Micro-seismic imaging using a source function independent full waveform inversion method

Hanchen Wang & Tariq Alkhalifah

SUMMARY
At the heart of micro-seismic event measurements is the task to estimate the location of the source micro-seismic events, as well as their ignition times. The accuracy of locating the sources is highly dependent on the velocity model. On the other hand, the conventional micro-seismic source locating methods require, in many cases manual picking of traveltime arrivals, which do not only lead to manual effort and human interaction, but also prone to errors. Using full waveform inversion (FWI) to locate and image micro-seismic events allows for an automatic process (free of picking) that utilizes the full wavefield. However, full waveform inversion of micro-seismic events faces incredible nonlinearity due to the unknown source locations (space) and functions (time). We developed a source function independent full waveform inversion of micro-seismic events to invert for the source image, source function and the velocity model. It is based on convolving reference traces with these observed and modeled to mitigate the effect of an unknown source ignition time. The adjoint-state method is used to derive the gradient for the source image, source function and velocity updates. The extended image for the source wavelet in Z axis is extracted to check the accuracy of the inverted source image and velocity model. Also, angle gathers is calculated to assess the quality of the long wavelength component of the velocity model. By inverting for the source image, source wavelet and the velocity model simultaneously, the proposed method produces good estimates of the source loca-
tion, ignition time and the background velocity for synthetic examples used here, like those corresponding to the Marmousi model and the SEG/EAGE overthrust model.

**Key words:** Full waveform inversion; Inverse theory; Joint inversion.

1 INTRODUCTION

Hydraulic fracturing is one of the most commonly used methods to enhance oil and gas extraction. This is especially prevalent in those reservoirs having dense rock such as shale. A typical application is to inject water or other liquid into the reservoir area to create high-pressure conditions that eventually crack the dense rock in order to make the oil and gas flow more freely. The fracking process causes micro-seismic events that can be monitored through sensors in a well or on the earth surface. Usually a treatment well is drilled into the reservoir so that engineers can inject fluid into the layer. Fractures are, then, formed and expanded. Micro-seismic events usually occur along these fractures. Receivers are set on the earth surface or in another monitoring well to record the signals of the micro-seismic waves. By processing the recorded data the locations of the micro-seismic events can be estimated. This allows us to monitor the fracturing process. Information about the development of fractures helps engineers optimize the injection strategy. Accurately locating the micro-seismic events is critical to engineer the hydraulic fracturing process (Warpinski, 2009).

There are many ways to find micro-seismic source locations. Some methods are based on traveltimes of both P-wave and S-wave arrivals (Eisner, et al., 2009; Waldhauser & Ellsworth, 2000). In the use of this technique the arrival times of the events in the recorded traces are manually determined. Through the relationship between distance, travel time and velocity, each arrival are smeared over a circle (in 2-D homogeneous case, for instance) with a radius that is related to velocity and travelt ime. By applying such process to several traces the circles of all traces, granted the velocity is accurate, intersect at one point. This is defined as the source location. However, travelt ime based methods suffer in noisy data resulting in errors in determining the arrival time that will eventually lead to incorrect micro-seismic source locations. Other methods are based on migration techniques. An example is time reverse migration (McMechan, George A., 1982; Gajewski, D. & E. Tessmer, 2005; Artman, et al., 2010). These methods use not only the traveltimes but also the waveform information. Using a velocity model one can use the wave equation to back propagate the observed data and find the sources in the time reversed wavefield. Although other modeling parameters,
like attenuation Zhu, Tieyuan (2014), have been taken into consideration, migration based methods require a good velocity model to correctly focus the energy as well as a dense and uniform acquisition system. Save (2011) used an interferometry based method to push the limits of these assumptions. However, he assumed the onsets of the sources are known which is not satisfied in micro-seismic imaging.

Another sort of micro-seismic imaging method is based on full waveform inversion. Kamei, et al. (2017) updated the velocity model via an FWI algorithm but they assumed only small time-lapse velocity change happened and a good prior information was available. The rest of the velocity model remained the same given by the initial velocity model. Sun, et al. (2016) used the conventional FWI technique to update both the source and the velocity model. By cross-correlation the different parts of decomposed back-propagated wavefields, the onset of source could be defined. However, this method still requires a relatively good starting model. Kaderli, et al. (2015) used an FWI based technique to update the spatial and temporal components of the source. In their paper a homogeneous model is assumed. This is not a realistic assumption in real world applications. Behura, J. (2016) also separated the source into spatial and temporal terms. But he only updated the source spatial distribution, which might have errors if the source onset or function is unknown. In this research, a gradient based FWI method is used to update the spatial component of the source and a source function independent FWI approach is used to update the velocity model. This FWI approach works well with any given source wavelet, as it is based on convolved wavefields (Choi & Alkhalifah, 2011 ). An extended image is also used to check whether the inverted velocity and source image are good. By inverting for the source spatial and time components and the velocity model, the method is able to find the source correctly even with an initial velocity model that is very different from the true one. Applying the method on a selected portion of the Marmousi model elucidates the features of inverting for the source image, source function and velocity iteratively. The angle domain common image gather (ADCIG) is computed to show the effectiveness of the approach.

2 SOURCE FUNCTION INDEPENDENT FWI

2.1 Conventional FWI On Source Terms

Acoustic wave propagation in a constant density 2D medium obeys the following wave equation:

$$\frac{1}{c^2(x,z)} \frac{\partial^2 u(x,z,t)}{\partial t^2} - \nabla^2 u(x,z,t) = f(x,z)w(t)$$

(1)
where \( u(x, z, t) \) is the acoustic wavefield in 2D media described location wise by \( x \) laterally, and \( z \) in depth, \( c(x, z) \) is the velocity, while \( f(x, z) \) and \( w(t) \) are source components in space and time, \( t \), respectively. This decomposition of the source assumes that the time and space dependency are decoupled, which may be a reasonable approximation considering our source function independent implementation.

We usually start our investigation with a crude guess or evaluation of the initial velocity model, \( c_0(x, z) \). We can estimate an initial \( w(t) \) with the available initial model using time reversal migration. Thus, through an iterative inversion we update these initial guesses of the micro-seismic event and medium. Considering the observed data of the micro-seismic events on the earth surface (or a well), \( d_{\text{obs}} \), we can devise an iterative inversion by defining the optimization problem based on the classic least-square minimization of

\[
E(f, w, c) = \sum_{i} \| \delta d_i \|^2 = \sum_{i} \| d_{\text{pred}}^i - d_{\text{obs}}^i \|^2,
\]

with respect to the source image, the source wavelet and the velocity model, where \( \delta d_i \) is the data residual and \( d_{\text{pred}}^i \) is the predicted data generated from the current velocity model and source. \( nr \) is the total receiver number and \( i \) means the \( i \)th trace.

Kaderli, et al. (2015) have derived the general update direction, given first by the gradient with respect to \( f(x, z) \), where we scaled the gradient by the pseudo Hessian inverse of \( f(x, z) \), \( H_f^{-1} \), courtesy of the adjoint-state method (Plessix, 2006), is as follows:

\[
H_f^{-1} \frac{\partial E}{\partial f} = \int_{0}^{T} \frac{w(t) \cdot u_r(x, z, t)}{|w(t)|^2} dt,
\]

where \( u_r(x, z, t) \) is the adjoint (residual) wavefield. The denominator \( |w(t)|^2 \) works as the pseudo Hessian to the gradient. \( T \) is defined as the total propagating time and we ensure is long enough so that the source onset is included. The update direction with respect to \( w(t) \) is,

\[
H_w^{-1} \frac{\partial E}{\partial w} = \int_{x} \int_{z} \frac{f(x, z) \cdot u_r(x, z, t)}{|f(x, z)|^2} dx dz,
\]

In this algorithm, we assume \( w_0(t) \) is given by the time reversed method (specifically its time). The observed data is simply back propagated into the initial velocity model, in which we apply the maximum energy imaging condition to locate the source. The maximum energy imaging
condition can be described as

$$I_{\text{max}}(x, z) = \int u_r(x, z, t) \delta(t' ; A = A_{\text{max}}) dt,$$  \hspace{1cm} (5)$$

where $\delta(t')$ is the delta function, and $t'$ is the time step when the wavefield amplitude $A$ equals to the maximum amplitude $A_{\text{max}}$.

### 2.2 Source Independent Velocity Gradient Derivation

We extract the wavefield at the maximum energy point, and it is our initial guess of the wavelet $w_0(t)$. However, the initial guess of $w(t)$, depending on the velocity, can be far from the true one, especially in its ignition time. If we directly use it in the inversion, we will have a big chance of converging to a local minima, as the resulting time shift will cause cycle skipping in the data fit. To avoid this problem, we utilize a convolution objective function (Choi & Alkhalifah, 2011), which is reasonably source function independent. In this case, the objective function is given by:

$$E_{\text{ind}}(c) = \sum_{i} ||r_i||^2 = \sum_{i} \left| d_{i}^{\text{pred}} \ast d_{\text{ref}}^{\text{obs}} - d_{i}^{\text{obs}} \ast d_{\text{ref}}^{\text{pred}} \right|^2,$$  \hspace{1cm} (6)$$

where $d_i$ means the data recorded at the $i$th receiver both in the predicted data and in the observed data; $d_{\text{ref}}^{\text{obs}}$ means a reference trace, which can be obtained simply by picking one of the traces in both the observed and synthetic data; and $r_i$ is the convolved data residual at the $i$th receiver location. In practice, we choose a near offset trace to be the reference trace. The symbol $\ast$ is the convolution operator. In this new objective function, the source wavelet is convolved in both terms equally, which means the effect of the delay in the source wavelet $w(t)$ is muted. The gradient of the convolution objective function is calculated by taking the derivative of Equation 6 with respect to the model parameters, $m(x, z) = \{ f(x, z), w(t), c(x, z) \}$. The gradient can then be written as:

$$\frac{\partial E_{\text{ind}}}{\partial c} = \sum_{i} \left[ \left( \frac{\partial d_{i}^{\text{pred}}}{\partial c} \ast d_{\text{ref}}^{\text{obs}} \right) \cdot r_i - \left( d_{i}^{\text{obs}} \ast \frac{\partial d_{\text{ref}}^{\text{pred}}}{\partial c} \right) \cdot r_i \right].$$  \hspace{1cm} (7)$$

In Choi and Alkhalifah’s paper, the gradient of this new objective function with respect to model parameters has already been derived in the case of conventional surface experiment. For the micro-seismic case, the derivation is quite similar to the surface one. The partial derivative of the new objective function in Equation 6 with respect to model parameters $m$ can be written as Equation 7.
For convenience, let’s calculate the gradient term by term. The first convolution term \( \left( \frac{\partial d_{i}^{\text{pred}}}{\partial m} * d_{i}^{\text{obs}} \right) \cdot r_{i} \) in Equation 7 can be rewritten in an integral form,

\[
\left( \frac{\partial d_{i}^{\text{pred}}}{\partial m} * d_{i}^{\text{obs}} \right) \cdot r_{i} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\partial d_{i}^{\text{pred}}}{\partial m} (t - \tau) d_{i}^{\text{obs}} (\tau) r_{i} (t) d\tau dt.
\]

Now let \( \eta = t - \tau \), thus, equation 8 can be written as,

\[
\left( \frac{\partial d_{i}^{\text{pred}}}{\partial m} * d_{i}^{\text{obs}} \right) \cdot r_{i} = - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\partial d_{i}^{\text{pred}}}{\partial m} (\eta) d_{i}^{\text{obs}} (t - \eta) r_{i} (t) d\eta dt.
\]

Since \( \frac{\partial d_{i}^{\text{pred}}}{\partial m} (\eta) \) is independent of \( t \) now, we may move it outside of the integration of \( t \), which gives,

\[
\left( \frac{\partial d_{i}^{\text{pred}}}{\partial m} * d_{i}^{\text{obs}} \right) \cdot r_{i} = - \int_{-\infty}^{+\infty} \frac{\partial d_{i}^{\text{pred}}}{\partial m} (\eta) \left[ \int_{-\infty}^{+\infty} d_{i}^{\text{obs}} (t - \eta) r_{i} (t) dt \right] d\eta.
\]

To obtain the partial derivative \( \frac{\partial d_{i}^{\text{pred}}}{\partial m} (\eta) \), we calculate the convolution between \( U_{x,i} \) and \( S_{x} \), where \( U_{x,i} \) is the adjoint wavefield of the \( i \)th backpropagated residual, \( S_{x} \) is the source wavefield ignited at a virtual source location (since we do not know the true source location) and \( x \) represents the model space location (Pratt, et al., 1998). Using the convolution form to rebuild Equation 8, we obtain

\[
\left( \frac{\partial d_{i}^{\text{pred}}}{\partial m} * d_{i}^{\text{obs}} \right) \cdot r_{i} = - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S_{x} (\eta - \tau) U_{x,i} (\tau) d\tau dt \left[ \int_{-\infty}^{+\infty} d_{i}^{\text{obs}} (t - \eta) r_{i} (t) dt \right] d\eta.
\]

Again let \( \eta - \tau = t \), then equation 11 becomes

\[
\left( \frac{\partial d_{i}^{\text{pred}}}{\partial m} * d_{i}^{\text{obs}} \right) \cdot r_{i} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S_{x} (t) U_{x,i} (\eta - t) dt \left[ \int_{-\infty}^{+\infty} d_{i}^{\text{obs}} (t - \eta) r_{i} (t) dt \right] d\eta = \int_{-\infty}^{+\infty} S_{x} (t) \left[ \int_{-\infty}^{+\infty} U_{x,i} (\eta - t) d_{i}^{\text{obs}} (t - \eta) r_{i} (t) d\eta \right] dt.
\]

Because of the reciprocity of \( U_{x,i} \), the inner integration \( \int_{-\infty}^{+\infty} U_{x,i} (\eta - t) d_{i}^{\text{obs}} (t - \eta) r_{i} (t) d\eta \) can be considered as a back propagation of the new data residual \( R_{i}^{(1)} = \int_{-\infty}^{+\infty} d_{i}^{\text{obs}} (t - \eta) r_{i} (t) dt \). Thus, the first term of the gradient is nothing but the equivalent imaging condition for reverse-time migration (RTM), except for the different back-propagating data \( R_{i}^{(1)} \).

A similar derivation can be applied to the second term of Equation 7, which is \( \left( d_{i}^{\text{obs}} * \frac{\partial d_{i}^{\text{ref}}}{\partial m} \right) r_{i} \).
We finally end up with

\[- \left( \frac{\partial d_{\text{obs}}}{\partial m} \right) \cdot r_i = \int_{-\infty}^{+\infty} S_x(t) \left[ \int_{-\infty}^{+\infty} U_{x,\text{ref}}(\eta - t) d_{\text{pred}}(t - \eta) r_i(t) \, d\eta \right] \, dt. \quad (13)\]

In Equation 13, the second term of the gradient is nothing but the backpropagation of the second residual

\[ R_i^{(2)} = - \int_{-\infty}^{+\infty} d_{\text{obs}}(t - \eta) r_i(t) \, dt , \]

which is similar to the backpropagation procedure in reverse time migration.

Thus, the gradient involves calculating the adjoint wavefields with the adjoint sources:

\[ R_i^{(1)} = d_{\text{obs}} \otimes \left( d_{\text{pred}} \ast d_{\text{obs}} - d_{\text{pred}} \ast d_{\text{obs}} \right) , \quad (14)\]

at the \( i \)-th receiver position, and

\[ R_i^{(2)} = - d_{\text{obs}} \otimes \left( d_{\text{pred}} \ast d_{\text{obs}} - d_{\text{pred}} \ast d_{\text{obs}} \right) , \quad (15)\]

at the reference traces position, where \( \otimes \) is the cross-correlation operator. The two cross-correlated seismograms are back-propagated at the same time in order to reduce the computational cost (equivalent to conventional FWI). Although the velocity update is independent of the source wavelet, we can still update the source wavelet with equation 4 in order to image the source.

Considering the multi nature of the inverted parameters, the inversion utilizes a nested approach, i.e. 1. inverting for the source image \( f(x,z) \) for several iterations using equation 3; 2. inverting for the background velocity \( c(x,z) \) for several iterations using equation 7, and 3. updating the source function \( w(t) \) using equation 4. Repeating the steps above until the gradients with respect to the three unknowns are small enough or the residual (Equation 2) meets our threshold. The whole workflow is shown in Figure 1, where the residual corresponds to \( \delta d \). In addition, the extended image is a powerful tool to check if the migration velocity is correct (Sava & Fomel, 2006). Here, a similar extended imaging condition is applied. Instead of extending the time axis in the imaging condition, the depth axis \( Z \) is extended when calculating for source wavelet \( w(t) \). Then Equation 4 can be modified as,

\[ \nabla_w E_{\text{ext}}(t, h) = \int_{x} \int_{z} f(x, z - h) \cdot u_r(x, z + h, t) \, dx \, dz , \quad (16)\]

where \( h \) is the offset in the \( Z \) direction (Symes, W. W., & James J. Carazzone, 1991). The reason of extending the depth axis is that the earth is mainly varying in the \( Z \) direction and in our case, measurement are done at the mostly horizontal surface of the Earth. This will
Figure 1. Workflow of the nested inversion, each iteration contains 3 steps, $f(x,z)$ update (Eq. 3), $c(x,z)$ update (Eq. 7) and $w(t)$ update (Eq. 4 with adjoint sources shown as Eq. 14 and Eq. 15).

lead to larger imaging errors in the $Z$ direction than that in $X$ direction. So extension in $Z$ induces larger difference between the images for a correct and incorrect velocity model. For demonstration, a homogeneous medium example is tested to show the extended image for $w(t)$ with both correct and incorrect velocity models, as shown in Figure 2. We can see that the maximum amplitude point, which is also the crossing point, is located at $h = 0$ (zero offset trace) when the velocity is correct. However, the maximum amplitude point is shifted left and upward when the velocity is lower and right and downward when the velocity is higher. That is to say, we may consider the velocity model to be correct if the maximum amplitude point is at the zero offset position. With this modified extended imaging condition, we can assess whether both the velocity model and the source image are correct.

Angle domain common image gathers (ADCIG) is another tool that can be utilized to

Figure 2. The extended images of $w(t,h)$ in a homogeneous model with (a) the true velocity; (b) 10% lower velocity and (c) 10% higher velocity.
check whether the migration velocity is correct (Sava & Fomel, 2003; Biondi & Tisserant, 2004). If the migration velocity is correct, the reflectors in the ADCIG will be flat and located at the right depth. If the migration velocity is incorrect, the reflectors will be curved upward (in the case of a lower velocity) or downward (in the case of a higher velocity). In micro-seismic cases, the unknown sources can be considered as scatters, in which they share a similar behavior to reflectors in migration. So by simply focusing on the common image point at the source location, the corresponding angle gather refers to the open angle between the transmission ray and the vertical axis. If the velocity is good enough, the ADCIG should be flat as well.

3 NUMERICAL EXAMPLES

We implement the source function independent full waveform inversion method on part of the Marmousi model, for demonstration. In the following examples, the initial $w(t)$ is obtained by the time reversed imaging method and the maximum energy imaging condition, which means simply back propagating the recorded surface data and extracting the $w(t)$ corresponding to the maximum energy time and location, as mentioned above. The initial $f(x, z)$ is set to 0 at all points in order to be consistent with the source gradient, which is equation 3. The same modeling code is used to generate and invert data.

3.1 Marmousi Model With Single Source

In the first experiment, a single shot case is assumed. A piece of the Marmousi model is used to represent what we refer to as the true velocity model, shown in Figure 3a. The model has 400 x 200 grid points with a 10m spatial interval in both $x$ and $z$ directions. The initial velocity model is linearly increasing with depth, shown in Figure 3b. The time interval is 1 ms. Every
Figure 4. Comparison between source functions (temporal component), red: initial, blue: true and pink: inverted, for the single source. True source locates at \{2.0km, 1.25km\}.

grid point on the top surface is assumed to be a receiver. A free surface is assumed on the top of the model while absorbing boundary condition is assumed on the other three boundaries.

A Ricker wavelet of 10Hz peak frequency is set to be the unknown source function. The true source position is located at \{2.0km, 1.25km\} in the model.

The starting \(f_0(x, z)\) is zero everywhere. In another word, no priori information of the source spatial component is given. Using the time reversal method mentioned above, the initial source wavelet \(w(t)\) is obtained by reversing the wavefield at the maximum energy point along time axis, shown as the red line in Figure 4. We can see the onset is later than the true one (blue).

With the nested approach, the misfit reached its minimum after repeating the steps for 150

Figure 5. The inverted (part of) Marmousi model at the 150th iteration with a single source.
iterations and in each iteration, each source term and the velocity model have been updated for 3 times. The inverted $c(x, z)$ results at 150th iteration is shown in Figure 5.

The inverted $f(x, z)$ result is shown in Figure 6. We have normalized the source image amplitudes as we combine them, because we separately invert for each source. The highest value in each source image indicates the source location. The initial $f(x, z)$ is set to zero, as was mentioned earlier. The inverted velocity model shows generally low resolution reflecting the peak frequency of the data and the limited illumination with a single source. The microseismic events are buried in the earth while the receivers are on the earth surface. Most of the energy in the observed data are transmission waves. In FWI, reflection waves update the high wavenumber parts of the model, while transmissions update the low wave number parts. So the gradient for the velocity updates is mostly focused on the smooth background. The artifacts in the velocity model are mostly caused by the limited information and aperture of the single shot case. If more sources are used in the inversion, the artifacts may be suppressed.

Figure 6. The inverted source image (spatial component) for the single source case.

Figure 7. Sketch depicting the ray paths of the recorded reflection energy, red lines: multi scattering (from top layers); blue lines: single scattering (from bottom layers). Explosion symbol refers to the source, and triangular symbols refer to the receivers on the earth surface.
Figure 8. The data observed (a), generated from initial model (b) and generated from inverted model. The data misfit $\delta d$ before inversion (d) and after inversion (e).

We may also find that the part which is below the source seems to have higher resolution than the part above the source. This is reasonable because the reflection energy is received after being reflected once, e.g. the blue line in Fig 7, while the reflection energy above the source can only be recorded after multi reflections, e.g. the red lines in Fig 7. That is to say, the energy from bottom reflections is higher than that from upper reflections. So the upper part of the model has less high wavenumber updates (reflections) compared to the lower part, which will result in the higher resolution in lower part model and poor resolution in upper part model.

Figure 9. Convergence curve of the data residual $\delta d$ in percentage.
Unlike the conventional surface survey having enough sources covering the whole survey area, micro-seismic sources are close to each other which means only certain area can be covered by the surface records. And this causes the huge difficulties in perfectly updating the whole model. As a result, we can only have the kinematic inversion (smooth) result accurately. We may obtain higher resolution for higher frequency events, beyond our 10 Hz peak frequency Ricker wavelet. Although the inverted velocity model is generally smooth, it still shows reasonably good structure of the model, such as the high velocity layers, which is located between 2.4 – 3.2 km in x and between 1.0 – 1.7 km in z. Also the border between the water layer and and the sediments is accurate in depth and shape, though it is a little bit smoothed. The inverted source image in Figure 6 shows an accurate source location at (2.0 km, 1.25 km), where we assume the center of the blue dot in the source image to be the inverted source location. There are some weak energy in the source image that is not focused to the source point. It is mainly caused by the misfit remained in the objective function and mostly by the limited aperture. One can either choose to ignore them or set some kind of constraint, for example a Laplacian filter, to reduce the influence. If higher frequency wavelet is used, the image should be sharpened in space because the width of the source image depends on the wavelength of the source wavelet under the assumption that the attenuation of medium is negligible, although in reality, medium attenuation will also affect the sharpness of the source image, which results in less resolution for deeper sources.

The inverted source wavelet is shown as the pink line in Figure 4. The true one (blue) and the initial one (red) are also shown in the figure for comparison. We can easily find that
Figure 11. Common image gather at the source horizontal location in (a) the initial model and (b) the inverted model.

The ignition time of the initial source wavelet is not accurate, which is shifted backward to 0.35 s while the true ignition time is 0.2 s (the max amplitude of the source Ricker wavelet). After the inversion, the inverted source wavelet (pink) is ignited at exactly the correct time as the true one. Although there are some small differences in the shape of the wavelet between the inverted one and the true one, it is still good enough for the calculation of $f(x, z)$ and $c(x, z)$. The difference may be also caused by the remaining misfit of the model parameter $m$. Figure 8 shows the data residuals at both the first iteration and the final (150) iteration. The original observed data (Figure 8a), the initial (Figure 8b) and final (Figure 8c) generated data are also shown. Pay attention to the first arrivals in all data sets. The first arrivals are cycle skipped between observed and initial data, which leads to the two separated first arrival events in initial data misfit. Since they are cycle skipped, the conventional FWI may not converge. Clearly, the data residual at the first iteration is much larger than the final iteration. After 150 iterations, the figure on the right shows a very small amplitude of data misfit. That is to say the approach we have proposed provides a velocity, source image and wavelet that predicts the data. The first arrivals of observed and inverted data are no longer cycle skipped.
That is why there is only one weak single first arrival event in the data misfit after the inversion. Figure 9 shows the convergence curve of the data residual. The residual drops about fifty-five percentage from the original one. The residual still going downward which means we may obtain better results if we run more iterations. Figure 10 shows the vertical extended image for $w(t)$ after the inversion. The maximum amplitude point is at the right place $(h = 0$ and $t = 0.2s)$, which demonstrates the accuracy of the kinematic component of the inverted velocity model.

To further demonstrate the role of this inversion, the ADCIG at the source position is extracted, as mentioned in the previous section. Figure 11a and Figure 11b show the ADCIG of the initial and the inverted velocity models, respectively. In Figure 11a the scatters moveout with angle is highly curved upward indicating a generally lower velocity of the initial model. On the other hand, the gather in Figure 11b is generally flat and the depth is correctly located at 1.25km, which means the inverted model is much more accurate.

### 3.2 Marmousi Model With Multi Sources

If we have many induced sources, the model should be better illuminated. Thus, we repeat the experiment with including ten sources randomly placed in the model between $1.7 - 2.3km$ in $x$ direction and $1.1 - 1.4km$ in $x$ direction, in an effort to depict a reservoir region. The source peak frequencies are randomly selected from $10 - 20Hz$, and the sources are ignited at random times, as well. In real case, micro-seismic data is hour-long continuous record with lots of events shown on it. We assume that we can manually separate these different events from each other. In another word, we cut clips of events from the continuous record. Thus, we do not know the true source onset since the cut time is arbitrary. So we have to set the source ignition time to random. The true and initial velocity model remain the same as the ones in the previous experiment, i.e. Figure 3a and Figure 3b. All the conditions of the sources and the medium are also the same. Like in FWI, the sources are separately inverted, which means the observed data of different sources can be identified and picked separately from the raw data. This is a reasonable assumption because it is usually satisfied in real micro-seismic monitoring data and there are many approaches to deblend sources. Also, by handling the sources separately, we have the liberty to have different $f(x,z)$ and $w(t)$ for the different sources, which is usually expected. The inverted source images are shown in Figure 12. To compare the source images, we sum the individual $f(x,z)$ for each source, like what is usually done in TRM. The red dots are the spatial locations of the sources and the sources are inverted to their reasonably accurate positions. The size of the source images reflect the peak frequency...
Figure 12. The inverted 10 source images (spatial component) for the part of the Marmousi model, zoomed into the source area, and summed for demonstration. Showed the initial positions (white dots) and the true source locations (black dots) for comparison.

of the source. Also we may find that the artifacts are much weaker compared to the single shot case. The whole inversion took 30 iterations, in which there are 10 sources contributing to the gradient. So the misfit decreases faster for each iteration than with the single source situation, although each iteration may take more computation time. The inverted velocity model is shown in Figure 13. The result is generally cleaner than the since shot case. The high velocity layers are also clearly identified. Since we separately invert for the various sources, the source strength will not affect the inversion results. Each source has its own source image \( f(x, z) \) and source function \( w(t) \). Although velocity update is a gradient summation of all sources, we may balance the contributions of different sources to the same level.

Figure 13. The inverted (part of the) Marmousi model after 30th iterations using the ten random sources.
Figure 14. (a) The true velocity model (SEG/EAGE overthrust model) and (b) the initial model for inversion.

3.3 SEG/EAGE Overthrust Model With Single Sources

We also tested our approach on the SEG/EAGE overthrust model, which is seems more realistic than the part of the Marmousi we choose. The true velocity model is shown in Figure 14a, where we modified the original model by setting a sea layer up top. The initial velocity model is also linearly increasing with depth, which is shown in Figure 14b. The model is $8\text{km}$ wide and $3.6\text{km}$ deep with the grid spacing of $20\text{m}$. The true single source locations is $(4\text{km}, 2.2\text{km})$. The inverted velocity model after 150 iterations is shown in Figure 15. It is also a quite good result because it highlights the main features of the fault zone and the high velocity layer, which is approximately located at a depth of $2.5 - 3.0\text{km}$. In addition, the sea water bottom is also well inverted at the right depth. Figure 16 shows the inverted source image of the single source. Although it has some weak artifacts around the source, it is still showing the right source location. As we can see in this example, the velocity update is mainly along the transmission wavepaths in this single source situation. In the multi-sources situation,
the source distribution will affect the update area. If the passive sources are distributed over a wide area, the illuminated area will be larger than the situation when the sources are close to each other, because when the sources are distributed in a small area, that transmission wave paths are similar for the sources, which means the update area is limited. Unfortunately, in reality, the injection processes are executed in a certain area and cause passive sources to be very close to each other. That makes the velocity update mainly focused in a certain region instead of a large area in the model space. Although this reality restricts the velocity inversion, it is still good enough to locate the passive source locations, since the un-updated parts do not affect the inversion results a lot. The resulting velocity is the velocity that helped us fit the modeled data to the observed ones.

4 CONCLUSIONS

We develop an approach to invert for the source location image, the source wavelet and the velocity model using recorded micro-seismic events from the Earth surface. To mitigate the role of the source wavelet, which is usually unknown in shape and time, we utilize a source function independent objective function based on the convolution of our recorded traces with a reference trace, through which we can loose the necessity of knowing the source onset. The inversion for all three unknowns yielded convergent results on a model inspired by the Marmousi model and the SEG/EAGE overthrust model, in spite starting with an initial linearly increasing velocity model that is far away from the true one. Both single and multi sources situations are considered. The extended image and angle gathers of the source is used to verify the accuracy of the kinematic components of the velocity model. The proposed method needs little knowledge of the subsurface structure, which releases the difficulty of
locating the sources. But on the other hand, it requires more computational cost comparing to the conventional migration or traveltime based methods.

ACKNOWLEDGMENTS

We thank KAUST for sponsoring this research. We also thank to the team of SWAG for their help during the research.

REFERENCES


B. Biondi and T. Tisserant 2004. 3D angle-domain common-image gathers for migration velocity analysis, Geophysical Prospecting, 52, 575–591, Wiley Online Library.


