Preconditioned tomographic full waveform inversion by wavelength continuation

Ali Almomin and Biondo Biondi

ABSTRACT

Tomographic full waveform inversion (TFWI) provides a framework for inverting seismic data which is immune to cycle-skipping problems. This is achieved by extending the wave equation and adding a spatial or temporal axis to the velocity model. For computational efficiency, the inversion is performed using a nested scheme. However, TFWI requires a large number of iterations because of its slow convergence rate. We analyze the Born and tomographic operators and find two major sources of this slow convergence. The first source is kinematic artifacts in the extended model due to a biased AVA behavior in the acoustic two-way wave-equation. The second source is early imposition of short wavelength updates in the velocity model that are difficult to move. We provide two modifications of the nested inversion scheme of TFWI that mitigate these sources of slow convergence. The first modification is preconditioning of the extended model to balance the AVA behavior of the acoustic wave-equation. The second modification is imposing wavelength continuation by filtering the gradient in the outer loop. We test the new algorithm on synthetic examples. The results of the modified algorithm on the BP model show a great improvement in convergence rate while maintaining the high accuracy of TFWI.

INTRODUCTION

Tomographic Full Waveform Inversion (TFWI) (Symes, 2008; Sun and Symes, 2012; Biondi and Almomin, 2012) provides a way to overcome cycle-skipping problems by combining both FWI and wave-equation migration velocity analysis (WEMVA) techniques in a generalized framework. This generalized approach utilizes all components of the seismic data to invert for the medium parameters. 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. The velocity model was first extended with subsurface offsets but later Biondi and Almomin (2013) presented an alternative extension using time lags that can handle both reflection and transmission effects. In either setting, this velocity model extension makes the propagation considerably more expensive because each multiplication by velocity becomes a convolution over the extended axis.

In a previous abstract (Almomin and Biondi, 2013), we presented an approximation that significantly reduced the computational cost of TFWI by breaking the extended velocity model into a background component and a perturbation component without sacrificing the accuracy of the method. We achieved this in two steps. First, we set up a nested inversion scheme that utilizes the nonlinear modeling operator to update the residuals. Second, the two components of the gradient were first mixed and then separated based on a Fourier domain scale separation.

One of the disadvantages of TFWI in our previous implementation is the large number of iterations required. By examining the Born and tomographic operators, we find two main sources of this slow convergence. The first source are the kinematic artifacts in the extended model. These artifacts are due to a strongly biased AVA behavior in the acoustic two-way wave-equation. To eliminate these artifacts, we examine the Shuey equations for reflection coefficients and find a preconditioner for the acoustic wave-equation. The second source of slow convergence is the early imposition of short wavelengths in the velocity model. By testing the Born operator, we find that short wavelength components require many iteration to be repositioned. Therefore, we redesign the nested inversion scheme to allow wavelength continuation of the model updates by restricting the total gradient in the outer loop. Finally, we test the new algorithm on the synthetic BP model. The convergence rate is greatly improved, without sacrificing any of the high accuracy level of TFWI.

KINEMATICS ARTIFACTS

We start by testing the Born modeling operator on a constant velocity model that has a single, flat reflector at 900m depth. We start with two background models, the first is slower and the second is faster than the correct model. Figure 1(a) and Figure 1(b) show the extended RTM images as a function of subsurface offset for the slow and fast background models. Although the events have the correct curvature, their energy at larger dips is significantly more dominant than the smaller dips. Next we use the DSO and tomographic operators to estimate the tomographic updates needed for the background velocity as shown in Figures 2(a) and 2(b). Both updates are overwhelmed by artifacts and are largely positive where we expected them to point in the opposite directions. Moreover the updates seem to be dominated by a single dip.

These artifacts can be explained by the Shuey equations (Shuey, 1985) for reflection coefficients of the elastic wave-equation:

$$R(\theta) = R(0) + G\sin^2(\theta) + \mathrm{F}\left(\tan^2(\theta) - \sin^2(\theta)\right), \tag{1}$$

where $R(\theta)$ is the reflection coefficient and θ is the reflection angle. $R(0)$ is the normal reflection coefficient which can be written as:

$$R(0) = \frac{1}{2}\left(\frac{\Delta V_p}{V_p} + \frac{\Delta \rho}{\rho}\right), \tag{2}$$

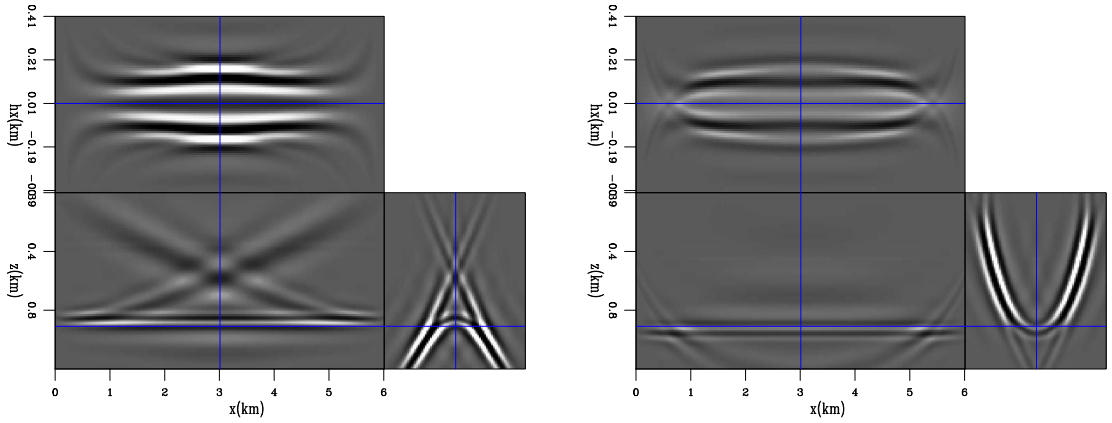


Figure 1: Extended RTM image with (a) a slow background model, (b) a fast background model. [**CR**]

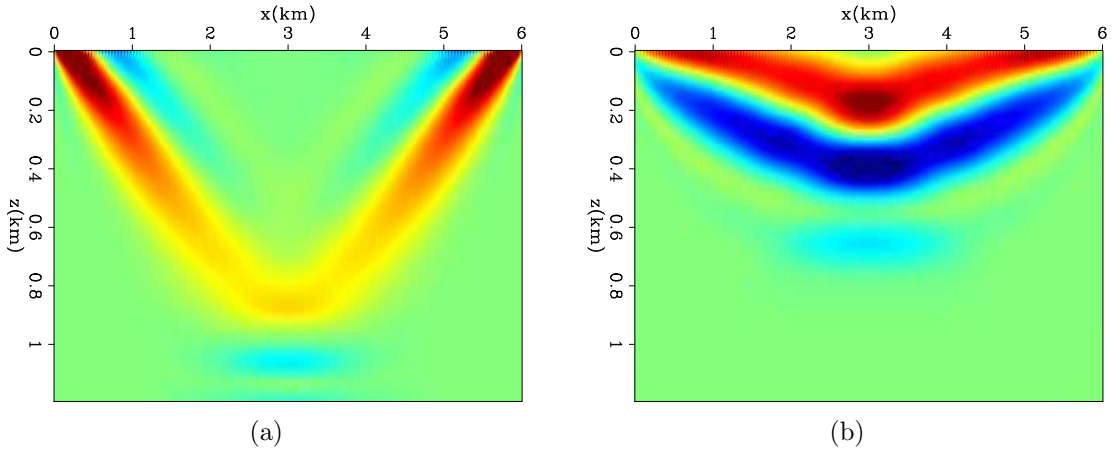


Figure 2: Tomographic gradient using DSO operator with (a) a slow background model and (b) a fast background model. [**CR**]

where V_p is the P-wave velocity, V_s is the S-wave velocity and ρ is the density. The coefficient G is:

$$G = \frac{1}{2}\frac{\Delta V_p}{V_p} - 2\frac{V_s^2}{V_p^2}\left(\frac{\Delta\rho}{\rho} + 2\frac{\Delta V_s}{V_s}\right), \tag{3}$$

and the coefficient F is:

$$F = \frac{1}{2}\frac{\Delta V_p}{V_p}. \tag{4}$$

Notice that for the acoustic case we set $V_s = 0$, reducing the reflection coefficient equation to:

$$R(\theta) = R(0) + F\tan^2(\theta). \tag{5}$$

The previous equation shows that the reflection coefficients always increases in magnitude by a tangent square as a function of angle. This explains the increase in amplitude with larger angles that causes and artifacts in the extended RTM images and the tomographic updates.

In order to reduce the AVA behavior that we observed, we propose preconditioning the image (or gradient of perturbation $\mathbf{g_{p^2}}$) by a stable inverse of the tangent square function such that:

$$\mathbf{q_{p^2}}(\theta) = \frac{\mathbf{g_{p^2}}(\theta)}{\tan^2(\theta) + \alpha}, \tag{6}$$

where $\mathbf{q_{p^2}}(\theta)$ is the preconditioned gradient in angle domain and α is a stabilizing term. Figures 3(a) and 3(b) show the extended RTM images for the slow and fast background models after preconditioning. The artifacts are vastly reduced and the image dips are more balanced. The tomographic updates using the preconditioned images are shown in Figures 4(a) and 4(b). The updates are now pointing in the directions we expect them to, and are more balanced compared to the results shown in Figures 2(a) and 2(b).

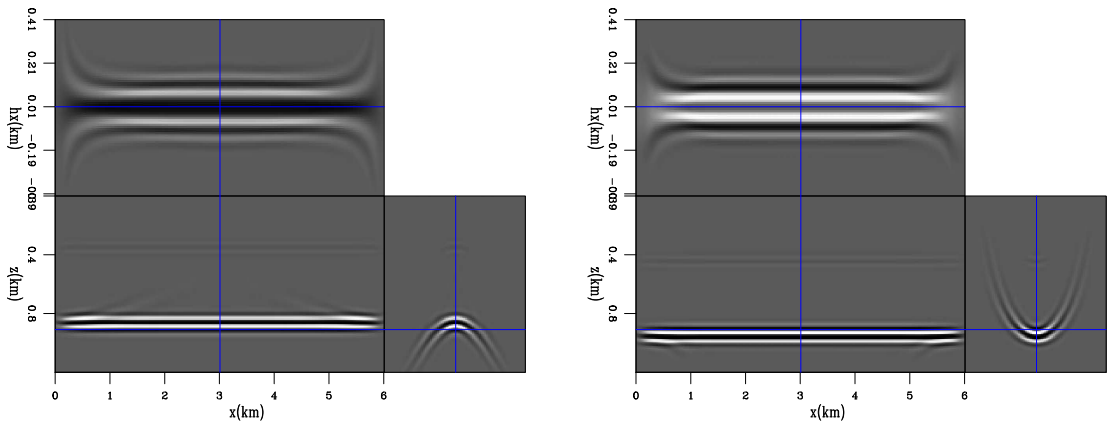
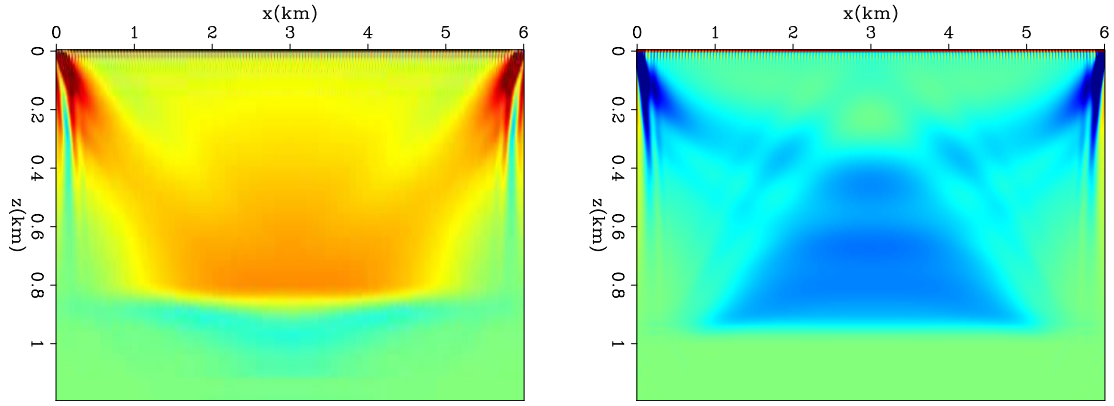


Figure 3: Preconditioned extended RTM image with (a) a slow background model, (b) a fast background model. [**CR**]



(a) (b)

Figure 4: Preconditioned tomographic gradient using DSO operator with (a) a slow background model and (b) a fast background model. [**CR**]

WAVELENGTH CONTINUATION

The direct approach in TFWI is to allow both the Born and tomographic operators to update the velocity model. The short wavelengths start by being positioned at the wrong depth but then get moved to the correct depth. I now examine the rate which Born operator can remove these short wavelength updates by running a least-squares RTM inversion in reverse, i.e., we use the first gradient of a conventional least-squares RTM as the initial model and try to fit data that is all zeros. In other words, the inversion will try to erase the initial model. We run this "reversed" LSRTM for 20 iterations. Figure 5 shows the residual norm as a function of iterations. Although the initial model was created by one iteration, the reverse process required more than 20 iterations to remove it using the same operator. Since this process happen repeatedly in TFWI, it explains the slow convergence rate, since every time the background model changes the perturbation has to be repositioned.

Conventional FWI algorithms avoid this problem by using a frequency continuation approach where the low frequencies are used first. The same approach is not possible in the previous TFWI algorithm because it requires adding the short wavelength updates in two parts. First, as a component of a tomographic operator. Second, it is required to properly minimize the nonlinear objective function in the outer loop.

To overcome these limitations, we propose two modifications to the algorithm. First, we removed the nonlinear line search and simply added the total update directly to the velocity model. The justification is that the updates coming out of the inner iteration are already scaled to fit the data so, further scaling is not necessary. Second, we add a low-pass filter to the total update that follows the inner loop. These modifications allow the tomographic operator to have the short wavelength updates required to build the tomographic gradient, while delaying them from being added to

the velocity model. The modified algorithm can be written as follows:

iterate {
$\Delta\mathbf{d} \longleftarrow \mathbf{d}_{\text{obs}} - \mathcal{L}(\mathbf{s})$
$\mathbf{b}^2 \longleftarrow \mathbf{s}^2$
$\mathbf{p}^2 \longleftarrow \mathbf{0}$
iterate {
$\mathbf{r_d} \longleftarrow \mathbf{L}(\mathbf{b})\mathbf{p}^2 - \Delta\mathbf{d}$
$\mathbf{r_m} \longleftarrow \mathbf{A}\mathbf{p}^2$
$\mathbf{r}_{\text{norm}} \longleftarrow 0.5\|\mathbf{r_d}\| + 0.5\epsilon\|\mathbf{r_m}\|$
$\mathbf{g}_{\mathbf{p}^2} \longleftarrow \mathbf{L}'(\mathbf{b})\mathbf{r_d} + \epsilon\mathbf{A}'\mathbf{r_m}$
$\mathbf{q}_{\mathbf{p}^2} \longleftarrow \text{precondition}(\mathbf{g}_{\mathbf{p}^2}, \alpha)$
$\mathbf{g}_{\mathbf{b}^2} \longleftarrow \mathbf{T}'(\mathbf{b}, \mathbf{p})\mathbf{r_d}$
$(\Delta\mathbf{p}^2, \Delta\mathbf{b}^2) \longleftarrow \text{mix}(\mathbf{q}_{\mathbf{p}^2}, \mathbf{g}_{\mathbf{b}^2})$
$(\mathbf{p}^2, \mathbf{b}^2) \longleftarrow \text{stepper}(\mathbf{p}^2, \mathbf{b}^2, \Delta\mathbf{p}^2, \Delta\mathbf{b}^2)$
}
$\Delta\mathbf{s}^2 \longleftarrow \text{low} - \text{pass}(\mathbf{p}^2 + \mathbf{b}^2 - \mathbf{s}^2)$
$\mathbf{s}^2 \longleftarrow \mathbf{s}^2 + \Delta\mathbf{s}^2$
}

where $\mathcal{L}$ is the wave-equation modeling operator, $\mathbf{s}$ is the slowness model, $\mathbf{b}$ and $\mathbf{p}$ are the background and perturbation components, $\mathbf{r}$ is the residual, $\mathbf{A}$ is the regularization operator and ϵ is a regularization weight term.

This new algorithm performs a wavelength continuation by relaxing the low-pass filter as the number of iteration increase. Hence, the long wavelength updates will be added in early iterations to the velocity model, whereas the short wavelength updates will be added in later iteration. It is important to distinguish between this algorithm and the conventional FWI frequency continuation algorithms. In FWI, there is a direct relationship between the frequency of reflected data and the resulting wavelength of the model update. Therefore, using low frequencies results in smooth updates. However, the relationship between the reflected data frequency and wavelength of the updates is more complicated in TFWI because the tomographic operator can produce long wavelength updates from high frequency data. This property allows wavelength continuation to always start from smooth updates, even when the low frequencies are not present in the data.

SYNTHETIC EXAMPLES

To test the new algorithm, we run a synthetic TFWI example on the BP model. We use a bandpassed wavelet with a frequency range between 5 Hz to 25 Hz and a small taper on both ends. The purpose of using this wavelet is to completely eliminate unrealistically low frequencies in the data. Figure 6(a) shows the correct velocity model. There are 1484 fixed receivers with a spacing of 20 m and 297 sources with a

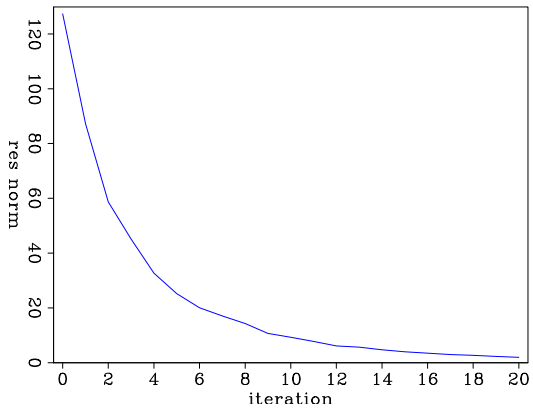


Figure 5: Residual norm of reversed LSRTM as a function of iterations. [**CR**]

spacing of 100 m. The initial 1D model is shown in Figure 6(b) which is obtained by taking the horizontal average of the correct model after removing the high velocity and low velocity anomalies.

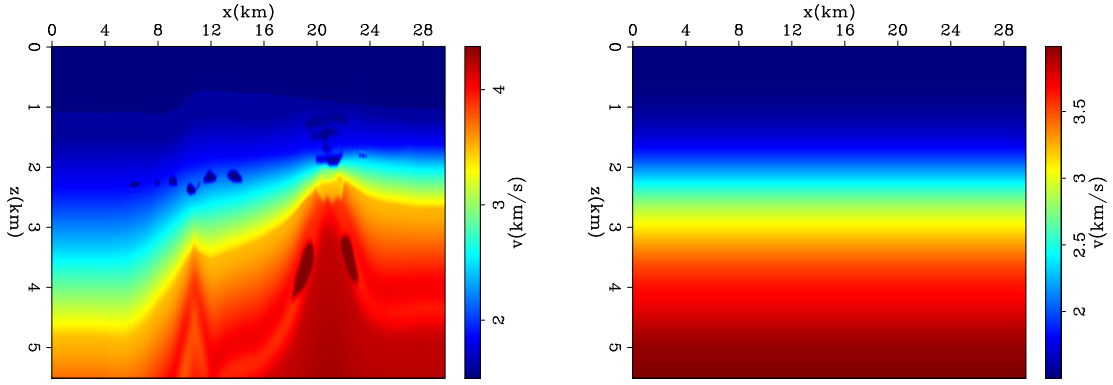
The inversion results after 38 outer loop iterations are shown in Figure 7(a). Each outer loop consists of 10 inner loop iterations. The inversion shows a remarkable reconstruction of most features of the velocity model. Moreover, the number of iterations is reduced by more than an order of magnitude, compared to 500 outer loop iterations in the previous implementation in Almomin and Biondi (2013). Figure 7(b) shows the data-fitting residual norm as a function of total iterations. The new algorithm achieved convergence at a much faster rate than the previous algorithm. The jumps in the objective function correspond to the iterations where the cutoff of the low-pass filter was increased, which slightly increased the data-fitting term of the objective function. In a companion abstract, Biondi and Almomin (2014) show more synthetic examples and present an efficient workflow that combines TFWI and FWI to further reduce the cost of inversion.

CONCLUSIONS

We introduced a modified inversion algorithm that significantly improved the convergence rate of TFWI. This was achieved by preconditioning the extended model to reduce the kinematic artifacts of the acoustic wave-equation and by implementing a low-pass filter that allows for an efficient wavelength continuation without compromising the tomographic operator. The synthetic tests show the fast convergence of the new algorithm even when starting from a far initial model. Currently, the low-pass filter is heuristically relaxed as a function of iterations. More sophisticated and deterministic methods need to be further investigated.

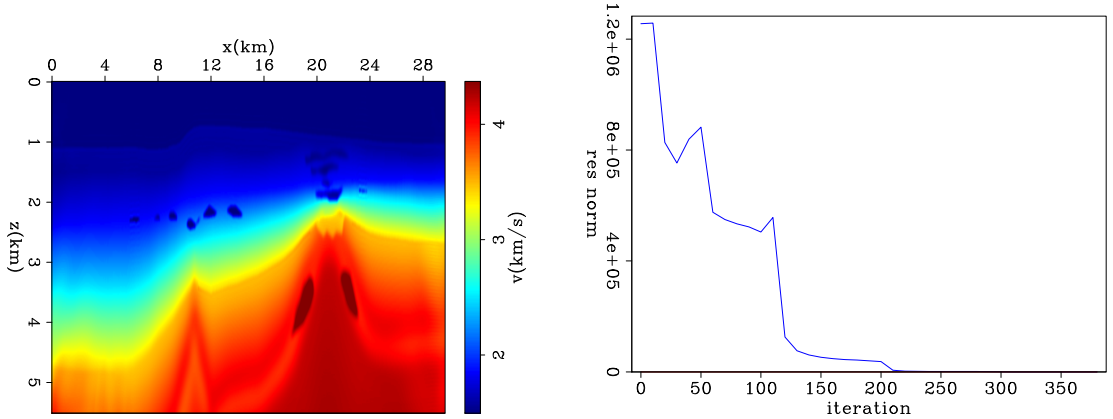
ACKNOWLEDGMENTS

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.



(a) (b)

Figure 6: (a) The correct velocity of BP model and (b) the initial velocity. [**CR**]



(a) (b)

Figure 7: (a) the TFWI results of BP model and (b) residual norm as a function of iterations. [**CR**]

REFERENCES

Almomin, A. and B. Biondi, 2013, Tomographic full waveform inversion (TFWI) by successive linearizations and scale separations: SEG Expanded Abstracts, 1048–1052.

Biondi, B. and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis: SEG Expanded Abstracts, **31**, 275–279.

———, 2013, Tomographic full waveform inversion (TFWI) by extending the velocity model along the time-lag axis: SEG Expanded Abstracts, 1031–1036.

———, 2014, Efficient and robust waveform-inversion workflow: tomographic FWI followed by FWI: SEG Expanded Abstracts, Submitted for publication.

Shuey, R., 1985, A simplification of the Zoeppritz equations: Geophysics, **50**, 609–614.

Sun, D. and W. Symes, 2012, Waveform Inversion via Nonlinear Differential Semblance Optimization: SEG Technical Program Expanded Abstracts, **31**, 497–502.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.