Interpretation®

Skeletonized Wave Equation Inversion in VTI media without too much Math

Journal:	<i>Interpretation</i>
Manuscript ID	INT-2016-0212.R2
Manuscript Type:	2016-05 Skeletonized/sparse/multiscale geophysical inversion for the interpreter
Date Submitted by the Author:	17-Apr-2017
Complete List of Authors:	Feng, Shihang; King Abdullah University of Science and Technology, Earth Science Department Schuster, Jerry; King Abdullah University for Science and Technology, Earth Science Department;
Keywords:	anisotropy, tomography, VTI, full-waveform inversion, travelttime
Subject Areas:	Other

SCHOLARONE™
Manuscripts

Skeletonized Wave Equation Inversion in VTI Media without
too much Math

Shihang Feng*,Gerard T. Schuster*,

** Department of Earth Science and Engineering,
King Abdullah University of Science and Technology,
Thuwal, Saudi Arabia, 23955-6900.*

Email: shihang.feng@kaust.edu.sa; Gerard.schuster@kaust.edu.sa.

(May 15, 2017)

Skeletonized VTI Inversion

Running head: *Skeletonized VTI Inversion*

ABSTRACT

We present a tutorial for skeletonized inversion of pseudo-acoustic anisotropic VTI data. We first invert for the anisotropic models using wave equation traveltimes inversion. Here, the skeletonized data are the traveltimes of transmitted and/or reflected arrivals that lead to simpler misfit functions and more robust convergence compared to full waveform inversion. This provides a good starting model for waveform inversion. The effectiveness of this procedure is illustrated with synthetic data examples and a marine data set recorded in the Gulf of Mexico.

INTRODUCTION

Conventional full waveform inversion (FWI) inverts for a velocity model by using an objective function that minimizes the L2 norm of the differences between the predicted and the observed traces (Tarantola, 2005; Virieux and Operto, 2009). However, such a misfit function is highly non-linear and the iterations often get stuck in a local minimum. In order to mitigate this problem, a skeletonized representation of the data such as first-arrival traveltimes (Luo and Schuster, 1991a,b; Zhou et al., 1995) can be inverted to obtain the low-to-intermediate wavenumber details of the background velocity model. The misfit function for skeletonized inversion is quasi-linear with respect to the model parameters and enjoys better convergence properties than conventional full waveform inversion. In ray-based traveltime tomography (Chapman and Pratt, 1992; Van Leeuwen and Mulder, 2010), the high-frequency approximation might fail when the Earth's velocity variations are characterized by the same wavelength as in the source wavelet. To overcome this problem, solutions to the wave equation are used to invert the traveltimes, and this is known as the wave equation traveltime inversion (WT) method (Luo and Schuster, 1991a,b).

WT was previously implemented under the isotropic approximation (Zhou et al., 1995, 1997; Van Leeuwen and Mulder, 2010; Ma and Hale, 2013; Zhang et al., 2015). However, the first arrivals may be strongly influenced by anisotropy at the near-surface with the consequence that isotropic migration will mis-position reflectors in depth (Shen et al., 2012). Thus, it is necessary to take anisotropy into account when applying WT to near-surface data. In anisotropic media, the P- and SV-waves are coupled to one another. To avoid undesired SV-waves, the pseudo-acoustic approximation is made to separate the P- and SV-waves (Alkhalifah, 1998; Pestana et al., 2011).

We now present a mostly non-mathematical tutorial for applying WT to data recorded in a pseudo-acoustic transversely isotropic medium with a vertical symmetry axis (VTI). The traveltimes of first-arrival diving waves are inverted to give the low-wavenumber estimate of the anisotropic parameters in the shallow portion of the model. The diving waves typically are characterized by an offset-to-penetration-depth ratio of 5:1, which says that for every 5 km of source-receiver offset we can invert to a depth of about 1 km. To reach deeper depths we invert the traveltimes and waveforms of reflection arrivals, which can be reliably inverted

for depths that are no deeper than $\approx \frac{1}{5}$ the maximum source-receiver offset in the recording.

array. Thus, data recorded with a maximum source-receiver offset of 5 km suggests that the reflection tomogram will be reliable to a depth no deeper than approximately 3 km. The velocity models from WT can then be used as accurate starting models for FWI, which can estimate the high-wavenumber details of the model. We denote this hybrid inversion strategy as wave equation traveltime and waveform inversion (WTW) of seismic data (Zhou et al., 1995, 1997). We demonstrate the effectiveness of this strategy with both synthetic data and marine data from the Gulf of Mexico.

THEORY

Seismic data usually contain many wiggles so that the corresponding waveform misfit functions will suffer from the local minimum problem (Bunks et al., 1995). These local minima often prevent complete convergence of FWI to the actual velocity model. As an example, Figure 1 depicts diagrams of seismic traces and the associated misfit functions. The waveform misfit function ζ is defined as

$$\zeta = \frac{1}{2} \sum_{s,g,t} [\Delta p(\mathbf{g}, \mathbf{s}, t)]^2, \quad (1)$$

where $\Delta p(\mathbf{g}, \mathbf{s}, t) = p_{obs}(\mathbf{g}, \mathbf{s}, t) - p_{cal}(\mathbf{g}, \mathbf{s}, t)$ is the data residual. Here, $p_{obs}(\mathbf{g}, \mathbf{s}, t)$ is the observed trace for a point source at $\mathbf{s}$ and a geophone at $\mathbf{g}$, and the model in this example is a two-layer medium so that the scattered data contain both primaries and multiples from the reflector and free surface. The velocity of the top layer is $V=1.0$ km/s. A gradient optimization method seeks the global minimum of ζ where $V=1.0$ km/s for the top layer by using a starting velocity model (in this case the starting model is $V=1.5$ km/s on the RHS of Figure 1b) to compute the calculated trace $p_{cal}(\mathbf{g}, \mathbf{s}, t)$ that lowers the value of the misfit function. The velocity model is updated by a gradient optimization method such as the steepest descent method:

$$m(\mathbf{x})^{(k+1)} = m(\mathbf{x})^{(k)} - \alpha \sum_{s,g,t} \Delta p(\mathbf{g}, \mathbf{s}, t) \frac{\partial \Delta p(\mathbf{g}, \mathbf{s}, t)}{\partial m(\mathbf{x})}, \quad (2)$$

where α is the step length and the superscript k indicates the iteration number. The step length is calculated by a numerical line-search method, where Δm is the search direction.

Several trial steps length α_1 and α_2 are chosen, e.g., $\alpha_1 = 0.5$ and $\alpha_2 = 1$, and the misfit

function values $\zeta(m + \alpha_1 \Delta m)$, $\zeta(m + \alpha_2 \Delta m)$ are computed. If both of them are less than $\zeta(m)$ and $\zeta(m + \alpha_1 \Delta m) < \zeta(m + \alpha_2 \Delta m)$, then the third trial step length is defined as

$$\alpha_3 = \frac{\alpha_1^2 * (\zeta(m) - \zeta(m + \alpha_2 \Delta m)) - \alpha_2^2 * (\zeta(m) - \zeta(m + \alpha_1 \Delta m))}{2\zeta(m + \alpha_2 \Delta m) * (\alpha_1 - \alpha_2) + 2\zeta(m + \alpha_1 \Delta m) * \alpha_2 - 2\zeta(m + \alpha_2 \Delta m) * \alpha_1}. \quad (3)$$

The smallest values of $\zeta(m + \alpha_1 \Delta m)$, $\zeta(m + \alpha_2 \Delta m)$ and $\zeta(m + \alpha_3 \Delta m)$ are used as the step length (Nocedal and Wright, 2006).

Unfortunately the gradient methods stop at the nearest local minima (red circle in Figure 1b) and cannot proceed further unless special adjustments are made. To mitigate this local minima problem we need to begin our iterations at a starting velocity model that is close to the true model at $V=1$ km/s. An alternative is to simplify the misfit function to eliminate most of its bumpiness. One way to do this is to find the velocity model that minimizes the sum of the squared traveltimes as depicted in Figure 1c. Here there is only one simple traveltime for each trace, rather than thousands of amplitudes per trace that need to be explained by the velocity model. In this case the inverted velocity model will be much simpler so that the traveltime misfit function,

$$\zeta = \frac{1}{2} \sum_{s,g} [\Delta\tau(\mathbf{g}, \mathbf{s})]^2, \quad (4)$$

will be much simpler than the waveform misfit function, where $\Delta\tau(\mathbf{g}, \mathbf{s}) = \tau_{obs}(\mathbf{g}, \mathbf{s}) - \tau_{cal}(\mathbf{g}, \mathbf{s})$ is the traveltime residual. Figure 1d depicts the much simpler traveltime misfit function associated with the Figure 1c traveltimes, where the associated steepest descent formula is

$$m(\mathbf{x})^{(k+1)} = m(\mathbf{x})^{(k)} - \alpha \sum_{s,g} \Delta\tau(\mathbf{g}, \mathbf{s}) \frac{\partial \Delta\tau(\mathbf{g}, \mathbf{s})}{\partial m(\mathbf{x})}. \quad (5)$$

There are no local minima in this case so a gradient optimization method will quickly converge to the global minima at $V=1$ km/s. In fact, this inverted velocity model, also known as a velocity tomogram, can be used as a starting velocity model for FWI to mitigate the local minima problem. Another simplification, i.e. skeletonization, of the traces, is to low-pass filter the data as shown in Figure 1e. In this case the waveform misfit function in Figure 1f is much simpler than that for the original data in Figure 1b. We can also mute all but the early arrivals and get an even simpler misfit function. This strategy of simplifying

the data (traveltimes and/or early arrivals or low frequencies) is defined as skeletonization or multiscaling of the data (Luo and Schuster, 1991a,b; Sun and Schuster, 1993; Bunks et al., 1995) which will be adapted for WTW of anisotropic data.

Cigars versus Rabbit Ears

In skeletonized inversion, we first invert the transmission traveltimes associated with the cigars in Figure 2, and then invert the reflection traveltimes along the rabbit ears. The transmission arrivals are associated with diving waves with an offset-to-penetration-depth ratio of about 5:1. This means that for data recorded with a maximum source-receiver offset of 5 km we can only invert for the velocity down to a depth of about 1 km. Deeper velocities are estimated by inverting the reflection arrivals, which are reliably inverted at depths no deeper than about 2/3 the maximum source-receiver offset. This means that for a 5 km source-receiver offset we can only invert for the velocity down to a depth of about 3 km.

The diving-wave raypath is along the center portion of the cigar in Figure 2 and the region of the cigar defines the first Fresnel zone, also denoted as a wavepath. A diving wave that only propagates along the diving-wave raypath will be the earliest arrival at **g** for a source at **s**. Any ray from the source to the receiver that lies within the first-Fresnel zone wavepath will arrive no later than $T/2$ of the first arrival and thereby influence its arrival time, where T is the dominant period of the source wavelet.

If the predicted traveltime of the diving wave does not agree with the observed traveltime at **g**, then this indicates that there is a mistake in the estimated velocity model somewhere within the cigar. Thus, it makes sense to only update the velocity for portions of the velocity within the cigar and the velocity update will be proportional to the magnitude of the traveltime residual. This updating procedure is sometimes known as "smearing" the traveltime residual along the wavepath, where the Fréchet derivative $\partial\tau/\partial v_{p0}(\mathbf{x})$ in the gradient of equation 5 is given by

$$\frac{\partial\tau}{\partial v_{p0}(x)} = \int_0^T dt \left[\frac{-2}{Ev_{p0}^3(\mathbf{x})} \dot{g}(\mathbf{x}, t|\mathbf{g}, 0) \star \dot{p}(\mathbf{x}, t|\mathbf{s}, 0) \right] \times \dot{p}(\mathbf{g}, t - \Delta\tau|\mathbf{s}, 0)^{obs}, \quad (6)$$

where $g(\mathbf{x}, t|\mathbf{g}, 0)$ is the Green's function and

$$E = \int dt \dot{p}(\mathbf{g}, t|\mathbf{s}, 0)^{obs} \dot{p}(\mathbf{g}, t + \Delta\tau|\mathbf{s}, 0)^{cal}. \quad (7)$$

Inserting the Fréchet derivative in equation 6 into equation 5 yields the traveltimes gradient:

$$\gamma(\mathbf{x}) = \sum_{s,g} \int_0^T dt \left[\frac{2}{v_{p0}^3(\mathbf{x})} \dot{g}(\mathbf{x}, t|\mathbf{g}, 0) \star \dot{p}(\mathbf{x}, t|\mathbf{s}, 0) \right] \times \Delta\tilde{\tau}(\mathbf{g}, t|\mathbf{s}, 0), \quad (8)$$

where $\Delta\tilde{\tau}$ is the pseudo-traveltime residual,

$$\Delta\tilde{\tau}(\mathbf{g}, t|\mathbf{s}, 0) = -\frac{1}{E} \dot{p}(\mathbf{g}, t - \Delta\tau|\mathbf{s}, 0)^{obs} \Delta\tau(\mathbf{g}, t|\mathbf{s}, 0). \quad (9)$$

This pseudo-residual is only non-zero within the region of the wavepath that connects the source at $\mathbf{s}$ with the geophone at $\mathbf{g}$. If the residual is a waveform residual, then the velocity is updated by "smearing" the waveform residual within the cigar.

Ray-tracing tomography updates the velocity by smearing residuals along skinny ray-paths, but is restricted to high-frequency data. This restriction means that the reconstructed velocity model will be much smoother than the source wavelength so that the velocity tomograms lack high-wavenumber details. To overcome this approximation, we use wave-equation tomography which updates the velocity by smearing residuals along the wavepaths.

Workflow

The workflow for multiscale WTW is shown in Figure 3. The wave-equation traveltimes and waveform inversion (WTW) misfit function is a blend of WT and FWI misfit functions,

$$\epsilon = \frac{1-a}{2} \sum_{s,g} [\Delta\tau(\mathbf{g}, \mathbf{s})]^2 + \frac{a}{2} \sum_{s,g,t} [\Delta p(\mathbf{g}, \mathbf{s}, t)]^2, \quad (10)$$

where a is a scalar used to adjust the weights of the two misfit functions. In this workflow, we first set $a = 0$ for the early iterations to invert traveltimes for the low-intermediate wavenumbers of the velocity model. Then a is gradually set equal to 1 as the iterations

proceed. Here, it is assumed that each of the misfit functions is normalized to take a

maximum magnitude if 1. To avoid local minima in the misfit function, the data are low-pass filtered and then higher frequencies are gradually added to the traces to get more detail in the velocity model (Bunks et al., 1995; Operto et al., 2004). Various strategies are used for increasing the frequency content (Boonyasiriwat et al., 2009). Also, diving-wave traveltimes are first inverted, then reflection traveltimes are inverted to get velocity estimates of deeper sediments. In the last step, the details of the background models are reconstructed with waveform inversion using high-frequency reflections.

NUMERICAL RESULTS

Anisotropic WTW is now applied to synthetic data. The observed data are simulated by the rapid expansion method (REM) (Pestana et al., 2011) applied to the approximate P-wave pseudo-acoustic wave equation,

$$\frac{1}{v_{p0}^2} \frac{\partial^2 P}{\partial t^2} = - \left[(1 + 2\epsilon)k_x^2 + k_z^2 - \frac{2(\epsilon - \delta)k_x^2 k_z^2}{k_x^2 + k_z^2} \right] P, \quad (11)$$

where v_{p0} , ϵ , δ are Thomsen anisotropy parameters (Thomsen, 1986), k_x , k_z are spatial wavenumbers and P is the wavefield function in the time-wavenumber domain. The steepest decent method is used for WT and the conjugate gradient method is employed for FWI.

Transmission WTW and Anticline Model

An anisotropic anticline model is shown in Figure 4, where the model size is 750 m in the X direction and 250 m in the Z direction with a grid spacing of 6.25 m. Fifty-nine sources are located on the surface with a spacing of 12.5 m, and the vertical-component traces are recorded by 120 receivers spaced at an interval of 6.25 m on the free surface. The source wavelet is a Ricker wavelet with a peak frequency of 30 Hz.

Surface seismic data have very low sensitivity to variations in the anisotropic parameters δ , so δ cannot be accurately recovered from such data (Cheng et al., 2014). Therefore, only the parameters v_{p0} and ϵ are inverted simultaneously and δ is set equal to 0. The initial models are shown in Figure 4. Anisotropic transmission WT is associated with a highly under-determined system of equations since only traveltime information is used, where

the tomograms obtained after 10 iterations of transmission WT are shown in Figure 5.

Although the values of the misfit function converges to less than 0.1% of the original value at the 1st iteration, the shape of the anticline is not inverted and the three layers are poorly determined.

To reduce the non-uniqueness associated with the WT solution, we use transmission WTW. The first 10 iterations are characterized by $a = 0$ in the equation 10. After 10 iterations, the transmission WT gradient in equation 10 is turned off by setting $a = 1$ and the waveform gradient is used to give the WTW inversion result in Figure 5. We notice that transmission WTW provides a much more accurate tomogram compared to transmission WT. Both the v_{p0} and ϵ transmission WTW tomograms are consistent with a low-wavenumber rendering of the true models.

Transmission WTW and GOM Field Data

The proposed methods are also tested on a 2D marine data set. There are 100 shots with a shot interval of 37.5 m, and each shot is recorded by a 6 km long cable with 480 receivers spaced at a 12.5 m receiver interval. The nearest receiver offset from the source is 200 m and the data are processed by a 25-Hz Wiener filter (Boonyasiriwat et al., 2009). Here the source wavelet is extracted from the raw data by stacking the time-shifted reflection events together from 200 to 250 m offset in the shot gather. The reflection traveltimes are then used to time shift the traces so the reflection events from the same interface are flattened. These flattened reflection events are stacked together to get an estimate of the source wavelet. This method is only valid when near-offset reflections event can be clearly identified in the data and only one reflection event is used in the calculation of the source wavelet. The amplitude spectrum needs to be smoothed and the minimum phase assumption is required. To overcome these limitations, multichannel blind deconvolution algorithms (Xu et al., 1995; Kazemi and Sacchi, 2014) can be used for a more reliable estimation of the source wavelet.

Traveltime tomography is used to provide an initial v_{p0} model, and the ϵ initial model is set as a homogeneous model with $\epsilon = 0.02$ below the sea bottom as shown in Figure 6. Only far-offset first arrivals (>4 km) are used for transmission WT and only reflection waves are used in multiscale FWI. The kernel size of the Gaussian filter, which is used to smooth the

inverted velocity model, decreases as the number of FWI iterations increases. Figures 7 and 8 show the isotropic and anisotropic transmission WT and WTW tomograms. The transmission WT and WTW tomograms are similar except there is a high v_{p0} zone around $x = 4.2$ km and $z = 0.6$ km in the isotropic tomograms. The misfit curves for both cases converge to around 10% for transmission WT and around 60% for FWI. Figure 10 shows the RTM images constructed from the initial model, isotropic transmission WTW tomograms, and anisotropic transmission WTW tomograms. Both the isotropic and anisotropic RTM images show more detail than the initial image. From the zoom views of Figures 11 and 12, the anisotropic images are more focused and the reflectors are more continuous. The angle-domain common image gathers (ADCIG) are compared in Figures 13 and 14, where the yellow circles show the advantages of anisotropic WTW compared to isotropic WTW. In this case the events in the anisotropic ADCIGs are flatter than those in the isotropic ADCIGs.

Reflection WTW and Gaussian Model

A significant limitation of transmission WTW method is that long offset data are required to obtain information in the deeper regions of the model. This problem can be overcome by applying this methodology to both transmission and reflection data rather than just transmission events.

To validate this chain, data are generated for a two-layered anisotropic model with a Gaussian anomaly embedded in the center of the Figure 15 model. The maximum value of the Gaussian anomaly for each anisotropic parameter is the same as the value in the second layer of the true model. Ninety-nine sources are located on the surface with a spacing of 25 m. The data are recorded by 400 receivers spaced at an interval of 6.25 m and the source wavelet is a Ricker wavelet with a peak frequency of 30 Hz. The initial anisotropic model for reflection WT is homogeneous with $v_{p0} = 2000$ m/s and $\epsilon = 0.2$, and the reflector is at the depth of 4 km; this reflector is the reference interface used to calculate the wavepaths of the reflection events. The area close to the source and the reflector are not updated because of near-source artifacts.

The WT tomograms obtained after 10 iterations are shown in Figures 16a and 16b. The <https://mc.manuscriptcentral.com/interpretation>

location of the Gaussian anomaly is accurately inverted and the values of the anisotropic parameters are less than the true values since only low-intermediate wavenumbers are reconstructed. Here, the strong artifacts in the ϵ tomogram are consistent with the theory that near-offset reflection data are weakly sensitive to changes in ϵ (Cheng et al., 2014). After 10 iterations of FWI, the high-wavenumber portions of the model, i.e. the geometry of the reflector, are recovered. However, the update in the ϵ WTW tomogram at the location of the reflector is polluted by the crosstalk between v_{p0} and ϵ .

Transmission + Reflection WTW and Marmousi Model

Transmission + reflection WTW is now applied to data generated for the Marmousi model, as described by the workflow diagram in Figure 3. The true v_{p0} and ϵ models are shown and the initial models are given as the horizontal layered model with the anisotropic parameter values linearly increasing in depth as depicted in Figure 17. The model size is 600×160 gridpoints, and the grid interval is 6.25 m, with 160 shot gathers having a shot interval of 6.25 m. Each shot gather is recorded with receivers fixed on the surface with a line length of 3.6 km. The source wavelet is a Ricker wavelet with a peak frequency of 30 Hz. The black dashed line in Figure 17a is chosen as the reference reflector.

In the first 10 iterations, the low- and intermediate-wavenumber components in the shallow region of the model are recovered using anisotropic WT with diving waves. Then the smooth portions above the reference reflector are reconstructed after 5 iterations of reflection WT. In the later 20 iterations, multiscale FWI is applied to reconstruct the high-wavenumber parts of the models. A bandpass filter with the center frequency of 16 Hz is applied to the gradient in the first 10 iterations of FWI, then higher frequencies are admitted in the latter iterations. For the area above the reference reflector, the WTW tomograms are consistent with the true models as shown in Figure 19 except some regions in the ϵ WTW tomogram are contaminated by crosstalk (red circles in Figures 19b and 19d). However, the v_{p0} and ϵ tomograms below the reference reflector can't be updated accurately due to the absence of reflection wavepaths. This problem can be remedied by choosing a deeper reference reflector and using the deepest reference reflections to update the deeper part of the model.

CONCLUSIONS

We presented a mostly non-mathematical tutorial for inverting skeletonized data. In our examples, the anisotropic parameters for a VTI medium were estimated by inverting both the traveltimes and waveforms in the traces. The long-wavelength features of the anisotropic models are reconstructed by WT and the short-wavelength features are then recovered by waveform inversion. In the synthetic tests, transmission WTW tomograms show accurate and detailed structural information, which is important in reservoir characterization. In the field data test, anisotropic transmission WTW provides flatter images in the ADCIG domain and more focused RTM images than isotropic WTW. The main problem with transmission WTW is the requirement for a large offset-to-penetration-depth ratio with diving waves. By combining transmission and reflection WTW, this problem can be largely eliminated. The reflection and transmission WTW can update deeper areas in the model and also gives accurate and detailed models of anisotropic parameters as shown in the synthetic examples.

A significant limitation of the transmission + reflection WTW method is that the location of reference reflectors should be accurately known before inversion. This problem can be mitigated by updating the reference reflector position with RTM images at every iteration.

ACKNOWLEDGEMENTS

The research reported in this publication is supported by funding from King Abdullah University of Science and Technology (KAUST). We also thank the sponsors for supporting the Consortium of Subsurface Imaging and Fluid Modeling (CSIM). For computer time, this research used the resources of the Supercomputing Laboratory at KAUST. We thank them for providing the computational resources required for carrying out this work.

REFERENCES

- Alkhalifah, T., 1998, Acoustic approximations for processing in transversely isotropic media: Geophysics, **63**, 623–631.
- Boonyasiriwat, C., P. Valasek, P. Routh, W. Cao, G. T. Schuster, and B. Macy, 2009, An efficient multiscale method for time-domain waveform tomography: Geophysics, **74**, WCC59–WCC68.
- Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.
- Chapman, C. t. and R. Pratt, 1992, Traveltime tomography in anisotropic media. theory: Geophysical Journal International, **109**, 1–19.
- Cheng, X., K. Jiao, D. Sun, D. Vigh, et al., 2014, Anisotropic parameter estimation with full-waveform inversion of surface seismic data: Presented at the 2014 SEG Annual Meeting.
- Kazemi, N. and M. D. Sacchi, 2014, Sparse multichannel blind deconvolution: Geophysics, **79**, V143–V152.
- Luo, Y. and G. T. Schuster, 1991a, Wave equation inversion of skeletalized geophysical data: Geophysical Journal International, **105**, 289–294.
- , 1991b, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.
- Ma, Y. and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion: Geophysics, **78**, R223–R233.
- Nocedal, J. and S. Wright, 2006, Numerical optimization: Springer.
- Operto, S., C. Ravaut, L. Improta, J. Virieux, A. Herrero, and P. Dell'Aversana, 2004, Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study: Geophysical prospecting, **52**, 625–651.
- Pestana, R. C., B. Ursin, and P. L. Stoffa, 2011, Separate p-and sv-wave equations for vti media: Presented at the 12th International Congress of the Brazilian Geophysical Society.
- Schuster, G. T., 2017, Seismic inversion: SEG Publishing, Tulsa, Oklahoma.
- Shen, X. et al., 2012, Early-arrival waveform inversion for near-surface velocity and anisotropic parameter: Parameterization study: Presented at the 2012 SEG Annual Meeting.
- Sun, Y. and G. T. Schuster, 1993, Time-domain phase inversion, *in* SEG Technical Program

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: *siam*.

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

Van Leeuwen, T. and W. Mulder, 2010, A correlation-based misfit criterion for wave-equation travelttime tomography: *Geophysical Journal International*, **182**, 1383–1394.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC1–WCC26.

Xu, G., H. Liu, L. Tong, and T. Kailath, 1995, A least-squares approach to blind channel identification: *IEEE Transactions on signal processing*, **43**, 2982–2993.

Zhan, G., W. Dai, M. Zhou, Y. Luo, and G. T. Schuster, 2014, Generalized diffraction-stack migration and filtering of coherent noise: *Geophysical Prospecting*, **62**, 427–442.

Zhang, S., Y. Luo, and G. Schuster, 2015, Shot-and angle-domain wave-equation travelttime inversion of reflection data: *Theory: Geophysics*, **80**, U47–U59.

Zhou, C., W. Cai, Y. Luo, G. T. Schuster, and S. Hassanzadeh, 1995, Acoustic wave-equation travelttime and waveform inversion of crosshole seismic data: *Geophysics*, **60**, 765–773.

Zhou, C., G. T. Schuster, S. Hassanzadeh, and J. M. Harris, 1997, Elastic wave equation travelttime and waveform inversion of crosswell data: *Geophysics*, **62**, 853–868.

Figure 1. Synthetic data in left column and right column depicts associated plots of misfit functions vs hypothetical velocity values of the first layer in the two-layer model. The correct value of V is 1.0 km/s, where the misfit functions tend to zero. Notice the fewer local minima (red circles) for the 2-Hz data compared to that for the 10-Hz data.

321x334mm (300 x 300 DPI)

Figure 2. a) Horizontal-reflector models along with their associated wavepaths to the right; the wavepaths in b) take the shape of a banana (i.e, cigar for a homogenous model), rabbit ears, and an elliptical smile. The wavepaths were computed by migrating a single trace excited by a wideband point source (star) and recorded at the geophone (quadrilateral).

135x41mm (300 x 300 DPI)

Figure 3. Left column: the workflow for multiscale inversion. Right column: examples of the data used in each step of multiscale inversion.

254x254mm (300 x 300 DPI)

Figure 4. a) The true v_{p0} and b) ϵ models for the anticline tests. c) The initial v_{p0} and d) ϵ models for the anticline tests.

200x88mm (300 x 300 DPI)

Figure 5. a) The transmission WT v_{p0} and b) ϵ tomograms after 10 iterations. c) The transmission WTW v_{p0} and d) ϵ tomograms after 10 iterations of FWI.

200x88mm (300 x 300 DPI)

Figure 6. a) The initial v_{p0} and b) initial ϵ models for the GOM data.

140x82mm (300 x 300 DPI)

Figure 7. a) The isotropic transmission WT v_{p0} tomograms after 7 iterations with the GOM data. b) The anisotropic transmission WT v_{p0} and c) ϵ tomograms after 4 iterations for the GOM data.

220x202mm (300 x 300 DPI)

Figure 8. a) The isotropic transmission WTW v_{p0} tomogram after 9 iterations of multiscale FWI iterations for the GOM data. b) The anisotropic transmission WTW v_{p0} and c) ϵ tomograms after 9 iterations of multiscale FWI with the GOM data.

220x202mm (300 x 300 DPI)

Figure 9. The RTM images constructed from the a) initial v_{p0} model, b) isotropic transmission WT tomogram and c) anisotropic transmission WT tomogram using the GOM data as the input traces.

220x202mm (300 x 300 DPI)

Figure 10. Same as Figure 9 except the WTW tomogram, not the WT tomograms, are used for migration.

220x202mm (300 x 300 DPI)

Figure 11. The zoom views of the yellow boxes in the GOM RTM images computed with the a) initial v_{p0} model, b) isotropic transmission WT tomogram and c) anisotropic transmission WT tomogram.

80x17mm (300 x 300 DPI)

Figure 12. The zoom views of the red boxes in the RTM image computed with the a) initial v_{p0} model, b) isotropic transmission WTW tomogram and c) anisotropic transmission WTW tomogram using the GOM data as the input traces.

220x220mm (300 x 300 DPI)

Figure 13. The ADCIGs computed with the a) initial v_{p0} model, b) isotropic transmission WTW tomogram and c) anisotropic transmission WTW tomogram using the GOM data at $x=3$ km.

150x68mm (300 x 300 DPI)

Figure 14. The ADCIGs computed with the a) initial v_{p0} model, b) isotropic transmission WTW tomogram and c) anisotropic transmission WTW tomogram using the GOM data at $x=4.5$ km.

150x68mm (300 x 300 DPI)

Figure 15. a) The true v_{p0} and b) ϵ models for the Gaussian tests. c) The initial v_{p0} and d) ϵ models for the reflection WTW tests.

200x88mm (300 x 300 DPI)

Figure 16. a) The reflection WT v_{p0} and b) ϵ tomograms after 10 iterations. c) The reflection WTW v_{p0} and d) ϵ tomograms after 10 iterations of FWI.

200x88mm (300 x 300 DPI)

Figure 17. a) The true v_{p0} and b) ϵ models for the Mamousi tests (black dashed line is the reference reflector for reflection WT). c) The initial v_{p0} and d) ϵ models for the Mamousi tests.

200x88mm (300 x 300 DPI)

Figure 18. a) The transmission WT v_{p0} and b) ϵ tomograms after 10 iterations. c) The reflection WT v_{p0} and d) ϵ tomograms after 5 iterations.

200x88mm (300 x 300 DPI)

Figure 19. a) The transmission+reflection WTW v_{p0} and b) ϵ tomograms after 10 iterations FWI with a 16 Hz bandpass filter. c) The transmission+reflection WTW v_{p0} and d) ϵ tomograms after 10 iterations of FWI with 30 Hz data. e) The FWI v_{p0} and f) ϵ tomograms after 35 iterations of multi-scale FWI using initial model. Red circles indicate the regions contaminated by crosstalks.

232x146mm (300 x 300 DPI)