GEOPHYSICS®

Multiscale Phase Inversion of Seismic Data

Journal:	<i>Geophysics</i>
Manuscript ID:	GEO-2017-0353.R1
Manuscript Type:	Technical Paper
Date Submitted by the Author:	15-Oct-2017
Complete List of Authors:	Fu, Lei; King Abdullah University of Science and Technology, Earth Science and Engineering Program Guo, Bowen; King Abdullah University of Science and Technology, Earth science and engineering Sun, Yonghe; Chevron Corp Schuster, Jerry; King Abdullah University for Science and Technology, Earth Science Department;
Keywords:	full-waveform inversion, integration, phase
Area of Expertise:	Seismic Inversion
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
table_1.tex table_2.tex	

SCHOLARONE™
Manuscripts

Multiscale Phase Inversion of Seismic Data

Lei Fu¹ , Bowen Guo¹ , Yonghe Sun² and Gerard T. Schuster¹

¹ King Abdullah University of Science and Technology, Division of Physical Science and Engineering, Thuwal 23955-6900, Saudi Arabia

² Chevron Corporation

Running Head: 2D-MPI

ABSTRACT

We present a scheme for multiscale phase inversion (MPI) of seismic data that is less sensitive to the unmodeled physics of wave propagation and a poor starting model than standard full waveform inversion (FWI). To avoid cycle-skipping, the multiscale strategy temporally integrates the traces several times, i.e. high-order integration, to produce low-boost seismograms that are used as input data for the initial iterations of MPI. As the iterations proceed, higher frequencies in the data are boosted by using integrated traces of lower order as the input data. The input data are also filtered into different narrow frequency bands for the MPI implementation. At low frequencies, we show that MPI with windowed reflections approximates wave equation inversion of the reflection traveltimes, except no traveltime picking is needed. Numerical results with synthetic acoustic data show that MPI is more robust than conventional multiscale FWI when the initial model is moderately far from the true model. Results from synthetic viscoacoustic and elastic data show that MPI is less sensitive than FWI to some of the unmodeled physics. Inversion of marine data shows that MPI is more robust and produces modestly more accurate results than FWI for this data set.

INTRODUCTION

Full waveform inversion (FWI) inverts for the high-resolution velocity model by minimizing the misfit between the predicted and observed data (Tarantola, 1984; Virieux and Operto, 2009). Despite its growing success, FWI suffers from the problem of cycle skipping (Bunks et al., 1995; Virieux and Operto, 2009; Warner and Guasch, 2016) when the predicted and observed events have a traveltime difference of more than half a period. Since FWI

is a gradient-based method, the iterative solution tends to get stuck in a local minimum associated with the cycle skipped events.

One of the reasons for cycle skipping is that an inaccurate starting model gives rise to large kinematic errors between the observed and predicted data (Bunks et al., 1995). To begin with a more accurate starting model, traveltime inversion of first arrivals (Luo and Schuster, 1991; Pratt and Goult, 1991; Hole, 1992; Nemeth et al., 1997; Min and Shin, 2006) is a robust method for estimating the velocity model at shallow and intermediate depths. The deeper regions of the velocity model cannot be estimated by first-arrival traveltime inversion because such depths are not accessed by the recorded diving waves. Another drawback of first-arrival traveltime inversion is that it is not suitable in the presence of low-velocity zones that create shadow zones for first arrivals (Virieux and Operto, 2009).

As an alternative, migration velocity analysis (MVA) can estimate the deep parts of the starting velocity model (Stork, 1992; Tieman, 1995; Sava et al., 2005; Sava and Vlad, 2008), as long as the offset-to-depth ratio is greater than about 1.5. MVA estimates the vertical shifts that flatten the migrated reflections from a common reflector and back-projects them to update the velocity model. However, it is still challenging to robustly extract these shifts for complex geology such as salt models. Moreover, the velocity model inverted by MVA can be distorted by multiples (Van Leeuwen and Mulder, 2008).

A multiscale strategy can also mitigate the cycle skipping and local minima problems of FWI (Bunks et al., 1995). Multiscale FWI first applies a low-pass filter to the data, so that the objective function has fewer local minima. After a suitable number of FWI iterations, higher frequencies are then added into the filtered data to start the next stage of FWI iterations. A problem with multiscale FWI is that the low-frequency components in

the recorded data are sometimes significantly weaker than the high-frequency components.

In addition to cycle skipping, another drawback of FWI is that the data misfit function is very sensitive to the amplitude mismatch between the recorded and predicted data. This mismatch often occurs because all of the physics of wave propagation are not modeled in the predicted data. For example, viscoacoustic and anisotropic effects are not fully modeled in the predicted data computed by solutions to the acoustic wave equation. A possible remedy to this problem is phase inversion, where the objective function is the sum of the squared phase differences between the predicted and observed data (Schuster, 1991; Woodward, 1992; Shin and Min, 2006; Maharramov et al., 2017) or a correlation-based functional (Luo and Schuster, 1991; De Hoop and van Der Hilst, 2005; Van Leeuwen and Mulder, 2010; Baumstein et al., 2011; Routh et al., 2011; Masoni et al., 2013; Luo et al., 2016). Some phase-like inversion methods invert the data in the space-Laplace domain (Shin and Min, 2006) or in the $\tau - p$ domain (Zhang and Wang, 2009). However the penalty of phase inversion is a moderate loss of resolution in the inverted tomogram because the inversion ignores the magnitude spectrum in the data. As we will show in our data examples, this reduction in resolution is modest and often quite tolerable for multiscale phase inversion.

To combine the multiscale FWI and phase inversion methods, Sun and Schuster (1993) proposed the multiscale phase inversion (MPI) method, where the recorded data are integrated several times in the time domain to boost the low-frequency components. This low-frequency boosting tends to mitigate the cycle-skipping problem so that the solution does not get stuck in a local minimum. Then a multiscale phase inversion method is applied to the integrated data, with the misfit function as the l_2 -norm of the difference between the modified-predicted data and the observed data. Here, the modified predicted trace has the same phase spectrum as the predicted data but it has the amplitude spectra of the observed

trace. This modification relieves MPI of having to explain the magnitude spectrum in the recorded data. Sun and Schuster (1993) showed a successful example of applying MPI to synthetic crosswell data. We now implement the MPI method on synthetic data and field data recorded by surface seismic experiments.

After the introduction, the next section introduces the theory of MPI. This is followed by the numerical results section which presents the results of applying MPI to synthetic data and field data. The Appendix shows that, at low frequencies, MPI approximates wave equation travelttime inversion of the reflections, except no travelttime picking is needed. It also relates MPI to other wave equation inversion methods. The last section provides a summary and conclusions.

THEORY

The calculated $d(\mathbf{g}, t|\mathbf{s})$ and observed $d(\mathbf{g}, t|\mathbf{s})^{obs}$ traces are Fourier-transformed to obtain

$$\mathcal{F}[d(\mathbf{g}, t|\mathbf{s})] = |D(\mathbf{g}|\mathbf{s})| e^{i\phi(\mathbf{g}|\mathbf{s})}, \tag{1}$$

$$\mathcal{F}[d(\mathbf{g}, t|\mathbf{s})^{obs}] = |D(\mathbf{g}|\mathbf{s})^{obs}| e^{i\phi(\mathbf{g}|\mathbf{s})^{obs}}, \tag{2}$$

where $|D|$ is the magnitude spectrum, ϕ is the phase spectrum, $\mathbf{s}$ is the location of the source and $\mathbf{g}$ is the location of the geophone. It is recognized that we can model the phase information, i.e. the kinematics of wave propagation, more accurately than the amplitude spectrum, so we modify the calculated trace $\bar{d}(\mathbf{g}, t|\mathbf{s})$ by replacing $|D(\mathbf{g}|\mathbf{s})|$ with $|D(\mathbf{g}|\mathbf{s})^{obs}|$ and apply the inverse Fourier transform to get the modified data:

$$\bar{d}(\mathbf{g}, t|\mathbf{s}) = \mathcal{F}^{-1} \left\{ L(\omega) |D(\mathbf{g}|\mathbf{s})^{new}| e^{i\phi(\mathbf{g}|\mathbf{s})} \right\}, \quad (3)$$

$$\bar{d}(\mathbf{g}, t|\mathbf{s})^{obs} = \mathcal{F}^{-1} \left\{ L(\omega) |D(\mathbf{g}|\mathbf{s})^{obs}| e^{i\phi(\mathbf{g}|\mathbf{s})^{obs}} \right\}, \quad (4)$$

where $|D(\mathbf{g}|\mathbf{s})^{new}| = |D(\mathbf{g}|\mathbf{s})^{obs}|$, $L(\omega)$ is a low-pass filter applied to the data, and the bar indicates the modified traces.

Misfit function. The modified traces are time-integrated and their difference is used for the MPI misfit function,

$$\epsilon^{mpi} = \sum_{\mathbf{s}, \mathbf{g}} \int dt \left[I^n \bar{d}(\mathbf{g}, t|\mathbf{s}) - I^n \bar{d}(\mathbf{g}, t|\mathbf{s})^{obs} \right]^2, \quad (5)$$

where I^n is an integration operator $I \equiv \int dt$, I^n indicates integration performed n times, and $\bar{d}(\mathbf{g}, t|\mathbf{s})$ and $\bar{d}(\mathbf{g}, t|\mathbf{s})^{obs}$ are the modified traces in equations (3) and (4). If we set $|D(\mathbf{g}|\mathbf{s})^{new}| = |D(\mathbf{g}|\mathbf{s})|$ in equation (3), then the MPI misfit function becomes that for full waveform inversion, except that the traces have been modified by the bandpass filter $L(\omega)$ and the integration operator.

Figure ??-a1 shows the calculated trace p_{cal} with one reflection event. The observed trace p_{obs} with three reflection events is shown in Figure ??-a2, and the corresponding modified calculated and observed traces are shown in Figures ??-a3 and ??-a4, respectively. Figures ??-b1 - ??-b4 show the calculated, observed, modified-calculated, and modified-observed traces with one integration, respectively. The events in the traces have the shape of a Gaussian wavelet. Finally, Figures ??-c1 - ??-c4 show the calculated, observed, modified-calculated, and modified-observed traces after two integrations, respectively. The events

now resemble Heaviside functions.

The Appendix shows that, at low frequencies, MPI applied to reflections approximates wave equation travelttime inversion of the reflections, except no travelttime picking is needed. It also relates MPI to other wave equation inversion methods.

Gradient. The gradient of the MPI misfit function ϵ^{mpi} w.r.t. the velocity field $c(\mathbf{x})$ is

$$\begin{aligned}\gamma^{mpi}(\mathbf{x}) &= \frac{\partial \epsilon^{mpi}}{\partial c(\mathbf{x})} \\ &= \frac{1}{c(\mathbf{x})^3} \sum_{\mathbf{s}, \mathbf{g}} \int dt \left[I^n \dot{\bar{d}}(\mathbf{x}, t|\mathbf{s}) \right] \left[I^n \dot{\bar{d}}'(\mathbf{x}, t|\mathbf{g}) \right],\end{aligned}\tag{6}$$

where dot indicates time differentiation, $\bar{d}(\mathbf{x}, t|\mathbf{s})$ is the forward pressure wavefield (Sun and Schuster, 1993)

$$\bar{d}(\mathbf{x}, t|\mathbf{s}) = g(\mathbf{x}, t|\mathbf{s}) * w(t),\tag{7}$$

$w(t)$ is the source wavelet and $g(\mathbf{x}, t|\mathbf{s})$ is the Green's function associated with the acoustic wave equation $\nabla^2 d(\mathbf{x}, t) - \frac{1}{v(\mathbf{x})^2} \frac{\partial^2 d(\mathbf{x}, t)}{\partial t^2} = f(\mathbf{x}, t)$. The term $\bar{d}'(\mathbf{x}, t|\mathbf{g})$ is the wavefield computed by backprojecting the seismogram residual $\delta \bar{d}$ (Luo and Schuster, 1991) :

$$\bar{d}'(\mathbf{x}, t|\mathbf{g}) = g(\mathbf{x}, -t|\mathbf{g}, 0) * \delta \bar{d},\tag{8}$$

$$\text{where} \quad \delta \bar{d} = \bar{d}(\mathbf{g}, t|\mathbf{s})^{obs} - \bar{d}(\mathbf{g}, t|\mathbf{s}).\tag{9}$$

As an alternative, the integration operator I^n applied to $\bar{d}$ and $\bar{d}'$ in equation 6 can be applied to the source-wavelet functions that generate them without changing the misfit gradient. For example, $I^n \dot{\bar{d}}(\mathbf{x}, t|\mathbf{s})$ is equivalent to computing a wavefield using a source

wavelet integrated $n - 1$ times. As a comparison, the conventional full waveform inversion (FWI) gradient is

$$\begin{aligned}\gamma^{fwi}(\mathbf{x}) &= \frac{\partial \epsilon^{fwi}}{\partial c(\mathbf{x})}, \\ &= \frac{1}{c(\mathbf{x})^3} \sum_{\mathbf{s}, \mathbf{g}} \int dt \left[\dot{d}(\mathbf{x}, t|\mathbf{s}) \right] \left[\dot{d}'(\mathbf{x}, t|\mathbf{g}) \right],\end{aligned}\quad (10)$$

where the FWI gradient is the dot product between the source forward wavefield $\dot{p}$ of the source and the wavefield computed by backprojecting the data residual $\delta d = d(\mathbf{g}, t|\mathbf{s})^{obs} - d(\mathbf{g}, t|\mathbf{s})$.

Figure ?? shows the comparison of the FWI and MPI gradients for data generated in the Figure ??a model, where Figures ??a and ??b are the true and initial models, respectively. The sources and receivers are evenly distributed on the surface. We use a first-order derivative of the Gaussian function with the peak frequency of 15 Hz as the source wavelet. Figure ??c shows the FWI gradient after the first iteration, and Figures ??d-??f depict the MPI gradients with different orders of integrations applied to the traces at the first iteration. If no integration is applied to these traces, the MPI gradient is quite similar to the FWI gradient. When $n = 2$ integrations are applied to the traces, the MPI gradient provides low-wavenumber information about the model, and the higher-wavenumber model information will be gradually incorporated as n decreases to zero.

Frequency multiscale strategy. The data are band-pass filtered into different frequency bands with different peak frequencies. The low-frequency data are used at the early iterations and then later iterations utilize higher frequency data. A Wiener filter (Boonyasiriwat

et al., 2009) is computed by

$$L_{wiener}(\omega) = \frac{W_{target}(\omega)W_{original}^{\dagger}(\omega)}{|W_{original}(\omega)|^2 + \epsilon^2}, \tag{11}$$

where, $L_{wiener}(\omega)$ is the Wiener filter, $W_{original}(\omega)$ is the original wavelet, $W_{target}(\omega)$ is the target wavelet, $\dagger$ denotes complex conjugation, ω is the angular frequency, and ϵ is the damping factor to prevent numerical instability. The formula for selecting optimal frequency bands proposed by Sirgue and Pratt (2004) is

$$f_{n+1} = \frac{f_n}{\alpha_{min}}, \tag{12}$$

where f_n is the current frequency, f_{n+1} is the next frequency to be chosen, and $\alpha_{min} = z/\sqrt{h^2 + z^2}$ is the parameter that depends on the maximum half offset h and the maximum depth z to be imaged.

Workflow of MPI. The multiscale frequency strategy is used with FWI and MPI, and their workflows are described in Figure ???. The MPI strategy (Sun and Schuster, 1993) temporally integrates the traces n times to produce low-boost seismograms that are used as input data for the initial iterations. These low-boost traces are inverted for several iterations until the decrease in residual (ϵ) is below a specified threshold (ϵ_1). Then the data are high-boosted by using the $n - 1$ integrated traces as the input data and the low-boosted tomogram as the new starting model. We repeat this process until the integration order is $n = 0$ and the data residual falls below a specified threshold. The data will then be appended with higher frequencies and the MPI procedure is repeated. See the dashed box in the MPI workflow shown in Figure ??b.

NUMERICAL RESULTS

To demonstrate the effectiveness of MPI and its advantages, we invert synthetic data from a modified Marmousi model and field data recorded by a marine survey. The modeling kernels for both MPI and FWI are based on solutions to the constant-density acoustic wave equation, while the synthetic input data are the numerical solutions to the constant-density acoustic wave equation (Alford et al., 1974), viscoacoustic wave equation (Operto et al., 2007) or elastic wave equation (Levander, 1988).

Marmousi Model

The modified Marmousi model (Figure ??) is uniformly discretized into 243×767 gridpoints with a node spacing of 10 m. There are 192 sources with a 40 m interval and 383 receivers with a 20 m interval located along the free surface at the top of the model.

Acoustic Data

The input acoustic data are generated by solving the constant-density acoustic equation, where the source wavelet is a Ricker wavelet with the peak frequency of 15 Hz. A common-shot gather for the source at $(x, z) = (0, 0)$ m is shown in Figure ??a, where different bandpass filters (table ??) are applied to the original data. For the frequency-multiscale strategy used with MPI and FWI, low-frequency data are inverted at an early stage, which provides low-wavenumber model information. Higher-frequency data are gradually incorporated at later stages, which provide the high-wavenumber information about the model.

Figure ??a depicts the smoothed initial velocity model with a velocity error of 12%, and

Figure ??b is a $v(z)$ initial velocity model with a velocity error of 22%. The FWI and MPI

tomograms with the smoothed initial model (Figure ??a) are shown in Figures ??c and ??e. The FWI and MPI tomograms with the $v(z)$ initial model (Figure ??b) are shown in Figures ??d and ??f. Figure ?? shows the velocity profile comparisons for the true, initial, FWI and MPI velocity models at different positions. We can see that both the FWI and MPI tomograms are in good agreement with the true model when the initial model is not too far from the true model. However, when the initial model is far away from the true model, traditional FWI gets stuck in a local minimal, while MPI provides a more accurate tomogram.

Figure ?? shows the data residual at each iteration for the FWI and MPI methods when the smoothed initial model (Figure ??a) is used. The original data are filtered into 4 different frequency bands, so there are 4 stages of convergence for both methods. In addition, the maximum integration number $N_{max} = 2$ gradually decreases to 0 for each frequency band, so there are 3 minor stages of convergence for each frequency band in the MPI method. The computational cost of MPI is about twice that of FWI because of the integration operations.

Visco-acoustic Data

We now use viscoacoustic data as input traces to the acoustic FWI and MPI algorithms. The goal is to test the sensitivity of each method to the unmodeled attenuation effects in the predicted data. The observed viscoacoustic data are generated by solving the viscoacoustic equation, where the source wavelet and acquisition geometry are the same as in the acoustic case. A pressure source is injected in the water, and the pressure field is recorded. The true v_p model is shown in Figure ??a and the Q model is shown in Figure ??b, where the

minimum Q is 15. The viscoacoustic data for the shot location at $(x, z) = (0, 0)$ m are shown in Figure ??b, where the reflections and refractions are attenuated due to the highly attenuative medium. The amplitudes of the viscoacoustic traces are quite different than those in the acoustic data (Figure ??a).

We apply the acoustic FWI and MPI methods to these synthetic viscoacoustic data, where the smoothed initial velocity model is shown in Figure ??a. The FWI and MPI tomograms are shown in Figures ??a and ??c, respectively, and the magnified views of the dashed box 'A' is shown in Figures ??c and ??e. The velocity profile comparisons at different locations are shown in Figure ??a, and the final model misfits of the FWI and MPI tomograms with respect to the true velocity model are 11.5% and 10.8%, respectively. Both the FWI and MPI tomograms contain significant errors in the deep areas, but the MPI velocity profiles tend to be more accurate. The MPI tomogram is also more accurate than the FWI tomogram in the shallow parts of the model because the MPI method is less sensitive than the FWI method to errors in computing amplitudes.

Elastic Data

We now use elastic data as input traces to the acoustic FWI and MPI algorithms. The goal is to test the sensitivity of each method to the unmodeled elastic effects in the data. The elastic data are generated by solving the isotropic elastic wave equation, where the source wavelet and acquisition geometry are the same as in the acoustic case. A pressure pulse is injected in the water and the pressure field is recorded as the negative of the average of the normal stresses. The true v_p model is shown in Figure ??a, the density is given by Gardner's relation $\rho = 0.31v_p^{0.25}$, and $v_s = v_p/\sqrt{3}$, except the shear velocity of the water layer is set to

0 m/s. The elastic data at shot location $(x, z) = (0, 0)$ m are shown in Figure ??c, where the three red arrows indicate the converted 'PSP' arrivals.

The FWI and MPI tomograms are shown in Figures ??b and ??d, respectively, and the magnified views of the dashed box 'B' is shown in Figures ??d and ??f. The associated velocity profiles are shown in Figure ??b. We can see that the MPI tomogram is noticeably more accurate than the FWI tomogram. In addition, the final model misfits of the FWI and MPI tomograms with respect to the true velocity model are 11.1% and 10.1%, respectively.

Marine Streamer Data

The MPI method is applied to a streamer data set recorded by a marine survey, which employed 515 shots with a shot interval of 37.5 m. The time-sampling interval is 2 ms, the recording time is 10 s, and there are 480 hydrophones per shot. The hydrophone interval is 12.5 m, with the minimum and maximum source-receiver offsets of 198 m and 6 km, respectively.

Prior to applying FWI and MPI to this marine data set, we apply four data preprocessing steps:

- The observed data are corrected from 3D to 2D format by scaling the amplitude by $\sqrt{t}$ to approximate geometrical spreading. A phase correction is applied by multiplying the spectrum of the observed seismogram with the filter $\sqrt{i/\omega}$ (Zhou et al., 1995).
- Mute the noise before the first arrivals.
- Re-sample the data with the new sampling interval of 0.5 ms.
- The source wavelet is estimated by stacking the aligned water-bottom reflections, and

the resulting wavelet is shown in Figure ??a and the spectrum is shown in Figure ??b.

The frequency band ranges from 1 Hz to 45 Hz.

The original data are filtered into different frequency bands with the different center filters displayed in table ?. The low-frequency data are inverted in the early iterations, and then higher frequency data are gradually incorporated into the later iterations. The $v(z)$ velocity model shown in Figure ??a is used as the initial model for the multiscale FWI and MPI. The initial model is discretized into 401×3006 gridpoints with a grid spacing of 6.25 m in both directions.

Figures ??b and ??c depict the FWI tomogram and MPI tomogram, respectively. Both the FWI and MPI tomograms have a higher resolution compared with the initial velocity model. In addition, the resolution of the MPI tomogram is slightly higher than that seen in the FWI tomogram. In order to verify the reconstructed FWI and MPI tomograms, we compare the migration images and angle-domain common image gathers (ADCIGs).

The original data are migrated using reverse time migration (RTM) based on the same initial velocity model, where the FWI tomogram, the MPI tomogram, and the RTM images are shown in Figures ??a, ??b, and ??c, respectively. Using either the FWI or MPI tomograms as the migration velocity, the resulting migration images are better focused than those obtained by using the initial velocity model. The magnified views in Figure ?? show that the reflectivities in the RTM image computed from the FWI tomogram might be considered more continuous than those computed from the MPI tomogram in box ‘A’, while the reflectors in the RTM images computed from the MPI tomogram are more continuous than those computed from the FWI tomogram in boxes ‘B’ and ‘C’.

tomogram are shown in Figures ??a, ??b and ??c, respectively. The magnified views of ADCIGs in boxes ‘A’ and ‘B’ are shown in Figures ??a and ??b, respectively. We can see that the ADCIGs associated with the MPI tomogram are slightly flatter than those from the FWI tomogram.

DISCUSSION AND CONCLUSIONS

The multiscale strategy temporally integrates the traces several times to boost the low-frequency component of the seismograms. The MPI procedure avoids predicting amplitudes and finds the velocity tomogram that best models the phases of the observed data. Thus, MPI is similar to wave-equation travelttime inversion without having to pick traveltimes. In fact, we show that MPI reduces to that of wave equation reflection travelttime inversion under the small phase assumption. This assumes that we are exclusively using the tomographic component of the RTM operator (Mora, 1989; Tang et al., 2013) to backproject the phases of the reflection residuals. In this case the gradient takes the form of a rabbit ear for a single source-receiver pair (Schuster, 2017). For the examples in our paper we employ the wavepaths for transmission (cigar-like gradient), reflection (rabbit-ear gradient), and migration (ellipse-like gradient).

Synthetic examples show that both the MPI and FWI methods can obtain similar tomograms when the initial velocity model is close to the true model. However, for our example, the MPI method provides a more accurate tomogram than FWI when the initial model is far from the true model. We also show examples where MPI generates more accurate tomograms than FWI when the input traces are viscoacoustic or elastic data from the Marmousi model. Synthetic examples suggest that MPI can be more robust than FWI for inverting seismic data.

In the marine data case, both the FWI and MPI methods successfully invert the data to obtain tomograms that are more accurate than the initial velocity model. However, the MPI tomogram has modestly higher resolution than that of the FWI tomogram. Comparing the RTM images, the MPI tomogram appears to be moderately more accurate than the FWI tomogram. The limitation of MPI is that it does not account for the amplitude spectrum in the data, which can modestly lower its resolution capabilities. Therefore, if the modeling algorithm perfectly captures the actual amplitudes in the data then the MPI tomogram will be a good starting model for FWI.

ACKNOWLEDGMENTS

The research reported in this publication was supported by the King Abdullah University of Science and Technology (KAUST) in Thuwal, Saudi Arabia. We are grateful to the sponsors of the Center for Subsurface Imaging and Modeling (CSIM) Consortium for their financial support. For computer time, this research used the resources of the Supercomputing Laboratory at KAUST and the IT Research Computing Group. We thank them for providing the computational resources required for carrying out this work. We greatly appreciate the constructive comments and suggestions from two anonymous reviewers, which helped improve this paper.

APPENDIX A: MPI $\approx$ WT

We will now show that the MPI misfit function reduces to a scaled replica of the wave equation travelttime misfit function for small perturbations in the phase. For each event, the scaling factor is the magnitude spectrum of the observed trace.

MPI Misfit Function for a Single Reflection/Trace

The starting point is the MPI misfit function

$$\epsilon = \frac{1}{2} \sum_{\omega} \sum_s \sum_g |D(\mathbf{g}|\mathbf{s})|^2 |e^{i\phi(\mathbf{s}|\mathbf{g})} - e^{i\phi(\mathbf{s}|\mathbf{g})^{obs}}|^2, \quad (\text{A-1})$$

where $|D(\mathbf{g}|\mathbf{s})| = |D(\mathbf{g}|\mathbf{s})^{obs}|$, and we initially assume only one reflection event in the trace. For small perturbations $|\delta\phi| \ll 1$ in the calculated phase $\phi = \phi^{obs} + \delta\phi$, the exponential difference in equation A-1 becomes

$$\begin{aligned} e^{i\phi(\mathbf{s}|\mathbf{g})} - e^{i\phi(\mathbf{s}|\mathbf{g})^{obs}} &= e^{i\phi(\mathbf{s}|\mathbf{g})^{obs}} (\cos \delta\phi(\mathbf{s}|\mathbf{g}) + i \sin \delta\phi(\mathbf{s}|\mathbf{g}) - 1), \\ &\approx i e^{i\phi(\mathbf{s}|\mathbf{g})^{obs}} \delta\phi(\mathbf{s}|\mathbf{g}). \end{aligned} \quad (\text{A-2})$$

If the source wavelet is zero phase and the trace only has one reflection event such that $\delta\phi(\mathbf{s}|\mathbf{g}) = \omega\delta\tau_{sg}$, then plugging equation A-2 into equation A-1 gives

$$\epsilon \approx \frac{1}{2} \sum_{\omega} \omega^2 \sum_s \sum_g |D(\mathbf{g}|\mathbf{s})|^2 \delta\tau_{sg}^2, \quad (\text{A-3})$$

where $\delta\tau_{sg} = \tau_{sg} - \tau_{sg}^{obs}$ is the travelttime difference between the predicted and observed reflections for the source at $\mathbf{s}$ and geophone at $\mathbf{g}$. For the initial iterations of MPI the data

are low-boosted so the traces are dominated by the low-frequency components where ω is very small compared to the traveltime residuals. Thus, the assumption that $|\omega\delta\tau_{sg}| \ll 1$ is satisfied at the early iterations even for large traveltime residuals. This supports our empirical results that suggest MPI has robust convergence properties because it is a type of traveltime inversion.

Equation A-3 is the weighted misfit function similar to those for the wave equation traveltime inversion methods described by Schuster and Quintus-Bosz (1993). For example, equation A-3 reduces exactly to that of ray-based reflection tomography if $\omega^2|D(\mathbf{g}|\mathbf{s})| = 1$ and ray tracing is used to compute the Fréchet derivative $\partial\tau/\partial c(\mathbf{x})$. If $|D(\mathbf{g}|\mathbf{s})| = 1$ and the wave equation is used to compute $\partial\tau/\partial c(\mathbf{x})$, then this is the misfit function for wave equation phase inversion (Schuster, 1991) in the space-frequency domain or the wave equation traveltime inversion method (Luo and Schuster, 1991) in the space-time domain. If the wavepath eikonal equation is used to predict the traveltimes in the phase-based Fréchet derivative of (Schuster and Quintus-Bosz, 1993) then this is known as eikonal traveltime inversion. Similarly, selecting $D(\mathbf{g}|\mathbf{s})$ to be consistent with the Rytov equation reduces the misfit to that of Rytov inversion (Woodward, 1992). In summary, MPI is similar to these classical traveltime- or phase-inversion methods where the only difference is the choice of the weighting function and how the Fréchet derivative $\partial\tau/\partial c(\mathbf{x})$ is computed.

MPI Misfit Function for Multiple Reflections/Trace

To extend this unification of wave equation traveltime inversion methods to traces that consist of a multitude of reflections, assume that $|D(\mathbf{g}|\mathbf{s})|e^{i\phi(\mathbf{g}|\mathbf{s})}$ can be replaced by

$$|D(\mathbf{g}|\mathbf{s})|e^{i\phi(\mathbf{g}|\mathbf{s})} = \sum_{j=1}^N |D(\mathbf{g}|\mathbf{s})_j|e^{i\phi(\mathbf{g}|\mathbf{s})_j}, \quad (\text{A-4})$$

where $|D(\mathbf{g}|\mathbf{s})_j|e^{i\phi(\mathbf{g}|\mathbf{s})_j}$ is the spectrum of the j^{th} reflection, and there are N reflections in a trace. Assuming that the predicted and observed magnitude spectra are the same as one another then, for small $|\delta\phi(\mathbf{s}|\mathbf{g})_j|$, the residual $r_{sg} = D(\mathbf{g}|\mathbf{s}) - D(\mathbf{g}|\mathbf{s})^{obs}$ becomes

$$\begin{aligned} r_{sg} &= \sum_{j=1}^N |D(\mathbf{g}|\mathbf{s})_j|(e^{i\phi(\mathbf{g}|\mathbf{s})_j} - e^{i\phi(\mathbf{g}|\mathbf{s})_j^{obs}}), \\ &\approx i \sum_{j=1}^N \overbrace{|D(\mathbf{g}|\mathbf{s})_j|e^{i\phi(\mathbf{g}|\mathbf{s})_j^{obs}}}^{D(\mathbf{g}|\mathbf{s})_j^{obs}} \overbrace{\delta\phi(\mathbf{s}|\mathbf{g})_j}^{\omega\delta\tau(\mathbf{g}|\mathbf{s})_j}, \\ &= i\omega \sum_{j=1}^N D(\mathbf{g}|\mathbf{s})_j^{obs} \delta\tau(\mathbf{s}|\mathbf{g})_j, \end{aligned} \quad (\text{A-5})$$

where $\delta\tau(\mathbf{g}|\mathbf{s})_j$ is the traveltime residual for the j^{th} reflection. For traces with N reflections, the misfit function $\epsilon = \frac{1}{2} \sum_{\omega} \sum_s \sum_g |r_{sg}|^2$ becomes

$$\epsilon = \frac{1}{2} \sum_{\omega} \omega^2 \sum_s \sum_g \sum_{j=1}^N |D(\mathbf{g}|\mathbf{s})_j^{obs}|^2 \delta\tau(\mathbf{s}|\mathbf{g})_j^2 + \text{cross-talk}, \quad (\text{A-6})$$

where the cross-talk term is

$$\text{cross-talk} = \frac{1}{2} \sum_{\omega} \omega^2 \sum_{s=1}^{N_s} \sum_{g=1}^{N_g} \sum_j \sum_{k \neq j} D(\mathbf{g}|\mathbf{s})_j D(\mathbf{g}|\mathbf{s})_k^* \delta\tau(\mathbf{s}|\mathbf{g})_j \delta\tau(\mathbf{s}|\mathbf{g})_k. \quad (\text{A-7})$$

The cross-talk can be negligible compared to the summation over the $|D(\mathbf{g}|\mathbf{s})_j^{obs}|^2 \delta\tau(\mathbf{s}|\mathbf{g})_j^2$ terms if the crosstalk phase for each pair of j, k indices is a random variable (Schuster et al., 2011) and $N, N_g, N_s \gg 0$ are large. In this case the misfit function reduces to that for wave equation reflection tomography:

$$\epsilon \approx \frac{1}{2} \sum_{\omega} \omega^2 \sum_s \sum_g \sum_{j=1}^N |D(\mathbf{g}|\mathbf{s})_j^{obs}|^2 \delta\tau(\mathbf{s}|\mathbf{g})_j^2, \quad (\text{A-8})$$

as if the reflection traveltimes have been picked. However, in the case of MPI no such picking is needed. If all of the events have the same signal to noise ratio, then the weighting of each event should be identical, which suggests that a normalized deconvolution or an AGC should be applied to the trace prior to inversion.

To completely eliminate the crosstalk terms one can window about the trajectory of reflection arrivals from the same reflector. Then the misfit function can be reformulated as a sum of the misfit functions for each reflector (Maharromov et al., 2017). This might expedite the convergence of the algorithm to that of a rapid form of reflection traveltime tomography.

Exact MPI Misfit Function

The squared magnitude of the phase residual in equation A-2 is

$$\begin{aligned} |e^{i\phi(\mathbf{s}|\mathbf{g})} - e^{i\phi(\mathbf{s}|\mathbf{g})^{obs}}|^2 &= |(\cos \delta\phi(\mathbf{s}|\mathbf{g}) + i \sin \delta\phi(\mathbf{s}|\mathbf{g}) - 1)|^2, \\ &= 2 - 2 \cos \delta\phi, \\ &= 4 \sin^2(\delta\phi/2). \end{aligned} \quad (\text{A-9})$$

Assume $|D(\omega)|^2 = 1/\omega^2$ for $-\omega_o \leq \omega \leq \omega_o$, otherwise $|D(\omega)| = 0$. The spectrum $|D(\omega)|$ divided by $i\omega$ is equivalent to integrating the original trace once, so $|D|^2 = 1/\omega^2$ is proportional to the squared magnitude of the integrated trace's spectrum.

Inserting $|D|^2 = 1/\omega^2$ and equation A-9 into equation A-1 gives

$$\begin{aligned}\epsilon &= 2 \sum_s \sum_g \int_{-\omega_o}^{\omega_o} d\omega \frac{\sin^2(\omega \delta\tau(\mathbf{g}|\mathbf{s})/2)}{\omega^2}, \\ &= \sum_s \sum_g \frac{[\tau(\mathbf{g}|\mathbf{s})]^2}{2} \int_{-\omega_o}^{\omega_o} d\omega \frac{\sin^2(\omega \delta\tau(\mathbf{g}|\mathbf{s})/2)}{[\omega \tau(\mathbf{g}|\mathbf{s})/2]^2}.\end{aligned}\tag{A-10}$$

A change in variables $\omega \delta\tau(\mathbf{g}|\mathbf{s})/2 \rightarrow x$ and $d\omega = 2dx/\delta\tau(\mathbf{g}|\mathbf{s})$ transforms equation A-10 to

$$\epsilon = \sum_s \sum_g |\tau(\mathbf{g}|\mathbf{s})| \int_{-0.5\omega_o\tau(\mathbf{g}|\mathbf{s})}^{0.5\omega_o\tau(\mathbf{g}|\mathbf{s})} dx \frac{\sin^2 x}{x^2}.\tag{A-11}$$

For a sufficiently large $\omega_o \gg \delta\tau$ the integration approximates the constant α so that the misfit is that of a l_1 norm:

$$\epsilon = \alpha \sum_s \sum_g |\tau(\mathbf{g}|\mathbf{s})|.\tag{A-12}$$

We conclude that for a single event with a wideband flat spectrum that the MPI misfit function is that of the l_1 norm of the traveltime residual. It is not realistic for a source to have DC components.

However, if the source wavelet is a time derivative of the delta function and damping is applied when integrating the traces, the misfit function resembles the L_1 -norm of the traveltime residuals when the residuals are not excessively large. For source wavelets that have a finite width, the misfit function will behave as a L_2 -norm of the traveltime residuals

on the order of, or smaller than, the wavelet width and will transition into an L_1 -norm for larger residuals. The misfit function will eventually lose sensitivity to larger traveltimes residuals due to the need of damping inevitable in realistic trace processing.

MPI and Recent FWI Methods

The MPI misfit function under the small phase approximation at low frequencies can be compared to some recent wave equation inversion methods.

- Full Waveform Inversion with a Bump Functional. For the waveform inversion method of Bharadwaj et al. (2016), each event in the seismic record is replaced with a positive-valued Gaussian function. These Gaussian functions can have nearly the same shape and amplitudes if the trace is gained so that roughly $|D(\mathbf{g}|\mathbf{s})_j| \approx |D(\mathbf{g}|\mathbf{s})_j^{obs}|$ in equation A-4, except $D(\mathbf{g}|\mathbf{s})_j$ is replaced by a positive Gaussian with the arrival time of the j^{th} reflection. Their multiscale strategy is to low-pass filter the data, but they do not low-boost it by integrating the traces. Forcing all amplitudes to be positive does not help in reducing the value of the crosstalk term in equation A-6.
- Adaptive Waveform Inversion. Warner and Guasch (2016) developed a FWI method where an adaptive matching filter is used to fit both the amplitude and phase of the predicted arrivals to the observed arrivals at each iteration. If the adaptive filter is constrained to just fit the magnitude spectrums then we can say $|D(\mathbf{g}|\mathbf{s})_j^{pred}| \approx |D(\mathbf{g}|\mathbf{s})_j|$ so that the squared residual is $|D(\mathbf{g}|\mathbf{s})_j - D(\mathbf{g}|\mathbf{s})_j^{obs}|^2 = |D(\mathbf{g}|\mathbf{s})_j|^2 |e^{i\phi(\mathbf{g}|\mathbf{s})_j} - e^{i\phi(\mathbf{g}|\mathbf{s})_j^{obs}}|^2$. Thus, the misfit function reduces to that of equation A-6 under the small phase approximation.

- Full Traveltime Inversion. Luo et al. (2016) developed a wave equation method that

inverts the traveltimes of reflection events. In their case they used a gradient method based on the Rytov approximation while the Fréchet derivative employed in the MPI scheme is based on the Born approximation. Their scheme employs a multiscale strategy where the early iterations are applied to low-frequency portions of the data, and later iterations are applied to data where the higher frequencies are added into the input data.

APPENDIX B: REFLECTION MPI

Mora (1989) showed that the gradient for the FWI misfit function can be decomposed into three different sub-gradients that take the shapes in model space of an ellipse, a cigar, and a pair of rabbit ears. The tomographic update to the velocity model refers to the rabbit-ear gradient that updates the long-wavelength component of the velocity model by smearing the reflection residuals along the associated rabbit-ear wavepaths. In contrast, the high-wavenumber part of the velocity model is updated by the ellipse-like gradient with foci at the source and receiver locations. Finally, the cigar-like gradient updates the long-wavelength part of the velocity model at relatively shallow depths by smearing the transmission residuals along cigar-like wavepaths. These gradients are weighted differently so it is often necessary to separate the gradients from one another and initially update the weakest one, the tomographic one with the gradient shape of a pair of rabbit ears. We now compute the tomographic gradient using the method described by Chi et al. (2017). Here, the the total field is decomposed into upgoing and downgoing parts (Fei et al., 2015), and we only correlate upgoing source fields with the backpropagated reflections to get the rabbit-ear portion of the gradient. However, we now compute the tomographic gradients with both the FWI and the MPI methods. The goal is to show that the tomographic MPI gradient is less susceptible to amplitude errors and incorrect modeling of physics than velocity updates computed by the FWI tomographic gradient.

Figure ??a shows the acoustic velocity model, which is discretized into 201×101 grid-points with a node spacing of 10 m. The upper layer velocity is 2.0 km/s, the bottom layer velocity is 2.5 km/s, and the velocity inside the circle is 1.8 m/s. There are 26 sources and 201 receivers evenly distributed on the surface, with a source interval of 80 m and a receiver

interval of 10 m. The peak frequency of the source wavelet is 15 Hz and the initial velocity model is shown in Figure ??b.

Figure ??c is the MPI tomogram after 50 iterations, which only reconstructs the high-wavenumber parts of the velocity model. The low-wavenumber parts are not estimated because the tomographic gradient is an order of magnitude weaker than the gradients associated with cigar-like and ellipse-like gradients (Schuster, 2017). However, after 50 iterations the MPI tomogram associated with the tomographic gradient is shown in Figure ??d and, as expected, reconstructs the low-wavenumber parts of the velocity model. This tomogram is denoted as the T-MPI tomogram, and the associated FWI tomogram (not shown) is denoted as T-FWI tomogram.

The input data are now computed by a finite-difference solution to the visco-acoustic wave equation, where the Q values for the model are shown in Figure ??a. The initial velocity model is the same as the one in Figure ??b, and the low-wavenumber parts of the velocity model are updated using the tomographic gradient. The T-FWI and T-MPI tomograms after 50 iterations are shown in Figures ??b and ??d, respectively. Figure ??c shows the velocity comparisons for profiles $A - A'$ and $B - B'$. It is obvious that the T-MPI tomogram is moderately more accurate than the T-FWI tomogram because MPI is less sensitive to lossy input data than FWI.

REFERENCES

Alford, R., K. Kelly, and D. M. Boore, 1974, Accuracy of finite-difference modeling of the acoustic wave equation: *Geophysics*, **39**, 834–842.

Baumstein, A., W. Ross, and S. Lee, 2011, Simultaneous source elastic inversion of sur-

- face waves: Presented at the 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC 2011.
- Bharadwaj, P., W. Mulder, and G. Drijkoningen, 2016, Full waveform inversion with an auxiliary bump functional: *Geophysical Journal International*, **206**, 1076–1092.
- 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.
- Chi, B., K. Gao, and L. Huang, 2017, Least-squares reverse time migration guided full-waveform inversion, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 1471–1475.
- De Hoop, M. V., and R. D. van Der Hilst, 2005, On sensitivity kernels for wave-equation transmission tomography: *Geophysical Journal International*, **160**, 621–633.
- Fei, T. W., Y. Luo, J. Yang, H. Liu, and F. Qin, 2015, Removing false images in reverse time migration: The concept of de-primary: *Geophysics*, **80**, S237–S244.
- Hole, J., 1992, Nonlinear high-resolution three-dimensional seismic travel time tomography: *Journal of Geophysical Research: Solid Earth*, **97**, 6553–6562.
- Levander, A. R., 1988, Fourth-order finite-difference p-sv seismograms: *Geophysics*, **53**, 1425–1436.
- Luo, Y., Y. Ma, Y. Wu, H. Liu, and L. Cao, 2016, Full-traveltime inversion: *Geophysics*, **81**, R261–R274.
- Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: *Geophysics*, **56**, 645–653.

- Maharramov, M., A. I. Baumstein, Y. Tang, P. S. Routh, S. Lee, and S. K. Lazaratos, 2017, Time-domain broadband phase-only full-waveform inversion with implicit shaping, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 1297–1301.
- Masoni, I., R. Brossier, J. Virieux, and J. Boelle, 2013, Alternative misfit functions for FWI applied to surface waves: Presented at the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013.
- Min, D.-J., and C. Shin, 2006, Refraction tomography using a waveform-inversion back-propagation technique: *Geophysics*, **71**, R21–R30.
- Mora, P., 1989, Inversion= migration+ tomography: *Geophysics*, **54**, 1575–1586.
- Nemeth, T., E. Normark, and F. Qin, 1997, Dynamic smoothing in crosswell traveltime tomography: *Geophysics*, **62**, 168–176.
- Operto, S., J. Virieux, P. Amestoy, J.-Y. L'Excellent, L. Giraud, and H. B. H. Ali, 2007, 3d finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study: *Geophysics*, **72**, SM195–SM211.
- Pratt, R. G., and N. R. Goulty, 1991, Combining wave-equation imaging with traveltime tomography to form high-resolution images from crosshole data: *Geophysics*, **56**, 208–224.
- Routh, P., J. Krebs, S. Lazaratos, A. Baumstein, S. Lee, Y. H. Cha, I. Chikichev, N. Downey, D. Hinkley, and J. Anderson, 2011, Encoded simultaneous source full-wavefield inversion for spectrally shaped marine streamer data, *in* SEG Technical Program Expanded Abstracts 2011: Society of Exploration Geophysicists, 2433–2438.
- Sava, P., and I. Vlad, 2008, Numeric implementation of wave-equation migration velocity analysis operators: *Geophysics*, **73**, VE145–VE159.
- Sava, P. C., B. Biondi, and J. Etgen, 2005, Wave-equation migration velocity analysis by

focusing diffractions and reflections: *Geophysics*, **70**, U19–U27.

Schuster, G., 2017, *Seismic inversion*: Society of Exploration Geophysicists.

Schuster, G. T., 1991, Wave-equation phase inversion in the frequency domain, *in* SEG Technical Program Expanded Abstracts 1991: Society of Exploration Geophysicists, 909–912.

Schuster, G. T., and A. Quintus-Bosz, 1993, Wavepath eikonal traveltime inversion: Theory: *Geophysics*, **58**, 1314–1323.

Schuster, G. T., X. Wang, Y. Huang, W. Dai, and C. Boonyasirawat, 2011, Theory of multi-source crosstalk reduction by phase-encoded statics: *Geophysical Journal International*, **184**, 1289–1303.

Shin, C., and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: *Geophysics*, **71**, R31–R42.

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

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

Sun, Y., and G. T. Schuster, 1993, Time-domain phase inversion, *in* SEG Technical Program Expanded Abstracts 1993: Society of Exploration Geophysicists, 684–687.

Tang, Y., S. Lee, A. Baumstein, and D. Hinkley, 2013, Tomographically enhanced full wavefield inversion, *in* SEG Technical Program Expanded Abstracts 2013: Society of Exploration Geophysicists, 1037–1041.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.

Tieman, H. J., 1995, Migration velocity analysis: Accounting for the effects of lateral ve-

locity variations: *Geophysics*, **60**, 164–175.

Van Leeuwen, T., and W. Mulder, 2008, Velocity analysis based on data correlation: *Geophysical Prospecting*, **56**, 791–803.

———, 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.

Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: Theory: *Geophysics*, **81**, R429–R445.

Woodward, M. J., 1992, Wave-equation tomography: *Geophysics*, **57**, 15–26.

Zhang, Y., and D. Wang, 2009, Traveltime information-based wave-equation inversion: *Geophysics*, **74**, WCC27–WCC36.

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.

LIST OF FIGURES

Figure 1: The original data without integration, a1) calculated, a2) observed, a3) modified-calculated, and a4) modified-observed traces. The traces with one ($n = 1$) integration are in b1 - b4, and the double ($n = 2$) integration results are in c1 - c4.

Figure 2: a) The true model, b) initial model, and c) FWI gradient. The MPI gradients with d) $n = 2$, e) $n = 1$, and f) $n = 0$ integrations of traces. The white dashed lines indicate the layer interfaces.

Figure 3: The multiscale phase inversion workflow. Here, n is the integration number, N_{max} is the maximum integration number, ϵ_0 and ϵ_1 are the predefined thresholds.

Figure 4: a) The Marmousi velocity and b) Q models used for computing the visco-acoustic data.

Figure 5: Synthetic seismic data: a) constant-density acoustic data, b) visco-acoustic data, and c) elastic data for the shot location at $x = 0$ km. The data are plotted with the same amplitude scale.

Figure 6: Inversion results for acoustic data. a) The smoothed initial velocity model with a velocity error of 12%, c) multiscale FWI and e) MPI tomograms based on a); b) the $v(z)$ initial velocity model with the velocity error of 22%, d) multiscale FWI and f) MPI tomograms based on b).

Figure 7: Velocity profiles for the true, initial, FWI and MPI velocity $v(z)$ values for a) smoothed and b) $v(z)$ initial-velocity models. Here, the black, blue, green, and red curves represent the true, initial, FWI and MPI velocity values, respectively.

Figure 8: Normalized data misfit at different iterations, for the multiscale a) FWI and b) MPI methods, where the red, blue and black curves represent the input data with twice, once, and no integrations, respectively.

Figure 9: Inversion results for visco-acoustic and elastic data. a) The FWI and c) MPI tomograms for visco-acoustic data, and b) the FWI and d) MPI tomograms for elastic data.

Figure 10: Magnified views of the dashed box ‘A’ in Figure ?? for the a) true model, c) FWI and e) MPI tomograms; Magnified views of the dashed box ‘B’ in Figure ?? for the b) true model, d) FWI and f) MPI tomograms. The MPI tomograms more closely resemble the Marmousi model compared with the FWI tomograms.

Figure 11: Velocity profile $v(z)$ extracted from the FWI and MPI tomograms for a) visco-acoustic and b) elastic data. Here, the black, blue, green, and red curves represent the true, initial, FWI and MPI velocity values, respectively.

Figure 12: a) The estimated source wavelet and b) the spectrum of the source wavelet.

Figure 13: Inversion results from the marine streamer data. a) The $v(z)$ initial model, b) multiscale FWI, and c) multiscale MPI tomogram.

Figure 14: RTM migration images from the marine streamer data that uses the migration velocity from a) the initial model, b) FWI tomogram, and c) MPI tomogram.

Figure 15: Magnified views of the boxes ‘A’, ‘B’ and ‘C’ in Figure ??.

Figure 16: Angle domain common image gathers (ADCIGs) from -55° to 0° based on a) the initial model, b) FWI tomogram, and c) MPI tomogram.

Figure 17: Magnified views of the boxes ‘A’ and ‘B’ in Figure ??.

Figure B-1: a) Acoustic velocity model, b) starting velocity model for MPI, c) MPI tomogram after 50 iterations using the total gradient to update the velocity model, and d) T-MPI tomogram after 50 iterations using the rabbit-ear-like gradient.

Figure B-2: Results for visco-acoustic data, a) the Q model, and the corresponding velocity model are shown in Figure ??a, b) T-FWI tomogram after 50 iterations, c) velocity comparisons for profiles $A - A'$ and $B - B'$, where the black, blue, green and red curves

are velocities for the true model, initial model, T-FWI and T-MPI tomograms, respectively, and d) T-MPI tomogram after 50 iterations.

Figure 1: The original data without integration, a1) calculated, a2) observed, a3) modified-calculated, and a4) modified-observed traces. The traces with one $(n=1)$ integration are in b1 - b4, and the double $(n=2)$ integration results are in c1 - c4.

139x118mm (300 x 300 DPI)

Figure 2: a) The true model, b) initial model, and c) FWI gradient. The MPI gradients with d) $n=2$, e) $n=1$, and f) $n=0$ integrations of traces. The white dashed lines indicate the layer interfaces.

118x80mm (300 x 300 DPI)

Figure 3: The multiscale phase inversion workflow. Here, n is the integration number, $N_{\max}$ is the maximum integration number, ϵ_0 and ϵ_1 are the predefined thresholds.

111x90mm (300 x 300 DPI)

Figure 4: a) The Marmousi velocity and b) Q models used for computing the visco-acoustic data.

95x99mm (300 x 300 DPI)

Figure 5: Synthetic seismic data: a) constant-density acoustic data, b) visco-acoustic data, and c) elastic data for the shot location at $x=0$ km. The data are plotted with the same amplitude scale.

114x67mm (300 x 300 DPI)

Figure 6: Inversion results for acoustic data. a) The smoothed initial velocity model with $\{a\}$ velocity error of 12%, c) multiscale FWI and e) MPI tomograms based on a); b) the $\{v(z)\}$ initial velocity model with the velocity error of 22%, d) multiscale FWI and f) MPI tomograms based on b).

143x116mm (300 x 300 DPI)

Figure 7: Velocity profiles for the true, initial, FWI and MPI velocity $v(z)$ values for a) smoothed and b) $v(z)$ $\textcolor{black}{\text{initial-velocity}}$ models. Here, the black, blue, green, and red curves represent the true, initial, FWI and MPI velocity values, respectively.

112x86mm (300 x 300 DPI)

Figure 8: Normalized data misfit at different iterations, for the multiscale a) FWI and b) MPI methods, where the red, blue and black curves represent the input data with twice, once, and no integrations, respectively.

129x116mm (300 x 300 DPI)

Figure 9: Inversion results for visco-acoustic and elastic data. a) The FWI and c) MPI tomograms for visco-acoustic data, and b) the FWI and d) MPI tomograms for elastic data.

104x56mm (300 x 300 DPI)

Figure 10: Magnified views of the dashed box 'A' in Figure~\ref{tomogram_visco_e} for the a) true model, c) FWI and e) MPI tomograms; Magnified views of the dashed box 'B' in Figure~\ref{tomogram_visco_e} for the b) true model, d) FWI and f) MPI tomograms. The MPI tomograms more closely resemble the Marmousi model compared with the FWI tomograms.

128x95mm (300 x 300 DPI)

Figure 11: Velocity profile $v(z)$ extracted from the FWI and MPI tomograms for a) visco-acoustic and b) elastic data. Here, the black, blue, green, and red curves represent the true, initial, FWI and MPI velocity values, respectively.

112x86mm (300 x 300 DPI)

Figure 12: a) The estimated source wavelet and b) the spectrum of the source wavelet.

137x131mm (300 x 300 DPI)

Figure 13: Inversion results from the marine streamer data. a) The $v(z)$ initial model, b) multiscale FWI, and c) multiscale MPI tomogram.

142x162mm (300 x 300 DPI)

Figure 14: RTM migration images from the marine streamer data that uses the migration velocity from a) the initial model, b) FWI tomogram, and c) MPI tomogram.

142x160mm (300 x 300 DPI)

Figure 15: Magnified views of the boxes 'A', 'B' and 'C' in Figure~\ref{fig9}.

144x131mm (300 x 300 DPI)

Figure 16: Angle domain common image gathers (ADCIGs) from -55° to 0° based on a) the initial model, b) FWI tomogram, and c) MPI tomogram.

141x159mm (300 x 300 DPI)

Figure 17: Magnified views of the boxes 'A' and 'B' in Figure~ \ref{adcig}.

65x27mm (300 x 300 DPI)

Figure B-1: a) Acoustic velocity model, b) starting velocity model for MPI, c) MPI tomogram after 50 iterations using the total gradient to update the velocity model, and d) T-MPI tomogram after 50 iterations using the rabbit-ear-like gradient.

111x62mm (300 x 300 DPI)

Figure B-2: Results for visco-acoustic data, a) the Q model, and the corresponding velocity model are shown in Figure~\ref{Tomo_TMPI}a, b) T-FWI tomogram after 50 iterations, c) velocity comparisons for profiles \$A-A'\$ and \$B-B'\$, where the black, blue, green and red curves are velocities for the true model, initial model, T-FWI and T-MPI tomograms, respectively, and d) T-MPI tomogram after 50 iterations.

110x62mm (300 x 300 DPI)