Full Waveform Inversion Algorithm Based on a Time-shift Nonlinear Operator
Zhaoqi Gao1,2,3, Ru-shan Wu2, Zhibin Pan1 and Jinghuai Gao1,3.
1: National Engineering Laboratory for Offshore Oil Exploration, Xi’an Jiaotong University, China; 2: University of California at Santa Cruz, USA; 3: School of Electronic and Information Engineering, Xi’an Jiaotong University, China.

SUMMARY
An accurate initial model is crucial for full waveform inversion (FWI). In this paper, we propose a new way to build an initial model for conventional FWI by developing a new FWI algorithm based on a new time-shift nonlinear operator. We apply the new time-shift nonlinear operator to the waveform, define a new misfit function and derive the corresponding gradient operator. Numerical results using synthetic data from the Overthrust model demonstrate that compared with traditional FWI, the new algorithm is less sensitive to the traveltimes error. Using the inverted model of the new algorithm as the initial model, conventional FWI can obtain much better results.

INTRODUCTION
Full waveform inversion (FWI) becomes increasingly popular in the exploration geophysical industry due to its ability to achieve high-resolution quantitative models of the subsurface by exploiting the full information of pre-stack seismic data. Nowadays, gradient based algorithms are most widely used to solve such problems (Tarantola, 1984; Belina et al., 2009; Liu et al., 2012). In practice, FWI method is an ill-posed and highly nonlinear inverse problem. One key limitation of FWI is that it requires an initial model to be sufficiently close to the true model to prevent cycle-skipping (Virieux and Operto, 2009) especially when low frequency data are missing and the acquisition aperture is narrow. In order to overcome this problem, different methods have been proposed to derive the initial model. Global optimization methods (Padhi et al., 2010, Wang et al., 2011; Tran et al., 2012; Gao et al., 2014) have been used to search for the global minimum of FWI. However, applications of these methods are limited due to tremendous computation cost. Bunks et al., (1995) introduced multi-scale full waveform inversion method, where the inversion is solved from low to high frequency. Brenders and Pratt (2007) used complex-valued frequencies in frequency domain full waveform inversion. Shin and Cha (2008) developed Laplace domain full waveform inversion which can give a smooth background model from an inaccurate initial model. Laplace-Fourier domain full waveform inversion is then proposed by Shin and Cha (2009). Wu et al., (2013, 2014) proposed envelope inversion algorithm which defines its objective function based on the envelope operator and showed its effectiveness for recovering the long wavelength component of the subsurface model. In this paper, we proposed a new FWI algorithm based on a new time-shift nonlinear operator to derive an initial model for traditional FWI. Numerical results on the Overthrust model demonstrated the validity of the new FWI algorithm.

FULL WAVEFORM INVERSION
Full waveform inversion is a challenging data-fitting algorithm. It can be represented as an optimization problem to find the best model that can minimize the misfit function defined as:

$$J(m) = \sum_{t,r} \int_{0}^{T} [u_{\text{obs}}(t,s,r) - u_{\text{syn}}(t,s,r,m)]^2 dt$$ (1)

where $u_{\text{obs}}(t,s,r)$ and $u_{\text{syn}}(t,s,r,m)$ are the observed data and the synthetic data at source $s$ receiver $r$, respectively. Equation (1) is generally solved using a gradient based method. The gradient can be formulated as:

$$\frac{\partial J}{\partial m} = -2 \sum_{t,r} \int_{0}^{T} [u_{\text{obs}} - u_{\text{syn}}] \cdot \frac{\partial u_{\text{syn}}}{\partial m} dt$$ (2)

where $\partial u_{\text{syn}}/\partial m$ is the partial derivative wavefield and the term in the square brackets serves as the effective residual.

Conventional FWI is based on the Born approximation that requires the initial model allows matching the observed traveltimes with an error less than half the period (Virieux and Operto, 2009). If not, the so-called cycle-skipping artifacts will lead to convergence toward a local minimum. Pratt et al. (2008) translates this condition in terms of relative time error $\Delta t/T_L$ as a function of the number of propagated wavelengths $N_\lambda$. The condition is expressed as:

$$\Delta t / T_L < \frac{1}{N_\lambda}$$ (3)

where $T_L$ denotes the duration of the simulation. From condition in equation (3) we can see that if one wants to design a FWI algorithm which is less sensitive to the traveltimes error, the only way is to have a smaller $N_\lambda$. Condition in equation (3) is very hard to be satisfied in conventional FWI because the misfit function measures the difference between the observed data and the synthetic data which has no very low frequencies information. However, if we do not use the original data $u$ but a transformed data $F(u)$ to define the misfit function, we may have the opportunity to decrease the sensitivity of the algorithm on the traveltimes error if the operator $F$ has the following features:

1: $F$ is a nonlinear operator and $F(u)$ can focus its energy on the low frequency components;
2: $F(u)$ does not change the amplitude and traveltimes information contained in the original data $u$. In other words, the position and the reflection energy in $F(u)$ must be consistent with that in $u$.

The first feature can leads to a relative small $N_\lambda$ while the second feature can guarantee that when using $F(u)$ to set up a
FWI based on a time-shift nonlinear operator

FWI algorithm, if the synthetic data match the observed data well, the inverted model has correct long wavelength components. With these two features being satisfied, using $F(u)$ instead of $u$ to define the misfit function, FWI algorithm has more opportunity to converge compared with conventional FWI when given an initial model with large traveltimes error. Here, we also have to point out that there is no guarantee that the FWI algorithm based on $F(u)$ can also obtain a reliable short wavelength component of the subsurface model, so the best way is only use the new FWI algorithm to build a more accurate long wavelength initial model for conventional FWI.

Figure 1: Comparison of $u$ (Red lines) and $P(u)$ (Blue lines). (a) $t_0 = 0 ms$; (b) $t_0 = 23 ms$; (c) $t_0 = 46 ms$.

NEW NONLINEAR OPERATOR

Based on the idea proposed above, in this part, we propose a new nonlinear operator and use numerical tests to demonstrate its properties. First, we define a linear operator $D$ as follow:

$$D(f) = f \ast d(t)$$

where "$\ast$" is the convolution in the time domain. $d(t)$ is defined as:

$$d(t) = \delta(t - t_0) - \delta(t + t_0)$$

where $t_0$ is a time-shift parameter. Operator $D$ acting on a signal is equivalent to first shift the signal with $\pm t_0$ along the time axis and then linearly combine these two signals together. Based on the operator $D$, we define a time-shift nonlinear operator $P$ as follow:

$$P(u) = \sqrt{u^2 + D^2(u)}$$

Operator $P$ non-linearly combines the original data $u$ with the shifted data $D(u)$. Choosing a reasonable time-shift parameter $t_0$, high frequency components of $P(u)$ can be eliminated during combination.

Next, one trace from the synthetic data of Marmousi model has been adopted to analyze properties of operator $P$. Three different values of $t_0$ has been chosen in this test and the results in both time and frequency domain are shown in Figure 1, from which we can see that:

1: Different value of $t_0$ lead to different $P(u)$. When $t_0 = 23ms$, $P(u)$ is the smoothest one in this test.

2: The traveltimes of reflection energy in these three situations are all consist with the original data.

Based on the above experiment results, we can reach the following conclusions. First, operator $P$ satisfied two conditions of operator "$F$" as listed above; Second, given a reasonable $t_0$, the energy of $P(u)$ is mainly located in the low frequency component. Using $P(u)$ to define the misfit function of FWI may lead to less sensitivity on the traveltimes error.

To test our idea, based on the nonlinear operator $P$, we set up a new FWI algorithm called time shift full waveform inversion (TSFWI). Detailed information of TSWFI is shown in the next part.

TIME SHIFT FULL WAVEFORM INVERSION

The misfit function of time shift full waveform inversion is defined as:

$$J(m) = \frac{1}{2} \sum_{x,r} \int_0^T \left[ P(u_{obs}) - P(u_{syn}) \right]^2 dt$$

The gradient of TSWFI can be formulated as:

$$\frac{\partial J}{\partial m} = \frac{1}{2} \sum_{x,r} \int_0^T \left\{ R \frac{\partial u_{syn}}{\partial m} \left[ \frac{\partial u_{syn}^2}{\partial m} + \frac{\partial D^2(u_{syn})}{\partial m} \right] \right\} dt$$

$$= \sum_{x,r} \int_0^T \left( \frac{R u_{syn}}{P(u_{syn})} \frac{\partial u_{syn}}{\partial m} + RD(u_{syn}) \frac{\partial D(u_{syn})}{\partial m} \right) dt$$

where $R = P(u_{syn}) - P(u_{obs})$. Considering that:

$$d(-t) = \frac{\delta(-t - t_0) - \delta(-t + t_0)}{2} = \frac{\delta(-t + t_0) - \delta(-t - t_0)}{2}$$

$$= -d(t)$$

DOI http://dx.doi.org/10.1190/segam2015-5838205.1
FWI based on a time-shift nonlinear operator

and the property of convolution, we obtain the following result:

\[
\frac{\partial J}{\partial m} = \sum_{x,t} \int_{0}^{T} \left\{ R u_{syn} - D \left( \frac{\partial u_{syn}}{\partial m} \right) \right\} dt
\]

where \( \frac{\partial u_{syn}}{\partial m} \) is known as the partial derivative of the wavefield. The gradient of TSFWI can also be calculated by using back-propagation method. The term in the square brackets of equation (10) serves as the effective residual.

\[
\frac{\partial J}{\partial m} = \sum_{x,t} \int_{0}^{T} \left\{ R u_{syn} - D \left( \frac{\partial u_{syn}}{\partial m} \right) \right\} dt
\]

NUMERICAL EXAMPLES

We tested time shift full waveform inversion algorithm on the 2D Overtrust model using synthetic data. The wave equation we adopt in our experiment is shown as below:

\[
\left( \frac{1}{\rho(x)} \frac{\partial}{\partial t} - \Delta \right) u(x,t) = f(t) \delta(x-x_0)
\]

The forward modeling is performed with a high-order finite-difference method in the time domain. The true model (m_1) is shown in Figure 2(a). The whole model has 8000m along the x-direction and 1540m along the z-direction. The grid interval for horizontal and vertical directions are 20m and 10m, respectively. There are 40 shots evenly distributed along the surface and 401 receivers across the surface for each shot. A ricker wavelet with the dominant frequency being 10Hz has been used as the source in the experiment. A linear gradient model (m_2) has been used as the initial model as shown in Figure 2(b). We have conducted the following experiments to test TSFWI algorithm.

Comparison of TSFWI with different time-shift \( t_0 \)

In this experiment, we investigate how parameter \( t_0 \) influences the sensitivity of TSFWI algorithm to the traveltime error. We empirically choose \( t_0 = 0ms, 36ms, 68ms \) for comparison. First, TSFWI algorithm with different \( t_0 \) has been used to do inversion. For three different \( t_0 \), we separately test wavefield \( P(u) \) corresponding to the true model, initial model and their own inverted model. One trace of these wavefields has been plotted in Figure 3 to show the results. From the experimental results, we can see that:

1: Given \( m_2 \) as the initial model, TSFWI algorithm also suffers from cycle-skipping artifacts when \( t_0 = 0ms \) or \( t_0 = 68ms \).

2: Given \( m_2 \) as the initial model, TSFWI with \( t_0 = 36ms \) can converge to the global minimum. At this time, \( P(u) \) corresponding to the inverted model and the true model are almost the same.

From these experimental results, we can conclude that the performance of TSFWI is sensitive to parameter \( t_0 \), this conclusion is consistent with the theoretical analysis. For a specific application, we should choose a reasonable value of \( t_0 \) which can lead to a smooth \( P(u) \).

Comparison of TSFWI+FWI with FWI

In this experiment, we compare the performance of TSFWI+FWI with conventional FWI. TSFWI+FWI means conventional FWI algorithm using the inverted model of TSFWI as the initial model. We choose parameter \( t_0 = 36ms \) in this experiment. Firstly, we directly use conventional FWI algorithm to do inversion, the inverted model can be found in Figure 4(a). Secondly, we use TSFWI algorithm to do inversion and the result of TSFWI is shown in Figure 4(b). Finally, we smooth the result of TSFWI and use this model as the initial model for conventional FWI, the final result is shown in Figure 4(c). One trace located in 3400m along the horizontal direction of the in-
FWI based on a time-shift nonlinear operator

Figure 4: Comparison the inverted models. (a) FWI; (b) TSFWI; (c) TSFWI+FWI.

Figure 5: Comparison of inversion results along one vertical profile from velocity models.

CONCLUSION

We proposed a new time-shift nonlinear operator $P$ and based on operator $P$, we developed a new full waveform inversion algorithm called time shift full waveform inversion (TSFWI). Numerical examples on Overthrust model demonstrated that compared with conventional FWI, TSFWI algorithm is less sensitive to the traveltimes error. Given an inaccurate initial model, TSFWI algorithm can successfully recover the long wavelength component of the subsurface structures. Using the inverted model of TSFWI as the initial model, conventional FWI can obtain a significant improvement on the inverted model.

ACKNOWLEDGMENTS

The work is supported by WTOPI (Wavelet Transform On Propagation and Imaging for seismic exploration) Project at University of California, Santa Cruz. We greatly appreciate the Major Programs of National Natural Science Foundation of China under grant No. 41390450 and 41390454, the Major Research Plan of the National Natural Science Foundation of China under grant No. 91330204, and Beijing Center for Mathematics and Information Interdisciplinary Science (BCMIIS) for their financial support.
REFERENCES


