Inverting reflections using full-waveform inversion with inaccurate starting models

Abdullah AlTheyab and G. T. Schuster

King Abdullah University of Science and Technology

SEG 2015 Annual Meeting
Outline

Introduction

Theory

Numerical Examples

Conclusions
Outline

Introduction

Theory

Numerical Examples

Conclusions
Objective: Seismic reflections $\rightarrow$ subsurface model

Source-receiver offset

Reflection Data

Model Wavenumber Spectrum

(Claerbout, 1985)
Objective: Seismic reflections $\rightarrow$ subsurface model

Source-receiver offset

Time

Reflection Data

moveout

Model Wavenumber

Spectrum

(Claerbout, 1985)

Can FWI invert both the move-out and the wiggles?

It could, because "Inversion = migration + tomography" [Mora, 1989].

BUT ... not with cycle-skipping!!
Objective: Seismic reflections $\rightarrow$ subsurface model

Can FWI invert both the move-out and the wiggles?

It could, because “Inversion = migration + tomography” [Mora, 1989].

BUT ..... not with cycle-skipping!!

(Claerbout, 1985)
Objective: Seismic reflections → subsurface model

Can FWI invert both the move-out and the wiggles?

Objective: Seismic reflections → subsurface model

- Can FWI invert both the move-out and the wiggles?
- It could, because “Inversion = migration + tomography” [Mora, 1989].
Can FWI invert both the move-out and the wiggles?

- It could, because “Inversion = migration + tomography” [Mora, 1989].
- BUT ..... not with cycle-skipping!!
Objective: Seismic reflections → subsurface model

- Extended domain approaches are robust in handling cycle-skipped reflection data [Symes, 2008, Almomin and Biondi, 2013, Zhang and Schuster, 2013], but with a high computational cost.

An alternative approach is to invert uncycle-skipped reflections using the FWI and progressively include new uncycle-skipped reflections.
Objective: Seismic reflections $\rightarrow$ subsurface model

- Extended domain approaches are robust in handling cycle-skipped reflection data [Symes, 2008, Almomin and Biondi, 2013, Zhang and Schuster, 2013], but with a high computational cost.

- An alternative approach is to invert uncycle-skipped reflections using the FWI and progressively include new uncycle-skipped reflections.
FWI of reflections data without cycle-skipping

\[ J_i (\delta m_i) = \frac{1}{2} \| W_i \triangle d (m_i + \delta m_i) \|_2^2. \]

- The re-weighting operator \( W_i \) excludes cycle-skipped events.
FWI of reflections data without cycle-skipping

\[ J_i(\delta m_i) = \frac{1}{2} \| W_i \triangle d (m_i + \delta m_i) \|_2^2. \]

- The re-weighting operator \( W_i \) excludes cycle-skipped events.

- Examples:
  - Multi-scale approach of FWI, where \( W_i \) is a low-pass filter [Bunks et al., 1995].
  - Hierarchical approach of seismic full-waveform inversion [Asnaashari et al., 2012].
  - Offset selection to avoid cycle-skipping [Al-Yaqoobi and Warner, 2013].
  - Adaptive data selection [Bi and Lin, 2014].
If we avoid cycle-skipping,

- Can we updated the background model using reflections?
If we avoid cycle-skipping,

- Can we updated the background model using reflections?
- 3 layers test: 6 km offset, 4-10 Hz
If we avoid cycle-skipping,

- Can we updated the background model using reflections?
- 3 layers test: 6 km offset, 4-10 Hz
If we avoid cycle-skipping,

- Can we updated the background model using reflections?
- 3 layers test: 6 km offset, 4-10 Hz

- 1 m/s update per iteration. The problem is \textit{slow convergence}. 
The issue: slow convergence

- It is recognized that the main cause is the weakness of the tomographic component of FWI update kernels.

Proposed solutions include:
- Enhance the tomographic update using scale-separation and single-scattering assumption [Zhou et al., 2012, Xu et al., 2012, Wang et al., 2013].
- Enhance the tomographic update by up- and down-going wavefield decomposition [Tang et al., 2013].
- Use Gauss-Newton optimization to enhance the tomographic updates [AlTheyab et al., 2013].

Successful cases so far show the ability to recover localized (mid-wavenumber) anomalies!!
The issue: slow convergence

- It is recognized that the main cause is the weakness of the tomographic component of FWI update kernels.

- Proposed solutions include:
  - Enhance the tomographic update using scale-separation and single-scattering assumption [Zhou et al., 2012, Xu et al., 2012, Wang et al., 2013].
  - Enhance the tomographic update by up- and down-going wavefield decomposition [Tang et al., 2013].
  - Use Gauss-Newton optimization to enhance the tomographic updates [AlTheyab et al., 2013].
The issue: slow convergence

- It is recognized that the main cause is the weakness of the tomographic component of FWI update kernels.

- Proposed solutions include:
  - Enhance the tomographic update using scale-separation and single-scattering assumption [Zhou et al., 2012, Xu et al., 2012, Wang et al., 2013].
  - Enhance the tomographic update by up- and down-going wavefield decomposition [Tang et al., 2013].
  - Use Gauss-Newton optimization to enhance the tomographic updates [AlTheyab et al., 2013].

- Successful cases so far show the ability to recover localized (mid-wavenumber) anomalies!!
The issue: slow convergence

▶ Is there another problem? (besides the weakness of tomographic terms)
▶ Data from different frequencies and offsets are sensitive the same area near the origin of the wavenumber spectrum.
▶ We need a solution where the coupling is reduced!!
The issue: slow convergence

- Is there another problem? (besides the weakness of tomographic terms)
The issue: slow convergence

- Is there another problem? (besides the weakness of tomographic terms)

Data from different frequencies and offsets are sensitive the same area near the origin of the wavenumber spectrum.

Conflicting tomographic updates!!
The issue: slow convergence

- Is there another problem? (besides the weakness of tomographic terms)

Conflicting tomographic updates!!

- Data from different frequencies and offsets are sensitive the same area near the origin of the wavenumber spectrum.
- We need a solution where the coupling is reduced!!
Decoupling and domain decomposition

- The **brute-force** approach for decoupling is to solve for single $v(k_z, k_x)$ at a time, simultaneously with the tomographic low wavenumber components of the model.
Decoupling and domain decomposition

- The **brute-force** approach for decoupling is to solve for single $v(k_z, k_x)$ at a time, simultaneously with the tomographic low wavenumber components of the model.
The brute-force approach for decoupling is to solve for single $v(k_z, k_x)$ at a time, simultaneously with the tomographic low wavenumber components of the model.
The **brute-force** approach for decoupling is to solve for single $v(k_z, k_x)$ at a time, simultaneously with the tomographic low wavenumber components of the model.
Decoupling and domain decomposition

- The brute-force approach for decoupling is to solve for single \( v(k_z, k_x) \) at a time, simultaneously with the tomographic low wavenumber components of the model.

- This is a domain-decomposition approach where the Gauss-Seidelization is performed on wavenumber components covered by the diffraction terms.
From domain decomposition to functional decomposition

- **Domain decomposition** can be achieved efficiently by **functional decomposition** (i.e. by regrouping the terms in the objective function and then solving individual terms sequentially).

\[
\begin{align*}
k_z &= 2\omega v \cos \theta \rightarrow J_i(\delta m_i) = \frac{1}{2N} \sum \theta \|W_{\theta i}\triangle d\theta_i(m_i + \delta m_i)\|^2, \\
k_z &= 2\omega c z \sqrt{h_2 + z_2} \rightarrow J_i(\delta m_i) = \frac{1}{2N} \sum h \|W_{h i}\triangle d\theta_i(m_i + \delta m_i)\|^2.
\end{align*}
\]

- We will proceed to Gauss-Seidel iterations with the constant-offset formulation...
From domain decomposition to functional decomposition

- **Domain decomposition** can be achieved efficiently by **functional decomposition** (i.e. by regrouping the terms in the objective function and then solving individual terms sequentially).

- Relations between the wavenumbers and the data/image coordinates [Sirgue and Pratt, 2004] give the basis for functional decomposition

\[
k_z = 2 \frac{\omega}{v} \cos \theta \quad \rightarrow \quad J_i (\delta m_i) = \frac{1}{2} \sum_{\theta} \left\| W_{i}^{\theta} \delta d_{i}^{\theta} (m_i + \delta m_i) \right\|^2,
\]

\[
k_z = 2 \frac{\omega}{c} \frac{z}{\sqrt{h^2 + z^2}} \quad \rightarrow \quad J_i (\delta m_i) = \frac{1}{2} \sum_{h} \left\| W_{i}^{h} \delta d_{i}^{h} (m_i + \delta m_i) \right\|^2.
\]
From domain decomposition to functional decomposition

- **Domain decomposition** can be achieved efficiently by **functional decomposition** (i.e. by regrouping the terms in the objective function and then solving individual terms sequentially).

- Relations between the wavenumbers and the data/image coordinates [Sirgue and Pratt, 2004] give the basis for functional decomposition

\[
k_z = 2 \frac{\omega}{v} \cos \theta \quad \rightarrow \quad J_i (\delta m_i) = \frac{1}{2} \sum_\theta \| W^\theta_i \Delta d^\theta_i (m_i + \delta m_i) \|^2,
\]

\[
k_z = 2 \frac{\omega}{c} \frac{z}{\sqrt{h^2 + z^2}} \quad \rightarrow \quad J_i (\delta m_i) = \frac{1}{2} \sum_h \| W^h_i \Delta d^h_i (m_i + \delta m_i) \|^2.
\]

- We will proceed to Gauss-Seidel iterations with the constant-offset formulation...
FWI with Gauss-Seidel-Newton iterations

1. **Regroup** the terms (functional decomposition) in the objective function based on offsets

\[
J_i(\delta m_i) = \frac{1}{2} \sum_{h}^{N} \|W_i^h \triangle d_i^h (m_i + \delta m_i)\|^2.
\]
FWI with Gauss-Seidel-Newton iterations

1. **Regroup** the terms (functional decomposition) in the objective function based on offsets

   $$J_i(\delta m_i) = \frac{1}{2} \sum_{h}^{N} \| W_i^h \Delta d_i^h (m_i + \delta m_i) \|^2.$$

2. Perform **constant-offset FWI sequentially** (offset-rolling)

   $$\delta m_i^h = \delta m_i^{h-1} + \arg \min_{x} \| W_i^h \Delta d_i^h (m_i + \delta m_i^{h-1} + x) \|^2.$$
FWI with Gauss-Seidel-Newton iterations

1. **Regroup** the terms (functional decomposition) in the objective function based on offsets

   \[ J_i(\delta m_i) = \frac{1}{2} \sum_{h} \| W_i^h \Delta d_i^h (m_i + \delta m_i) \|^2. \]

2. Perform **constant-offset FWI sequentially** (offset-rolling)

   \[ \delta m_i^h = \delta m_i^{h-1} + \arg\min_x \| W_i^h \Delta d_i^h (m_i + \delta m_i^{h-1} + x) \|^2. \]

3. **Update** the initial model for the next stage

   \[ m_{i+1} = m_i + S_i \delta m_i^N \]

   ▶ **S** is the under-relaxation operator, which performs Gaussian smoothing over the commutative updates from the constant-offset FWIs.
1. Start with a simple smooth velocity model.
2. Find the anchor reflections; i.e., portion of the data containing the apex of a reflection move-out (i.e., $\nabla \tau(h, x) \approx 0$).
3. For layered media, anchor reflections can be found in near-offset traces.
4. Why? Anchor reflections are unlikely to be cycle-skipped.

**Workflow**

- Select un-cycle-skipped data
- Select a narrow window of source-receiver **offsets**
- **Gauss-Newton FWI**
- **Shift** the offset-window
  - Yes: Are there more offsets?
  - No: **Smooth** cumulative update from GN-FWI
- Is all data inverted with acceptable mode?
  - Yes: end
  - No: shift the offset-window

- Offset-rolling
Start with a simple smooth velocity model.

- Find the anchor reflections; i.e. portion of the data containing the apex of a reflection move-out (i.e. $\nabla \tau (h, x) \approx 0$).
- For layered media, anchor reflections can be found in near-offset traces.
- Why? Anchor reflections are unlikely to be cycle-skipped.

**Workflow**

- Select uncycle-skipped data
- Select a narrow window of source-receiver offsets
- Gauss-Newton FWI
- Shift the offset-window
- Smooth cumulative update from GN-FWI
- Is all data inverted with acceptable mode?
- Yes
- No

Offset-rolling

- Yes
- No

Choose unskipped data

Gauss-Newton FWI

Shift the offset-window

Yes

Are there more offsets?

No

Smooth cumulative update from GN-FWI

Is all data inverted with acceptable mode?

end
- Start with a simple smooth velocity model.
- Find the anchor reflections; i.e. portion of the data containing the apex of a reflection move-out (i.e. $\nabla \tau (h, x) \approx 0$).
Start with a simple smooth velocity model.

Find the **anchor reflections**; i.e. portion of the data containing the apex of a reflection move-out (i.e. $\nabla \tau(h, x) \approx 0$).

For layered media, anchor reflections can be found in near-offset traces.
Start with a simple smooth velocity model.

Find the **anchor reflections**; i.e. portion of the data containing the apex of a reflection move-out (i.e. $\nabla \tau (h, x) \approx 0$).

For layered media, anchor reflections can be found in near-offset traces.

Why? anchor reflections are unlikely to be cycle-skipped.
For the sake of simplicity, we mask cycle-skipped data based on offset.
Workflow

For the sake of simplicity, we mask cycle-skipped data based on offset.

Smooth cumulative update from GN-FWI

Select uncycle-skipped data

Select a narrow window of source-receiver offsets

Gauss-Newton FWI

Shift the offset-window

Are there more offsets?

Yes

No

Smooth cumulative update from GN-FWI

Is all data inverted with acceptable mode?

Yes

No

end
For the sake of simplicity, we mask cycle-skipped data based on offset.

Is all data inverted with acceptable mode?

Selecting a narrow window of source-receiver offsets.

Gauss-Newton FWI

Smooth cumulative update from GN-FWI

Select un-cycle-skipped data

Offset-rolling

Start

Are there more offsets?

Shift the offset-window

Offset-rolling

end
Workflow

For the sake of simplicity, we mask cycle-skipped data based on offset.

Select uncycle-skipped data

Select a narrow window of source-receiver offsets

Gauss-Newton FWI

Shift the offset-window

Are there more offsets?

Yes

No

Smooth cumulative update from GN-FWI

Is all data inverted with acceptable mode?

No

Yes

end
Workflow

For the sake of simplicity, we mask cycle-skipped data based on offset.

Select uncycle-skipped data

Select a narrow window of source-receiver offsets

Gauss-Newton FWI

Shift the offset-window

Yes

Are there more offsets?
No

Smooth cumulative update from GN-FWI

Is all data inverted with acceptable mode?

Yes
end
No

Offset-rolling

Start
For the sake of simplicity, we mask cycle-skipped data based on offset.

**Select un-cycle-skipped data**

Select a narrow window of source-receiver **offsets**

**Gauss-Newton FWI**

*Shift the offset-window*

Are there more offsets?

**Smooth cumulative update from GN-FWI**

Is all data inverted with acceptable mode?

**end**
3 layers test

The proposed FWI method significantly enhances the low-wavenumber updates.
The proposed FWI method significantly enhances the low-wavenumber updates.
The proposed FWI method significantly enhances the low-wavenumber updates.
The proposed FWI method significantly enhances the low-wavenumber updates.
Outline

Introduction

Theory

Numerical Examples

Conclusions
Synthetic data test 1

2D full aperture acquisition parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot spacing</td>
<td>20 m</td>
</tr>
<tr>
<td>Receiver spacing</td>
<td>10 m</td>
</tr>
<tr>
<td>Offset range</td>
<td>0 - 6 km</td>
</tr>
<tr>
<td>Frequency range</td>
<td>4 - 10 Hz</td>
</tr>
<tr>
<td>Modeling</td>
<td>constant-density acoustic TDFD</td>
</tr>
<tr>
<td>Inversion</td>
<td>constant-density acoustic TDFD (inversion crime)*</td>
</tr>
</tbody>
</table>

True Model

Initial Model
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI without the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
Synthetic data test 2

2D streamer acquisition parameters:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot spacing</td>
<td>40 m</td>
<td>Receiver spacing</td>
</tr>
<tr>
<td>Offset range</td>
<td>0 - 6 km</td>
<td>Frequency range</td>
</tr>
<tr>
<td>Modeling</td>
<td>variable-density elastic TDFD</td>
<td>Inversion</td>
</tr>
</tbody>
</table>
FWI with the proposed strategy

Proposed FWI

Depth [m]

Distance [m]

Velocity [m/s] 1500 3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Depth [m]
0 1200 800 400 0

Distance [m]
0 4000 8000

Velocity [m/s] 1500 3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Depth [m]
0
1200 800 400

Distance [m]
0 4000 8000

Velocity [m/s] 1500 3000
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Depth [m]
0 1200 800 400 0

Distance [m]
0 4000 8000

Velocity [m/s] 1500 3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Depth [m]

Distance [m]

Velocity [m/s] 1500 to 3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Depth [m]

Distance [m]

Velocity [m/s] 1500 - 3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Distance [m]

Depth [m]

Velocity [m/s] 1500

3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy

Proposed FWI

Depth [m] 0 1200 800 400 0

Distance [m] 0 4000 8000

Velocity [m/s] 1500 3000
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
FWI with the proposed strategy
Synthetic data test 2 results
Synthetic data test 2 results

FWI stage 1

Depth [m]

Distance [m]

Velocity [m/s] 1351.05 3980.66

500 m
Synthetic data test 2 results

FWI stage 42

Depth [m]

Distance [m]

Velocity [m/s] 1351.05 3980.66
Synthetic data test 2 results

Normalized residual

RMS

0.2

1

Standard FWI Iteration

Equivalent
GOM field-data parameters

2D streamer acquisition parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shot spacing</td>
<td>37.5 m</td>
</tr>
<tr>
<td>Receiver spacing</td>
<td>12.5 m</td>
</tr>
<tr>
<td>Number of receivers</td>
<td>480</td>
</tr>
<tr>
<td>Number of shots</td>
<td>266</td>
</tr>
<tr>
<td>Offset range</td>
<td>200m - 6.2 km</td>
</tr>
</tbody>
</table>

Offset: 198 - 6,184 m

Time [s] Sea bottom

2.8

0 -1 Hz  0 -2 Hz  0 -3 Hz

Time [s] Offset [km]
GOM field-data preprocessing

1) 0-10 Hz Bandpass

2) Mute transmitted waves

▶ Assume a minimum phase source wavelet.
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results
GOM field-data results

Initial model

Distance [km]

Depth [km]

Velocity [m/s]:

0 10 0 3

0 10

1400 2500

Initial model

2500
GOM field-data results

Proposed FWI (40 Stages)

Depth [km]

Distance [km]

Velocity [m/s]: 1400 - 2500
GOM field-data results
GOM field-data results
Outline

Introduction

Theory

Numerical Examples

Conclusions
Liabilities and limitations

- Because we invert over small data subsets:
  - Good signal-to-noise ratio is required for a stable inversion.
  - Current implementation shows sensitivity to poor illumination.
Liabilities and limitations

- Because we invert over small data subsets:
  - Good signal-to-noise ratio is required for a stable inversion.
  - Current implementation shows sensitivity to poor illumination.
- The design of the relaxation operator $S$ is empirical and relay on prior information (i.e. smoothing along dips).
Liabilities and limitations

- Because we invert over small data subsets:
  - Good signal-to-noise ratio is required for a stable inversion.
  - Current implementation shows sensitivity to poor illumination.
- The design of the relaxation operator $S$ is empirical and relay on prior information (i.e. smoothing along dips).
- Starting from fast models is better than starting from slow models.
Liabilities and limitations

- Because we invert over small data subsets:
  - Good signal-to-noise ratio is required for a stable inversion.
  - Current implementation shows sensitivity to poor illumination.
- The design of the relaxation operator $S$ is empirical and relay on prior information (i.e. smoothing along dips).
- Starting from fast models is better than starting from slow models.
- The strategy can be 3x-10x more costly than standard FWI depending on how wrong the initial velocity model.
Conclusions

▶ We analyzed the slow convergence and its relation to the problem of coupling.

▶ We proposed a Gauss-Seidel-Newton FWI for inaccurate smooth starting models.

▶ Synthetic data tests and preliminary application to field data illustrate the ability of FWI to update the low-wavenumber components of the velocity model.

▶ Constant-offset formulation is easy to understand and implement, but it's not suitable for complex media (functional decomposition \( \neq \) domain decomposition).

▶ Implement alternatives to constant-offset formulation (maybe single scattering angle formulation).
Conclusions

- We analyzed the slow convergence and its relation to the problem of coupling.

- We proposed a Gauss-Seidel-Newton FWI for inaccurate smooth starting models.
Conclusions

- We analyzed the slow convergence and its relation to the problem of coupling.

- We proposed a Gauss-Seidel-Newton FWI for inaccurate smooth starting models.

- Synthetic data tests and preliminary application to field data illustrate the ability of FWI to update the low-wavenumber components of the velocity model.
Conclusions

▶ We analyzed the slow convergence and its relation to the problem of coupling.

▶ We proposed a Gauss-Seidel-Newton FWI for inaccurate smooth starting models.

▶ Synthetic data tests and preliminary application to field data illustrate the ability of FWI to update the low-wavenumber components of the velocity model.

▶ Constant-offset formulation is easy to understand and implement, but its **not suitable for complex media** (functional decomposition ≠ domain decomposition).
  
  ▶ Implement alternatives to constant-offset formulation (maybe single scattering angle formulation).
Acknowledgments

- We thank the sponsors of the Center of Subsurface Imaging and Fluid Modeling (CSIM) consortium.
- ATTheyab thanks Saudi Aramco for sponsoring his studies at KAUST.
- KAUST for the support.
- Thank you ...

Questions?
Key References


Relation to extended-domain FWI

- The non-linear Gauss-Seidel iterations
  \[ x_i^{k+1} = \arg\min_y J(x_1^{k+1}, \ldots, y, \ldots, x_n^k). \]

- The non-linear Jacobi iterations
  \[ x_i^{k+1} = \arg\min_y J(x_1^k, \ldots, y, \ldots, x_n^k). \]
Relation to extended-domain FWI

- The non-linear Gauss-Seidel iterations

\[ x_i^{k+1} = \arg \min_y J(x_1^{k+1}, \ldots, y, \ldots, x_n^k). \]

- The non-linear Jacobi iterations

\[ x_i^{k+1} = \arg \min_y J(x_1^k, \ldots, y, \ldots, x_n^k). \]

- The Gauss-Seidel-Newton FWI

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

- The Jacobi-Newton FWI

\[ \delta \mathbf{m}_i^h = \arg \min_x \| \mathbf{W}_i^h \triangle d_i^h (\mathbf{m}_i + \mathbf{x}) \|^2. \]
Relation to extended-domain FWI

- The non-linear Gauss-Seidel iterations
  \[ x_i^{k+1} = \arg \min_y J(x_1^{k+1}, \ldots, y, \ldots, x_n^k). \]

- The non-linear Jacobi iterations
  \[ x_i^{k+1} = \arg \min_y J(x_1^k, \ldots, y, \ldots, x_n^k). \]

- The Gauss-Seidel-Newton FWI
  \[ \delta m_i^h = \delta m_i^{h-1} + \arg \min_x \| W_i^h \Delta d_i^h (m_i + \delta m_i^{h-1} + x) \|^2. \]

- The Jacobi-Newton FWI
  \[ \delta m_i^h = \arg \min_x \| W_i^h \Delta d_i^h (m_i + x) \|^2. \]

- The gradient of Jacobi-Newton FWI is equivalent to the gradient FWI with surface-offset extension, at the 1st non-linear iteration.
Relation to extended-domain FWI

However,

- Gauss-Seidel-Newton FWI is **NOT** immune to cycle-skipping. That’s why we need $W_i h$.

- Sequencing the inversion eliminates the need for a focusing/annihilation term in the objective function.

- Gauss-Seidel-Newton FWI has lower memory requirements.