Sensor Optimisation in Seismic Imaging Via Bilevel Learning

Shaunagh Downing
with Silvia Gazzola, Ivan Graham, Euan Spence

University of Bath

Outline

- **Part I**
  - Seismic Imaging

- **Part II**
  - Sensor Placement Optimisation

Part I: Seismic Imaging

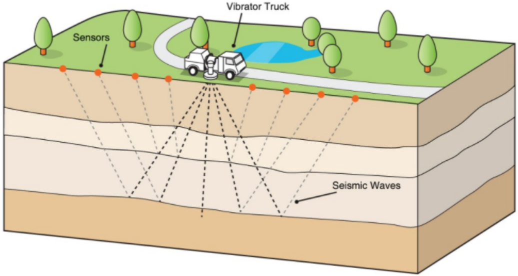
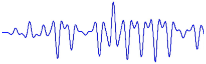


Figure: Seismic Acquisition

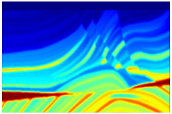
- A **source** creates a disturbance in the form of a wave.
- Wave travels through earth and reflects off material properties interfaces.
- **Sensors** record the returning wave.

Seismic Imaging

Aim: Use the measured data to create an image of the subsurface (image = map of geological properties of subsurface)

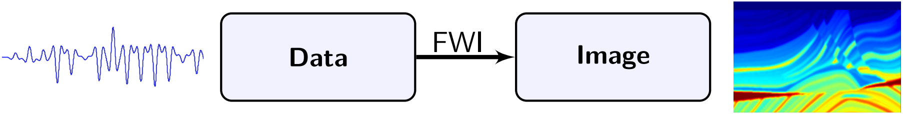


Data → **Image**



Seismic Imaging

Aim: Use the measured data to create an image of the subsurface (image = map of geological properties of subsurface)



Method: **Full Waveform Inversion**

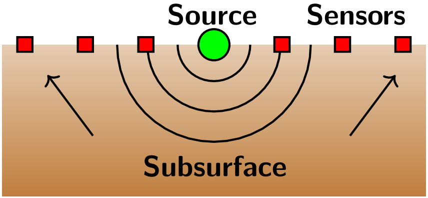
- Goal: Estimation of subsurface properties (**model** $\mathbf{m}$)
- How? Minimising the *data misfit* $\Delta \mathbf{d}$
- *Data Misfit* $\Delta \mathbf{d}$ = difference between **measured data** ($\mathbf{d}$) and **modelled data** ($\mathbf{d}^{mod}$)

FWI

Goal: Find the model $\mathbf{m}$ that predicts the measured data $\mathbf{d}$ best

$$\mathbf{m}^{FWI}(\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m},\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \frac{1}{2}\|\underbrace{\mathbf{d} - \mathbf{d}^{mod}(\mathbf{m},\mathbf{p})}_{\Delta\mathbf{d}}\|_2^2$$

- ϕ: Misfit function $\frac{1}{2}\|\Delta\mathbf{d}\|_2^2$
- $\mathbf{d}$: Measured data
- $\mathbf{m}$: Model - geological property we want to find
- $\mathbf{d}^{mod}$: Modelled data $= R(\mathbf{p})\mathbf{u}(\mathbf{m})$
  - $\mathbf{u}$: Wavefield - numerical solution of wave equation
  - R: Sampling operator at sensor positions $\mathbf{p}$

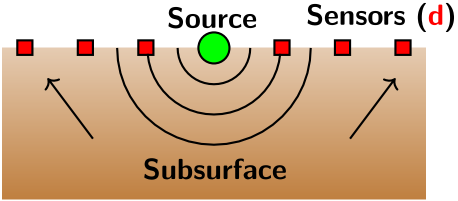


FWI

Goal: Find the model $\mathbf{m}$ that predicts the measured data $\mathbf{d}$ best

$$\mathbf{m}^{FWI}(\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{mod}(\mathbf{m}, \mathbf{p})||_2^2$$

- ϕ: Misfit function
- $\mathbf{d}$: Measured data
- $\mathbf{m}$: Model - geological property we want to find
- $\mathbf{d}^{mod}$: Modelled data $= R(\mathbf{p})\mathbf{u}(\mathbf{m})$
  - $\mathbf{u}$: Wavefield - numerical solution of wave equation
  - R: Sampling operator at sensor positions $\mathbf{p}$

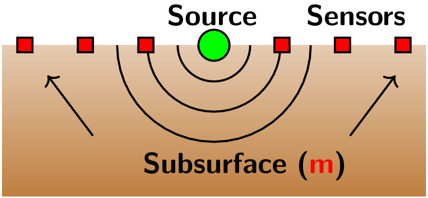


FWI

Goal: Find the model $\mathbf{m}$ that predicts the measured data $\mathbf{d}$ best

$$\mathbf{m}^{FWI}(\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{mod}(\mathbf{m}, \mathbf{p})||_2^2$$

- ϕ: Misfit function
- $\mathbf{d}$: Measured data
- $\mathbf{m}$: Model - geological property we want to find
- $\mathbf{d}^{mod}$: Modelled data $= R(\mathbf{p})\mathbf{u}(\mathbf{m})$
  - $\mathbf{u}$: Wavefield - numerical solution of wave equation
  - R: Sampling operator at sensor positions $\mathbf{p}$

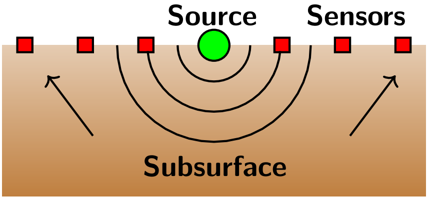


FWI

Goal: Find the model $\mathbf{m}$ that predicts the measured data $\mathbf{d}$ best

$$\mathbf{m}^{FWI}(\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{mod}(\mathbf{m}, \mathbf{p})||_2^2$$

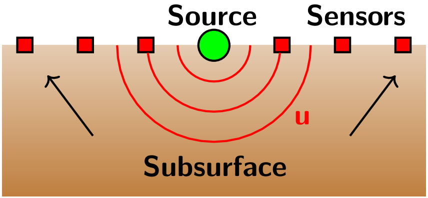
- ϕ: Misfit function
- $\mathbf{d}$: Measured data
- $\mathbf{m}$: Model - geological property we want to find
- $\mathbf{d}^{mod}$: Modelled data $= R(\mathbf{p})\mathbf{u}(\mathbf{m})$
  - $\mathbf{u}$: Wavefield - numerical solution of wave equation
  - R: Sampling operator at sensor positions $\mathbf{p}$



Goal: Find the model $\mathbf{m}$ that predicts the measured data $\mathbf{d}$ best

$$\mathbf{m}^{FWI}(\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{mod}(\mathbf{m}, \mathbf{p})||_2^2$$

- ϕ: Misfit function
- $\mathbf{d}$: Measured data
- $\mathbf{m}$: Model - geological property we want to find
- $\mathbf{d}^{mod}$: Modelled data $= R(\mathbf{p})\mathbf{u}(\mathbf{m})$
  - $\mathbf{u}$: Wavefield - numerical solution of wave equation
  - R: Sampling operator at sensor positions $\mathbf{p}$

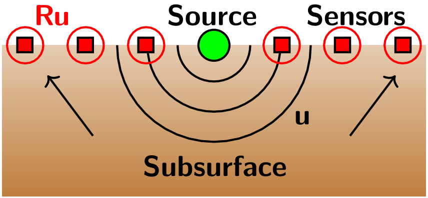


FWI

Goal: Find the model $\mathbf{m}$ that predicts the measured data $\mathbf{d}$ best

$$\mathbf{m}^{FWI}(\mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) = \underset{\mathbf{m}}{\operatorname{argmin}}\, \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{mod}(\mathbf{m}, \mathbf{p})||_2^2$$

- ϕ: Misfit function
- $\mathbf{d}$: Measured data
- $\mathbf{m}$: Model - geological property we want to find
- $\mathbf{d}^{mod}$: Modelled data $= R(\mathbf{p})\mathbf{u}(\mathbf{m})$
  - $\mathbf{u}$: Wavefield - numerical solution of wave equation
  - R: Sampling operator at sensor positions $\mathbf{p}$



Full FWI Formulation

Multiple source, multiple frequency problem

Find the model $\mathbf{m}^{FWI}$ such that

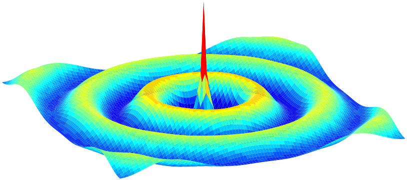
$$\begin{aligned}\mathbf{m}^{FWI}(\mathbf{p}) &= \operatorname*{argmin}_{\mathbf{m}} \phi(\mathbf{m}, \mathbf{p}, s, \omega) \\ &= \operatorname*{argmin}_{\mathbf{m}} \sum_{s} \sum_{\omega} \frac{1}{2} ||\mathbf{d}(s, \omega) - R(\mathbf{p})\mathbf{u}(\mathbf{m}, s, \omega)||_2^2 + \text{Regularisation}\end{aligned}$$

Forward Modelling

- Solving wave equation to compute modelled data $R\mathbf{u}$
- Model **acoustic waves** in the **frequency** domain $\Longrightarrow$ **Helmholtz Equation**

$$\left(-\omega^2 m - \nabla^2\right) u = q$$

- Model $\mathbf{m} = \dfrac{1}{\textbf{wavespeed}^2}$
- Solve discretised equation numerically $A(\mathbf{m},\omega)\mathbf{u}(\mathbf{m},\omega) = \mathbf{q}$



FWI Process

Initial model

↓

Forward Modelling

Find $\mathbf{u}(\mathbf{m})$ by solving the wave equation

FWI Process

Initial model

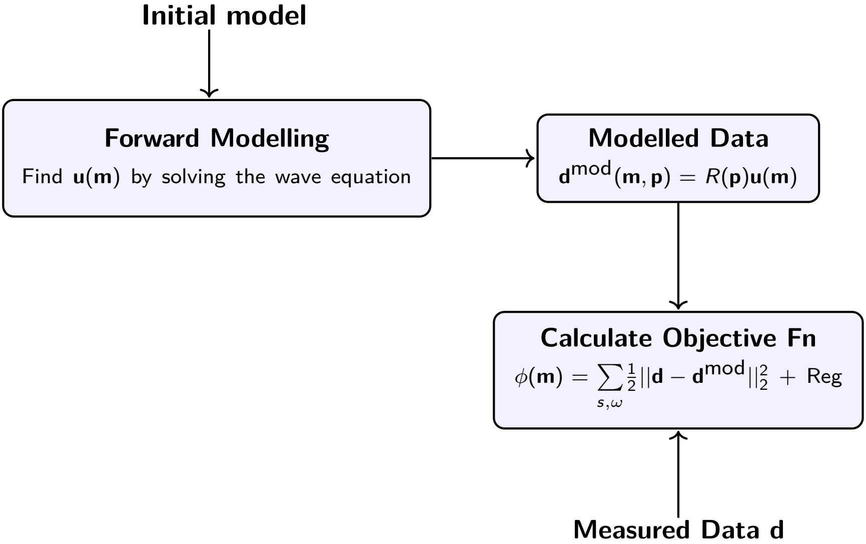
Forward Modelling

Find $\mathbf{u}(\mathbf{m})$ by solving the wave equation

Modelled Data

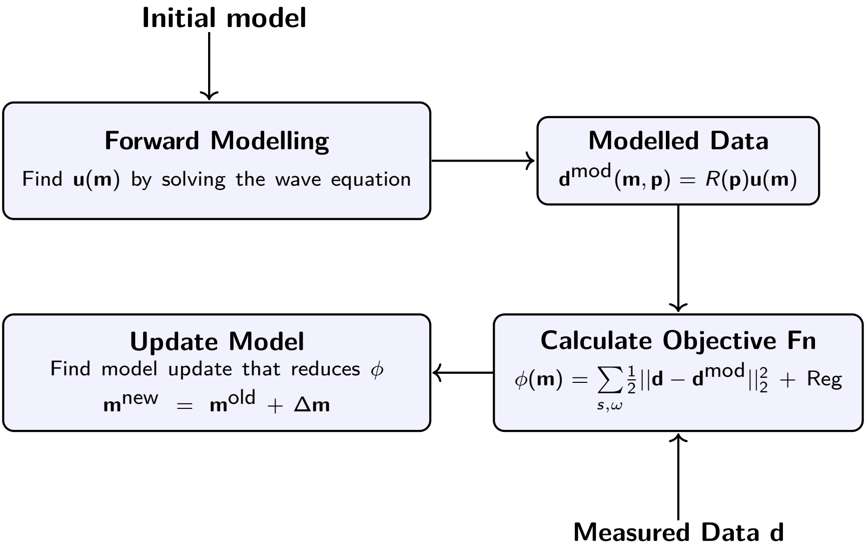
$\mathbf{d}^{\text{mod}}(\mathbf{m}, \mathbf{p}) = R(\mathbf{p})\mathbf{u}(\mathbf{m})$

FWI Process

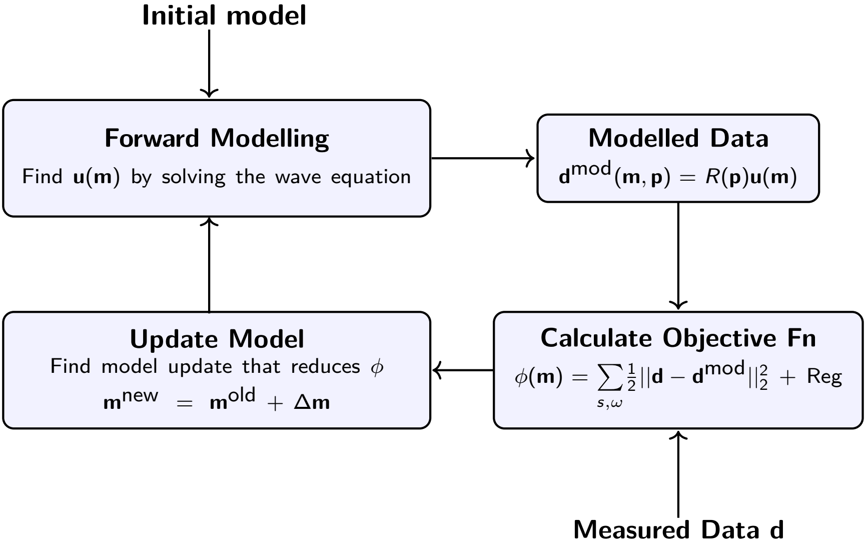


Initial model
Forward Modelling
Find $\mathbf{u}(\mathbf{m})$ by solving the wave equation
Modelled Data
$\mathbf{d}^{\mathrm{mod}}(\mathbf{m},\mathbf{p}) = R(\mathbf{p})\mathbf{u}(\mathbf{m})$
Calculate Objective Fn
$\phi(\mathbf{m}) = \sum_{s,\omega} \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{\mathrm{mod}}||_2^2 + \text{Reg}$
Measured Data d

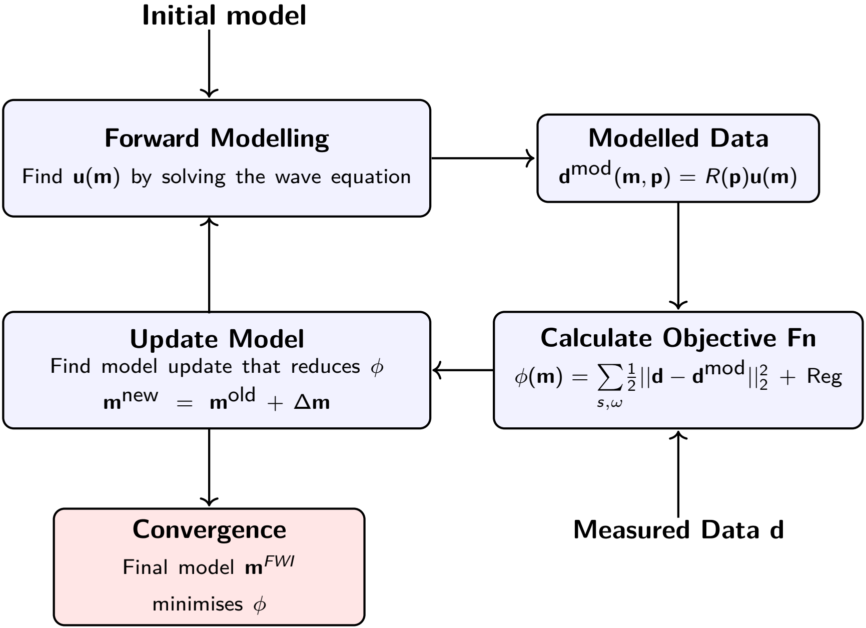
FWI Process


Initial model
Forward Modelling
Find $\mathbf{u}(\mathbf{m})$ by solving the wave equation
Modelled Data
$\mathbf{d}^{\text{mod}}(\mathbf{m}, \mathbf{p}) = R(\mathbf{p})\mathbf{u}(\mathbf{m})$
Update Model
Find model update that reduces ϕ
$\mathbf{m}^{\text{new}} = \mathbf{m}^{\text{old}} + \Delta\mathbf{m}$
Calculate Objective Fn
$\phi(\mathbf{m}) = \sum_{s,\omega} \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{\text{mod}}||_2^2 + \text{Reg}$
Measured Data $\mathbf{d}$

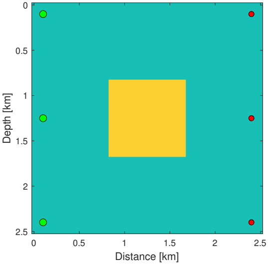
FWI Process


Initial model
Forward Modelling
Find $\mathbf{u}(\mathbf{m})$ by solving the wave equation
Modelled Data
$\mathbf{d}^{\text{mod}}(\mathbf{m}, \mathbf{p}) = R(\mathbf{p})\mathbf{u}(\mathbf{m})$
Calculate Objective Fn
$\phi(\mathbf{m}) = \sum_{s,\omega} \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{\text{mod}}||_2^2 + \text{Reg}$
Update Model
Find model update that reduces ϕ
$\mathbf{m}^{\text{new}} = \mathbf{m}^{\text{old}} + \Delta\mathbf{m}$
Measured Data $\mathbf{d}$

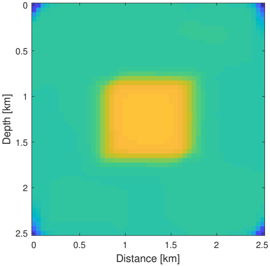
FWI Process


Initial model
Forward Modelling
Find $\mathbf{u}(\mathbf{m})$ by solving the wave equation
Modelled Data
$\mathbf{d}^{\text{mod}}(\mathbf{m}, \mathbf{p}) = R(\mathbf{p})\mathbf{u}(\mathbf{m})$
Update Model
Find model update that reduces ϕ
$\mathbf{m}^{\text{new}} = \mathbf{m}^{\text{old}} + \Delta\mathbf{m}$
Calculate Objective Fn
$\phi(\mathbf{m}) = \sum_{s,\omega} \frac{1}{2} ||\mathbf{d} - \mathbf{d}^{\text{mod}}||_2^2 + \text{Reg}$
Convergence
Final model $\mathbf{m}^{FWI}$
minimises ϕ
Measured Data d

Example



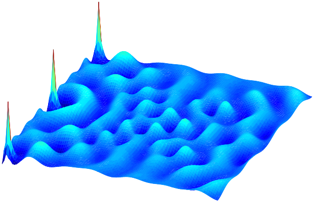
(a) Ground Truth



(b) FWI Reconstruction

● = Sources

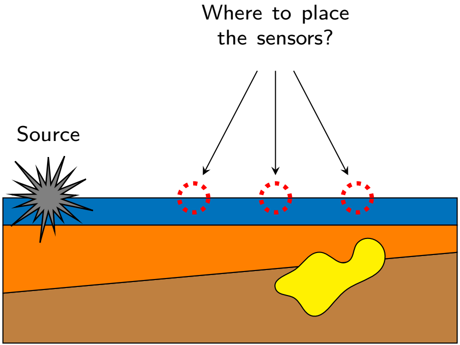
● = Sensors



Part II: Sensor Optimisation

Goal

Optimise the placement of sensors in the seismic imaging process



Problem Formulation

Problem Statement

Given a training set of N models $\mathbf{m}'$, we learn the sensor positions $\mathbf{p}$, such that the FWI output model $\mathbf{m}^{FWI}$ is as close to the ground truth as possible.

- **Objective function:**

$$\psi(\mathbf{p}) = \frac{1}{2N} \sum_{\mathbf{m}'} ||\mathbf{m}' - \mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}')||_2^2$$

- ψ measures average error in FWI reconstruction

Problem Formulation

Problem Statement

Given a training set of N models $\mathbf{m}'$, we learn the sensor positions $\mathbf{p}$, such that the FWI output model $\mathbf{m}^{FWI}$ is as close to the ground truth as possible.

- **Objective function:**

$$\psi(\mathbf{p}) = \frac{1}{2N} \sum_{\mathbf{m}'} ||\mathbf{m}' - \mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}')||_2^2$$

- ψ measures average error in FWI reconstruction

- Solution to FWI optimisation problem where ‘measured’ data $\mathbf{d}$ is produced from $\mathbf{m}'$

Bilevel Optimisation

- 2 levels of optimisation

- **Upper Level**: Sensor Placement Optimisation
- **Lower Level**: Seismic Imaging (FWI)

Bilevel Problem Statement

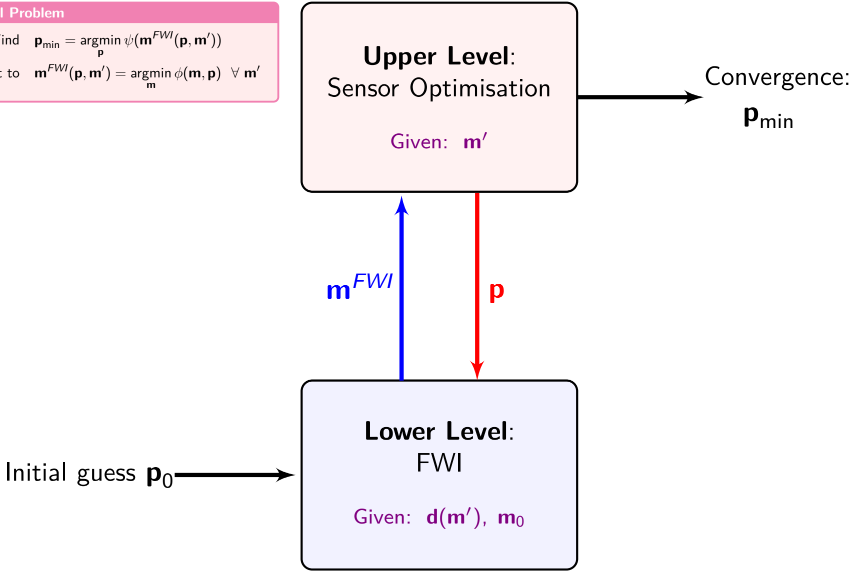
$$\begin{aligned} \text{Find} \quad & \mathbf{p}_{\text{min}} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}')) \\ \text{subject to} \quad & \mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}') = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) \quad \forall\ \mathbf{m}' \end{aligned}$$

Bilevel Optimisation Process

Bilevel Problem

$$\text{Find} \quad \mathbf{p}_{\min} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p},\mathbf{m}'))$$

$$\text{subject to} \quad \mathbf{m}^{FWI}(\mathbf{p},\mathbf{m}') = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m},\mathbf{p}) \quad \forall\ \mathbf{m}'$$



Solution Approach

Bilevel Problem

$$\text{Find} \quad \mathbf{p}_{\min} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'))$$

$$\text{subject to} \quad \mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}') = \underset{\mathbf{m}}{\operatorname{argmin}}\, \phi(\mathbf{m}, \mathbf{p}) \quad \forall\ \mathbf{m}'$$

Gradient

$$\nabla_{\mathbf{p}}\psi = -\frac{1}{N}\sum_{\mathbf{m}'}\left(\frac{d\mathbf{m}^{FWI}}{d\mathbf{p}}\right)^T \left(\mathbf{m}' - \mathbf{m}^{FWI}\right)$$

- Quasi-Newton method $\Longrightarrow$ need to compute the gradient

Solution Approach

Bilevel Problem

$$\text{Find} \quad \mathbf{p}_{\min} = \operatorname*{argmin}_{\mathbf{p}} \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'))$$

$$\text{subject to} \quad \mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}') = \operatorname*{argmin}_{\mathbf{m}} \phi(\mathbf{m}, \mathbf{p}) \quad \forall\ \mathbf{m}'$$

Gradient

$$\nabla_{\mathbf{p}} \psi = -\frac{1}{N} \sum_{\mathbf{m}'} \left(\frac{d\mathbf{m}^{FWI}}{d\mathbf{p}} \right)^T \left(\mathbf{m}' - \mathbf{m}^{FWI} \right)$$

- Quasi-Newton method $\implies$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation

Solution Approach

Bilevel Problem

$$\text{Find} \quad \mathbf{p}_{\text{min}} = \underset{\mathbf{p}}{\text{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'))$$

$$\text{subject to} \quad \nabla_{\mathbf{m}} \phi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'), \mathbf{p}) = \mathbf{0} \quad \forall\ \mathbf{m}'$$

Gradient

$$\nabla_{\mathbf{p}} \psi = -\frac{1}{N} \sum_{\mathbf{m}'} \left(\frac{d\mathbf{m}^{FWI}}{d\mathbf{p}} \right)^T \left(\mathbf{m}' - \mathbf{m}^{FWI} \right)$$

- Quasi-Newton method $\Longrightarrow$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation
- **Instead:** replace the lower-level with its optimality condition

Solution Approach

Bilevel Problem

$$\begin{aligned} \text{Find} \quad & \mathbf{p}_{\min} = \operatorname*{argmin}_{\mathbf{p}} \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}')) \\ \text{subject to} \quad & \nabla_{\mathbf{m}} \phi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'), \mathbf{p}) = \mathbf{0} \quad \forall\ \mathbf{m}' \end{aligned}$$

Gradient

$$\nabla_{\mathbf{p}} \psi = -\frac{1}{N} \sum_{\mathbf{m}'} \left(\frac{d\mathbf{m}^{FWI}}{d\mathbf{p}} \right)^T \left(\mathbf{m}' - \mathbf{m}^{FWI} \right)$$

- Quasi-Newton method $\implies$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation
- **Instead:** replace the lower-level with its optimality condition
- Chain rule $\implies$ $\frac{d\mathbf{m}^{FWI}}{d\mathbf{p}} = -H^{-1}(\mathbf{m}^{FWI}) \nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$

Solution Approach

Bilevel Problem

$$\text{Find} \quad \mathbf{p}_{\min} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'))$$
$$\text{subject to} \quad \nabla_{\mathbf{m}} \phi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'), \mathbf{p}) = \mathbf{0} \quad \forall\ \mathbf{m}'$$

Gradient

$$\nabla_{\mathbf{p}} \psi = \frac{1}{N} \sum_{\mathbf{m}'} \left(\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})\right)^T \delta(\mathbf{m}^{FWI})$$

- Quasi-Newton method $\implies$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation
- **Instead:** replace the lower-level with its optimality condition
- Chain rule $\implies$ $\dfrac{d\mathbf{m}^{FWI}}{d\mathbf{p}} = -H^{-1}(\mathbf{m}^{FWI}) \nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$
- Computing the gradient:
  - $\delta(\mathbf{m}^{FWI})$: Solve the linear system $H(\mathbf{m}^{FWI})\delta(\mathbf{m}^{FWI}) = \mathbf{m}' - \mathbf{m}^{FWI}$
  - $\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$: Solve $dim \times \#\, sensors \times \#\, s \times \#\, \omega$ PDEs

Solution Approach

Bilevel Problem

Find $\quad \mathbf{p}_{\min} = \operatorname*{argmin}_{\mathbf{p}} \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'))$

subject to $\quad \nabla_{\mathbf{m}} \phi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'), \mathbf{p}) = \mathbf{0} \quad \forall\ \mathbf{m}'$

Gradient

$$\nabla_{\mathbf{p}} \psi = \frac{1}{N} \sum_{\mathbf{m}'} \left(\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI}) \right)^T \boldsymbol{\delta}(\mathbf{m}^{FWI})$$

- Quasi-Newton method $\implies$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation
- **Instead:** replace the lower-level with its optimality condition
- Chain rule $\implies$ $\frac{d\mathbf{m}^{FWI}}{d\mathbf{p}} = -H^{-1}(\mathbf{m}^{FWI}) \nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$
- Computing the gradient:
  - $\boldsymbol{\delta}(\mathbf{m}^{FWI})$: Solve the linear system $H(\mathbf{m}^{FWI})\delta(\mathbf{m}^{FWI}) = \mathbf{m}' - \mathbf{m}^{FWI}$
  - $\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$: Solve $dim \times \# sensors \times \# s \times \# \omega$ PDEs

Solution Approach

Bilevel Problem

$$\begin{aligned} \text{Find} \quad & \mathbf{p}_{\min} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}')) \\ \text{subject to} \quad & \nabla_{\mathbf{m}} \phi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'), \mathbf{p}) = \mathbf{0} \quad \forall\ \mathbf{m}' \end{aligned}$$

Gradient

$$\nabla_{\mathbf{p}} \psi = \frac{1}{N} \sum_{\mathbf{m}'} \left(\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI}) \right)^T \delta(\mathbf{m}^{FWI})$$

- Quasi-Newton method $\implies$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation
- **Instead:** replace the lower-level with its optimality condition
- Chain rule $\implies \dfrac{d\mathbf{m}^{FWI}}{d\mathbf{p}} = -H^{-1}(\mathbf{m}^{FWI}) \nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$
- Computing the gradient:
  - $\delta(\mathbf{m}^{FWI})$: Solve the linear system $H(\mathbf{m}^{FWI})\delta(\mathbf{m}^{FWI}) = \mathbf{m}' - \mathbf{m}^{FWI}$
  - $\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$: Solve $dim \times \#\, sensors \times \#\, s \times \#\, \omega$ PDEs

Solution Approach

Bilevel Problem

Find $\quad \mathbf{p}_{\min} = \underset{\mathbf{p}}{\operatorname{argmin}}\, \psi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'))$

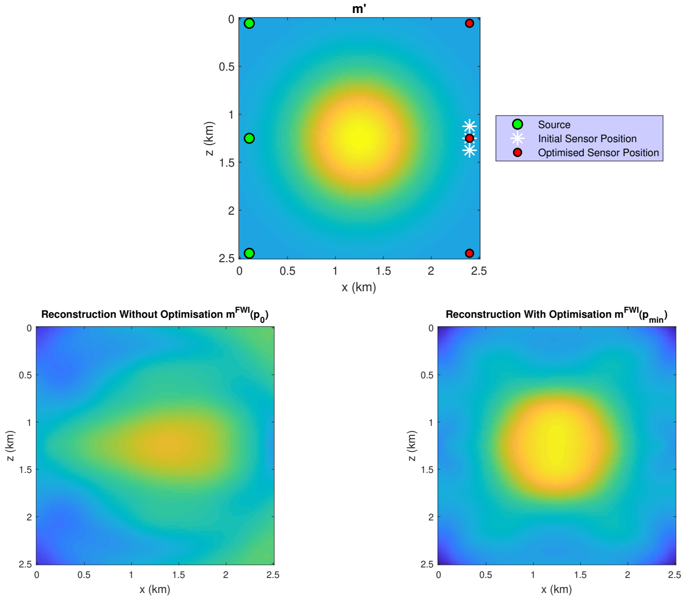
subject to $\quad \nabla_{\mathbf{m}} \phi(\mathbf{m}^{FWI}(\mathbf{p}, \mathbf{m}'), \mathbf{p}) = \mathbf{0} \quad \forall\ \mathbf{m}'$

Gradient

$$\nabla_{\mathbf{p}} \psi = \frac{1}{N} \sum_{\mathbf{m}'} \left(\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI}) \right)^T \boldsymbol{\delta}(\mathbf{m}^{FWI})$$

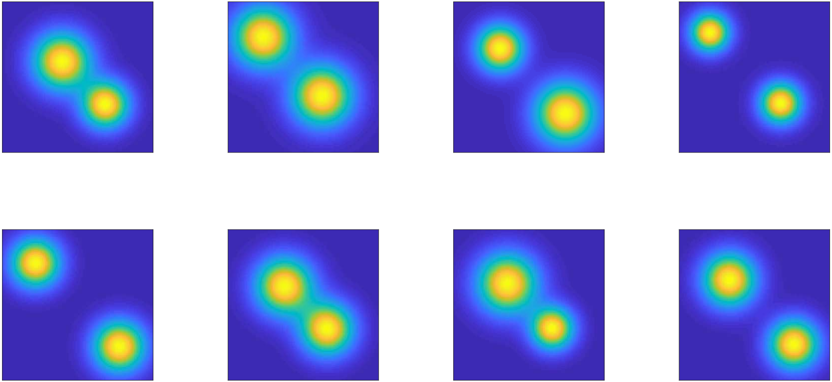
- Quasi-Newton method $\implies$ need to compute the gradient
- **Difficulty:** differentiating the argmin operation
- **Instead:** replace the lower-level with its optimality condition
- Chain rule $\implies$ $\dfrac{d\mathbf{m}^{FWI}}{d\mathbf{p}} = -H^{-1}(\mathbf{m}^{FWI}) \nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$
- Computing the gradient:
  - $\boldsymbol{\delta}(\mathbf{m}^{FWI})$: Solve the linear system $H(\mathbf{m}^{FWI}) \boldsymbol{\delta}(\mathbf{m}^{FWI}) = \mathbf{m}' - \mathbf{m}^{FWI}$
  - ~~$\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI})$: Solve $dim \times \#\, sensors \times \#\, s \times \#\, \omega$ PDEs~~
  - $\left(\nabla_{\mathbf{m},\mathbf{p}} \phi(\mathbf{m}^{FWI}) \right)^T \boldsymbol{\delta}(\mathbf{m}^{FWI})$: $\#$ PDEs independent of $dim \times \#\, sensors$

Example 1

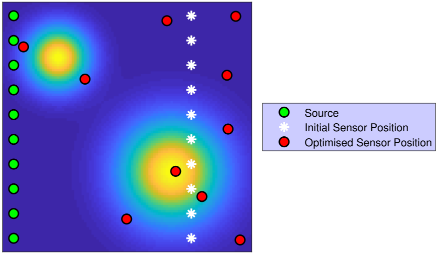

m'
0
0.5
1
1.5
2
2.5
z (km)
0
0.5
1
1.5
2
2.5
x (km)
Source
Initial Sensor Position
Optimised Sensor Position
Reconstruction Without Optimisation $m^{FWI}(p_0)$
Reconstruction With Optimisation $m^{FWI}(p_{min})$
z (km)
x (km)
z (km)
x (km)

Example 2 - Setup

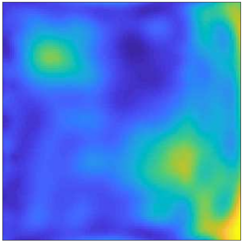
- Learn optimal sensor positions AND **level of regularisation**
- Set of **training models** of certain class - smooth circles of different size, wavespeeds and positions along diagonal
- Subset of training models:



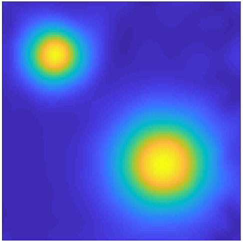
Example 2 - Results



Reconstruction Without Optimisation $m^{FWI}(p_0)$

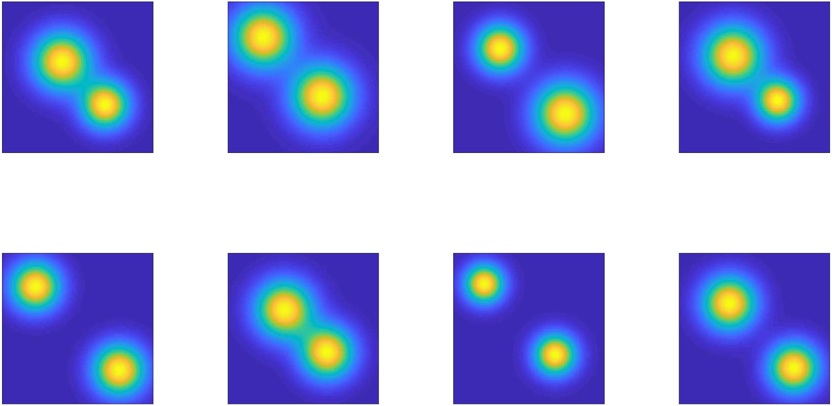


Reconstruction With Optimisation $m^{FWI}(p_{min})$

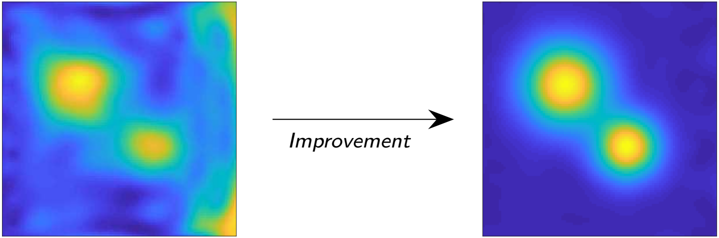


Example 2 - Testing

- How well do these results work on other models in this class?
- Subset of 50 random <u>testing models</u>:



Example 2 - Testing Results

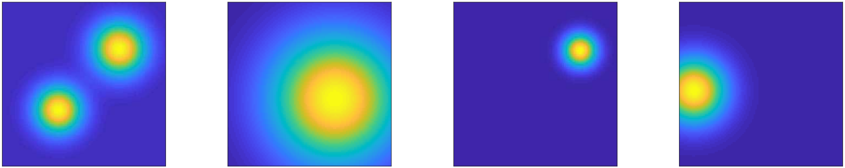


ψ_{start} ψ_{optim}

- **Improvement factor** $= \dfrac{\psi_{start}}{\psi_{optim}}$
- Over 50 testing models:
  - Average improvement factor $= 536$
  - Range of improvement factors $= 287$ - 842

Example 2 - Further Testing

- How well do these results work on models *outside* this class?
- Subset of 150 random testing models:

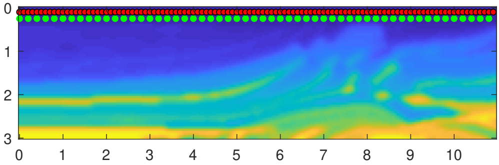


- Over 150 testing models:
  - Average improvement factor $= 533$
  - Range of improvement factors $= 42$ - 1171

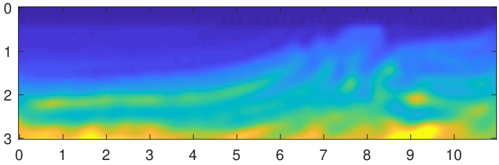
Future Work

- Extend current bilevel framework to learn the optimal source placement and the number of sensors and sources
- **Preliminary Results:**

- Starting Setup $\Longrightarrow$ 110 sensors



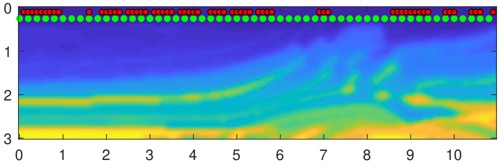
- Reconstruction



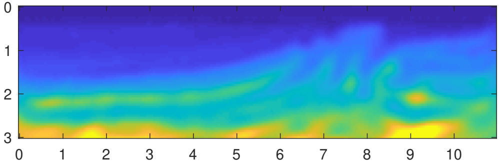
Future Work

- Extend current bilevel framework to learn the optimal source placement and the number of sensors and sources
- **Preliminary Results:**

- Optimised Weights $\implies$ 62 sensors



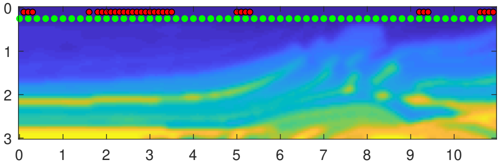
- Reconstruction $\approx 5\%$ worse



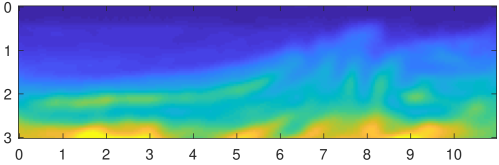
Future Work

- Extend current bilevel framework to learn the optimal source placement and the number of sensors and sources
- **Preliminary Results:**

- Larger Sparsity Parameter $\Longrightarrow$ 33 sensors



- Reconstruction $\approx$ 22% worse



Summary

- Seismic Imaging can be performed with Full Waveform Inversion
- Sensor placement can be optimised using a bilevel learning framework
- Bilevel learning results in significant improvements in FWI images


THANK
YOU