Microseismic Image-Domain Elastic Wavefield Tomography

Can Oren & Jeffrey Shragge
Center for Wave Phenomena and Dept. of Geophysics, Colorado School of Mines, Golden CO 80401
coren@mines.edu

ABSTRACT
Elastic time-reverse imaging offers a robust wavefield-based approach for locating microseismic events; however, event location accuracy greatly depends on the veracity of the elastic velocity models (i.e., V_P and V_S) used for wave propagation. In this study, we propose a methodology for microseismic image-domain wavefield tomography using the elastic wave equation and zero-lag and extended source images, the focusing of which is used as a quality control tool for velocity models. The objective function is designed to measure the focusing of time-reversed microseismic energy in zero-lag and extended event images. The function applies penalty operators to source images to highlight poorly focused residual energy caused by backpropagation through erroneous velocity models. Minimizing the objective function leads to a model optimization problem aimed at improving the image-focusing quality. P- and S-wave velocity model updates are computed using the adjoint-state method and build on the zero-lag and extended image residuals that satisfy the differential semblance optimization criterion. Synthetic experiments demonstrate that one can construct accurate elastic velocity models using the proposed method, which can significantly improve the focusing of imaged events leading to, e.g., enhanced fluid-injection programs.

Key words: wavefield tomography, elastic, extended images, microseismic

1 INTRODUCTION

Understanding the distribution and properties of passive sources is one of the main objectives in seismic monitoring at a variety of different scales. In unconventional oil and gas fields, inferring the locations and mechanisms of (micro) earthquakes caused by induced or triggered seismicity is crucial for evaluating the success and safety of fluid injection programs such as hydraulic fracturing and waste-water disposal (Maxwell and Urbancic, 2001). Accurate event locations provide useful information when characterizing fracture lengths and heights for hydraulic stimulation as well as analysing reservoir stress change and potential hazards (Maxwell, 2014; Weingarten et al., 2015). Similarly, at regional and global scales, accurately determining event hypocenters is of particular interest to many geoscientists looking to gain better insight into the spatio-temporal characteristics of faults and their potentially devastating effects (Kiser and Ishii, 2017).

A traditional approach for locating (micro)-earthquakes is based on picking the P- and S-wave traveltimes. Given V_P and V_S velocity models, the traveltime picks can then be inverted to estimate the spatial location of the recorded event. However, the arrival time information can also be used for tomographic updates to improve velocity models. For instance, the double-difference method commonly used in earthquake seismology relies on traveltime differences between pairs of events or stations to find optimal velocity models and event locations (Waldhauser and Ellsworth, 2000; Zhang and Thurber, 2003). In microseismic monitoring, first-break traveltime tomography is commonly used to calibrate isotropic/anisotropic velocity models along raypaths between perforation shots with known spatial and temporal origin and downhole receiver arrays (Bardainne and Gaucher, 2010; Grechka and Yaskevich, 2013). The success of these phase-based picking approaches is, though, heavily dependent on the robustness of traveltime picking process, which is a challenging task for low signal-to-noise ratio (S/N) data.

Alternatively, migration-based event location algorithms have been widely used for passive source scenarios. Like pick-based methods, Kirchhoff-migration techniques require computing ray-traced traveltimes from each candidate event location in the velocity model to each receiver (Kao and Shan, 2004; Baker et al., 2005). The data are then stacked across traveltime trajectories and over all possible source origin times to find the maximum stack power, which is assumed to represent the event location. Although

Kirchhoff techniques are computationally efficient, they struggle for noisy data scenarios, which can affect the accuracy of traveltime picks analogous to the aforementioned methods. Moreover, ray-based methods typically rely on simplifying assumptions (e.g., infinite frequency approximation) and thereby have limited ability to handle complex subsurface models (e.g., neglecting multipathing).

Full-wavefield time-reverse imaging (TRI) techniques are a second type of migration approach that recently have gained popularity (Artman et al., 2010; Chambers et al., 2014; Douma and Snieder, 2015; Witten and Shragge, 2015; Nakata and Beroza, 2016; Rocha et al., 2019; Oren and Shragge, 2021a). By solving the acoustic or elastic wave equation, this class of imaging methods accurately handles wave propagation through complex models, and offers numerous imaging conditions that stack wavefield energy over the time axis, ideally obtaining a well-focused event image. However, TRI methods require computationally expensive wavefield extrapolation as well as accurate velocity models to obtain reliable event location estimates. In most cases, though, such information is not available at monitoring sites due to limited well information (e.g., well logs, perf shots) or other direct or indirect subsurface constraints. This lack of knowledge can lead to inaccurate velocity models that significantly degrade the accuracy of event location estimates. Possible failures in such geophysical observations (e.g., inaccurate location estimates) may comprehensively affect the process of assessing and derisking fluid-injection programs as well as understanding the cause of seismicity.

In TRI applications, the full-wavefield migration-based imaging process typically involves two steps: (1) extrapolating recorded microseismic event data in reverse time through subsurface velocity models (McMechan, 1982), and (2) evaluating an imaging condition (i.e., zero-lag auto/crosscorrelation) that generates a zero-lag image (Artman et al., 2010). If the subsurface model is satisfactorily accurate and a judicious preprocessing is applied to the recorded event data (e.g., source radiation pattern mitigation and/or noise elimination), the peak amplitude of the resulting zero-lag image can be inferred as the correct spatial event location. In such ideal conditions, one can produce complementary PP, SS, and PS zero-lag images by autocorrelating and crosscorrelating the extrapolated P and S wavefields reconstructed from (multicomponent) event data. When these conditions are not met, though, the zero-lag images exhibit erroneous misfocusing energy at incorrect locations. In these instances, a notable shortcoming of zero-lag images is that they rarely provide sufficient information about how to overcome model inaccuracy through velocity updating.

To address the limited sensitivity of zero-lag images to velocity errors, different types of extended imaging conditions have been proposed that may be useful for updating migration velocities for passive seismic scenarios (Witten and Shragge, 2015; Rocha et al., 2019; Oren and Shragge, 2021a). Extended images are an extension of zero-lag images and are typically calculated at the estimated event location. Velocity accuracy information can be directly extracted from extended image volumes by examining whether energy focuses at or away from spatial/temporal zero correlation lag (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011; Witten and Shragge, 2015). Given sufficiently accurate elastic velocity models, extended images should focus at zero lag and the complementary PP, SS, and PS zero-lag images should produce a focal maximum at the same spatial location due to a self-consistency principle (Witten and Shragge, 2015).

Extended images are successfully used as a quality control tool to update velocity models using image misfocusing criteria for active and passive seismic surveys (Shen and Symes, 2008; Burdick et al., 2013; Oren and Shragge, 2020; Shabelansky et al., 2015; Yang and Sava, 2015; Witten and Shragge, 2017a). Velocity updating can be achieved by differential semblance optimization (DSO) (Symes and Carazzone, 1991), which is one of the most common migration velocity analysis (MVA) model building approaches in exploration seismology. The principle of DSO is to minimize differences between neighboring lags or angles associated with a given reflection (Shen and Symes, 2008). In reflection seismology, common-image-point space- and time-lag gathers are effective at reconstructing P-wave velocity models for complex geologic structures (Yang and Sava, 2015; Díaz and Sava, 2017). Similarly, space-lag common-image-gathers can be constructed through a crosscorrelation-based converted-phase imaging condition to perform elastic MVA to jointly update P- and S-wave models (Shabelansky et al., 2015). For microseismic data, Witten and Shragge (2017a) develop a pseudo-acoustic image-domain inversion method to invert for P- and S-wave velocity models and successfully apply it to a 3-D field data set (Witten and Shragge, 2017b).

There are several advantages of full-wavefield microseismic image-domain tomography methods. They require neither first-arrival picking nor the origin time of events, which can be problematic to determine for surface-recorded data exhibiting low S/N. Also, unlike the recently published full waveform inversion (FWI) approaches (Sun et al., 2016; Wang and Alkhalifah, 2018), image-domain inversion methods hold less stringent requirements for initial velocity model accuracy to achieve a successful optimization, and thereby is typically less sensitive to the well-known FWI cycle skipping problem.

In this paper, we present a methodology for computing the isotropic elastic model gradients of an image-domain objective function for surface-recorded microseismic data. To achieve this goal, we first use the kinetic energy term of the PS energy imaging condition (Oren and Shragge, 2021a) to form zero-lag and extended source images. We apply several penalty operators to these images to generate image-domain residuals, which form the basis of a multiterm objective function. Finally, we use the adjoint-state formalism to form the gradients and simultaneously invert for P- and S-wave velocity model parameters. Distinct from the

pseudo-acoustic approach proposed by Witten and Shragge (2017a), our methodology is fully elastic and explicitly generalizes to multiparameter anisotropic inversion, though this extension is not explored here.

We begin by presenting the imaging condition along with the image-domain penalty functions, the combination of which forms the multiterm objective function to be minimized. Next, we derive expressions for the elastic model gradients through adjoint-state tomography (Plessix, 2006). To demonstrate the effectiveness of our inversion methodology, we present a number of synthetic numerical examples. We first highlight the key ingredients of the proposed method by estimating smooth background perturbations on 2D velocity models. We then present a 3D example using structurally complex models with a realistic source and receiver geometry as more challenging and "semi-realistic" settings for estimating elastic model updates. Finally, we conclude with a discussion of the advantages and shortcomings of the proposed inversion methodology as well as its possible applications across scales from time-lapse monitoring programs (e.g., carbon capture and storage projects) to regional/global tomography.

2 THEORY

The image-domain adjoint-state elastic inversion of microseismic data presented herein is comprised of an iterative application of the following steps: (1) inject and individually backpropagate separated multicomponent P- and S-wave data of a single event to reconstruct the wavefields (state variables) using the elastic wave equation; (2) apply the elastic TRI process using the back-propagated P and S wavefields to generate zero-lag and extended source images; (3) evaluate the objective function by summing the image-domain residuals formed by applying penalty operators to annihilate energy consistently focused amongst the zero-lag images or in the vicinity of zero lag in extended images; (4) calculate the adjoint-state variable using the residual energy; (5) form the gradients by correlating the state and adjoint-state variables; (6) use a line-search method to find an optimal step-length to scale the model gradients and determine the magnitude of the multiparameter model updates; and (7) update the model parameters. We detail each inversion process step below.

2.1 Steps 1 & 2: Elastic time-reverse propagation and imaging

We consider the source-free elastic wave equation (EWE) in a slowly varying isotropic medium in a unbounded domain:

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

where $\mathbf{u}(\mathbf{x},t)$ is the displacement field as a function of space $(\mathbf{x})$ and time (t); $\hat{\alpha}(\mathbf{x})$ and $\hat{\beta}(\mathbf{x})$ are defined through the P- and S-wave velocities (i.e., V_P and V_S) as $\hat{\alpha} = V_P^2 = (\lambda+2\mu)/\rho$ and $\hat{\beta} = V_S^2 = \mu/\rho$; $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$, and $\rho(\mathbf{x})$ are the two Lamé parameters and density, respectively; ∇, $\nabla\cdot$, and $\nabla\times$ are the gradient, divergence, and curl operators; and two superscript dots on the displacement field $\mathbf{u}$ denote second-order time differentiation.

To generate microseismic source images from which we derive event location estimates, we use an elastic TRI procedure involving backpropagating injected P- and S-wave data components separated beforehand through data-domain preprocessing (Oren and Shragge, 2021a) and applying an imaging condition. Because we consider isotropic elastic inversion in this study, we find it sufficient to compute zero-lag and extended source images using only the kinetic energy term of the PS energy imaging condition (Oren and Shragge, 2021a) to derive the gradient terms:

$$I_{\alpha\beta}(\mathbf{x},\boldsymbol{\lambda},e) = \int_{t_{min}}^{t_{max}} \rho(\mathbf{x}+\boldsymbol{\lambda})\,\dot{\mathbf{u}}_\alpha^\dagger(\mathbf{x}+\boldsymbol{\lambda},t,e)\cdot\dot{\mathbf{u}}_\beta^\dagger(\mathbf{x}-\boldsymbol{\lambda},t,e)\,\mathrm{d}t, \tag{2}$$

where $\dot{\mathbf{u}}_\alpha^\dagger(\mathbf{x},t,e)$ and $\dot{\mathbf{u}}_\beta^\dagger(\mathbf{x},t,e)$ approximately represent P- and S-wave particle velocity fields where the † symbol denotes adjoint; $\boldsymbol{\lambda} = (\lambda_x,\lambda_y,\lambda_z)$ is the vector space-lag extension (Sava and Vasconcelos, 2011); and e is the event index. The extended imaging condition is described in a generic form and reduces to the zero-lag imaging condition when $(\lambda_x,\lambda_y,\lambda_z) = (0,0,0)$ m. Due to the unknown source origin time, the imaging condition involves a summation over time to marginalize the temporal dependence. While equation 2 is specific for computing PS crosscorrelation zero-lag and extended images; it can also be used to generate zero-lag PP and SS autocorrelation images $\Big($i.e,. $I_{\alpha\alpha}(\mathbf{x},e)$ and $I_{\beta\beta}(\mathbf{x},e)\Big)$ by calculating the dot product of the corresponding adjoint particle velocity field with itself and summing over time.

To attenuate imaging artifacts (i.e., fake modes and/or truncation wavefronts that are correlated with the true wave modes (Yan and Sava, 2008; Rocha et al., 2019)) that arise due to elastic time-reversal wavefield extrapolation, we apply the imaging condition

within a time window $t = [t_{min}, t_{max}]$ that is narrower than the original data window (i.e., $t = [0, T]$) (Oren and Shragge, 2021a). Therefore, the integral evaluation in equation 2 in practice starts from a maximum time ($t_{max} < T$ s) and progresses in reverse time back to a minimum time ($t_{min} \geq 0$ s) using the adjoint isotropic EWE operator. Finally, this imaging condition precludes the costly wave-mode decomposition during backpropagation, which offers an advantage over its crosscorrelation counterpart outlined in Shabelansky et al. (2015). In the elastic TRI procedure, we first compute the zero-lag PP, SS, and PS image volumes using the current velocity models for each event. Because the zero-lag PS image produces higher resolution results compared to its autocorrelation counterparts, we base our event location estimates on the spatial location of maximum absolute amplitude of the PS image. The TRI process is finalized by computing the extended image at the estimated event location. We exclude the negative values in the extended images, which degrade the image residuals as well as the gradient computation procedure in practical implementations.

2.2 Step 3: Image-domain residuals and objective function evaluation

For the inversion procedure, we define our multiterm objective function $\mathcal{J}$ over all selected events to be minimized as

$$\mathcal{J} = \frac{1}{2} \sum_{e} \int \Big[\epsilon_1 \int_{-\boldsymbol{\lambda}}^{\boldsymbol{\lambda}} P_{\lambda}^2(\boldsymbol{\lambda})\, I_{\alpha\beta}^2(\mathbf{x}, \boldsymbol{\lambda}, e)\, \mathrm{d}\boldsymbol{\lambda} + P_{\alpha\beta}^2(\mathbf{x}, e) \Big(\epsilon_2\, I_{\alpha\alpha}^2(\mathbf{x}, e) + \epsilon_3\, I_{\beta\beta}^2(\mathbf{x}, e) \Big) \Big] \mathrm{d}\mathbf{x}, \tag{3}$$

where ϵ_1, ϵ_2, and ϵ_3 are scalar weights that determine the relative contribution of each term in the objective function, and $P_{\lambda}(\boldsymbol{\lambda})$ and $P_{\alpha\beta}(\mathbf{x}, e)$ are to-be-specified extended and zero-lag image-domain penalty operators. These operators are designed to respectively annihilate the focal energy at and about zero lag and at the estimated event location; whatever energy remains is considered to be the image-domain residual. A number of different extended penalty operators have been proposed to serve this purpose (Shragge et al., 2013; Yang and Sava, 2015; Díaz and Sava, 2017), but herein we use a Gaussian function centered at zero lag to penalize the extended image volumes:

$$P_{\lambda}(\boldsymbol{\lambda}) = 1 - \exp\left(- \frac{\lambda_x^2}{2\sigma_x^2} - \frac{\lambda_y^2}{2\sigma_y^2} - \frac{\lambda_x^2}{2\sigma_x^2} \right), \tag{4}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ controls the variances of the Gaussian function in the shift dimensions. Because $I_{\alpha\beta}(\mathbf{x}, e)$ produces the highest resolution among all the zero-lag images due to the P and S wavefields backpropagating at different velocities and coexisting over a narrower zone in space and time, we create the zero-lag penalty operator $P_{\alpha\beta}(\mathbf{x}, e)$ based upon $I_{\alpha\beta}(\mathbf{x}, e)$ to measure the inconsistency in the zero-lag autocorrelation images (Shragge et al., 2013):

$$P_{\alpha\beta}(\mathbf{x}, e) = \mathrm{sech}\left[\frac{w\, I_{\alpha\beta}(\mathbf{x}, e)}{\max\big(I_{\alpha\beta}(\mathbf{x}, e)\big)} \right], \tag{5}$$

where w is a dimensionless parameter used to adjust the penalty width. This multiplicative penalty operator removes the maximum focal energy where the PS and PP or SS zero-lag images are consistent, and tends to unity elsewhere. The energy remaining in the penalized images forms the zero-lag image-domain residuals.

2.3 Steps 4 & 5: Adjoint-state variables and gradient computation

We calculate the gradients (sensitivity kernels) of the objective function in equation 3 with respect to model parameters $\hat{\alpha}$ and $\hat{\beta}$ (see equation 1) by using the perturbation theory with the goal of obtaining an expression like

$$\delta\mathcal{J} = \int \Big(\delta\hat{\alpha}\, \mathcal{K}^{\hat{\alpha}}(\mathbf{x}) + \delta\hat{\beta}\, \mathcal{K}^{\hat{\beta}}(\mathbf{x}) \Big) \mathrm{d}\mathbf{x}, \tag{6}$$

where $\mathcal{K}^{\hat{\alpha}}(\mathbf{x})$ and $\mathcal{K}^{\hat{\beta}}(\mathbf{x})$ are the gradients of the objective function, which is perturbed with respect to the model parameters $\hat{\alpha}$ and $\hat{\beta}$, respectively. Following the adjoint-state method (Plessix, 2006), we obtain the gradient terms as

$$\mathcal{K}^{\hat{\alpha}}(\mathbf{x}) = -\sum_{e} \int_0^T \mathbf{u}_{\alpha}^{\dagger}(\mathbf{x}, t, e) \cdot \boldsymbol{v}^{\alpha}(\mathbf{x}, t, e)\, \mathrm{d}t \tag{7}$$

and

$$\mathcal{K}^{\hat{\beta}}(\mathbf{x}) = \sum_{e} \int_0^T \mathbf{u}_{\beta}^{\dagger}(\mathbf{x}, t, e) \cdot \boldsymbol{v}^{\beta}(\mathbf{x}, t, e)\, \mathrm{d}t, \tag{8}$$

where $\boldsymbol{v}^{\alpha}(\mathbf{x}, t, e)$ and $\boldsymbol{v}^{\beta}(\mathbf{x}, t, e)$ denote the adjoint-state variables, respectively. As seen in equations 7 and 8, the gradients are computed by correlating the state- and adjoint-state variables that are calculated using the forward propagation of an adjoint source, which is a function of backpropagated P- and S-wavefield energy and the penalized zero-lag and extended images. Appendix A presents the full derivation and definitions of the adjoint-state variables as well as the gradient terms. In practice, after calculating the individual gradients for each event, we apply illumination compensation (Warner et al., 2013; Yang et al., 2013) leading to more accurate results by reducing the artifacts at and around source locations. This approach requires dividing the gradient by a stabilized measure of the total adjoint-state wavefield energy. This procedure is followed by the application of a smoothing operator in both vertical and horizontal directions to the individual gradients, which are then summed over the selected events.

2.4 Steps 6 & 7: Step length determination and velocity model updating

After calculating the gradients, we determine the final step length through a multiparameter line-search approach (Tang and Ayeni, 2015) based on the gradient-descent optimization scheme (Nocedal and Wright, 2006). We first search for individual P- and S-wave step lengths and then perform a second line search along a 2D vector whose components are defined by the determined individual P- and S-wave step lengths (Witten and Shragge, 2017a). We then scale the negative of gradients by the final step length to determine the magnitude of the update and add it to the current velocity model:

$$m_{i+1}^{k} = m_i^{k} - h_i\, \mathcal{K}_i^{k}, \quad k = \hat{\alpha}, \hat{\beta}, \tag{9}$$

where m is the velocity model, h is the step length, and i is the iteration number. We iterate the given steps until reaching convergence or meeting a stopping criterion.

3 NUMERICAL EXPERIMENTS

This section describes two synthetic numerical experiments that illustrate the ability of the developed method to reconstruct velocity models leading to improved location estimates. The first numerical experiment presents a 2D example, in which we use elastic velocity models consisting of a smooth 1D gradient background with Gaussian low and high velocity anomalies that represent the inversion targets. The second experiment presents a more realistic example that includes complex 3D subsurface models with sparse acquisition geometry. In our experiments, we model microseismic sources using an Ormbsy wavelet specified by the four corner frequencies $[f_1, f_2, f_3, f_4] = [2, 3, 20, 25]$ Hz (experiment 1) and a 15 Hz Ricker wavelet (experiment 2) with events characterized by a moment-tensor stress-source mechanism. In forward modeling, we use a graphics processing unit (GPU)-based finite-difference time-domain (FDTD) solver with the second-order temporal and eighth-order spatial accuracy stencil (Weiss and Shragge, 2013).

Using the true velocity models, we first forward model synthetic elastic multicomponent microseismic data. Prior to imaging, we separate the direct P- and S-wave arrivals by applying hyperbolic mute functions to the data (Oren and Shragge, 2021a). We then individually image all events by following the elastic TRI strategy (equation 2) using the inaccurate initial velocity models. This step is followed by calculating the model gradients for each event and stacking over all sources to obtain the final results (equations 7 and 8). Finally, using the iterative approach described above, we update the velocity models, which are then used in the TRI procedure to obtain enhanced event locations. To stabilize the gradient computation and balance the resolution of the resulting gradients, we apply a low-pass filter below 9 Hz and 6 Hz to the P- and S-wave mode data, respectively, prior to the forward and adjoint wavefield propagations.

3.1 Experiment 1: : Gradient background model

In the first numerical experiment, we illustrate our method using the P- and S-wave velocity models shown in Figure 1. The models share the same kinematics and include smooth background P- and S-wave velocities, $V_P(z) = 3 + z/2$ km/s and $V_S(z) = \sqrt{3} + z/2$ km/s shown in Figure 1c and 1d, along with two Gaussian low and high velocity variations with maximum and minimum

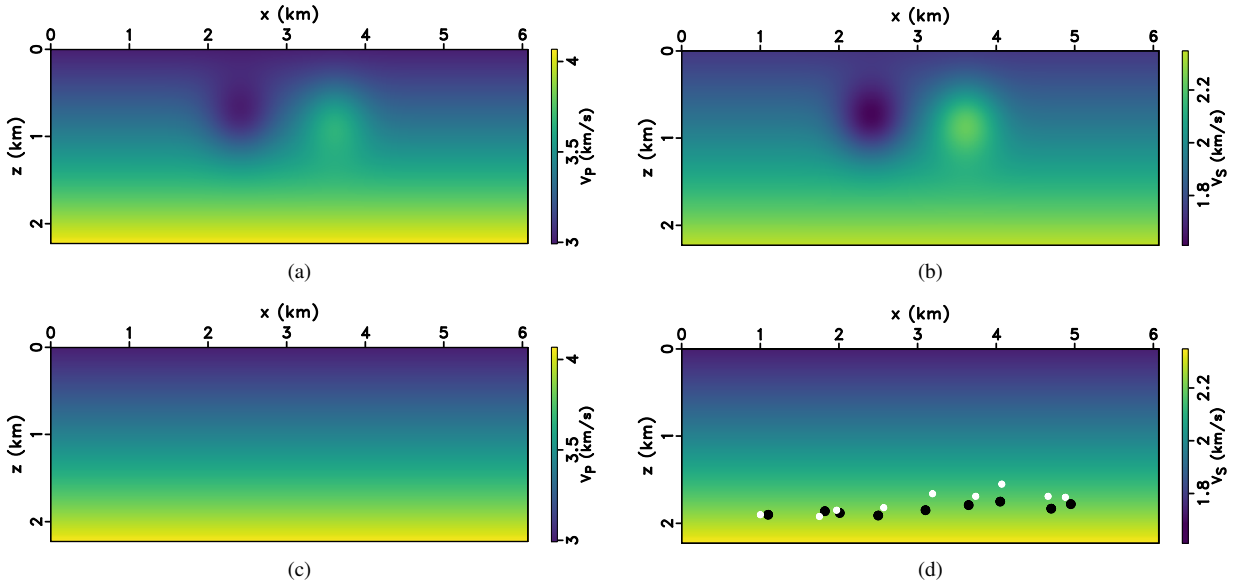


Figure 1. True (a) P- and (b) S-wave velocity models with two Gaussian anomalies in homogeneous background gradients. Background (c) V_P and (d) V_S model overlain with the true (black dots) and initial estimated (white dots) event locations.

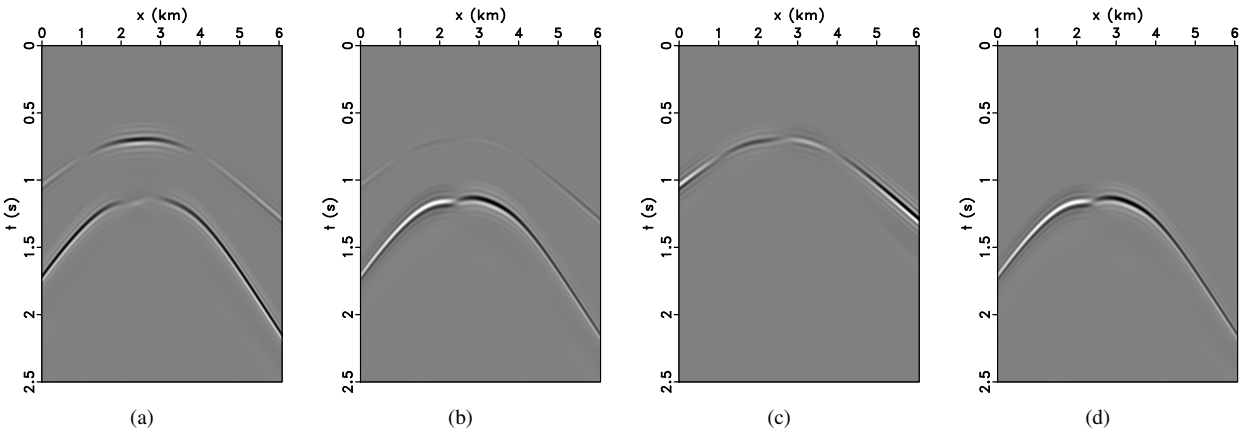


Figure 2. Simulated (a) u_z and (b) u_x components of 2D microseismic data associated with a single event located at $[x, z] = [2.50, 1.91]$ km computed using the true velocity models in Figure 1. (c) P- and (d) S-wave modes in the u_x component separated in the data domain using hyperbolic time muting. Although not shown here, the corresponding wave modes in the u_z component are similarly separated.

perturbations of $\Delta V_P = \Delta V_S = \pm 0.3$ km/s visible in Figure 1a and 1b. We also use a constant density model of $\rho = 2.0$ g/cm^3. The numerical setup consists of a computational domain of dimension $[N_x, N_z] = [608, 224]$, $N_t = 5001$ time steps, temporal and spatial sampling intervals of $\Delta t = 0.5$ ms and $\Delta x = \Delta z = 0.01$ km, respectively, with no free-surface boundary condition applied. The simulated data are recorded at multicomponent (vertical and horizontal motions) receivers placed on the surface at each computation grid point ($\Delta r = 0.01$ km).

Figure 2a and 2b respectively displays the vertical and horizontal components of the surface-recorded 2D elastic data simulated from a single event at $[x, z] = [2.50, 1.91]$ km using the true velocity models in Figure 1. To enhance the S/N of the imaging condition results, we apply a mask around the direct P- and S-wave arrivals in the data domain prior to injection. The wave-mode separation procedure (completed using hyperbolic mute operators) is repeated for all events. This operation is followed by the individual backpropagation of the separated wave modes (Figure 2c and 2d) using the EWE operator (Step 1), and evaluating the imaging condition in equation 2 (Step 2).

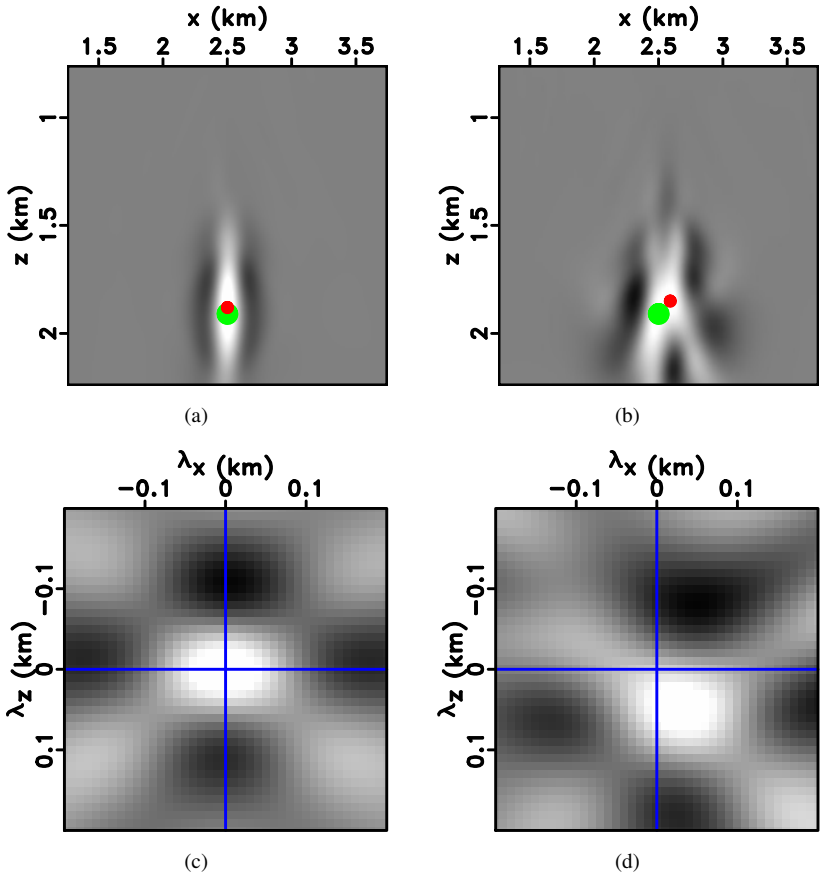


Figure 3. (a)-(b) Zero-lag and (c)-(d) space-lag extended images computed using the true (left column) and background (right column) velocity models. The larger green and smaller red dots respectively denote the true and estimated source locations. The blue crosshairs in the extended images highlight the zero lag. Extended images are evaluated only at the estimated source locations, which correspond to the zero-lag image maxima indicated by the red dots.

As shown in Figure 1c and 1d, we generate the initial velocity models by simply removing the two Gaussian anomalies from the background linear gradient trends. This experiment uses nine irregularly distributed microseismic sources centered about 1.85 km depth, the original locations of which are denoted by black dots in Figure 1d. Using the initial models, we construct the zero-lag images for each event to estimate the event locations as the maximum value in each image. To attenuate the imaging artifacts discussed in the previous section, we generate the zero-lag images computed in the time window $[t_{min}, t_{max}] = [0, 0.4]$ s.

Figure 3a and 3b respectively show the zero-lag images calculated using the true and initial velocity models for the multicomponent data shown in Figure 2. The energy (red dot) in Figure 3a is collapsed at the true source location (larger green dot), which is expected when imaging with the correct velocity models, whereas Figure 3b shows a focus (smaller red dot) smeared due to the inaccurate imaging velocities. We also compute the corresponding extended images constructed for $[\,|\lambda_x|, |\lambda_z|\,] \leq 0.2$ km at the estimated event locations (i.e., maximum amplitude points) in the zero-lag images (Figure 3a and 3b) using the true and background velocity models for evaluating their sensitivity to velocity inaccuracy. As observed in Figure 3c, the extended image is well-focused around zero lag due to the correct velocities while the extended image in Figure 3d exhibits energy shifted away from zero lag due to the inaccurate velocities. The extended image in Figure 3d also reveals the directions of the required velocity updates. Moreover, the event locations estimated from the zero-lag images generated using the initial models are shown as white dots in Figure 1d, and have a mean spatial root-mean-square (RMS) error of 128 m.

For the inversion parameterization, we choose the weight parameters that adjust the contribution of each term in the objective function as $[\epsilon_1, \epsilon_2, \epsilon_3] = [1.0, 0.01, 0.0001]$; the broadness of the zero-lag and extended penalty functions that effectively remove the well-focused energy in the images are $w = 5.0$ and $[\sigma_x, \sigma_z] = [0.05, 0.05]$, respectively, determined through trial and error. The values for ϵ_2 and ϵ_3 balance the influence of the zero-lag terms as well as decrease their contributions one order of magnitude

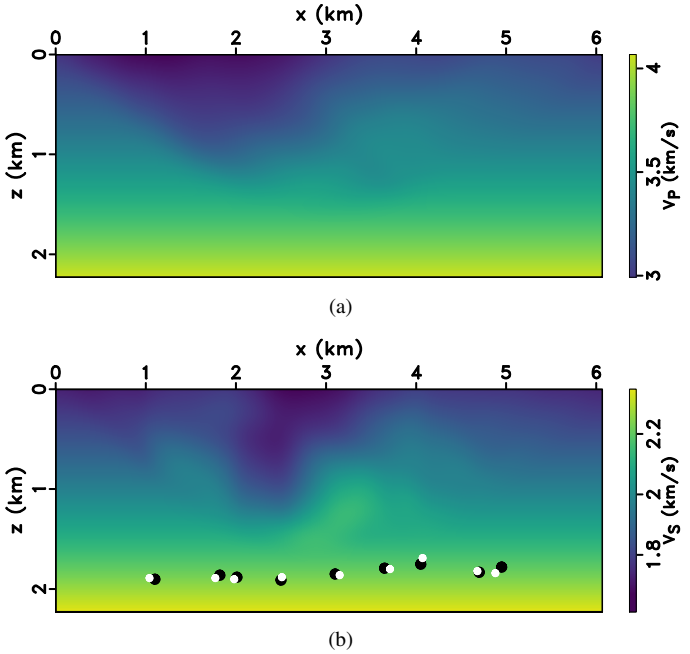


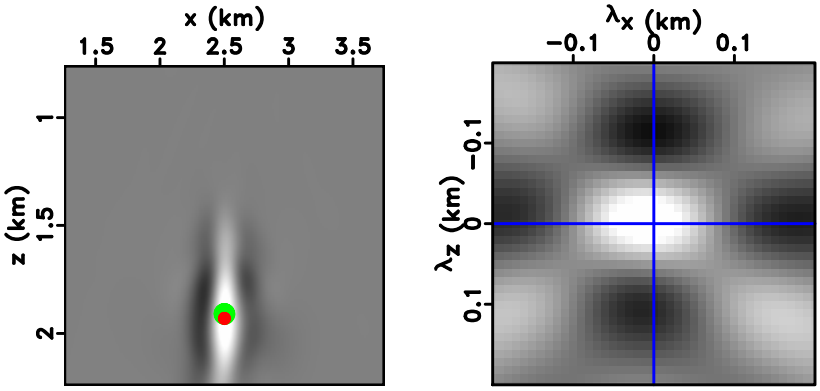
Figure 4. Inverted (a) V_P and (b) V_S model overlain with the true (black dots) and final estimated (white dots) event locations.

compared to the extended term, whose optimization is the primary focus of the inversion procedure due to its highest velocity sensitivity amongst the images. Moreover, because the magnitude of the backpropagated S wavefield $\dot{\mathbf{u}}_\beta^\dagger$ is typically higher than that of the P wavefield $\dot{\mathbf{u}}_\alpha^\dagger$, we downweight the image $I_{\beta\beta}$ more strongly to equalize their relative energy in the waveform.

Figure 4 depicts the V_P and V_S velocity models recovered after 12 iterations of the adjoint-state tomography framework described above. Both the inverted V_P and V_S models show the target Gaussian perturbations with different spatial resolutions. Considering the shorter wavelength of S waves relative to P waves, the inverted V_S model typically features a higher-resolution recovery compared to the inverted V_P model. The inversion results also show that the velocity updates are quite smooth due to the low spatial wavenumbers of the images and forward-scattering nature of the microseismic direct-wave image-domain tomography problem. The white dots in Figure 4b indicate the imaged event locations obtained using the inverted models, which are now closer to the true event locations denoted by the black dots. Figure 5 shows the zero-lag and extended images computed using the inverted models for the multicomponent data shown in Figure 2. When compared with the initial imaging results shown in Figure 3b and 3d, the optimized zero-lag image is better focused and at the correct location while the extended image exhibits energy more tightly focused around zero lag. Finally, Figure 6 presents a crossplot between the initial and final RMS location errors for all events. The dashed line implies no change in location, while points falling below or above the dashed line indicate decrease or increase in location errors, respectively. These results demonstrate that the inversion procedure has reasonably improved the event location estimates and has reduced the average spatial RMS error from 128 m to 52 m or about a 60% reduction in misposition error.

3.2 Experiment 2: SEG/EAEG 3D Overthrust model

The second numerical experiment applies our proposed inversion method to a more complex velocity model – SEG/EAEG 3D Overthrust model (Aminzadeh et al., 1994) displayed in Figure 7a. Because the original 3D Overthrust model includes only the P-wave velocity, we define an S-wave velocity model by assuming an oscillatory PS velocity ratio as a function of depth (Figure 7b) as well as an accompanying density model (not shown) that ranges between $\rho = 2.0 - 3.0\,\text{g/cm}^3$. We place 12 microseismic events (larger blue dots in Figure 7c) with an irregular distribution at depths ranging between $z = 2.4$ and $z = 2.55$ km. Our surface receiver geometry is extracted from a field experiment (Witten and Shragge, 2017b) and consists of 192 non-uniformly distributed 3C receivers (smaller red dots in Figure 7c) covering an approximately 3×3 km^2 area. Using the true models for each event, we forward model synthetic elastic 3C microseismic data using a 3D computational domain of dimension $[N_x, N_y, N_z] = [140, 140, 140]$, $N_t = 2100$ time steps, temporal and spatial sampling intervals of $\Delta t = 1$ ms, and $\Delta x = \Delta y = \Delta z = 0.025$ km with no free-surface boundary condition applied.



(a) (b)

Figure 5. (a) Zero-lag and (b) space-lag extended images computed using the recovered velocity models. The green and red dots respectively show the true and estimated source locations. The extended image is calculated only at the estimated source location, which correspond to the zero-lag image maximum indicated by the red dot. Note how the image focusing in (a) is improved and the energy in (b) is shifted towards zero lag after using the inverted models compared to the imaging results (Figure 3b and 3d) obtained using the initial models.

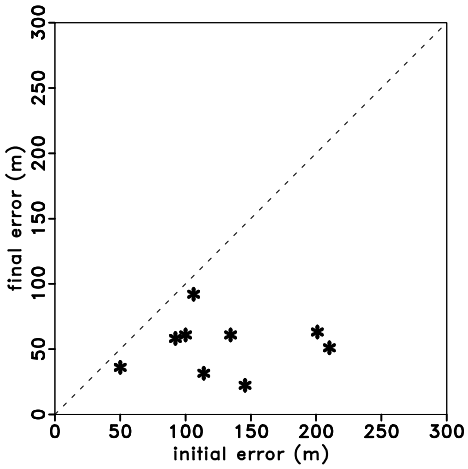
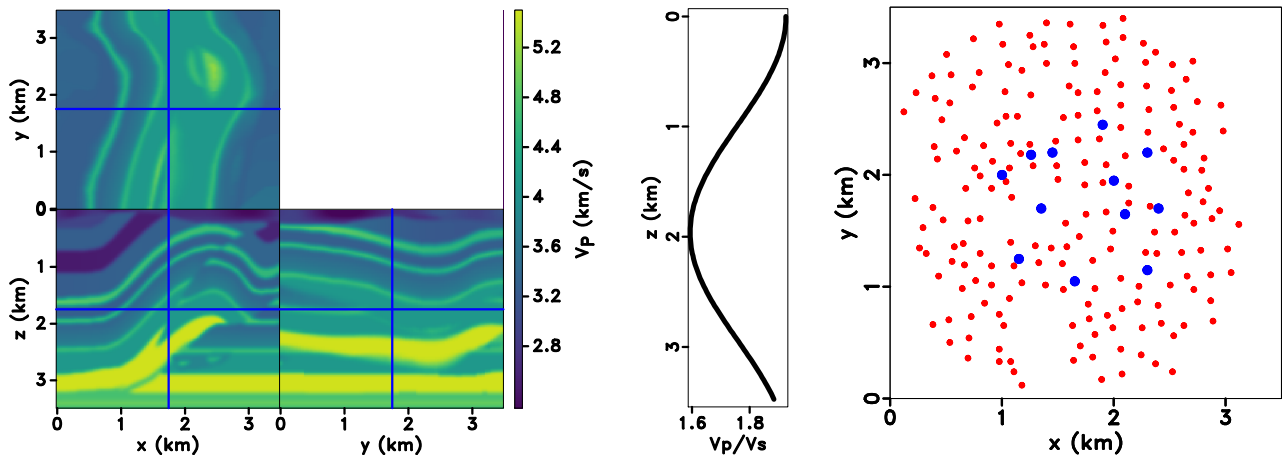


Figure 6. Crossplot of the initial and final RMS location errors of nine forward-modeled events. Points falling along the dashed line indicates no change in error whereas those below the line indicate improved event location estimates. The results show an approximately 60% decrease in location errors for all events.

Our initial model building strategy follows a scenario where we heavily smooth the true V_P model along all the spatial axes to generate an initial V_P model as shown in Figure 8a. Using this information, we also assume an erroneous V_P/V_S ratio of $\sqrt{3}$ when constructing the initial V_S model. Figure 8b and 8c respectively depicts the true percentage V_P and V_S model perturbations (i.e., the difference between the true and initial models), which are the targets of the inversion process.

Following a similar strategy to the previous 2D inversion experiment, we choose ϵ_i values as $[\epsilon_1, \epsilon_2, \epsilon_3] = [1.0, 0.001, 0.0001]$ to favor the optimization of the extended term while balancing the relative contributions of the zero-lag terms in the objective function. Moreover, through trial and error, we select the zero-lag and extended penalty parameterizations to be $w = 5.0$ and $[\sigma_x, \sigma_y, \sigma_z] = [0.1, 0.1, 0.1]$, respectively, to adequately penalize the well-focused energy in images.

Figure 9 shows the inverted V_P and V_S models after applying 11 iterations of the proposed tomography formalism. Most of the update is applied to increase the velocities in both models and corresponds to the smooth part of the high-velocity thrust structure present in the true models. The V_S model updates appear to be more significant compared to those in the V_P model likely due to our stress-source mechanism radiating stronger S-wave energy that dominates the resulting imaged events (Oren and Shragge, 2021a). Furthermore, because of the events that are located relatively deeper in the model (i.e., $z = 2.4 - 2.55$ km), the updates are likely affected by this narrow aperture angle between sources and receivers.



(a) (b) (c)

Figure 7. (a) True P-wave velocity from the 3D Overthrust model and (b) V_P/V_S ratio used to generate the true S-wave velocity model. (c) Illustration of source and receiver geometry employed in the 3D inversion experiment. The larger blue dots indicate the projected source coordinates while the smaller red dots indicate the locations of the multicomponent receivers deployed at the surface.

Figure 10a displays the normalized objective function exhibiting a reduction of $\sim 25\%$ after 11 iterations, beyond which point the convergence curve becomes flat and little-to-no improvement is achieved in the model updates. Figure 10b shows a crossplot of the initial and final RMS location errors calculated based upon the maximum amplitude of the $I_{\alpha\beta}$ images for all sources. We observe a reasonable decrease ($\sim 45\%$) in error with the events plotting closer to the horizontal axis.

The zero-lag and extended images constructed using the initial, inverted, and true V_P and V_S models are displayed in Figure 11 where the crosshairs in the zero-lag panels indicate the true location of a single event at $[x, y, z] = [1.26, 2.18, 2.47]$ km. We construct the zero-lag images in the time window $[t_{min}, t_{max}] = [0, 0.3]$ s to suppress the imaging artifacts and evaluate the extended image gathers for $[\,|\lambda_x|, |\lambda_y|, |\lambda_z|\,] \leq 0.25$ km. Comparing these results, we note defocused energy in the initial image shifted away from the correct event location, whereas the final image exhibits relatively better focused energy centralized around the true location, which is a similar pattern to the image obtained with the true models. Similarly, the final extended image volume extracted at the maximum location of the relative zero-lag image exhibits a lower degree of moveout pattern in comparison with the initial extended imaging result.

4 DISCUSSION

The proposed inversion methodology is typically less prone to the well-known cycle skipping problem of FWI, but generally produces lower resolution inversion results relative to microseismic FWI (Shabelansky et al., 2015; Witten and Shragge, 2017a). However, our inversion results demonstrate that generating model updates that exhibit relatively lower resolution can still produce sufficiently accurate event locations. Additionally, because stacking-based migration methods enhance the S/N of microseismic data, image-domain inversion generally does not suffer from low S/N data as severely as data-domain methods. Furthermore, our inversion approach can be extended to anisotropic media by incorporating the potential wavefield energy term in the PS energy imaging condition as described by Oren and Shragge (2021b). Perturbing the stiffness tensor present in the potential energy term with respect to anisotropy parameters can allow for such an extension. To determine a prospectus for image-domain anisotropic elastic inversion, Oren and Shragge (2021b) present a 3D sensitivity analysis of the PS energy imaging condition to the Thomsen (1986) anisotropy parameters for various transversely isotropic symmetries (e.g., transversely isotropic with a vertical and horizontal symmetry axes as well as orthorhombic). Their analysis demonstrates that the imaged events mostly are affected by errors in ϵ and δ, but show almost no sensitivity to γ for the SEG Advanced Modeling (SEAM) Barrett Unconventional model (Regone et al., 2017).

Because the developed inversion method relies on the forward and adjoint numerical solutions of the elastic wave equation for each seismic event, its computational cost remains as a primary drawback, particularly for 3D applications. Each imaging step requires one P- and one S-wave propagation to construct zero-lag and extended images. The number of space lags used to build extended images is one of the main factors affecting the computational expense. To increase the computational efficiency, we com-

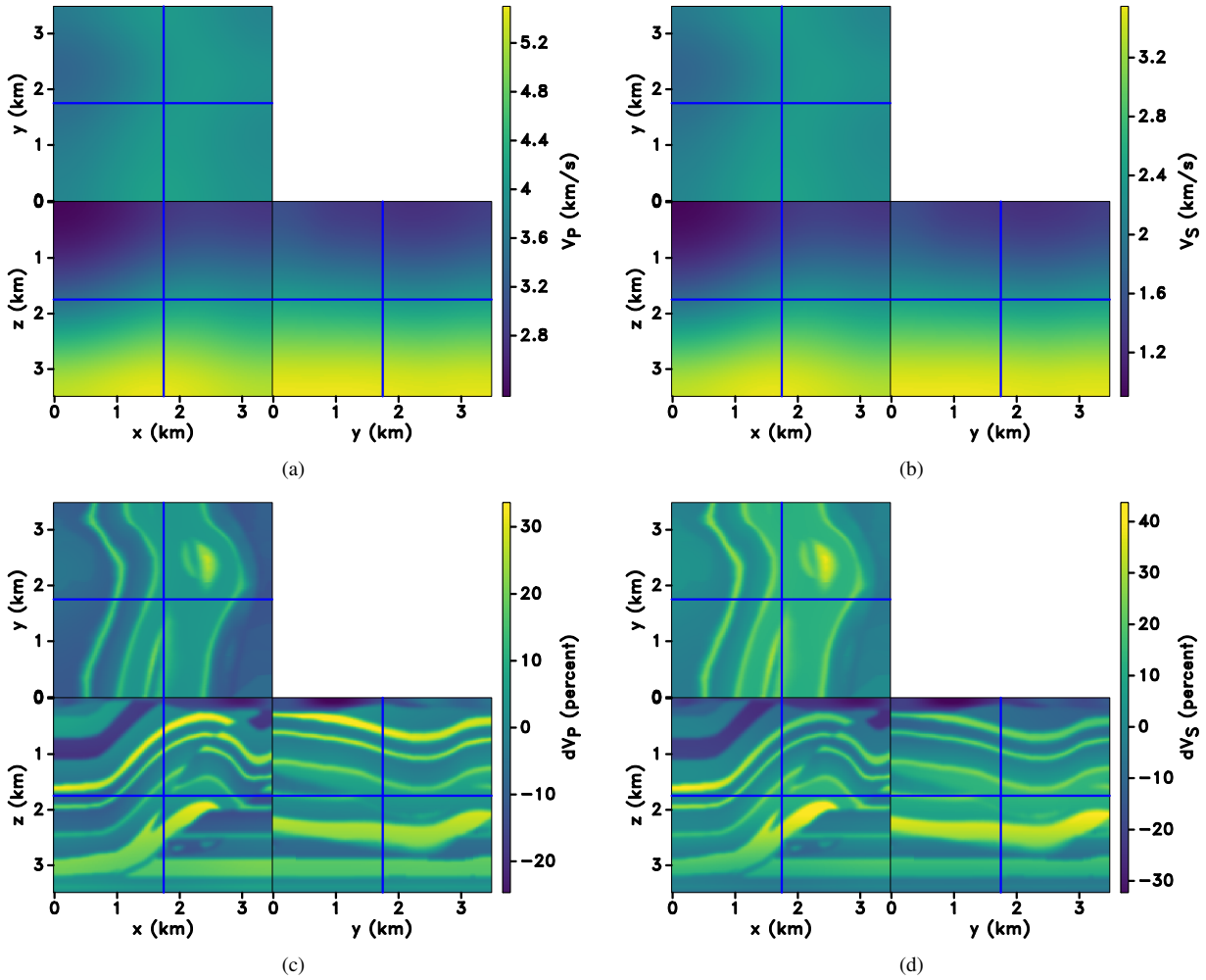
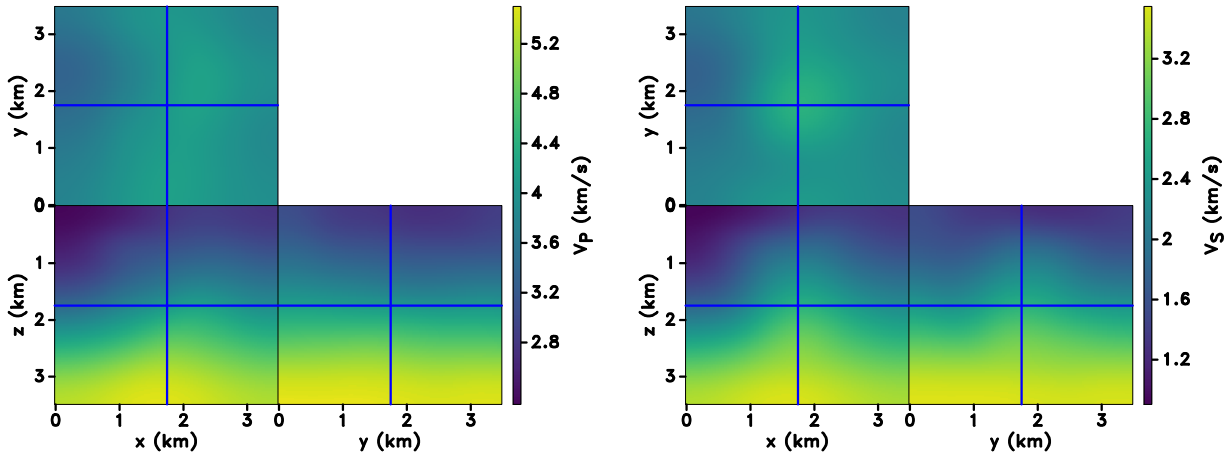


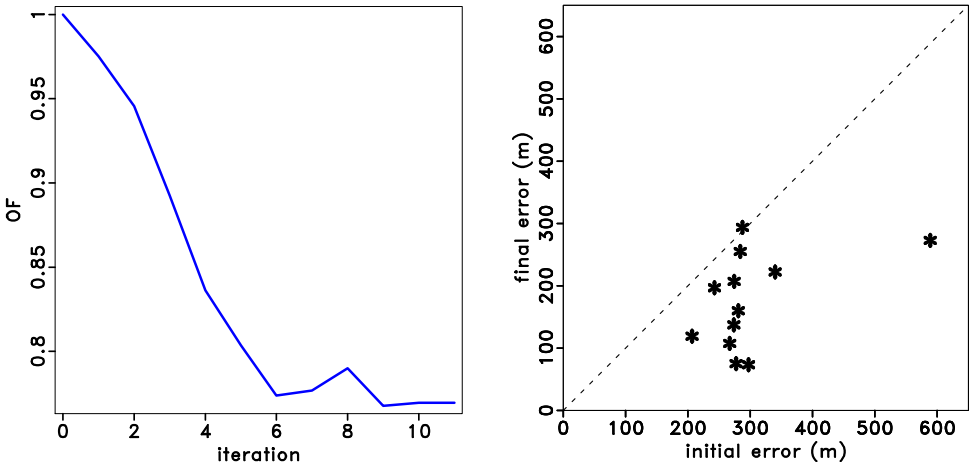
Figure 8. Initial (a) V_P and (b) V_S models along with the true model perturbations for (c) V_P and (d) V_S models, which are the targets of the inversion. The initial V_S model in (b) is constructed by assuming $V_S = V_P/\sqrt{3}$.

pute the zero-lag and extended images during a single imaging experiment. We first independently propagate the P and S wavefields backward in time within the original data window (i.e., $t = [0, T]$) to form the zero-lag image between t_{min} and t_{max}, which is then used to extract the maximum amplitude point for extended imaging. This step is followed by propagating the P and S wavefields at $t = 0$ forward in time until $t = t_{max}$ and evaluating the extended imaging condition at the maximum zero-lag image point between t_{min} and t_{max}. By doing so, we avoid two extra P- and S-wave propagations for extended image construction, which appears to be a substantial increase in efficiency considering the high computational expense of iterative inversion process. The gradient computation requires four elastic propagations (i.e., two backward propagations for calculating P- and S-wave adjoint sources and two forward propagations for calculating P- and S-wave adjoint-state wavefields). The computational cost also depends on the number of spatial locations at which the extended images are calculated for the adjoint source computation (see equations A.17 and A.26). For extended image calculation, we use the same number of spatial locations and space lags per event in our numerical experiments. Finally, the 2D step length determination process requires three imaging steps (i.e., six propagations for P- and S-wave data) for each event, but we opt to perform the first objective function evaluation within the first imaging step to reduce this requirement to two imaging steps. Therefore, we must compute ten elastic wavefield propagations for a single iteration, which, for instance, takes approximately two hours for the previously presented 3D synthetic example when using a single device on a NVIDIA V100 GPU card. However, the aforementioned computational cost can be further reduced using a domain-decomposition strategy across multiple GPU nodes.



(a) (b)

Figure 9. Inverted (a) V_P and (b) V_S models, which exhibit the main high-velocity thrust structure.



(a) (b)

Figure 10. (a) Normalized objective function and (b) crossplot of the initial and final RMS location errors of 12 forward-modeled events.

The image-domain tomography approach presented herein could be applied for a variety of different (time-lapse) monitoring programs (e.g., gas hydrates, geothermal, and CO_2 sequestration projects) (Deichmann and Giardini, 2009; Kim et al., 2018; Kumar et al., 2018) using surface-recorded passive data or active-source vertical seismic profiling (VSP) measurements. The active-source VSP acquisition configuration allows one to exploit the principle of reciprocity by exchanging the locations of sources and receivers to set up a computationally efficient imaging and tomography problem with the sources and receivers located at the borehole and surface, respectively. This configuration mimics a conventional surface microseismic monitoring investigation, which is a very similar scenario to the one we have successfully validated our imaging and inversion algorithms. The proposed inversion method may have practical applications not only in exploration seismology, but also in earthquake seismology to obtain regional or global tomographic models, which could be helpful to further reduce the uncertainty associated with the Earth's heterogeneous structure and dynamics.

5 CONCLUSIONS

We develop a fully automatic adjoint-state wavefield tomography method for multicomponent passive data to jointly reconstruct elastic velocity models. We define a multiterm objective function that optimizes the focusing of different types of event images. We

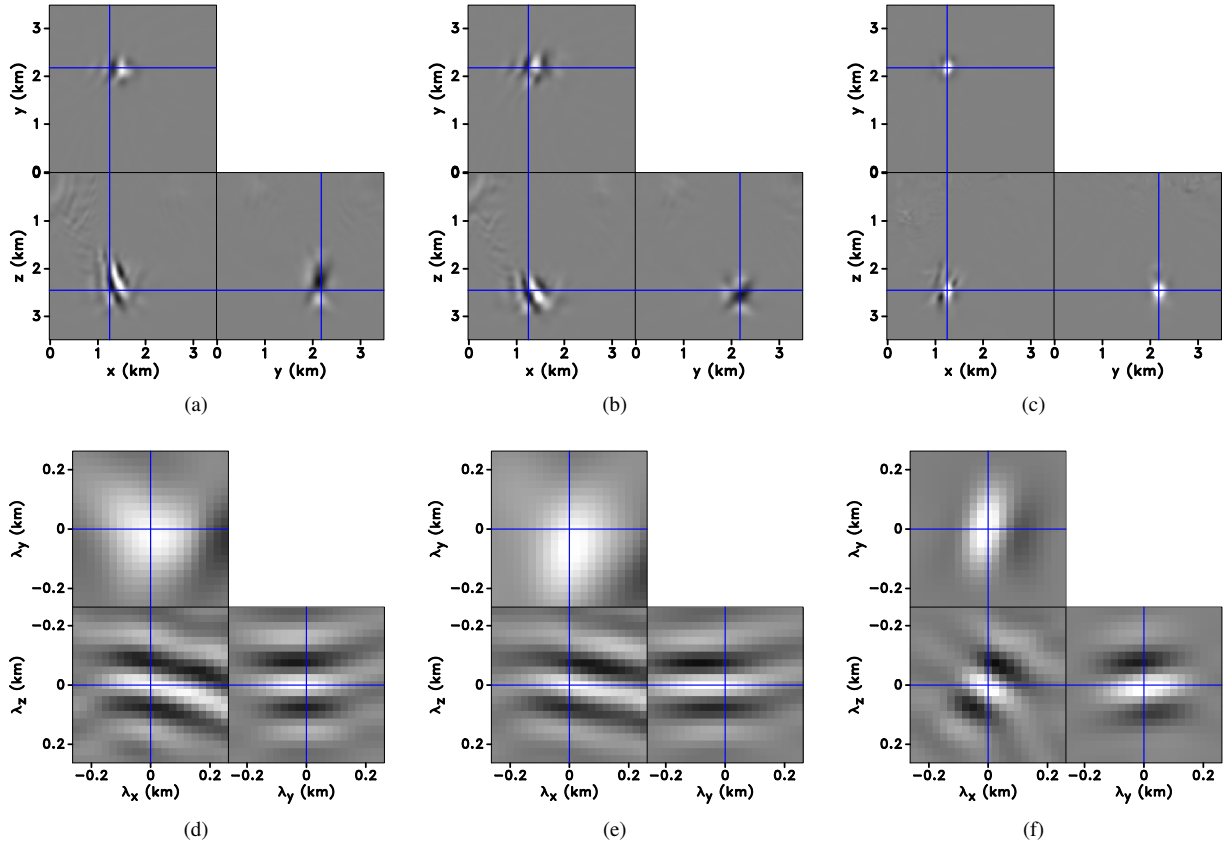


Figure 11. Zero-lag image volumes associated with a single event located at $[x, y, z] = [1.26, 2.18, 2.47]$ km computed using the (a) initial, (b) recovered, and (c) true models. The blue crosshairs in the zero-lag images denote the true source location. Note how the maximum image amplitude shown in white in (b) gets closer to the true event location. The extended image volumes extracted at the corresponding zero-lag image maxima calculated with the (d) initial, (e) recovered, and (f) true models. Note how the final extended image in (e) shows less moveout relative to (d).

conduct 2D and 3D synthetic numerical experiments that gradually present more complex models along with more realistic source and receiver configurations. The numerical experiments show that one can generate accurate elastic models that result in location errors considerably reduced relative to the initial model estimates. In addition to microseismic monitoring, the proposed inversion technique can be possibly applied in other fields including regional/global-scale seismology as well as (time-lapse) gas hydrates, geothermal, and CO_2 sequestration monitoring projects.

6 ACKNOWLEDGEMENTS

We thank the Center for Wave Phenomena consortium sponsors for their financial support. The reproducible examples in this paper were generated on the CSM *Wendian* HPC facilities using the Madagascar software package (http://www.ahay.org).

APPENDIX A - GRADIENT DERIVATION

To derive the P- and S-wave model gradients $\mathcal{K}^{\hat{\alpha}}(\mathbf{x})$ and $\mathcal{K}^{\hat{\beta}}(\mathbf{x})$ in equations 7 and 8, we perturb each term in the objective function (equation 3) with respect to the image for a single event e. We separately present the derivations for the extended and zero-lag terms of the objective functional in the subsections below.

Term 1: Extended image

We introduce $\mathcal{J}_1$ as the extended term of equation 3

$$\mathcal{J}_1 = \frac{\epsilon_1}{2} \iint P_\lambda^2(\boldsymbol{\lambda})\, I_{\alpha\beta}^2(\mathbf{x}, \boldsymbol{\lambda})\, \mathrm{d}\boldsymbol{\lambda}\, \mathrm{d}\mathbf{x}, \tag{A.1}$$

and perturb equation A.1 with respect to the state variables $\mathbf{u}_\alpha^\dagger$ and $\mathbf{u}_\beta^\dagger$ to obtain

$$\delta\mathcal{J}_1 = \epsilon_1 \iint \delta I_{\alpha\beta}(\mathbf{x}, \boldsymbol{\lambda})\, R(\mathbf{x}, \boldsymbol{\lambda})\, \mathrm{d}\boldsymbol{\lambda}\, \mathrm{d}\mathbf{x}, \tag{A.2}$$

where $R(\mathbf{x}, \boldsymbol{\lambda}) = P_\lambda^2(\boldsymbol{\lambda})\, I_{\alpha\beta}(\mathbf{x}, \boldsymbol{\lambda})$ and

$$\delta I_{\alpha\beta}(\mathbf{x}, \boldsymbol{\lambda}) = \int \rho(\mathbf{x}+\boldsymbol{\lambda}) \Big(\delta\dot{\mathbf{u}}_\alpha^\dagger(\mathbf{x}+\boldsymbol{\lambda}, t) \cdot \dot{\mathbf{u}}_\beta^\dagger(\mathbf{x}-\boldsymbol{\lambda}, t) + \dot{\mathbf{u}}_\alpha^\dagger(\mathbf{x}+\boldsymbol{\lambda}, t) \cdot \delta\dot{\mathbf{u}}_\beta^\dagger(\mathbf{x}-\boldsymbol{\lambda}, t) \Big) \mathrm{d}t. \tag{A.3}$$

Substituting equation A.3 into equation A.2 obtains

$$\delta\mathcal{J}_1 := \delta\mathcal{J}_1^{\hat{\alpha}} + \delta\mathcal{J}_1^{\hat{\beta}}, \tag{A.4}$$

where

$$\delta\mathcal{J}_1^{\hat{\alpha}} := \epsilon_1 \iiint \Big(\rho(\mathbf{x}+\boldsymbol{\lambda})\, \delta\dot{\mathbf{u}}_\alpha^\dagger(\mathbf{x}+\boldsymbol{\lambda}, t) \cdot \dot{\mathbf{u}}_\beta^\dagger(\mathbf{x}-\boldsymbol{\lambda}, t)\, R(\mathbf{x}, \boldsymbol{\lambda}) \Big) \mathrm{d}t\, \mathrm{d}\boldsymbol{\lambda}\, \mathrm{d}\mathbf{x} \tag{A.5}$$

and

$$\delta\mathcal{J}_1^{\hat{\beta}} := \epsilon_1 \iiint \Big(\rho(\mathbf{x}+\boldsymbol{\lambda})\, \dot{\mathbf{u}}_\alpha^\dagger(\mathbf{x}+\boldsymbol{\lambda}, t) \cdot \delta\dot{\mathbf{u}}_\beta^\dagger(\mathbf{x}-\boldsymbol{\lambda}, t)\, R(\mathbf{x}, \boldsymbol{\lambda}) \Big) \mathrm{d}t\, \mathrm{d}\boldsymbol{\lambda}\, \mathrm{d}\mathbf{x}. \tag{A.6}$$

To find $\delta\dot{\mathbf{u}}_\alpha^\dagger$ and $\delta\dot{\mathbf{u}}_\beta^\dagger$ in equations A.5 and A.6, we rewrite the isotropic elastic wave equation (equation 1) using linear operator notation (Shabelansky et al., 2015):

$$\mathcal{L}^\dagger\, \mathbf{u}_\alpha^\dagger = \mathbf{d}_\alpha, \tag{A.7}$$

where $\mathbf{d}_\alpha$ is the separated P-wave data vector and $\mathcal{L}^\dagger$ is the adjoint isotropic EWE operator:

$$\mathcal{L}^\dagger = \hat{\alpha}\, \nabla\nabla\cdot - \hat{\beta}\, \nabla\times\nabla\times - \partial_{tt}. \tag{A.8}$$

We perturb equation A.7 with respect to the model parameter $\hat{\alpha}$ and obtain

$$\delta\mathcal{L}^\dagger\, \mathbf{u}_\alpha^\dagger + \mathcal{L}^\dagger\, \delta\mathbf{u}_\alpha^\dagger = 0, \tag{A.9}$$

and solve for $\delta\mathbf{u}_\alpha^\dagger$

$$\delta\mathbf{u}_\alpha^\dagger = -(\mathcal{L}^\dagger)^{-1} \delta\mathcal{L}^\dagger\, \mathbf{u}_\alpha^\dagger, \tag{A.10}$$

where

$$\delta\mathcal{L}^{\dagger} = \delta\hat{\alpha}\,\nabla\nabla\cdot - \delta\hat{\beta}\,\nabla\times\nabla\times . \tag{A.11}$$

Introducing $\delta\mathcal{L}^{\dagger}$ into equation A.10 yields

$$\delta\mathbf{u}_{\alpha}^{\dagger} = -(\mathcal{L}^{\dagger})^{-1}(\delta\hat{\alpha}\,\nabla\nabla\cdot - \delta\hat{\beta}\,\nabla\times\nabla\times)\,\mathbf{u}_{\alpha}^{\dagger}. \tag{A.12}$$

Because the adjoint displacement vector field $\mathbf{u}_{\alpha}^{\dagger}$ is curl-free in isotropic elastic media and if we take the first-order time derivative of each side, equation A.12 becomes

$$\delta\dot{\mathbf{u}}_{\alpha}^{\dagger} = -\partial_t\Big((\mathcal{L}^{\dagger})^{-1}\delta\hat{\alpha}\,\nabla\nabla\cdot\mathbf{u}_{\alpha}^{\dagger}\Big). \tag{A.13}$$

If we substitute this expression into equation A.5

$$\delta\mathcal{J}_1^{\hat{\alpha}} = -\,\epsilon_1\iiint\rho(\mathbf{x}+\boldsymbol{\lambda})\,\partial_t\Big((\mathcal{L}^{\dagger})^{-1}\delta\hat{\alpha}\,\nabla\nabla\cdot\mathbf{u}_{\alpha}^{\dagger}(\mathbf{x}+\boldsymbol{\lambda},t)\Big)\cdot\dot{\mathbf{u}}_{\beta}^{\dagger}(\mathbf{x}-\boldsymbol{\lambda},t)\,R(\mathbf{x},\boldsymbol{\lambda})\mathrm{d}t\,\mathrm{d}\boldsymbol{\lambda}\,\mathrm{d}\mathbf{x}. \tag{A.14}$$

We recognize an inner product and rearrange the integral in equation A.14 using the inner product rule $\Big[$i.e., $\Big\langle\,\rho\,\partial_t\,(\mathcal{L}^{\dagger})^{-1}\,\nabla\nabla\cdot\mathbf{u}_{\alpha}^{\dagger}\,,\,\dot{\mathbf{u}}_{\beta}^{\dagger}\,R\Big\rangle = \Big\langle\,\mathbf{u}_{\alpha}^{\dagger}\,,\,\nabla\nabla\cdot\mathcal{L}^{-1}\,\partial_t\,\rho\,\dot{\mathbf{u}}_{\beta}^{\dagger}\,R\Big\rangle\Big]$ to remove the operator dependence on $\delta\hat{\alpha}$ (Shabelansky et al., 2015; Witten and Shragge, 2017a):

$$\delta\mathcal{J}_1^{\hat{\alpha}} = -\,\epsilon_1\int\delta\hat{\alpha}\iint\mathbf{u}_{\alpha}^{\dagger}(\mathbf{x}+\boldsymbol{\lambda},t)\cdot\Big(\nabla\nabla\cdot\mathcal{L}^{-1}\rho(\mathbf{x}+\boldsymbol{\lambda})\,\ddot{\mathbf{u}}_{\beta}^{\dagger}(\mathbf{x}-\boldsymbol{\lambda},t)\,R(\mathbf{x},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}\Big)\mathrm{d}t\,\mathrm{d}\mathbf{x}, \tag{A.15}$$

where $\mathcal{L}^{-1}$ is the inverse of the forward isotropic elastic wave equation operator. To further simplify the calculation of equation A.15, we apply a shift in the spatial coordinates (Shen and Symes, 2008) and rearrange the terms as the following:

$$\begin{aligned}\delta\mathcal{J}_1^{\hat{\alpha}} &= -\,\epsilon_1\int\delta\hat{\alpha}\int\mathbf{u}_{\alpha}^{\dagger}(\mathbf{x},t)\cdot\Big(\nabla\nabla\cdot\mathcal{L}^{-1}\int\rho(\mathbf{x})\,\ddot{\mathbf{u}}_{\beta}^{\dagger}(\mathbf{x}-2\boldsymbol{\lambda},t)\,R(\mathbf{x}-\boldsymbol{\lambda},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}\Big)\mathrm{d}t\,\mathrm{d}\mathbf{x}\\ &= -\,\epsilon_1\int\delta\hat{\alpha}\int\mathbf{u}_{\alpha}^{\dagger}(\mathbf{x},t)\cdot\boldsymbol{v}_1^{\alpha}(\mathbf{x},t)\,\mathrm{d}t\,\mathrm{d}\mathbf{x}\\ &= \int\delta\hat{\alpha}\,\mathcal{K}_1^{\hat{\alpha}}(\mathbf{x})\,\mathrm{d}\mathbf{x},\end{aligned} \tag{A.16}$$

where

$$\boldsymbol{v}_1^{\alpha}(\mathbf{x},t) = \epsilon_1\,\nabla\nabla\cdot\mathcal{L}^{-1}\int\rho(\mathbf{x})\;\ddot{\mathbf{u}}_{\beta}^{\dagger}(\mathbf{x}-2\boldsymbol{\lambda},t)\,R(\mathbf{x}-\boldsymbol{\lambda},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}, \tag{A.17}$$

is the P-wave adjoint wavefield.

The spatial shift that we apply to the residual extended image $R(\mathbf{x},\boldsymbol{\lambda})$ suggests that we compute numerous extended images in the vicinity of the estimated event location $\mathbf{x}$, which are then used in the calculation of the adjoint source $\boldsymbol{v}_1^{\alpha}(\mathbf{x},t)$.

Similarly, we rewrite the isotropic EWE (equation 1) using linear operator notation to find $\dot{\mathbf{u}}_{\beta}^{\dagger}$ in equation A.6:

$$\mathcal{L}^{\dagger}\,\mathbf{u}_{\beta}^{\dagger} = \mathbf{d}_{\beta}, \tag{A.18}$$

where $\mathbf{d}_{\beta}$ is the separated S-wave data vector. We perturb equation A.18 with respect to the model parameter $\hat{\beta}$ and obtain

$$\delta\mathcal{L}^{\dagger}\,\mathbf{u}_{\beta}^{\dagger}+\mathcal{L}^{\dagger}\,\delta\mathbf{u}_{\beta}^{\dagger}=0, \tag{A.19}$$

and solve for $\delta\mathbf{u}_{\beta}^{\dagger}$

$$\delta\mathbf{u}_{\beta}^{\dagger}=-(\mathcal{L}^{\dagger})^{-1}\delta\mathcal{L}^{\dagger}\,\mathbf{u}_{\beta}^{\dagger}. \tag{A.20}$$

Introducing $\delta\mathcal{L}^{\dagger}$ in equation A.11 into equation A.20 obtains

$$\delta\mathbf{u}_{\beta}^{\dagger}=-(\mathcal{L}^{\dagger})^{-1}(\delta\hat{\alpha}\,\nabla\nabla\cdot-\,\delta\hat{\beta}\,\nabla\times\nabla\times)\,\mathbf{u}_{\beta}^{\dagger}. \tag{A.21}$$

Because the adjoint S-wavefield $\mathbf{u}_{\beta}^{\dagger}$ is divergence-free in isotropic elastic media, if we take the first-order time derivative of each side equation A.21 becomes

$$\delta\dot{\mathbf{u}}_{\beta}^{\dagger}=\partial_t\Big((\mathcal{L}^{\dagger})^{-1}\delta\hat{\beta}\,\nabla\times\nabla\times\,\mathbf{u}_{\beta}^{\dagger}\Big). \tag{A.22}$$

Substituting this expression into equation A.6

$$\delta\mathcal{J}_1^{\hat{\beta}}=\,\epsilon_1\iiint\Big(\rho(\mathbf{x}+\boldsymbol{\lambda})\,\dot{\mathbf{u}}_{\alpha}^{\dagger}(\mathbf{x}+\boldsymbol{\lambda},t)\cdot\partial_t\Big((\mathcal{L}^{\dagger})^{-1}\delta\hat{\beta}\,\nabla\times\nabla\times\,\mathbf{u}_{\beta}^{\dagger}(\mathbf{x}-\boldsymbol{\lambda},t)\Big)\,R(\mathbf{x},\boldsymbol{\lambda})\Big)\mathrm{d}t\,\mathrm{d}\boldsymbol{\lambda}\,\mathrm{d}\mathbf{x}. \tag{A.23}$$

Using the inner product rule, we remove the operator dependence on $\delta\hat{\beta}$:

$$\delta\mathcal{J}_1^{\hat{\beta}}=\,\epsilon_1\iiint\Big(\nabla\times\nabla\times\mathcal{L}^{-1}\rho(\mathbf{x}+\boldsymbol{\lambda})\,\ddot{\mathbf{u}}_{\alpha}^{\dagger}(\mathbf{x}+\boldsymbol{\lambda},t)\cdot\delta\hat{\beta}\,\mathbf{u}_{\beta}^{\dagger}(\mathbf{x}-\boldsymbol{\lambda},t)\,R(\mathbf{x},\boldsymbol{\lambda})\Big)\mathrm{d}t\,\mathrm{d}\boldsymbol{\lambda}\,\mathrm{d}\mathbf{x}. \tag{A.24}$$

To further simplify the calculation of equation A.24, we similarly apply a shift in the spatial coordinates and rearrange the terms as the following:

$$\begin{aligned}\delta\mathcal{J}_1^{\hat{\beta}}&=\,\epsilon_1\iint\Big(\nabla\times\nabla\times\mathcal{L}^{-1}\int\rho(\mathbf{x}+2\boldsymbol{\lambda})\,\ddot{\mathbf{u}}_{\alpha}^{\dagger}(\mathbf{x}+2\boldsymbol{\lambda},t)R(\mathbf{x}+\boldsymbol{\lambda},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}\Big)\cdot\delta\hat{\beta}\,\mathbf{u}_{\beta}^{\dagger}(\mathbf{x},t)\mathrm{d}t\,\mathrm{d}\mathbf{x}\\&=\,\epsilon_1\int\delta\hat{\beta}\int\mathbf{u}_{\beta}^{\dagger}(\mathbf{x},t)\cdot\boldsymbol{v}_1^{\beta}(\mathbf{x},t)\,\mathrm{d}t\,\mathrm{d}\mathbf{x}\\&=\int\delta\hat{\beta}\,\mathcal{K}_1^{\hat{\beta}}(\mathbf{x})\,\mathrm{d}\mathbf{x},\end{aligned} \tag{A.25}$$

where

$$\boldsymbol{v}_1^{\beta}(\mathbf{x},t)=\epsilon_1\,\nabla\times\nabla\times\mathcal{L}^{-1}\int\rho(\mathbf{x}+2\boldsymbol{\lambda})\;\ddot{\mathbf{u}}_{\alpha}^{\dagger}(\mathbf{x}+2\boldsymbol{\lambda},t)\,R(\mathbf{x}+\boldsymbol{\lambda},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}, \tag{A.26}$$

is the S-wave adjoint wavefield.

Terms 2 and 3: Zero-lag autocorrelation images penalized by crosscorrelation image

We derive the adjoint-state variables for the zero-lag autocorrelation images $I_{\alpha\alpha}(\mathbf{x})$ and $I_{\beta\beta}(\mathbf{x})$ penalized by the zero-lag cross-correlation image $I_{\alpha\beta}(\mathbf{x})$ by perturbing

$$\mathcal{J}_2=\frac{\epsilon_2}{2}\int P_{\alpha\beta}^2(\mathbf{x})\,I_{\alpha\alpha}^2(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{A.27}$$

to obtain

$$\delta\mathcal{J}_2 = \epsilon_2 \int \Big(P^2_{\alpha\beta}(\mathbf{x})\, I_{\alpha\alpha}(\mathbf{x})\, \delta I_{\alpha\alpha}(\mathbf{x}) + P_{\alpha\beta}(\mathbf{x})\, \delta P_{\alpha\beta}(\mathbf{x})\, I^2_{\alpha\alpha}(\mathbf{x}) \Big)\, \mathrm{d}\mathbf{x}, \tag{A.28}$$

where

$$\delta P_{\alpha\beta}(\mathbf{x}) = -w\, P_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, \delta I_{\alpha\beta}(\mathbf{x}), \tag{A.29}$$

where $\delta I_{\alpha\beta}(\mathbf{x})$ is defined in equation A.3 with $(\lambda_x, \lambda_y, \lambda_z) = (0, 0, 0)$ m, and $T_{\alpha\beta}(\mathbf{x}) = \tanh\big(w\, I_{\alpha\beta}(\mathbf{x})\big)$. Based on the imaging condition in equation 2, one can form the PP autocorrelation image using

$$I_{\alpha\alpha}(\mathbf{x}) = \int \rho(\mathbf{x})\, \dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t) \cdot \dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)\, \mathrm{d}t, \tag{A.30}$$

to obtain

$$\delta I_{\alpha\alpha}(\mathbf{x}) = 2 \int \rho(\mathbf{x})\, \dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t) \cdot \delta\dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)\, \mathrm{d}t. \tag{A.31}$$

If we substitute $\delta P_{\alpha\beta}(\mathbf{x})$ and $\delta I_{\alpha\alpha}(\mathbf{x})$ into equation A.28, and separate $\mathcal{J}_2$ into component functions $\delta\dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)$ and $\delta\dot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x}, t)$, we find

$$\delta\mathcal{J}_2^{\hat{\alpha}} = 2\,\epsilon_2 \int \Big(P^2_{\alpha\beta}(\mathbf{x})\, I_{\alpha\alpha}(\mathbf{x}) \int \rho(\mathbf{x})\, \dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t) \cdot \delta\dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)\, \mathrm{d}t - w\, P^2_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, I^2_{\alpha\alpha}(\mathbf{x}) \int \rho(\mathbf{x})\, \delta\dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t) \cdot \dot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x}, t)\, \mathrm{d}t \Big)\, \mathrm{d}\mathbf{x}, \tag{A.32}$$

and

$$\delta\mathcal{J}_2^{\hat{\beta}} = -\,\epsilon_2\, w \int P^2_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, I^2_{\alpha\alpha}(\mathbf{x}) \int \rho(\mathbf{x})\, \dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)\ \cdot\ \delta\dot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x}, t)\, \mathrm{d}t\, \mathrm{d}\mathbf{x}. \tag{A.33}$$

If we substitute $\delta\dot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)$ (equation A.12) and $\delta\dot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x}, t)$ (equation A.22) into equations A.32 and A.33 and rearrange the terms as done in equations A.16 and A.25, we find

$$\begin{aligned} \delta\mathcal{J}_2^{\hat{\alpha}} &= \epsilon_2 \int \delta\hat{\alpha} \int \mathbf{u}^{\dagger}_{\alpha}(\mathbf{x}, t) \cdot \boldsymbol{v}_2^{\alpha}(\mathbf{x}, t)\, \mathrm{d}t\, \mathrm{d}\mathbf{x} \\ &= \int \delta\hat{\alpha}\, \mathcal{K}_2^{\hat{\alpha}}(\mathbf{x})\, \mathrm{d}\mathbf{x}, \end{aligned} \tag{A.34}$$

and

$$\begin{aligned} \delta\mathcal{J}_2^{\hat{\beta}} &= \epsilon_2 \int \delta\hat{\beta} \int \mathbf{u}^{\dagger}_{\beta}(\mathbf{x}, t) \cdot \boldsymbol{v}_2^{\beta}(\mathbf{x}, t)\, \mathrm{d}t\, \mathrm{d}\mathbf{x} \\ &= \int \delta\hat{\beta}\, \mathcal{K}_2^{\hat{\beta}}(\mathbf{x})\, \mathrm{d}\mathbf{x}, \end{aligned} \tag{A.35}$$

where the adjoint-state variable contributions for this term are as the following:

$$\begin{aligned} \boldsymbol{v}_2^{\hat{\alpha}}(\mathbf{x}, t) = \epsilon_2\, \nabla\nabla \cdot \mathcal{L}^{-1} \Big(&-2\, P^2_{\alpha\beta}(\mathbf{x})\, I_{\alpha\alpha}(\mathbf{x})\, \rho(\mathbf{x}) \int \ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}, t)\, \mathrm{d}t \\ &+ w\, P^2_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, I^2_{\alpha\alpha}(\mathbf{x})\, \rho(\mathbf{x}) \int \ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x}, t)\, \mathrm{d}t \Big), \end{aligned} \tag{A.36}$$

and

$$\boldsymbol{v}_2^{\hat{\beta}}(\mathbf{x},t) = -\epsilon_2\, w\, \nabla\times\nabla\times\mathcal{L}^{-1}\Big(\, P^2_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, I^2_{\alpha\alpha}(\mathbf{x})\, \rho(\mathbf{x}) \int \ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x},t)\,\mathrm{d}t\Big). \tag{A.37}$$

Following a similar strategy used in the derivation for term 2, we find the expressions for term 3 $\left(\mathcal{J}_3 = \frac{\epsilon_3}{2}\int P^2_{\alpha\beta}(\mathbf{x})\, I^2_{\beta\beta}(\mathbf{x})\mathrm{d}\mathbf{x}\right)$ to be

$$\begin{aligned}\delta\mathcal{J}_3^{\hat{\alpha}} &= \epsilon_3 \int\delta\hat{\alpha} \int\mathbf{u}^{\dagger}_{\alpha}(\mathbf{x},t)\cdot\boldsymbol{v}_3^{\alpha}(\mathbf{x},t)\,\mathrm{d}t\,\mathrm{d}\mathbf{x}\\ &= \int\delta\hat{\alpha}\,\mathcal{K}_3^{\hat{\alpha}}(\mathbf{x})\,\mathrm{d}\mathbf{x},\end{aligned} \tag{A.38}$$

and

$$\begin{aligned}\delta\mathcal{J}_3^{\hat{\beta}} &= \epsilon_3 \int\delta\hat{\beta} \int\mathbf{u}^{\dagger}_{\beta}(\mathbf{x},t)\cdot\boldsymbol{v}_3^{\beta}(\mathbf{x},t)\,\mathrm{d}t\,\mathrm{d}\mathbf{x}\\ &= \int\delta\hat{\beta}\,\mathcal{K}_3^{\hat{\beta}}(\mathbf{x})\,\mathrm{d}\mathbf{x},\end{aligned} \tag{A.39}$$

where the adjoint-state variables are defined as

$$\boldsymbol{v}_3^{\hat{\alpha}}(\mathbf{x},t) = \epsilon_3\, w\, \nabla\nabla\cdot\mathcal{L}^{-1}\Big(P^2_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, I^2_{\beta\beta}(\mathbf{x})\, \rho(\mathbf{x}) \int \ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x},t)\,\mathrm{d}t\Big), \tag{A.40}$$

and

$$\begin{aligned}\boldsymbol{v}_3^{\hat{\beta}}(\mathbf{x},t) = \epsilon_3\,\nabla\times\nabla\times\mathcal{L}^{-1}\Big(\,&2\,P^2_{\alpha\beta}(\mathbf{x})\, I_{\beta\beta}(\mathbf{x})\,\rho(\mathbf{x}) \int \ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x},t)\,\mathrm{d}t\\ &- w\, P^2_{\alpha\beta}(\mathbf{x})\, T_{\alpha\beta}(\mathbf{x})\, I^2_{\beta\beta}(\mathbf{x})\,\rho(\mathbf{x}) \int \ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x},t)\,\mathrm{d}t\Big).\end{aligned} \tag{A.41}$$

Total gradient

We form the total perturbation of the objective functional using the terms above

$$\delta\mathcal{J} = \delta\mathcal{J}^{\hat{\alpha}} + \delta\mathcal{J}^{\hat{\beta}}, \tag{A.42}$$

where

$$\begin{aligned}\delta\mathcal{J}^{\hat{\alpha}} &= \int\delta\hat{\alpha}\,\mathcal{K}^{\hat{\alpha}}(\mathbf{x})\,\mathrm{d}\mathbf{x}\\ &= \int\delta\hat{\alpha}\int\mathbf{u}^{\dagger}_{\alpha}(\mathbf{x},t)\cdot\sum_{i=1}^{3}\boldsymbol{v}_i^{\alpha}(\mathbf{x},t)\,\mathrm{d}t\,\mathrm{d}\mathbf{x},\end{aligned} \tag{A.43}$$

and

$$\begin{aligned}\delta\mathcal{J}^{\hat{\beta}} &= \int\delta\hat{\beta}\,\mathcal{K}^{\hat{\beta}}(\mathbf{x})\,\mathrm{d}\mathbf{x}\\ &= \int\delta\hat{\beta}\int\mathbf{u}^{\dagger}_{\beta}(\mathbf{x},t)\cdot\sum_{i=1}^{3}\boldsymbol{v}_i^{\beta}(\mathbf{x},t)\,\mathrm{d}t\,\mathrm{d}\mathbf{x},\end{aligned} \tag{A.44}$$

where $\boldsymbol{v}^{\alpha}(\mathbf{x},t)$ and $\boldsymbol{v}^{\beta}(\mathbf{x},t)$ are the adjoint-state variables constructed by forward propagating the summation of the individual adjoint sources derived in the previous subsections. The final adjoint-state variables in equations A.43 and A.44 can be written as

$$\begin{aligned}\boldsymbol{v}^{\alpha}(\mathbf{x},t) = \nabla\nabla\cdot\mathcal{L}^{-1}\Big[&\epsilon_1\int\rho(\mathbf{x})\;\ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x}-2\boldsymbol{\lambda},t)\,R(\mathbf{x}-\boldsymbol{\lambda},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}\\ &+\epsilon_2\Big(-2\,P^2_{\alpha\beta}(\mathbf{x})\,I_{\alpha\alpha}(\mathbf{x})\,\rho(\mathbf{x})\int\ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x},t)\,\mathrm{d}t + w\,P^2_{\alpha\beta}(\mathbf{x})\,T_{\alpha\beta}(\mathbf{x})\,I^2_{\alpha\alpha}(\mathbf{x})\,\rho(\mathbf{x})\int\ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x},t)\,\mathrm{d}t\Big)\\ &+\epsilon_3\,w\,P^2_{\alpha\beta}(\mathbf{x})\,T_{\alpha\beta}(\mathbf{x})\,I^2_{\beta\beta}(\mathbf{x})\,\rho(\mathbf{x})\int\ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x},t)\,\mathrm{d}t\Big],\end{aligned}\tag{A.45}$$

and

$$\begin{aligned}\boldsymbol{v}^{\beta}(\mathbf{x},t) = \nabla\times\nabla\times\mathcal{L}^{-1}\Big[&\epsilon_1\int\rho(\mathbf{x}+2\boldsymbol{\lambda})\;\ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x}+2\boldsymbol{\lambda},t)\,R(\mathbf{x}+\boldsymbol{\lambda},\boldsymbol{\lambda})\,\mathrm{d}\boldsymbol{\lambda}\\ &-\epsilon_2\,w\,P^2_{\alpha\beta}(\mathbf{x})\,T_{\alpha\beta}(\mathbf{x})\,I^2_{\alpha\alpha}(\mathbf{x})\,\rho(\mathbf{x})\int\ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x},t)\,\mathrm{d}t\\ &+\epsilon_3\Big(2\,P^2_{\alpha\beta}(\mathbf{x})\,I_{\beta\beta}(\mathbf{x})\,\rho(\mathbf{x})\int\ddot{\mathbf{u}}^{\dagger}_{\beta}(\mathbf{x},t)\,\mathrm{d}t - w\,P^2_{\alpha\beta}(\mathbf{x})\,T_{\alpha\beta}(\mathbf{x})\,I^2_{\beta\beta}(\mathbf{x})\,\rho(\mathbf{x})\int\ddot{\mathbf{u}}^{\dagger}_{\alpha}(\mathbf{x},t)\,\mathrm{d}t\Big)\Big].\end{aligned}\tag{A.46}$$

REFERENCES

Aminzadeh, F., N. Burkhard, L. Nicoletis, F. Rocca, and K. Wyatt, 1994, SEG/EAEG 3-D modeling project: 2nd update: The Leading Edge, **13**, 949–952.

Artman, B., I. Podladtchikov, and B. Witten, 2010, Source location using time-reverse imaging: Geophysical Prospecting, **58**, 861–873.

Baker, T., R. Granat, and R. W. Clayton, 2005, Real-time earthquake location using Kirchhoff reconstruction: Bulletin of the Seismological Society of America, **95**, 699–707.

Bardainne, T., and E. Gaucher, 2010, Constrained tomography of realistic velocity models in microseismic monitoring using calibration shots: Geophysical Prospecting, **58**, 739–753.

Burdick, S., M. V. de Hoop, S. Wang, and R. D. van der Hilst, 2013, Reverse-time migration-based reflection tomography using teleseismic free surface multiples: Geophysical Journal International, **196**, 996–1017.

Chambers, K., B. D. Dando, G. A. Jones, R. Velasco, and S. A. Wilson, 2014, Moment tensor migration imaging: Geophysical Prospecting, **62**, 879–896.

Deichmann, N., and D. Giardini, 2009, Earthquakes induced by the stimulation of an enhanced geothermal system below Basel (Switzerland): Seismological Research Letters, **80**, no. 5, 784–798.

Díaz, E., and P. Sava, 2017, Cascaded wavefield tomography and inversion using extended common-image-point gathers: A case study: Geophysics, **82**, no. 5, S391–S401.

Douma, J., and R. Snieder, 2015, Focusing of elastic waves for microseismic imaging: Geophysical Journal International, **200**, 390–401.

Grechka, V., and S. Yaskevich, 2013, Inversion of microseismic data for triclinic velocity models: Geophysical Prospecting, **61**, 1159–1170.

Kao, H., and S. J. Shan, 2004, The source-scanning algorithm: Mapping the distribution of seismic sources in time and space: Geophysical Journal International, **157**, 589–594.

Kim, K.-H., J.-H. Ree, Y. Kim, S. Kim, S. Y. Kang, and W. Seo, 2018, Assessing whether the 2017 M_w5.4 Pohang earthquake in South Korea was an induced event: Science, **360**, 1007–1009.

Kiser, E., and M. Ishii, 2017, Back-projection imaging of earthquakes: Annual Review of Earth and Planetary Sciences, **45**, 271–299.

Kumar, A., K. Chao, R. Hammack, W. Harbert, W. Ampomah, R. Balch, and L. Garcia, 2018, Surface seismic monitoring of an active CO_2-EOR operation in the Texas Panhandle using broadband seismometers: 88th Annual International Meeting, SEG, Expanded Abstracts, 3027–3031.

Maxwell, S., 2014, Microseismic imaging of hydraulic fracturing: Society of Exploration Geophysicists.

Maxwell, S. C., and T. I. Urbancic, 2001, The role of passive microseismic monitoring in the instrumented oil field: Leading Edge, **20**, 636–639.

McMechan, G. A., 1982, Determination of source parameters by wavefield extrapolation: Geophysical Journal International, **71**, 613–628.

Nakata, N., and G. C. Beroza, 2016, Reverse time migration for microseismic sources using the geometric mean as an imaging condition: Geophysics, **81**, no. 2, KS51–KS60.

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

Oren, C., and J. Shragge, 2020, Image-domain elastic wavefield tomography for passive data: 90th Annual International Meeting, SEG, Expanded Abstracts, 3669–3673.

———, 2021a, PS energy imaging condition for microseismic data – Part 1: Theory and applications in 3D isotropic media: Geophysics, **86**, no. 2, KS37–KS48.

———, 2021b, PS energy imaging condition for microseismic data – Part 2: Sensitivity analysis in 3D anisotropic media: Geophysics, **86**, no. 2, KS49–KS62.

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.

Regone, C., J. Stefani, P. Wang, C. Gerea, G. Gonzalez, and M. Oristaglio, 2017, Geologic model building in SEAM Phase II-Land seismic challenges: Leading Edge, **36**, 738–749.

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

Rocha, D., P. Sava, J. Shragge, and B. Witten, 2019, 3D passive wavefield imaging using the energy norm: Geophysics, **84**, no. 2, KS13–KS27.

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

Sava, P., and I. Vasconcelos, 2011, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

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**, 952–966.

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

Shragge, J., T. Yang, and P. Sava, 2013, Time-lapse image-domain tomography using adjoint-state methods: Geophysics, **78**, no. 4, A29–A33.

Sun, J., Z. Xue, S. Fomel, T. Zhu, and N. Nakata, 2016, Full waveform inversion of passive seismic data for sources and velocities: 86th Annual International Meeting, SEG, Expanded Abstracts, **35**, 1405–1410.

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

Tang, Y., and G. Ayeni, 2015, Efficient line search methods for multiparameter full wave-field inversion: U.S. Patent 20, 150, 323, 689.

Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966.

Waldhauser, F., and W. L. Ellsworth, 2000, A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California: Bulletin of the Seismological Society of America, **90**, 1353–1368.

Wang, H., and T. Alkhalifah, 2018, Microseismic imaging using a source function independent full waveform inversion method: Geophysical Journal International, **214**, 46–57.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Štekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, no. 2, R59–R80.

Weingarten, M., S. Ge, J. W. Godt, B. A. Bekins, and J. L. Rubinstein, 2015, High-rate injection is associated with the increase in u.s. mid-continent seismicity: Science, **348**, 1336–1340.

Weiss, R. M., and J. Shragge, 2013, Solving 3D anisotropic elastic wave equations on parallel GPU devices: Geophysics, **78**, no. 2, F7–F15.

Witten, B., and J. Shragge, 2015, Extended wave-equation imaging conditions for passive seismic data: Geophysics, **80**, no. 6, WC61–WC72.

———, 2017a, Image-domain velocity inversion and event location for microseismic monitoring: Geophysics, **82**, no. 5, KS71–KS83.

———, 2017b, Microseismic image-domain velocity inversion: Marcellus Shale case study: Geophysics, **82**, no. 6, KS99–KS112.

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

Yang, T., and P. Sava, 2015, Image-domain wavefield tomography with extended common-image-point gathers: Geophysical Prospecting, **63**, 1086–1096.

Yang, T., J. Shragge, and P. Sava, 2013, Illumination compensation for image-domain wavefield tomography: Geophysics, **78**, no.

5, U65–U76.

Zhang, H., and C. H. Thurber, 2003, Double-difference tomography: The method and its application to the Hayward fault, California: Bulletin of the Seismological Society of America, **93**, 1875–1889.