Check for updates

Elastic wave-equation migration velocity analysis preconditioned through mode decoupling

Chenlong Wang[1], Jiubing Cheng[1], Wiktor Waldemar Weibull[2], and Børge Arntsen[3]

ABSTRACT

Multicomponent seismic data acquisition can reveal more information about geologic structures and rock properties than single component acquisition. Full elastic wave seismic imaging, which uses multicomponent seismic to its full potential, is promising because it provides more opportunities to understand the material properties of the earth by the joint use of P- and S-waves. A prerequisite of seismic imaging is the availability of a reliable macrovelocity model. Migration velocity analysis for P-waves, which can fill that requirement for the P-wave velocity, has been well-studied, especially under the acoustic approximation. However, a reliable estimation of the S-wave velocities remains troublesome. Elastic wave-equation migration velocity analysis has the potential to build P- and S-wave velocity models together, but it inevitably suffers from the effects of mode coupling and conversion in the forward and adjoint wavefield reconstructions. We have developed a differential semblance optimization approach to sequentially invert the background P- and S-wave velocity models from extended PP- and PS-images in the subsurface offset domain. Preconditioning of the gradients with respect to the S-wave velocity through mode decoupling can improve the reliability of the optimization. Numerical investigations with synthetic examples demonstrate the effectiveness of gradient preconditioning and the feasibility of our migration velocity analysis approach for elastic wave imaging.

INTRODUCTION

In the scales of seismic exploration, the earth can be approximated accurately as an elastic solid, in which seismic waves propagate as a superposition of P- and S-waves. Multicomponent receivers allow recording of both wave modes and provide important complementary information when compared to conventional P-wave exploration. These data can, for example, be used to produce improved subsurface images (e.g., in gas-bearing zones) and provide more information about the distribution of lithologies, pore-fluid saturation, fracture, and stress field in the subsurface (Thomsen, 1999; Stewart et al., 2003; Hardage et al., 2011; Gaiser, 2016). In contrast to conventional imaging schemes, which still use methods derived from the acoustic wave equation and single component data, full elastic wave imaging uses multicomponent data and methods derived from the elastic wave equation. These methods have

become more popular thanks to progress in multicomponent data acquisition, processing, and interpretation (e.g., Sun et al., 2006; Criss, 2007; Yan and Sava, 2008; Olofsson et al., 2012; Chaveste et al., 2013; Ravasi and Curtis, 2013; Reiser et al., 2015; Wang et al., 2016; Amundsen et al., 2017). As the industry turns toward higher-risk conventional and unconventional sources of hydrocarbons, multicomponent seismic imaging has gained further importance (Atkinson and Davis, 2011).

The potential of elastic wave imaging can only be achieved with high-quality P- and S-wave macrovelocity models. P-wave velocity model building is relatively mature and robust, whereas S-wave velocity estimation is more challenging and less well-studied. For the widely acquired converted-wave (PS) data, S-wave velocity is often estimated in the time domain by tuning the V_P/V_S ratio, e.g., Gaiser (1996), or through velocity analysis based on the asymptotic moveout equation of the reflected and converted events

---

[1]Tongji University, State Key Laboratory of Marine Geology, Shanghai, China. E-mail: clwang88@gmail.com; cjb1206@tongji.edu.cn (corresponding author).
[2]University of Stavanger, Department of Petroleum Engineering, Stavanger, Norway. E-mail: wiktor.w.weibull@uis.no.
[3]Norwegian University of Science and Technology, Department of Petroleum Engineering and Applied Geophysics, Trondheim, Norway. E-mail: borge.arntsen@ntnu.no.

(Dai and Li, 2008). However, not all boundaries in the medium necessarily produce the PP and PS reflected energy. Interval velocity models in depth derived from root-mean-square (rms) velocities obtained through time-domain velocity analysis may not be adequate in a geologic setting with strong lateral velocity variations. The velocity model for prestack depth migration is typically estimated with ray-based tomography methods (Stork, 1992; Adler et al., 2008). However, most tomographic inversion approaches for PS data impose the restriction that the V_P and V_S models have the same topology; i.e., they have the same layer or block boundaries (Broto et al., 2003; Ursin et al., 2005; Du et al., 2012a). As we know, discontinuities in the P-wave velocity may not be colocated with those of the S-wave velocity under certain lithologic or fluid-bearing conditions. Most importantly, reflection traveltime tomography requires extensive event picking and has limited capacity in sharp velocity variations (e.g., salt, thrust, and foothills).

Wave-equation-based velocity inversion can tackle complex wave phenomena and thus provide more realistic sensitivity for the velocity perturbations. It can be implemented in the data and image domains, with or without event picking. The data-domain approach could be formulated by finding a model that produces data that resemble the observed data either in kinematics (Luo and Schuster, 1991) or in the full waveform (Tarantola, 1986). Full-waveform inversion (FWI) relies on the kinematic and dynamic consistency between the predicted and observed data. Due to the nonlinearity associated with matching full waveforms, the success of FWI mainly relies on the use of only the transmission information of the data and through the use of multiscale inversion strategies that update the velocity models from long wavelengths to short wavelengths gradually (Sirgue and Pratt, 2004; Shin and Ho Cha, 2009; Alkhalifah, 2016). To use the reflection component of the data, and improve the accuracy of the background velocity model (especially for the deeper part), many authors use wave-equation modeling or demigration operators (Xu et al., 2012; Ma and Hale, 2013; Chi et al., 2015; Luo et al., 2016). Along with this approach, Wang et al. (2018) propose approaches to invert the V_P and V_S models with elastic wave mode decomposition-based preconditioning in the data domain.

The image-domain approach seeks kinematic consistency of the wavefield at an image location and aims to improve the image focusing (Sava and Biondi, 2004; Shen and Symes, 2008; Zhang et al., 2015; Chauris and Cocher, 2017). Therefore, it allows convergence from a poor initial model and the deficiency of low frequencies in the data. This means that the image-domain approach is more robust than data-domain methods. However, image-domain methods are in general unable to provide models with the same resolution methods. Chavent and Jacewitz (1995) propose the stack-power maximization (SPM) criteria to estimate the background velocity. Shen et al. (2003) apply the differential semblance optimization (DSO) (Symes and Carazzone, 1991) to depth imaging results to obtain a migration velocity model. Several wave-equation migration velocity analysis (WEMVA) methods using extended PP-images (Rickett and Sava, 2002; Sava and Fomel, 2006) have been proposed for isotropic or anisotropic (pseudo-)acoustic media (e.g., Mulder, 2008; Shen and Symes, 2008; Yang and Sava, 2011; Li et al., 2014, 2016; Weibull and Arntsen, 2014). Compared with the well-discussed P-wave velocity estimation, the study of the S-wave velocity is limited. Given the P-wave velocity model, Yang et al. (2015) propose an image registration guided wavefield tomography for extracting S-wave velocity error information from the PS-images using the PP-images as references. Yan and Sava (2010) point out that the DSO objective function using the extended PS-images has good convexness for estimation of background S-wave velocities, but they do not show any example with inverted models. Shabelansky et al. (2015) apply the DSO approach to invert the P- and S-wave velocity models simultaneously for passive seismic data.

In this study, we focus on S-wave velocity model building with extended PS-images for active-source seismic data. We first derive the gradient of the DSO-based objective function using the adjoint-state (AS) method (Plessix, 2006). Then, we demonstrate that the source-side gradient term is negligible because we assume that the P-wave velocity model is known and the background S-wave velocity has no effect on the source-side kinematics. To avoid S-to-P mode conversion when injecting the adjoint sources related to the image residual along the subsurface offsets, we introduce the Levi-Civita tensor in the AS equation to generate a pure S-wave virtual (secondary) source for the receiver-side wavefields (Wang et al., 2015). Then, we propose to precondition the gradients by suppressing the artifacts through elastic wave mode decoupling. Finally, we use synthetic seismic data examples to demonstrate the effectiveness of the proposed approach.

ELASTIC WEMVA

Elastic reverse time migration with an extended imaging condition

For image-domain velocity analysis, we rely on the "semblance principle," which states that the model is accurate when the images constructed from different experiments are consistent with one another. Therefore, we can use common-image gathers (CIGs) or the extended images for WEMVA. To provide CIGs for elastic wave velocity analysis, we use elastic reverse time migration (ERTM) with an extended imaging condition (Rickett and Sava, 2002):

$$I_{mn}(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{s}} \int \hat{u}_m^s(\mathbf{x} - \mathbf{h}, t; \mathbf{s}) \hat{u}_n^r(\mathbf{x} + \mathbf{h}, t; \mathbf{s}) dt, \quad (1)$$

where $\mathbf{x}$ is the imaging location, $\mathbf{h}$ is the spatial lag, $m, n \in \{P, S\}$ represent the body-wave modes, t is the traveltime, $\mathbf{s}$ is the source index, and $\hat{u}_m^s$ and $\hat{u}_n^r$ are the single-mode source and receiver wavefields, respectively. To clarify the notation in this paper, the $\mathbf{s}$ with bold type signify the source index, the P and S as subscripts denote the P- and S-wave modes, and the superscripts with s and r indicate the wavefield from the source and receiver side, respectively.

CIGs contain redundant offset information for velocity estimation (Biondi and Symes, 2004): Generally, a downward (or upward) curvature indicates a too-low (or too-high) velocity. However, migration smiles (an upward curvature) also appear due to limited acquisition even the migration velocity is exact. Tapers are classically applied to CIGs can partly reduce this artifact.

The elastic wavefield satisfies the second-order wave equation:

$$\rho(\mathbf{x}) \frac{\partial^2 u_i}{\partial t^2}(\mathbf{x}, t) - \frac{\partial}{\partial x_j}\left[C_{ijkl}(\mathbf{x}) \frac{\partial u_l}{\partial x_k}(\mathbf{x}, t)\right] = f(\mathbf{x}_s, t), \quad (2)$$

in which $\mathbf{x}$ is the spatial coordinate, t indicates the traveltime, u_i is the ith component of the displacement field, ρ denotes the mass density, C_{ijkl} is the stiffness tensor, f represents the external source

at the position of $\mathbf{x}_s$, and $i, j, k, l \in \{x, y, z\}$ are the indexes satisfying the summation convention.

For isotropic media, single-mode wavefields are separated via the divergence and curl operators (Aki and Richards, 2002), i.e.,

$$\hat{u}_p(\mathbf{x}, t) = V_P^2(\mathbf{x}) \nabla \cdot \mathbf{u}(\mathbf{x}, t) \qquad (3)$$

and

$$\hat{\mathbf{u}}_s(\mathbf{x}, t) = V_S^2(\mathbf{x}) \nabla \times \mathbf{u}(\mathbf{x}, t), \qquad (4)$$

where $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ and V_P and V_S are the P- and S-wave velocities, respectively. In a 2D case, we have

$$\hat{u}_p(\mathbf{x}, t) = V_P^2 \left(\frac{\partial u_x}{\partial x} + \frac{\partial u_z}{\partial z} \right) \qquad (5)$$

and

$$\hat{u}_s(\mathbf{x}, t) = V_S^2 \left(\frac{\partial u_x}{\partial z} - \frac{\partial u_z}{\partial x} \right). \qquad (6)$$

DSO-based misfit function

The CIGs obtained by equation 1 preserve the spatial lag between the source and receiver-side wavefield as the subsurface offset. At the scattering location, i.e., when the spatial lag $|\mathbf{h}|$ is zero, the wavefields collapse to a delta function. Away from the zero spatial lag, no energy should exist. This means that when the velocities are accurate, the CIGs should be focused around the subsurface zero offset.

Therefore, we can update the migration velocity model with a DSO-based approach by penalizing the energy at nonzero offsets. For retrieving the P-wave velocity, we could combine the DSO and SPM criteria to make the inversion more robust (Soubaras and Gratacos, 2007). Because the polarity reversal exists in the PS-images, we only use DSO misfit function to update the S-wave velocity model. The following DSO misfit function represents a direct way to quantify the focusing of CIGs:

$$\mathcal{J} = \frac{1}{2} \int d\mathbf{x} \int d\mathbf{h} \left[\mathbf{h} \frac{\partial I_{mn}(\mathbf{x}, \mathbf{h})}{\partial z} \right]^2. \qquad (7)$$

Compared with the DSO misfit function in Shen et al. (2003), an additional vertical derivative is applied to remove the low-wavenumber artifacts in the ERTM result (Guitton et al., 2007; Weibull and Arntsen, 2013).

To check the sensitivity of the misfit to the parameter perturbation, we evaluate the normalized misfit function for the PP and PS CIGs in a simple 1D model (see Figure 1). First, we apply the elastic wave equation to synthesize common-shot gathers with a maximum offset of 0.8 km using a pressure source. Then, we perturb V_P and V_S in the first layer and apply ERTM to generate the PP and PS CIGs, respectively. The PS CIGs are obtained using correct P-wave velocity. Finally, we calculate the normalized misfit functions with these CIGs. As shown in Figure 1, the

well-behaved convex misfit functions imply that the DSO penalty functions can be used for the P- and S-wave migration velocity analysis. It is noticeable that the misfit function using PS CIGs is asymmetric and has a slight deviation from the correct value due to the amplitude and waveform effects. The spatial wavelength of the reflector in the CIGs is dependent on the velocities. It causes the difference between positive and negative perturbations of the velocity fields. Positive perturbation implies an increase of the wavelength; thus, the sensitivity curve of the misfit is potentially more linear than a negative perturbation.

Gradient calculation

An efficient unconstrained optimization algorithm, the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) quasi-Newton algorithm (Byrd et al., 1995; Nocedal, 1999), is used to minimize the misfit function. Using the AS method (Chavent, 1974; Plessix, 2006), we can derive the gradients of the misfit function with respect to V_P and V_S (see Appendix A). The gradient with respect to P-wave velocity is given by Weibull and Arntsen (2014) as

$$
\begin{aligned}
\nabla_{V_P} \mathcal{J}(\mathbf{x}) = & \sum_s \int dt \frac{\partial C_{ijkl}}{\partial V_P}(\mathbf{x}) \frac{\partial u_l^s}{\partial x_k}(\mathbf{x}, t; \mathbf{s}) \frac{\partial \psi_i^s}{\partial x_j}(\mathbf{x}, t; \mathbf{s}) \\
& + \sum_s \int dt \frac{\partial C_{ijkl}}{\partial V_P}(\mathbf{x}) \frac{\partial u_l^r}{\partial x_k}(\mathbf{x}, t; \mathbf{s}) \frac{\partial \psi_i^r}{\partial x_j}(\mathbf{x}, t; \mathbf{s}) \\
& + \sum_s \int dt 2 V_P(\mathbf{x}) \hat{u}_p^r(\mathbf{x}, t; \mathbf{s}) \\
& \times \int d\mathbf{h} \mathbf{h}^2 \frac{\partial^2 I_{pp}}{\partial z^2}(\mathbf{x} - \mathbf{h}) \hat{u}_p^s(\mathbf{x} - 2\mathbf{h}, t; \mathbf{s}) \\
& + \sum_s \int dt 2 V_P(\mathbf{x}) \hat{u}_p^s(\mathbf{x}, t; \mathbf{s}) \\
& \times \int d\mathbf{h} \mathbf{h}^2 \frac{\partial^2 I_{pp}}{\partial z^2}(\mathbf{x} - \mathbf{h}) \hat{u}_p^r(\mathbf{x} - 2\mathbf{h}, t; \mathbf{s}), \qquad (8)
\end{aligned}
$$

in which $\mathbf{s}$ indicates the source index, u is the regular or state wavefields, and ψ represents the adjoint wavefields simulated with residual (secondary) sources, which are constructed with PP CIGs (for more details, see equation 9 in Weibull and Arntsen, 2014). Note that the last two terms in equation 8 are introduced because
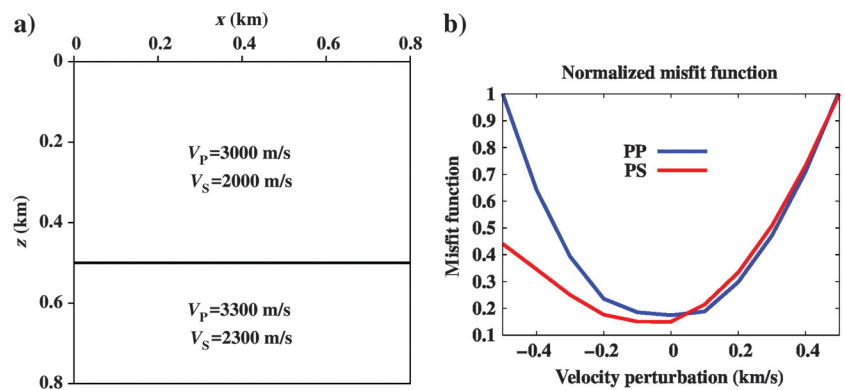

Figure 1. Misfit function analysis: (a) model structure and (b) normalized misfit functions with respect to perturbations of V_P (the blue line) and V_S (the red line).

the divergence operation is scaled by the squared P-wave velocity (see equation 3). We derive the gradient with respect to S-wave velocity using the extended PS-images as follows:

$$\nabla_{V_S}\mathcal{J}(\mathbf{x}) = \sum_{\mathbf{s}}\int dt \frac{\partial C_{ijkl}}{\partial V_S}(\mathbf{x})\frac{\partial u_l^s}{\partial x_k}(\mathbf{x},t;\mathbf{s})\frac{\partial \psi_i^s}{\partial x_j}(\mathbf{x},t;\mathbf{s})$$

$$+ \sum_{\mathbf{s}}\int dt \frac{\partial C_{ijkl}}{\partial V_S}(\mathbf{x})\frac{\partial u_l^r}{\partial x_k}(\mathbf{x},t;\mathbf{s})\frac{\partial \psi_i^r}{\partial x_j}(\mathbf{x},t;\mathbf{s})$$

$$+ \sum_{\mathbf{s}}\int dt 2V_S(\mathbf{x})\hat{u}_s^r(\mathbf{x},t;\mathbf{s})$$

$$\times \int d\mathbf{h}\mathbf{h}^2 \frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h},t;\mathbf{s}), \quad (9)$$

where ψ represents the AS wavefields controlled by the following equations:

$$\psi_i^s(\mathbf{x},t,\mathbf{s}) = \int d\mathbf{x}'\frac{\partial G_{ij}}{\partial x^j}(\mathbf{x},0;\mathbf{x}',t) * \mathcal{F}_{ps}^s(\mathbf{x}',t;\mathbf{s}) \quad (10)$$

and

$$\psi_i^r(\mathbf{x},t,\mathbf{s}) = \int d\mathbf{x}'\varepsilon_{ijk}\frac{\partial G_{ij}}{\partial x^k}(\mathbf{x},t;\mathbf{x}',0) * \mathcal{F}_{ps}^r(\mathbf{x}',t;\mathbf{s}), \quad (11)$$

where G_{ij} denotes the Green's function at point $\mathbf{x}$ due to a source at $\mathbf{x}'$, i denotes the wavefield component, and j indicates a delta force source in the jth direction. Notation "$*$" represents a time convolution. We introduce the Levi-Civita tensor ε_{ijk} to construct a pure S-wave source to prevent generating any P-wave component when injecting the adjoint sources ($\mathcal{F}_{ps}^r$) due to the extended-domain image residual. The last term in equation 9 is introduced because the curl operation is scaled by the squared S-wave velocity (see equa-

Figure 2. Model and synthetic common-shot records used to verify the gradient calculation: (a) V_P and (b) V_S models and (c) x- and (d) z-component of the displacements.

tion 6). The adjoint sources $\mathcal{F}_{ps}^s$ and $\mathcal{F}_{ps}^r$ are formulated by taking derivatives of the misfit function with respect to the state variables (u_i^s and u_i^r). They depend on the extended PS-images as follows:

$$\mathcal{F}_{ps}^s(\mathbf{x}',t;\mathbf{s}) = \int d\mathbf{h}\mathbf{h}^2 V_P^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}+\mathbf{h})\hat{u}_s^r(\mathbf{x}+2\mathbf{h}) \quad (12)$$

and

$$\mathcal{F}_{ps}^r(\mathbf{x}',t;\mathbf{s}) = \int d\mathbf{h}\mathbf{h}^2 V_S^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h}). \quad (13)$$

To validate the gradients calculated by the AS method, the finite-difference (FD) approximation is used to calculate the gradients for a small-scale model (Figure 2a and 2b), and the results are compared with those obtained by the AS method. As shown in Figure 2c and 2d, with a pressure source on the surface, we generate a synthetic shot recording without taking account of the free surface. For the FD approximation, the gradients with respect to P- and S-wave velocities are calculated by evaluating the misfit function values with a very small velocity perturbation at every model grid. Taking a model with 100×100, for example, it means that we need to estimate the misfit function for 10,000 times in this simple model to calculate the gradient. We observe that the gradients calculated with the two methods are comparable, although those obtained by the AS method look relatively smooth (Figures 3 and 4). It is worth mentioning that FD can be used for small problems only, due to the huge requirement of memory storage and computational time when the problem is large.

Gradient preconditioning

Different from DSO-based WEMVA using the acoustic approximation, e.g., Li et al. (2016), elastic wave migration velocity analysis requires V_P and V_S to be provided. However, if only the kinematics of P-wave propagation is considered, the S-wave velocities are of minor importance (Alkhalifah and Tsvankin, 1995). Weibull and Arntsen (2014) use DSO-based WEMVA to build models for the P-wave velocities and two Thomsen parameters. Here, we focus on building the S-wave velocity model with a similar approach but pay more attention to mitigating the artifacts due to crosstalk between the wave modes in the calculated gradients. To ensure a fast convergence, we will introduce the following preconditioning steps for the inversion.

First, we ignore the contribution of the source-side term in the gradient with respect to the S-wave velocity. This is because we assume that the P-wave velocity model is known before estimating the S-wave velocity and the source-side wavefields are controlled by only the P-wave velocity. In other words, the kinematics of the extended PS-images is only affected by the background S-wave velocity assuming we give a correct background P-wave velocity model. Yang et al. (2015) use a similar strategy in their image registration guided wavefield tomography for S-wave velocity model building owing to the same reason. As shown in Figure 4, observable changes only appear in the vicinity of the source location when we ignore the source term for the gradients with respect to the S-wave velocity. In addition, Weibull and Arntsen (2013) precondition the gradients in this area (e.g., by muting) to tackle the singularity at the source location.
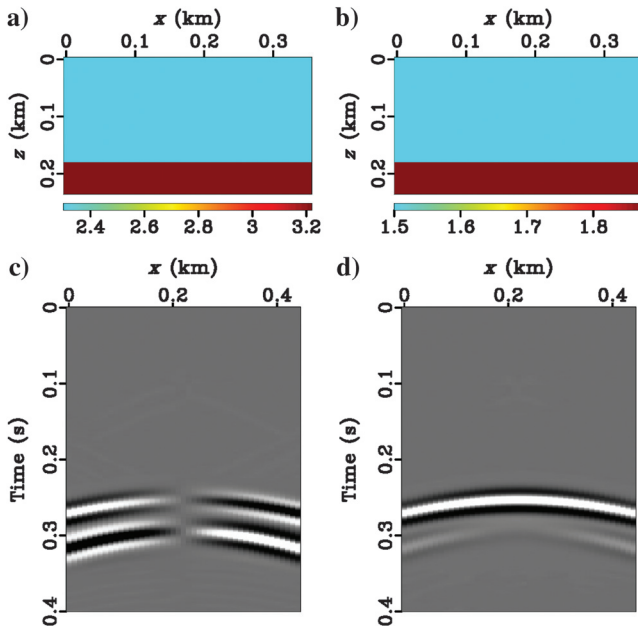
As we know, the background S-wave velocities do not affect the kinematics of the P-wave. Therefore, on the receiver-side wavepaths, only the converted S-wave wavefields have the positive contribution for estimating the background S-wave velocities. So we apply elastic wave mode decomposition to the receiver-side forward and adjoint wavefields to get a preconditioned gradient:

$$\nabla_{V_S} \mathcal{J}(\mathbf{x}) = \sum_{\mathbf{s}} \int dt \frac{\partial C_{ijkl}}{\partial_{V_S}}(\mathbf{x}) \frac{\partial u_{sl}^r}{\partial x_k}(\mathbf{x}, t; \mathbf{s}) \frac{\partial \psi_{si}^r}{\partial x_j}(\mathbf{x}, t; \mathbf{s})$$

$$+ \sum_{\mathbf{s}} \int dt 2_{V_S}(\mathbf{x}) \hat{u}_s^r(\mathbf{x}, t; \mathbf{s})$$

$$\times \int d\mathbf{h} \mathbf{h}^2 \frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x} - \mathbf{h}) \hat{u}_p^s(\mathbf{x} - 2\mathbf{h}, t; \mathbf{s}), \qquad (14)$$

in which u_{sl} and ψ_{si} indicate the S-wave data separated from the forward and adjoint wavefields using

$$\mathbf{u}_s(\mathbf{x}) = - \int e^{i\mathbf{k}\mathbf{x}} \bar{\mathbf{k}} \times \bar{\mathbf{k}} \times \tilde{\mathbf{u}}(\mathbf{k}) d\mathbf{k}, \qquad (15)$$

where $\bar{\mathbf{k}}$ is the normalized wavenumber vector and $\tilde{\mathbf{u}}$ is the displacement wavefield in the wavenumber domain.

The implementation of WEMVA is more complicated than that of RTM (as shown in Figure 5), which contains three crosscorrelations of four different wavefields for a single shot record (two regular wavefields and two adjoint wavefields). It means that one must perform four times of wavefield simulation to calculate the corresponding wavefields in equation 9. The involved regular wavefields u^s and u^r are calculated by the forward and backward wavefield simulation, which are exactly the same as in RTM. Whereas, to determine the adjoint wavefields, ψ^s and ψ^r, one must first know the adjoint sources, which are virtual sources representing the interaction of the regular wavefields with the extended-domain image residual. Meanwhile, we omit the source-side term in the gradient with respect to the S-wave velocity (the dashed box in Figure 5) and apply wave mode decoupling to the corresponding wavefields as preconditioning method in equation 14. A straightforward solution for time-domain methods would store the whole regular wavefields to disk at each time step during the state-equation simulation and then read it back during the AS equation simulation to calculate the interaction of these four fields. In the 2D case, this approach is feasible, but in the 3D case, some techniques (such as the checkpointing) should be applied to handle the issue of disk-memory explosion (Griewank and Walther, 2000; Clapp, 2008).
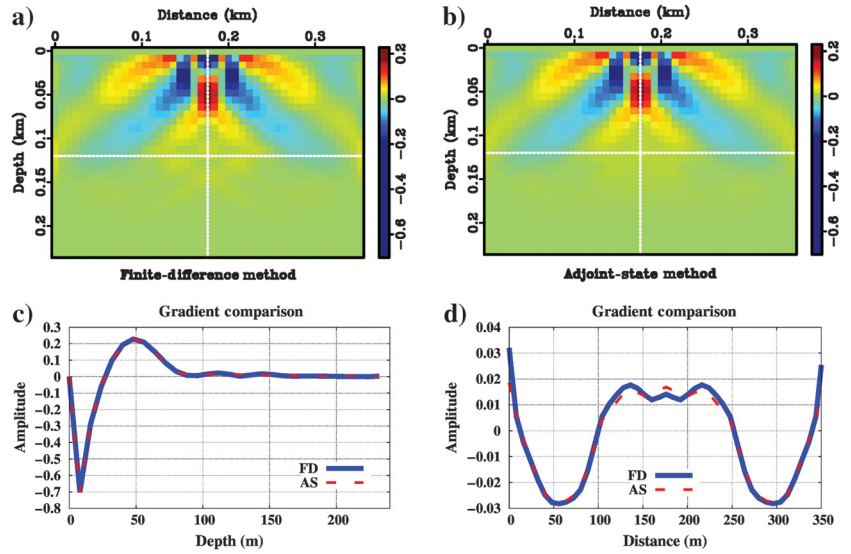

Figure 3. The gradients with respect to V_P calculated by the (a) FD and (b) AS methods. A detailed comparison is displayed at (c) $x = 0.176$ km and (d) $z = 0.12$ km, with the solid blue and dashed red lines corresponding to the FD and AS methods, respectively.
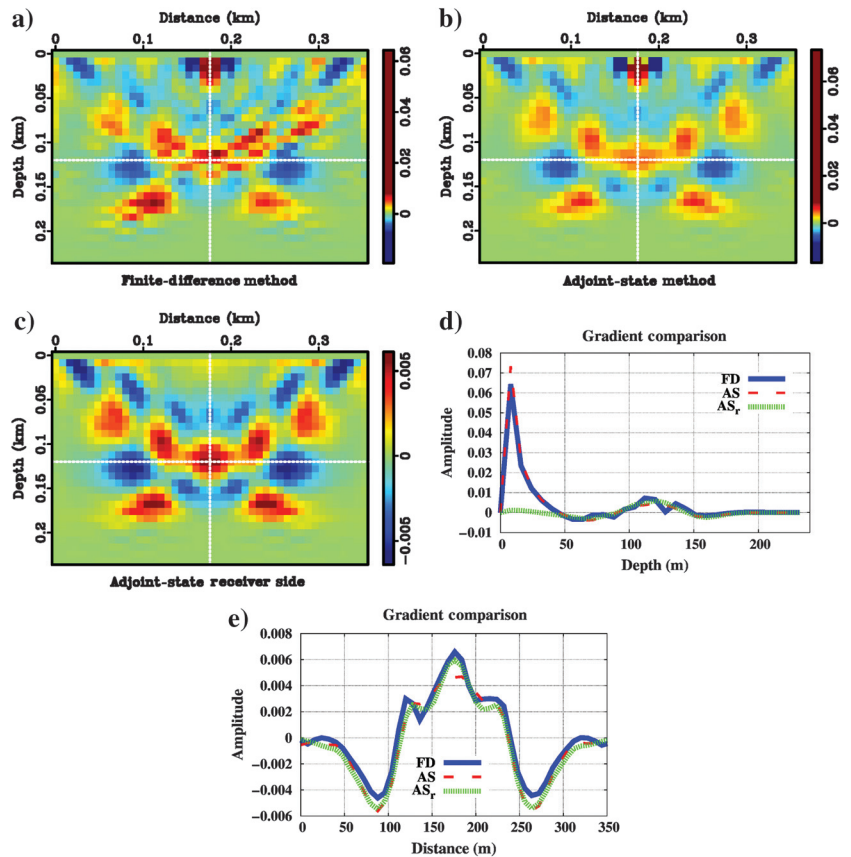

Figure 4. The gradients with respect to V_S calculated by (a) the FD and the AS methods (b) with and (c) without the source term in equation 9. Along with the solid blue and dashed red lines, the solid green lines denote the calculated gradients without the source term.

A simple model is used to demonstrate the proposed preconditioning methods. As shown in Figure 6, the model size is 2.0×0.6 km, and an interface is located at the depth of 0.4 km. The true P- and S-wave velocities of the first layer are 3.0 and 2.1 km/s, respectively. In total, 101 shots are evenly triggered on the surface with an interval of 20 m. Given the correct P-wave velocity model, Figure 7 displays the space-lag CIGs of PS-reflections at $x = 1.0$ km when we set the S-wave migration velocities with 2.1, 1.8, and 2.4 km/s, respectively. We observe that the PS-images are symmetrically focused around the zero offset at the true depth when the S-wave migration velocity is correct. A downward and upward curvature indicates a too-low and a too-high velocity, respectively. The unfocused energy relatively far away from the zero offset is regarded as the image residual due to the errors of the migration velocity.

As shown in Figure 8, for the 51st shot, we observe a remarkable difference between the gradients with respect to the S-wave velocity before and after preconditioning based on mode decomposition. In the fast and slow velocity cases, the preconditioning has mitigated the oscillations caused by the crosstalk from the P-wavefields along the converted S-wave paths. Figure 9 displays the corresponding gradients of the total 101 shots. Because the original gradients with slow velocity have too many artifacts, their stacking still has strong oscillations and produces a wrong updating direction for the inversion. In both cases, the stacked gradients after mode decomposition-based preconditioning not only provide the correct sign for the model update but also are almost free of artifacts.

Note that DSO-based WEMVA has its own shortcomings. It suffers from defocusing with subsurface offset-domain CIGs in complex regions with poor illumination. Like many previously published wave-equation-based methods, which attempt to automatically invert traveltime or kinematic information in seismic data or migrated gathers, the proposed approach will be affected by amplitude information. Nowadays, there are some works that exploit the strengths and avoid the weakness, e.g., Lameloise et al. (2015), Zhang et al. (2015), and Luo et al. (2016), which is beyond the scope of this study. The DSO approach using an L-BFGS algorithm iteratively updates the velocity model through annihilating the non-zero subsurface-offset energy in the CIGs. In each iteration, four times of wavefield simulation are required to calculate the gradient for a single shot record. In general, the misfit function will be efficiently reduced after tens of iterations.

NUMERICAL EXAMPLES

We demonstrate the proposed approach with 2D synthetic data sets. A 10th-order explicit staggered grid FD algorithm is used to extrapolate the source and receiver wavefields. A perfectly matched layer absorbing boundary is used around the calculation area to avoid reflections from the boundaries of the models. We use explosive sources to generate wavefields of pure P-wave sources, of which the source function is a Ricker wavelet with a peak frequency of 25 Hz. Although the analysis with an S-wave source is an easy extension, we do not consider the S-wave source problem in this paper. The data sets are preprocessed by muting the direct and refracted waves on the receiver side.

Optimization is carried out with an L-BFGS method (Byrd et al., 1995). In the examples, we first build the V_P model using the extended PP-images because the S-wave velocity does not affect the kinematics of the P-wave. Once the P-wave migration velocity model is known, we inverted the S-wave migration velocity using the extended PS-images.

Gauss anomaly model

The first example is based on a simple model (Figure 10). Two smooth Gauss-shape negative perturbations (with maximum magnitude 0.3 km/s) are added into constant P- and S-wave
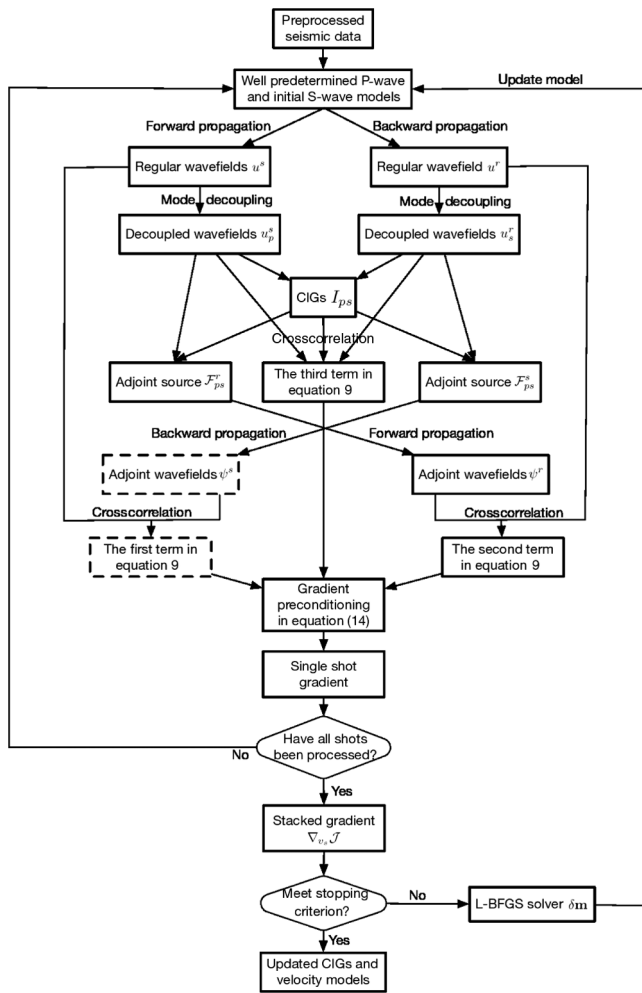

Figure 5. The flowchart of the converted wave WEMVA algorithm. The box marked with dashed lines is neglected when the proposed gradient preconditioning has been used.
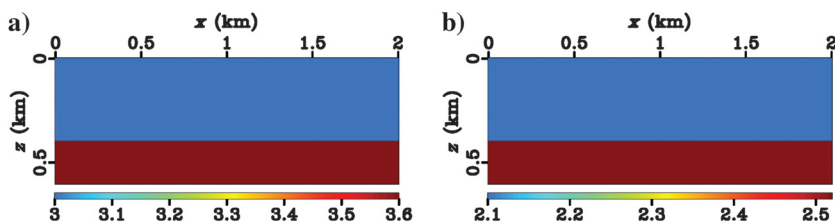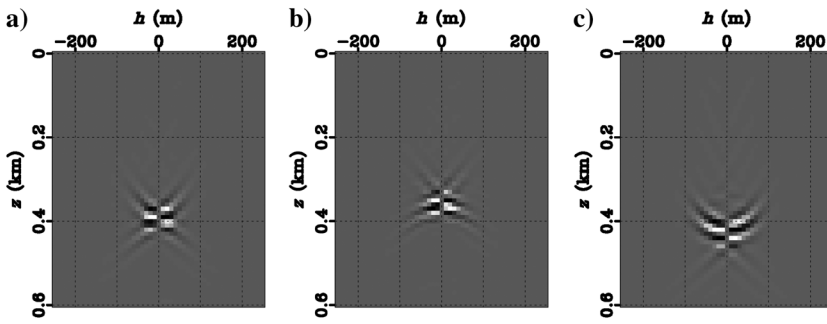

Figure 6. A simple model: (a) V_P and (b) V_S.

Figure 7. The subsurface-offset domain CIGs generated with (a) the true (V_S = 2.1 km/s), (b) slow (1.8 km/s), and (c) fast (2.4 km/s) velocities for the S-wave.
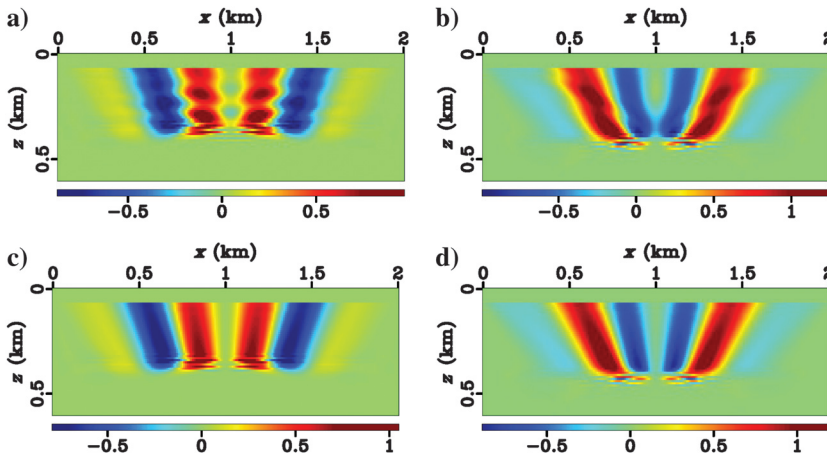

Figure 8. The receiver-side gradients of the 51th shot (top) before and (bottom) after the preconditioning step (equation 14). Left: a slow velocity (V_S = 1.8 km/s) and right: a fast velocity (V_S = 2.4 km/s).
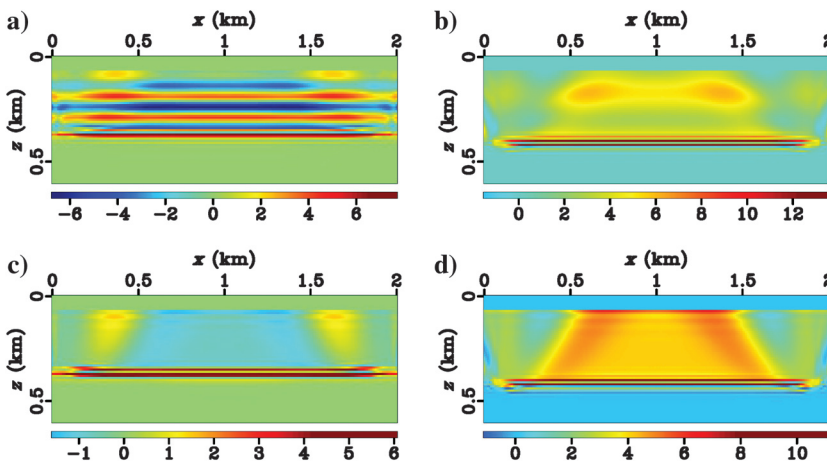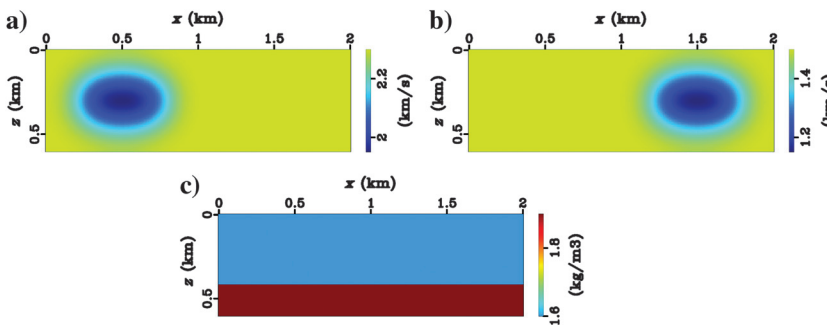

Figure 9. Stacked gradients of total 101 shots (top) before and (bottom) after the preconditioning step. Left: a slow velocity (V_S = 1.8 km/s) and right: a fast velocity (V_S = 2.4 km/s).

Figure 10. A Gauss anomaly model: (a) V_P, (b) V_S, and (c) ρ.

velocities around $x = 0.5$ and $x = 1.5$ km, respectively. A horizontal interface is introduced by density perturbation to generate reflections. In total, 51 shots are simulated at the surface with a shot spacing of 40 m. This experiment with two spatially isolated P- and S-wave velocity abnormals is designed to demonstrate that the proposed method is free of the crosstalk between two velocities'

estimation. Figure 11 shows the updated P- and S-wave velocity models. It demonstrates the accuracy and robustness of the proposed method, which breaks the restriction that V_P and V_S models have the same topology. The ERTM images and CIGs demonstrate that the DSO approach provides good velocity models for prestack depth migration (see Figures 12 and 13). Note that a polarity rever-

Figure 11. The updated model using the proposed DSO approach: (a) V_P and (b) V_S.
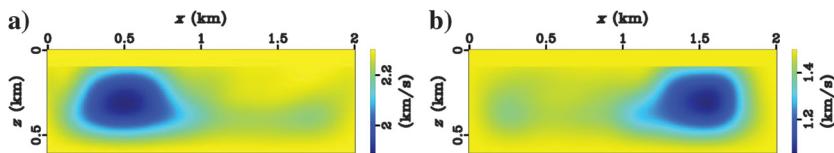
Figure 12. The PP-images with (a) initial, (b) updated, and (c) true models and PS-images with (d) initial, (e) updated, and (f) true models.
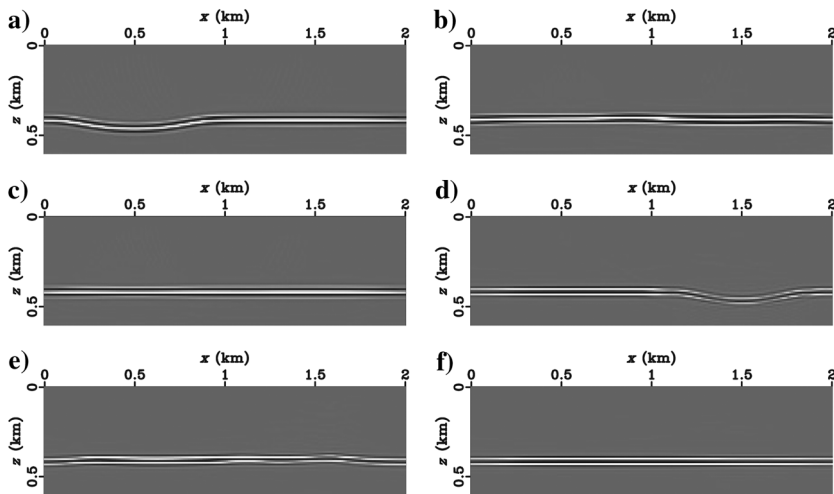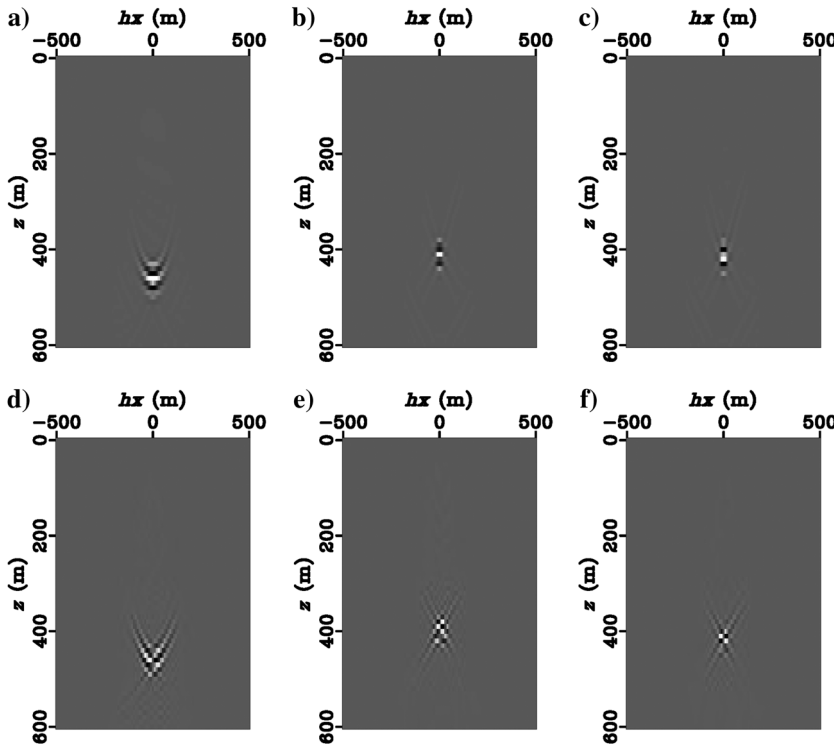

Figure 13. Subsurface offset-domain CIGs: PP CIGs at $x = 0.5$ km with (a) initial, (b) updated, and (c) true models; PS CIGs at $x = 1.5$ km with (d) initial, (e) updated, and (f) true models.

sal correction is not applied in the CIGs for velocity estimation, whereas it is demanded in conventional converted wave imaging with zero spatial lag as shown in Figure 12 (Du et al., 2012b).

Sigbee2A model

We apply the method to a complex data set simulated on part of the Sigbee2A model (Figure 14). In total, 48 shots are triggered on the surface with the maximum offset of 1.5 km. Figure 15 shows the starting and updated models, respectively. We observe that the inverted models reasonably represent low-wavenumber components of the true models even though we start from linearly increasing velocities for V_P and V_S after 38 iterations. The comparison of the ERTM images (Figure 16) with the initial, updated, and true models demonstrates the improvement of migration velocity models for P- and S-waves. It is noticeable that the scatters in the converted image with the updated velocity model are not focused completely. This issue has two contributors: The first is that the slight inaccuracy in the predetermined P-wave velocity affects the estimation of the S-wave velocity. Meanwhile, the converted wave imaging needs the polarity reversal correction, which is parameter dependent. It will magnify the inaccuracy in the PS-image. The CIGs at three locations (Figure 17) further show the validity of the velocity updating.

DISCUSSION

ERTM-based WEMVA using the DSO misfit function provides an automatic way of improving the quality of depth-domain P- and S-velocit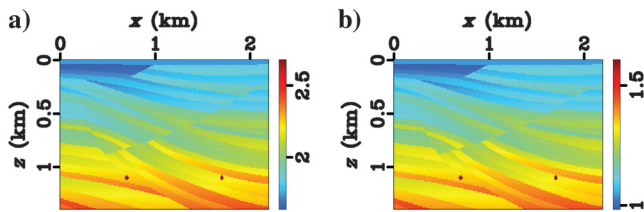y models. To mitigate the nonlinearity and parameter trade-offs, we estimate P- and S-wave velocity stepwise. This two-step strategy has been widely used in the estimation of elastic wave velocities and shows its robustness (Yang et al., 2015; Wang et al., 2018). Updating the S-wave velocity model requires a well-predetermined P-wave velocity model; otherwise, the kinematic errors on the source side of the PS-reflection prevent a robust convergence. Theoretically, one could update the P- and S-wave velocities simultaneously (Shabelansky et al., 2015). However, better strategies are required to mitigate the trade-off between the two velocity models and to guarantee the convergence.
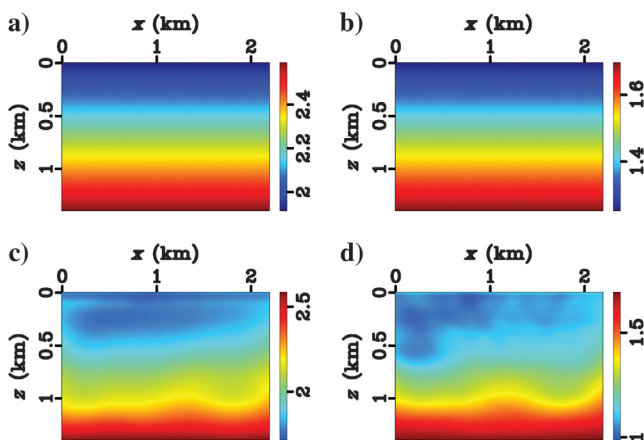
The algorithm of elastic wave mode decoupling used in this paper only works in isotropic media. An extension to consider anisotropy leads to prohibitively expensive computation. The effectiveness of gradient preconditioning through mode decoupling in dealing with anisotropy is worthy of future investigation. In the 3D case, the decoupled S-wave by the curl operator (equation 4) is a vector wavefield. Therefore, it is hard to process via the conventional imaging condition. To solve this issue, some authors proposed new imaging conditions for 3D ERTM (Wang et al., 2016; Du et al., 2017). If these new imaging conditions are used, we will change the third and fourth terms in equation 8, and the last term in equation 9. In the meantime, the adjoint wavefields will be computed with the corresponding adjoint sources for the new imaging condition (see Appendix B). The surface wave and multiples are also inevitable in field data. The surface wave is regarded as a noise that must be suppressed before application of ERTM because the DSO only concerns the reflected waves. Meanwhile, the objective function in equation 7 is strictly valid under the single-scattering assumption. One solution to this problem is to include surface wave and multiple attenuation as a part of the processing of the data used for velocity analysis and use the absorbing boundary condition at all sides for wavefield reconstruction.

Figure 14. True models: (a) V_P and (b) V_S.

Figure 15. (Top) Starting and (bottom) updated velocity models. Left: V_P and right: V_S.
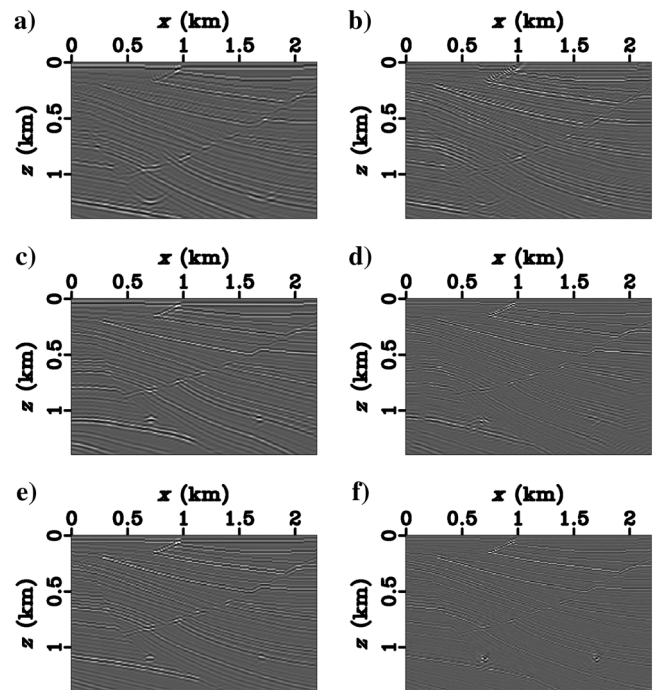

Figure 16. ERTM images with (top) initial, (middle) updated, and (bottom) true models. Left: PP-images and right: PS-image.
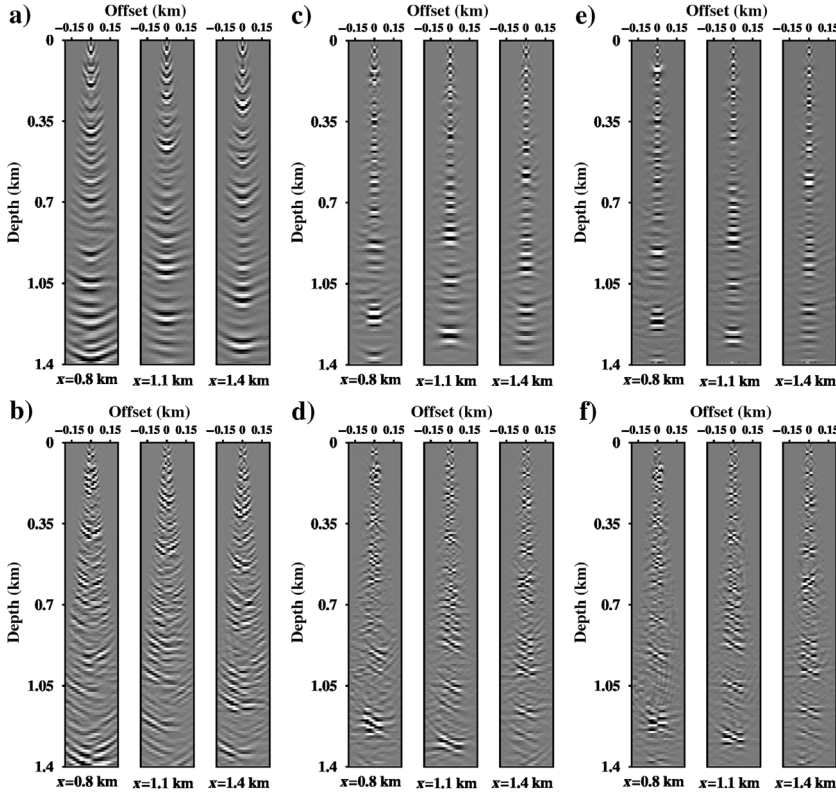
Figure 17. Subsurface offset-domain CIGs at $x = 0.8$, 1.1, and 1.4 km with (left) initial, (middle) updated, and (right) true models. Top: PP-images and bottom: PS-images.

CONCLUSION

Reliable macrovelocity estimation of P- and S-waves is important for multicomponent seismic exploration. Based on the DSO misfit function measuring the residual in the extended PS-images, we have derived the gradient with respect to the S-wave velocity using the AS method. We found that the source-side term in the gradient can be neglected to reduce the computational cost once the P-wave velocity model is given. To avoid the artifacts due to the crosstalk from the P-wavefields, we have introduced the Levi-Civita tensor in the AS equation when we inject the PS-image residual as adjoint sources and applied P/S mode decoupling before the crosscorrelation of the receiver-side forward and adjoint wavefields. The numerical investigation has shown that this preconditioning significantly suppresses the oscillations and provides a reliable gradient for S-wave velocity updating. The two synthetic examples have demonstrated that the proposed approach has good potential for building P- and S-wave migration velocity models for elastic full-wave imaging. Future works include mitigating the amplitude effect on the DSO-based gradients and 3D applications.

ACKNOWLEDGMENTS

toil), T. Wang (CGG), and B. Ursin (NTNU) for their helpful discussions. We would also like to show our gratitude to the associate editor J. D. De Basabe, Carlos and two anonymous reviewers for their careful review. We appreciate the support of the *Madagascar* software package.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

APPENDIX A

AS METHOD FOR DSO-BASED ELASTIC WEMVA

In this appendix, we derive the gradient of elastic DSO with respect to S-wave velocity using the AS method (Chavent, 1974; Plessix, 2006) in the 2D case. The objective function is given by

$$\mathcal{J} = \frac{1}{2} \int d\mathbf{x} \int d\mathbf{h} \left[\mathbf{h} \frac{\partial I_{ps}(\mathbf{x}, \mathbf{h})}{\partial z} \right]^2, \quad \text{(A-1)}$$

where the $I_{ps}(\mathbf{x}, \mathbf{h}) = \int \hat{u}_p^s(\mathbf{x} - \mathbf{h}) \hat{u}_s^r(\mathbf{x} + \mathbf{h}) dt$, $\hat{u}_p^s$ and $\hat{u}_s^r$ are defined in equations 5 and 6, respectively. For simplicity, we rewrite equation 2 as

$$\rho \partial_t^2 \mathbf{u} - \nabla \cdot (\mathbf{c} : \nabla \mathbf{u}) = \mathbf{f}, \quad \text{(A-2)}$$

where ∂_t^2 is the second time derivative. Equation A-2 is the state equation in the elastic DSO inversion. Our objective is to minimize the misfit function, equation A-1, with respect to V_S and subject to the constraints that the background wavefield satisfies the state equation A-2. Therefore, we define the Lagrangian $\mathcal{L}$:

$$\mathcal{L} = \frac{1}{2} \int d\mathbf{x} \int d\mathbf{h} \left[\mathbf{h} \frac{\partial I_{ps}(\mathbf{x}, \mathbf{h})}{\partial z} \right]^2$$
$$+ \int_0^T dt \int_\Omega d\mathbf{x} [\rho \partial_t^2 \mathbf{u}^s - \nabla \cdot (\mathbf{c} : \nabla \mathbf{u}^s) - \mathbf{f}] \boldsymbol{\psi}^s$$
$$+ \int_0^T dt \int_\Omega d\mathbf{x} [\rho \partial_t^2 \mathbf{u}^r - \nabla \cdot (\mathbf{c} : \nabla \mathbf{u}^r)] \boldsymbol{\psi}^r, \quad \text{(A-3)}$$

in which Ω is the integration area, $\mathbf{u}^s$ and $\mathbf{u}^r$ are the state variables, and $\boldsymbol{\psi}^s$ and $\boldsymbol{\psi}^r$ are the Lagrange multipliers (or AS variables) that remain to be determined. Using the Gauss theorem, we take the variation of equation A-3, this gives

$$\delta\mathcal{L} = \int_0^T dt \int_\Omega d\mathbf{x} \int d\mathbf{h}\mathbf{h}^2 V_P^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}+\mathbf{h})\hat{u}_s^r(\mathbf{x}+2\mathbf{h})\delta\mathbf{u}^s$$

$$+ \int_0^T dt \int_\Omega d\mathbf{x}\boldsymbol{\varepsilon} \int d\mathbf{h}\mathbf{h}^2 V_S^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h})\delta\mathbf{u}^r$$

$$+ \int_0^T dt 2V_S\hat{u}_s^r \int d\mathbf{h}\mathbf{h}^2\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h},t)\delta v_s$$

$$- \int_0^T dt \int_\Omega d\mathbf{x}\left(\frac{\partial\mathbf{c}}{\partial V_S}\delta V_S\nabla\boldsymbol{\psi}^s\nabla\mathbf{u}^s + \delta\mathbf{u}^s[\rho\partial_t^2-\nabla\cdot\mathbf{c}:\nabla]\boldsymbol{\psi}^s + \boldsymbol{\psi}^s\delta\mathbf{f}\right)$$

$$+ \int_0^T dt \oint_{\partial\Omega} d\mathbf{x}\left(\boldsymbol{\psi}^s\cdot\mathbf{c}:\frac{\partial\delta\mathbf{u}^s}{\partial\mathbf{n}} - \mathbf{c}:\frac{\partial\boldsymbol{\psi}^s}{\partial\mathbf{n}}\cdot\delta\mathbf{u}^s\right)$$

$$- \int_0^T dt \int_\Omega d\mathbf{x}\left(\frac{\partial\mathbf{c}}{\partial V_S}\delta V_S\nabla\boldsymbol{\psi}^r\nabla\mathbf{u}^r + \delta\mathbf{u}^r[\rho\partial_t^2-\nabla\cdot\mathbf{c}:\nabla]\boldsymbol{\psi}^r\right)$$

$$+ \int_0^T dt \oint_{\partial\Omega} d\mathbf{x}\left(\boldsymbol{\psi}^r\cdot\mathbf{c}:\frac{\partial\delta\mathbf{u}^r}{\partial\mathbf{n}} - \mathbf{c}:\frac{\partial\boldsymbol{\psi}^r}{\partial\mathbf{n}}\cdot\delta\mathbf{u}^r\right)$$

$$- \int_\Omega d\mathbf{x}\rho[\boldsymbol{\psi}^s\partial_t\delta\mathbf{u}^s - \delta\mathbf{u}^s\partial_t\boldsymbol{\psi}^s + \boldsymbol{\psi}^r\partial_t\delta\mathbf{u}^r - \delta\mathbf{u}^r\partial_t\boldsymbol{\psi}^r]_0^T, \quad \text{(A-4)}$$

where we omit the space and time variables in the integrand for notation simplicity unless the space variable need to be emphasized, $\boldsymbol{\varepsilon}$ is the Levi-Civita tensor, and $\mathbf{n}$ is an outward-pointing unit vector normal to the surface $\partial\Omega$. The regular wavefield is subject to the initial and boundary condition:

$$\mathbf{u}^s(\mathbf{x},0) = 0, \partial_t\mathbf{u}^s(\mathbf{x},0) = 0, \mathbf{u}^s(\mathbf{x},t)|_{\mathbf{x}\to\infty} \to 0, \quad \text{(A-5)}$$

$$\mathbf{u}^r(\mathbf{x},0) = 0, \partial_t\mathbf{u}^r(\mathbf{x},0) = 0, \mathbf{u}^r(\mathbf{x},t)|_{\mathbf{x}\to\infty} \to 0, \quad \text{(A-6)}$$

whereas the adjoint wavefields satisfy the final (at the time of T) and boundary condition:

$$\boldsymbol{\psi}^s(\mathbf{x},T) = 0, \partial_t\boldsymbol{\psi}^s(\mathbf{x},T) = 0, \boldsymbol{\psi}^s(\mathbf{x},t)|_{\mathbf{x}\to\infty} \to 0, \quad \text{(A-7)}$$

$$\boldsymbol{\psi}^r(\mathbf{x},T) = 0, \partial_t\boldsymbol{\psi}^r(\mathbf{x},T) = 0, \boldsymbol{\psi}^r(\mathbf{x},t)|_{\mathbf{x}\to\infty} \to 0, \quad \text{(A-8)}$$

on $\partial\Omega$. Therefore, the surface integrals can be neglected in equation A-4. Then, we have

$$\delta\mathcal{L} = \int_0^T dt \int_\Omega d\mathbf{x} \int d\mathbf{h}\mathbf{h}^2 V_P^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}+\mathbf{h})\hat{u}_s^r(\mathbf{x}+2\mathbf{h})\delta\mathbf{u}^s$$

$$+ \int_0^T dt \int_\Omega d\mathbf{x}\boldsymbol{\varepsilon} \int d\mathbf{h}\mathbf{h}^2 V_S^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h})\delta\mathbf{u}^r$$

$$+ \int_0^T dt 2V_S\hat{u}_s^r \int d\mathbf{h}\mathbf{h}^2\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h},t)\delta V_S$$

$$- \int_0^T dt \int_\Omega d\mathbf{x}\left(\frac{\partial\mathbf{c}}{\partial V_S}\delta V_S\nabla\boldsymbol{\psi}^s\nabla\mathbf{u}^s + \delta\mathbf{u}^s[\rho\partial_t^2-\nabla\cdot\mathbf{c}:\nabla]\boldsymbol{\psi}^s + \boldsymbol{\psi}^s\delta\mathbf{f}\right)$$

$$- \int_0^T dt \int_\Omega d\mathbf{x}\left(\frac{\partial\mathbf{c}}{\partial V_S}\delta V_S\nabla\boldsymbol{\psi}^r\nabla\mathbf{u}^r + \delta\mathbf{u}^r[\rho\partial_t^2-\nabla\cdot\mathbf{c}:\nabla]\boldsymbol{\psi}^r\right). \quad \text{(A-9)}$$

In the absence of perturbations in the model parameters δV_S and $\delta\mathbf{f}$, the variation in the Lagrangian given in equation A-9 is stationary with respect to perturbations $\delta\mathbf{u}^s$ and $\delta\mathbf{u}^r$ provided the Lagrange multipliers, $\boldsymbol{\psi}^s$ and $\boldsymbol{\psi}^r$, satisfy the two AS equations:

$$[\rho\partial_t^2 - \nabla\cdot\mathbf{c}:\nabla]\boldsymbol{\psi}^s = \int d\mathbf{h}\mathbf{h}^2 V_P^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}+\mathbf{h})\hat{u}_s^r(\mathbf{x}+2\mathbf{h})$$

$$[\rho\partial_t^2 - \nabla\cdot\mathbf{c}:\nabla]\boldsymbol{\psi}^r = \boldsymbol{\varepsilon}\int d\mathbf{h}\mathbf{h}^2 V_S^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h}). \quad \text{(A-10)}$$

The first and second equations indicate that the adjoint wavefields $\boldsymbol{\psi}^s$ and $\boldsymbol{\psi}^r$ are determined by the AS equations with the residual source at all imaging spaces. Especially, the Levi-Civita tensor in the second equation implies that the residual source is triggered as a pure shear source without injecting any P-wave energy. Under these conditions, equation A-9 gives the gradient of the objective function with respect to the V_S:

$$\nabla_{V_S}\mathcal{J} = \int_0^T dt 2V_S\hat{u}_s^r \int d\mathbf{h}\mathbf{h}^2\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h},t)$$

$$+ \int_0^T dt \int_\Omega d\mathbf{x}\frac{\partial\mathbf{c}}{\partial V_S}(\nabla\boldsymbol{\psi}^s\nabla\mathbf{u}^s + \nabla\boldsymbol{\psi}^r\nabla\mathbf{u}^r), \quad \text{(A-11)}$$

or in a more detailed manner:

$$\nabla_{V_S}\mathcal{J} = \int_0^T dt 2V_S\hat{u}_s^r \int d\mathbf{h}\mathbf{h}^2\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{u}_p^s(\mathbf{x}-2\mathbf{h},t)$$

$$+ \int_0^T dt \frac{\partial c_{ijkl}}{\partial V_S}\left(\frac{\partial\mathbf{u}_i^s}{\partial x_j}\frac{\partial\boldsymbol{\psi}_k^s}{\partial x_l} + \frac{\partial\mathbf{u}_i^r}{\partial x_j}\frac{\partial\boldsymbol{\psi}_k^r}{\partial x_l}\right). \quad \text{(A-12)}$$

APPENDIX B

DSO-BASED ELASTIC WEMVA WITH VECTOR IMAGING CONDITION

In this appendix, we demonstrate the gradient of elastic DSO with respect to the S-wave velocity using the vector imaging condition (Wang et al., 2016; Du et al., 2017), which can be easily applied in the 3D case. We substitute the imaging result in equation A-1 using the vector imaging condition in Wang et al.'s (2016) paper without the angle-dependent scale factor (for details, see equation 14 in Wang et al., 2016) and yield

$$I_{ps}(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{s}} \int \hat{\mathbf{u}}_p^s(\mathbf{x}-\mathbf{h},t;\mathbf{s})\cdot\hat{\mathbf{u}}_s^r(\mathbf{x}+\mathbf{h},t;\mathbf{s})dt, \quad \text{(B-1)}$$

in which the "$\cdot$" indicates the inner product and $\hat{\mathbf{u}}_p^s$ and $\hat{\mathbf{u}}_s^r$ denote the single-mode vector wavefields that are separated by

$$\hat{\mathbf{u}}_p^s = V_P^2\nabla\nabla\cdot\mathbf{u}^s \quad \text{and} \quad \hat{\mathbf{u}}_s^r = V_S^2\nabla\times\nabla\times\mathbf{u}^r, \quad \text{(B-2)}$$

where the V_P and V_S are the P- and S-wave velocities, respectively. Using the same procedure as in Appendix A, the gradient of the objective function with respect to V_S can be written as a similar expression with equation A-11, namely,

$$\nabla_{V_S}\mathcal{J} = \int_0^T dt 2V_S\hat{\mathbf{u}}_s^r \cdot \int d\mathbf{h}\mathbf{h}^2\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{\mathbf{u}}_p^s(\mathbf{x}-2\mathbf{h},t)$$

$$+ \int_0^T dt \int_\Omega d\mathbf{x}\frac{\partial\mathbf{c}}{\partial V_S}(\nabla\boldsymbol{\psi}^s\nabla\mathbf{u}^s + \nabla\boldsymbol{\psi}^r\nabla\mathbf{u}^r), \quad \text{(B-3)}$$

where the Lagrange multipliers, $\boldsymbol{\psi}^s$ and $\boldsymbol{\psi}^r$, satisfy the two AS equations with newly derived residual sources:

$$[\rho\partial_t^2 - \nabla \cdot \mathbf{c} : \nabla]\boldsymbol{\psi}^s = \nabla\nabla \cdot \int d\mathbf{h}\mathbf{h}^2 V_P^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}+\mathbf{h})\hat{\mathbf{u}}_s^r(\mathbf{x}+2\mathbf{h}),$$

$$[\rho\partial_t^2 - \nabla \cdot \mathbf{c} : \nabla]\boldsymbol{\psi}^r = \nabla\times\nabla\times \int d\mathbf{h}\mathbf{h}^2 V_S^2(\mathbf{x})\frac{\partial^2 I_{ps}}{\partial z^2}(\mathbf{x}-\mathbf{h})\hat{\mathbf{u}}_p^s(\mathbf{x}-2\mathbf{h}). \quad \text{(B-4)}$$

REFERENCES

Adler, F., R. Baina, M. A. Soudani, P. Cardon, and J.-B. Richard, 2008, Nonlinear 3D tomographic least-squares inversion of residual moveout in Kirchhoff prestack-depth-migration common-image gathers: Geophysics, **73**, no. 5, VE13–VE23, doi: 10.1190/1.2956427.

Aki, K., and P. G. Richards, 2002, Quantitative seismology, 2nd ed.: W. H. Freeman.

Alkhalifah, T., 2016, Full-model wavenumber inversion: An emphasis on the appropriate wavenumber continuation: Geophysics, **81**, no. 3, R89–R98, doi: 10.1190/geo2015-0537.1.

Alkhalifah, T., and I. Tsvankin, 1995, Velocity analysis in transversely isotropic media: Geophysics, **60**, 1550–1566, doi: 10.1190/1.1443888.

Amundsen, L., Ø. Pedersen, A. Osen, J. O. A. Robertsson, and M. Landr, 2017, Broadband seismic over/under sources and their designature-deghosting: Geophysics, **82**, no. 5, P61–P73, doi: 10.1190/geo2016-0512.1.

Atkinson, J., and T. Davis, 2011, Multicomponent time-lapse monitoring of two hydraulic fracture stimulations in an unconventional reservoir, Pouce Coupe Field, Canada: 81st Annual International Meeting, SEG, Expanded Abstracts, 4097–4101, doi: 10.1190/1.3628062.

Biondi, B., and W. W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298, doi: 10.1190/1.1801945.

Broto, K., J. Kommedal, and P. G. Folstad, 2003, Anisotropic traveltime tomography for depth consistent imaging of PP and PS data: The Leading Edge, **22**, 114–119, doi: 10.1190/1.1559037.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995, A limited memory algorithm for bound constrained optimization: SIAM Journal on Scientific Computing, **16**, 1190–1208, doi: 10.1137/0916069.

Chauris, H., and E. Cocher, 2017, From migration to inversion velocity analysis: Geophysics, **82**, no. 3, S207–S223, doi: 10.1190/geo2016-0359.1.

Chavent, G., 1974, Identification of function parameters in partial differential equations: Joint Automatic Control Conference, 155–156.

Chavent, G., and C. A. Jacewitz, 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490, doi: 10.1190/1.1443785.

Chaveste, A., Z. Zhao, S. Altan, and J. Gaiser, 2013, Robust rock properties through PP-PS processing and interpretation — Marcellus Shale: The Leading Edge, **32**, 86–92, doi: 10.1190/tle32010086.1.

Chi, B., L. Dong, and Y. Liu, 2015, Correlation-based reflection full-waveform inversion: Geophysics, **80**, no. 4, R189–R202, doi: 10.1190/geo2014-0345.1.

Clapp, R. G., 2008, Reverse time migration: Saving the boundaries: Stanford Exploration Project, 137.

Criss, J., 2007, Another look at full-wave seismic imaging: First Break, **25**, 109–116.

Dai, H., and X. Y. Li, 2008, Effect of errors in the migration velocity model of PS-converted waves on traveltime accuracy in prestack Kirchhoff time migration in weak anisotropic media: Geophysics, **73**, no. 5, S195–S205, doi: 10.1190/1.2957926.

Du, Q., C. Guo, Q. Zhao, X. Gong, C. Wang, and X.-Y. Li, 2017, Vector-based elastic reverse time migration based on scalar imaging condition: Geophysics, **82**, no. 2, S111–S127, doi: 10.1190/geo2016-0146.1.

Du, Q., F. Li, J. Ba, Y. Zhu, and B. Hou, 2012a, Multicomponent joint migration velocity analysis in the angle domain for PP-waves and PS-waves: Geophysics, **77**, no. 1, U1–U13, doi: 10.1190/geo2010-0423.1.

Du, Q., Y. Zhu, and J. Ba, 2012b, Polarity reversal correction for elastic reverse time migration: Geophysics, **77**, no. 2, S31–S41, doi: 10.1190/geo2011-0348.1.

Gaiser, J., 2016, 3C seismic and VSP: Converted waves and vector wavefield application: Springer-Verlag.

Gaiser, J. E., 1996, Multicomponent Vp/Vs correlation analysis: Geophysics, **61**, 1137–1149, doi: 10.1190/1.1444034.

Griewank, A., and A. Walther, 2000, Algorithm 799: Revolve: An implementation of checkpointing for the reverse or adjoint mode of computational differentiation: ACM Transactions on Mathematical Software, **26**, 19–45, doi: 10.1145/347837.347846.

Guitton, A., B. Kaelin, and B. Biondi, 2007, Least square attenuation of reverse time migration artifacts: Geophysics, **72**, no. 1, S19–S23, doi: 10.1190/1.2399367.

Hardage, B., M. V. Deangelo, P. E. Murray, and D. Sava, 2011, Multicomponent seismic technology: SEG.

Lameloise, C.-A., H. Chauris, and M. Noble, 2015, Improving the gradient of the image domain objective function using quantitative migration for a more robust migration velocity analysis: Geophysical Prospecting, **63**, 391–404, doi: 10.1111/1365-2478.12195.

Li, V., I. Tsvankin, and T. Alkhalifah, 2016, Analysis of RTM extended images for VTI media: Geophysics, **81**, no. 3, S139–S150, doi: 10.1190/geo2015-0384.1.

Li, Y., B. Biondi, R. Clapp, and D. Nichols, 2014, Wave-equation migration velocity analysis for VTI models: Geophysics, **79**, no. 3, WA59–WA68, doi: 10.1190/geo2013-0338.1.

Luo, Y., Y. Ma, Y. Wu, H. Liu, and L. Cao, 2016, Full-traveltime inversion: Geophysics, **81**, no. 5, R261–R274, doi: 10.1190/geo2015-0353.1.

Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653, doi: 10.1190/1.1443081.

Ma, Y., and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion: Geophysics, **78**, no. 6, R223–R233, doi: 10.1190/geo2013-0004.1.

Mulder, W., 2008, Automatic velocity analysis with the two-way wave equation: 70th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: 10.3997/2214-4609.20147941.

Nocedal, J., 1999, Numerical optimization: Springer, Springer Series in Operations Research.

Olofsson, B., P. Mitchell, and R. Doychev, 2012, Decimation test on an ocean-bottom node survey: Feasibility to acquire sparse but full-azimuth data: The Leading Edge, **31**, 457–464, doi: 10.1190/tle31040457.1.

Plessix, R. E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: 10.1111/j.1365-246X.2006.02978.x.

Ravasi, M., and A. Curtis, 2013, Elastic imaging with exact wavefield extrapolation for application to ocean-bottom 4C seismic data: Geophysics, **78**, no. 6, S265–S284, doi: 10.1190/geo2013-0152.1.

Reiser, C., T. Bird, and M. Whaley, 2015, Reservoir property estimation using only dual-sensor seismic data — A case study from the West of Shetlands, UKCS: First Break, **33**, 93–101.

Rickett, J., and P. Sava, 2002, Offset and angle domain common image point gathers for shot profile migration: Geophysics, **67**, 883–889, doi: 10.1190/1.1484531.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, **52**, 593–606.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, no. 6, S209–S217, doi: 10.1190/1.2338824.

Shabelansky, A. H., A. E. Malcolm, M. C. Fehler, X. Shang, and W. L. Rodi, 2015, Source independent full wavefield converted-phase elastic migration velocity analysis: Geophysical Journal International, **200**, 954–968, doi: 10.1093/gji/ggu450.

Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, no. 5, VE49–VE59, doi: 10.1190/1.2972021.

Shen, P., W. W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: 73rd Annual International Meeting, SEG, Expanded Abstracts, 2132–2135, doi: 10.1190/1.1817759.

Shin, C., and Y. Ho Cha, 2009, Waveform inversion in the Laplace-Fourier domain: Geophysical Journal International, **177**, 1067–1079, doi: 10.1111/j.1365-246X.2009.04102.x.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248, doi: 10.1190/1.1649391.

Soubaras, R., and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: Geophysics, **72**, no. 5, U67–U73, doi: 10.1190/1.2743612.

Stewart, R. R., J. E. Gaiser, R. J. Brown, and D. C. Lawton, 2003, Converted-wave seismic exploration: Applications: Geophysics, **67**, 1348–1363, doi: 10.1190/1.1512781.

Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, 680–692, doi: 10.1190/1.1443282.

Sun, R., G. A. McMechan, C.-S. Lee, J. Chow, and C.-H. Chen, 2006, Prestack scalar reverse-time depth migration of 3D elastic seismic data: Geophysics, **71**, no. 5, S199–S207, doi: 10.1190/1.2227519.

Symes, W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663, doi: 10.1190/1.1443082.

Tarantola, A., 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903, doi: 10.1190/1.1442046.

Thomsen, L., 1999, Converted-wave reflection seismology over inhomogeneous, anisotropic media: Geophysics, **64**, 678–690, doi: 10.1190/1.1444577.

Ursin, B., F. Maarten, V. de Hoop, S.-K. Foss, and B.-D. Sverre, 2005, Seismic angle tomography: The Leading Edge, **24**, 628–634, doi: 10.1190/1.1946220.

Wang, C., J. Cheng, and B. Arntsen, 2016, Scalar and vector imaging based on wave mode decoupling for elastic reverse time migration in isotropic and transversely isotropic media: Geophysics, **81**, no. 5, S383–S398, doi: 10.1190/geo2015-0704.1.

Wang, C. L., J. B. Cheng, and B. Arntsen, 2015, Numerical pure wave source implementation and its application to elastic reverse time migration in anisotropic media: 77th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: 10.3997/2214-4609.201412834.

Wang, T., J. Cheng, Q. Guo, and C. Wang, 2018, Elastic wave-equation-based reflection kernel analysis and traveltime inversion using wave mode decomposition: Geophysical Journal International, **215**, 450–470, doi: 10.1093/gji/ggy291.

Weibull, W., and B. Arntsen, 2014, Anisotropic migration velocity analysis using reverse time migration: Geophysics, **79**, no. 1, R13–R25, doi: 10.1190/geo2013-0108.1.

Weibull, W. W., and B. Arntsen, 2013, Automatic velocity analysis with reverse-time migration: Geophysics, **78**, no. 4, S179–S192, doi: 10.1190/geo2012-0064.1.

Xu, S., D. Wang, F. Chen, G. Lambaré, and Y. Zhang, 2012, Inversion on reflected seismic wave: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-1473.1.

Yan, J., and P. Sava, 2008, Isotropic angle-domain elastic reverse-time migration: Geophysics, **73**, no. 6, S229–S239, doi: 10.1190/1.2981241.

Yan, J., and P. Sava, 2010, Analysis of converted-wave extended images for migration velocity analysis: 80th Annual International Meeting, SEG, Expanded Abstracts, 1666–1671, doi: 10.1190/1.3513161.

Yang, D., X. Shang, A. Malcolm, M. Fehler, and H. Baek, 2015, Image registration guided wavefield tomography for shear-wave velocity model building: Geophysics, **80**, no. 3, U35–U46, doi: 10.1190/geo2014-0360.1.

Yang, T., and P. Sava, 2011, Wave-equation migration velocity analysis with time-shift imaging: Geophysical Prospecting, **59**, 635–650, doi: 10.1111/j.1365-2478.2011.00954.x.

Zhang, S., Y. Luo, and G. Schuster, 2015, Shot- and angle-domain wave-equation traveltime inversion of reflection data: Synthetic and field data examples: Geophysics, **80**, no. 4, U47–U59, doi: 10.1190/geo2014-0178.1.