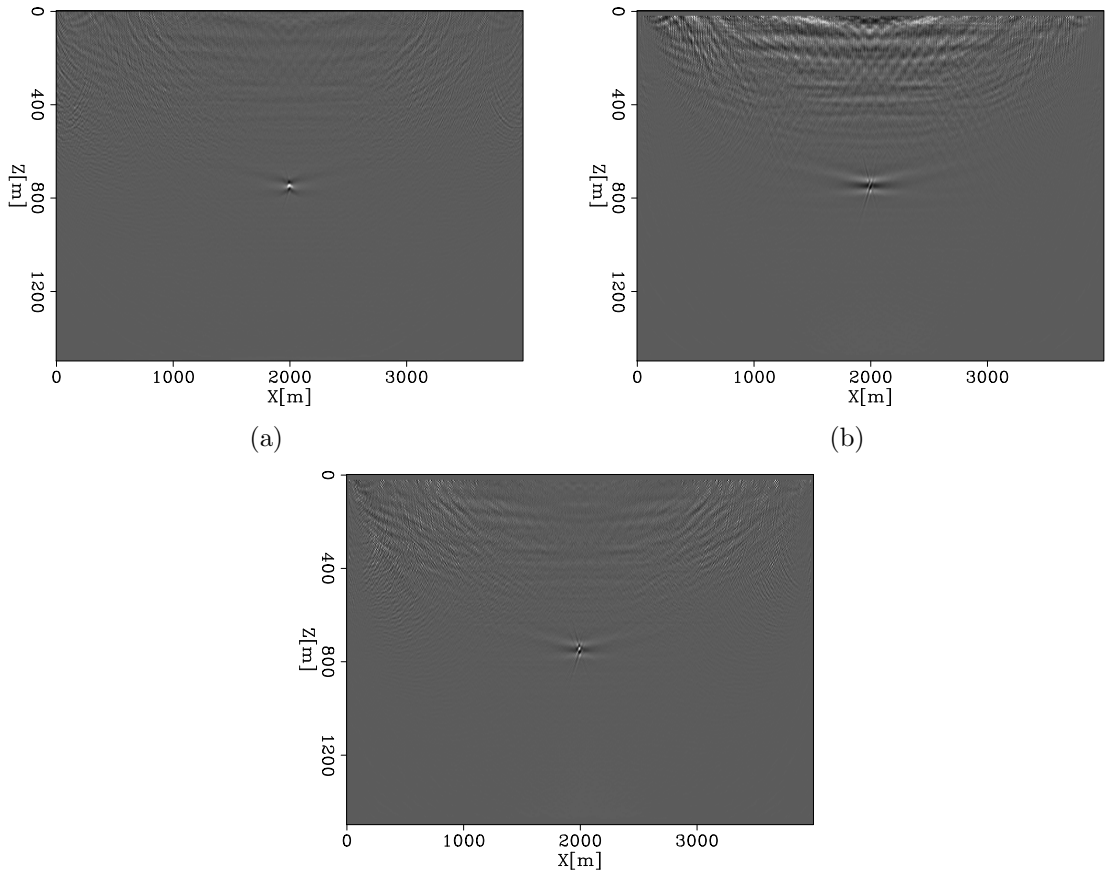


Figure 8: Gradients of the objective function with respect to: (a) squared vertical P-wave slowness, (b) ϵ, and (c) δ, for a simple model with a point perturbation in a homogeneous VTI medium. These gradients focus at the perturbation points. [**CR**]

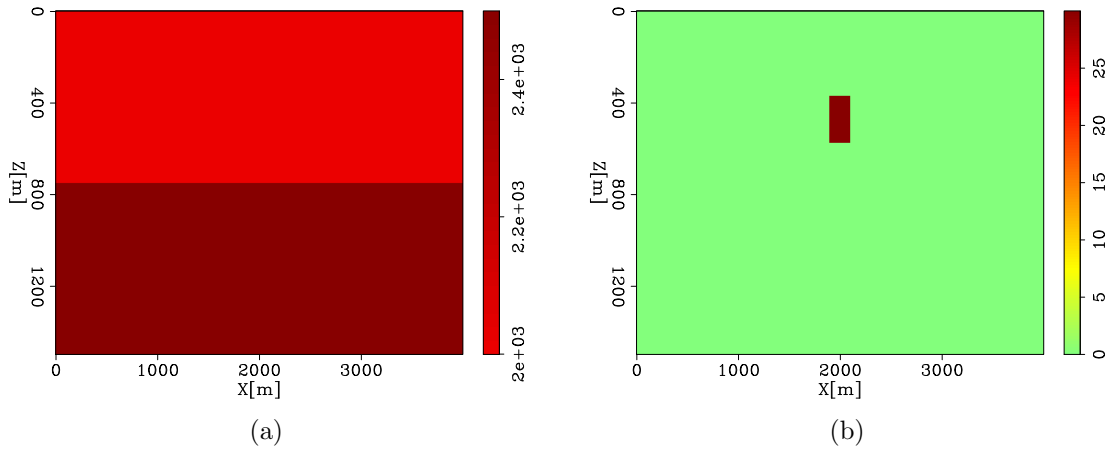


Figure 9: Model for testing gradient calculation: (a) background and (b) perturbed vertical P-wave velocity. Models for Thomsen parameters are similar. [**ER**]

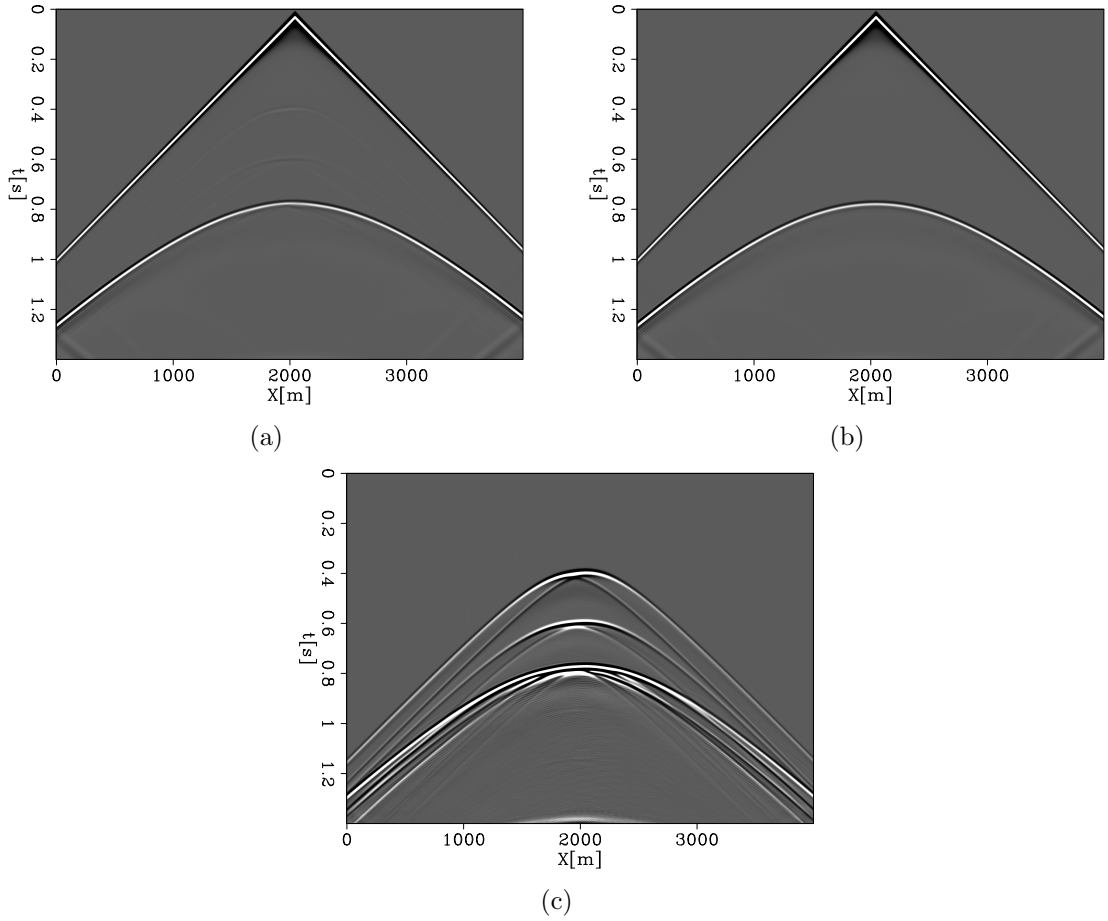


Figure 10: (a) Full data, (b) background data, and (c) residual for a model with two VTI layers and a rectangular perturbation. [**ER**]

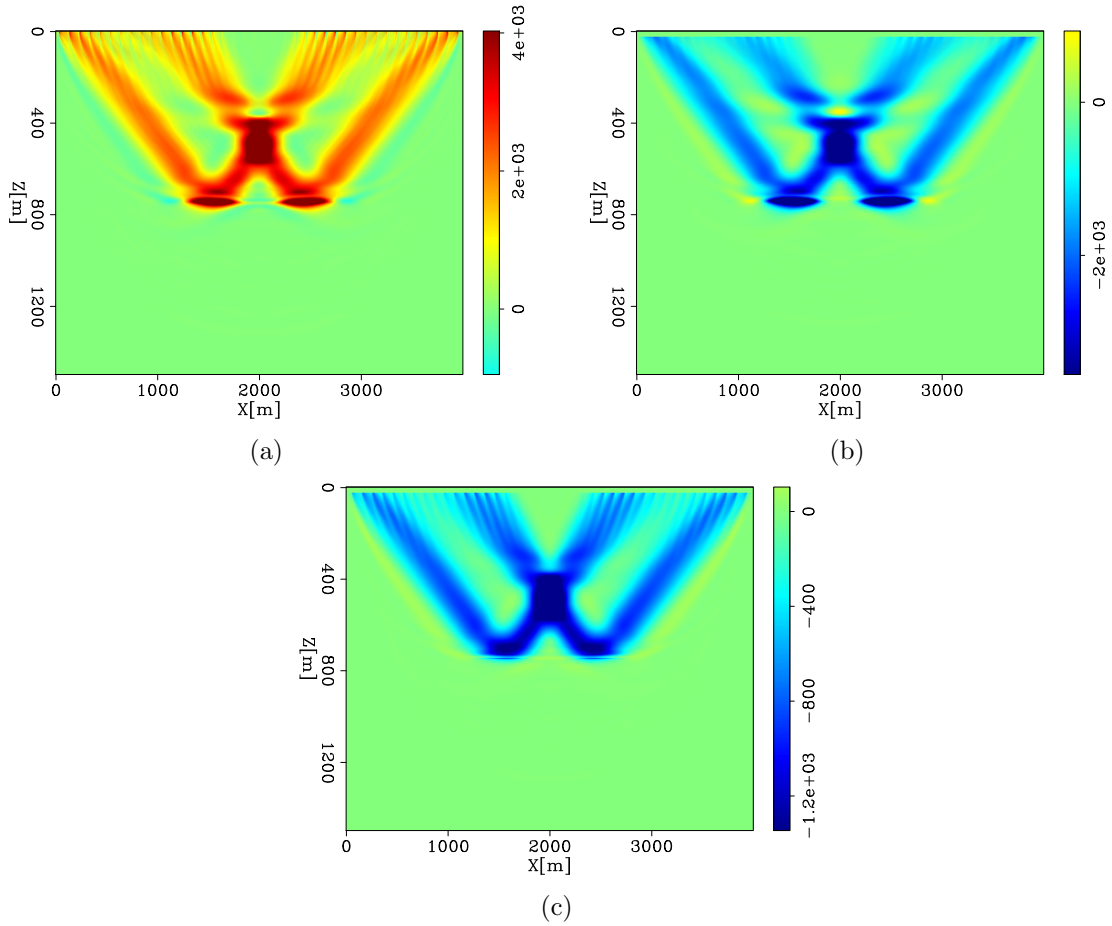


Figure 11: Gradients of the objective function with respect to: (a) squared vertical P-wave slowness, (b) ϵ, and (c) δ, for a model with two VTI layers and a rectangular perturbation. These gradients accurately locate the perturbation and have correct update directions. [**CR**]

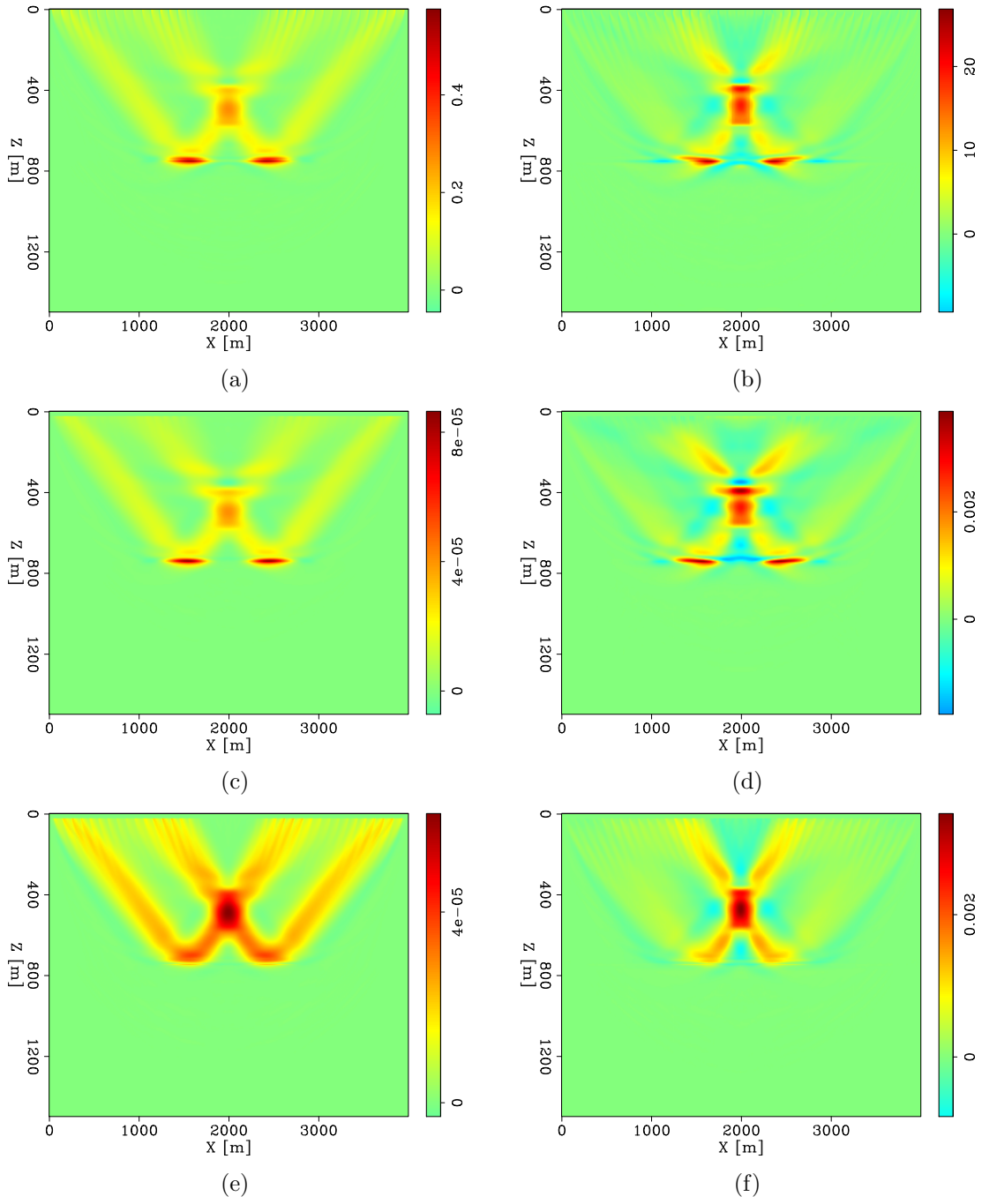


Figure 12: Model updates after the first iteration (left column) and after 100 iterations (right column) in: velocity (m/s) (top row), ϵ (middle row), and δ (bottom row). The updates show improvements after 100 iterations compared to the first-iteration updates. [**CR**]

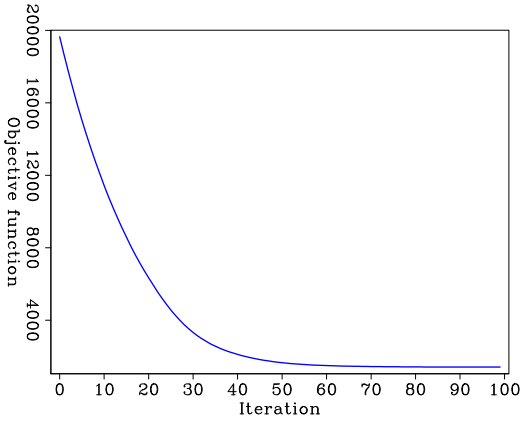


Figure 13: Plot of the objective function with iteration. [**CR**]

Consequently, to the first order:

$$\delta S = \left(\frac{\partial S}{\partial \nabla u}\right)^T \nabla \delta u + \sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i, \tag{A-3}$$

which is subtituted into equation A-2 to obtain:

$$m_1 \partial_t^2 \delta u - \nabla \cdot (S \nabla \delta u) = -\partial_t^2 u \delta m_1 + \nabla \cdot \left\{ \left[\left(\frac{\partial S}{\partial \nabla u}\right)^T \nabla \delta u + \sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i \right] \nabla u \right\}.$$

Collecting the terms that involve δu to obtain the linearized equation 5:

$$m_1 \partial_t^2 \delta u - \nabla \cdot \left\{ \left[S\mathbf{I} + \nabla u \left(\frac{\partial S}{\partial \nabla u}\right)^T \right] \nabla \delta u \right\} = -\partial_t^2 u \delta m_1 + \nabla \cdot \left[\left(\sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i \right) \nabla u \right].$$

In VTI (Xu and Zhou, 2014):

$$S = \frac{1}{2} \left[n_a^2 + \sqrt{n_a^4 - 8\left(\epsilon - \delta\right) n_x^2 n_z^2} \right], \tag{A-4}$$

where n_x and n_z are components of the normalized wave vector, $\mathbf{n}$, and $n_a^2 = (1 + 2\epsilon)\, n_x^2 + n_z^2$.

In orthorhombic media (Song and Alkhalifah, 2013):

$$S = \frac{1}{3} \left[a - \frac{d}{\sqrt[3]{2}} - \frac{\sqrt[3]{2}\left(a^2 + 3b\right)}{d} \right], \tag{A-5}$$

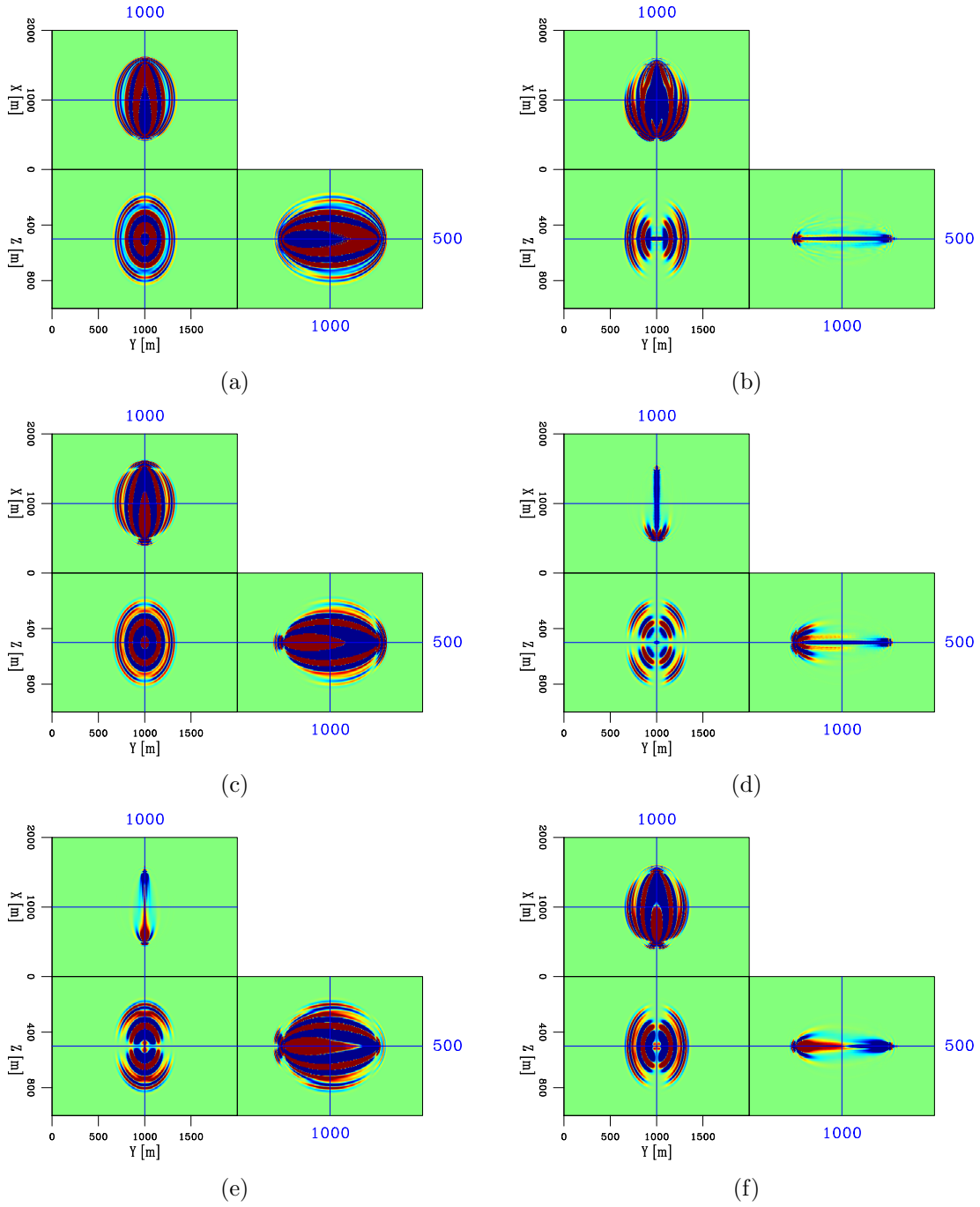


Figure 14: Sensitivity kernels for one shot and one receiver, separated in the x-direction, in a homogeneous orthorhombic subsurface of: (a) squared vertical P-wave slowness, (b) ϵ_1, (c) ϵ_2, (d) δ_1, (e) δ_2, and (f) δ_3. [**CR**]

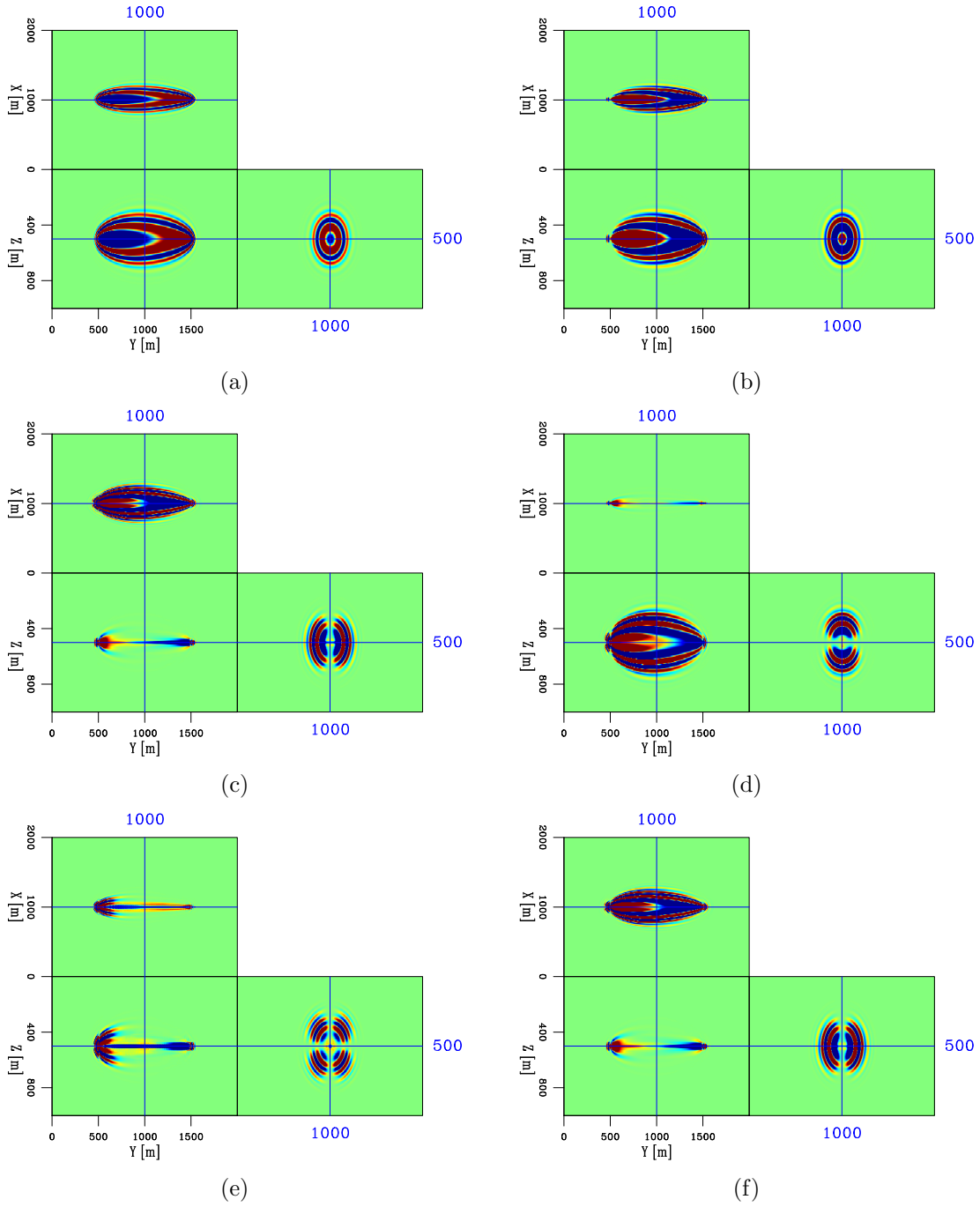


Figure 15: Sensitivity kernels for one shot and one receiver, separated in the y-direction, in a homogeneous orthorhombic subsurface of: (a) squared vertical P-wave slowness, (b) ϵ_1, (c) ϵ_2, (d) δ_1, (e) δ_2, and (f) δ_3. [**CR**]

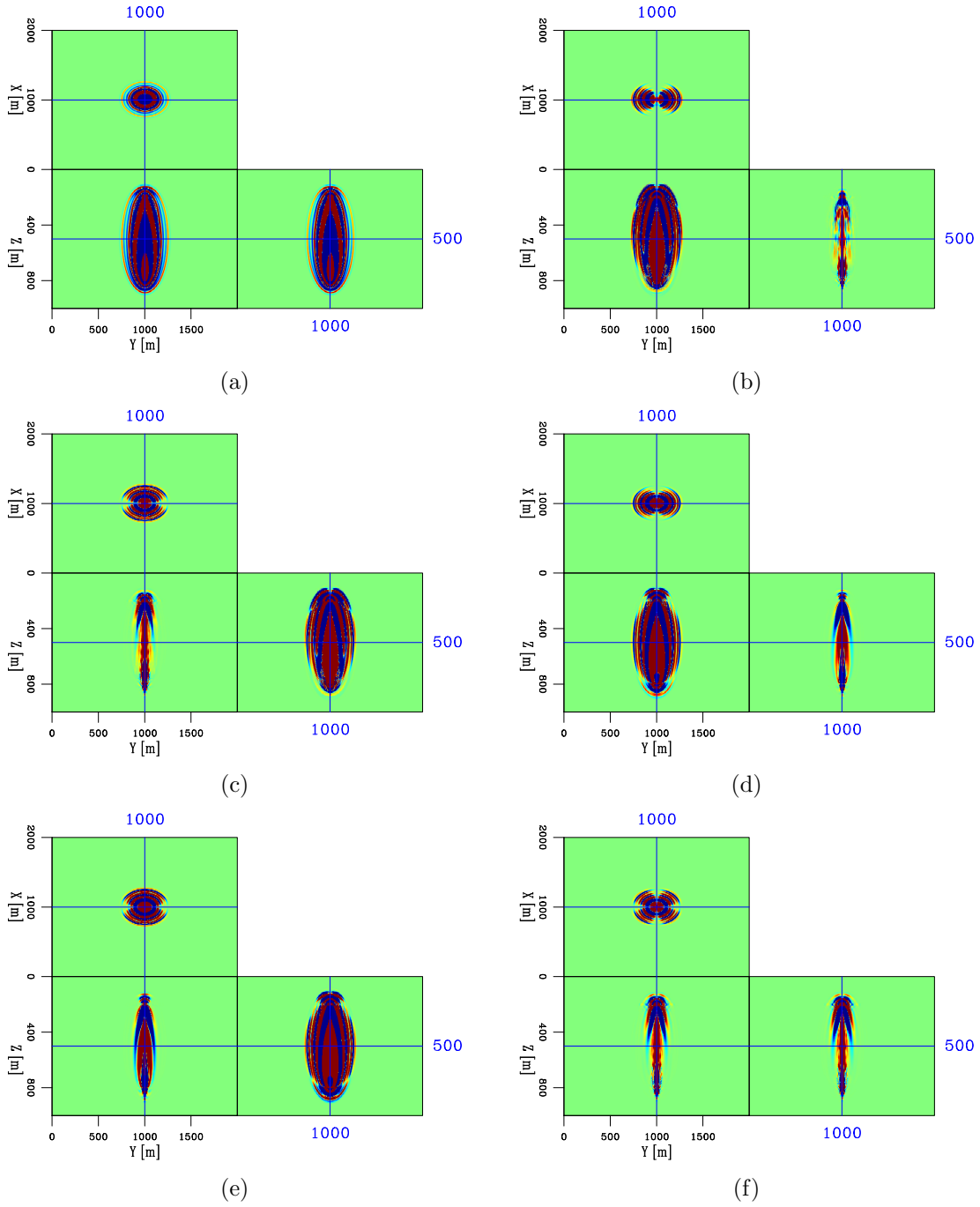


Figure 16: Sensitivity kernels for one shot and one receiver, separated in the z-direction, in a homogeneous orthorhombic subsurface of: (a) squared vertical P-wave slowness, (b) ϵ_1, (c) ϵ_2, (d) δ_1, (e) δ_2, and (f) δ_3. [**CR**]

where:

$$a = (2\epsilon_2 + 1)\, n_x^2 + (2\epsilon_1 + 1)\, n_y^2 + n_z^2,$$
$$b = \left[(2\epsilon_2 + 1)^2 (2\delta_3 + 1) - (2\epsilon_2 + 1)(2\epsilon_1 + 1)\right] n_x^2 n_y^2 - 2(\epsilon_1 - \delta_1)\, n_y^2 n_z^2 - 2(\epsilon_2 - \delta_2)\, n_x^2 n_z^2,$$
$$d = \sqrt[3]{-2a^3 + 3(e - 9c) - 9ab},$$
$$c = \left[2(2\epsilon_2 + 1)\sqrt{(2\delta_1 + 1)(2\delta_2 + 1)(2\delta_3 + 1)} - (2\epsilon_2 + 1)^2 (2\delta_3 + 1) + \epsilon_1\epsilon_2 - \epsilon_1\delta_2 - \epsilon_2\delta_1\right] n_x^2 n_y^2 n_z^2,$$
$$e = \sqrt{3b^2(a^2 + 4b) - 6ac(2a^2 + 9b) - 81c^2}.$$

APPENDIX B

In this appendix, I present a step-by-step derivation of equation 11. Substitute perturbation in S (equation A-3) into equation 10 to obtain:

$$\begin{aligned}\delta\chi = &\int_0^T \int_{\mathbf{\Omega}} \sum_r \left[u(\mathbf{x}_r, t) - d(\mathbf{x}_r, t)\right] \delta(\mathbf{x} - \mathbf{x}_r)\delta u dV dt \\ &+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left\{ \delta m_1 \partial_t^2 u - \nabla \cdot \left[\left(\sum_{i>1} \frac{\partial S}{\partial m_i} \delta m_i \right) \nabla u \right] - \delta f \right\} dV dt \qquad \text{(B-1)} \\ &+ \int_0^T \int_{\mathbf{\Omega}} \lambda \left[m_1 \partial_t^2 \delta u - \nabla \cdot (\mathbf{A} \nabla \delta u) \right] dV dt.\end{aligned}$$

Integration by parts gives, for $i > 1$:

$$\begin{aligned}\int_0^T \int_{\mathbf{\Omega}} \lambda \nabla \cdot \left(\frac{\partial S}{\partial m_i} \delta m_i \nabla u \right) dV dt = &\int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \left(\lambda \frac{\partial S}{\partial m_i} \delta m_i \nabla u \right) ds dt \\ &- \int_0^T \int_{\mathbf{\Omega}} \nabla \lambda \cdot \left(\frac{\partial S}{\partial m_i} \delta m_i \nabla u \right) dV dt,\end{aligned}$$

$$\begin{aligned}\int_0^T \int_{\mathbf{\Omega}} \lambda m_1 \partial_t^2 \delta u dV dt &= \int_{\mathbf{\Omega}} \lambda m_1 \partial_t \delta u |_0^T dV - \int_0^T \int_{\mathbf{\Omega}} m_1 \partial_t \lambda \partial_t \delta u dV dt \\ &= \int_{\mathbf{\Omega}} m_1 \left(\lambda \partial_t \delta u - \partial_t \lambda \delta u \right) |_0^T dV + \int_0^T \int_{\mathbf{\Omega}} m_1 \partial_t^2 \lambda \delta u dV dt \\ &= \int_{\mathbf{\Omega}} m_1 \left(\lambda \partial_t \delta u - \partial_t \lambda \delta u \right) |_T dV + \int_0^T \int_{\mathbf{\Omega}} m_1 \partial_t^2 \lambda \delta u dV dt,\end{aligned}$$

and:

$$\begin{aligned}
\int_0^T \int_{\mathbf{\Omega}} \lambda \nabla \cdot (\mathbf{A} \nabla \delta u) dV dt &= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \lambda \mathbf{A} \nabla \delta u ds dt - \int_0^T \int_{\mathbf{\Omega}} \nabla \lambda \cdot (\mathbf{A} \nabla \delta u) dV dt \\
&= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \lambda \mathbf{A} \nabla \delta u ds dt - \int_0^T \int_{\mathbf{\Omega}} (\mathbf{A}^T \nabla \lambda) \cdot \nabla \delta u dV dt \\
&= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \left(\lambda \mathbf{A} \nabla \delta u - \mathbf{A}^T \nabla \lambda \delta u \right) ds dt \\
&\quad + \int_0^T \int_{\mathbf{\Omega}} \nabla \cdot \left(\mathbf{A}^T \nabla \lambda \right) \delta u dV dt \\
&= \int_0^T \oint_{\partial \mathbf{\Omega}} \hat{\mathbf{n}} \cdot \lambda \mathbf{A} \nabla \delta u ds dt + \int_0^T \int_{\mathbf{\Omega}} \nabla \cdot \left(\mathbf{A}^T \nabla \lambda \right) \delta u dV dt.
\end{aligned}$$

where I have exploited the perturbed initial and boundary conditions (equations 5b and 5c). Substituting the above three integrations into equation B-1 results in equation 11.

REFERENCES

Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.

Le, H. and S. A. Levin, 2014, Removing shear artifacts in acoustic wave propagation in orthorhombic media: SEG Annual International Meeting, Expanded Abstracts, 486–490.

Liu, Q. and J. Tromp, 2006, Finite-frequency kernels based on adjoint methods: Bulletin of the Seismological Society of America, **96**, 2383–2397.

Ozdenvar, T. and G. A. McMechan, 1996, Causes and reduction of numerical artefacts in pseudo-spectral wavefield extrapolation: Geophysical Journal International, **126**, 819–828.

Saenger, E. H., N. Gold, and S. A. Shapiro, 2000, Modeling the propagation of elastic waves using a modified finite-difference grid: Wave Motion, **31**, 77–92.

Song, X. and T. Alkhalifah, 2013, Modeling of pseudoacoustic p-waves in orthorhombic media with a low-rak approximation: Geophysics, **78**, no. 4, C33–C40.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Vigh, D. and E. W. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, no. 5, VE135–VE144.

Xu, S. and H. Zhou, 2014, Accurate simulations of pure quasi-P-waves in complex anisotropic media: Geophysics, **79**, no. 6, T341–T348.