

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE

10.1002/2017JB015207

Key Points:

- A new method has been proposed to separate the migration and tomography modes of waveform inversion for the study of seismology
- The method has been verified through theory analyses and numerical examples
- This new method has been successfully applied on field data sets

Correspondence to:

G. Yao,
gyao@imperial.ac.uk

Citation:

Yao, G., da Silva, N. V., Warner, M., & Kalinicheva, T. (2018). Separation of migration and tomography modes of full-waveform inversion in the plane wave domain. *Journal of Geophysical Research: Solid Earth*, 123, 1486–1501.
<https://doi.org/10.1002/2017JB015207>

Received 7 NOV 2017

Accepted 15 JAN 2018

Accepted article online 20 JAN 2018

Published online 7 FEB 2018

Separation of Migration and Tomography Modes of Full-Waveform Inversion in the Plane Wave Domain

Gang Yao¹ , Nuno V. da Silva¹, Michael Warner¹, and Tatiana Kalinicheva¹

¹Department of Earth Science and Engineering, Imperial College London, London, UK

Abstract Full-waveform inversion (FWI) includes both migration and tomography modes. The migration mode acts like a nonlinear least squares migration to map model interfaces with reflections, while the tomography mode behaves as tomography to build a background velocity model. The migration mode is the main response of inverting reflections, while the tomography mode exists in response to inverting both the reflections and refractions. To emphasize one of the two modes in FWI, especially for inverting reflections, the separation of the two modes in the gradient of FWI is required. Here we present a new method to achieve this separation with an angle-dependent filtering technique in the plane wave domain. We first transform the source and residual wavefields into the plane wave domain with the Fourier transform and then decompose them into the migration and tomography components using the opening angles between the transformed source and residual plane waves. The opening angles close to 180° contribute to the tomography component, while the others correspond to the migration component. We find that this approach is very effective and robust even when the medium is relatively complicated with strong lateral heterogeneities, highly dipping reflectors, and strong anisotropy. This is well demonstrated by theoretical analysis and numerical tests with a synthetic data set and a field data set.

1. Introduction

Full-waveform inversion (FWI) is an effective technique for the recovery of Earth properties from seismic data. It is being used widely for both oil and gas exploration (e.g., Cheng et al., 2016, 2017; Silva et al., 2016; Virieux & Operto, 2009; Warner et al., 2013) and global seismology (Chen et al., 2015; Tao et al., 2017; Zhu et al., 2012). It is achieved by optimizing an objective function constrained by a wave equation (Tarantola, 1984). Recent advances in computer hardware and numerical modeling algorithms (Y. Wang et al., 2014; Yao et al., 2016) have significantly reduced the computational burden of FWI.

Theoretical analyses suggest that FWI includes two modes: tomography and migration (Mora, 1989). Successful examples of using the tomographic mode of FWI to improve background models all seem to be related to data with large offsets, which are rich in refractions and wide angle reflections. However, conventional acquisition geometries, for example, towed streamer, are dominated by reflections. Applying standard FWI on the reflections from this type of data sets is essentially equivalent to carrying out nonlinear least squares migration (e.g., D. Wu et al., 2016; Yao & Jakubowicz, 2016; Yao & Wu, 2015). Thus, the migration mode of FWI dominates the inversion. Although the inversion also includes the tomographic mode, it is much weaker than the migration mode (Yao et al., 2014).

To emphasize one of the modes, it is necessary to separate the two modes of FWI. There are several existing methods to achieve this goal. Parameterizing the model into a long-wavelength background velocity and short-wavelength impedance (Zhou et al., 2015) or reflectivity (Yao & Warner, 2015; Yao & Wu, 2017) is an effective method to separate the two modes of FWI. This method generates two wavefields, one of which is the full wavefield modeled with background velocity and impedance or reflectivity, and the other of which is the background wavefield generated with background velocity only. The full wavefield includes both the incident wavefield and the scattering wavefield, while, ideally, the background wavefield includes only the incident wavefield. As a result, the subtraction of the two wavefields gives the scattering wavefields. Alternatively, Born-based modeling can also separate the scattering wavefield from the incident wavefield (Sun et al., 2016; Xu et al., 2012; Yao & Wu, 2017; Zhang et al., 2011). Cross-correlating the corresponding wavefields then can produce the gradient of long-wavelength background velocity and the gradient of short-wavelength impedance or reflectivity. The former is dominated by the tomography component,

while the latter is mainly the migration component. However, the scattering wavefield produced by the subtraction or Born-based modeling is not pure backward scattered energy but includes forward scattered energy as well. Furthermore, if the background velocity is not smooth enough, then the background wavefield includes scattering energy. As a result, this method may be unable to achieve a high-quality separation of the two modes.

Another method is to split the wavefields along the up and down, and along the left and right propagation directions, and then cross-correlate each one of these components separately (Irabor & Warner, 2016; F. Wang et al., 2013; Z. Wu & Alkhalifah, 2016). Similarly to the parameterization method, this method achieves the mode separation by separating the backward scattering wavefields from the incident wavefields. But it assumes the backward scattering waves propagate along a different direction to the incident waves. However, if only the left-right and/or up-down separation is applied, it might not be possible to distinguish the backward scattering waves from the incident waves. Consequently, the mode separation may be incomplete.

Alternatively, we can directly extract the components of the two modes from the normal gradient according to the opening angle of the source wavefield and the residual wavefield, which are cross-correlated to generate the gradient. The fundamental principle of this type of methods is that the waves forming the tomography component have an opening angle close to 180°, while the waves with other angles mainly contribute to the migration component. As a result, it is crucial to accurately estimate the opening angle.

The Poynting vectors determine the propagation directions of the wavefields; thus, they can be used in principle to measure the opening angle of the source wavefield and the residual wavefield (Xie et al., 2005). This measurement is accurate for events that are separate from each other. However, in places where multiple wave events overlap with each other, the Poynting vector points to the same direction for all the overlapping events. Thus, it is impossible to measure the opening angle of two events from the source and residual wavefields in these areas accurately, unless the two wavefields can be separated effectively with some sort of decomposition technique. Local slant stack is a way to decompose the overlapped waves into plane waves, which travel in different directions (apparent slowness), by applying slant stack inside the sliding windows. Consequently, the opening angle can be measured after decomposition (Xie, 2015). Generally, small size windows are chosen to balance the accuracy and computational cost. However, a small window is prone to give a low resolution.

The inverse scattering imaging condition is another method to separate the two components through angle filtering (Whitmore & Crawley, 2012). The method produces the gradient of bulk modulus and the gradient of density. Subtraction of the two gradients after scaling gives the tomography component enhanced gradient, while the summation of them produces the migration component enhanced gradient. The enhancement is equivalent to filtering with the cosine square of the opening angle, which has a range from 0 to 180°. Obviously, the filtering function with a wide transition band could lead to the leakage of one component to the other. To mitigate the leakage, Ramos-Martinez et al. (2016) used an adaptive weighting before the subtraction and summation.

Khalil et al. (2013) proposed a method to attenuate the low-wave number noise in reverse-time migration (RTM) with time-shift gathers under the assumption of isotropic velocity. Since the low-wave number noise of RTM is the tomography component in the gradient of FWI effectively, this method was then further used to separate the modes of FWI with angle filtering (Alkhalifah, 2015; Z. Wu & Alkhalifah, 2015; Z. Wu & Alkhalifah, 2017). This method has to apply a coordinate transform for adapting to variable velocity.

Herein, we propose an easy-to-implement and effective method to achieve this separation by an angle-dependent filtering technique in the plane wave domain. In our method, we apply a spatial Fourier transform to decompose the waves into plane waves and then filter the gradient by angle.

2. Methodology

The wave equation can be expressed symbolically as

$$\mathbf{A}\mathbf{p} = \mathbf{s}, \quad (1)$$

where A denotes the wave equation operator, \mathbf{p} is the wavefield, and \mathbf{s} is the source wavelet. Conventional FWI then minimizes an objective function ϕ

$$\phi = \frac{1}{2}(\mathbf{d} - \mathbf{d}_0)^T(\mathbf{d} - \mathbf{d}_0) = \frac{1}{2}\delta\mathbf{d}^T\delta\mathbf{d}, \quad (2)$$

where \mathbf{d} and \mathbf{d}_0 represent the predicted and observed data, respectively, and $\delta\mathbf{d}$ is the residual. The gradient of the objective function with respect to the model \mathbf{m} can then be derived as

$$\frac{\partial\phi}{\partial\mathbf{m}} = -\mathbf{p}^T\left(\frac{\partial A}{\partial\mathbf{m}}\right)^T A^{-T} D^T \delta\mathbf{d}, \quad (3)$$

where D is the picking matrix so that $\mathbf{d} = D\mathbf{p}$, \mathbf{p} is the source wavefield, and $A^{-T}D^T\delta\mathbf{d}$ is the residual wavefield. Equation (3) means that the model gradient in FWI is formed by the zero-lag cross correlation of the source wavefield and the back-propagated residual wavefield, scaled by $\frac{\partial A}{\partial\mathbf{m}}$. If the velocity of P waves is the inversion parameter, then the gradient is expressed as

$$\frac{\partial\phi}{\partial v} = g(\mathbf{x}) = \frac{2}{v^3} \sum_t \ddot{p}_s(\mathbf{x}, t) p_r(\mathbf{x}, t), \quad (4)$$

where \ddot{p}_s represents the second-order time derivative of the source wavefield and p_r indicates the back propagated residual wavefield.

Here we develop an angle filtering technique to isolate the migration and tomography components of FWI, which requires knowledge of the opening angle between the source wavefield and the residual wavefield. In the case of two plane waves, it is conceivable to use the respective wave vectors to define the opening angle. Thus, we first convert the wavefields from the space domain into the wave number domain by the Fourier transform. Applying the spatial Fourier transform on the wavefields along the spatial axes, \mathbf{x} , yields

$$\tilde{g}(\mathbf{k}) = \frac{2}{v^3} \sum_t \tilde{p}_s(\mathbf{k}_s, t) * \tilde{p}_r(\mathbf{k}_r, t) = \frac{2}{v^3} \sum_t \sum_i \sum_j \tilde{p}_s(\mathbf{k}_{si}, t) \tilde{p}_r(\mathbf{k}_{rj}, t), \quad (5)$$

where \tilde{p}_s and \tilde{p}_r denote the counterpart of \ddot{p}_s and p_r in the wave number domain, \mathbf{k}_s , \mathbf{k}_r , and \mathbf{k} are the wave number vectors of the source wavefield, the residual wavefield, and the gradient, i and j are the indices of the wave number of \mathbf{k}_s and \mathbf{k}_r , respectively, t is the time, and asterisk represents a convolution. The convolution in the wave number domain is the counterpart of the multiplication in the space domain. In this analysis, we assume that the velocity is constant, but the method is valid for variable velocity because we can take the inverse Fourier transformation after the summation over time and before the scaling with $\frac{2}{v^3}$. In the Fourier domain, each element represents a plane wave with a particular wave number, amplitude, and phase. The gradient is formed by multiplication and summation described in equation (5). For example, the multiplication of two plane waves, $A_s e^{i\theta_s} e^{i\mathbf{k}_s \cdot \mathbf{x}}$ and $A_r e^{i\theta_r} e^{i\mathbf{k}_r \cdot \mathbf{x}}$, from the source and the residual wavefields can be expressed as

$$A_s A_r e^{i(\theta_s + \theta_r)} e^{i(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}}. \quad (6)$$

The wave number vectors of the gradient, therefore, satisfies the following relationship

$$\mathbf{k} = \mathbf{k}_s + \mathbf{k}_r. \quad (7)$$

The wave number mapping of the multiplication is sketched in Figure 1. Using equations (5) and (7), we can compute explicitly the FWI gradient in the wave number domain.

In general, the tomography component of the gradient is formed when waves from the source wavefield and the residual wavefield have opposite directions, which means that the opening angle, α , shown in Figure 1, is 180° . By considering the finite frequency of seismic data, the opening angle of waves forming the tomography component is around 180° . In contrast, the waves producing the migration component have much smaller angles since it is formed by reflections. The two modes, therefore, can be isolated based on the opening angle between the two wave number vectors, which can be used to design a filter, $f(\alpha)$, to extract one mode from the FWI gradient,

$$\hat{g}(\mathbf{k}) = \frac{2}{v^3} \sum_t \sum_i \sum_j f(\alpha) \tilde{p}_s(\mathbf{k}_{si}, t) \tilde{p}_r(\mathbf{k}_{rj}, t). \quad (8)$$

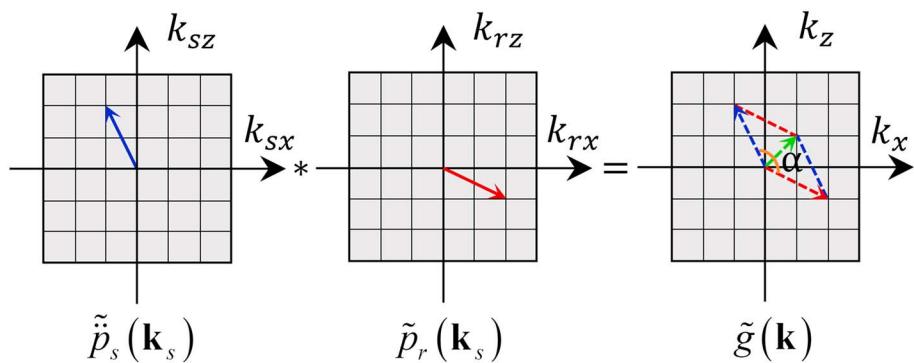


Figure 1. Sketch of the wave number mapping in the multiplication process to form the gradient in the plane wave domain.

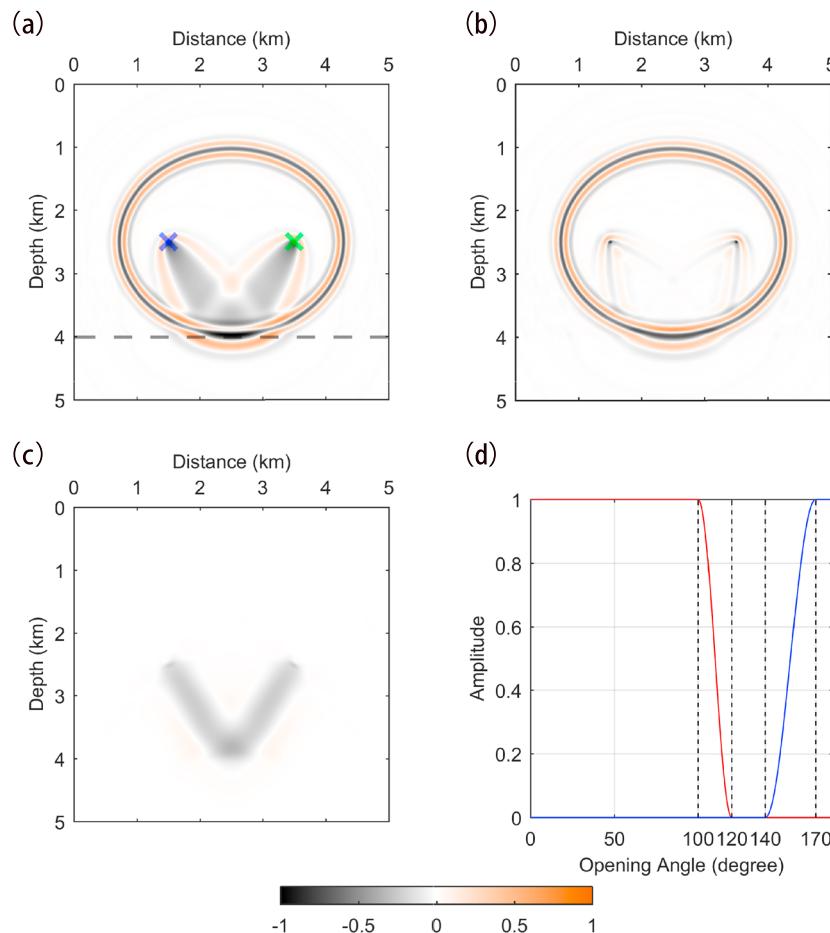


Figure 2. The mode separation of the gradient of the two-layer model with a one-source-one-receiver configuration. The interface of the model is at the depth of 4 km indicated by the dashed line in Figure 2a. The true velocity is 2 km/s above the interface, while it is 3 km/s below the interface. The initial model is also a two-layer model with the interface at the same depth as the true model, but the velocity is 1.9 km/s for the top layer and 2.925 km/s for the bottom layer. The source indicated by the blue cross is located at the depth of 2.5 km and at the distance of 1.5 km. The source signature is a 5 Hz Ricker wavelet. The receiver represented by the green cross is positioned at the depth of 2.5 km and at the distance of 3.5 km. (a) The normal gradient. (b) Low-angle-pass filtered gradient with a transition band from 100° to 120°. (c) High-angle-pass filtered gradient with a transition band from 140° to 170°. Figures 2b and 2c show the migration and tomography components of the gradient, respectively. (d) The angle filters used for Figures 2b and 2c. The red curve shows the filter for Figure 2b, while the blue curve represents the filter for Figure 2c. The vertical dash lines indicate the filters' transition bands, which have a shape of the square of cosine.

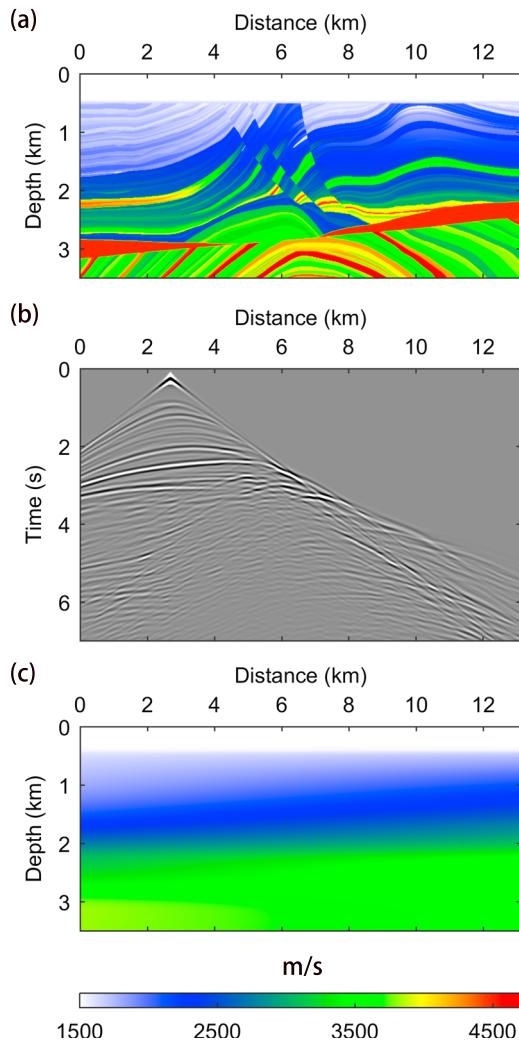


Figure 3. (a) The true velocity model of Marmousi. (b) One shot record from Figure 3a. The source is located at the distance of 2.69 km and the depth of 25 m. The source signature is a 10 Hz Ricker wavelet. (c) The initial velocity model for this FWI test.

Figure 2a. The migration component appears as an ellipse; each point of which has an identical travel time summation to the source and the receiver. Stacking this component from all sources and receivers builds the gradient information for FWI to correct the interface but not the background velocity. The remainder of the gradient is the tomography component, which gives the information required by FWI to update the background velocity. This tomography component is colloquially referred to as “rabbit ears” in the exploration geophysics community.

Figure 2b shows the migration component extracted by low-angle-pass filtering with a transition band between 100° and 120°. The filter is represented by the red curve in Figure 2d. One can observe in Figure 2b that the ellipse-like migration component is kept but the rabbit-ears-like tomography component is removed. On the contrary, high-angle-pass filtering with a transition band from 140° to 170° extracts the tomography component of the FWI gradient (Figure 2c) but eliminates the migration component. The filter is represented by the blue curve in Figure 2d. This test demonstrates that the proposed method is effective in separating the migration and tomography components. This test illustrates the principle of our method in a configuration with one source and one receiver. In the next section, we demonstrate the application of our method to a synthetic data set generated with a complex velocity model for multiple sources and receivers, as well as, to a real data set.

The opening angle α here can be computed from the two wave number vectors of the source and residual wavefields:

$$\cos\alpha = \frac{\mathbf{k}_s \cdot \mathbf{k}_r}{|\mathbf{k}_s||\mathbf{k}_r|}. \quad (9)$$

If the filter is high-angle pass which keeps the plane waves having opening angles around 180°, then the gradient is controlled by the tomography component. On the contrary, if the filter is low-angle pass which rejects the plane waves having opening angles close to 180°, then the calculated gradient is dominated by the migration component.

If we only look at a single temporal snapshot of the wavefields, it is essentially impossible to distinguish whether a plane wave propagates along the direction indicated by the wave number vector or in its opposite direction. This 180° ambiguity could seemly undermine the performance of the angle-filtering technique proposed here. However, as shown in Appendix A, the ambiguity can be resolved when multiple time steps of data are used since only along the correct directions the plane waves are expected to have coherent phases and therefore can be superposed constructively. In contrast, the unwanted components in the final stacked gradient are generally incoherent and therefore are eliminated by destructive stacking.

Herein, we apply the new mode separation method to a synthetic two-layer model, which is described in Figure 2, for demonstrating the effectiveness of the proposed method to the most elementary acquisition configuration: one source and one receiver. A source and a receiver are placed within the top layer. The initial model described in Figure 2 is also a two-layer model but has a lower velocity than the true model. The direct arrivals are removed from the observed and predicted data. More details can be found in the caption of Figure 2.

Since the initial model has an interface, the reflection exists in the source and residual wavefields. Consequently, the gradient of FWI includes both tomography and migration components as depicted in

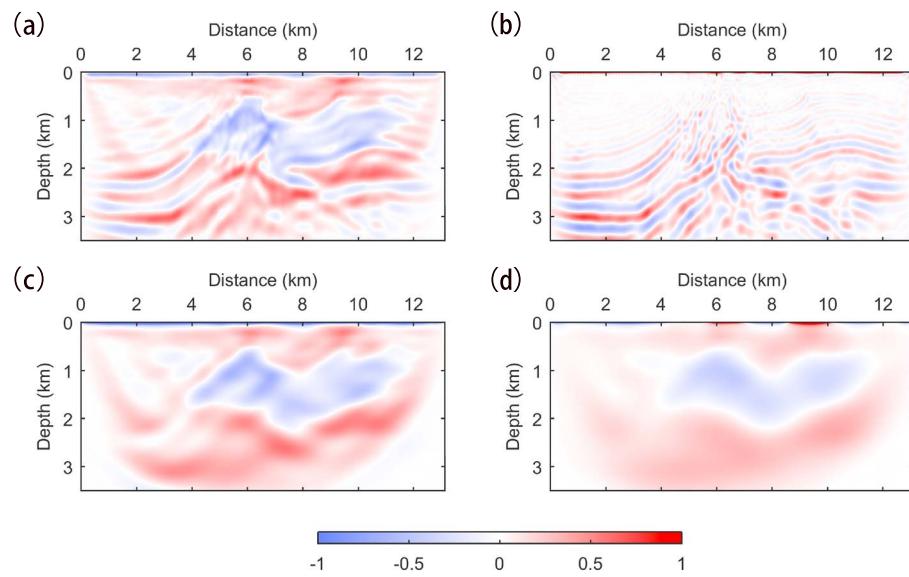


Figure 4. The gradient from the Marmousi model. (a) Normal gradient. (b) Migration component of the gradient with low-angle-pass filtering with a transition band from 100° to 120° . (c) Tomography component of the gradient with high-angle-pass filtering with a transition band from 140° to 170° . (d) As Figure 4c but the transition band is from 160° to 180° .

3. Examples

In this section, we apply this new approach on both synthetic and field data sets to extract the tomography component of the gradient and further update the background velocity model with the tomography component.

3.1. Marmousi Model

The first example is based on the Marmousi model (Figure 3a). We generated 128 shots with a 10 Hz Ricker wavelet at the depth of 25 m. The receiver array is fixed at the same depth as the sources, and it extends across the whole model with a spacing of 12.5 m, which is the cell size of finite difference modeling in this

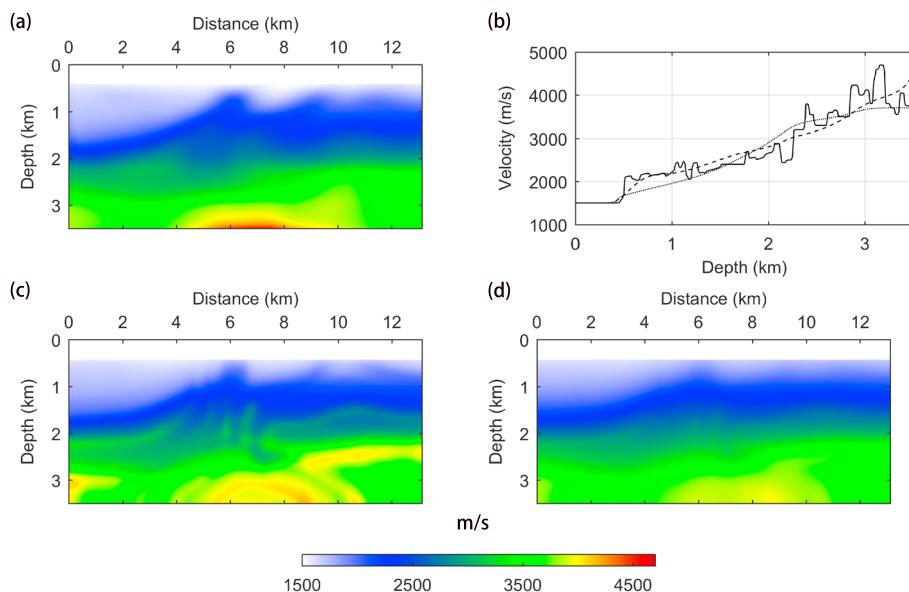


Figure 5. (a) The inverted model by using tomographic mode only. (b) A velocity profile at the distance of 6.25 km. The solid curve represents the true velocity, the dotted curve denotes the initial velocity, while the dashed curve is the recovered velocity. (c and d) The smoothed true velocity models with the cosine-square window size of 30 and 60 cells, respectively.

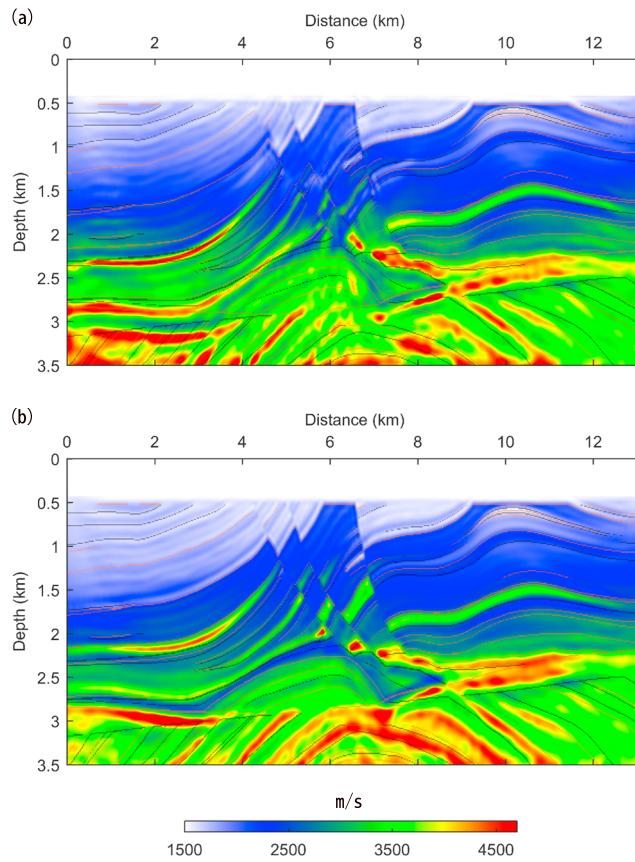


Figure 6. The inverted models with FWI by using normal gradient started from (a) the initial velocity model, which is shown in Figure 3c, and (b) the inverted model with tomography mode only, which is shown in Figure 5a. Top on the inverted velocity model, the thin curves indicate the main velocity interfaces of the true model. The red curves represent the positive reflectivity, while the black curves are for the negative reflectivity.

test. Free surface is applied to generate the record. A shot gather for the source located at 2.69 km is shown in Figure 3b. The starting model used in the inversion is a simple and very smooth model (Figure 3c), which has significant errors of background velocity and needs to use the tomography mode of FWI to fix them.

Figure 4a shows the normal gradient, which includes both the tomography and migration components. The data used to generate the gradient have a narrow frequency band centered at 3 Hz. One can observe in Figure 4a that the left and bottom regions are dominated by the red color while the middle-right region has a blue background color. This means that the velocity of the left and bottom regions should be decreased, while the velocity of the middle-right region should be increased for decreasing the value of the objective function. On top of the background color, we can also see the short-wavelength perturbations, which correspond to the migration component. With low-angle-pass filtering using a transition band of 100–120°, we extract the migration component of the gradient (Figure 4b). The migration component shown in Figure 4b apparently exhibits the structure of the Marmousi model, but the reflectors are placed at the wrong depths because the starting model has a wrong background velocity. To correct the background velocity, FWI needs to use the tomography components of the gradient, which is generated by cross-correlating waves having an opening angle around 180°.

Using high-angle-pass filtering on the gradient, with a transition band from 140° to 170°, we obtain the tomography component (Figure 4c). It can be seen that the gradient is smooth and that it has the correct sign at the left and right parts of the model. However, one can still observe some residual of the migration component. We, therefore, push the transition band up to 160–180°. The resulting tomography component is shown in Figure 4d. Compared to Figure 4c, the gradient in Figure 4d has the same trend but is smoother, which means the faint residual of migration component is removed. This example demonstrates that the angle-dependent plane wave filtering method is also effective for complex wavefields.

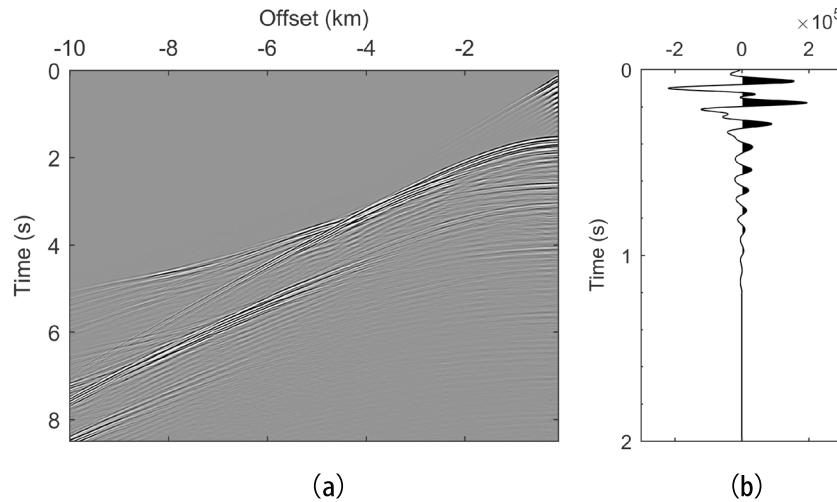


Figure 7. (a) One preprocessed shot gather and (b) the extracted source wavelet.

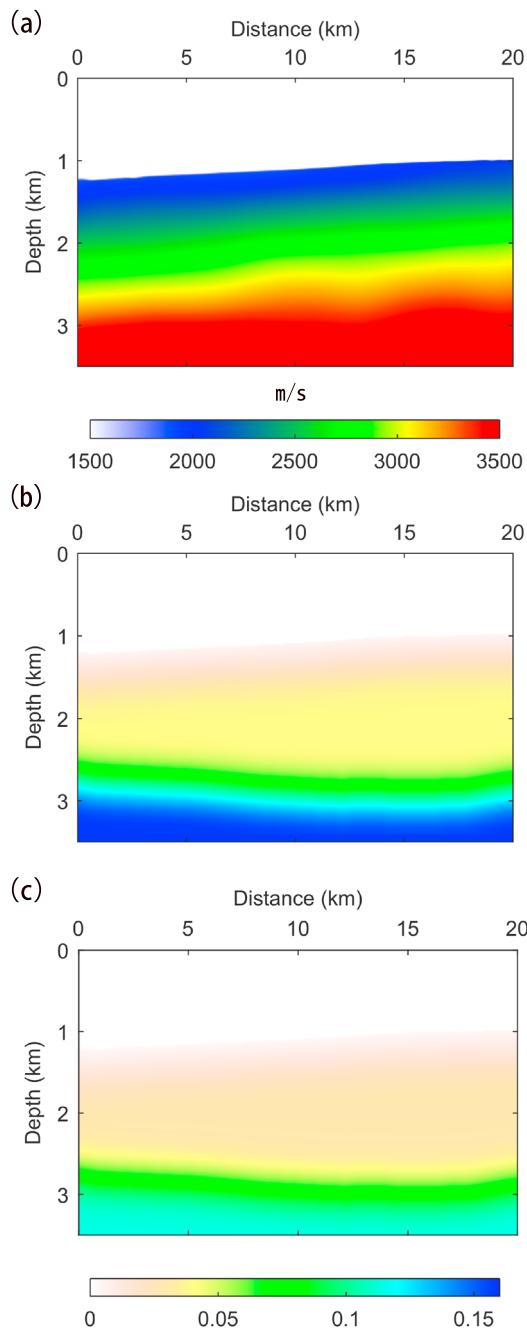


Figure 8. The initial models of (a) vertical velocity, (b) epsilon, and (c) delta.

FWI. The length of the streamer cables is 10 km, which is longer than the typical length of streamer cables. This extended length is more suitable for recording refraction events, which are the key information for a successful application of FWI. The streamers and air gun arrays are fixed at the depth of 25 m and 10 m, respectively, which is deeper than conventional acquisitions (e.g., 7 m for streamers and 2 m for air gun arrays). This setting can boost low-frequency signals as compared to conventional acquisitions. Because the air gun arrays are deployed deeper, the volume of air guns was increased to 5040 cubic inches in order to maintain the air bubble size similar to that at the regular (shallower) depth, resulting in a source signal with more low-frequency contents. The shot and receiver intervals are 50 m and 12.5 m, respectively.

We first applied a low-cut filter to the raw data to remove contents with frequencies lower than 1.9 Hz, which are contaminated by the swell noise. We also filtered out signals and noises with frequencies higher than

If we iterate the inversion with high-angle-pass filtering of 160–180°, the tomography mode dominates the FWI, resulting in a rather smoothed velocity model. The final model after 20 iterations is shown in Figure 5a, which has the general structural features of the true model. For comparison, we plot out a velocity profile located at the distance of 6.25 km in Figure 5b. One can see that the recovered velocity is smooth and has a much closer trend to the true velocity than the initial velocity. Furthermore, we generate two smoothed versions of the true model with the cosine-square window size of 30 and 60 cells, which are shown in Figures 5c and 5d, respectively. A quick comparison of the three models shown in Figure 5 suggests that the recovered model (Figure 5a) has roughly the similar resolution as Figure 5c at the top section and the similar resolution as Figure 5d at the bottom section. This implies that the resolution of the tomography mode decreases with depth. This is because the resolution of tomography is inversely proportional to the width of the first Fresnel zone, which increases with depth (Pratt, 1999; Williamson, 1991).

Figures 6a and 6b show the result of FWI without mode selection starting from the initial model (Figure 3c) and the recovered model with tomography mode only (Figure 5a), respectively. The inversion here employs the so-called multiscale strategy (Bunks et al., 1995), which starts from a relatively low frequency of 5 Hz and gradually increases up to 15 Hz with an increment of 3 Hz. Note that each frequency in the inversion is selected by narrow band-pass filtering, so each frequency indicates a narrow frequency band. Comparing with the true model, one can observe that the depth of the reflectors in Figure 6b matches better than that of the reflectors in Figure 6a, to the depth of the reflectors in the true model. The strong interfaces of the true model (Figure 3a) are indicated by the thin black curves in Figures 6a and 6b to aid comparison. In addition, the reconstruction of the velocity structures in the shallow part of Figure 6b show a stronger sharpness, which is also much closer to the true velocity model (Figure 3a).

This result demonstrates the effectiveness of our method of angle filtering to correctly extract the tomography component of the FWI gradient for improving the kinematics of the background model, which is key for correct positioning of the anomalies in depth, thus explaining the improvement of the inverted model in Figure 6b.

3.2. A Marine Towed Stream Field Data Set

We also applied the proposed technique to an offshore field data set acquired within the Carnarvon basin located on the northwest coast of Australia. Here we use a ~20 km section of a 2-D survey, which has a total length of ~80 km. The 2-D survey was designed especially for

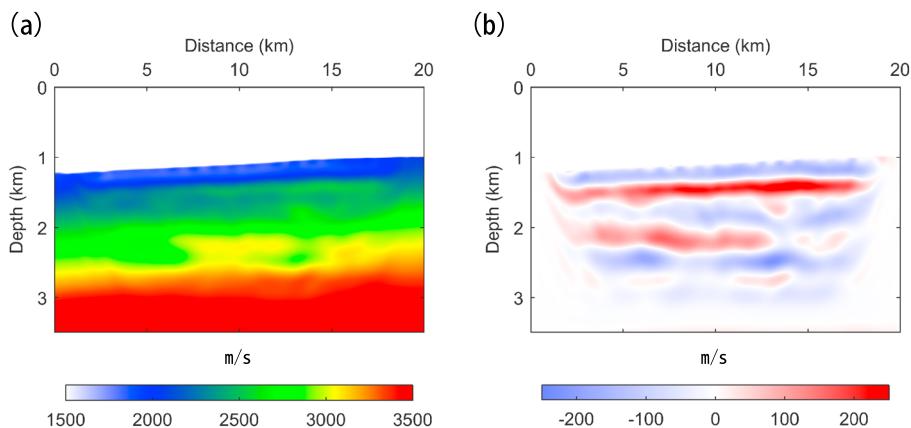


Figure 9. (a) The inverted model of vertical velocity. (b) The update by the inversion, which is the difference between the inverted model and the initial model.

24 Hz. An example of the preprocessed shot gather is shown in Figure 7a. The wavelet used for this inversion shown in Figure 7b is extracted from the direct arrivals. More details about acquisition and preprocessing can be found in Kalinicheva et al. (2017).

In this example, we test how the new angle filtering method extracts the tomography component of the FWI gradient to achieve tomography with FWI. Figure 8 depicts the initial models for velocity and Thomsen's parameters epsilon and delta (Thomsen, 1986). In this inversion, we considered the anisotropic effect by using an acoustic VTI wave equation to model the wave propagation. However, the anisotropic parameters, epsilon (Figure 8b) and delta (Figure 8c), were fixed in the inversion as in Warner et al. (2013). It must be noticed that the data set is acquired in a 3-D setting, while we used a 2-D modeling engine in the inversion. To compensate the difference of wave propagation between 2D and 3D, we carried out two processes: first, the source wavelet was multiplied with $\sqrt{j\omega}$ to compensate the phase difference between 2D and 3D (Y. H. Wang & Rao, 2009); second, the coordinates of receivers were projected onto the sail line to ignore the feathering effects of steamers (Kalinicheva et al., 2017). The frequency multiscale strategy is also applied here. We performed the inversion from the low frequency of 2 Hz and then gradually increased the frequency content by an increment of 1 Hz up to 6 Hz, following by a step of 2 Hz till 24 Hz. Thus, the number of frequency blocks is 14 in total. Each block includes 4 iterations, so there are 56 iterations in total. We applied a high-angle-pass filtering on the gradient with a transition band of 160–180° in each iteration. As the result, the inversion only uses the tomography mode updating the background velocity only. The final velocity model with the tomography model inversion is shown in Figure 9a. Figure 9b shows the model update, which is the difference between the final and initial models. The final model and the model update are smooth because the filtered gradient updates only the long wavelengths of background perturbations. These results demonstrate that the angle-filtering technique proposed here is also capable to separate the tomography and migration modes of FWI for anisotropic media.

To verify the effectiveness of the inversion, we plotted out the value of the objective function against the number of iterations in Figure 10. It shows that the functional value decreases gradually in each inversion block. This means more iterations improve the data fit. The improved data fit can also be seen from the alternating display of the observed data with predicted data from the initial model and the final recovered model (Figure 11). To further assess the quality of the inversion, we applied Kirchhoff prestack depth migration with the initial and inverted velocity models. The migrated profiles are shown in Figures 12a and 12b, respectively. The most noticeable difference between the two profiles is the shift of the depth of reflectors due to the velocity update.

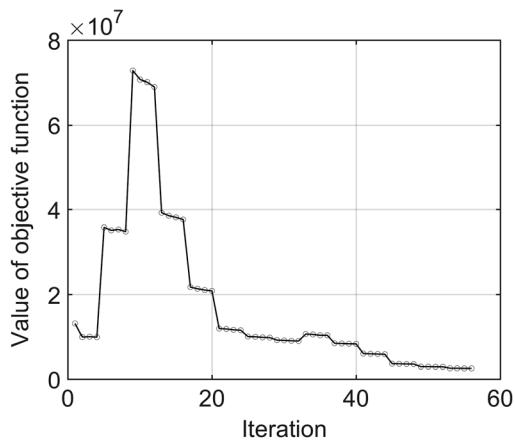


Figure 10. The value of the objective function versus the number of iterations.

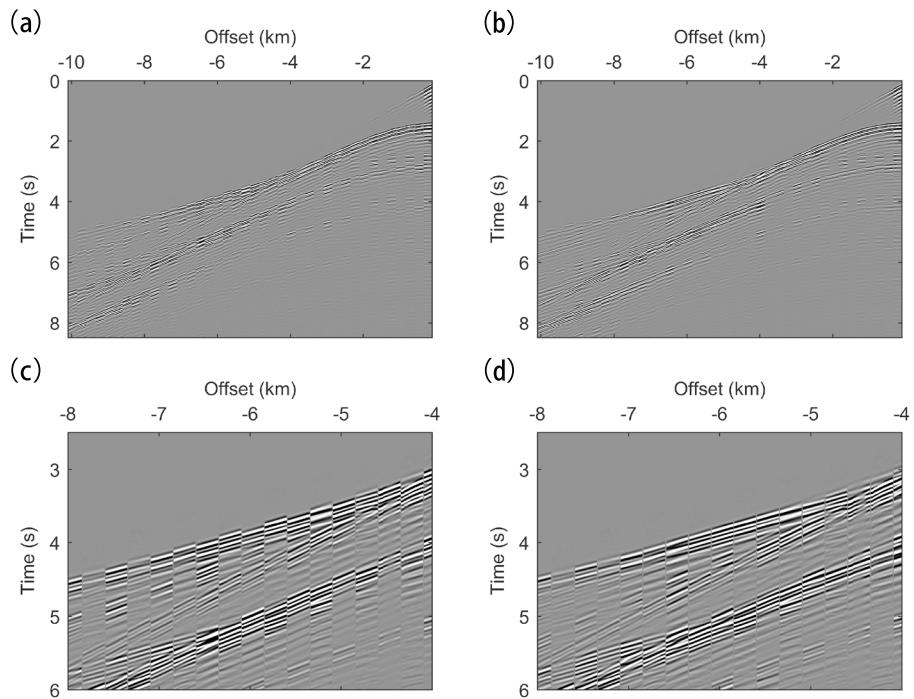


Figure 11. Alternating display of one shot observed data and the predicted data with (a) the initial model and (b) the observed model. In the display, the predicted data are shown first and then followed by the observed data. (c and d) The enlarged view of Figures 11a and 11b, respectively. The source of this shot is located at the distance of 17.625 km.

Furthermore, we can see that the inverted velocity model improves the continuity and focusing of the reflectors; some of which are indicated by the red and black arrows, respectively. These features can be seen in Figure 13 clearly. The improved migration image demonstrates that the tomography mode of FWI does improve the accuracy of the background velocity model.

To further verify this statement, we generate the common imaging-point gathers (CIGs) shown in Figure 14. The flatness of CIGs is a reliable and robust indicator of the accuracy of the migration velocity. The flatter the CIGs, the more accurate the migration velocity. Comparing the CIGs, we can see that the flatness of the CIGs from the inverted velocity model is improved relative to that of the CIGs obtained with the initial velocity model. The improved flatness of the CIGs is indicated by the dotted red lines in Figure 14. All the assessments prove that both our FWI and the new angle filtering method work with field data sets affected by seismic anisotropy.

4. Discussion

4.1. Parameter Selection in the Angle Filtering

The proposed method of angle filtering separates the tomography and migration modes of FWI by essentially using the fact that if the source wavefield travels opposite to the residual wavefield, meaning an opening angle about 180° , the gradient formed by them will be dominated by the tomography component; otherwise, the gradient is mainly the migration component. Since the seismic data have limited bandwidth, which means the wiggles of seismic waves have a certain width, the opening angle for extracting the tomography component should not be limited to 180° . To demonstrate this, we plotted out the tomography and migration components of Figures 2a in 15. Figure 15a shows the tomography component. As can be seen, the tomography component produced by the 5 Hz Ricker wavelet has an oval shape. The center black part is the first Fresnel zone, which contributes to the stacked gradient of the tomography component, while the others are the high-order Fresnel zones, which are canceled out when stacking the gradients from different sources and receivers. Thus, we are only interested in the first Fresnel zone for extracting the tomography component. The angle α_{\min}^t indicates the maximum angle of a point on the edge of the first Fresnel zone

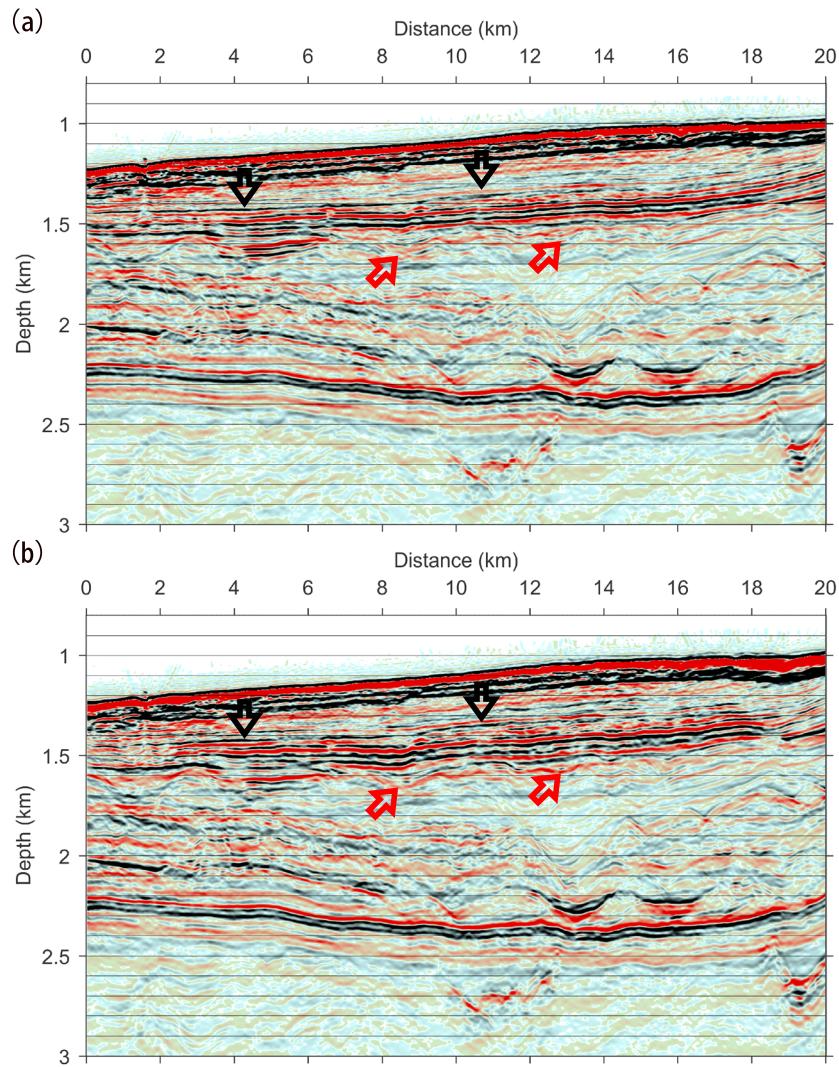


Figure 12. The images of Kirchhoff prestack depth migration with (a) initial vertical velocity and (b) inverted vertical velocity. The black arrows indicate the reflectors which are focused differently in the two migration images, while the red arrows point out the reflectors which have different continuity. Each pair of the arrows in Figures 12a and 12b are located in the same position.

to the source and the receiver. Thus, any two plane waves from the source and residual wavefields having an opening angle larger than α_{\min}^t will form the tomography component inside the first Fresnel zone. For a constant velocity, the relation between α_{\min}^t , the depth of the reflector, d , and the wavelength, λ , can be expressed as

$$\sin \frac{\alpha_{\min}^t}{2} = \frac{4d}{\lambda + 4d}. \quad (10)$$

As a result, we can use the α_{\min}^t as the cutoff angle in the high-angle-pass filter for extracting the tomography component.

Similarly, we illustrate the maximum opening angle, α_{\max}^m for extracting the migration component in Figure 14b. For a constant velocity, the relation between α_{\max}^m , the depth of the reflector, d , and the offset, h , can be expressed as

$$\tan \frac{\alpha_{\max}^m}{2} = \frac{h}{2d}. \quad (11)$$

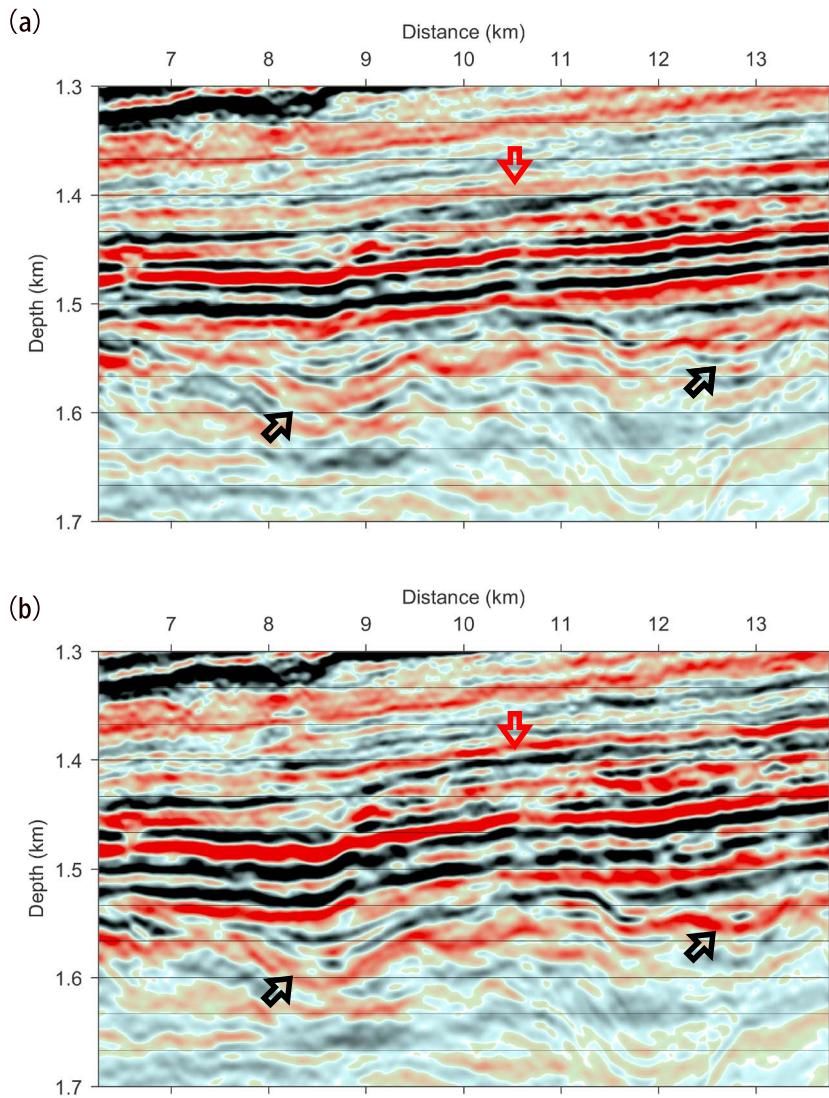


Figure 13. Enlarged view of the images of Kirchhoff prestack depth migration shown in Figure 12. The arrows have the same indication as those in Figure 12.

As a result, any two plane waves from the source and residual wavefields having an opening angle smaller than α_{\max}^m will contribute to the migration component only. Thus, we can use the α_{\max}^m as the cutoff angle in the low-angle-pass filter for extracting the migration component.

However, since the velocity usually is not a constant or reflectors are not located at the same depth, therefore, the cutoff angle α_{\min}^t and α_{\max}^m cannot be calculated analytically, as shown in equations (10) and (11). A rule of thumb of the angle selection is that we can select a narrow transition band around 180°, for example, 160–180°, to extract the tomography component or to remove the tomography component for keeping the migration component.

4.2. Reduction in the Computational Cost

The proposed angle filtering method is achieved by using equation (8), which has three loops of multiplication and thus is computationally intensive. As shown in the previous section of angle selection, only a small portion of angles contributes to the tomography component; therefore, the multiplication in equation (8) only needs to be implemented on these angles. Furthermore, the maximum wave number of the wavefields is much smaller than the Nyquist wave number of the grids of finite difference modeling. The multiplications

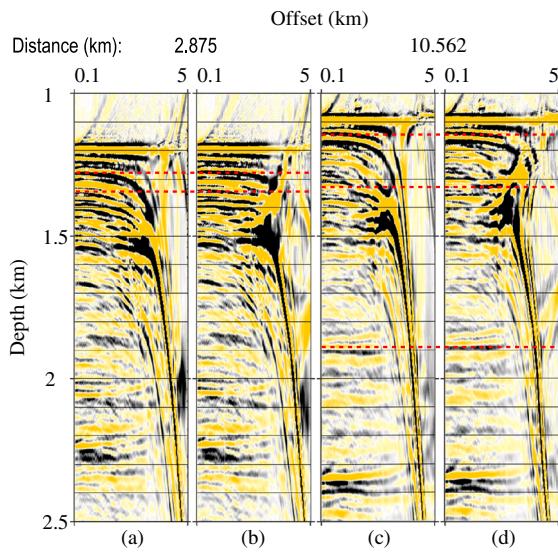


Figure 14. Two pairs of the common imaging-point gathers after Kirchhoff pre-stack depth migration with (a, c) initial vertical velocity and (b, d) inverted vertical velocity. The imaging point of Figures 14a and 14b are located at the distance of 2.875 km, while Figures 14c and 14d are at the distance of 10.562 km. The horizontal red dashed lines are used to indicate the flatness of the gathers.

into the plane wave domain and then by selecting the opening angle of the plane waves from the two wavefields to distinguish the tomography and migration components. The tomography component corresponds to opening angles around 180° , whereas the migration component corresponds to smaller opening angles. The tests on the 2-D synthetic and field data sets have demonstrated the effectiveness of this method for extracting the tomography component to fix the background velocity, which is crucial for correctly positioning velocity anomalies and reflectors in depth, as well as for improving the flatness of image gathers. These are aspects of importance for a more automated and improvement of seismic processing and imaging techniques. This angle-dependent filtering in the plane wave domain can be used not just for extracting the tomography and migration components of the gradient of FWI but can also be further applied to removing low-wave number noise from RTM images. Extension of this technique to 3-D applications is also straightforward.

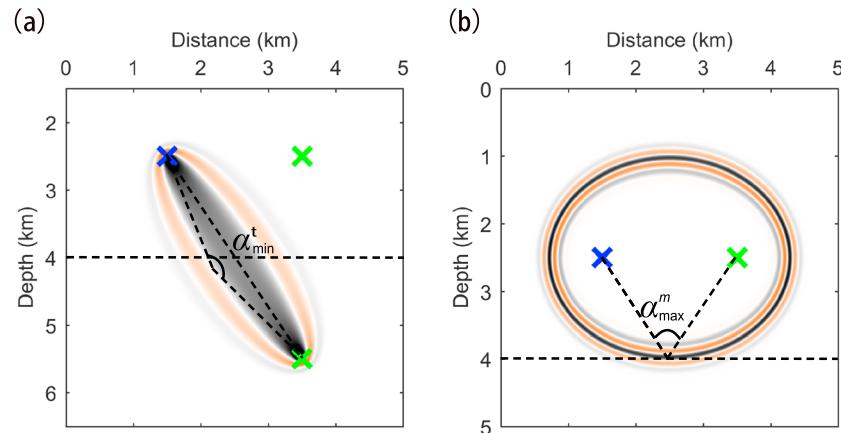


Figure 15. Illustration of the angle selection for extracting (a) the tomography component and (b) the migration component of the FWI gradient. The geometry setting is identical to the test in Figure 2. The horizontal dashed lines indicate the interface of the two-layer model. The blue and green crosses represent the source and the receiver, respectively. The bottom green cross in Figure 15a represents the mirror receiver. α_{\min}^t in Figure 15a shows the maximum opening angle between the source wavefield and the residual wavefield on the edge of the first Fresnel zone. α_{\max}^m in Figure 15b shows the maximum opening angle of the central rays of the source and residual wavefields for the migration component.

in equation (8), therefore, can be only carried on from zero to the maximum wave number. Consequently, as illustrated in Figure 16, an element of $\tilde{p}_s(\mathbf{k}_s, t)$ just needs to multiply the elements of $\tilde{p}_r(\mathbf{k}_r, t)$ in the top part of the cone, which is indicated by the red triangle. In addition, since seismic wavefields are real valued, equation (8) can be calculated on half elements of $\tilde{p}_s(\mathbf{k}_s, t)$, and the other half elements are the complex conjugate of the first half elements. Finally, the loop over time in equation (8) can be as sparse as the half period of the highest frequency of wavefields, which is much larger than the time sampling interval of finite difference modeling. These four strategies can be used to significantly improve the efficiency of the angle-filtering technique proposed here. In addition, as the algorithm described in equation (8) is very simple but computational intensive, the efficiency can be further accelerated with fast computing hardware, such as graphics processing unit (GPU) and field-programmable gate array (FPGA).

5. Conclusions

We have presented a new effective method for separating the tomography and migration components of the FWI gradient. Our method is robust for heterogeneous velocity, the dip angle of reflectors, the direction of wave propagation, and anisotropy. The method is implemented by Fourier transforming the source and residual wavefields

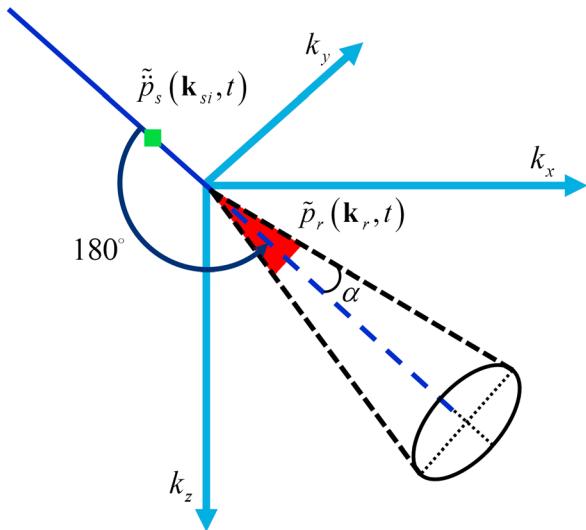


Figure 16. Schematic illustration of reduction in the computational cost by selecting a small region of $\tilde{p}_r(\mathbf{k}_r, t)$, which is indicated by the red triangle, to multiply with $\tilde{p}_s(\mathbf{k}_s, t)$, which is represented by the green square. The angle α indicates the range of the passing angle.

Appendix A: A Proof That Imaging in the Wave Number Domain Is Capable to Distinguish Two Plane Waves Traveling in Opposite Directions

Assuming that there are two plane waves, one from the source wavefield and the other from the residual wavefield. After the Fourier transform along space axes, the plane wave of the source wavefield and its complex conjugate field can be expressed as

$$\begin{cases} p_{s0} = A_s e^{j\theta_s} e^{j\mathbf{k}_s \cdot \mathbf{x}} \\ p_{s0}^\dagger = A_s e^{-j\theta_s} e^{-j\mathbf{k}_s \cdot \mathbf{x}}. \end{cases} \quad (\text{A1})$$

Here “dagger” represents complex conjugate. Similarly, the plane wave of the residual wavefield and its complex conjugate are

$$\begin{cases} p_{r0} = A_r e^{j\theta_r} e^{j\mathbf{k}_r \cdot \mathbf{x}} \\ p_{r0}^\dagger = A_r e^{-j\theta_r} e^{-j\mathbf{k}_r \cdot \mathbf{x}}. \end{cases} \quad (\text{A2})$$

If we take the product of one source and one residual wavefields to form one component of the image, then we can obtain a total of four components, which are

$$\begin{cases} I_{10} = p_{s0} p_{r0} = A_s A_r e^{j(\theta_s + \theta_r)} e^{j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{20} = p_{s0} p_{r0}^\dagger = A_s A_r e^{j(\theta_s - \theta_r)} e^{j(\mathbf{k}_s - \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{30} = p_{s0}^\dagger p_{r0} = A_s A_r e^{j(-\theta_s + \theta_r)} e^{j(-\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{40} = p_{s0}^\dagger p_{r0}^\dagger = A_s A_r e^{-j(\theta_s + \theta_r)} e^{-j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}}. \end{cases} \quad (\text{A3})$$

In the next time step, the phase of the two plane waves increases by a factor of $\Delta\theta$ when they travel in the direction indicated by their wave number. Consequently, the new plane waves become

$$\begin{cases} p_{s1} = A_s e^{j(\theta_s + \Delta\theta)} e^{j\mathbf{k}_s \cdot \mathbf{x}} \\ p_{s1}^\dagger = A_s e^{j(-\theta_s - \Delta\theta)} e^{-j\mathbf{k}_s \cdot \mathbf{x}} \end{cases} \quad (\text{A4})$$

and

$$\begin{cases} p_{r1} = A_r e^{j(\theta_r + \Delta\theta)} e^{j\mathbf{k}_r \cdot \mathbf{x}} \\ p_{r1}^\dagger = A_r e^{j(-\theta_r - \Delta\theta)} e^{-j\mathbf{k}_r \cdot \mathbf{x}}. \end{cases} \quad (\text{A5})$$

The four components of the image are

$$\begin{cases} I_{11} = p_{s1} p_{r1} = A_s A_r e^{j(\theta_s + \theta_r + 2\Delta\theta)} e^{j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{21} = p_{s1} p_{r1}^\dagger = A_s A_r e^{j(\theta_s - \theta_r)} e^{j(\mathbf{k}_s - \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{31} = p_{s1}^\dagger p_{r1} = A_s A_r e^{j(-\theta_s + \theta_r)} e^{j(-\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{41} = p_{s1}^\dagger p_{r1}^\dagger = A_s A_r e^{-j(\theta_s + \theta_r + 2\Delta\theta)} e^{-j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}}. \end{cases} \quad (\text{A6})$$

Assuming that the plane wave of the residual wavefield travels opposite to the direction of its wave number, then in the next time step, the phase of this plane wave is shifted by $-\Delta\theta$. However, the plane wave of the source wavefield will get a phase shift, $\Delta\theta$. Consequently, the plane waves in the next time step become

$$\begin{cases} p_{s1} = A_s e^{j(\theta_s + \Delta\theta)} e^{j\mathbf{k}_s \cdot \mathbf{x}} \\ p_{s1}^\dagger = A_s e^{j(-\theta_s - \Delta\theta)} e^{-j\mathbf{k}_s \cdot \mathbf{x}} \end{cases} \quad (\text{A7})$$

and

$$\begin{cases} p_{r1} = A_r e^{j(\theta_r - \Delta\theta)} e^{j\mathbf{k}_r \cdot \mathbf{x}} \\ p_{r1}^\dagger = A_r e^{j(-\theta_r + \Delta\theta)} e^{-j\mathbf{k}_r \cdot \mathbf{x}}. \end{cases} \quad (\text{A8})$$

The four components of the image are

$$\begin{cases} I_{11} = p_{s1}p_{r1} = A_s A_r e^{j(\theta_s + \theta_r)} e^{j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{21} = p_{s1}p_{r1}^\dagger = A_s A_r e^{j(\theta_s - \theta_r + 2\Delta\theta)} e^{j(\mathbf{k}_s - \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{31} = p_{s1}^\dagger p_{r1} = A_s A_r e^{j(-\theta_s + \theta_r - 2\Delta\theta)} e^{j(-\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_{41} = p_{s1}^\dagger p_{r1}^\dagger = A_s A_r e^{-j(\theta_s + \theta_r)} e^{-j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \end{cases} . \quad (\text{A9})$$

Stacking all the N time steps for the first situation, in which the two plane waves have a phase increase of $\Delta\theta$, the image becomes

$$\begin{cases} I_1 = \sum_{i=1}^N p_{si}p_{ri} = A_s A_r e^{j(\theta_s + \theta_r)} \left(\sum_{i=1}^N e^{j(2i\Delta\theta)} \right) e^{j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_2 = \sum_{i=1}^N p_{si}p_{ri}^\dagger = N A_s A_r e^{j(\theta_s - \theta_r)} e^{j(\mathbf{k}_s - \mathbf{k}_r) \cdot \mathbf{x}} \\ I_3 = \sum_{i=1}^N p_{si}^\dagger p_{ri} = N A_s A_r e^{j(-\theta_s + \theta_r)} e^{j(-\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_4 = \sum_{i=1}^N p_{si}^\dagger p_{ri}^\dagger = A_s A_r e^{-j(\theta_s + \theta_r)} \left(\sum_{i=1}^N e^{-j(2i\Delta\theta)} \right) e^{-j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \end{cases} . \quad (\text{A10})$$

Since summation of a series of complex numbers having a unit amplitude but different phases can be canceled out, the magnitude of I_2 and I_3 having the wave number of $\mathbf{k}_s - \mathbf{k}_r$ is much greater than that of I_1 and I_4 having the wave number of $\mathbf{k}_s + \mathbf{k}_r$.

Stacking all the N time steps for the second situation, in which one plane wave has a phase increase of $\Delta\theta$ but the other has a phase increase of $-\Delta\theta$, the image becomes

$$\begin{cases} I_1 = \sum_{i=0}^{N-1} p_{si}p_{ri} = N A_s A_r e^{j(\theta_s + \theta_r)} e^{j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_2 = \sum_{i=0}^{N-1} p_{si}p_{ri}^\dagger = A_s A_r e^{j(\theta_s - \theta_r)} \left(\sum_{i=0}^{N-1} e^{j(2i\Delta\theta)} \right) e^{j(\mathbf{k}_s - \mathbf{k}_r) \cdot \mathbf{x}} \\ I_3 = \sum_{i=0}^{N-1} p_{si}^\dagger p_{ri} = A_s A_r e^{j(-\theta_s + \theta_r)} \left(\sum_{i=0}^{N-1} e^{-j(2i\Delta\theta)} \right) e^{j(-\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \\ I_4 = \sum_{i=0}^{N-1} p_{si}^\dagger p_{ri}^\dagger = N A_s A_r e^{-j(\theta_s + \theta_r)} e^{-j(\mathbf{k}_s + \mathbf{k}_r) \cdot \mathbf{x}} \end{cases} . \quad (\text{A11})$$

Equation (A11) shows that the magnitude of I_1 and I_4 having the wave number of $\mathbf{k}_s + \mathbf{k}_r$ is much larger than that of I_2 and I_3 having the wave number of $\mathbf{k}_s - \mathbf{k}_r$. This is opposite to the first situation. As a result, we have proved that imaging in the plane wave domain can distinguish the plane waves traveling in one direction or its opposite direction.

Acknowledgments

We thank the sponsors of the FULLWAVE research consortium for supporting this work and are grateful to Woodside Energy Limited and Mitsui E&P Australia for providing the field data set. Authors would like to thank the Editor, two anonymous reviewers for their very meaningful comments and suggestions, which helped to improve and clarify the manuscript significantly. In this study, the Marmousi model, which is a standard test model for seismic study, can be downloaded from Allied Geophysical Laboratories (<http://www.agl.uh.edu/downloads/downloads.htm>); the field data are available from the corresponding author upon request.

References

- Alkhalifah, T. (2015). Scattering-angle based filtering of the waveform inversion gradients. *Geophysical Journal International*, 200(1), 363–373. <https://doi.org/10.1093/gji/ggu379>
- Bunks, C., Saleck, F., Zaleski, S., & Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5), 1457–1473. <https://doi.org/10.1190/1.1443880>
- Chen, M., Niu, F., Liu, Q., Tromp, J., & Zheng, X. (2015). Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons. *Journal of Geophysical Research: Solid Earth*, 120, 1762–1786. <https://doi.org/10.1002/2014JB011638>
- Cheng, X., Jiao, K., Sun, D., & Vigh, D. (2016). Multiparameter estimation with acoustic vertical transverse isotropic full-waveform inversion of surface seismic data. *Interpretation*, 4(4), SU1–SU16. <https://doi.org/10.1190/INT-2016-0029.1>
- Cheng, X., Jiao, K., Sun, D., Xu, Z., Vigh, D., & El-Emam, A. (2017). High-resolution Radon preconditioning for full-waveform inversion of land seismic data. *Interpretation*, 5(4), SR23–SR33. <https://doi.org/10.1190/INT-2017-0020.1>
- Irabor, K., & Warner, M. (2016). Reflection FWI. In *SEG technical program expanded abstracts* (pp. 1136–1140). Dallas, TX. <https://doi.org/10.1190/segam2016-13944219.1>
- Kalinicheva, T., Warner, M., Ashley, J., & Mancini, F. (2017). Two- vs three-dimensional full-waveform inversion in a 3D world. In *SEG technical program expanded abstracts 2017* (pp. 1383–1387). Houston, TX: Society of Exploration Geophysicists. <https://doi.org/10.1190/segam2017-17737603.1>

- Khalil, A., Sun, J., Zhang, Y., & Poole, G. (2013). RTM noise attenuation and image enhancement using time-shift gathers. In *SEG technical program expanded abstracts 2013* (pp. 3789–3793). Houston, TX: Society of Exploration Geophysicists. <https://doi.org/10.1190/segam2013-0600.1>
- Mora, P. (1989). Inversion = migration + tomography. *Geophysics*, 54(12), 1575–1586. <https://doi.org/10.1190/1.1442625>
- Pratt, R. (1999). Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3), 888–901. <https://doi.org/10.1190/1.1444597>
- Ramos-Martinez, J., Crawley, S., Zou, Z., Valenciano, A. A., Qiu, L., & Chemingui, N. (2016). A robust gradient for long wavelength FWI updates, in 78th EAGE Conference and Exhibition 2016, European Association of Geoscientists and Engineers. <https://doi.org/10.3997/2214-4609.201601536>
- Silva, N. V. D., Ratcliffe, A., Vinje, V., & Conroy, G. (2016). A new parameter set for anisotropic multiparameter full-waveform inversion and application to a North Sea data set. *Geophysics*, 81(4), U25–U38. <https://doi.org/10.1190/geo2015-0349.1>
- Sun, D., Jiao, K., Cheng, X., & Vigh, D. (2016). Reflection-based waveform inversion. In *SEG technical program expanded abstracts 2016*, (pp. 1151–1156). Dallas, TX: Society of Exploration Geophysicists. <https://doi.org/10.1190/segam2016-13966097.1>
- Tao, K., Grand, S. P., & Niu, F. (2017). Full-waveform inversion of triplicated data using a normalized-correlation-coefficient-based misfit function. *Geophysical Journal International*, 210(3), 1517–1524. <https://doi.org/10.1093/gji/ggx249>
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8), 1259–1266. <https://doi.org/10.1190/1.1441754>
- Thomsen, L. (1986). Weak elastic anisotropy. *Geophysics*, 51(10), 1954–1966. <https://doi.org/10.1190/1.1442051>
- Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74, WCC1–WCC26. <https://doi.org/10.1190/1.3238367>
- Wang, Y. H., & Rao, Y. (2009). Reflection seismic waveform tomography. *Journal of Geophysical Research*, 114, B03304. <https://doi.org/10.1029/2008JB005916>
- Wang, F., Chauris, H., Donno, D., & Calandra, H. (2013). Taking advantage of wave field decomposition in full waveform inversion, in 75th EAGE Conference and Exhibition 2013, European Association of Geoscientists and Engineers. <https://doi.org/10.3997/2214-4609.20130415>
- Wang, Y., Liang, W., Nashed, Z., Li, X., Liang, G., & Yang, C. (2014). Seismic modeling by optimizing regularized staggered-grid finite-difference operators using a time-space-domain dispersion-relationship-preserving method. *Geophysics*, 79(5), T277–T285. <https://doi.org/10.1190/geo2014-0078.1>
- Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., ... Bertrand, A. (2013). Anisotropic 3D full-waveform inversion. *Geophysics*, 78(2), R59–R80. <https://doi.org/10.1190/geo2012-0338.1>
- Whitmore, N., & Crawley, S. (2012). Applications of RTM inverse scattering imaging conditions. In *SEG technical program expanded abstracts 2012* (pp. 1–6). Las Vegas, Nevada: Society of Exploration Geophysicists. <https://doi.org/10.1190/segam2012-0779.1>
- Williamson, P. (1991). A guide to the limits of resolution imposed by scattering in ray tomography. *Geophysics*, 56(2), 202–207. <https://doi.org/10.1190/1.1443032>
- Wu, Z., & Alkhalifah, T. (2015). Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion. *Geophysics*, 80(6), R317–R329. <https://doi.org/10.1190/geo2014-0365.1>
- Wu, Z., & Alkhalifah, T. (2016). The optimized gradient method for full waveform inversion and its spectral implementation. *Geophysical Journal International*, 205(3), 1823–1831. <https://doi.org/10.1093/gji/ggw112>
- Wu, Z., & Alkhalifah, T. (2017). Efficient scattering-angle enrichment for a nonlinear inversion of the background and perturbations components of a velocity model. *Geophysical Journal International*, 210(3), 1981–1992. <https://doi.org/10.1093/gji/ggx283>
- Wu, D., Yao, G., Cao, J., & Wang, Y. (2016). Least-squares RTM with L1 norm regularisation. *Journal of Geophysics and Engineering*, 13(5), 666–673. <https://doi.org/10.1088/1742-2132/13/5/666>
- Xie, X. (2015). An angle-domain wavenumber filter for multi-scale full-waveform inversion. In *SEG technical program expanded abstracts 2015*, (pp. 1132–1137). New Orleans, LA: Society of Exploration Geophysicists. <https://doi.org/10.1190/segam2015-5877023.1>
- Xie, X., Wu, R., Fehler, M., & Huang, L. (2005). Seismic resolution and illumination: A wave-equation-based analysis. In *SEG technical program expanded abstracts 2005*, (pp. 1862–1865). Houston, TX: Society of Exploration Geophysicists. <https://doi.org/10.1190/1.2148066>
- Xu, S., Wang, D., Chen, F., Zhang, Y., & Lambare, G. (2012). Full waveform inversion for reflected seismic data, 74th EAGE Conference & Exhibition - Copenhagen, 74, w024.
- Yao, G., & Jakubowicz, H. (2016). Least-squares reverse-time migration in a matrix-based formulation. *Geophysical Prospecting*, 64(3), 611–621. <https://doi.org/10.1111/1365-2478.12305>
- Yao, G., & Warner, M. (2015). Bootstrapped waveform inversion: Long-wavelength velocities from pure reflection data, in 77th EAGE Conference and Exhibition 2015, European Association of Geoscientists and Engineers. Madrid, Spain. <https://doi.org/10.3997/2214-4609.201413208>
- Yao, G., & Wu, D. (2015). Least-squares reverse-time migration for reflectivity imaging. *Science China Earth Sciences*, 58(11), 1982–1992. <https://doi.org/10.1007/s11430-015-5143-1>
- Yao, G., & Wu, D. (2017). Reflection full waveform inversion. *Science China Earth Sciences*, 60(10), 1783–1794. <https://doi.org/10.1007/s11430-016-9091-9>
- Yao, G., Warner, M., & Silverton, A. (2014). Reflection FWI for both reflectivity and background velocity. In 76th EAGE Conference and Exhibition 2014, European Association of Geoscientists and Engineers. Amsterdam, Netherlands. <https://doi.org/10.3997/2214-4609.20141089>
- Yao, G., Wu, D., & Debens, H. A. (2016). Adaptive finite difference for seismic wavefield modelling in acoustic media. *Scientific Reports*, 6(1), 30,302. <https://doi.org/10.1038/srep30302>
- Zhang, S., Schuster, G., & Luo, Y. (2011). Wave-equation reflection traveltime inversion, in *SEG technical program expanded abstracts 2011*, edited, pp. 2705–2710, Society of Exploration Geophysicists. San Antonio, TX. <https://doi.org/10.1190/1.3627756>
- Zhou, W., Brossier, R., Operto, S., & Virieux, J. (2015). Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation. *Geophysical Journal International*, 202(3), 1535–1554. <https://doi.org/10.1093/gji/ggv228>
- Zhu, H., Bozdağ, E., Peter, D., & Tromp, J. (2012). Structure of the European upper mantle revealed by adjoint tomography. *Nature Geoscience*, 5(7), 493–498. <https://doi.org/10.1038/ngeo1501>