Multi-source multi-scale source-independent full waveform inversion

Yundong Guo1,2, Jianping Huang1,2,*, Chao Cui1,2, Zhenchun Li1,2, Liyun Fu1,2 and Qingyang Li1,3

1 School of Geosciences, China University of Petroleum (East China), Qingdao, 266580, China
2 Laboratory for Marine Mineral Resources, Qingdao National Laboratory for Marine Science and Technology, Qingdao, 266071, China
3 Geophysical Exploration Research Institute, Zhongyuan Oilfield Company, Puyang 457001, China

*Corresponding author: Jianping Huang. E-mail: jphuang@upc.edu.cn

Received 5 September 2019, revised 8 January 2018
Accepted for publication 2 April 2019

Abstract
Multi-scale full waveform inversion (FWI) in the time domain requires an accurate source signature estimation, which is difficult to obtain from the field data. As an alternative, we present a source-independent inversion algorithm that modifies Wiener filter calculations. A low-pass Wiener filter can be computed by using information from a reference trace rather than the true source wavelet. A Wiener filter was designed to handle observed and modeled wave fields separately, with a misfit function consisting of the filtered wave field. The source wavelets for the observed and modeled wave fields are both eliminated in the misfit function, thus offsetting any adverse effects from a directly observed or adjoint source wavelet. As the filtered data can be expressed as a convolution of the reflection coefficient and the target source wavelet, it is straightforward to use a low to high frequency selection strategy. Although this new method uses the same forward-modeling and inversion schemes as conventional FWI, it differs in how it calculates the misfit function and corresponding adjoint source. In order to achieve more faithful inversion results and promote computational efficiency, multi-scale source-independent inversion and multi-source inversion strategies were also adopted. In particular, we realized that a weighted average of the zero-offset traces in the super gather constitutes a better choice for reference traces. Aside from this, the noise resistance capability of the proposed method was tested and synthetic data evaluations of both the source-independent and multi-source inversion strategies were undertaken to assess the efficiency of the proposed method.

Keywords: full waveform inversion, source-independent, multi-scale inversion, multi-source inversion

1. Introduction
Providing efficient inversions of velocity distribution is of significant importance for both the accurate imaging and characterization of reservoirs (Sun et al. 2017; Liao et al. 2009). However, conventional velocity analysis and tomography methods based on the travel time associated with picked events no longer meet the demands of exploration environments (Rao & Wang 2005; Tao & Sen 2015). Full waveform inversion (FWI), which focuses on minimizing misfit functions, measures the difference between synthetic and observed data to get the best possible data reproduction. This ensures more accurate velocity estimation (Tarantola 1984; Virieux & Operto 2009; Liao et al. 2017). As the computational ability of supercomputers and quality of data acquisition has improved, FWI has been used with increasing success for transmission data and wide-angle
reflection/refraction seismic data (Luo & Schuster 1991; Yu et al. 2014; Operto et al. 2015).

Despite this success, the effective use of FWI is hindered by a number problems. First of all, gradient-based algorithms are the most popular choice for inversion. They search on the basis of an initial model that is already fairly accurate, then update the model with gradient information obtained by using an adjoint state method (Plessix 2006). Ultimately, this results in a minimized misfit function (Vireux & Operto 2009; Brossier 2011). However, when using this approach, the FWI misfit function is poorly defined and the calculations can produce numerous local minima, especially if there are low wavenumber components in the velocity model. Without a sufficiently good initial model and suitable observed data, the inversion process can easily become locked in the local minima and obtain results that are not in line with the actual situation. The non-linearity of the misfit function is closely related to the frequency of the data and misfit functions associated with low-frequency data display better global convergence. A frequency continuation strategy, which inverts the low frequencies first by filtering the data according to the frequency or time domain, has been proposed as a way of avoiding the local minima issue by controlling the wavenumber admitted to the gradient in a sequence from low to high (Bunks et al. 1995; Boonyasiriwat et al. 2009). Sirgue & Pratt (2004) managed to significantly improve the computational efficiency of multi-scale FWI by defining a frequency discretization strategy that depended on the maximum effective offset. FWI in the frequency domain naturally follows a multi-scale approach because the data is already decomposed into separate frequency components (Pratt 1999).

Boonyasiriwat et al. (2009) have also proposed a multi-scale method for time domain FWI where the frequency selection strategy is extended to time domain waveform inversion. Although multi-scale FWI in the time domain achieves higher computational efficiency and more linear convergence, it remains limited by the fact that the Wiener filter computation still needs information from a true source wavelet (Boonyasiriwat et al. 2009). In real cases, especially for land datasets, the source wavelet is not directly available and estimation of the source wavelet is complicated.

As FWI applies a wave equation to delineate the subsurface velocity structure according to the best fit between the synthetic and observed data, the source signature is one of the key factors that affects the non-linearity of the misfit function (Hu et al. 2017). In real cases, the source signature of the observed data can be obtained by windowing a portion of the early arrivals (Yu et al. 2014, 2016) or it can be estimated along with the velocity structure in the waveform inversion algorithm (Zhou et al. 1997). However, because of the poor knowledge of the coupling between the source and medium, it is difficult to invert for a good source wavelet without knowing the exact near-surface velocity distribution. The estimated source signature can differ substantially from the real source and the final tomogram may become severely polluted with unacceptable artifacts. Frazer & Sun (1998) have suggested a new objective function for the inversion of sonic waveforms with an unknown source where receiver functions can be used to interpret well-logged sonic waveform data. To avoid cumbersome source signature estimation, some researchers have proposed using source-independent FWI methods in both the frequency and time domains (Lee & Kim 2003; Choi et al. 2005; Choi & Alkhalifah 2011; Xie et al. 2018). These algorithms can be divided into two groups: convolution-based approaches and de-convolution-based approaches. For convolution-based approaches, the effects of the source wavelets can be eliminated by convolving the gathered data with a reference trace from a modeled wave-field and the modeled data with a reference trace from an observed wave field (Choi et al. 2005; Choi & Alkhalifah 2011; Li et al. 2017). For de-convolution-based approaches, the seismic data is normalized with a reference trace selected from the seismic data in the frequency domain (Lee & Kim 2003). Out of the two, convolution-based methods are better-suited to FWI in the time domain because they avoid the additional Fourier transforms and division necessary in the frequency domain. Choi & Alkhalifah (2011) have applied convolution-based theory to encoded multi-source FWI in the time domain, where it was found to be a more efficient and stable inversion strategy.

Another key factor restricting the application of FWI is its high computational burden. With the advent of high-density 3D survey systems, FWI has come to bear an even greater computational cost (Virieux & Operto 2009). Unlike FWI in the frequency domain, where the computational burden is somewhat independent of the number of shots (Pratt 1999), the computational burden of FWI in the time domain is proportionally related to the number of shots and length of recording over time. This makes it unsuitable for high-density or 3D data (Anagaw & Sacchi 2014; Kwon et al. 2015). Following on from the successful application of pre-stack migration schemes, simultaneous shooting techniques that generate super-shots by summing individual sources with a random encoding function have been turned to as a way to reduce the computational burden (Herrmann et al. 2009; Krebs et al. 2009; Ben-Hadj-Ali et al. 2011; Rao & Wang 2017; Rao et al. 2018). Here, only two forward-modeling procedures are needed to calculate the gradient of the misfit function with respect to the model parameters. This eliminates the proportional relationship between the computational cost of FWI and the number of shots. However, the simulation of simultaneous sources introduces crosstalk artifacts, which negatively influence the inversion results. In order to suppress the crosstalk noise effectively, multiple sources can be randomly encoded with a phase encoding function or with a time-shift function (Romero et al.
An L2-norm-based FWI misfit function can be defined as:

$$\chi(m) = \sum_{x_s} \sum_{x_r} \int (u_{\text{obs}}(x_s, x_r, t) - u_{\text{cal}}(x_s, x_r, t))^2 \, dt$$

(1)

where $u_{\text{obs}}$ and $u_{\text{cal}}$ represent the seismic field and modeling data recorded at time $t$ for source point $x_s$ and receiver point $x_r$, respectively. $m$ stands for the velocity model parameters.

Using an adjoint state method (Plessix 2006; Fichtner & Trampert 2011), the gradient of the misfit function (1) with respect to velocity model parameters can be obtained as follows:

$$V_m \chi \delta m = \langle \delta u | V_m B(u) \delta m \rangle$$

(2)

where $B$ is the forward-modeling operator. The source wave field $u$ and adjoint wave field $\delta u$, which make up the wave equation and adjoint wave equation, can be expressed as:

$$B(m)u = S(x_s, t)$$

(3)

and

$$\hat{B}(m)\hat{u} = S_{\text{adj}}(x_s, x_r, t)$$

(4)

where $S$ and $S_{\text{adj}}$ represent the source and adjoint source, respectively, and $\hat{B}$ represents the adjoint forward-modeling operator. The adjoint source $S_{\text{adj}}$ of the L2-norm-based misfit function can be expressed as:

$$S_{\text{adj}}(x_s, x_r, t) = u_{\text{obs}}(x_s, x_r, t) - u_{\text{cal}}(x_s, x_r, t)$$

(5)

Gradient-based algorithms are optimization algorithms that are commonly used to update the velocity during FWI. They start by searching on the basis of reasonably accurate initial input, then update the model using gradient information at a current point. The velocity is updated using the following:

$$m_{k+1} = m_k - \alpha_k H_k V_m \chi$$

(6)

where $k$ refers to the number of iterations and $H_k$ represents the modification of the gradient term, e.g. a conjugate gradient or a Limited Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm. $\alpha_k$ is the step length, which can be determined by a quadratic line-search method (Köhn 2011).

As mentioned above, conventional time-domain FWI tends to get stuck in the local minima of the misfit function when there is predominantly high-frequency data. Multi-scale inversion, proceeding from low to high frequencies sequentially, improves the chances of reaching a global minimum for the misfit function, because low-frequency data is more linear with respect to model perturbations. Using a Wiener filter, Boonyasiriwat et al. (2009) developed an efficient multi-scale FWI strategy in the time domain. The Wiener filter in the frequency domain can be obtained as follows:

$$W_{\text{Wiener}}(w) = \frac{S_{\text{target}}(w)S_{\text{original}}^\dagger(w)}{|S_{\text{original}}(w)|^2 + \varepsilon^2}$$

(7)

where $S_{\text{original}}$ refers to the spectrum of the original true source wavelet and $S_{\text{target}}$ represents the spectrum of the low-frequency target wavelet at the angular frequency $w$, adding the small amount $\varepsilon$ to prevent the denominator from becoming 0. The Wiener filter can transform original high-frequency data into more desirable low-frequency data under the condition of knowing the true source wavelet. The transformed data can be obtained from:

$$U'(w) = W_{\text{Wiener}}(w)U(w)$$

(8)

where $U(w)$ is the spectrum of the original seismic data.

A Wiener filter with extended frequency selection (Boonyasiriwat et al. 2009) can be used in combination...
with multi-scale FWI in the time domain to obtain higher computational efficiency and better convergence performance than is typically found when using conventional FWI. However, low-pass Wiener filters need information from the original source wavelet and this is usually unavailable in real cases. As a result, complicated source wavelet estimation has to be used to get the source wavelet (Zhou et al. 1997; Yu et al. 2014, 2016). In order to avoid this cumbersome and costly source wavelet estimation issue, we have modified the low-pass Wiener filter so as to be able to use a reference trace based on a convolution model. A reference trace chosen from the seismic data can be expressed as a product of the source wavelet and a reflection coefficient in the frequency domain:

\[ U_{\text{ref}}(w) = S(w)R_{\text{ref}}(w) \]  

(9)

where \( S \) represents the source wavelet and \( R_{\text{ref}} \) represents the reflection coefficient of the reference trace. Similarly the seismic data can be written as follows:

\[ U(w) = S(w)R(w) \]  

(10)

where \( R \) represents the reflection coefficient of the seismic data in the frequency domain. A Wiener filter using the reference trace from the seismic data can then be defined as:

\[ W_{\text{ref}}(w) = \frac{S_{\text{target}}(w)S^*(w)R_{\text{ref}}^*(w)}{|S(w)R_{\text{ref}}(w)|^2 + \epsilon^2} \]  

(11)

Using the Wiener filter seismic data yields the following:

\[ U'(w) = \frac{S_{\text{target}}(w)S^*(w)R_{\text{ref}}^*(w)}{|S(w)R_{\text{ref}}(w)|^2 + \epsilon^2} R_{\text{ref}}(w) \]

\[ S(w)R(w) = \frac{S_{\text{target}}(w)R_{\text{ref}}^*(w)}{|R_{\text{ref}}(w)|^2 + \epsilon^2} R(w) \]  

(12)

In this way, the source wavelet term is offset in the filtered data. We can compute the Wiener filter for the calculated and observed data separately. Thus, the filtered data becomes:

\[ U'_{\text{cal}}(w) = \frac{S_{\text{target}}(w)R_{\text{ref}}^*(w)}{|R_{\text{ref}}(w)|^2 + \epsilon^2} R_{\text{cal}}(w) \]

\[ U'_{\text{obs}}(w) = \frac{S_{\text{target}}(w)R_{\text{ref}}^*(w)}{|R_{\text{ref}}(w)|^2 + \epsilon^2} R_{\text{obs}}(w) \]  

(13)

where the subscripts \( \text{cal} \) and \( \text{obs} \) represent the calculated and observed data, respectively.

The misfit function can be defined as the L2-norm of the distance between the filtered observed and synthetic data, i.e.:

\[ J = \| U'_{\text{cal}} - U'_{\text{obs}} \|_2^2 \]  

(14)

This removes the effect of the sources from the misfit function. Using an adjoint state method (Plessix 2006), the corresponding adjoint source of the misfit function can be expressed as:

\[ S_{\text{adj}} = W_{\text{cal}}^\dagger \frac{\partial W_{\text{cal}}}{\partial U_{\text{cal}}} \frac{U_{\text{cal}}^\dagger (U'_{\text{cal}} - U'_{\text{obs}})}{U_{\text{cal}}^\dagger (U'_{\text{cal}} - U'_{\text{obs}})} \]  

(15)

where \( W_{\text{cal}} \) represents the modified Wiener filter for the calculated data. In this paper, the second term of the adjoint source in (9) can be dropped for the following reasons: (i) If we choose a single trace as the reference trace, the second term is back-propagated only at the reference trace point. The influence of this on the gradient is comparatively small and plays the role of a singular wave field. (ii) If an average trace reference is chosen in a multi-source inversion experiment, the effect of the second adjoint source term is much weaker than that of the first term because the super-shot is summed across the receiver and is divided by the number of receivers. In addition, as shown by Choi & Alkhalifah (2011), the choice of reference trace can have a significant effect upon the inversion result. For independent shot inversion, we suggest choosing the nearest offset trace as the reference (Choi & Alkhalifah 2011). The choice of the reference trace for multi-source experiments can be more complex. We will be looking at this separately in the next section.

### 2.2. Encoded multi-source waveform inversion

A randomly encoded multi-source inversion strategy is an effective way of improving the computational efficiency of FWI. The objective function of an encoded multi-source waveform inversion is:

\[ \chi(m) = \left( \sum_{n=1}^{N_e} u_{\text{cal}}(m : e_n \otimes s_n) - \sum_{n=1}^{N_e} e_n \otimes u_{\text{obs}} \right)^2 \]  

(16)

where \( e_n \) is the sequence of the coding matrices, \( \otimes \) denotes convolution in the time domain and \( N_e \) represents the number of shots per super-gather. The polarity coding sequence \( e_n \) satisfies or approximately satisfies the following condition: \( ee^T = I \). With an effective cross-talk noise suppression strategy, the inversion result of multi-source FWI is comparable to that of independent source FWI, while the computational efficiency is significantly improved. However, it should be noted that source estimation is more difficult for multi-source FWI for the following reasons: (i) The common shot data are summed up by a random encoding function, making the source signature of early arrivals more obscure. (ii) The back-propagated source is different at each iteration step and cross-talk noise is present during the process of random summation. This complicates the inversion of the source wavelet.
In general, a few additional separate source inversions are needed for the source estimation. This is time-consuming when used for a multi-source inversion scheme. Therefore, in this paper, we adopt a multi-scale source independent strategy with encoded multi-source FWI to avoid complicated and time-consuming source estimation and to enhance the accuracy of the result.

3. Example

A modified Marmousi model (see figure 1a), which is complex and rich in reflectors, was used to test the effectiveness of the proposed method. The target area was $4.6 \times 1.5$ km with 10 m cells and the velocity varied from 1.5 to 5.5 km s$^{-1}$. Forty-six shots were spaced out at 100 m intervals to avoid aliasing artifacts and receivers were placed at every grid point on the surface. A constant gradient velocity model, increasing from 1.5 to 4 km s$^{-1}$ (which is lower than a true model on average), was used as the initial model for the inversion (figure 1b). The velocity in the water layer was 1500 m s$^{-1}$ during the inversion. A staggered grid and explicit finite-difference method, with eighth-order accuracy in the space and second-order accuracy in the time domain, was adopted for the forward modeling and adjoint calculation.

In order to verify the effects of the source wavelet on the inversion result, observed and modeled data using different source wavelets were generated by different source wavelets. The observed data was generated with a Gaussian wavelet and true model. A Ricker wavelet with a peak frequency of 10 Hz was chosen as the source wavelet for the modeled data. A comparison of the two source wavelets is shown in figure 2.

To assess some of the features of the proposed source independent FWI strategy, we took the source term of the Ricker wavelet and compared it to the adjoint source term produced by a conventional strategy. Figure 3a shows the observed common shot-gather for the true model using the Gaussian wavelet and figure 3b shows the modeled common shot-gather for the initial model using the Ricker wavelet. Figure 3c shows the adjoint source for the conventional FWI, which contains a strong direct wave residual even though the velocity for the water layer in the initial model was the same as it was in the real model. This is because the source signature difference led to a direct wave residual in the adjoint source. As the adjoint source term in the proposed method (figure 3d) does not contain direct waves, just reflected and refracted waves, the difference between the source wavelets in the modeled and observed data does not matter. The zero offset trace was chosen as the reference trace and the target wavelet was the Ricker wavelet with the peak frequency of 10 Hz, which is the same as the source wavelet term for the modeled data. There was also some noise in the adjoint source because of the effect of the Wiener filter. However, the energy of the noise was too small to affect the gradient term. A comparison of the adjoint sources shows that the proposed method can eliminate the source signature effect, despite the absence of any knowledge about the source.

The inversion result for a simulation using conventional FWI with the Ricker wavelet as the source term is shown in figure 4a. There were some strong inversion noises in the shallower region, resulting from the strong direct waves in the adjoint source (see figure 3c). For the deeper region, only a few shadows of the structure were added to the background model because the energy of the direct waves in the adjoint source was much stronger than that of the reflections. To verify that our method was independent of the source wavelet term, we applied the inversion strategy outlined above. Once again, the zero offset trace was chosen as the reference trace and the target wavelet was the Ricker wavelet with the peak frequency of 10 Hz, as with the
modeled data. The inversion result is shown in figure 4b. As the initial linear constant-gradient model was significantly different from the true model, the FWI suffered from a cycle-skip issue, which is reflected by the low-velocity artifacts, so the inversion result for the deeper structure was not good enough. However, when compared to the inversion result for the conventional FWI, it can be seen that the strong inversion noise was eliminated and the inversion result in the shallower
Figure 4. The inversion results for: (a) conventional FWI and (b) FWI using a source-independent inversion strategy with a zero offset trace as the reference trace.

Figure 5. The inversion results for conventional multi-scale FWI using a Ricker wavelet as the original source wavelet: (a) where a 5 Hz Ricker wavelet was the target wavelet and (b) the final inversion result.

Figure 6. The inversion results for the proposed multi-scale source-independent FWI using a zero offset trace as the reference trace; (a) where a 5 Hz Ricker wavelet was the target wavelet and (b) the final inversion result.

We realized that the nonlinearity of the misfit function for the high-frequency data and the inaccuracy of the initial model were the causes of the failure of the inversion. To improve the inversion result, we therefore employed a time domain multi-scale inversion strategy. A Ricker wavelet with a peak frequency ranging from 3 to 11 Hz in 2-Hz intervals was chosen as the target wavelet for different stages. An inverted model of the different frequency bands was obtained after five iterations and the inversion result of the current stage was used as the input for the following stage. The difference between the conventional method and the proposed method lies how the Wiener filter is calculated, with a source wavelet being used for the conventional method, while the proposed method uses a reference trace.

First of all, it was assumed that the original wavelet was a Ricker wavelet with a peak frequency of 10 Hz. The observed data was low-pass filtered using a conventional Wiener filter and the model was updated using the proposed time domain multi-scale FWI method. Figure 5 parts a and b show the inverted models obtained by replacing the target source wavelet with 5 and 11 Hz Ricker wavelets, respectively. The result after the second stage was dominated by inversion noise in the shallower region, which was the result of a strong direct wave residual in the adjoint source. For the deeper region, virtually nothing was updated. As a consequence, the inversion result after the last stage still produced an unsatisfactory velocity estimation. Although there were some updates in the middle to the deeper region, many layers were clearly dislocated and the resolution was very poor. However, when both the observed and modeled data were filtered by the new Wiener filter using the zero offset trace as the reference trace, the source wavelet effect was effectively eliminated, as shown in figure 3. We then used the multi-scale source independent inversion strategy to invert the velocity model. Figure 6 parts a and b show the inverted models after using the 5- and 11-Hz Ricker wavelet as the
target source wavelet. The inversion result after the second stage (figure 6a) proved to have inverted the larger part of the structure of the true model and the final inversion result (figure 6b) revealed even more detailed information. Thus, the inverted model managed to converge smoothly to the right answer. Calculating the low-pass Wiener filter with the reference trace successfully avoided inversion noise. By recovering process more wave-number components of the velocity structure the multi-scale inversion strategy is able to refine the inversion result and slowly reveal all of the data.

To further improve the computational efficiency of the proposed method, we combined the multi-scale source independent FWI with a multi-source strategy. Separateshotgathers were summed into a super-shot-gather by a random encoding function. The source functions were also gathered into a super-shot function by using the same encoding function. Following the approach advocated by Krebs et al. (2009), the encoding function was established as a matrix containing $-1$ or $+1$ randomly. The encoding function was randomly varied across each iteration to suppress cross-talk noise. The inversion parameters were the same as they were for the earlier multi-scale inversion, while the number of iterations for each frequency band became 20.

The adjoint source for the proposed method was then compared with the conventional one. The target source wavelet of the Wiener filter was a Ricker wavelet with a peak frequency of 3 Hz. Figure 7a shows the adjoint source corresponding to the L2-norm based misfit function where it was assumed that the wavelet was a Ricker wavelet with a peak frequency of 10 Hz. As the original source wavelet used by the Wiener filter was different from the real source term (which was a Gaussian wavelet), the adjoint source contained strong direct wave residuals. However, when a reference trace was used to calculate the Wiener filter, the obtained adjoint source (figure 7c) was similar to that produced by using the real source term (figure 7a). The adjoint source term for the proposed method did not contain direct waves, only reflected and refracted waves. The reference trace was chosen to be the weighted average of the traces corresponding to the source points, hereafter termed the weighted average of the zero offset traces. The weighting function was chosen to be the encoding function. Some noise still remained in the adjoint source after use of the Wiener filter, but its energy was relatively small so it could be ignored. The frequency component of the adjoint source calculated by equation (15) was also lower than it was in figure 7a.

The multi-source multi-scale FWI strategy was also used with the exact source term being set as the reference for comparison. The inversion result for this is shown in figure 8. Overall, the inversion result recovered most of the structure of the velocity model successfully, which is comparable to the inversion result obtained when using conventional independent source FWI. However, we found the inversion result was noisier than it was for the separate source FWI, with a poorer resolution. This was largely a result of the cross-talk between the different sources and an insufficient number of iterations. By using a more sophisticated encoding function or some kind of preconditioning such as smoothing, it would be
possible to eliminate this noise effectively. As we are mainly concerned with the influence of the source term on the inversion result, a discussion of cross-talk noise suppression is outside the scope of this paper.

After this, conventional multi-source multi-scale FWI was used to update the model. The original source used for the Wiener filter was a Ricker wavelet with a peak frequency of 10 Hz. The inversion result is shown in figure 9. Here, we can observe that a strong direct wave residual in the adjoint source resulted in inversion noise in the shallower region. There was also insufficient updating from the middle to the deeper region. By contrast, figure 10 shows an updated velocity model using the multi-source multi-scale source-independent FWI proposed in this paper. Here, instead of using the source wavelet, a reference trace was used to calculate the Wiener filter, thus avoiding the difficulty of source signature estimation. The reference trace was chosen to be the weighted average of the zero offset traces. Without the presence of a strong direct wave residual in the adjoint source term, the three main faults were largely resolved and the reflectors were closer to their real position. The middle to deeper region in the inversion result also appears to be smoother than it was for the conventional multi-source inversion using the exact source (figure 8). Overall the inversion result for the multi-source multi-scale FWI was significantly better than the source independent inversion strategy.

Figure 11 shows the data errors for the different encoded multi-source FWI methods using the different source wavelets. It can be seen that the convergence rate for the multi-source encoded FWI using the true source wavelet was much faster than the others in the earlier stages, while the error curve begins to fluctuate for the later iterations. To synthesize the different super-shot-gathers, the encoding function was different for each iteration, which had a significant impact on the convergence curves. The error curve for the FWI using the wrong source wavelet did not converge at all. Although, the convergence speed for the proposed multi-source multi-scale source independent FWI was slower than the method using true source wavelets, it finally converged to 20% of the initial error, illustrating its stability and effectiveness. As the initial error in the proposed source independent FWI was quite large, the final absolute error was larger than the FWI using a true source wavelet.

Table 1 shows a comparison of the calculation time for various inversion methods (the CPU parameters were: an Intel(R) Xeon(R) CPU E5–2650 v.2 at 2.60 GHz; a double CPU with 32 threads; 64 Gb of RAM). The multi-source FWI displayed notably better computational efficiency than a traditional multi-scale FWI, with it also using fewer threads and less memory. The computational time for the source independent multi-source FWI was largely comparable with the multi-source FWI using a known wavelet.

4. Discussion

In the previous section, we illustrated the accuracy and efficiency of the proposed method when applied to synthetic data generated by a modified Marmousi model. However, we realize that there are many factors that can affect the quality of the final reconstructed model. As with conventional source independent inversion strategies, the choice of the reference trace can significantly influence the inversion result (Choi & Alkhalifah 2011). In this section, we will examine the influence of the choice of reference trace on the inversion result and the robustness of the proposed inversion strategy in relation to noise.

Apart from the weighted average of the zero offset traces, there are several other alternatives that might be used for reference. These include an average trace of the super-gather
across the receivers or a single trace. Figure 12 parts a and b show the inversion results using a single trace corresponding to the first shot position and the average trace of the super-shot as references, respectively. Both of these inversion results are inferior to the result when using the weighted average of the zero offset traces as a reference. To further study the potential reasons for this, we compared the different reference traces, the results of which can be seen in figure 13. Figure 13a shows the use of a single trace at the first source point. Note how there were fluctuations after the direct wave, resulting from the random summation of the super-shots. The reflectivity of the reference trace, which is used in the denominator of our proposed Wiener filter, is very complicated, which risks making the Wiener filter unstable. The average trace of the whole super-gather across receivers (figure 13b) is even more complicated and reveals little information about the source signature. The waveform during averaging can be canceled out, because all traces are summed up by random polarizations determined by the encoding function. However, because the weighting function can correct the polarization of the traces corresponding to the shot position, the weighted average of the zero offset traces (figure 13c) not only shows the character of the direct wave but also suppresses later noise.

The choice of frequency has a large impact on the multi-scale inversion effect. We used 6 Hz in the numerical experiments as the initial peak frequency. The inversion results for this are shown in figure 15. The independent source method had good applicability and was able to inverse the general structure, but a local extremum problem began to appear because the background field had an inaccurate inversion. The method proposed in this paper can eliminate the influence of the source wavelet, but it cannot inverse the background field effectively if there is high peak frequency.

In real cases, a zero offset trace of the different sources is not always available because the scope for acquiring it is limited. We therefore also investigated the inversion result when
Figure 13. Comparison of different reference traces: (a) the trace at the first shot position; (b) the average trace of the super-gather; (c) the weighted average of the zero offset traces and (d) the weighted average of the 50 m offset traces.
5. Conclusion

In order to avoid complex source signature estimation and improve inversion stability, we developed a multi-scale
source-independent inversion method by modifying the Wiener filter and by using a reference trace. Here, the source wavelets for observed and modeled wave-fields are both eliminated in the misfit function. When combined with a multi-source inversion strategy, the proposed method was found to effectively decrease computational burden. Numerical tests based on a synthetic model that used data generated by a modified Marmousi model verified that the proposed method is capable of recovering the velocity structure without source wavelet information. When analyzing the influence of the reference trace on the inversion result, we realized that the weighted average of the common offset traces is a better choice for the reference trace than a single or average trace. We have also found that the proposed method is able to resist noise.

As we mainly focused on the influence of the accuracy of the source signature on the inversion result, a relatively simple encoding strategy was adopted, leading to some cross-talk noise in the inversion result. To enhance the quality of the inversion result, a more complicated encoding strategy would need to be adopted.

Acknowledgements

We appreciate comments from editors and reviewers that significantly improve the manuscript. We appreciate the open source SU software for providing convenient research platforms. This research is supported by the National Natural Science Foundation of China (41720104006), the National Gas and Oil project (2016ZX05002-005-07HZ, 2016ZX05014-001-008HZ), the Fundamental Research Funds for the Central Universities (17CX06033), the Sinopec Project (G5800-17-ZS-WX004), and the Zhongyuan Oilfield Science and Technology Project (2018XQ01-13).

Conflict of interest statement. None declared.

References

Anagaw, A.Y. & Sacchi, M.D., 2014. Comparison of multifrequency selection strategies for simultaneous-source full-waveform inversion, Geophysics, 79, R165–181.

Ben-Hadj-Ali, H., Operto, S. & Virieux, J., 2011. An efficient frequency domain full waveform inversion method using simultaneous encoded sources, Geophysics, 76, R109–124.

Berkhout, A. & Blacquire, G., 2011. Blended acquisition with dispersed source arrays, the next step in seismic acquisition?, 81st Annual International Meeting, SEG, Expanded Abstracts, 16–19.

Boonyasiriwat, C., Valasek, P., Routh, P., Weiping, C., Gerard, T.S. & Brian, M., 2009. An efficient multiscale for time-domain waveform tomography, Geophysics, 74, WCC59–68.

Brossier, R., 2011. Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance, Computers & Geosciences, 37, 444–455.

Bunks, C., Saleck, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, Geophysics, 60, 1457–1473.

Choi, Y. & Alkhalifah, T., 2011. Source-independent time-domain waveform inversion using convolved wavefields: Application to the encoded multisource waveform inversion, Geophysics, 76, R125–134.

Choi, Y., Shin, C., Min, D.J. & Ha, T., 2005. Efficient calculation of the steepest descent direction for source-independent seismic waveform inversion: An amplitude approach, Journal of Computational Physics, 208, 455–468.

Fichtner, A. & Trampert, J., 2011. Hessian kernels of seismic data functionals based upon adjoint techniques, Geophysical Journal International, 185, 775–798.

Frazer, L.N. & Sun, X.H., 1998. New objective functions for waveform inversion, Geophysics, 63, 213–222.

Guitton, A. & Diaz, E., 2012. Attenuating crosstalk noise with simultaneous source full waveform inversion, Geophysical Prospecting, 60, 759–768.

Herrmann, F.J., Erlangga, Y.A. & Lin, T.T., 2009. Compressive simultaneous full-waveform simulation, Geophysics, 74, A35–40.

Hu, Y., Han, L.G., Xu, Z. & Zhang, T.Z., 2017. Demodulation envelope multi-scale full waveform inversion based on precise seismic source function, Chinese Journal Of Geophysics (in Chinese), 60, 1088–1105.

Köhnn, D., 2011. Time Domain 2D Elastic Full Waveform Tomography, Kiel: Kiel University.
Krebs, J.R., Anderson, J.E., Hinkley, D., Neelamani, R., Lee, S., Baumstein, A. & Lacasse, M.D., 2009. Fast full-wavefield seismic inversion using encoded sources, Geophysics, 74, WCC177–188.

Kwon, T., Seol, S.J. & Byun, J., 2015. Efficient full-waveform inversion with normalized plane-wave data, Geophysical Journal International, 201, 53–60.

Lee, K.H. & Kim, H.J., 2003. Source-independent full-waveform inversion of seismic data, Geophysics, 68, 2010–2015.

Li, Q.Y., Huang, J.P. & Li, Z.C., 2017. Cross-correlation least-squares reverse time migration in the pseudo-time domain, Journal of Geophysics and Engineering, 14, 841–851.

Liao, J.P., Wang, H.Z. & Ma, Z., 2009. 2-D elastic wave modeling with frequency-space 25-point finite-difference operators, Applied Geophysics, 6, 259–266.

Liao, J.P. et al., 2017. Application of frequency-dependent traveltime tomography to 2D crosswell seismic field data, Journal of Environmental and Engineering Geophysics, 22, 421–426.

Luo, Y. & Schuster, G.T., 1991. Wave-equation traveltime inversion, Geophysics, 56, 645–653.

Operto, S. et al., 2015. Efficient 3-d frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation, Geophysical Journal International, 202, 1362–1391.

Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, Geophysical Journal International, 167, 495–503.

Pratt, R.G., 1999. Seismic wavefield inversion in the frequency domain—Part 1: Theory and verification in a physical scale model, Geophysics, 64, 888–901.

Rao, Y. & Wang, Y., 2005. Crosshole seismic tomography: working solutions to issues in real data travel time inversion, Journal of Geophysics and Engineering, 2, 139.

Rao, Y. & Wang, Y., 2017. Seismic waveform tomography with shot-encoding using a restarted L-BFGS algorithm, Scientific Reports, 7, 8494.

Rao, Y., Wang, Y. & Han, D., 2018. Seismic waveform tomography with simplified restarting scheme, IEEE Geoscience and Remote Sensing Letters, 99, 1–5.

Romero, L.A., Ghiglia, D.C., Ober, C.C. & Morton, S.A., 2000. Phase encoding of shot records in prestack migration, Geophysics, 65, 426–436.

Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies, Geophysics, 69, 231–248.

Sun, M., Yang, J., Dong, L. & Huang, C., 2017. Density reconstruction in multiparameter elastic full-waveform inversion, Journal of Geophysics and Engineering, 14, 1445–1462.

Tao, Y. & Sen, M.K., 2015. Frequency-domain full waveform inversion with a scattering-integral approach and its sensitivity analysis, Journal of Geophysics and Engineering, 10, 5008.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, Geophysics, 49, 1259–1266.

Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, Geophysics, 74, WCC1–26.

Xie, F., Huang, J.P. & Li, Z.C., 2018. Full waveform inversion in the frequency domain with an efficient iterative solver, Chinese Journal Of Geophysics (in Chinese), 61, 3346–3355.

Yu, H., Zhang, D. & Huang, Y., 2016. Application of early arrival waveform inversion with pseudo-deconvolution misfit function by source convolution, Inverse Problems in Science and Engineering, 25, 57–72.

Yu, H., Zhang, D. & Wang, X., 2014. Application of weighted early-arrival waveform inversion to shallow land data, Journal of Applied Geophysics, 102, 68–76.

Zhou, C., Schuster, G.T., Hassanzadeh, S. & Harris, J.M., 1997. Elastic wave equation traveltime and waveform inversion of Crosswell data, Geophysics, 62, 853–868.