Multi-scattering inversion for low model wavenumbers
Tariq Alkhalifah and Zedong Wu, KAUST

SUMMARY

A successful full wavenumber inversion (FWI) implementation updates the low wavenumber model components first for proper wavefield propagation description, and slowly adds the high-wavenumber potentially scattering parts of the model. The low-wavenumber components can be extracted from the transmission parts of the recorded data given by direct arrivals or the transmission parts of the single and double-scattering wavefields developed from a predicted scatter field. We develop a combined inversion of data modeled from the source and those corresponding to single and double scattering to update both the velocity model and the component of the velocity (perturbation) responsible for the single and double scattering. The combined inversion helps us access most of the potential model wavenumber information that may be embedded in the data. A scattering angle filter is used to divide the gradient of the combined inversion so initially the high wavenumber (low scattering angle) components of the gradient is directed to the perturbation model and the low wavenumber (high scattering angle) components to the velocity model. As our background velocity matures, the scattering angle divide is slowly lowered to allow for more of the higher wavenumbers to contribute the velocity model.

INTRODUCTION

The recorded seismic data are inherently dependent on the short wavelength components of the Earth model for the presence of the majority of the recorded events, and dependent on the long wavelength components for the geometrical features of such events. In inverting for a velocity model, as the inversion process demands, we use the geometrical features of the wavefield to extract the long wavelength components first, which is in turn used to place the scattered events in their accurate position in what we typically refer to as imaging. Likewise, in FWI we use the extracted model from the geometrical behavior or the transmission components to place the perturbations in their accurate locations in the velocity model in a frequency continuation process to accommodate inaccuracies in the propagator velocity (Bunks et al., 1995; Pratt et al., 1996; Virieux and Operto, 2009). The key element in the inversion strategy is the model wavenumber build up.

Recently, many have explored the wavenumber continuation theme from the model stance point, instead of the data (Tang et al., 2013; Almomin and Biondi, 2013; Albertin et al., 2013). Despite the value of data decimation and selection, the real objective of this process is far more apparent in the model domain and specifically at the gradient level (Sirgue and Pratt, 2004). Such developments resulted into a more direct control of the model wavenumber update, independent of frequency, given by a scattering angle filter (Alkhalifah, 2015). Unlike a classic smoothing operator, the scattering angle filter is governed by a physical quantity of the scattering process providing low wavenumbers along even complex ray paths, resulting from a complex background velocity. However, a proper wavenumber continuation also requires the presence of the wavenumbers, or their proper extraction from the data. To do so, we need to utilize conventionally modeled data, as well as those that rely on the image (Xu et al., 2012; Ma et al., 2012; Almomin and Biondi, 2012; Fleury and Perrone, 2012; Wang et al., 2013; Albertin et al., 2013). Low wavenumber information extracted directly from the data residual has typically mild depth penetration limited by the penetration of the diving waves. However, we can also extract low wavenumbers from the data by imaging the reflections. We can either use extended images for classic MVA (Symes and Kern, 1992) or apply a demigration to implement what is referred to as reflection waveform inversion (RWI) (Chavent and Plessix, 1999; Clement et al., 2001; Xu et al., 2012).

In this abstract, we model the source, as well as the single and double scattered wavefields, to contribute to the predicted data to be compared with the observed data using a least square objective functional. While the source modeled wavefield depends only on the velocity model, the single and double scattered wavefields depend on the velocity and perturbation models. Using a combined objective functional, we invert for both a velocity and perturbation models. The perturbation model produces reflections and double scattering only when they are not predicted by the modeled wavefield from the source.

MULTI SCATTERING INVERSION

Alkhalifah and Wu (2015) combined FWI and RWI into a single objective functional that naturally adhered to both objectives. Accordingly, we add the, reasonably important for Salt flanks, double-scattering component to suggest the following objective functional:

\[ E_{MWI} = \frac{1}{2} |d_0 - d_s(m) - d_{m1}(m, \delta m) - d_{m2}(m, \delta m)|^2 , \]

where \( d_0 \) are the observed data, \( d_s \) are data corresponding to the modeled wavefield, \( u_0(x_r) \), from the source, \( d_{m1} \) are data corresponding to the modeled wavefield, \( u_0(x_r) \), from the image, and \( d_{m2}(m, \delta m) \) is the double scattering data extracted from the double scattered wavefield, \( \delta^2 u_s(x_r) \), all at the receiver positions, \( x_r \).
Low model wavenumbers

These wavefields satisfy the following wave equations (Zhang and Duan, 2012):

\[
L(m)u_s = f(x)\delta(x - x_s),
\]

\[
L(m)\delta u_s = \omega^2 \delta m(\omega)u_s(x),
\]

\[
L(m)\delta^2 u_s = \omega^2 \delta m(\omega)\delta u_s(x),
\]

in which \(m\) is the background model, \(\delta m\) is the perturbation (scattering) model, and Helmholtz operator,

\[
L(m) = L'(m) = \nabla^2 + \omega^2 m(x),
\]

which in this case is self adjoint, with the adjoint operator given by \(L^*\). In addition, \(f\) is the source function, \(\omega\) is the angular frequency, \(x\) is a vector describing the location in space, and \(x_s\) is the source location. Similar to its single scattering counterpart (Alkhalifah and Wu, 2015), any reflections or double scattered multiples not predicted by the model \(m\) will be generated in \(d_{m1}\) and \(d_{m2}\), respectively.

To obtain the corresponding gradients, we use the extended Lagrangian formulation (Plessix, 2006)

\[
E(m, \delta m, \lambda_1, \lambda_2, \lambda_3) = E_{MW1} + A_1 + A_2 + A_3,
\]

where

\[
A_1(m, \lambda_3) = < \lambda_3, Lu - s >,
\]

\[
A_2(m, \delta m, \lambda_2) = < \lambda_2, L\delta u - \delta m >,
\]

\[
A_3(m, \delta m, \lambda_1) = < \lambda_1, L\delta^2 u - \delta m\delta u >,
\]

and \(< . >\) stands for the dot product over space (since our wavefields are in the frequency domain). The reason for the reverse order in the index of \(\lambda\) will be apparent later. Our objective is given by the stationary of the Lagrangian with respect to the state and adjoint state variables, as well as our model, and thus, we set the gradient of \(E\) with respect to all these variables to zero.

The state equations are obvious and given by equations 2 and 3, and thus, we evaluate the adjoint-state formulas obtained by setting the following to zero

\[
\frac{\partial E}{\partial u} = L^\dagger \lambda_3 - \delta m \lambda_2 + \Delta d(x_s)\delta(x - x_s) = 0,
\]

\[
\frac{\partial E}{\partial \delta u} = L^\dagger \lambda_2 - \delta m \lambda_1 + \Delta d(x_s)\delta(x - x_s) = 0,
\]

\[
\frac{\partial E}{\partial \delta^2 u} = L^\dagger \lambda_1 + \Delta d(x_s)\delta(x - x_s) = 0,
\]

where \(\Delta d = d_{m1} - d_{m1}(m) - d_{m2}(m, \delta m) - d_{m2} - (m, \delta m)\).

Reorganizing these equations in the order they would be evaluated yields:

\[
L^T \lambda_1 = -\Delta d(x_s)\delta(x - x_s),
\]

\[
L^T \lambda_2 = \delta m \lambda_1 - \Delta d(x_s)\delta(x - x_s)
\]

\[
\delta m \lambda_1 + L^T \lambda_1,
\]

\[
L^T \lambda_3 = \delta m \lambda_2 - \Delta d(x_s)\delta(x - x_s)
\]

\[
\delta m \lambda_2 + L^T \lambda_1.
\]

Thus, we can define \(\lambda_2' = \lambda_2 - \lambda_1\) and \(\lambda_3' = \lambda_3 - \lambda_1 - \lambda_2'\), and as a result

\[
L^T \lambda_1 = -\Delta d(x_s)\delta(x - x_s),
\]

\[
L^T \lambda_2' = \delta m \lambda_1,
\]

\[
L^T \lambda_3' = \delta m \lambda_2',
\]

which represents the adjoint wavefields corresponding to a single source type.

Finally, taking the derivative of \(E\) with respect to the model parameters, where the state and adjoint state variables are considered fixed, and set them to zero, we obtain

\[
\nabla_m E_{MW1} = \Re \{-\frac{\partial L}{\partial m}(u\lambda_3 + \delta u\lambda_2 + \delta_2 u\lambda_1)\}
\]

\[
\nabla_m E_{MW1} = \Re \{-\frac{\partial L}{\partial m}(u\lambda_1 + u\lambda_2 + \delta u\lambda_1 + u\lambda_2' + \delta_2 u\lambda_1)\},
\]

\[
\nabla_m E_{MW1} = \Re \{-\frac{\partial L}{\partial m}(u\lambda_2 + \delta u\lambda_1)\}
\]

\[
\Re \{-\frac{\partial L}{\partial m}(u\lambda_1 + u\lambda_2' + \delta u\lambda_1)\},
\]

where \(\Re\) stands for the real part. If we look at the gradient with respect to the model carefully, we realize it includes the gradients corresponding to FWI and RWI.

FILTERING THE GRADIENT

Filtering the gradient using extended images does not only provide us with an opportunity to control the wavenumbers admitted to the velocity model, independent of frequency, but also provides the opportunity to clean up our gradients for the model and perturbation. In multi-scattering, the large scattering angle may not correspond to large offsets any more, depending on the scattering path, but still it will provide the smoother wave path elements we seek.

Thus, to improve the action of the gradient and specifically to allow it to focus on the appropriate long wavelength components, we can utilize a slightly modified time lag (velocity scaled, \(\zeta\), (Khalil et al., 2013)) to our gradients for FWI (Alkhalifah, 2015),

\[
R(x, \zeta) = R(x)e^{-4i\omega \zeta},
\]

where \(R\) stands for any one of the gradients discussed above. Here, \(v\) is the velocity \(\left(\frac{1}{\sqrt{\mu}}\right)\), and \(\zeta = \frac{\tau}{\sqrt{\mu}}v(x)\), where \(\tau\) is the classic time lag. An inherent feature of this modified time-lag (distance units) representation is that the relationship between the scattering angle and the wavenumber of the gradient is free of a velocity (space) dependency. In fact, the scattering angle, \(\theta\), is then given by \(\cot^2 \frac{\theta}{2} = \frac{|k_m|}{k}\), where \(k_m\) is the model wavenumber vector and \(k\) is the wavenumber (Fourier transform) corresponding to \(\zeta\). The scattering angle filter corresponding to a particular angle is linear and, thus, we can denote the operation by the symbol \(L_{\zeta}\), in which

\[
L_{\zeta}(R_1 + R_2) = L_{\zeta}R_1 + L_{\zeta}R_2.
\]
Since the action of a Hessian, $H$, on a desired update, $x$, is given by the gradient; specifically

$$ L_x H x = L_x R(x), \quad (17) $$

then the application of the scattering angle filter on formula 17 can result in an update $x(x)$ corresponding to the filter gradient. The action of the Hessian on the desired update is given by Mativier et al. (2014), as

$$ H v(x) = \Re \{- \frac{\partial L}{\partial m} (\delta u_\lambda_1 + u_\lambda_2 + u_\lambda_4)\}, \quad (18) $$

where

$$ L^j \lambda_4 = \sum_j d_m(x_j) \delta(x - x_j). \quad (19) $$

So the application of the filter is distributive to all the components in equation 18.

**COMBINING THE PIECES**

In the following, we describe an approach that utilizes both the FWI/RWI/MWI combination with the power of a scattering angle filter. The approach is based on manipulating the gradient using a scattering angle filter. The process starts with an initial velocity model that is used to generate synthetic seismograms using, as much as we can, the conditions and experiment parameters used in the field. We will, however, restrict the description here to the acoustic constant density case, for simplicity. The residual between the modeled and field data provides the source function for the adjoint state method. We will, however, restrict the conditions and experiment parameters used in the field. We will, however, restrict the description here to the acoustic constant density case, for simplicity. The residual between the modeled and field data provides the source function for the adjoint state method. We will, however, restrict the conditions and experiment parameters used in the field. We will, however, restrict the description here to the acoustic constant density case, for simplicity. The residual between the modeled and field data provides the source function for the adjoint state method.

We then, divide the gradient into two parts, one corresponding to low scattering angles, $R_l$, and the other to high scattering angles, $R_h$, as follows:

$$ R_l(x, \zeta) = R_l(x, \zeta) + R_h(x, \zeta). \quad (20) $$

Thus, the gradients are given by $R_l(x) = R_l(x, 0)$ and $R_h(x) = R_h(x, 0)$. If the cut off angle is high (near 180 degrees), $R_h$, will be dominated by only low wavenumber information that can be added to the background velocity model. The residual, $R_l$, is then used in the RWI as the initial $\delta m$. This perturbation model should be free of low frequency artifacts and possibly other noise (Khalil et al., 2013). This will result in a cleaner scattered field $\delta u_{\lambda_2}$ and a cleaner RWI gradient, $R_2(x)$. However, $R_2(x)$ may contain high wavenumber components from multi scattering information in our data. It also, may include additional high wavenumber information if the adjoint state method was used to develop the gradient instead of least square inversion. In this case, we apply another scattering angle filter, but now, to the RWI gradient, using probably the same cut off angle used in the first filter. We can, then, add the two smooth gradients, as well as, the sharp ones to update our velocity and perturbation models:

$$ m(x) = R_{l_1}(x, 0) + R_{l_2}(x, 0) \quad (21) $$

$$ \delta m(x) = R_{l_1}(x, 0) + R_{l_2}(x, 0). \quad (22) $$

We can also use $\delta m(x)$ to obtain a smooth gradient corresponding to wave paths associated with smooth scattering energy, and extract the smooth part in the same way.

![Figure 1](https://example.com/figure1.png)

**A SIMPLE GRADIENT TEST**

In this example, we demonstrate how we can split the gradient to a velocity and perturbation model components using a scattering angle filter. We consider the velocity model shown in Figure 1(a). We acquired data from 40 shots on the surface with shot interval of 140 m. The receivers are distributed along the surface at 20 meters interval. Our background velocity model is constant throughout given by 2000 m/s. Conventional single scattering least square imaging is not capable of imaging the vertical component of the perturbation as shown in Figure 1(c). However, our perturbation model given by the double scattering component (equation 14)
Low model wavenumbers

as well, managed to image the vertical piece of the interface. A feature most useful for Salt flanks. In both cases, we apply a scattering angle filter that attenuates large scattering, which explains the clean nature of the images.

If we impose on the velocity gradients (equation 13) a high pass scattering angle filter, we obtain smooth updates corresponding to the direct arrival (Figure 1(b)), corresponding to single scattering (Figure 1(d)), and corresponding to double scattering (Figure 1(f)). Since a high pass scattering angle cut off near 180 degrees is applied the gradients are considerably smooth. Thus, we can split our computed gradients to a background model part and a perturbation part.

THE MARMOUSI MODEL

We next test the approach on the versatile Marmousi model (Figure 2(a)). We filter out of the assumed observed data frequencies below 4.5 Hz to make the problem more challenging (a shot gather is shown in Figure 2(b)). The velocity, \( m \), and perturbation, \( \delta m \), models are sampled at 32 by 32 meters. We consider shots every 256 m and receivers at the location of all the grid points on the surface. The initial linearly increasing velocity model is shown in Figure 3(a), and the initial \( \delta m = 0 \), and as a result the residual data (Figure 3(b)) contain a lot of the reflection events not predicted by the initial model. The update of \( m \) will contain only smooth components thanks to the scattering angle filter, and initial \( \delta m \) is mainly given by the double scattering image, also cleaned by the scattering angle filter. A perfect implementation of FWI would yield \( m \) equal to the true velocity and \( \delta m = 0 \) at the end. However, \( \delta m \) will store the model perturbation components necessary to produce reflections that are not produced by \( m \) due to the inability of the classic FWI to constrain such reflections.

After 100 iterations, we end up with the inverted velocity model shown in Figure 4(a). The residual shot gather in Figure 4(b) shows limited energy. The match of data is provided by the reflections and double scattering produced by the inverted model in Figure 4(a) and those corresponding to the final \( \delta m \) (image), shown in Figure 5(a). This image contains the energy we did not manage to generate using the velocity model. It will usually include the high frequencies and deeper reflections that FWI did not manage to couple with. For comparison, the inverted model using conventional FWI applied to this dataset is shown in Figure 5(b), and it suffers from missing the low frequencies (below 4.5 Hz) necessary to avoid falling into a local minima with the initial model.

CONCLUSIONS

Utilizing the double scattering component of the wavefield in inverting for a velocity and a perturbation model allows us to utilize more low wavenumbers to build the background model, as well as include the double scattering energy in fitting the data. In the combined objective functional, the perturbation model provides an image that fits the residual data, and thus, reflection or double scattering energy not predicted by the velocity model will be generated by the perturbation model. A scattering angle filter is administrated to insure that we initially admit smooth velocity updates in the initial iterations to the velocity, and clean high wavenumber components, free of low frequency artifacts, to the perturbation model.

ACKNOWLEDGMENTS

We thank KAUST for it’s support.
EDITED REFERENCES
Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2015 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES


Fleury, C., and F. Perrone, 2012, 2012, Bi-objective optimization for the inversion of seismic reflection data: Combined FWI and MVA: Presented at the 82nd Annual International Meeting, SEG.


Zhang, Y., and L. Duan, 2012, 2012, Predicting multiples using a reverse time demigration: Presented at the 82nd Annual International Meeting, SEG.