Elastic wavefield tomography with physical model constraints

Yuting Duan & Paul Sava
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT
We propose elastic isotropic wavefield tomography formulated with the misfit of observed and recorded data as the objective function. The adjoint-state method enables efficient calculation of the gradient of this data misfit function. We solve this optimization problem with the gradient-descent method in order to update the model parameters iteratively. Using the elastic isotropic wave-equation formulated for slowly varying media, we invert for the squared velocities of compressional (P) and shear (S) modes. Physical relationships between P- and S-velocities, which are assumed to be generally linear, can be derived from laboratory measurements, well logs, and seismic data. Poor illumination of P- and S-waves often prevents reliable update of the model parameters, while preserving their intrinsic physical relationships. Thus, we introduce a model constraint term in the objective function, which sets the ratio of P- and S-wave velocities to a chosen range, assumed to be generally linear as suggested by laboratory measurements or well logs. Examples demonstrate that this constraint yields models that are more physically plausible than models obtained using only the data misfit as the objective function.

Key words: elastic wavefield, tomography, physical constraint

1 INTRODUCTION

Seismic tomography is a commonly used tool for building models of subsurface parameters from recorded seismic data. The various approaches to seismic tomography generally fall into two categories. The first is traveltime tomography, which seeks to match the traveltimes of recorded seismic data with those of synthetic seismic data simulated in a trial model by using ray tracing (Zhu et al., 1992; Zelt et al., 2006; Taillandier et al., 2009) or by solving a wave-equation (Woodward, 1992; Schuster and Quintus-Bosz, 1993). However, models computed with traveltime tomography tend to be less accurate when the subsurface contains complicated geologic structures, which results in wavefield multipathing. The second category consists of methods for wavefield tomography, which use both the amplitude and phase information of recorded seismic data to invert for the subsurface parameter model. Although more expensive than traveltime tomography, wavefield tomography is more effective in recovering model parameters that are sensitive to waveform amplitudes (Tarantola, 1986; Mora, 1988; Pratt, 1999; Prieux et al., 2010), and has greater capability to accurately recover parameters at large depths (Bae et al., 2012; McNeely et al., 2012).

However, one issue with wavefield tomography is its non-uniqueness, i.e., multiple models that fit the data equally well. Ivanov et al. (2005) summarize several factors that contribute to non-uniqueness, for example, insufficient acquisition and the presence of noise in the data. Methods for regularization have been proposed to stabilize the results and reduce the number of possible solutions (Tikhonov and Arsenin, 1977; Sun and Shuster, 1992; Zhou et al., 2002). Regularization is defined in the model domain, and has the purpose of adding prior information to the model. Regularization can be formulated in many different, possibly complementary, ways. Tikhonov and Arsenin (1977) propose a regularization term that encourages smoothness in the computed model. Xiang and Zhang (2014) use an edge-preserving regularization term in order to preserve sharp interfaces within the model. Asnaashari et al. (2013) use an a priori model to constrain the inversion. These and other examples demonstrate that proper regularization or preconditioning can effectively reduce the number of non-unique solutions and, in many cases, improve the accuracy of the reconstructed models.

Wavefield tomography is well-developed under the acoustic assumption (Tarantola, 1984; Pratt, 1999; Operto et al., 2004; Biondi and Almomin, 2014); however, recorded seismic data include shear waves in addition to the compressional waves. Because all wave modes contain useful information about the subsurface, elastic wavefield tomography may allow for better characterization of the subsurface (Tarantola, 1986; Pratt, 1990; Guasch et al., 2012; Vigh et al., 2012). There are many possible model parameterizations, which lead to different inversion schemes. For example, Tarantola (1986) shows wavefield tomography for com-

pressional (P) impedance, shear (S) impedance, and density, while Mora (1988) and Guasch et al. (2012) compute P- and S-velocity models using wavefield tomography.

Though inversion for multi-parameters adds more physical information to the updated model compared to single-parameter inversion, different parameters in the updated model may not be physically realistic (Plessix, 2006), i.e. inversion might not be able to resolve the model parameters while preserving their intrinsic physical relationships. Such physical relationships need to be enforced explicitly because their action on the inverted model differs from the alternative constraints provided by data or by shaping regularizations. Physical relationships between model parameters in elastic media can be derived from well logs, seismic data, and laboratory measurements, but they can also be derived based on first-principle physical relationships (Tsuneyama, 2006; Compton and Hale, 2013). Experiments show that the relationship between P- and S-velocities is generally linear (Castagna et al., 1985; Katahara, 1999; Tsuneyama, 2006); therefore, enforcing a range of constant ratios between the two velocities could guide the model update, thus increasing the robustness of wavefield tomography.

In this paper, we perform wavefield tomography using the isotropic elastic wave-equation, and our mathematical derivation shows that the simplest approach is to invert for the squared velocities of P- and S-waves. We propose an objective function for wavefield tomography with a logarithmic term in order to constrain the relationship of P- and S-wave velocities. The constraint term only dominants the gradient when inverted model parameters approach the boundaries of the constraints. We formulate the tomographic problem using the adjoint-state method, which enables us to efficiently derive the gradient. Examples demonstrate that the model constraint helps enforce appropriate physical relationships between the model parameters and speeds up convergence.

2 ELASTIC WAVEFIELD MODELING

We consider the elastic isotropic wave-equation, in which model parameters $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are slowly varying (Lay and Wallace, 1995):

$$\ddot{\mathbf{u}} - \alpha\nabla(\nabla\cdot\mathbf{u}) + \beta\nabla\times(\nabla\times\mathbf{u}) = \mathbf{d}\ , \tag{1}$$

where $\mathbf{u}(e,\mathbf{x},t) = [u_x \ \ u_y \ \ u_z]^T$ is the displacement vector, and $\mathbf{d}(e,\mathbf{x},t)$ is the source function. Both $\mathbf{u}$ and $\mathbf{d}$ are functions of experiment e, spatial location $\mathbf{x}$, and time t. As indicated earlier, model parameters $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are squared P-wave and S-wave velocities, respectively. Equation 1 describes a linear relationship between displacement vector $\mathbf{u}$ and the source function $\mathbf{d}$:

$$\mathbf{L}\mathbf{u} = \mathbf{d}\ , \tag{2}$$

where $\mathbf{L}$ is the elastic wave operator corresponding to equation 1, with model parameters α and β, and with adjoint $\mathbf{L}^{\mathsf{T}}$.

For the wavefield tomography problem, one can directly apply the adjoint-state method to compute the gradient (Plessix, 2006). In the case of elastic wavefield tomography, we impose a physical constraint $\boldsymbol{\mathcal{F}}_s$ corresponding to equation 2:

$$\boldsymbol{\mathcal{F}}_s = \mathbf{L}\mathbf{u}_s - \mathbf{d}_s = \mathbf{0}\ . \tag{3}$$

Equation 3 simply indicates that we compute the source wavefield $\mathbf{u}_s$ from a given source function $\mathbf{d}_s$, i.e., $\mathbf{u}_s$ is a solution to the given wave-equation.

We illustrate the elastic wavefield tomography method with the synthetic model shown in Figures 1(a) and 1(b). We consider a source located at $(0.2, 2.55)$ km and a receiver at $(4.9, 2.55)$ km. Both α (Figure 1(a)) and β (Figure 1(b)) models contain negative Gaussian anomalies embedded within constant background models. We use the constant background model as the starting model for inversion. We compute the source wavefield $\mathbf{u}_s$ using a 3 Hz Ricker wavelet as the displacement source function. Both the x- and z-components of the source wavefield $\mathbf{u}_s$ consist of P- and S-waves, as shown in Figure 2. The recorded data have two components as shown in Figure 3(a). Because the source and receiver are located at the same depth and the model parameters are slowly varying, the P wave on the x-component and the S wave is recorded primarily on the z-component because the P- and S-wave polarizations are parallel and perpendicular to the direction of wave propagation, respectively.

3 OBJECTIVE FUNCTION

For wavefield tomography, one typically updates a model by minimizing an objective function. We consider an objective function $\mathcal{J}(\mathbf{u}_s,\alpha,\beta)$ consisting of three terms: a data term $\mathcal{J}_D(\mathbf{u}_s,\alpha,\beta)$ that quantifies the data misfit, a regularization term $\mathcal{J}_M(\alpha,\beta)$ that allows for the use of prior models $\bar{\alpha}(\mathbf{x})$ and $\bar{\beta}(\mathbf{x})$ as well as model shaping, and a constraint term $\mathcal{J}_C(\alpha,\beta)$ which restricts the relationship of α and β within a feasible region:

$$\mathcal{J}(\mathbf{u}_s,\alpha,\beta) = \mathcal{J}_D(\mathbf{u}_s,\alpha,\beta) + \mathcal{J}_M(\alpha,\beta) + \mathcal{J}_C(\alpha,\beta)\ . \tag{4}$$

We define the data misfit term using the difference between the predicted and observed data:

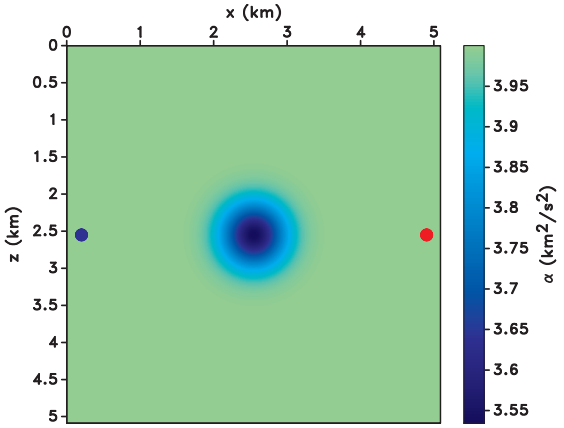
$$\mathcal{J}_D = \sum_e \frac{1}{2}\|\mathbf{W}_u\mathbf{u}_s - \mathbf{d}_o\|^2\ , \tag{5}$$

where $\mathbf{d}_o(e,\mathbf{x},t)$ are the observed data recorded at the receiver locations and weights $\mathbf{W}_u(e,\mathbf{x},t)$ restrict the source wavefield $\mathbf{u}_s(e,\mathbf{x},t)$ to the receiver locations.

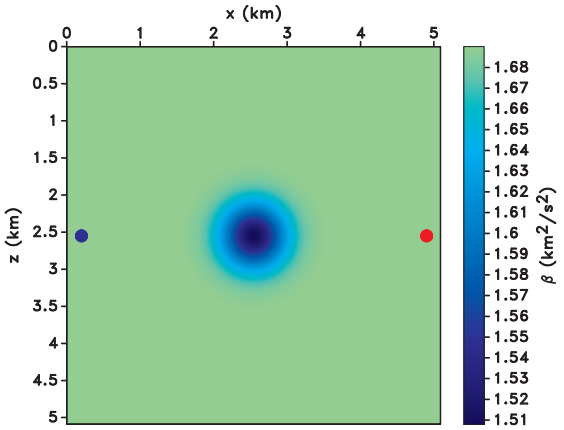
The regularization term $\mathcal{J}_M$ in the objective function contains prior models $\bar{\alpha}(\mathbf{x})$ and $\bar{\beta}(\mathbf{x})$ as reference:

$$\mathcal{J}_M = \frac{1}{2}\|\mathbf{W}_\alpha(\alpha-\bar{\alpha})\|^2 + \frac{1}{2}\|\mathbf{W}_\beta(\beta-\bar{\beta})\|^2\ , \tag{6}$$

where $\mathbf{W}_\alpha(\mathbf{x})$ and $\mathbf{W}_\beta(\mathbf{x})$ are model shaping operators, whose inverses are related to the model covariance matrices. Such operators can describe either simple and space-invariant smoothing, e.g. by a Laplacian filter, or they can describe



(a)



(b)

Figure 1. 2D synthetic (a) α and (b) b models with a negative Gaussian velocity anomaly with constant background model. The source is at $(0.2, 2.55)$ km and the receiver is at $(4.9, 2.55)$ km.

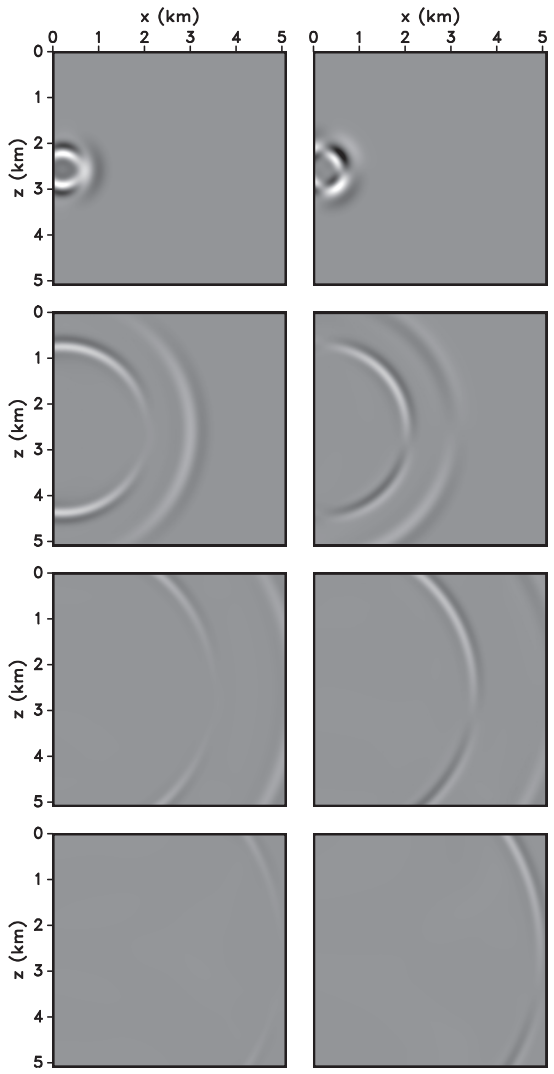


Figure 2. From top to bottom: snapshots in chronological order of the source wavefield $\mathbf{u}_s$. The left and right columns are the x- and z-components of the wavefield, respectively.

image-shaping with non-stationary filtering (Hale, 2007; Guitton et al., 2010).

Due to differences in illumination and amplitude between P and S modes, models α and β computed using physical wavefield tomography update independently, and this may lead to unphysical solutions. One can constrain model updating using physical relationships between α and β within a region, for example defined using an upper boundary $h_u(\alpha, \beta) = 0$ and a lower boundary $h_l(\alpha, \beta) = 0$:

$$h_u > 0 \,, \tag{7}$$

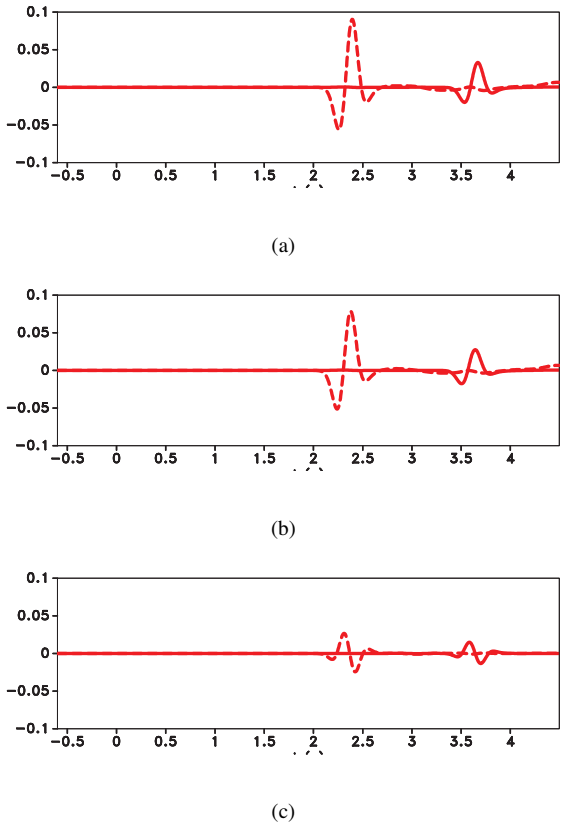
$$h_l > 0 \,. \tag{8}$$

In order to keep the updated model within the boundaries, we include in the objective function a constraint term $\mathcal{J}_C$ that uses a logarithmic penalty function (Peng et al., 2002; Nocedal and Wright, 2006; Gasso et al., 2009):

$$\mathcal{J}_C = -\eta \sum_{\mathbf{x}} \left[\log(h_u) + \log(h_l)\right] \,. \tag{9}$$

The term $\mathcal{J}_C$ relies on the fact that $\mathcal{J}_C$ tends to $-\infty$ as h_u or h_l tends to 0, i.e., term $\mathcal{J}_C$ penalizes violations of inequalities 7 and 8 . The parameter η weighs the constraint term $\mathcal{J}_C$ against other terms in the objective function $\mathcal{J}$. The starting model is required to be in-between the boundaries, and this constraint term mainly contributes to model updating when the updated model approaches the boundaries.

For elastic wavefield inversion, because the relationship of P- and S-velocities is generally linear (Castagna et al., 1985; Zimmer et al., 2002; Rojas et al., 2005), we set the boundaries



(a)

(b)

(c)

Figure 3. (a) Observed data, (b) predicted data, and (c) data difference. The dashed and solid lines are the x- and z-components of the data, respectively.

to be

$$h_u = -\sqrt{\alpha} + c_u\sqrt{\beta} + b_u = 0\ , \tag{10}$$

$$h_l = \quad \sqrt{\alpha} - c_l\sqrt{\beta} - b_l = 0\ , \tag{11}$$

where the user-defined parameters c_l, c_u, b_l, and b_u characterize the specific boundaries.

4 OBJECTIVE FUNCTION GRADIENT

For model optimization, we use a gradient descent method (Lailly, 1983; Tarantola, 1984) and minimize the objective function $\mathcal{J}$ to iteratively update the model. The gradient of the objective function can be computed efficiently using the adjoint-state method (Plessix, 2006). The augmented functional $\mathcal{H}$ is defined as

$$\mathcal{H} = \mathcal{J} - \sum_e \boldsymbol{\mathcal{F}}^{\mathsf{T}}\mathbf{a}_s\ . \tag{12}$$

The gradient of $\mathcal{H}$ indicates the searching direction toward minimum of $\mathcal{J}$, constrained by $\boldsymbol{\mathcal{F}}_s = 0$. Vector $\mathbf{a}_s\,(e,\mathbf{x},t)$ is the adjoint variable and is computed by solving the adjoint equations obtained by setting the partial derivatives of the augmented functional, relative to the state variables $\mathbf{u}_s$, to zero:

$$\mathbf{L}^{\mathsf{T}}\mathbf{a}_s = \mathbf{g}_s\ . \tag{13}$$

We compute the adjoint sources $\mathbf{g}_s\,(e,\mathbf{x},t)$, which depend on the objective function $\mathcal{J}$:

$$\mathbf{g}_s = \frac{\partial\mathcal{J}}{\partial\mathbf{u}_s} = \mathbf{W}_u^{\mathsf{T}}\left(\mathbf{W}_u\mathbf{u}_s - \mathbf{d}_o\right). \tag{14}$$

Because terms $\mathcal{J}_M$ and $\mathcal{J}_C$ are not functions of $\mathbf{u}_s$, they do not contribute to the adjoint source.

The gradient of the augmented functional $\mathcal{H}$ is given by

$$\begin{bmatrix}\frac{\partial\mathcal{H}}{\partial\alpha}\\ \frac{\partial\mathcal{H}}{\partial\beta}\end{bmatrix} = \begin{bmatrix}\frac{\partial\mathcal{J}}{\partial\alpha}\\ \frac{\partial\mathcal{J}}{\partial\beta}\end{bmatrix} - \sum_e \begin{bmatrix}\frac{\partial\boldsymbol{\mathcal{F}}_s^{\mathsf{T}}}{\partial\alpha}\mathbf{a}_s\\ \frac{\partial\boldsymbol{\mathcal{F}}_s^{\mathsf{T}}}{\partial\beta}\mathbf{a}_s\end{bmatrix}\ . \tag{15}$$

We obtain the gradient of $\mathcal{J}_D$ by solving $\mathcal{H} = 0$:

$$\begin{bmatrix}\frac{\partial\mathcal{J}_D}{\partial\alpha}\\ \frac{\partial\mathcal{J}_D}{\partial\beta}\end{bmatrix} = \sum_e \begin{bmatrix}\frac{\partial\boldsymbol{\mathcal{F}}_s^{\mathsf{T}}}{\partial\alpha}\mathbf{a}_s\\ \frac{\partial\boldsymbol{\mathcal{F}}_s^{\mathsf{T}}}{\partial\beta}\mathbf{a}_s\end{bmatrix} = \sum_e \begin{bmatrix}-[\nabla(\nabla\cdot\mathbf{u}_s)]^T\star\mathbf{a}_s\\ [\nabla\times(\nabla\times\mathbf{u}_s)]^T\star\mathbf{a}_s\end{bmatrix}\ , \tag{16}$$

where the symbol $\star$ denotes zero-lag crosscorrelation.

We illustrate the components of the method with a simple synthetic model. In this model, the adjoint source is the difference, Figure 3(c), between the predicted (Figure 3(b)) and observed data Figure 3(a). Note that both P and S waves in the adjoint source generate additional P and S waves in the adjoint variable (Figure 4). However, only the modes generated from the same wave mode in the recorded data provide effective contribution to their corresponding gradient. The other modes (*fake* modes) in the adjoint variable correlate with the state variable and generate artifacts as seen in Figures 5(a) and 5(b). Note that the artifacts could be attenuated with better coverage of sources and receivers. To improve the model illumination, we design sources and receivers to surround the model, thus depicting an ideal acquisition in a best-case scenario. The source and receivers for each experiment are at opposite sides of the velocity anomalies so that the receivers record transmissions that travel through the model. In this experiment we use one source and 46 receivers, and the α and β models are shown in Figures 6(a) and 6(b), respectively, while the gradients for α and β are shown in Figures 7(a) and 7(b). The artifacts near the receiver locations are weaker than those in the case with one receiver. The gradient for β is weak in the lower part because the shear wave in the recorded data changes polarity at the receiver located around $(4.8, 3.5)$ km. If we increase the number of sources and then stack the gradient over experiments, we attenuate artifacts around source locations. Figures 8(a) and 8(b) show the stacked gradient over 20 experiments, sampled every 18°, for α and β, respectively. The artifacts around source and receiver positions, which are caused by the *fake* modes, are attenuated.

The gradient of terms $\mathcal{J}_M$ and $\mathcal{J}_C$, with respect to the model parameters α and β, are

$$\begin{bmatrix}\frac{\partial\mathcal{J}_M}{\partial\alpha}\\ \frac{\partial\mathcal{J}_M}{\partial\beta}\end{bmatrix} = \begin{bmatrix}\mathbf{W}_\alpha^{\mathsf{T}}\mathbf{W}_\alpha\left(\alpha-\bar{\alpha}\right)\\ \mathbf{W}_\beta^{\mathsf{T}}\mathbf{W}_\beta\left(\beta-\bar{\beta}\right)\end{bmatrix}\ , \tag{17}$$

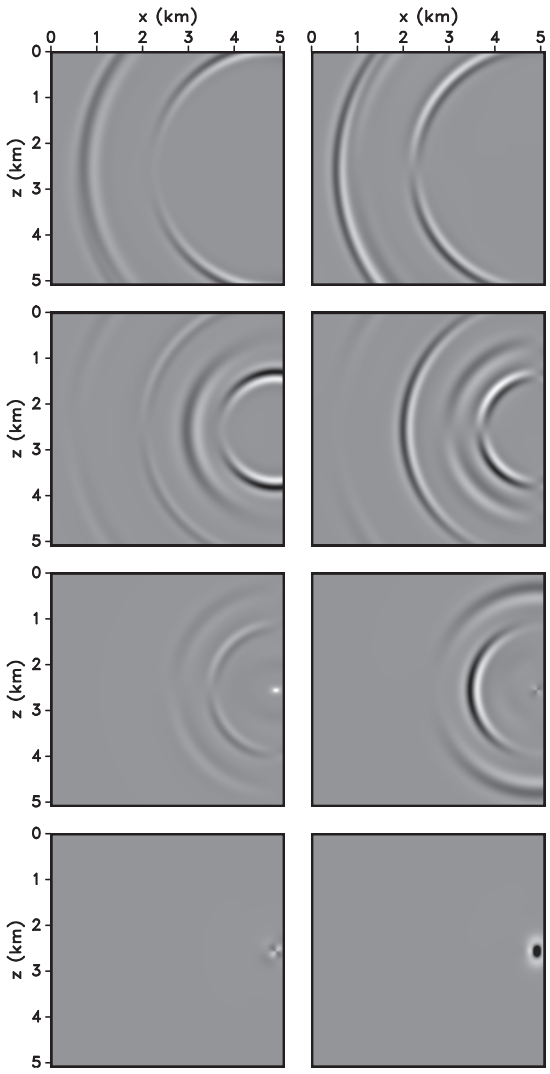


Figure 4. From top to bottom: snapshots in chronological order of the adjoint variable $\mathbf{a}_s$. The left and right columns are the x- and z-components of the wavefield, respectively. The snapshots are taken at the same time step as the corresponding panels in Figure 2.

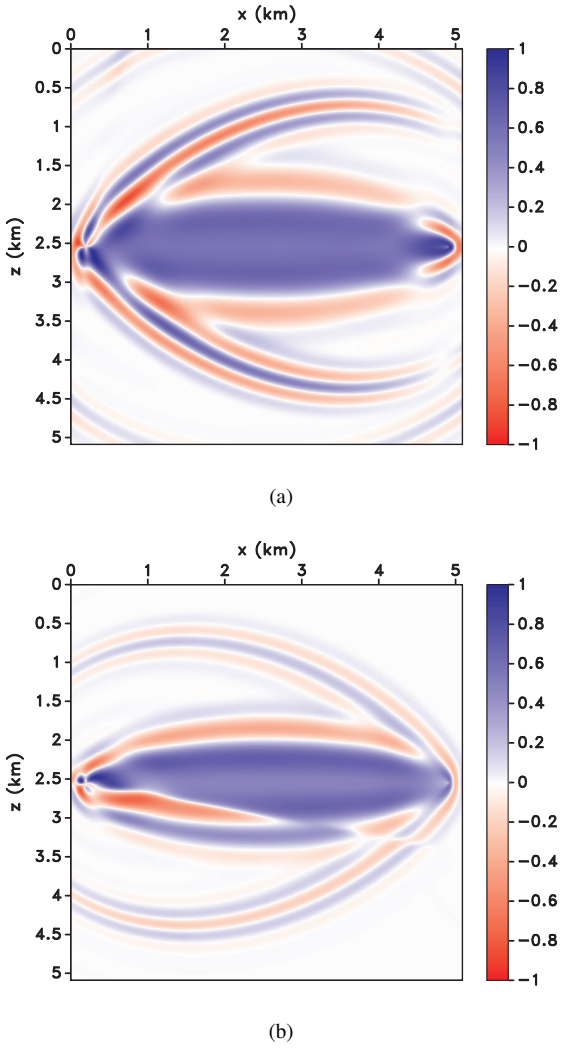


Figure 5. Normalized gradient for (a) α and (b) β. Strong and asymmetric artifacts are caused by the *fake* modes.

and

$$\begin{bmatrix} \frac{\partial \mathcal{J}_C}{\partial \alpha} \\ \frac{\partial \mathcal{J}_C}{\partial \beta} \end{bmatrix} = \begin{bmatrix} -\frac{\eta}{2\sqrt{\alpha}} \frac{1}{\sqrt{\alpha}-c_l\sqrt{\beta}-b_l} + \frac{\eta}{2\sqrt{\alpha}} \frac{1}{b_u-\sqrt{\alpha}+c_u\sqrt{\beta}} \\ +\frac{c_l\eta}{2\sqrt{\beta}} \frac{1}{\sqrt{\alpha}-c_l\sqrt{\beta}-b_l} - \frac{c_u\eta}{2\sqrt{\beta}} \frac{1}{b_u-\sqrt{\alpha}+c_u\sqrt{\beta}} \end{bmatrix}, \tag{18}$$

respectively. The gradient of term $\mathcal{J}_M$ is linear to model parameters, as shown in equation 17, while the gradient of term $\mathcal{J}_C$ is strongly nonlinear with respect to model parameters (equation 18). When updated models approach the boundaries defined in term $\mathcal{J}_C$, the gradient of $\mathcal{J}_C$ dominate the total gradient $\mathcal{J}$, thus pushing the updated model away from the boundaries; otherwise, term $\mathcal{J}_C$ has less influence on the total gradient $\mathcal{J}$, and the data term controls the inversion.

In the synthetic example, we take $c_l = c_u$ to be the ratio of the true P and S velocity model, and $b_l = -b_u = 0.02$ km^2/s^2. We compare the inversion results with objective functions $\mathcal{J}_D$ and $\mathcal{J}_D + \mathcal{J}_C$. By including term $\mathcal{J}_C$, the updated α and β are forced away from the upper and lower boundaries. Because the ratio of the true α and β is in the middle of the constraint region, the global minimum of term $\mathcal{J}_C$ is also the true model. The updated models after 10 iterations are shown in Figures 9(a) and 9(b). The bottom panels in Figures 9(a) and 9(b) show a subset from the update model at depth $z = 2.55$ km, indicated by the solid lines on the top panels; the dashed lines show the corresponding traces from the true model. Compared to inversion using term $\mathcal{J}_D$, the updated models using $\mathcal{J}_D + \mathcal{J}_C$ are closer to the true model, as seen in Figures 9(a) and 9(b).

To visualize the effect of term $\mathcal{J}_C$ on model updating, we plot the model parameters at coordinates $(2.55, 2.55)$ km for each iteration (Figure 11). The red and blue dots are the

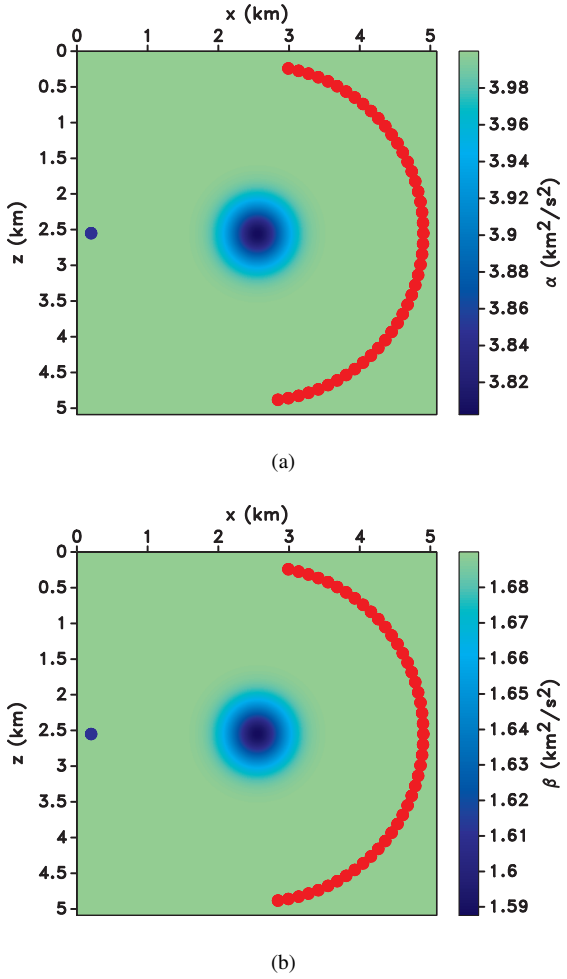


Figure 6. 2D synthetic (a) α and (b) b models with a negative Gaussian velocity anomaly with a constant background model. The source is at $(0.2, 2.55)$ km. There are 46 receivers on a circle of radius 2.35 km. The source and receivers are on opposite sides of the anomaly in order to record waves transmitted through the anomaly.

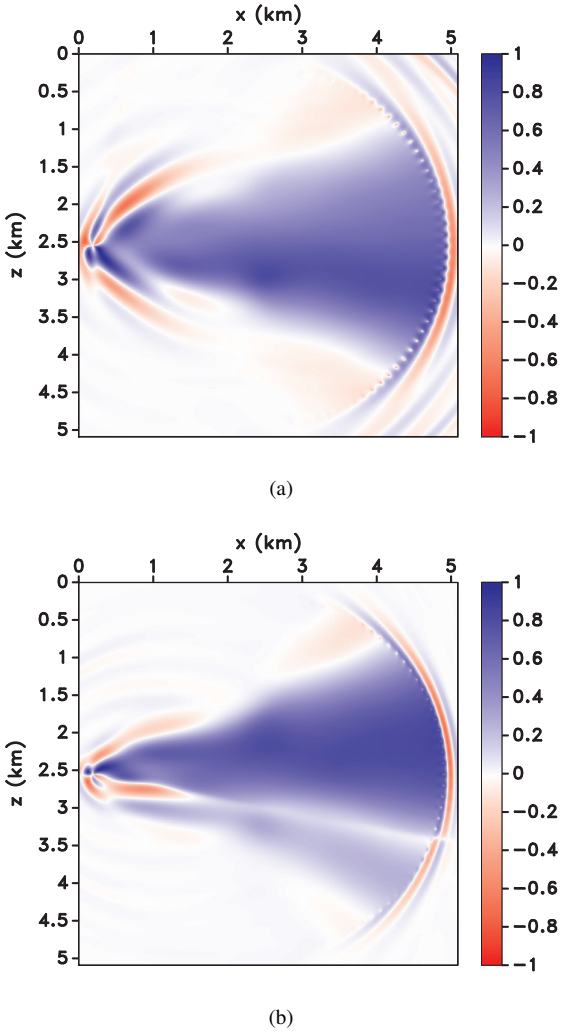


Figure 7. Normalized gradient for (a) α and (b) β. The artifacts around receivers are attenuated by increasing the number of receivers. The gradient for β is weak in the lower part due to the polarity change in the shear wave fronts.

updated models with and without physical constraints, respectively. Symbols S and T indicate the starting model and true model, respectively. The three lines from top to bottom are upper boundary, true α–β relationship, and lower boundary. We observe that in both inversions the model updates toward the true model in each iteration. The constraint term $\mathcal{J}_C$ affects model updating in the first iteration, and it enforces the updated model to be close to the line that indicates the relationship of true α and β models. In the last several iterations, the model updates are small, thus indicating that we approach convergence. With constraint term $\mathcal{J}_C$, the updated models of α and β after 25 iterations are closer to the true model compared to the updated model using only $\mathcal{J}_D$.

5 EXAMPLE

We illustrate our elastic wavefield tomography with a more complex synthetic model, and compare inversion with only the data misfit term $\mathcal{J}_D$ to inversion using data misfit with physical constraints $\mathcal{J}_D + \mathcal{J}_C$.

The synthetic model contains two negative Gaussian anomalies centered at $(1.5, 2.0)$ and $(1.5, 5.0)$ km. There are 60 shots in a well at $x = 0.2$ km and a line of receivers at $x = 2.8$ km. The α and β models are shown in Figures 12. The vertical component of a shot gather is shown in Figure 13(a), which contains both P- and S-waves. To illustrate the influence of the physical constraint term $\mathcal{J}_C$ on model updating when P- and S-waves have different illuminations, we mute the S-waves in both observed and predicted data for the receivers from $z = 0$ to 3.5 km, and the P-waves in both observed and

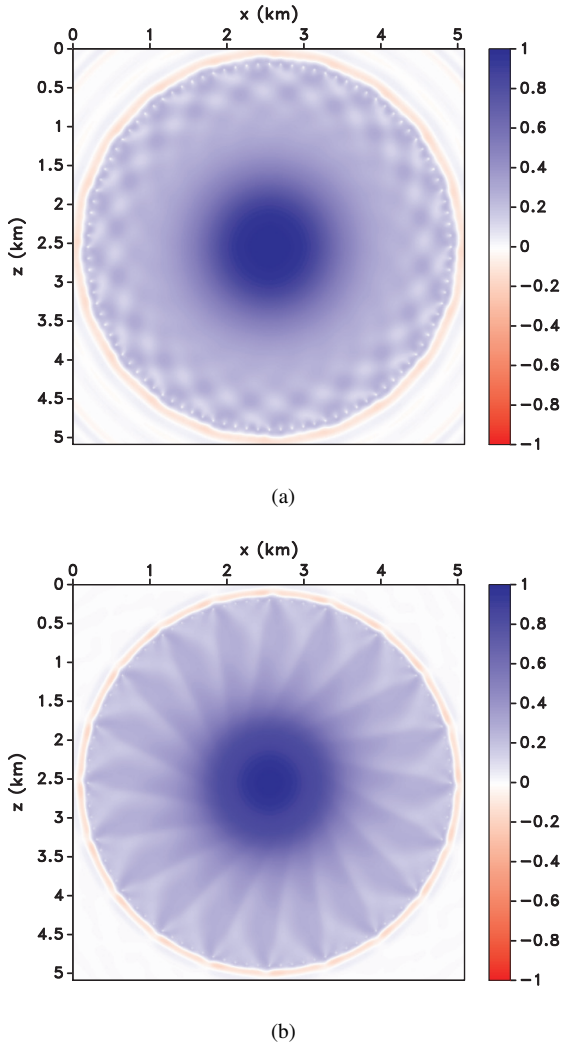


Figure 8. Normalized gradient for (a) α and (b) β. The acquisition geometry includes 20 shots, sampled every 18°, with 46 receivers for each shot. Both the source- and receiver-side artifacts are attenuated and the gradient are uniform in the center.

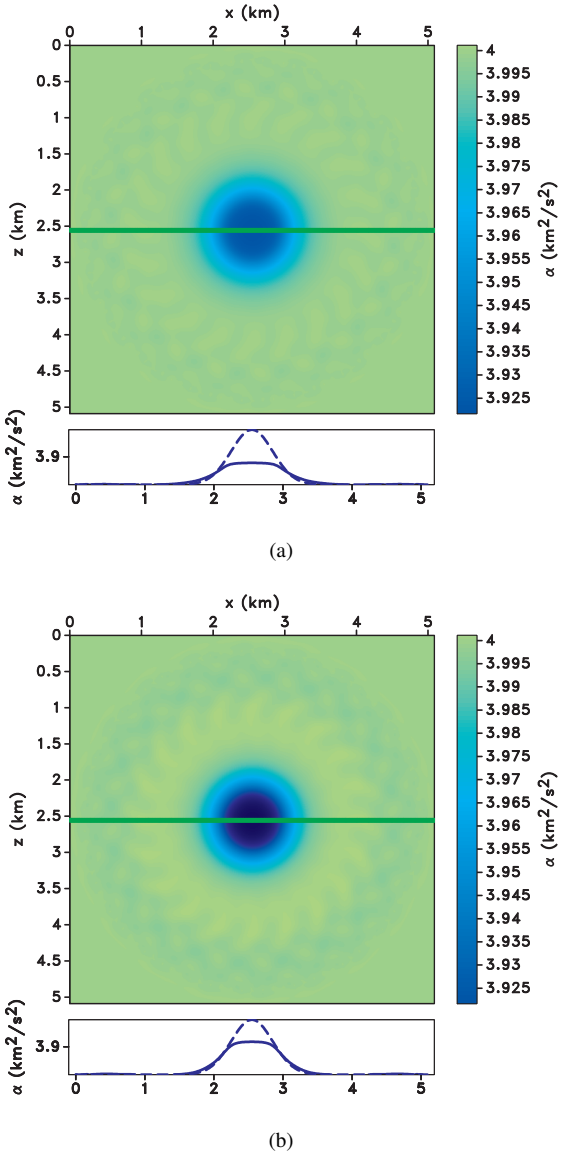
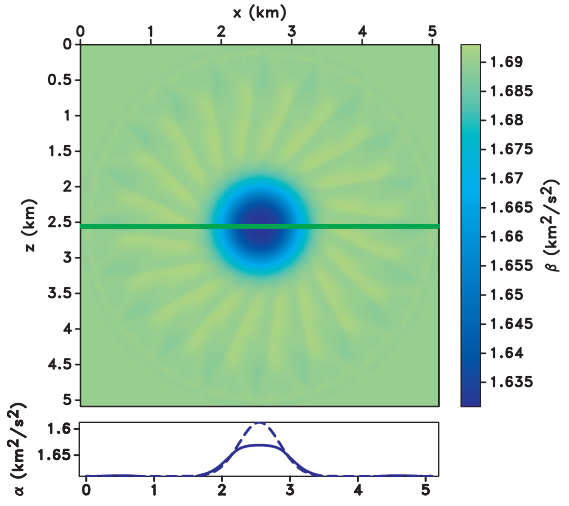


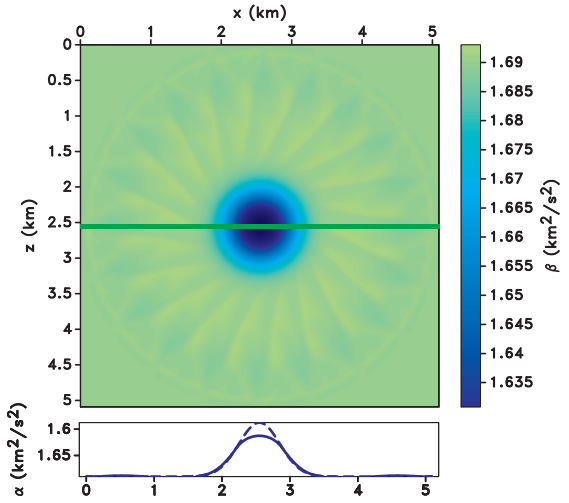
Figure 9. Updated α models after 10 iterations, using (a) objective function $\mathcal{J}_D$ and (b) $\mathcal{J}_D + \mathcal{J}_C$. The bottom panels are subsets of the model at depth $z = 2.55$ km. The solid and dashed lines are the updated and true models, respectively. The updated α using objective function $\mathcal{J}_D + \mathcal{J}_C$ is closer to the true model.

predicted data for the receivers from $z = 3.5$ to 7 km, as shown in Figure 13(b); i.e., in elastic wavefield tomography, we use primarily the P-waves going through the upper Gaussian anomaly and S-waves going through the lower Gaussian anomaly. This is, of course, an artificial construction meant to simulate partial illumination and to highlight the influence of the model constraint.

Using the data misfit term $\mathcal{J}_D$ as the objective function, we obtain the updated models for α and β after 21 iterations, shown in Figures 14(a) and 15(a), respectively. Notice that the model update of α focuses on the upper Gaussian anomaly, while that of β focuses on the lower Gaussian anomaly. We include the physical constraint term $\mathcal{J}_C$ in the objective function, and obtain the updated α and β models shown in Figures 14(b) and 15(b), respectively. We choose the weighting parameter η to be 0.6 in order to balance the gradients of $\mathcal{J}_D$ and $\mathcal{J}_C$. With physical constraints, both Gaussian anomalies are better recovered in α and β than the corresponding updates obtained without constraints. Figures 16(a) and 16(b) are, respectively, the values of the updated models at $(1.5, 2)$ and $(1.5, 5)$ km as a function of iteration. The blue and red dots show the models updates without and with physical constraints, respectively. Using only $\mathcal{J}_D$ as the objective function, we obtain α and β model updates that are non-physical. In Figure 16(a), the recovered α and β models tend to move toward the upper boundary in later iterations. In contrast, including



(a)



(b)

Figure 10. Updated β models after 10 iterations, using (a) objective function $\mathcal{J}_D$ and (b) $\mathcal{J}_D + \mathcal{J}_C$. The bottom panels are subsets of the model at depth $z = 2.55$ km. The solid and dashed lines are the updated and true models, respectively. The two updated β models are similar to each other.

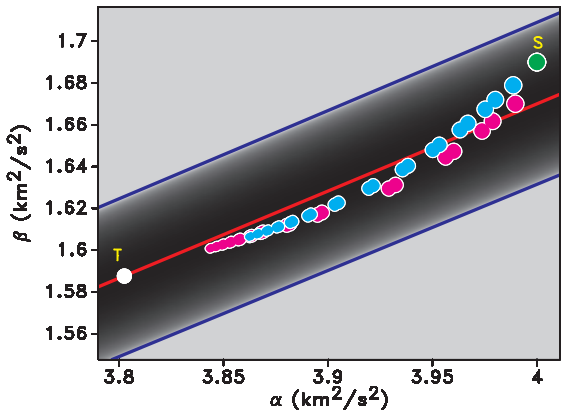


Figure 11. Model updates as a function of iterations. The dots represent the model at (2.55,2.55) km. Symbols S and T are the initial and true models at this location, respectively. The middle line depicts the true ratio between α and β. The upper and lower lines are the boundaries for the relationship of α and β. The red and blue dots are the updated models with and without constraints, respectively. The size of the dots decreases with increasing iteration number.

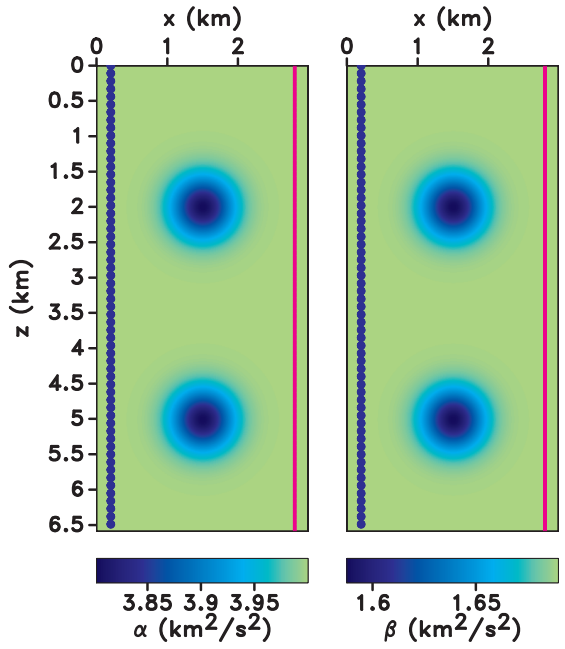


Figure 12. The α (left panel) and β (right panel) models with two negative Gaussian anomalies. The dots are the source locations at $x = 0.2$ km, and the vertical line shows the receivers at $x = 2.8$ km.

the physical constraint term $\mathcal{J}_C$ in the objective function, we enforce appropriate physical relationships of α and β.

6 CONCLUSIONS

We demonstrate an elastic wavefield tomography method using the isotropic elastic wave-equation formulated to reduce the misfit between observed and predicted data, as well as to obtain models that are physically plausible. We invert for a model of multiple elastic material parameters, specifically the squared velocities of P- and S-waves. Due to differences in illumination and amplitude between P- and S-waves, the model updates for the two parameters may differ in amplitude and location, thus leading to nonphysical models. To obtain a physically realistic model, we introduce a constraint term that enforces model updates within a feasible range. We use a logarithmic function as the constraint term; this term only impacts

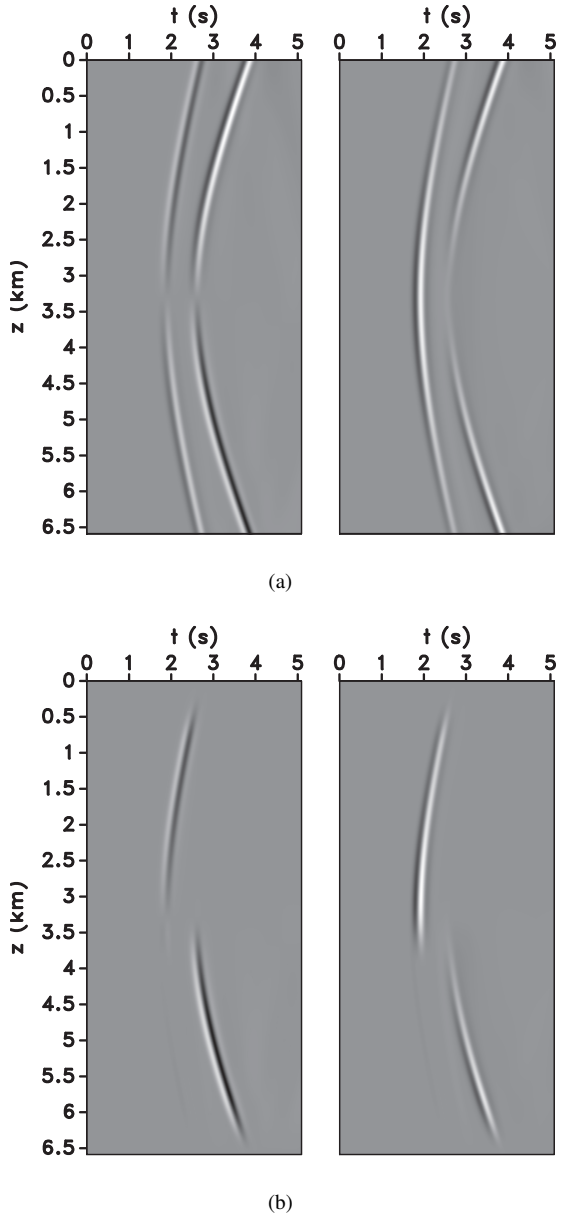


Figure 13. (a) The horizontal (left panels) and vertical (right panels) components of a shot gather with the source at $z = 3.5$ km. The first arrival is the P-wave and second arrival is the S-wave. (b) The processed data for inversion with P-waves only from $z = 0$ to 3.5 km and S-waves only from $z = 3.5$ to 7 km.

model updating when the inverted model parameters are close to the boundaries of the constraints.

7 ACKNOWLEDGMENTS

We thank sponsor companies of the Consortium Project on Seismic Inverse Methods for Complex Structures, whose support made this research possible. The reproducible numeric

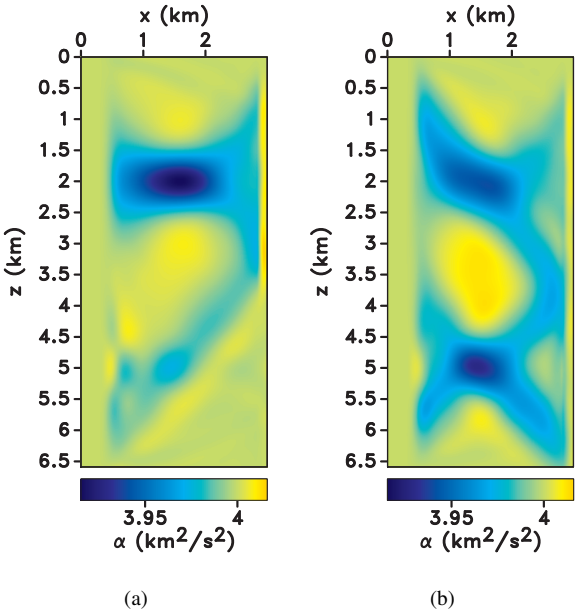


Figure 14. Updated α models after 21 iterations, using the data misfit objective function (a) without and (b) with model constraints.

examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org.

REFERENCES

Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2013, Regularized seismic full waveform inversion with prior model information: Geophysics, **78(2)**, R25–R36.

Bae, H. S., S. Pyun, C. Shin, K. J. Marfurt, and W. Chung, 2012, Laplace-domain waveform inversion versus refraction-traveltime tomography: Geophysical Journal International, **190**, 595606.

Biondi, B., and A. Almomin, 2014, Efficient and robust waveform-inversion workflow: SEG Technical Program Expanded Abstracts, 917–921.

Castagna, J. P., M. L. Batzle, and R. L. Eastwood, 1985, Relationships between compressional-wave and shear-wave velocities in elastic silicate rocks: Geophysics, **50(4)**, 571–581.

Compton, S., and D. Hale, 2013, Estimating Vp/Vs rratio using smooth dynamic image warping: SEG Technical Program Expanded Abstracts, 1639–1643.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8.

Gasso, G., A. Rakotomamonjy, and S. Canu, 2009, Recovering sparse signals with a certain family of non-convex

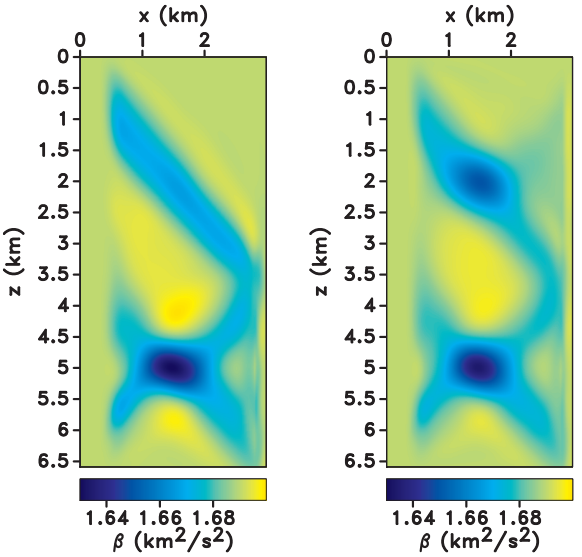


Figure 15. Updated α models after 21 iterations, using (a) objective function $\mathcal{J}_D$ and (b) $\mathcal{J}_D + \mathcal{J}_C$. Note that the recovered Gaussian anomaly in (b) at $(1.5, 2)$ km is better recovered, and its amplitude and shape are closer to the true model compared to that in (a).

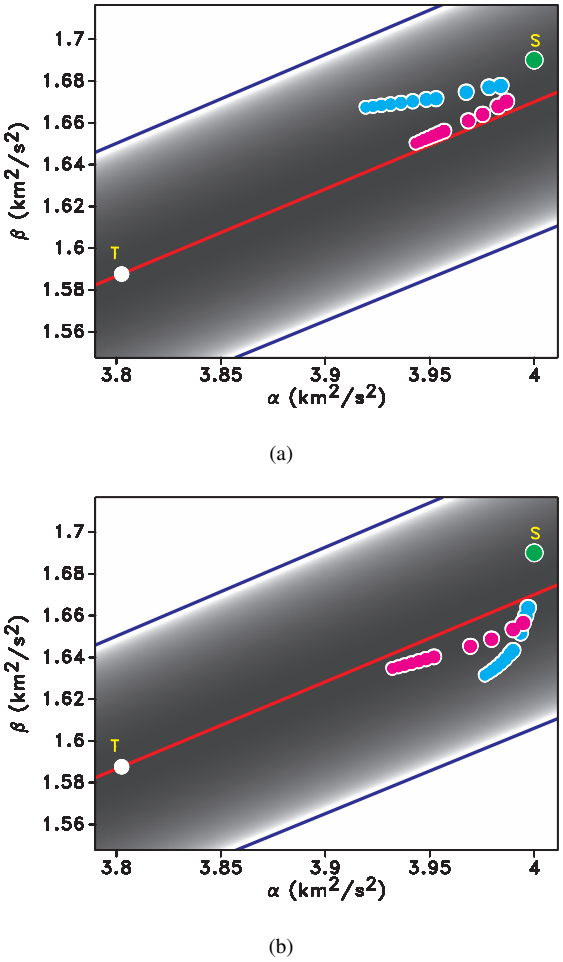


Figure 16. Model updates as a function of iterations. The dots represent the model at (a) $(1.5, 2)$ km and (b) $(1.5, 5)$ km. The red and blue dots are the updated models with and without constraints, respectively. Note that without constraints, the α and β models update toward the boundaries, while with constraints, the relationship of α and β models is enforced to be close to the true model.

penalties and DC programming: IEEE Transactions on Signal Processing, **57(12)**, 4686–4698.

Guasch, L., M. Warner, T. Nangoo, J. Morgan, A. Umpleby, I. Stekl, and N. Shah, 2012, Elastic 3D full-waveform inversion: SEG Technical Program Expanded Abstracts, 1–5.

Guitton, A., G. Ayeni, and G. Gonzales, 2010, A preconditioning scheme for full waveform inversion: SEG Technical Program Expanded Abstracts, 1008–1012.

Hale, D., 2007, Local dip filtering with directional laplacians: CWP Consortium Project on Seismic Inverse Methods for Complex Structures, 91–102.

Ivanov, J., R. D. Miller, J. Xia, D. Steeples, and C. B. Park, 2005, The Inverse Problem of Refraction Travel Times, Part i: Types of Geophysical Nonuniqueness Through Minimization: Pure and applied geophysics, **162(3)**, 447–459.

Katahara, K. W., 1999, Effect of mineral Vp/Vs on rock vp/Vs: SEG Technical Program Expanded Abstracts, 1029–1032.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering: Theory and Application, Society of Industrial and Applied Mathematics, Expanded Abstracts, 206–220.

Lay, T., and T. Wallace, 1995, Modern global seismology: Academic Press, San Diego.

McNeely, J., T. Keho, T. Tonellot, R. Ley, and J. Chen, 2012, 3d acoustic waveform inversion of land data: a case study from saudi arabia: SEG Technical Program Expanded Abstracts, 1–5.

Mora, P., 1988, Inversion of Seismic Wavefield Data for P- and S-wave Velocity: Exploration Geophysics, **19(1/2)**, 312–316.

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

Operto, S., C. Ravaut, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana, 2004, Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study: Geophysical Prospecting, **52**, 625–651.

Peng, J., C. Roos, and T. Terlaky, 2002, Self-regular functions and new search directions for linear and semidefinite optimization: Mathematical Programming, **93**, 129–171.

Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysics Journal International, **167**, 495–503.

Pratt, R. G., 1990, Inverse theory applied to multi-source cross-hole tomography: Geophysical Prospecting, **38**, 311–

329.

———, 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64(3)**, 888–901.

Prieux, V., S. Operto, R. Brossier, J. Virieux, J. Kommendal, and O. Barkved, 2010, Application of 2d acoustic frequency-domain full-waveform inversion to obc wide-aperture data from the valhall field: SEG Technical Program Expanded Abstracts, 920–924.

Rojas, E., T. L. Davis, M. Batzle, M. Prasad, and R. J. Michelena, 2005, Vp-vs ratio sensitivity to pressure, fluid, and lithology changes in tight gas sandstones: SEG Technical Program Expanded Abstracts, 1401–1404.

Schuster, G. T., and A. Quintus-Bosz, 1993, Wavepath eikonal traveltime inversion: Theory: Geophysics, **58(9)**, 1314–1323.

Sun, Y., and G. Shuster, 1992, Hierarchic optimizations for smoothing and cooperative inversion: SEG Technical Program Expanded Abstracts, 745–748.

Taillandier, C., M. Noble, H. Chauris, and H. Calandra, 2009, First-arrival traveltime tomography based on the adjoint-state method: Geophysics, **74(6)**, WCB1–WCB10.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49(8)**, 1259–1266.

———, 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51(10)**, 1893–1903.

Tikhonov, A., and V. Arsenin, 1977, Solution of ill-posed problems: Winston & Sons.

Tsuneyama, F., 2006, Exploring Vp-Vs relations: Approach from effective medium theories: Proceedings of the 8th SEGJ International Symposium, Kyoto, Japan, 1–6.

Vigh, D., K. Jiao, and D. Watts, 2012, Elastic full-waveform inversion using 4C data acquisition: SEG Technical Program Expanded Abstracts, 1–6.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57(1)**, 15–26.

Xiang, S., and H. Zhang, 2014, A new adaptive edge-preserving regularization scheme for seismic full waveform inversion: SEG Technical Program Expanded Abstracts, 1238–1242.

Zelt, C. A., A. Azaria, and A. Levander, 2006, 3D seismic refraction traveltime tomography at a groundwater contamination site: Geophysics, **71(5)**, H67–H78.

Zhou, H., Y. Zhang, and S. Gray, 2002, Regularization algorithms for seismic inverse problems: SEG Technical Program Expanded Abstracts, 942–945.

Zhu, X., D. Sixta, and B. Angstman, 1992, Tomostatics: Turning-ray tomography + tatics corrections: The Leading Edge, **11**, 15–23.

Zimmer, M., M. Prasad, and G. Mavko, 2002, Pressure and porosity influences on vp-vs ratio in unconsolidated sands: The Leading Edge, **21**, 178–183.