Balancing amplitude and phase in TFWI

Ali Almomin and Biondo Biondi

ABSTRACT

Tomographic Full Waveform Inversion (TFWI) provides a robust and accurate method to invert the seismic data by simultaneously inverting all scales of the model using both amplitude and phase information. However, one shortcoming of TFWI is the large number of iteration required to achieve accurate results due to its slow convergence. In this report, we analyze the source of its slow convergence and propose two modifications to mitigate the problem. First by modifying the formulation of the regularization term to focus more on the phase information, and second by using an alternative enhancing operator that is less sensitive to the amplitudes in the extended model. Then, we test the modified TFWI on the marmousi 2 model. The results show a significant improvement in the convergence rate.

INTRODUCTION

TFWI, similar to other data-space inversion method, produces highly accurate results due to matching both the amplitude and phase of the data. This is achieved in two steps: first, extending the wave equation and adding an additional axis to the velocity model, and second, adding a regularization term that drives the solution towards a non-extended model. However, one limitation to TFWI in its earlier formation, as shown in Almomin and Biondi (2013), is the large number of iterations required to achieve accurate results. This large number of iterations causes TFWI to be very expensive, especially for large-scale 3D datasets. In Almomin and Biondi (2014), I proposed a wavelength continuation scheme that preconditioned the gradient and smoothed the updates in the early iterations. While preconditioning did improve the convergence rate, the number of iterations was still in the hundreds.

In this report, we analyze the source of TFWI's slow convergence and find that it is due to the unbalanced effects of amplitudes and phase both in the formulation of the regularization term and the enhancing operator. This imbalance resulted in a strong dependence of the kinematic updates on the amplitude fitting, causing it to take more iterations. To mitigate the problem and speed up convergence, we propose two modifications to TFWI. First by modifying the formulation of the regularization term to focus more on the phase information, and second by using an alternative enhancing operator that is less sensitive to the amplitudes in the extended model. Finally, we test the modified TFWI on the marmousi 2 synthetic model.

MODIFIED REGULARIZATION

The conventional L2 objective function for TFWI can be written as follows:

$$J_{\text{TFWI}}(\widetilde{\mathbf{m}}) = \frac{1}{2}\|\mathbf{f}(\widetilde{\mathbf{m}}) - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\epsilon}{2}\|\mathbf{A}\widetilde{\mathbf{m}}\|_2^2 = \frac{1}{2}\|\mathbf{r_d}\|_2^2 + \frac{\epsilon}{2}\|\mathbf{r_m}\|_2^2, \tag{1}$$

where $\widetilde{\mathbf{m}}$ is the extended model, $\mathbf{f}$ is the forward modeling operator, $\mathbf{d}_{\text{obs}}$ is the observed surface data, $\mathbf{A}$ is the defocusing operator, $\mathbf{r_d}$ is the data-fitting residual and $\mathbf{r_m}$ is the model regularization residual. The gradient $\mathbf{g}$ can be written as:

$$\mathbf{g}(\widetilde{\mathbf{m}}) = \mathbf{L}^*(\widetilde{\mathbf{m}})\mathbf{r_d} + \epsilon\mathbf{A}^*\mathbf{r_m}, \tag{2}$$

where * denotes an adjoint and $\mathbf{L}$ is the linearized modeling operator. To avoid cycle-skipping, the modeled data needs to be very close to the observed data such that the data residual is basically a 90 degree phase rotation of the observed data. In TFWI the data-fitting term can fit the observed data regardless of the initial model because the model is extended. The regularization term will adjust the model such that the modeled data in the following iterations will slightly shift from the observed data to create a kinematic difference that results in a meaningful tomographic update.

To better understand the process, we can think of the different stages that TFWI cycles through as it iterates. First, the data-fitting term fits the observed data by creating reflectors in the extended model. Then, the regularization term slightly focuses the model which results in a small shift of the modeled data and a decrease in the data-fitting. Finally, the data-fitting term fits the shifted modeled data to the observed data by creating smooth tomographic updates in the model and adjusting the reflectors locations. Notice that, except for the first few iterations, all these stages happen simultaneously in every iteration.

There are a few convergence issues that take place in practice. First, the data-fitting term creates both the reflectors and tomographic updates which makes it more difficult to balance or emphasize either amplitude or phase fitting. Second, the data-fitting term takes several iterations until it becomes sufficiently small mainly due to amplitude differences resulting from using an adjoint instead of an inverse. If the regularization term focuses the model before the model data fits the observed data, then the data residual will not produce tomographic updates. Third, the regularization term will continue to focus the model at a rate that only depends on epsilon, regardless the data-fitting term. This results in a strong sensitivity to epsilon. If epsilon is too small, the focusing will be too slow and the inversion might take thousands of iterations before producing any useful tomographic updates. If epsilon is too large, the model will be focused too fast and will result in cycle-skipping.

To solve the previous issues, we want to modify the regularization term such that it directly produces the tomographic updates instead of indirectly through the data-fitting term. We can start by rewriting the regularization term as follows:

$$\frac{\epsilon}{2}\|\mathbf{A}\widetilde{\mathbf{m}}\|_2^2 = \frac{\epsilon}{2}\|\widetilde{\mathbf{m}} - (\mathbf{I} - \mathbf{A})\widetilde{\mathbf{m}}\|_2^2 = \frac{\epsilon}{2}\|\widetilde{\mathbf{m}} - \mathbf{E}\widetilde{\mathbf{m}}\|_2^2, \tag{3}$$

where $\mathbf{E}$ is an enhancing or focusing operator which is the complement of $\mathbf{A}$. It is easier to see that the regularization term minimizes the difference between the extended model and an enhanced version of the extended model by focusing the model. This formulation makes it similar to the data-fitting term, however, it is still missing the wave-equation operators. Therefore, we propose a new regularization term as written in the following objective function:

$$J_{\text{TFWI}}(\widetilde{\mathbf{m}}) = \frac{1}{2}\|\mathbf{f}(\widetilde{\mathbf{m}}) - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\epsilon}{2}\|\mathbf{f}(\widetilde{\mathbf{m}}) - \mathbf{f}(\mathbf{E}\widetilde{\mathbf{m}})\|_2^2. \tag{4}$$

This new regularization term completely changes the behavior of TFWI because it does not modify the model. Instead, it directly calculates a residual between the modeled data and the focused modeled data and back-projects it into a tomographic update. This regularization residual is guaranteed to have the correct amount of kinematic shift compared to the modeled data regardless of how well we fit the observed data. Therefore, the data-fitting term only produces reflectors while the regularization term only produces tomographic updates. Moreover, epsilon now balances amplitude fitting and phase fitting without any danger of cycle-skipping at any value, making the inversion process less sensitive to epsilon.

ENHANCING OPERATORS

The enhancing operator is required to create a kinematic shift in the regularization term that can be back-projected into tomographic updates. In the case of DSO, the enhancing operator, i.e. the complement of the focusing operator, is described in equation (3), is:

$$\mathbf{E}_{\text{DSO}}\widetilde{\mathbf{m}}(\tau) = \left(1 - \frac{\tau}{\tau_{\max}}\right)\widetilde{\mathbf{m}}(\tau), \tag{5}$$

where τ is the extended axis value, either in subsurface offset or time lag. As described in the previous section, a proper enhancing operator should result in a residual that is a 90 degree phase rotation of the modeled data. However, DSO operator creates the tomographic update without creating such phase rotated residual. Instead, I scales the amplitude such that the combined effect of all lags (or all offsets in data space) result in the desired tomographic update.

This DSO approach has a few shortcomings. First, DSO assumes the amplitudes will not change significantly along the reflectors or along offsets. This assumption breaks down in many cases such as the presence of AVO effects, complex geometry or irregular acquisition. Second, DSO enhances the image by scaling down the amplitudes along the extended axis. Since the total energy in the data is conserved, this approach assumes that the energy reduction in the image will be converted to tomographic updates in the velocity model, so it indirectly moves focuses the reflectors energy towards the zero-lag. However, this assumption ignores the possibility that the energy can simply be converted reflectors in the null-space of the modeling operator, which are usually present at the edge of the model.

We illustrate the effects of these shortcomings in a simple synthetic example with a single reflector and a constant velocity. Figure 1 shows a shot profile from the modeled data with a slow background velocity using an extended image. The extension of the image preserves all the kinematic information in the observed data. Figure 2(a) shows a shot profile from the modeled data with a slow background velocity using DSO operator regularization on the extended image. Figure 3 compares a trace at 2km offset of the modeled data (bottom) with the DSO regularization residual (middle). We can see that there is no significant change in the phase when we compare the DSO regularization residual to the modeled data and only a small amplitude change.

We propose using a different enhancing operator that shifts the energy towards the zero-lag of the extended axis as follows:

$$\mathbf{E}_{\text{shift}}\widetilde{\mathbf{m}}(\tau) = \widetilde{\mathbf{m}}(\tau + 1 * \text{sign}(\tau)). \tag{6}$$

This shifting operator directly forces the energy to move towards the zero-lag. The main advantage to this enhancing operator compared to the DSO enhancing operator is that it rotates each traces by 90 degrees, therefore, the tomographic update does not depend on how the amplitude of different traces affect each other. In other words, we have no removed the amplitude assumptions. Furthermore, since the energy is directly focused, there are no artifacts at the edge of the model.

We recalculate the regularization residuals with the shifting operator. The shot profile is shown in Figure 2(b) and the trace at 2km is shown in Figure 3 (top). We can see the phase rotation at every offset of the residual.

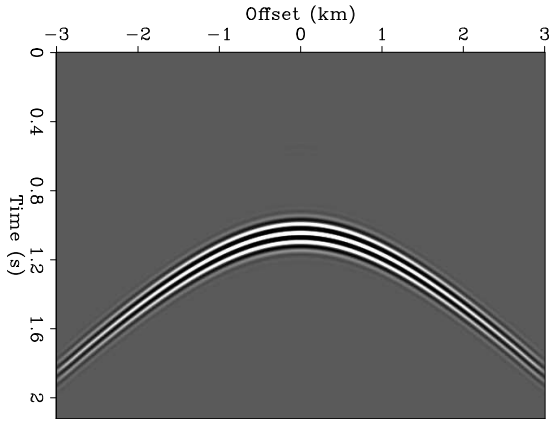
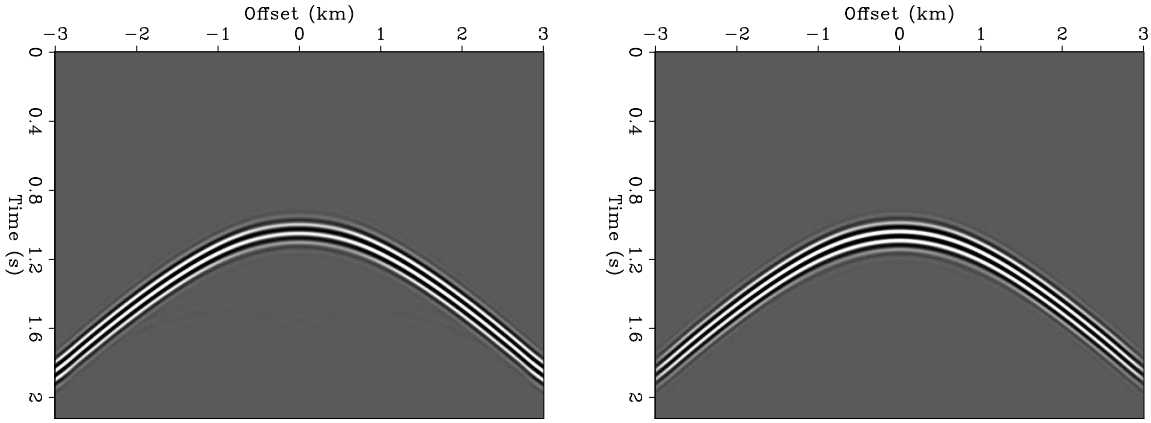


Figure 1: A shot profile from the modeled data with a slow background velocity using an extended image. [**CR**]

We repeat the previous example but with a fast background velocity instead of a slow background velocity. Figure 4 shows a shot profile from the modeled data with a slow background velocity using an extended image. Figures 5(a) and 5(b) show a shot profile from the modeled data with a slow background velocity using DSO operator regularization and shifting operator regularization, respectively, on the extended image. Figure 6 compares a trace at 2km offset of the modeled data (bottom) with



(a) (b)

Figure 2: A shot profile from the modeled data with a slow background velocity using (a) DSO operator regularization on the extended image and (b) shifting operator regularization on the extended image. [**CR**]

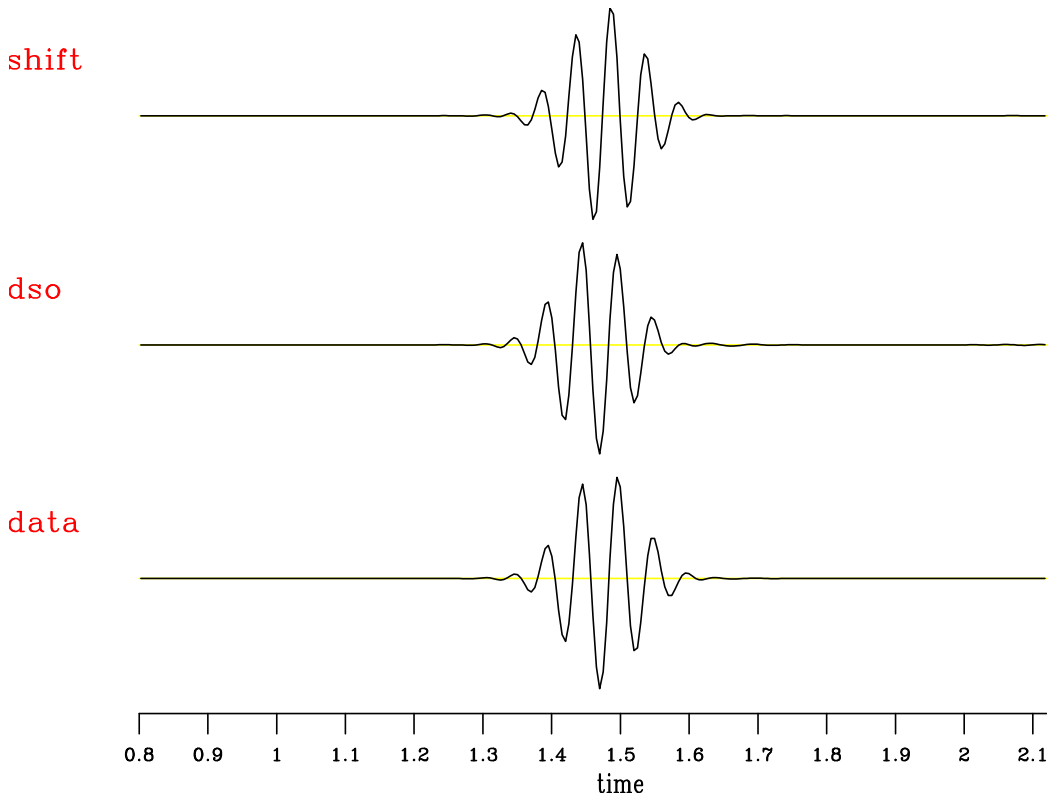


Figure 3: A trace at 2km offset from the modeled data with a slow background velocity using an extended image (bottom), DSO operator regularization on the extended image (middle) and shifting operator regularization on the extended image (top). [**CR**]

the DSO regularization residual (middle) and shifting operator regularization (top). Again, we can see that there is no significant change in the phase when we compare the DSO regularization residual to the modeled data and only a small amplitude change. On the other hand, the shifting operator shows a clear phase rotation (and in the opposite direction to the previous example).

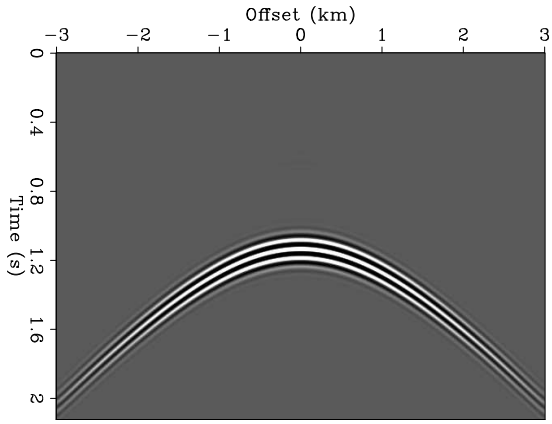
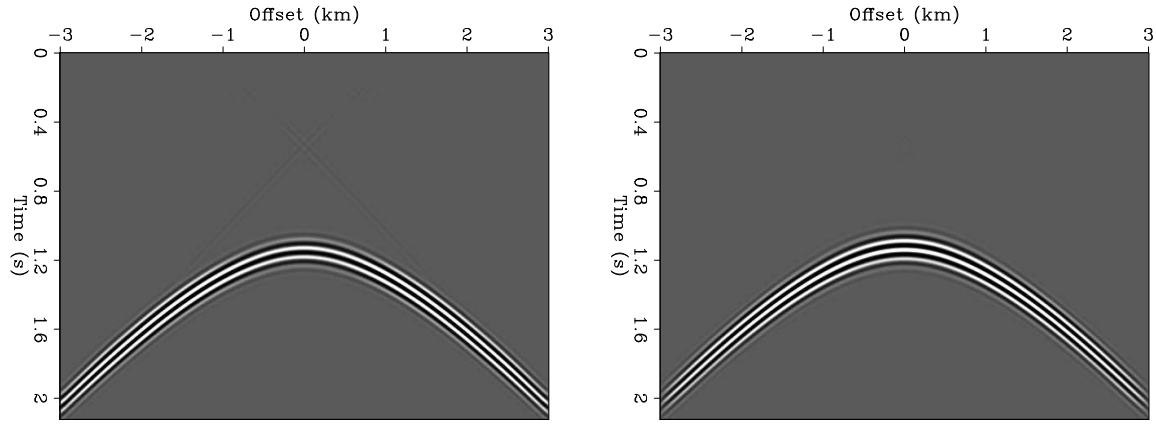


Figure 4: A shot profile from the modeled data with a fast background velocity using an extended image. [**CR**]



(a) (b)

Figure 5: A shot profile from the modeled data with a fast background velocity using (a) DSO operator regularization on the extended image and (b) shifting operator regularization on the extended image. [**CR**]

Next, we compare the two enhancing operator in model space by calculate the tomographic update of two Gaussian anomaly, one faster than the background velocity and one slow than the background velocity with a flat reflector below. The flat and constant-amplitude reflector is the best possible scenario for DSO. Figures 7(a) and 7(b) show the tomographic update of the fast and slow anomalies, respectively, using DSO regularization. Figures 7(c) and 7(d) show the tomographic update of the fast and slow anomalies, respectively, using shifting operator regularization. There is a

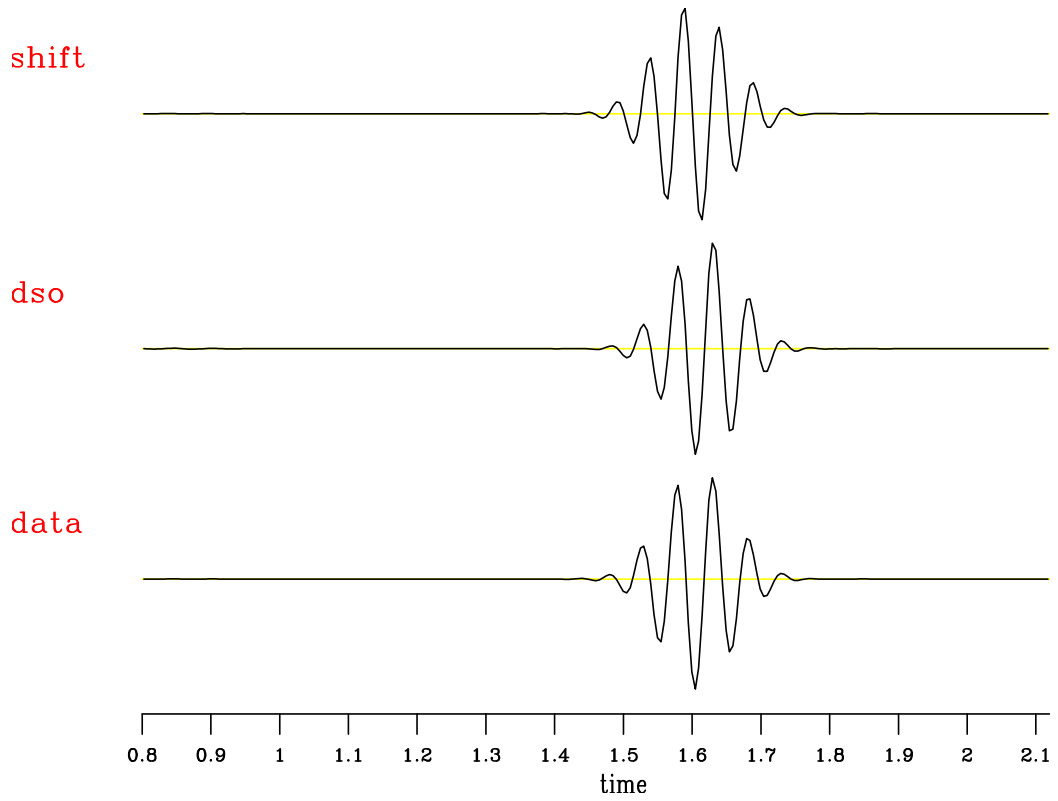


Figure 6: A trace at 2km offset from the modeled data with a fast background velocity using an extended image (bottom), DSO operator regularization on the extended image (middle) and shifting operator regularization on the extended image (top). [**CR**]

significant reduction in kinematic artifacts of the shifting operator around the edges of the model. However, since DSO uses the amplitudes of the image, it has a better focusing of the location of the anomaly.

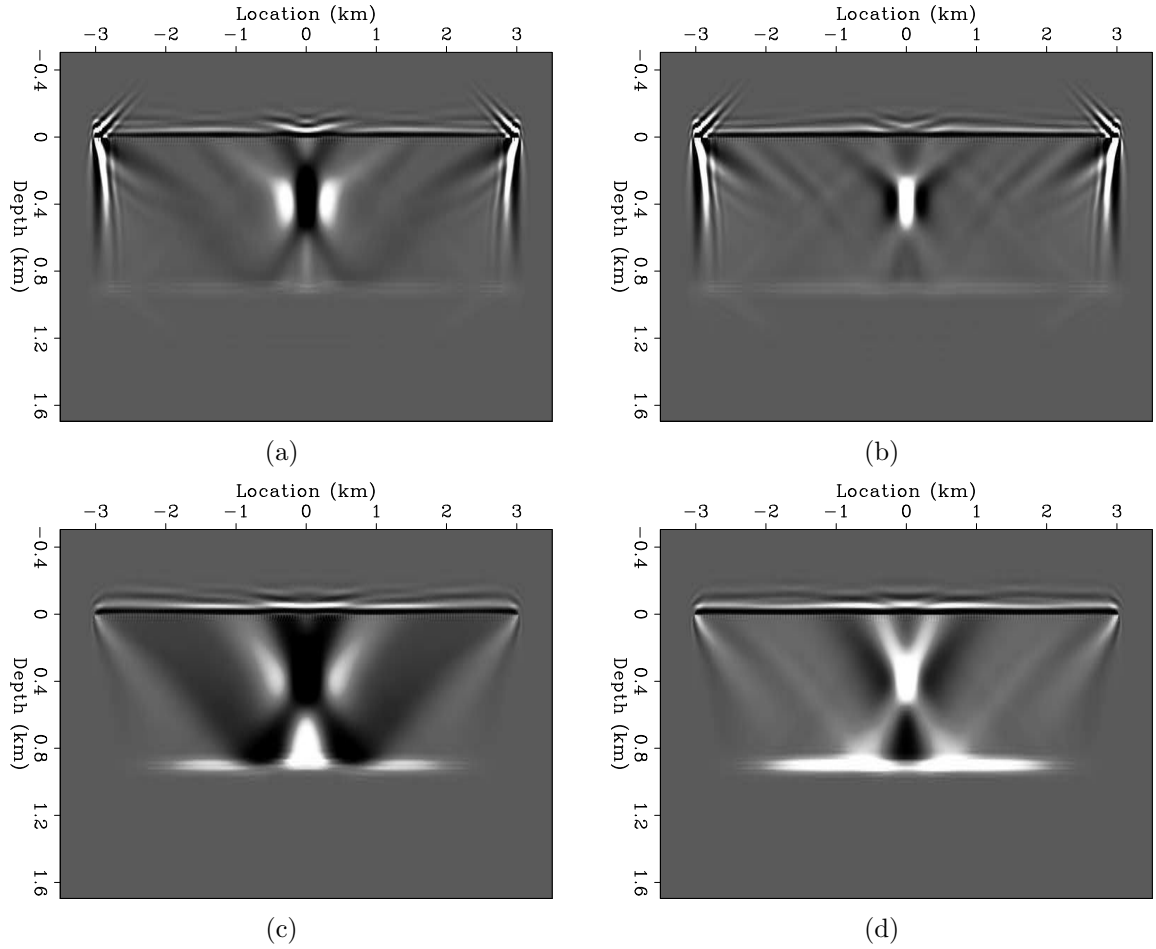


Figure 7: The tomographic update of (a) fast anomaly using DSO regularization, (b) slow anomaly using DSO regularization, (c) fast anomaly using the shifting operator regularization and (d) slow anomaly using the shifting operator regularization. [**CR**]

SYNTHETIC EXAMPLES

To test the new algorithm, we run a synthetic TFWI example on the marmousi 2 model. We use a Ricker wavelet with a frequency range between 5 Hz to 25 Hz and a small taper on both ends. Figure 8 shows the correct velocity model. There are 851 fixed receivers with a spacing of 20 m and 171 sources with a spacing of 100 m. The initial 1D model is shown in Figure 9(a) which is obtained by taking the horizontal average of the correct model after removing the high velocity and low velocity anomalies. The RTM image obtained using the initial model is shown in Figure 9(b).

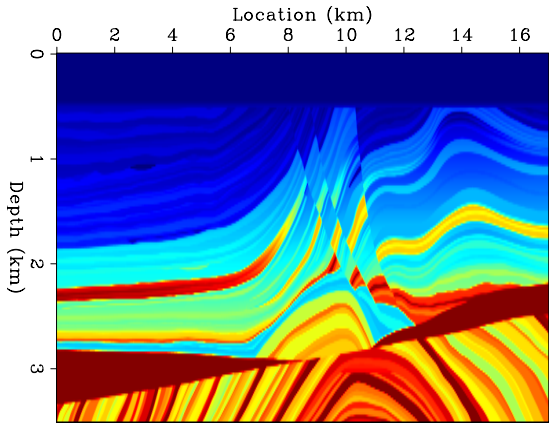
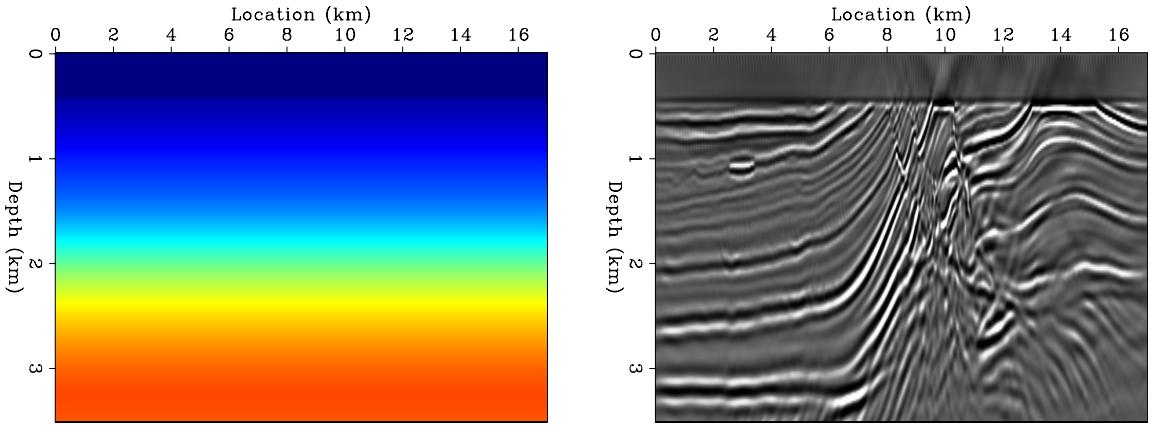


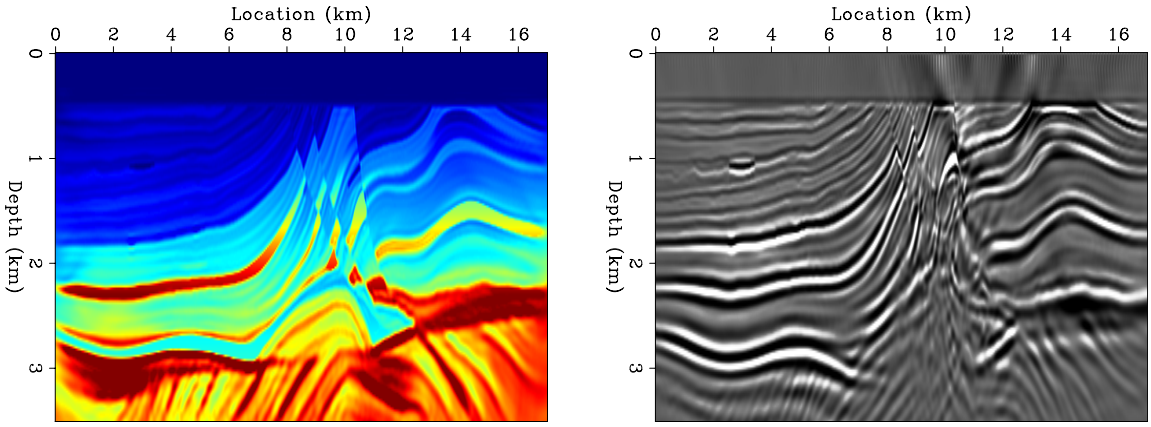
Figure 8: The marmousi 2 velocity model. [**ER**]



(a) (b)

Figure 9: (a) The initial velocity model. (b) The RTM image using the initial velocity model. [**CR**]

We followed the workflow proposed by Biondi and Almomin (2014) by first running TFWI on the low frequencies (up to 10Hz) and then running FWI on the higher frequencies. The inverted model is shown in Figure 10(a) after only 15 TFWI iterations and 30 FWI iterations. The RTM image obtained using the inverted model is shown in Figure 10(b). We see that TFWI successfully inverted the velocity model with high accuracy except around the edges where the illumination is insufficient.



(a) (b)

Figure 10: (a) The inverted velocity model. (b) The RTM image using the inverted velocity model. [**CR**]

CONCLUSIONS

We introduced a modified inversion algorithm that significantly improved the convergence rate of TFWI. This was achieved by introducing a new regularization term that separates the amplitude and phase fitting in the inversion and by changing the enhancing operator to reduce the kinematic artifacts. The synthetic tests show the fast convergence of the new algorithm even when starting from a far initial model.

ACKNOWLEDGMENTS

We would like to thank the Stanford Exploration Project affiliate companies for financial support. Almomin would like to thank Saudi Aramco for supporting his graduate studies at Stanford University.

REFERENCES

Almomin, A. and B. Biondi, 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations: SEP-Report, **149**, 51–58.

———, 2014, Preconditioned tomographic full waveform inversion by wavelength continuation : SEP-Report, **152**, 11–18.

Biondi, B. and A. Almomin, 2014, Efficient and robust waveform-inversion workflow: Tomographic FWI followed by FWI : SEP-Report, **152**, 1–10.