Illumination Compensation In Linearized Waveform Inversion Images By Frequency Domain Extension

Rahul Sarkar[1] Biondo Biondi[1,2]

[1]Institute for Computational and Mathematical Engineering (ICME)
Stanford University

[2]Department of Geophysics
Stanford University

SEP Sponsors Meeting, 2019
SEP 176 (Pages 49-61)

A brief recap...

In previous reports and SEP talks we discussed how the 2D TFWI problem benefits from the frequency domain formulation.

Tomographic Full Waveform Inversion (TFWI)

Optimization scheme based on the time-lag extension (Biondi and Almomin, 2014).

Objective function

$$J(\mathbf{c}, \delta\mathbf{c}) = \underbrace{\sum_{i=1}^{N_s} \frac{1}{2} ||\mathbf{R}_i(\mathbf{u}_i + \delta\mathbf{u}_i) - \mathbf{d}_i||_2^2}_{\text{data fitting}} + \underbrace{\frac{\epsilon^2}{2} ||\; t\delta\mathbf{c}||_2^2}_{\text{regularization}}$$

The terms in the above expression are:

- N_s, ϵ, t : Number of sources, regularization strength, and time
- $\mathbf{c}, \delta\mathbf{c}$: Background and extended model
- $\mathbf{d}_i, \mathbf{R}_i$: Recorded data and sampling operator for i^{th} source
- $\mathbf{u}_i$: Primary wave field for i^{th} source (depends on $\mathbf{c}$)
- $\delta\mathbf{u}_i$: Extended wave field for i^{th} source (depends on $\mathbf{c}, \delta\mathbf{c}$)

A slight caveat

In the original TFWI paper, Biondi & Almomin thinks of $\mathbf{c}, \delta\mathbf{c}$ as follows:

There is a velocity function $\widetilde{\mathbf{c}}(\mathbf{x}, t)$ in the extended time lag space and $\mathbf{c}, \delta\mathbf{c}$ are built out of this function as

$$\mathbf{c}(\mathbf{x}) = \widetilde{\mathbf{c}}|_{t=0}$$
$$\delta\mathbf{c}(\mathbf{x}, t) = \delta\widetilde{\mathbf{c}}(\mathbf{x}, t).$$

Original objective:

$$J_{\text{TFWI}}(\widetilde{\mathbf{c}}) = \underbrace{\sum_{i=1}^{N_s} \frac{1}{2} ||\mathbf{R}_i(\mathbf{u}_i + \delta\mathbf{u}_i) - \mathbf{d}_i||_2^2}_{\text{data fitting}} + \underbrace{\frac{\epsilon^2}{2} ||\, t\widetilde{\mathbf{c}}||_2^2}_{\text{regularization}} .$$

Easier thing to do:

Consider $\mathbf{c}, \delta\mathbf{c}$ as completely independent objects.
(consistent with FWIME, and other extended schemes)

Computational bottlenecks

The major computational bottlenecks of the TFWI algorithm are:

Repeated wave equation solves: Both the forward and adjoint wave equations need to be solved per shot, whenever the gradient of the objective function needs to be computed.

Computing convolutions: The forward and adjoints of the extended Born modeling operator involves expensive convolutions, which must be carried out separately for each shot. This happens per gradient computation.

Objective function in frequency domain

We first Fourier transform (in time) the TFWI objective function.

Objective function

- Slightly modified time domain objective function :

$$J(\mathbf{c}, \delta\mathbf{c}) = \sum_{i=1}^{N_s} \frac{1}{2} ||\mathbf{R}_i(\mathbf{u}_i + \delta\mathbf{u}_i) - \mathbf{d}_i||_2^2 + \frac{\epsilon^2}{2} || \, t\delta\mathbf{c}||_2^2 + \frac{\gamma^2}{2} ||\delta\mathbf{c}||_2^2$$

- Frequency domain objective function (exactly equivalent) :

$$J_{\epsilon,\gamma}(\mathbf{c}, \delta\hat{\mathbf{c}}) = \sum_{i=1}^{N_s} \frac{1}{2} ||\mathbf{R}_i(\hat{\mathbf{u}}_i + \delta\hat{\mathbf{u}}_i) - \hat{\mathbf{d}}_i||_2^2 + \frac{\epsilon^2}{2} \left\| \frac{\partial}{\partial\omega} \delta\hat{\mathbf{c}} \right\|_2^2 + \frac{\gamma^2}{2} ||\delta\hat{\mathbf{c}}||_2^2$$

where all "hat" quantities are Fourier transforms of time domain quantities.

We will consider the "variable projection" method for solving this problem (Barnier and Biondi, 2018).

Helmholtz equation

We next obtain the following Helmholtz equations upon Fourier transforming the hyperbolic wave equation PDEs:

- Primary wavefield

$$\left(\nabla^2 + \frac{\omega^2}{\mathbf{c}^2(\mathbf{x})}\right) \hat{\mathbf{u}}_i(\mathbf{x}, \omega) = -\hat{\mathbf{f}}_i(\mathbf{x}, \omega)$$

- Extended Born modeled wavefield

$$\left(\nabla^2 + \frac{\omega^2}{\mathbf{c}^2(\mathbf{x})}\right) \delta\hat{\mathbf{u}}_i(\mathbf{x}, \omega) = \underbrace{\frac{2\omega^2}{\mathbf{c}^3(\mathbf{x})} \hat{\mathbf{u}}_i(\mathbf{x}, \omega) \delta\hat{\mathbf{c}}(\mathbf{x}, \omega)}_{\text{pointwise multiplication}}$$

where ω denotes the angular frequency.

In this talk...

- We will talk about solving the extended linear inverse problem in the first order Born scattering regime.
- A nice application of the regularized linear inverse problem is illumination compensation.

Causes of illumination artifacts in seismic images

The presence of “shadow zones” is extremely common in seismic imaging.

Imperfect acquisition geometry:

- Economic constraints (time and money).
- Physical constraints (rigs, human settlements).
- Algorithmic advancements to mitigate some of the acquisition related artifacts.

Complex velocity models:

- Presence of salt bodies makes subsalt imaging extremely difficult.
- These effects are present even if the acquisition geometry is perfect.

Mitigation of illumination artifacts in seismic images

Several strategies exist to mitigate the illumination artifacts:

Imperfect acquisition geometry: Mitigated before imaging.

- Interpolation to fill missing holes in the data.
- Fold normalization (e.g. Voronoi, or simply regridding the data).

Complex velocity models: Cannot be solved with simple processing algorithms.

- Typical idea — use regularization during imaging.
- Most easily done in the linearized waveform inversion (LWI) setting.
- Fancy stacking algorithms can be interpreted as a “mild” form of regularization.

We will discuss...

Regularization of the extended linearized waveform inversion problem, and how it helps to mitigate illumination artifacts caused by

- Presence of complex velocity models
- Imperfect acquisition geometry.

Prior work on regularized extended LWI

Not a new idea...

- Remove acquisition footprints due to missing data (Nemeth et al., 1999).
- Suppression of cross-talk and missing data artifacts in LWI of blended data (Tang and Biondi, 2009; Xue et al., 2015).
- To improve subsalt imaging typical idea is to perform extended LWI with regularization; e.g. *use geophysical regularizers along angles or offset ray parameter axis* (Prucha and Biondi, 2002; Clapp,2005; Clapp et al., 2005).

The Born scattering regime

In this talk we will consider the regime where first order Born scattering is valid.

- Decomposition of velocity into **background** and **reflectivity** components.
- Reflectivity is small in magnitude.
- Reflectivity is localized and / or averages to zero locally.

Similarities between Born and extended Born scattering

- Extended Born scattering

$$\left(\nabla^2 + \frac{\omega^2}{\mathbf{c}^2(\mathbf{x})}\right) \delta\hat{\mathbf{u}}_i(\mathbf{x},\omega) = \frac{2\omega^2}{\mathbf{c}^3(\mathbf{x})}\hat{\mathbf{u}}_i(\mathbf{x},\omega)\delta\hat{\mathbf{c}}(\mathbf{x},\omega)$$

- Born scattering

$$\left(\nabla^2 + \frac{\omega^2}{\mathbf{c}^2(\mathbf{x})}\right) \delta\hat{\mathbf{v}}_i(\mathbf{x},\omega) = \frac{2\omega^2}{\mathbf{c}^3(\mathbf{x})}\hat{\mathbf{u}}_i(\mathbf{x},\omega)\delta\mathbf{m}(\mathbf{x})$$

Note: No dependence of $\delta\mathbf{m}$ on frequency ω.

The variable projection step

Solve for fixed $\mathbf{c}$ (the variable projection step):

$$J_{\epsilon,\gamma}(\delta\hat{\mathbf{c}}) = \sum_{i=1}^{N_s} \frac{1}{2} ||\mathbf{R}_i(\hat{\mathbf{u}}_i + \delta\hat{\mathbf{u}}_i) - \hat{\mathbf{d}}_i||_2^2 + \underbrace{\frac{\epsilon^2}{2} \left\| \frac{\partial}{\partial\omega} \delta\hat{\mathbf{c}} \right\|_2^2}_{\text{promotes constancy across frequency}} + \frac{\gamma^2}{2} ||\delta\hat{\mathbf{c}}||_2^2.$$

Ideal result in Born regime

If we are in the Born scattering regime, one should not need to vary $\delta\hat{\mathbf{c}}$ with ω to fit the data. So we are ideally looking for a solution $\delta\hat{\mathbf{c}}$ that does not change with frequency.

Some terminology

- **Inverted images**

$$\delta\hat{\mathbf{c}}^*(\mathbf{x},\omega) = \underset{\delta\hat{\mathbf{c}}}{\text{argmin}} \;\; J_{\epsilon,\gamma}(\delta\mathbf{c}).$$

- **Stacked inverted image**

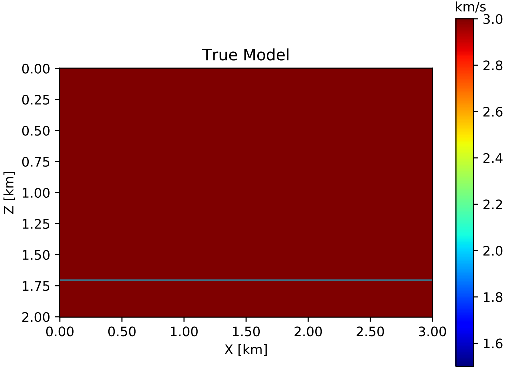
$$\sum_{\omega} \delta\hat{\mathbf{c}}^*(\mathbf{x},\omega).$$

Experiment 1:

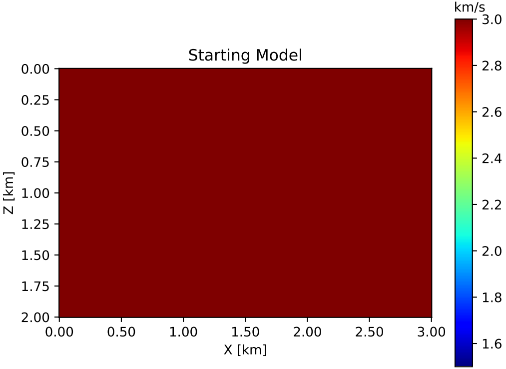
Illumination compensation when shadow zones are created by a complex velocity model.

($\gamma = 0$ for all experiments)

Velocity models without Gaussian anomaly

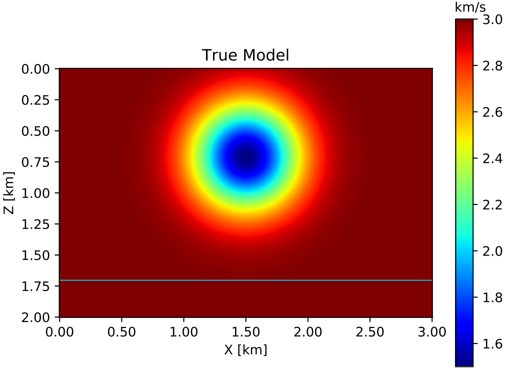


True velocity model

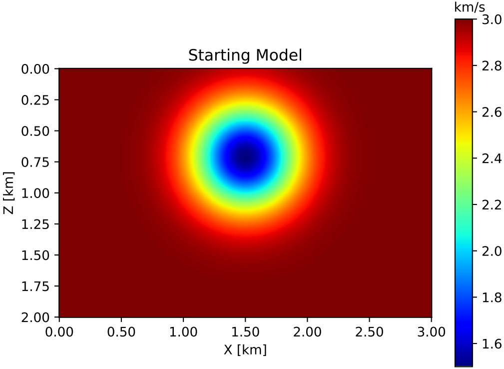


Starting velocity model

Velocity models with Gaussian anomaly

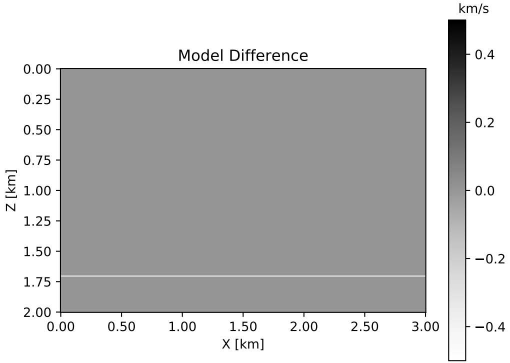


True velocity model

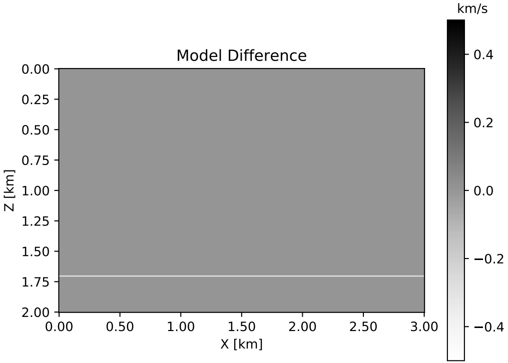


Starting velocity model

Model differences



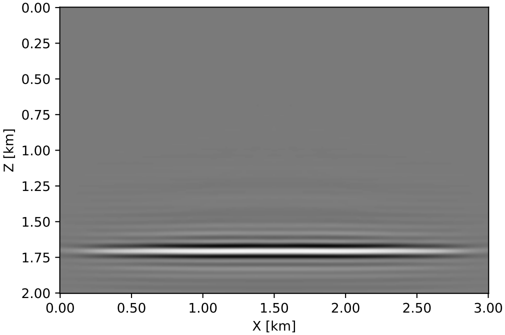
No Gaussian anomaly case



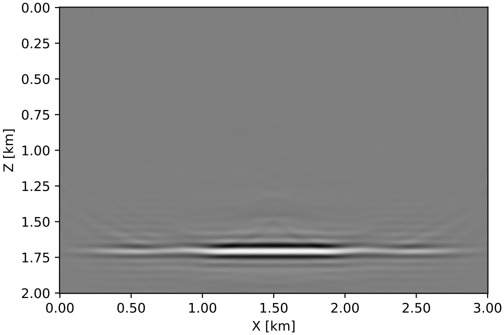
Gaussian anomaly case

Inversion results — stacked images without regularization

Stacked inverted images without regularization ($\epsilon = 0$)



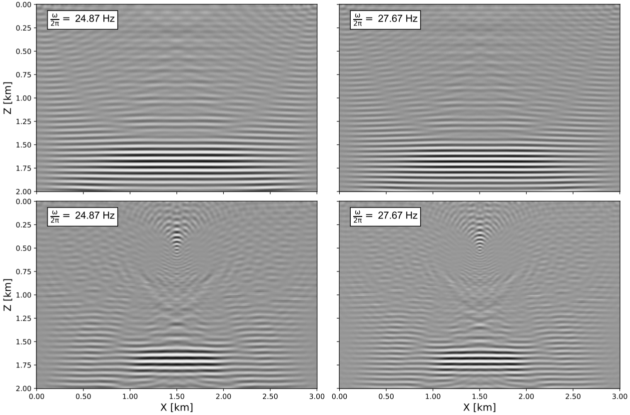
No Gaussian anomaly case



Gaussian anomaly case

Inversion results — frequency image comparisons without regularization

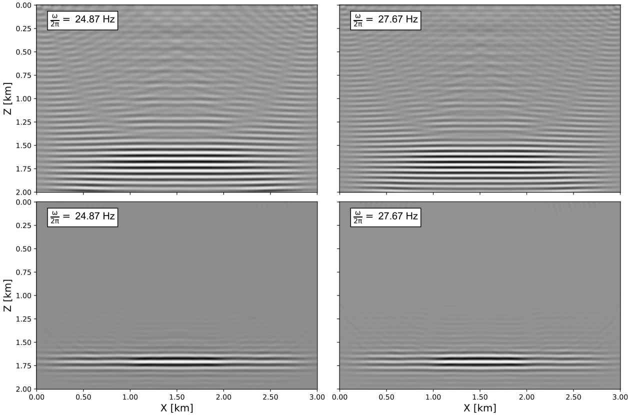
Real part of the inverted image without regularization ($\epsilon = 0$)



Top row: No Gaussian anomaly case
Bottom row: Gaussian anomaly case

Inversion results — frequency image comparisons with regularization

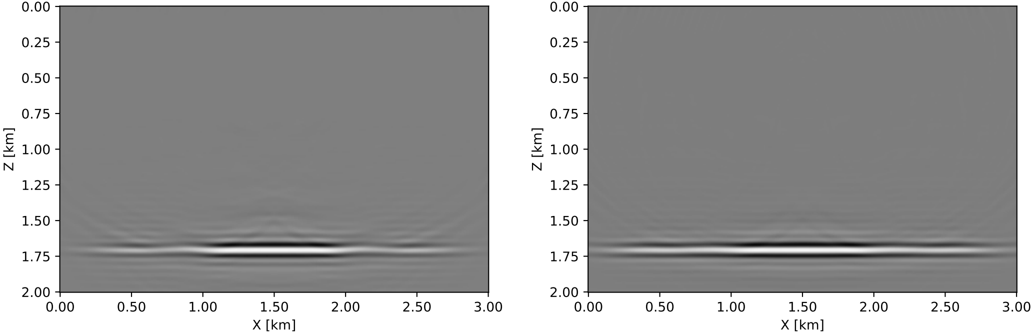
Real part of the inverted image



Top row: No Gaussian anomaly case ($\epsilon = 0$)
Bottom row: Gaussian anomaly case ($\epsilon \neq 0$)

Inversion results — comparison of stacked images with and without regularization

Stacked inverted image comparison for the Gaussian anomaly case



With regularization

Without regularization

Numerical Examples

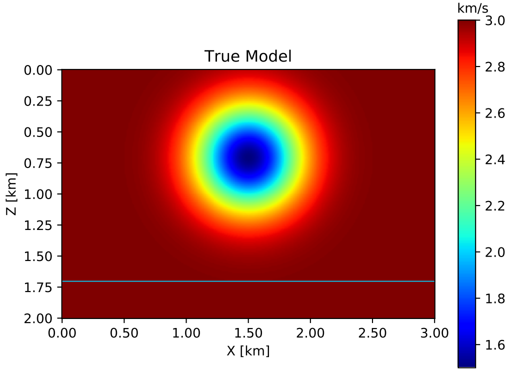
Experiment 2:

Illumination compensation when shadow zones are created by a complex velocity model and acquisition hole.

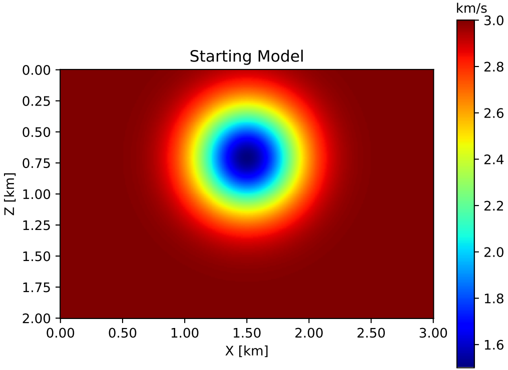
($\gamma = 0$ for all experiments)

Velocity models with Gaussian anomaly

Sources and receivers removed in a region of width 0.5 km over the Gaussian anomaly.



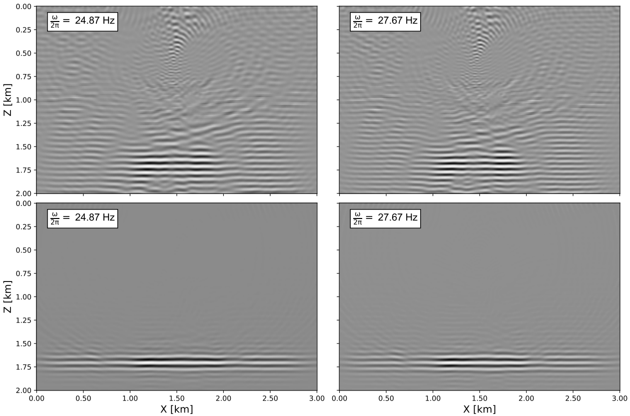
True velocity model



Starting velocity model

Inversion results — frequency images with and without regularization

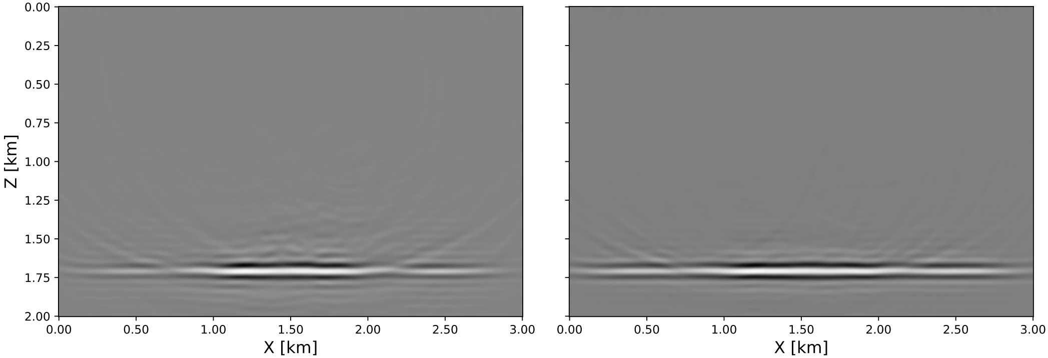
Real part of the inverted image



Top row: No regularization
Bottom row: Regularized

Inversion results — comparison of stacked images with and without regularization

Stacked inverted image comparison



Left: Without regularization; **Right:** With regularization

Where I'm headed with the project...

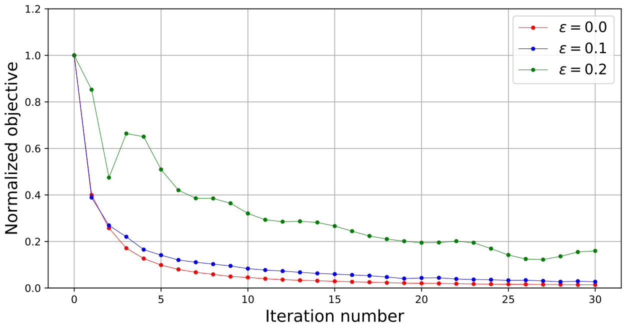
- Faster solvers for the least squares problem
- Methods to implement non-differentiable regularizers
  - Total-variation in frequency
  - L1 constraints
  - Convex constraints
- Randomized linear solvers (including shot encoding, randomization in frequency)

Faster linear solvers for the least squares problem

- Currently the solver solves the normal equations using CG.
- Better solvers available - **CGLS**, **LSQR**, LSMR. They are better behaved for ill-conditioned problems. The first two have been implemented (under testing).
- Implement preconditioners

Convergence rates comparison

Normalized objective function with respect to initial (Intially $\delta\hat{\mathbf{c}} = 0$)



Non-differentiable regularizers and convex constraints

Consider other forms of regularization, not necessarily L2 type.

Example: TV regularization

Solve for fixed $\mathbf{c}$ (no longer variable projection, but this is **alternating minimization**):

$$J_\epsilon(\delta\hat{\mathbf{c}}) = \sum_{i=1}^{N_s} \frac{1}{2} ||\mathbf{R}_i(\hat{\mathbf{u}}_i + \delta\hat{\mathbf{u}}_i) - \hat{\mathbf{d}}_i||_2^2 + \frac{\epsilon^2}{2} \| \nabla \delta\hat{\mathbf{c}}\|_1 + \text{others}.$$

Comments and acknowledgments

- I originally started this project with the broad goal of speeding up the solution to the variable projection step in TFWI.
- Guillaume and Ettore pointed out that this work could be more broadly applicable to other forms of linearized extended inversion. For example, matching data modeled with elastic model using extended acoustic modeling (same context as Ettore's work).
- The particular application discussed in this talk is Biondo's idea, based on prior work with Marie Clapp.

Questions?