

Elastic least-squares one-way wave-equation migration

Aaron Stanton¹ and Mauricio D. Sacchi²

ABSTRACT

Least-squares migration seeks a reflectivity model that fits the observed data. It is used to compensate for acquisition noise, poor sampling of sources and receivers on the surface, as well as poor illumination of the subsurface. To date, least-squares migration has been mainly restricted to the imaging of acoustic wavefields. We have developed an extension of one-way wave-equation least-squares migration for elastic wavefields in isotropic media. Least-squares migration is an iterative method that requires a forward and an adjoint operator. In elastic least-squares one-way wave-equation migration, the forward operator generates data components from multiparameter images by recursive wavefield decomposition, extrapolation, and recombination. Conversely, the adjoint operator generates multiparameter images from data components by recursively applying the adjoint of the wavefield recombination, extrapolation, and wavefield decomposition operators. We use an extended imaging condition and regularize the inversion by applying a smoothing filter on the depth-angle axes of each common image point gather to reduce the effect of source/receiver sampling, noise, and crosstalk artifacts. Elastic least-squares migration is able to compensate for irregular subsurface illumination in elastic imaging and provides an alternative approach to interpolation and wavefield separation of multicomponent seismic data.

INTRODUCTION

Converted waves are a useful tool for characterizing the elastic properties of a reservoir for many reasons. The comparison of PP and PS reflectivities can help to identify fluid contacts, and S-wave raypaths provide a tool to illuminate targets hidden beneath gas-filled zones. Perhaps most importantly, P- and S-waves provide

complementary information that is helpful in inverting for elastic parameters (Stewart, 1990).

A critical step in processing converted waves is prestack migration. For P-wave seismology, it is common to simulate wave propagation using the scalar-wave equation. This approximation is often sufficient to produce reasonable P-wave structural images. Extending this approximation to converted wave imaging disregards the polarization of wave modes within data components. Migration algorithms based on the elastic-wave equation model wave propagation more accurately and allow multiple wavefields to be imaged simultaneously. There are a variety of approaches to elastic migration. Kuo and Dai (1984) develop an elastic implementation of Kirchhoff migration. Chang and McMechan (1987) apply reverse time migration (RTM) to multicomponent data using an elastic finite-difference algorithm, resulting in horizontal and vertical component images. Dellinger and Etgen (1990) propose the application of Helmholtz decomposition to elastic data using a Fourier-domain operator. Later implementations of elastic RTM use this idea to provide distinct PP- and PS-images (e.g., see Yan and Sava, 2008). Cheng et al. (2016) use a low-rank approximation to simulate decoupled elastic waves in anisotropic media, and Guo and Alkhalifah (2016) perform reflection waveform inversion using the elastic data. Extensive work has also been done on one-way wave-equation-based methods to migrate elastic data. Xie and Wu (2005) use an elastic version of split step migration to downward continue elastic data, whereas Bale (2006) modifies phase-shift migration to handle anisotropic elastic wavefields. One-way wave-equation-based methods are an attractive option because they perform more accurately than ray-based methods, and they are much less computationally expensive than methods that use finite differences. For example, in time-domain RTM, many wavefield snapshots for all time samples must be generated and saved prior to imaging, whereas in one-way wave-equation migration, frequency slices are treated independently and the wavefield is recursively updated with depth. One-way wave-equation migration operators are also computationally efficient in memory and operation count, perform accurately in moderately complex geologic settings, and result in relatively few numerical

Manuscript received by the Editor 22 July 2016; revised manuscript received 26 January 2017; published online 12 May 2017.

¹Key Seismic Solutions Ltd., Calgary, Alberta, Canada. E-mail: aaron.stanton@keyseismic.com.

²University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: msacchi@ualberta.ca.

© 2017 Society of Exploration Geophysicists. All rights reserved.

artifacts. There are also many promising methods to improve on traditional one-way wave-equation migration that can make its results competitive with RTM (e.g., [Shan and Biondi, 2008](#)).

Whereas in acoustic migration, the image represents the partitioning of energy between incident and reflected waves that share a common mode, in elastic migration, a multiparameter image is created that represents the partitioning of energy between all combinations of incident and reflected wave modes. The relative strength of these images as a function of angle of incidence is related to the physical properties of the interface by the Zoeppritz equations ([Aki and Richards, 2002](#)). Because P-waves propagate by compression and dilation, their corresponding reflection strengths are greatly influenced by pore fluids. On the other hand, S-wave reflection strengths are less influenced by the presence of pore fluids and instead depend more heavily on changes in the rock matrix. As an example, [MacLeod et al. \(1999\)](#) delineate a sandstone reservoir surrounded by shale by comparing PP- and PS-reflection strengths. A counteraction of impedance associated with a change in lithology and an oil-water contact made the reservoir difficult to identify in PP-images, whereas the PS-image was able to isolate the change in lithology necessary to identify the reservoir.

The migration of converted waves comes with some additional complications compared with the P-wave migration. Typically, PS data have a lower signal-to-noise ratio than PP data, and the lower propagation velocity of the receiver side wavefield moves the point of reflection away from the midpoint and toward the receiver. Even though lower propagation velocity implies more restrictive aliasing criteria, it also presents an opportunity: Given the same acquisition geometry, the converted waves can image the subsurface with a wider source side aperture than P-wave data. Furthermore, ignoring the effects of attenuation, a given frequency of converted wave data images the earth with a higher resolution than P-wave data. Realistically, attenuation is always a factor (especially in the near surface), which challenges the resolution of PP- and PS-depth images ([Bale and Stewart, 2002](#)). Some of these challenges can be dealt with before migration. For example, regularization can be applied to reduce the effects of the acquisition footprint on the migrated image ([Cary, 2011](#)). In this paper, we aim to address some of the challenges faced while using converted wave migration via least-squares inversion.

Least-squares migration, in its various forms, has been an active field of research for many years. [Lambaré et al. \(1992\)](#) use an iterative data-fitting approach to solve for reflectivity. [Nemeth et al. \(1999\)](#) apply least-squares migration with a Kirchhoff operator to image in the presence of poor spatial sampling, whereas [Kühl and Sacchi \(2003\)](#) use regularized least-squares one-way wave-equation migration to generate angle gathers that are unaffected by the acquisition footprint. [Wang et al. \(2005\)](#) and [Wang \(2005\)](#) explore the application of 3D one-way wave-equation least-squares migration with constraints to a real data set of the Western Canadian Sedimentary Basin. In a subsequent contribution, [Wang and Sacchi \(2007\)](#) also explore the incorporation of a sparsity constraint to least-squares migration to the ubiquitous problem of vertical resolution enhancement. Least-squares migration of structurally complex data has also been investigated via the use of preconditioning operators synthesized using prediction error filters ([Wang and Sacchi, 2009](#)). More recently, there has been a great deal of research incorporating two-way wave-equation migration operators into least-squares migration ([Ji, 2009; Dai and Schuster, 2013; Dutta and Schuster, 2014; Wong et al., 2015; Zhang et al., 2015](#)). A

common theme of all least-squares migration algorithms is data fitting, which depends greatly on the ability of the migration operator to accurately propagate energy from source to receiver. Extending least-squares migration from the acoustic to the elastic case is a natural progression in data fitting.

This paper is organized as follows. First, we review the theory of elastic-wave propagation and derive the necessary elements for elastic one-way wave propagation as eigensolutions to the Christoffel equation. Next, we integrate these elements into the first-order Born approximation, setting up the forward problem to simulate elastic wavefields from the multiparameter images. We then pose elastic imaging as an inverse problem in which the ideal model best fits the observed data in a least-squares sense. Finally, we demonstrate the ability of elastic least-squares migration (ELSM) to improve wavefield separation, mitigate the effects of poor spatial sampling, and compensate for nonuniform illumination.

THEORY

Elastic-wave propagation

The mechanics necessary for wavefield separation and extrapolation of vector data are provided by the elastic-wave equation:

$$\rho \ddot{u}_i = C_{ijkl} \partial_j \partial_l u_k, \quad (1)$$

where u_i is the displacement in the direction i , ρ is the density, and C_{ijkl} is the tensor of elastic moduli ([Bale, 2006](#)). In this section, Einstein notation is used, implying summation over repeated indices. Displacement can be represented as a complex harmonic (plane wave), $u_k = U_k e^{i\omega(t-s_j x_j)}$, where s_j is the slowness vector that is the inverse of the phase velocity. The phase velocity is related to the slowness via $s_j = n_j/V$, where n_j indicates the direction of wave propagation. The direction of wave propagation is just the normalized wavenumber vector, $n_j = k_j/k$, where $k = \sqrt{k_1^2 + k_2^2 + k_3^2}$. Substituting a plane wave into the wave equation, we obtain the Christoffel equation:

$$\rho U_i = C_{ijkl} s_j s_l U_k, \quad (2)$$

more commonly written in the form

$$\begin{bmatrix} G_{11} - \rho V^2 & G_{12} & G_{13} \\ G_{12} & G_{22} - \rho V^2 & G_{23} \\ G_{13} & G_{23} & G_{33} - \rho V^2 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad (3)$$

where the 3×3 Christoffel matrix is defined by

$$G_{ij} = C_{ijkl} n_j n_i. \quad (4)$$

The Christoffel equation is a 3×3 eigenvalue — eigenvector problem. Given a tensor of elastic moduli, the eigenvalues correspond to the phase velocities of three distinct wave modes and the eigenvectors correspond to the polarizations of these modes ([Tsvankin, 2012](#)). In the context of elastic migration, the eigenvectors provide the polarization information necessary for wavefield separation and the eigenvalues provide the phase velocity necessary for extrapolation. Assuming isotropy and substituting into equa-

tion 3 considering propagation in the $[x_1, x_3]$ plane, we obtain the following Christoffel equation:

$$\begin{bmatrix} (\lambda+2\mu)n_1^2+\mu n_3^2-\rho V^2 & 0 & (\lambda+\mu)n_1n_3 \\ 0 & \mu(n_1^2+n_3^2)-\rho V^2 & 0 \\ (\lambda+\mu)n_1n_3 & 0 & \mu n_1^2+(\lambda+2\mu)n_3^2-\rho V^2 \end{bmatrix} \times \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad (5)$$

where λ and μ are the Lamé's first and second parameters, respectively. To find the eigenvalues corresponding to the three wave modes, we set the determinant equal to zero and substitute $\mathbf{n} = [0 \ 0 \ 1]$ (i.e., assuming a plane wave traveling vertically downward) leading to the cubic equation:

$$0 = (\mu - \rho V^2)(\mu - \rho V^2)(\lambda + 2\mu - \rho V^2) \quad (6)$$

with solutions $V_1 = \sqrt{(\lambda + 2\mu)/\rho}$, $V_2 = \sqrt{\mu/\rho}$, and $V_3 = \sqrt{\mu/\rho}$, which are the familiar relations for P-, SV-, and SH-phase velocities, respectively. Solving for the corresponding eigenvectors, we obtain

$$\mathbf{U}^P = \begin{bmatrix} \hat{n}_1 \\ \hat{n}_2 \\ \hat{n}_3 \end{bmatrix}, \quad \mathbf{U}^{S1} = \begin{bmatrix} -\hat{n}_2/a \\ \hat{n}_1/a \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{U}^{S2} = \begin{bmatrix} -\hat{n}_1\hat{n}_3/a \\ -\hat{n}_2\hat{n}_3/a \\ a \end{bmatrix}, \quad (7)$$

where $a^2 = \hat{n}_1^2 + \hat{n}_2^2$.

Even though we simplify our calculations greatly by assuming isotropy, lower symmetry material parameters can also be incorporated into the Christoffel equation. For example, Bale (2006) solves the Christoffel equation assuming transverse isotropy.

Considering one-way wave propagation in laterally homogeneous media, we can propagate elastic wavefields vertically by wavefield decomposition, a vertical phase shift, and wavefield composition. This is represented by

$$\mathbf{u}(\omega, x, y, z + \Delta z) = \mathbf{P}\mathbf{u}(\omega, x, y, z), \quad (8)$$

where

$$\mathbf{P} = \mathbf{QRQ}^{-1}. \quad (9)$$

In laterally homogeneous media, the operator \mathbf{R} corresponds to a vertical phase shift (Gazdag, 1978) applied to each wavefield potential

$$\mathbf{R} = \begin{bmatrix} \mathcal{F}_x^{-1} & 0 & 0 \\ 0 & \mathcal{F}_x^{-1} & 0 \\ 0 & 0 & \mathcal{F}_x^{-1} \end{bmatrix} \begin{bmatrix} e^{i\omega k_z(V_p)\Delta z} & 0 & 0 \\ 0 & e^{i\omega k_z(V_S)\Delta z} & 0 \\ 0 & 0 & e^{i\omega k_z(V_S)\Delta z} \end{bmatrix} \times \begin{bmatrix} \mathcal{F}_x & 0 & 0 \\ 0 & \mathcal{F}_x & 0 \\ 0 & 0 & \mathcal{F}_x \end{bmatrix}, \quad (10)$$

and the wavefield composition operator is given by

$$\mathbf{Q} = \begin{bmatrix} \mathcal{F}_x^{-1} & 0 & 0 \\ 0 & \mathcal{F}_x^{-1} & 0 \\ 0 & 0 & \mathcal{F}_x^{-1} \end{bmatrix} [\mathbf{U}^P \ \mathbf{U}^{S1} \ \mathbf{U}^{S2}] \begin{bmatrix} \mathcal{F}_x & 0 & 0 \\ 0 & \mathcal{F}_x & 0 \\ 0 & 0 & \mathcal{F}_x \end{bmatrix}, \quad (11)$$

where \mathcal{F}_x and \mathcal{F}_x^{-1} represent the forward and inverse Fourier transforms over spatial axes. For isotropic media, the composition operator becomes

$$\mathbf{Q} = \begin{bmatrix} \mathcal{F}_x^{-1} & 0 & 0 \\ 0 & \mathcal{F}_x^{-1} & 0 \\ 0 & 0 & \mathcal{F}_x^{-1} \end{bmatrix} \begin{bmatrix} \hat{n}_1 & -\hat{n}_2/a & -\hat{n}_1\hat{n}_3/a \\ \hat{n}_2 & \hat{n}_1/a & -\hat{n}_2\hat{n}_3/a \\ \hat{n}_3 & 0 & a \end{bmatrix} \times \begin{bmatrix} \mathcal{F}_x & 0 & 0 \\ 0 & \mathcal{F}_x & 0 \\ 0 & 0 & \mathcal{F}_x \end{bmatrix}, \quad (12)$$

with wavefield decomposition given by

$$\mathbf{Q}^{-1} = \begin{bmatrix} \mathcal{F}_x^{-1} & 0 & 0 \\ 0 & \mathcal{F}_x^{-1} & 0 \\ 0 & 0 & \mathcal{F}_x^{-1} \end{bmatrix} \begin{bmatrix} \hat{n}_1 & \hat{n}_2 & \hat{n}_3 \\ -\hat{n}_2/a & \hat{n}_1/a & 0 \\ -\hat{n}_1\hat{n}_3/a & -\hat{n}_2\hat{n}_3/a & a \end{bmatrix} \times \begin{bmatrix} \mathcal{F}_x & 0 & 0 \\ 0 & \mathcal{F}_x & 0 \\ 0 & 0 & \mathcal{F}_x \end{bmatrix}, \quad (13)$$

where \hat{n}_1 , \hat{n}_2 , and \hat{n}_3 can be replaced by the normalized wavenumbers \hat{k}_1 , \hat{k}_2 , and \hat{k}_3 .

To demonstrate the action of the composition \mathbf{Q} and decomposition \mathbf{Q}^{-1} operators, we consider the decomposition and recomposition of an elastic wavefield in a homogeneous, isotropic 2D medium as shown in Figure 1. The model consists of a P-wave velocity of 1700 m/s and an S-wave velocity of 981 m/s. Here, the x - and z -components are shown for a time slice at 0.35 s after a displacement

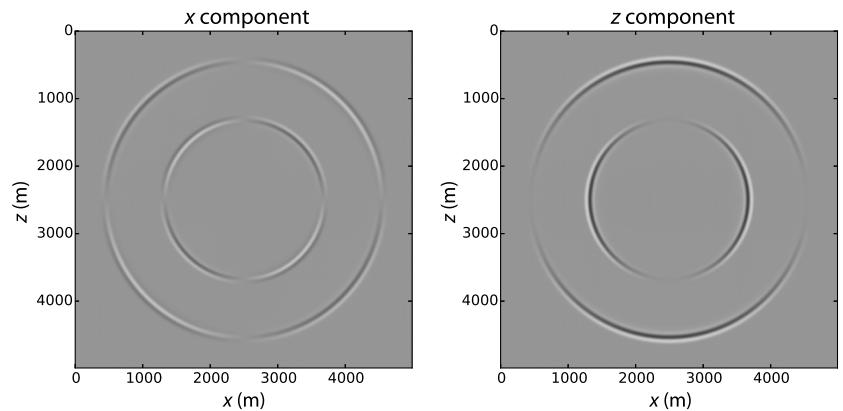


Figure 1. The x and z input data components for an elastic-wave propagating in a homogeneous isotropic medium. The records were generated using elastic finite-difference modeling.

source was fired in the center of the model. In Figure 2, the separated P- and S-wavefields are shown after application of the decomposition operator, whereas the recomposed wavefields are shown in Figure 3.

It is interesting to compare the operator \mathbf{Q}^{-1} with the Helmholtz decomposition operator, which separates an elastic wavefield into curl-free ($\partial u_1/\partial x_1 + \partial u_2/\partial x_2 + \partial u_3/\partial x_3$) and divergence-free components $((\partial u_3/\partial x_2 - \partial u_2/\partial x_3)i + (\partial u_1/\partial x_3 - \partial u_3/\partial x_1)j + (\partial u_2/\partial x_1 - \partial u_1/\partial x_2)k$). Writing this as a linear operator, we have

$$\mathbf{H}^{-1} = \mathcal{F}_x^{-1} \begin{bmatrix} ik_1 & ik_2 & ik_3 \\ 0 & -ik_3 & ik_2 \\ ik_3 & 0 & -ik_1 \\ -ik_2 & ik_1 & 0 \end{bmatrix} \mathcal{F}_x, \quad (14)$$

which (aside from being a rectangular matrix) is somewhat similar to an unnormalized version of the decomposition operator derived from the Christoffel equation. This normalization is important for stability when we consider recursive application of the operator and its inverse during propagation. Sun et al. (2011) examine the changes in amplitude introduced by separating elastic wavefields via Helmholtz decomposition. This amplitude change is a significant problem be-

cause it alters the amplitude ratio between PP- and PS-images, which can greatly affect any subsequent interpretation. The fact that the Christoffel formulation results in a separation operator with columns that form an orthonormal basis means that decomposition and recombination steps are unitary operations that leave the relative amplitudes between wave modes unaltered. Another important distinction is that for 3C data in three spatial dimensions, Helmholtz decomposition provides a single component curl-free wavefield and a 3C divergence free wavefield, whereas the Christoffel-based formulation provides three mutually orthogonal scalar wavefields, even for anisotropic media. This property makes the imaging condition straightforward to design — We can correlate every combination of source and receiver sides wave mode to form a 9C image or, as is done in this study, we can choose to only correlate the source-side P-wavefield with the receiver-side P- and S-wavefields to form a 2C image.

Elastic shot-profile modeling and migration

We incorporate the wavefield decomposition, extrapolation, and recombination operators derived in the previous section to model elastic wavefields. Our approach closely follows the derivation of the forward and adjoint one-way wave-equation operators for acoustic least-squares migration given by Rickett (2003) and Kaplan et al. (2010). Here, we extend shot-profile modeling and migration to the elastic case. We begin with the first-order Born approximation:

$$\mathbf{u}(\mathbf{x}^g, \omega) \sim \omega^2 \times \int_{-\infty}^{\infty} G_0(\mathbf{x}^g | \mathbf{x}) \mathbf{M}(\mathbf{x}) G_0(\mathbf{x} | \mathbf{x}^s) d\mathbf{x}, \quad (15)$$

where \mathbf{x} indicates all subsurface positions, \mathbf{x}^g indicates all receiver positions, and \mathbf{x}^s indicates the position of a source. Here, $G_0(\mathbf{x} | \mathbf{x}^s)$ is the 3×1 Green's function response at all subsurface positions due to an impulse at the position of the source (representing all modes of incident wavefield), $\mathbf{M}(\mathbf{x})$ is a 3×3 matrix of scattering potentials, and $G_0(\mathbf{x}^g | \mathbf{x})$ is the 3×3 Green's function response at the receiver positions due to an impulse generated at all subsurface positions (scattering all combinations of incident and scattered wave modes). The recorded data components at all receiver locations for a given frequency ω are denoted by $\mathbf{u}(\mathbf{x}^g, \omega)$. A complete description of the perturbations in physical properties that comprise the 3×3 scattering potential $\mathbf{M}(\mathbf{x})$ is provided by Stolt and Weglein (2012). Practically, equation 15 implies downward continuation of the source wavefield, multiplication with the reflectivity, followed by upward continuation of the scattered wavefield to the receiver datum.

It is perhaps more helpful to represent equation 15 in terms of a series of discrete linear operations that can be readily programmed in a computer. To derive the forward operator for a single source and single frequency, we begin by downward continuing the source wavefield by

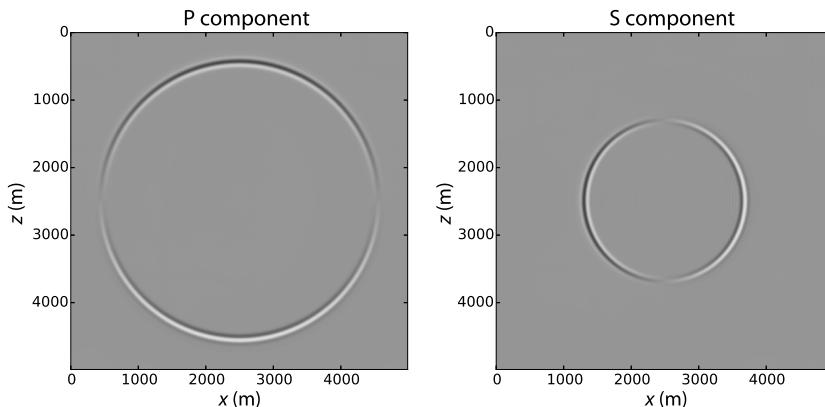


Figure 2. The decomposed P- and S-components after applying the wavefield decomposition operator \mathbf{Q}^{-1} to the data components U^x and U^y .

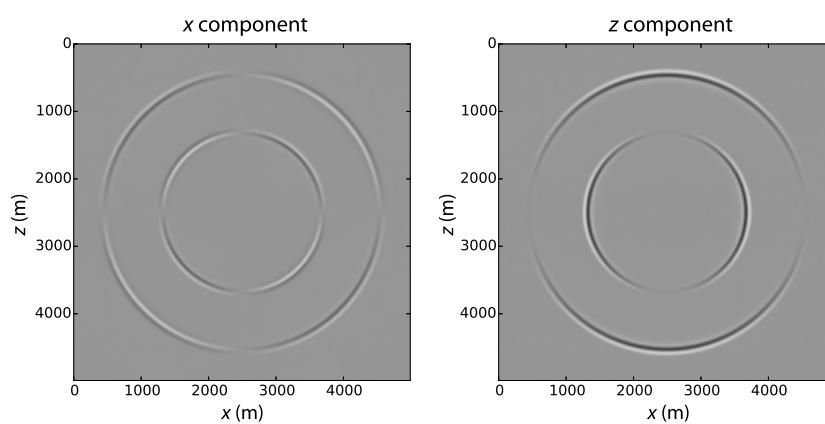


Figure 3. The recomposed x and y components after applying the wavefield recombination operator \mathbf{Q} to the wavefield components U^P and U^S .

$$\begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \dots \\ \Psi_{N_z} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1^{-1} & 0 & 0 & \dots & 0 \\ 0 & \mathbf{Q}_2^{-1}\mathbf{P}_1 & 0 & \dots & 0 \\ 0 & 0 & \mathbf{Q}_3^{-1}\mathbf{P}_2\mathbf{P}_1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mathbf{Q}_{N_z}^{-1}\mathbf{P}_{N_z-1}\dots\mathbf{P}_1 \end{bmatrix}$$

$$\times \begin{bmatrix} \mathbf{u}_1^s \\ \mathbf{u}_1^s \\ \mathbf{u}_1^s \\ \dots \\ \mathbf{u}_1^s \end{bmatrix}, \quad (16)$$

where $\mathbf{P}_i = \mathbf{Q}_i \mathbf{R}_i \mathbf{Q}_i^{-1}$ represents the elastic propagation using the medium properties of layer i , \mathbf{u}_1^s represents one frequency of the vector-valued source wavelet at the surface of the earth, and $\Psi_i = [\Psi_i^p \Psi_i^{s1} \Psi_i^{s2}]^T$ are the source-side wavefield potentials for a single frequency at depth interval i of the model. To obtain data at the surface of the earth, we propagate the scattered wavefield upward by

$$\mathbf{u}_1^g = [\mathbf{Q}_1 \quad \mathbf{P}_1^{-1}\mathbf{Q}_2 \quad \dots \quad \mathbf{P}_1^{-1}\dots\mathbf{P}_{N_z-1}^{-1}\mathbf{Q}_{N_z}] \begin{bmatrix} \mathbf{M}_1 \Psi_1 \\ \mathbf{M}_2 \Psi_2 \\ \dots \\ \mathbf{M}_{N_z} \Psi_{N_z} \end{bmatrix}, \quad (17)$$

where

$$\mathbf{M}_i = \begin{bmatrix} \mathbf{m}_i^{PP} & \mathbf{m}_i^{S_1P} & \mathbf{m}_i^{S_2P} \\ \mathbf{m}_i^{PS_1} & \mathbf{m}_i^{S_1S_1} & \mathbf{m}_i^{S_2S_1} \\ \mathbf{m}_i^{PS_2} & \mathbf{m}_i^{S_1S_2} & \mathbf{m}_i^{S_2S_2} \end{bmatrix} \quad (18)$$

are the multiparameter image amplitudes at depth interval i of the model. For the adjoint operator, we begin by propagating the receiver wavefield (data) into the earth by

$$\begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \dots \\ \Phi_{N_z} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1^\dagger & 0 & 0 & 0 \\ 0 & \mathbf{Q}_2^\dagger(\mathbf{P}_1^{-1})^\dagger & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \mathbf{Q}_{N_z}^\dagger(\mathbf{P}_{N_z-1}^{-1})^\dagger\dots(\mathbf{P}_1^{-1})^\dagger \end{bmatrix}$$

$$\times \begin{bmatrix} \mathbf{u}_1^g \\ \mathbf{u}_1^g \\ \mathbf{u}_1^g \\ \dots \\ \mathbf{u}_1^g \end{bmatrix}, \quad (19)$$

where $\Phi_i = [\phi_i^p \phi_i^{s1} \phi_i^{s2}]^T$ are the receiver-side wavefield potentials for a single frequency at depth interval i of the model. We then correlate using the outer product of the source and receiver wavefields as

$$\begin{bmatrix} \tilde{\mathbf{M}}_1 \\ \tilde{\mathbf{M}}_2 \\ \dots \\ \tilde{\mathbf{M}}_{N_z} \end{bmatrix} = \begin{bmatrix} \Psi_1^* & 0 & 0 & 0 \\ 0 & \Psi_2^* & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \Psi_{N_z}^* \end{bmatrix} \begin{bmatrix} \Phi_1^T \\ \Phi_2^T \\ \dots \\ \Phi_{N_z}^T \end{bmatrix} \quad (20)$$

to obtain a multiparameter image $\tilde{\mathbf{m}}$. To avoid confusing notation equations 16–20, we only consider vertical propagation in a 1D earth. Extending these equations in the x - and y -dimensions is achieved by extending each subvector by a factor of N_{xy} , and block diagonalizing each submatrix increasing their rank by a factor of N_{xy} . Furthermore, to extend the equations to the case of multiple frequencies and multiple shots, the forward operator begins by spraying the reflectivities to N_ω frequencies and N_{shot} shots, whereas the adjoint ends by summing over all frequencies and shots (Rickett, 2003).

To extend the extrapolation operator \mathbf{R} to the case of laterally inhomogeneous velocity, we incorporate the phase-shift-plus interpolation (PSPI) and split-step corrections (Gazdag and Sguazzero, 1984; Stoffa et al., 1990; Xie and Wu, 2005). An important consideration for one-way extrapolation is the accuracy of the operator versus propagation angle. Biondi (2002) provides an analysis of the phase error for several pseudodifferential operators as a function of propagation angle. Notably, the split step correction is accurate (<0.2 radians phase error) to approximately 45° for a reference velocity with an error of 200 m/s. By combining the split step and PSPI corrections, we are able to narrow the range of reference velocity error significantly. For the examples, we use five reference velocities. A detailed derivation of the forward and adjoint acoustic shot profile split-step migration operators is provided by Kaplan et al. (2010). To apply the separation operator \mathbf{Q}^{-1} in laterally inhomogeneous media, we use the average P- and S-wave velocities for a given layer. The accuracy of the separation could be further improved via a spatial windowing approach following Bale and Margrave (2004).

It is important to note that in this work, we do not consider the effect of the free surface. In our numerical examples, we use data generated using elastic finite differences with an absorbing boundary condition above the recording surface. Effectively, this assumes that the recorded wavefield consists only of upgoing energy. To compensate for the effects of upgoing and downgoing energies at the recording surface, a projection onto the downgoing wavefield must be made as an initial step (Bale, 2006). A further complication in least-squares migration is the estimation of the source wavelet. In our numerical examples, we consider the case of a known source wavelet. A method to estimate an unknown source wavelet from the data can be found in Zhang et al. (2016).

In the following discussion of the least-squares formulation, we refer to the forward and adjoint operations more simply as $\mathbf{u} = \mathbf{L}\mathbf{m}$ and $\tilde{\mathbf{m}} = \mathbf{L}^\dagger\mathbf{u}$, respectively, in which it is understood that vector \mathbf{m} consists of submatrices and vector \mathbf{u} consists of subvectors.

Finally, to model the complex amplitude behavior of elastic data, including polarity reversal of PS-images at an incidence angle of zero, we use an extended imaging condition. Of the numerous methods to compute extended images, we selected a mapping from subsurface offset image gathers to opening angle image gathers via the radial trace transform as outlined in Rickett and Sava (2002) as we find it provides smooth, artifact-free angle gathers.

The method uses the Fourier transform to map from subsurface offset to the half-opening angle γ via

$$\tan \gamma = \frac{k_h}{k_z}, \quad (21)$$

where k_h and k_z are the subsurface offset and depth wavenumbers, respectively.

Incorporating an angle-dependent reflectivity into the adjoint operator \mathbf{L}^\dagger involves

- 1) computing the imaging condition for discrete spatial lags between the source and receiver wavefields,
- 2) taking the Fourier transform of the subsurface offset image along the depth and offset axes,
- 3) performing the mapping from subsurface offset wavenumber to angle via equation 21, and
- 4) taking the inverse Fourier transform along the depth axis,

whereas the forward operator \mathbf{L} involves

- 1) taking the Fourier transform of the image along the depth axis,
- 2) performing the adjoint mapping from the angle to the subsurface offset wavenumber via equation 21,
- 3) taking the inverse Fourier transform of the subsurface offset image along the depth and offset axes, and

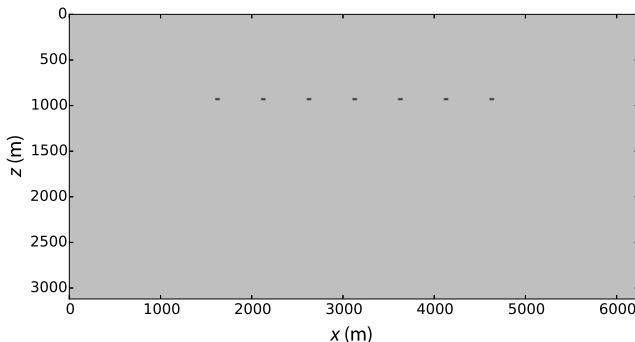


Figure 4. The P-wave velocity model used to create finite-difference synthetic data. The background P-wave velocity is 2200 m/s, whereas the seven diffractors have a velocity of 2500 m/s. The S-wave velocity model corresponds to this model by a scale factor of $1/\sqrt{3}$.

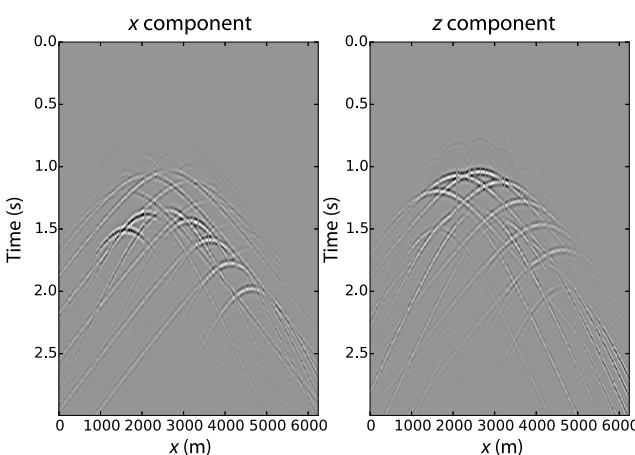


Figure 5. The x - and z -components for a synthetic shot gather at $x = 2500$ m generated using elastic finite-difference modeling. The data have been randomly decimated by 30% to simulate the missing receivers.

- 4) computing the receiver wavefields at depth for discrete spatial lags between the source and the image point.

The above method parameterizes the images as a function of the half-opening angle between source and receiver wavefields. For PP-reflections, the equivalence of source- and receiver-side velocities means that the opening angle is bisected by the reflector normal vector, making the half-opening angle equivalent to the source-side incidence angle with respect to the reflector normal. For PS-reflections, there is no such equivalence and the half-opening angle does not adequately describe the zero incidence angle in which polarity reversal is expected (except for the special case of a flat-lying reflector). Even though a postprocessing step can be used to correctly position the polarity reversal for PS-reflections prior to stacking (Rosales et al., 2007), we observe that simply using the half-opening angle provides an adequate extended imaging condition for ELSM.

Least-squares formulation

We pose elastic imaging as an inverse problem with the objective function

$$J = \|\mathbf{T}(\mathbf{L}\mathbf{m} - \mathbf{u})\|_2^2 + \|\boldsymbol{\mu}^T \mathbf{m}\|_2^2, \quad (22)$$

where \mathbf{T} is a diagonal sampling operator with the ones in place of observations and the zeroes in place of missing traces, \mathbf{m} represents the multicomponent image, \mathbf{u} represents the multicomponent data, and $\boldsymbol{\mu}$ is a multicomponent trade-off parameter to control the level to fit the data versus the level to minimize the $l - 2$ norm of the model. The components of $\boldsymbol{\mu}$ can be fixed for all components of the image (effectively treating $\boldsymbol{\mu}$ as a scalar), or it can be tailored to individually constrain components.

Because wavefield crosstalk artifacts are migrated to an incorrect velocity, they will appear as dipping energy in angle gathers. For this reason, we can regularize the inversion with an operator that weights up this energy in each angle gather

$$J = \|\mathbf{T}(\mathbf{L}\mathbf{m} - \mathbf{u})\|_2^2 + \|\boldsymbol{\mu}^T \mathbf{D}\mathbf{m}\|_2^2, \quad (23)$$

or by a change of variables $\mathbf{z} = \mathbf{D}\mathbf{m}$, we write

$$J = \|\mathbf{T}(\mathbf{L}\mathbf{S}\mathbf{z} - \mathbf{u})\|_2^2 + \|\boldsymbol{\mu}^T \mathbf{z}\|_2^2 \quad (24)$$

such that \mathbf{S} ($\approx \mathbf{D}^{-1}$) is a lateral smoother that removes strongly dipping energy in the angle gathers (Equation 24 is an approximate form of equation 23). We minimize equation 24 using CG with the forward operator \mathbf{TLS} and the adjoint operator $\mathbf{S}^\dagger \mathbf{L}^\dagger \mathbf{T}$. After the final iteration, we substitute $\mathbf{m} = \mathbf{Sz}$.

EXAMPLES

Diffractor example

As a first example, we demonstrate the ability of ELSM to interpolate and separate wavefield components using the velocity model shown in Figure 4. A single shot gather, shown in Figure 5, was generated using finite-difference modeling, and 30% of the traces were zeroed to simulate an irregular receiver geometry. In all examples, a compressional source wavelet was used for the numerical modeling and the imaging. Next, we applied the adjoint operator to obtain the migrated images shown in Figures 6 and 7. The adjoint

operator is not able to perfectly separate elastic wavefields, leaving some wavefield crosstalk artifacts in images and artifacts related to the irregular receiver grid. In the PP-image, the PS crosstalk artifacts are propagated with too high of a velocity, sending them too deep into the earth. Conversely, in the PS-image, the PP crosstalk artifacts are propagated with too low of a velocity, giving them an imaging depth that is too shallow. For the next step, we performed 20 iterations of ELSM for one shot with a purely data fitting objective function ($\mu = \mathbf{0}$ in equation 24). The results are shown in Figures 8 and 9. Even without any regularization, the data fitting is able to reduce crosstalk artifacts in the images, especially in the PS-image. The ELSM algorithm presents an opportunity to simultaneously interpolate and separate elastic-wave modes. As an example, after first

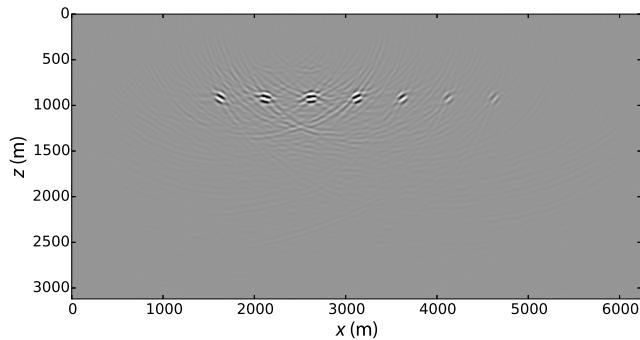


Figure 6. The PP-image obtained by elastic migration (the adjoint operator) for a single shot at $x = 2500$ m.

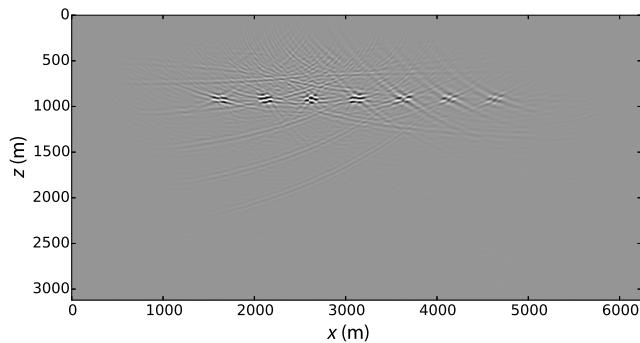


Figure 7. The PS-image obtained by elastic migration (the adjoint operator) for a single shot at $x = 2500$ m.

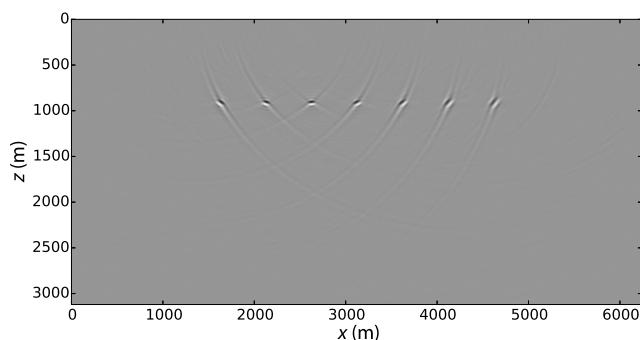


Figure 8. The PP-image obtained by 20 iterations of ELSM for a single shot at $x = 2500$ m.

fitting the data in the exercise above, we can then forward model a pure reflected P-wavefield by $\mathbf{L}[\mathbf{m}_{\text{PP}} \quad \mathbf{0} \quad \mathbf{0}]^T$ as shown in Figure 10, or similarly forward model a pure reflected S-wavefield by $\mathbf{L}[\mathbf{0} \quad \mathbf{0} \quad \mathbf{m}_{\text{PS}}]^T$ as shown in Figure 11. Interestingly, a purely data-fitting-based approach such as this is not able to perfectly sep-

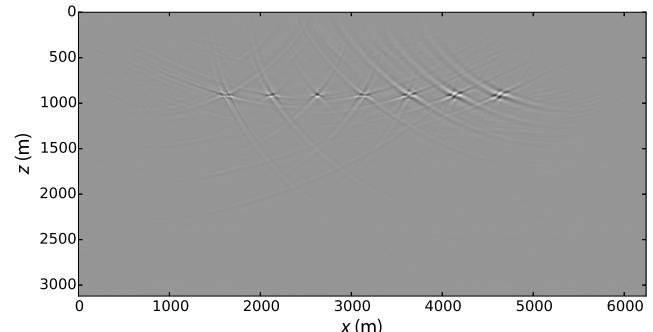


Figure 9. The PS-image obtained by 20 iterations of ELSM for a single shot at $x = 2500$ m.

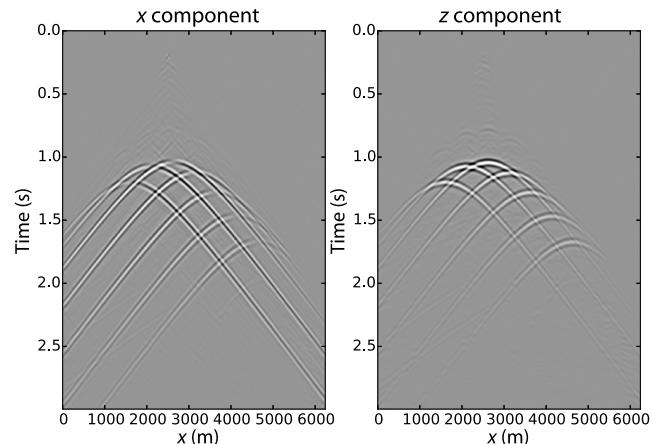


Figure 10. The predicted x - and z -components of the interpolated and wavefield separated PP-wavefield for a shot gather at $x = 2500$ m generated by the forward operator following 20 iterations of ELSM.

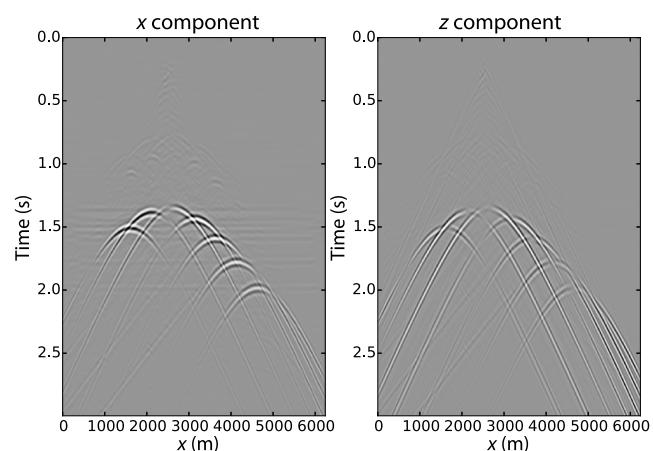


Figure 11. The predicted x - and z -components of the interpolated and wavefield separated PS-wavefield for a shot gather at $x = 2500$ m generated by the forward operator following 20 iterations of ELSM.

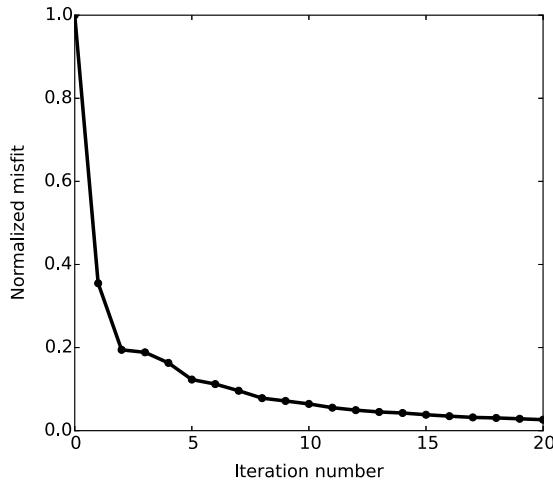


Figure 12. A normalized misfit versus iteration number for 20 iterations of ELSM.

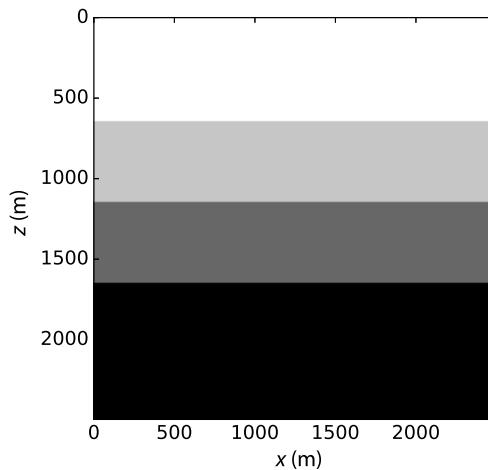


Figure 13. The P-wave velocity model used to create finite-difference synthetic data.

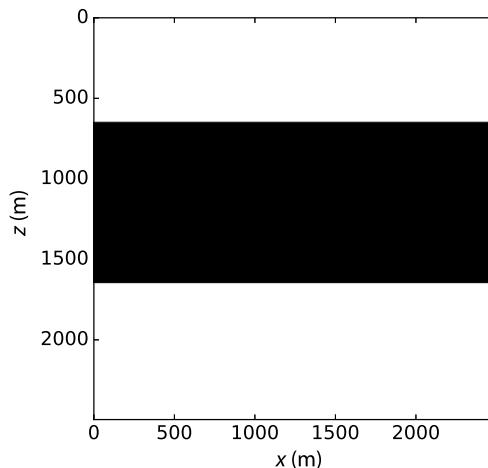


Figure 14. The S-wave velocity model used to create finite-difference synthetic data.

erate the wavefields. Additional constraints that penalize crosstalk energy might benefit the solution. Finally, Figure 12 shows the normalized misfit versus the iteration number.

Flat layer example

Next, we consider a model consisting of flat layers with a depth-variable V_P/V_S ratio. Figures 13 and 14 show the P- and S-wave velocity models, respectively. A shot gather, shown in Figure 15, was generated using finite-difference modeling and random noise was added to the data. Next, the adjoint operator was applied to obtain the PP- and PS-images as shown in Figures 16 and 17. Crosstalk artifacts are clearly visible in the adjoint images. In the PP-image, the PS-energy has been extrapolated using too high of a velocity (placing it too deep), whereas in the PS-image, the PP-energy has been extrapolated using too low of a velocity (placing it too shallow). The results of 10 iterations of ELSM using $\mu = [0.001 \ 0.001 \ 0.001]^T$ are shown in Figures 18 and 19. In both images, the resolution has been improved and the crosstalk artifacts have been largely attenu-

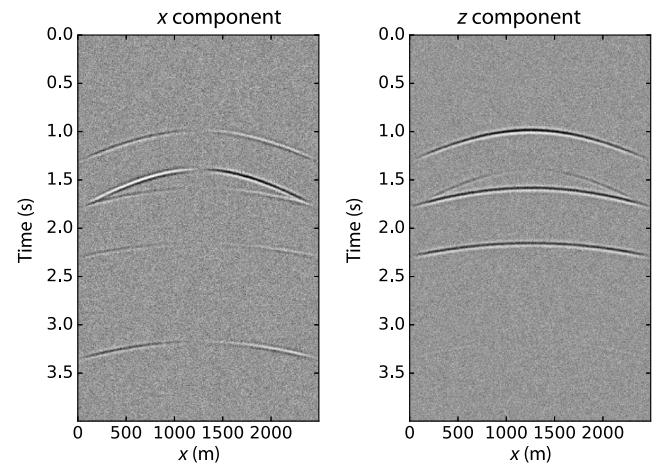


Figure 15. The x - and z -components for a synthetic shot gather at $x = 1250$ m generated using elastic finite-difference modeling with random noise added.

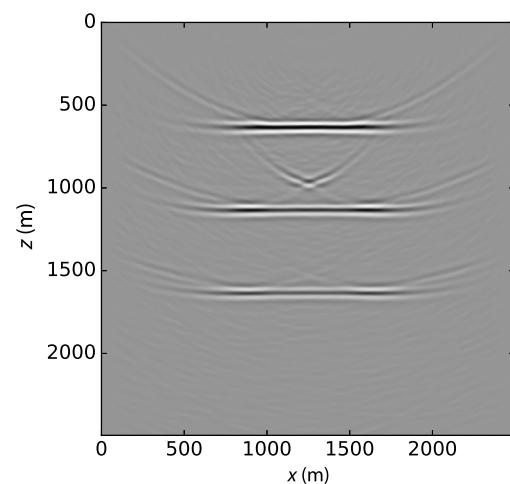


Figure 16. The PP-image obtained by elastic migration (the adjoint operator) for a single shot at $x = 1250$ m.

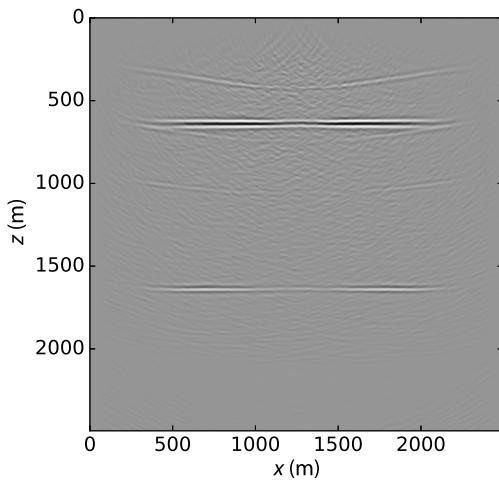


Figure 17. The PS-image obtained by elastic migration (the adjoint operator) for a single shot at $x = 1250$ m.

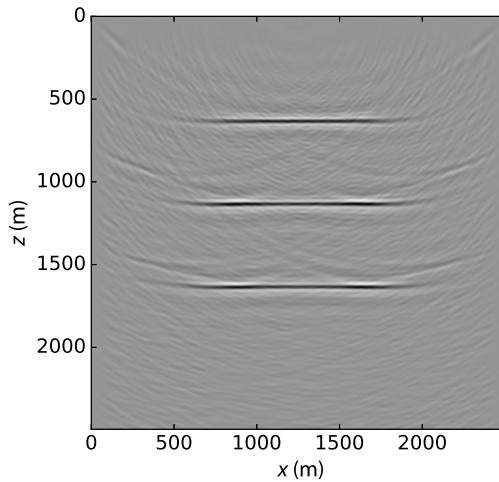


Figure 18. The PP-image obtained by 10 iterations of ELSM for a single shot at $x = 1250$ m.

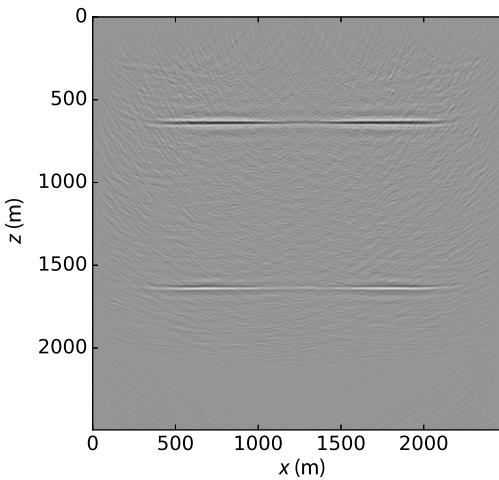


Figure 19. The PS-image obtained by 10 iterations of ELSM for a single shot at $x = 1250$ m.

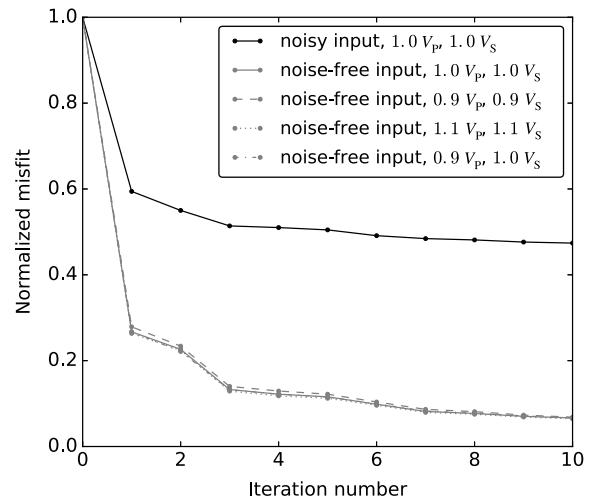


Figure 20. A normalized misfit versus iteration number for 10 iterations of ELSM. The solid black line shows convergence when added random noise is present on the input data. The gray curves display trials of the algorithm using different combinations of velocity error. The solid gray line shows convergence using the true velocity, whereas as the dashed lines show convergence when using different combinations of incorrect velocities. For a single shot experiment, the convergence is relatively insensitive to small velocity errors.

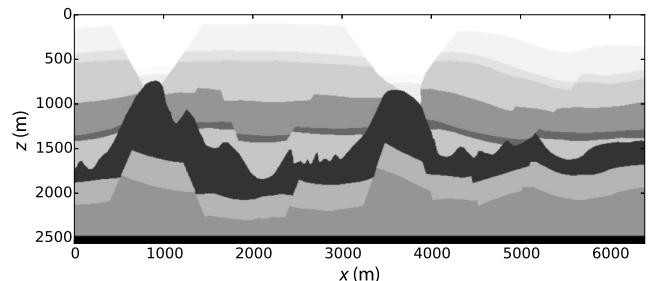


Figure 21. A portion of the BP 2.5D P-wave velocity model. The S-wave velocity corresponds to this model by a scale factor of $1/\sqrt{3}$.

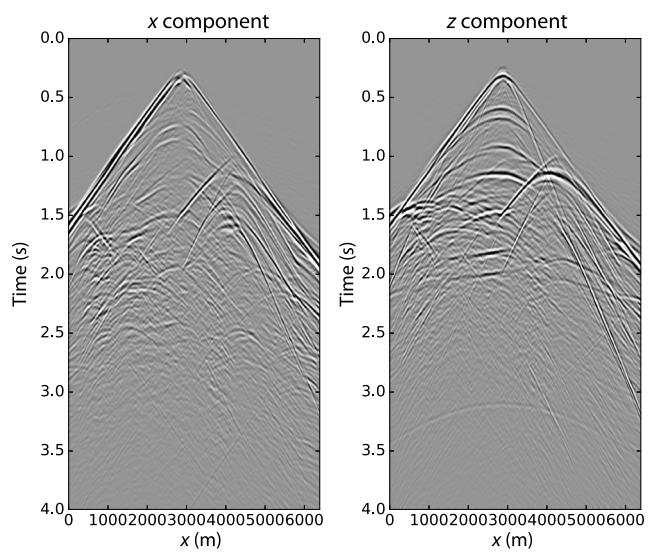


Figure 22. The x - and z -components for a synthetic shot gather at $x = 2880$ m generated using elastic finite-difference modeling.

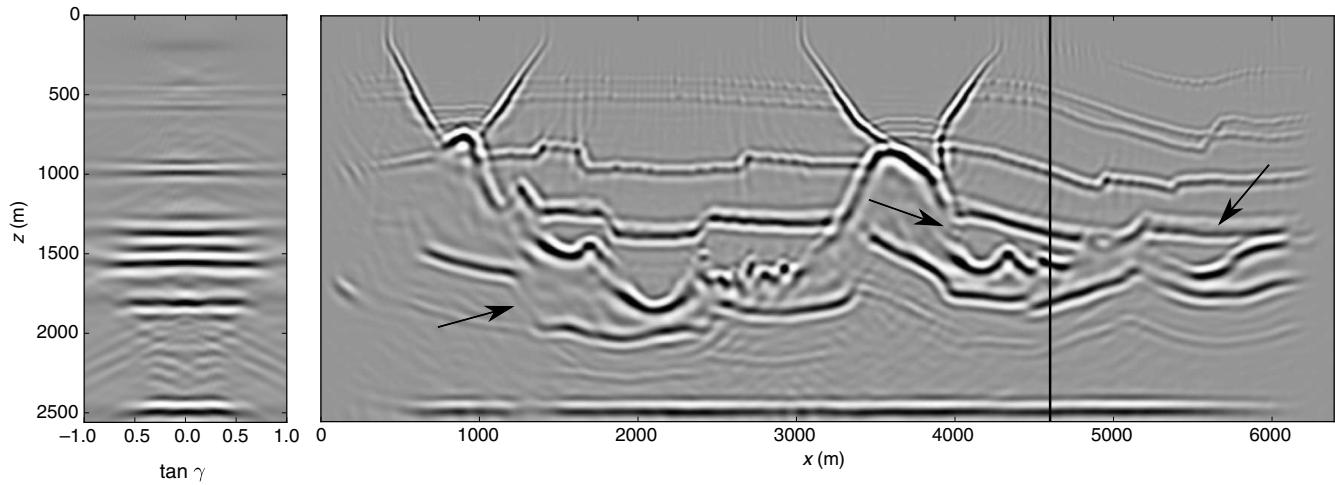


Figure 23. The PP-image and an angle gather from $x = 4600$ m obtained by elastic migration (the adjoint operator). The arrows indicate regions of the model that have been poorly imaged.

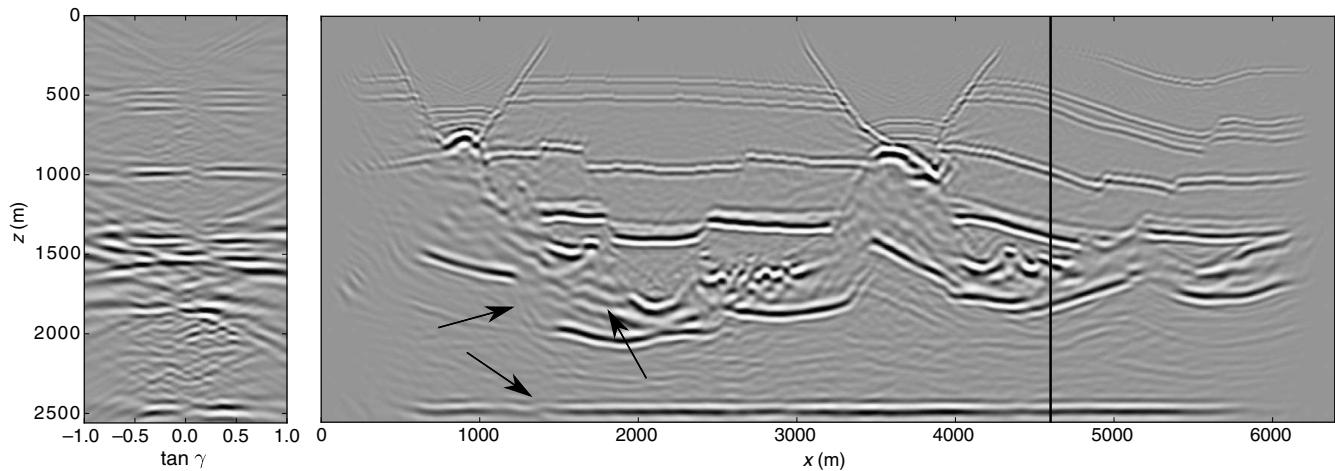


Figure 24. The PS-image and an angle gather from $x = 4600$ m obtained by elastic migration (the adjoint operator). The arrows indicate regions of the model that have been poorly imaged.

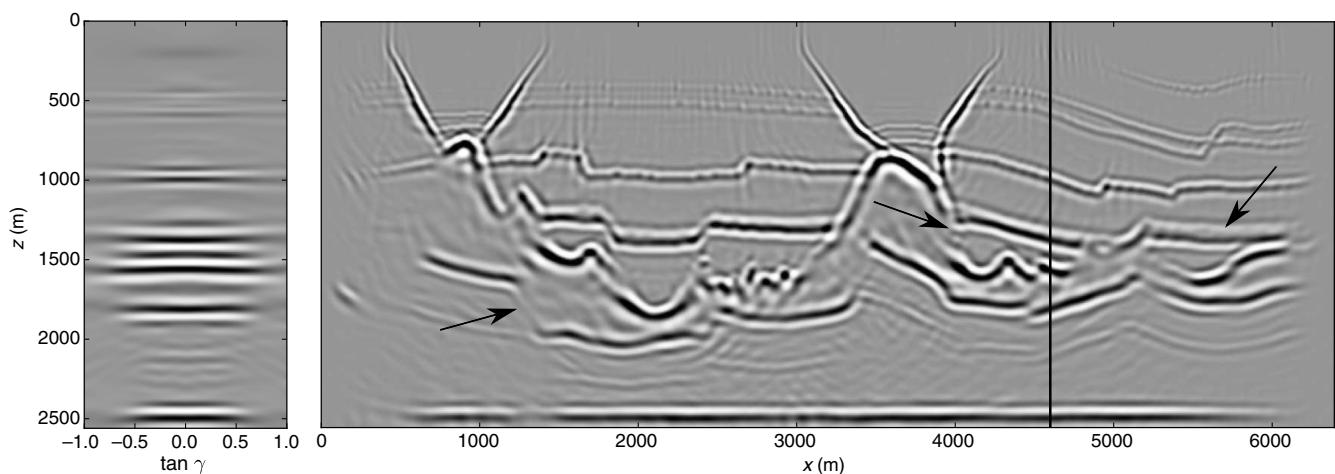


Figure 25. The PP-image and an angle gather from $x = 4600$ m obtained by elastic migration (the adjoint operator). A smoothing filter has been applied across angles to attenuate dipping crosstalk energy. Notice that this filter has little effect on the overall stack. The arrows indicate regions of the model that have been poorly imaged.

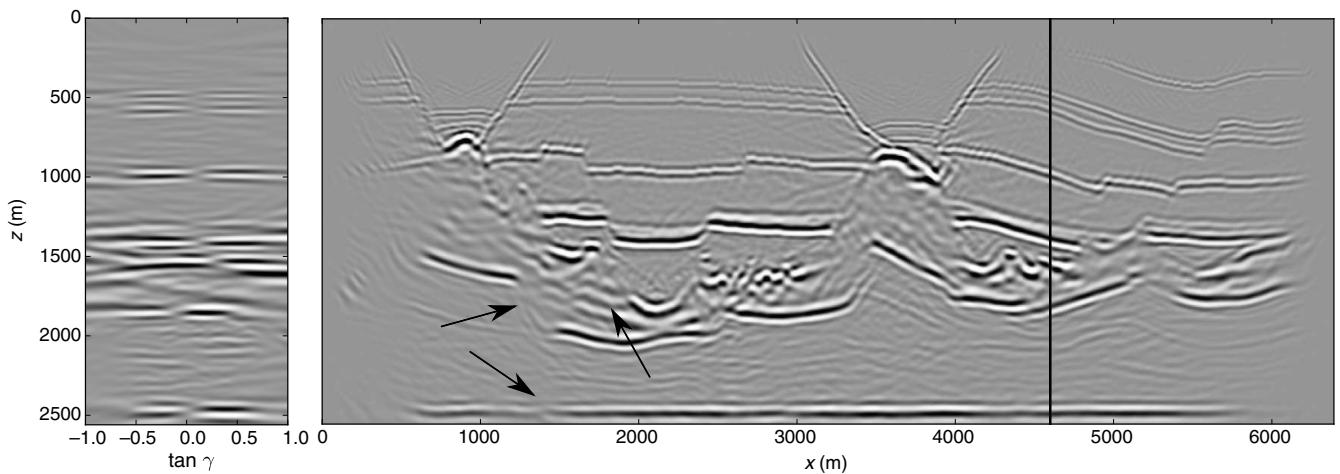


Figure 26. The PS-image and an angle gather from $x = 4600$ m obtained by elastic migration (the adjoint operator). A smoothing filter has been applied across angles to attenuate dipping crosstalk energy. Notice that this filter has little effect on the overall stack. The arrows indicate regions of the model that have been poorly imaged.

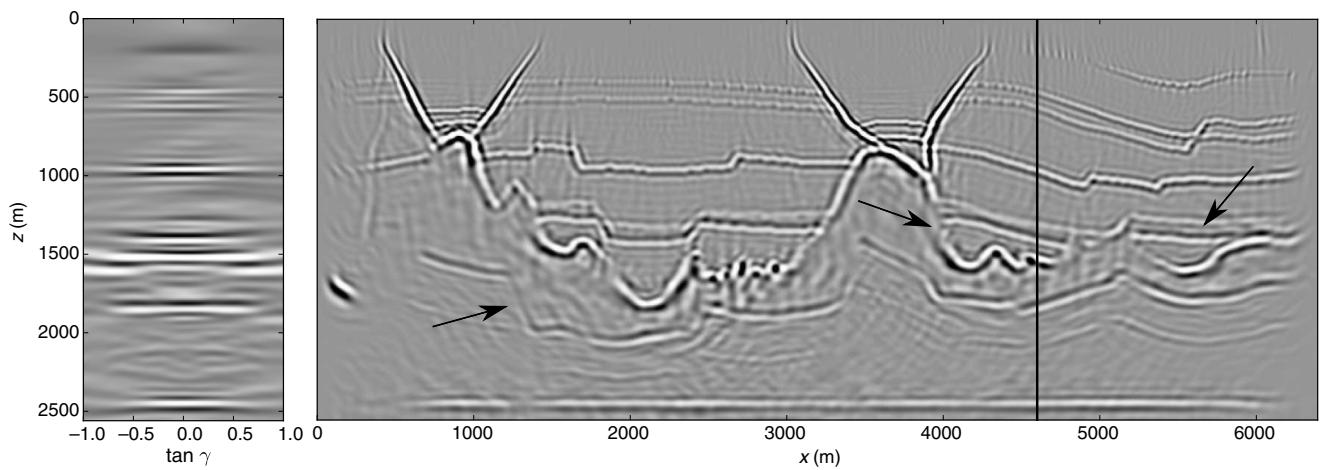


Figure 27. The PP-image and an angle gather from $x = 4600$ m obtained by 10 iterations of ELSM. The arrows indicate regions of the model in which the image has been improved.

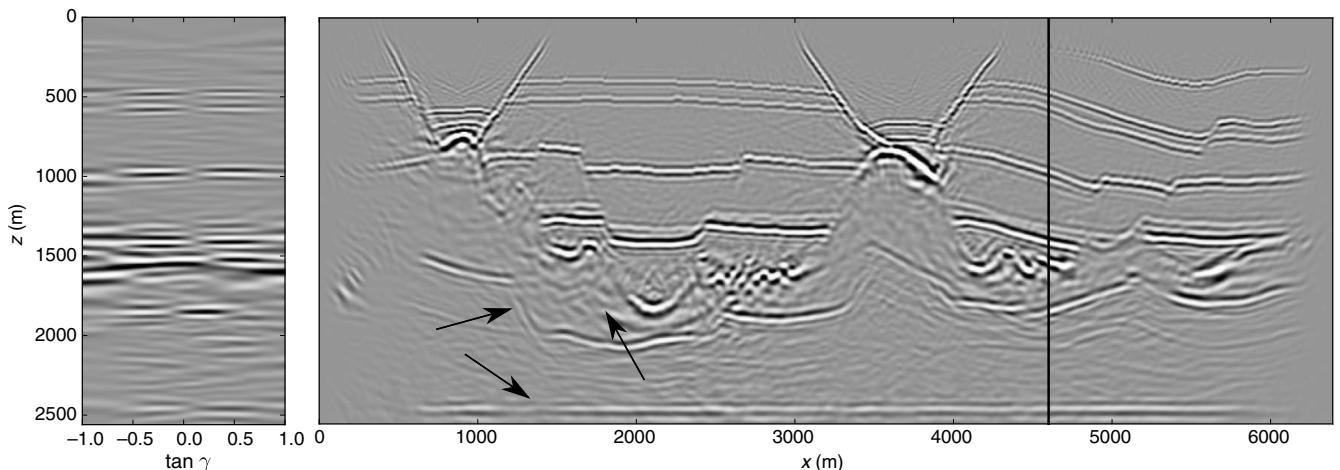


Figure 28. The PS-image and an angle gather from $x = 4600$ m obtained by 10 iterations of ELSM. The arrows indicate regions of the model in which the image has been improved.

ated. The PP-crosstalk artifact at approximately $z = 1000$ m in the adjoint PS-image (Figure 17) could be mistakenly interpreted as a true PS-reflection corresponding to the event at a similar depth in the PP-adjoint image. This example demonstrates the ability of ELSM to mitigate crosstalk artifacts that could potentially harm an interpretation.

Figure 20 shows the misfit as a function of iteration number for the case of added random noise (the solid black line), as well as for four trials considering velocity perturbations from the true models using noise-free input data. In the first trial (the solid gray line), the correct velocity was used, whereas in the other trials, the velocities were varied by $\pm 10\%$ from the true models. It is interesting to note that for a single shot experiment, the convergence is largely insensitive to these perturbations in the velocity models. When multiple sources are used in shot-profile LSM, the convergence is more sensitive to the accuracy of the velocity model as a consequence of stacking/spraying multiple improperly migrated/demigrated shot gathers. This dependence is relaxed through the use of an extended image.

BP 2.5D model example

For a more complicated example, we next consider a modified version of the BP 2.5D model (Etgen and Regone, 1998) (the model increments were changed from $dx = 12.5$ m, $dz = 12.5$ m to $dx = 8$ m, $dz = 8$ m prior to modeling, and an S-wave velocity model was created as $1/\sqrt{3}$ times the P-wave velocity). The P-wave velocity model is shown in Figure 21. This model contains high-velocity contrasts that challenge the accuracy of our wavefield separation and propagation operators, and contains highly variable illumination patterns that make least-squares migration an ill-posed inverse problem. We generated 130 shot gathers using elastic finite-difference modeling. The horizontal and vertical components for a common shot gather at $x = 2880$ m are shown in Figure 22. The results of applying the adjoint operator and stacking are shown in Figures 23 and 24. To demonstrate the effect of the preconditioning operator written in equation 24, we applied it to the adjoint image volumes. The result of this filter is shown in Figures 25 and 26. It is interesting to note that even though the smoothing filter was able to attenuate much of the dipping energy within each angle gather, it

produced little effect on the overall stack. The results of applying 10 iterations of preconditioned ELSM with $\mu = [0.1 \quad 0.1 \quad 0.1]^T$ are shown in Figures 27 and 28. Ideally, the amplitude of the bottom-most reflector should be constant, but variations in illumination cause the amplitude to fluctuate. Comparing the adjoint and ELSM PP-images (Figures 23 and 27), we see that the larger amplitude variations of this bottom reflector have been partially corrected, and improvements to the spatial resolution can be seen throughout the image. Comparing the adjoint and ELSM PS-images (Figures 24 and 28), we see a similar improvement to the illumination of the base reflector, especially the continuity of the bottom-most reflector at $x = 1200$ m. Finally, Figure 29 shows the normalized cost versus iteration number for 10 iterations of ELSM. Other improvements include the attenuation of wavefield crosstalk artifacts within the high-velocity body, as well as improvements to the spatial resolution of the surrounding structure. Rickett (2003) applies a purely data-fitting-based-least-squares migration to the BP 2.5D model, showing an increase in the noise level with least-squares migration and instead advocates using an approximation to the diagonal of the Hessian using the forward and adjoint operators common to least-squares migration. Illumination compensation such as this could also be applied using the elastic forward and adjoint operators as described in this paper.

CONCLUSION

Least-squares migration algorithms attempt to fit recorded data with predictions generated from a migrated image. By improving the accuracy of the migration operator to include elastic-wave propagation, we expect to improve the ability of least-squares migration to fit reflection amplitudes. Our examples demonstrate that least-squares migration can improve the imaging of multicomponent seismic data and that it can also be used for missing trace interpolation and wavefield decomposition. A significant challenge in the migration of elastic wavefields is the mitigation of crosstalk artifacts. Even though pure data fitting is able to partially resolve the problem, incorporating additional constraints could further improve the performance of the method.

ACKNOWLEDGMENTS

We thank the sponsors of the Signal Analysis and Imaging Group (SAIG) at the University of Alberta for their financial support and BP for sharing the 2.5D velocity model. We are also grateful for the thoughtful suggestions of J. Etgen, S.-K. Foss, and three anonymous reviewers.

REFERENCES

- Aki, K., and P. G. Richards, 2002, Quantitative seismology: University Science Books 1.
- Bale, R., 2006, Elastic wave-equation depth migration of seismic data for isotropic and azimuthally anisotropic media: Ph.D. thesis, University of Calgary.
- Bale, R., and G. Margrave, 2004, Elastic wave-equation migration for laterally varying isotropic and HTI media: Technical report, CREWES Research Report.
- Bale, R., and R. Stewart, 2002, The impact of attenuation on the resolution of multicomponent seismic data: Technical report, CREWES Research Report.
- Biondi, B., 2002, Stable wide-angle Fourier finite-difference downward extrapolation of 3-d wavefields: Geophysics, **67**, 872–882, doi: [10.1190/1.1484530](https://doi.org/10.1190/1.1484530).
- Cary, P., 2011, True-amplitude ps prestack time migration via 5D interpolation: Presented at the CSEG Conference and Exhibition.

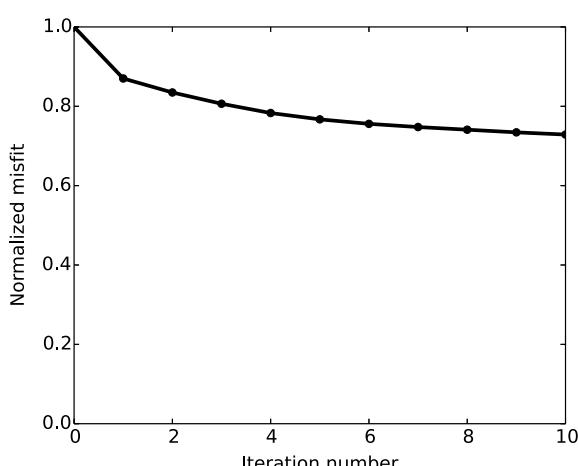


Figure 29. A normalized misfit versus iteration number for 10 iterations of ELSM.

- Chang, W.-F., and G. A. McMechan, 1987, Elastic reverse-time migration: *Geophysics*, **52**, 1365–1375, doi: [10.1190/1.1442249](https://doi.org/10.1190/1.1442249).
- Cheng, J., T. Alkhalifah, Z. Wu, P. Zou, and C. Wang, 2016, Simulating propagation of decoupled elastic waves using low-rank approximate mixed-domain integral operators for anisotropic media: *Geophysics*, **81**, no. 2, T63–T77, doi: [10.1190/geo2015-0184.1](https://doi.org/10.1190/geo2015-0184.1).
- Dai, W., and G. T. Schuster, 2013, Plane-wave least-squares reverse-time migration: *Geophysics*, **78**, no. 4, S165–S177, doi: [10.1190/geo2012-0377.1](https://doi.org/10.1190/geo2012-0377.1).
- Dellinger, J., and J. Etgen, 1990, Wave field separation in two dimensional anisotropic media: *Geophysics*, **55**, 914–919, doi: [10.1190/1.1442906](https://doi.org/10.1190/1.1442906).
- Dutta, G., and G. T. Schuster, 2014, Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation: *Geophysics*, **79**, no. 6, S251–S262, doi: [10.1190/geo2013-0414.1](https://doi.org/10.1190/geo2013-0414.1).
- Etgen, J., and C. Regone, 1998, Strike shooting, dip shooting, widepatch shooting — Does prestack migration care? A model study: 68th Annual International Meeting, SEG, Expanded Abstracts, 66–69.
- Gazdag, J., 1978, Wave equation migration with the phase-shift method: *Geophysics*, **43**, 1342–1351, doi: [10.1190/1.1440899](https://doi.org/10.1190/1.1440899).
- Gazdag, J., and P. Sguazzero, 1984, Migration of seismic data by phase shift plus interpolation: *Geophysics*, **49**, 124–131, doi: [10.1190/1.1441643](https://doi.org/10.1190/1.1441643).
- Guo, Q., and T. Alkhalifah, 2016, Elastic reflection based waveform inversion in isotropic media: 78th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201601194](https://doi.org/10.3997/2214-4609.201601194).
- Ji, J., 2009, An exact adjoint operation pair in time extrapolation and its application in least-squares reverse-time migration: *Geophysics*, **74**, no. 5, H27–H33, doi: [10.1190/1.3173894](https://doi.org/10.1190/1.3173894).
- Kaplan, S. T., P. S. Routh, and M. D. Sacchi, 2010, Derivation of forward and adjoint operators for least-squares shot-profile split-step migration: *Geophysics*, **75**, no. 6, S225–S235, doi: [10.1190/1.3506146](https://doi.org/10.1190/1.3506146).
- Kühl, H., and M. Sacchi, 2003, Least-squares wave-equation migration for AVP/AVA inversion: *Geophysics*, **68**, 262–273, doi: [10.1190/1.1543212](https://doi.org/10.1190/1.1543212).
- Kuo, J., and T. Dai, 1984, Kirchhoff elastic wave migration for the case of noncoincident source and receiver: *Geophysics*, **49**, 1223–1238, doi: [10.1190/1.1441751](https://doi.org/10.1190/1.1441751).
- Lambaré, G., J. Virieux, R. Madariaga, and S. Jin, 1992, Iterative asymptotic inversion in the acoustic approximation: *Geophysics*, **57**, 1138–1154, doi: [10.1190/1.1443328](https://doi.org/10.1190/1.1443328).
- MacLeod, M., M. Hadley, K. Reynolds, and A. Tura, 1999, Multicomponent analysis of OBC data: Presented at the Offshore Technology Conference, Offshore Technology Conference.
- Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64**, 208–221, doi: [10.1190/1.1444517](https://doi.org/10.1190/1.1444517).
- Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: *Geophysics*, **68**, 1371–1379, doi: [10.1190/geo2003-0461.1](https://doi.org/10.1190/geo2003-0461.1).
- Rickett, J. E., and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: *Geophysics*, **67**, 883–889, doi: [10.1190/1.1484531](https://doi.org/10.1190/1.1484531).
- Rosales, D. A., S. Fomel, B. L. Biondi, and P. C. Sava, 2007, Wave-equation angle-domain common-image gathers for converted waves: *Geophysics*, **73**, no. 1, S17–S26, doi: [10.1190/1.2821193](https://doi.org/10.1190/1.2821193).
- Shan, G., and B. Biondi, 2008, Plane-wave migration in tilted coordinates: *Geophysics*, **73**, no. 5, S185–S194, doi: [10.1190/1.2957891](https://doi.org/10.1190/1.2957891).
- Stewart, R., 1990, Joint P and P-SV inversion: Technical Report, CREWES Research Report.
- Stoffa, P. L., J. T. Fokkema, R. M. de Luna Freire, and W. P. Kessinger, 1990, Split-step Fourier migration: *Geophysics*, **55**, 410–421, doi: [10.1190/1.1442850](https://doi.org/10.1190/1.1442850).
- Stolt, R. H., and A. B. Weglein, 2012, Seismic imaging and inversion — Application of linear inverse theory: Cambridge University Press 1.
- Sun, R., G. A. McMechan, and H.-H. Chuang, 2011, Amplitude balancing in separating P- and S-waves in 2D and 3D elastic seismic data: *Geophysics*, **76**, no. 3, S103–S113, doi: [10.1190/1.3555529](https://doi.org/10.1190/1.3555529).
- Tsvankin, I., 2012, Seismic signatures and analysis of reflection data in anisotropic media (3rd ed.): SEG.
- Wang, J., 2005, Least-squares wave-equation AVP/AVA migration of common azimuth data: Ph.D. thesis, University of Alberta.
- Wang, J., H. Kuehl, and M. Sacchi, 2005, High-resolution wave-equation AVA imaging: Algorithm and tests with a data set from the western Canadian sedimentary basin: *Geophysics*, **70**, no. 5, S91–S99, doi: [10.1190/1.2076748](https://doi.org/10.1190/1.2076748).
- Wang, J., and M. D. Sacchi, 2007, High-resolution wave-equation amplitude-variation-with-ray-parameter (AVP) imaging with sparseness constraints: *Geophysics*, **72**, no. 1, S11–S18, doi: [10.1190/1.2387139](https://doi.org/10.1190/1.2387139).
- Wang, J., and M. D. Sacchi, 2009, Structure constrained least-squares migration: 79th Annual International Meeting, SEG, Expanded Abstracts, 2763–2767.
- Wong, M., B. L. Biondi, and S. Ronen, 2015, Imaging with primaries and free-surface multiples by joint least-squares reverse time migration: *Geophysics*, **80**, no. 6, S223–S235, doi: [10.1190/geo2015-0093.1](https://doi.org/10.1190/geo2015-0093.1).
- Xie, X., and R. Wu, 2005, Multicomponent prestack depth migration using the elastic screen method: *Geophysics*, **70**, no. 1, S30–S37, doi: [10.1190/1.1852787](https://doi.org/10.1190/1.1852787).
- Yan, J., and P. Sava, 2008, Isotropic angle-domain elastic reverse-time migration: *Geophysics*, **73**, no. 6, S229–S239, doi: [10.1190/1.2981241](https://doi.org/10.1190/1.2981241).
- Zhang, Q., H. Zhou, H. Chen, and J. Wang, 2016, Least-squares reverse time migration with and without source wavelet estimation: *Journal of Applied Geophysics*, **134**, 1–10, doi: [10.1016/j.jappgeo.2016.08.003](https://doi.org/10.1016/j.jappgeo.2016.08.003).
- Zhang, Y., L. Duan, and Y. Xie, 2015, A stable and practical implementation of least-squares reverse time migration: *Geophysics*, **80**, no. 1, V23–V31, doi: [10.1190/geo2013-0461.1](https://doi.org/10.1190/geo2013-0461.1).