Source estimation and uncertainty quantification for wave-equation based seismic imaging and inversion

Zhilong Fang

PhD Defense - April 27, 2018

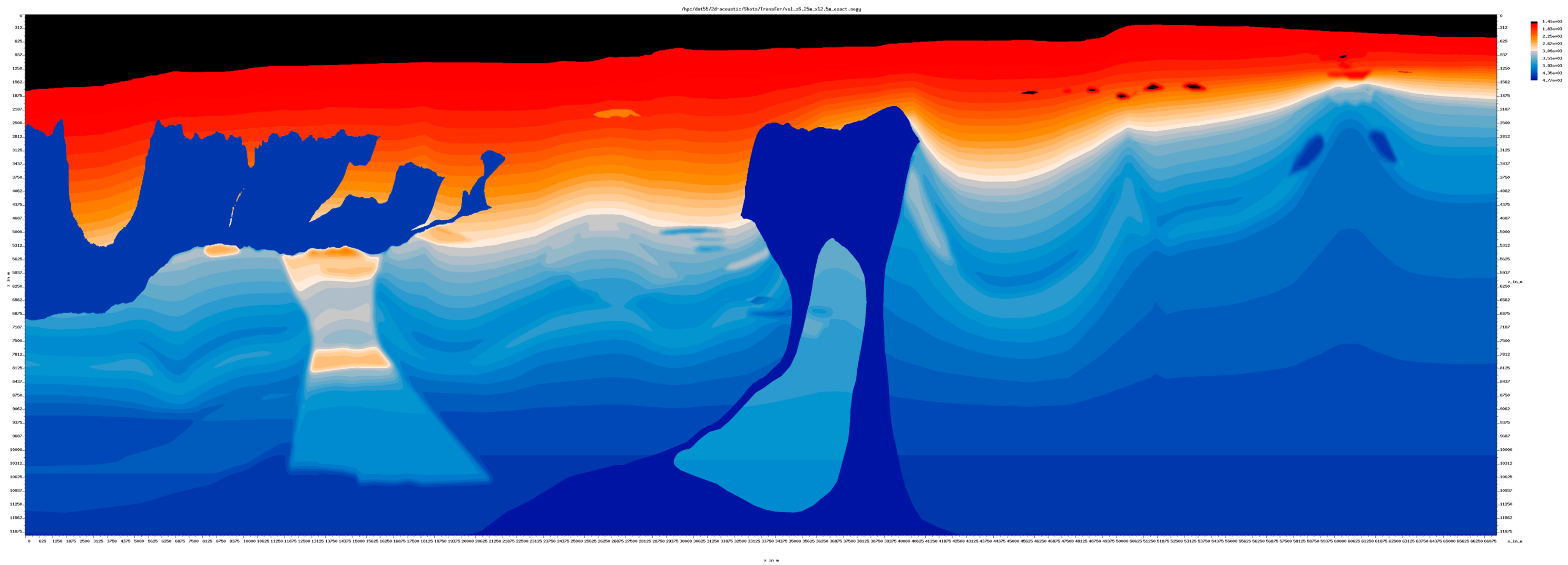
SLIM
University of British Columbia

https://wiki.seg.org/wiki/2004_BP_velocity_estimation_benchmark_model

Seismic exploration

Main questions:

1. Where is the oil and gas reservoir?

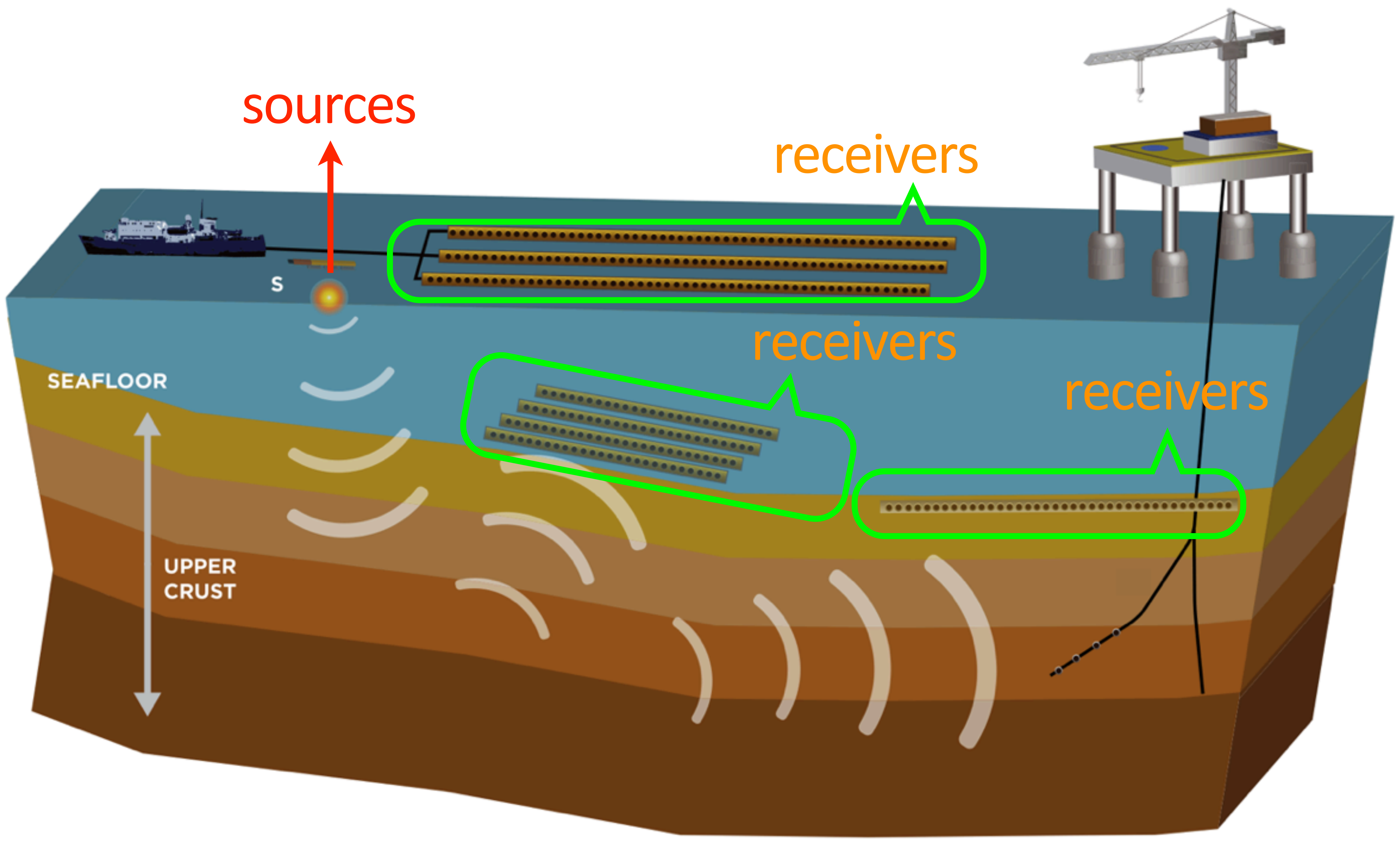


Where?

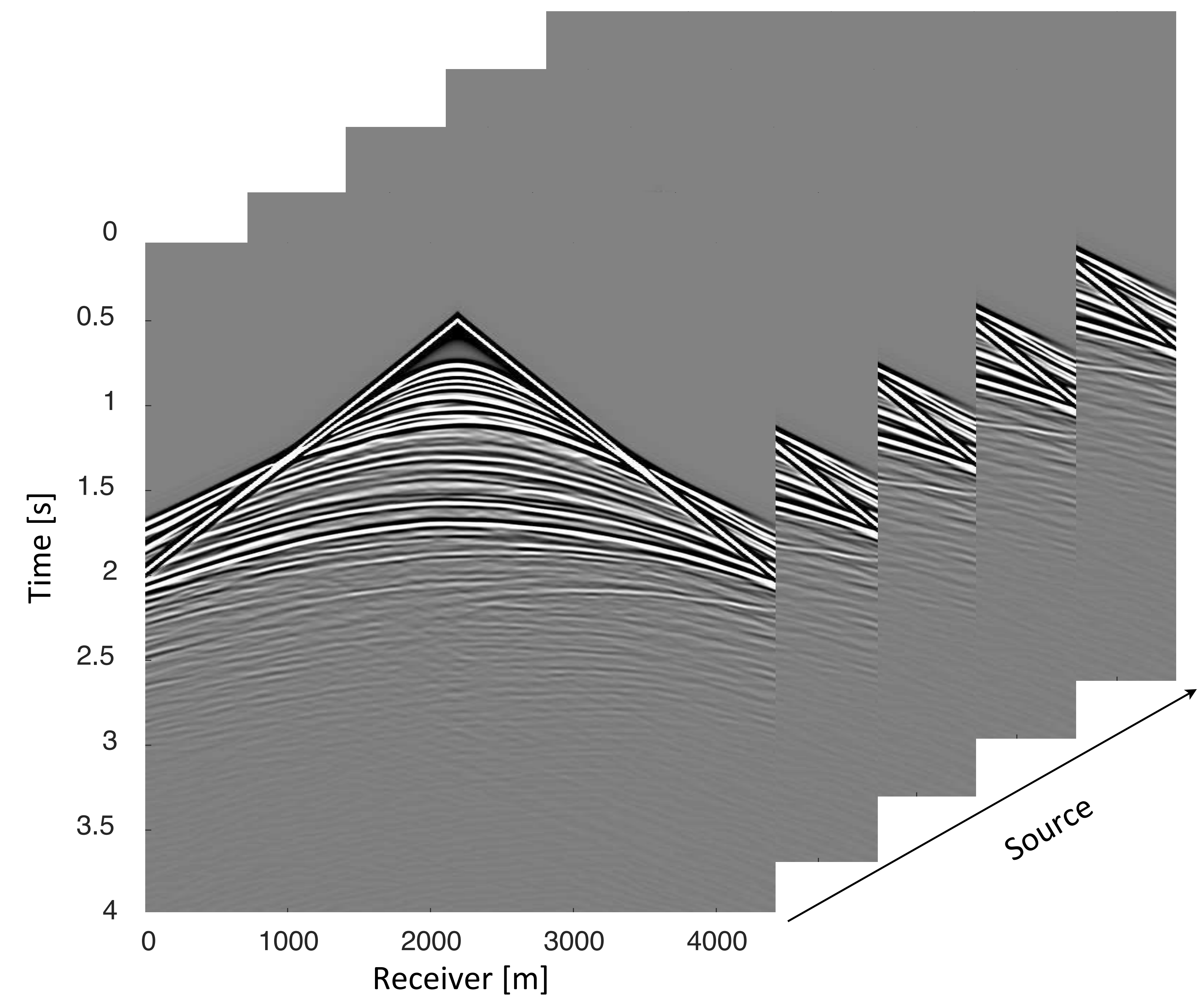
2. What is the volume of the oil and gas reservoir?

How much?

Seismic exploration



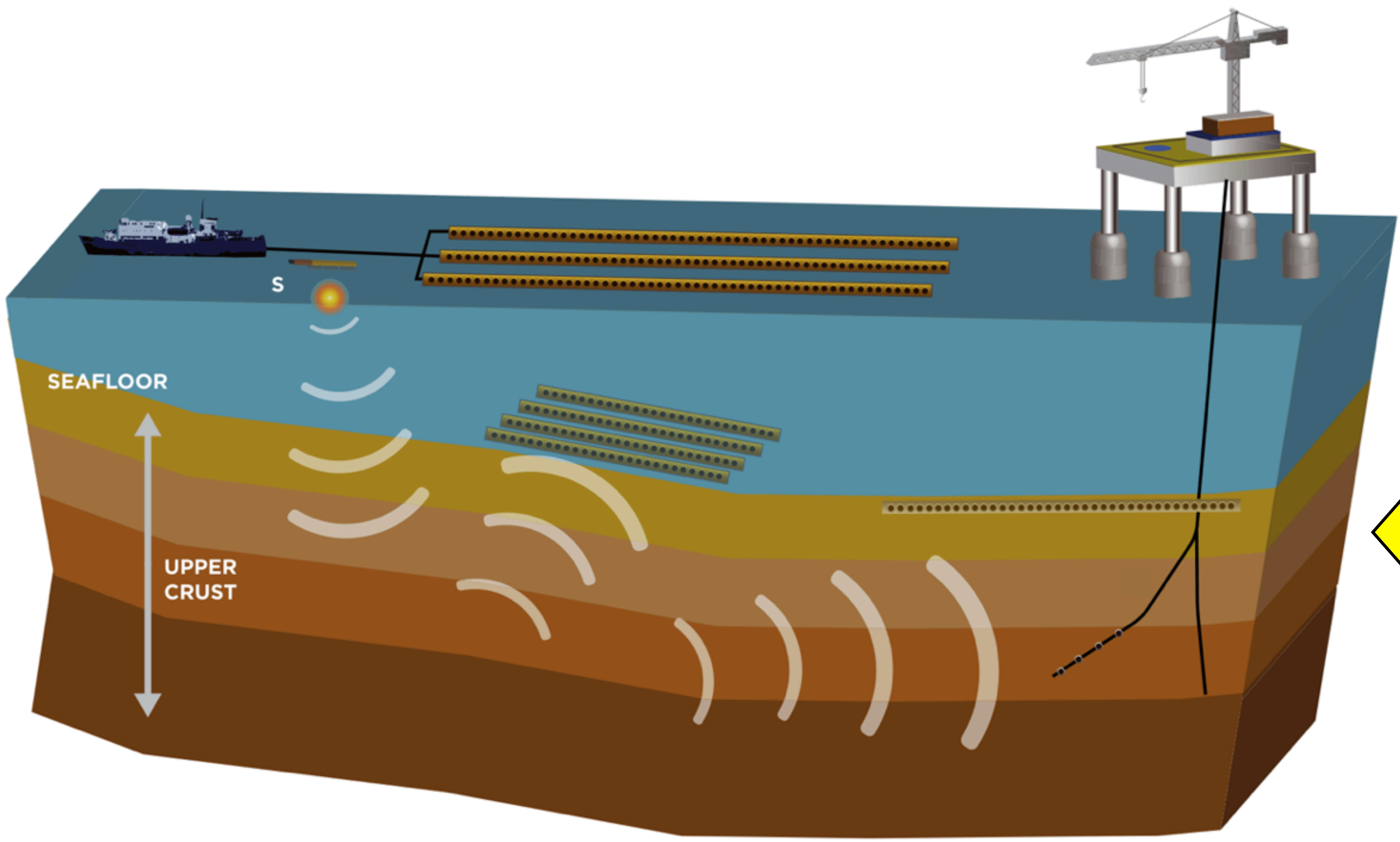
Seismic exploration survey



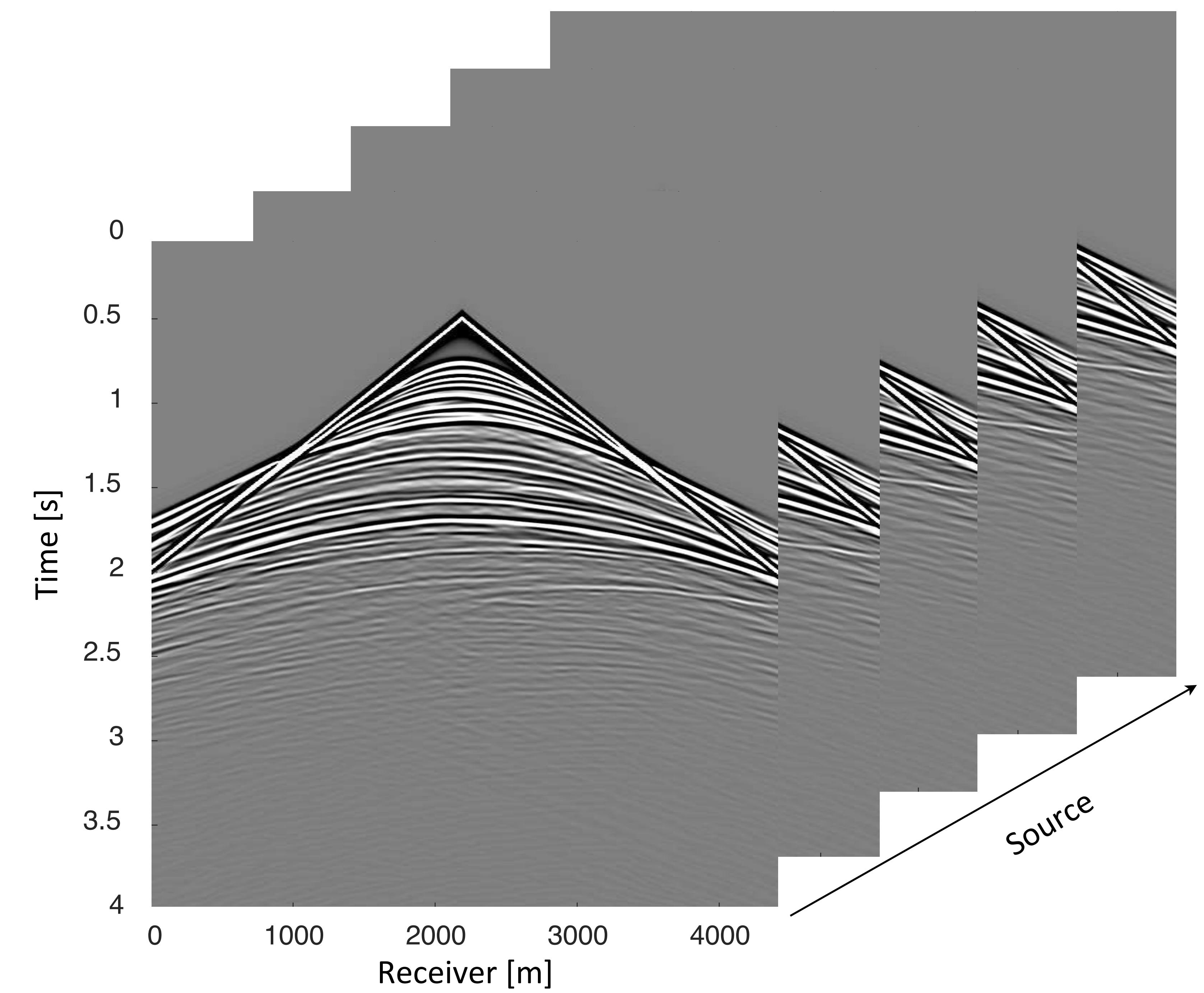
Seismic data volume

Seismic exploration

Objective

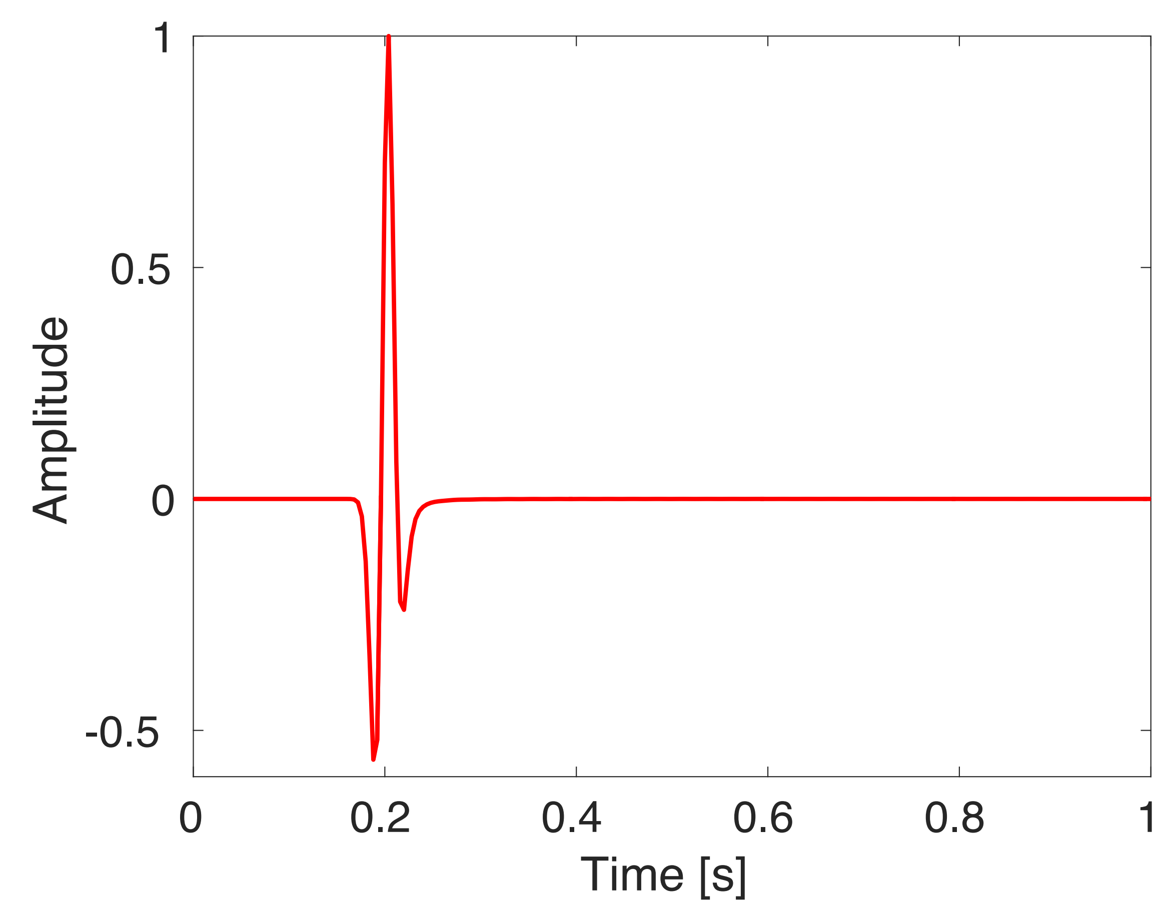


Seismic exploration survey

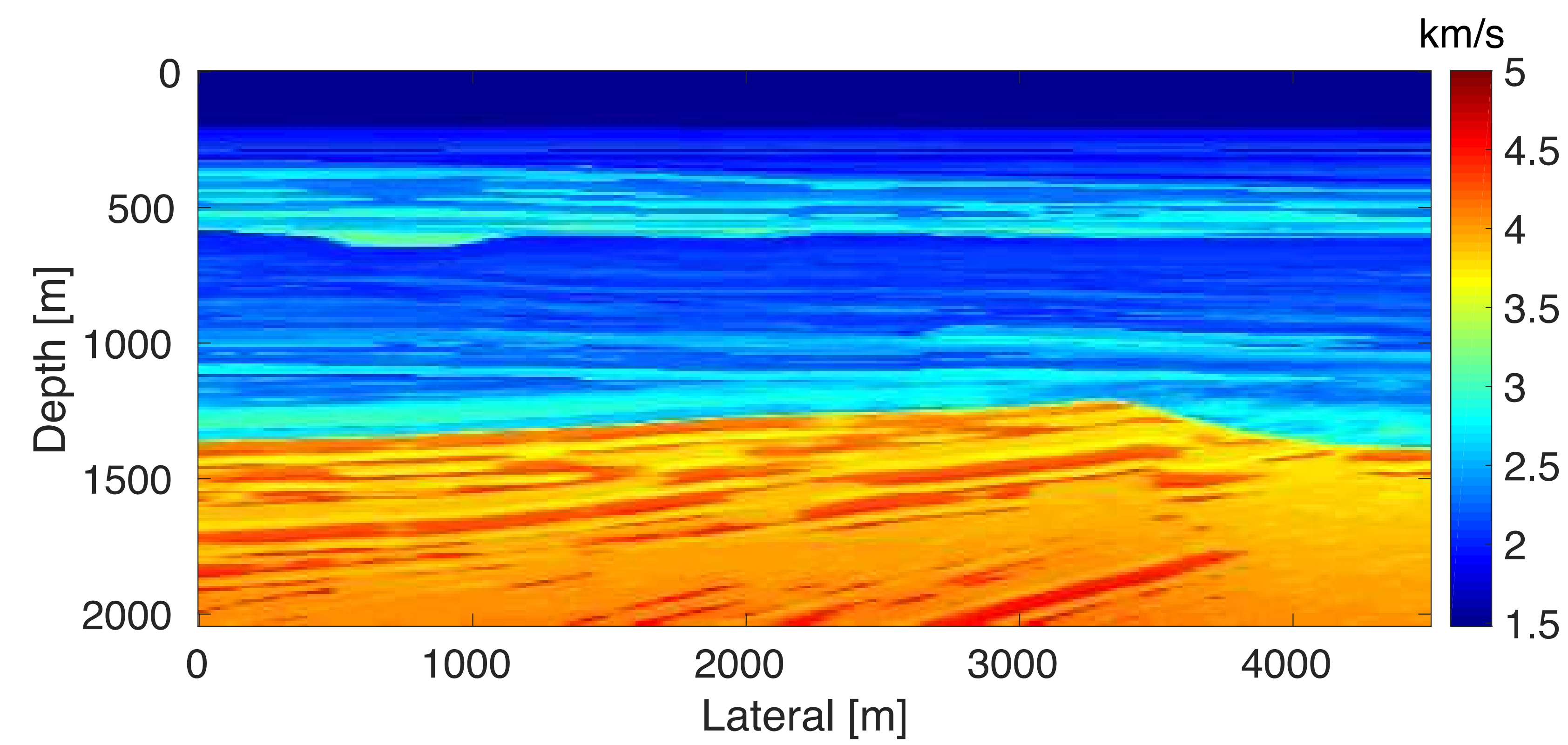


Seismic data volume

Seismic exploration



$\mathbf{q}$ - source function



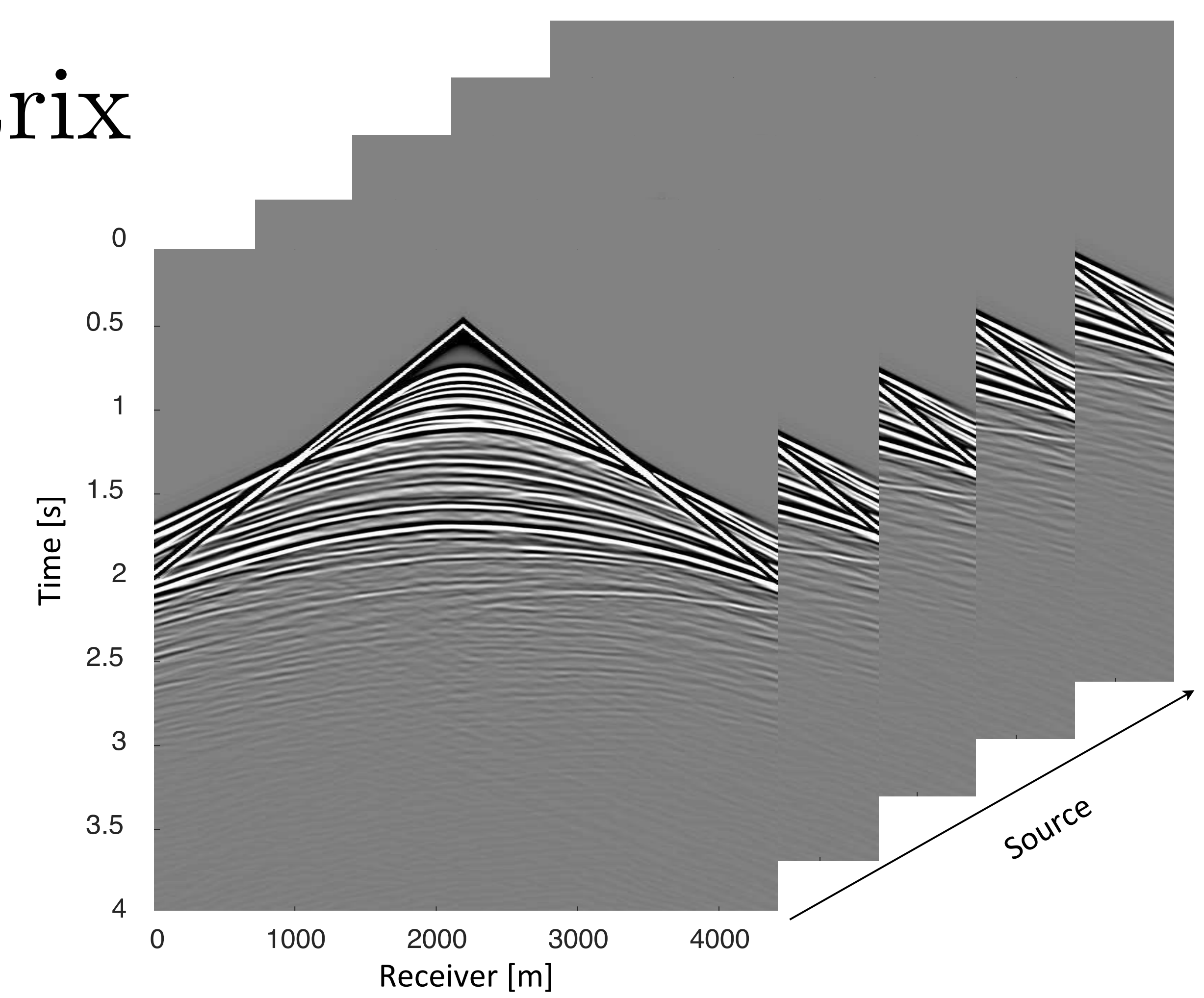
$\mathbf{m}$ - model

Forward modeling:

$$\mathbf{d} = F(\mathbf{m}) = \mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q}$$

$\mathbf{A}$: discretized wave equation stencil matrix

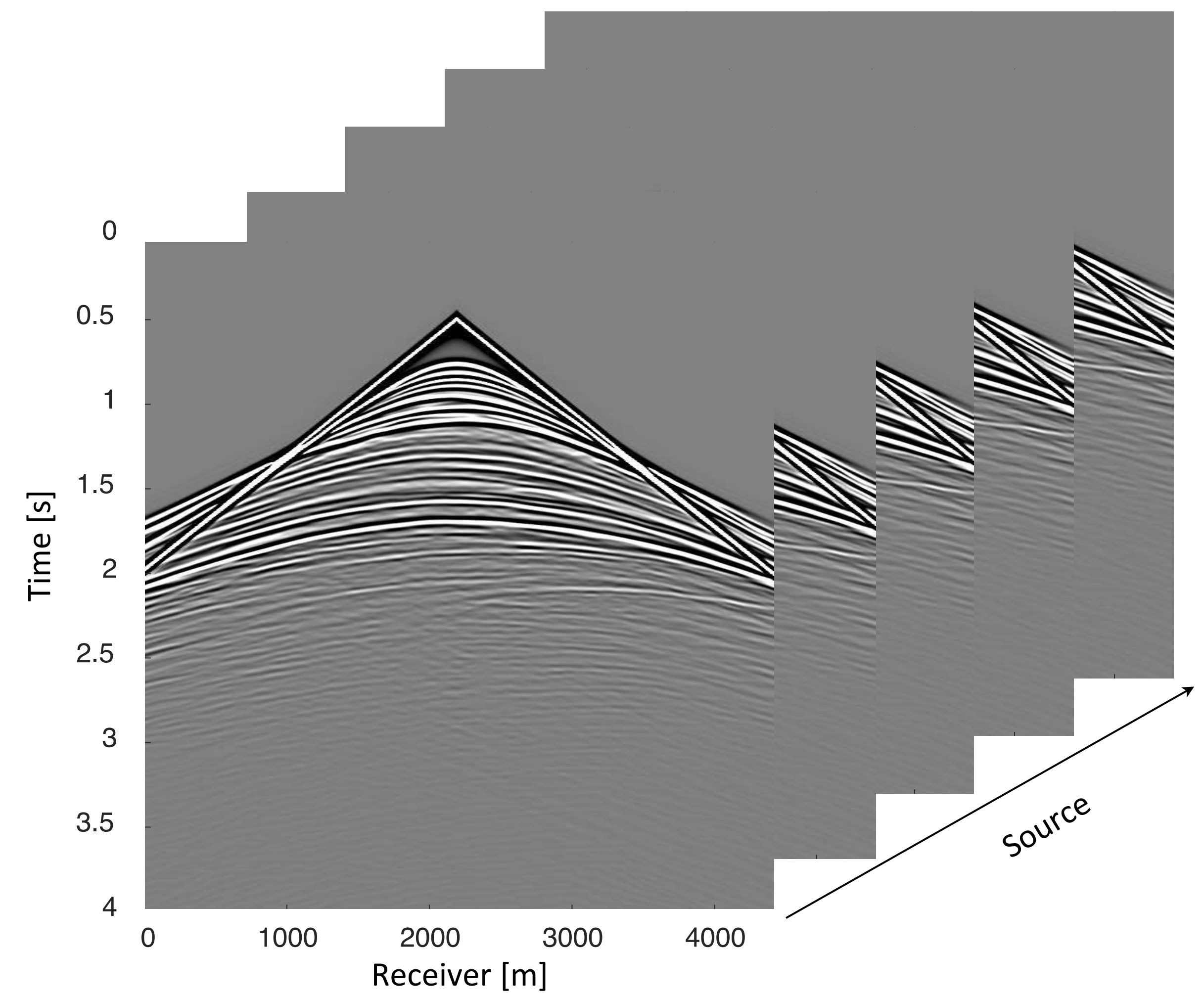
$\mathbf{P}$: receiver restriction operator



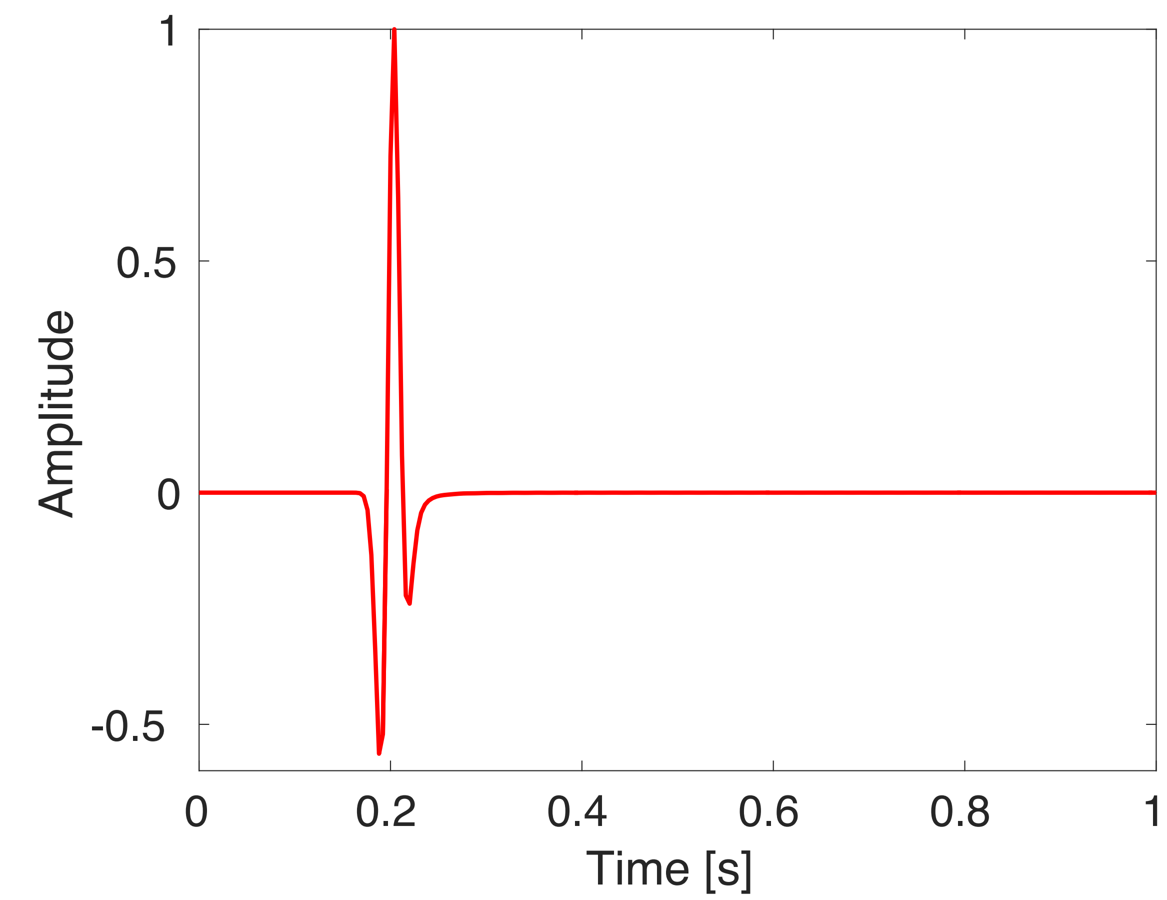
$\mathbf{d}$ - data

5

Seismic exploration



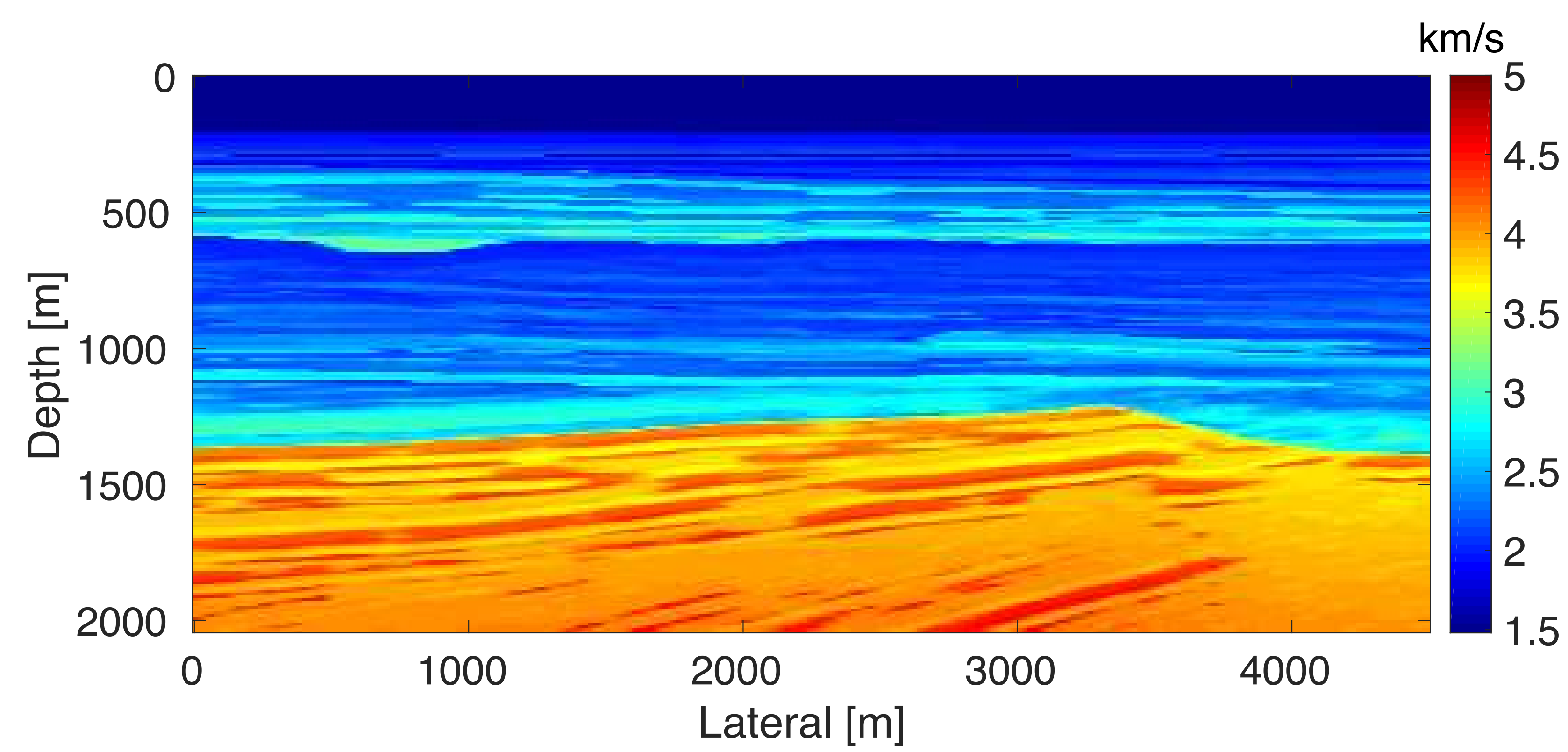
$\mathbf{d}$ - data



$\mathbf{q}$ - source function

Inversion or imaging:

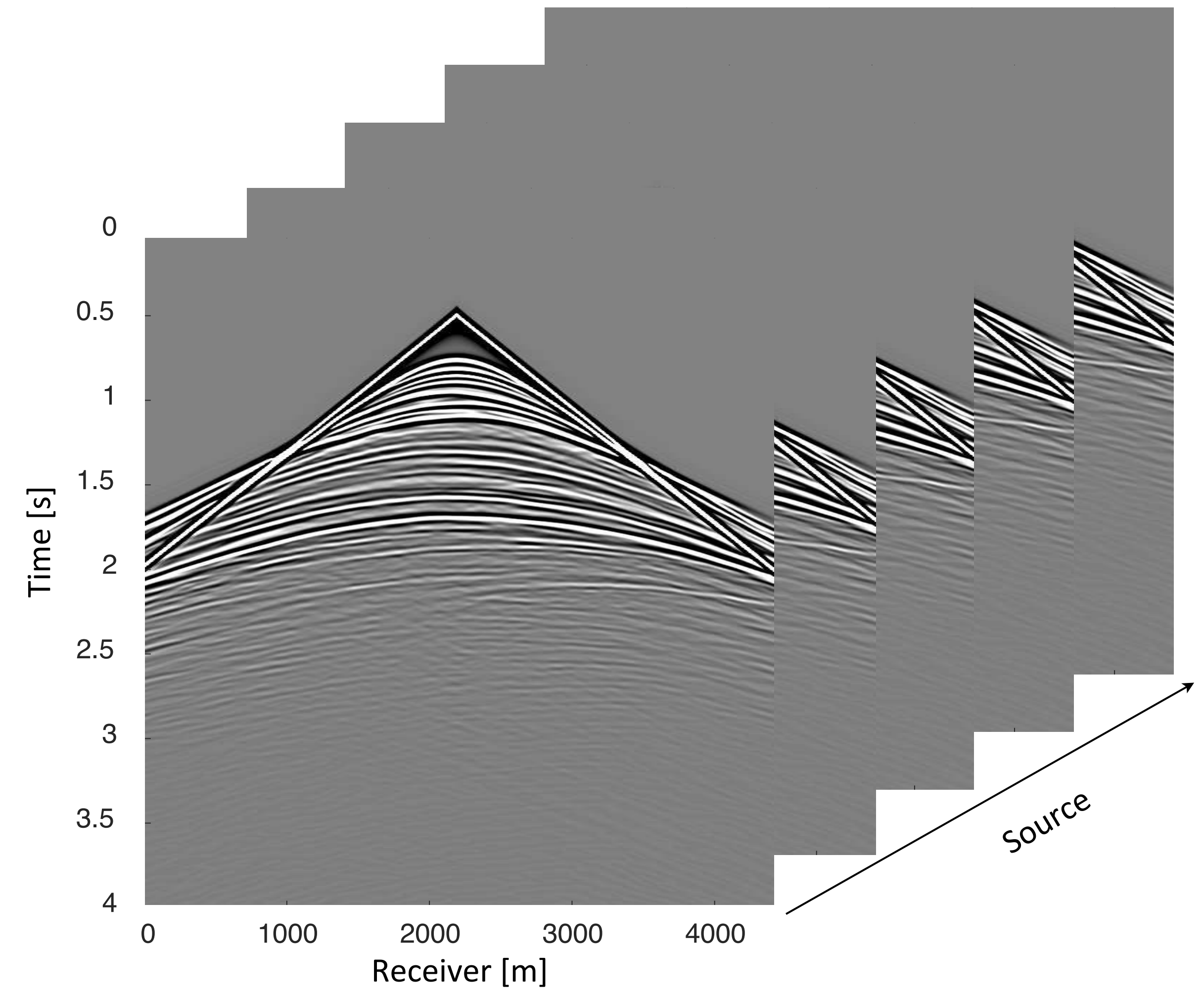
$$\text{“}\mathbf{m} = F^{-1}(\mathbf{d})\text{”}$$



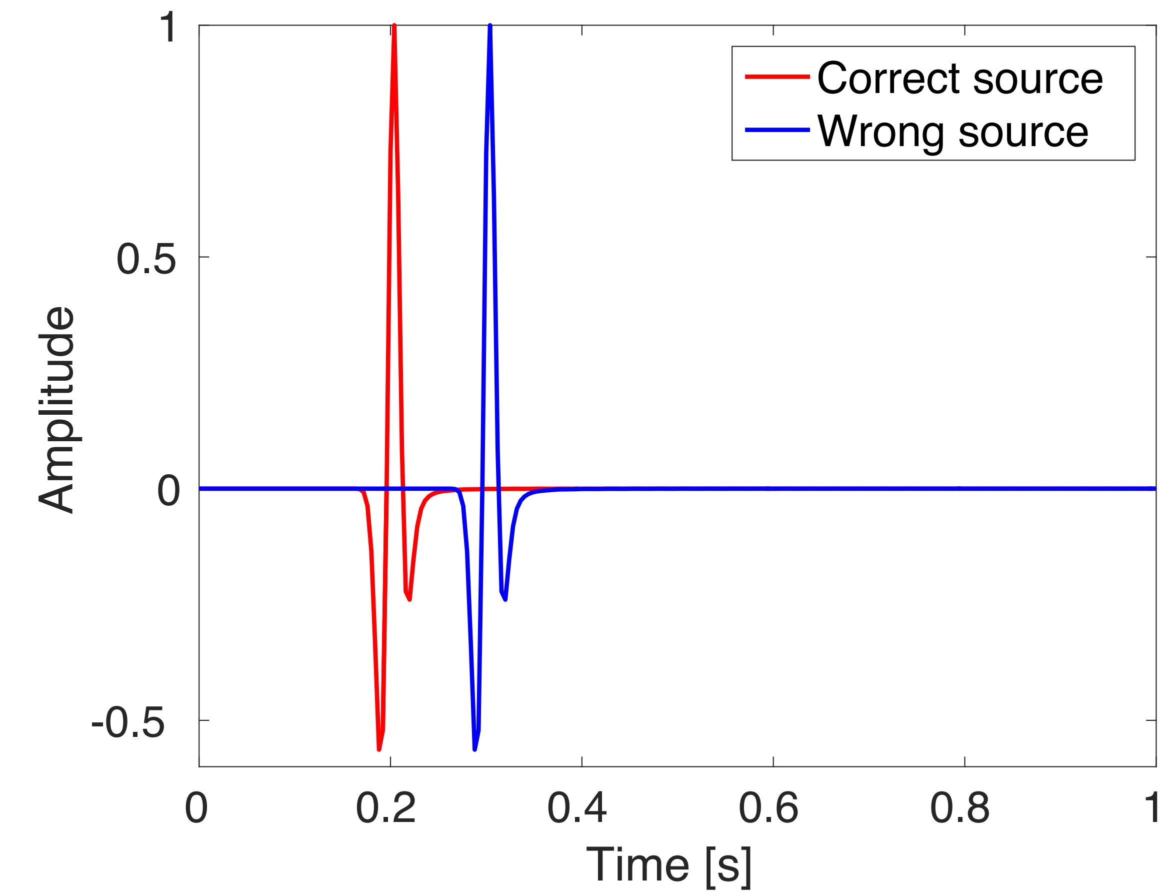
$\mathbf{m}$ - model

6

Seismic exploration
— unknown source



$\mathbf{d}$ - data

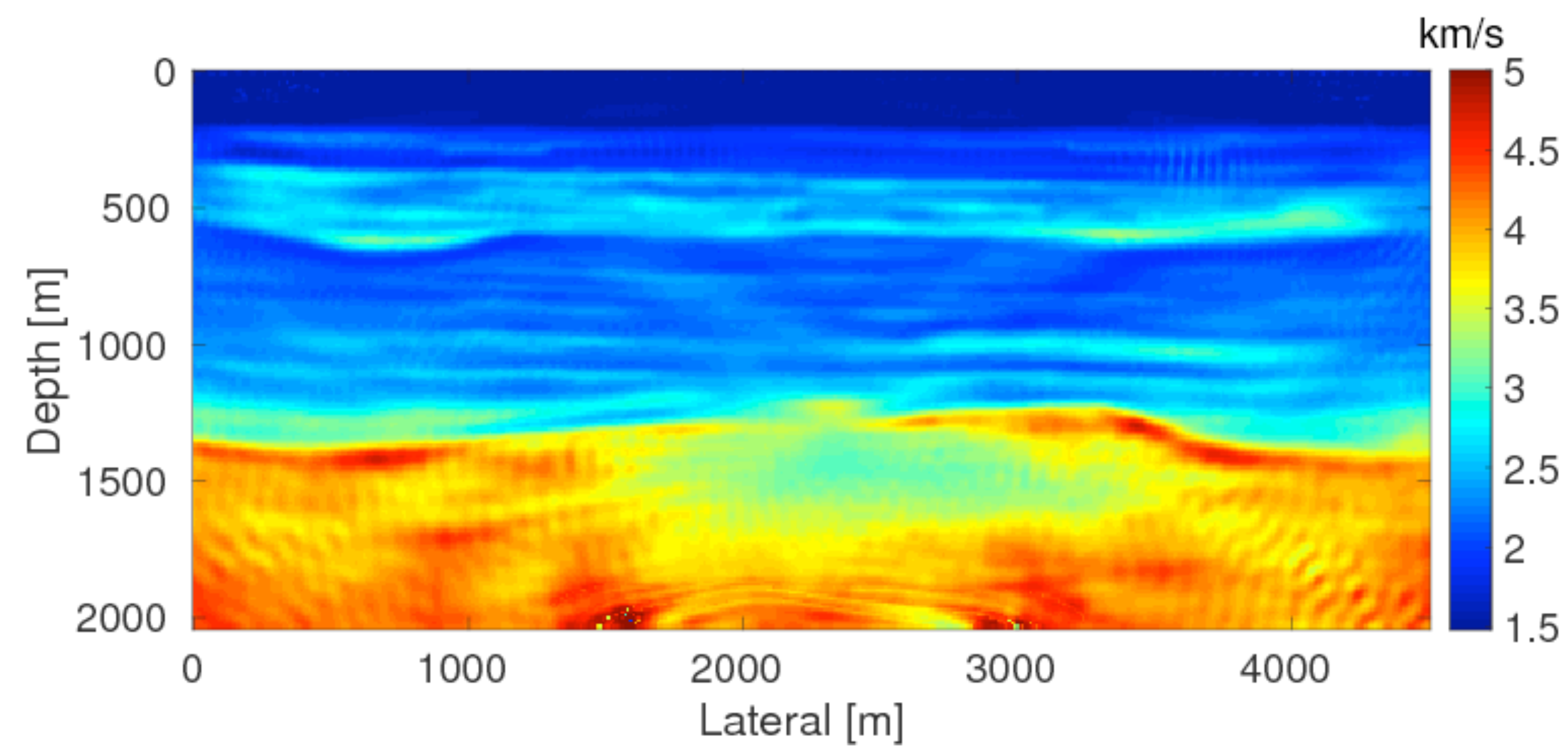


$\mathbf{q}$ - source function

Inversion or imaging:

$$\text{“}\mathbf{m} = F^{-1}(\mathbf{d}, \underline{\mathbf{q}})\text{”}$$

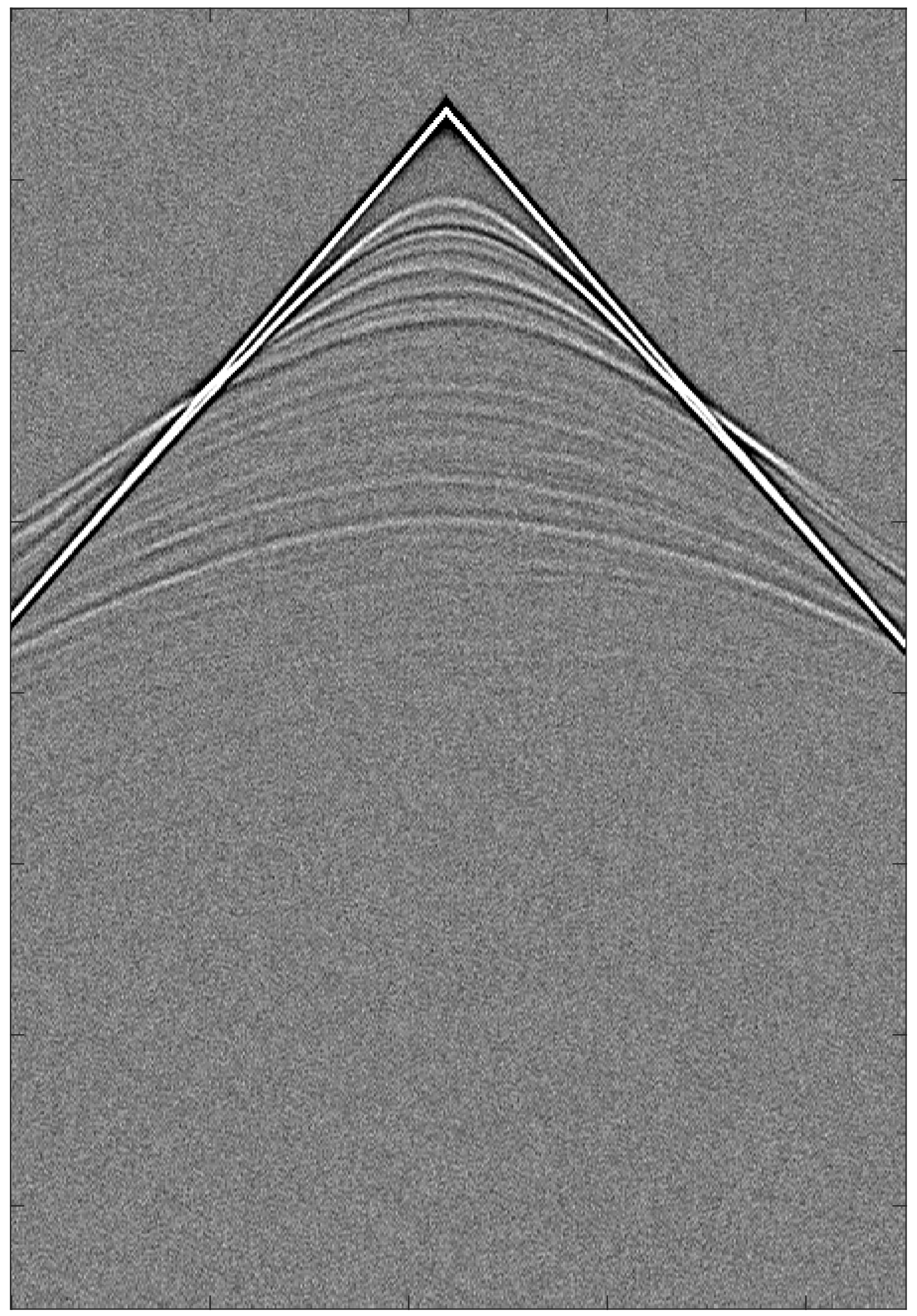
if source is wrong ?

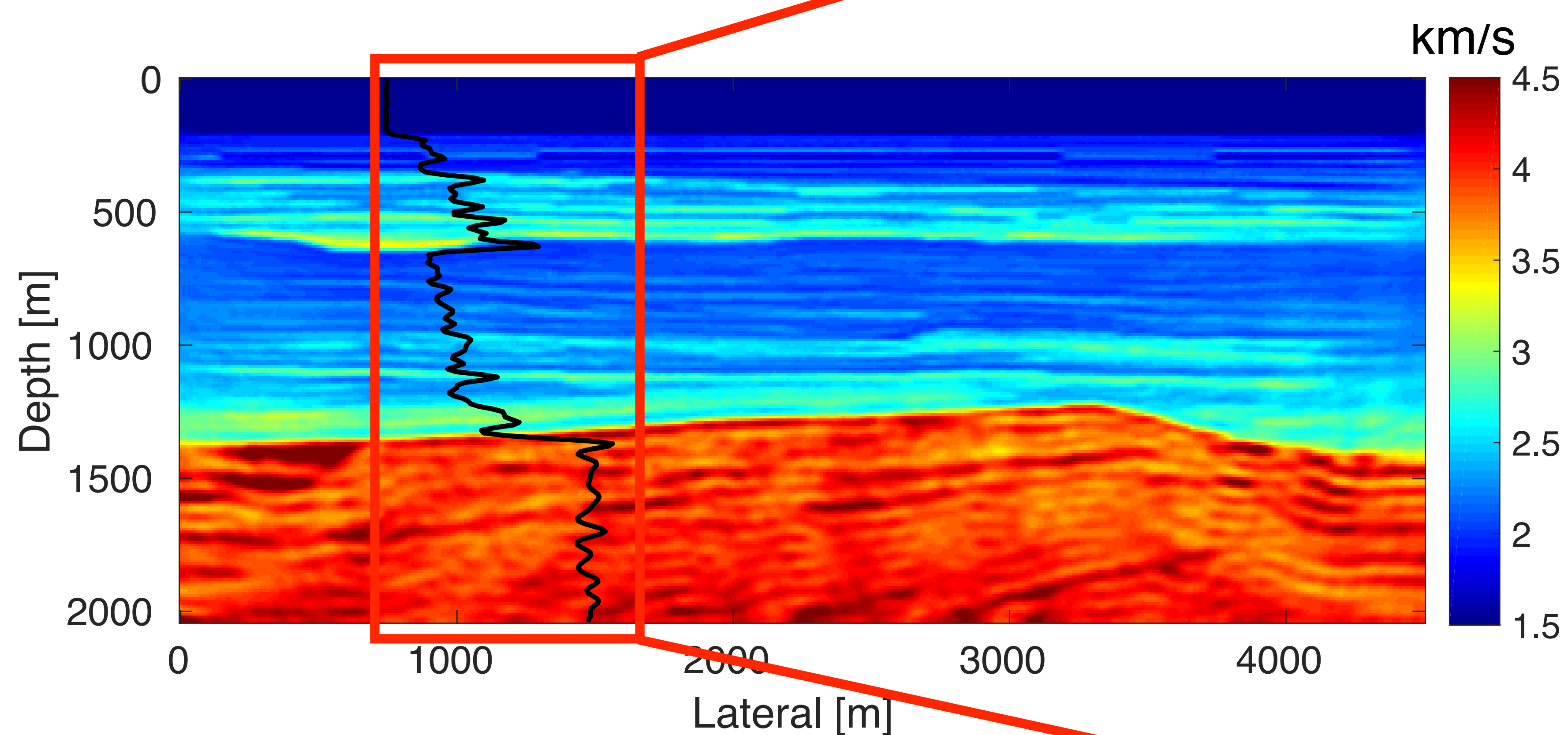


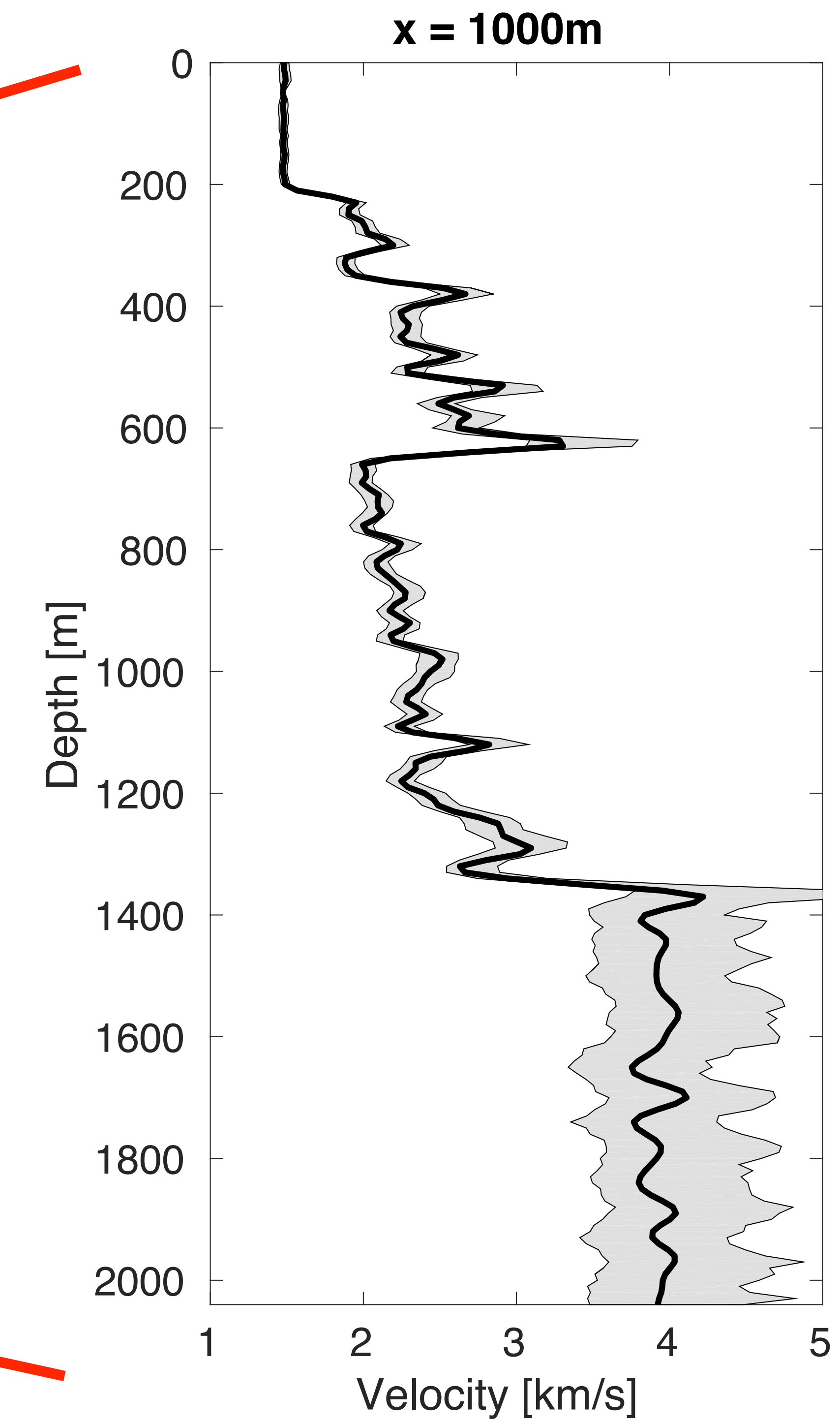
$\mathbf{m}$ - model

7

Seismic exploration
— uncertainty quantification







Uncertainty in data

Uncertainty in model

Confidence of the model

Purposes of thesis

Outline:

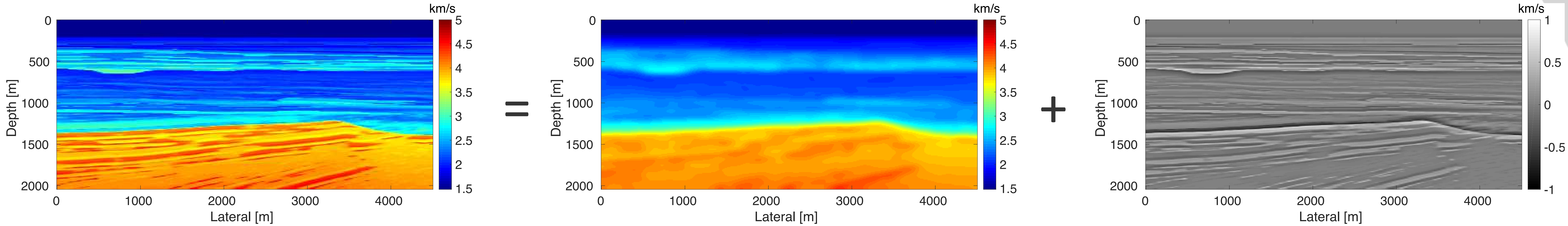
- **Chapter 2:** Source estimation for time-domain sparsity promoting least-squares reverse-time migration (LS-RTM);
- **Chapter 3-5:** Source estimation for wavefield reconstruction inversion (WRI);
- **Chapter 6:** Uncertainty quantification for inverse problems with weak wave-equation constraints.

Chapter 2

Source estimation for time-domain least-squares reverse-time migration

Time domain LS-RTM

Model separation

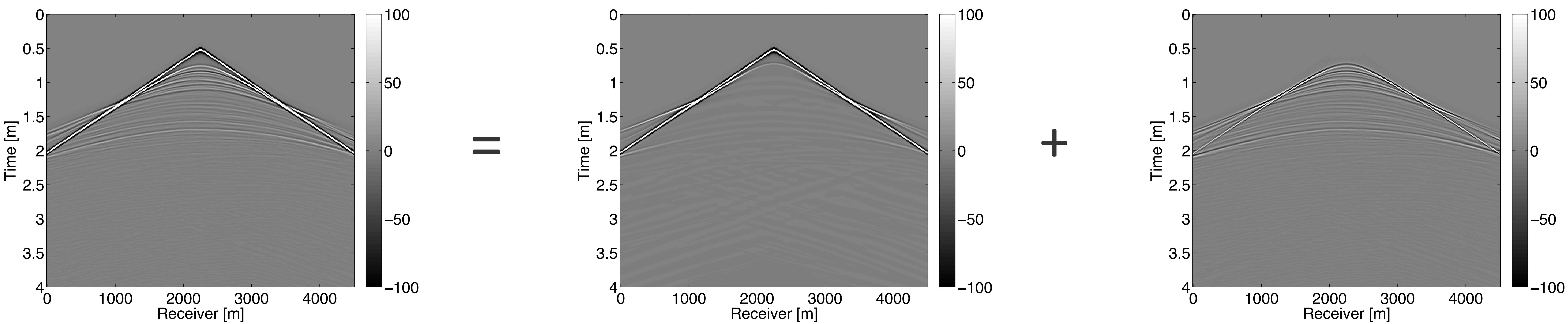


True model $\mathbf{m}_{\mathrm{t}}$

Smooth background model $\mathbf{m}_{\mathrm{s}}$

Model perturbation $\delta\mathbf{m}$

Data separation



$F(\mathbf{m}_{\mathrm{t}})$

$F(\mathbf{m}_{\mathrm{s}})$

$\delta\mathbf{d}$

[N. Aoki and G. T. Schuster, 2009] [X. Li and F. Herrmann, 2012], [N. Tu and F. Herrmann, 2015]

SLIM

Time domain LS-RTM

Taylor expansion:

$$F(\mathbf{m}_\mathrm{t}) = F(\mathbf{m}_\mathrm{s}) + \nabla F(\mathbf{m}_\mathrm{s})\delta\mathbf{m} + \mathcal{O}(\delta\mathbf{m}^2)$$

Linearized approximation or Born modeling:

$$F(\mathbf{m}_\mathrm{t}) \approx F(\mathbf{m}_\mathrm{s}) + \nabla F(\mathbf{m}_\mathrm{s})\delta\mathbf{m} \Rightarrow \delta\mathbf{d} \approx \nabla F(\mathbf{m}_\mathrm{s})\delta\mathbf{m}$$

Forward operator:

$$F(\mathbf{m}) = \mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q},$$

it is related to the acoustic wave equation:

$$\underbrace{\boxed{(\mathbf{m}\frac{\partial^2}{\partial t^2} + \Delta)}}_{\mathbf{A}(\mathbf{m})}\mathbf{u} = \mathbf{q}.$$

Born modeling operator:

2 PDE solves

$$\nabla F(\mathbf{m}) = -\mathbf{P}\boxed{\mathbf{A}(\mathbf{m})^{-1}}(\nabla\mathbf{A}(\mathbf{m})\boxed{\mathbf{u}}),$$

where $\nabla\mathbf{A}(\mathbf{m})\mathbf{u}$ is the Jacobian matrix.

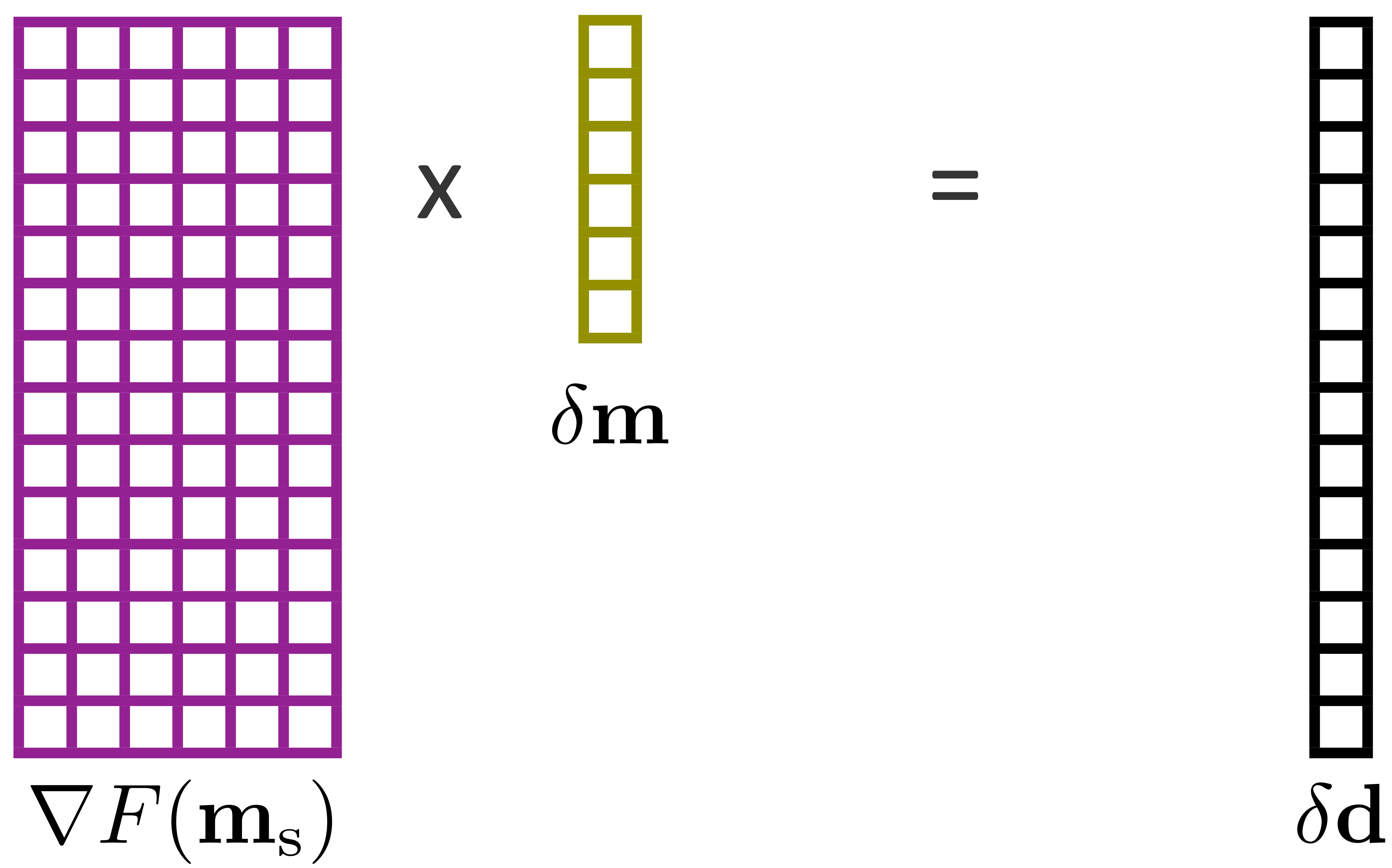
Time domain LS-RTM

Obtain the model perturbation $\delta\mathbf{m}$ from the data perturbation $\delta\mathbf{d}$:

$$\delta\mathbf{m} \approx \nabla F(\mathbf{m}_{\mathrm{s}})^{-1}\delta\mathbf{d}.$$

In practice, we have to solve the following optimization problem:

$$\min_{\delta\mathbf{m}} \frac{1}{2n_{\mathrm{src}}} \sum_{i=1}^{n_{\mathrm{src}}} \|\nabla F_i(\mathbf{m}_{\mathrm{s}})\delta\mathbf{m} - \delta\mathbf{d}_i\|^2$$ **(LS-RTM)**



1. Very large overdetermined system
2. Computationally expensive

SLIM

Time domain LS-RTM

Dimensional reduction + sparsity promoting optimization:

from $$\min_{\delta\mathbf{m}} \frac{1}{2n_{\mathrm{src}}} \sum_{i=1}^{n_{\mathrm{src}}} \|\nabla F_i(\mathbf{m}_{\mathrm{s}})\delta\mathbf{m} - \delta\mathbf{d}_i\|^2$$

to $$\min_{\mathbf{x}} \|\boxed{\mathbf{x}}\|_1$$

Curvelet coefficients

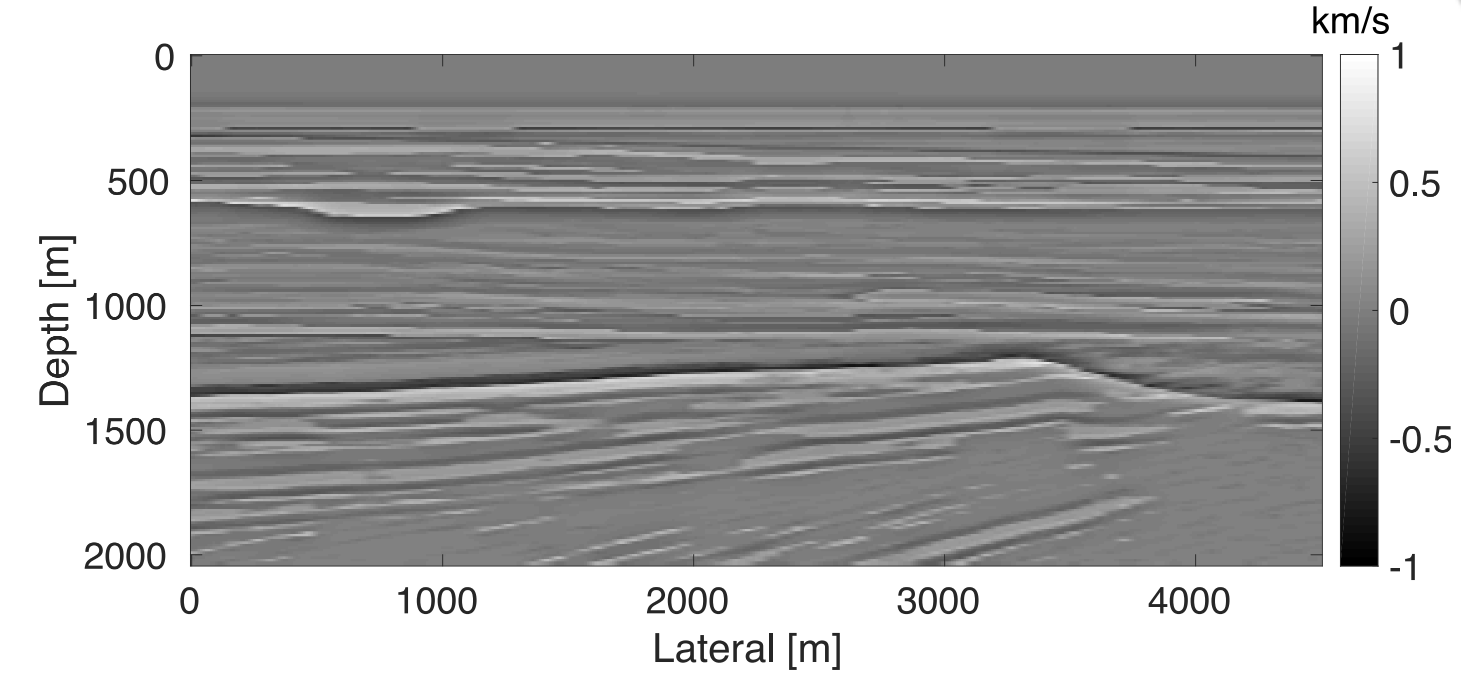
$$\text{s.t.} \quad \frac{1}{2|\mathcal{I}|} \sum_{i\in\mathcal{I}} \|\nabla F_i(\mathbf{m}_{\mathrm{s}})\boxed{\mathbf{C}}^\top \mathbf{x} - \delta\mathbf{d}_i\|^2 \leq \boxed{\sigma}$$

random subset

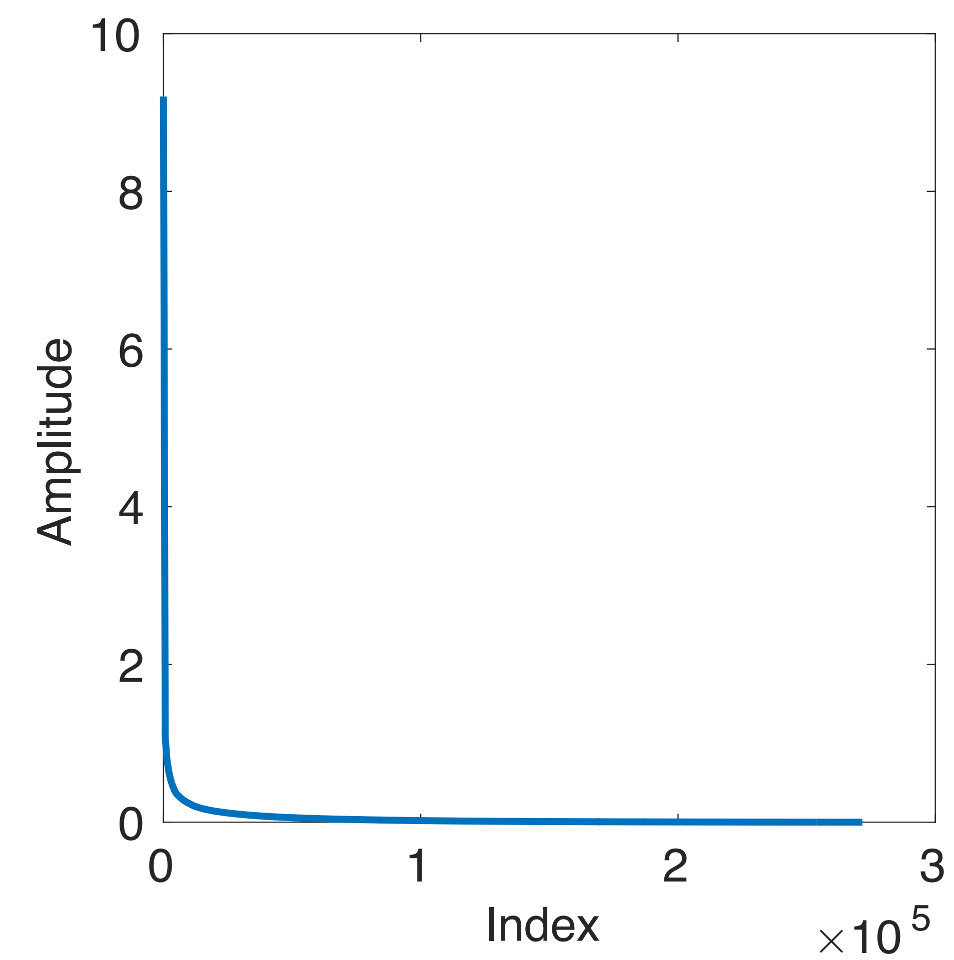
Curvelet transform operator

Tolerance for noise

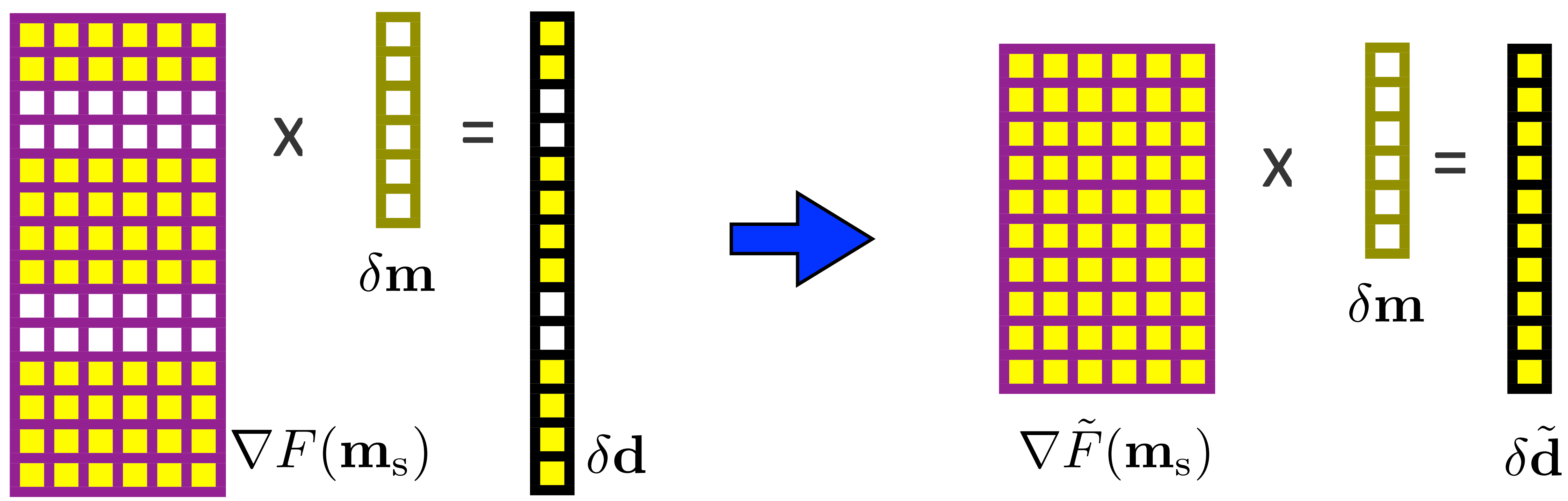
where $\mathcal{I} \subset [1, 2, \cdots, n_{\mathrm{src}}]$, and $|\mathcal{I}| << n_{\mathrm{src}}$.



$\delta\mathbf{m}$



$\mathbf{C}\delta\mathbf{m}$



Time domain LS-RTM

Linearized Bregman method:

from $\min_{\mathbf{x}} \|\mathbf{x}\|_1$ to $\min_{\mathbf{x}} \lambda_1 \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2$

$$\text{s.t.} \quad \frac{1}{2|\mathcal{I}|} \sum_{i \in \mathcal{I}} \|\nabla F_i(\mathbf{m}_\mathrm{s})\mathbf{C}^\top \mathbf{x} - \delta \mathbf{d}_i\|^2 \leq \sigma$$

Algorithm:

1. $\quad$ **Initialize** $\mathbf{x}_0 = \mathbf{0}$, $\mathbf{z}_0 = \mathbf{0}$, $\mathbf{q}$, λ_1, batchsize $n'_\mathrm{src} \ll n_\mathrm{src}$
2. $\quad$ **for** $k = 0, 1, \cdots$
3. $\quad\quad$ **Randomly choose shot subsets** $\mathcal{I} \in [1, \cdots, n_\mathrm{src}]$, $|\mathcal{I}| = n'_\mathrm{src}$
4. $\quad\quad$ $\mathbf{A}_k = \{\nabla F_i(\mathbf{m}_\mathrm{s})\mathbf{C}^\top\}_{i \in \mathcal{I}}$
5. $\quad\quad$ $\mathbf{b}_k = \{\delta \mathbf{d}_i\}_{i \in \mathcal{I}}$
6. $\quad\quad$ $\mathbf{z}_{k+1} = \mathbf{z}_k - t_k \mathbf{A}_k^\top \mathcal{P}_\sigma(\mathbf{A}_k \mathbf{x}_k - \mathbf{b}_k)$
7. $\quad\quad$ $\mathbf{x}_{k+1} = S_{\lambda_1}(\mathbf{z}_{k+1})$
8. $\quad$ **end**

note: $S_{\lambda_1}(\mathbf{z}_{k+1}) = \mathrm{sign}(\mathbf{z}_{k+1}) \max\{0, |\mathbf{z}_{k+1}| - \lambda_1\}$

$\mathcal{P}_\sigma(\mathbf{A}_k \mathbf{x}_k - \mathbf{b}_k) = \max\{0, 1 - \frac{\sigma}{\|\mathbf{A}_k \mathbf{x}_k - \mathbf{b}_k\|}\} \cdot (\mathbf{A}_k \mathbf{x}_k - \mathbf{b}_k)$

Time domain LS-RTM

Unknown source function:

$$\min_{\mathbf{x},\mathbf{q}} \lambda_1 \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2$$

$$\text{s.t.} \quad \frac{1}{2|\mathcal{I}|}\sum_{i\in\mathcal{I}} \|\nabla F_i(\mathbf{m}_{\mathrm{s}},\mathbf{q})\mathbf{C}^\top\mathbf{x} - \delta\mathbf{d}_i\|^2 \leq \sigma$$

$$\begin{aligned}\mathbf{q} \xrightarrow{\textit{linear}} & \nabla F_i(\mathbf{m}_{\mathrm{s}},\mathbf{q}) \\ & = -\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}(\nabla\mathbf{A}(\mathbf{m})\mathbf{A}(\mathbf{m}_{\mathrm{s}})^{-1}\mathbf{q})\end{aligned}$$

Subproblem:

$$\min_{\mathbf{q}} \frac{1}{2|\mathcal{I}|}\sum_{i\in\mathcal{I}} \|\nabla F_i(\mathbf{m}_{\mathrm{s}},\mathbf{q})\mathbf{C}^\top\mathbf{x} - \delta\mathbf{d}_i\|^2$$

(Linear data fitting problem)

Cons:

1. Expensive;

2. Unstable.

Time domain LS-RTM

Solutions: deconvolution + regularization

Let: $\mathbf{q} = \mathbf{w} * \mathbf{q}_0$, $\mathbf{q}_0$ is well defined.

Property 1:

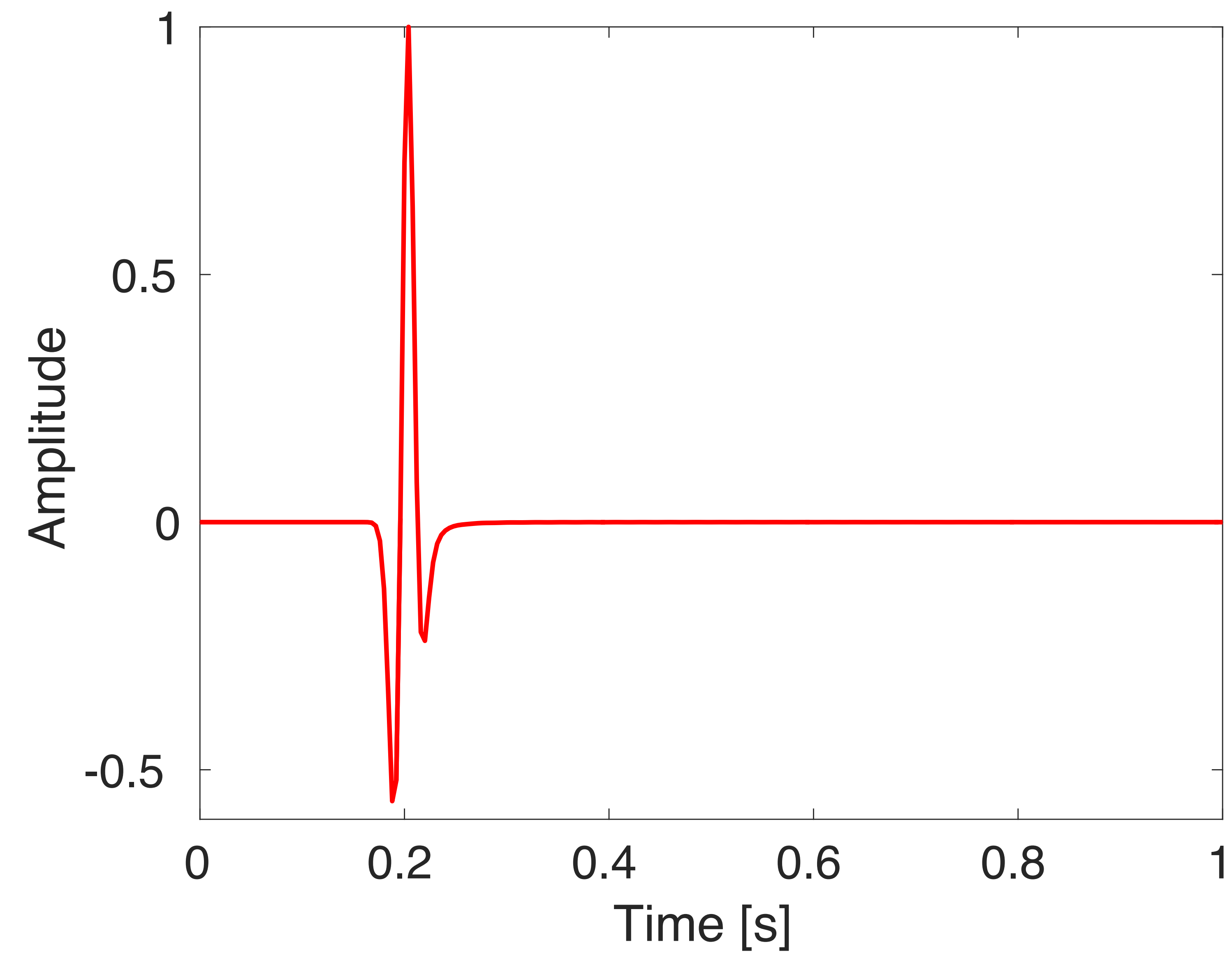
$$\nabla F_i(\mathbf{m}_\mathrm{s}, \mathbf{w} * \mathbf{q}_0)\mathbf{C}^\top \mathbf{x} \approx \mathbf{w} * \nabla F_i(\mathbf{m}_\mathrm{s}, \mathbf{q}_0)\mathbf{C}^\top \mathbf{x};$$

Property 2:

Source function should decay smoothly to zero with few oscillations within a short duration of time.

Property 3:

The energy of the source function can not explode.



Source function

17

[L. Rosasco *et al*, 2004]

Time domain LS-RTM

New subproblem:

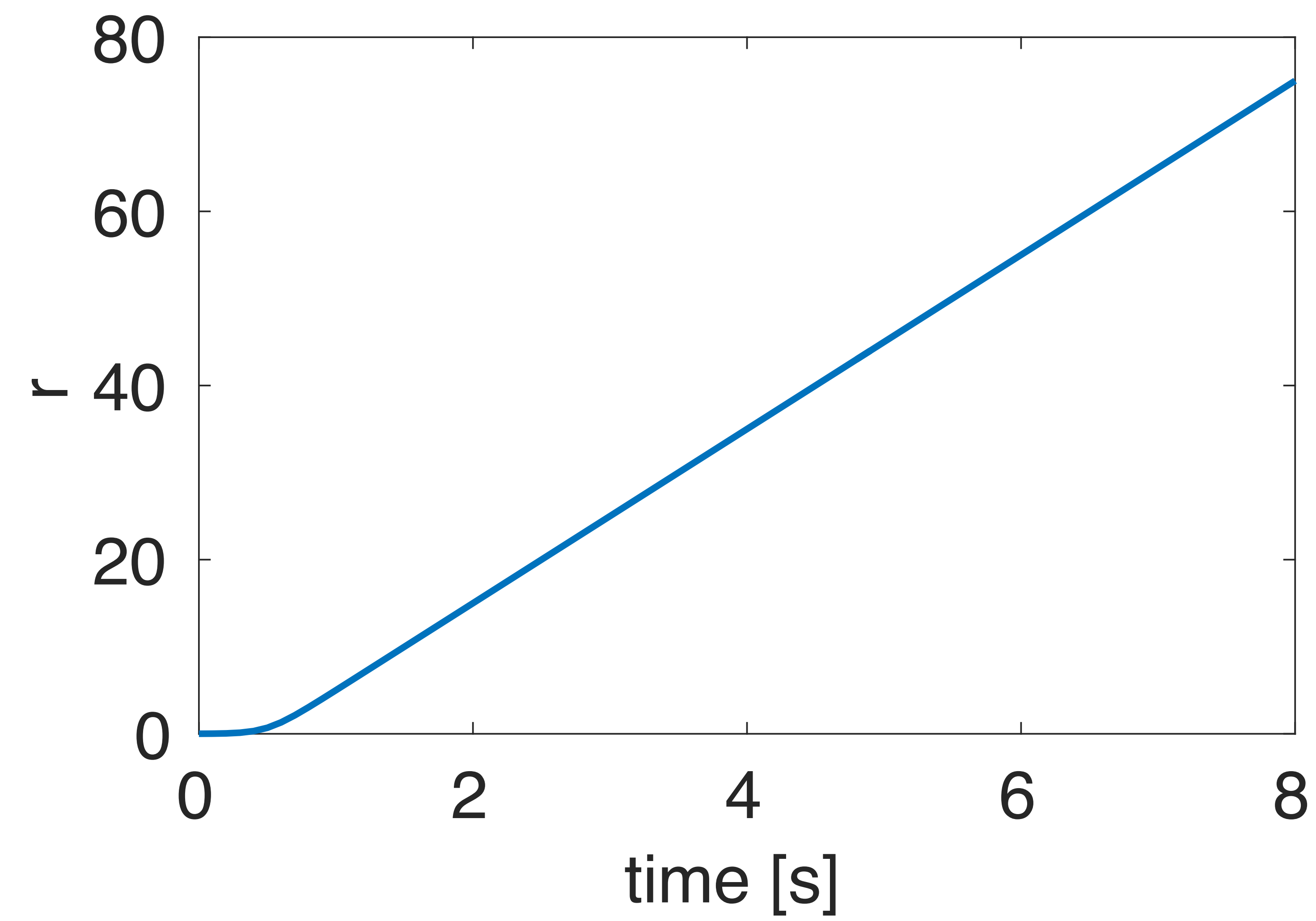
$$\min_{\mathbf{w}} \frac{1}{2|\mathcal{I}|} \sum_{i \in \mathcal{I}} \|\mathbf{w} * \nabla F_i(\mathbf{m}, \mathbf{q}_0) - \delta \mathbf{d}_i\|^2$$
$$+ \|\mathrm{diag}(\mathbf{r})(\mathbf{w} * \mathbf{q}_0)\|^2 + \lambda_2 \|\mathbf{w} * \mathbf{q}_0\|^2.$$

Penalty function:

$$\mathbf{r}(t) = \log(1 + e^{\alpha(t - t_0)})$$

t_0: the approximate duration of the source signature

α :controls the steepness of weight curve



Logistic loss function

Workflow for sparsity LS-RTM via LB with source estimation

1. **Initialize** $\mathbf{x}_0 = \mathbf{0}$, $\mathbf{z}_0 = \mathbf{0}$, $\mathbf{q}_0$, λ_1, λ_2, **batchsize** $n'_{\mathrm{src}} \ll n_{\mathrm{src}}$, **weights** $\mathbf{r}$
2. **for** $k = 0, 1, \cdots$
3. **Randomly choose shot subsets** $\mathcal{I} \in [1, \cdots, n_{\mathrm{src}}], |\mathcal{I}| = n'_{\mathrm{src}}$
4. $\mathbf{A}_k = \{\nabla F_i(\mathbf{m}_{\mathrm{s}}, \mathbf{q}_0)\mathbf{C}^\top\}_{i\in\mathcal{I}}$
5. $\mathbf{b}_k = \{\delta\mathbf{d}_i\}_{i\in\mathcal{I}}$
6. $\tilde{\mathbf{d}}_k = \mathbf{A}_k\mathbf{x}_k$
7. $\mathbf{w}_k = \arg\min_{\mathbf{w}} \|\mathbf{w} * \tilde{\mathbf{d}}_k - \mathbf{b}_k\|^2 + \|\mathrm{diag}(\mathbf{r})(\mathbf{w} * \mathbf{q}_0)\|^2 + \lambda_2\|\mathbf{w} * \mathbf{q}_0\|^2$
8. $\mathbf{z}_{k+1} = \mathbf{z}_k - t_k\mathbf{A}_k^\top\Big(\mathbf{w}_k \star P_\sigma(\mathbf{w}_k * \tilde{\mathbf{d}}_k - \mathbf{b}_k)\Big)$
9. $\mathbf{x}_{k+1} = S_\lambda(\mathbf{z}_{k+1})$
10. **end**

Experiments

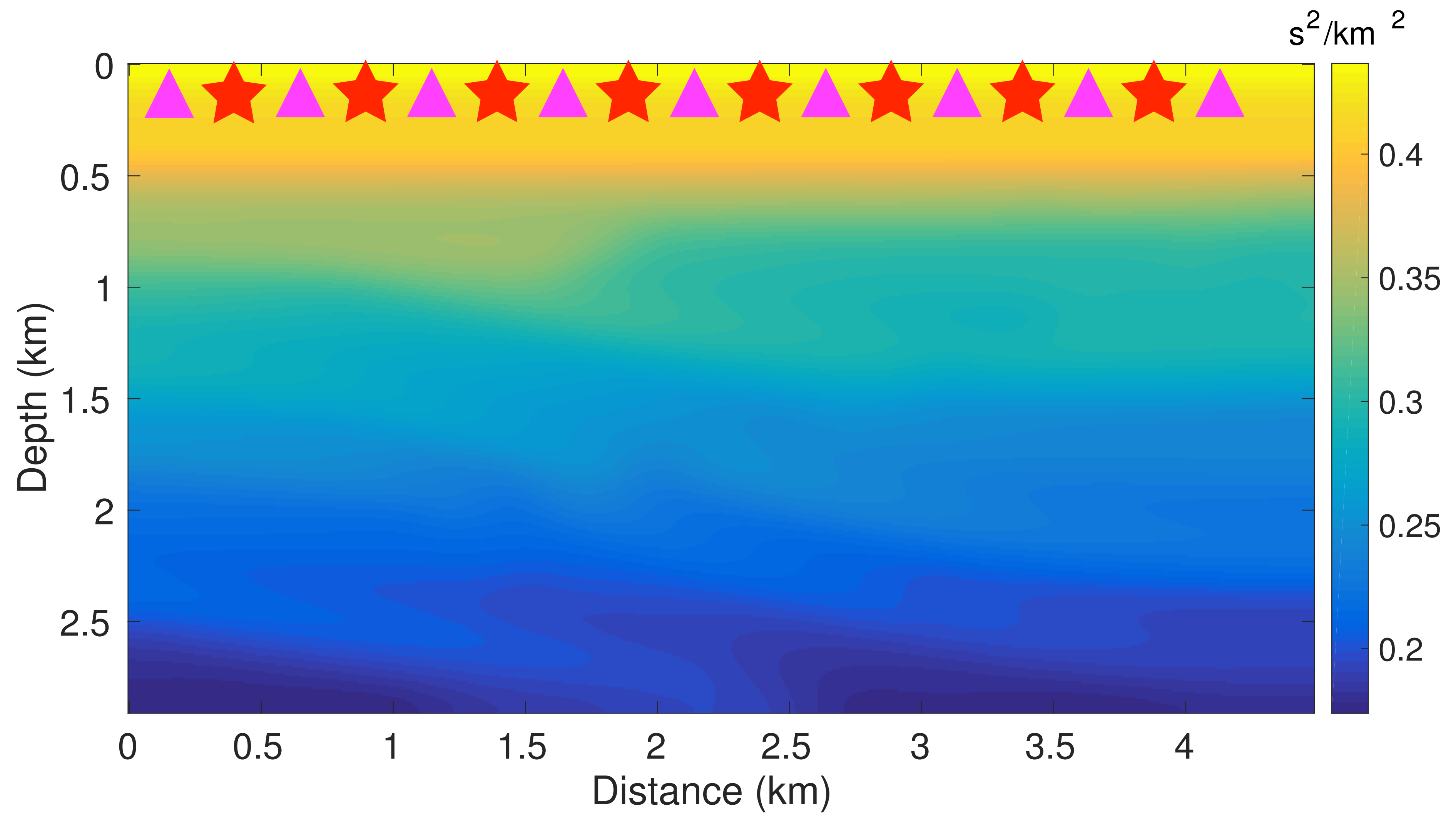
Data:

- 295 shots with shot interval 15m
- 295 receivers with receiver interval 15m
- 4s record, 15Hz peak frequency designed wavelet
- synthetic linearized data

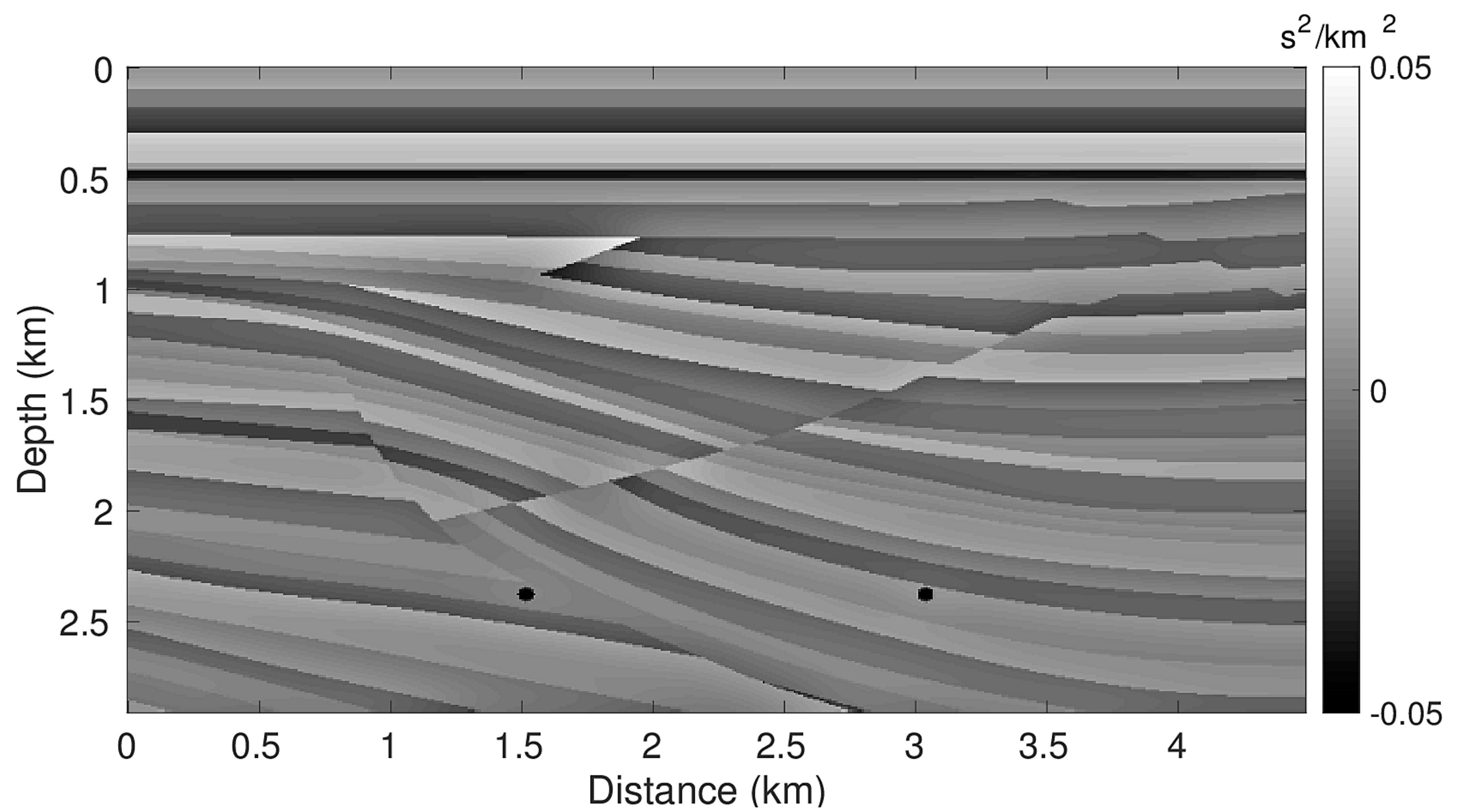
Experiments:

- one pass through the data with batch sizes 2.5% data
- randomized subset of shots
- true source wavelet & initial guessed wavelet

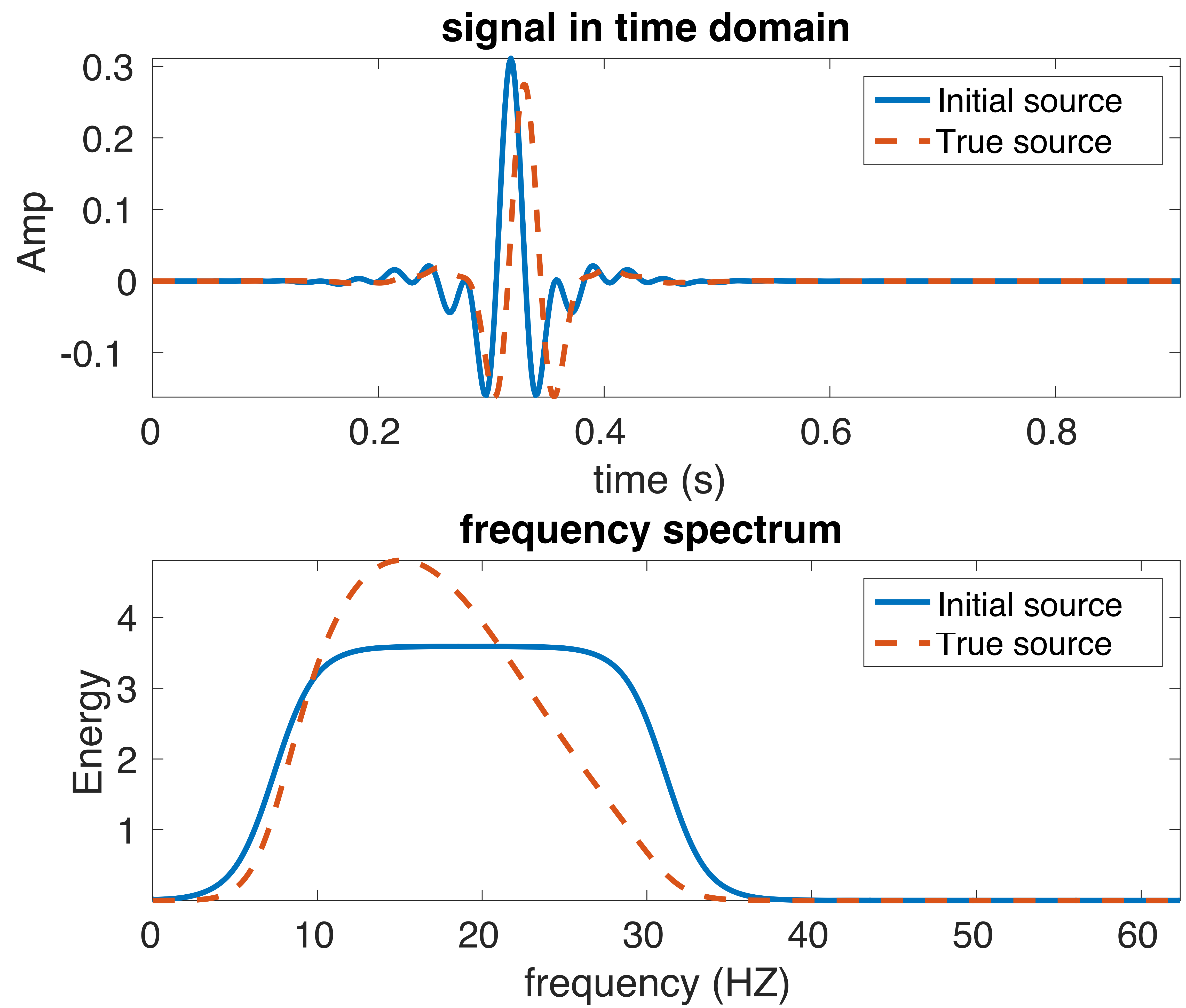
Models and source functions



(a) Background model

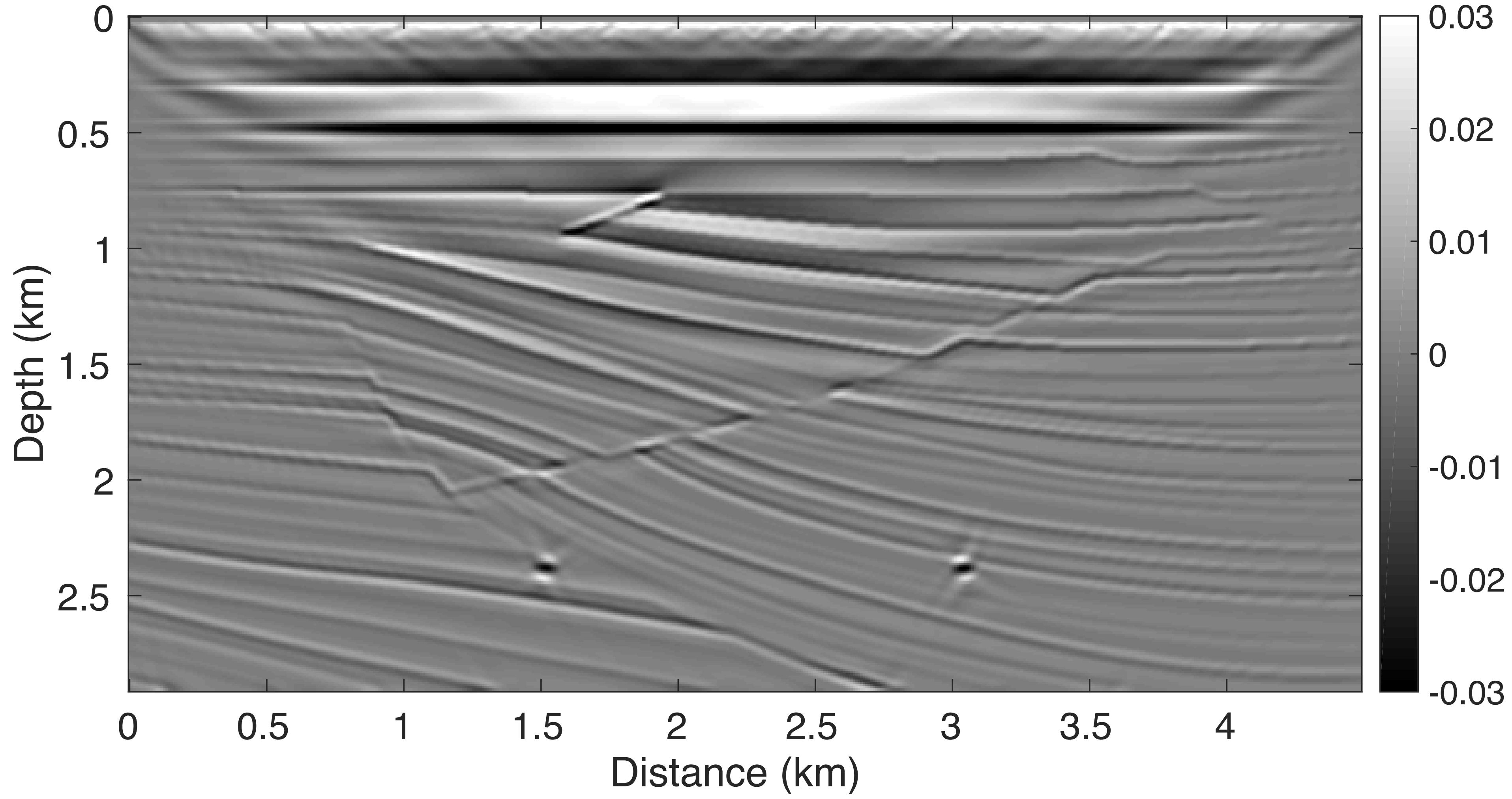


(b) True model perturbation

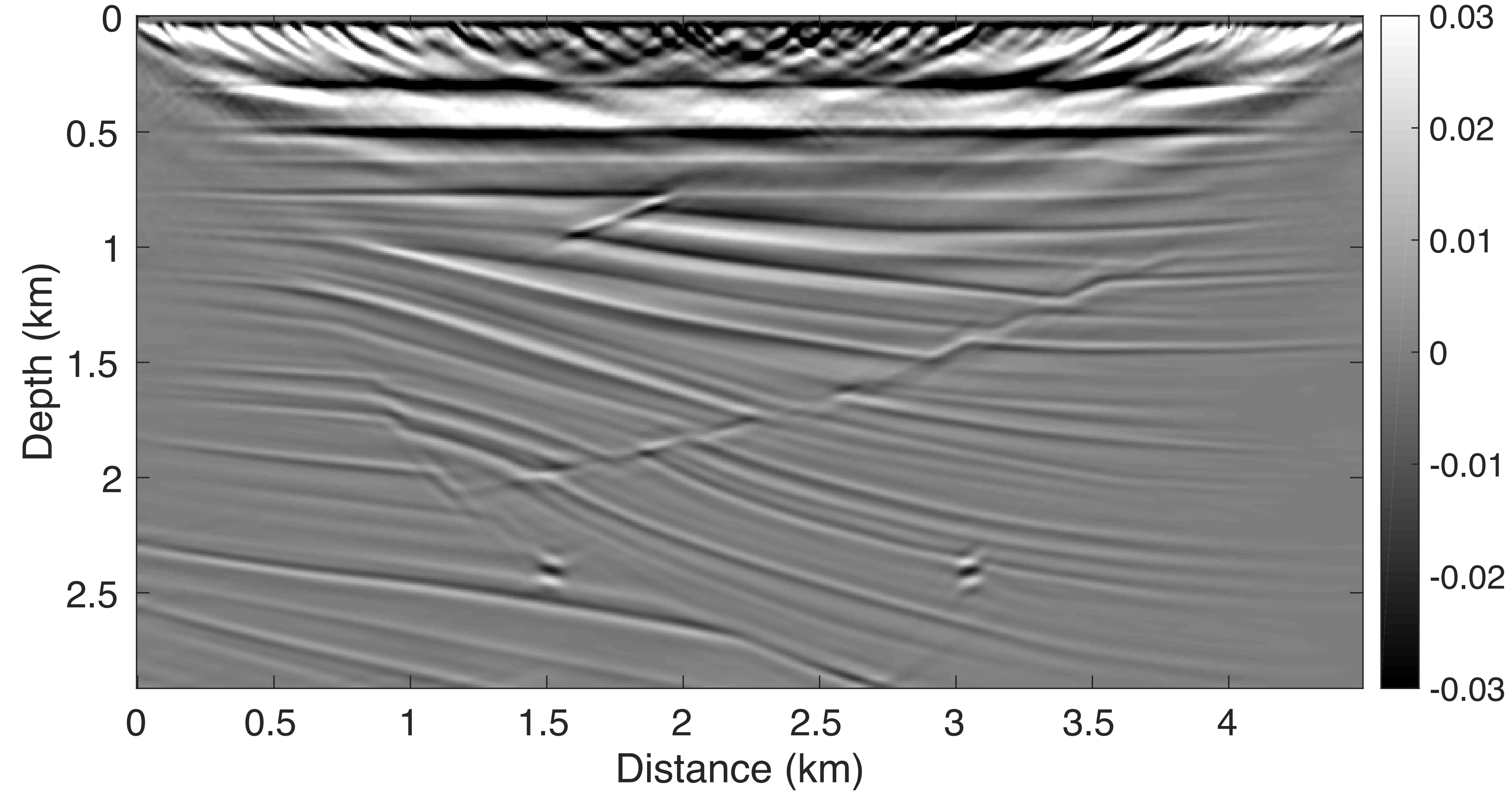


(c) Source function

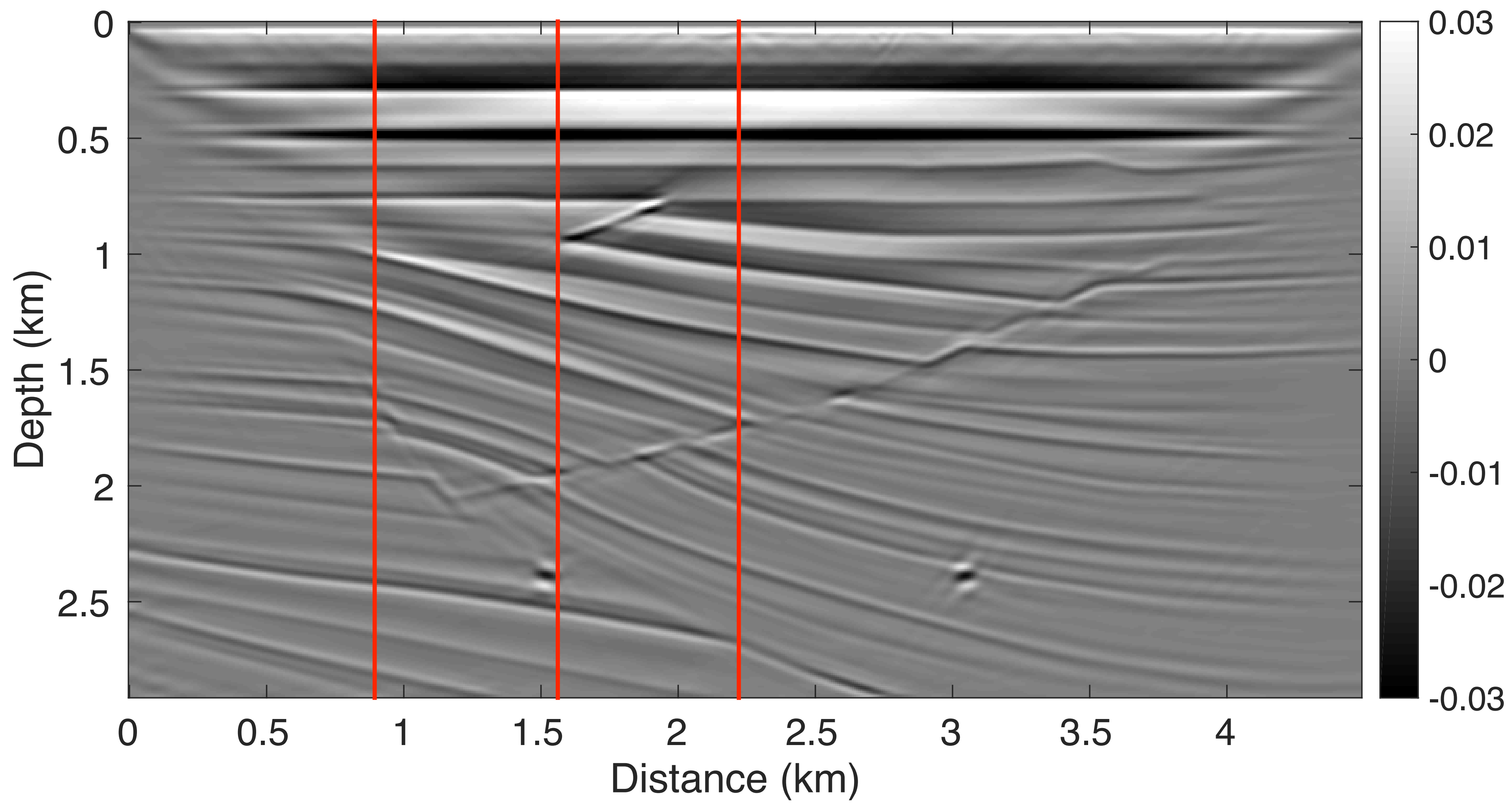
Results



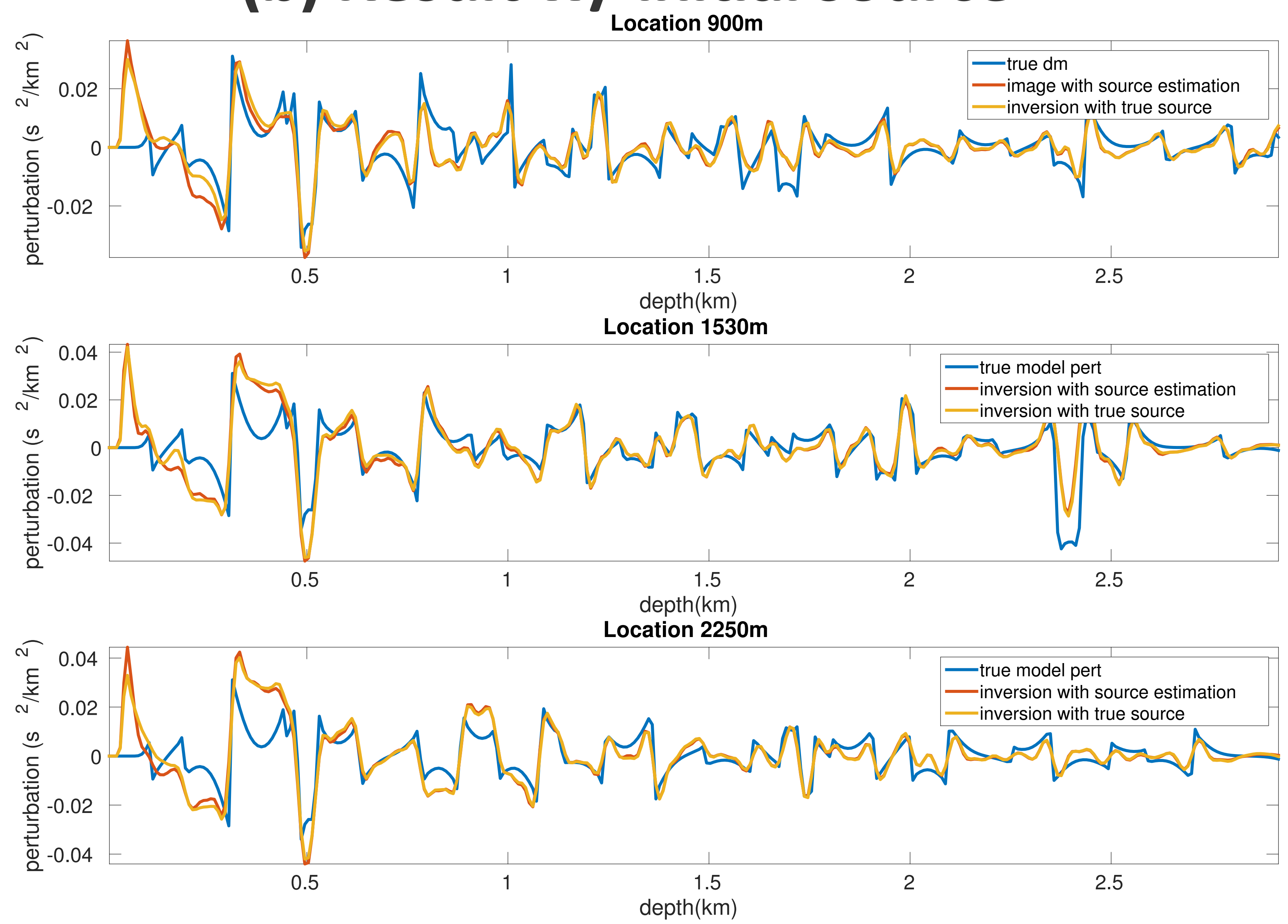
(a) Result w/ correct source



(b) Result w/ initial source

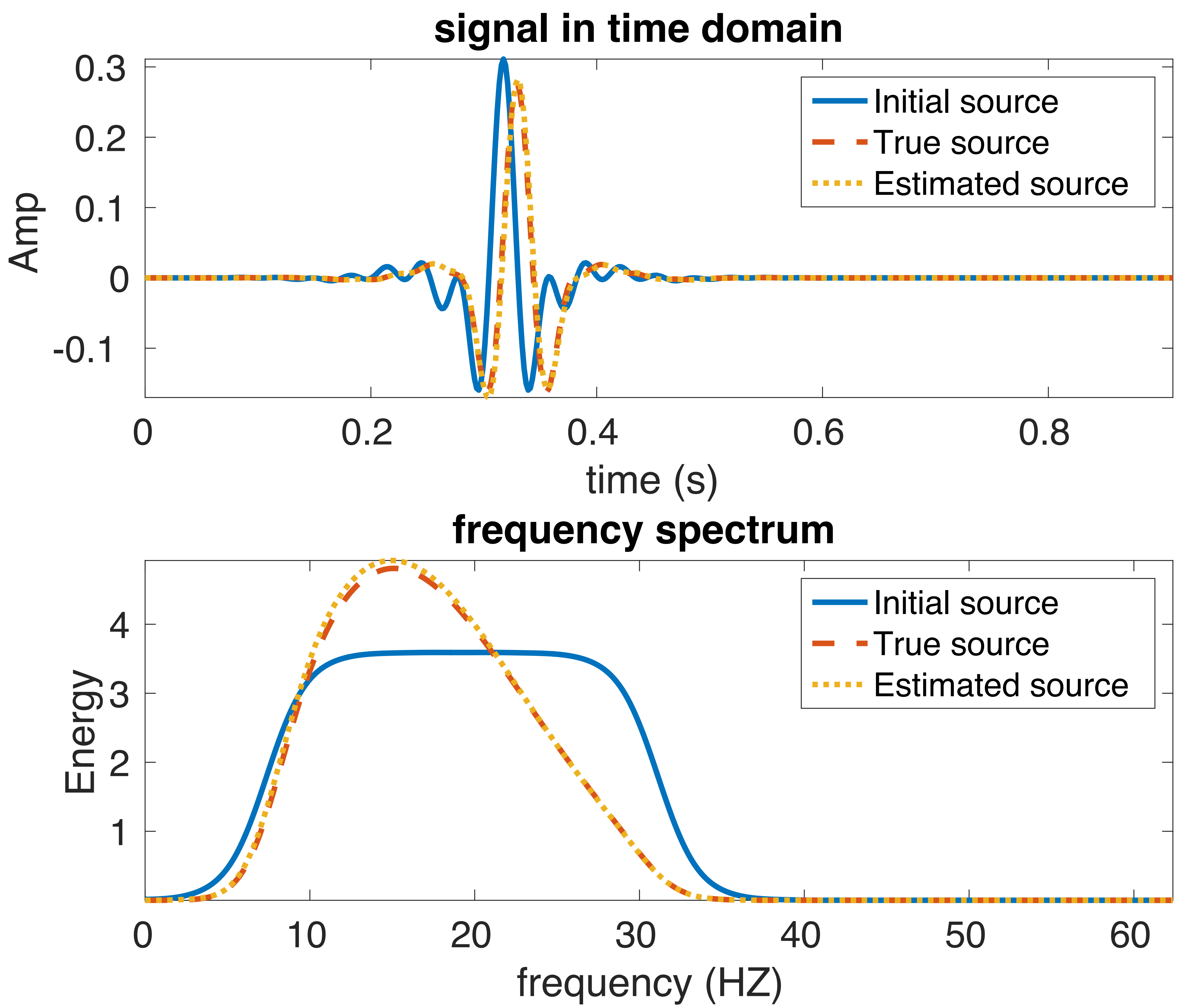


(c) Result w/ estimated source

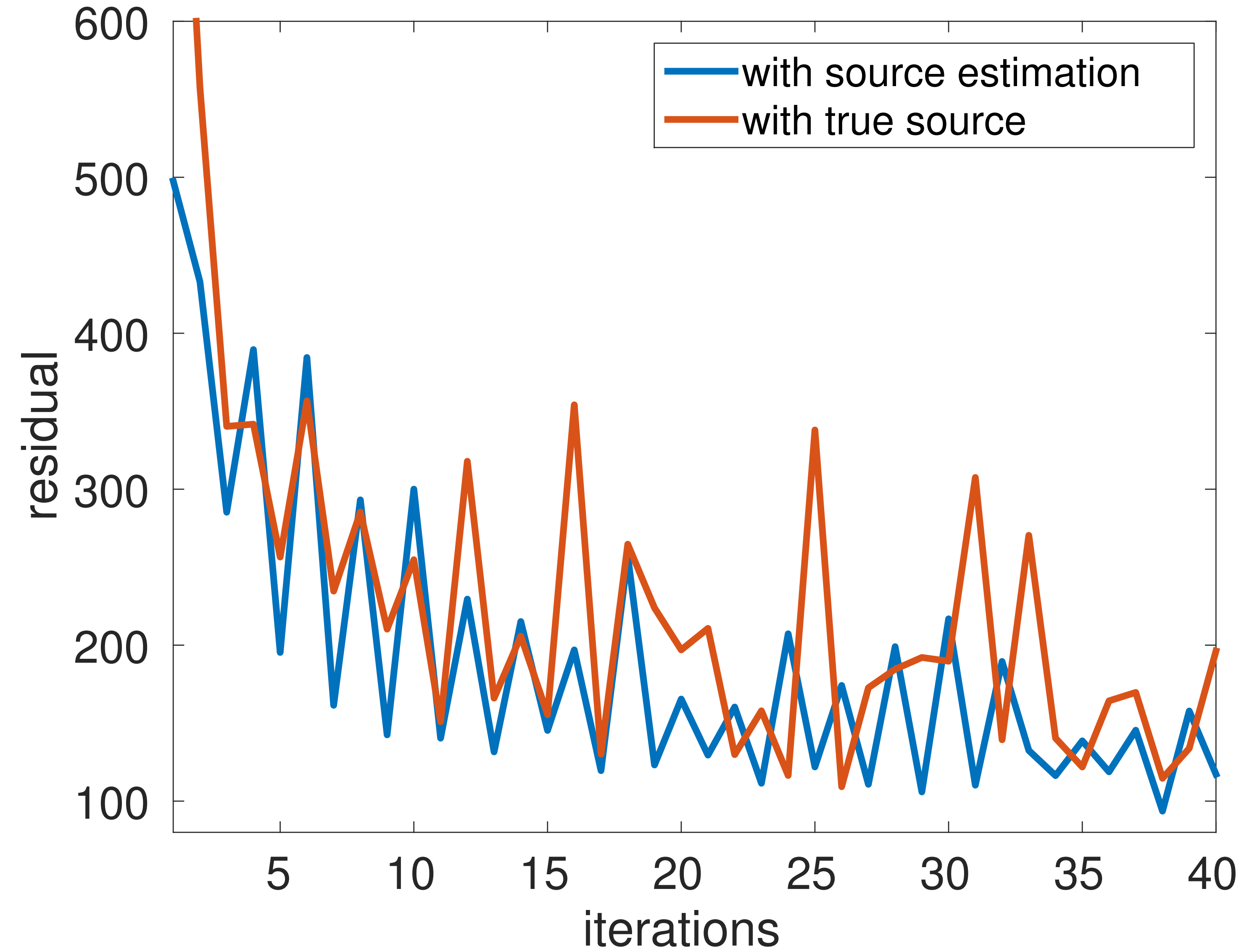


(d) Vertical trace comparisons

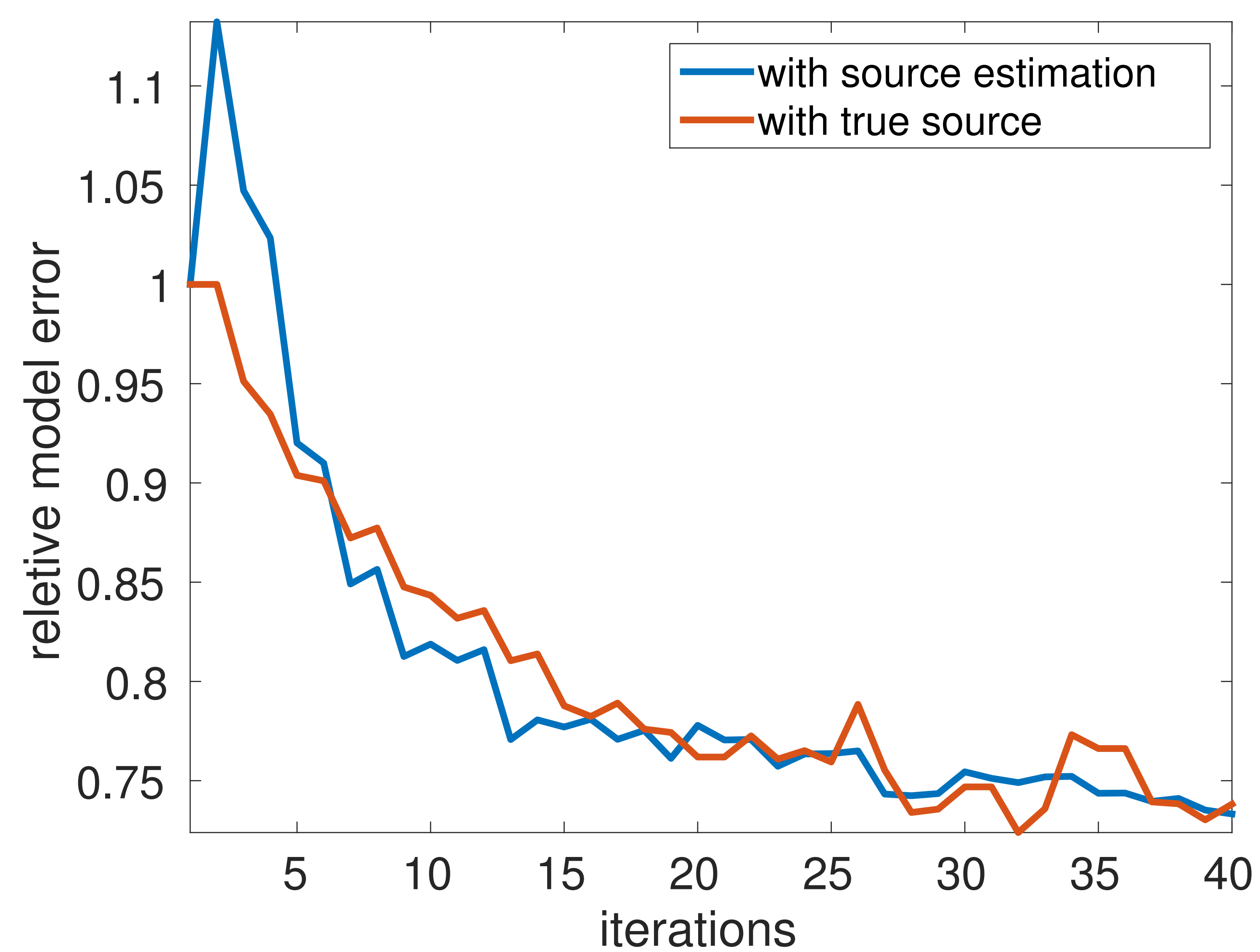
Results



(a) Source comparisons



(b) Residual decays



(c) Relative model error decays

Chapters 3,4,5

Source estimation for wavefield reconstruction inversion

Full-waveform inversion (FWI)

[Tarantola, 1984] [Peters, B, Herrmann, F J and van Leeuwen, T, 2014]
[J. Virieux and S. Operto, 2009] [Golub, G H and Pereyra, V , 1973]
[van Leeuwen, T and Herrmann, F J , 2013]

SLIM

Original problem:

$$\min_{\mathbf{u},\mathbf{m}} \frac{1}{2}\sum_{k,l} \|\mathbf{P}_k\mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2$$

$$\text{subject to } \mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} = \mathbf{q}_{k,l},$$

where,

$\mathbf{u}_{k,l}$ – Wavefield of the kth shot at lth frequency
$\mathbf{d}_{k,l}$ – Observed data of the kth shot at lth frequency
$\mathbf{q}_{k,l}$ – Source of the kth shot at lth frequency
$\mathbf{A}_{k,l}$ – Helmholtz of the kth shot at lth frequency
$\mathbf{P}_k$ – Receiver projection operator of the kth shot
$\mathbf{m}$ – Squared-slowness

Reduced/adjoint-state method:

$$\min_{\mathbf{u},\mathbf{m}} \frac{1}{2}\sum_{k,l} \|\mathbf{P}_k\mathbf{A}_{k,l}(\mathbf{m})^{-1}\mathbf{q}_{k,l} - \mathbf{d}_{k,l}\|_2^2$$

with the gradient given by

$$\mathbf{g} = \sum_{k,l} \mathbf{u}_{k,l}^\top \frac{\partial \mathbf{A}_{k,l}^\top}{\partial \mathbf{m}} \mathbf{v}_{k,l}$$

$$\mathbf{u}_{k,l} = \mathbf{A}_{k,l}(\mathbf{m})^{-1}\mathbf{q}_{k,l}$$

$$\mathbf{v}_{k,l} = \mathbf{A}_{k,l}^{-\top}(\mathbf{m})\mathbf{P}_k^\top\mathbf{r}_{k,l}$$

$$\mathbf{r}_{k,l} = \mathbf{P}_k\mathbf{A}_{k,l}(\mathbf{m})^{-1}\mathbf{q}_{k,l} - \mathbf{d}_{k,l}$$

2 PDE solves are required !

Wavefield reconstruction inversion (WRI):

$$\min_{\mathbf{u},\mathbf{m}} \frac{1}{2}\sum_{k,l} \left(\|\mathbf{P}_k\mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2\|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \mathbf{q}_{k,l}\|_2^2\right)$$

Eliminating $\mathbf{u}$ w/ variable projection:

$$\overline{\mathbf{u}} = \arg\min_{\mathbf{u}} \frac{1}{2}\sum_{k,l} \Big(\|\mathbf{P}_k\mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2\|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \mathbf{q}_{k,l}\|_2^2\Big)$$

1 linear system solve is required !

Analytical solution:

$$\left(\lambda^2 {\mathbf{A}_{k,l}}^\top\mathbf{A}_{k,l} + \mathbf{P}_k^\top\mathbf{P}_k\right)\overline{\mathbf{u}}_{k,l} = \lambda^2\mathbf{A}_{k,l}^\top\mathbf{q}_{k,l} + \mathbf{P}_k^\top\mathbf{d}_{k,l}$$

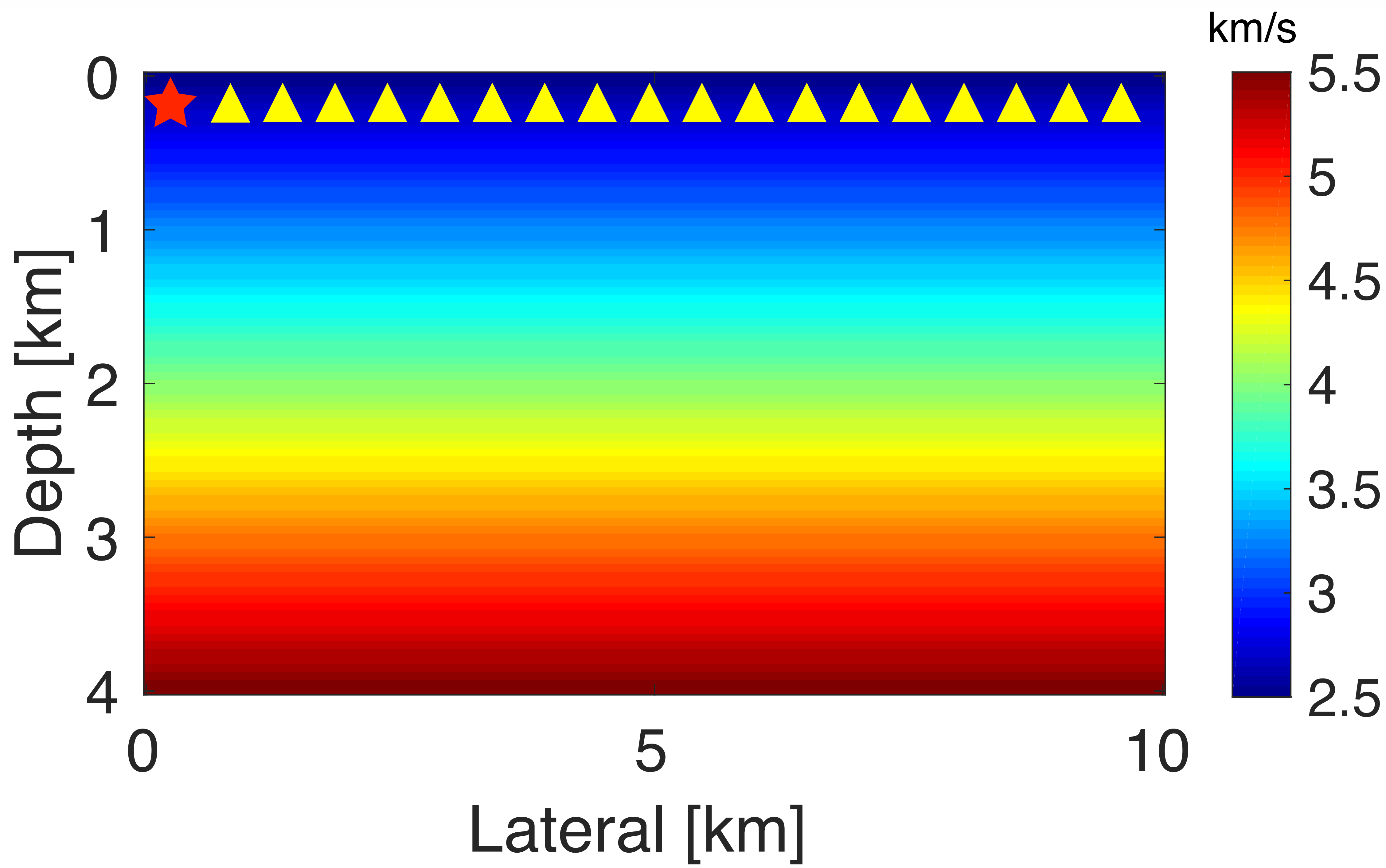
Gradient:

$$\mathbf{g} = \sum_{k,l} \overline{\mathbf{u}}_{k,l}^\top \frac{\partial \mathbf{A}_{k,l}^\top}{\partial \mathbf{m}} \overline{\mathbf{v}}_{k,l}$$

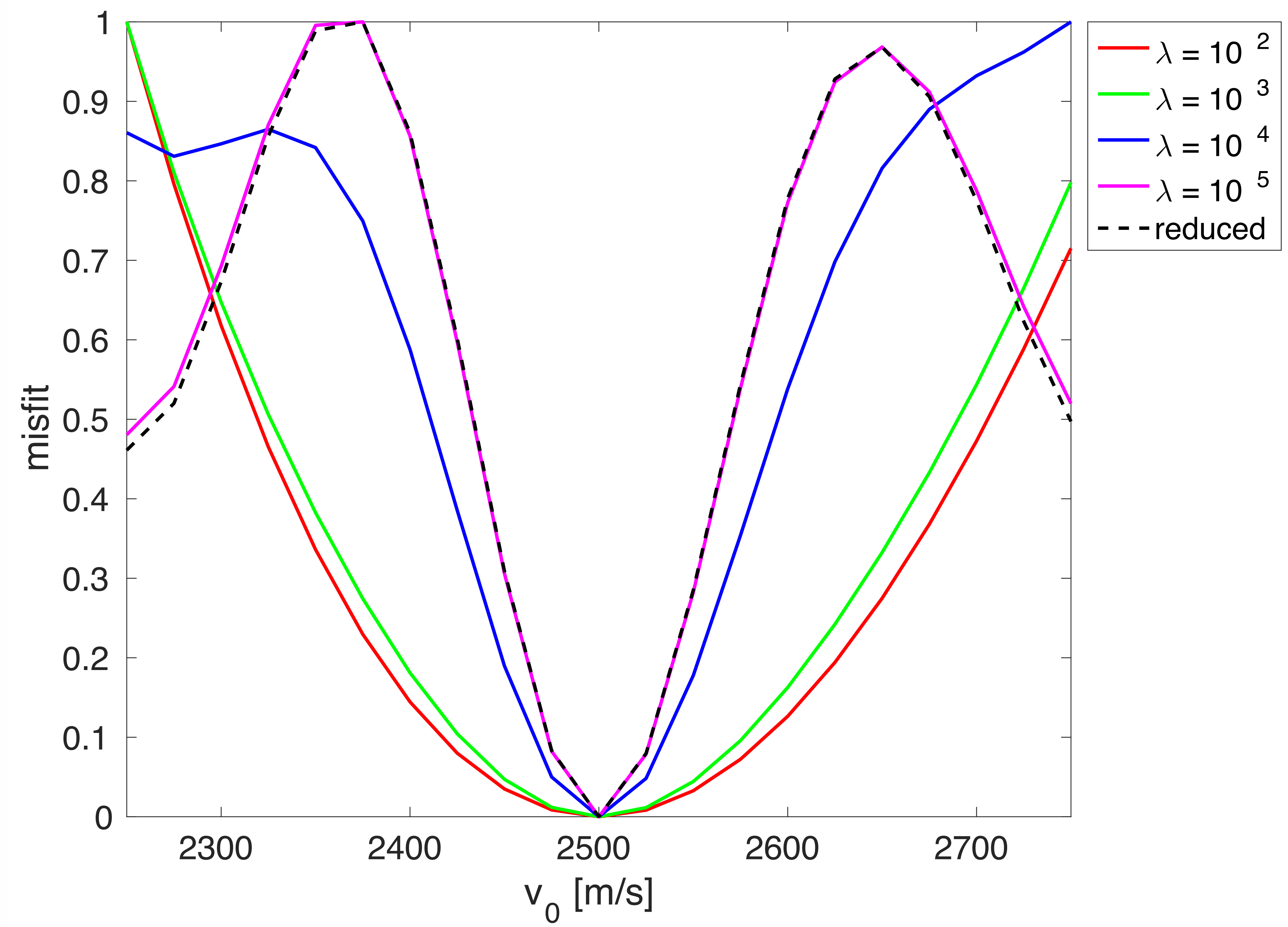
$$\overline{\mathbf{v}}_{k,l} = \mathbf{A}_{k,l}(\mathbf{m})\overline{\mathbf{u}}_{k,l} - \mathbf{q}_{k,l}$$

FWI vs WRI

1D velocity model: $v(z) = v_0 + 0.75z\,\mathrm{km/s}, \quad \text{with} \quad v_0 = 2.5\,\mathrm{km/s}.$



(a) 1D model



(b) Misfit function comparison

WRI with source estimation

Triple parameters optimization problem:

$$\min_{\mathbf{u},\mathbf{m},\alpha} \frac{1}{2} \sum_{k,l} \left(\|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2 \|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \alpha_{k,l}\mathbf{e}_{k,l}\|_2^2 \right)$$

Two solutions:

- eliminate $\mathbf{u}$ first, and update $(\mathbf{m}, \alpha)$ simultaneously (WRI-SE-MS);
  - **Con:** the amplitudes of $\mathbf{m}$ and α are different and the same for the gradients.

- simultaneously eliminate $(\mathbf{u}, \alpha)$ and then update $\mathbf{m}$ (WRI-SE-WS).
  - **Pro:** only need to update $\mathbf{m}$, so that it does not suffer from the different amplitudes.

WRI with source estimation

Eliminate $\mathbf{u}$ and α jointly w/ variable projection:

$$[\overline{\mathbf{u}}, \overline{\alpha}] = \arg\min_{\mathbf{u},\alpha} \frac{1}{2} \sum_{k,l} \left(\|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2 \|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \alpha_{k,l}\mathbf{e}_{k,l}\|_2^2 \right)$$

Analytical solution:

$$\begin{bmatrix} \overline{\mathbf{u}}_{k,l}(\mathbf{m}) \\ \overline{\alpha}_{k,l}(\mathbf{m}) \end{bmatrix} = \begin{bmatrix} \lambda^2 \mathbf{A}_{k,l}^\top \mathbf{A}_{k,l} + \mathbf{P}_k^\top \mathbf{P}_k & -\lambda^2 \mathbf{A}_{k,l}^\top \mathbf{e}_{k,l} \\ -\lambda^2 \mathbf{e}_{k,l}^\top \mathbf{A}_{k,l} & \lambda^2 \mathbf{e}_{k,l}^\top \mathbf{e}_{k,l} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{P}_k \mathbf{d}_{k,l} \\ \mathbf{0} \end{bmatrix}.$$

[D. S. Bernstein, 2005]

Fast solver

When $\mathbf{P}_k = \mathbf{P},\ \mathbf{A}_{k,l} = \mathbf{A}_l,$

1. Differ from source to source;
2. Require different factorization and preconditioning matrix.

$$\begin{bmatrix} \overline{\mathbf{u}}_{k,l}(\mathbf{m}) \\ \overline{\alpha}_{k,l}(\mathbf{m}) \end{bmatrix} = \boxed{\begin{bmatrix} \lambda^2 \mathbf{A}_l^\top \mathbf{A}_l + \mathbf{P}^\top \mathbf{P} & -\lambda^2 \mathbf{A}_l^\top \mathbf{e}_{k,l} \\ -\lambda^2 \mathbf{e}_{k,l}^\top \mathbf{A}_l & \lambda^2 \mathbf{e}_{k,l}^\top \mathbf{e}_{k,l} \end{bmatrix}}^{-1} \begin{bmatrix} \mathbf{P}^\top \mathbf{d}_{k,l} \\ \mathbf{0} \end{bmatrix}.$$

Solution —— block matrix inversion formula:

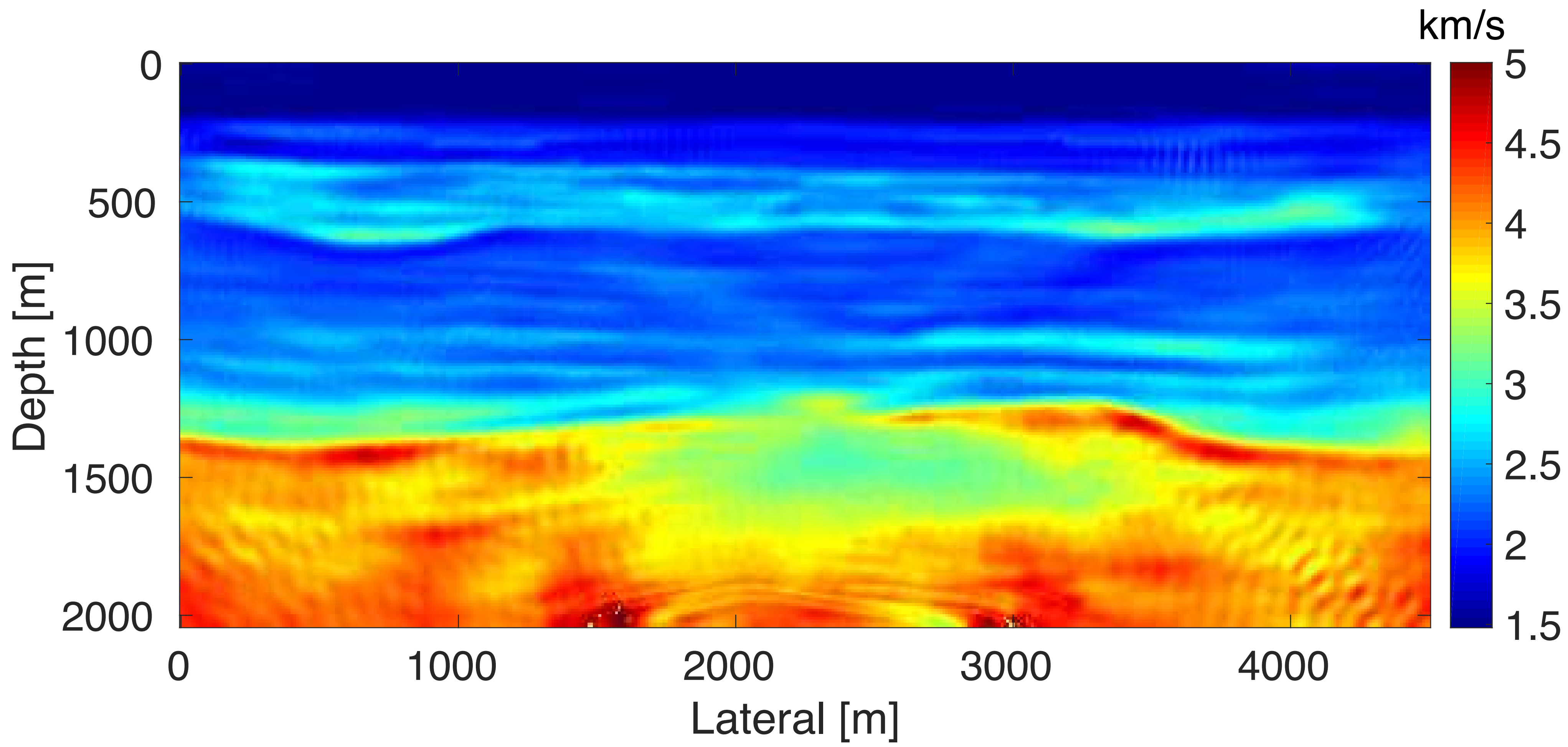
$$\mathbf{C}_{k,l}(\mathbf{m}) = \begin{bmatrix} \boxed{\lambda^2 \mathbf{A}_l^\top \mathbf{A}_l + \mathbf{P}^\top \mathbf{P}} & -\lambda^2 \mathbf{A}_l^\top \mathbf{e}_{k,l} \\ -\lambda^2 \mathbf{e}_{k,l}^\top \mathbf{A}_l & \lambda^2 \mathbf{e}_{k,l}^\top \mathbf{e}_{k,l} \end{bmatrix} = \begin{bmatrix} \boxed{\mathbf{M}_1} & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix}$$

Analytical solution:

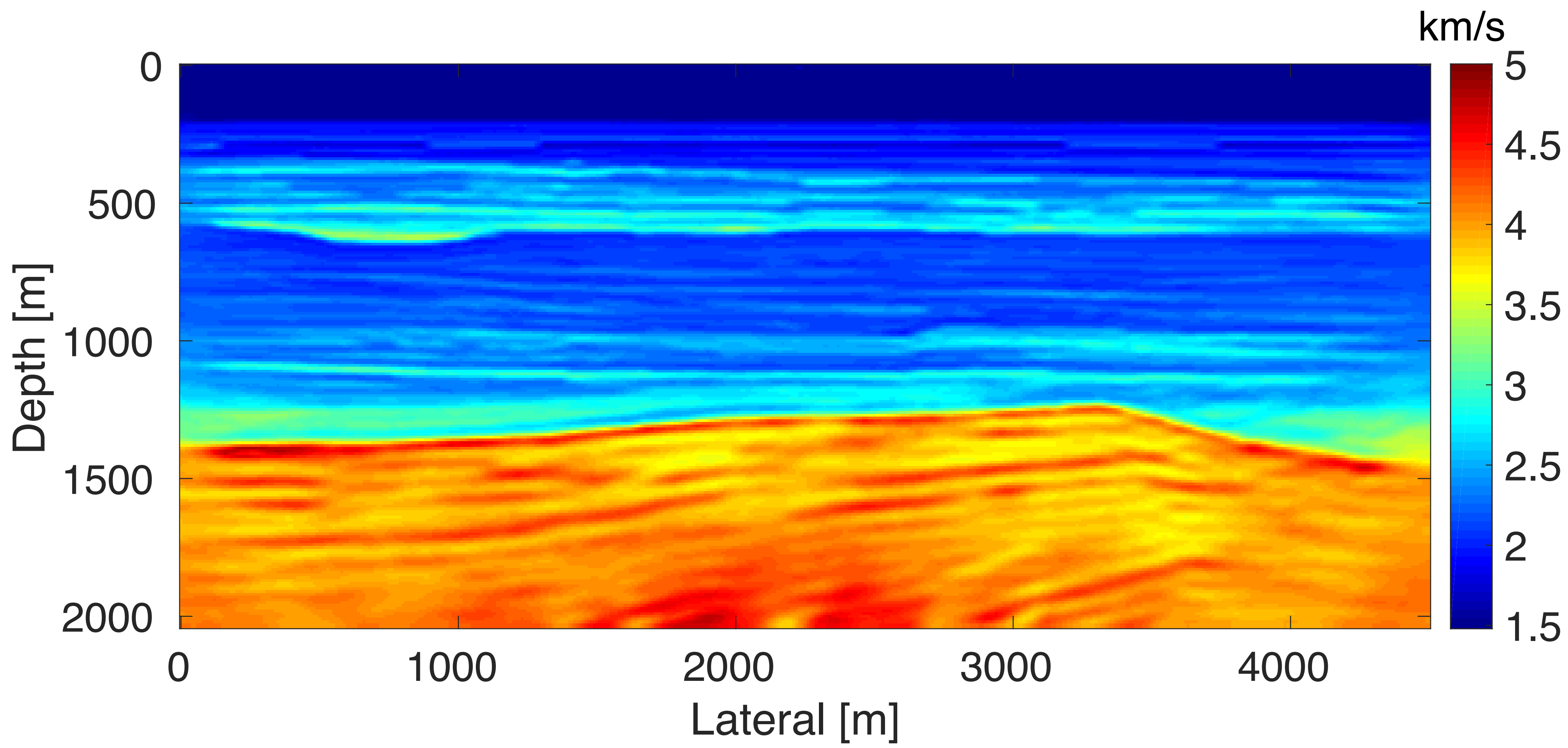
$$\begin{aligned} \overline{\mathbf{x}}_{k,l}(\mathbf{m}) &= \mathbf{C}_{k,l}^{-1}(\mathbf{m}) \begin{bmatrix} \mathbf{P}^\top \mathbf{d}_{i,j} \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{I} + \boxed{\mathbf{M}_1^{-1}\mathbf{M}_2}(\mathbf{M}_4 - \mathbf{M}_3\mathbf{M}_1^{-1}\mathbf{M}_2)^{-1}\mathbf{M}_3)\boxed{\mathbf{M}_1^{-1}\mathbf{P}^\top \mathbf{d}_{i,j}} \\ -(\mathbf{M}_4 - \mathbf{M}_3\mathbf{M}_1^{-1}\mathbf{M}_2)^{-1}\mathbf{M}_3\mathbf{M}_1^{-1}\mathbf{P}^\top \mathbf{d}_{i,j} \end{bmatrix}. \end{aligned}$$

Example on BG Compass model

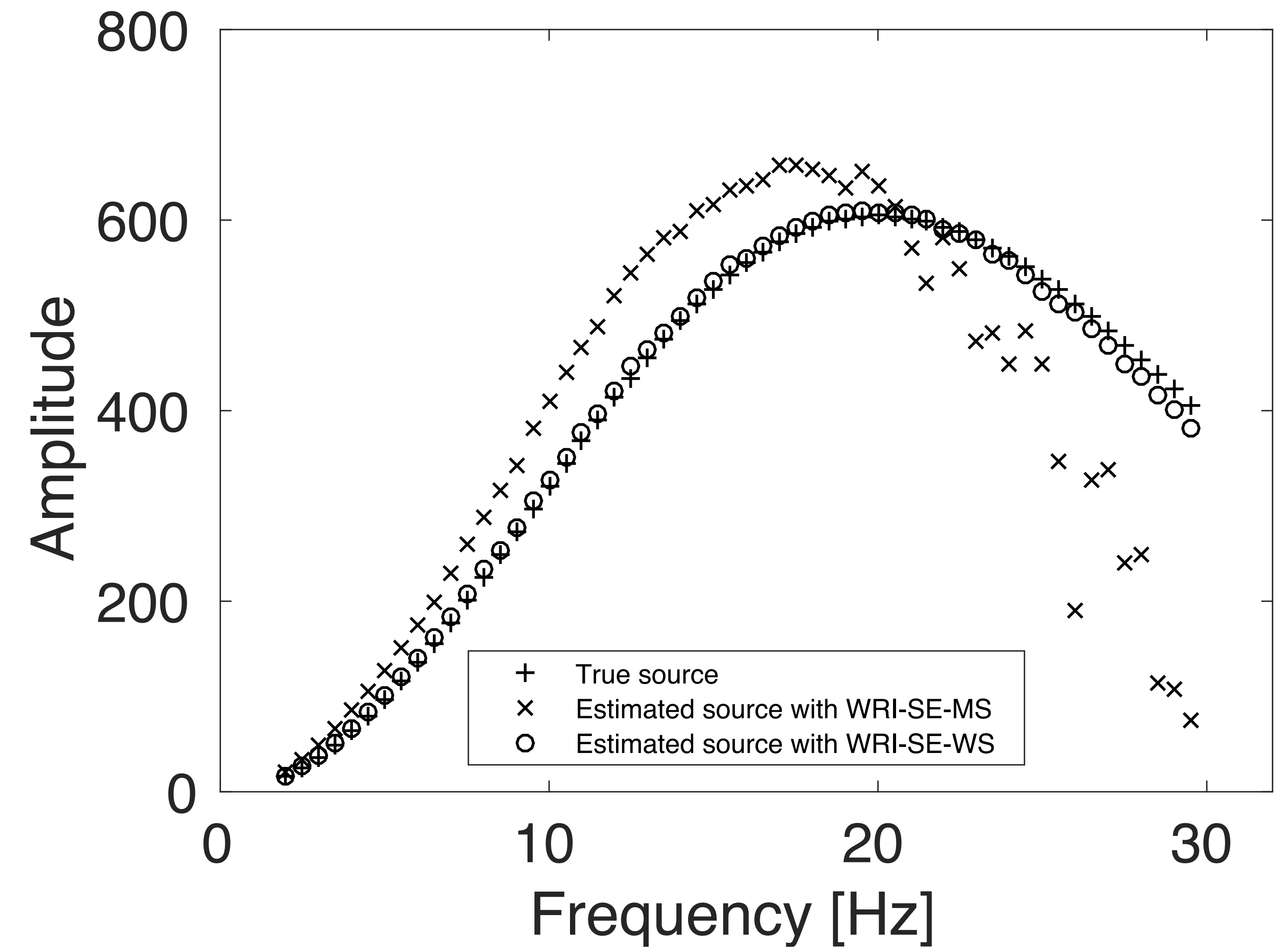
—— WRI-SE-WS vs WRI-SE-MS



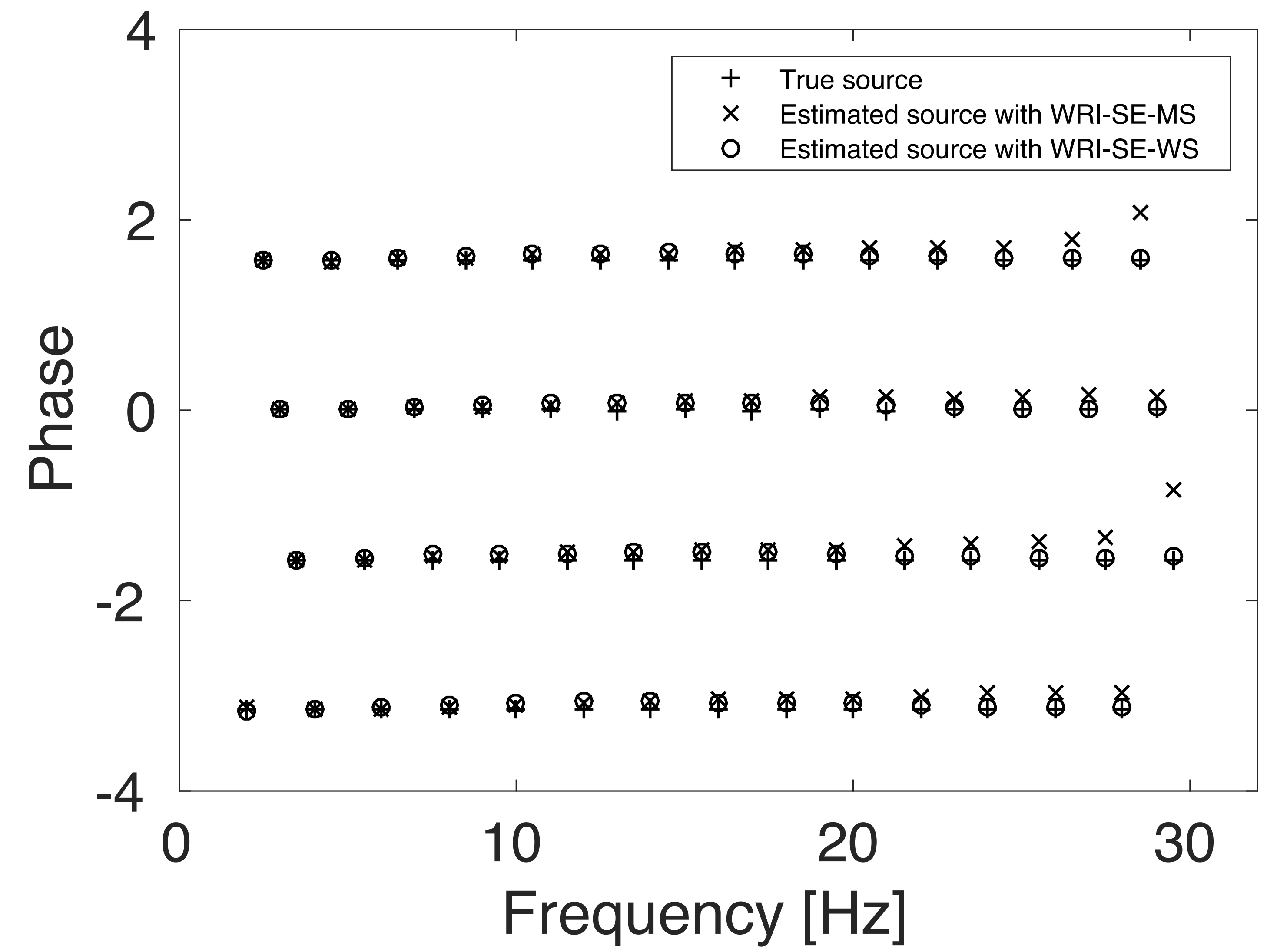
(a) WRI-SE-MS



(b) WRI-SE-WS

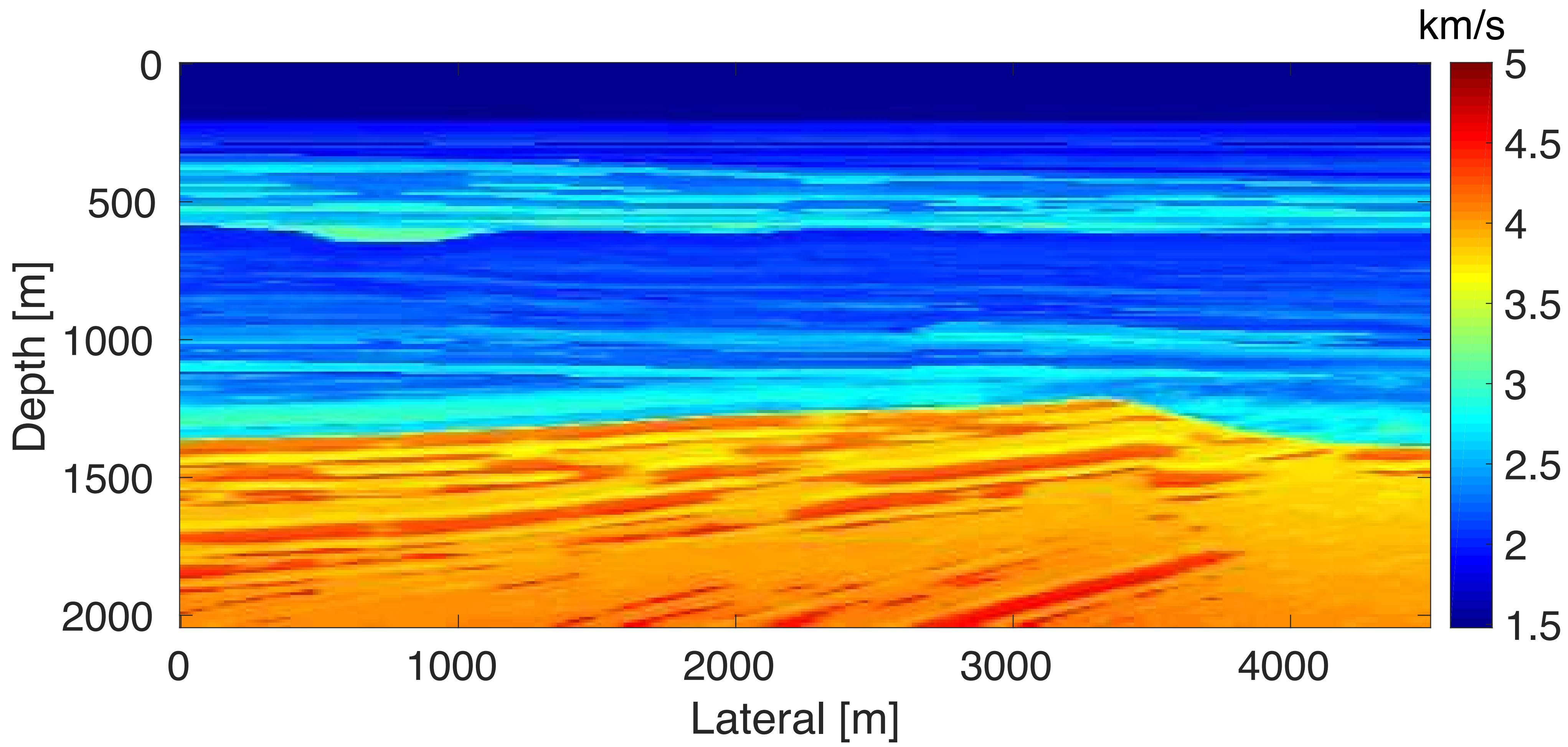


(c) Amplitude

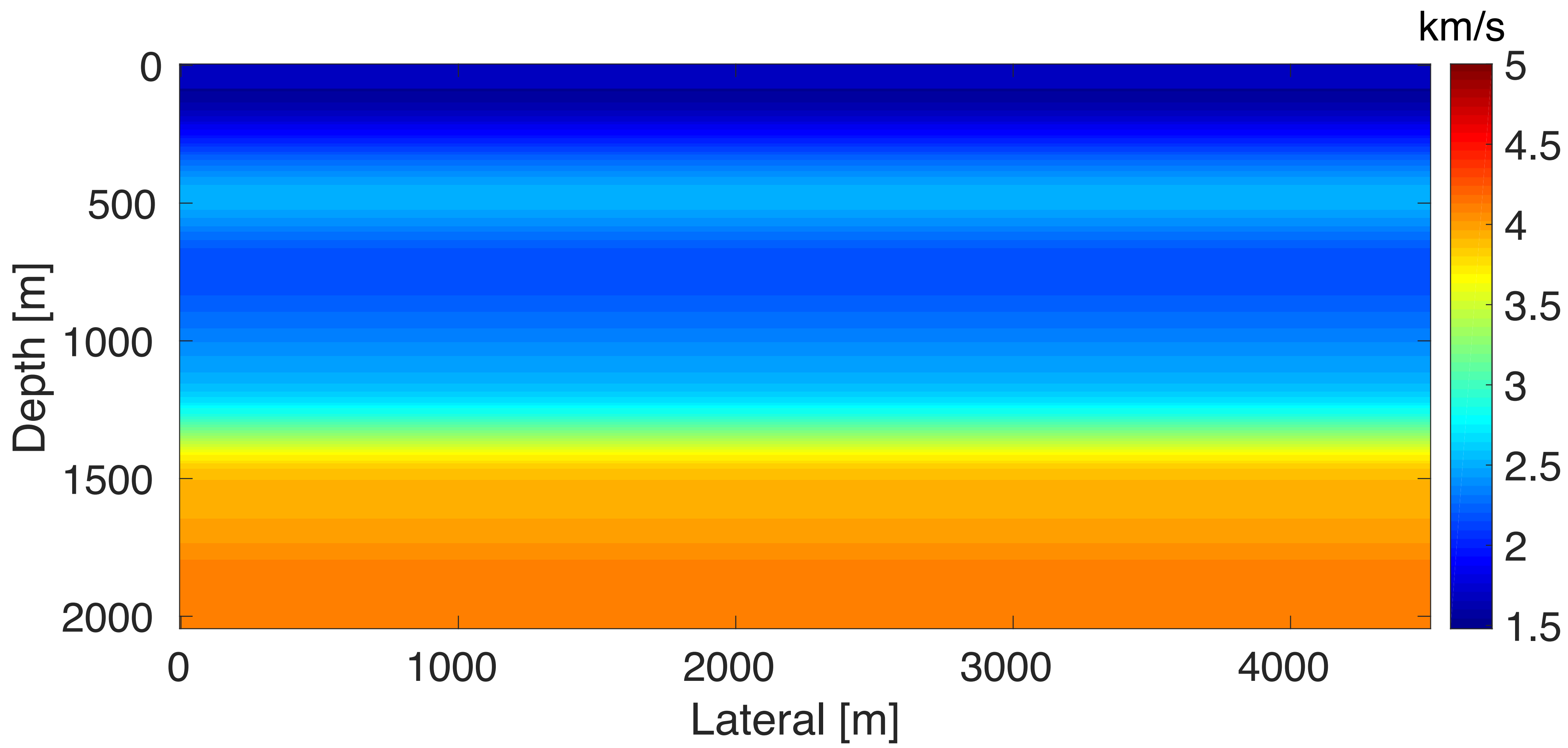


(d) Phase

Example on BG Compass model
—— WRI-SE-WS vs FWI-SE



(a) True model



(b) Bad initial model

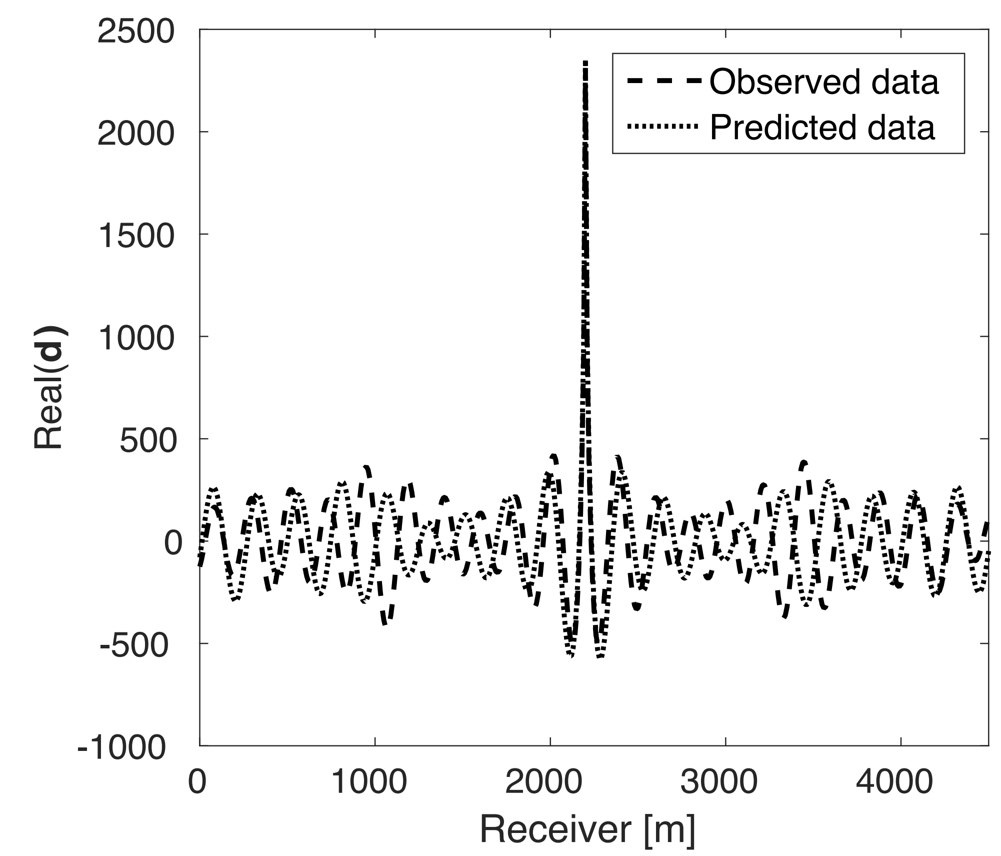
Inversion information:

Frequency : 7~30 Hz

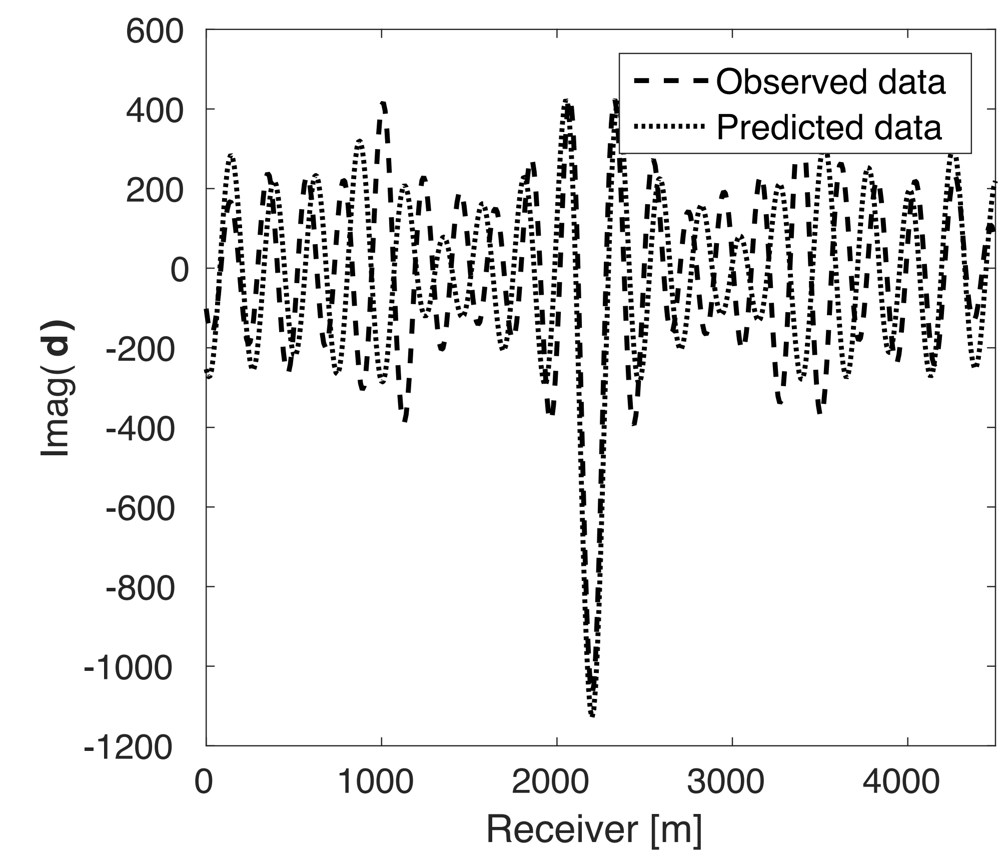
Optimization Solver: l-BFGS

Iterations per frequency band: 20

Penalty parameter λ **:** 1e0



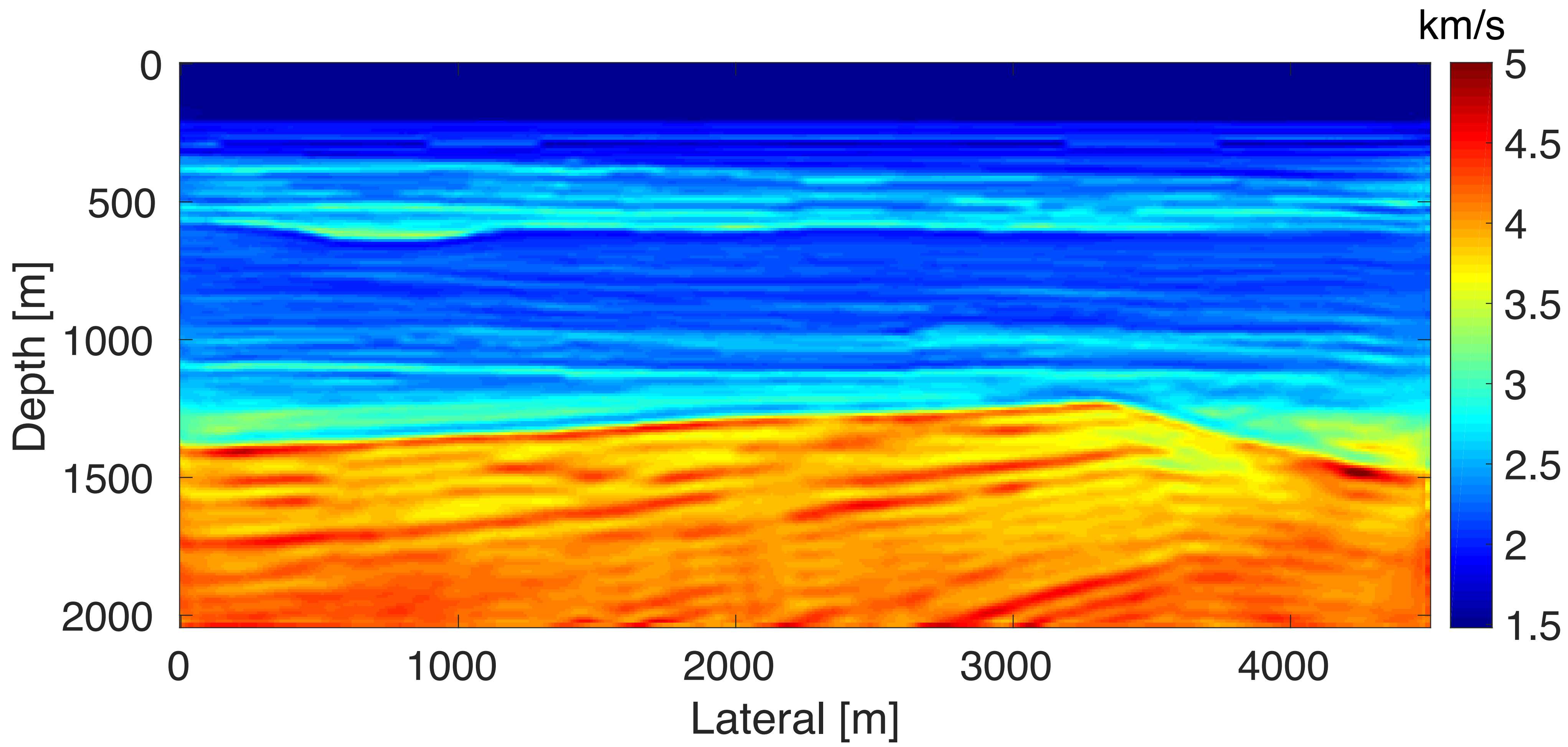
(c) Real part



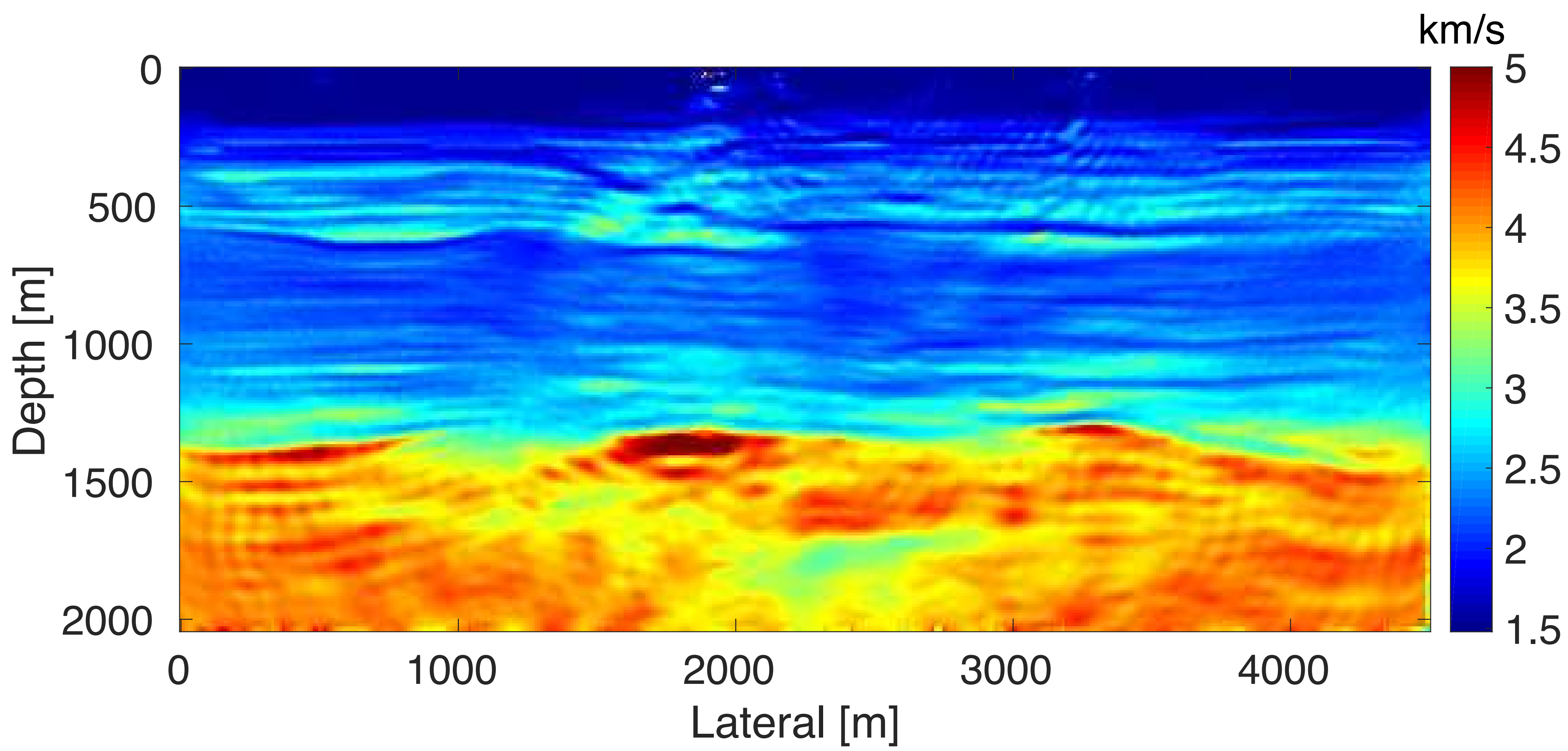
(d) Imaginary part

Example on BG Compass model

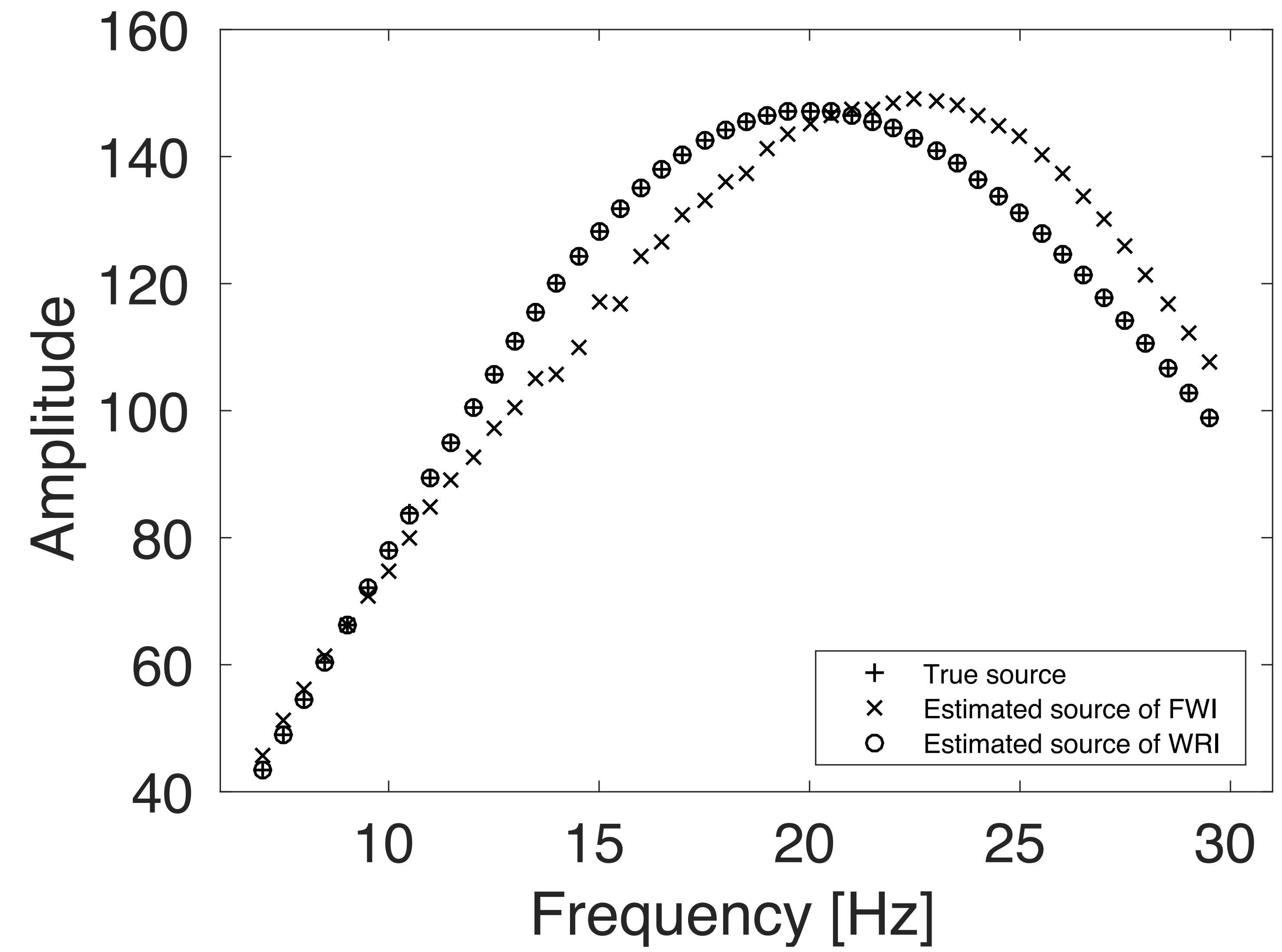
—— WRI-SE-WS vs FWI-SE



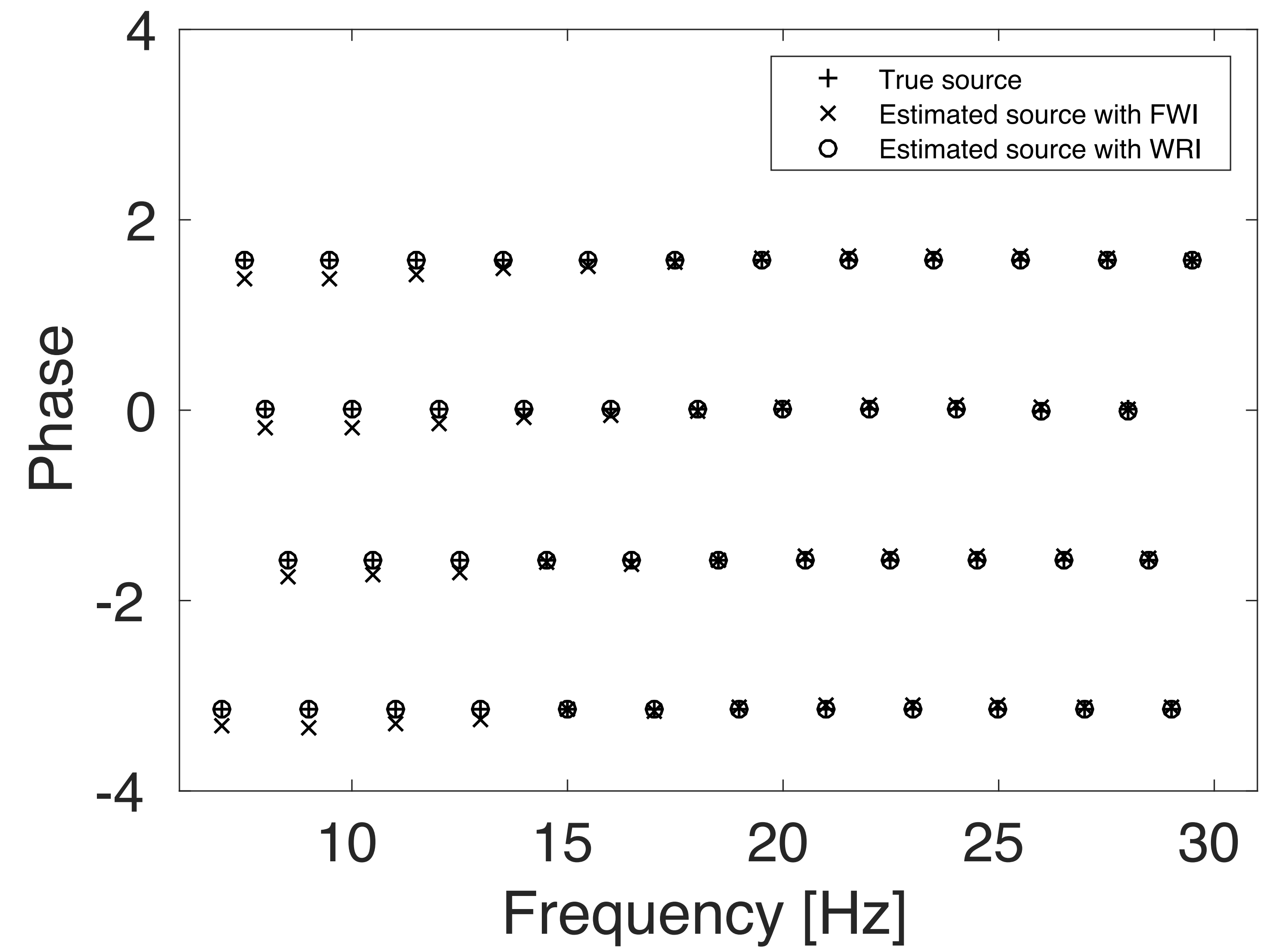
(a) WRI-SE-WS



(b) FWI-SE



(c) Amplitude

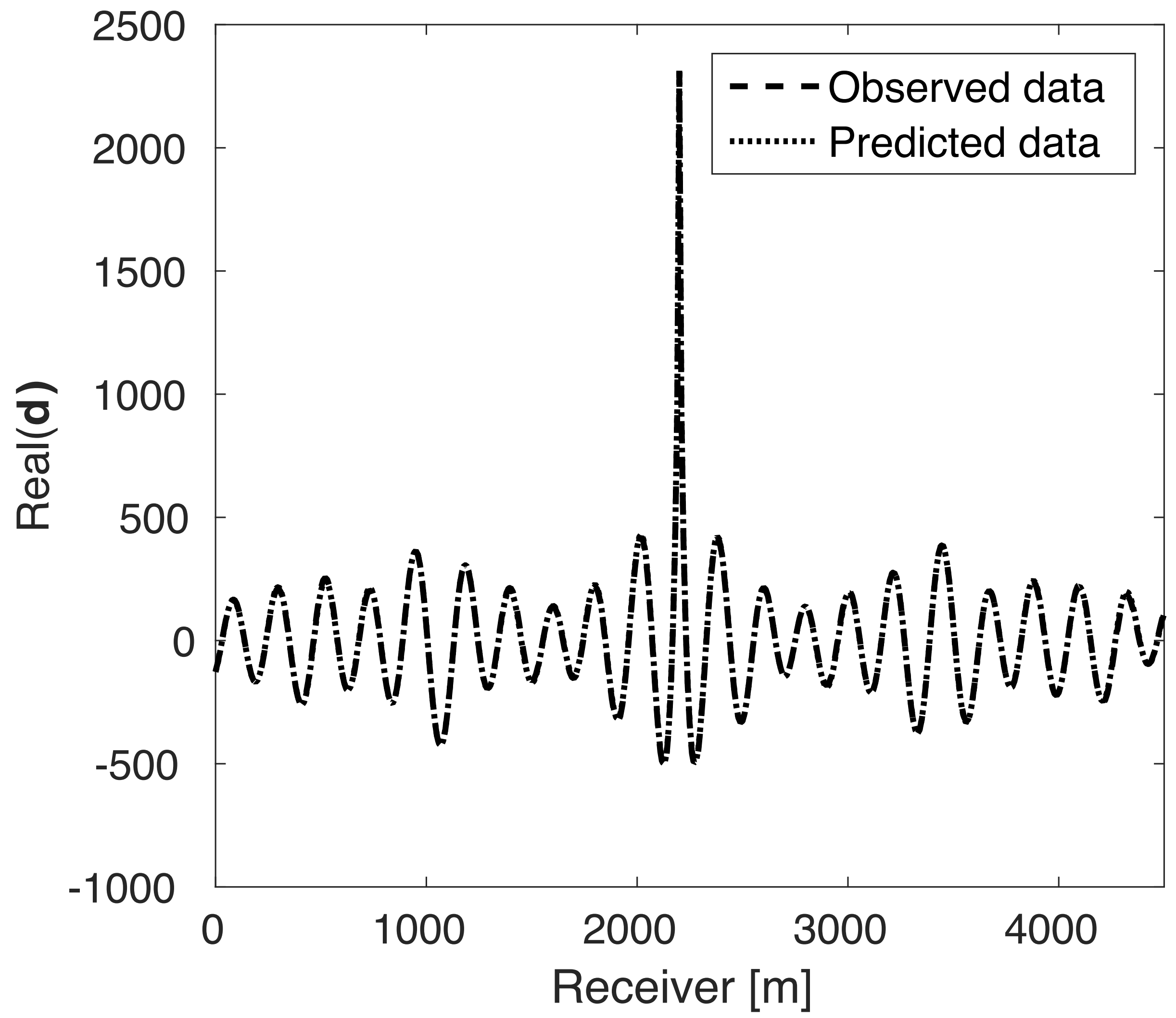


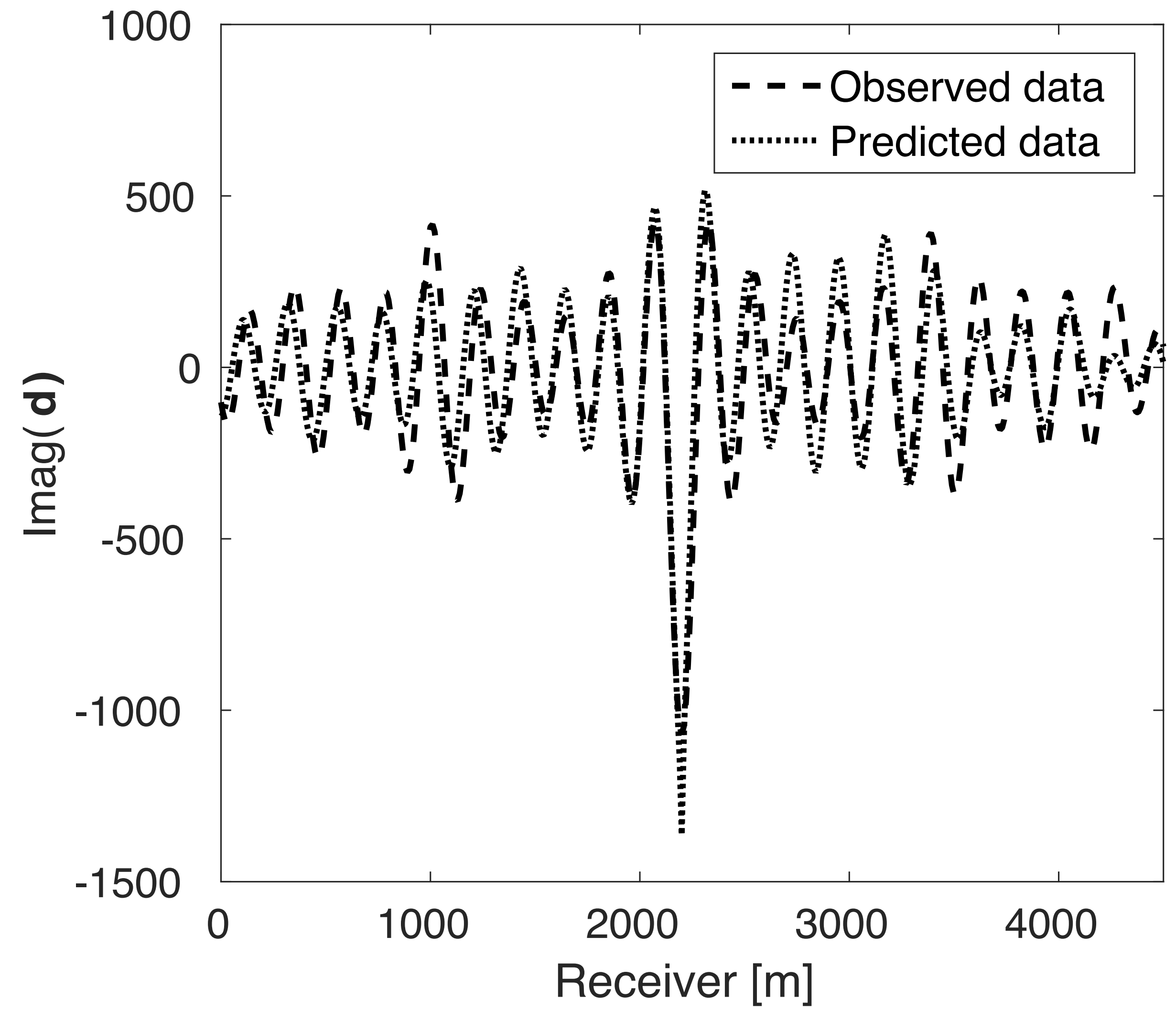
(d) Phase

Example on BG Compass model

—— WRI-SE-WS vs FWI-SE (Data comparison)

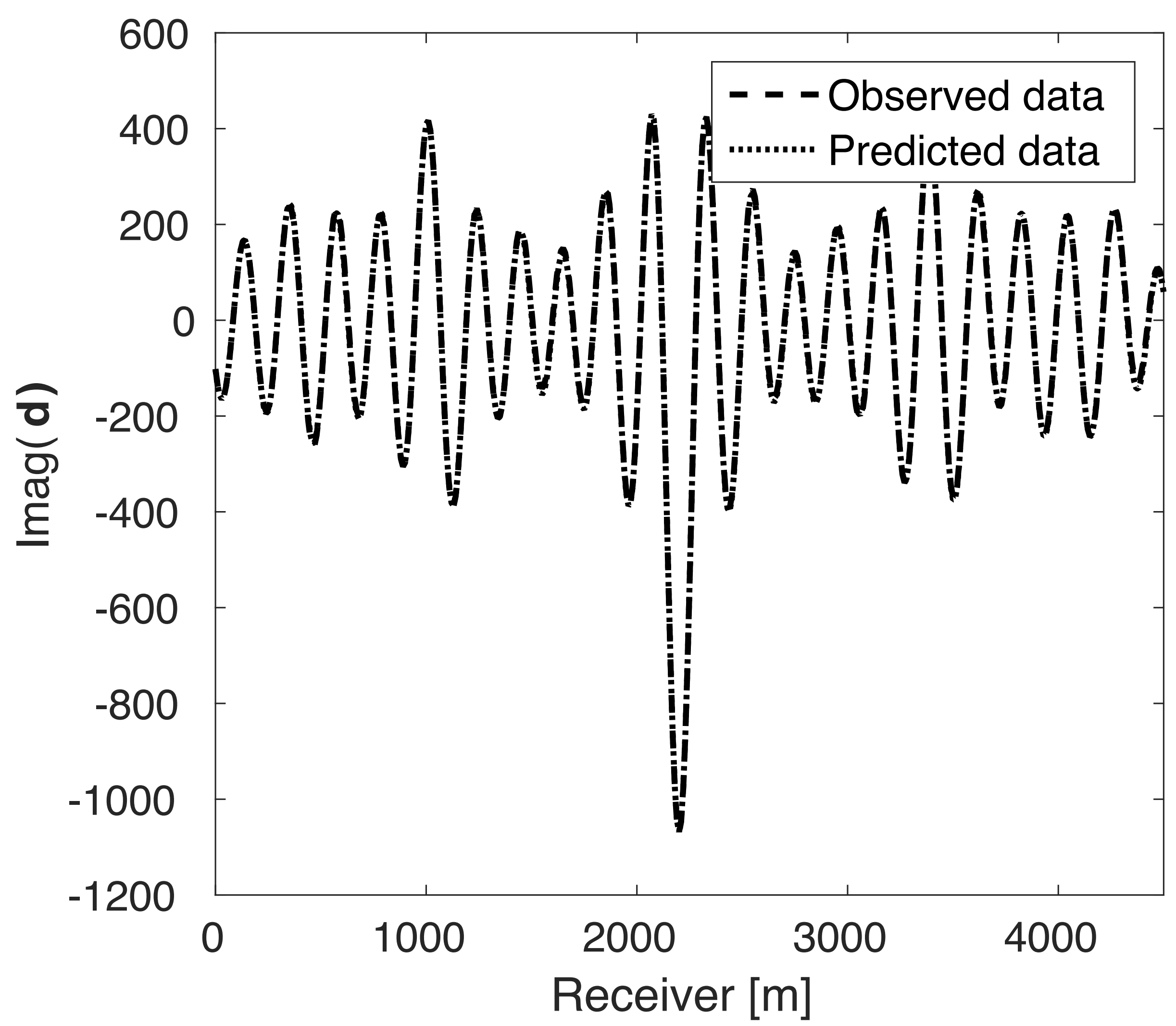
(a) Real part
WRI-SE-WS

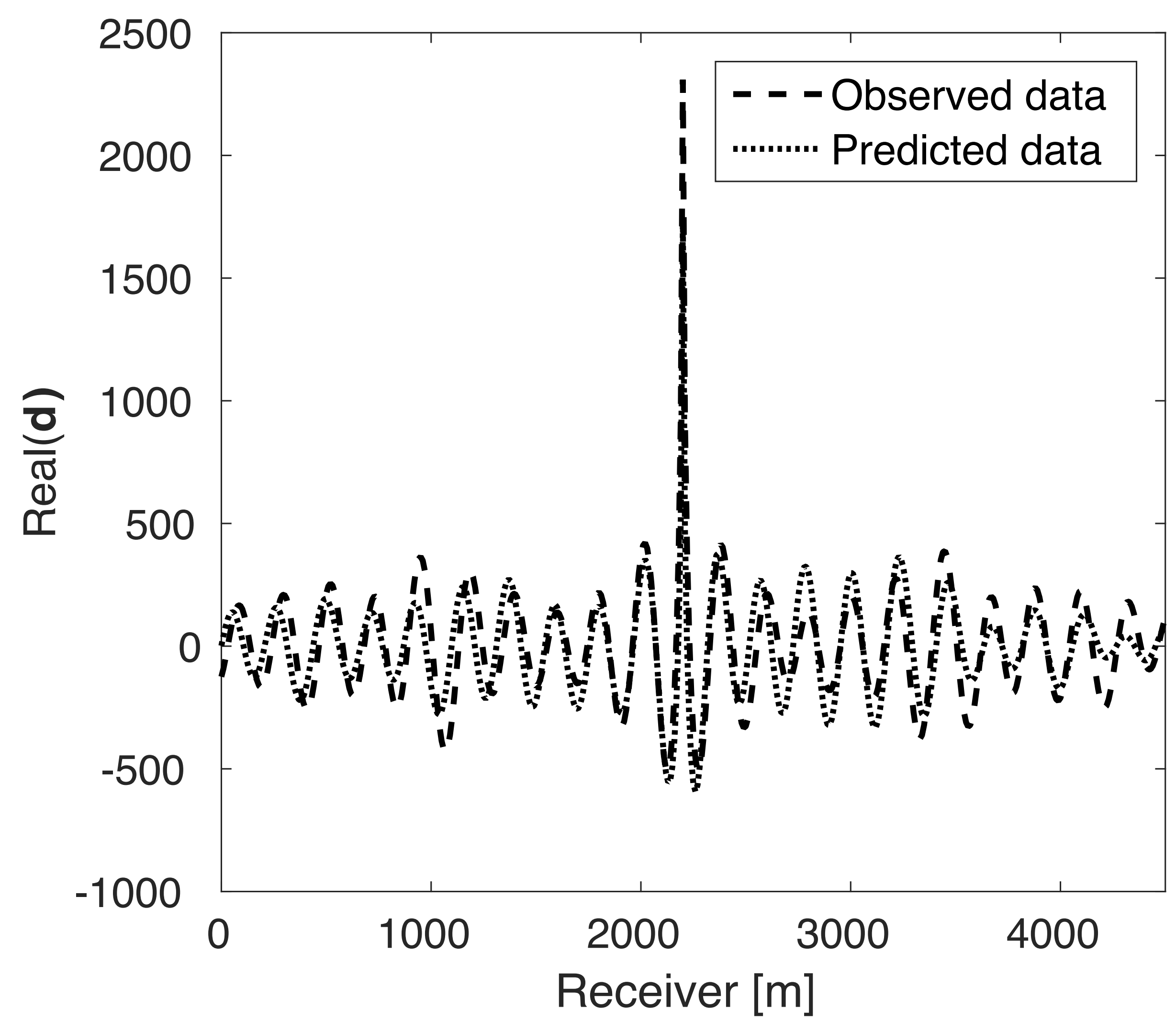




(c) Real part
FWI-SE

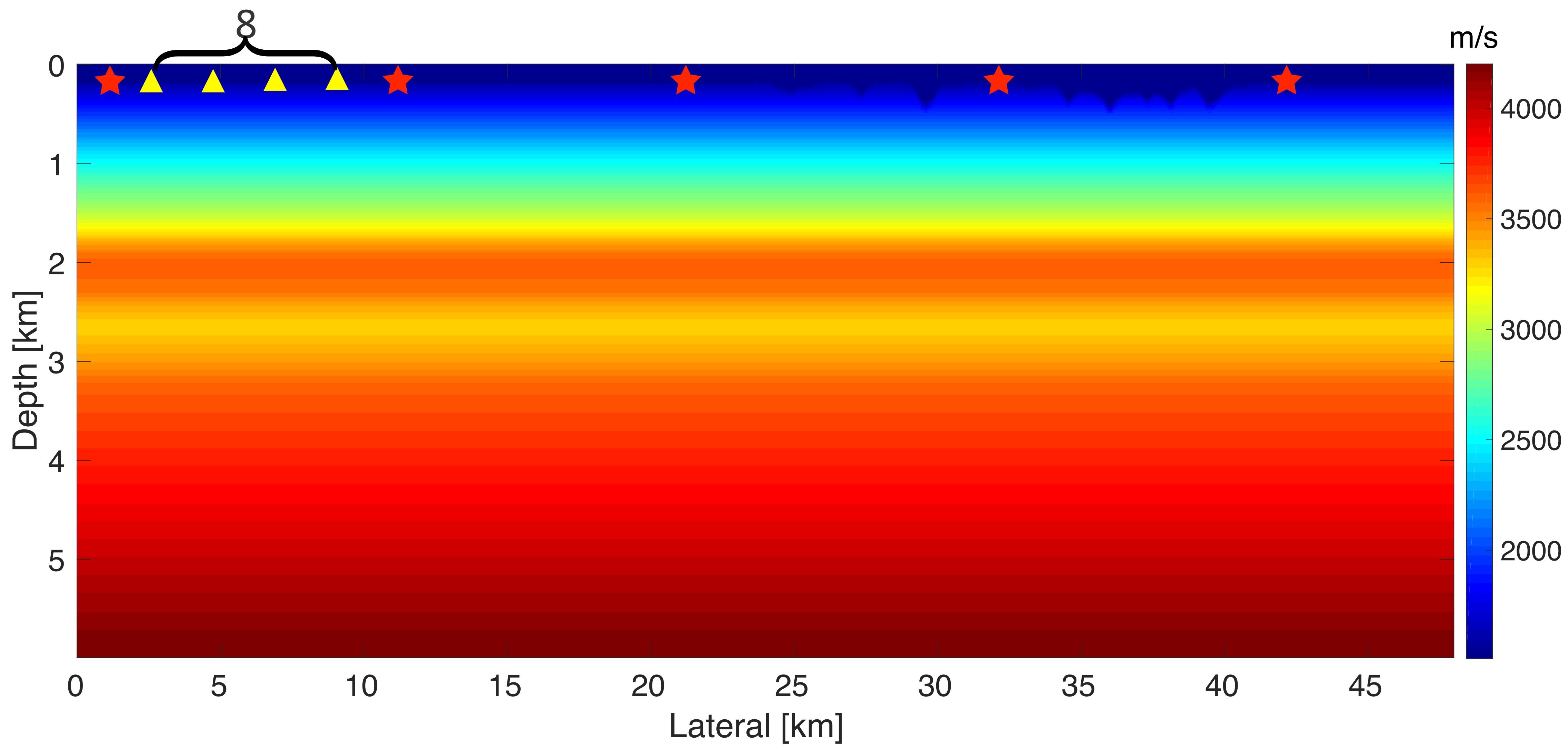
(b) Imaginary part
WRI-SE-WS



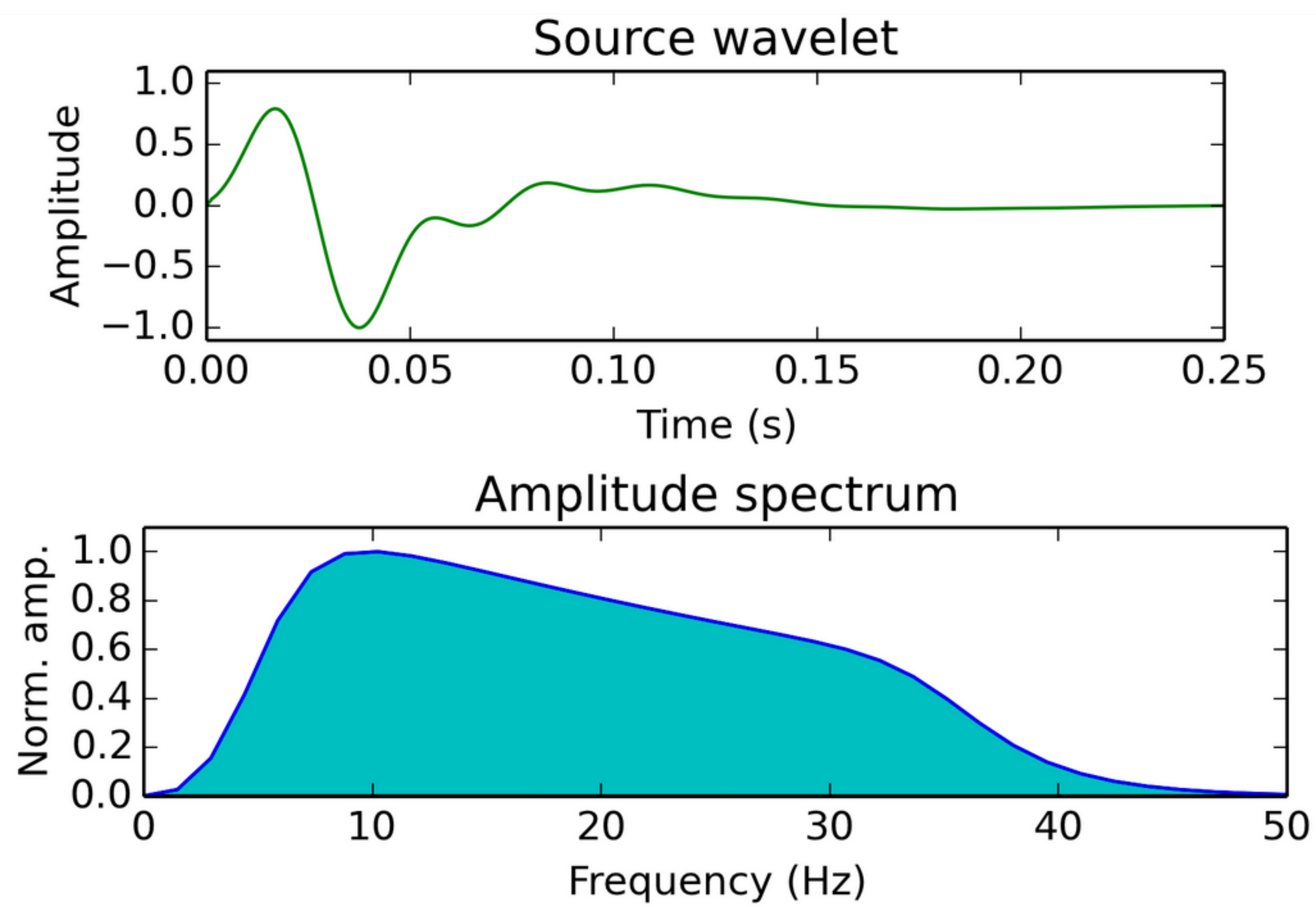


(d) Imaginary part
FWI-SE

Chevron blind test data



(a) Initial model



(b) Source information

Data-set information:

1. 1600 shots: ds = 25 m, Source depth = 15 m;
2. 321 recs/shot: dr = 25 m, Receiver depth = 15 m;
3. Maximum offset = 8000 m;
4. Record time = 8.0 s, sample rate 4 ms;
5. Vp water = constant = 1510 m/s;
6. With free surface multiples present in the data;
7. Isotropic Elastic.

SLIM

WRI with minimum smoothness constraint

$$\min_{\mathbf{u}(\mathbf{m}),\mathbf{m},\alpha(\mathbf{m})} \sum_{k,l} \|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2 \|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \alpha_{k,l}\mathbf{e}_{k,l}\|_2^2$$

$$\text{suject to} \quad \mathbf{m} \in \mathcal{C}_1 \bigcap \mathcal{C}_2$$

$$\mathcal{C}_1 \equiv \{\mathbf{m} \mid \mathbf{b}_l \leq \mathbf{m} \leq \mathbf{b}_u\}$$

$$\mathcal{C}_2 \equiv \{\mathbf{m} \mid \mathbf{E}^\top \mathbf{F}^\top (\mathbf{I} - \mathbf{S})\mathbf{FEm} = 0\}$$

SLIM

WRI with minimum smoothness constraint

$$\min_{\mathbf{u}(\mathbf{m}),\mathbf{m},\alpha(\mathbf{m})} \sum_{k,l} \|\mathbf{P}_k \mathbf{u}_{k,l} - \mathbf{d}_{k,l}\|_2^2 + \lambda^2 \|\mathbf{A}_{k,l}(\mathbf{m})\mathbf{u}_{k,l} - \alpha_{k,l}\mathbf{e}_{k,l}\|_2^2$$

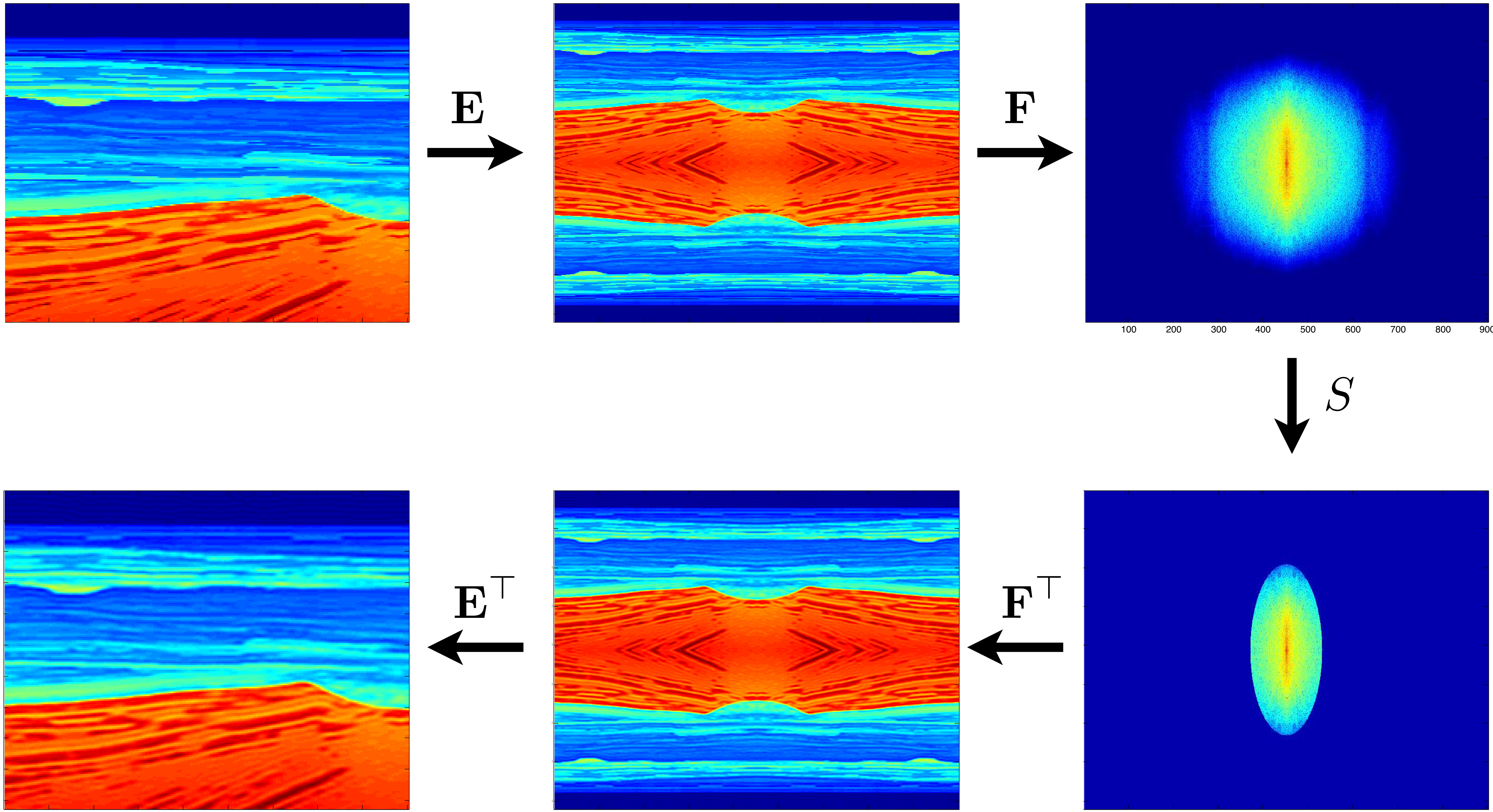
$$\text{suject to} \quad \mathbf{m} \in \mathcal{C}_1 \bigcap \mathcal{C}_2$$

$$\mathcal{C}_1 \equiv \{\mathbf{m} \mid \mathbf{b}_l \leq \mathbf{m} \leq \mathbf{b}_u\}$$

$$\mathcal{C}_2 \equiv \{\mathbf{m} \mid \mathbf{E}^\top \mathbf{F}^\top (\mathbf{I} - \mathbf{S})\mathbf{F}\mathbf{E}\mathbf{m} = 0\}$$

1. 2D mirror extension of the model (to avoid periodic boundaries)
2. 2D DFT
3. Remove coefficients outside ellipse (highest spatial frequencies)
4. 2D inverse DFT

WRI with minimum smoothness constraint



[M. Schmidt *et al*, 2012]

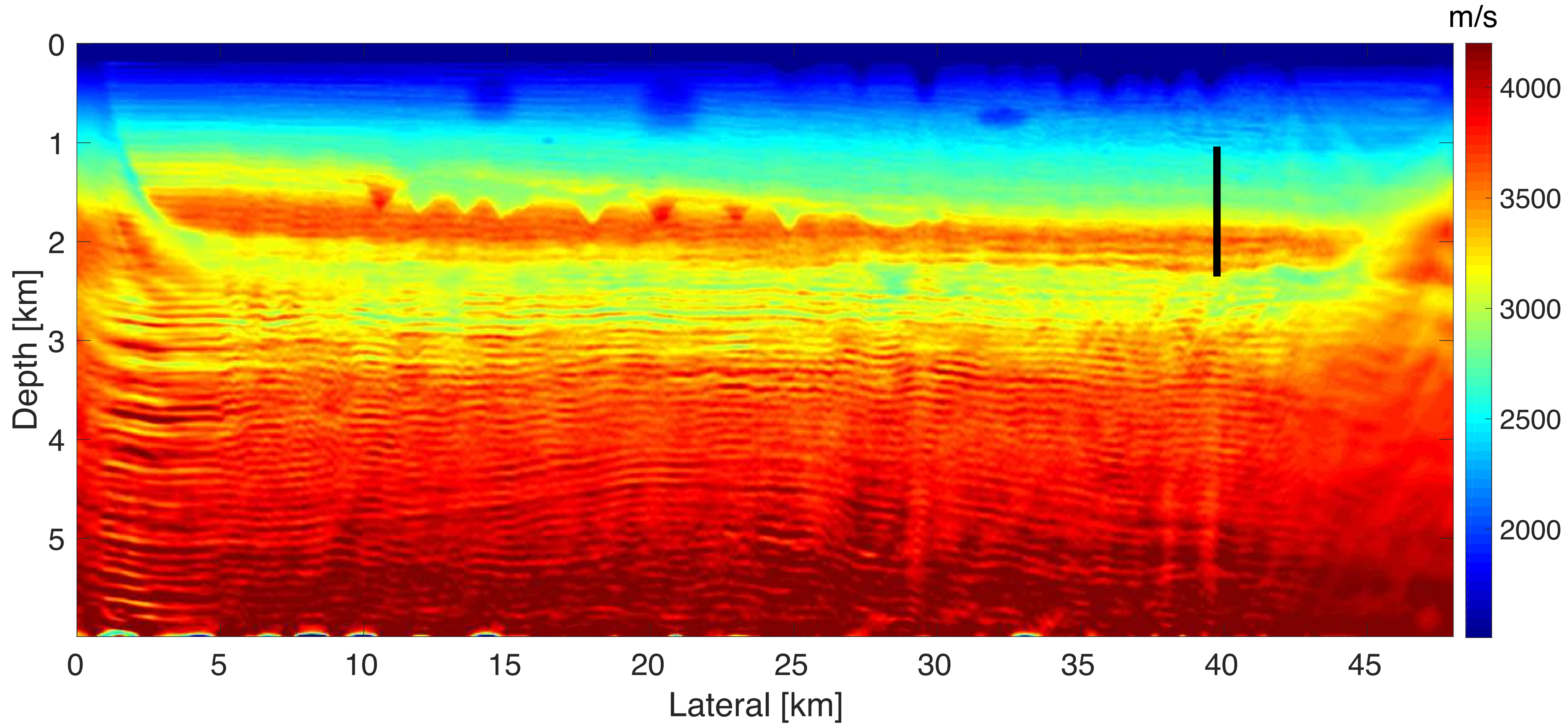
SLIM

Chevron blind test data

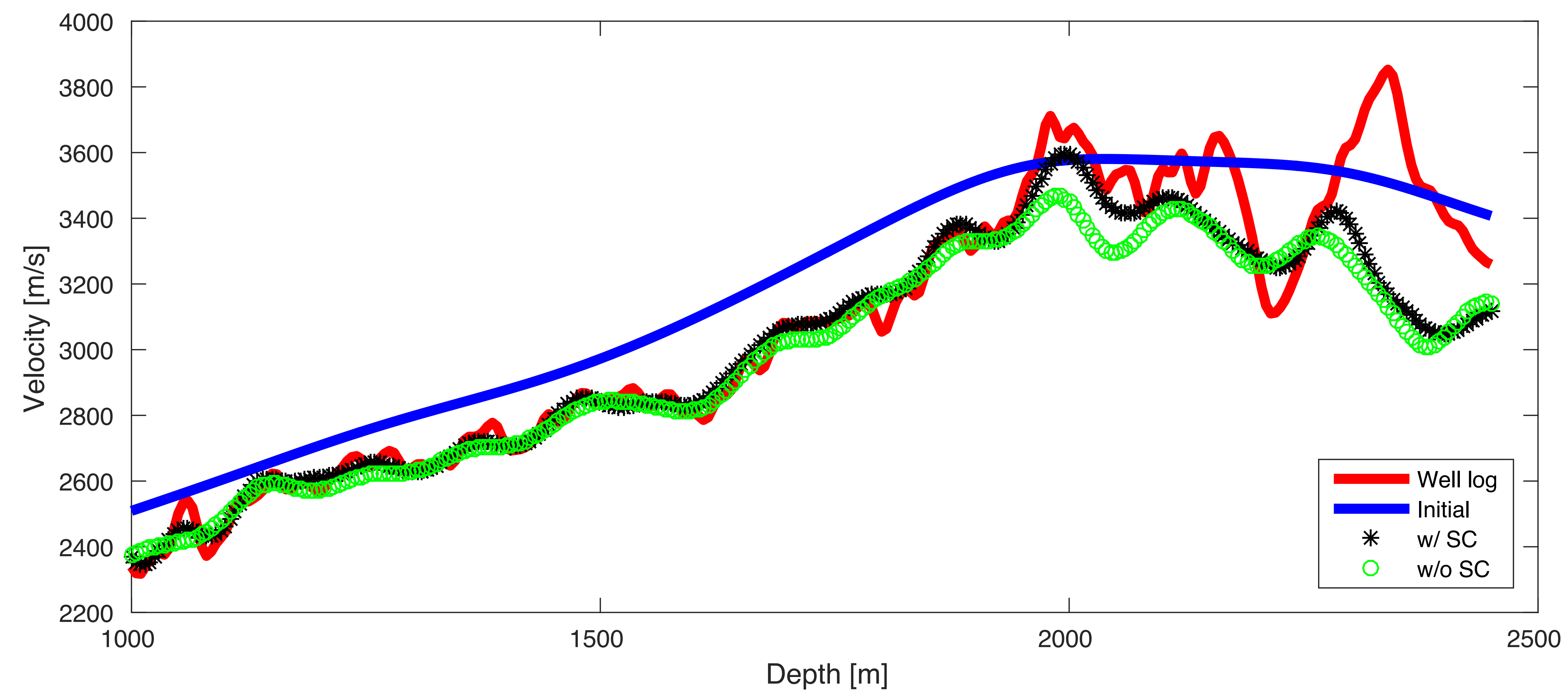
Inversion strategy:

1. Frequency domain WRI with Source estimation and smoothness constraint;
2. Frequency bands: [3:0.2:5]Hz, [3:0.2:7]Hz, [3:0.2:9]Hz, [3:0.2:11], [3:0.2:15], [3:0.2:19]Hz;
3. Batch sizes of random frequency subsets: 3, 6, 10, 10, 15, 15;
4. Batch size of random source subsets: 300;
5. Optimization solver: Projected quasi Newton with 20 iterations per frequency band;
6. 2 passes of WRI from frequency 3-19Hz;
7. Grid size: 20m for 3-11Hz and 12m for 3-19Hz;
8. No pre-processing !!!

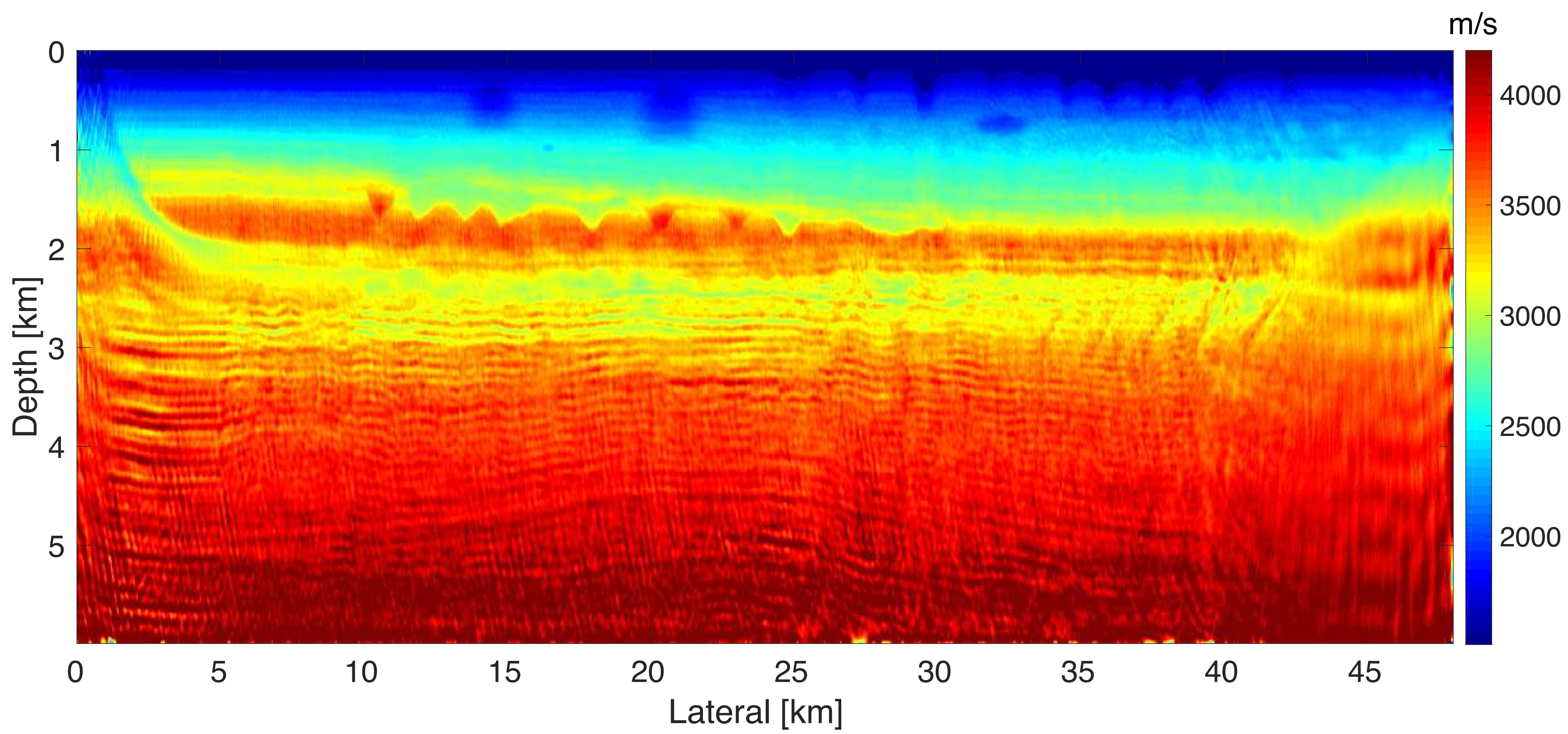
Inversion results



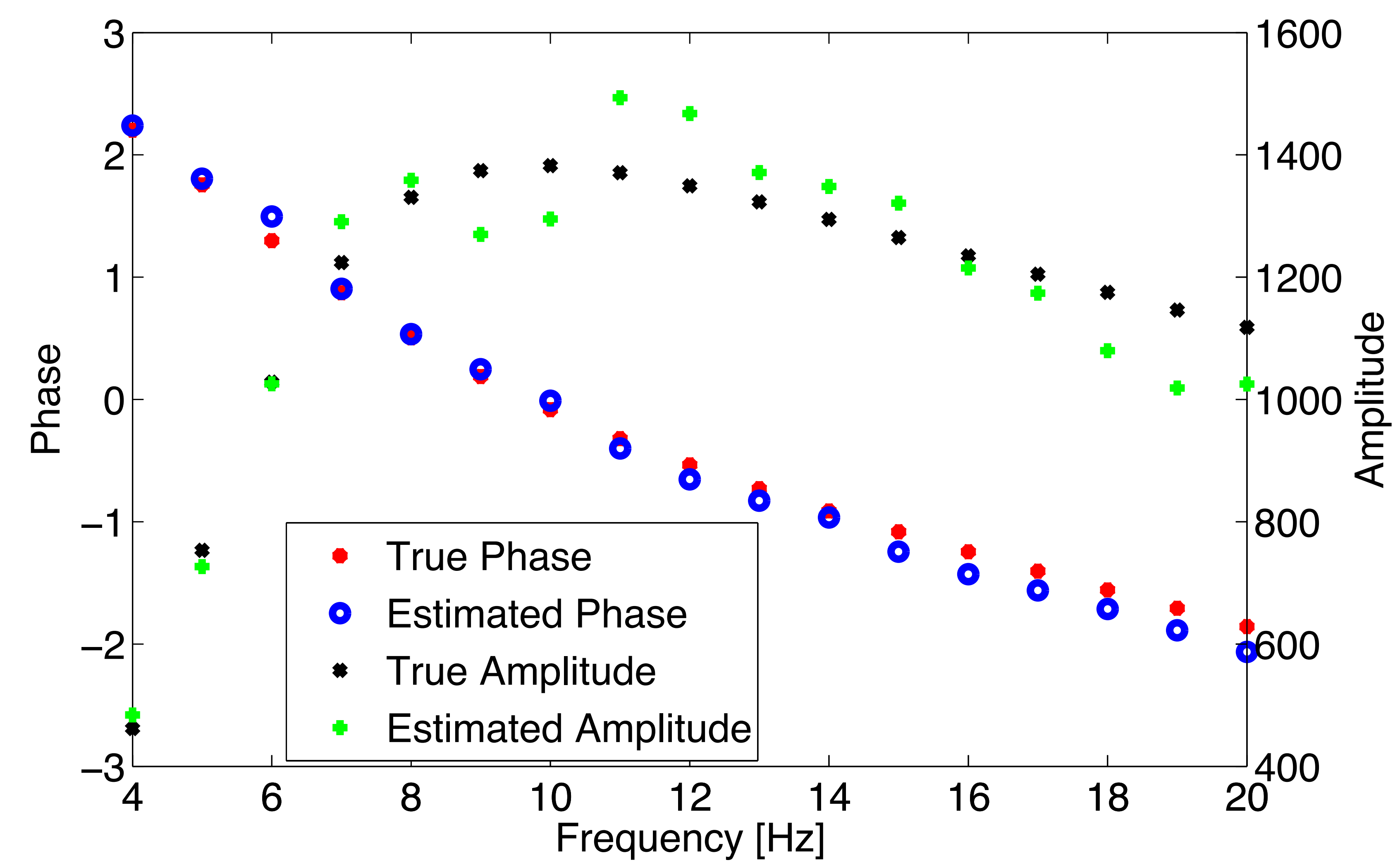
(a) Result w/ smoothness constraint



(c) Well log comparison



(b) Result w/o smoothness constraint

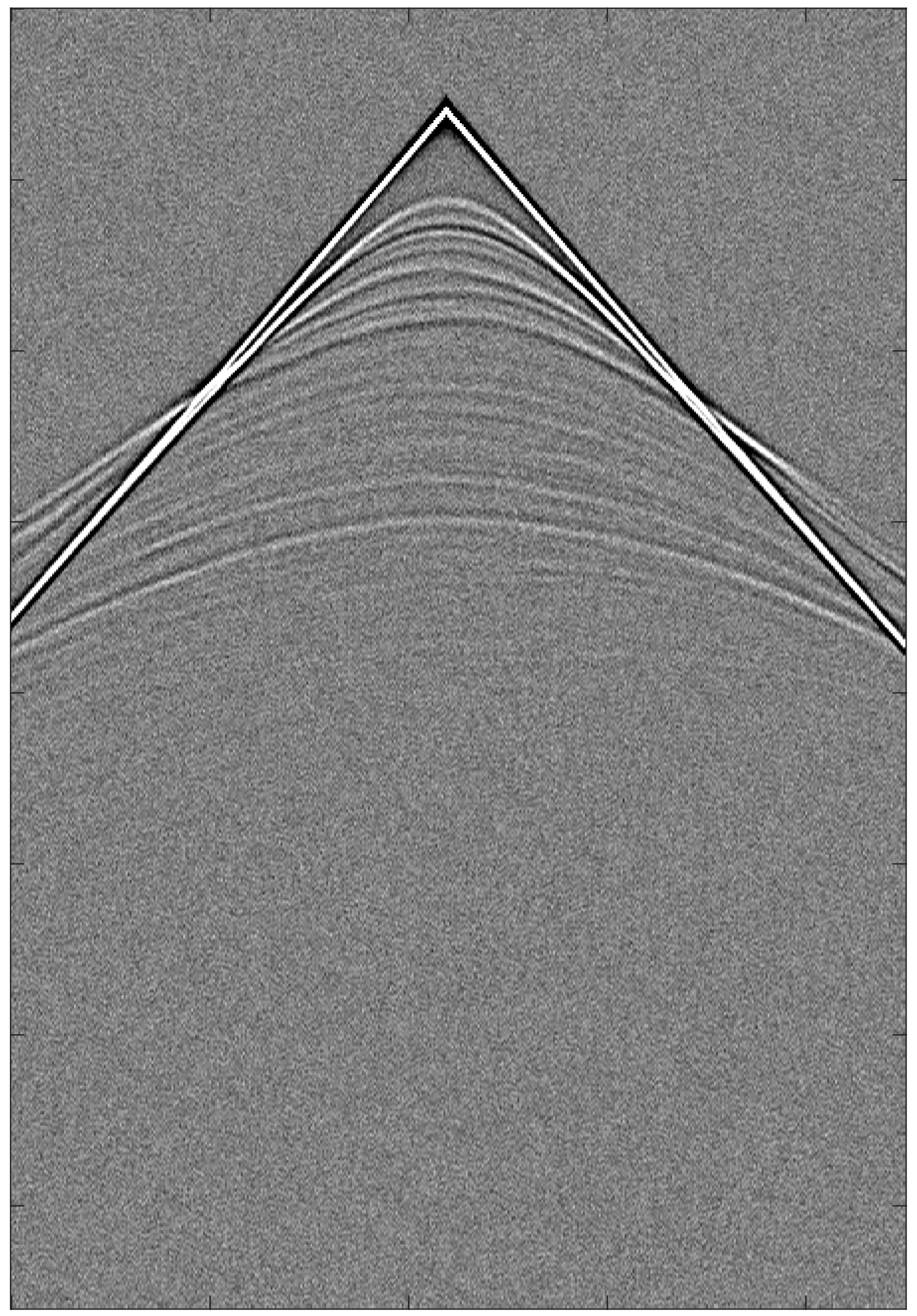


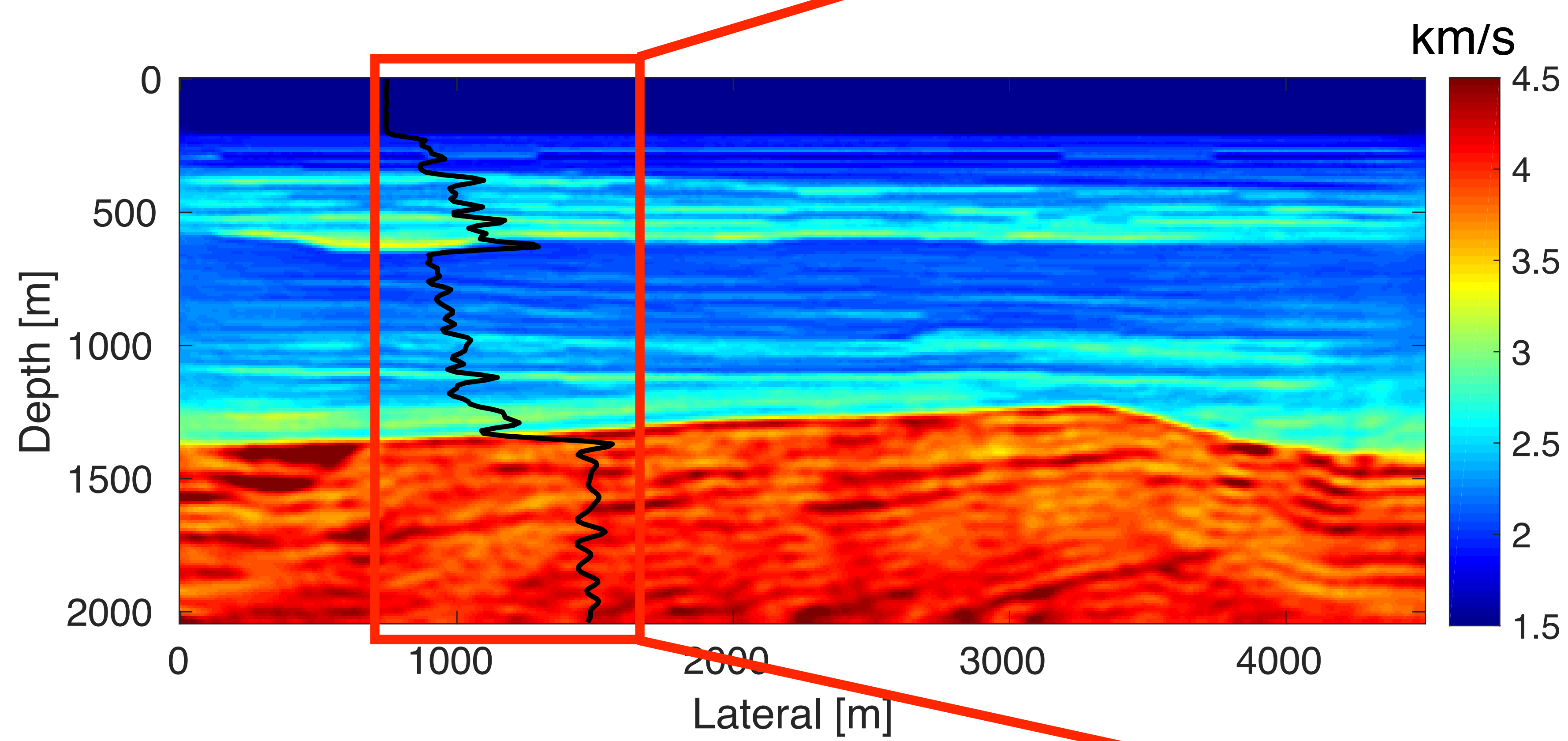
(d) Source function comparison

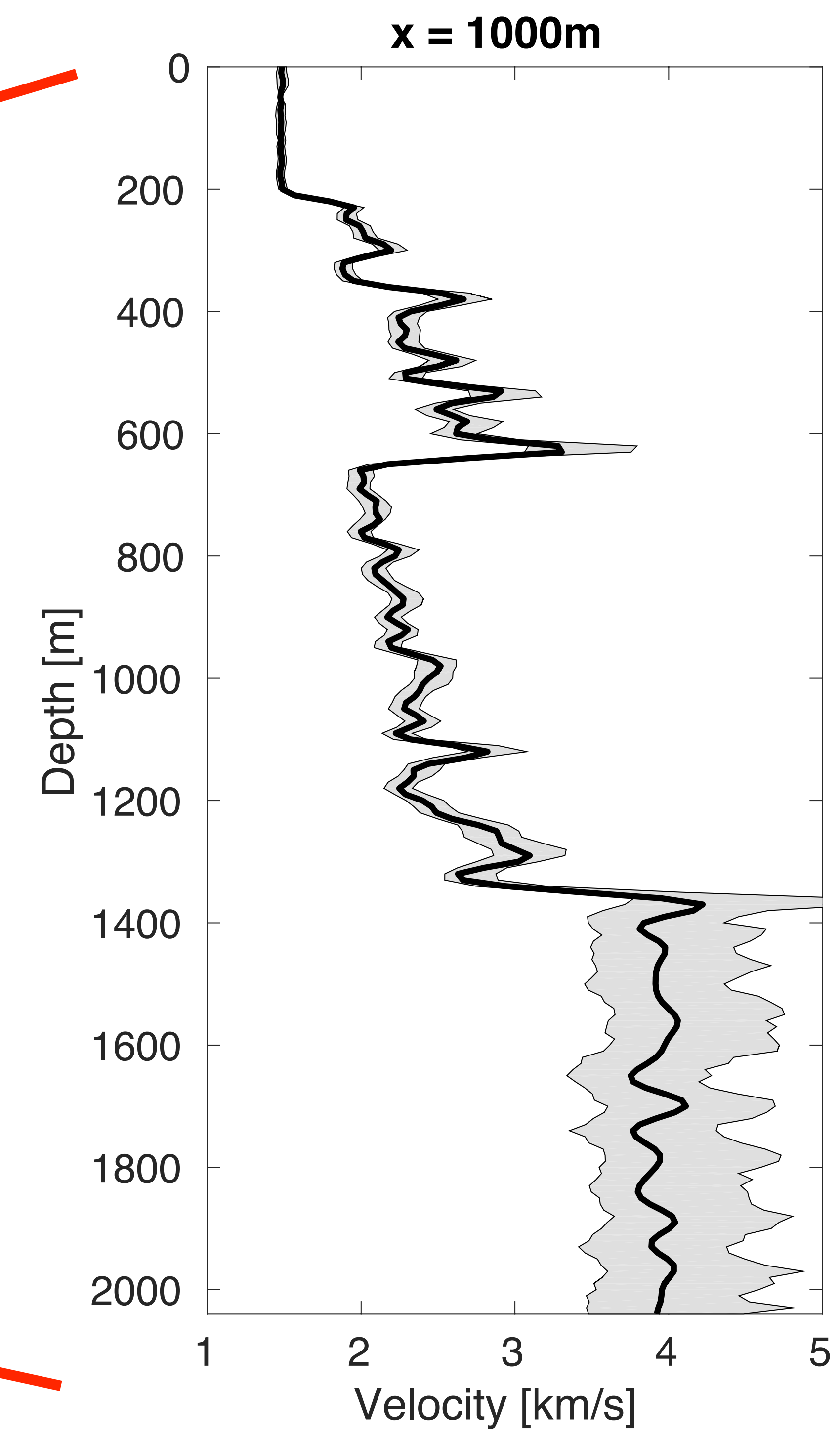
Chapter 6

Uncertainty quantification for inverse problems with weak wave-equation constraints

Motivation







Uncertainty in data **Uncertainty in model** **Confidence of the model**

[A. Tarantola and B. Valette, 1982]
[J. Kaipio and E. Somersalo, 2004]

SLIM

Bayesian inference

Prior probability density function (PDF):

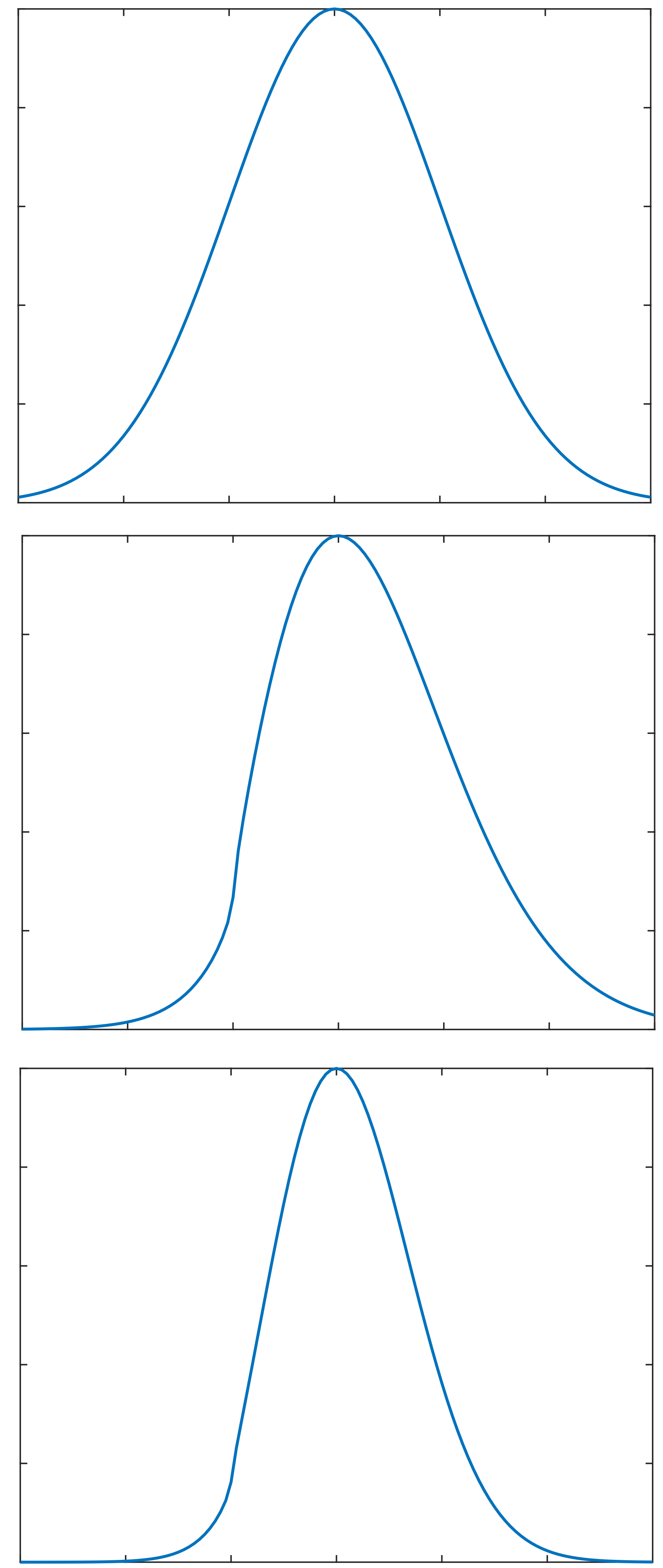
$$\mathbf{m} \longrightarrow \rho_{\text{prior}}(\mathbf{m})$$

Likelihood PDF: given data $\mathbf{d}$

$$\mathbf{m} \longrightarrow \rho_{\text{like}}(\mathbf{d}|\mathbf{m})$$

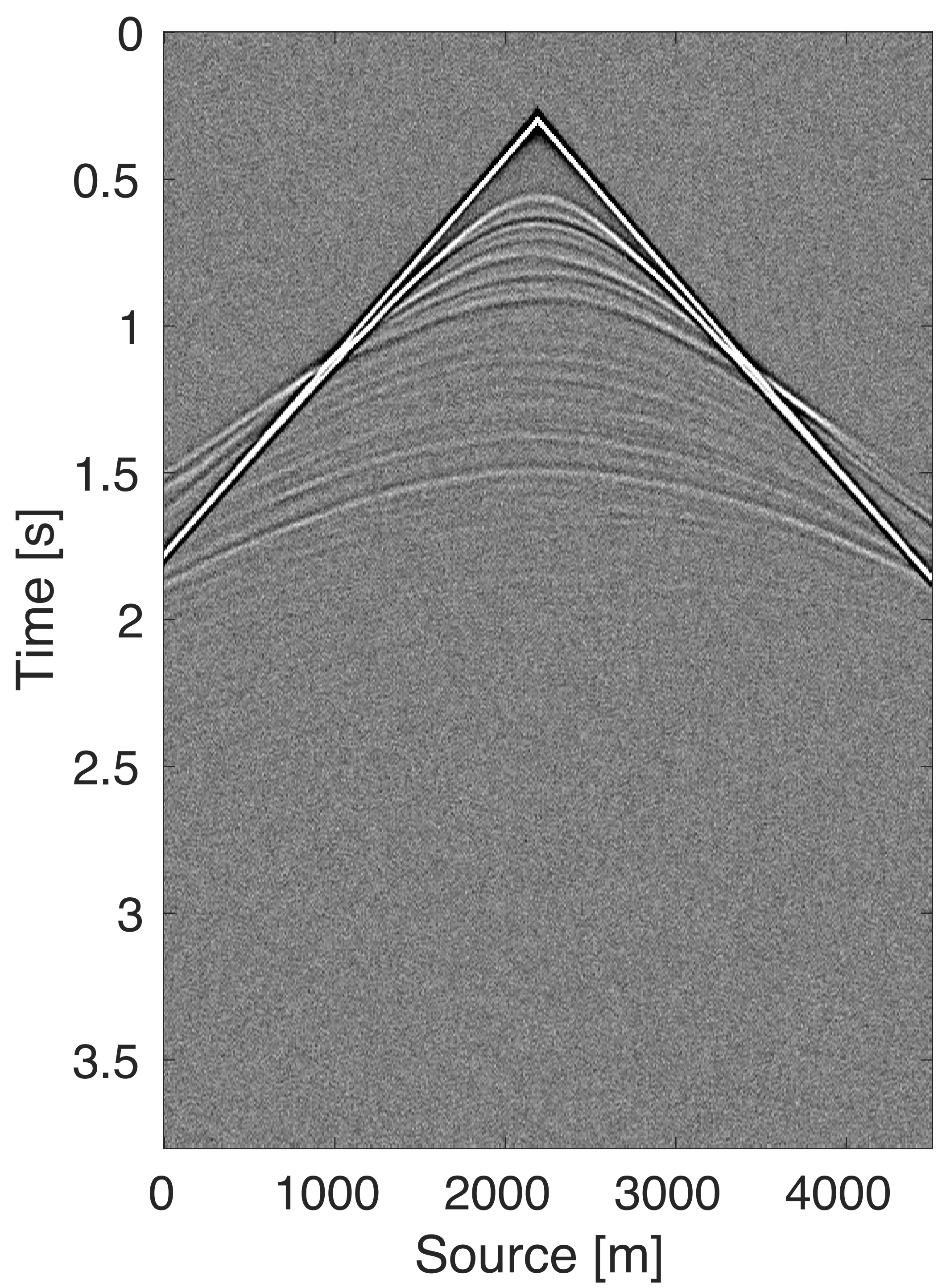
Posterior PDF (Bayes' rule):

$$\rho_{\text{post}}(\mathbf{m}|\mathbf{d}) \propto \rho_{\text{like}}(\mathbf{d}|\mathbf{m})\rho_{\text{prior}}(\mathbf{m})$$

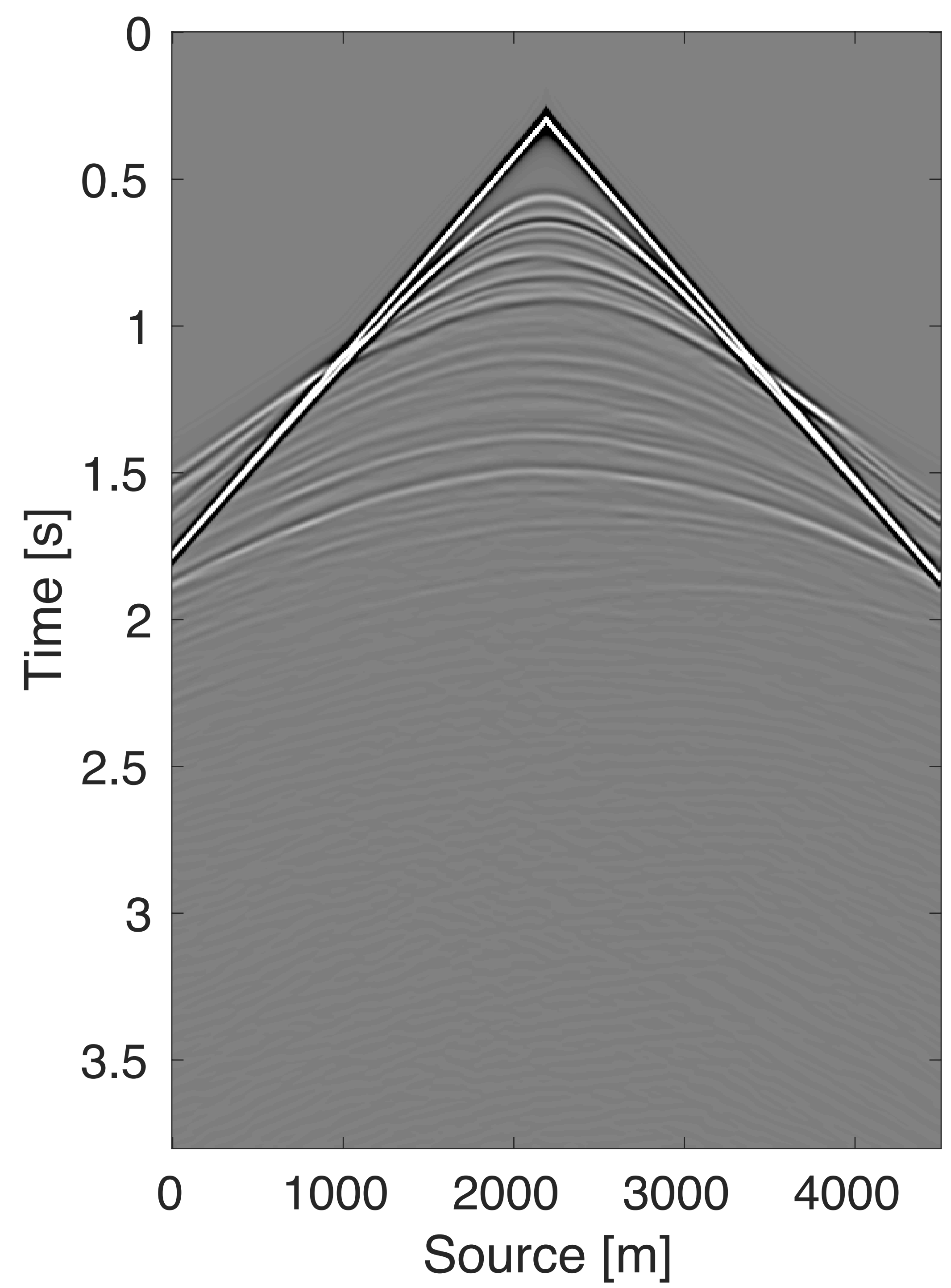


[James Martin *et al*, 2012]

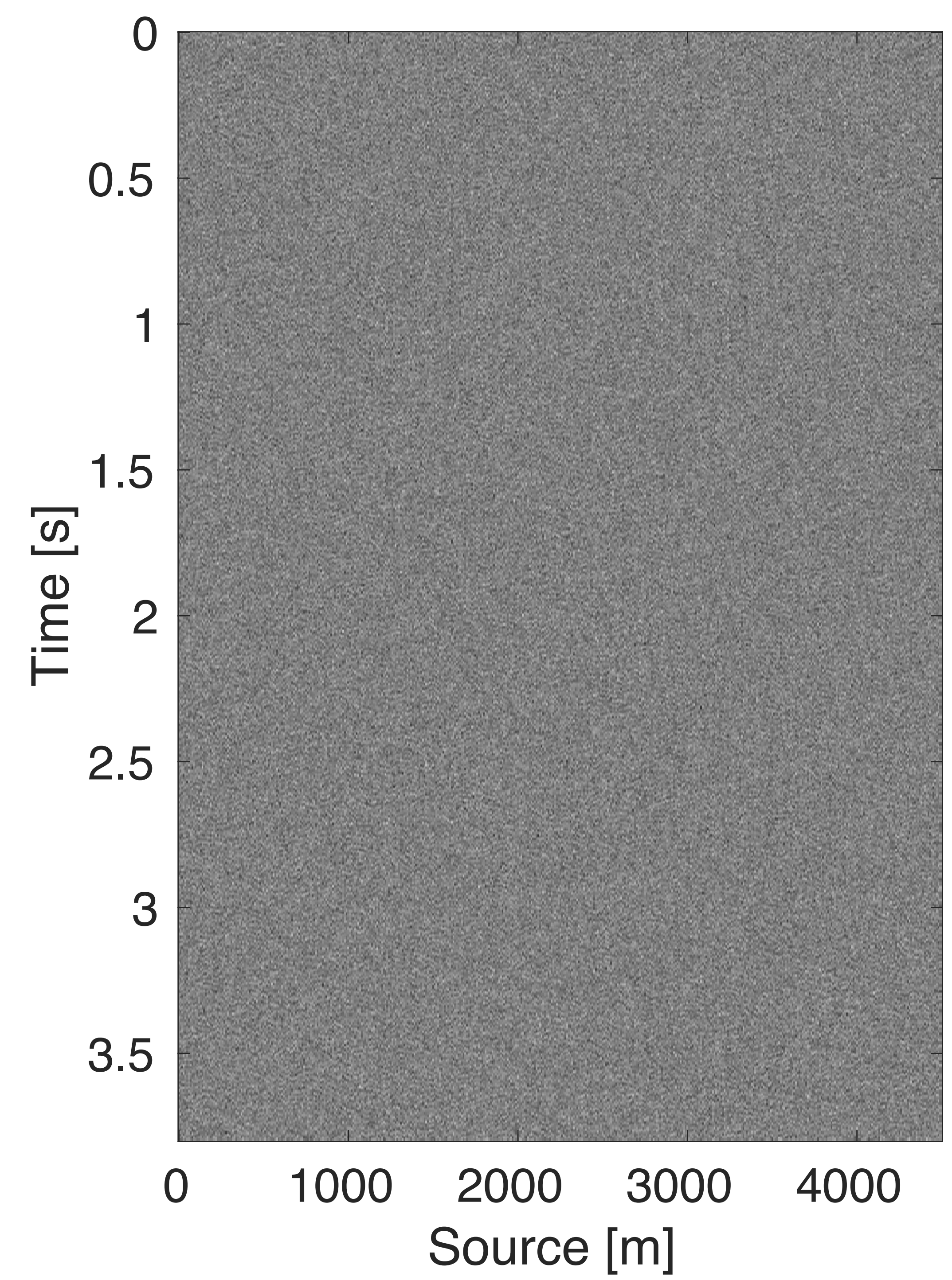
Bayesian w/ FWI



=



+



$$\mathbf{d}_{\mathrm{obs}} \qquad F(\mathbf{m}) \qquad \epsilon \sim \mathcal{N}(0, \Sigma_{\mathrm{noise}})$$

[J. Virieux and S. Operto, 2009]
[James Martin *et al*, 2012]

Bayesian w/ FWI

Posterior PDF of FWI:

$$\rho_{\mathrm{post}}(\mathbf{m}|\mathbf{d}) \propto \exp\left(-\frac{1}{2}\|\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q}-\mathbf{d}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{noise}}} - \frac{1}{2}\|\mathbf{m}-\mathbf{m}_{\mathrm{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\mathrm{prior}}}\right)$$

[J. Virieux and S. Operto, 2009]
[James Martin *et al*, 2012]

SLIM

Bayesian w/ FWI

Posterior PDF of FWI:

$$\rho_{\text{post}}(\mathbf{m}|\mathbf{d}) \propto \exp\left(-\frac{1}{2}\underline{\|\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q}-\mathbf{d}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}} - \frac{1}{2}\|\mathbf{m}-\mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}\right)$$

Strong nonlinearity
Many local minima

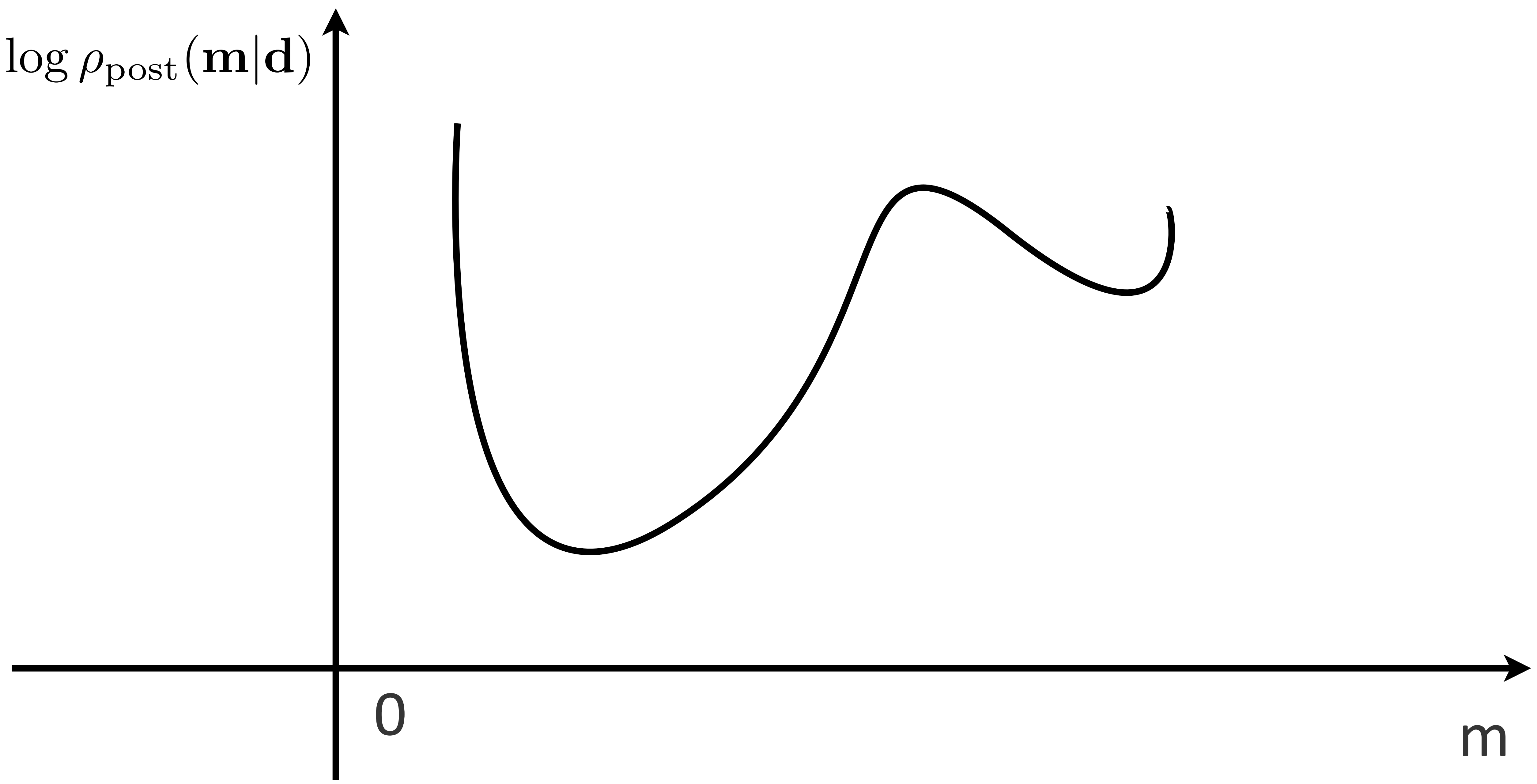
[J. Virieux and S. Operto, 2009]
[James Martin *et al*, 2012]

Bayesian w/ FWI

Posterior PDF of FWI:

$$\rho_{\text{post}}(\mathbf{m}|\mathbf{d}) \propto \exp\left(-\underline{\frac{1}{2}\|\mathbf{PA}(\mathbf{m})^{-1}\mathbf{q}-\mathbf{d}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}} - \frac{1}{2}\|\mathbf{m}-\mathbf{m}_{\text{prior}}\|^2_{\mathbf{\Sigma}^{-1}_{\text{prior}}}\right)$$

Strong nonlinearity
Many local minima



Extend the search space

Introduce auxiliary variable:

$$\rho_{\text{post}}(\mathbf{u},\mathbf{m}|\mathbf{d}) \propto \rho(\mathbf{d}|\mathbf{u},\mathbf{m})\rho(\mathbf{u},\mathbf{m})$$

$$\rho(\mathbf{u},\mathbf{m}) = \rho(\mathbf{u}|\mathbf{m})\rho_{\text{prior}}(\mathbf{m})$$

Strict PDE constraints:

$$\rho(\mathbf{d}|\mathbf{u},\mathbf{m}) \propto \exp\left(-\frac{1}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}\right)$$

$$\rho(\mathbf{u}|\mathbf{m}) = \delta\big(\mathbf{A}(\mathbf{m})\mathbf{u}-\mathbf{q}\big)$$

Extend the search space

Introduce auxiliary variable:

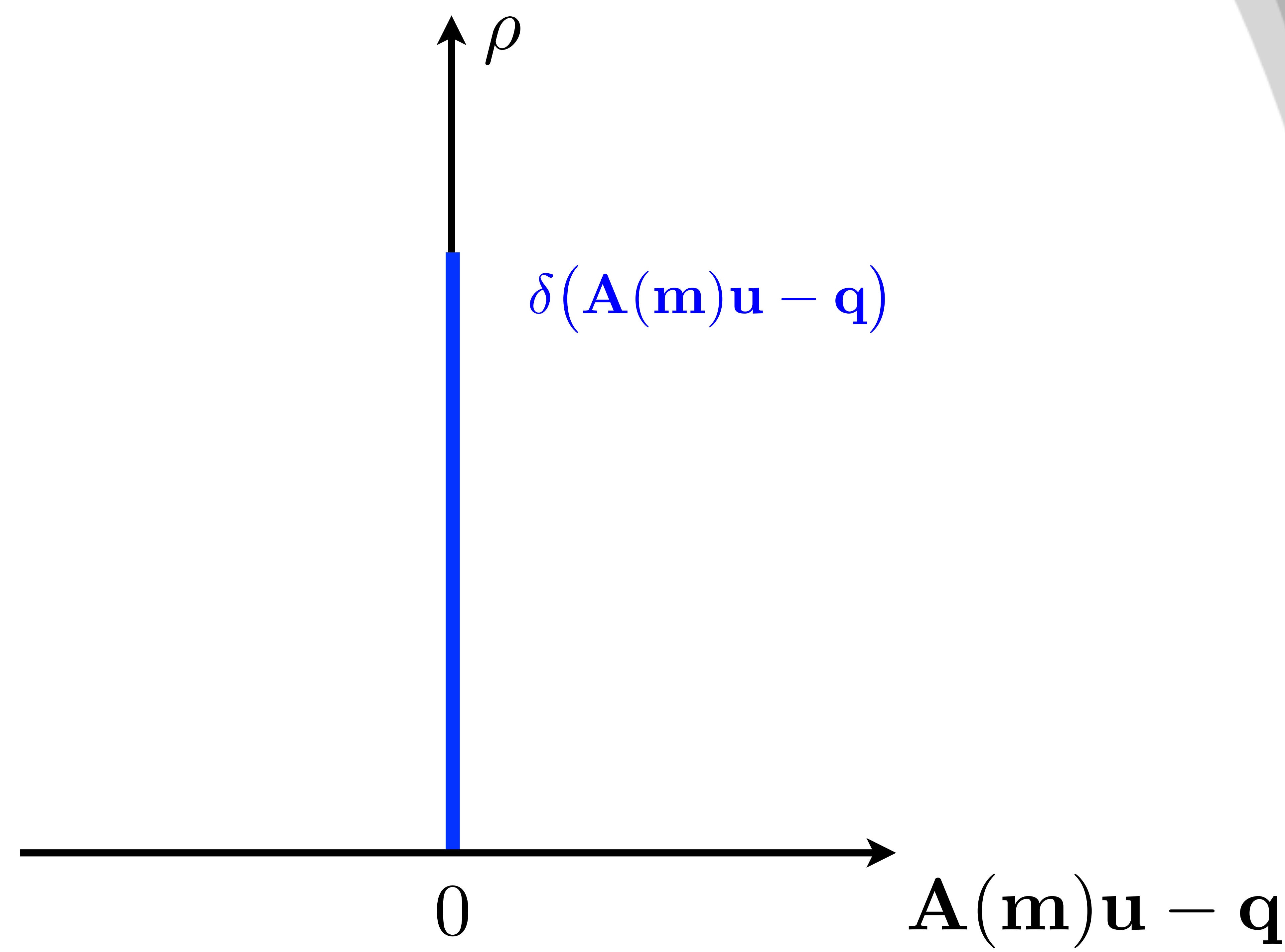
$$\rho_{\text{post}}(\mathbf{u},\mathbf{m}|\mathbf{d}) \propto \rho(\mathbf{d}|\mathbf{u},\mathbf{m})\rho(\mathbf{u},\mathbf{m})$$

$$\rho(\mathbf{u},\mathbf{m}) = \rho(\mathbf{u}|\mathbf{m})\rho_{\text{prior}}(\mathbf{m})$$

Strict PDE constraints:

$$\rho(\mathbf{d}|\mathbf{u},\mathbf{m}) \propto \exp\left(-\frac{1}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|^2_{\mathbf{\Sigma}^{-1}_{\text{noise}}}\right)$$

$$\rho(\mathbf{u}|\mathbf{m}) = \delta\big(\mathbf{A}\left(\mathbf{m}\right)\mathbf{u}-\mathbf{q}\big)$$

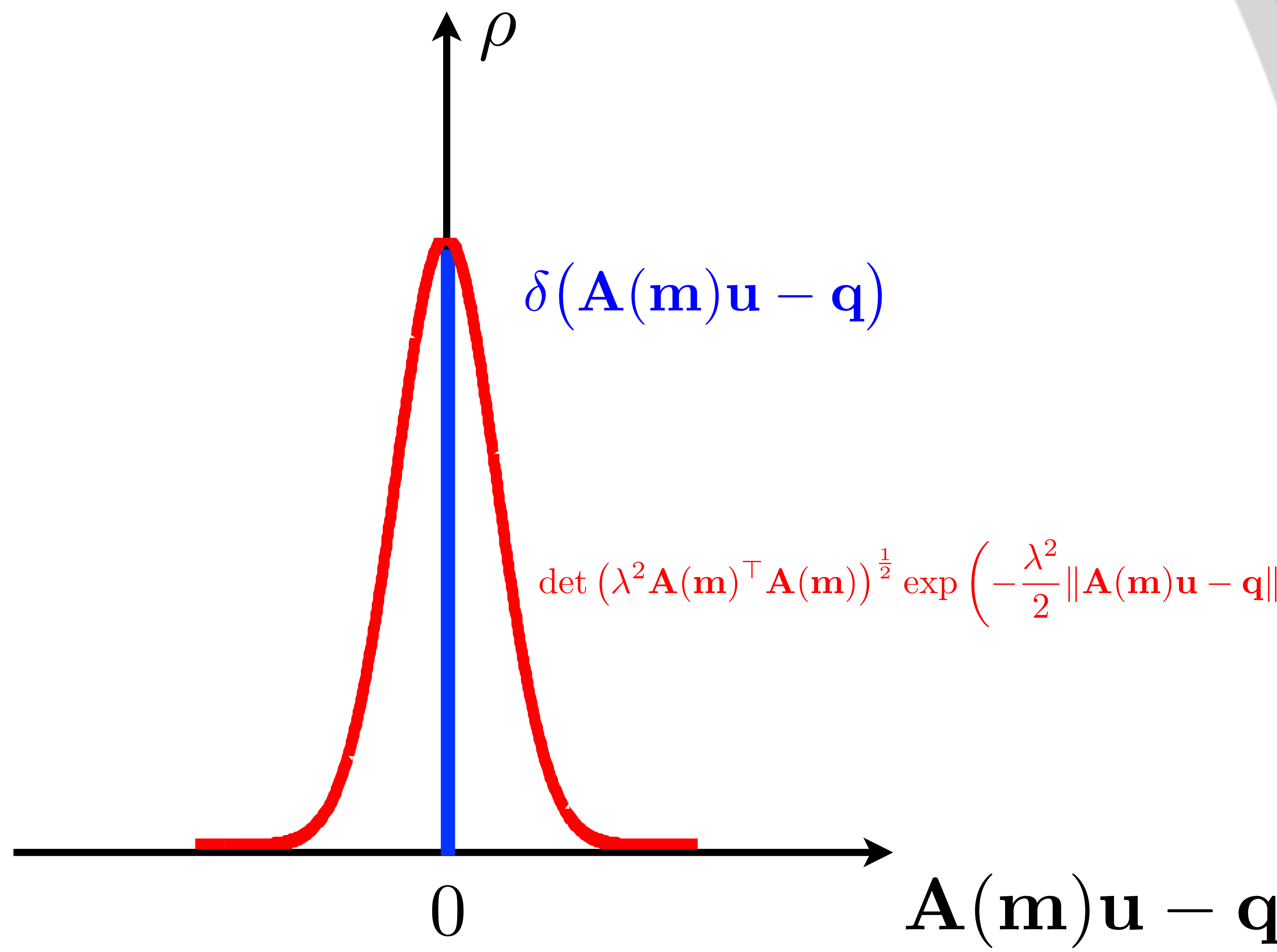


Extend the search space

Relax PDE constraints:

$$\rho(\mathbf{u}|\mathbf{m}) = \delta\big(\mathbf{A}\left(\mathbf{m}\right)\mathbf{u} - \mathbf{q}\big)$$

$$\rho(\mathbf{u}|\mathbf{m}) \propto \det\left(\lambda^2 \mathbf{A}(\mathbf{m})^\top \mathbf{A}(\mathbf{m})\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda^2}{2}\|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{q}\|^2\right)$$



Posterior distribution with weak PDE constraints

Joint posterior distribution:

$$\rho_{\text{post}}(\mathbf{u},\mathbf{m}|\mathbf{d}) \propto \rho(\mathbf{d}|\mathbf{u},\mathbf{m})\rho(\mathbf{u},\mathbf{m}) \quad \text{with}$$

$$\rho(\mathbf{u},\mathbf{m}) = \rho(\mathbf{u}|\mathbf{m})\rho_{\text{prior}}(\mathbf{m}),$$

$$\rho(\mathbf{d}|\mathbf{u},\mathbf{m}) \propto \exp\left(-\frac{1}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|^2_{\boldsymbol{\Sigma}^{-1}_{\text{noise}}}\right),$$

$$\rho(\mathbf{u}|\mathbf{m}) \propto \det\left(\lambda^2\mathbf{A}(\mathbf{m})^\top\mathbf{A}(\mathbf{m})\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda^2}{2}\|\mathbf{A}(\mathbf{m})\mathbf{u}-\mathbf{q}\|^2\right).$$

Marginal distribution:

$$\begin{aligned}\rho_{\text{post}}(\mathbf{m}|\mathbf{d}) &= \int \rho_{\text{post}}(\mathbf{u},\mathbf{m}|\mathbf{d})\mathrm{d}\mathbf{u} \\ &\propto \det\big(\mathbf{H}(\mathbf{m})\big)^{-\frac{1}{2}} \det\left(\lambda^2\mathbf{A}(\mathbf{m})^\top\mathbf{A}(\mathbf{m})\right)^{\frac{1}{2}} \\ &\quad\times \exp\Big(-\frac{1}{2}\big(\lambda^2\|\mathbf{A}(\mathbf{m})\overline{\mathbf{u}}(\mathbf{m})-\mathbf{q}\|^2 + \|\mathbf{P}\overline{\mathbf{u}}(\mathbf{m})-\mathbf{d}\|^2_{\Gamma^{-1}_{\text{noise}}} + \|\mathbf{m}-\tilde{\mathbf{m}}\|^2_{\Gamma^{-1}_{\text{prior}}}\big)\Big)\end{aligned}$$

where $\mathbf{H}(\mathbf{m}) = -\nabla^2_{\mathbf{u}} \log \rho_{\text{post}}(\mathbf{u},\mathbf{m}|\mathbf{d})\big|_{\mathbf{u}=\overline{\mathbf{u}}(\mathbf{m})}$

$$= \lambda^2\mathbf{A}(\mathbf{m})^\top\mathbf{A}(\mathbf{m}) + \mathbf{P}^\top\Gamma^{-1}_{\text{noise}}\mathbf{P},$$

$$\overline{\mathbf{u}}(\mathbf{m}) = \mathbf{H}(\mathbf{m})^{-1}\big(\lambda^2\mathbf{A}(\mathbf{m})^\top\mathbf{q} + \mathbf{P}^\top\Gamma^{-1}_{\text{noise}}\mathbf{d}\big).$$

[van Leeuwen, T and Herrmann, F J , 2015]

Selection of λ

Negative logarithm function of posterior distribution:

$$\phi(\mathbf{m}) = -\log \rho_{\text{post}}(\mathbf{m}|\mathbf{d}) = \phi_1(\mathbf{m}) + \phi_2(\mathbf{m}), \quad \text{with}$$

$$\phi_1(\mathbf{m}) = \frac{1}{2}\log\det\left(\mathbf{I} + \frac{1}{\lambda^2}\Gamma_{\text{noise}}^{-\frac{1}{2}}\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{A}(\mathbf{m})^{-\top}\mathbf{P}^{\top}\Gamma_{\text{noise}}^{-\frac{\top}{2}}\right)$$

$$\phi_2(\mathbf{m}) = \frac{1}{2}\left(\lambda^2\|\mathbf{A}(\mathbf{m})\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{q}\|^2 + \|\mathbf{P}\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{d}\|^2_{\Gamma_{\text{noise}}^{-1}} + \|\mathbf{m} - \tilde{\mathbf{m}}\|^2_{\Gamma_{\text{prior}}^{-1}}\right).$$

when $\lambda \rightarrow 0$, $\overline{\mathbf{u}}(\mathbf{m})$ tends to fit the data and

$$\phi_2(\mathbf{m}) \rightarrow \frac{1}{2}\|\mathbf{m} - \tilde{\mathbf{m}}\|^2_{\Gamma_{\text{prior}}^{-1}},$$
$$\phi_1(\mathbf{m}) \rightarrow \infty.$$

when $\lambda \rightarrow \infty$,

$$\phi_1(\mathbf{m}) \rightarrow 0$$
$$\phi_2(\mathbf{m}) \rightarrow \underbrace{\frac{1}{2}\|\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}\mathbf{q} - \mathbf{d}\|^2_{\Gamma_{\text{noise}}^{-1}} + \frac{1}{2}\|\mathbf{m} - \tilde{\mathbf{m}}\|^2_{\Gamma_{\text{prior}}^{-1}}}_{\text{Strict constraint}}$$

[van Leeuwen, T and Herrmann, F J , 2015]

SLIM

Selection of λ

Select λ according to the largest eigenvalue μ_1 of the matrix

$\mathbf{A}(\mathbf{m})^{-\top}\mathbf{P}^{\top}\Gamma_{\text{noise}}^{-1}\mathbf{P}\mathbf{A}(\mathbf{m})^{-1}$:

(1) $\lambda^2 \gg \mu_1$, λ is large;

(2) $\lambda^2 \ll \mu_1$, λ is small.

Selection of λ :

$$\lambda^2 = 0.01\mu_1.$$

Approximate the posterior distribution by:

$$\rho_{\text{post}}(\mathbf{m}|\mathbf{d}) \approx \overline{\rho}_{\text{post}}(\mathbf{m}|\mathbf{d})$$
$$\propto \exp\Big(\underbrace{-\frac{\lambda^2}{2}\|\mathbf{A}(\mathbf{m})\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{q}\|^2 - \frac{1}{2}\|\mathbf{P}\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{d}\|^2_{\Gamma_{\text{noise}}^{-1}}}_{\textbf{Likelihood } \rho_{\text{like}}} \underbrace{- \frac{1}{2}\|\mathbf{m} - \tilde{\mathbf{m}}\|^2_{\Gamma_{\text{prior}}^{-1}}}_{\textbf{Prior } \rho_{\text{prior}}}\Big).$$

Gaussian Approximation

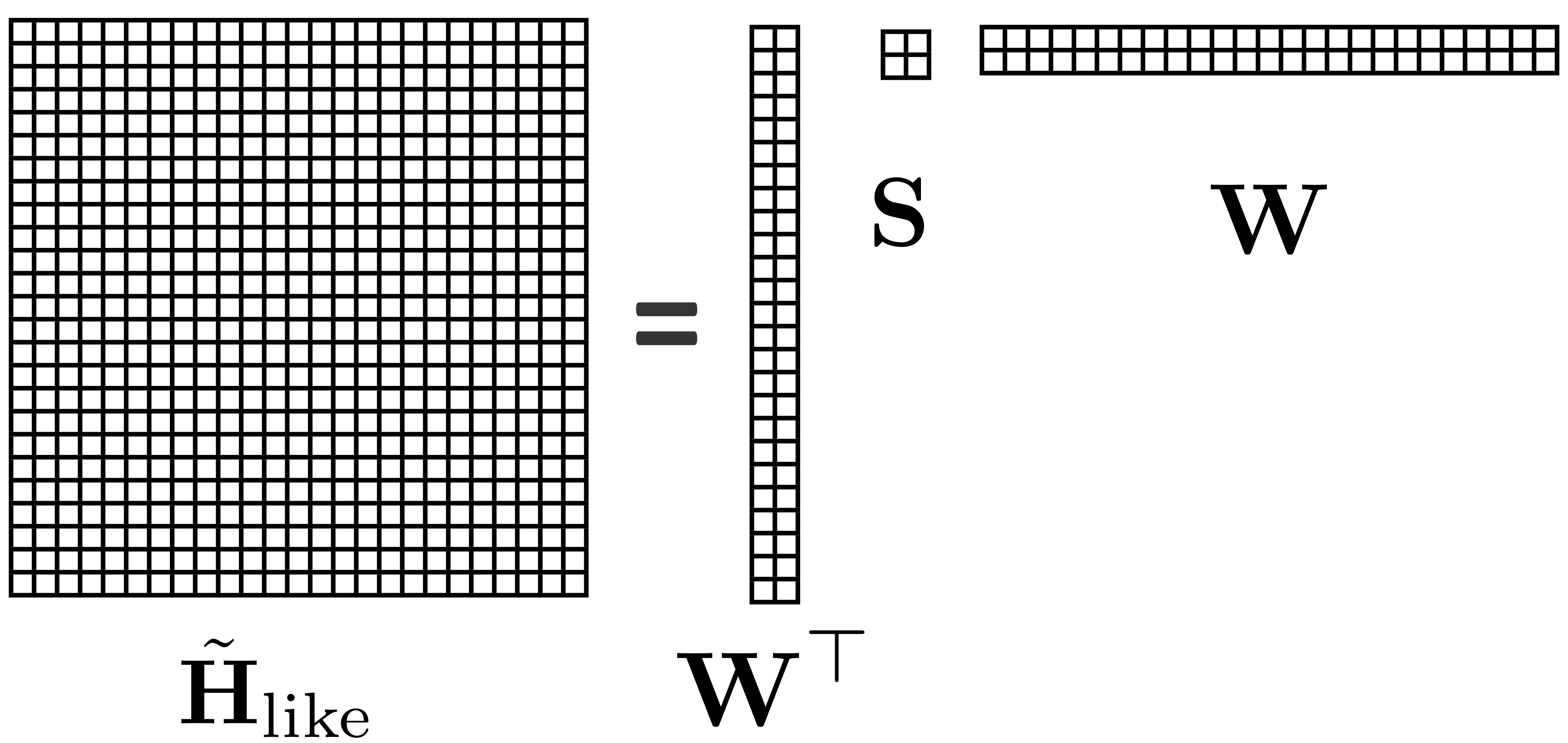
Approximate the posterior distribution by a Gaussian distribution:

$$\overline{\rho}_{\mathrm{post}}(\mathbf{m}|\mathbf{d}) \approx \rho_{\mathrm{Gauss}}(\mathbf{m}) = \mathcal{N}(\mathbf{m}_*, \tilde{\mathbf{H}}_{\mathrm{post}}^{-1}) = \mathcal{N}(\mathbf{m}_*, (\tilde{\mathbf{H}}_{\mathrm{like}} + \Gamma_{\mathrm{prior}}^{-1})^{-1}),$$

with the maximum a posterior estimate $\mathbf{m}_*$, and the Gauss-Newton Hessian $\tilde{\mathbf{H}}_{\mathrm{like}}$ of the negative logarithm likelihood function $-\log \rho_{\mathrm{like}}(\mathbf{m})$ given by:

$$\tilde{\mathbf{H}}_{\mathrm{like}} = \underbrace{\boxed{\mathbf{G}^\top}\boxed{\mathbf{A}^{-\top}\mathbf{P}^\top}}_{\mathbf{W}^\top}\underbrace{(\Gamma_{\mathrm{noise}} + \frac{1}{\lambda^2}\mathbf{P}\mathbf{A}^{-1}\mathbf{A}^{-\top}\mathbf{P}^\top)^{-1}}_{\mathbf{S}}\mathbf{P}\mathbf{A}^{-1}\mathbf{G}.$$

where $\mathbf{G} = \dfrac{\partial \mathbf{A}\overline{\mathbf{u}}}{\partial \mathbf{m}}$.



$$\tilde{\mathbf{H}}_{\mathrm{like}} = \mathbf{W}^\top \mathbf{S} \mathbf{W}$$

$$= (\mathbf{S}^{\frac{1}{2}}\mathbf{W})^\top \mathbf{S}^{\frac{1}{2}}\mathbf{W}$$

$$= \mathbf{R}_{\mathrm{like}}^\top \mathbf{R}_{\mathrm{like}}$$

Computational cost:

$n_{\mathrm{freq}} \times (n_{\mathrm{src}} + n_{\mathrm{rcv}})$

storage cost:

$n_{\mathrm{freq}} \times n_{\mathrm{grid}} \times (n_{\mathrm{src}} + n_{\mathrm{rcv}})$

in parallel !!

[J. M. Bardsley *et al*, 2014]

Sampling method

Sample Gaussian distribution:

$$\overline{\rho}_{\text{post}}(\mathbf{m}|\mathbf{d}) \approx \rho_{\text{Gauss}}(\mathbf{m}) = \mathcal{N}(\mathbf{m}_*, \tilde{\mathbf{H}}_{\text{post}}^{-1}) = \mathcal{N}(\mathbf{m}_*, (\tilde{\mathbf{H}}_{\text{like}} + \Gamma_{\text{prior}}^{-1})^{-1}).$$

Cholesky factorization method:

$$\tilde{\mathbf{H}}_{\text{post}} = \mathbf{R}_{\text{post}}^{\top}\mathbf{R}_{\text{post}},$$

$$\mathbf{m}_{\text{s}} = \mathbf{m}_* + \mathbf{R}^{-1}\mathbf{r}, \mathbf{r} \sim \mathcal{N}(0, \mathcal{I}).$$

Cons: 1. Explicit formulation of $\tilde{\mathbf{H}}_{\text{post}}$ requires a storage cost of $\mathcal{O}(n_{\text{grid}}^2)$.

2. Computational cost is $\mathcal{O}(n_{\text{grid}}^3)$.

Randomize then optimize (RTO) method:

Let $\tilde{\mathbf{H}}_{\text{like}} = \mathbf{R}_{\text{like}}^{\top}\mathbf{R}_{\text{like}}$, and $\Gamma_{\text{prior}} = \mathbf{R}_{\text{prior}}^{\top}\mathbf{R}_{\text{prior}}$, we have:

$$\mathbf{m}_{\text{s}} = \arg\min_{\mathbf{m}} \|\mathbf{R}_{\text{like}}\mathbf{m} - \mathbf{R}_{\text{like}}\mathbf{m}_* - \mathbf{r}_{\text{like}}\|^2 + \left\|\mathbf{R}_{\text{prior}}^{-\top}\mathbf{m} - \mathbf{R}_{\text{prior}}^{-\top}\mathbf{m}_* - \mathbf{r}_{\text{prior}}\right\|^2,$$

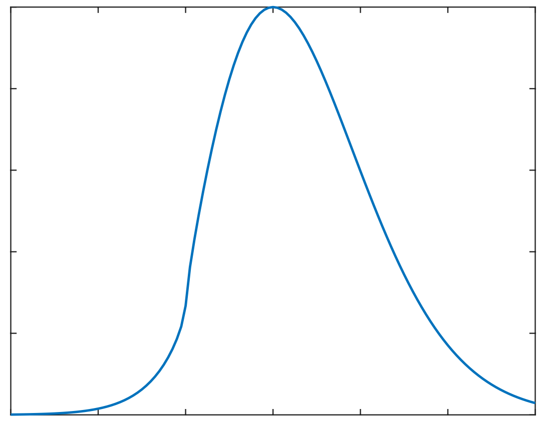
$$\mathbf{r}_{\text{like}} \sim \mathcal{N}(\mathbf{0}, \mathcal{I}_{n_{\text{data}} \times n_{\text{data}}}),$$

$$\mathbf{r}_{\text{prior}} \sim \mathcal{N}(\mathbf{0}, \mathcal{I}_{n_{\text{grid}} \times n_{\text{grid}}}).$$

Pros: 1. Do not need the explicit Hessian matrix;

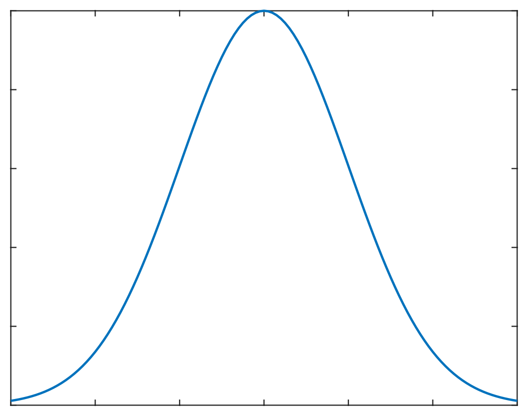
2. Can use iterative methods to solve it.

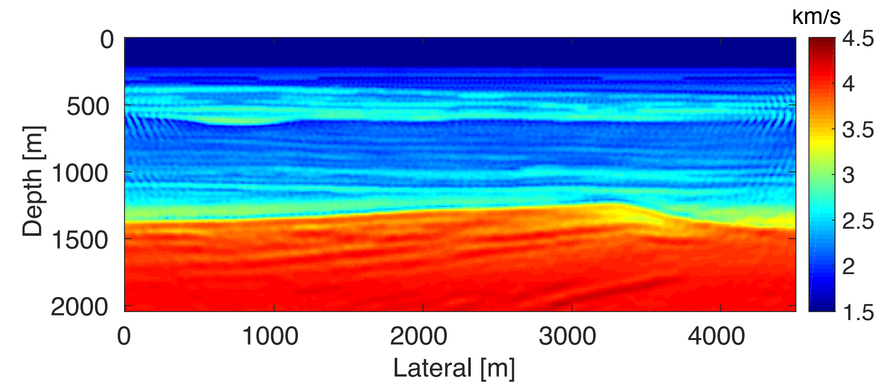
Workflow



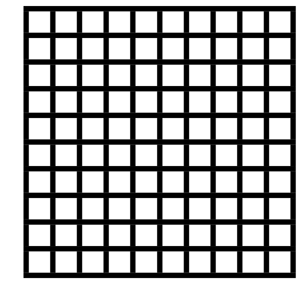
Likelihood distribution

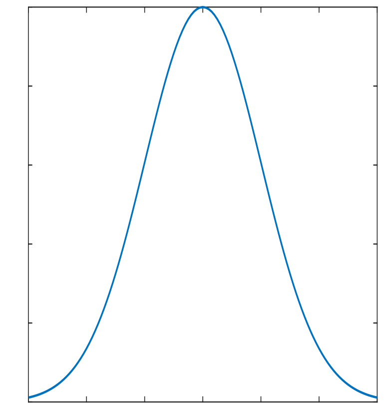
Prior distribution





MAP estimate

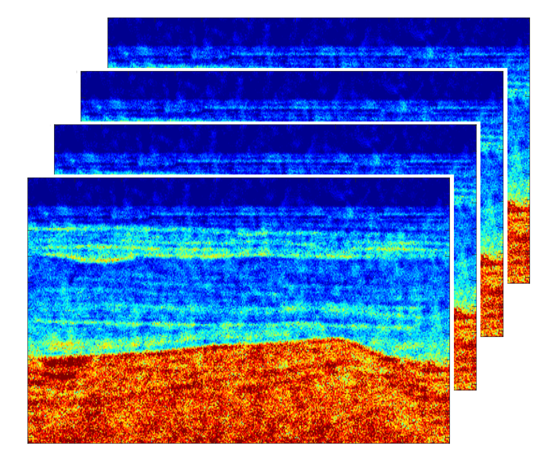




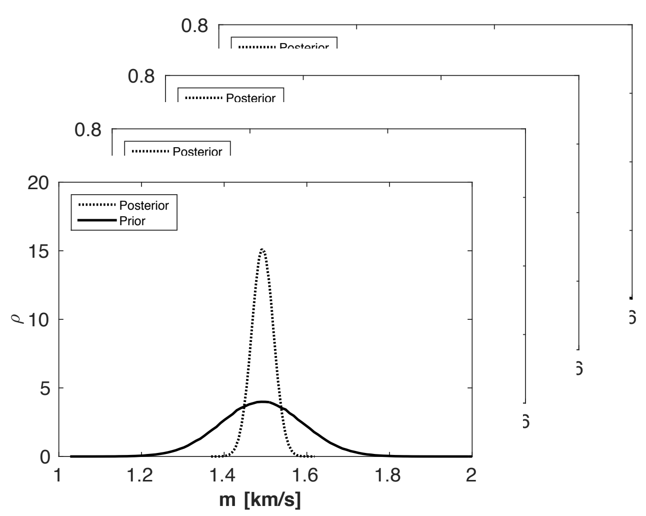
GN Hessian and Gaussian distribution

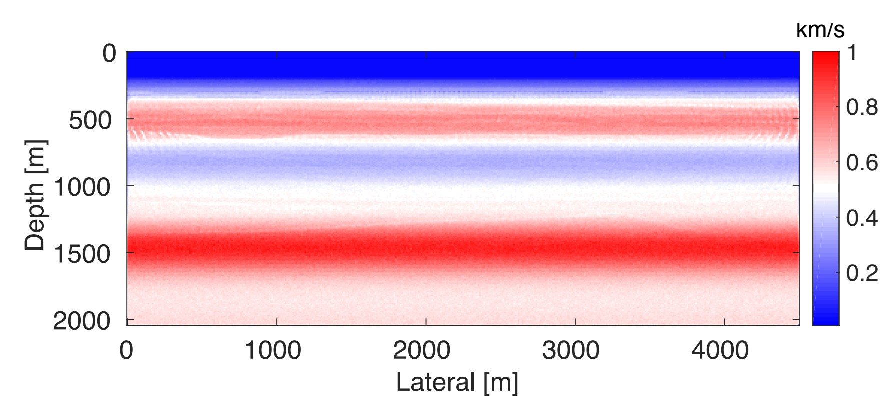
Gaussian approximation + RTO (GA-RTO)

Sample the Gaussian distribution by RTO



Compute statistics from samples

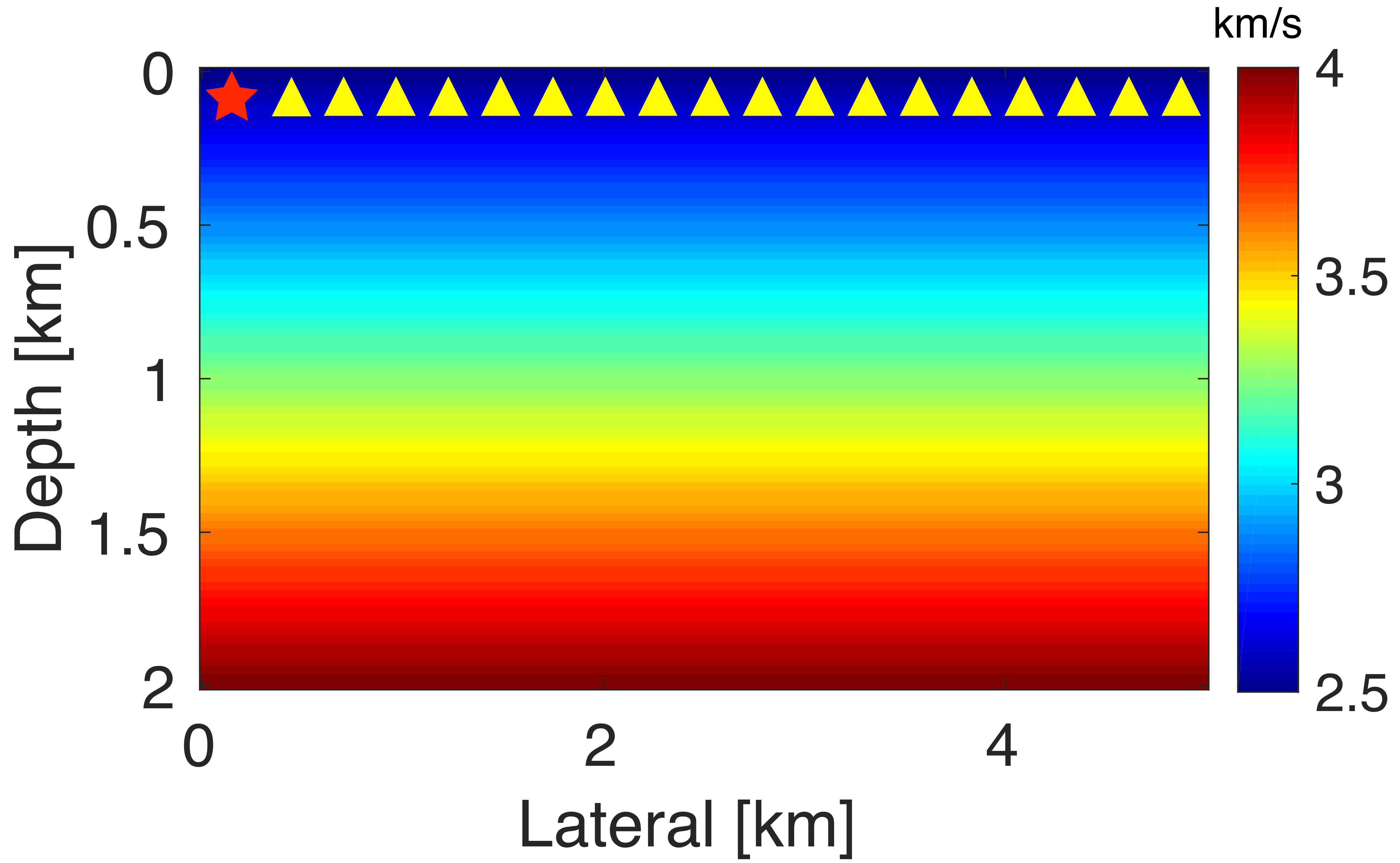




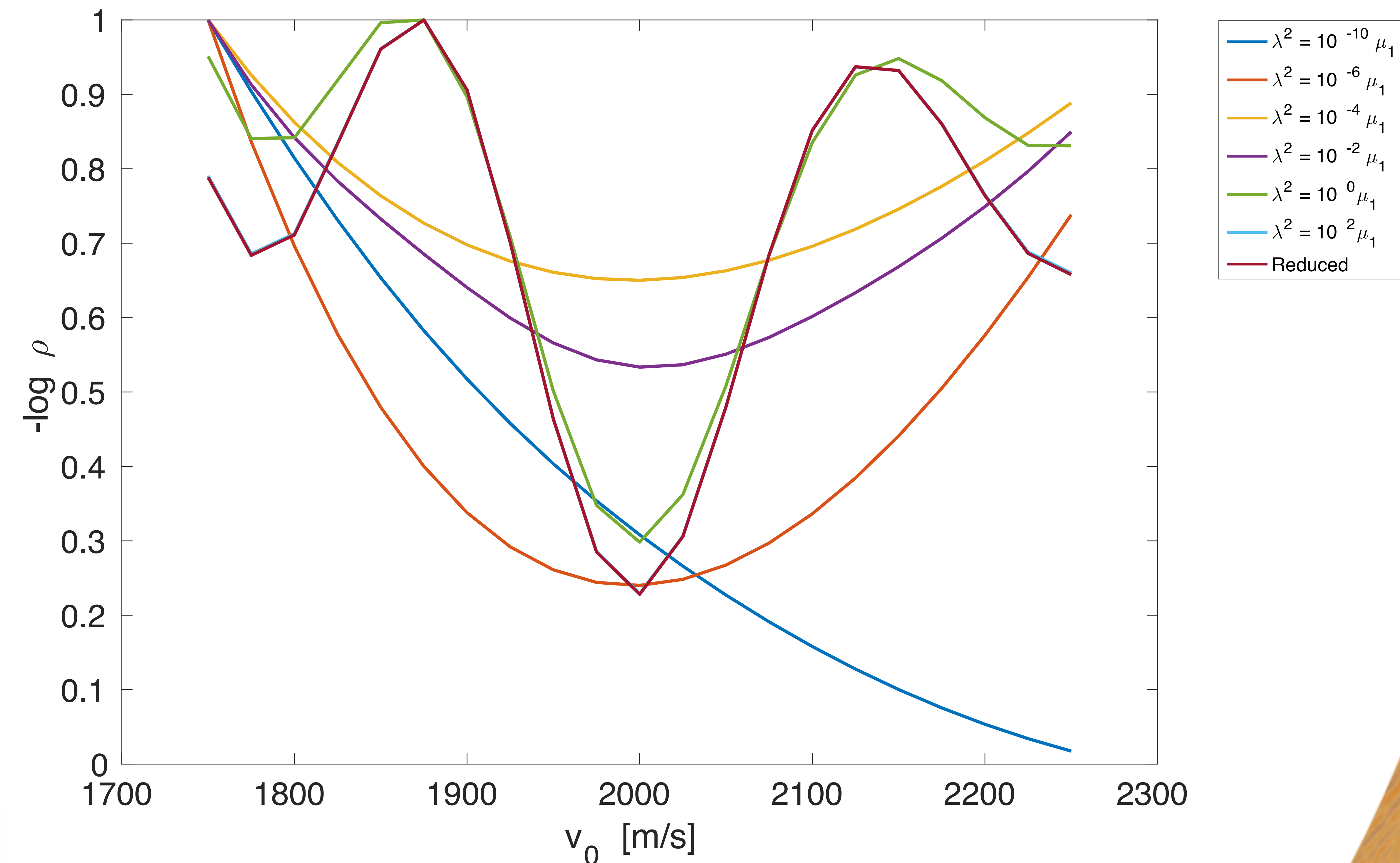
Numerical examples
— investigate λ

1D velocity model: $v(z) = v_0 + 0.75z\,\mathrm{km/s}, \quad \text{with} \quad v_0 = 2.5\,\mathrm{km/s}.$

10% Gaussian noise.



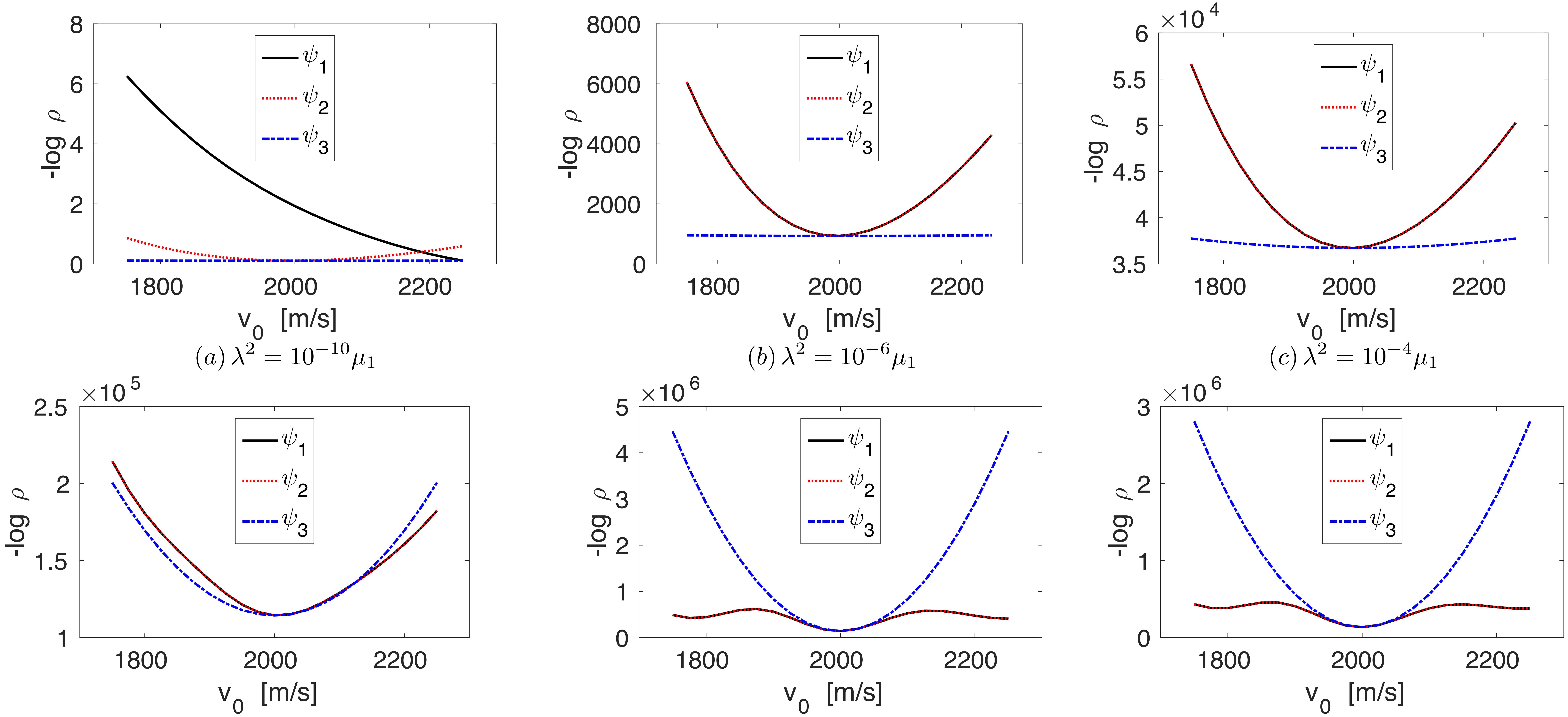
(a) 1D model



(b) Negative logarithm likelihood function

Numerical examples

— investigate λ



(a) $\lambda^2 = 10^{-10}\mu_1$ (b) $\lambda^2 = 10^{-6}\mu_1$ (c) $\lambda^2 = 10^{-4}\mu_1$

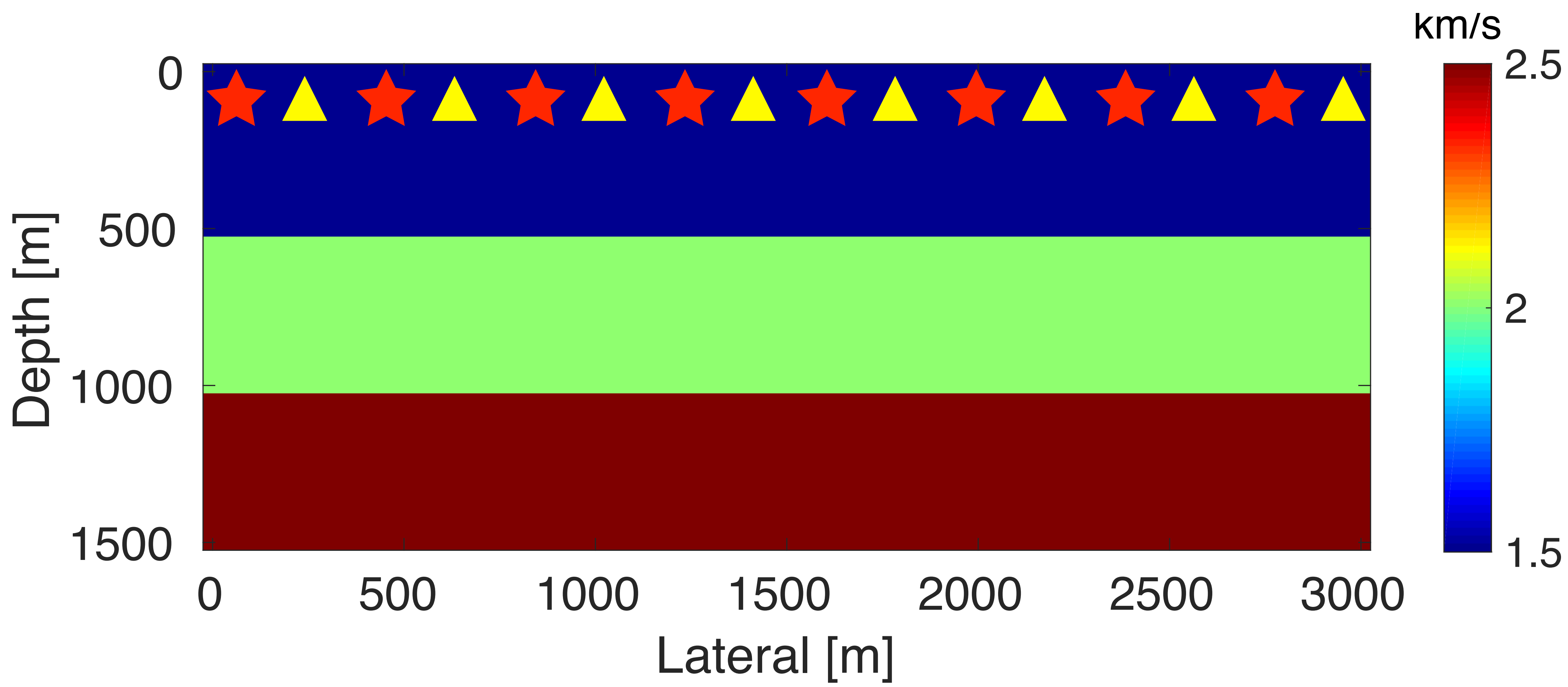
(d) $\lambda^2 = 10^{-2}\mu_1$ (e) $\lambda^2 = 10^{0}\mu_1$ (f) $\lambda^2 = 10^{2}\mu_1$

$$\text{Note} : \psi_1 = -\log \rho_{\text{like}},\ \psi_2 = -\log \overline{\rho}_{\text{like}},\ \psi_3 = -\log \rho_{\text{Gauss,like}}$$

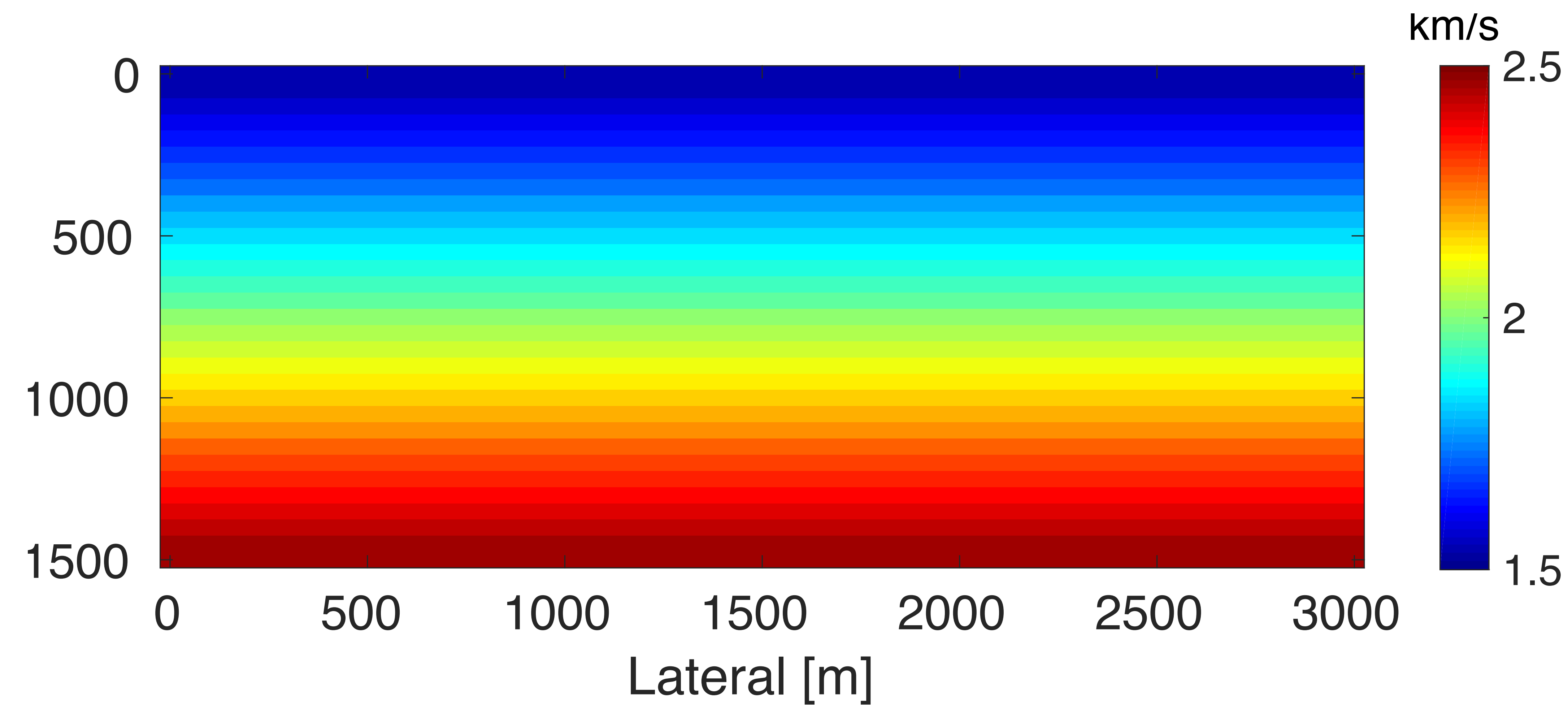
Numerical examples
— Layer model

Depth of sources and receivers: 50 m
Number of sources and receivers: 61
Penalty parameter: $\lambda^2 = 0.01\mu_{\mathbf{1}}$

Frequency: 5,6 and 7 Hz
15% Gaussian noise: $\Gamma_{\text{noise}}=175^2\mathbf{I}$



(a) True model



(b) Prior model

[J. Martin *et al*, 2012] [G. Matheron, 2012]

Numerical examples
—— Layer model

Covariance matrix of the prior distribution —— Gaussian smoothness:

$$\Gamma_{\text{prior}}(k,l) = a\exp\left(\frac{-\|\mathbf{s}_k - \mathbf{s}_l\|^2}{2b^2}\right) + c\delta_{k,l},$$

where the vectors $\mathbf{s}_k = (z_k, x_k)$ and $\mathbf{s}_l = (z_l, x_l)$ denote the k^{th} and l^{th} elements in the vector $\mathbf{m}$. Parameters a, b, and c control the correlation strength, variance, and spatial correlation distance.

In this example, our selections are as follows:

$$\begin{aligned} a &= 0.1\,\mathrm{km}^2/\mathrm{s}^2, \\ b &= 0.65, \\ c &= 0.01\,\mathrm{km}^2/\mathrm{s}^2. \end{aligned}$$

[J. M. Bardsley *et al*, 2014]

SLIM

Benchmark method

Randomized Maximum Likelihood - RML

Generate independent samples from $\tilde{\rho}_{\text{post}}(\mathbf{m})$ by solving:

$$\min_{\mathbf{m}} \frac{1}{2}(\|\mathbf{P}\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{d} - \mathbf{r}_{\text{d}}\|^2_{\Gamma^{-1}_{\text{noise}}} + \lambda^2\|\mathbf{A}(\mathbf{m})\overline{\mathbf{u}}(\mathbf{m}) - \mathbf{q} - \mathbf{r}_{\text{s}}\|^2) + \frac{1}{2}\|\mathbf{m} - \mathbf{m}_{\text{p}} - \mathbf{r}_{\text{p}}\|^2_{\Sigma^{-1}_{\text{prior}}},$$

where

$$\mathbf{r}_{\text{d}} \sim \mathcal{N}(0, \sigma^2\mathcal{I}_{n_{\text{rcv}}\times n_{\text{rcv}}}),$$

$$\mathbf{r}_{\text{s}} \sim \mathcal{N}(0, \lambda^{-2}\mathcal{I}_{n_{\text{grid}}\times n_{\text{grid}}}),$$

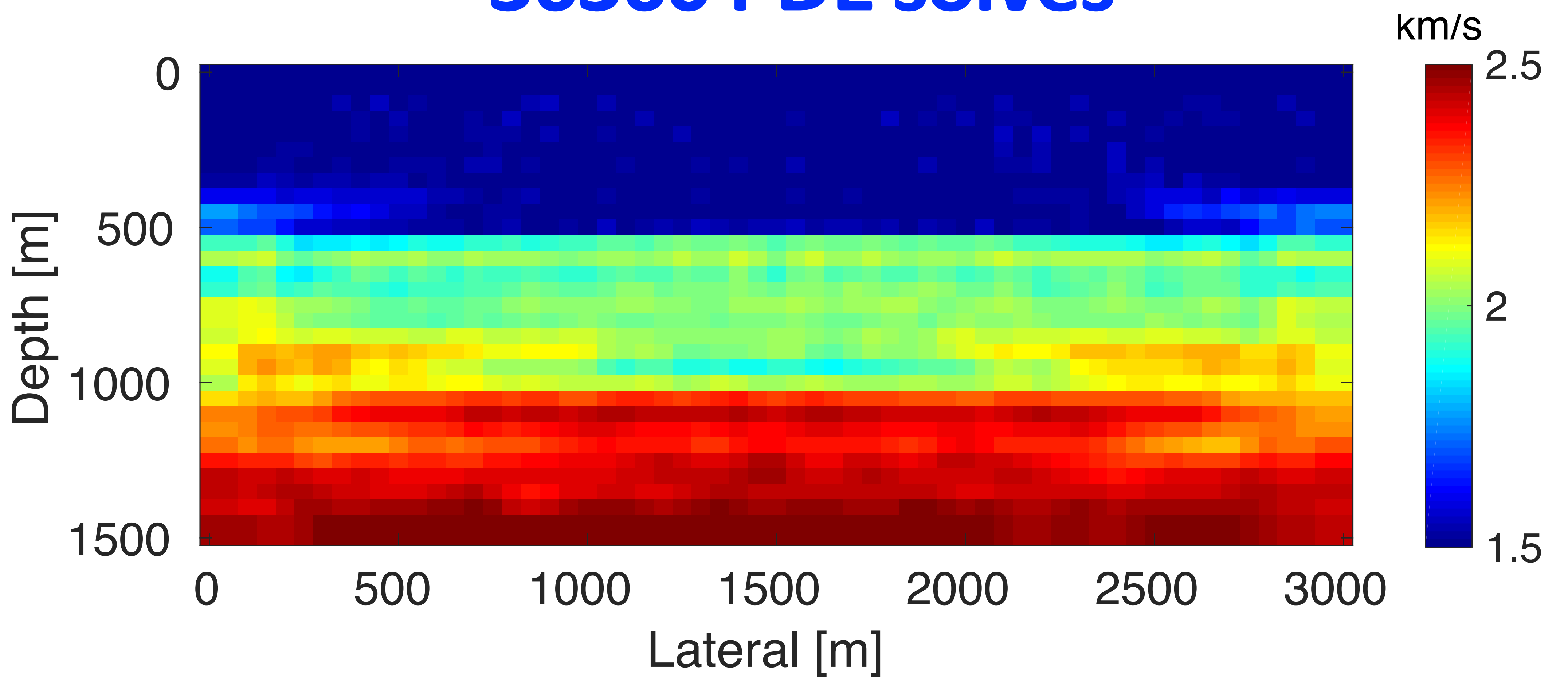
$$\mathbf{r}_{\text{p}} \sim \mathcal{N}(0, \Sigma_{\text{prior}}).$$

Require solving an FWI type problem to draw one sample.

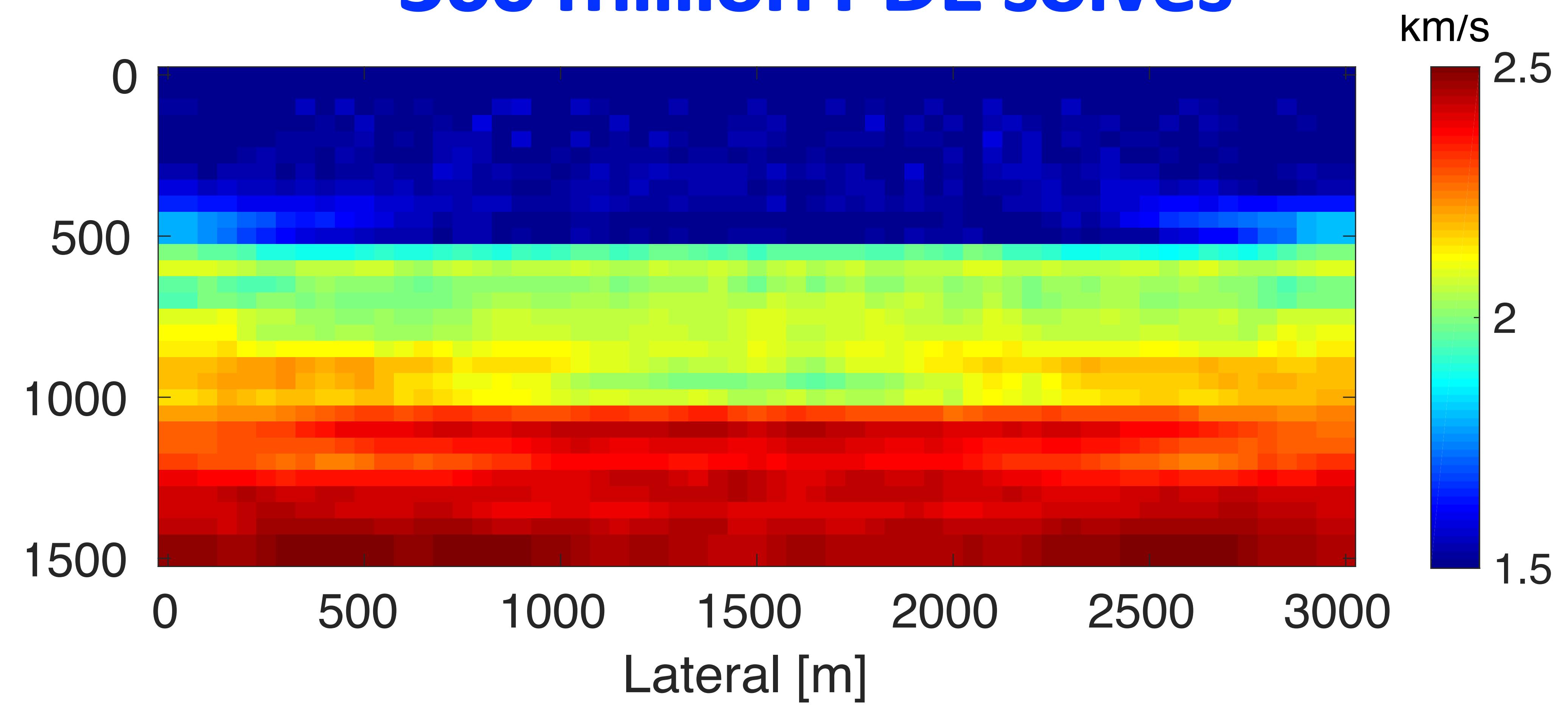
Posterior mean models and standard deviations of 10000 samples

SLIM

36360 PDE solves | **360 million PDE solves**

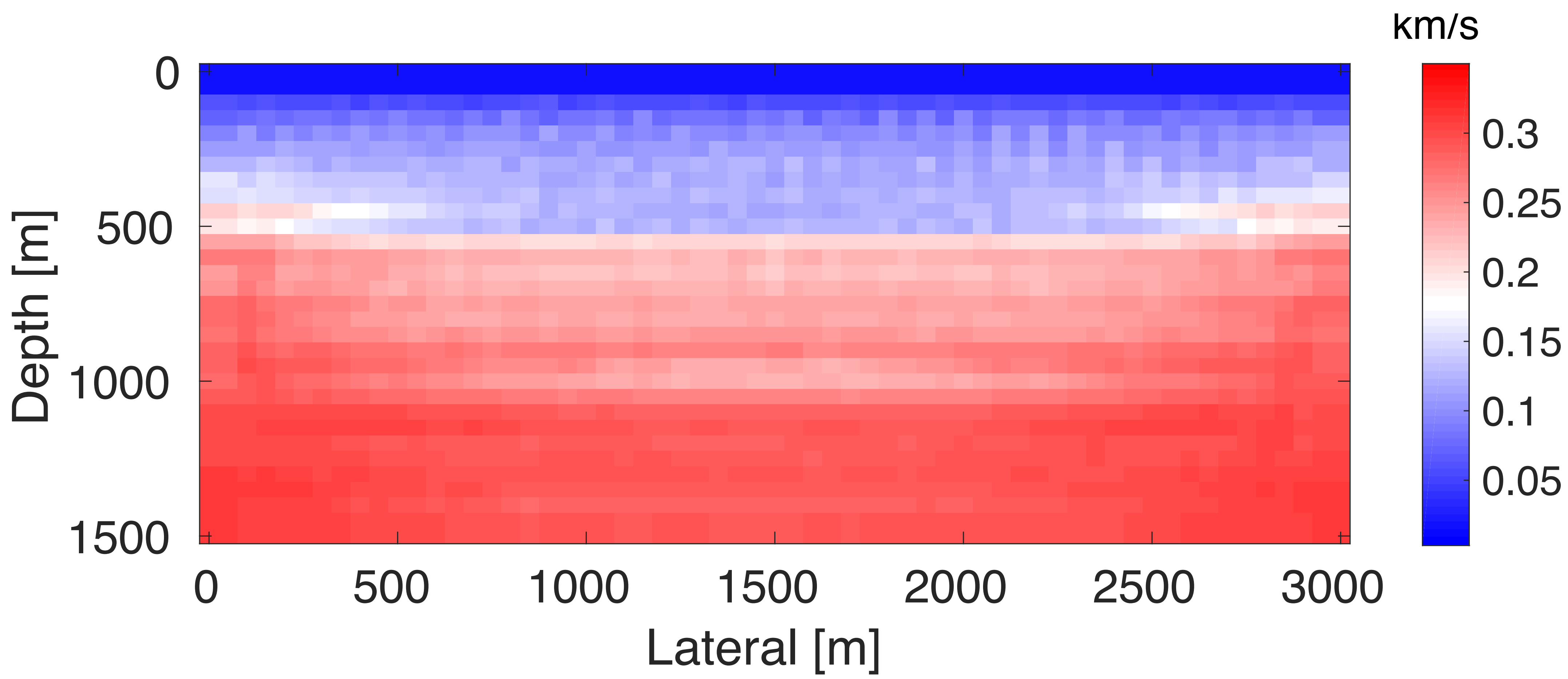


(a) Mean model by GA-RTO

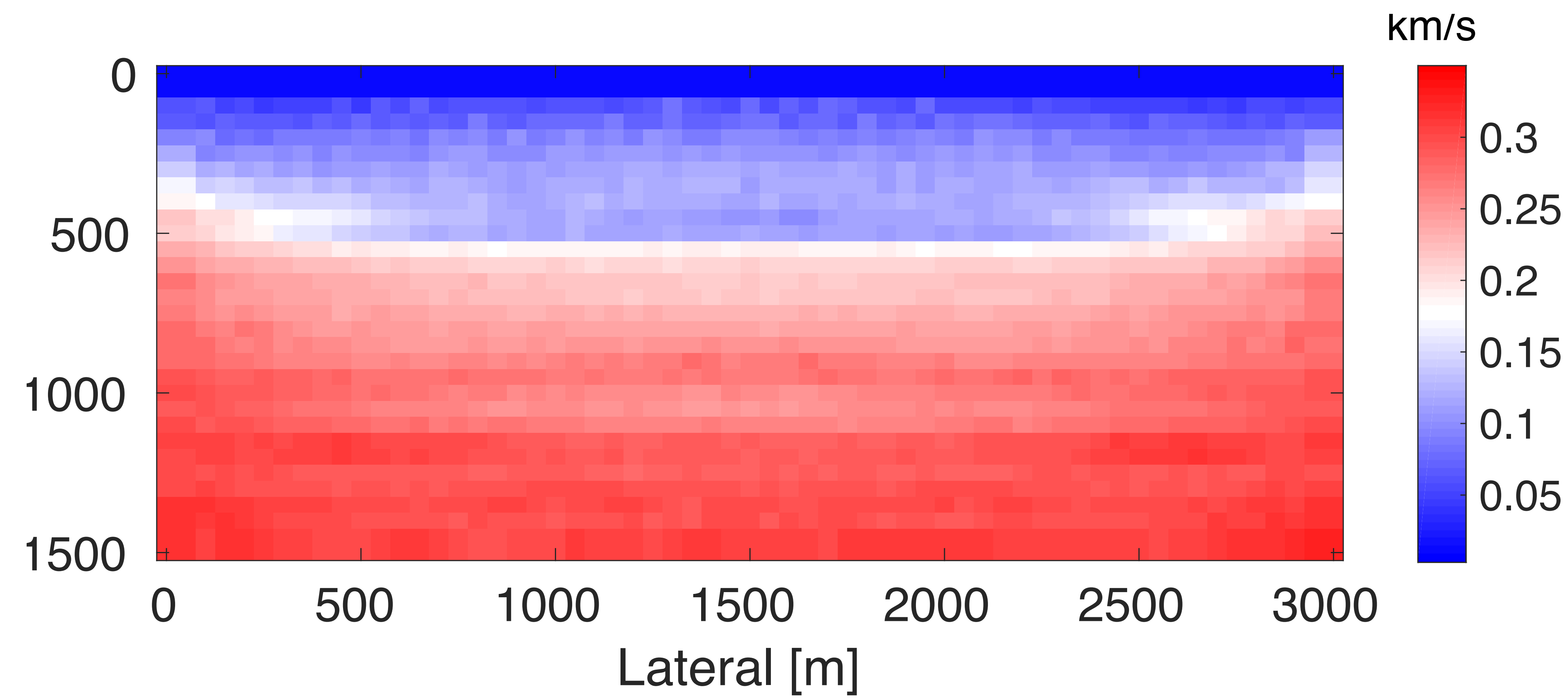


(b) Mean model by RML

1.5%



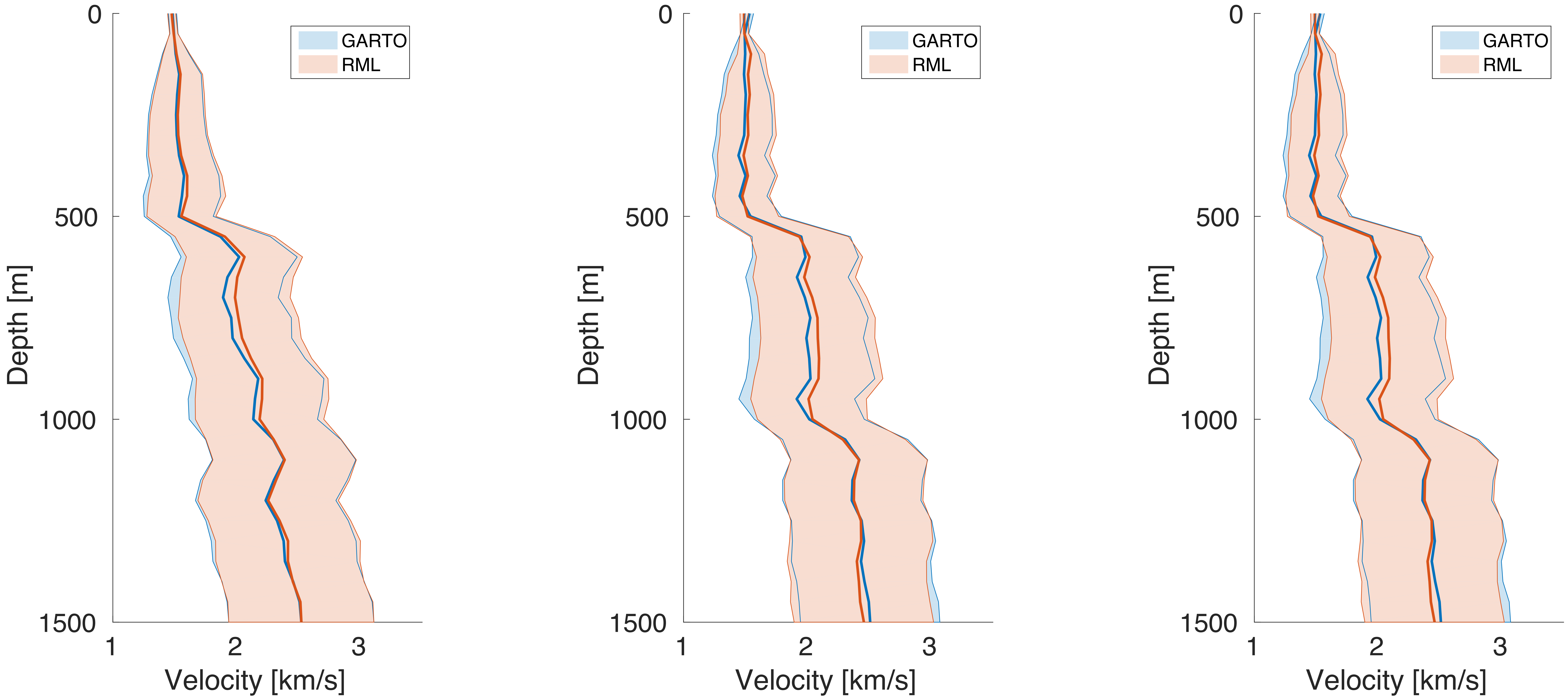
(c) Standard deviation by GA-RTO



(d) Standard deviation by RML

6%

95% Confidence intervals of 10000 samples



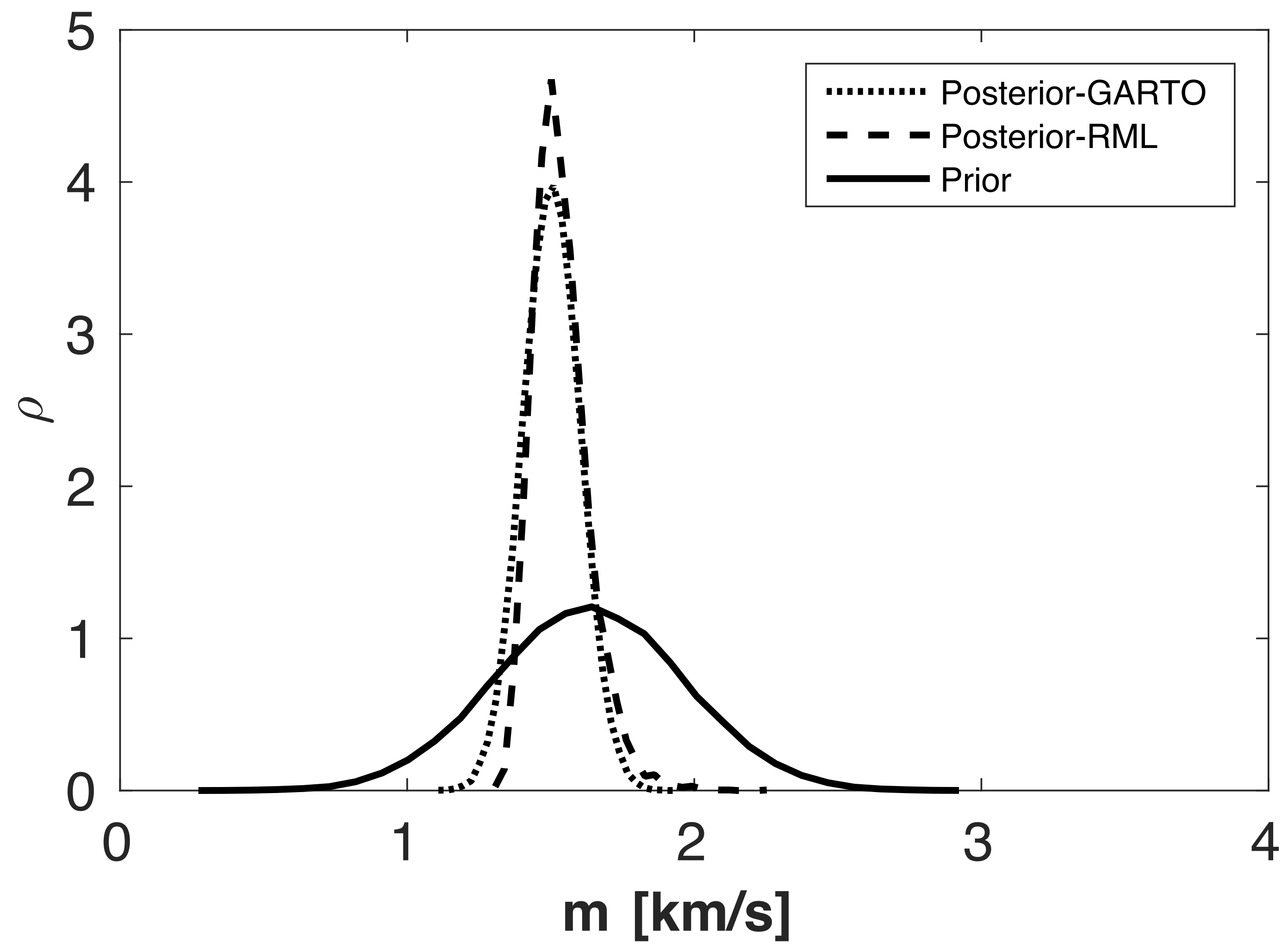
(a) x = 500 m

(b) x = 1500 m

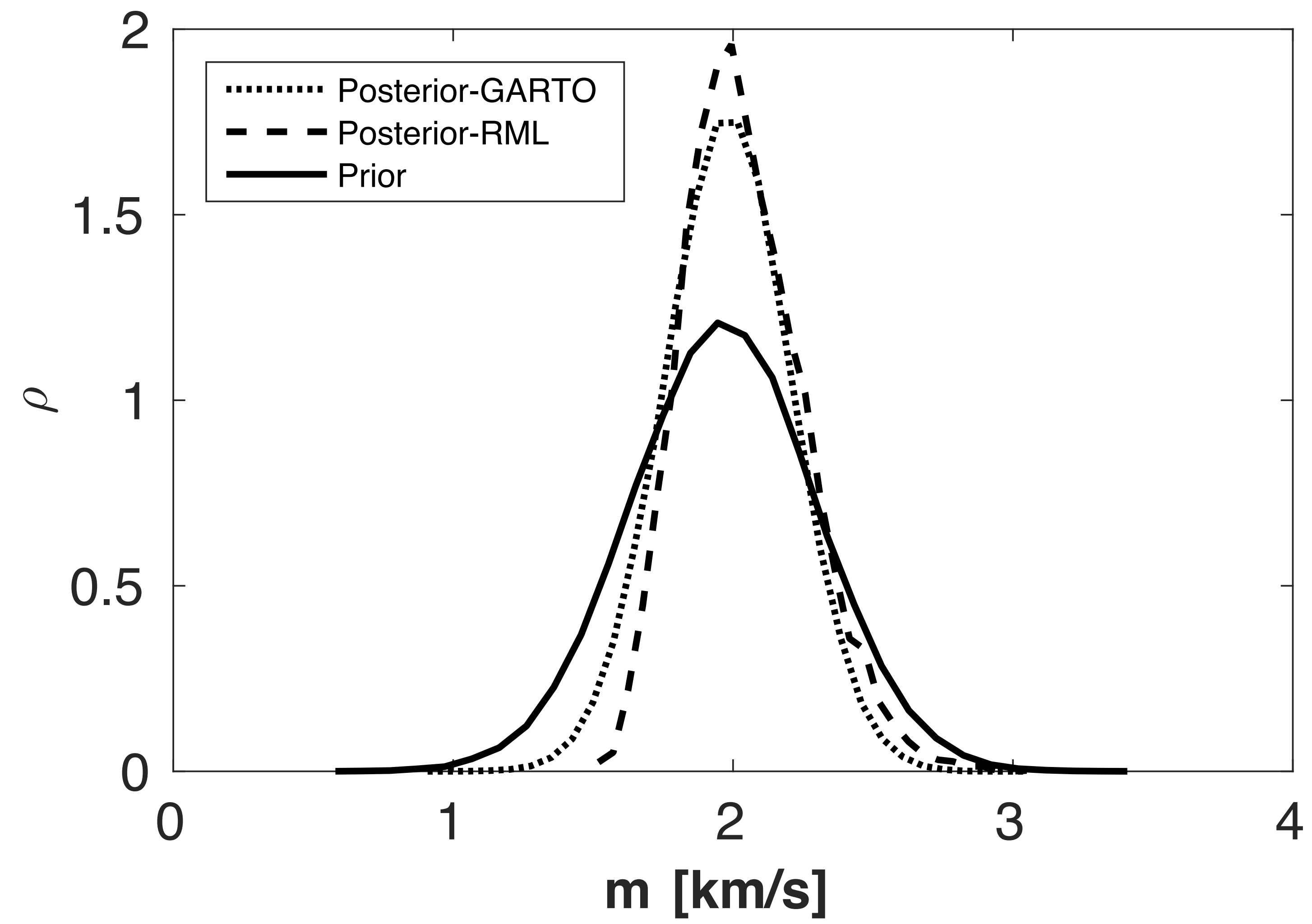
(c) x = 2500 m

58

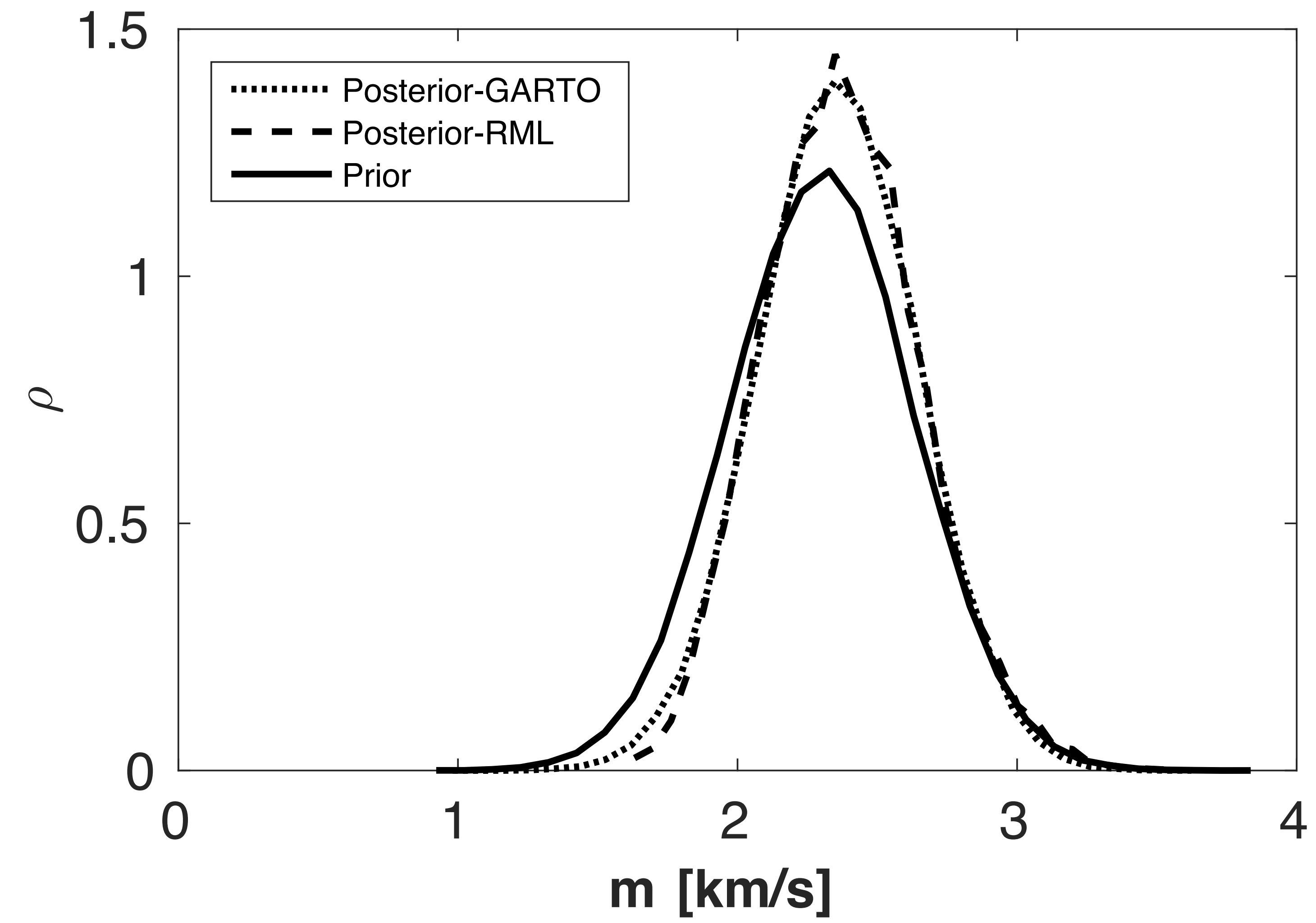
Marginal distributions of 10000 samples



(a) x = 1500 m, z = 200 m



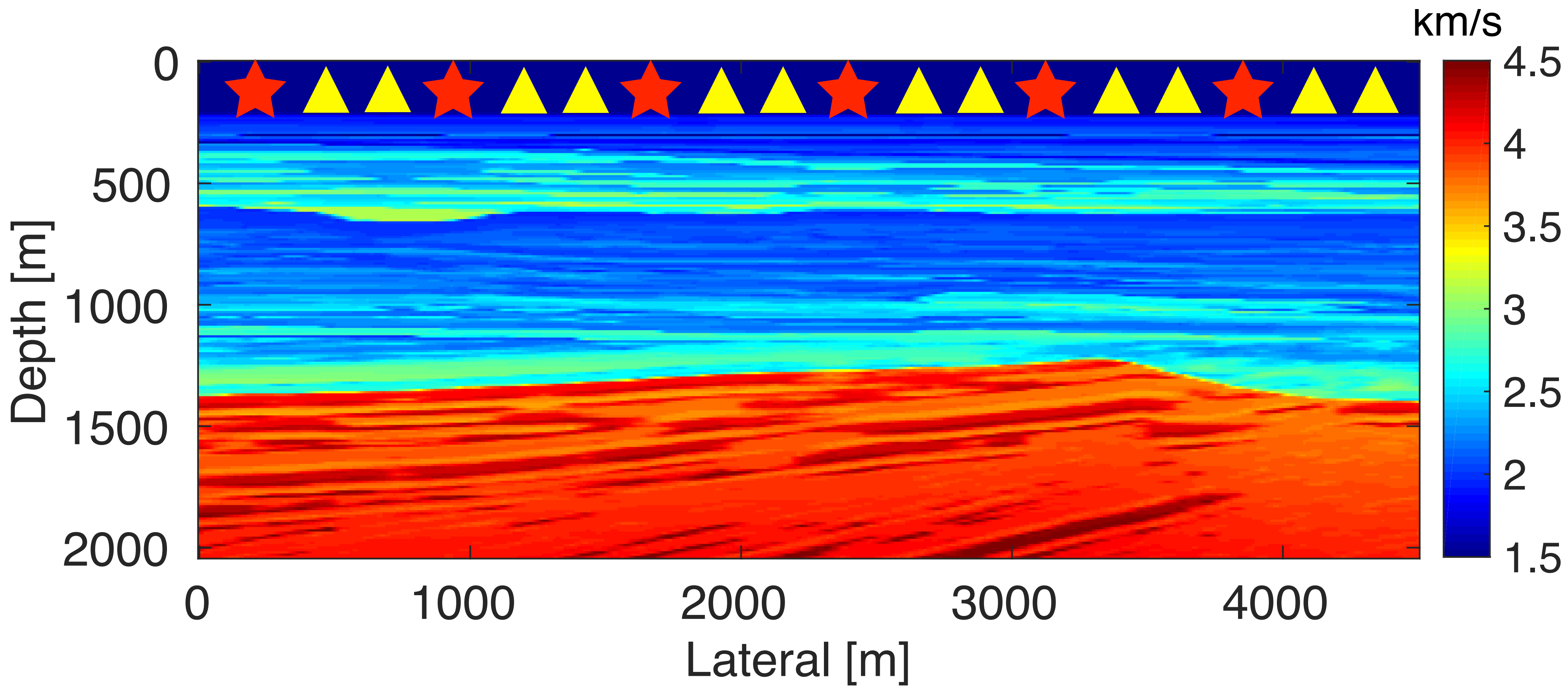
(b) x = 1500 m, z = 700 m



(c) x = 1500 m, z = 1200 m

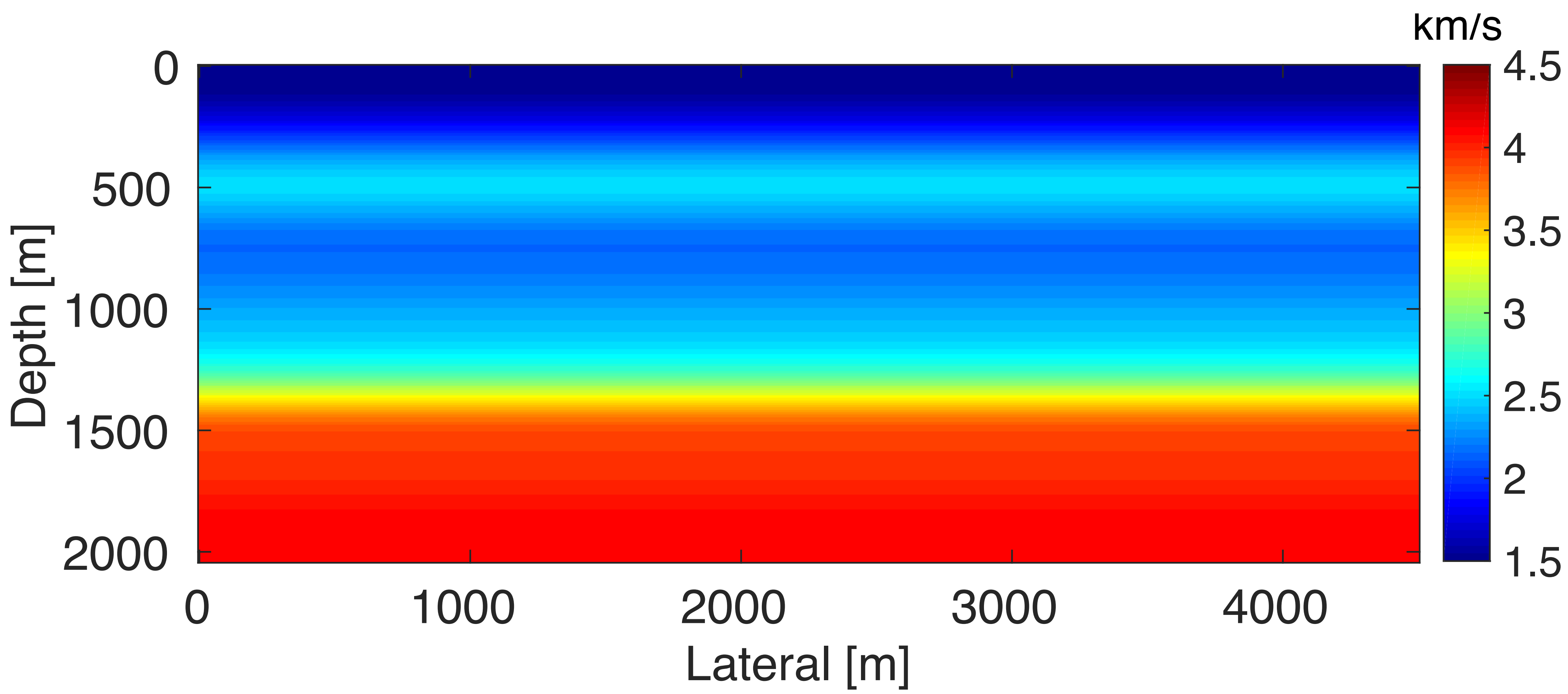
59

Numerical examples
— BG compass model

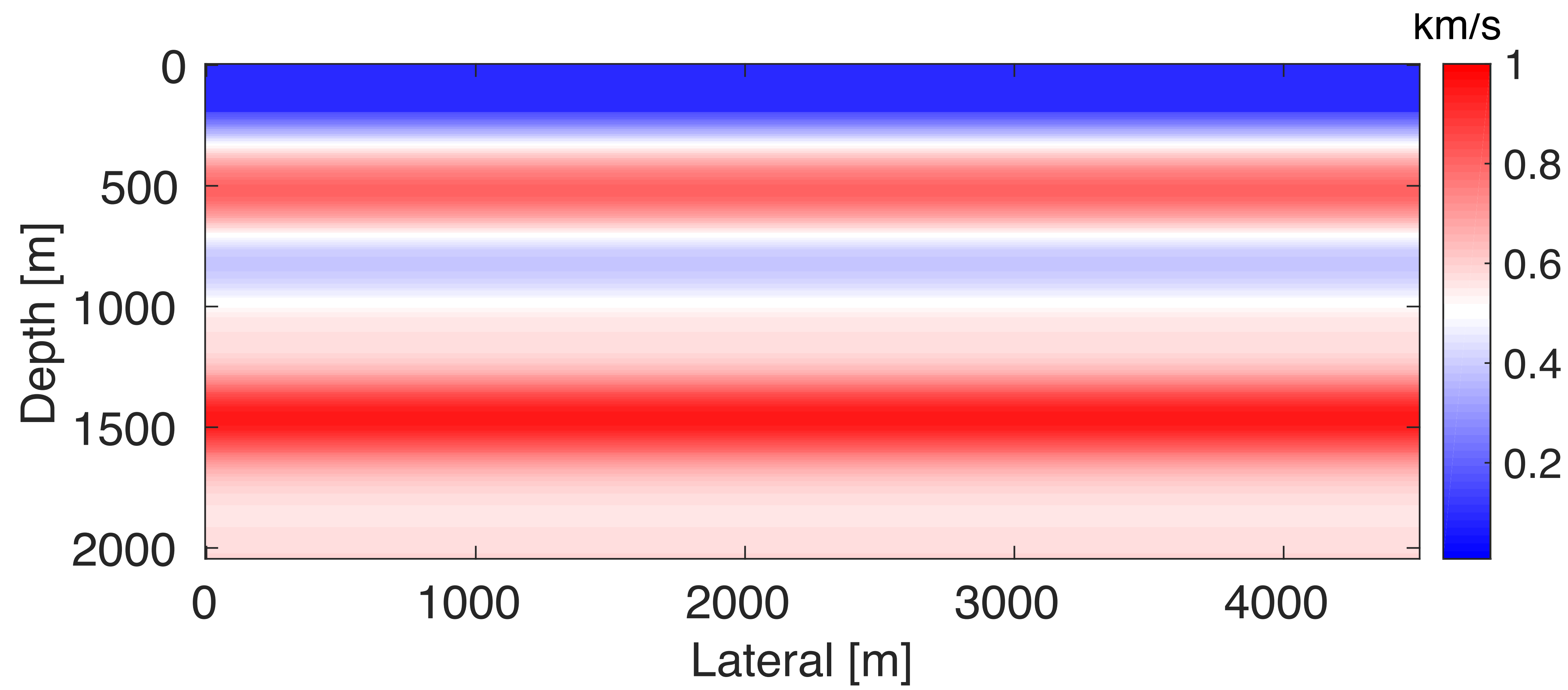


(a) True model

Depth of sources and receivers: 50
Number of sources and receivers: 91 / 451
Central frequency: 15 Hz
Frequency: 2-31 Hz
15% Gaussian noise: $\Gamma_{\text{noise}} := 46^2\mathbf{I}$
Penalty parameter: $\lambda^2 = 0.01\mu_{\mathbf{1}}$

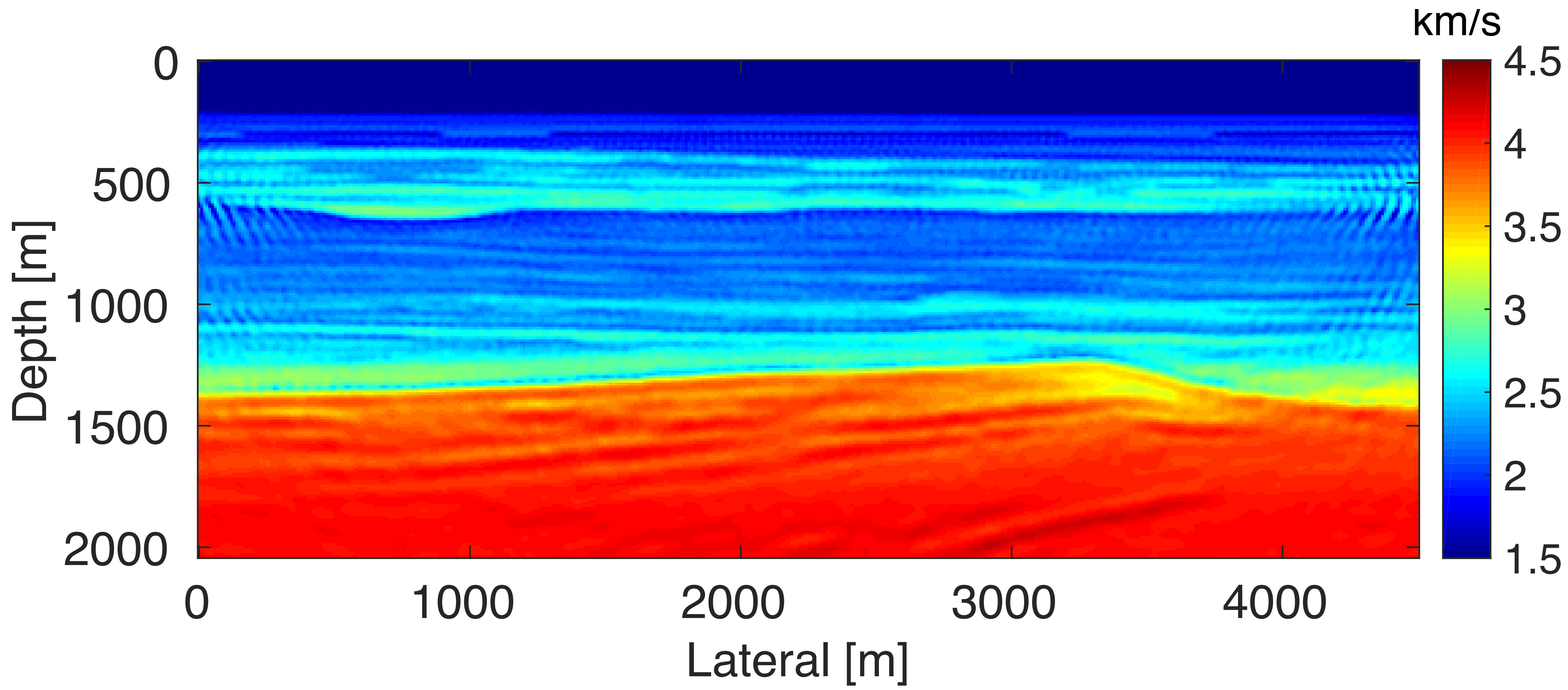


(b) Prior model



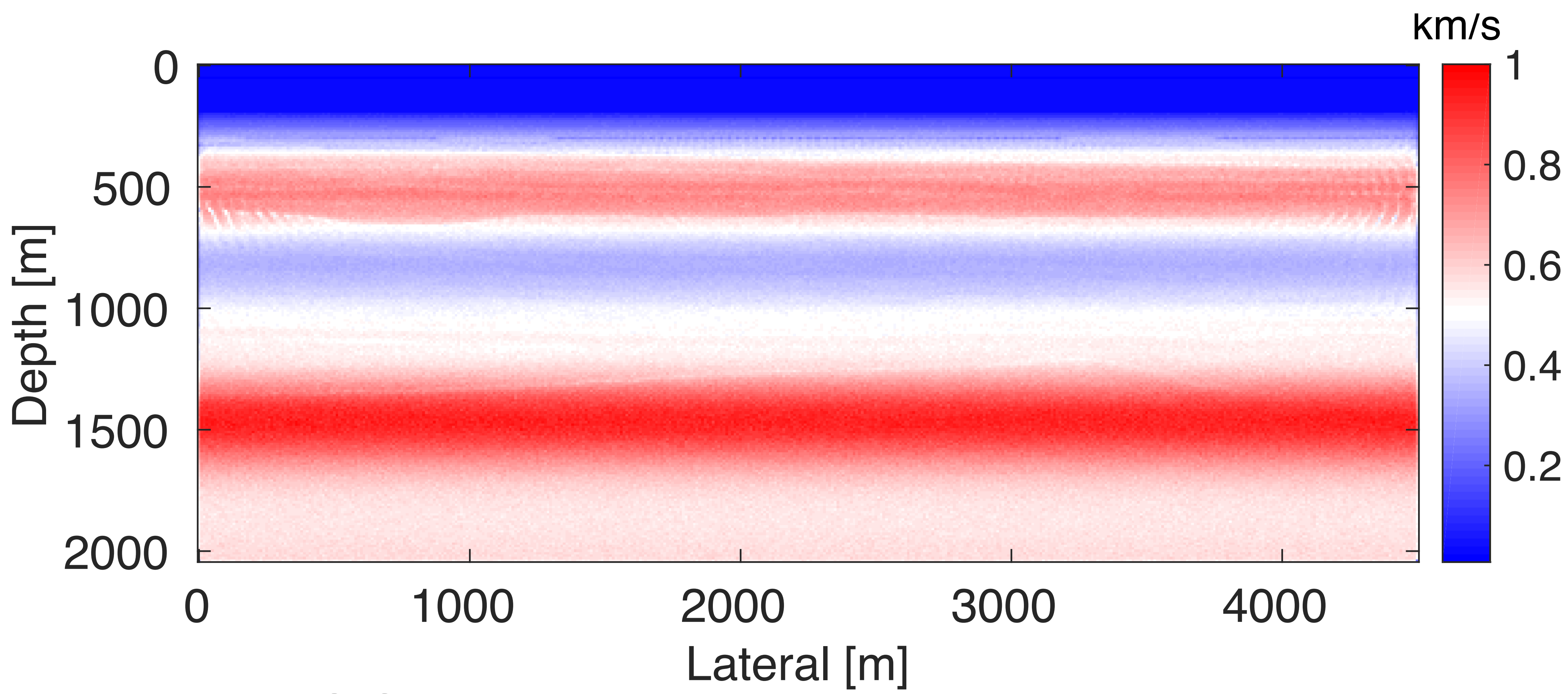
(c) STD of prior distribution

MAP and standard deviation of 2000 samples

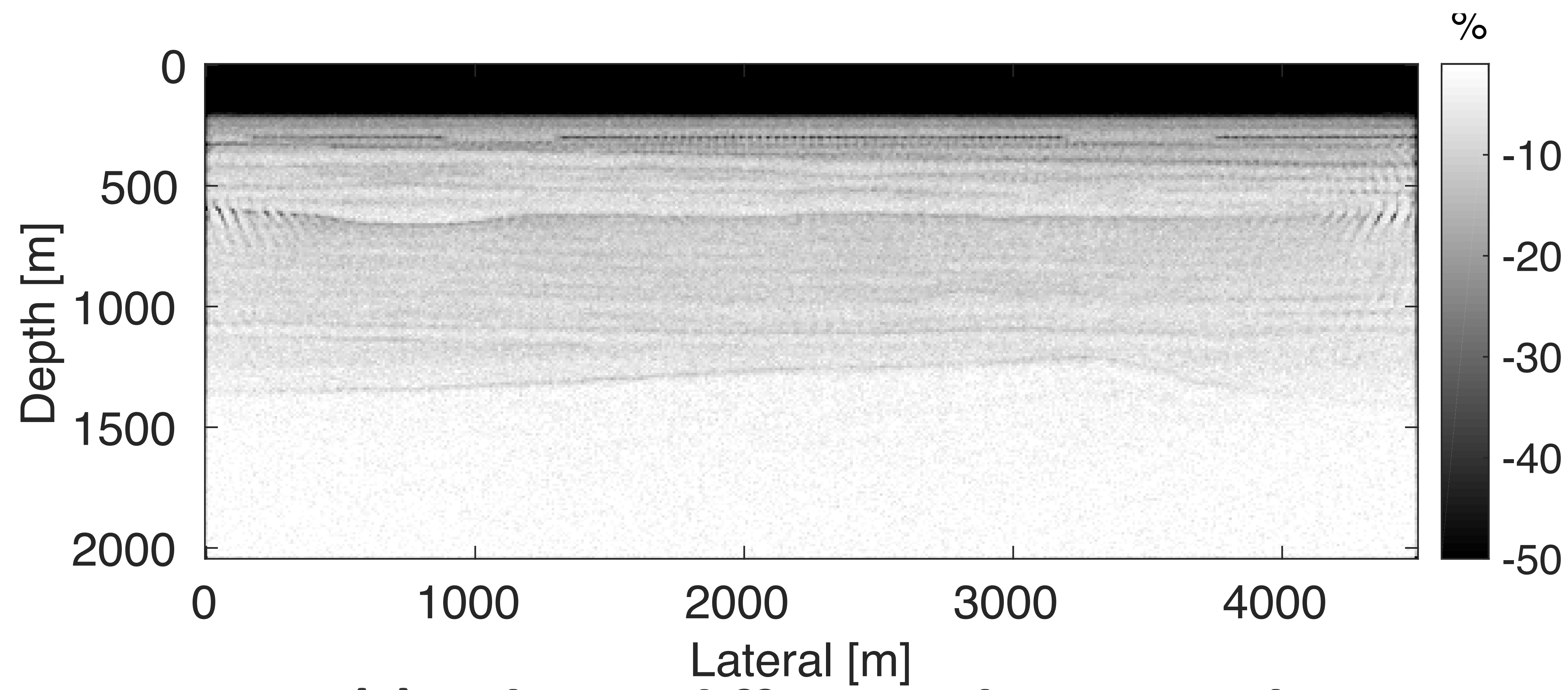


(a) MAP estimate

We first compute the MAP estimate, and then generate 2000 samples.

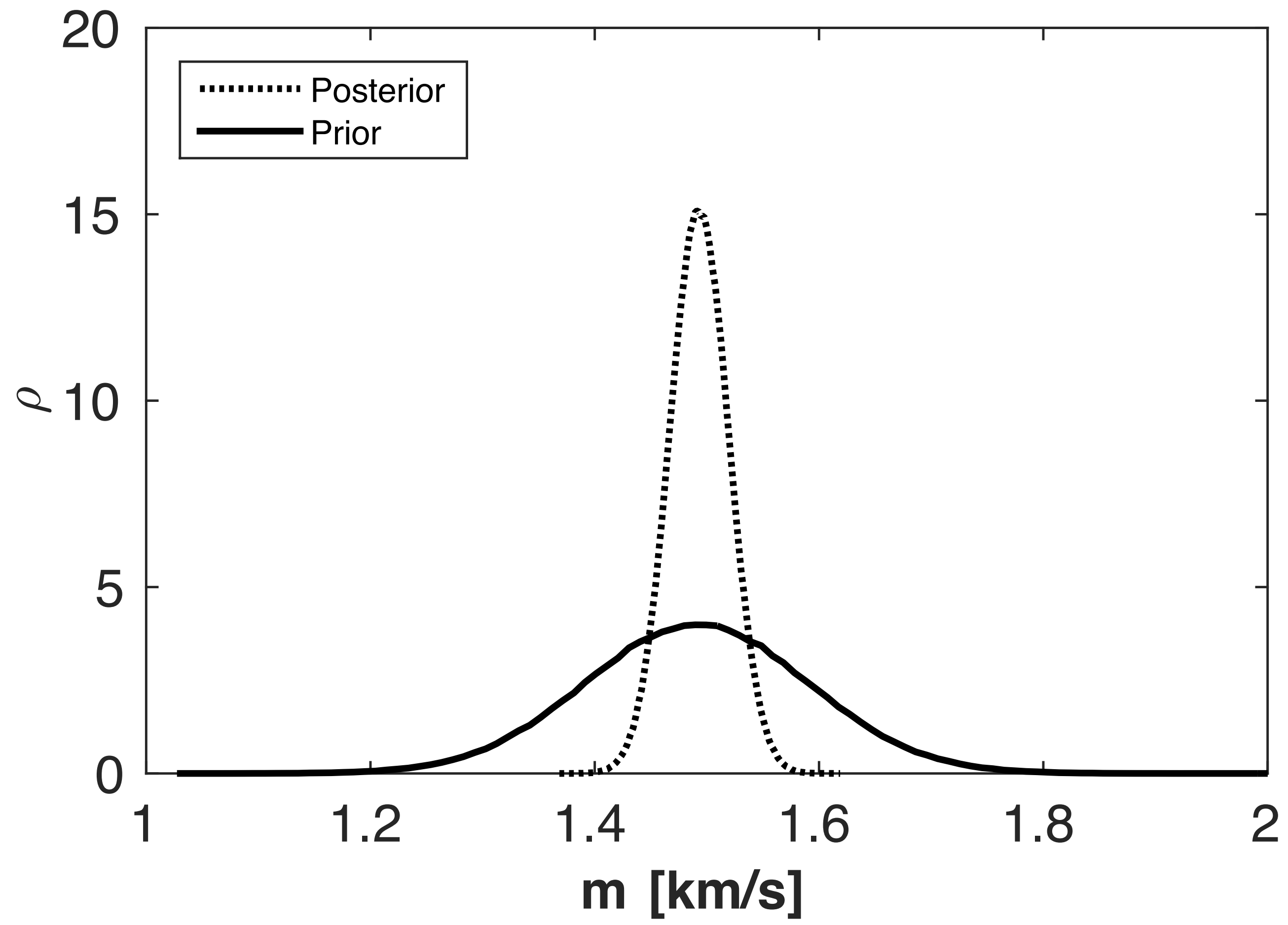


(b) Posterior standard deviation

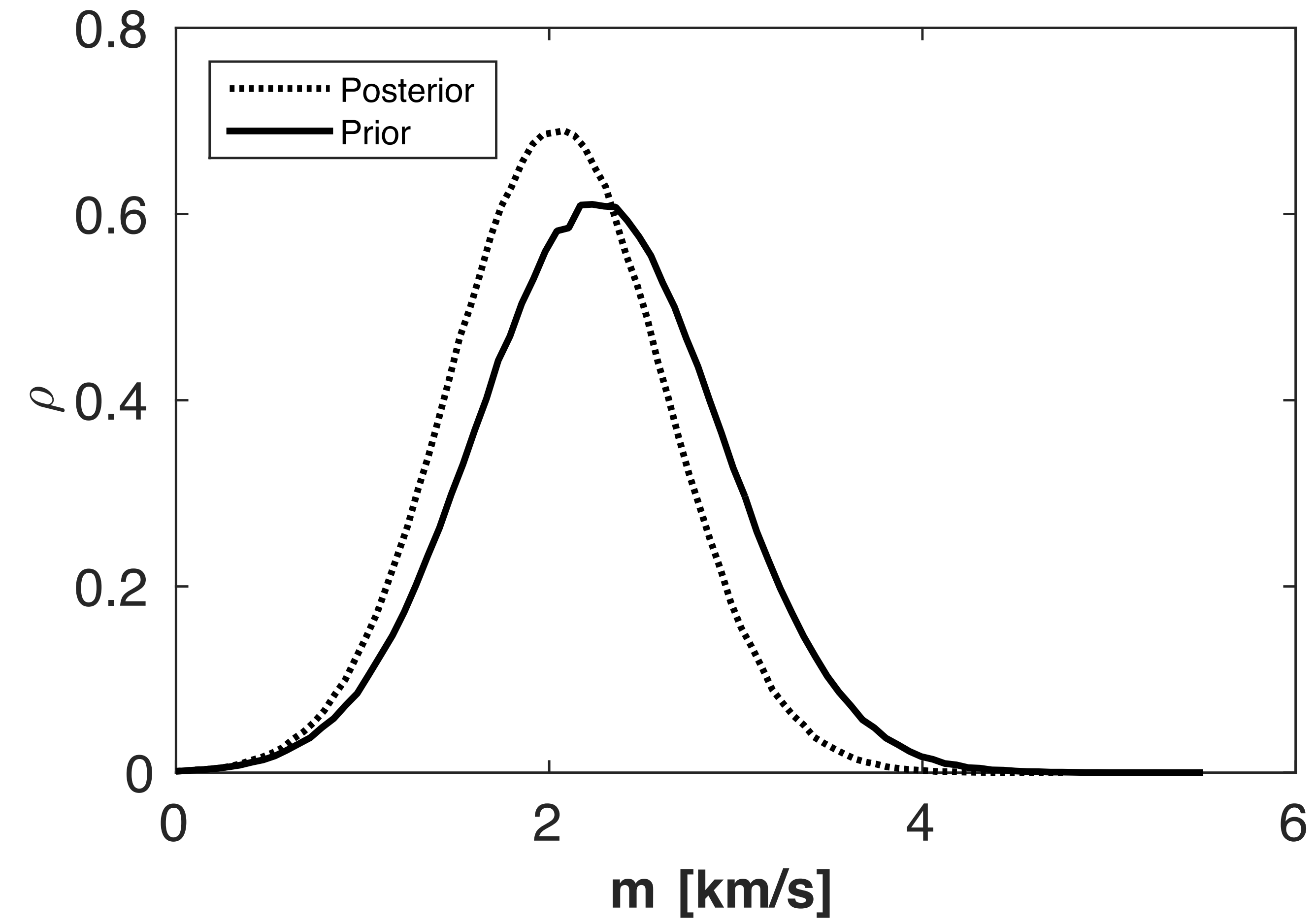


(c) Relative difference between the posterior and prior standard deviation

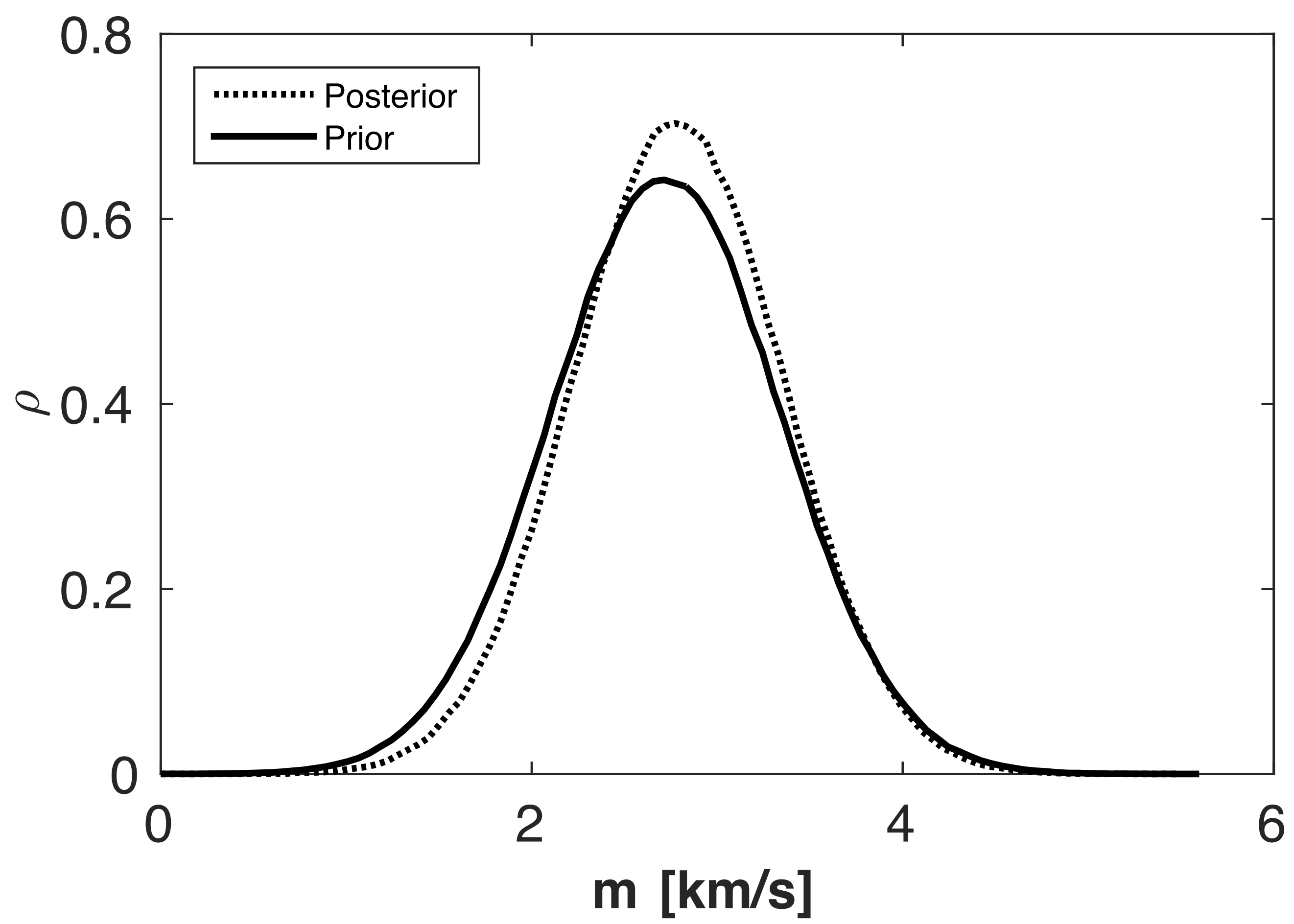
Marginal distributions of 2000 samples



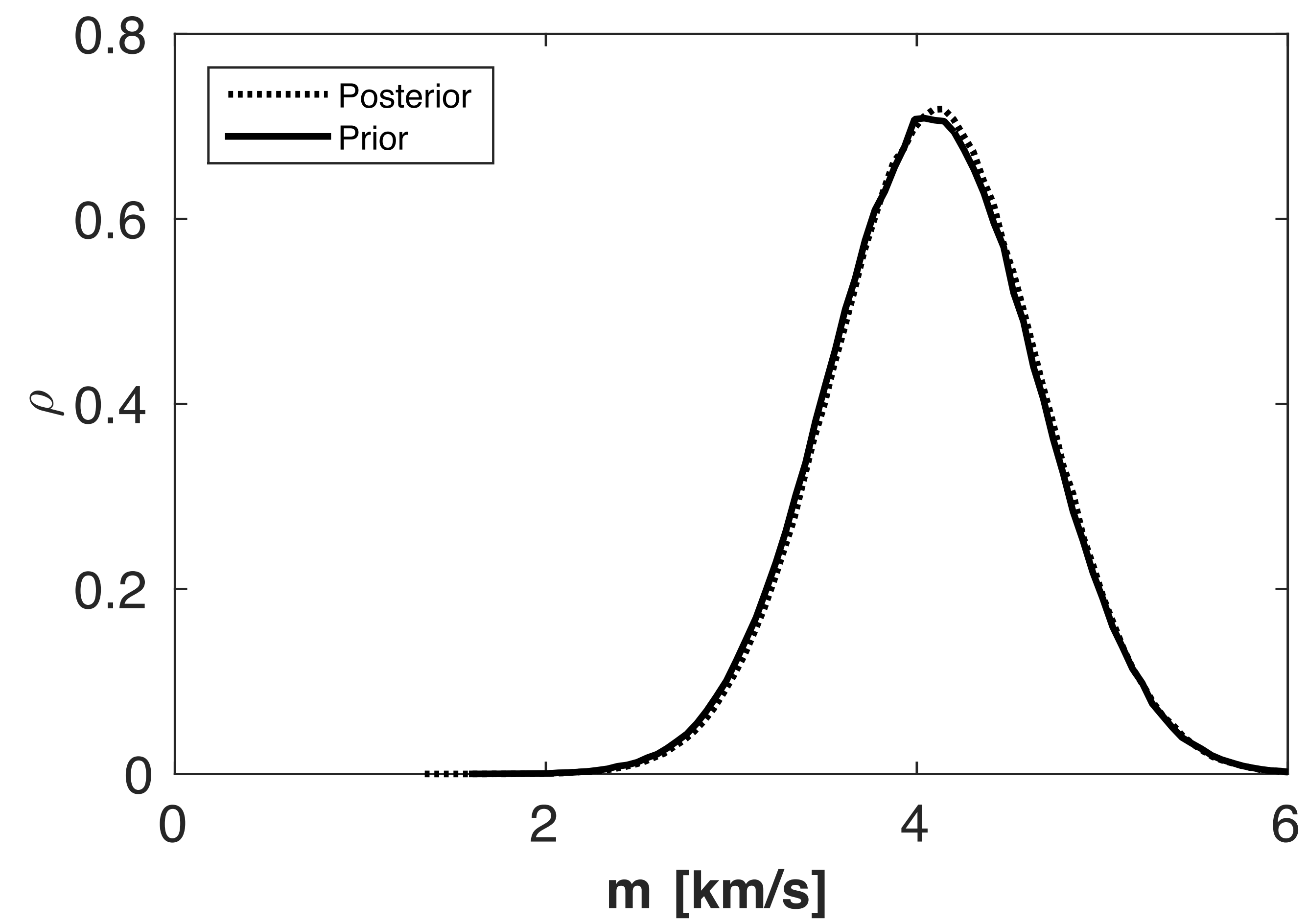
(a) x = 2240 m, z = 40 m



(b) x = 2240 m, z = 640 m

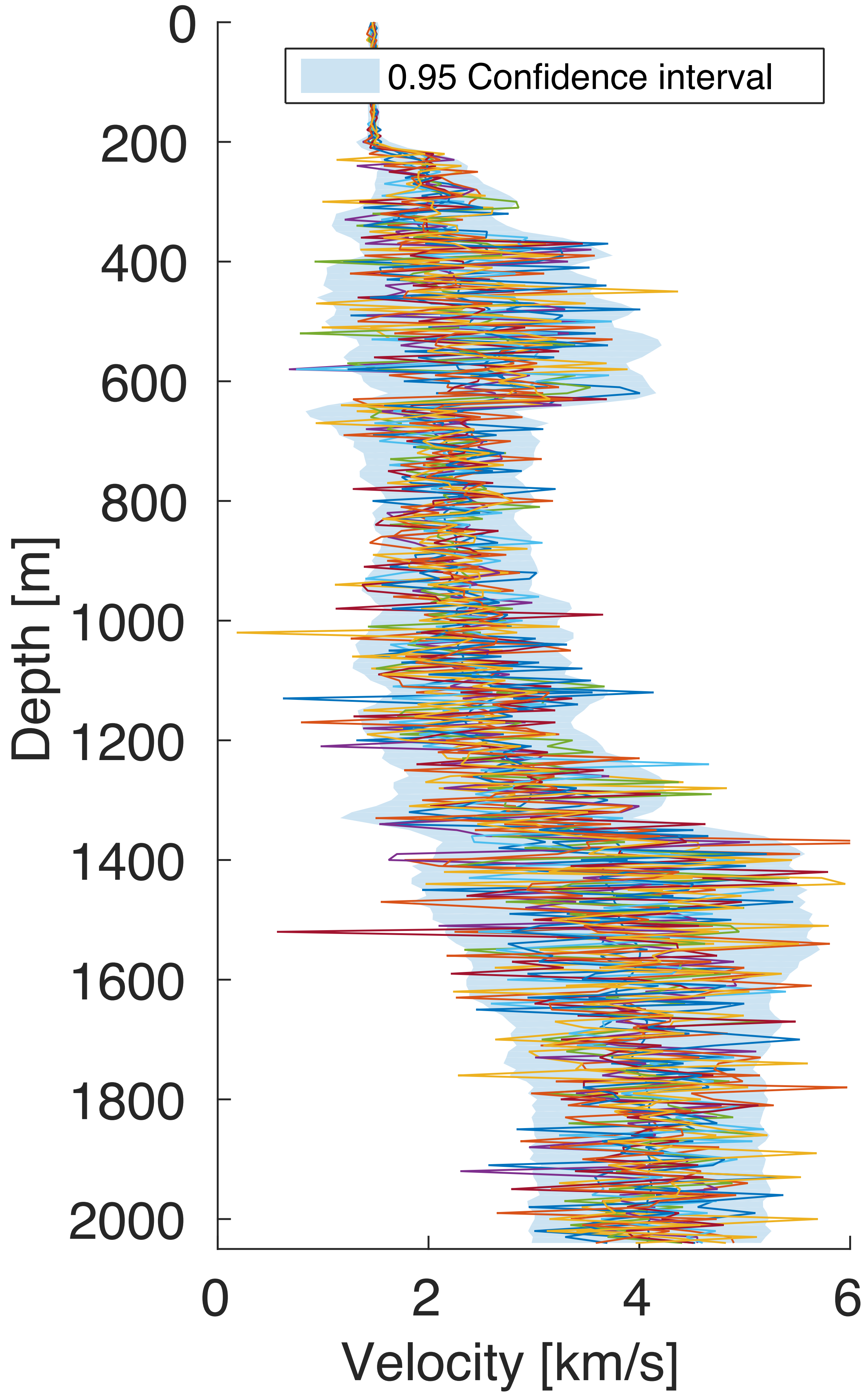


(c) x = 2240 m, z = 1240 m

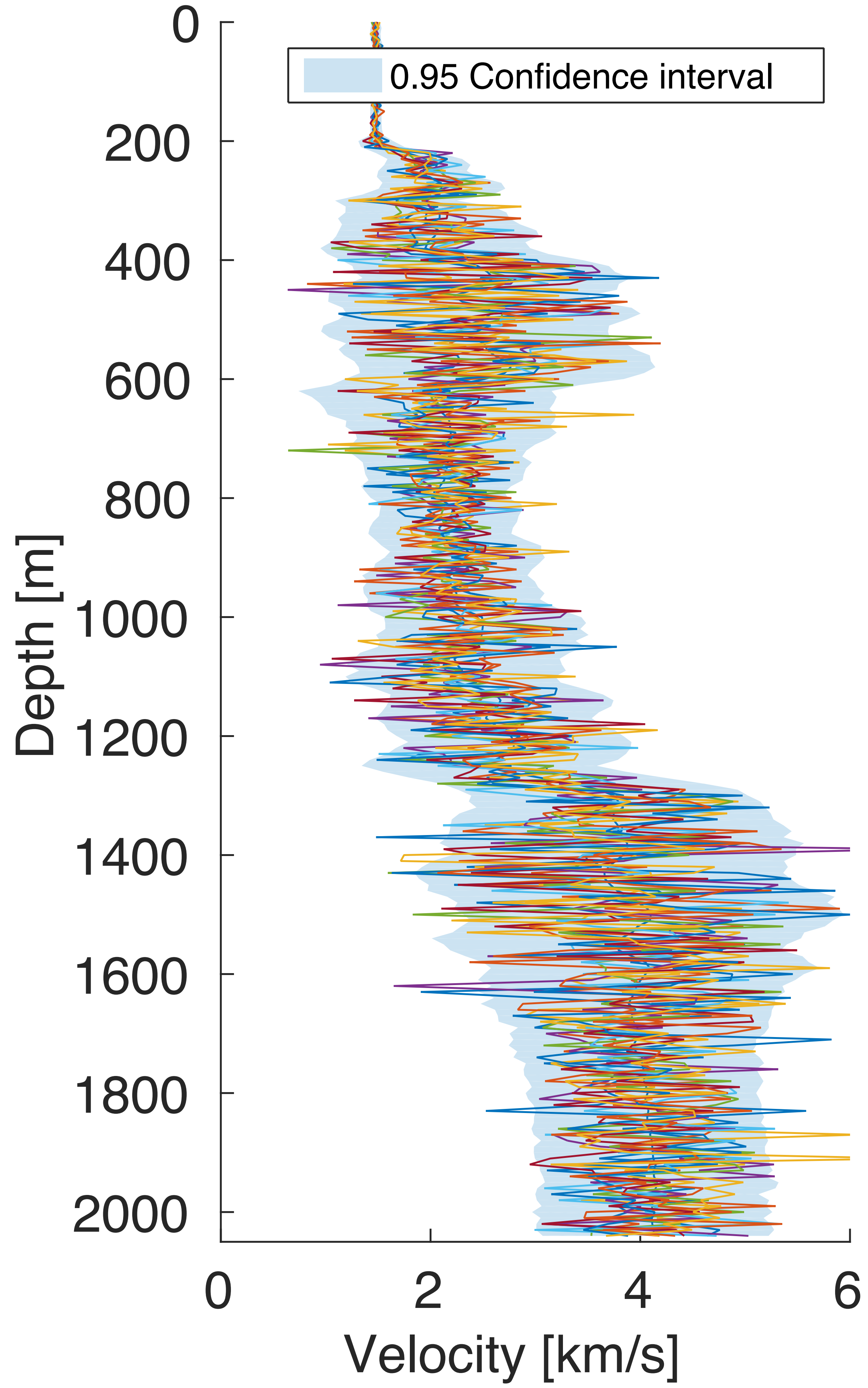


(d) x = 2240 m, z = 1840 m

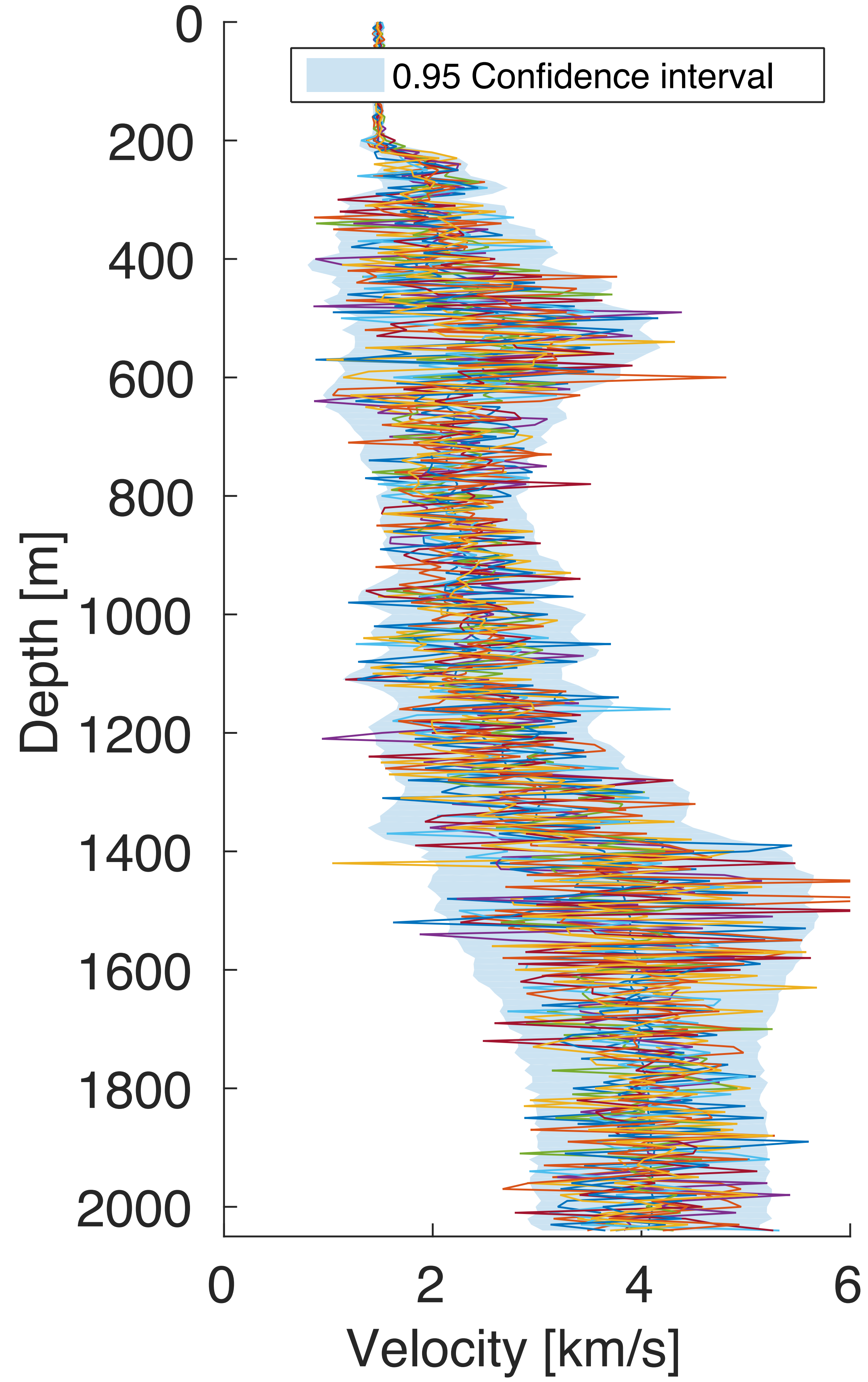
Cross section comparison
– 95% confidence interval vs 10 realizations by RML



(a) x = 1000 m



(b) x = 2500 m



(c) x = 4000 m

Conclusion

In this thesis, I have developed

- a source estimation algorithm for time-domain sparsity promoting LS-RTM approaches

- a source estimation algorithm for wavefield reconstruction inversion

- an algorithm to quantify uncertainties for inverse problems with weak wave-equation constraints

Publications

[1] Z. Fang, R. Wang, F. Herrmann, “Source estimation for wavefield-reconstruction inversion”. Accepted by Geophysics, 2018.

[2] Z. Fang, C. Da Silva, R. Kuske, F. Herrmann, “Uncertainty quantification for inverse problems with weak PDE constraints”. Submitted, 2018.

[3] Z. Fang, C. Da Silva and F. J. Herrmann, 2017, An efficient penalty method for PDE-constrained optimization problem with source estimation and stochastic optimization, Applied Inverse Problems 2017.

[4] Z. Fang, C. Da Silva, R. Kuske and F. Herrmann, 2017, Uncertainty quantification for inverse problems with a weak wave-equation constraint, 13th International Conference on Mathematical and Numerical Aspects of Wave Propagation.

[5] M. Yang, P. A. Witte, Z. Fang and F. Herrmann, 2016, Time-domain sparsity-promoting least-squares migration with source estimation, 86th SEG Annual Meeting 2016.

[6] X. Chai, M. Yang, P. A. Witte, R. Wang, Z. Fang and F. Herrmann, 2016, A linearized Bregman method for compressive waveform inversion, 86th SEG Annual Meeting 2016.

[7] Z. Fang, C. Y. Lee, C. Da Silva, T. van Leeuwen and F. Herrmann, 2016, Uncertainty quantification for Wavefield Reconstruction Inversion using a PDE free semidefinite Hessian and randomize-then-optimize method, 86th SEG Annual Meeting 2016.

[8] Z. Fang and F. Herrmann, 2015, Source estimation for Wavefield Reconstruction Inversion, 77th EAGE Conference and Exhibition 2015.

[9] Z. Fang, C. Y. Lee, C. Da Silva, F. Herrmann and R. Kuske, 2015, Uncertainty quantification for Wavefield Reconstruction Inversion, 77th EAGE Conference and Exhibition 2015.

[10] R. Lago, A. Petrenko, Z. Fang and F. Herrmann, 2014, Fast solution of time-harmonic wave-equation for full-waveform inversion, 76th EAGE Conference and Exhibition 2014.

[11] Z. Fang, C. Da Silva and F. Herrmann, 2014, Fast uncertainty quantification for 2D full-waveform inversion with randomized source subsampling, 76th EAGE Conference and Exhibition 2014.

Thank you !

PhD Advisory Committee members

PhD Exam Committee members

SLIM Group Members

SINBAD sponsors & NSERC

Family and Friends