Reflection Full-waveform Inversion for Inaccurate Starting Models
Abdullah AlTheyab* and G. T. Schuster, King Abdullah University of Science and Technology (KAUST)

SUMMARY
We propose a seismic-reflection full-waveform inversion (FWI) workflow to obtain a velocity model when the starting velocity model contains low-wavenumber errors. This workflow is based on Gauss-Seidel iterations where we apply Gauss-Newton FWI over a small subset of the seismic data. The subset is selected such that each subset invert a narrow range of the high-wavenumber components of the model space. The different subsets are inverted in sequence where the final model from inverting a subset is the starting model for the next subset. An under-relaxation operator is applied to the cumulative update from one pass of Gauss-Seidel iterations to accelerate convergence. The main advantage of the approach is that it enhances the low-wavenumber updates. Applications to synthetic and field data show significant low-wavenumber updates and flattening of common-image gathers after many iterations.

INTRODUCTION
FWI misfit function (Tarantola, 1984) is non-linear with many local minima, often caused by cycle-skipping (Gauthier et al., 1986). Therefore, local-gradient optimization algorithms often stagnate at a local minimum. A set of methods have been developed to avoid this issue, where cycle-skipped events are excluded by filtering in different transform domains (Bunks et al., 2015; Aman et al., 2012), or by evaluating the phase lag between the observed and calculated data for each event and muting cycle-skipped events (Bi and Lin, 2014). Such methods apply FWI in multistages, where at every stage, part of the data is inverted

\[
J_i(\delta m_i) = \frac{1}{2} \| W_i \delta d_i(\delta m_i) \|_2^2, \tag{1}
\]

where the subscript \(i\) denotes the stage number, \(\delta d_i\) is the difference between the observed and calculated data, \(W_i\) is a weighting operator that excludes (by multiplication by zero) the cycle-skipped events in the data, \(m_i\) is the initial model at the \(i\)th stage, and \(\delta m_i\) is the cumulative model updated to the initial model \(m_i\) needed to fit the data. After each stage, a new model is computed

\[
m_{i+1} = m_i + \delta m_i, \tag{2}
\]

where \(m_{i+1}\) is the initial model for the following stage. Note that at each stage the objective function is minimized with several iterations. At new stages, the weighting operator is updated accordingly to include new uncycle-skipped events. Such approaches are effective for inverting transmitted waves. However, they fail in updating the background velocity model below the reach of transmitted waves.

FWI failure in updating the low-wavenumber components of the model using reflection is often attributed to the weakness of the tomographic terms in the update kernels of FWI. Several approaches for enhancing the tomographic components were proposed for reflection FWI while minimizing the difference between the observed and calculated data (Zhou et al., 2012; Xu et al., 2012; Wang et al., 2013; Brossier et al., 2013; Yu et al., 2014). Such approaches use scale separation between high-wavenumber and low-wavenumber components of the model where they are inverted at every iteration in separate steps: the imaging step (where the reflectors are mapped in the subsurface) and the tomography step, where the low-wavenumber updates are computed based on the data misfit function. Unfortunately, the scale separation approach is based on the single scattering assumption, and therefore transmitted events and higher-order scattering can not be handled accurately during the inversion. To avoid the single scattering assumption, AlTheyab et al. (2013) proposed using Gauss-Newton optimization (Accev et al., 2002; Erlangga and Hermann, 2009; AlTheyab et al., 2013) where low-wavenumber updates along reflection-wavepaths are naturally enhanced within the FWI iterations. However, the success of all aforementioned solutions for updating the low-wavenumber components is limited to recovering localized mid-wavenumber anomalies when the reflectors are close to their accurate positions. This raises the question whether the weakness of reflection wavepaths is the sole cause for the failure in updating the low-wavenumber components of the model.

We believe that the main problem is the coupling between the low-wavenumber components of the model along all the high-wavenumber components, as well as the contradicting updates along reflection wavepaths from different angles of incidence. To solve this problem, we propose splitting the FWI problem at each stage such that we sequentially invert a narrow-offset range of traces, starting with the short offsets and ending at the far offsets. We will refer to FWI with a narrow-offset range as constant-offset FWI. In each constant-offset FWI, Gauss-Newton optimization is used, and the final velocity model of a constant-offset FWI is the initial model to the following constant-offset FWI. This approach is closely related to the non-linear block Gauss-Seidel iteration method with overlapping blocks (Tai, 1992; Gutiérrez et al., 2011). Here, each constant-offset FWI is naturally tuned for enhancing low-wavenumber updates along reflection wavepaths.

In the following sections, we will elaborate on the coupling problem and the proposed solution to enhance the low-wavenumber updates. Later, we show the results of applying the proposed method to synthetic and field data, where low-wavenumber errors in the starting model are corrected after many stages. Finally, we comment on the generalization of the method.

THEORY
There are a group of receivers along a planar surface where the travel-time gradient of a specific reflection event is zero
overlapping at the small wavenumbers near the origin. This occurrence, the corresponding tomographic terms are completely related to the model wavenumber via the relationship (Sirgue and Pratt, 2004) for a horizontally invariant model, is related to the model wavenumber via the relationship between the wavenumber coverage between the velocity model, the angle of incidence, for a horizontally invariant model, is related to reflection wavepaths. Even though there is literature overlap between the wavenumber coverage for a model with a single reflector (see Mora (1989) for details) where the tomographic terms are related to reflection wavepaths depending on the angles of incidence. When there are components of the model via reflection wavepaths. Figure 2 shows the wavenumber coverage for a model with a single reflector (see Mora (1989) for details) where the tomographic terms are related to reflection wavepaths. Even though there is little overlap between the wavenumber coverage between the diffraction terms from different frequencies and angles of incidence, the corresponding tomographic terms are completely overlapping at the small wavenumbers near the origin. This multistage FWI with automatic detection and exclusion of cycle-skipped is detected in an adaptive manner as in Bi and Lin (2014). This multistage FWI with automatic detection and exclusion of cycle-skipping is extremely slow, because subsurface reflectors will be mispositioned in the early iterations, and the positioning has to be gradually corrected with a large number of iterations. To understand the causes of slow convergence, we consider the following scenario. Consider a horizontally invariant two-layer model shown in Figure 1 which will give a single reflection from the deep interface between the two layers. For a homogeneous starting model, the angle of incidence, for a horizontally invariant model, is related to the model wavenumber via the relationship (Sirgue and Pratt, 2004)

\[ k_z = \frac{2\omega}{c}\cos\theta, \]

(3)

where \(\omega\), \(c\), and \(\theta\) are, respectively, the angular frequency, the initial velocity, and the angle of incidence. When there are low-wavenumber errors in the initial velocity model, the model phase \(\phi (k_z)\) will be a weighted average of the phases from different angles and frequencies that cover the same wavenumber \(k_z\) (i.e. the apparent depth of the reflector will be a weighted average of apparent depths from different angles). At the second iteration, phase delays of predicted specular events will vary depending on the angles of incidence \(\theta\). It follows, then, that there will be both positive and negative phase delays between the observed and calculated waveforms depending on the angle of incidence. When we invert residuals, the mispositioned reflector will act as a secondary source for updating the low-wavenumber components of the model via reflection wavepaths. Figure 2 shows the wavenumber coverage for a model with a single reflector (see Mora (1989) for details) where the tomographic terms are related to reflection wavepaths. Even though there is little overlap between the wavenumber coverage between the diffraction terms from different frequencies and angles of incidence, the corresponding tomographic terms are completely overlapping at the small wavenumbers near the origin. This illustrates the strong coupling between the high wavenumbers (from different angles and frequencies) and the low wavenumbers reconstructed by the tomographic terms in Figure 2. Therefore, the positive and negative phase errors result in conflicting updates in the tomographic terms (i.e. the overlapping zone in Figure 2), and, consequently, negligible low-wavenumber updates from reflection residuals.

To resolve the strong coupling and the conflicting updates, the objective function in equation 1 at the \(i\)–th stage, is regrouped into different terms based on source-receiver offset,

\[ J_i(\delta\mathbf{m}) = \frac{1}{2} \sum_h^N \left\| W^h \mathbf{\Delta d}_i^h (\mathbf{m}_i + \delta\mathbf{m}_i) \right\|^2, \]

(4)

where the superscript \(h\) denote the offset index and \(N\) is the number of offsets bins in the data. For the model in Figure 1, the source-receiver offset is directly related to angle of incidence \(\theta\) by the relationship (Sirgue and Pratt, 2004)

\[ \cos \theta = \frac{z}{\sqrt{q^2 + z^2}}, \]

(5)

where \(z\) is the depth of reflector, and \(q\) is the half the distance between the source and the receiver. Therefore, decomposing the objective function here is equivalent to a decomposition in the wavenumber domain based on the angles of incidence. For general media, the objective function is decomposed such that each term of the decomposed objective function has data that is sensitive to a limited area of the model’s wavenumber spectrum \(\delta\mathbf{m}\) using the diffraction terms (see Figure 2). With this direct mapping between terms in the decomposed objective function and the high-wavenumber components of the model, we can solve this system using Gauss–Seidel iterations starting from a zero update vector \(\delta\mathbf{m}_0 = \mathbf{0}\), and the model update at every \(h\)–th iteration is computed using

\[ \delta\mathbf{m}_i^h = \delta\mathbf{m}_i^{h-1} + \arg\min_x \left\| W^h \mathbf{\Delta d}_i^h (\mathbf{m}_i + \delta\mathbf{m}_i^{h-1} + x) \right\|^2. \]

(6)

Here, we apply a few iterations of FWI with Gauss-Newton optimization for the model update \(x\) which is added to the in-
tial model of the $i-$th stage $\mathbf{m}_i$ and the update from the previous Gauss-Seidel iteration $\delta \mathbf{m}_{i-1}^h$ to minimize the misfit function of the constant-offset data at the $h$-th offset bin. Note that low-wavenumbers components of the model, coupled by the tomographic terms, are freely updated during Gauss-Seidel iterations.

Due to the coupling, the low-wavenumber components of the model will oscillate between different passes through the Gauss-Seidel iterations. An under-relaxation scheme is used as a preconditioner to prevent this oscillatory behavior, where the under relaxation operator $\mathbf{S}_i$ is used during the update step between stages

$$\mathbf{m}_{i+1} = \mathbf{m}_i + \mathbf{S}_i \delta \mathbf{m}_i^N \quad (7)$$

where the $\mathbf{S}_i$ is a Gaussian smoothing operator applied to the cumulative update, the Gaussian smoothing filter is wide at early stages to allow updates to the very low-wavenumber components of the model, and the width of the filter is reduced gradually at later stages. We empirically find that smoothing along geological dip further accelerates the convergence process. Now, $\mathbf{m}_{i+1}$ is the initial model for the next stage of constant-offset inversions. Figure 3 demonstrate depicts the updates from standard FWI and the proposed Gauss-Seidel iterations using the same dataset. The latter shows enhanced low-wavenumber updates after 1 pass of Gauss-Seidel iteration, where FWI with comparable cost show negligible low-wavenumber updates.

FIELD DATA EXAMPLE

Figure 4 shows preliminary results for application to the 2D GOM streamer data with a 4 km maximum offset and 10 Hz maximum frequency after windowing. With a homogenous starting model of 2000 m/s the common-angle gathers have large moveout. The tomogram shown is after 40 Gauss-Seidel passes and the angle gathers are nearly flat as shown in Figure 4.

CONCLUSIONS

We proposed a method based on Gauss-Seidel iterations for inverting reflections starting from inaccurate velocity models. The Gauss-Seidel iterations have two advantages over the full-domain Newton inversion; it naturally enhances low-wavenumber updates, and it is easier to precondition with under-relaxation schemes. Here, we avoid the single scattering assumption and explicit scale separation, and the reflection data can still be inverted simultaneously with transmitted waves. Our approach does not require generating common image gathers, which will be an advantage in 3D applications.

ACKNOWLEDGMENTS

We thank the sponsors of the Center of Subsurface Imaging and Fluid Modeling consortium. AlTheyab thanks Saudi Aramco for sponsoring his studies at KAUST.

Figure 3: Low-wavenumber updates using Gauss-Newton FWI with the full data and with offset-rolling strategy (i.e. Gauss-Seidel iterations over offset).
Figure 4: (Top) final tomogram using the proposed method, (middle) angle gathers using the initial velocity model (2000 m/s), and (bottom) angle gathers using the final tomogram.
REFERENCES


Tai, X.-c., 1992, Parallel function decomposition and space decomposition methods with applications to optimization, splitting and domain decomposition: Presented at the BEIJING MATHEMATICS, Citeseer.


