Localized Wavefield Inversion (LWI): an Adaptation of Multi-Block ADMM for Localized FWI

Hossein S. Aghamiry

June 22, 2022

Solvers for frequency-domain wave problems and applications

WIND consortium

https://www.geoazur.fr/WIND

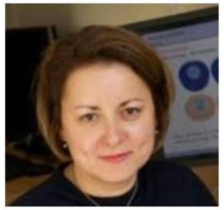
Permanent staff

S. Operto (UCA)

A. Ribodetti (UCA)

A. Gholami (UT)

D. Auroux (UCA)


V. Dolean(UCA)

P. Jolivet (UTO)

Young researchers

UCA: University Côte d'Azur
UT: University Tehran
UTO: University Toulouse
SU: Sorbonne University
UGA: University Grenoble Alpes
OCA: Observatoire Côte d'Azur
HIT: Harbin Institute of Technology
CAS: Chinese Academy of Sciences

H. Aghamiry (UCA)

S. Beller (UCA)

HPC engineers

A. Miniussi (OCA)

L. Combe (UCA)

P.-H. Tournier (SU)

PhD students

A. Aghazade (UT)

G. Guo (UCA)

F. Mamfoumbi (UCA)

N. Mohamadi (UCA)

Y. Wu (HIT)

X. Zhou (CAS)

WIND members

UNIVERSITÉ CÔTE D'AZUR
SCIENCES SORBONNE UNIVERSITÉ
UNIVERSITÉ TOULOUSE III PAUL SABATIER
University of TEHRAN
Geo AZUR
OBSERVATOIRE DE LA CÔTE D'AZUR
UNIVERSITÉ CÔTE D'AZUR
IRD Institut de Recherche pour le Développement FRANCE
cnrs
LJAD
IRIT
JJL
INP ENSEEIHT

WIND sponsors

Chevron
BR PETROBRAS
TOTAL
Shell

Outline

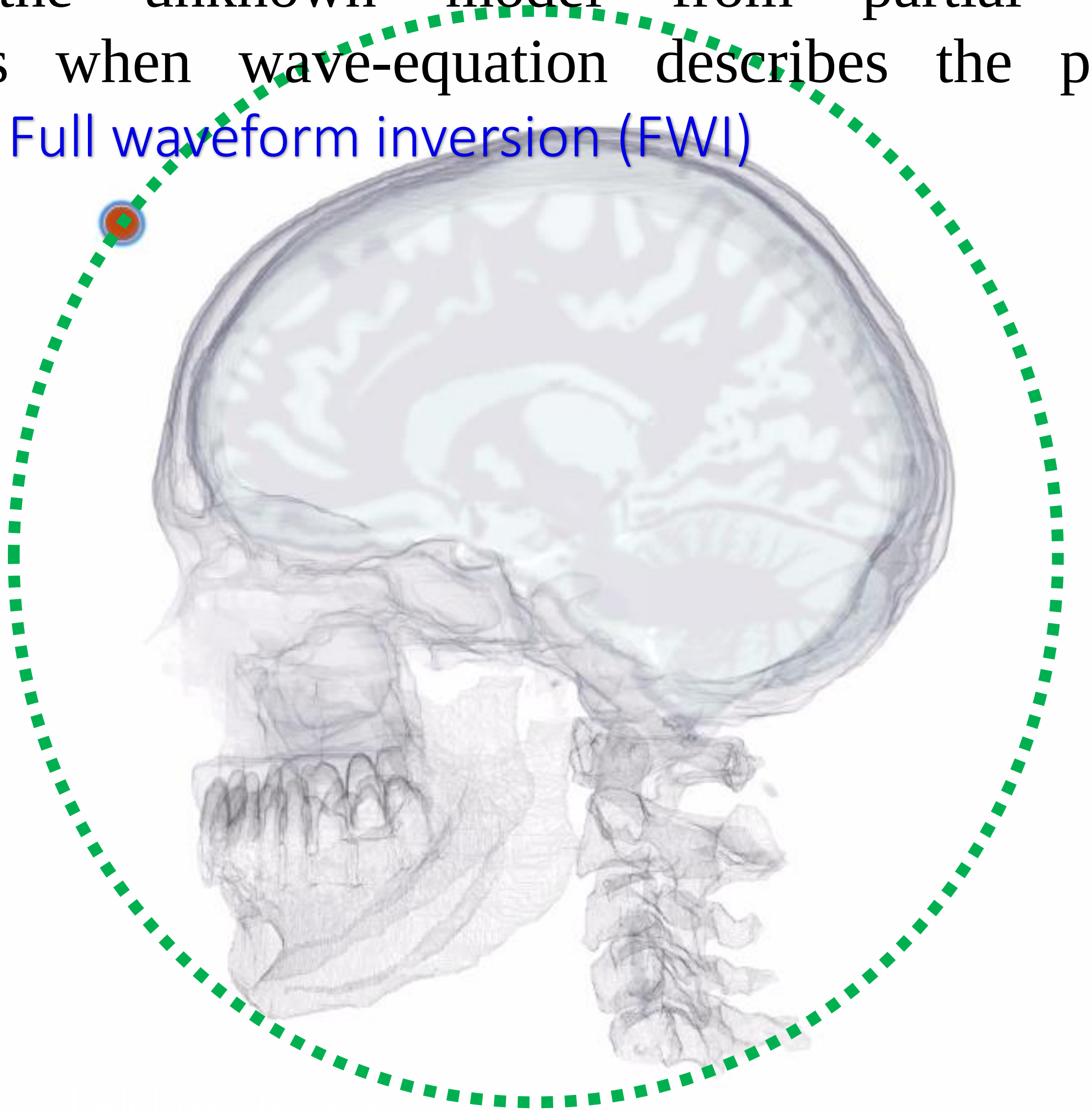
1. Full waveform inversion (FWI) and its challenges
2. Some FWI solution methods
   - Method of Lagrange multipliers
   - Penalty method
   - Method of multipliers (Augmented Lagrangian)
3. Localized FWI or target-oriented FWI

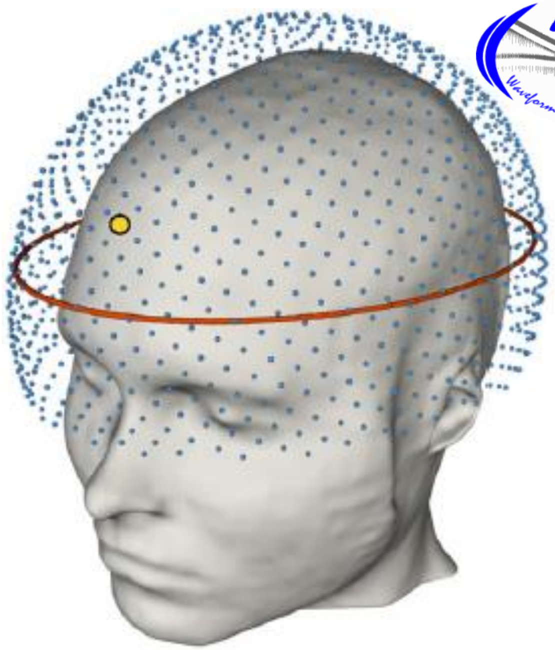
Problem statement: a medical imaging example

Identifying the unknown model from partial wavefield measurements when wave-equation describes the propagated wavefields ➲ Full waveform inversion (FWI)

Problem statement: a medical imaging example

Identifying the unknown model from partial wavefield measurements when wave-equation describes the propagated wavefields ➲ Full waveform inversion (FWI)



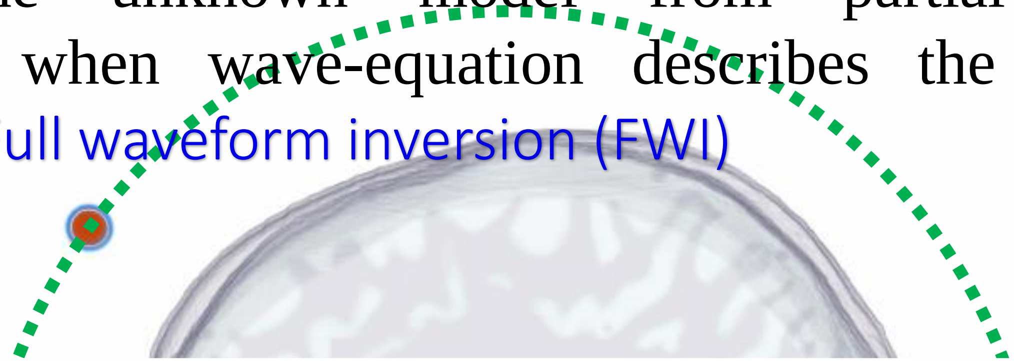


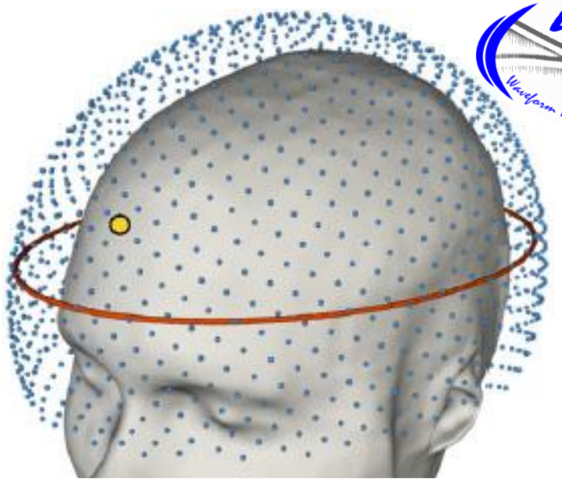
The array of transducers (sensors) used for data generation and recording.

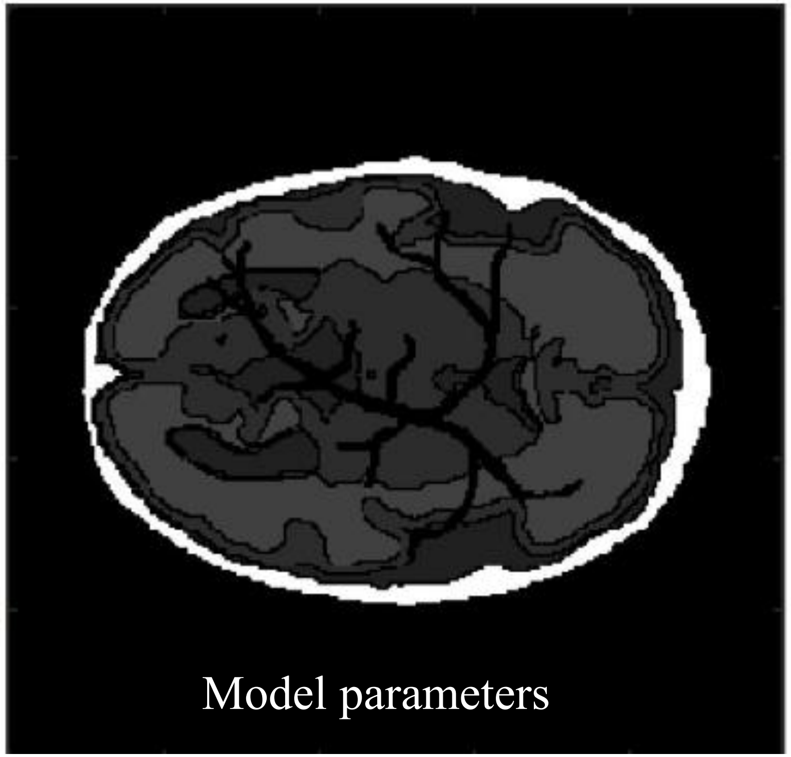
Adapted from Guasch, L., Agudo, O. C., Tang, M. X., Nachev, P., & Warner, M. (2020). Full-waveform inversion imaging of the human brain.

Problem statement: a medical imaging example

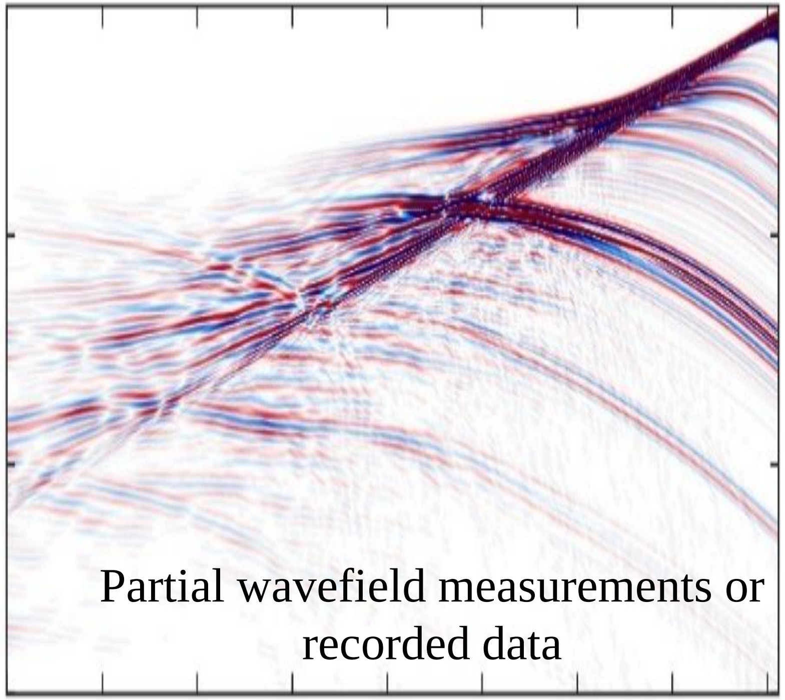
Identifying the unknown model from partial wavefield measurements when wave-equation describes the propagated wavefields ➲ Full waveform inversion (FWI)







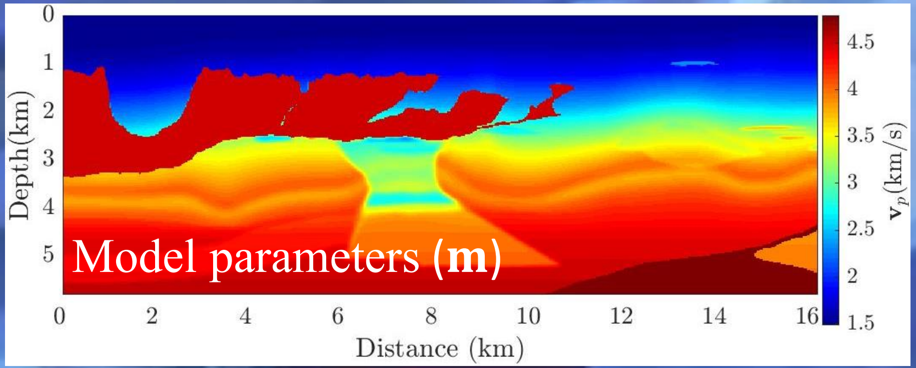
Model parameters



Partial wavefield measurements or recorded data

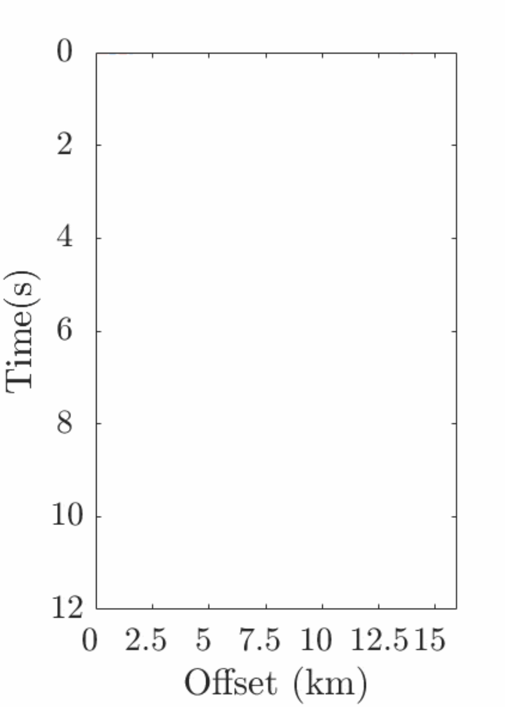
Adapted from Guasch, L., Agudo, O. C., Tang, M. X., Nachev, P., & Warner, M. (2020). Full-waveform inversion imaging of the human brain.

Problem statement: an exploration seismology example



Model parameters ($\mathbf{m}$)

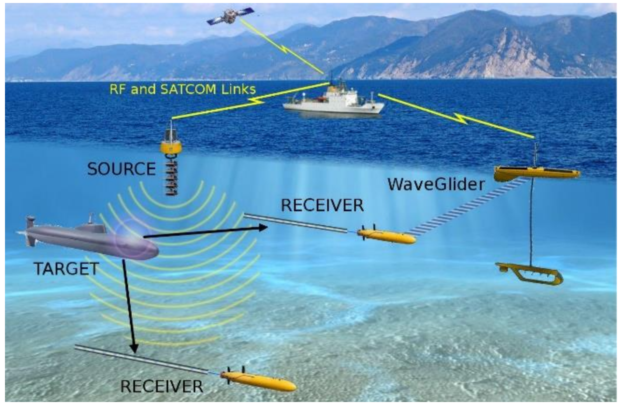
FWI



Recorded data ($\mathbf{d}$)

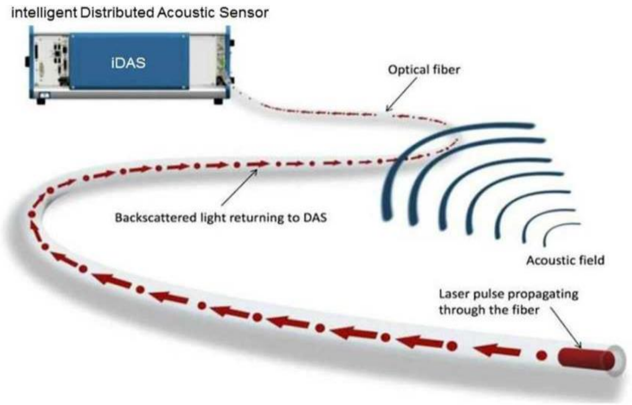
Applications of FWI (1/2)

FWI offers an important method for modelling and remote sensing from sparse measurements in different fields of applications.


RF and SATCOM Links
SOURCE
WaveGlider
RECEIVER
TARGET
RECEIVER

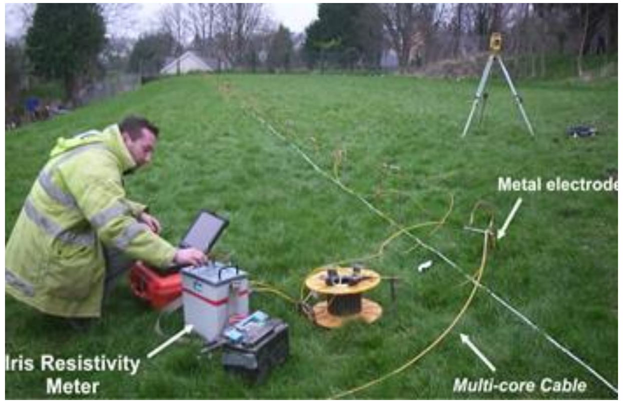
Oceanography (Understanding of the ocean turbulence phenomenon)

Nondestructive testing (Corrosion monitoring)

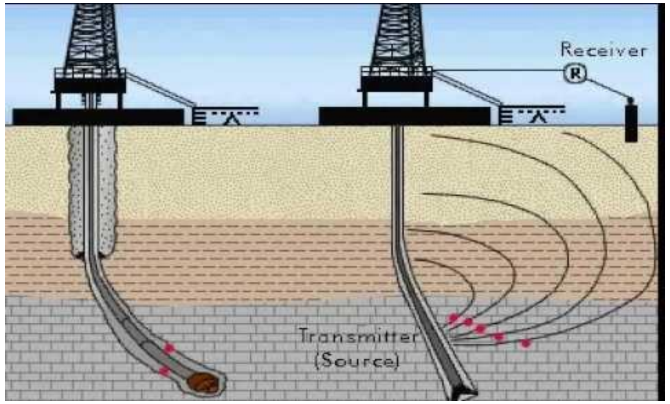

intelligent Distributed Acoustic Sensor
iDAS
Optical fiber
Backscattered light returning to DAS
Acoustic field
Laser pulse propagating through the fiber

Distributed Acoustic Sensing

Geotechnical investigation


Metal electrod
Iris Resistivity Meter
Multi-core Cable

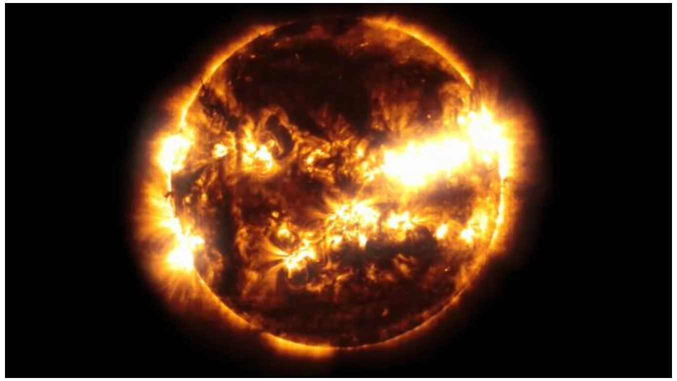
Electrical Resistivity Tomography (ERT)


Receiver
R
Transmitter (Source)

Seismic while drilling

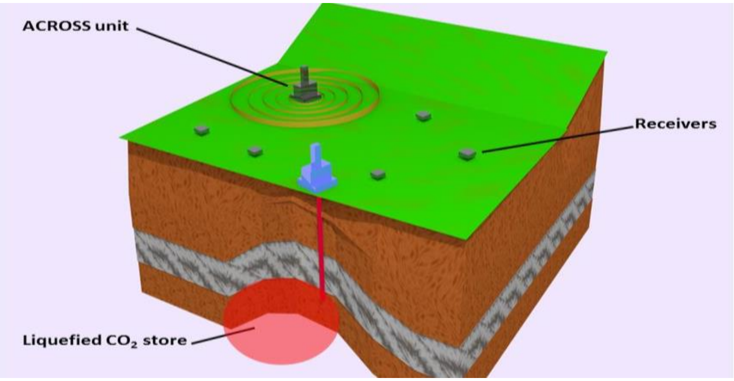
Applications of FWI (2/2)

FWI offers an important method for modelling and remote sensing from sparse measurements in different fields of applications.

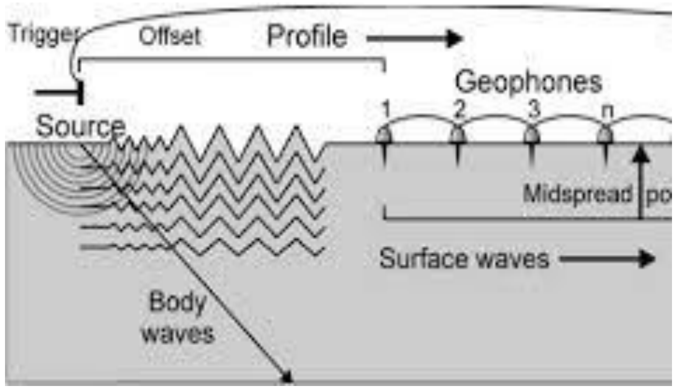


Helioseismology (study of the structure and dynamics of the sun)

Glaciology (the study of ice)


ACROSS unit
Receivers
Liquefied CO_2 store

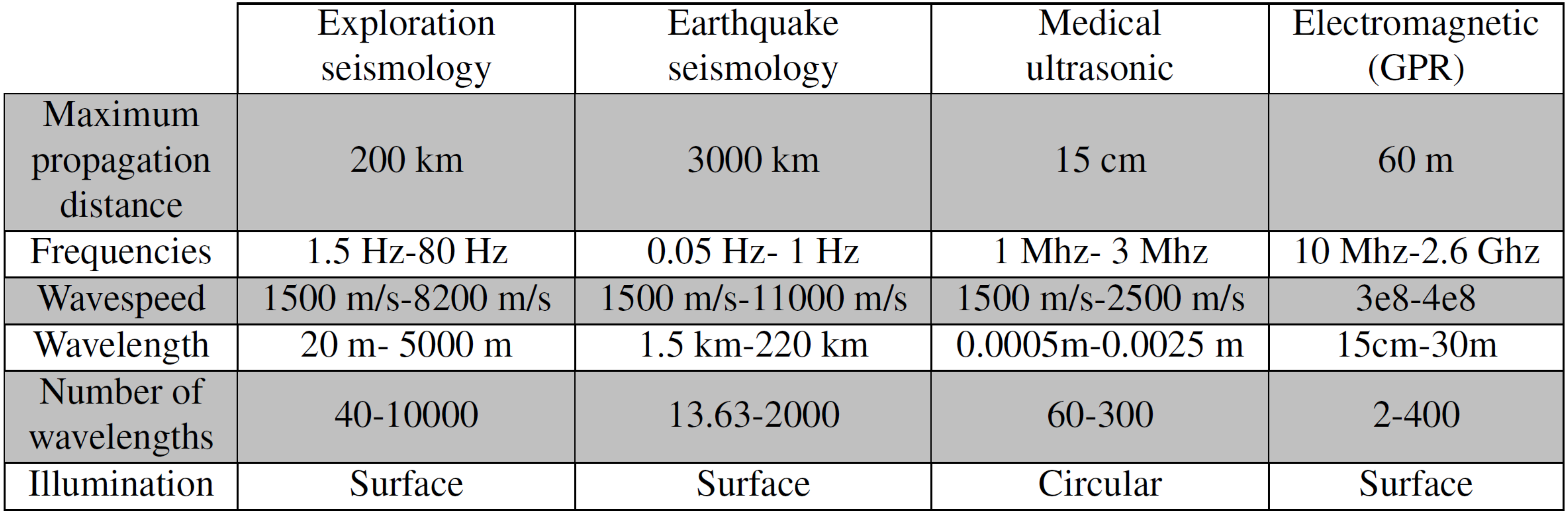
CO_2 monitoring


Trigger
Offset
Profile
Geophones
Source
1
2
3
n
Midspread po
Surface waves
Body waves

Surface waves inversion

Applications of FWI (2/2)

FWI offers an important method for modelling and remote sensing from sparse measurements in different fields of applications.

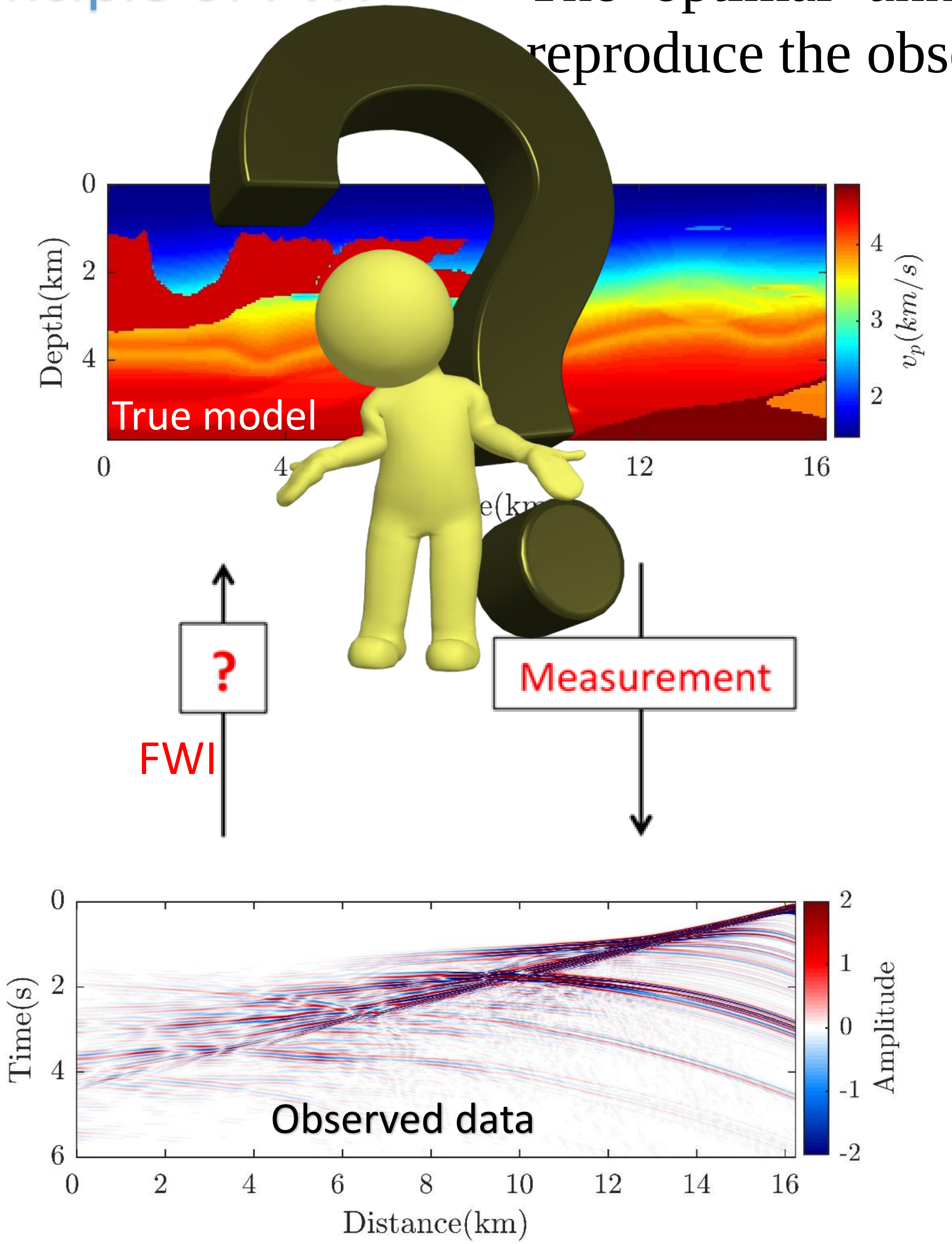


	Exploration seismology	Earthquake seismology	Medical ultrasonic	Electromagnetic (GPR)
Maximum propagation distance	200 km	3000 km	15 cm	60 m
Frequencies	1.5 Hz-80 Hz	0.05 Hz- 1 Hz	1 Mhz- 3 Mhz	10 Mhz-2.6 Ghz
Wavespeed	1500 m/s-8200 m/s	1500 m/s-11000 m/s	1500 m/s-2500 m/s	3e8-4e8
Wavelength	20 m- 5000 m	1.5 km-220 km	0.0005m-0.0025 m	15cm-30m
Number of wavelengths	40-10000	13.63-2000	60-300	2-400
Illumination	Surface	Surface	Circular	Surface

Table 1: A comparison of kinematic scaling parameters in Seismology, ultrasonic medical imaging and electromagnetic imaging. Inspired from [Pratt, 2018].

Principle of FWI

The optimal unknown model is the one that can reproduce the observed data.


0
2
4
Depth(km)
0
4
12
16
e(km
4
3
2
$v_p(km/s)$
0
2
4
6
Time(s)
0
2
4
6
8
10
12
14
16
Distance(km)
2
1
0
-1
-2
Amplitude

True model

?

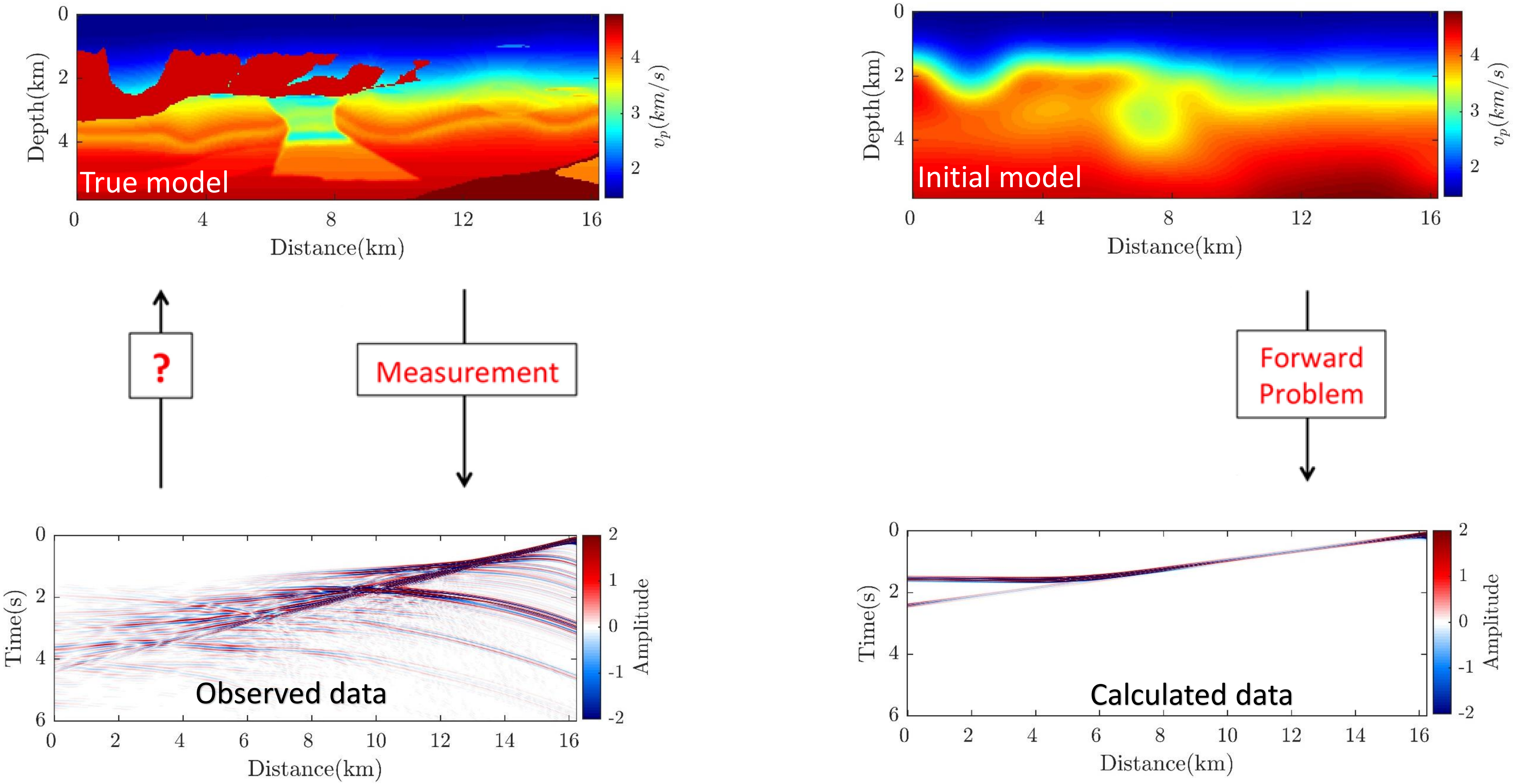
FWI

Measurement

Observed data

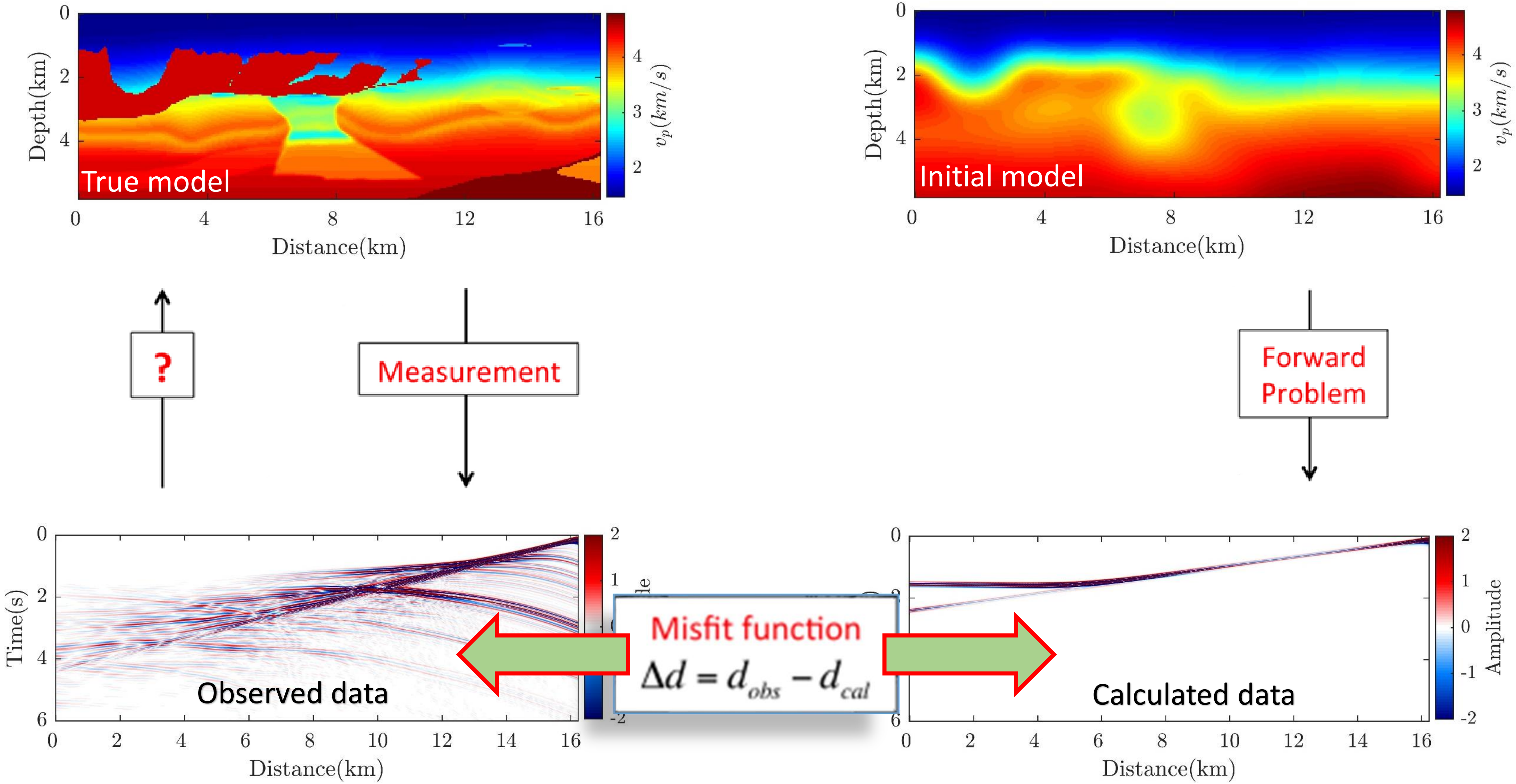
Principle of FWI

The optimal unknown model is the one that can reproduce the observed data.


True model
Depth(km)
Distance(km)
$v_p(km/s)$
0
2
4
0
4
8
12
16
2
3
4
Initial model
Depth(km)
Distance(km)
$v_p(km/s)$
?
Measurement
Forward Problem
Observed data
Time(s)
Distance(km)
Amplitude
0
2
4
6
0
2
4
6
8
10
12
14
16
2
1
0
-1
-2
Calculated data
Time(s)
Distance(km)
Amplitude

Principle of FWI

The optimal unknown model is the one that can reproduce the observed data.



True model
Depth(km)
Distance(km)
$v_p(km/s)$
0
2
4
4
8
12
16
3
Initial model
Observed data
Time(s)
6
10
14
1
-1
-2
Calculated data
Amplitude

?

Measurement

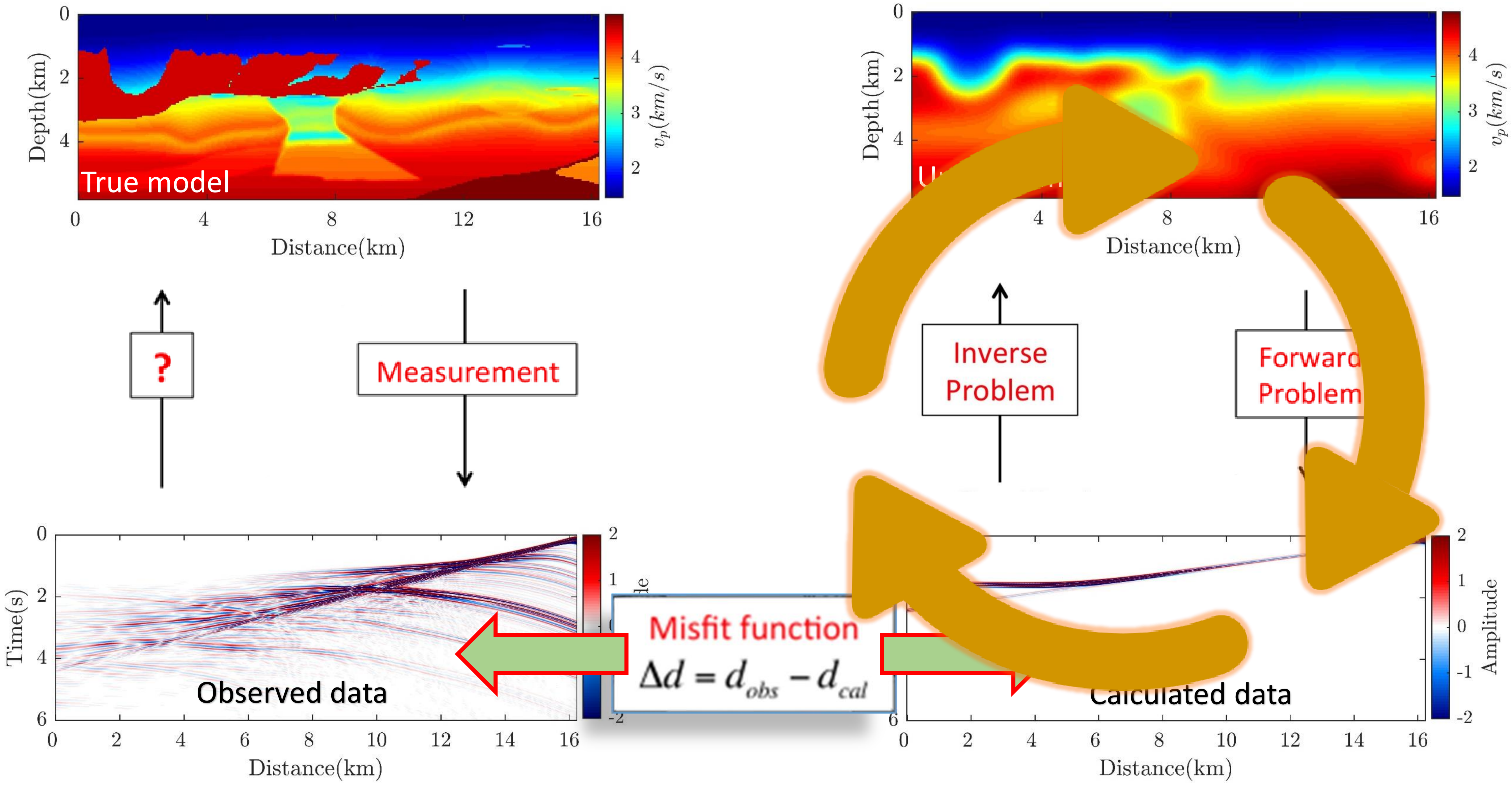
Forward Problem

Misfit function

$$\Delta d = d_{obs} - d_{cal}$$

Principle of FWI

The optimal unknown model is the one that can reproduce the observed data.

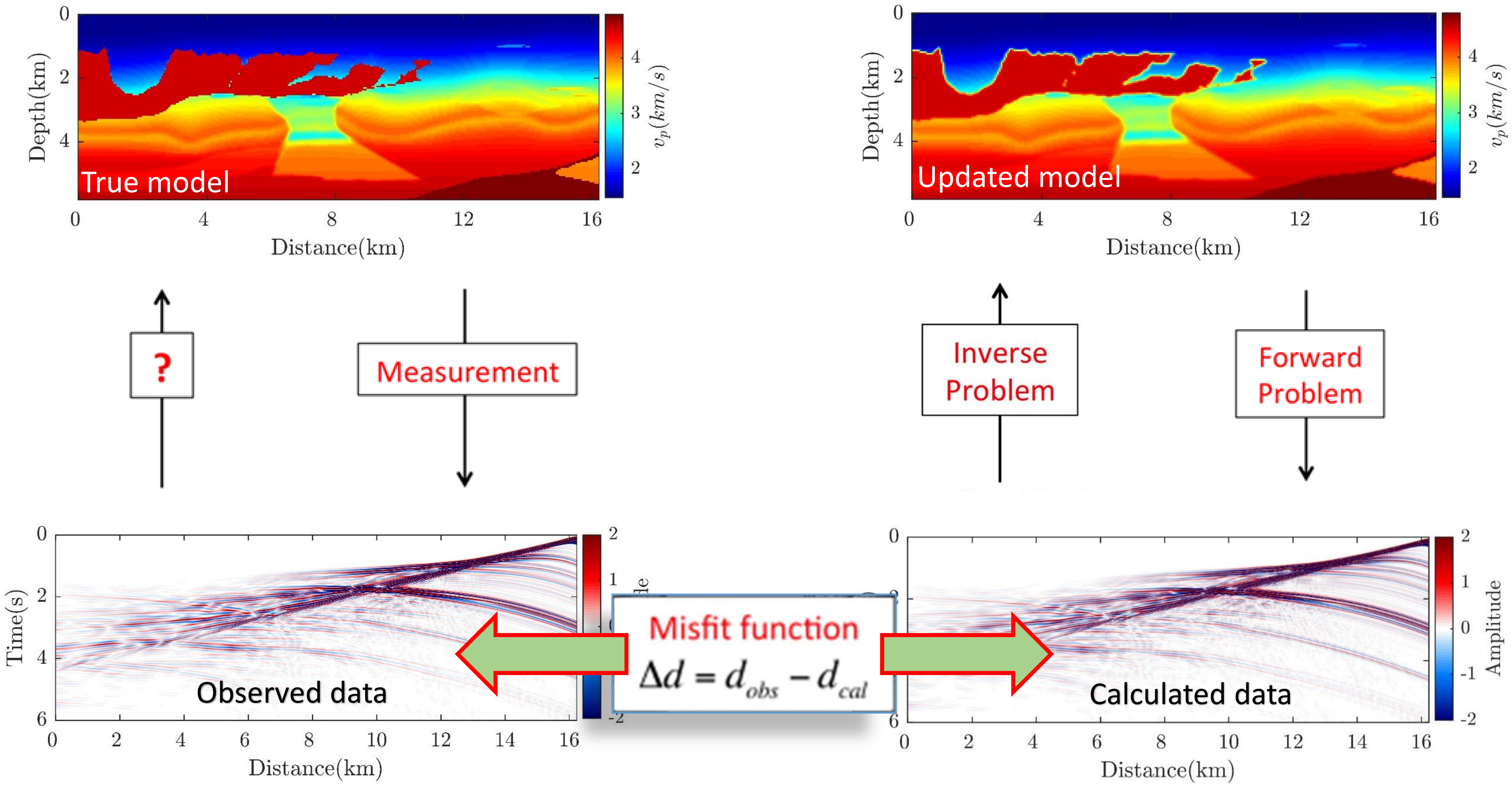

True model
?
Measurement
Inverse Problem
Forward Problem
Observed data
Calculated data
Depth(km)
Distance(km)
$v_p(km/s)$
Time(s)
Amplitude
0
2
4
8
12
16

Misfit function

$$\Delta d = d_{obs} - d_{cal}$$

Principle of FWI

The optimal unknown model is the one that can reproduce the observed data.



True model

Updated model

?

Measurement

Inverse Problem

Forward Problem

Misfit function

$$\Delta d = d_{obs} - d_{cal}$$

Observed data

Calculated data

Depth(km)
0
2
4
Distance(km)
0
4
8
12
16
$v_p(km/s)$
2
3
4
Time(s)
0
2
4
6
Distance(km)
0
2
4
6
8
10
12
14
16
Amplitude
-2
-1
0
1
2

FWI from mathematical point of view (Tarantola, 1984)

FWI is a non-linear (bi-linear) ill-posed PDE-constrained optimization problem:

$$\min_{\mathbf{m},\mathbf{u}} \underbrace{\|\mathbf{P}\mathbf{u} - \mathbf{d}\|_2^2}_{\text{Observation equation}} + \underbrace{\boldsymbol{\varphi}(\mathbf{m})}_{\text{Regularization function}} \quad \textit{Subject to} \quad \underbrace{\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}}_{\text{Wave equation (bilinear term)}}$$

$$\begin{cases} \mathbf{m} \in \mathbb{R}^{N\times 1} \rightarrow \text{Model parameters} \\ \mathbf{u} \in \mathbb{C}^{N\times 1} \rightarrow \text{Wavefield} \\ \mathbf{b} \in \mathbb{C}^{N\times 1} \rightarrow \text{Source} \\ \mathbf{d} \in \mathbb{C}^{n_r\times 1} \rightarrow \text{Recorded data} \end{cases}$$

Some remarks:

- ✓ This equation is quite general, and it is the typical form of imaging methods.
- ✓ Most of the time, **b** and **d** are approximately known and we try to find **m**.
- ✓ Sometimes only **d** is approximately known, and we try to find **b** and **m**.

FWI from mathematical point of view (Tarantola, 1984)

FWI is a non-linear (bi-linear) ill-posed PDE-constrained optimization problem:

$$\min_{\mathbf{m},\mathbf{u}} \underbrace{\|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2}_{\text{Observation equation}} + \underbrace{\varphi(\mathbf{m})}_{\text{Regularization function}} \quad Subject\ to \quad \underbrace{\mathbf{A}(\mathbf{m})\mathbf{u}=\mathbf{b}}_{\text{Wave equation (bilinear term)}}$$

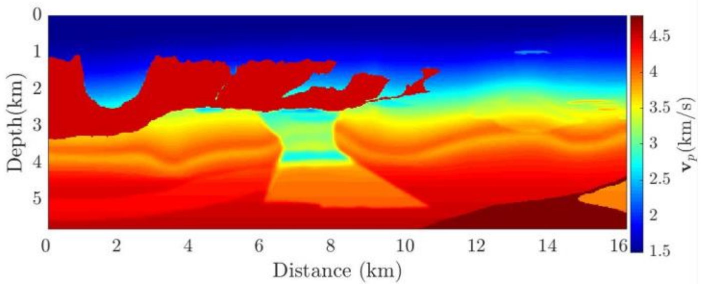
$$\begin{cases} \mathbf{m} \in \mathbb{R}^{N\times 1} \rightarrow \text{Model parameters} \\ \mathbf{u} \in \mathbb{C}^{N\times 1} \rightarrow \text{Wavefield} \\ \mathbf{b} \in \mathbb{C}^{N\times 1} \rightarrow \text{Source} \\ \mathbf{d} \in \mathbb{C}^{n_r\times 1} \rightarrow \text{Recorded data} \end{cases}$$

$$\begin{cases} \Delta \rightarrow \text{Laplacian} \\ \omega \rightarrow \text{Frequency} \\ diag(\blacksquare) \rightarrow \text{a diagonal matrix} \end{cases}$$

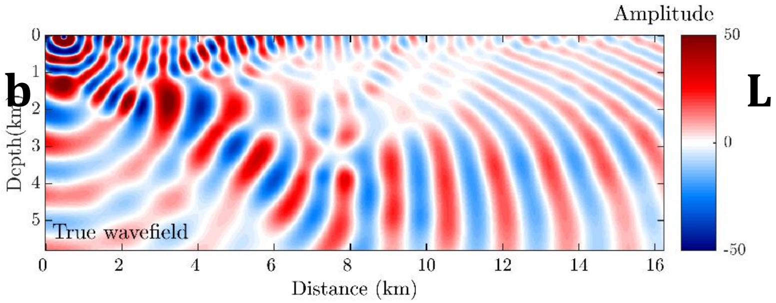
$$\mathbf{A}(\mathbf{m}) = \Delta + \omega^2 diag(\mathbf{m})$$

$$\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b} \rightarrow \Delta\mathbf{u} + \overbrace{\omega^2 diag(\mathbf{m})\mathbf{u}}^{\text{Bilinear}} = \mathbf{b} \rightarrow \overbrace{\omega^2 diag(\mathbf{u})}^{\mathbf{L}(\mathbf{u})}\mathbf{m} = \overbrace{\mathbf{b}-\Delta\mathbf{u}}^{\mathbf{y}}$$

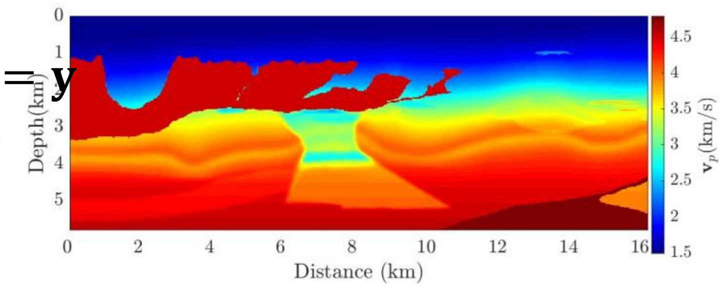
With fixed $\mathbf{m}$ → a linear system for $\mathbf{u}$, and, with fixed $\mathbf{u}$ → a linear system for $\mathbf{m}$



$\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}$



$\mathbf{L}(\mathbf{u})\mathbf{m} = \mathbf{y}$



FWI from mathematical point of view (Tarantola, 1984)

FWI is a non-linear (bi-linear) ill-posed PDE-constrained optimization problem:

$$\min_{\mathbf{m},\mathbf{u}} \underbrace{\|\mathbf{P}\mathbf{u} - \mathbf{d}\|_2^2}_{\text{Observation equation}} + \underbrace{\varphi(\mathbf{m})}_{\text{Regularization function}} \quad \textit{Subject to} \quad \underbrace{\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}}_{\text{Wave equation (bilinear term)}}$$

$\mathbf{m} \in \mathbb{R}^{N\times 1} \rightarrow$ Model parameters
$\mathbf{u} \in \mathbb{C}^{N\times 1} \rightarrow$ Wavefield
$\mathbf{b} \in \mathbb{C}^{N\times 1} \rightarrow$ Source
$\mathbf{d} \in \mathbb{C}^{n_r\times 1} \rightarrow$ Recorded data

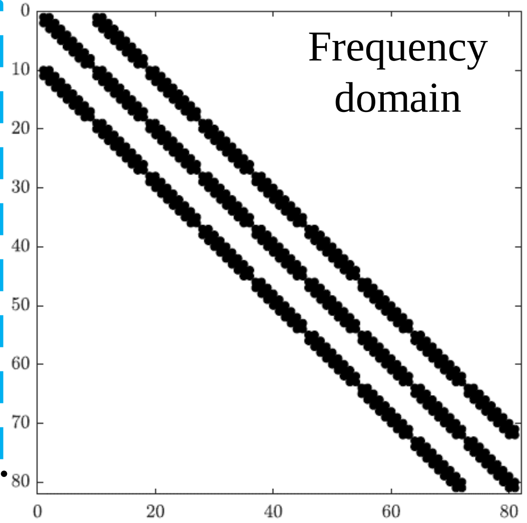
$\Delta \rightarrow$ Laplacian
$\omega \rightarrow$ Frequency
$diag(\blacksquare) \rightarrow$ a diagonal matrix

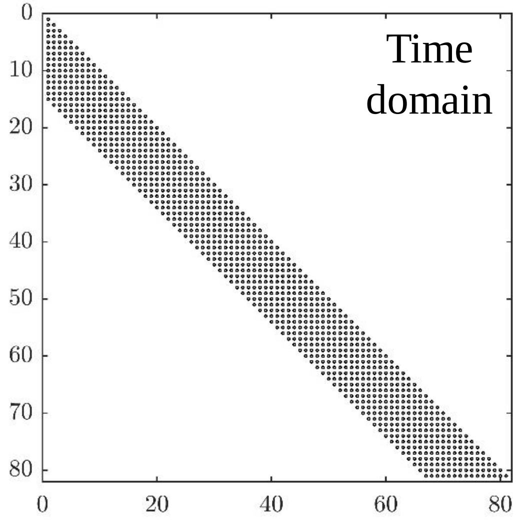
$$\mathbf{A}(\mathbf{m}) = \Delta + \omega^2 diag(\mathbf{m})$$

$$\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b} \rightarrow \Delta\mathbf{u} + \overbrace{\omega^2 diag(\mathbf{m})\mathbf{u}}^{\text{Bilinear}} = \mathbf{b} \rightarrow \overbrace{\omega^2 diag(\mathbf{u})}^{\mathbf{L}(\mathbf{u})}\mathbf{m} = \overbrace{\mathbf{b} - \Delta\mathbf{u}}^{\mathbf{y}}$$

With fixed $\mathbf{m}$ → a linear system for $\mathbf{u}$, and, with fixed $\mathbf{u}$ → a linear system for $\mathbf{m}$

- ✓ It can be solved in time or frequency domain.
- ✓ In time-domain, the explicit time-stepping (matrix-free) methods are used.
- ✓ It is separable in frequency domain → Solvable for each frequency separately.





FWI from mathematical point of view (Tarantola, 1984)

FWI is a non-linear (bi-linear) ill-posed PDE-constrained optimization problem:

Pitfalls and challenges

Nonlinearity → Non convexity and local minima (cycle skipping)

Ill-posedness → Incomplete illumination → Parameter cross-talk

Systematic errors → Approximate physics in forward problem

HPC issue → Computational burden

Subject to $\underbrace{A(m)u = b}_{\text{Wave equation (bilinear term)}}$

Observation Regularization

$m \in \mathbb{R}^{N\times 1} \rightarrow$ Model parameters

$u \in \mathbb{C}^{N\times 1} \rightarrow$ Wavefield

$b \in \mathbb{C}^{N\times 1} \rightarrow$ Source

$d \in \mathbb{C}^{n_r\times 1} \rightarrow$ Recorded data

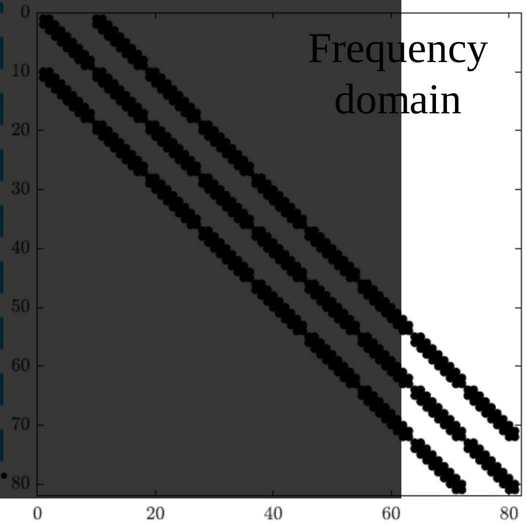
$A(m)u = b \rightarrow \Delta u + \overbrace{\omega^2 diag(m)u}^{\text{Bilinear}} = b \rightarrow$

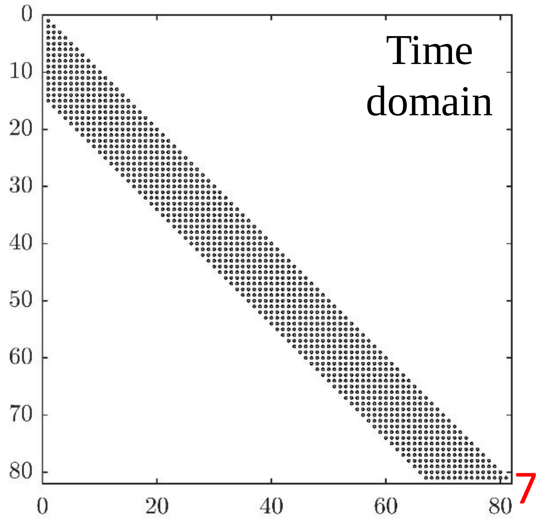
$$\overbrace{\omega^2 diag(u)}^{L(u)}m = \overbrace{b - \Delta u}^{y}$$

With fixed m → a linear system for u, and, with fixed u → a linear system for m

time-stepping (matrix-free) methods are used.

Solvable for each



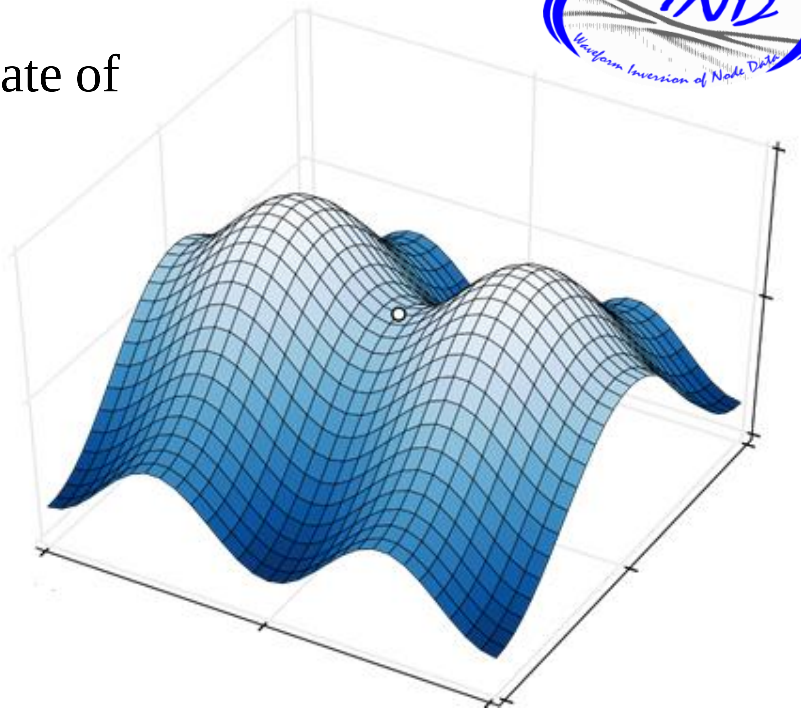


Outline

1. Full waveform inversion (FWI) and its challenges
2. Some FWI solution methods
3. Localized FWI or target-oriented FWI

How to solve FWI:

FWI is a parameter identification problem for PDE which requires the joint update of parameter and the state variable.



$$\min_{\mathbf{m},\mathbf{u}} \|\mathbf{P}\mathbf{u} - \mathbf{d}\|_2^2 + \varphi(\mathbf{m}) \quad \textit{Subject to} \quad \mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}$$

Some solution methods:

1- Method of Lagrange multipliers

2- Penalty method

3- Method of multipliers (Augmented Lagrangian)

1- Method of Lagrange multipliers

Define the Lagrangian function + Determine the partial derivatives (KKT conditions) and set them equal to zero.

Lagrangian →
(Nocedal & Wright, 2006)

$$\mathcal{L}(\mathbf{m},\mathbf{u},\mathbf{v}) = \|\mathbf{Pu} - \mathbf{d}\|_2^2 + \varphi(\mathbf{m}) + \mathbf{v}^{\mathrm{T}}\,[\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{b}]$$

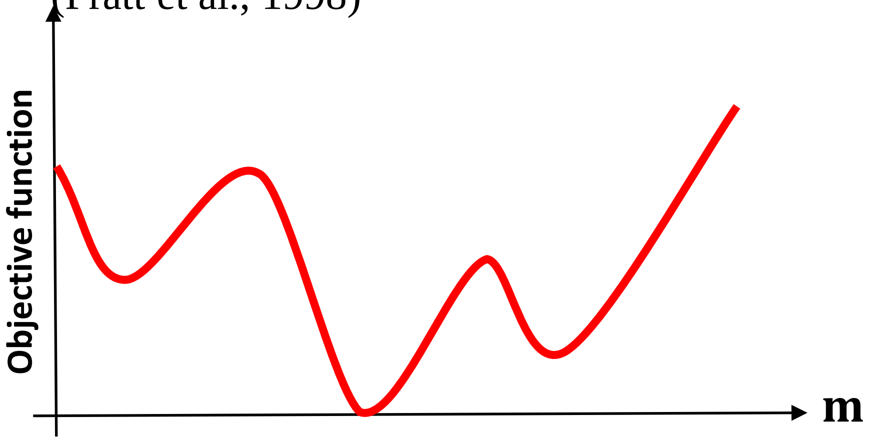
$\mathbf{v}$→ Lagrange multipliers

KKT conditions →

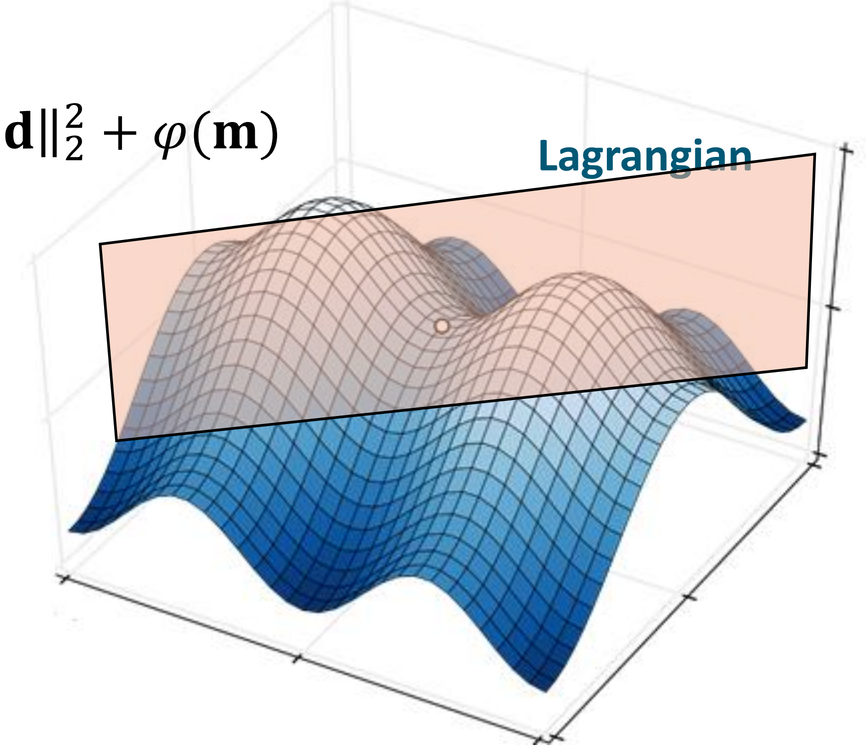
$$\frac{\delta\mathcal{L}(\mathbf{m},\mathbf{u},\mathbf{v})}{\delta\mathbf{m}} = 0 \qquad \frac{\delta\mathcal{L}(\mathbf{m},\mathbf{u},\mathbf{v})}{\delta\mathbf{u}} = 0 \qquad \frac{\delta\mathcal{L}(\mathbf{m},\mathbf{u},\mathbf{v})}{\delta\mathbf{v}} = 0$$

Full space approach → Solve for **m**, **u** and ~~**v**, jointly~~.
(Haber et al., 2000)

Reduced approach → $\mathbf{u} = \mathbf{A}(\mathbf{m})^{-1}\mathbf{b} \rightarrow \min_{\mathbf{m}} \|\mathbf{PA}(\mathbf{m})^{-1}\mathbf{b} - \mathbf{d}\|_2^2 + \varphi(\mathbf{m})$
(Pratt et al., 1998)

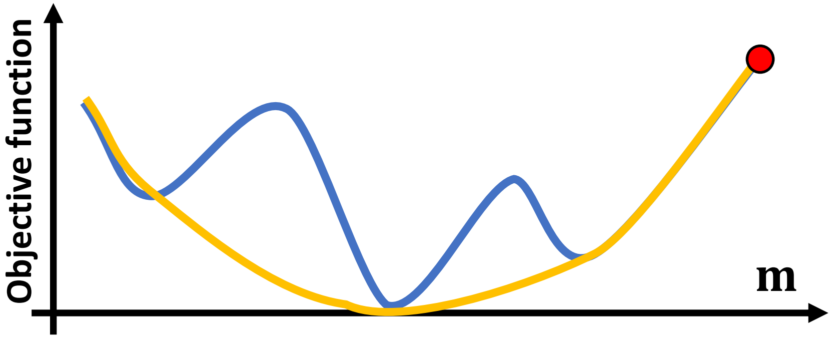


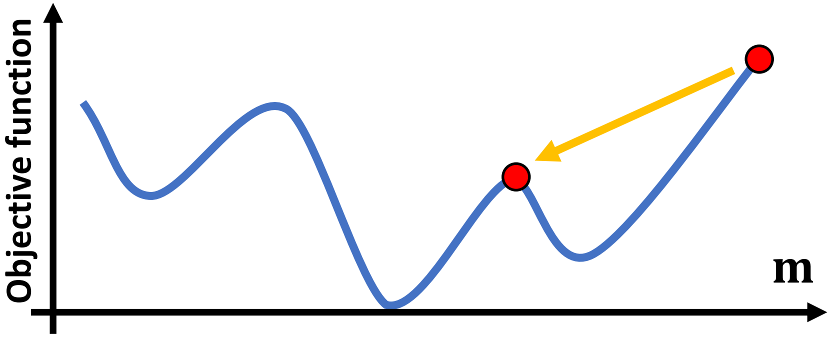
Reduced approach



It makes the objective function more oscillating.

Works done to solve the local minima issue of the reduced formulation





Modifying the cost function:

- ✓ Correlation-based misfit function (Luo & Schuster, 1991).
- ✓ Envelope and instantaneous phase-based misfit function (Fichtner et al., 2008; Luo and Wu, 2015).
- ✓ Dynamic wrapping-based misfit function (Ma and Hale, 2013).
- ✓ Normalized integration-based misfit function (Donno et al., 2013)
- ✓ Optimal transport misfit function (Engquist et al., 2016; Métivier et al., 2106)

Preparing a good initial model:

- ✓ Ray-based tomography (Tavakoli et al., 2017, Sambolian et al., 2019)
- ✓ Migration velocity analysis (MVA) (Symes and Kern, 1994)
- ✓ Reflection wavefield inversion (RWI) (Brossier et al., 2015)
- ✓ Global optimization for FWI (Shaw and Srivastava, 2007, Ely et al., 2015, Datta and Sen, 2016, Sajeva et al., 2016, Galuzzi et al., 2017)

2- Penalty method

(Abubakar et al., 2009, van Leeuwen&Herrmann, 2013)

Transform a constrained optimization problem into a sequence of unconstrained optimization subproblems that are easier to solve. By replacing the hard constraint with a soft constraint, we have:

Penalty form→ $$C(\mathbf{m},\mathbf{u}) = \|\mathbf{P}\mathbf{u} - \mathbf{d}\|_2^2 + \varphi(\mathbf{m}) + \lambda\, \|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{b}\|_2^2$$

λ→ Penalty parameter

Bi-convex optimizations (Gorski et al., 2007)

- ✓ Basically, FWI is a bi-convex optimization problem (Aghamiry et al, 2019).
- ✓ Such problems may have many local minima as generally they are non-convex optimization problems.

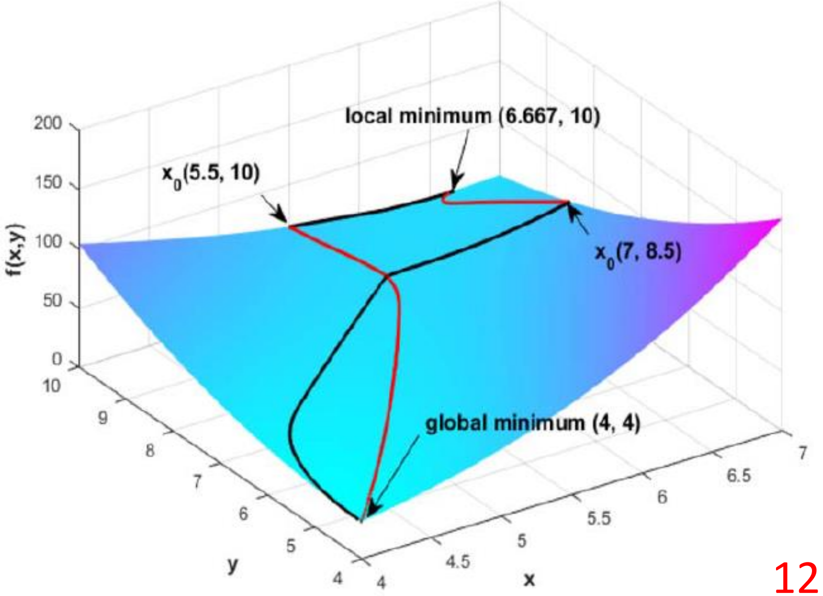
Is it possible to find *the global minimum of a bi-convex optimization*?

Methods for solving biconvex optimization problems

1. General methods that are used for non-convex optimizations.
2. The algorithms which exploit the convex subproblems of a biconvex optimization (Gorski et al., 2007).

2

- ✓ Block relaxation methods, e.g., Alternating Convex set (ACS) (de Leeuw, 1994), Alternating direction method of multipliers (ADMM) (Boyd et al., 2010) are common methods for solving although they don't have a convergence proof.
- ✓ There are some methods, e.g., Global OPtimization algorithm (GOP) (Floudas, 2000), with convergence proof.



2- Penalty method

(Abubakar et al., 2009, van Leeuwen&Herrmann, 2013)

Transform a constrained optimization problem into a sequence of unconstrained optimization subproblems that are easier to solve. By replacing the hard constraint with a soft constraint, we have:

Penalty form→ $C(\mathbf{m},\mathbf{u}) = \|\mathbf{Pu} - \mathbf{d}\|_2^2 + \varphi(\mathbf{m}) + \lambda\, \|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{b}\|_2^2$

λ→ Penalty parameter

Wavefield reconstruction inversion (WRI) (van Leeuwen&Herrmann, 2013)

Solving the penalty form with *Alternate Convex Search (ACS)* (Wendell & Hurter, 1976)

ACS →
1- Keep **m** fixed and solve the convex optimization for **u**.
2- Keep **u** fixed and solve the convex optimization for **m**.
3- Check the stopping criterion if it is satisfied

The order of 1 and 2 can be permuted.

$$\min_{\mathbf{u}} \; C(\mathbf{m}^k,\mathbf{u}) \rightarrow \min_{\mathbf{m}} \; C(\mathbf{m}\,,\mathbf{u}^k) \rightarrow \text{stopping criterion}$$

Superscript k→ mentions to the value of variables $\mathbf{m}^k$ and $\mathbf{u}^k$ at iteration k.

- ✓ Slow convergence.
- ✓ Selecting λ is challenging.
- ✓ WRI, and generally ACS, may fail and stuck in local-minima (Symes, 2020).
- ✓ The Lagrange multipliers are scaled versions of wave-equation residuals, i.e., $\mathbf{v} = \lambda(\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{b})$, which is not efficient.

3- Method of multipliers (Augmented Lagrangian)

The method was studied much in the 1970 and 1980s as a good alternative to penalty methods.

Augmented Lagrangian (AL) → Lagrangian + a penalty term (Nocedal & Wright, 2006)

$$\mathcal{L}_A(\mathbf{m},\mathbf{u},\mathbf{v}) = \|\mathbf{Pu}-\mathbf{d}\|_2^2 + \varphi(\mathbf{m}) + \mathbf{v}^T(\mathbf{A}(\mathbf{m})\mathbf{u}-\mathbf{b}) + \lambda\|\mathbf{A}(\mathbf{m})\mathbf{u}-\mathbf{b}\|_2^2$$

λ → Penalty parameter,
v → Lagrange multipliers or dual variables.

Scaled form of AL → Define scaled form of dual variables as $\tilde{\mathbf{b}} = -{}^{\mathbf{v}}/_{\lambda}$ (Boyd et al, 2010).

$$\mathcal{L}_A(\mathbf{m},\mathbf{u},\tilde{\mathbf{b}}) = \|\mathbf{Pu}-\mathbf{d}\|_2^2 + \varphi(\mathbf{m}) + \lambda\left\|\mathbf{A}(\mathbf{m})\mathbf{u}-\mathbf{b}-\tilde{\mathbf{b}}\right\|_2^2 - \lambda\left\|\tilde{\mathbf{b}}\right\|_2^2$$

We prefer the scaled form of AL because we can have a physical interpretation for dual variables.

ADMM-based FWI (Aghamiry et al., 2018)

ADMM is a recommended tool for biconvex optimizations (Boyd et al., 2010; Brás et al., 2012).

ADMM Solves AL in an alternating mode for primal and dual variables as

$$\underbrace{\min_{\mathbf{u}} \mathcal{L}_A(\mathbf{m}^k,\mathbf{u},\tilde{\mathbf{b}}^k) \rightarrow \min_{\mathbf{m}} \mathcal{L}_A(\mathbf{u}^{k+1},\mathbf{m},\tilde{\mathbf{b}}^k)}_{\text{Primals}} \rightarrow \underbrace{\max_{\tilde{\mathbf{b}}} \mathcal{L}_A(\mathbf{u}^{k+1},\mathbf{m}^{k+1},\tilde{\mathbf{b}})}_{\text{Dual}}$$

It uses partial updates of primal variables (similar to the Gauss–Seidel method for solving linear equations)

How does ADMM-based FWI work?

First subproblem →

$$\mathbf{u}^{k+1} = \underset{\mathbf{u}}{\operatorname{argmin}} \; \|\mathbf{P}\mathbf{u} - \mathbf{d}\|_2^2 + \lambda \left\|\mathbf{A}(\mathbf{m}^k)\mathbf{u} - \mathbf{b} - \tilde{\mathbf{b}}^k\right\|_2^2$$

$$\begin{bmatrix} \mathbf{P} \\ \sqrt{\lambda}\mathbf{A}(\mathbf{m}^k) \end{bmatrix} \mathbf{u}^{k+1} = \begin{bmatrix} \mathbf{d} \\ \sqrt{\lambda}\,[\mathbf{b} + \tilde{\mathbf{b}}^k] \end{bmatrix}$$

①

- ✓ Solve wave equation with a feedback term from data → *data assimilated* wavefield (Auroux & Blum, 2008).
- ✓ Extrapolation problem with a feedback term from physics.

Second subproblem →

$$\mathbf{m}^{k+1} = \underset{\mathbf{m}}{\operatorname{argmin}} \; \varphi(\mathbf{m}) + \lambda \left\|\mathbf{L}(\mathbf{u}^{k+1})\mathbf{m} - \mathbf{y}\right\|_2^2$$

②

- ✓ Push back $\mathbf{u}^{k+1}$ toward the wave equation + satisfy the regularization.

Third subproblem →

$$\tilde{\mathbf{b}}^{k+1} = \tilde{\mathbf{b}}^k + \mathbf{b} - \mathbf{A}(\mathbf{m}^{k+1})\mathbf{u}^{k+1}$$

③

- ✓ Updating the RHS by the running sum of wave-equation error.

Mechanism →

1- Estimate an accurate wavefield (as much as possible). ⇨ n_s forward modeling

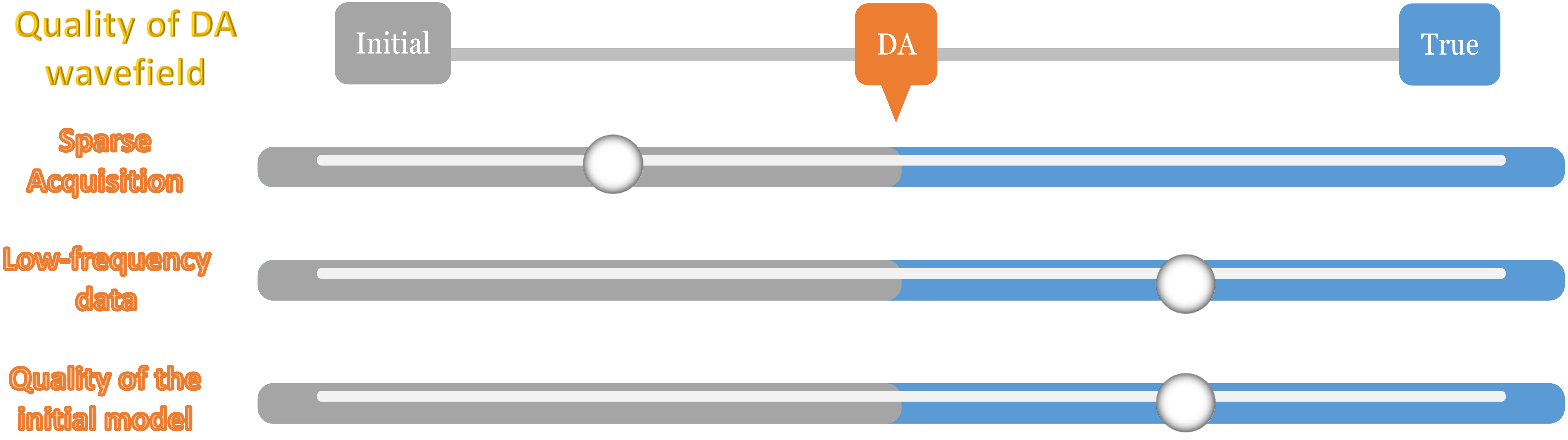
2- Reconstruct the model parameters using this wavefield. ⇨ Negligible

3- Iterative refinement/defect correction. ⇨ Negligible

The challenges of the method

- ✓ We don’t know under which conditions the method works.

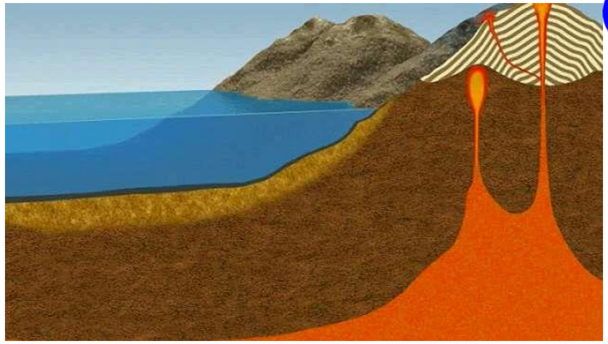
Qualitatively, we know the DA wavefield should be close to the true wavefield near the receivers.

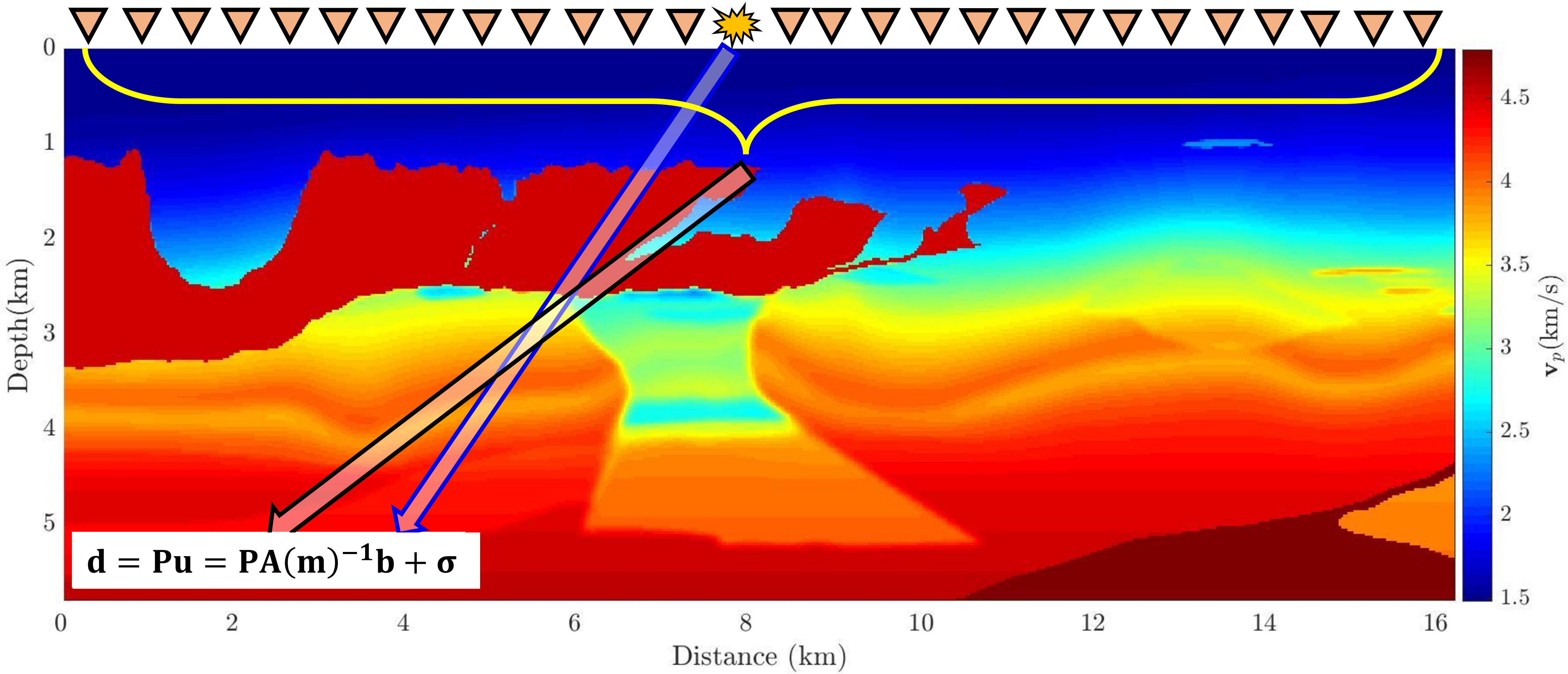

Quality of DA wavefield
Initial
DA
True
Sparse Acquisition
Low-frequency data
Quality of the initial model

- ✓ The dual variables, as well as wavefields, should be stored on the disk.
- ✓ Original ADMM has a linear convergence.
- ✓ Extracting DA wavefield is challenging for large scale (in time and frequency domain formulation).

2004 BP salt model

- ✓ It is representative of the geology of the deep offshore Gulf of Mexico.
- ✓ It has a simple background with a complex rugose multivalued salt body, sub-salt slow velocity anomalies related to over-pressure zones.

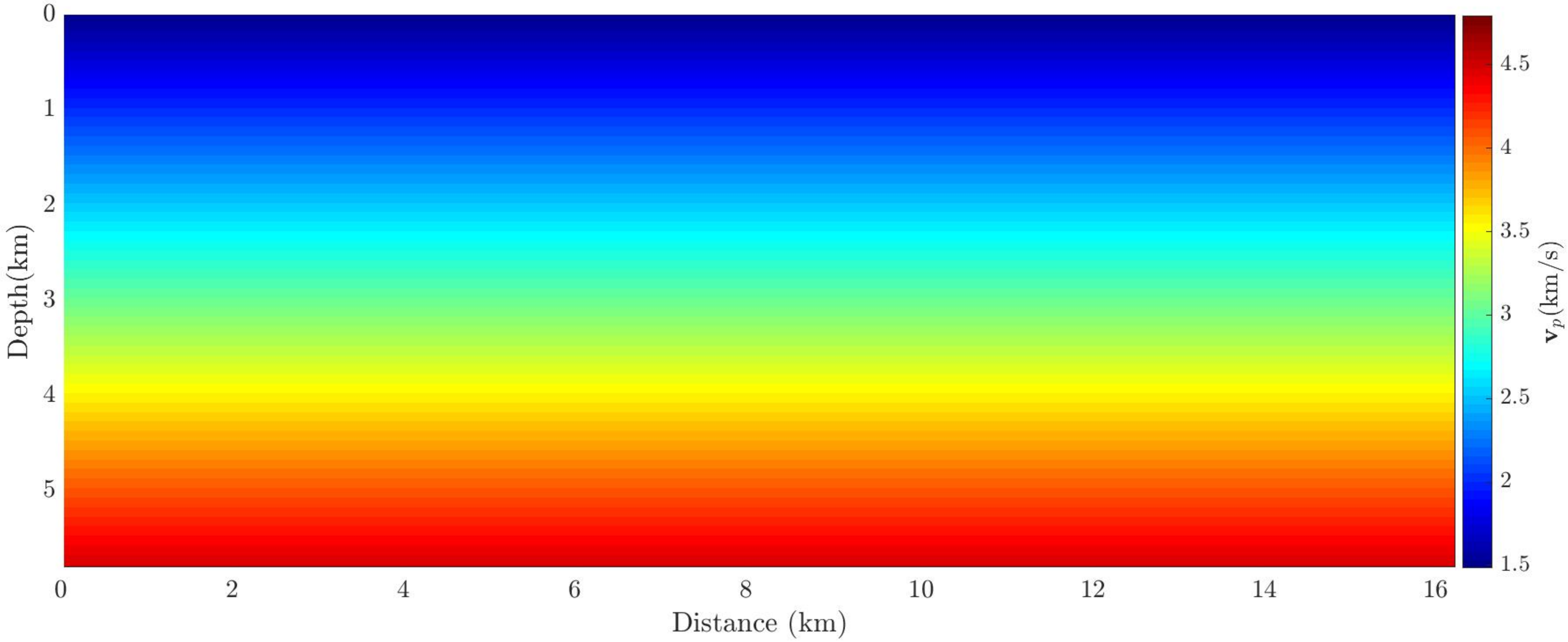




$$\mathbf{d} = \mathbf{Pu} = \mathbf{PA}(\mathbf{m})^{-1}\mathbf{b} + \boldsymbol{\sigma}$$

Setup and initial model

- ✓ Surface acquisition with 162 sources and 650 receivers.
- ✓ A 9-point finite-difference staggered-grid stencil with PML boundary condition and anti-lumped mass is used.
- ✓ Inverted frequencies are 3-13Hz with frequency continuation when batches of 2 frequencies with a 0.5Hz spacing are used and three paths over batches are used.

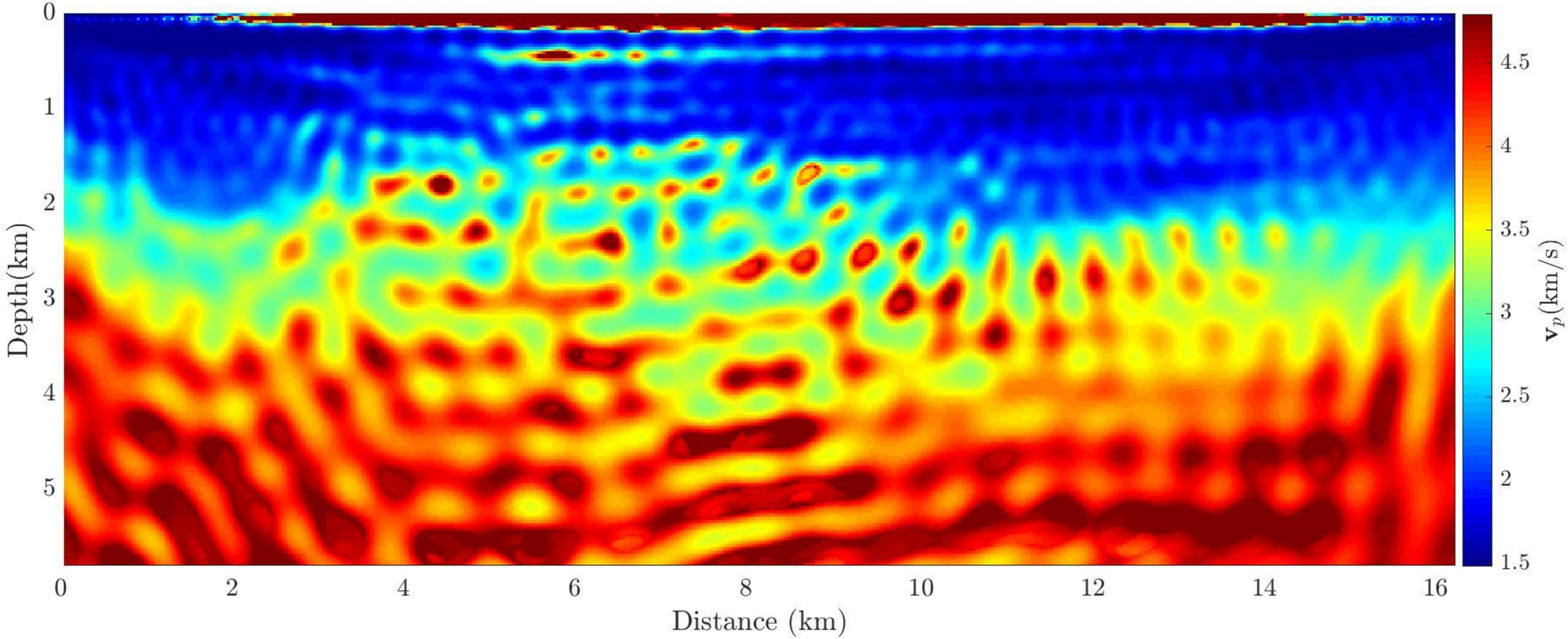


Reduced approach (Pratt et al., 1998)

$$\min_{\mathbf{m}} \|\mathbf{PA}(\mathbf{m})^{-1}\mathbf{b} - \mathbf{d}\|_2^2$$

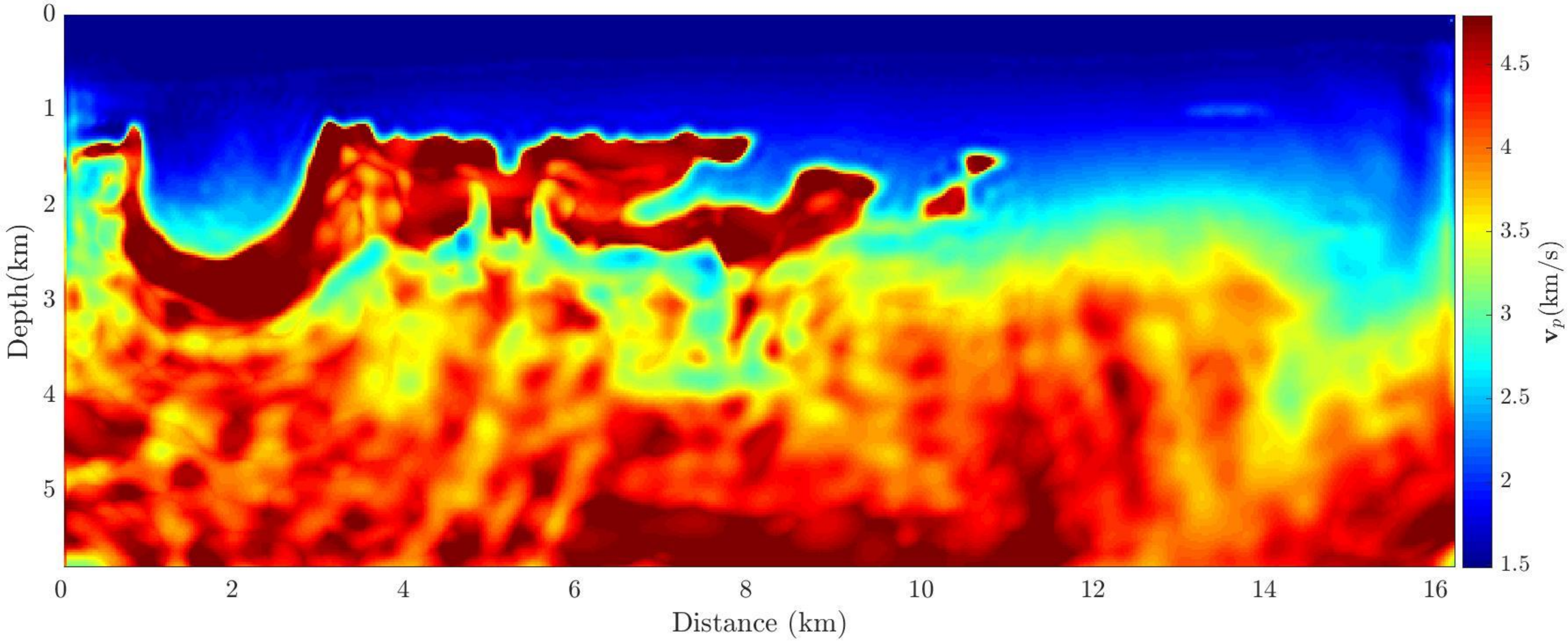
The model parameters are updated with the L-BFGS quasi-Newton optimization and a line search procedure for step length estimation (that satisfies the Wolfe conditions).

The reduced approach FWI is stuck in a local minimum during the inversion of the first batch.



WRI (van Leeuwen&Herrmann, 2013)

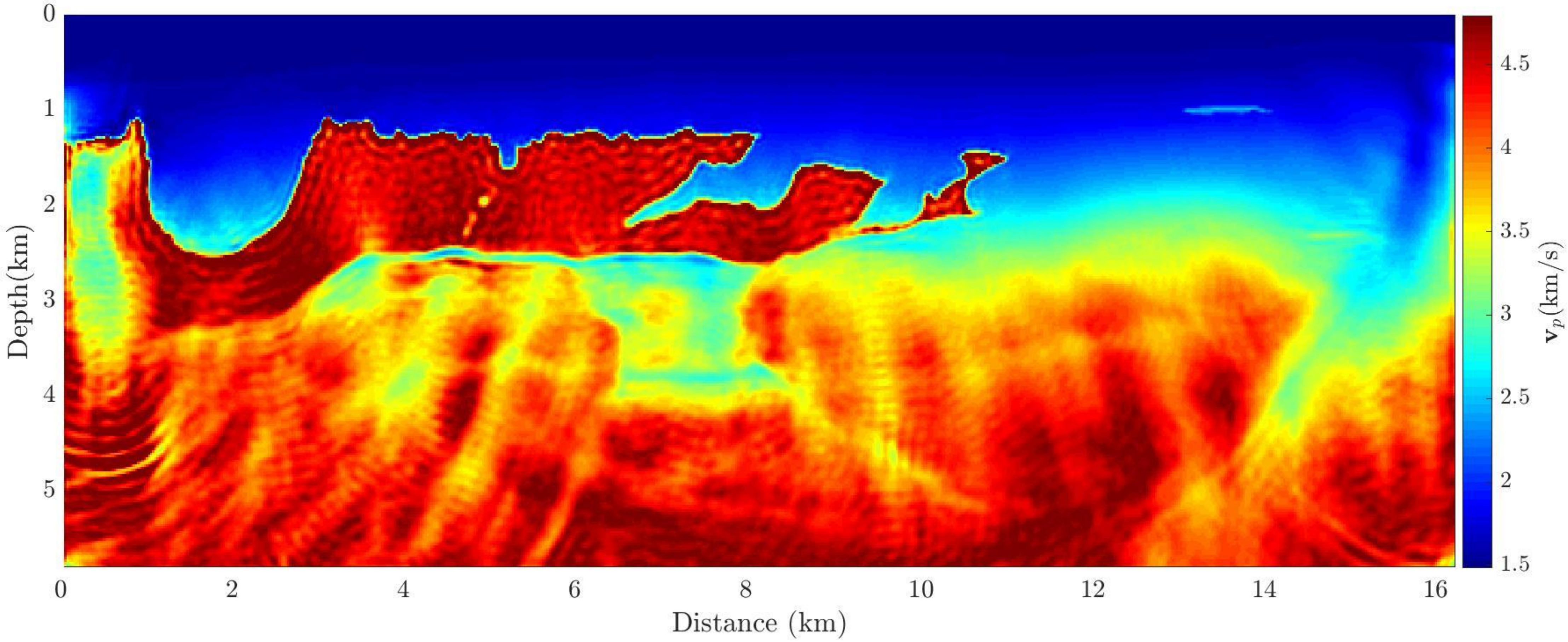
Number of iterations → 561



IR-WRI (Aghamiry et al., 2019a)

Number of iterations → 288 → A faster convergence to more accurate subsurface model.

The extracted model still is not acceptable because of the poor illumination of surface acquisition.



Pitfalls and challenges

Nonlinearity
Non convexity and local minima (cycle skipping)

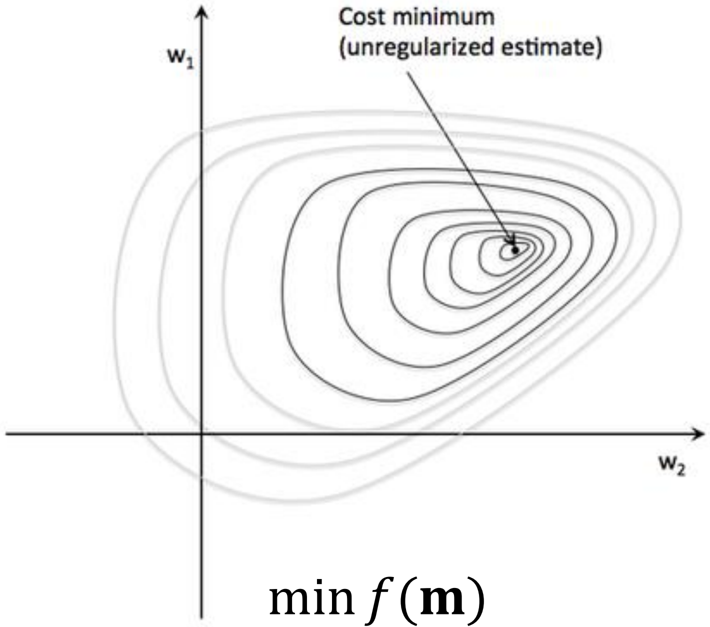
Ill-posedness
Incomplete illumination
Parameter cross-talk

Systematic errors
Approximate physics in forward problem

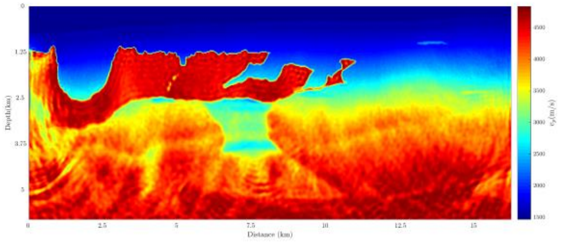
HPC issue
Computational burden

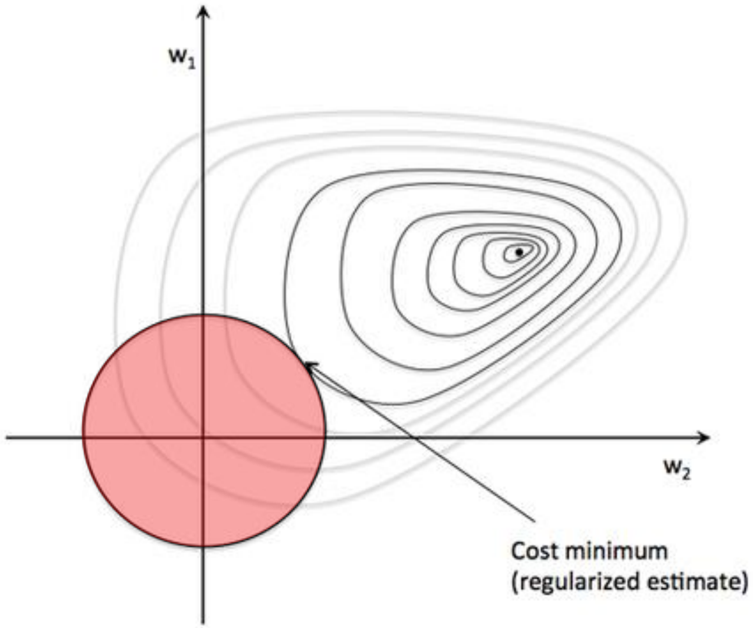
Regularization and adding prior information

- ✓ Because of the insufficient illumination of surface data acquisition, some parts of the model can’t be reconstructed (there are in the null space).
- ✓ Regularization is a process of introducing additional information in order to solve an ill-posed problem or to prevent overfitting.



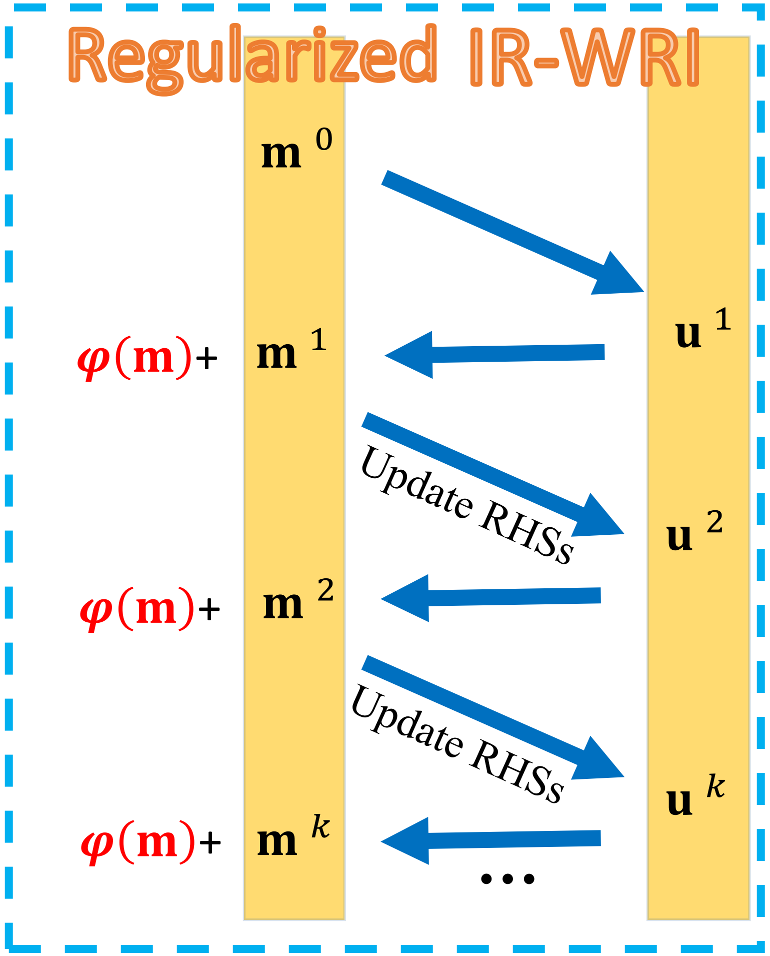
$$\min_{\mathbf{m}} f(\mathbf{m})$$





$$\min_{\mathbf{m}} f(\mathbf{m}) + \boldsymbol{\varphi}(\mathbf{m})$$

Drive the inversion toward priors of $\boldsymbol{\varphi}(\mathbf{m})$



Regularized IR-WRI

$\mathbf{m}^0$

$\mathbf{u}^1$

$\boldsymbol{\varphi}(\mathbf{m})+$ $\mathbf{m}^1$

Update RHSs

$\mathbf{u}^2$

$\boldsymbol{\varphi}(\mathbf{m})+$ $\mathbf{m}^2$

Update RHSs

$\mathbf{u}^k$

$\boldsymbol{\varphi}(\mathbf{m})+$ $\mathbf{m}^k$

…

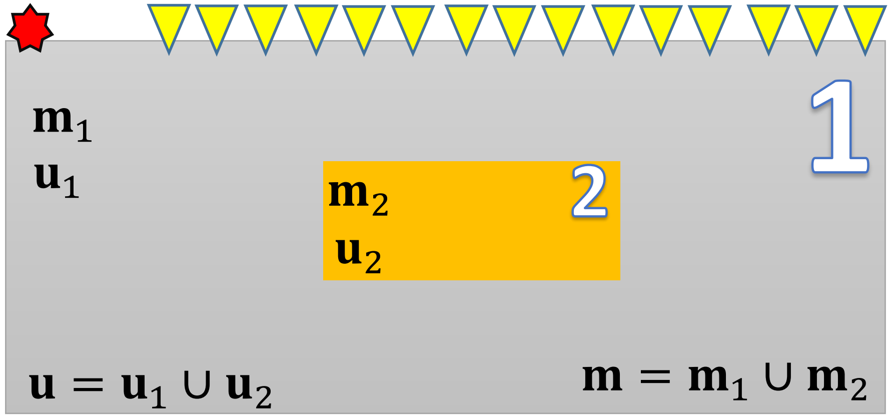
Aghamiry, et al. (2019). Compound regularization of full-waveform inversion for imaging piecewise media. *IEEE TGRS*, *58*(2), 1192-1204.

Outline

1. Full waveform inversion (FWI) and its challenges
2. Some FWI solution methods
3. Localized FWI or target-oriented FWI

Localized FWI

- ✓ FWI is a computationally intensive imaging technique to reconstruct **m** in the domain in which the waves propagate.
- ✓ For some specific applications like CO2 monitoring, reservoir monitoring, geothermal exploitation, and teleseismic imaging we only seek to update **m** within a localized region of interest.



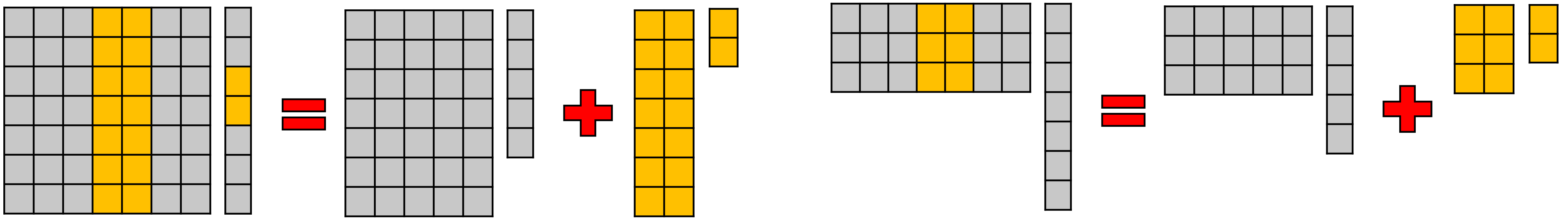
$\mathbf{m}_1$
$\mathbf{u}_1$
1
$\mathbf{m}_2$
$\mathbf{u}_2$
2
$\mathbf{u} = \mathbf{u}_1 \cup \mathbf{u}_2$
$\mathbf{m} = \mathbf{m}_1 \cup \mathbf{m}_2$

Here, we want to update $\mathbf{m}_2$ when a good approximation of $\mathbf{m}_1$ exist.

The conventional methods for localized FWI {

1. Data redatuming and then FWI.
2. FWI using local wave-equation solvers.

I propose to use block decomposition to decompose the equations as



$$\mathbf{A}(\mathbf{m})\ \mathbf{u} = \mathbf{A}_1(\mathbf{m}_1)\ \mathbf{u}_1 + \mathbf{A}_2(\mathbf{m}_2)\ \mathbf{u}_2 \qquad \mathbf{P}\ \mathbf{u} = \mathbf{P}_1\ \mathbf{u}_1 + \mathbf{P}_2\ \mathbf{u}_2$$

Multi-block FWI

FWI for sub-domains 1 and 2 can be written as:

$$\min_{\mathbf{m}_1,\mathbf{m}_2,\mathbf{u}_1,\mathbf{u}_2} \underbrace{\|\mathbf{P}_1\mathbf{u}_1+\mathbf{P}_2\mathbf{u}_2-\mathbf{d}\|_2^2}_{\text{Observation equation}} + \underbrace{\varphi_1(\mathbf{m}_1)}_{\text{Reg. for sub.1}} + \underbrace{\varphi_2(\mathbf{m}_2)}_{\text{Reg. for sub.2}} \quad \textit{subject to} \quad \underbrace{\mathbf{A}_1(\mathbf{m}_1)\mathbf{u}_1+\mathbf{A}_2(\mathbf{m}_2)\mathbf{u}_2=\mathbf{b}}_{\text{Wave equation}}$$

Here we have four primal variables, $\mathbf{m}_1, \mathbf{m}_2, \mathbf{u}_1$ and $\mathbf{u}_2$.

Multi-block ADMM uses the augmented Lagrangian to update primal and dual variables in an alternating mode.

$$\mathcal{L}_A(\mathbf{m}_1,\mathbf{m}_2,\mathbf{u}_1,\mathbf{u}_2)=\varphi_1(\mathbf{m}_1)+\varphi_2(\mathbf{m}_2)+\|\mathbf{P}_1\mathbf{u}_1+\mathbf{P}_2\mathbf{u}_2-\mathbf{d}\|_2^2+\lambda\left\|\mathbf{A}_1(\mathbf{m}_1)\mathbf{u}_1+\mathbf{A}_2(\mathbf{m}_2)\mathbf{u}_2-\mathbf{b}-\tilde{\mathbf{b}}\right\|_2^2-\lambda\left\|\tilde{\mathbf{b}}\right\|_2^2$$

ADMM-based multi-block FWI → Solve AL in an alternating mode for primal and dual variables.

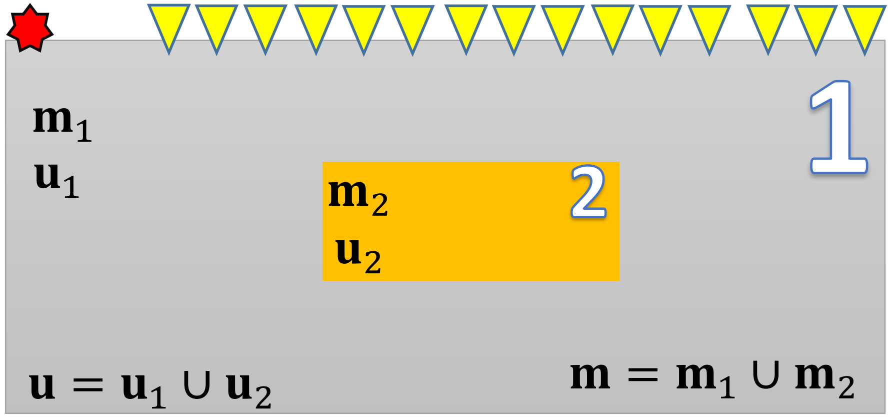
$$\min_{\mathbf{u}_1} \mathcal{L}_A\left(\mathbf{u}_1,\mathbf{u}_2^k,\mathbf{m}_1^k,\mathbf{m}_2^k,\tilde{\mathbf{b}}^k\right)$$

$$\min_{\mathbf{u}_2} \mathcal{L}_A\left(\mathbf{u}_1^{k+1},\mathbf{u}_2,\mathbf{m}_1^k,\mathbf{m}_2^k,\tilde{\mathbf{b}}^k\right)$$

$$\min_{\mathbf{m}_1} \mathcal{L}_A\left(\mathbf{u}_1^{k+1},\mathbf{u}_2^{k+1},\mathbf{m}_1,\mathbf{m}_2^k,\tilde{\mathbf{b}}^k\right)$$

$$\min_{\mathbf{m}_2} \mathcal{L}_A\left(\mathbf{u}_1^{k+1},\mathbf{u}_2^{k+1},\mathbf{m}_1^{k+1},\mathbf{m}_2,\tilde{\mathbf{b}}^k\right)$$

$$\max_{\tilde{\mathbf{b}}} \mathcal{L}_A\left(\mathbf{u}_1^{k+1},\mathbf{u}_2^{k+1},\mathbf{m}_1^{k+1},\mathbf{m}_2^{k+1},\tilde{\mathbf{b}}\right)$$


$\mathbf{m}_1$
$\mathbf{u}_1$
1
$\mathbf{m}_2$
$\mathbf{u}_2$
2
$\mathbf{u}=\mathbf{u}_1\cup\mathbf{u}_2$
$\mathbf{m}=\mathbf{m}_1\cup\mathbf{m}_2$

These subproblems converge to the solution of conventional IR-WRI, but with a slower convergence rate.

An Adaptation of Multi-Block ADMM for Localized FWI

In localized FWI, a good approximation of $\mathbf{m}_1$ is exist and the goal is to update $\mathbf{m}_2$.

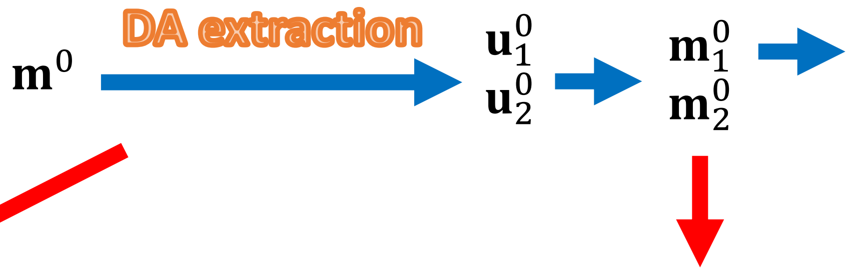
Localized Wavefield Inversion (LWI): Keep $\mathbf{m}_1$ and $\mathbf{u}_1$ as passive variables and solve the rest of subproblems:

① $\min_{\mathbf{u}_2} \mathcal{L}_A(\mathbf{u}_1^0, \mathbf{u}_2, \mathbf{m}_1^0, \mathbf{m}_2^k, \tilde{\mathbf{b}}^k) \rightarrow \mathbf{u}_2^{k+1} = \operatorname{argmin}_{\mathbf{u}_2} \left\|\mathbf{A}(\mathbf{m}_2^k)\mathbf{u}_2 + \mathbf{A}(\mathbf{m}_1^0)\mathbf{u}_1^0 - \mathbf{b} - \tilde{\mathbf{b}}^k\right\|_2^2$

② $\min_{\mathbf{m}_2} \mathcal{L}_A(\mathbf{u}_1^0, \mathbf{u}_2^{k+1}, \mathbf{m}_1^0, \mathbf{m}_2, \tilde{\mathbf{b}}^k) \rightarrow \mathbf{m}_2^{k+1} = \operatorname{argmin}_{\mathbf{m}_2} \varphi_2(\mathbf{m}_2) + \lambda\left\|\mathbf{L}(\mathbf{u}_2^{k+1})\mathbf{m}_2 + \mathbf{A}_1(\mathbf{m}_1^0)\mathbf{u}_1^0 - \mathbf{y}\right\|_2^2$

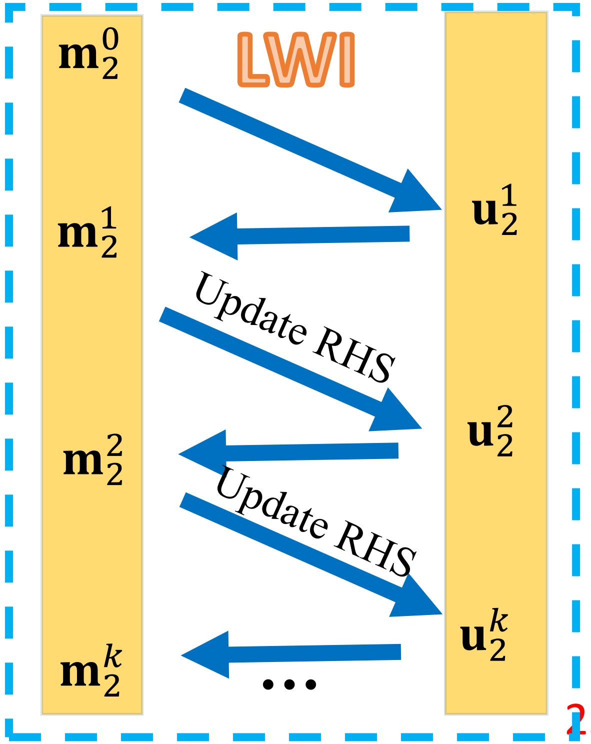
③ $\max_{\tilde{\mathbf{b}}} \mathcal{L}_A(\mathbf{u}_1^0, \mathbf{u}_2^k, \mathbf{m}_1^0, \mathbf{m}_2^k, \tilde{\mathbf{b}}) \rightarrow \tilde{\mathbf{b}}^{k+1} = \tilde{\mathbf{b}}^k + \mathbf{b} - \mathbf{A}(\mathbf{m}_2^{k+1})\mathbf{u}_2^{k+1} - \mathbf{A}(\mathbf{m}_1^0)\mathbf{u}_1^0$

How to extract $\mathbf{u}_1^0$? → We can extract the DA wavefield in the whole of the medium one-time before starting the inversion.


$\mathbf{m}^0$
DA extraction
$\mathbf{u}_1^0$
$\mathbf{u}_2^0$
$\mathbf{m}_1^0$
$\mathbf{m}_2^0$

It requires solving n_s augmented wave-equation.

We can update $\mathbf{m}_1^0$ and $\mathbf{m}_2^0$ since updating model parameters in IR-WRI doesn't have any computational cost.


LWI
$\mathbf{m}_2^0$
$\mathbf{m}_2^1$
$\mathbf{m}_2^2$
$\mathbf{m}_2^k$
$\mathbf{u}_2^1$
$\mathbf{u}_2^2$
$\mathbf{u}_2^k$
Update RHS
Update RHS
...

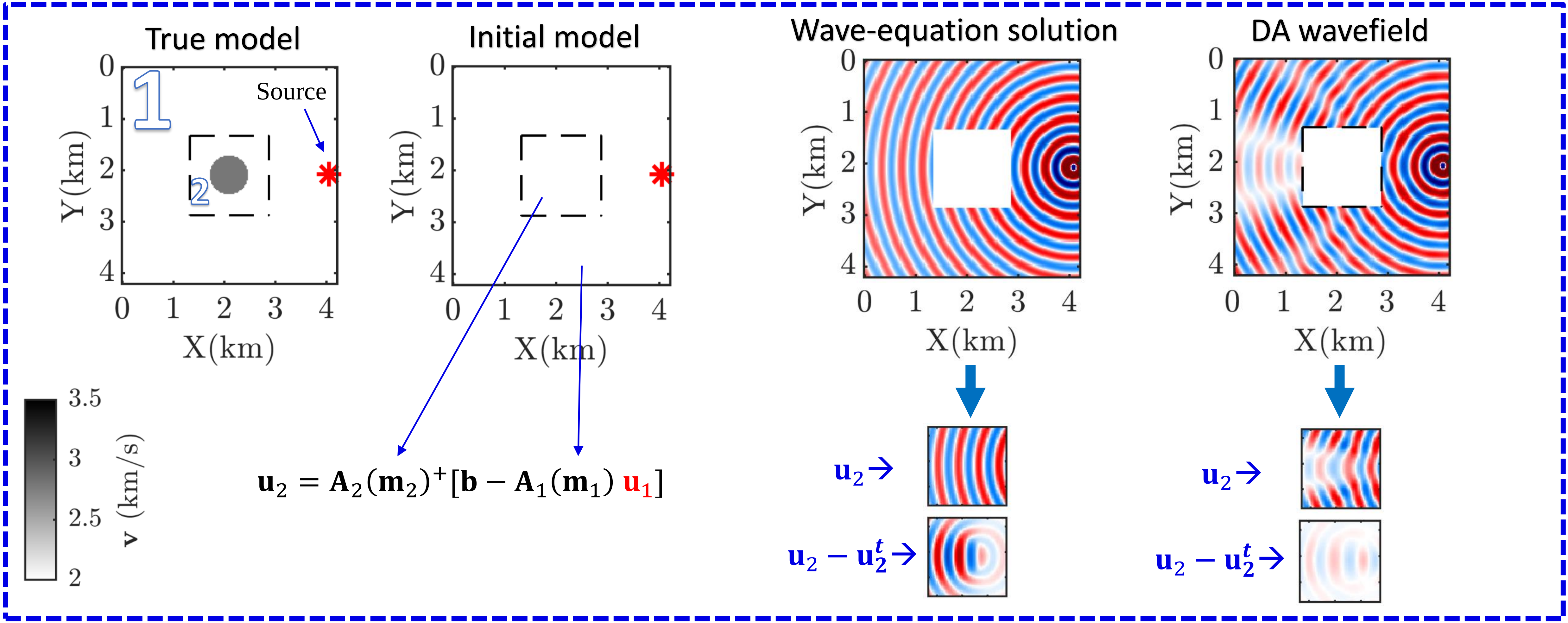
Computational cost of LWI → n_s augmented wave-equation in full domain + $n_{it} * n_s$ wave-equation in subdomain 2.

What is missed in LWI?

wave-equation → $\mathbf{A}_1(\mathbf{m}_1)\,\mathbf{u}_1 + \mathbf{A}_2(\mathbf{m}_2)\,\mathbf{u}_2 = \mathbf{b}$ → $\mathbf{u}_2 = \mathbf{A}_2(\mathbf{m}_2)^{\dagger}[\mathbf{b} - \mathbf{A}_1(\mathbf{m}_1)\,\underbrace{\mathbf{u}_1}_{\mathbf{u}_1(\mathbf{m}_1,\mathbf{m}_2)}]$

$■^{\dagger}$ → The generalized inverse of ■.

When $\mathbf{m}_2$ changes, we should update $\mathbf{u}_1$, otherwise the interaction of the wavefield between 2 and 1 are missed and extracted $\mathbf{u}_2$ is an approximation.



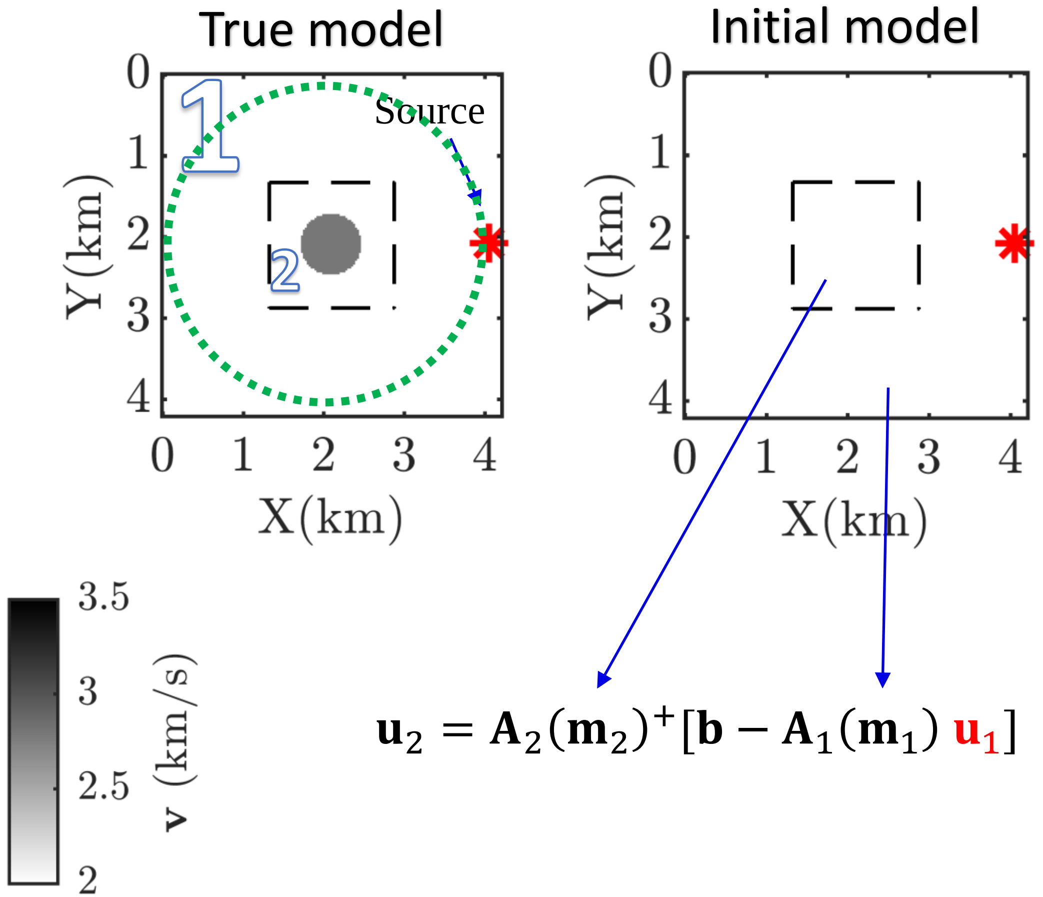
$$\mathbf{u}_2 = \mathbf{A}_2(\mathbf{m}_2)^{+}[\mathbf{b} - \mathbf{A}_1(\mathbf{m}_1)\,\mathbf{u}_1]$$

What is missed in LWI?

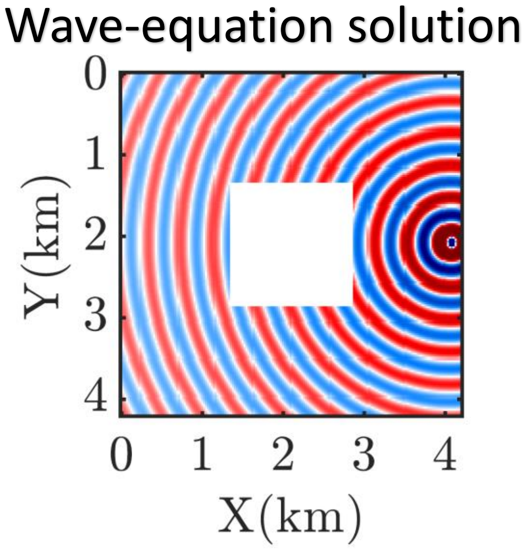
wave-equation → $\mathbf{A}_1(\mathbf{m}_1)\,\mathbf{u}_1 + \mathbf{A}_2(\mathbf{m}_2)\,\mathbf{u}_2 = \mathbf{b}$ → $\mathbf{u}_2 = \mathbf{A}_2(\mathbf{m}_2)^{\dagger}[\mathbf{b} - \mathbf{A}_1(\mathbf{m}_1)\,\underbrace{\mathbf{u}_1}_{\mathbf{u}_1(\mathbf{m}_1,\mathbf{m}_2)}]$

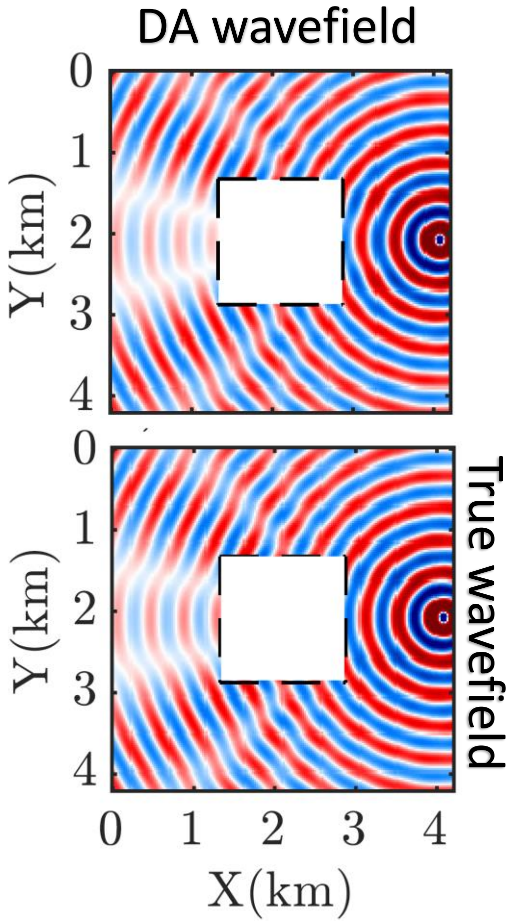
$■^{\dagger}$ → The generalized inverse of ■.

When $\mathbf{m}_2$ changes, we should update $\mathbf{u}_1$, otherwise the interaction of the wavefield between 2 and 1 are missed and extracted $\mathbf{u}_2$ is an approximation.



$$\mathbf{u}_2 = \mathbf{A}_2(\mathbf{m}_2)^{+}[\mathbf{b} - \mathbf{A}_1(\mathbf{m}_1)\,\mathbf{u}_1]$$





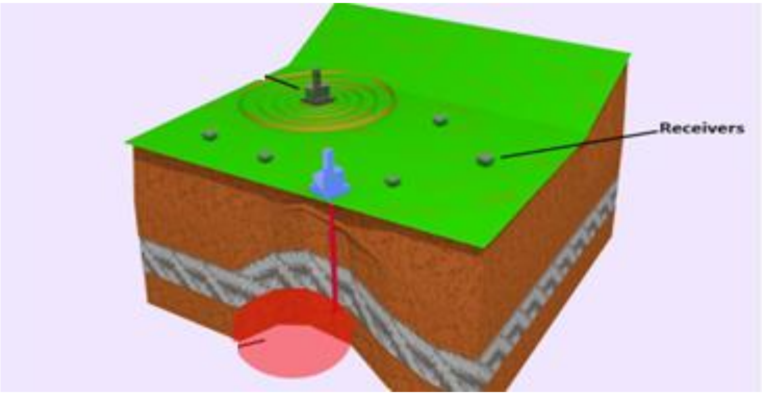
DA wavefield in the initial model is a good approximation of $\mathbf{u}_1^t$.

Marmousi II test: a 4D example

The term 4D reflects that calendar time represents the fourth dimension.

Here the goal is to rapidly estimate the local changes that happen because of injected fluids or gas in the subsurface between a baseline and monitor data.

- ✓ Surface acquisition with 57 sources and 650 receivers.
- ✓ Inverted frequencies are 3-13Hz with frequency.
- ✓ A 10 Hz Ricker is used as the wavelet.

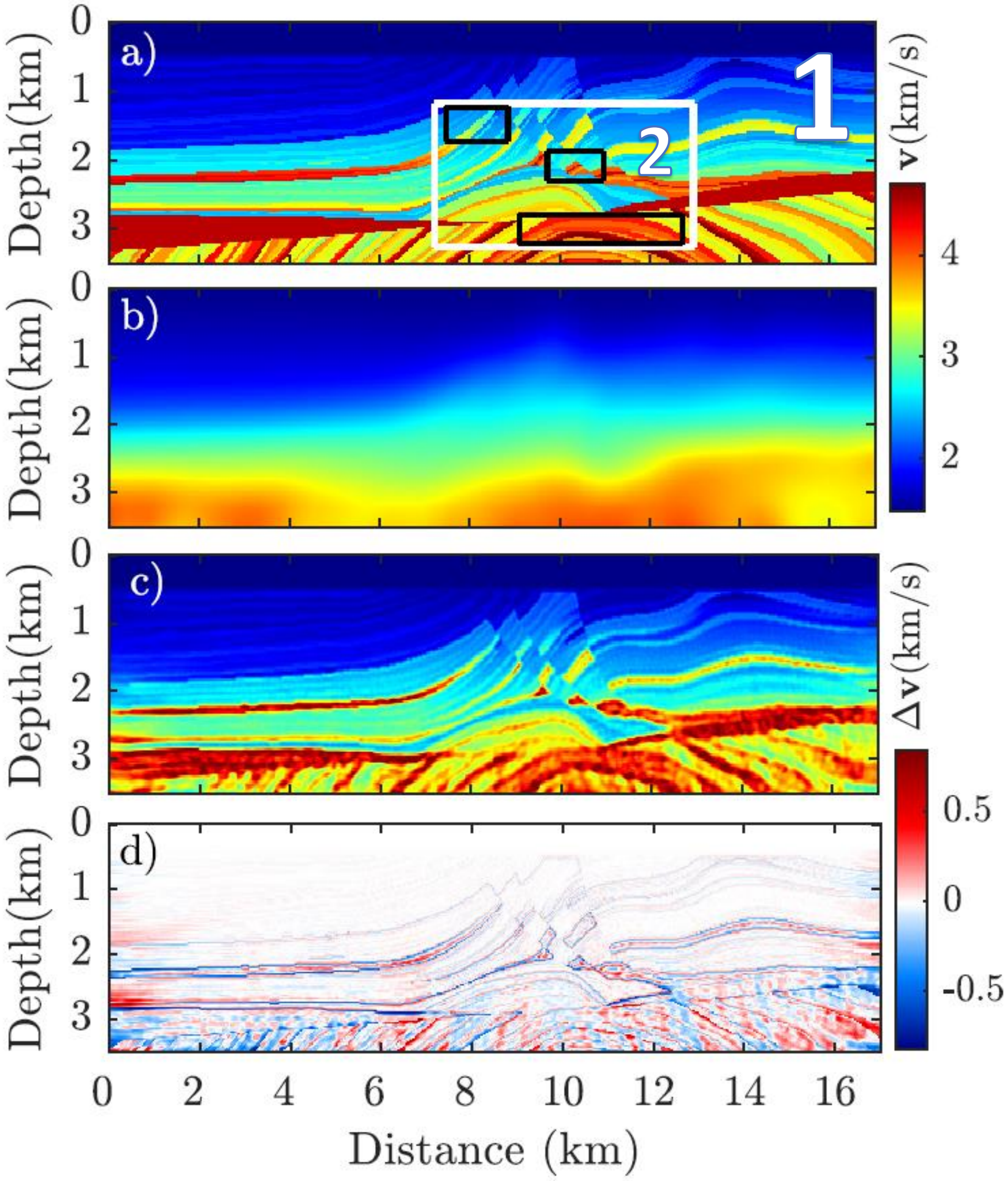


True baseline model

Initial model

IR-WRI applied on baseline data

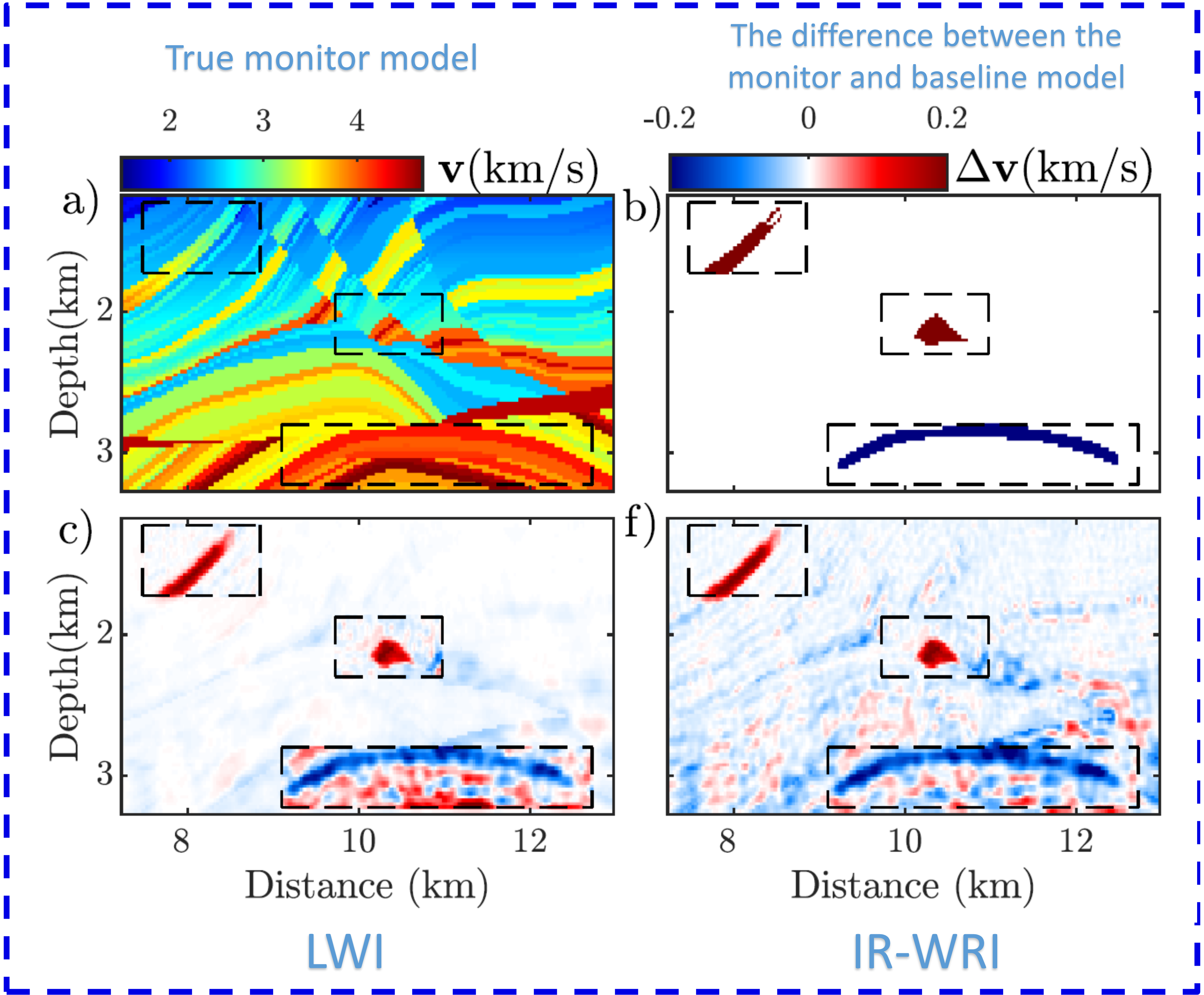
The difference between true and inverted models



Marmousi II test: a 4D example

- ✓ We use three frequencies for the monitor data inversion, [5, 10, 15]Hz, and a successive mono-frequency inversion.
- ✓ We use the baseline inverted model as the initial model.

LWI reaches approximately to the same model as IR-WR, but 24 times faster.



Pitfalls and challenges

- Nonlinearity
  - Non convexity and local minima (cycle skipping)
- Ill-posedness
  - Incomplete illumination
  - Parameter cross-talk
- Systematic errors
  - Approximate physics in forward problem
- HPC issue
  - Computational burden

Conclusions

- ✓ FWI is a high-resolution imaging technique that has a wide range of applications.
- ✓ We proposed to use augmented Lagrangian for FWI when it is solved using ADMM.
- ✓ We show this formulation can improve the difficulty of the classical formulation with the initial model as well as the difficulties of FWI based on penalty formulation.
- ✓ We show an adaption of multi-block ADMM-based wavefield inversion to reduce the computational cost of FWI for target-oriented applications.
- ✓ In this method, the subproblems related to the zone of interest are solved normally at each iteration, while the rest are solved only once.

THANKS

THANKS FOR YOUR ATTENTION

WIND

Waveform Inversion of Node Data

ICMS

Questions?

Comments?