Introduction

Full-waveform inversion (FWI) is a highly non-linear problem which aims at minimizing the misfit between predicted and observed data to retrieve a high-resolution velocity model (Tarantola, 1984). The knowledge of the true source function is an important prerequisite to achieve a successful wave equation based inversion. Thus, the source wavelet estimation becomes an essential step for practical FWI applications (Pratt, 1999). To mitigate the source estimation error influence, a convolution-based approach was developed (Choi and Alkhalifah, 2011). This approach utilizes a reference trace from the observed and predicted data. In their implementation, they convolve the observed data with the selected predicted data reference trace, and vice versa, they convolve the predicted data with the selected observed reference trace to formulate the objective function. With this source-independent objective function, the source function of both observed and predicted data are equally distributed over both terms, and thus, the effect of the source wavelet is mitigated. This objective function suffers from the cycle-skipping problem due to a poor initial velocity. Efficient wavefield inversion (EWI) is a new formulation of wavefield reconstruction inversion (WRI) to mitigate the cycle-skipping issue in FWI (Leeuwen and Herrmann, 2013; Alkhalifah and Song, 2019). We propose to combine the convolution-based source-independent approach and EWI to remove the source function effect in inverting the wavefield. This source-independent EWI (SIEWI) formulation replaces the data fitting term by a convolution-based source-independent misfit function represented by multiplications in the frequency domain. SIEWI not only maintains the advantage of the efficiency and cycle-skipping reduced features of EWI, but also mitigates the source wavelet effect. We test the proposed approach on an Overthrust model, and the results demonstrate the accuracy of the proposed method.

Theory

EWI formulates an optimization problem for the wavefield that is independent of the velocity perturbation in the wave equation operator, and given by (Alkhalifah and Song, 2019):

$$E(f_{ei}, u_i) = \min \frac{1}{2} \sum_i \| d_i - C u_i \|^2 + \frac{\varepsilon^2}{2} \| L_0 u_i - f_{ei} \|^2,$$

(1)

where, $i$ denotes the source index. $C$ acts as a mapping operator projecting the background wavefield $u_i$ vector onto the receiver positions. $d_i$ represents the observed data, and $L(V) = (V_0 + V)\omega^2 + V^2 = L_0 + \omega^2 V$ is the modelling operator, and $L_0 = V_0\omega^2 + V^2$ corresponds to the background modelling operator. $V$ denotes the squared slowness perturbation for an acoustic isotropic medium, and we set it to zero at the beginning of the inversion process. $f_{ei}$ represents the modified source function. In equation 1, there are two independent variables: the wavefield $u_i$ and the modified source function $f_{ei}$. We calculate the wavefield using a linear equation given by an augmented wave equation:

$$
\begin{pmatrix}
\varepsilon L_0 \\
C
\end{pmatrix} u_i = \begin{pmatrix}
\varepsilon f_{ei} \\
d_i
\end{pmatrix}.
$$

(2)

In each frequency, the modelling operator $L_0$ in equation 2 remains stationary. If the wavefield $u$ is accurately reconstructed, the modified source $f_{ei}$ can be derived using:

$$L u_i = f_i \rightarrow (L_0 + \omega^2 V) u_i = f_i \rightarrow L_0 u_i = f_{ei} = f_i - \omega^2 V u_i.
$$

(3)

As the objective function $E$ is a quadratic function of $f_{ei}$ in equation 1, by setting $\frac{\partial E}{\partial f_{ei}} = 0$, $f_{ei}$ can be evaluated using:

$$f_{ei} = L_0 u_i.
$$

(4)

Initially, the wavefield only has single scattering components from equation 2. As medium parameter perturbations are stored in $f_{ei}$, inner iterations between equations 2 and 4 are used to add the multiscattering components to the wavefield corresponding to secondary sources. This is discussed in details in Alkhalifah and Song (2019). After optimizing the two independent variables $u$ and $f_{ei}$ using inner iterations, $V$ can be directly calculated using a direct division process according to equation 3:

$$V = \sum_i \frac{(f_i - f_{ei})}{\omega^2 u_i} \approx \sum_i \frac{(f_i - f_{ei}) u_i^*}{\omega^2 u_i u_i^* + \lambda}.$$

(5)
where $\lambda$ is a small positive real number to avoid dividing by zero. Usually, we use one percentage of the maximum value of the wavefield energy $u_t^2 u_t$. However, if a wrong source wavelet is used in equation 2, the reconstructed wavefield will have significant errors compared to the true wavefield. In this case, the error in $u$ will propagate to $f_e$ and $V$ through equations 2 to 5, and hamper the convergence of the inversion process. We propose a source-independent EWI (SIEWI) to mitigate this issue. In the time domain, a source-independent based objective function is stated as (Choi and Alkhalifah, 2011):

$$J = \min_{f_e} \frac{1}{2} \sum_i \sum_j \left\| \hat{d}_{i,j} * d_{i,k} - d_{i,j} * \hat{d}_{i,k} \right\|^2,$$

(6)

where, $j$ denotes the receiver index. $\hat{d}_{i,j}$ and $d_{i,j}$ represent the predicted and the observed data, and $d_{i,k}$ and $\hat{d}_{i,k}$ are defined as the reference trace at the $k$th receiver position from the predicted and observed data, respectively. The symbol $*$ represents the convolution operation in time. For each shot gather, the predicted data are convolved with a selected observed data reference trace, and vice versa. By doing so, the source function is represented equally in both terms, and its effect is removed from the objective function. This new formulation aims to avoid the source estimation and mitigate the cycle-skipping problem, simultaneously. It can be formulated in the frequency domain using the same reference trace concept as:

$$E(f_e, u_i) = \min_{f_e} \frac{1}{2} \sum_i \sum_j \left\| C u_i \odot d_{i,k} - d_{i,j} \odot C_k u_i \right\|^2 + \frac{\epsilon^2}{2} \left\| L_0 u_i - f_e \right\|^2.$$

(7)

In equation 7, $C u_i$ is equivalent to $\hat{d}_{i,j}$, which denotes the predicted data. $C_k$ is a mapping operator which just takes the value from the wavefield $u_i$ at the reference trace location, $k$. So we have $\hat{d}_{i,k} = C_k u_i$. In the frequency domain, the convolution process becomes an element by element multiplication, and we use the symbol $\odot$ to denote this. The wavefield $u$ and the modified source wavelet $f_e$ are still two independent variables in SIEWI. As a result, we can calculate the wavefield by solving a source-independent augmented wave equation:

$$\left( C \odot d_{i,k} - d_{i,j} \odot C_k \right) u_i = \left( \epsilon f_{ei} \right).$$

(8)

This source-independent augmented wave equation is able to reconstruct the wavefield with limited source function influence. Then, we can calculate the modified source function and squared slowness perturbation using equations 4 and 5. Finally, the background model is updated by $V_0 = V_0 + V$.

**Examples**

We apply the proposed method on the Overthrust model. The true velocity is shown in Figure 1a, and for the initial velocity we use a highly smoothed version of the true velocity shown in Figure 1b. The true time-domain wavelet is shown in Figure 2a. The peak frequency of the ricker wavelet is 10 Hz, and its frequency spectrum is shown in Figure 2b. We transform the time-domain wavelet using fast Fourier transform (FFT) to generate synthetic frequency-domain data. In practical FWI implementations, the true wavelet information is usually unavailable, so we might end up using a wrong wavelet such as the one in Figure 2c to perform the wavefield reconstruction and inversion. The frequency spectrum of the wrong wavelet is shown in Figure 2d, and we can easily notice the phase distortion and amplitude change compared to the frequency spectrum of the original wavelet. The key point of wavefield reconstruction-based methods is to invert for a good wavefield at the beginning. Figures 3a and 3b are true wavefields calculated using the wave equation with the true and initial velocities, respectively, at 4 Hz. If we use the augmented wave equation 2 to invert the wavefield using the wrong wavelet and the initial velocity, we obtain the inverted wavefield shown in Figure 3c. As the black arrow points out, we see that there is obvious polarity change near the source location. By comparison, the wavefield calculated using the source-independent augmented wave equation using equation 8 is shown in Figure 3d, and it is obvious that no polarity change exists near the source location.

We use 21 sources uniformly distributed on the surface, and all the grid points on the surface act as receivers. We transform the true wavelet shown in Figure 2a into the frequency domain, and generate
Figure 1 The true (a), and initial (b) Overthrust models.

Figure 2 The true wavelet (a), its corresponding frequency spectrum (b), the wrong wavelet (c), and its corresponding frequency spectrum (d).

Figure 3 Wavefields constructed using the wave equation for the true (a), and initial (b) velocity models. Wavefields constructed using the original augmented wave equation (c), and the source-independent augmented wave equation (d), both with the wrong wavelet.

the synthetic data ranging from 3 Hz to 15 Hz with a sampling interval of 0.5 Hz. The $\varepsilon^2$ in this example is set $10^7$. We use the wrong wavelet to perform the inversion, the results using WRI and FWI are shown in Figures 4a and 4b, respectively. We observe that both methods fail to get a reasonable model. Using the wrong wavelet, EWI inverted velocity contains strong source-signature artifacts in the shallow part, as shown in Figure 4c. For SIEWI, we show the inversion result in Figure 4d. We see that SIEWI recovers detailed structure of the true model even with one iteration over all frequencies. After five outer iterations sweeping the full frequency band, the inverted velocity using SIEWI is shown in Figure 5, and all structures in the Overthrust model are reasonably inverted with high resolution.
**Figure 4** The inverted velocity using WRI (a), FWI (b), EWI (c), and SIEWI (d) after one sweep over all frequencies with one iteration per frequency with the wrong wavelet.

**Figure 5** The inverted velocity using SIEWI after five outer iterations with the wrong wavelet.

**Conclusions**

We proposed a source-independent efficient wavefield inversion (SIEWI) to mitigate the source wavelet effect by incorporating a source-independent objective into the original EWI. If we start with a wrong source wavelet in the original EWI, the source wavelet error will propagate into the reconstructed wavefield. In this case, the theory of EWI breaks down and the optimization problem diverges from the desired solution. The proposed method removes the source wavelet accuracy dependency while maintaining the EWI features. Application to synthetic data generated from the Overthrust model show that SIEWI is capable of obtaining reasonable inversion results regardless of the source wavelet accuracy.

**Acknowledgements**

We thank KAUST for its support and the SWAG group for the collaborative environment.

**References**


