

3D coupled acoustic-elastic migration with topography and bathymetry based on spectral-element and adjoint methods

Yang Luo¹, Jeroen Tromp², Bertrand Denel³, and Henri Calandra³

ABSTRACT

In the context of the adjoint method, we considered 3D coupled acoustic-elastic migration in the presence of surface topography and/or bathymetry. Isotropic elastic imaging involves three primary kernels, related to mass density and shear and bulk moduli, and various secondary kernels, for example, related to P-wave impedance and compressional and shear-wave speeds. Similar to reverse-time migration, these kernels reflect the constructive interference between a forward wavefield generated by active sources and an adjoint wavefield triggered by simultaneously back propagating recorded reflections from all receivers. Forward and adjoint wavefields were simulated using a spectral-element method, which, due to its weak nature, captures free-surface topography in land surveys and bathymetry in marine acquisition. To avoid storing the entire 3D forward wavefield, required for calculating its interaction with the adjoint wavefield, we only saved information on domain boundaries and reconstructed the forward wavefield while simulating the adjoint wavefield. Their interactions were calculated and integrated on the fly, thereby eliminating storage issues but doubling memory and CPU requirements. Our 3D images confirmed a previous conclusion based on 2D simulations, namely, that the impedance kernel best highlights reflectors, whereas wave-speed kernels constrain large-scale structures, i.e., the background model.

INTRODUCTION

Revealing geologic structures beneath the earth's surface, especially possible hydrocarbon reservoirs, is the main goal of explora-

tion seismology. Migration has played a central role in constructing subsurface images by projecting reflections recorded at the surface back to their origins. Claerbout (1971) articulates that reflectors exist when down- and upgoing wavefields coincide in space and time. Mathematically, the strength of a reflector may be approximated by crosscorrelations between the two wavefields, an insight that defines the classical "imaging principle" most commonly used in migration. Although the spatial distribution of reflectors may thus be imaged, their amplitudes may be biased (Zhang et al., 2003) due to the crude approximation of reflectivity and the effects of geometric spreading and attenuation.

Traditionally, Kirchhoff migration involves a high-frequency asymptotic solution to the wave equation, calculating traveltimes based on the eikonal equation (e.g., Červený, 2005). In terms of the imaging principle, down- and upgoing wavefields are treated as delta functions arriving at appropriate traveltimes. Therefore, their crosscorrelation has contributions only when their traveltimes are coincident. If required, amplitude corrections can be made by removing propagation effects through the transport equation. However, in complex media, these ray-based techniques break down due to multipathing effects. Some implementations, for example, those based on Gaussian beams (e.g., Wu, 1985; Hill, 1990, 2001), distribute wave energy around rays, and thus they allow certain kinds of finite-frequency effects to come into play. In contrast, reverse time migration (RTM) (McMechan, 1982, 1983; Baysal et al., 1983) incorporates the full wave equation and, hence, makes it possible to take into account complex wave propagation phenomena. In RTM, the down- and upgoing wavefields are usually termed "forward" (source-side) and "adjoint" (receiver-side) wavefields, and their interactions naturally conform with Claerbout's imaging principle.

Although the theory of RTM does not impose an acoustic approximation, traditional migration techniques have been associated with the acoustic-wave equation due to its simplicity. Even when considering anisotropic behavior, modifications to the acoustic-wave

Manuscript received by the Editor 31 October 2012; revised manuscript received 15 January 2013; published online 3 June 2013.

¹Princeton University, Department of Geosciences, Princeton, New Jersey, USA. E-mail: yangl.princeton@gmail.com.

²Princeton University, Department of Geosciences and Program in Applied & Computational Mathematics, Princeton, New Jersey, USA. E-mail: jtromp@princeton.edu.

³TOTAL Exploration & Production, Pau, France. E-mail: bertrand.denel@total.com; henri.calandra@total.com.

© 2013 Society of Exploration Geophysicists. All rights reserved.

equation are invoked to facilitate changes in wave speed along different directions. In a classical marine survey, seismic waves are generated in the water layer and reflections are recorded in the form of pressure fluctuations. In that case, using a purely acoustic-wave equation is a reasonable assumption because PP reflections contain a significant amount of information on subsurface structure beneath the ocean. However, with ocean-bottom cable (OBC) or land surveys, shear waves do play an important role. On one hand, incorporating elastic information in imaging may enhance the coherence of arrivals and thus may provide better images. On the other hand, recorded shear signals might contaminate the image if they are falsely interpreted as reflected compressional waves. In such a case, applying elastic wave-equation migration is essential. To account for elastic effects, several groups have developed modified imaging principles, using divergence and curl operators to separate compressional and shear motions (e.g., Xie and Wu, 2005; Yan and Sava, 2008).

Although various migration techniques have been consistent with Claerbout's generic definition, the imaging principle has also been related to the gradient of a waveform misfit functional in the context of an optimization problem, in which we seek to determine a model that minimizes a misfit function (Luo et al., 2009; Zhu et al., 2009). Migration can then be treated as a partial inversion or the initial step of an iterative inversion. It is widely known that the gradient of such a misfit function may be calculated via the adjoint method (Chavent, 1974; Lailly, 1983; Tarantola, 1984; Talagrand and Courtier, 1987; Tromp et al., 2005; Plessix, 2006), which has been applied to full waveform inversion in exploration seismology (see Virieux and Operto, 2009, for a review) as well as in regional tomography, e.g., southern California (Chen et al., 2007; Tape et al., 2009, 2010), Australia (Fichtner et al., 2009), and Europe (Zhu et al., 2012).

This paper is an extension of our previous 2D elastic imaging studies (Luo et al., 2009; Zhu et al., 2009), illustrating the importance of the “impedance kernel” in 3D cases. We use 3D elastic and coupled acoustic-elastic simulations to illustrate how the impedance kernel reveals clearer reflectors than other (modified) imaging principles. The spectral-element method (e.g., Peter et al., 2011) is used as the wave-equation solver, which, due to its weak formulation, is more suitable than the finite-difference method for accurate simulations involving significant topography and/or bathymetry and fluid-solid domain decomposition.

THEORY

As Nolet (1987) and Tromp et al. (2005) discuss, we may define a least-squares waveform misfit function to measure the quality of fit between observed \mathbf{d} and simulated \mathbf{s} seismograms as follows:

$$\chi(\mathbf{m}) = \frac{1}{2} \sum_r \int \|\mathbf{s}(\mathbf{x}_r, t; \mathbf{m}) - \mathbf{d}(\mathbf{x}_r, t)\|^2 dt, \quad (1)$$

where \mathbf{m} denotes the model and \mathbf{x}_r is the locations of receivers. Although there are other variations of this definition, in particular a discretized form in the frequency domain (Pratt et al., 1998), we prefer the above expression because of its flexibility in choosing arbitrary model basis functions and frequency bands. In principle, we may combine elastic signals (displacement/velocity/acceleration) with acoustic measurements (pressure), e.g., in a marine survey in which hydrophones and OBCs are available. The corresponding theory is slightly more complicated, requiring a more

careful implementation of fluid-solid interactions, as discussed in Appendix A. Connections between time- and frequency-domain inversions are explored in Appendix B, where we demonstrate how frequency- or Laplace-domain imaging may be accomplished based on a time-domain solver, thereby placing all approaches on the same footing.

In a generic elastic medium, variations in the misfit function $\delta\chi$ may be related to model perturbations $\delta\mathbf{m} = (\delta\rho, \delta\mathbf{c})$ via (Tromp et al., 2005)

$$\delta\chi = \int_{\Omega} (K_{\rho}\delta\rho + \mathbf{K}_{\mathbf{c}} :: \delta\mathbf{c}) d^3\mathbf{x}, \quad (2)$$

where ρ denotes the mass density, \mathbf{c} the fourth-order elastic tensor, and the “double-dot product” $::$ denotes successive contractions over four repeated indices. The primary sensitivity kernels K_{ρ} and $\mathbf{K}_{\mathbf{c}}$ are given by

$$K_{\rho}(\mathbf{x}) = \int \partial_t \mathbf{s}^{\dagger}(\mathbf{x}, -t) \cdot \partial_t \mathbf{s}(\mathbf{x}, t) dt, \quad (3)$$

$$\mathbf{K}_{\mathbf{c}}(\mathbf{x}) = - \int \nabla \mathbf{s}^{\dagger}(\mathbf{x}, -t) \nabla \mathbf{s}(\mathbf{x}, t) dt, \quad (4)$$

where the adjoint wavefield \mathbf{s}^{\dagger} is obtained by solving exactly the same wave equation as for the forward wavefield \mathbf{s} except using an adjoint source,

$$\mathbf{f}^{\dagger}(\mathbf{x}, t) = \sum_r [\mathbf{s}(\mathbf{x}_r, -t) - \mathbf{d}(\mathbf{x}_r, -t)] \delta(\mathbf{x} - \mathbf{x}_r). \quad (5)$$

This source involves back propagating time-reversed differences between observed and simulated seismograms simultaneously from all receivers on all components.

Equation 2 provides a straightforward inversion strategy to consider anisotropic problems with any prescribed symmetry. It may be reduced to simpler forms given certain symmetry considerations. For example, in an isotropic material, the elastic tensor is

$$c_{ijk\ell} = \left(\kappa - \frac{2}{3}\mu \right) \delta_{ij} \delta_{k\ell} + \mu (\delta_{ik} \delta_{j\ell} + \delta_{i\ell} \delta_{jk}), \quad (6)$$

where the bulk and shear moduli are denoted by κ and μ , respectively. Therefore, the misfit variation becomes

$$\delta\chi = \int_{\Omega} (K_{\rho}\delta\ln\rho + K_{\kappa}\delta\ln\kappa + K_{\mu}\delta\ln\mu) d^3\mathbf{x}, \quad (7)$$

where the corresponding primary kernels are just linear combinations of K_{ρ} and $\mathbf{K}_{\mathbf{c}}$, taking the forms,

$$K_{\rho}(\mathbf{x}) = \rho(\mathbf{x}) \int \partial_t \mathbf{s}^{\dagger}(\mathbf{x}, -t) \cdot \partial_t \mathbf{s}(\mathbf{x}, t) dt, \quad (8)$$

$$K_{\kappa}(\mathbf{x}) = -\kappa(\mathbf{x}) \int [\nabla \cdot \mathbf{s}^{\dagger}(\mathbf{x}, -t)] [\nabla \cdot \mathbf{s}(\mathbf{x}, t)] dt, \quad (9)$$

$$K_{\mu}(\mathbf{x}) = -2\mu(\mathbf{x}) \int \mathbf{D}^{\dagger}(\mathbf{x}, -t) : \mathbf{D}(\mathbf{x}, t) dt, \quad (10)$$

where $\mathbf{D} = (1/2)[\nabla \mathbf{s} + (\nabla \mathbf{s})^T] - (1/3)(\nabla \cdot \mathbf{s})\mathbf{I}$ and \mathbf{D}^\dagger denote the traceless strain deviator and its adjoint and $:$ is the contraction over two indices. Note that model parameters are expressed in terms of relative perturbations; e.g., $\delta \ln \rho = \delta \rho / \rho$. Also note that the adjoint wavefield is in reverse time, as indicated by the argument “ $-t$.” We can easily see that the density kernel K_ρ' is similar to Claerbout’s imaging principle, whereas the bulk and shear kernels K_κ and K_μ resemble modified imaging principles introduced in previous studies (Xie and Wu, 2005; Yan and Sava, 2008).

Equivalently, and physically more realistically, we may describe the model in terms of mass density ρ and compressional- and shear-wave speeds α and β , instead of mass density and moduli. For this model parameterization, the misfit perturbation becomes

$$\delta\chi = \int_{\Omega} (K_\rho' \delta \ln \rho + K_\alpha \delta \ln \alpha + K_\beta \delta \ln \beta) d^3x, \quad (11)$$

where K_ρ' , K_α , and K_β denote an alternative density kernel, a compressional wave-speed kernel, and a shear wave-speed kernel, respectively. These secondary kernels are linear combinations of the three primary kernels, namely,

$$K_\rho' = K_\rho + K_\kappa + K_\mu, \quad (12)$$

$$K_\alpha = 2 \left(1 + \frac{4\mu}{3\kappa} \right) K_\kappa, \quad (13)$$

$$K_\beta = 2 \left(K_\mu - \frac{4\mu}{3\kappa} K_\kappa \right). \quad (14)$$

Other model parameterizations such as $(Z_\alpha \equiv \rho\alpha, \alpha, \beta)$ or $(Z_\beta \equiv \rho\beta, \alpha, \beta)$ may also be used. The importance of the alternative density kernel K_ρ' lies in the fact that (Luo et al., 2009; Zhu et al., 2009)

$$K_\rho' = K_{Z_\alpha} = K_{Z_\beta}. \quad (15)$$

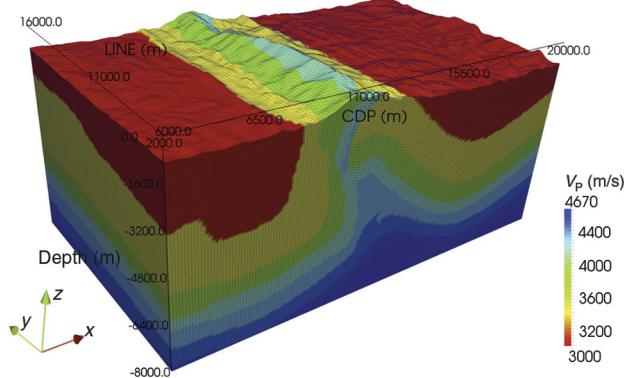


Figure 1. Compressional wave-speed model, with elements denoted by thin blue lines. Along the CDP direction (x -direction), the model ranges from 2 to 19 km. Along the LINE direction (y -direction), the model ranges from 6 to 16 km. The model is cut at a depth of 8 km, which leads to 7 km in height, taking into consideration around 1 km of topography on the free surface.

In other words, the alternative density kernel is the same as the compressional and shear impedance kernels. Therefore, from here on, we refer to K_ρ' as the impedance kernel. Mathematically, the impedance kernel determines how the misfit function changes due to density/impedance variations, holding the wave speeds fixed. This interpretation is consistent with the physical problem we need to solve, i.e., finding the locations of reflectors in a given background wave-speed model.

With regard to the numerical implementation of these expressions, we first carry out the forward simulations, which provide the synthetic seismograms we need to construct the adjoint sources, in this case, the time-reversed difference between observed and simulated seismograms. Adjoint wavefields are obtained by injecting the adjoint sources simultaneously at all receiver locations. A numerical difficulty in evaluating the sensitivity kernels is to access the forward and adjoint wavefields simultaneously for their convolution in the construction of the kernels. A straightforward but brute-force and costly approach is to save the entire forward wavefield as a function of space and time, reading it back during the construction of the kernels. Because of oversampling of seismograms, one may not need to save the forward wavefield at every time step, using interpolation to fill in the gaps (see, e.g., Symes,

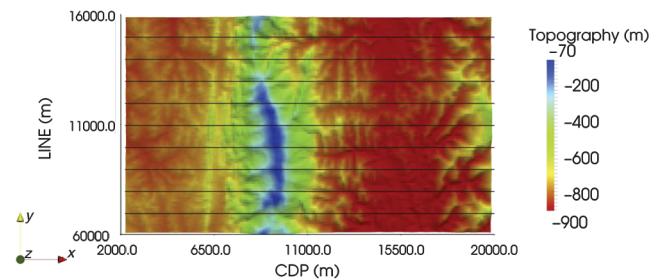


Figure 2. Map view of topography and the acquisition geometry. The topography is color coded, whereas the nine lines of sources used in the study are denoted by black lines. Geophones are assumed to be located everywhere on the surface with 40-m spacing (not plotted).

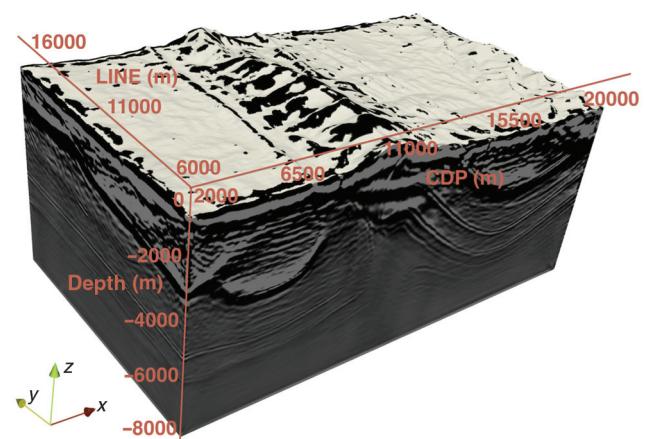


Figure 3. Image of the density kernel K_ρ , which corresponds to the classical imaging condition. Note the large-scale imaging artifacts.

2007). Our preferred approach is to save the last snapshot of the forward simulation and time-dependent information on the boundaries of the model domain, using this to reconstruct the forward wavefield while the adjoint wavefield is being simulated (Liu and Tromp, 2006). This approach does not require a large storage system, but it doubles the memory and CPU requirements.

NUMERICAL EXAMPLES

In the following three sections, we illustrate the nature of the primary and secondary kernels in elastic, acoustic, and coupled acoustic-elastic imaging. The open-source spectral-element software package SPECFEM3D is available for download via the Computational Infrastructure for Geodynamics (www.geodynamics.org).

Elastic imaging

We perform a 3D elastic imaging experiment, investigating which of the six primary and secondary kernels, that is, K_ρ , K_κ , K_μ , K'_ρ , K_α , and K_β , leads to the best subsurface image. The model consists of a $17 \times 10 \times 7$ km block with topographic variations of approximately 1 km on the free surface. As illustrated in Figure 1, the model contains layered structures with an overthrust fault in the middle. The compressional wave speed varies between 3 and 4.5 km s⁻¹, and the shear-wave speed is linearly scaled to the compressional wave speed by a factor of $1/\sqrt{3}$. The density model, a similar layered structure, is obtained from amplitude variation with offset analysis. To accommodate the free surface, we design a 1.2-million-element mesh that honors the topography, as illustrated by the dense grid in Figure 1. We neglect interfaces beneath the surface; i.e., these boundaries may cross spectral elements. However,

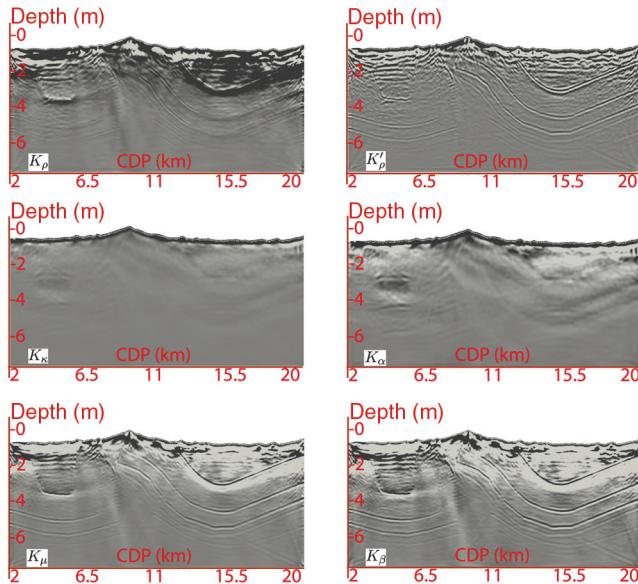


Figure 4. Comparison of six different kernels at a cross section in the center of the model. On the left are the three primary kernels K_ρ , K_κ , and K_μ , which suffer from large-scale artifacts due to diving waves. On the right are the impedance kernel K'_ρ and the two wave-speed kernels K_α and K_β . Although the wave-speed kernels behave similarly to the moduli kernels, the impedance kernel erases large-scale information, which is useless in migration, thereby highlighting the reflectors. From this perspective, the impedance kernel is the best choice for migration purposes.

note that we use 125 Gauss-Legendre-Lobatto integration points with variable model parameters per element, and that the spectral-element method behaves as any method based on a strong formulation, such as the finite-difference method, if a discontinuity crosses an element boundary. Tests show that honoring or not honoring subsurface interfaces does not affect reflections from these boundaries but may have a strong influence on the dispersion of surface waves. Because surface waves are irrelevant in migration, not honoring internal interfaces does not harm image quality, and it dramatically reduces the difficulty of mesh generation.

Figure 2 presents a map view of the model, outlining the acquisition geometry. Nine source lines are used, indicated by thick black lines, each involving approximately 200 shots (with an 80-m interval), whereas geophones are deployed everywhere on the surface, with a 40-m interval. Sources and receivers are 5 m below the free surface; i.e., they follow the topography.

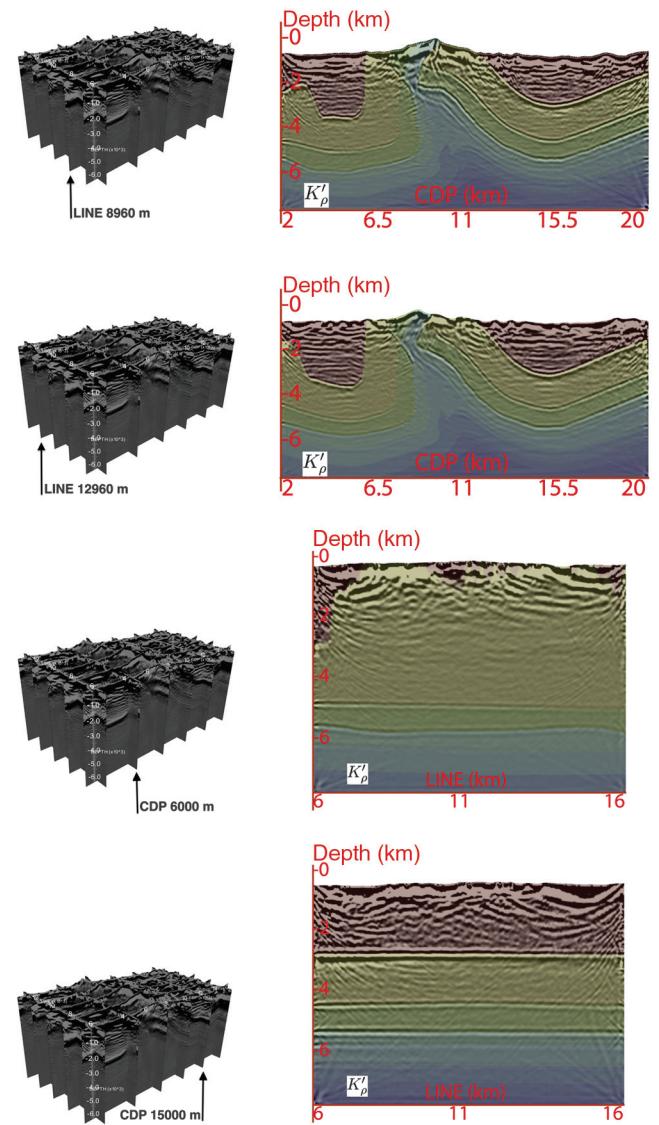


Figure 5. Images of the impedance kernel K'_ρ (equation 12) along several vertical profiles. Shown on the left are the locations of these profiles, and shown on the right are the impedance kernels in black and white plotted on top of the wave-speed model in color.

Figure 3 shows a migrated image based on the classical imaging principle, that is, our density kernel K_ρ . Surface waves are muted by a simple $f\text{-}k$ filter before migration. Although reflectors are revealed properly, long-wavelength artifacts exist, especially at shallow depths. Effectively, these large-scale artifacts come from the summation of “banana-doughnut” finite-frequency sensitivity kernels and contaminate images of reflectors. Although the source lines are along the CDP direction, reflectors in inline cross sections are imaged as well, thanks to the dense receiver geometry.

We compare the six different kernels at cross section $y = 10 \text{ km}$ (y -axis is for fixed CDP, as indicated by the notation in the bottom left corner of Figure 1), as illustrated in Figure 4. All three primary kernels K_ρ , K_κ , and K_μ behave similarly, showing reflectors superimposed on large-scale artifacts. It is noteworthy that the sign of these large-scale artifacts is opposite for density and the moduli. Because the wavelengths of the compressional waves are longer than the shear waves in the same frequency band, the bulk-modulus kernel K_κ gives a more blurred image than the density and shear-modulus kernels. For the impedance and wave-speed kernels, the images of K_α and K_β look similar to the corresponding moduli kernels, K_κ and K_μ , as expected, whereas the impedance kernel, K'_ρ , stands out by eliminating large-scale artifacts and highlighting only reflectors. This makes sense mathematically because the large-scale

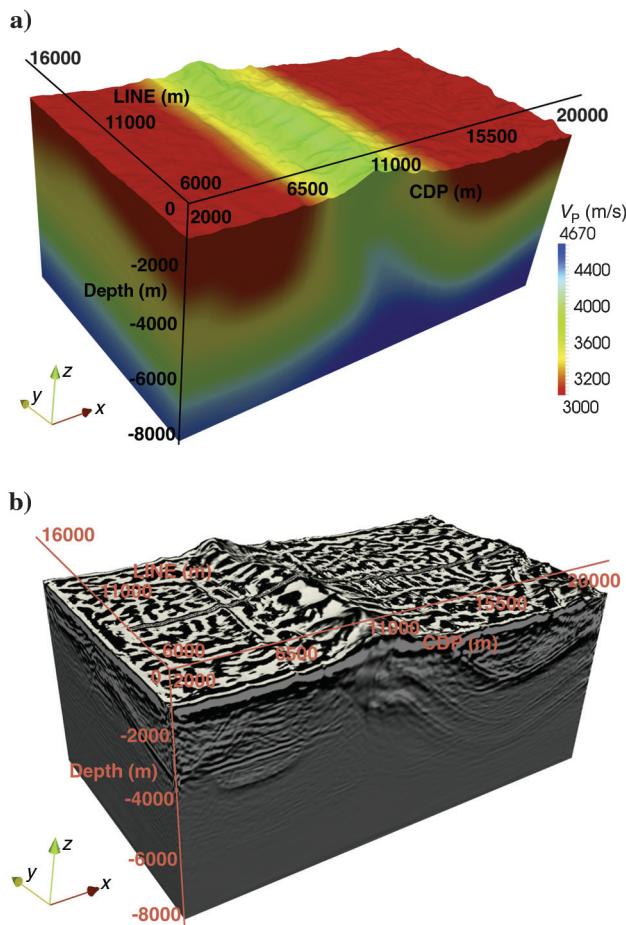


Figure 6. (a) Smoothed wave-speed background model and (b) the corresponding density kernel. The main reflectors are still clearly delineated despite the shortcomings of the background model.

artifacts are canceled due to the sum of the three primary kernels in equation 12. Physically, it means that the combination of density and wave speeds is a better model parameterization than the combination of density and moduli. Figure 5 shows images of the impedance kernel at other cross sections, plotted on top of the wave-speed model, to validate the locations of reflectors.

Thus far, we use the exact wave-speed model for migration in the construction of the images. To investigate effects of the background wave-speed model, we carry out migration in a smoothed model, obtained using a 3D Gaussian filter with a 500-m standard deviation. Figure 6 illustrates the smoothed wave-speed model and the

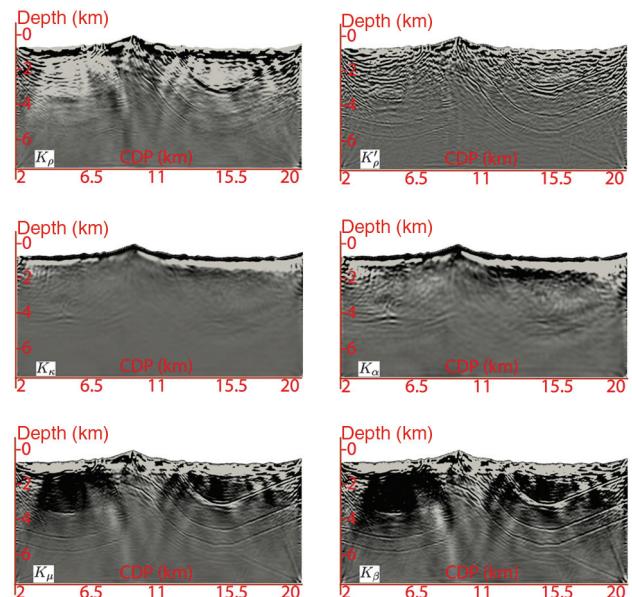


Figure 7. Similar to Figure 4, except that these kernels are obtained in a smoothed background model. This model introduces some artifacts into the images when compared with the results based on the exact background model shown in Figure 4. However, the basic characteristics of these kernels remain the same; i.e., the impedance kernel still provides the best imaging condition.

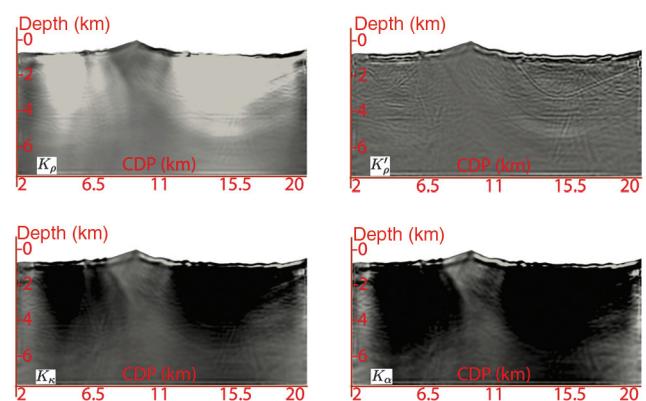


Figure 8. Four kernels from a purely acoustic migration, using vertical elastic displacements as pressures. As expected, the impedance kernel again stands out by erasing large-scale artifacts. By contrast to the impedance kernels shown in Figures 4 and 7, the acoustic impedance image has fewer near-surface artifacts, but the reflectors are not as crisp.

resulting density kernel. Compared to Figure 3, there are some artifacts due to the inaccurate wave-speed model, but overall, the two images are very consistent.

The six kernels along a cross section right beneath the source line are shown in Figure 7, which should be contrasted to Figure 4. Although the wave-speed model differs from the true model, the impedance kernel still highlights the reflectors properly.

Acoustic imaging

For the acquisition geometry shown in Figure 2, we also conduct a purely acoustic migration, treating vertical displacements in the elastic model as recorded pressure fluctuations. The two primary acoustic kernels K_ρ and K_κ and the two secondary acoustic kernels

K'_ρ and K_α shown in Figure 8 should be contrasted with the elastic kernels shown in Figures 4 and 7. Because no surface waves are involved in the acoustic simulations, noise in the images near the surface is greatly reduced. However, reflectors are more blurred, whereas some artificial reflectors emerge because of improper mapping of shear arrivals in the elastic data sets.

Coupled acoustic-elastic imaging

The flexibility of the spectral-element method allows us to deal with coupled acoustic-elastic simulations based on domain decomposition, honoring the free-slip fluid-solid boundary condition. To illustrate this point, we add a water layer on top of the elastic model and carry out a marine survey. The mesh for the elastic domain remains the same, whereas the mesh for the fluid domain is obtained by sweeping along the LINE direction, as shown in Figure 9. In Figure 9, a particular shot, denoted by the spike, is recorded by several lines of hydrophones, represented by spheres, which are moving with the shot. Prior to back projection, multiples are removed from the data set by subtracting a forward simulation with just the water layer on top of a layer with the actual bathymetry and constant material properties. Figure 10 compares the resulting six kernels in the elastic domain, which contain artificial reflectors when compared with the results of the land survey shown in Figure 4. These artifacts are due to converted waves at the ocean bottom, e.g., the pressure field recorded by hydrophones may be due to upgoing PP or SP conversions at the seafloor. Both conversions are contained in the images — one of which (PP) results in the correct reflector, whereas the other one (SP) creates artifacts.

Figure 11 shows four acoustic kernels using purely acoustic migration. Upon comparing the impedance kernels in Figures 10 and 11, we conclude that using the acoustic approximation in a marine survey reduces the quality of the image due to mapping shear signals into artificial reflectors. Such effects can be easily understood by comparing data sets from a coupled acoustic-elastic simulation with a purely acoustic simulation, which contains numerous converted shear waves, as shown in Figure 12.

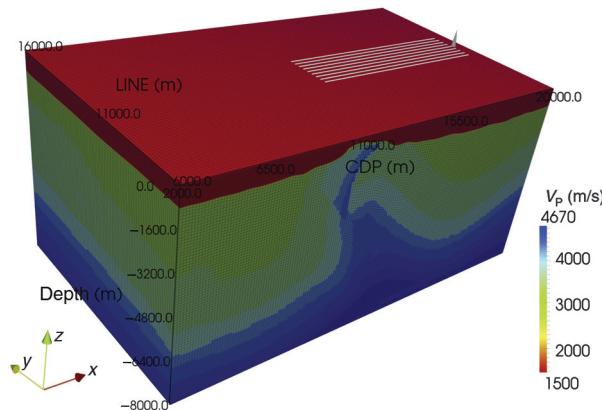


Figure 9. Illustration involving marine acquisition. A water layer (red) is added on top of the original elastic domain in Figure 1. Note that the color bar has changed. The mesh for the elastic domain is the same as that in Figure 1, whereas the mesh for the acoustic domain is created by sweeping along the y -axis. The spike denotes the shot, whereas the spheres denote hydrophones.

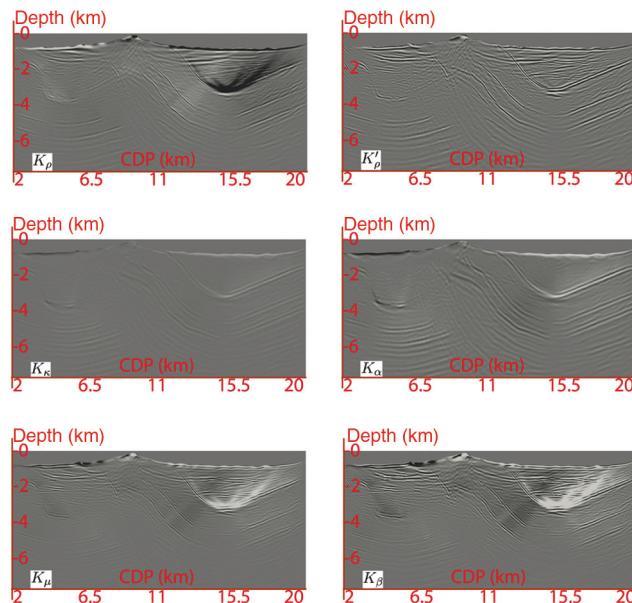


Figure 10. Six elastic kernels obtained for the marine survey with irregular bathymetry, for a configuration shown in Figure 9.

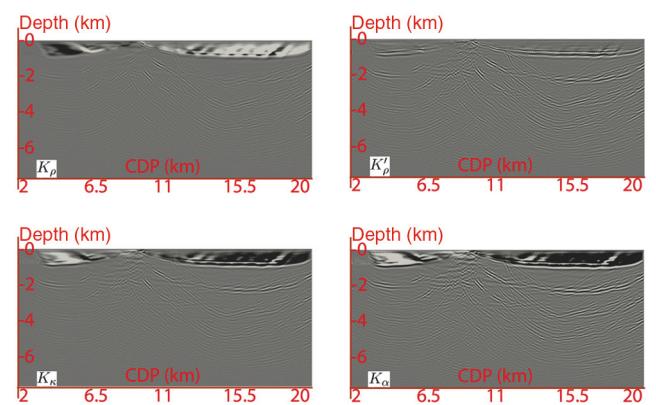


Figure 11. Four acoustic kernels obtained for the marine survey with irregular bathymetry using the acoustic approximation. Artificial reflectors are introduced as a result of the improper interpretation of shear motions in the coupled elastic-acoustic data set as compressional motions. Differential seismograms for one shot are shown in Figure 12.

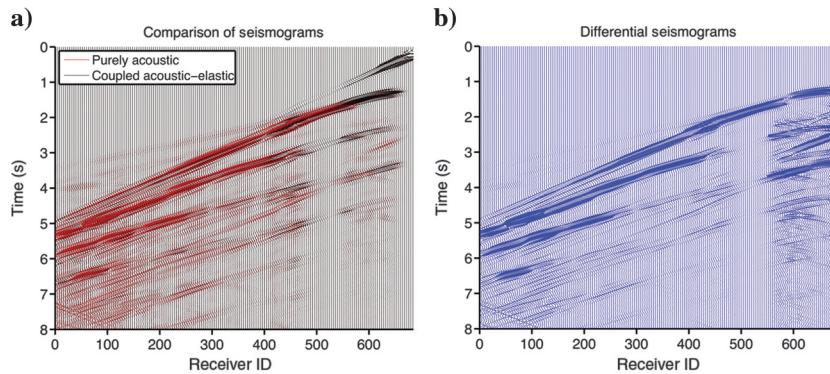


Figure 12. (a) Comparison between seismograms from purely acoustic (red) and coupled acoustic-elastic (black) simulations for a single shot. (b) Their differences, which are due to the lack of shear motions in the acoustic approximation. These differences are responsible for the reduction in image quality between Figures 10 and 11. Note that differential seismograms are normalized; i.e., the scale is not the same as that on the left.

CONCLUSIONS

We have extended our previous impedance imaging work from 2D to 3D, presenting impedance kernels for land and marine surveys with topography and/or bathymetry. Based on our experiments and experience in global and regional seismology, we conclude that a combination of density and wave speeds forms the most suitable parameterization of a seismic model. Although wave-speed kernels provide constraints on large-scale structures, useful in an iterative inversion, the impedance kernel identifies fine-scale features related to reflectors. The kernel expressions are a natural extension of Claerbout's classical imaging principle, and they smoothly connect migration with seismic inversions. Inaccurate wave-speed models may introduce noise in migrated images, but the main reflectors are still correctly mapped. To obtain an image with higher quality, shear- and compressional-wave-speed models need to be improved, e.g., by iterative inversions. Artificial reflectors related to shear arrivals may arise when using the acoustic wave equation to migrate elastic data sets, illustrating the advantages of applying elastic migration techniques.

ACKNOWLEDGMENTS

We thank J. Claerbout for encouraging us to validate self-adjointness of our forward and adjoint solvers, which led to a reformulation of the coupled elastic-acoustic simulations when there are hydrophones in the acoustic region, as discussed in Appendix A. All numerical simulations were performed on the "Sesame" Dell cluster built and maintained by the Princeton Institute for Computational Science and Engineering. This research was partly sponsored by TOTAL and by the United States National Science Foundation under grants EAR-1112906 and DMS-1025418.

APPENDIX A

SENSITIVITY KERNELS IN ACOUSTIC MEDIA

As a result of a discussion with J. Claerbout regarding self-adjointness of coupled acoustic-elastic simulations, herein we carefully derive the adjoint equations required for coupled acoustic-

elastic imaging, thereby correcting and extending results that Peter et al. (2011) obtain. The elastic wave equation is

$$\rho \partial_t^2 \mathbf{s} = \nabla \cdot \mathbf{T} + \mathbf{f}, \quad (\text{A-1})$$

where the stress tensor \mathbf{T} is related to the displacement gradient $\nabla \mathbf{s}$ via Hooke's law

$$\mathbf{T} = \mathbf{c} : \nabla \mathbf{s}. \quad (\text{A-2})$$

We may obtain the acoustic wave equation,

$$\frac{1}{\kappa} \partial_t^2 \phi = \nabla \cdot \left(\frac{1}{\rho} \nabla \phi \right) + f, \quad (\text{A-3})$$

by introducing a potential ϕ ; namely,

$$\mathbf{s} = \frac{1}{\rho} \nabla \phi. \quad (\text{A-4})$$

Consequently, the fluid pressure may be related to the potential via

$$p = -\kappa(\nabla \cdot \mathbf{s}) \approx -\partial_t^2 \phi. \quad (\text{A-5})$$

Note that the source term f is ignored in equation A-5. This is generally a valid assumption, except when a volumetric source, such as gravity, is involved, in which case, the source term is critical. For our purposes in this paper, the source term is irrelevant.

Previously, sensitivity kernels in the acoustic domain have been calculated using the same expressions as in the elastic domain (Peter et al., 2011), assuming the adjoint displacement \mathbf{s}^\dagger and adjoint potential ϕ^\dagger may be expressed in terms of similar relationships as equations A-4 and A-5; namely,

$$\mathbf{s}^\dagger = \frac{1}{\rho} \nabla \phi^\dagger, \quad (\text{A-6})$$

$$p^\dagger = -\kappa(\nabla \cdot \mathbf{s}^\dagger) \approx -\partial_t^2 \phi^\dagger. \quad (\text{A-7})$$

These equations are actually generally invalid. This can be understood by noting that equations A-6 and A-7 are derived from the adjoint elastic wave equation. In other words, the adjoint potential definition and the related adjoint acoustic-wave equation are not a consequence of the correct acoustic optimization problem, as we show below. When there are no receivers in the acoustic domain (that is, in the absence of adjoint sources), the adjoint potential definition (equation A-6) is suitable. Thus, on a global scale, when there are no receivers in the liquid core or oceans, equations A-6 and A-7 are valid. However, when the misfit function includes measurements made in fluids, i.e., hydrophone pressure records, the adjoint potential definition — and hence the adjoint acoustic wave equation — are determined uniquely by the acoustic optimization problem, rather than by reduction of the adjoint elastic wave equation. This appendix presents the correct relationships for the adjoint displacement and potential, following the Lagrange multiplier method introduced by Liu and Tromp (2006).

To deal with an acoustic domain properly, we define a waveform misfit function that assimilates acoustic and elastic measurements:

$$\chi = \frac{1}{2} \sum_{r \in \text{el}} \int \| \mathbf{s}(\mathbf{x}_r, t) - \mathbf{d}(\mathbf{x}_r, t) \|^2 dt + \frac{1}{2} \sum_{r \in \text{ac}} \int [p^{\text{syn}}(\mathbf{x}_r, t) - p^{\text{obs}}(\mathbf{x}_r, t)]^2 dt, \quad (\text{A-8})$$

where \mathbf{s} and \mathbf{d} are synthetic and observed displacements, and p^{syn} and p^{obs} are synthetic and observed pressures. The partial-differential-equation-constrained optimization problem is

$$\begin{aligned} \chi = & \frac{1}{2} \sum_{r \in \text{el}} \int \| \mathbf{s}(\mathbf{x}_r, t) - \mathbf{d}(\mathbf{x}_r, t) \|^2 dt \\ & - \int \int_{\Omega_{\text{el}}} \lambda \cdot (\rho \partial_t^2 \mathbf{s} - \nabla \cdot \mathbf{T} - \mathbf{f}) d^3 \mathbf{x} dt \\ & + \frac{1}{2} \sum_{r \in \text{ac}} \int [p^{\text{syn}}(\mathbf{x}_r, t) - p^{\text{obs}}(\mathbf{x}_r, t)]^2 dt \\ & - \int \int_{\Omega_{\text{ac}}} \mu \left[\frac{1}{\kappa} \partial_t^2 \phi - \nabla \cdot \left(\frac{1}{\rho} \nabla \phi \right) - f \right] d^3 \mathbf{x} dt, \end{aligned} \quad (\text{A-9})$$

where λ and μ are Lagrange multipliers for the elastic and acoustic wave equations, respectively. Taking the variation of equation A-9, we find

$$\begin{aligned} \delta \chi = & \int \int_{\Omega_{\text{el}}} \delta \mathbf{s} \cdot \sum_{r \in \text{el}} [(\mathbf{s}(\mathbf{x}_r, t) - \mathbf{d}(\mathbf{x}_r, t)) \delta(\mathbf{x} - \mathbf{x}_r)] d^3 \mathbf{x} dt \\ & - \int \int_{\Omega_{\text{el}}} \lambda \cdot [(\delta \rho \partial_t^2 \mathbf{s} + \rho \partial_t^2 \delta \mathbf{s}) - \nabla \cdot (\delta \mathbf{c} : \nabla \mathbf{s}) - \nabla \cdot (\mathbf{c} : \nabla \delta \mathbf{s}) - \delta \mathbf{f}] d^3 \mathbf{x} dt \\ & + \int \int_{\Omega_{\text{ac}}} \delta p \left\{ \sum_{r \in \text{ac}} [p^{\text{syn}}(\mathbf{x}_r, t) - p^{\text{obs}}(\mathbf{x}_r, t)] \delta(\mathbf{x} - \mathbf{x}_r) \right\} d^3 \mathbf{x} dt \\ & - \int \int_{\Omega_{\text{ac}}} \mu \left\{ \delta \left(\frac{1}{\kappa} \right) \partial_t^2 \phi + \frac{1}{\kappa} \partial_t^2 \delta \phi - \nabla \cdot \left[\delta \left(\frac{1}{\rho} \right) \nabla \phi \right] - \nabla \cdot \left(\frac{1}{\rho} \nabla \delta \phi \right) - \delta f \right\} d^3 \mathbf{x} dt. \end{aligned} \quad (\text{A-10})$$

Using integration by parts, we obtain, after some tedious algebra, the following expression:

$$\begin{aligned} \delta \chi = & - \int \int_{\Omega_{\text{el}}} \delta \mathbf{s} \cdot \left\{ \rho \partial_t^2 \lambda - \nabla \cdot (\mathbf{c} : \nabla \lambda) - \sum_{r \in \text{el}} [(\mathbf{s}(\mathbf{x}_r, t) - \mathbf{d}(\mathbf{x}_r, t)) \delta(\mathbf{x} - \mathbf{x}_r)] \right\} d^3 \mathbf{x} dt \\ & - \int \int_{\Omega_{\text{ac}}} \delta \phi \left\{ \frac{1}{\kappa} \partial_t^2 \mu - \nabla \cdot \left(\frac{1}{\rho} \nabla \mu \right) + \sum_{r \in \text{ac}} [(\partial_t^2 p^{\text{syn}}(\mathbf{x}_r, t) - \partial_t^2 p^{\text{obs}}(\mathbf{x}_r, t)) \delta(\mathbf{x} - \mathbf{x}_r)] \right\} \\ & \times d^3 \mathbf{x} dt - \int \int_{\Omega_{\text{el}}} \delta \rho \lambda \cdot \partial_t^2 \mathbf{s} d^3 \mathbf{x} dt - \int \int_{\Omega_{\text{el}}} (\nabla \lambda) : \delta \mathbf{c} : (\nabla \mathbf{s}) d^3 \mathbf{x} dt + \int \int_{\Omega_{\text{el}}} \lambda \cdot \delta \mathbf{f} d^3 \mathbf{x} dt \\ & - \int \int_{\Omega_{\text{ac}}} \delta \left(\frac{1}{\kappa} \right) \mu \partial_t^2 \phi d^3 \mathbf{x} dt - \int \int_{\Omega_{\text{ac}}} \delta \left(\frac{1}{\rho} \right) \nabla \mu \cdot \nabla \phi d^3 \mathbf{x} dt + \int \int_{\Omega_{\text{ac}}} \mu \delta \mathbf{f} d^3 \mathbf{x} dt \\ & - \int_{\Omega_{\text{el}}} \rho \lambda \cdot \partial_t \delta \mathbf{s}^3 \mathbf{x} \Big|_{t=0}^{t=T} + \int_{\Omega_{\text{el}}} \rho \partial_t \lambda \cdot \delta \mathbf{s}^3 \mathbf{x} \Big|_{t=0}^{t=T} - \int_{\Omega_{\text{ac}}} \frac{1}{\kappa} \mu \partial_t \delta \phi d^3 \mathbf{x} \Big|_{t=0}^{t=T} \\ & + \int_{\Omega_{\text{ac}}} \frac{1}{\kappa} \partial_t \mu \delta \phi d^3 \mathbf{x} \Big|_{t=0}^{t=T} + \int \int_{\partial \Omega_{\text{el}}} [\lambda \cdot (\hat{\mathbf{n}} \cdot \delta \mathbf{c} : \nabla \mathbf{s} + \hat{\mathbf{n}} \cdot \mathbf{c} : \nabla \delta \mathbf{s}) - \hat{\mathbf{n}} \cdot (\mathbf{c} : \nabla \lambda) \cdot \delta \mathbf{s}] d^2 \mathbf{x} dt \\ & + \int \int_{\partial \Omega_{\text{ac}}} \left\{ \mu \left[\hat{\mathbf{n}} \cdot \delta \left(\frac{1}{\rho} \right) \nabla \phi + \hat{\mathbf{n}} \cdot \left(\frac{1}{\rho} \nabla \delta \phi \right) \right] - \hat{\mathbf{n}} \cdot \left(\frac{1}{\rho} \nabla \mu \right) \delta \phi \right\} d^2 \mathbf{x} dt \\ & - \int \int_{\Sigma} \left\{ \delta \phi \left[\partial_t^2 \lambda + \frac{1}{\rho} \nabla \mu \right] \cdot \hat{\mathbf{n}} - \delta \mathbf{s} \cdot [\hat{\mathbf{n}} \cdot (\mathbf{c} : \nabla \lambda) + \mu \hat{\mathbf{n}}] \right\} d^2 \mathbf{x} dt, \end{aligned} \quad (\text{A-11})$$

where $\partial \Omega_{\text{el}}$ and $\partial \Omega_{\text{ac}}$ denote free surfaces of elastic and acoustic domains, respectively, Σ represents solid-fluid interfaces, and $\hat{\mathbf{n}}$ is the unit outward normal on boundaries.

When there are no perturbations in the model parameters, the variation in equation A-11 should be stationary with respect to $\delta \mathbf{s}$ and $\delta \phi$. Therefore, we obtain the following equations for the Lagrange multipliers λ and μ :

$$\rho \partial_t^2 \lambda = \nabla \cdot (\mathbf{c} : \nabla \lambda) + \sum_{r \in \text{el}} [\mathbf{s}(\mathbf{x}_r, t) - \mathbf{d}(\mathbf{x}_r, t)] \delta(\mathbf{x} - \mathbf{x}_r), \quad (\text{A-12})$$

$$\begin{aligned} \frac{1}{\kappa} \partial_t^2 \mu = & \nabla \cdot \left(\frac{1}{\rho} \nabla \mu \right) - \sum_{r \in \text{ac}} [\partial_t^2 p^{\text{syn}}(\mathbf{x}_r, t) \\ & - \partial_t^2 p^{\text{obs}}(\mathbf{x}_r, t)] \delta(\mathbf{x} - \mathbf{x}_r). \end{aligned} \quad (\text{A-13})$$

The ending conditions are

$$\lambda(\mathbf{x}, T) = \mathbf{0}, \quad \partial_t \lambda(\mathbf{x}, T) = \mathbf{0}, \quad (\text{A-14})$$

$$\mu(\mathbf{x}, T) = 0, \quad \partial_t \mu(\mathbf{x}, T) = 0, \quad (\text{A-15})$$

and the free-surface boundary conditions are

$$\hat{\mathbf{n}} \cdot (\mathbf{c} : \nabla \lambda) = \mathbf{0} \quad (\mathbf{x} \in \partial \Omega_{\text{el}}), \quad (\text{A-16})$$

$$\mu = 0 \quad (\mathbf{x} \in \partial \Omega_{\text{ac}}). \quad (\text{A-17})$$

Coupling between elastic domain and acoustic domains involves the boundary conditions,

$$\partial_t^2 \lambda \cdot \hat{\mathbf{n}} = - \frac{1}{\rho} \nabla \mu \cdot \hat{\mathbf{n}} \quad (\mathbf{x} \in \Sigma), \quad (\text{A-18})$$

$$\hat{\mathbf{n}} \cdot (\mathbf{c} : \nabla \lambda) = -\mu \hat{\mathbf{n}} \quad (\mathbf{x} \in \Sigma). \quad (\text{A-19})$$

To change the ending conditions into initial conditions, we define the adjoint displacement and adjoint potential as

$$\mathbf{s}^\dagger(\mathbf{x}, t) \equiv \lambda(\mathbf{x}, T-t), \quad (\text{A-20})$$

$$\phi^\dagger(\mathbf{x}, t) \equiv \mu(\mathbf{x}, T-t). \quad (\text{A-21})$$

The fluid-solid boundary conditions (equations A-18 and A-19) then become

$$\partial_t^2 \mathbf{s}^\dagger \cdot \hat{\mathbf{n}} = - \frac{1}{\rho} \nabla \phi^\dagger \cdot \hat{\mathbf{n}}, \quad (\text{A-22})$$

$$\hat{\mathbf{n}} \cdot (\mathbf{c} : \nabla \mathbf{s}^\dagger) = -\phi^\dagger \hat{\mathbf{n}}. \quad (\text{A-23})$$

We conclude that the relationship between \mathbf{s}^\dagger and ϕ^\dagger is quite different from that between \mathbf{s} and ϕ expressed by equation A-4. The adjoint pressure in fluids is related to the adjoint potential ϕ^\dagger , not $\partial_t^2 \phi^\dagger$, as indicated by equation A-23. In terms of numerical implementation, equations A-22 and A-23 facilitate an exchange of information between solids and fluids during adjoint simulations, whereas equations A-4 and A-5 accomplish this exchange for forward simulations.

Provided the Lagrange multipliers are determined by the adjoint equations A-12–A-19, the variation of the misfit function (equation A-11) reduces to

$$\delta\chi = \int_{\Omega_{\text{el}}} (\delta\rho K_{\rho}^{\text{el}} + \delta\mathbf{c} : \mathbf{K}_{\mathbf{c}}^{\text{el}}) d^3\mathbf{x} + \int_{\Omega_{\text{ac}}} \left[\delta\left(\frac{1}{\kappa}\right) K_{\kappa}^{\text{ac}} + \delta\left(\frac{1}{\rho}\right) K_{\rho}^{\text{ac}} \right] d^3\mathbf{x}, \quad (\text{A-24})$$

where we have ignored source perturbations $\delta\mathbf{f}$ and δf . The sensitivity kernels are

$$K_{\rho}^{\text{el}}(\mathbf{x}) = - \int \mathbf{s}^{\dagger}(\mathbf{x}, T-t) \cdot \partial_t^2 \mathbf{s}(\mathbf{x}, t) dt, \quad (\text{A-25})$$

$$\mathbf{K}_{\mathbf{c}}^{\text{el}}(\mathbf{x}) = - \int \nabla \mathbf{s}^{\dagger}(\mathbf{x}, T-t) \nabla \mathbf{s}(\mathbf{x}, t) dt, \quad (\text{A-26})$$

$$K_{\kappa}^{\text{ac}}(\mathbf{x}) = - \int \phi^{\dagger}(\mathbf{x}, T-t) \partial_t^2 \phi(\mathbf{x}, t) dt, \quad (\text{A-27})$$

$$K_{\rho}^{\text{ac}}(\mathbf{x}) = - \int \nabla \phi^{\dagger}(\mathbf{x}, T-t) \cdot \nabla \phi(\mathbf{x}, t) dt. \quad (\text{A-28})$$

Given the relationships in equations A-4 and A-22, it is noteworthy that

$$\delta\rho K_{\rho}^{\text{el}} = \delta\left(\frac{1}{\rho}\right) K_{\rho}^{\text{ac}}. \quad (\text{A-29})$$

Consistency between acoustic and elastic bulk modulus kernels is also straightforward to demonstrate: The elastic sensitivity kernels indeed reduce to the corresponding acoustic sensitivity kernels, using the correct definition of the adjoint potential.

APPENDIX B

TIME-DOMAIN VERSUS FREQUENCY-DOMAIN INVERSION

The inverse problem may be formulated either in the time (Tromp et al., 2005) or frequency domain (Pratt et al., 1998), depending on the kind of misfit function that is used. In this appendix, we connect frequency-, Laplace-, and time-domain inversions by defining an equivalent time-domain form of frequency- or Laplace-domain inversion.

Frequency-domain misfit function

The waveform misfit function may also be defined in the frequency domain; e.g., for a single frequency ω_k ,

$$\chi = \frac{1}{2} \sum_k \sum_r \|\tilde{\mathbf{s}}(\mathbf{x}_r, \omega_k) - \tilde{\mathbf{d}}(\mathbf{x}_r, \omega_k)\|^2, \quad (\text{B-2})$$

where $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{d}}$ are Fourier transforms of \mathbf{s} and \mathbf{d} , respectively:

$$\tilde{\mathbf{s}}(\mathbf{x}, \omega) = \int \mathbf{s}(\mathbf{x}, t) \exp(-i\omega t) dt, \quad (\text{B-3})$$

$$\tilde{\mathbf{d}}(\mathbf{x}, \omega) = \int \mathbf{d}(\mathbf{x}, t) \exp(-i\omega t) dt. \quad (\text{B-4})$$

Consequently, the corresponding adjoint source is

$$\begin{aligned} \mathbf{f}(\mathbf{x}, t) &= \text{Re} \sum_k \sum_r [\tilde{\mathbf{s}}(\mathbf{x}_r, \omega_k) - \tilde{\mathbf{d}}(\mathbf{x}_r, \omega_k)] \exp(i\omega_k t) \delta(\mathbf{x} - \mathbf{x}_r) \\ &= \text{Re} \sum_k \sum_r \left\{ \int [\mathbf{s}(\mathbf{x}_r, t') - \mathbf{d}(\mathbf{x}_r, t')] \exp(-i\omega_k t') dt' \right\} \\ &\quad \times \exp(i\omega_k t) \delta(\mathbf{x} - \mathbf{x}_r) \\ &= \sum_k \sum_r \left\{ \cos(\omega_k t) \int [\mathbf{s}(\mathbf{x}_r, t') - \mathbf{d}(\mathbf{x}_r, t')] \cos(\omega_k t') dt' \right. \\ &\quad \left. + \sin(\omega_k t) \int [\mathbf{s}(\mathbf{x}_r, t') - \mathbf{d}(\mathbf{x}_r, t')] \sin(\omega_k t') dt' \right\} \\ &\quad \times \delta(\mathbf{x} - \mathbf{x}_r). \end{aligned} \quad (\text{B-5})$$

The last equation demonstrates that frequency-domain inversion corresponds to using harmonic adjoint sources in time-domain inversion. Physically, using harmonic oscillations as adjoint sources excites motions related to select frequencies only, which is identical to solving frequency-domain wave equations at those specific frequencies. One may choose to minimize the real (or imaginary) part of $\tilde{\mathbf{s}} - \tilde{\mathbf{d}}$, in which case, the adjoint source takes only the first (or second) term in equation B-5.

Laplace-domain misfit function

A similar connection may be established between Laplace- and time-domain inversions, by minimizing differences in the coefficients of the Laplace transform:

$$\chi = \frac{1}{2} \sum_k \sum_r \|\hat{\mathbf{s}}(\mathbf{x}_r, \lambda_k) - \hat{\mathbf{d}}(\mathbf{x}_r, \lambda_k)\|^2. \quad (\text{B-6})$$

The Laplace transforms $\hat{\mathbf{s}}$ and $\hat{\mathbf{d}}$ are defined as

$$\hat{\mathbf{s}}(\mathbf{x}, \lambda) = \int \mathbf{s}(\mathbf{x}, t) \exp(-\lambda t) dt, \quad (\text{B-7})$$

$$\hat{\mathbf{d}}(\mathbf{x}, \lambda) = \int \mathbf{d}(\mathbf{x}, t) \exp(-\lambda t) dt. \quad (\text{B-8})$$

The corresponding adjoint source is

$$\begin{aligned} \mathbf{f}^{\dagger}(\mathbf{x}, t) &= \sum_k \sum_r [\hat{\mathbf{s}}(\mathbf{x}_r, \lambda_k) - \hat{\mathbf{d}}(\mathbf{x}_r, \lambda_k)] \exp(-\lambda_k t) \delta(\mathbf{x} - \mathbf{x}_r) \\ &= \sum_k \sum_r \left\{ \int [\mathbf{s}(\mathbf{x}_r, t') - \mathbf{d}(\mathbf{x}_r, t')] \exp(-\lambda_k t') dt' \right\} \\ &\quad \times \exp(-\lambda_k t) \delta(\mathbf{x} - \mathbf{x}_r). \end{aligned} \quad (\text{B-9})$$

The larger λ_k is, the more important early arrivals are in the inversion. Because the Laplace transform can be treated as a complex Fourier transform, one may use complex frequencies in the Fourier

transform to define a new misfit function that combines complementary advantages of frequency- and Laplace-domain inversions.

Time-domain misfit function — Effects of filtering

Another way to look at frequency selection is through filtering, which is applied most often in global and regional seismology. To take into account the filtering process explicitly, we may rewrite the waveform misfit function in the time domain as

$$\chi = \frac{1}{2} \sum_r \int \|F(t) * \mathbf{s}(\mathbf{x}_r, t) - F(t) * \mathbf{d}(\mathbf{x}, t)\|^2 dt, \quad (\text{B-10})$$

where the filtering process is denoted by convolution with the filter F . The adjoint source for measurements made on filtered seismograms is

$$\mathbf{f}^\dagger(\mathbf{x}, t) = \sum_r F(t) * \{F(-t) * [\mathbf{s}(\mathbf{x}_r, -t) - \mathbf{d}(\mathbf{x}_r, -t)]\} \delta(\mathbf{x} - \mathbf{x}_r). \quad (\text{B-11})$$

Effectively, to account for filtering processes, adjoint sources need to be filtered twice. Note that two-stage filtering introduces no phase shift, which is perfect for consistency in timing. Also note that measurements should be made between filtered synthetics and data, that is $F(t) * \mathbf{s}(\mathbf{x}, t)$ and $F(t) * \mathbf{d}(\mathbf{x}, t)$. Although this introduces phase shifts, the operation is consistently applied to data and synthetics.

REFERENCES

- Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: *Geophysics*, **48**, 1514–1524, doi: [10.1190/1.1441434](https://doi.org/10.1190/1.1441434).
- Červený, V., 2005, Seismic ray theory: Cambridge University Press.
- Chavent, G., 1974, Identification of function parameters in partial differential equations, in R. E. Goodman, and M. P. Polis, eds., Identification of parameter distributed systems: American Society of Mechanical Engineers.
- Chen, P., L. Zhao, and T. H. Jordan, 2007, Full 3D tomography for the crustal structure of the Los Angeles region: *Bulletin of the Seismological Society of America*, **97**, 1094–1120, doi: [10.1785/0120060222](https://doi.org/10.1785/0120060222).
- Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: *Geophysics*, **36**, 467–481, doi: [10.1190/1.1440185](https://doi.org/10.1190/1.1440185).
- Fichtner, A., B. Kennett, H. Igel, and H.-P. Bunge, 2009, Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods: *Geophysical Journal International*, **179**, 1703–1725, doi: [10.1111/j.1365-246X.2009.04368.x](https://doi.org/10.1111/j.1365-246X.2009.04368.x).
- Hill, R., 1990, Gaussian beam migration: *Geophysics*, **55**, 1416–1428, doi: [10.1190/1.1442788](https://doi.org/10.1190/1.1442788).
- Hill, R., 2001, Prestack Gaussian-beam depth migration: *Geophysics*, **66**, 1240–1250, doi: [10.1190/1.1487071](https://doi.org/10.1190/1.1487071).
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration, in J. Bednar, ed., Conference on Inverse Scattering: Theory and Application: SIAM, 206–220.
- Liu, Q., and J. Tromp, 2006, Finite-frequency kernels based on adjoint methods: *Bulletin of the Seismological Society of America*, **96**, 2383–2397, doi: [10.1785/0120060041](https://doi.org/10.1785/0120060041).
- Luo, Y., H. Zhu, T. Nissen-Meyer, C. Morency, and J. Tromp, 2009, Seismic modeling and imaging based upon spectral-element and adjoint methods: *The Leading Edge*, **28**, 568–574, doi: [10.1190/1.3124932](https://doi.org/10.1190/1.3124932).
- McMechan, G., 1982, Determination of source parameters by wavefield extrapolation: *Geophysical Journal International*, **71**, 613–628, doi: [10.1111/j.1365-246X.1982.tb02788.x](https://doi.org/10.1111/j.1365-246X.1982.tb02788.x).
- McMechan, G., 1983, Migration by extrapolation of time-dependent boundary values: *Geophysical Prospecting*, **31**, 413–420, doi: [10.1111/j.1365-2478.1983.tb01060.x](https://doi.org/10.1111/j.1365-2478.1983.tb01060.x).
- Nolet, G., 1987, Seismic wave propagation and seismic tomography, in G. Nolet, ed., *Seismic tomography: With applications in global seismology and exploration geophysics*: D. Reidel Publishing Company.
- Peter, D., D. Komatitsch, Y. Luo, R. Martin, N. Le Goff, E. Casarotti, P. Le Loher, F. Magnoni, Q. Liu, C. Blitz, T. Nissen-Meyer, P. Basini, and J. Tromp, 2011, Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes: *Geophysical Journal International*, **186**, 721–739, doi: [10.1111/j.1365-246X.2011.05044.x](https://doi.org/10.1111/j.1365-246X.2011.05044.x).
- Plessix, R., 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](https://doi.org/10.1111/j.1365-246X.2006.02978.x).
- Pratt, R., G. Shin, and G. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362, doi: [10.1046/j.1365-246X.1998.00498.x](https://doi.org/10.1046/j.1365-246X.1998.00498.x).
- Symes, W., 2007, Reverse time migration with optimal checkpointing: *Geophysics*, **72**, no. 5, SM213–SM221, doi: [10.1190/1.2742686](https://doi.org/10.1190/1.2742686).
- Talagrand, O., and P. Courtier, 1987, Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory: *Quarterly Journal of the Royal Meteorological Society*, **113**, 1311–1328, doi: [10.1002/qj.49711347812](https://doi.org/10.1002/qj.49711347812).
- Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2009, Adjoint tomography of the southern California crust: *Science*, **325**, 988–992, doi: [10.1126/science.1175298](https://doi.org/10.1126/science.1175298).
- Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2010, Seismic tomography of the southern California crust based on spectral-element and adjoint methods: *Geophysical Journal International*, **180**, 433–462, doi: [10.1111/j.1365-246X.2009.04429.x](https://doi.org/10.1111/j.1365-246X.2009.04429.x).
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266, doi: [10.1190/1.1441754](https://doi.org/10.1190/1.1441754).
- Tromp, J., C. Tape, and Q. Y. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: *Geophysical Journal International*, **160**, 195–216, doi: [10.1111/j.1365-246X.2004.02453.x](https://doi.org/10.1111/j.1365-246X.2004.02453.x).
- Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Wu, R., 1985, Gaussian beams, complex rays, and the analytic extension of the Green's function in smoothly inhomogeneous media: *Geophysical Journal of the Royal Astronomical Society*, **83**, 93–110, doi: [10.1111/j.1365-246X.1985.tb05158.x](https://doi.org/10.1111/j.1365-246X.1985.tb05158.x).
- Xie, X., and R. Wu, 2005, Multicomponent prestack depth migration using the elastic screen method: *Geophysics*, **70**, no. 1, S30–S37, doi: [10.1190/1.1852787](https://doi.org/10.1190/1.1852787).
- Yan, J., and P. Sava, 2008, Isotropic angle-domain elastic reverse-time migration: *Geophysics*, **73**, no. 6, S229–S239, doi: [10.1190/1.2981241](https://doi.org/10.1190/1.2981241).
- Zhang, Y., G. Zhang, and N. Bleistein, 2003, True amplitude wave equation migration arising from true amplitude one-way wave equations: *Inverse Problems*, **19**, 1113–1138, doi: [10.1088/0266-5611/19/5/307](https://doi.org/10.1088/0266-5611/19/5/307).
- Zhu, H., E. Bozdag, D. Peter, and J. Tromp, 2012, Structure of the European upper mantle revealed by adjoint tomography: *Nature Geoscience*, **5**, 493–498, doi: [10.1038/ngeo1501](https://doi.org/10.1038/ngeo1501).
- Zhu, H., Y. Luo, T. Nissen-Meyer, C. Morency, and J. Tromp, 2009, Elastic imaging and time-lapse migration based on adjoint methods: *Geophysics*, **74**, no. 6, WCA167–WCA177, doi: [10.1190/1.3261747](https://doi.org/10.1190/1.3261747).