Elastic wavefield tomography with physical model constraints

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

SUMMARY

We propose elastic isotropic wavefield tomography formulated with the misfit of observed and recorded data as the objective function. Using the elastic isotropic wave-equation formulated for slowly varying media, we invert for the squared velocities of compressional (P) and shear (S) waves. 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.

INTRODUCTION

Seismic tomography is a commonly used tool for building models of subsurface parameters from recorded seismic data. 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 better characterizes the subsurface (Tarantola, 1986; Pratt, 1990; Guasch et al., 2012; Vigh et al., 2012). There are many possible parameterizations for elastic models, which lead to different inversion schemes. For example, Tarantola (1986) shows wavefield tomography for compressional impedance, shear 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 may not be able to resolve the model parameters while preserving their intrinsic relationships. Therefore, 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 regularization. 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.

THEORY

We consider the elastic isotropic wave-equation, in which model parameters $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ representing the squared P and S velocities 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. 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 $\mathbf{L}^{\mathsf{T}}$. The source wavefield $\mathbf{u}_s$ is a solution to the given wave-equation.

Objective function

For wavefield tomography, one typically updates a model by minimizing an objective function. In our formulation, the objective function $\mathcal{J}(\mathbf{u}_s,\alpha,\beta)$ consists 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 $\mathcal{J}_C(\alpha,\beta)$ that restricts the relationship of α and β within a feasible region:

$$\mathcal{J}(\mathbf{u}_s,\alpha,\beta) = \mathcal{J}_D(\mathbf{u}_s,\alpha,\beta) + \mathcal{J}_M(\alpha,\beta) + \mathcal{J}_C(\alpha,\beta)\ . \tag{3}$$

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

$$\mathcal{J}_D = \sum_e \frac{1}{2}\|\mathbf{W}_u\mathbf{u}_s - \mathbf{d}_o\|^2\ , \tag{4}$$

where $\mathbf{d}_o(e,\mathbf{x},t)$ are the observed data recorded by receivers, and the weighting operator $\mathbf{W}_u(e,\mathbf{x},t)$ restricts the source wavefield $\mathbf{u}_s(e,\mathbf{x},t)$ to the receiver locations.

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

$$\mathcal{J}_M = \frac{1}{2}\|\mathbf{W}_\alpha(\alpha-\bar{\alpha})\|^2 + \frac{1}{2}\|\mathbf{W}_\beta(\beta-\bar{\beta})\|^2 , \quad (5)$$

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. a Laplacian filter, or they can describe image-shaping with non-stationary filtering (Hale, 2007; Guitton et al., 2010).

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

$$h_u > 0 , \quad (6)$$

$$h_l > 0 . \quad (7)$$

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):

$$\mathcal{J}_C = -\eta\sum_{\mathbf{x}}\left[\log(h_u)+\log(h_l)\right] . \quad (8)$$

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 6 and 7. 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 between the boundaries, and this constraint term mainly contributes to the inversion 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 as

$$h_u = -\sqrt{\alpha} + c_u\sqrt{\beta} + b_u = 0 , \quad (9)$$

$$h_l = \sqrt{\alpha} - c_l\sqrt{\beta} - b_l = 0 , \quad (10)$$

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

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. Using the adjoint-state method (Plessix, 2006), we derive the gradients of the three terms in the objective function. The gradient of $\mathcal{J}_D$ is

$$\begin{bmatrix}\frac{\partial\mathcal{J}_D}{\partial\alpha}\\ \frac{\partial\mathcal{J}_D}{\partial\beta}\end{bmatrix} = \sum_e \begin{bmatrix}-\left[\nabla(\nabla\cdot\mathbf{u}_s)\right]^T \star \mathbf{a}_s\\ \left[\nabla\times(\nabla\times\mathbf{u}_s)\right]^T \star \mathbf{a}_s\end{bmatrix} , \quad (11)$$

where the symbol $\star$ denotes zero-lag crosscorrelation. Vector $\mathbf{a}_s(e,\mathbf{x},t)$ is the adjoint variable and is computed by solving the adjoint equation

$$\mathbf{L}^\mathsf{T}\mathbf{a}_s = \mathbf{g}_s . \quad (12)$$

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

$$\mathbf{g}_s = \frac{\partial\mathcal{J}}{\partial\mathbf{u}_s} = \mathbf{W}_u^\mathsf{T}(\mathbf{W}_u\mathbf{u}_s - \mathbf{d}_o) . \quad (13)$$

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

$$\begin{bmatrix}\frac{\partial\mathcal{J}_M}{\partial\alpha}\\ \frac{\partial\mathcal{J}_M}{\partial\beta}\end{bmatrix} = \begin{bmatrix}\mathbf{W}_\alpha^\mathsf{T}\mathbf{W}_\alpha(\alpha-\bar{\alpha})\\ \mathbf{W}_\beta^\mathsf{T}\mathbf{W}_\beta(\beta-\bar{\beta})\end{bmatrix} , \quad (14)$$

and

$$\begin{bmatrix}\frac{\partial\mathcal{J}_C}{\partial\alpha}\\ \frac{\partial\mathcal{J}_C}{\partial\beta}\end{bmatrix} = \begin{bmatrix}-\frac{\eta}{2\sqrt{\alpha}}\frac{1}{\sqrt{\alpha}-c_l\sqrt{\beta}-b_l} + \frac{\eta}{2\sqrt{\alpha}}\frac{1}{b_u-\sqrt{\alpha}+c_u\sqrt{\beta}}\\ +\frac{c_l\eta}{2\sqrt{\beta}}\frac{1}{\sqrt{\alpha}-c_l\sqrt{\beta}-b_l} - \frac{c_u\eta}{2\sqrt{\beta}}\frac{1}{b_u-\sqrt{\alpha}+c_u\sqrt{\beta}}\end{bmatrix} , \quad (15)$$

respectively. The gradient of term $\mathcal{J}_M$ is linear with respect to model parameters, as shown in equation 14, while the gradient of term $\mathcal{J}_C$ is strongly nonlinear with respect to model parameters (equation 15). When the models updated during iterations approach the boundaries defined by $\mathcal{J}_C$, the gradient of $\mathcal{J}_C$ dominates 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.

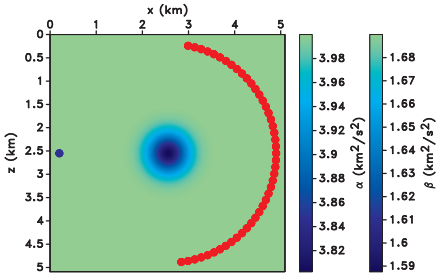

Figure 1: 2D synthetic α and β models containing a negative Gaussian velocity anomaly with a constant background model. The acquisition geometry includes 20 shots, sampled every 18°, with 46 receivers for each shot. The source and receivers are on opposite sides of the anomaly in order to record waves transmitted through the anomaly.

We illustrate the elastic wavefield tomography method with the synthetic model shown in Figure 1. For the model constraint term $\mathcal{J}_C$, we take $c_l = c_u$ to be the ratio of the true P and S velocities, and $b_l = -b_u = 0.02$ km^2/s^2. We compare the inversion results with objective functions $\mathcal{J}_D$ and $\mathcal{J}_D+\mathcal{J}_C$. The updated models after 10 iterations are shown in Figures 2a and 2b. The bottom panels 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.

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 3). We observe that, in both inversions, the model updates toward the true model. The constraint term $\mathcal{J}_C$ affects inversion primarily in the first iteration by forcing the

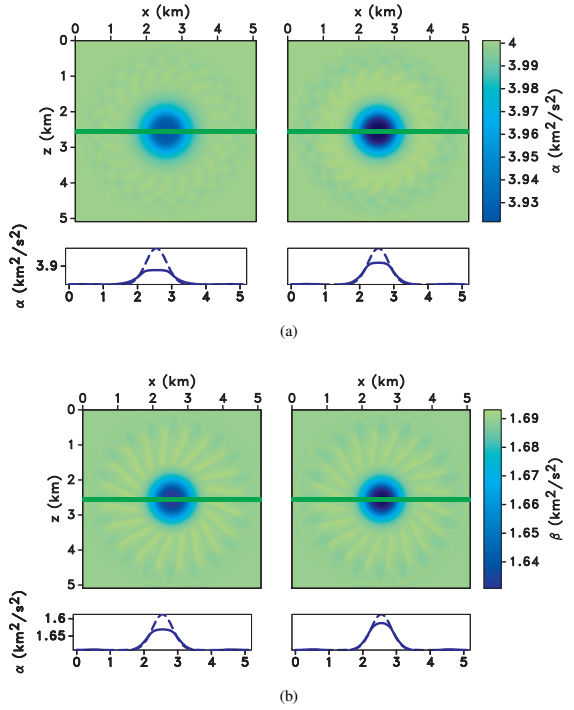

Figure 2: Updated (a) α and (b) β models after 10 iterations, using objective function $\mathcal{J}_D$ (left panels) and $\mathcal{J}_D + \mathcal{J}_C$ (right panels). The bottom panels are subsets of the model at depth $z = 2.55$ km. The solid and dashed lines are the updated and true models, respectively. The updated α and β using objective function $\mathcal{J}_D + \mathcal{J}_C$ is closer to the true model.

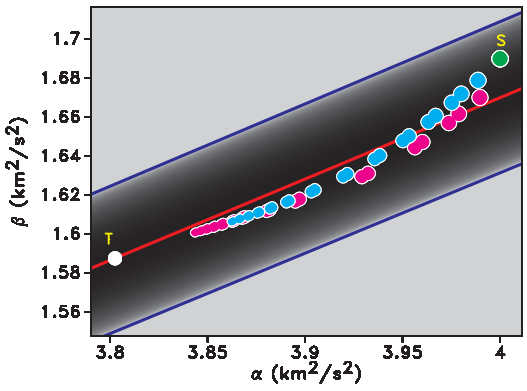

Figure 3: Model updates as a function of iterations. The dots represent the model at (2.55,2.55) km. Symbols S and T are the initial and true models at this location, respectively. The middle line depicts the true ratio between α and β. The upper and lower lines are the boundaries for the relationship of α and β. The red and blue dots are the updated models with and without constraints, respectively. For easier visualization, the size of the dots decreases as the number of iterations increases.

updated model toward the line indicating the true relationship between α and β. In the last several iterations, the model updates are small, indicating that we approach convergence. After 25 iterations, the updated models of α and β, using the constraint term $\mathcal{J}_C$, are closer to the true model, compared to the updated model using only $\mathcal{J}_D$.

EXAMPLE

We also 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 the 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, Figure 4. There are 60 shots in a well at $x = 0.2$ km and a line of receivers at $x = 2.8$ km. The horizontal component of a shot gather contains both P- and S-waves, left panel in Figure 5. To illustrate the influence of the physical constraint term $\mathcal{J}_C$ on model updating when the 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 predicted data for the receivers from $z = 3.5$ to 7 km, right panel in Figure 5; i.e., in elastic wavefield tomography, we use primarily the P-waves going through the upper Gaussian anomaly and the S-waves going through the lower anomaly. This is, of course, an artificial construction meant to simulate partial illumination and to highlight the influence of the model constraint.

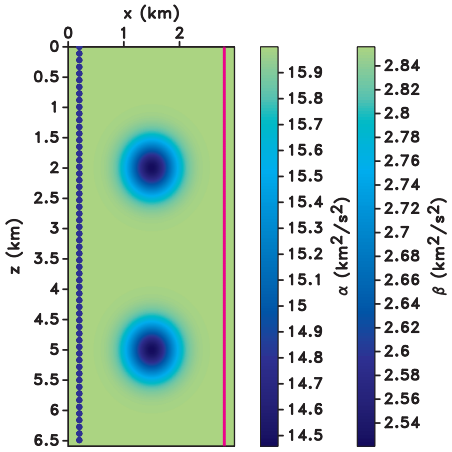

Figure 4: The α and β models with two negative Gaussian anomalies. The dots are the source locations at $x = 0.2$ km, and the vertical line shows the receivers at $x = 2.8$ km.

Using the data misfit term $\mathcal{J}_D$ as the objective function, we obtain the updated models for α and β after 21 iterations, left panels in Figures 6a and 6b, 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, right panels in Figures 6a and 6b, respectively. We choose the weighting parameter η to be 0.6 in order to balance the gradients of $\mathcal{J}_D$

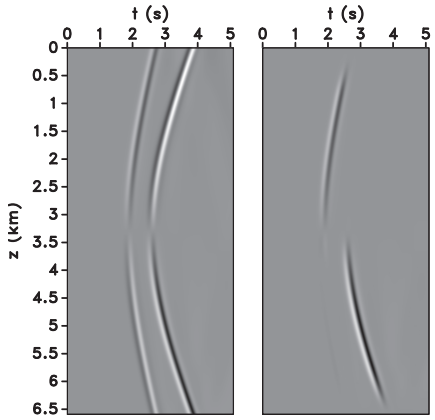

Figure 5: Left panel: the horizontal component of a shot gather with the source at $z = 3.5$ km. The early arrival is the P-wave and later arrival is the S-wave. Right panel: the processed data for inversion with P-waves only from $z = 0$ to 3.5 km and S-waves only from $z = 3.5$ to 7 km.

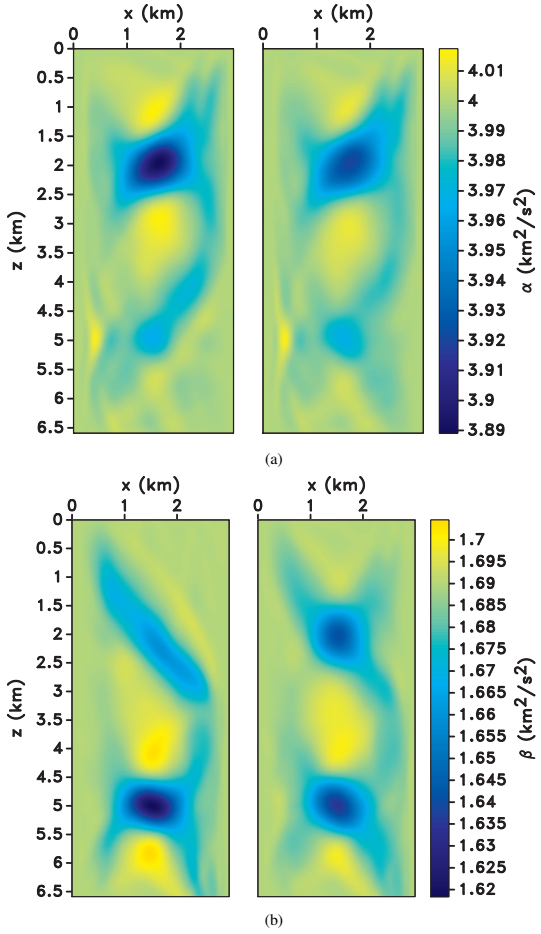

Figure 6: Updated (a) α and (b) β models after 21 iterations, using objective function $\mathcal{J}_D$ (left panels) and $\mathcal{J}_D + \mathcal{J}_C$ (right panels). Note that the Gaussian anomaly at $(1.5, 2)$ km is better recovered in right panel of (b).

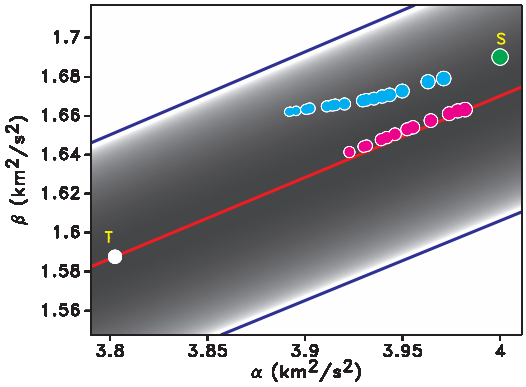

Figure 7: Model updates as a function of iterations for the example in Figure 4. The dots represent the model at $(1.5, 2)$ km. The red and blue dots are the updated models with and without constraints, respectively. Using the constraints forces the α and β models closer to the true model.

and $\mathcal{J}_C$. With physical constraints, both Gaussian anomalies are better recovered in α and β than the corresponding updates obtained without constraints. Figure 7 shows the values of the updated models at $(1.5, 2)$ km as a function of iteration. The blue and red dots depict the models without and with physical constraints, respectively. Using only $\mathcal{J}_D$ as the objective function, we obtain α and β model updates that are nonphysical, i.e. the models do not converge toward the center line corresponding to the true $\alpha - \beta$ relationship. In contrast, the physical constraint term $\mathcal{J}_C$ enforces the appropriate physical relationship between α and β.

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

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

REFERENCES

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 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.

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.

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-aperturesseismic 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.

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.

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.

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.

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.