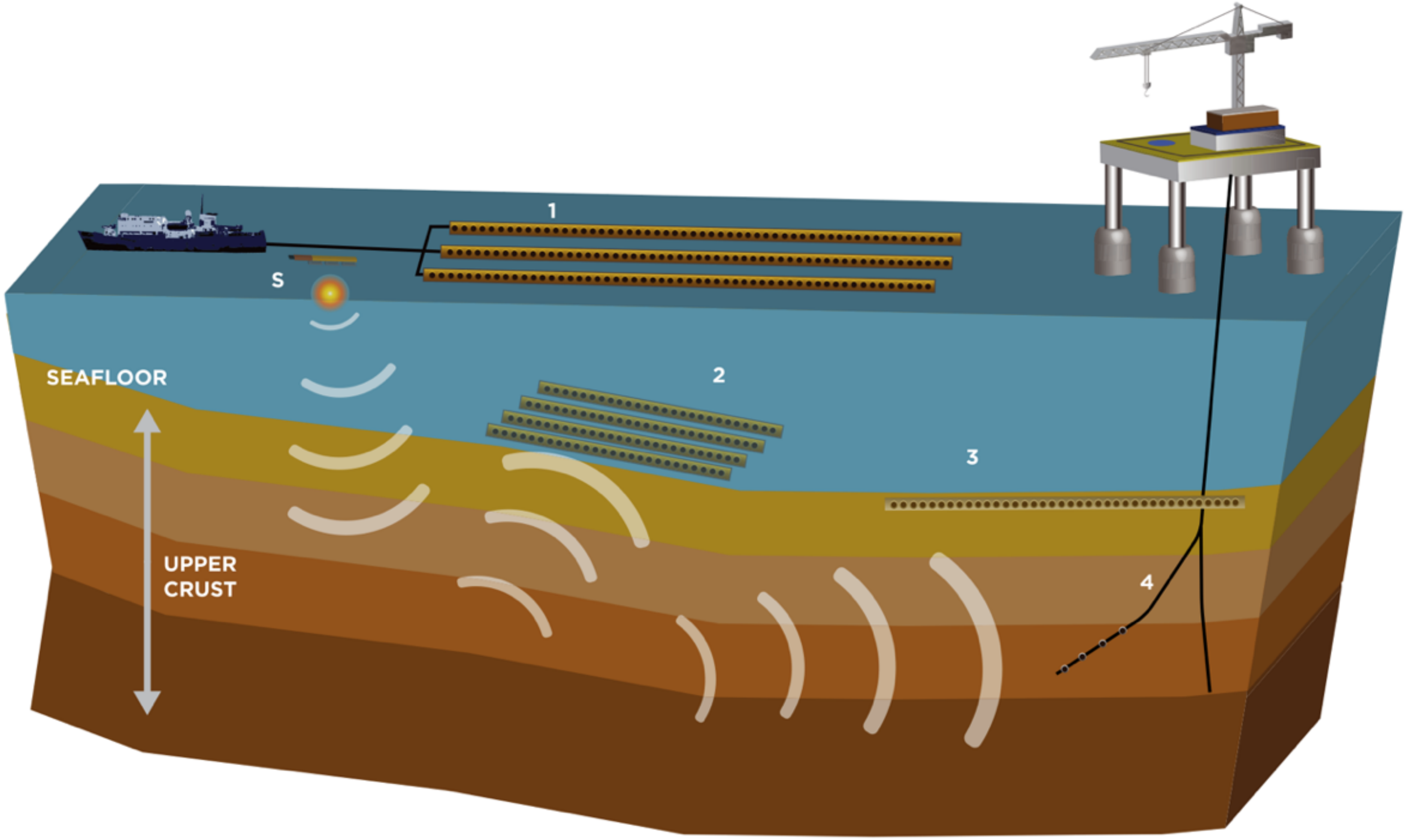
Large scale high-frequency seismic wavefield reconstruction, acquisition via rank minimization and sparsity-promoting source estimation

Shashin Sharan
Ph.D. Defense of Dissertation
October 27, 2020

SLIM
Georgia Institute of Technology

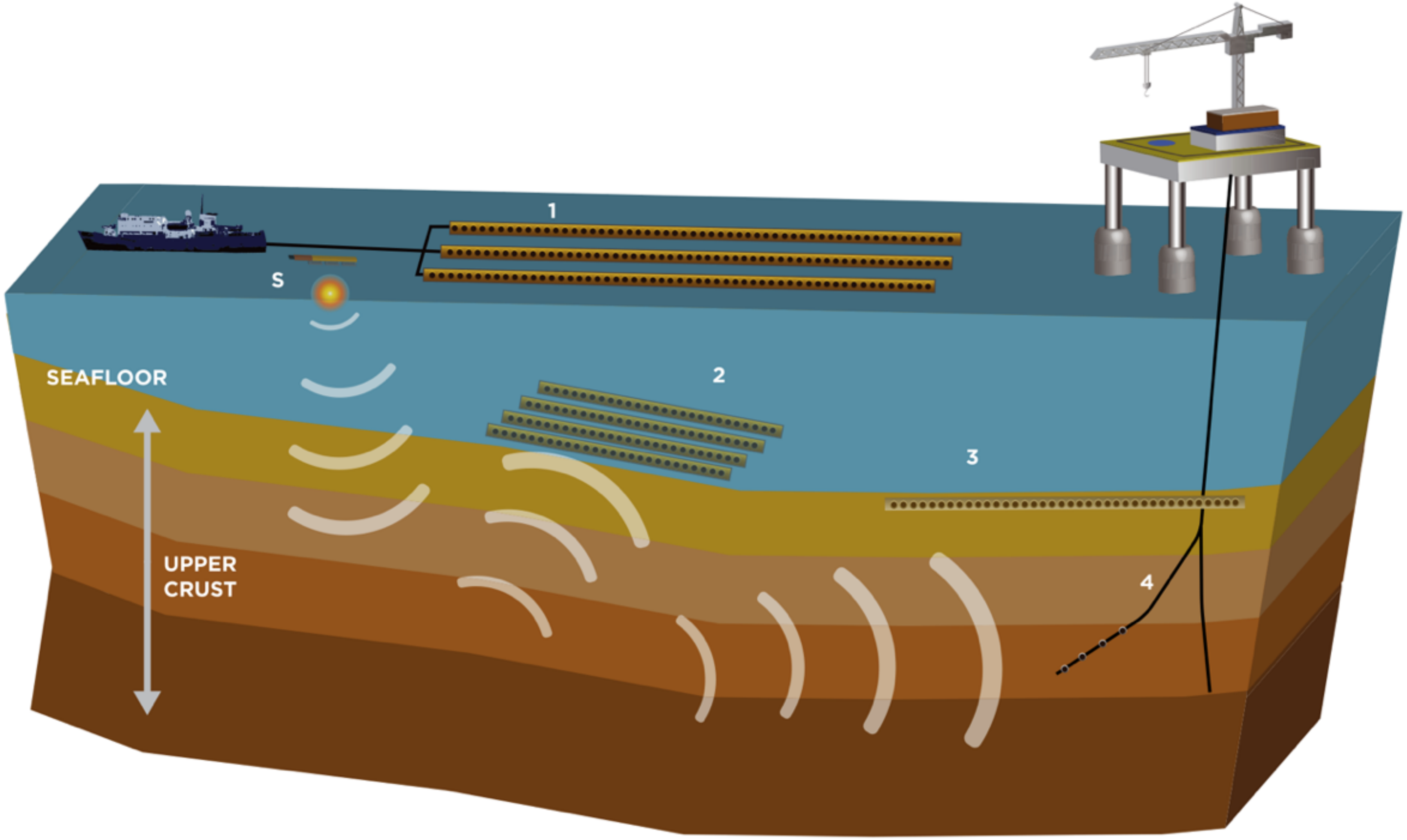
[Caldwell and Walker ,'11]

Seismic data acquisition



[Caldwell and Walker ,'11]

Seismic data acquisition

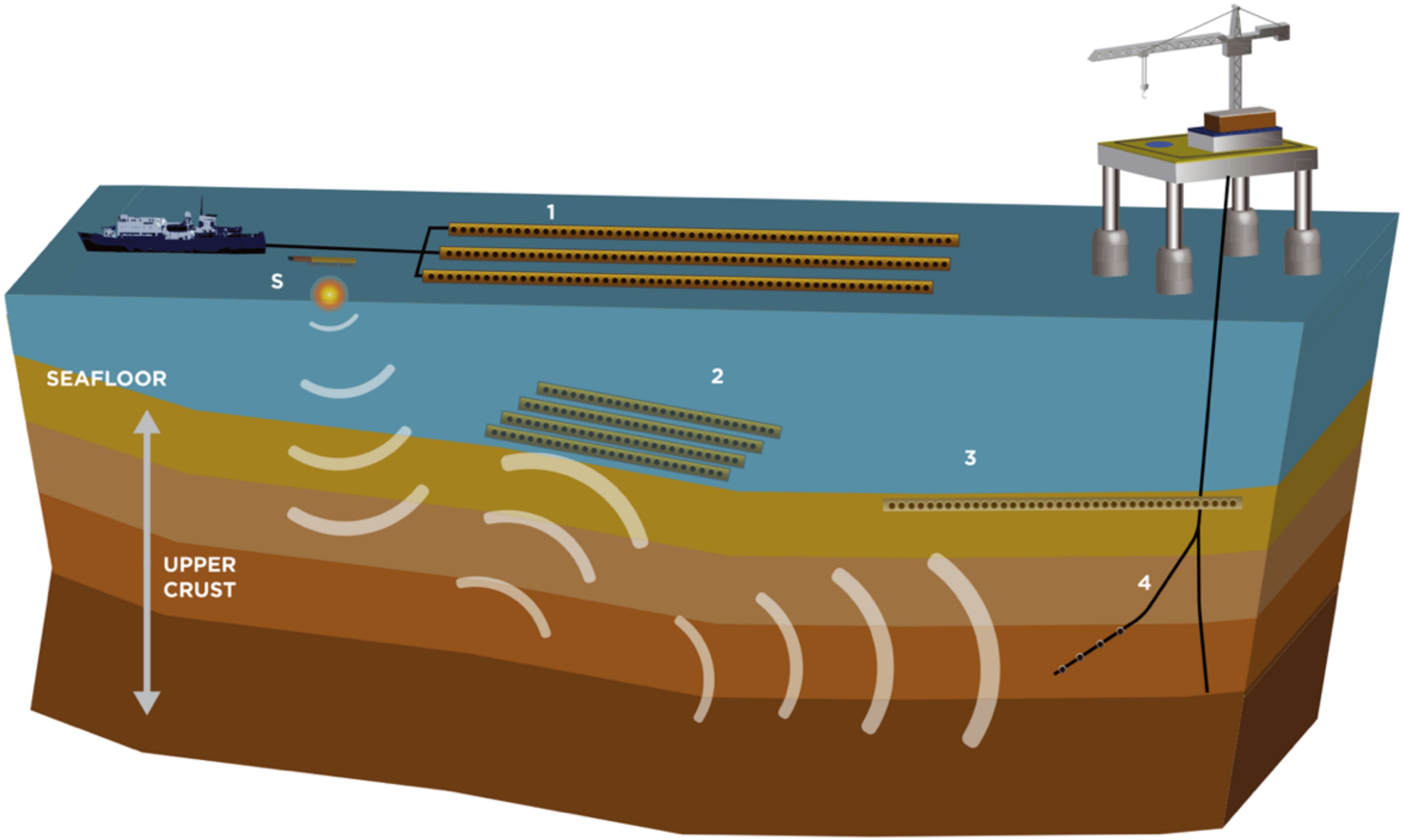


Objective

- Acquire Dense seismic data
- for high-resolution imaging

[Caldwell and Walker ,'11]

Seismic data acquisition



Objective

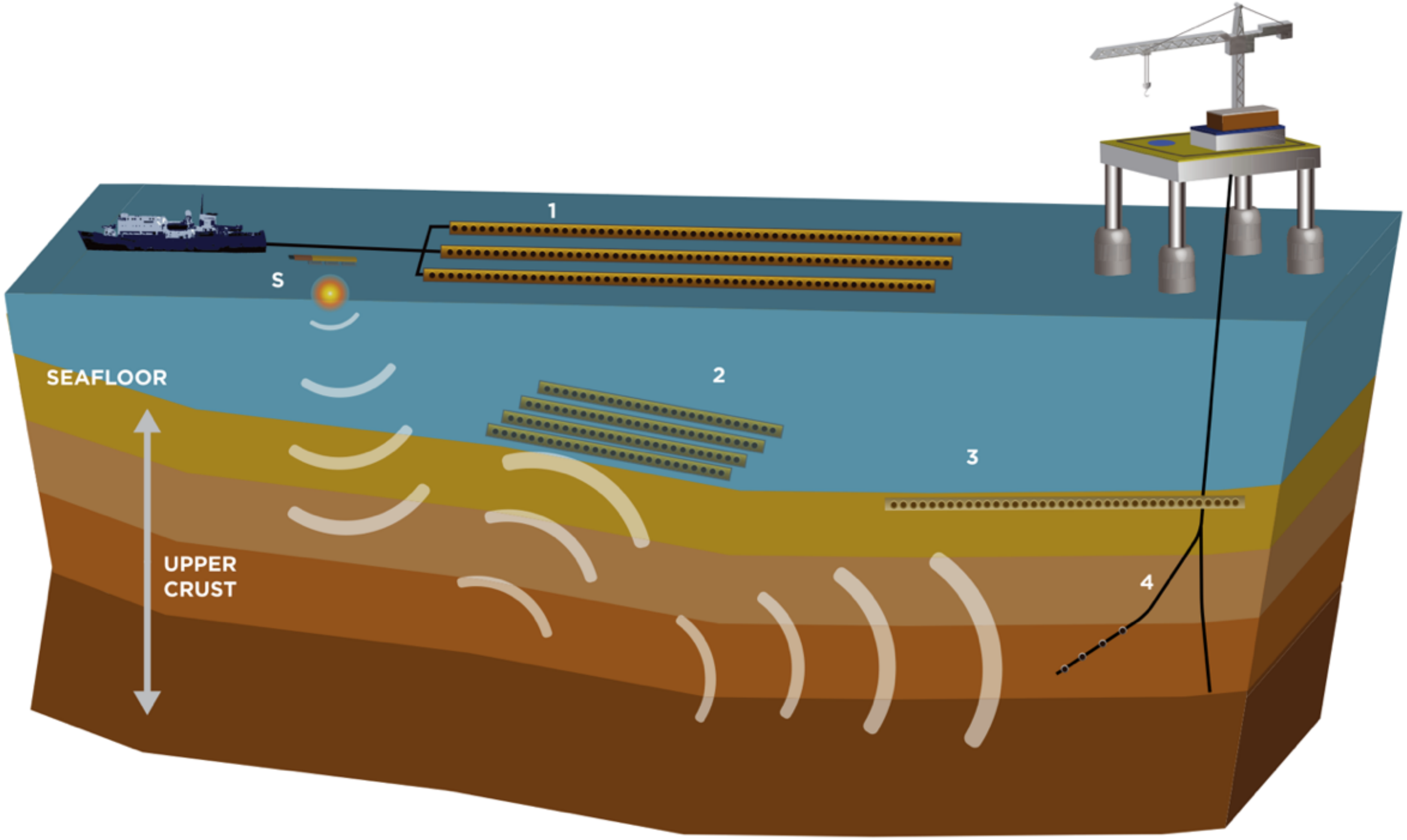
- Acquire Dense seismic data
- for high-resolution imaging

Challenge

- Operationally complex
- to acquire dense data

[Caldwell and Walker ,'11]

Seismic data acquisition



Objective

- Acquire Dense seismic data
- for high-resolution imaging

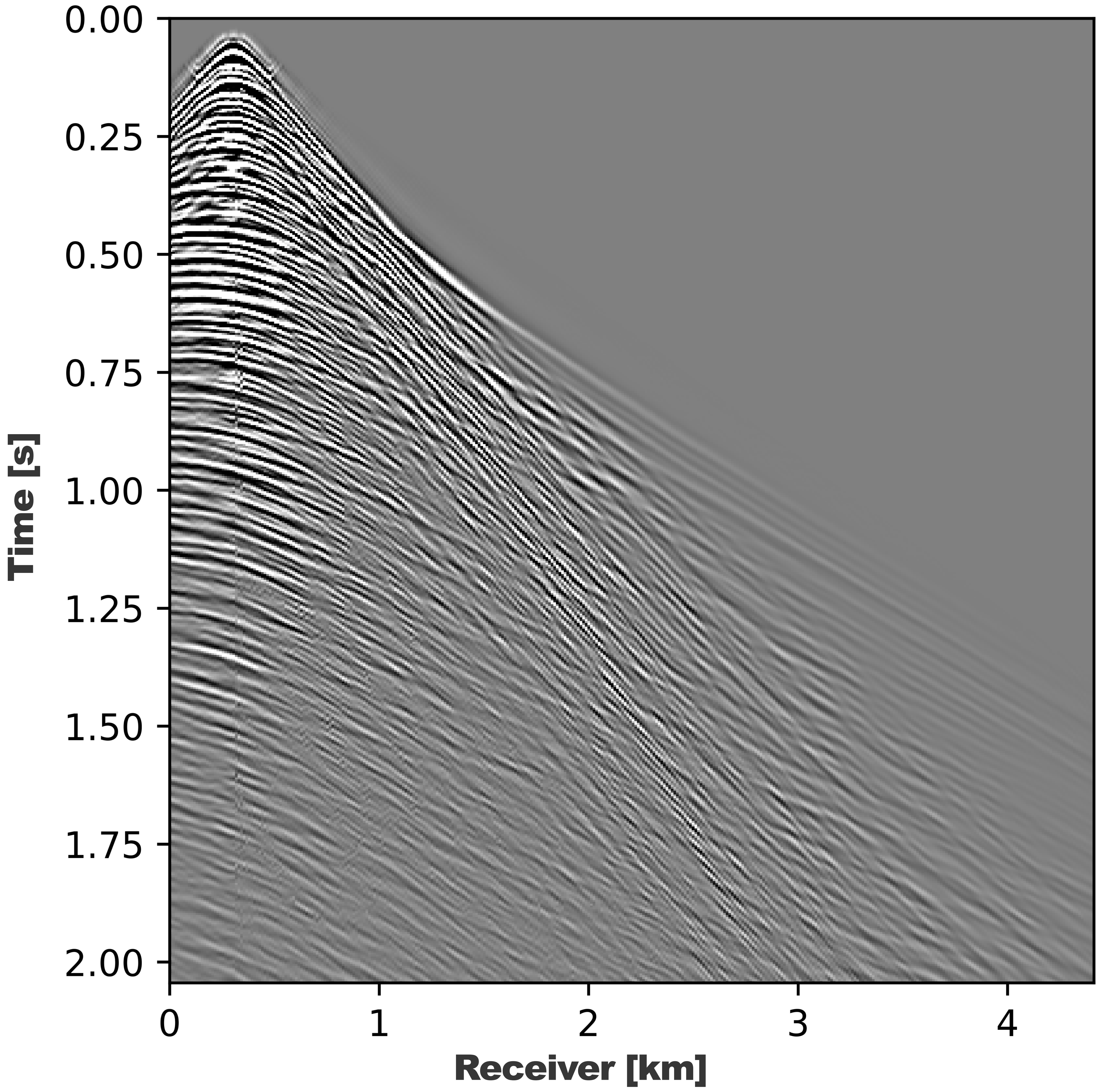
Challenge

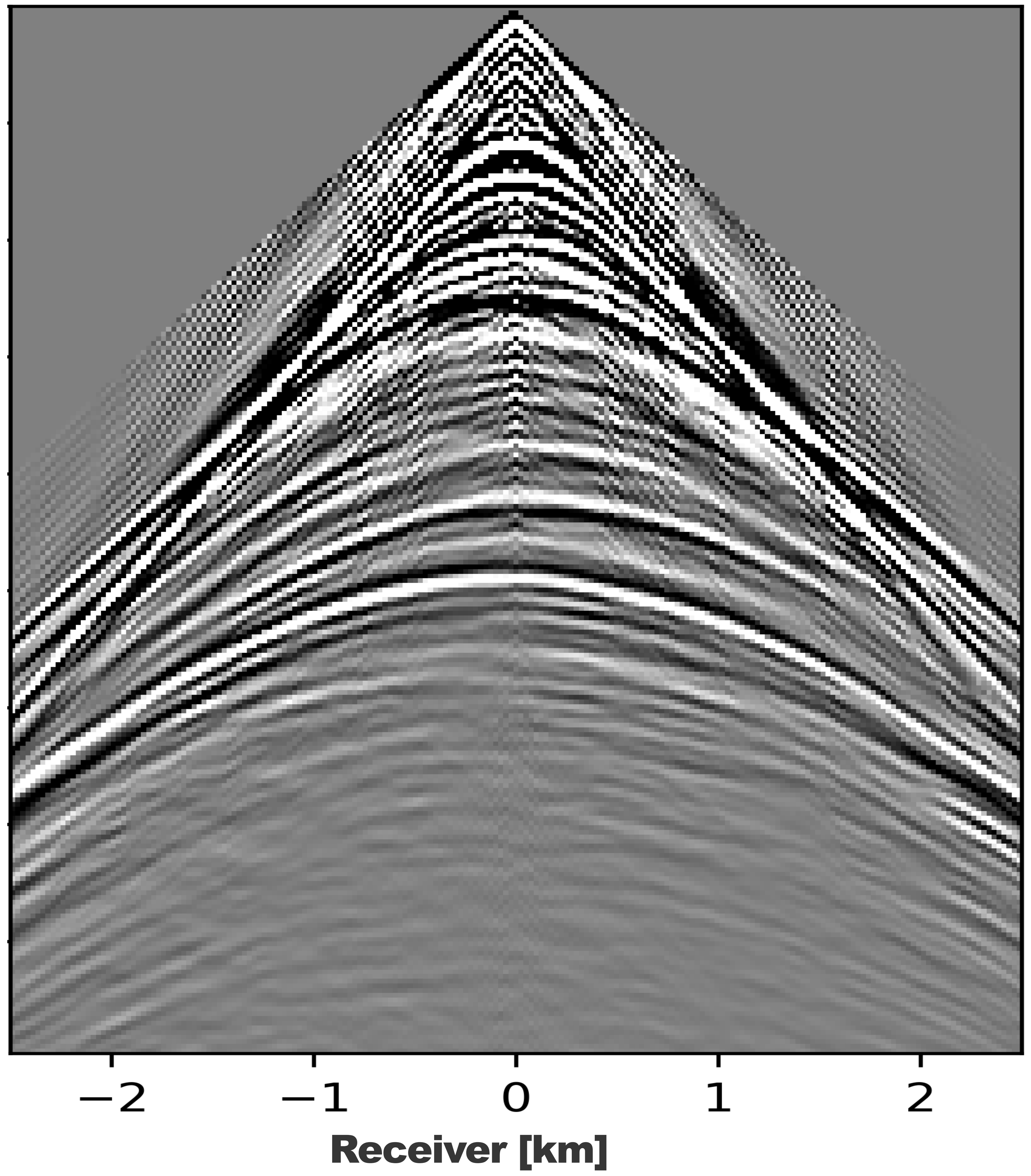
- Operationally complex
- to acquire dense data

Solution

- Acquisition on coarse grid
- followed by data reconstruction
- on periodic fine grid

Common Shot Gather Visualization





[Kumar et al. ,'16]

Seismic data reconstruction w/ Low-rank

Advantages

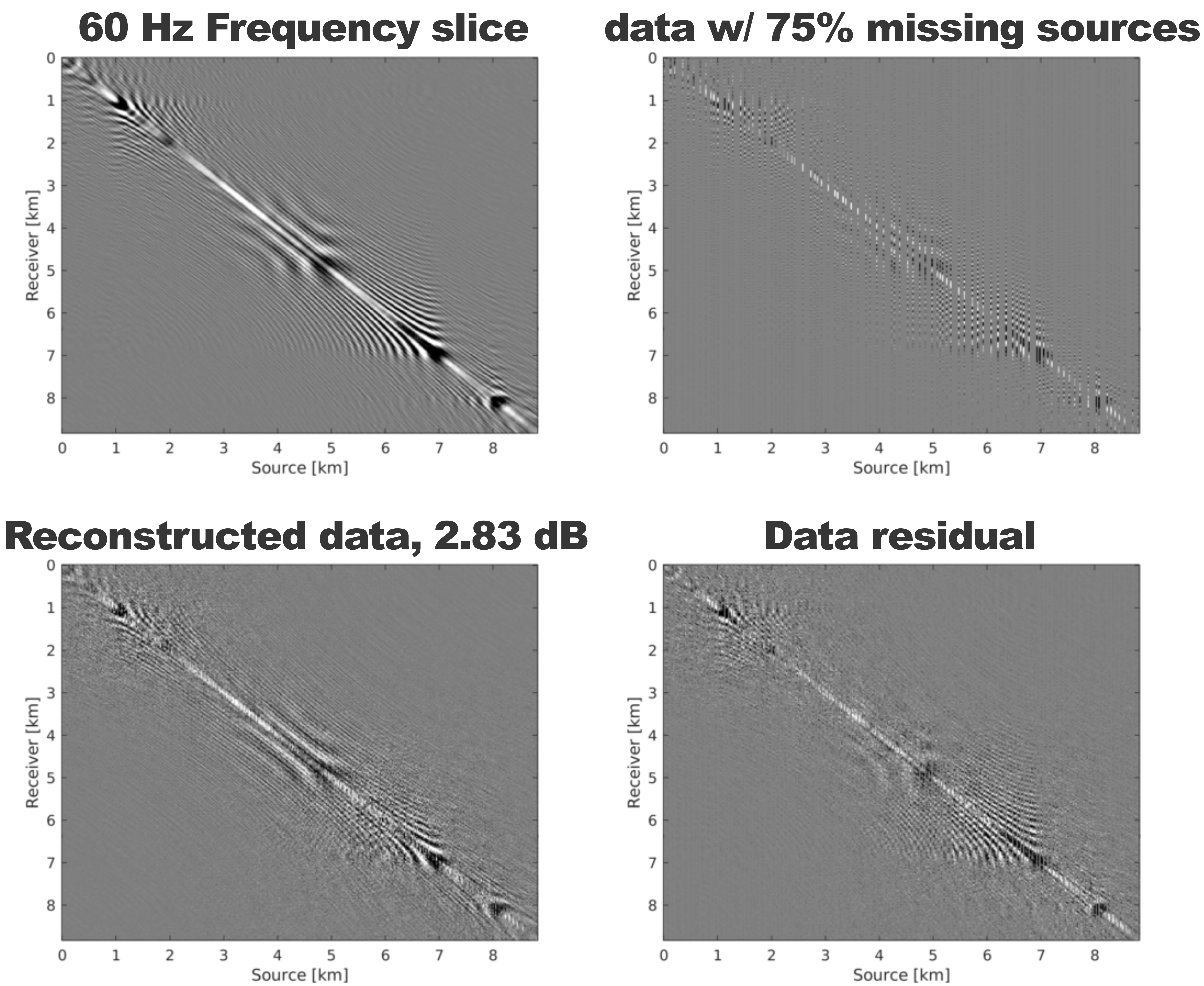
- Scalable for large scale 3D data
- Performs well at lower frequencies

Limitations

- Performs poorly at higher frequencies

[Kumar et al. ,'16]

Seismic data reconstruction w/ Low-rank



Advantages

- Scalable for large scale 3D data
- Performs well at lower frequencies

Limitations

- Performs poorly at higher frequencies

Key Contributions: Chapters 2 & 3

Recursively Weighted Matrix Completion Framework

- Improved data reconstruction at high frequency
- Scalable for large scale 3D data
- Computationally faster weighted method in comparison to traditional weighted method

5D Time-Jittered Marine Acquisition

- Simultaneous separation and reconstruction of sources from blended data
- Scalable for large scale 3D data

SLIM

Matrix completion

Successful reconstruction scheme

- exploit *structure*
  - *low-rank* / *fast decay* of singular values
- sampling
  - randomness *increases* rank in "transform domain"
- optimization
  - via *rank-minimization (nuclear norm-minimization)*

[Recht et al.,'16; Aravkin et al.,'14; Eftekhari et al.,'18]

Nuclear-norm minimization

convex relaxation of rank-minimization

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\|\mathbf{X}\|_*}_{\text{Sum of singular values of X}} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{B}\|_F \leq \epsilon$$

Sum of singular values of X

*where $\|.\|_F$ is the Frobenius norm

$\mathcal{A}$ is the Measurement operator

$\mathbf{B}$ is the observed data

ϵ is the noise level

[Recht et al.,'16; Aravkin et al.,'14; Eftekhari et al.,'18]

SLIM

Nuclear-norm minimization

convex relaxation of rank-minimization

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\|\mathbf{X}\|_*}_{} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

Sum of singular values of X

*where $\|.\|_F$ is the Frobenius norm

$\mathcal{A}$ is the Measurement operator

$\mathbf{B}$ is the observed data

ϵ is the noise level

Weighted Nuclear-norm minimization

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{QXW}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

*where $\mathbf{Q} = w_1\mathbf{U}\mathbf{U}^H + \mathbf{U}^\perp\mathbf{U}^{\perp^H}$ and $\mathbf{W} = w_2\mathbf{V}\mathbf{V}^H + \mathbf{V}^\perp\mathbf{V}^{\perp^H}$

$\mathbf{U},\mathbf{V}$ are row and column subspaces of adjacent frequency slice

scalars $w_1, w_2 \in (0,1]$ are weights

[Recht et al.,'16; Aravkin et al.,'14; Eftekhari et al.,'18]

SLIM

Nuclear-norm minimization

convex relaxation of rank-minimization

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\|\mathbf{X}\|_*}_{} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

Sum of singular values of X

*where $\|.\|_F$ is the Frobenius norm

$\mathcal{A}$ is the Measurement operator

$\mathbf{B}$ is the observed data

ϵ is the noise level

Weighted Nuclear-norm minimization

expensive projection operators

$$\underset{\mathbf{X}}{\text{minimize}} \quad \|\mathbf{QXW}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

*where $\mathbf{Q} = w_1\mathbf{U}\mathbf{U}^H + \mathbf{U}^\perp\mathbf{U}^{\perp^H}$ and $\mathbf{W} = w_2\mathbf{V}\mathbf{V}^H + \mathbf{V}^\perp\mathbf{V}^{\perp^H}$

$\mathbf{U}, \mathbf{V}$ are row and column subspaces of adjacent frequency slice

scalars $w_1, w_2 \in (0, 1]$ are weights

Efficient Weighted Matrix Completion

$$\underset{\bar{\mathbf{X}}}{\text{minimize}} \quad \|\bar{\mathbf{X}}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

*where $\bar{\mathbf{X}} = \mathbf{Q}\mathbf{X}\mathbf{W}$

and $\mathbf{X} = \mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}$

Efficient Weighted Matrix Completion

$$\underset{\bar{\mathbf{X}}}{\text{minimize}} \ \|\bar{\mathbf{X}}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

*where $\bar{\mathbf{X}} = \mathbf{Q}\mathbf{X}\mathbf{W}$

and $\mathbf{X} = \mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}$

Factorized Form

$$\underset{\bar{\mathbf{L}},\bar{\mathbf{R}}}{\text{minimize}} \ \frac{1}{2}\left\| \begin{bmatrix} \bar{\mathbf{L}} \\ \bar{\mathbf{R}} \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{L}}\bar{\mathbf{R}}^H\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

Efficient Weighted Matrix Completion

$$\underset{\bar{\mathbf{X}}}{\text{minimize}} \ \|\bar{\mathbf{X}}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

*where $\bar{\mathbf{X}} = \mathbf{QXW}$
and $\mathbf{X} = \mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}$

Factorized Form

$$\underset{\bar{\mathbf{L}},\bar{\mathbf{R}}}{\text{minimize}} \ \frac{1}{2}\left\|\begin{bmatrix}\bar{\mathbf{L}}\\ \bar{\mathbf{R}}\end{bmatrix}\right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{L}}\bar{\mathbf{R}}^H\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

*where $\bar{\mathbf{X}} = \bar{\mathbf{L}}\bar{\mathbf{R}}^H$
with $\bar{\mathbf{L}} \in \mathbb{C}^{m\times k}$ and $\bar{\mathbf{R}} \in \mathbb{C}^{n\times k}$
and $k << m, n$
and $\mathbf{L} = \mathbf{Q}^{-1}\bar{\mathbf{L}}$, $\mathbf{R} = \mathbf{W}^{-1}\bar{\mathbf{R}}$

Efficient Weighted Matrix Completion

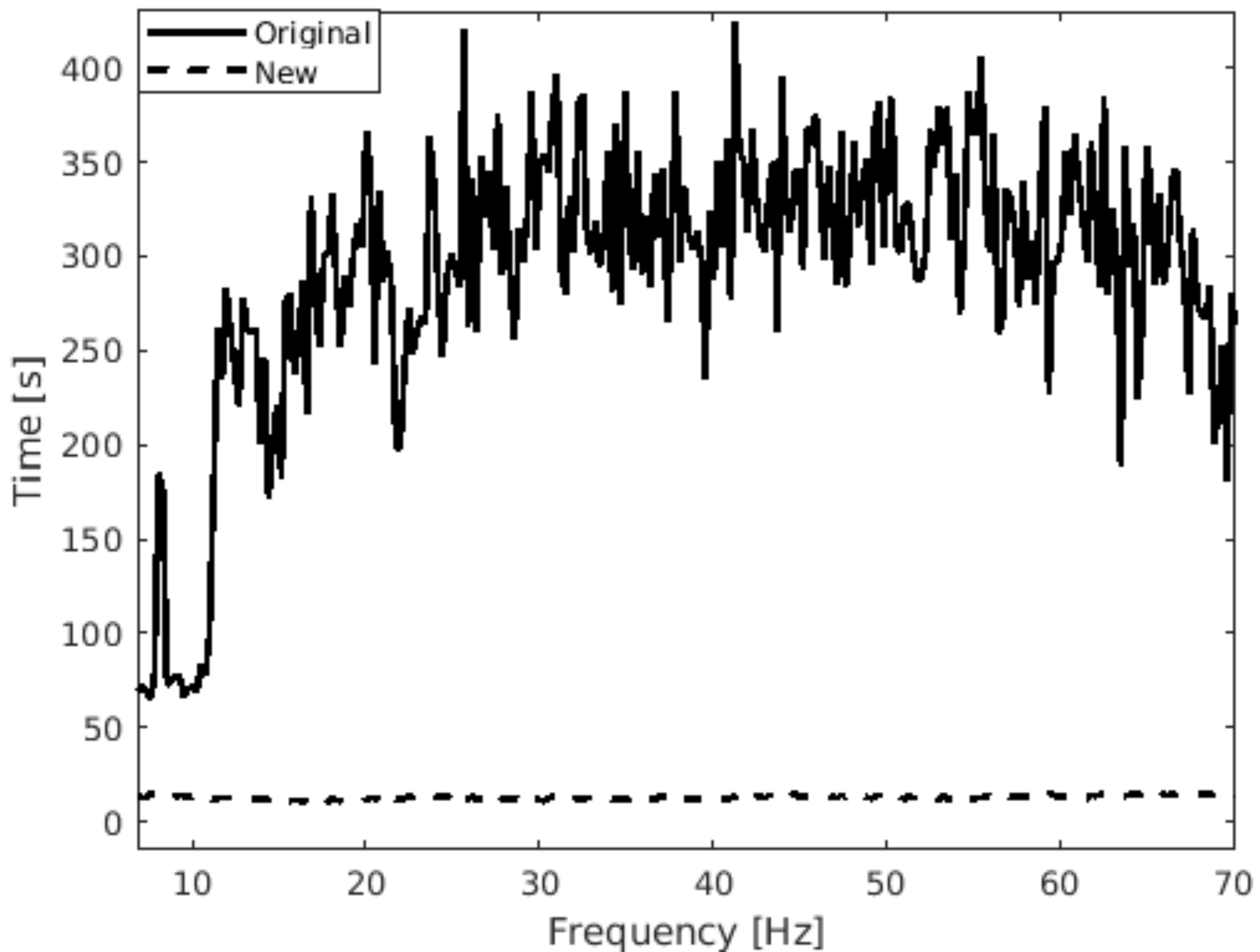
$$\underset{\bar{\mathbf{X}}}{\text{minimize}} \quad \|\bar{\mathbf{X}}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

*where $\bar{\mathbf{X}} = \mathbf{QXW}$
and $\mathbf{X} = \mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}$

Factorized Form

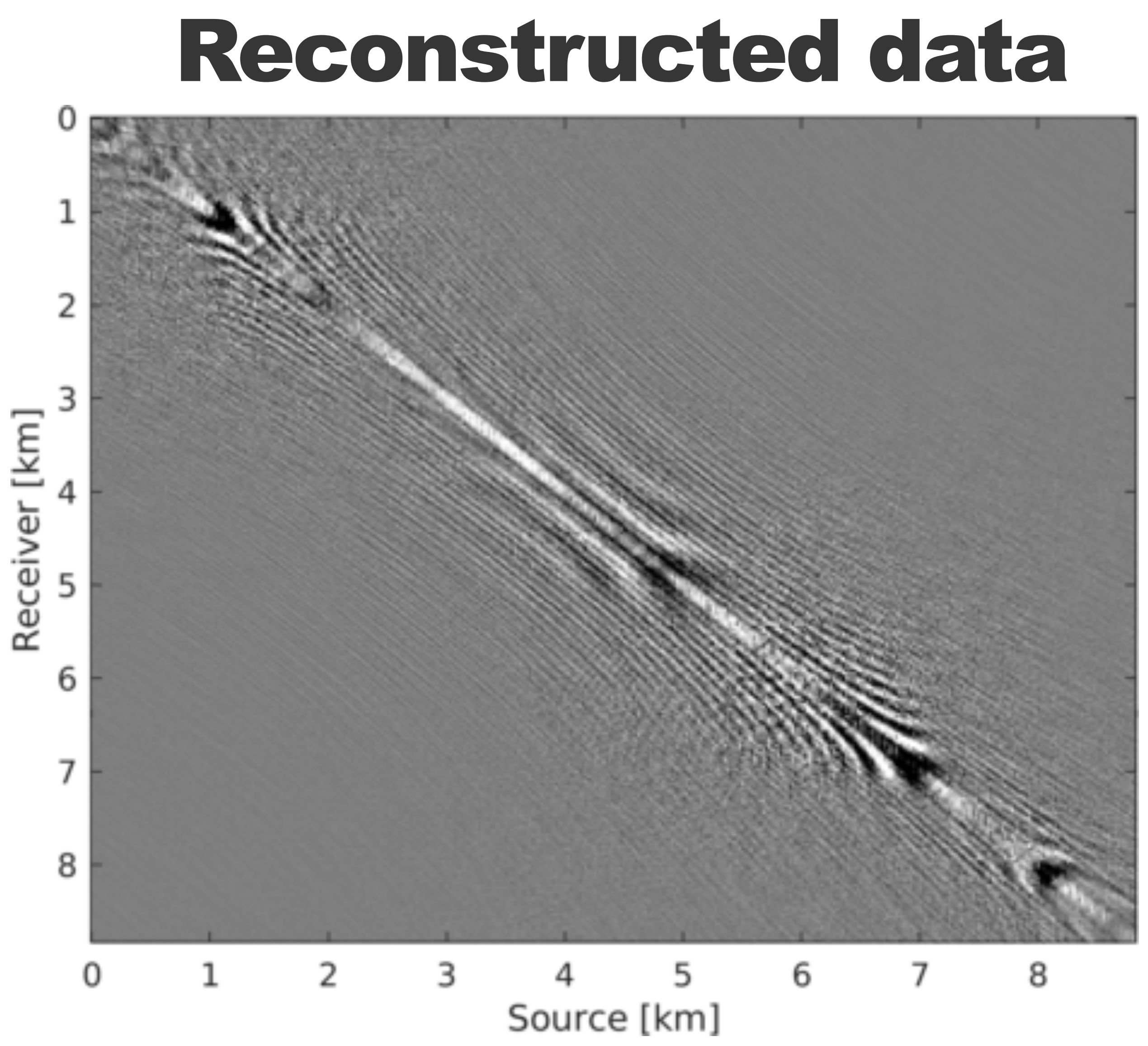
$$\underset{\bar{\mathbf{L}},\bar{\mathbf{R}}}{\text{minimize}} \quad \frac{1}{2}\left\|\begin{bmatrix}\bar{\mathbf{L}} \\ \bar{\mathbf{R}}\end{bmatrix}\right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{L}}\bar{\mathbf{R}}^H\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

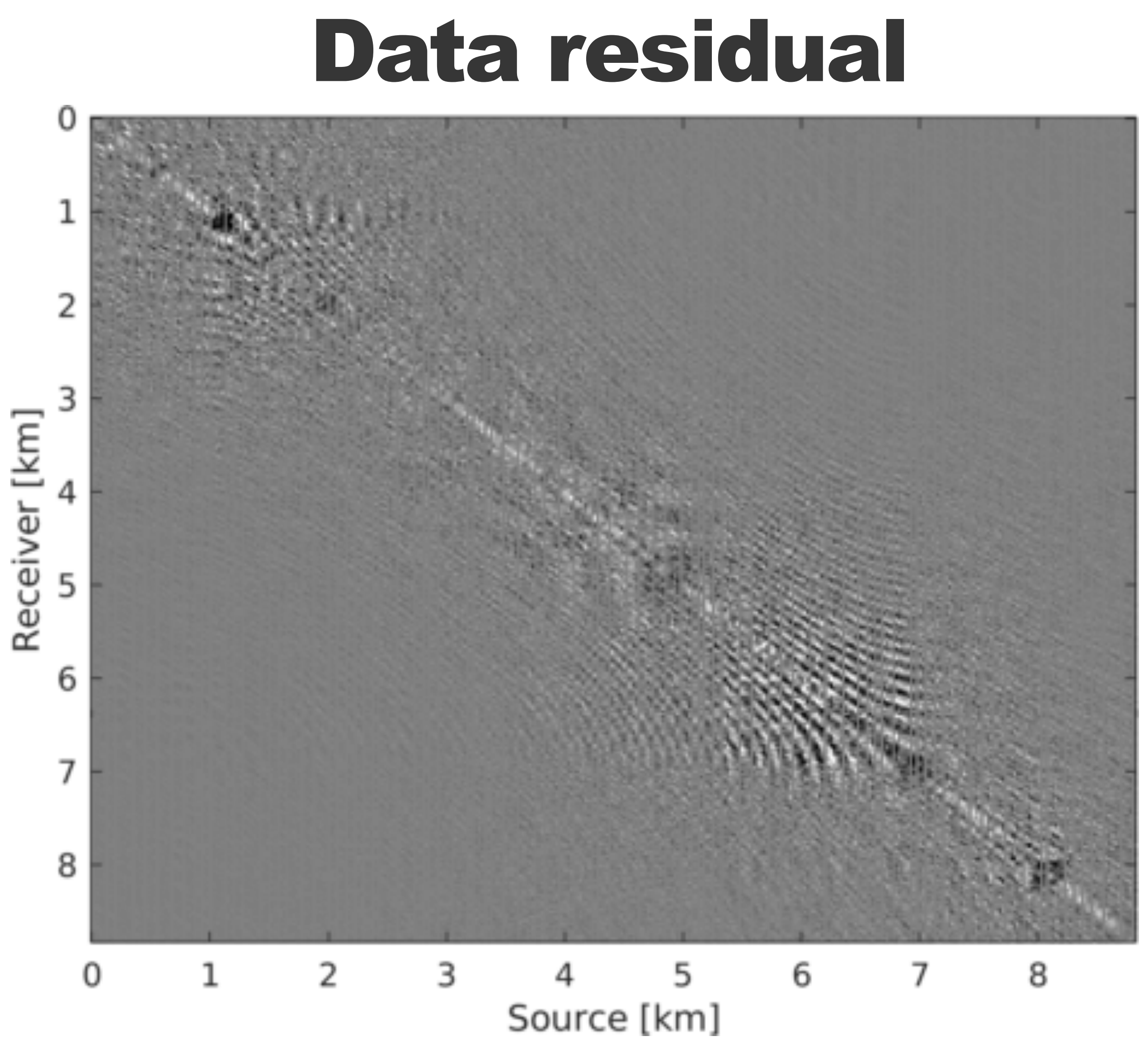
*where $\bar{\mathbf{X}} = \bar{\mathbf{L}}\bar{\mathbf{R}}^H$
with $\bar{\mathbf{L}} \in \mathbb{C}^{m\times k}$ and $\bar{\mathbf{R}} \in \mathbb{C}^{n\times k}$
and $k << m, n$
and $\mathbf{L} = \mathbf{Q}^{-1}\bar{\mathbf{L}}$, $\mathbf{R} = \mathbf{W}^{-1}\bar{\mathbf{R}}$



Runtime Comparison

Recursively Weighted vs Pair Weighted

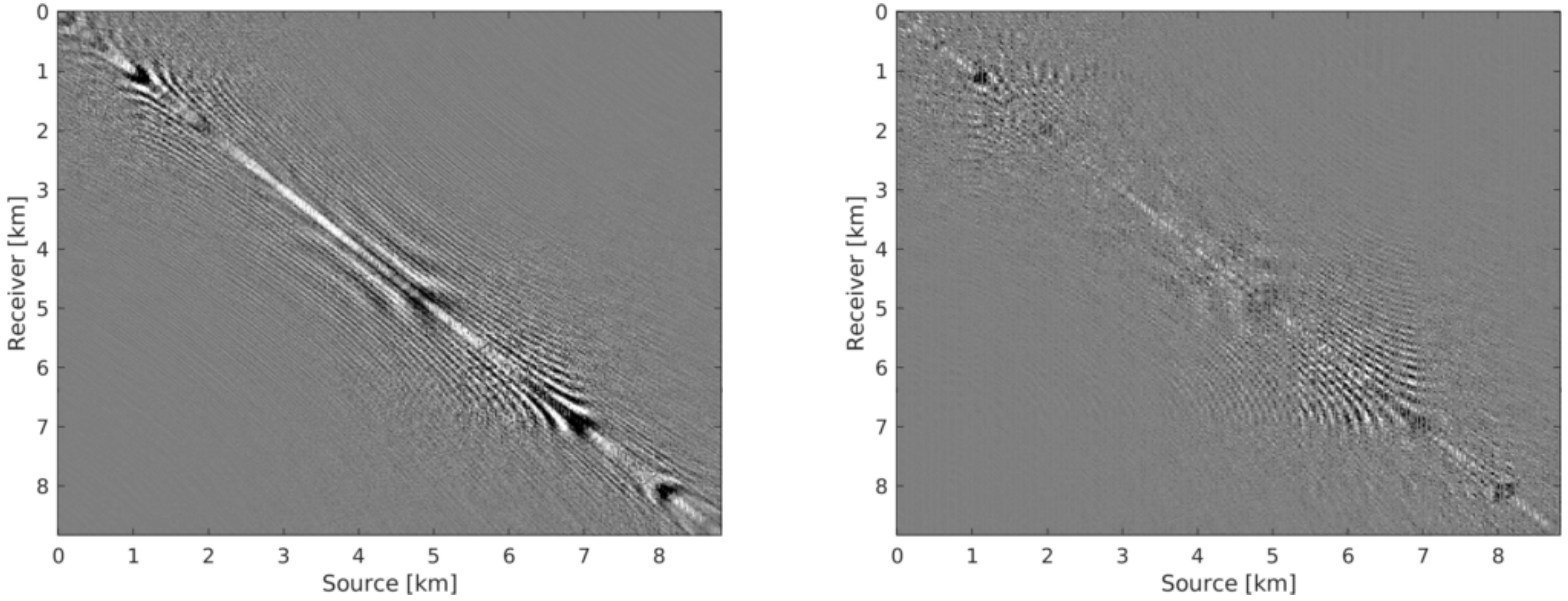




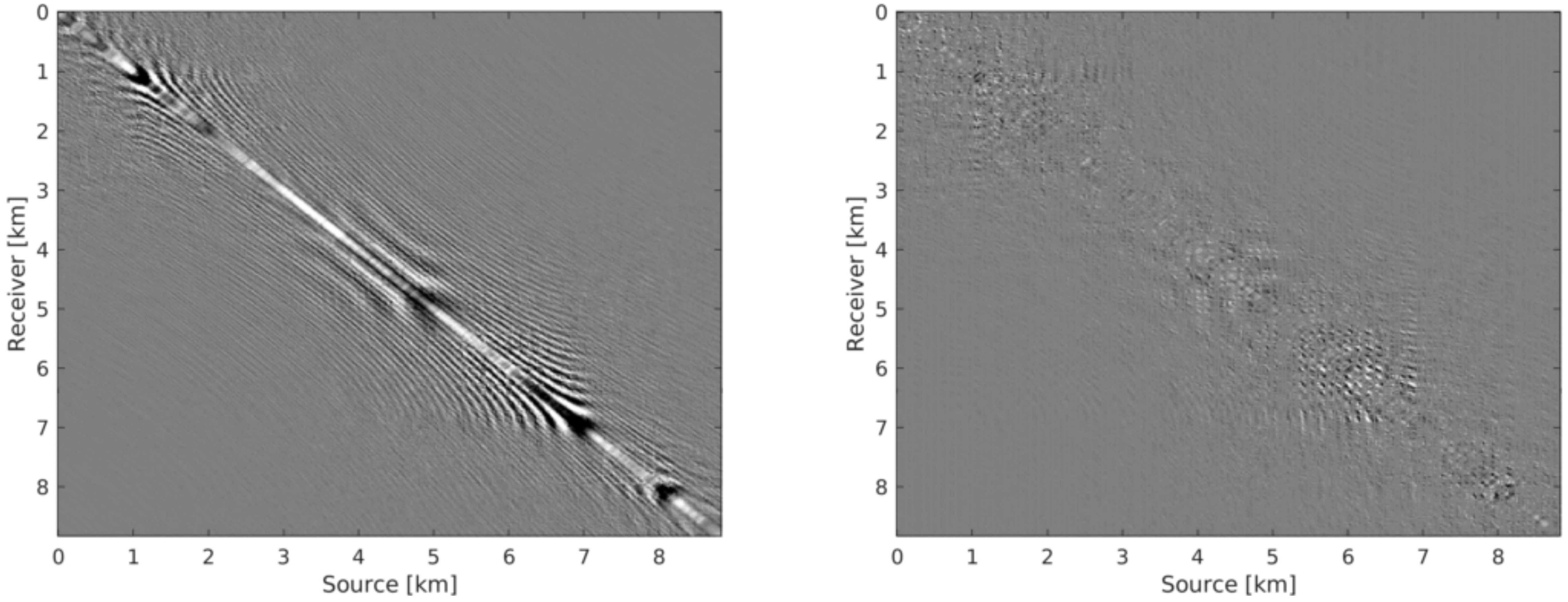
Pair Weighted, SNR = 5.08 dB

Recursively Weighted vs Pair Weighted

Reconstructed data | **Data residual**



Pair Weighted, SNR = 5.08 dB



Recursively Weighted, SNR = 8.72 dB

[Recht et al.,'13; Lopez et al.,'15]

Parallel Implementation

Recursively weighted method challenges

- need to wait for reconstruction of previous frequencies
- computationally demanding for large scale 3D data

Alternating minimization and decoupling strategies enable

- recursively framework scalable for large 3D data
- parallelized framework

Parallelization w/o weights

$$\mathbf{R}(l_1,:)^H := \arg\min_{\mathbf{v}} \quad \frac{1}{2}\|\mathbf{v}\|^2 \quad \text{subject to} \quad \|\mathcal{A}_{l_1}(\mathbf{L}\mathbf{v}) - \mathbf{B}(:,l_1)\| \leq \gamma$$

for $l_1 = 1, 2, \cdots, n$

$$\mathbf{L}(l_2,:)^H := \arg\min_{\mathbf{u}} \quad \frac{1}{2}\|\mathbf{u}\|^2 \quad \text{subject to} \quad \|\mathcal{A}_{l_2}((\mathbf{R}\mathbf{u})^H) - \mathbf{B}(l_2,:)\| \leq \gamma$$

for $l_2 = 1, 2, \cdots, m$

* where $\mathcal{A}_{l_1}$ is acquisition mask for l_1^{th} row of $\mathbf{B}$

and $\mathcal{A}_{l_2}$ is acquisition mask for l_2^{th} column of $\mathbf{B}$

Inclusion of weight matrices poses challenge in parallelizing

Parallelization w/ weights

For large weights:

$$\mathbf{Q}\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) \approx \mathcal{A}(\bar{\mathbf{X}}\mathbf{W}^{-1}) \quad \text{and} \quad \mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1})\mathbf{W} \approx \mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}})$$

Parallelization w/ weights

For large weights:

$$\mathbf{Q}\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) \approx \mathcal{A}(\bar{\mathbf{X}}\mathbf{W}^{-1}) \quad \text{and} \quad \mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1})\mathbf{W} \approx \mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}})$$

This commutation property allows parallelization

$$\bar{\mathbf{R}}(l_1,:)^H := \arg\min_{\bar{\mathbf{v}}} \quad \frac{1}{2}\|\bar{\mathbf{v}}\|^2 \quad \text{subject to} \quad \|\mathcal{A}_{l_1}(\widehat{\mathbf{Q}}\bar{\mathbf{L}}\bar{\mathbf{v}}) - w_1 w_2 \mathbf{B}(:,l_1)\| \leq w_1 w_2 \gamma$$

$$\bar{\mathbf{L}}(l_2,:)^H := \arg\min_{\bar{\mathbf{u}}} \quad \frac{1}{2}\|\bar{\mathbf{u}}\|^2 \quad \text{subject to} \quad \|\mathcal{A}_{l_2}((\bar{\mathbf{R}}\bar{\mathbf{u}})^H\widehat{\mathbf{W}}) - w_1 w_2 \mathbf{B}(l_2,:)\| \leq w_1 w_2 \gamma$$

*where $\widehat{\mathbf{Q}} = \mathbf{U}\mathbf{U}^H + w_1\mathbf{U}^{\perp}\mathbf{U}^{\perp^H} = w_1\mathbf{Q}^{-1}$

and $\widehat{\mathbf{W}} = \mathbf{V}\mathbf{V}^H + w_2\mathbf{V}^{\perp}\mathbf{V}^{\perp^H} = w_2\mathbf{W}^{-1}$

Case Study: BG Synthetic 3D Data

Data dimension: 501 x 201 x 201 x 41 x 41 (nt x nrx x nry x nsx x nsy)

Time sampling interval: 10 ms

Source sampling interval (x and y): 150 m

Receiver sampling interval (x and y): 25 m

Velocity model: BG Compass

Observed data: 90 % missing receivers

Optimization Information

Number of alternations per frequency: 4

Number of inner iterations per alternation: 40

Rank parameter: 228

Weight: 0.75

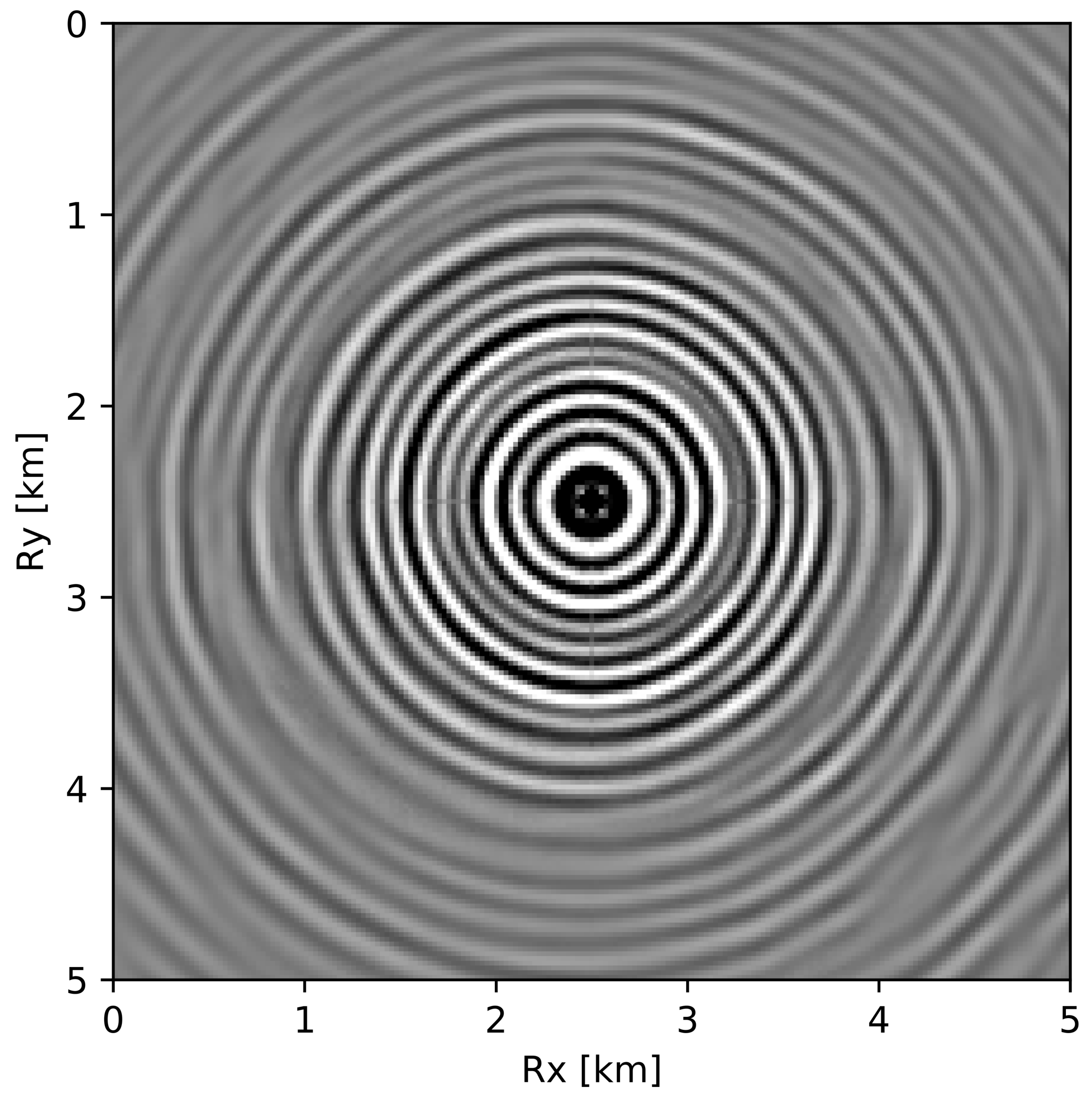
Computational resource: AWS cloud

Time per frequency: 8 minutes

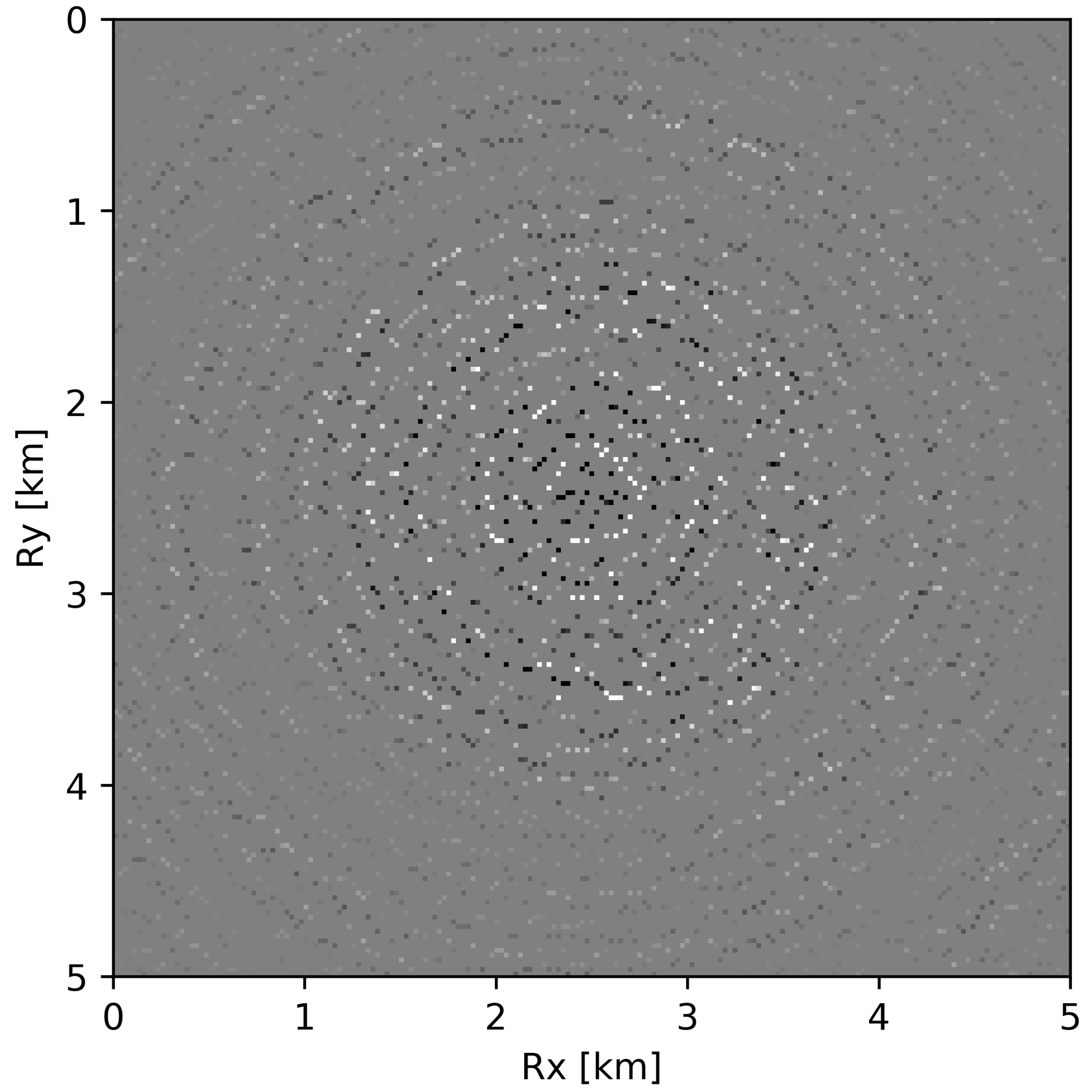
Final volume: 7 GB (95% compressed)

15 Hz Frequency Slice

Ground Truth

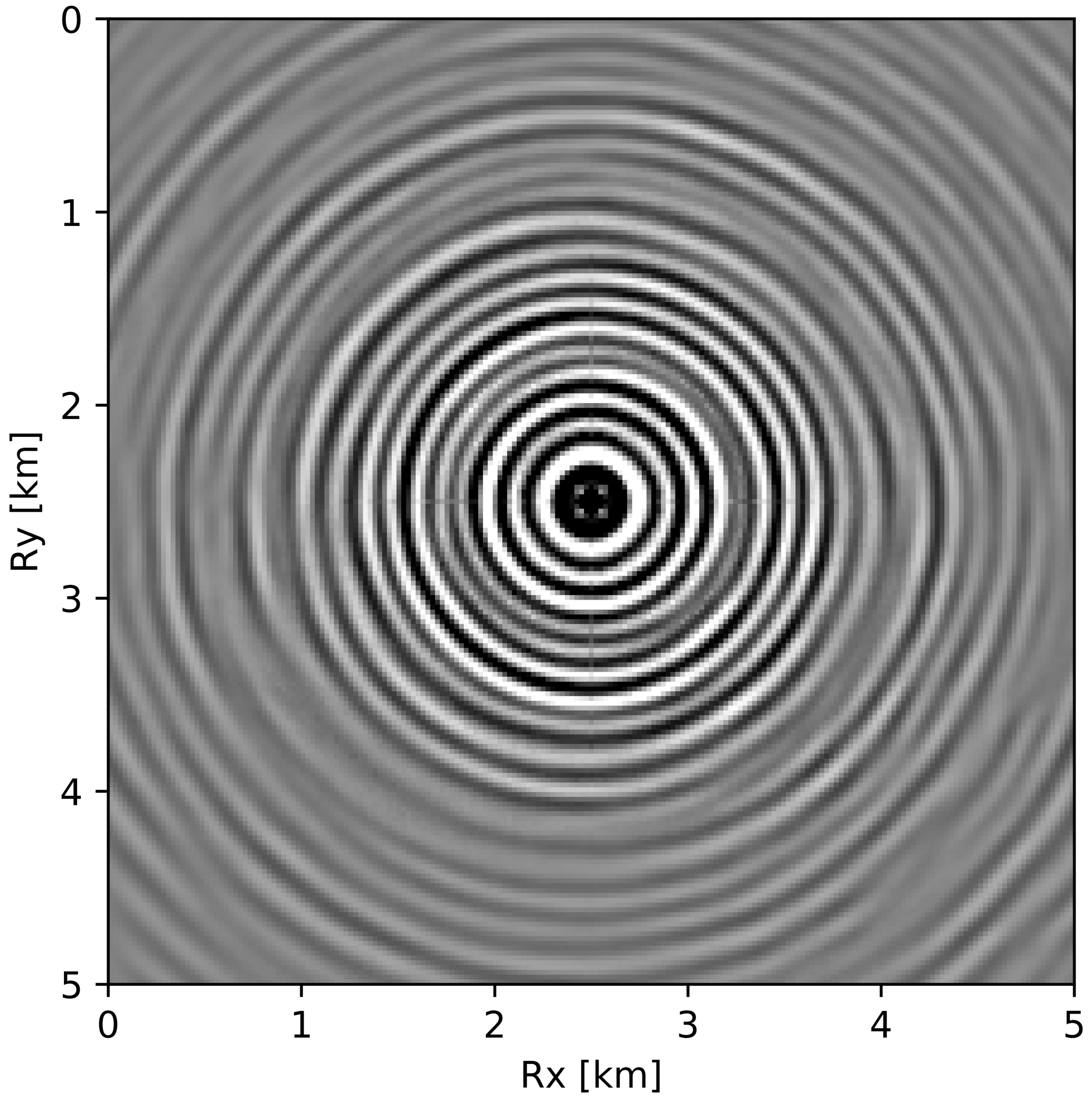


Observed w/ 90% missing receivers

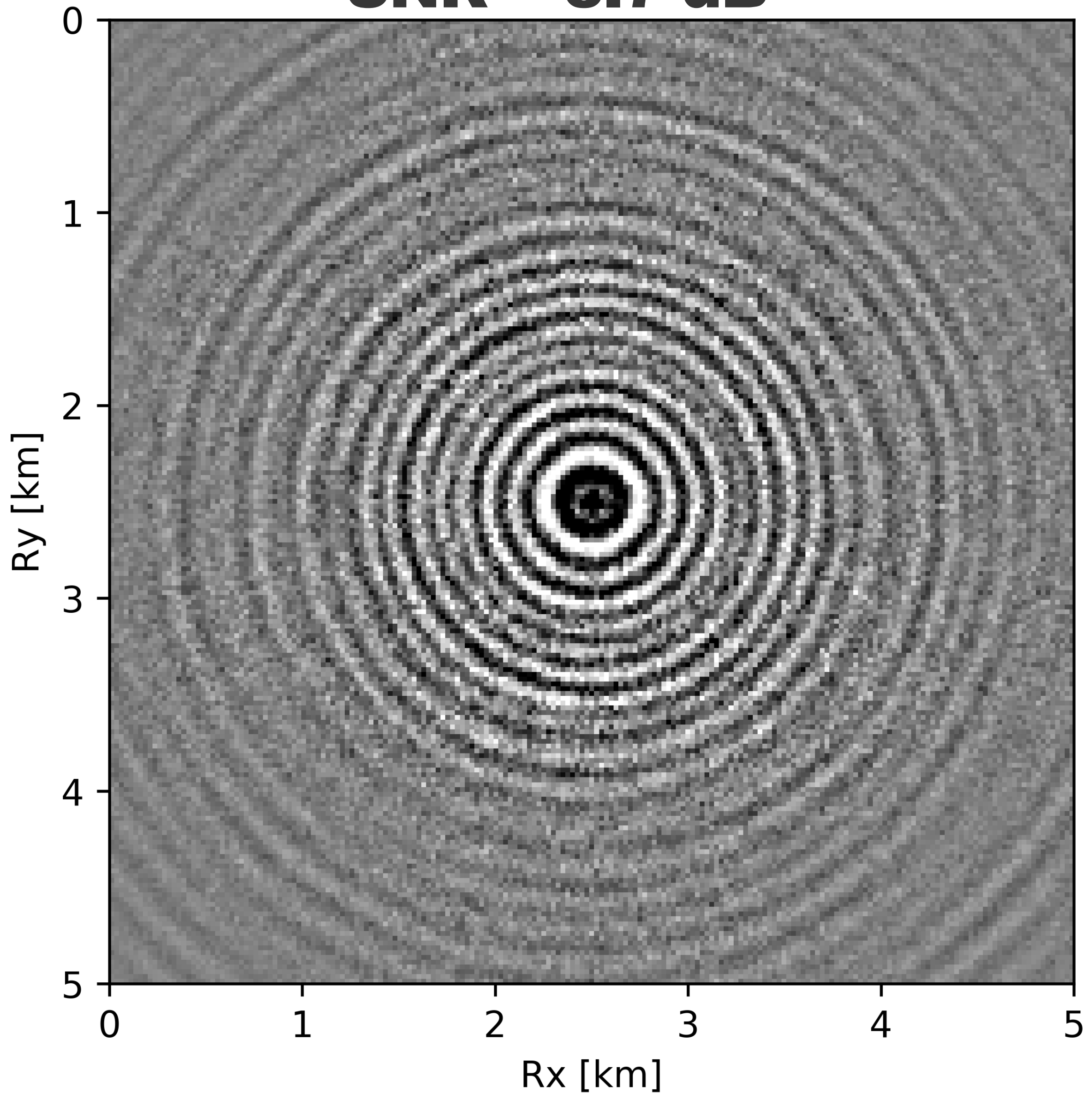


15 Hz Frequency Slice

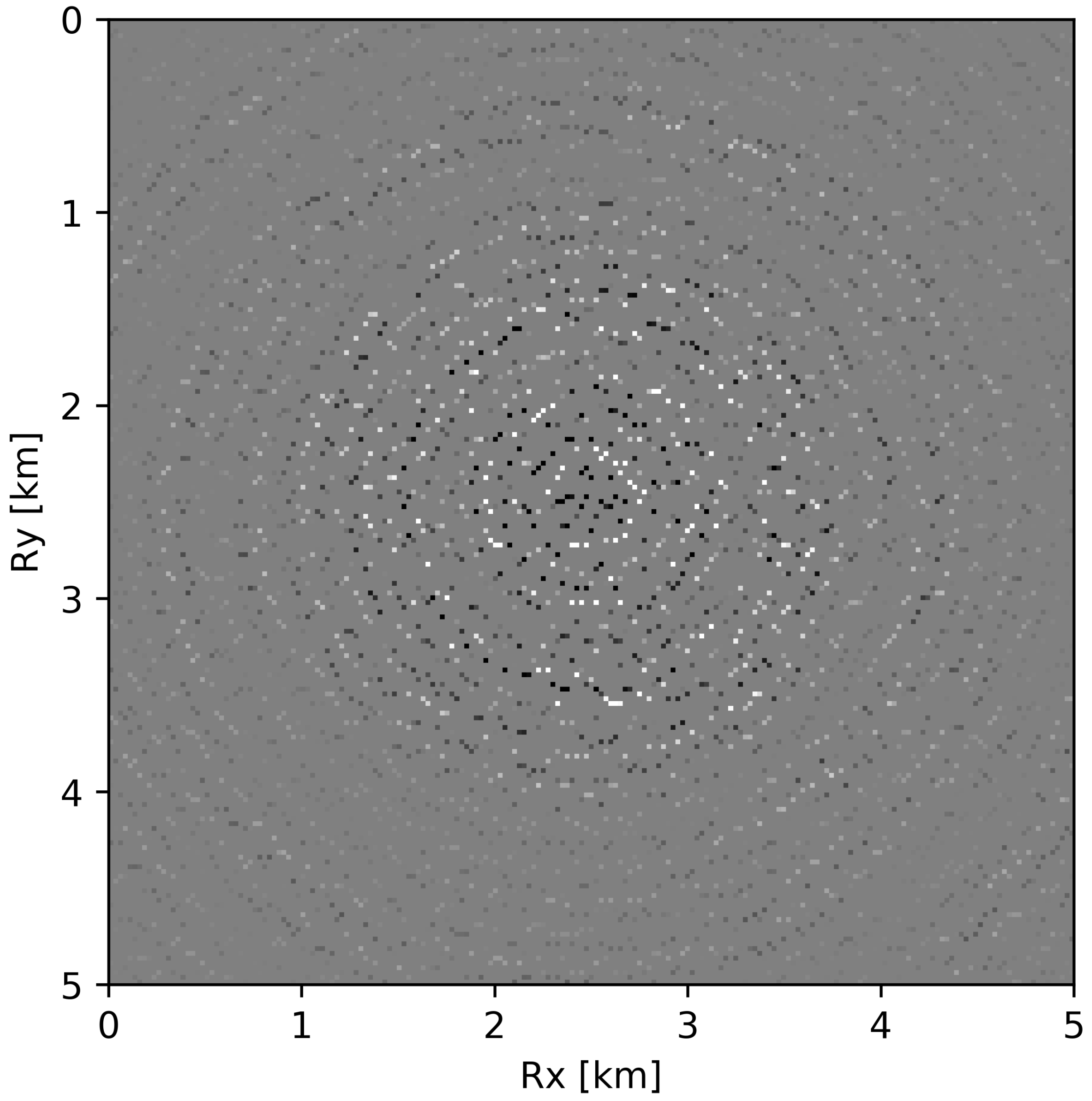
Ground Truth



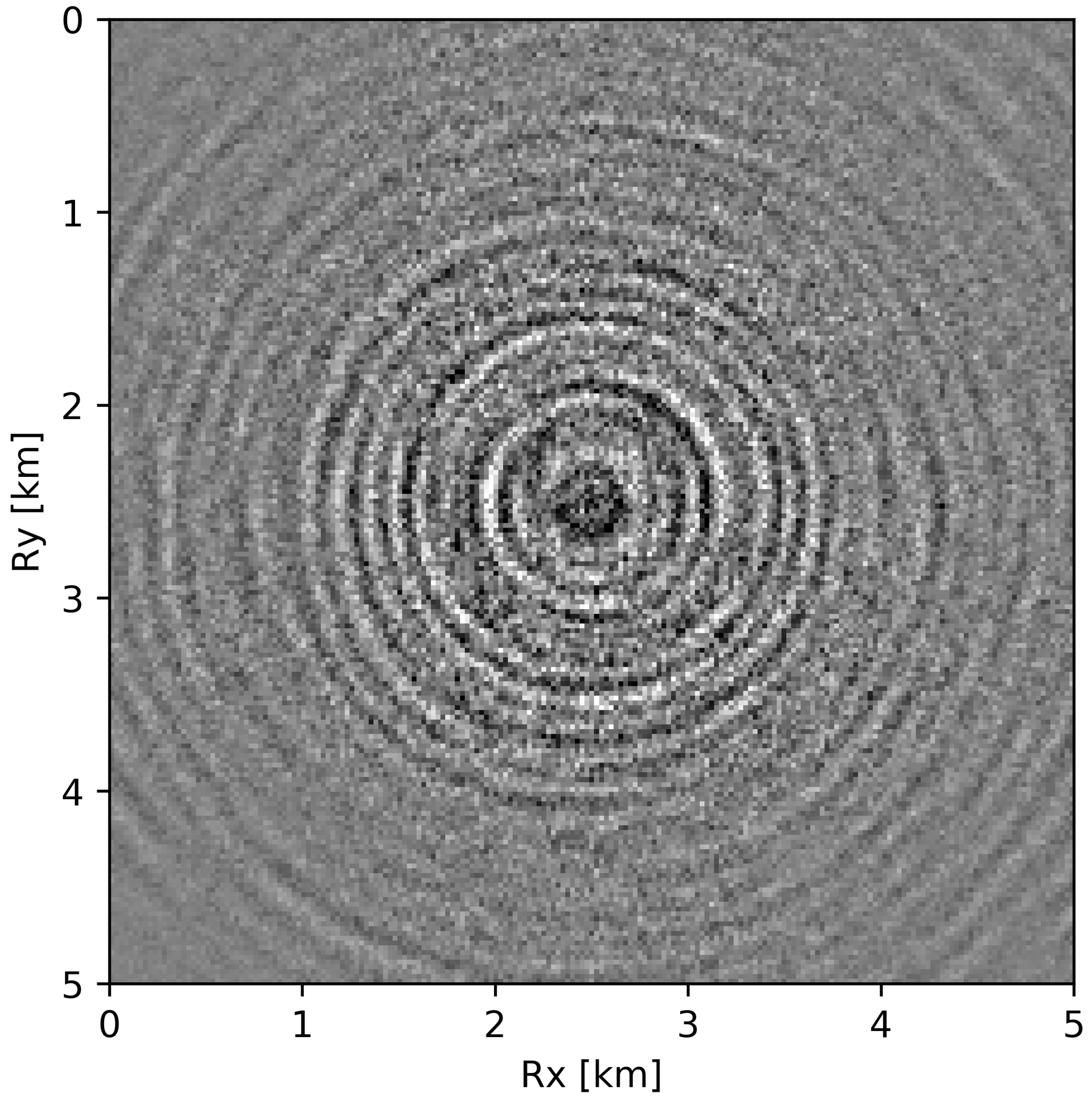
Reconstruction w/ conventional, SNR = 3.7 dB



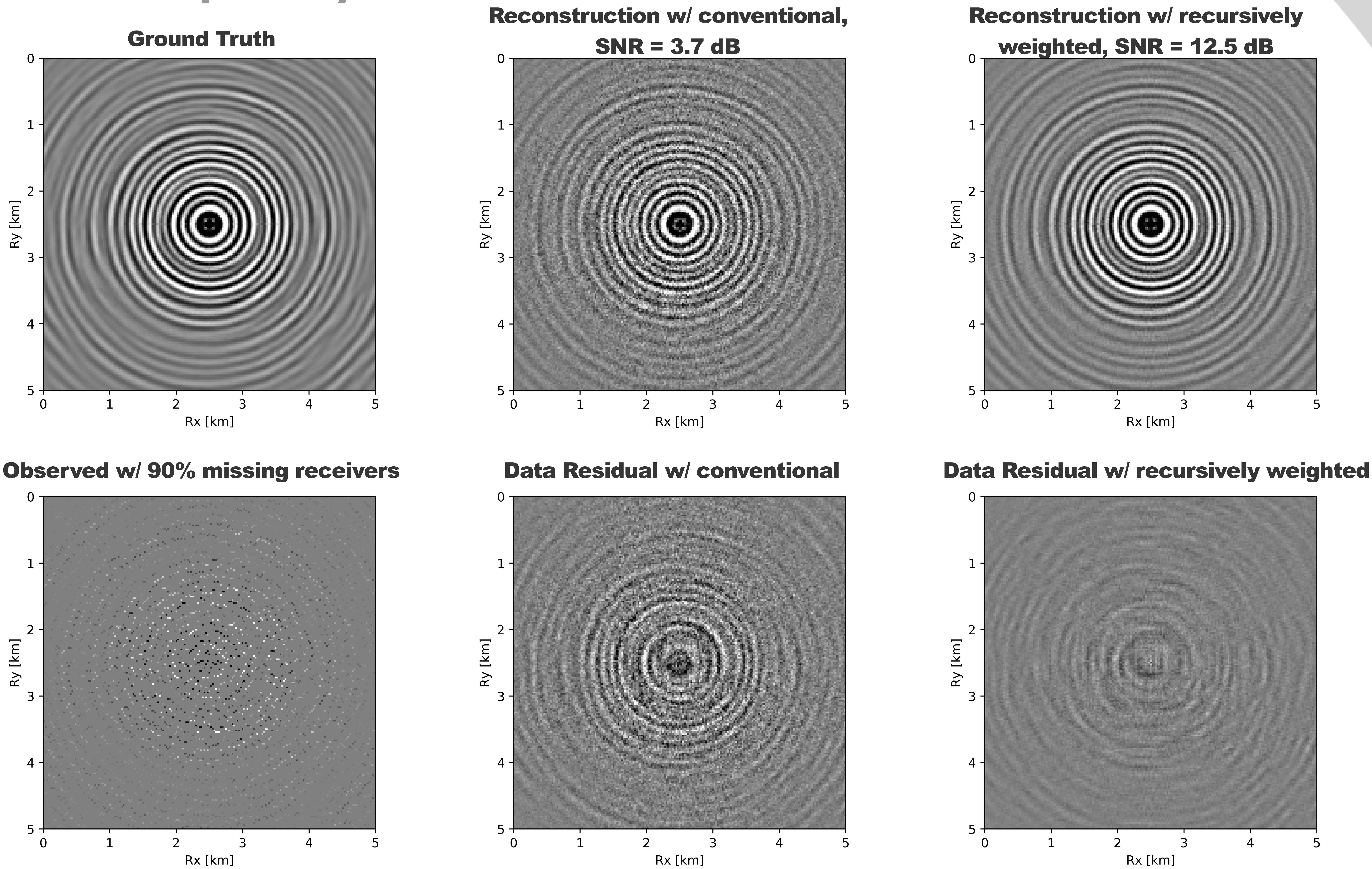
Observed w/ 90% missing receivers



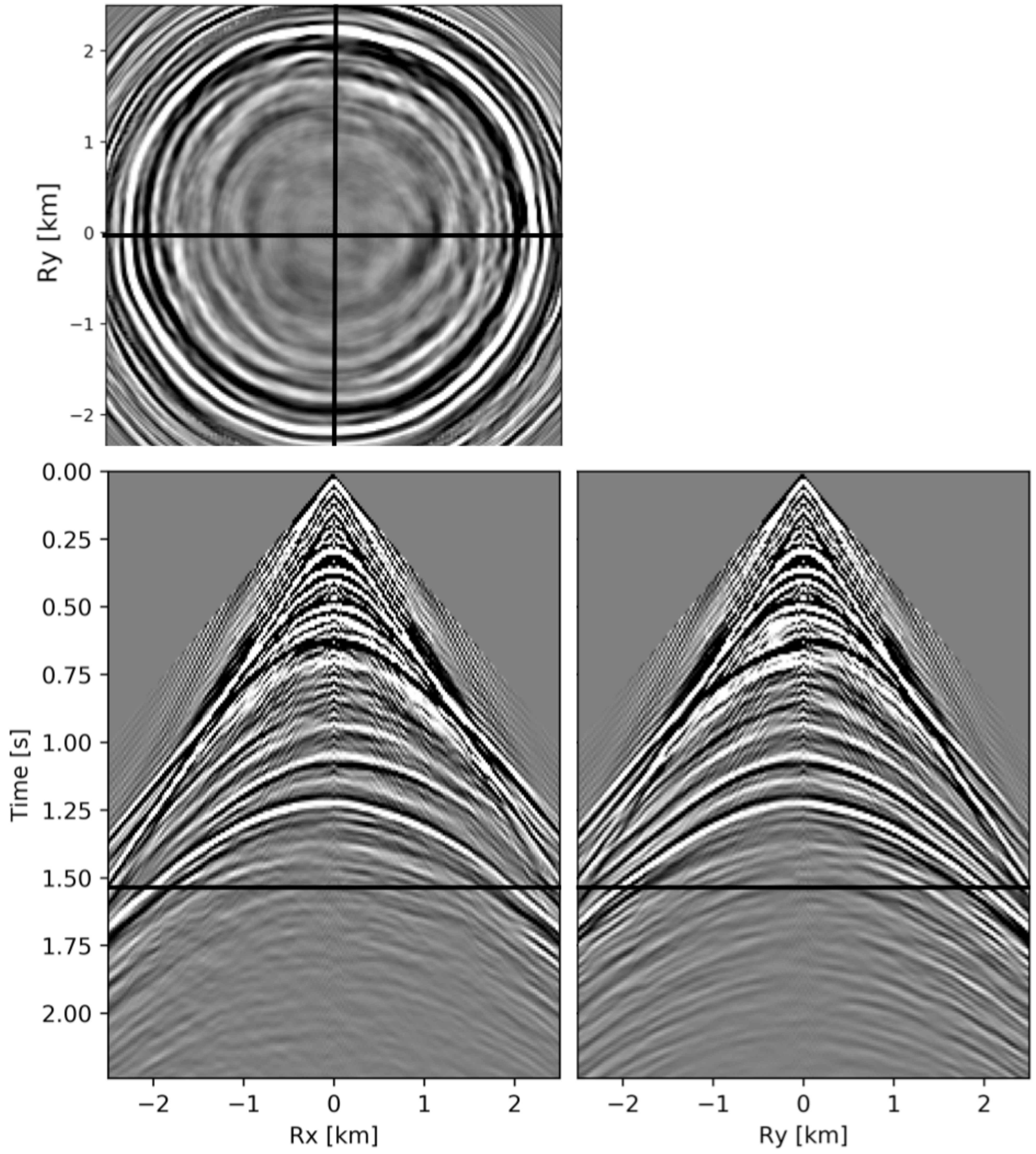
Data Residual w/ conventional



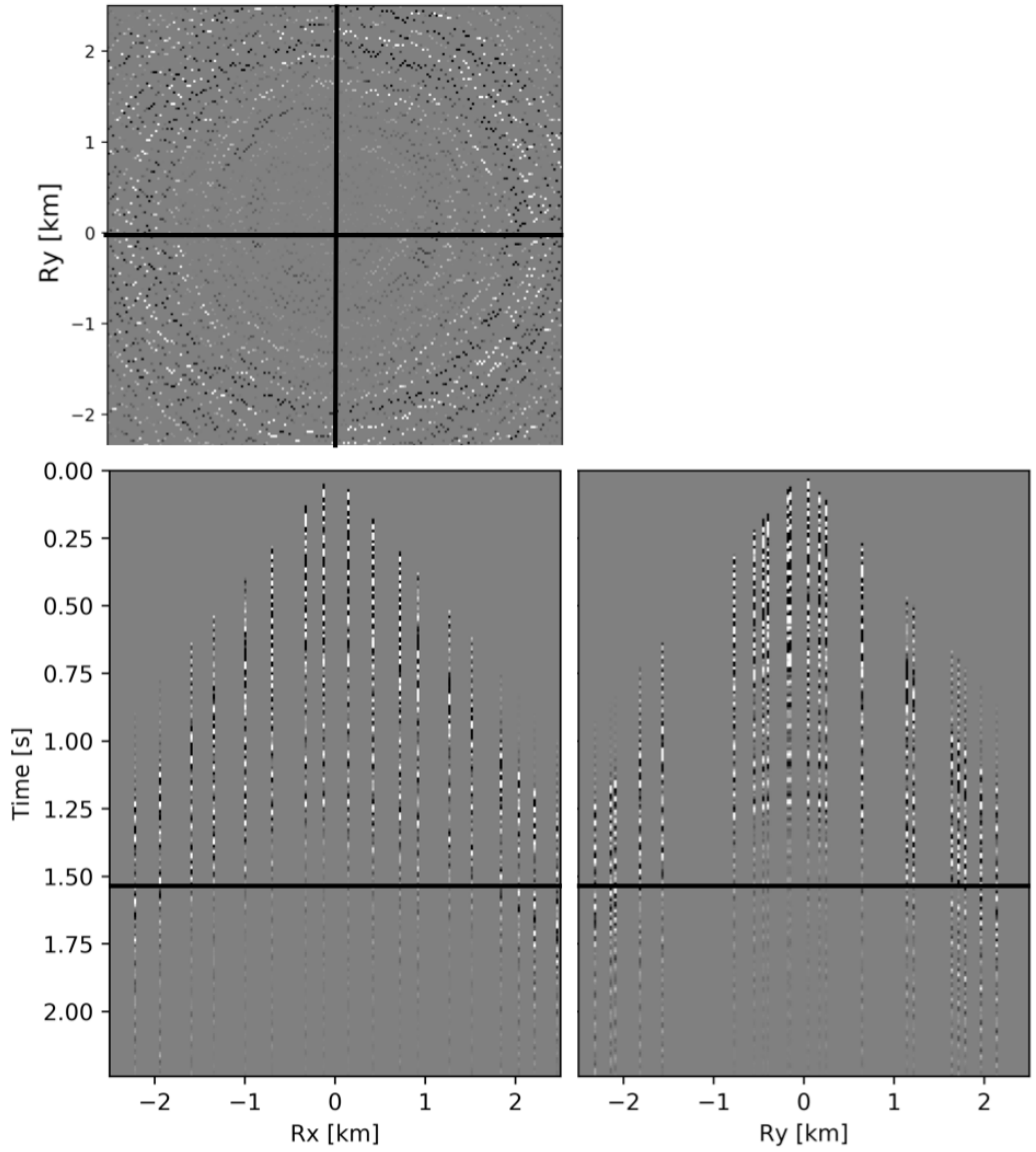
15 Hz Frequency Slice



Time Domain Results

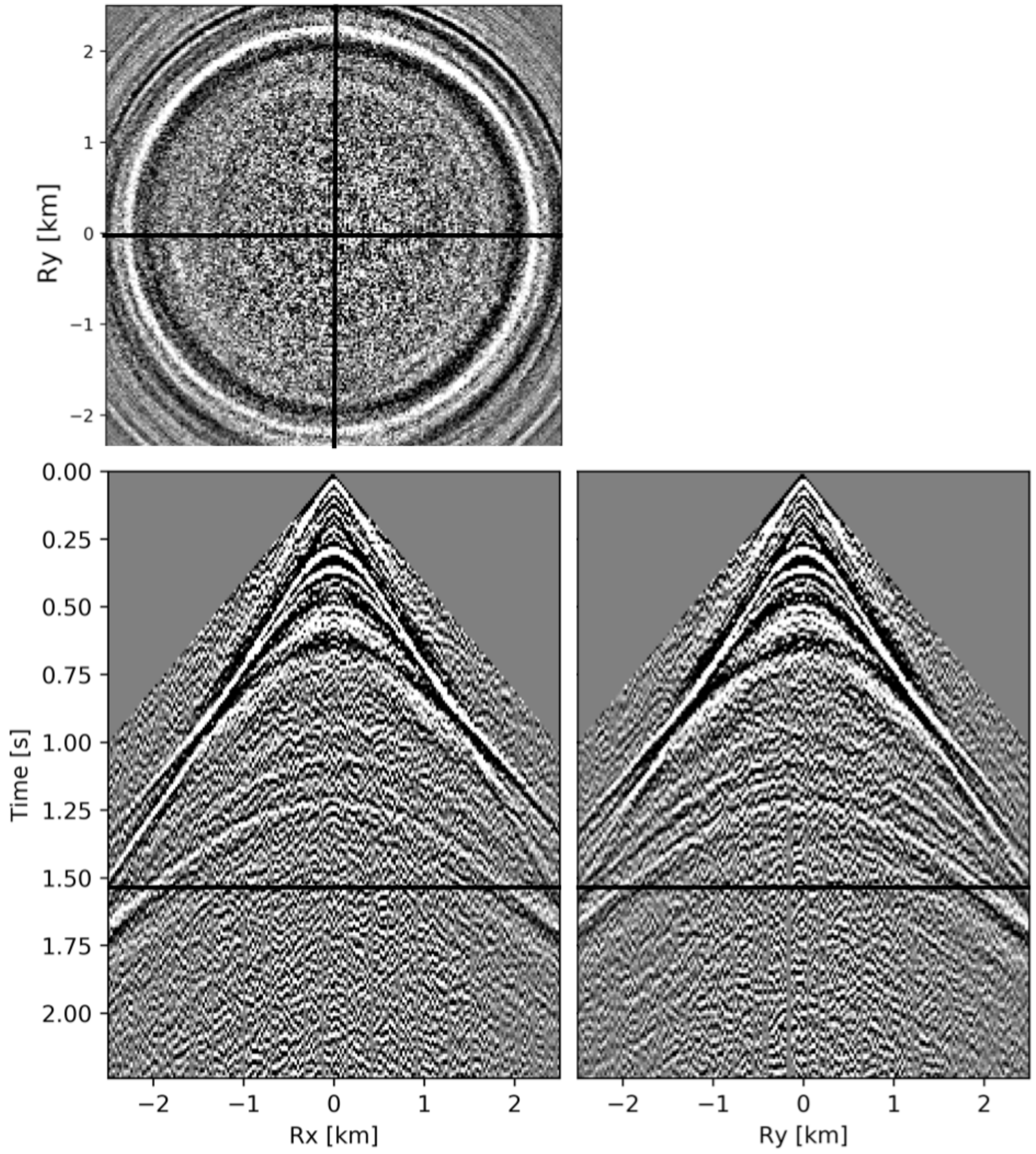


Ground Truth

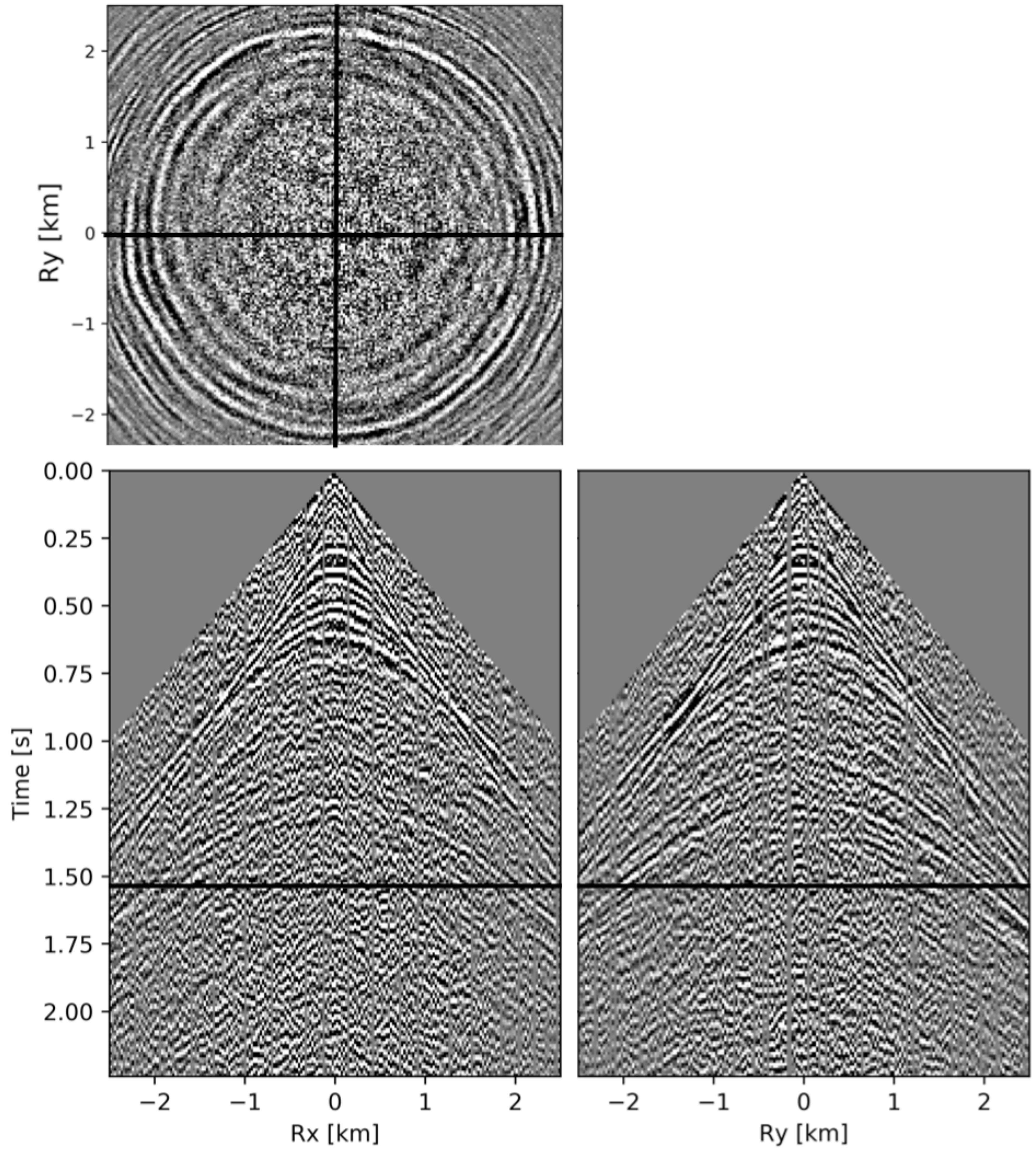


Observed Data w/ 90% missing receivers

Time Domain Results

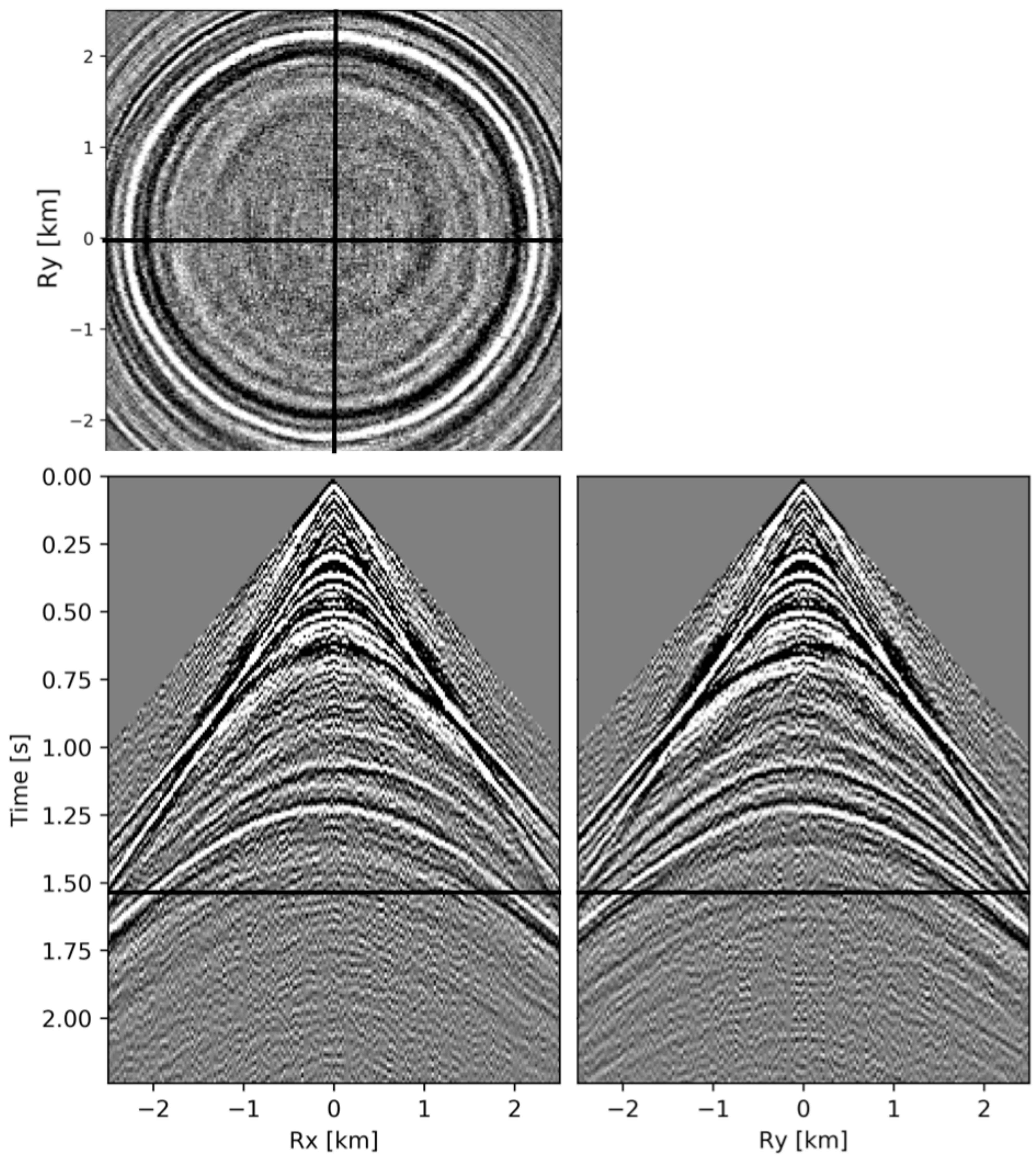


Reconstruction w/ conventional

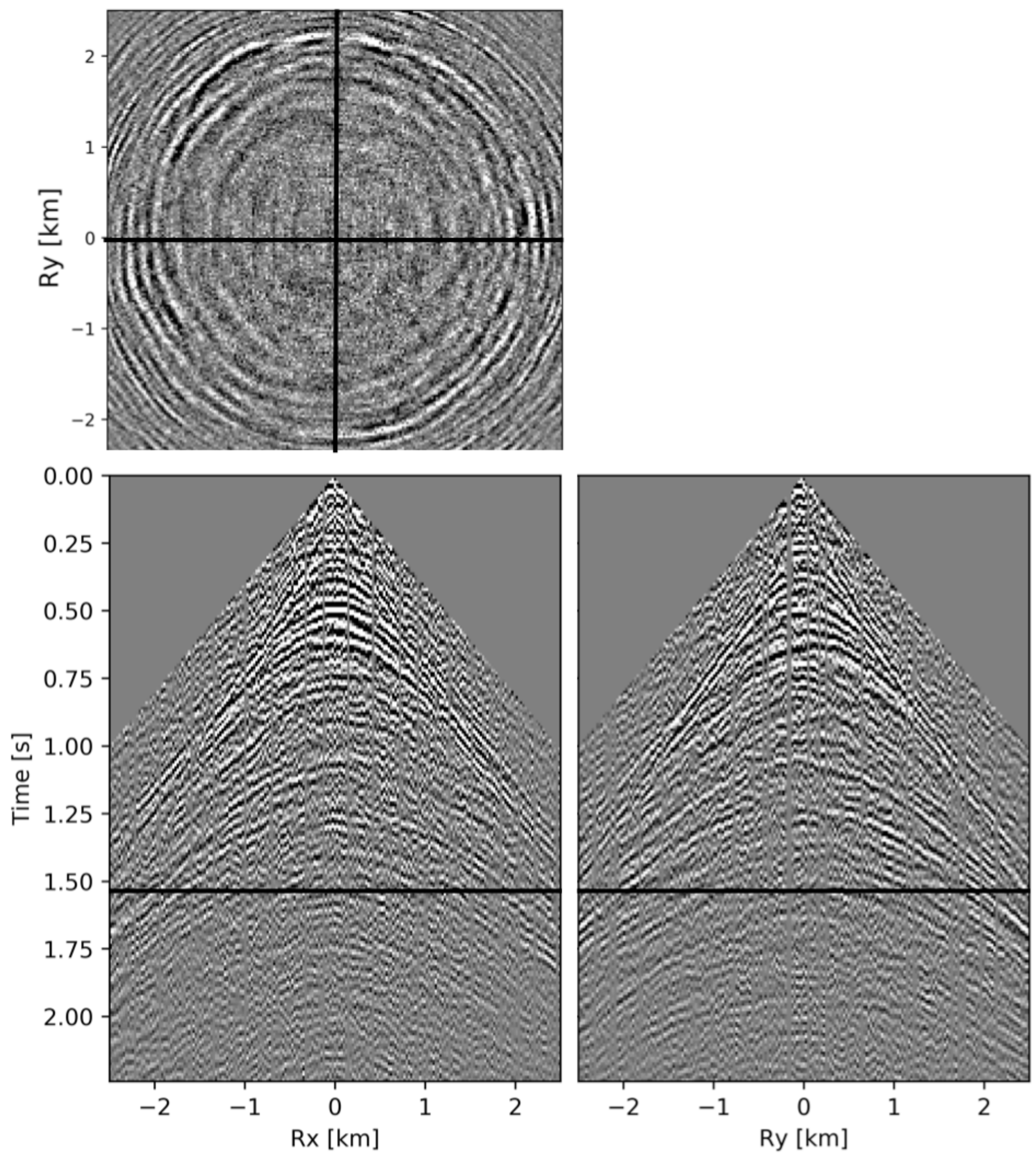


Data residual

Time Domain Results

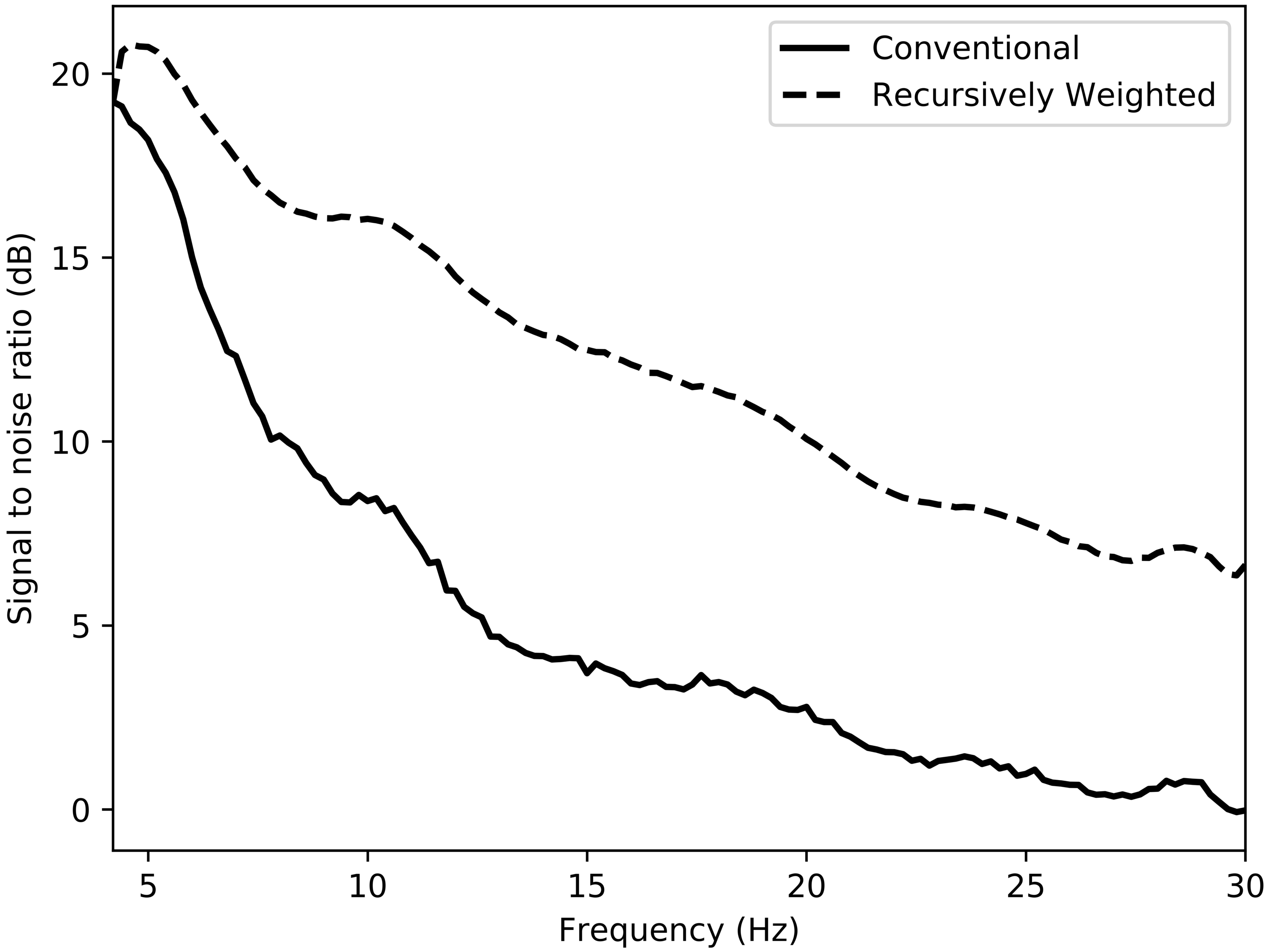


Reconstruction w/ recursively weighted

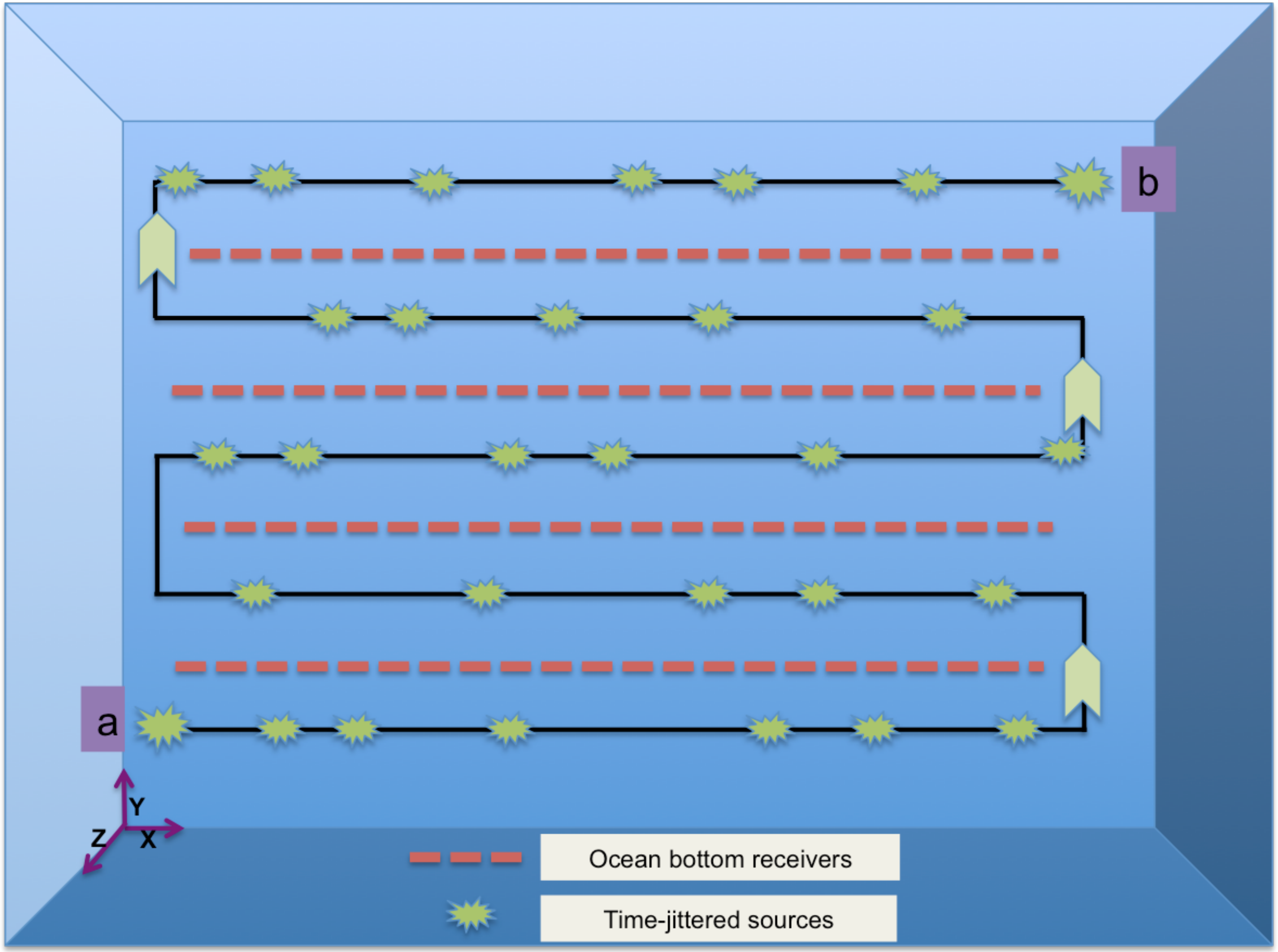


Data residual

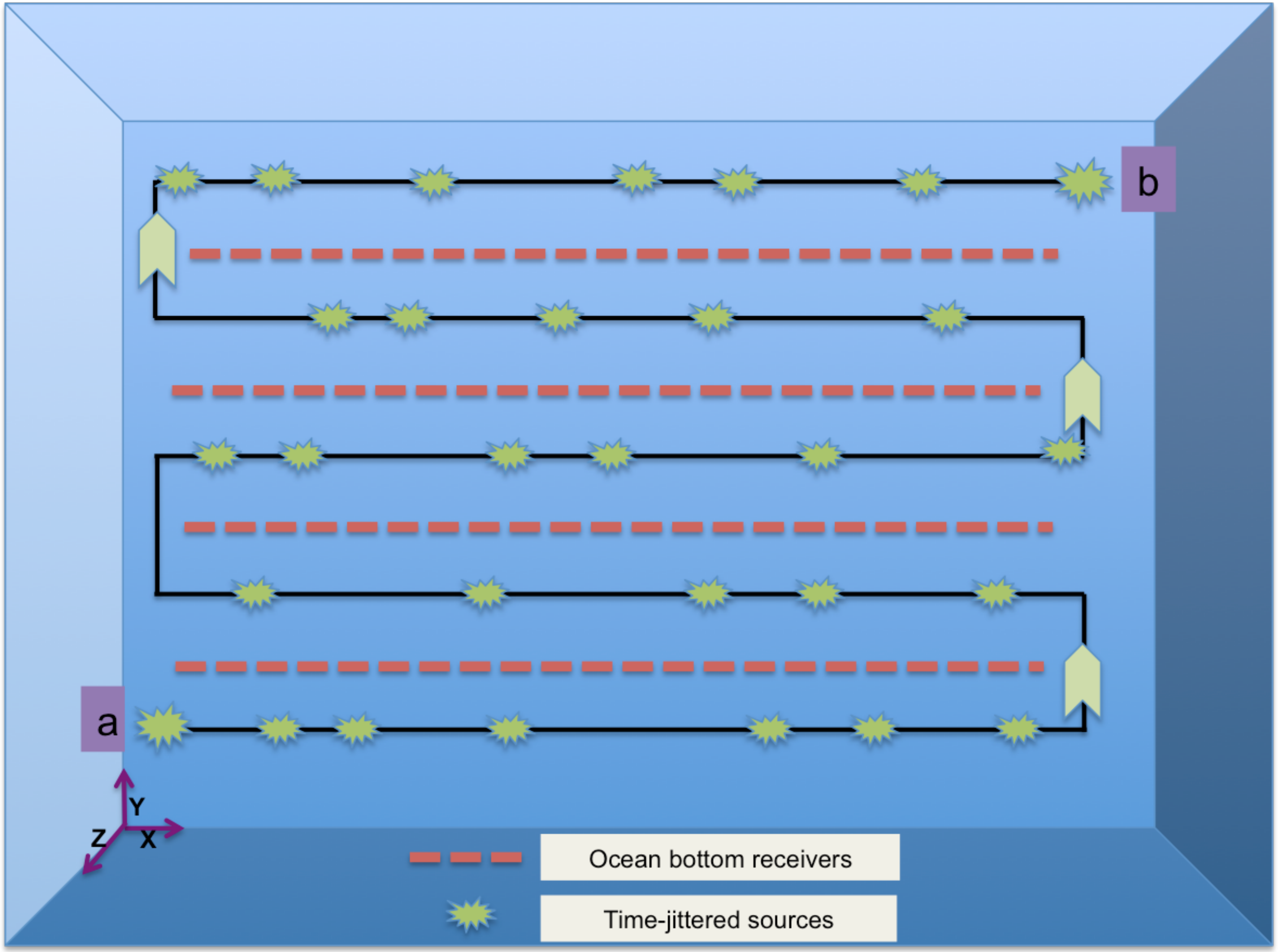
Signal to noise ratio comparison



5D Time-Jittered marine acquisition



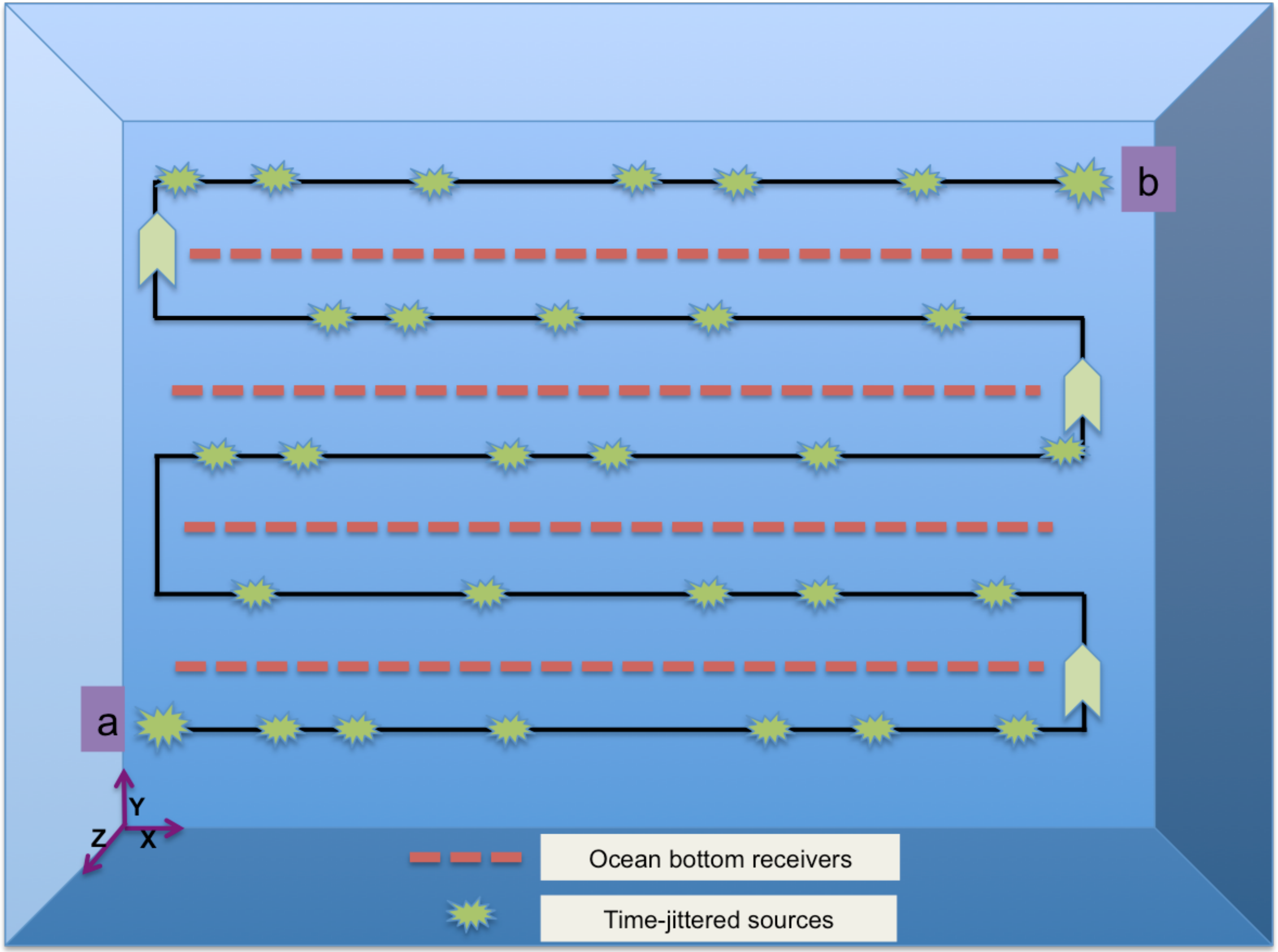
5D Time-Jittered marine acquisition



Objective

- Acquire blended seismic data using multiple sources
- Simultaneous separation and reconstruction of sources on dense grid

5D Time-Jittered marine acquisition



Objective

- Acquire blended seismic data using multiple sources
- Simultaneous separation and reconstruction of sources on dense grid

Benefits

- Reduction in overall cost of acquiring dense seismic data

Low-rank formulation

Restriction operator is non-separable

- combination of time-shifting and shot-jittered operator

Low-rank formulation

Restriction operator is non-separable

- combination of time-shifting and shot-jittered operator

$$\mathcal{A}(.) = \mathbf{M}\mathbf{F}^H \mathcal{S}^H (.)$$

$\mathbf{M}$	time-jittered operator
$\mathbf{F}^H$	inverse Fourier transform along frequency axis
$\mathcal{S}$	rank-revealing transform

Low-rank formulation

Restriction operator is non-separable

- combination of time-shifting and shot-jittered operator

$$\mathcal{A}(.) = \mathbf{M}\mathbf{F}^H\mathcal{S}^H(.)$$

$\mathbf{M}$	time-jittered operator
$\mathbf{F}^H$	inverse Fourier transform along frequency axis
$\mathcal{S}$	rank-revealing transform

Can't perform matrix-completion over independent frequencies

- reformulate low-rank factorization over temporal-frequency domain

Low-rank formulation

Restriction operator is non-separable

- combination of time-shifting and shot-jittered operator

$$\mathcal{A}(.) = \mathbf{M}\mathbf{F}^H\mathcal{S}^H(.)$$

$\mathbf{M}$	time-jittered operator
$\mathbf{F}^H$	inverse Fourier transform along frequency axis
$\mathcal{S}$	rank-revealing transform

Can't perform matrix-completion over independent frequencies

- reformulate low-rank factorization over temporal-frequency domain

$$\underset{\mathbf{L},\mathbf{R}}{\text{minimize}} \sum_{j}^{n_f} \frac{1}{2} \left\| \begin{bmatrix} \mathbf{L}_j \\ \mathbf{R}_j \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{L}\mathbf{R}^H) - \mathbf{b}\|_2 \leq \epsilon$$

Case Study: 3D BG Compass model

Temporal length

- 65 minutes

25 m flip-flop shooting

- source-sampling ranges from 25 m to 175 m
- acquired 400 sources

101 x 101 receivers (nrx x nry)

Ricker wavelet with central frequency 15 Hz

Dimensions of deblended/interpolated data volume on 6.25 m grid

- 2501 x 101 x 101 x 40 x 40 (nt x nrx x nry x nsx x nsy)

Optimization Information

Computational environment

- SENAI Yemoja cluster

Parallelized over

- receivers and frequencies

Computational information

- 200 iterations, 42 hours

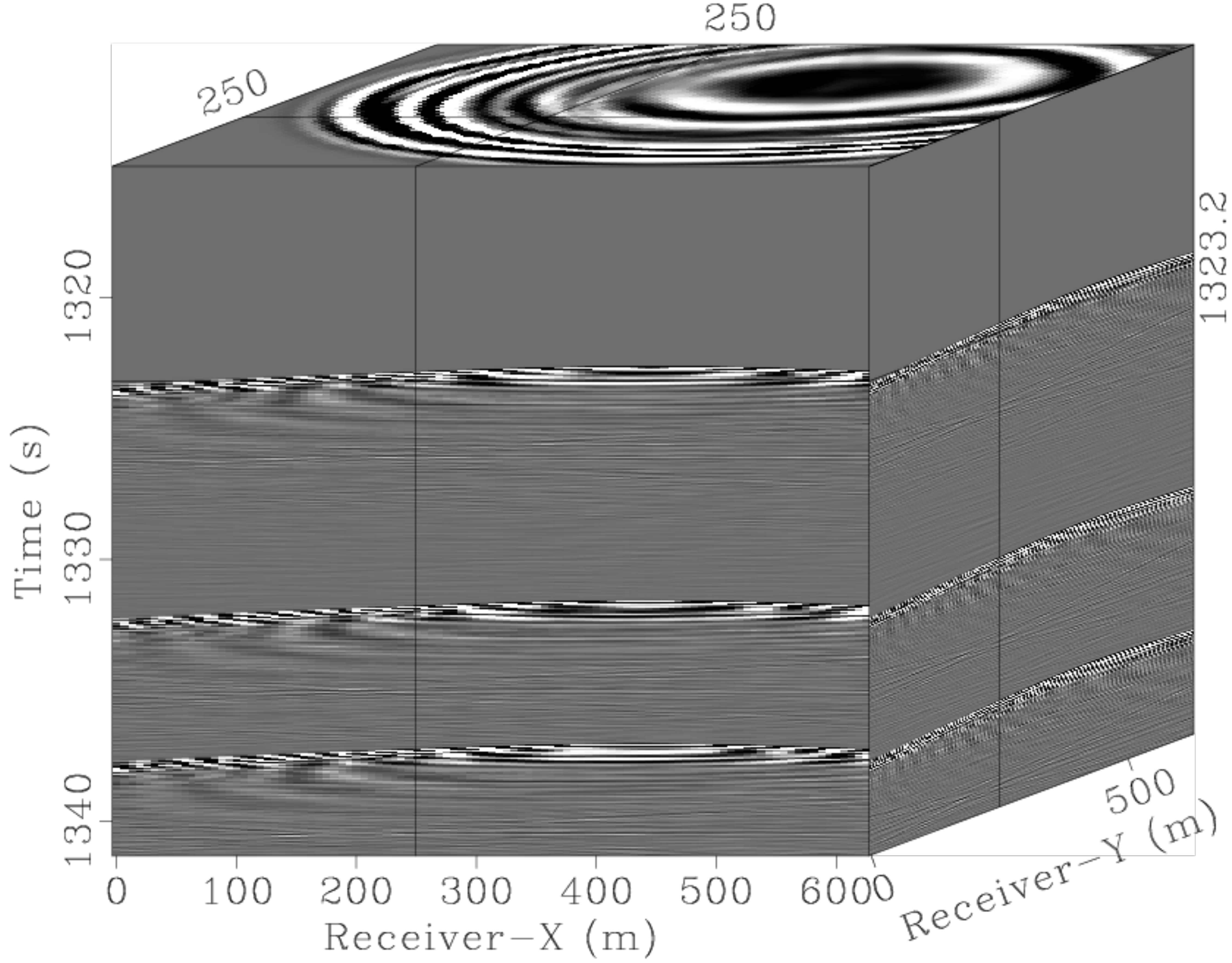
Final volume

- 13 GB (98% compression)

5D Time-Jittered marine acquisition

SLIM

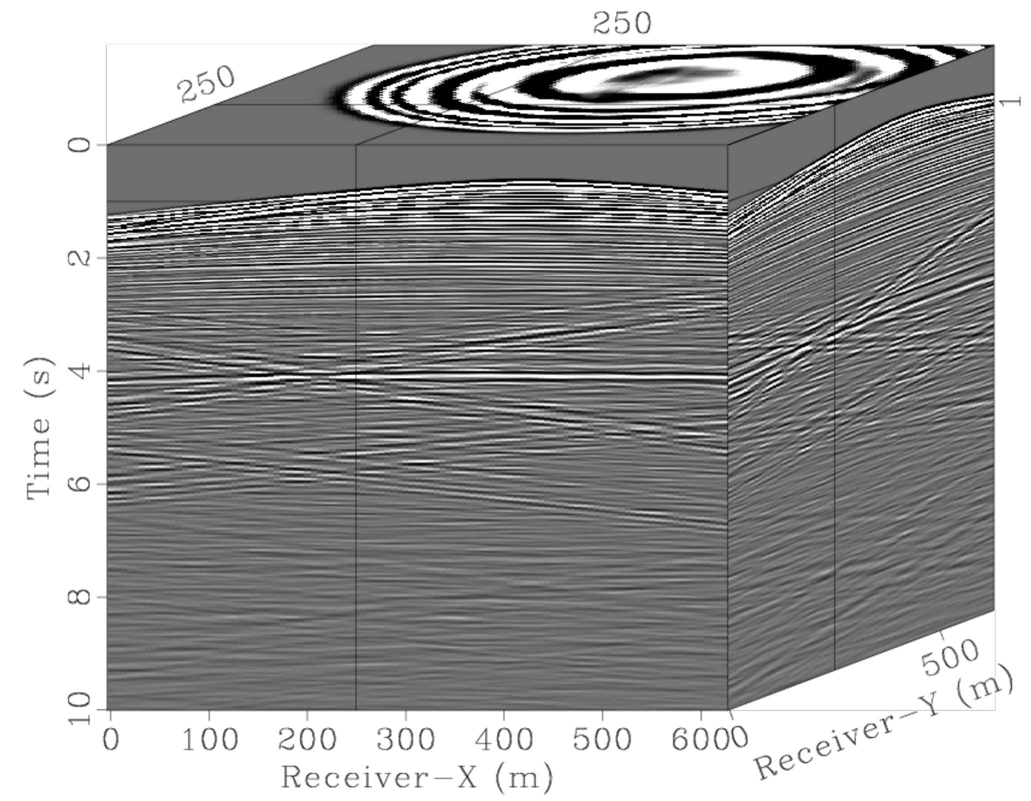
Blended data @ 25 m flip-flop (overlapping & missing shots)

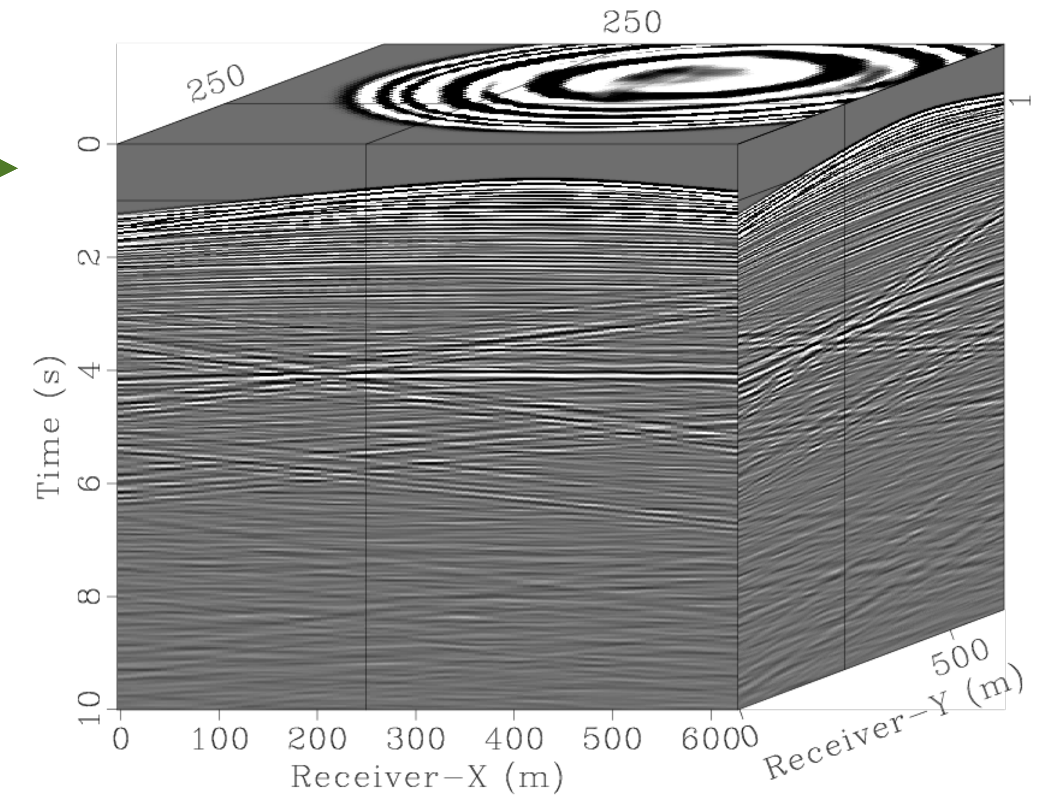


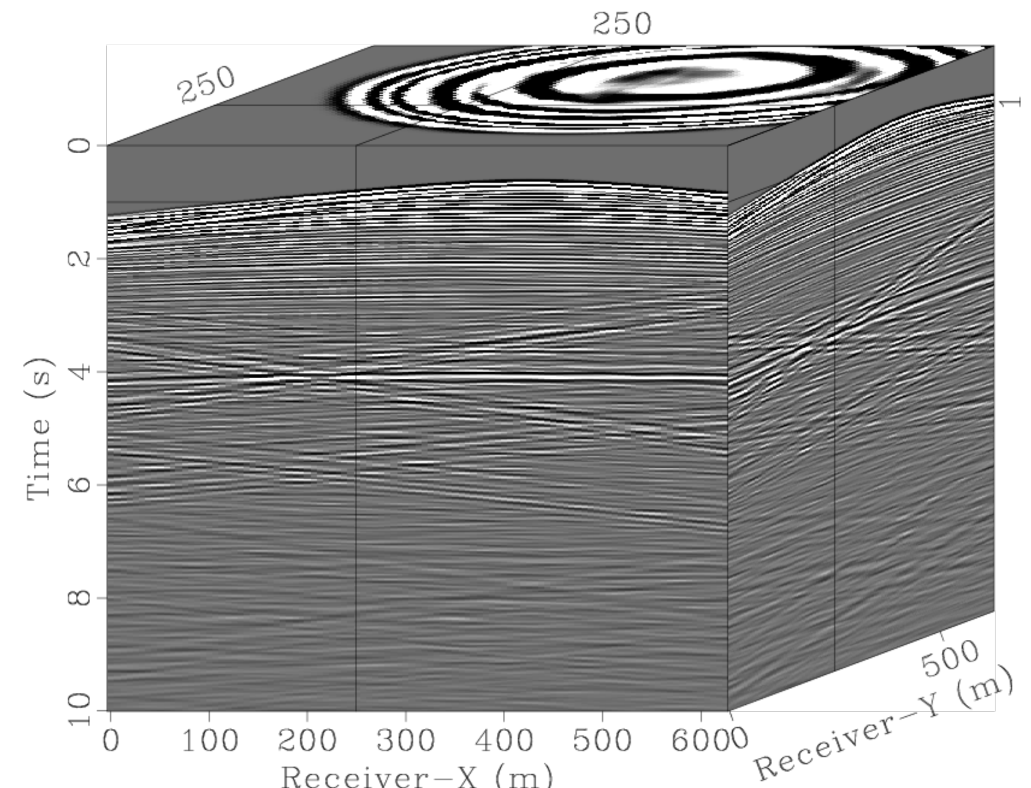
Recovery

SNR = 21 dB

Separation + Interpolation (recovered grid @ 6.25m)







Conclusions

Low-rank matrix factorization based wavefield reconstruction method

- performs poorly at higher frequencies
- recursively weighted method improves reconstruction at higher frequencies
- by including prior information from lower frequencies

Conclusions

Low-rank matrix factorization based wavefield reconstruction method

- performs poorly at higher frequencies
- recursively weighted method improves reconstruction at higher frequencies
- by including prior information from lower frequencies

Scaling for full azimuth industry-size data is achieved via

- shifting the weights from objective to data-misfit to avoid expensive projections
- using strategies of alternating minimization and decoupling
- parallelizing over rows of low-rank factors

SLIM

Conclusions

Low-rank matrix factorization based wavefield reconstruction method

- performs poorly at higher frequencies
- recursively weighted method improves reconstruction at higher frequencies
- by including prior information from lower frequencies

Scaling for full azimuth industry-size data is achieved via

- shifting the weights from objective to data-misfit to avoid expensive projections
- using strategies of alternating minimization and decoupling
- parallelizing over rows of low-rank factors

Factorization based time-jittered acquisition

- scalable for large scale 5D data
- by using parallel computation
- achieves data compression by saving low-rank factors

Future work

In recursively weighted framework

- include smaller weights
- for large scale data

Future work

In recursively weighted framework

- include smaller weights
- for large scale data

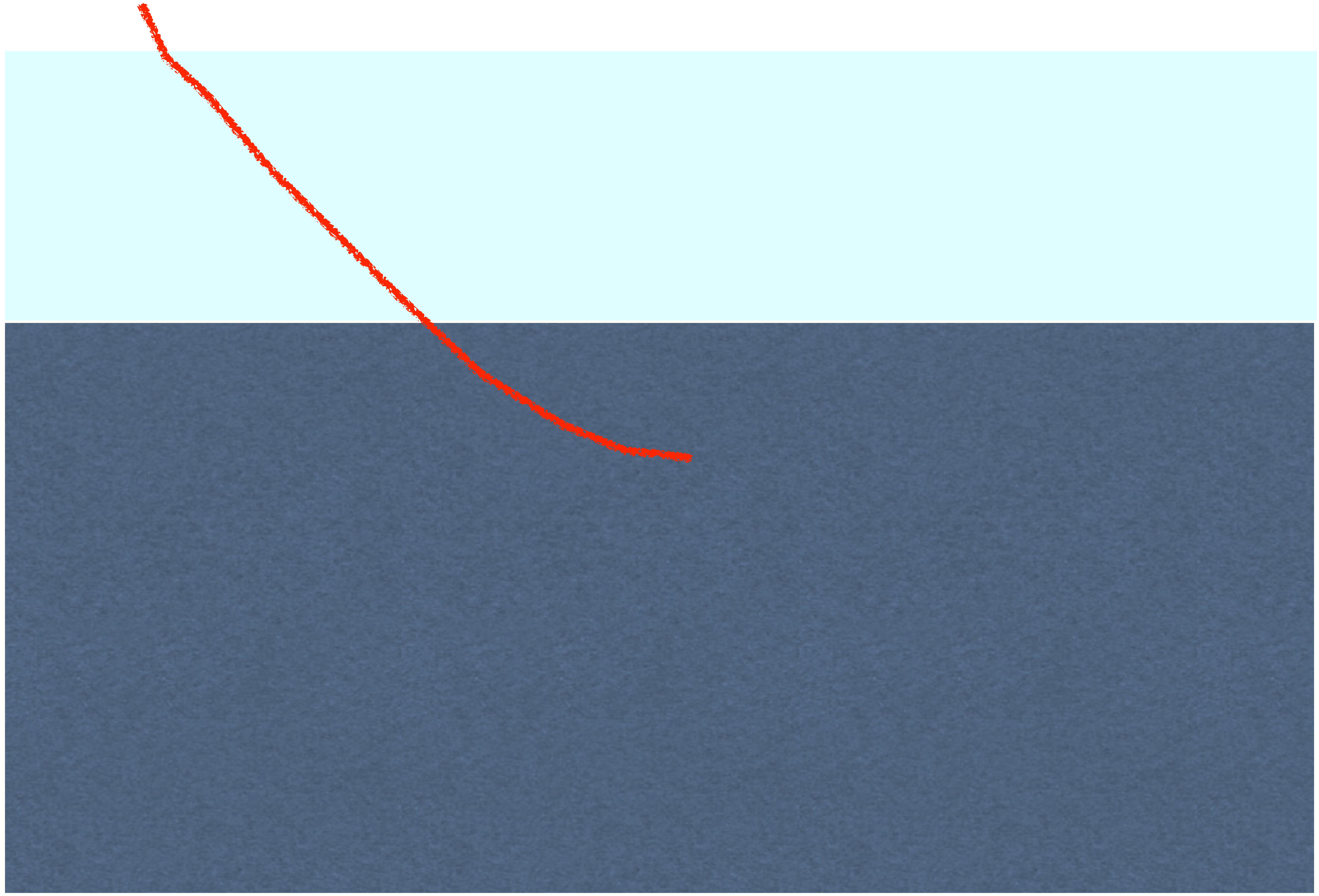
In time-jittered acquisition

- include weights
- to further improve reconstruction quality

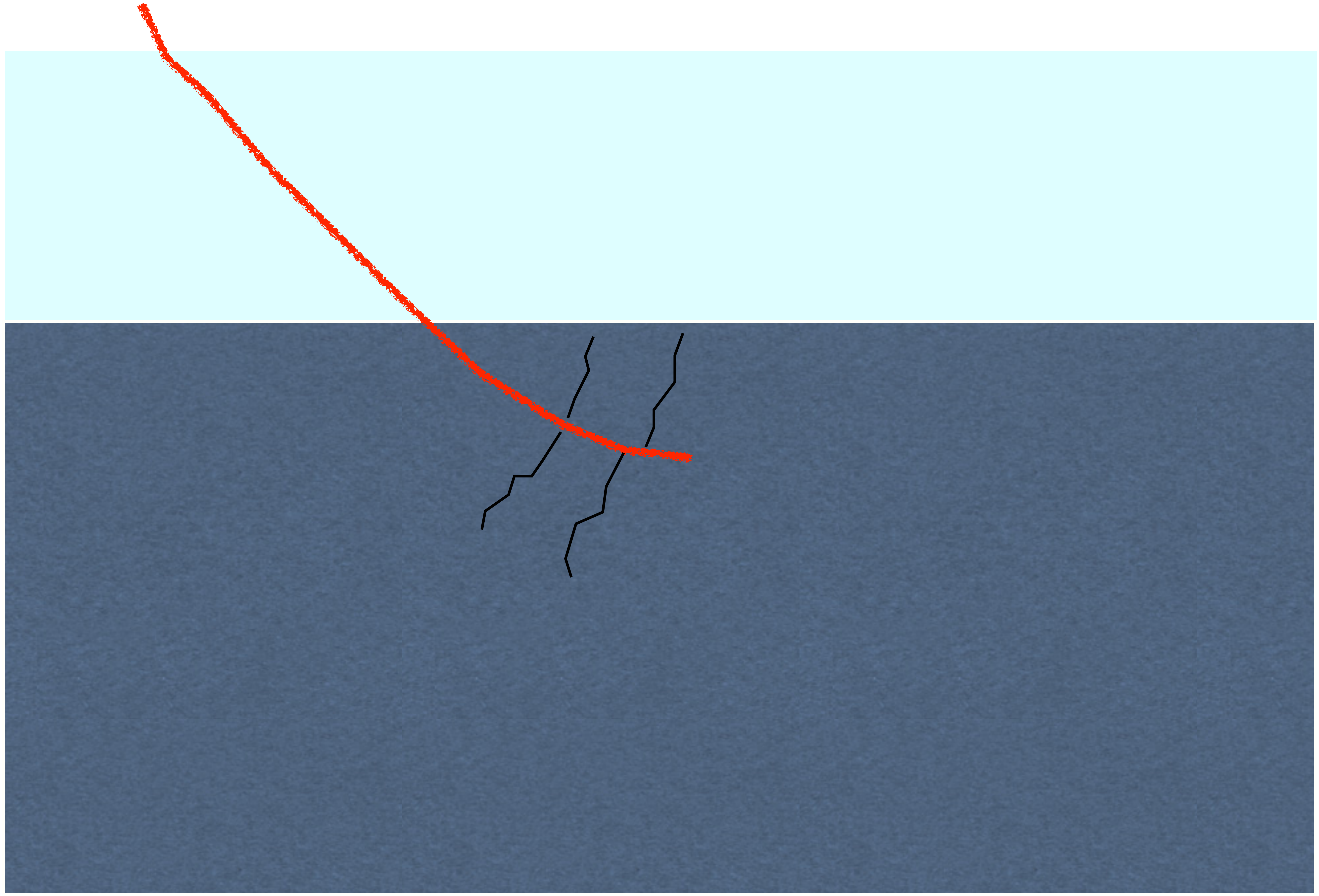
Sparsity-promoting source estimation

Unconventional Reservoir

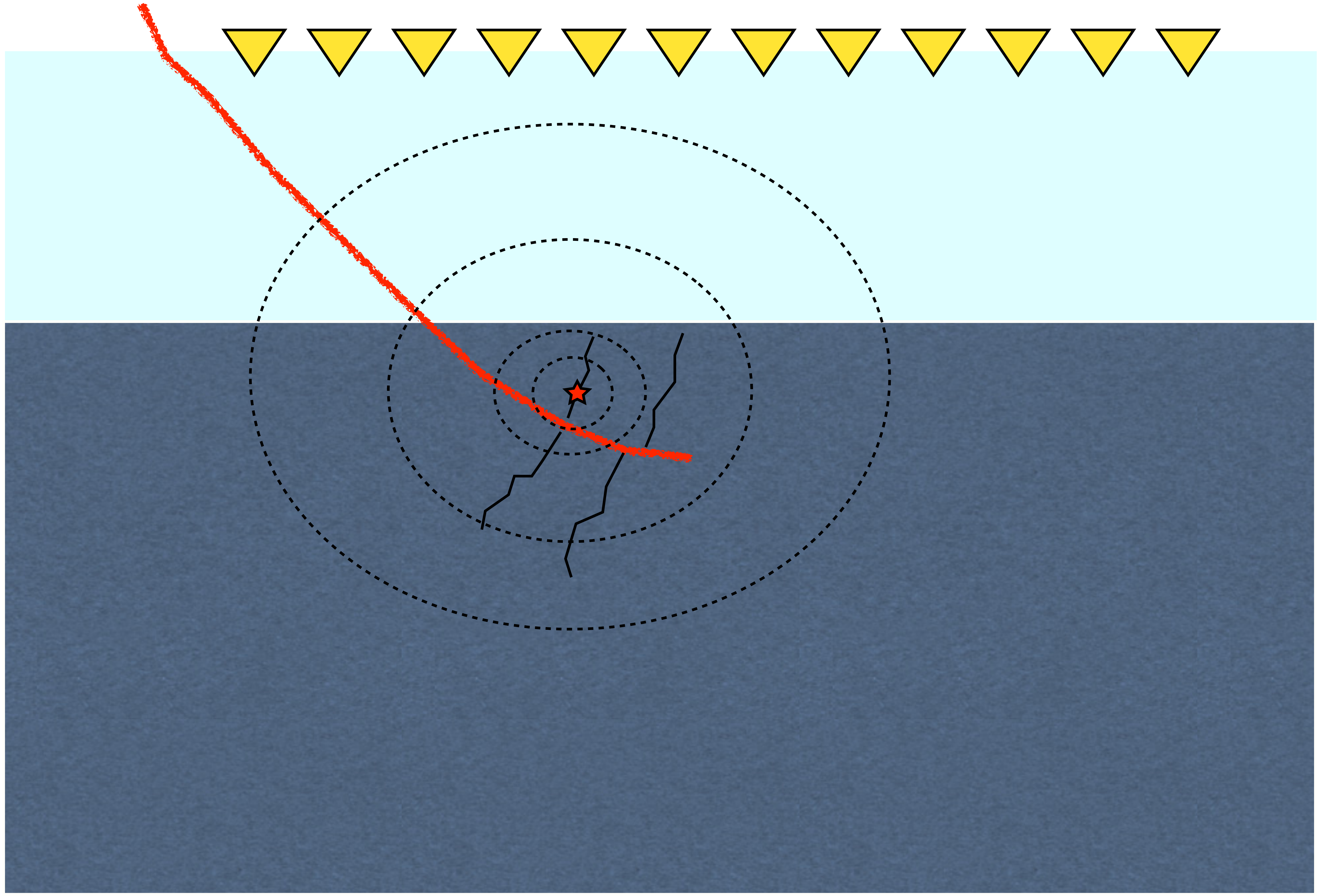
Unconventional Reservoir



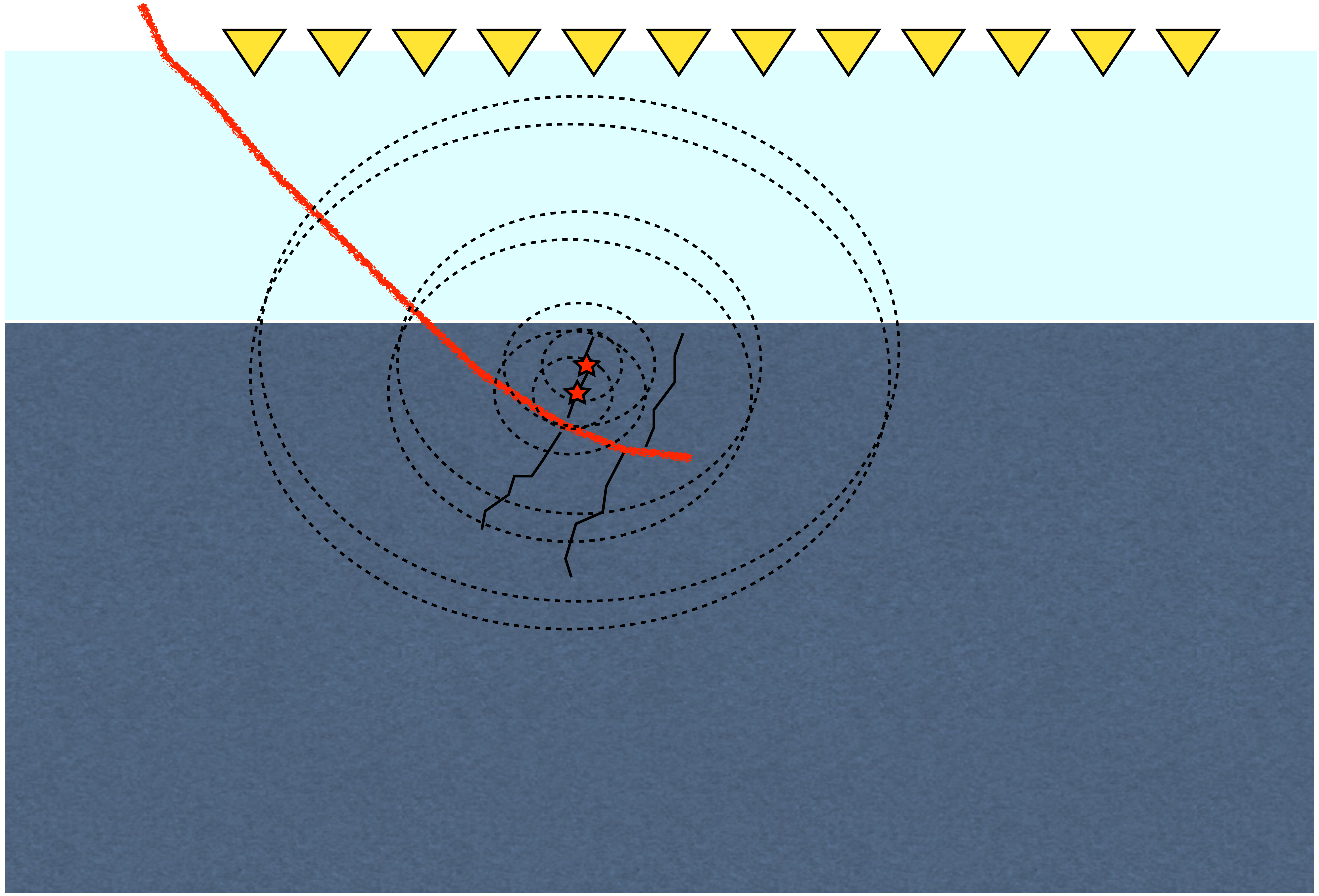
Unconventional Reservoir



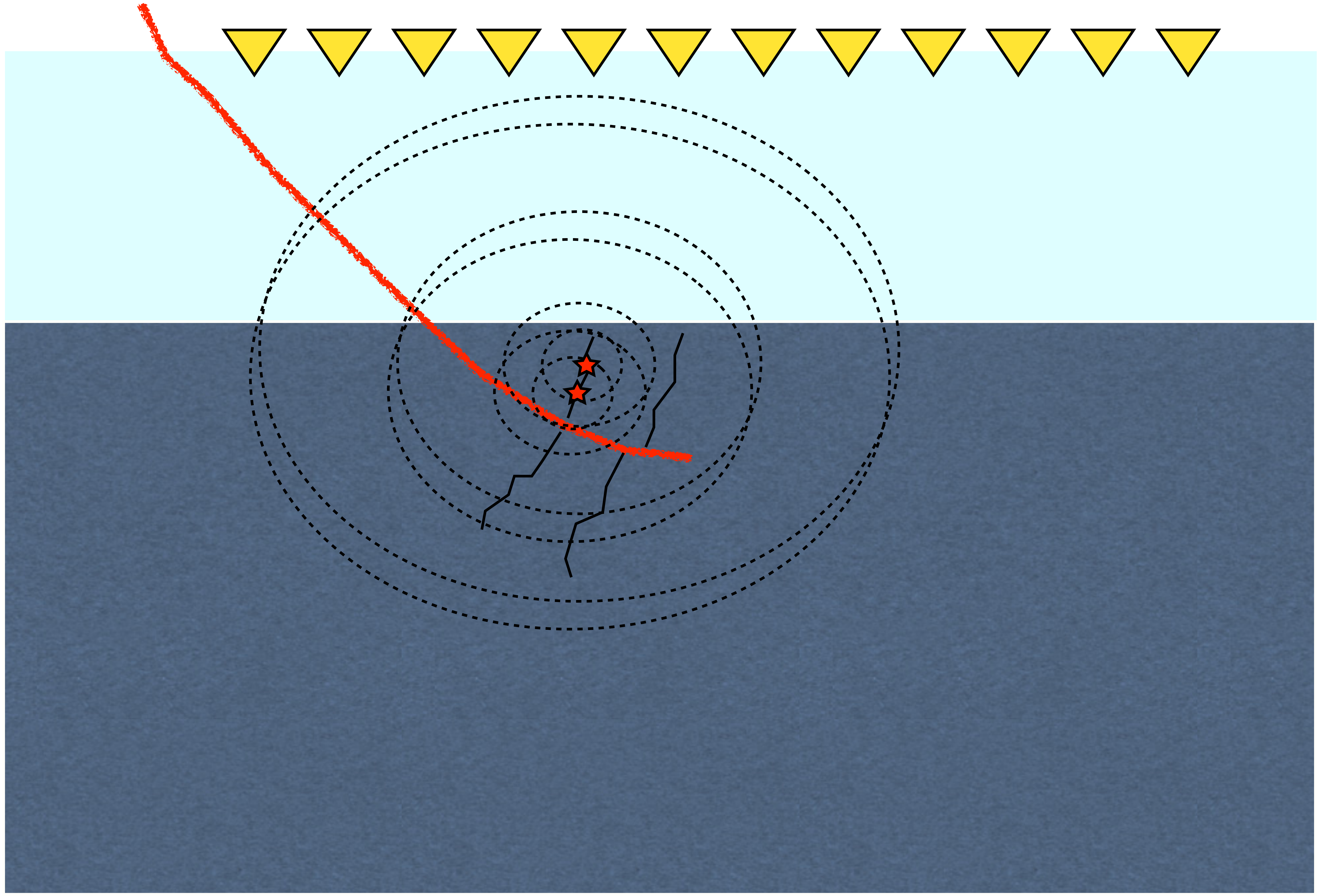
Unconventional Reservoir



Unconventional Reservoir



Unconventional Reservoir



Objectives

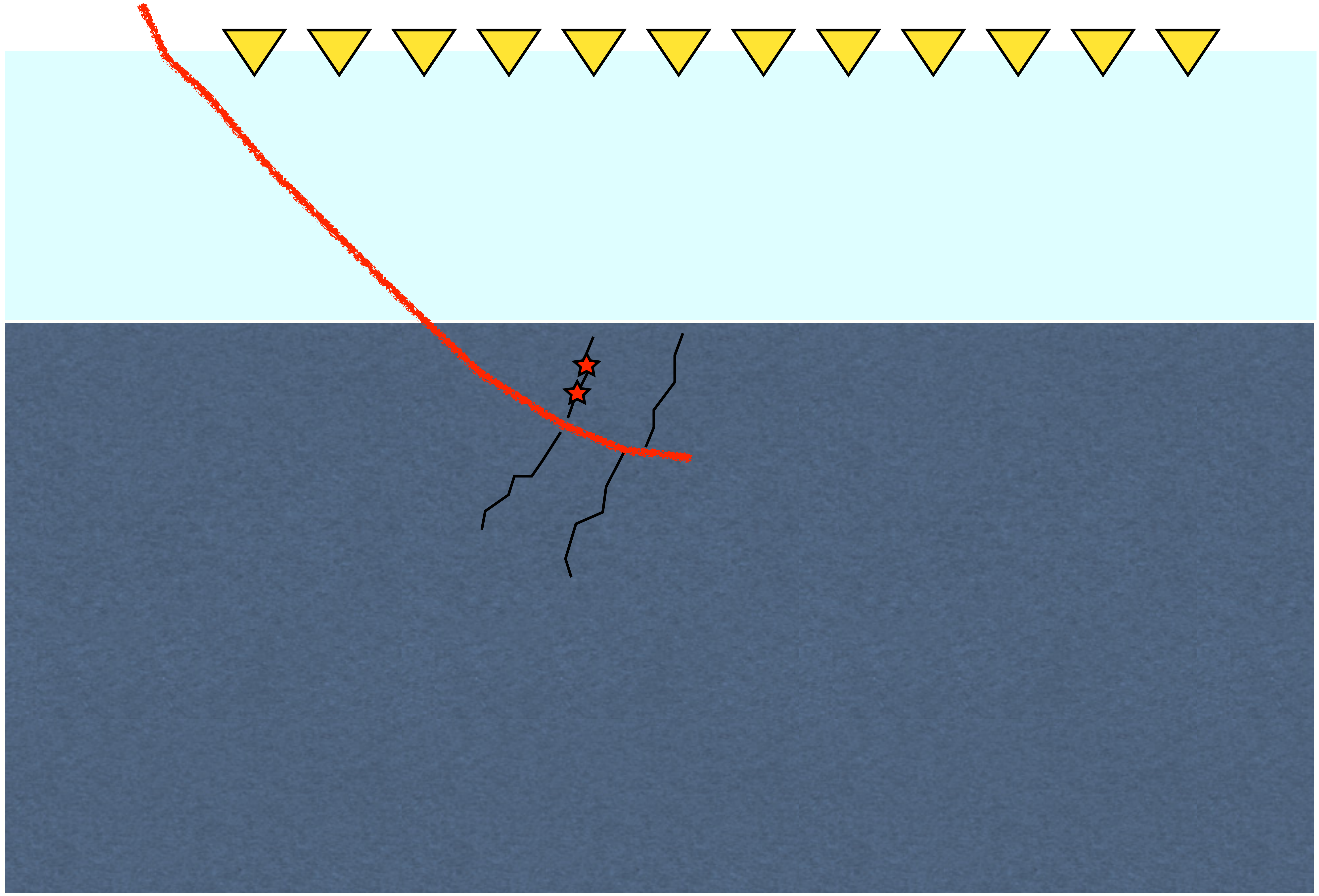
- detection of microseismic events in space and time
- estimation of source-time function

Key Contributions: Chapters 4 & 5

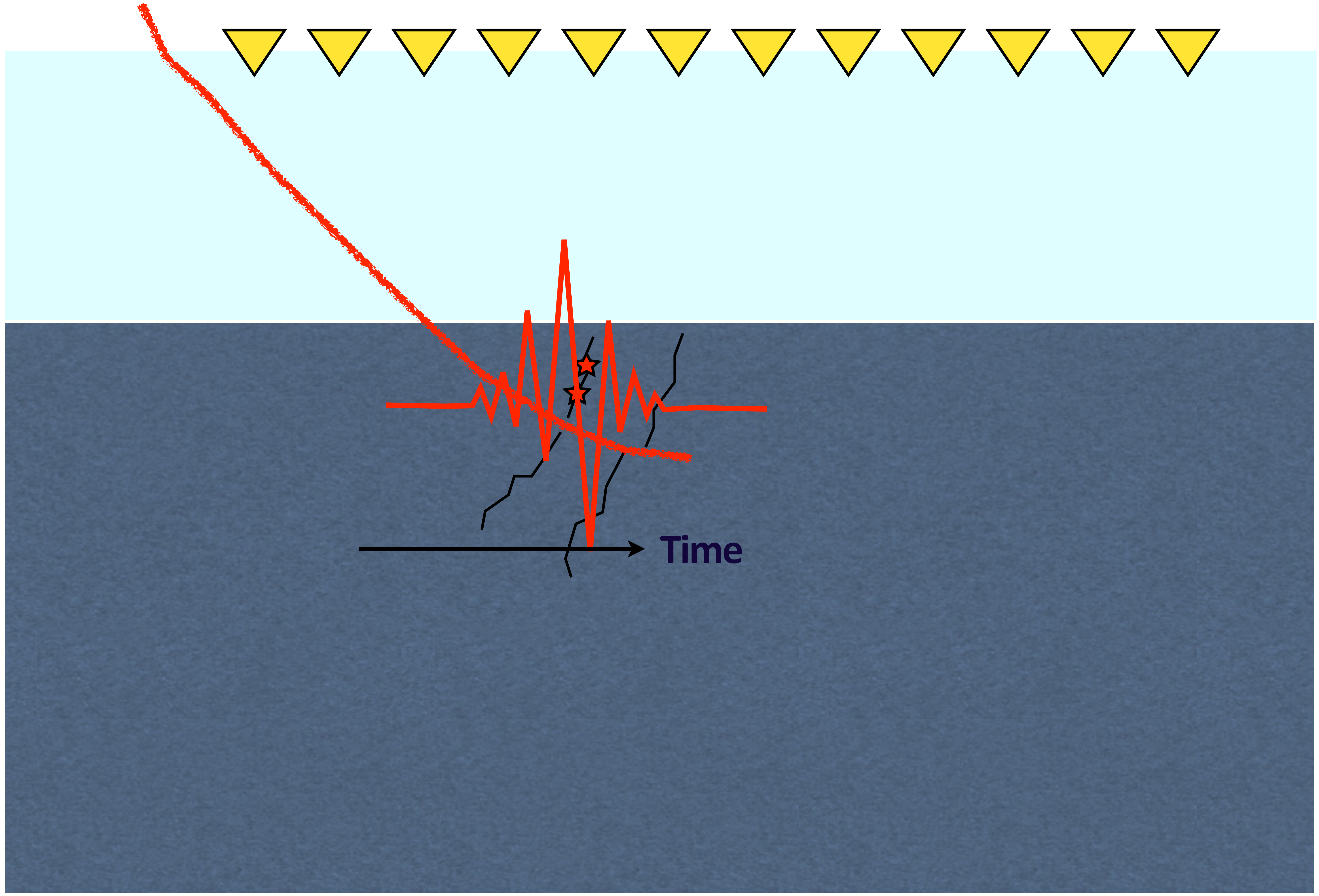
Sparsity-promoting microseismic estimation

- Detection of closely spaced microseismic sources from noisy data
- Estimation of associated source-time function

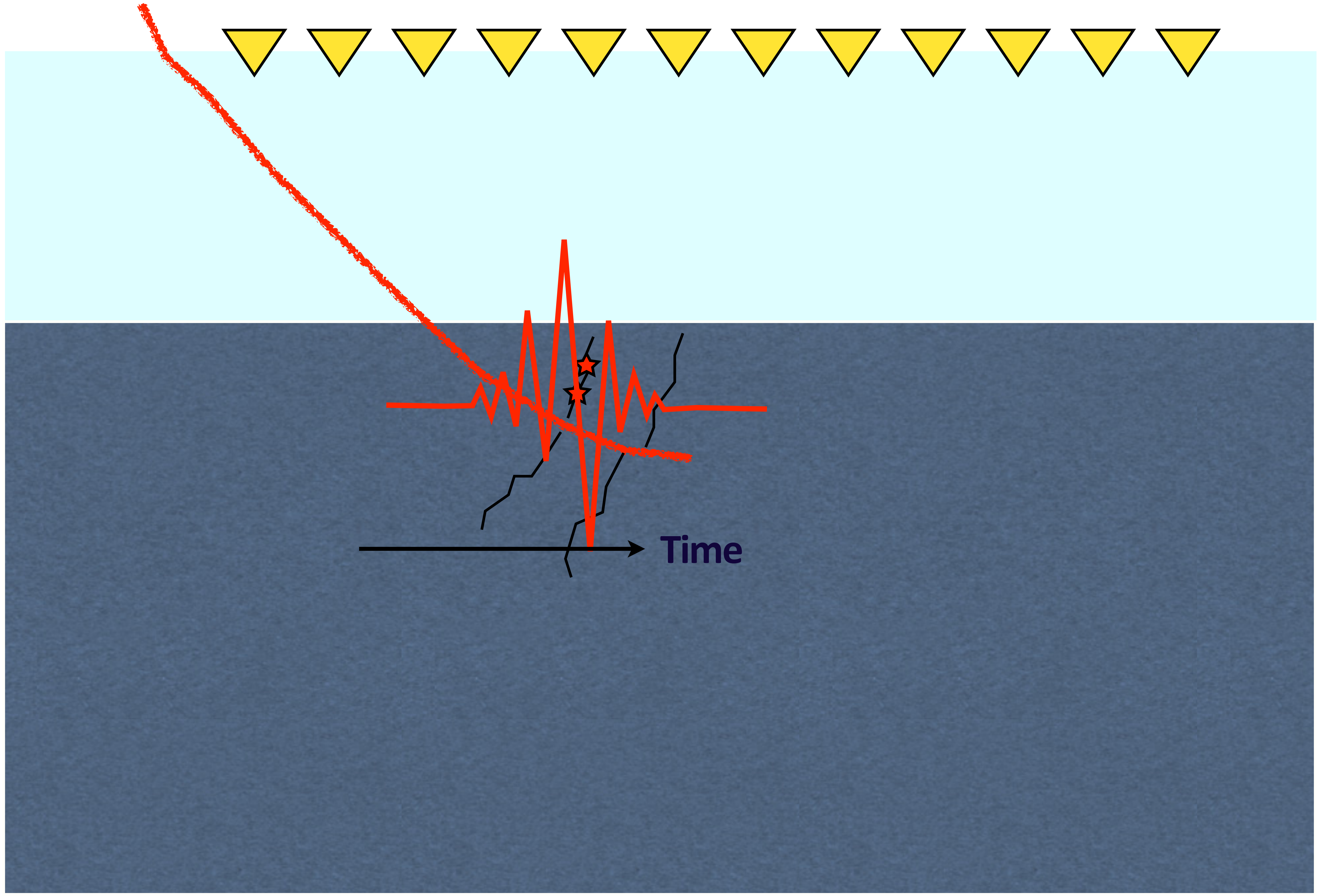
Proposed method w/ sparsity promotion



Proposed method w/ sparsity promotion



Proposed method w/ sparsity promotion



Assumptions

- localized in space
- finite energy along time

[Van Den Berg et al.,'08]

Proposed method w/ sparsity promotion

Source field

Noise level

$$\underset{\mathbf{Q}}{\text{minimize}} \quad \|\mathbf{Q}\|_{2,1}$$

$$\text{subject to} \quad \|\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathbf{d}\|_2 \leq \epsilon$$

Forward modeling operator

Slowness square

Observed data

$\mathbf{Q} \in \mathbb{R}^{n_x \times n_t}$

n_x: number of grid points

n_t: number of time samples

[Van Den Berg et al.,'08]

Proposed method w/ sparsity promotion

Source field **Noise level**

$$\begin{aligned}\underset{\mathbf{Q}}{\text{minimize}} \quad & \|\mathbf{Q}\|_{2,1} \\ \text{subject to} \quad & \|\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathbf{d}\|_2 \leq \epsilon\end{aligned}$$

Forward modeling operator **Slowness square** **Observed data**

$\mathbf{Q} \in \mathbb{R}^{n_x \times n_t}$

n_x: number of grid points

n_t: number of time samples

Similar to classic Basis pursuit denoising (BPDN)

[Lorentz et al.,’14; Herrmann et al.,’15; Sharan et al.,’18]

SLIM

Solving w/ Linearized Bregman

$$\underset{\mathbf{Q}}{\text{minimize}} \;\; \|\mathbf{Q}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{Q}\|_F^2$$

$$\text{subject to } \|\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathbf{d}\|_2 \leq \epsilon$$

*where $\|.\|_F$ is the Frobenius norm

- Recent successful application to seismic imaging problem
- Three-step algorithm simple to implement
- Choice of μ controls the trade off between sparsity and the Frobenius norm
- $\mu \uparrow \infty$ corresponds to solving original BPDN problem

[Lorentz et al.,'14; Combettes et al.,'11]

SLIM

Linearized Bregman algorithm

1. **Data d, slowness square m** //Input
2. **for** $k = 0, 1, \cdots$
3. $\mathbf{V}_k = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}))$ //adjoint solve
4. $\mathbf{Z}_{k+1} = \mathbf{Z}_k - t_k \mathbf{V}_k$ //auxiliary variable update
5. $\mathbf{Q}_{k+1} = \mathrm{Prox}_{\mu\ell_{2,1}}(\mathbf{Z}_{k+1})$ //sparsity promotion
6. **end**
7. $\mathbf{I}(\mathbf{x}) = \sum_t \mid \mathbf{Q}(\mathbf{x}, t) \mid$ //Intensity plot

[Lorentz et al.,'14; Combettes et al.,'11]

SLIM

Linearized Bregman algorithm

1. $\textbf{Data d, slowness square m}$ //Input
2. $\textbf{for } k = 0, 1, \cdots$
3. $\quad \mathbf{V}_k = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}))$ //adjoint solve
4. $\quad \mathbf{Z}_{k+1} = \mathbf{Z}_k - t_k \mathbf{V}_k$ //auxiliary variable update
5. $\quad \mathbf{Q}_{k+1} = \mathrm{Prox}_{\mu \ell_{2,1}}(\mathbf{Z}_{k+1})$ //sparsity promotion
6. $\textbf{end}$
7. $\mathbf{I}(\mathbf{x}) = \sum_t \mid \mathbf{Q}(\mathbf{x}, t) \mid$ //Intensity plot

* $\Pi_\epsilon(\mathbf{x}) = \max\{0, 1 - \frac{\epsilon}{\|\mathbf{x}\|}\}.(\mathbf{x})$ the projection on to ℓ_2 norm ball

[Lorentz et al.,'14; Combettes et al.,'11]

SLIM

Linearized Bregman algorithm

1. **Data d, slowness square m** //Input
2. **for** $k = 0, 1, \cdots$
3. $\mathbf{V}_k = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}))$ //adjoint solve
4. $\mathbf{Z}_{k+1} = \mathbf{Z}_k - t_k \mathbf{V}_k$ //auxiliary variable update
5. $\mathbf{Q}_{k+1} = \mathrm{Prox}_{\mu\ell_{2,1}}(\mathbf{Z}_{k+1})$ //sparsity promotion
6. **end**
7. $\mathbf{I}(\mathbf{x}) = \sum_t \mid \mathbf{Q}(\mathbf{x}, t) \mid$ //Intensity plot

* $\Pi_\epsilon(\mathbf{x}) = \max\{0, 1 - \frac{\epsilon}{\|\mathbf{x}\|}\}.(\mathbf{x})$ the projection on to ℓ_2 norm ball

*where $t_k = \frac{\|\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}\|^2}{\|\mathcal{F}^\top[\mathbf{m}](\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d})\|^2}$ is the dynamic step length

[Lorentz et al.,’14; Combettes et al.,’11]

SLIM

Linearized Bregman algorithm

1. **Data d, slowness square m** //Input
2. **for** $k = 0, 1, \cdots$
3. $\mathbf{V}_k = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}))$ //adjoint solve
4. $\mathbf{Z}_{k+1} = \mathbf{Z}_k - t_k \mathbf{V}_k$ //auxiliary variable update
5. $\mathbf{Q}_{k+1} = \text{Prox}_{\mu \ell_{2,1}}(\mathbf{Z}_{k+1})$ //sparsity promotion
6. **end**
7. $\mathbf{I}(\mathbf{x}) = \sum_t \mid \mathbf{Q}(\mathbf{x}, t) \mid$ //Intensity plot

* $\Pi_\epsilon(\mathbf{x}) = \max\{0, 1 - \frac{\epsilon}{\|\mathbf{x}\|}\}.(\mathbf{x})$ the projection on to ℓ_2 norm ball

*where $t_k = \frac{\|\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}\|^2}{\|\mathcal{F}^\top[\mathbf{m}](\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d})\|^2}$ is the dynamic step length

* $\text{Prox}_{\mu\ell_{2,1}}(\mathbf{C}) := \arg\min_{\mathbf{B}} \|\mathbf{B}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{C} - \mathbf{B}\|_F^2$ is the proximal mapping of the $\ell_{2,1}$ norm

[Lorentz et al.,'14; Combettes et al.,'11]

Linearized Bregman algorithm

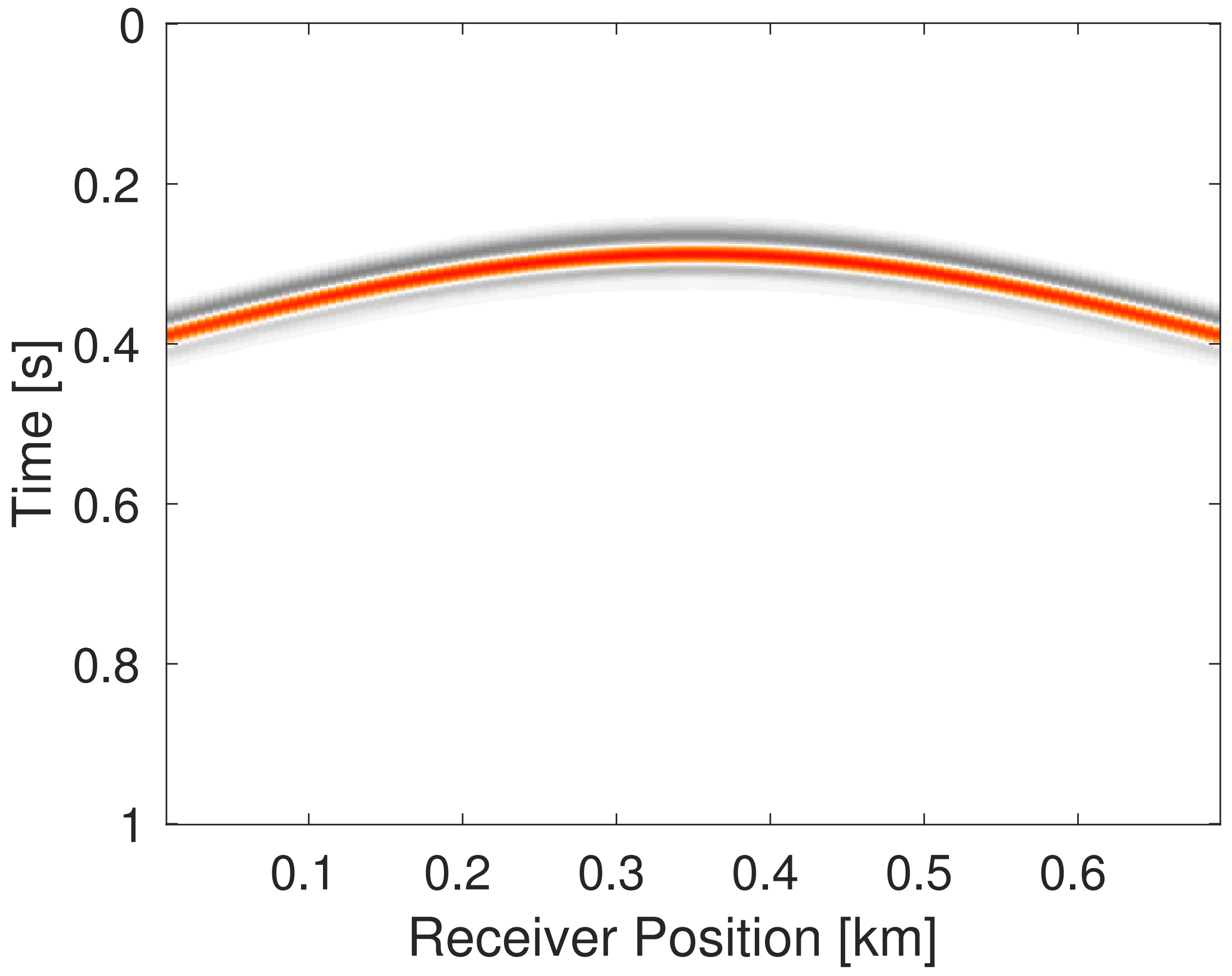
1. **Data d, slowness square m** //Input
2. **for** $k = 0, 1, \cdots$
3. $\mathbf{V}_k = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}))$ //adjoint solve
4. $\mathbf{Z}_{k+1} = \mathbf{Z}_k - t_k \mathbf{V}_k$ //auxiliary variable update
5. $\mathbf{Q}_{k+1} = \mathrm{Prox}_{\mu\ell_{2,1}}(\mathbf{Z}_{k+1})$ //sparsity promotion
6. **end**
7. $\mathbf{I}(\mathbf{x}) = \sum_t \mid \mathbf{Q}(\mathbf{x}, t) \mid$ //Intensity plot

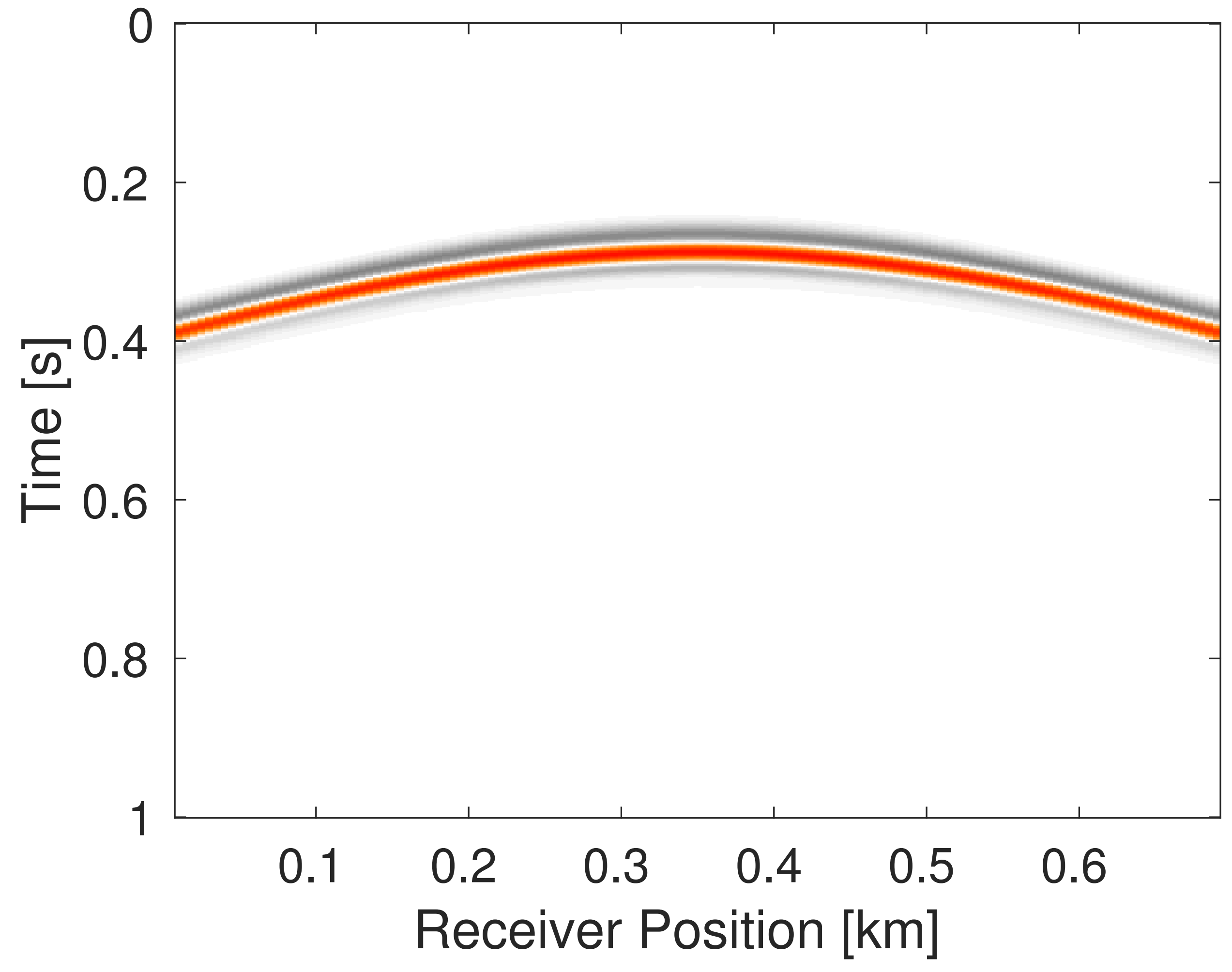
* $\Pi_\epsilon(\mathbf{x}) = \max\{0, 1 - \frac{\epsilon}{\|\mathbf{x}\|}\}.(\mathbf{x})$ the projection on to ℓ_2 norm ball

*where $t_k = \frac{\|\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d}\|^2}{\|\mathcal{F}^\top[\mathbf{m}](\mathcal{F}[\mathbf{m}](\mathbf{Q}_k) - \mathbf{d})\|^2}$ is the dynamic step length

* $\mathrm{Prox}_{\mu\ell_{2,1}}(\mathbf{C}) := \arg\min_{\mathbf{B}} \|\mathbf{B}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{C} - \mathbf{B}\|_F^2$ is the proximal mapping of the $\ell_{2,1}$ norm

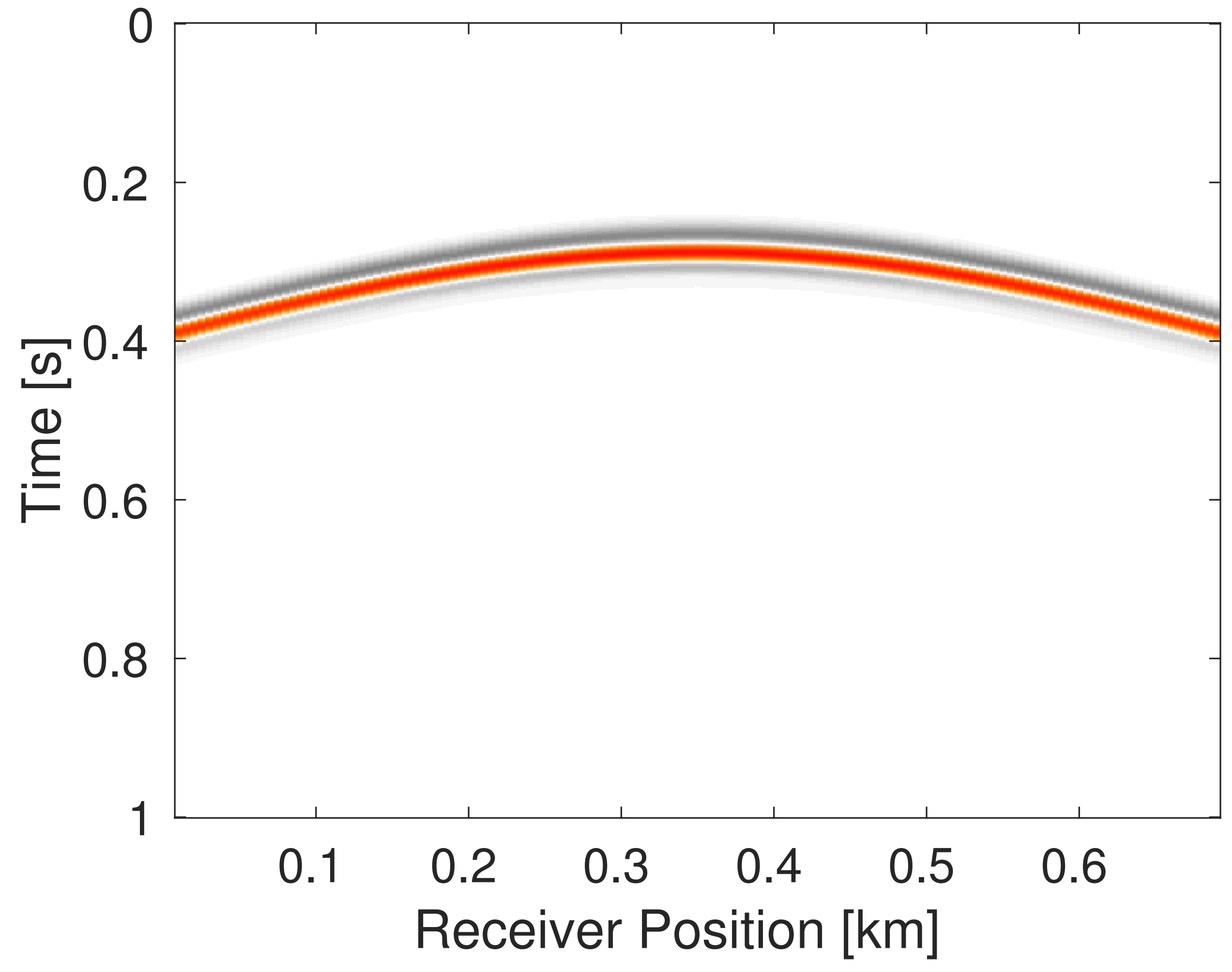
- **Source location**: estimated as outlier in intensity plot
- **Source-time function**: temporal variation of wavefield at estimated source location


0
0.2
0.4
0.6
0.8
1
Time [s]
0.1
0.2
0.3
0.4
0.5
0.6
Receiver Position [km]



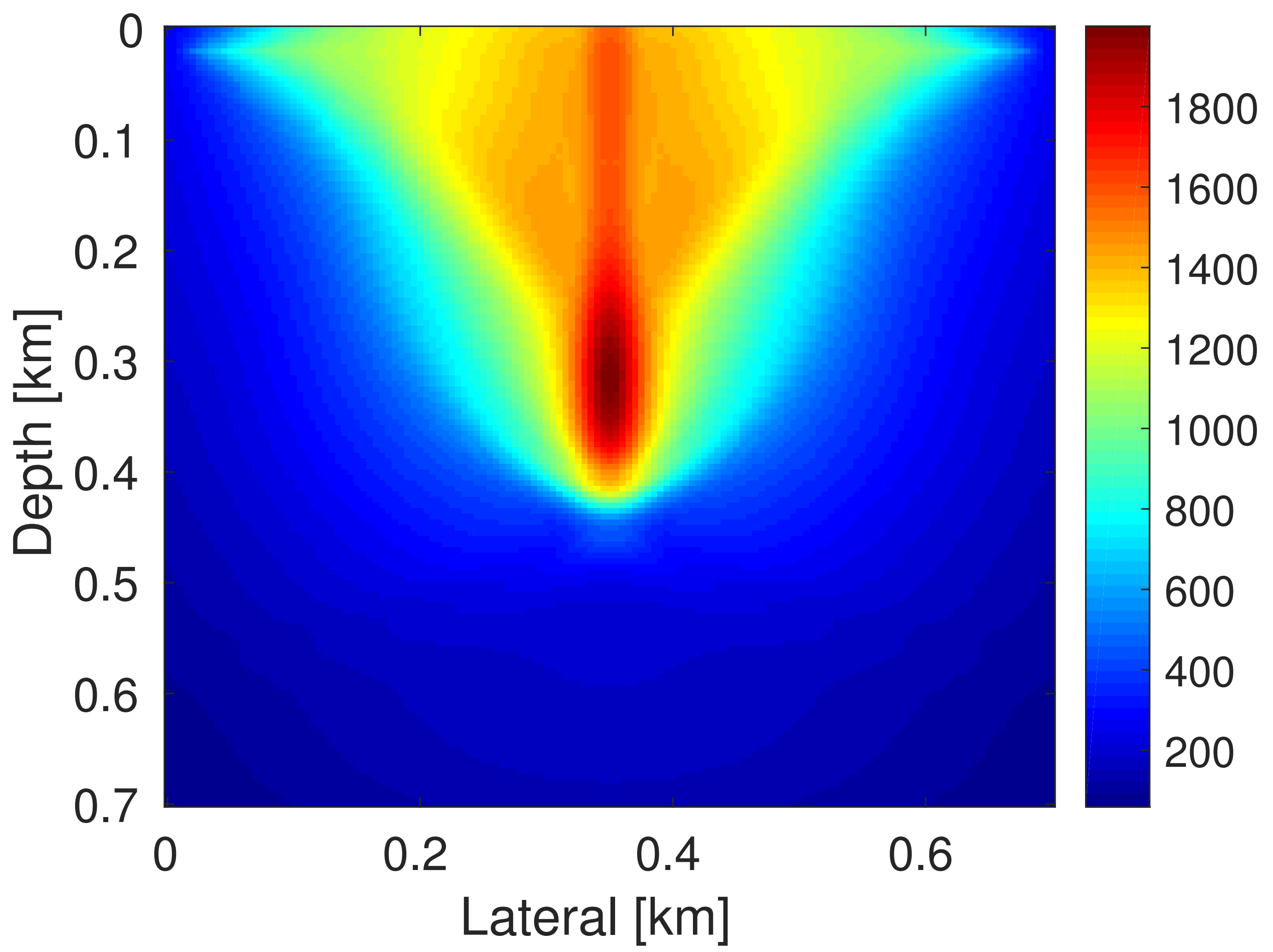
$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

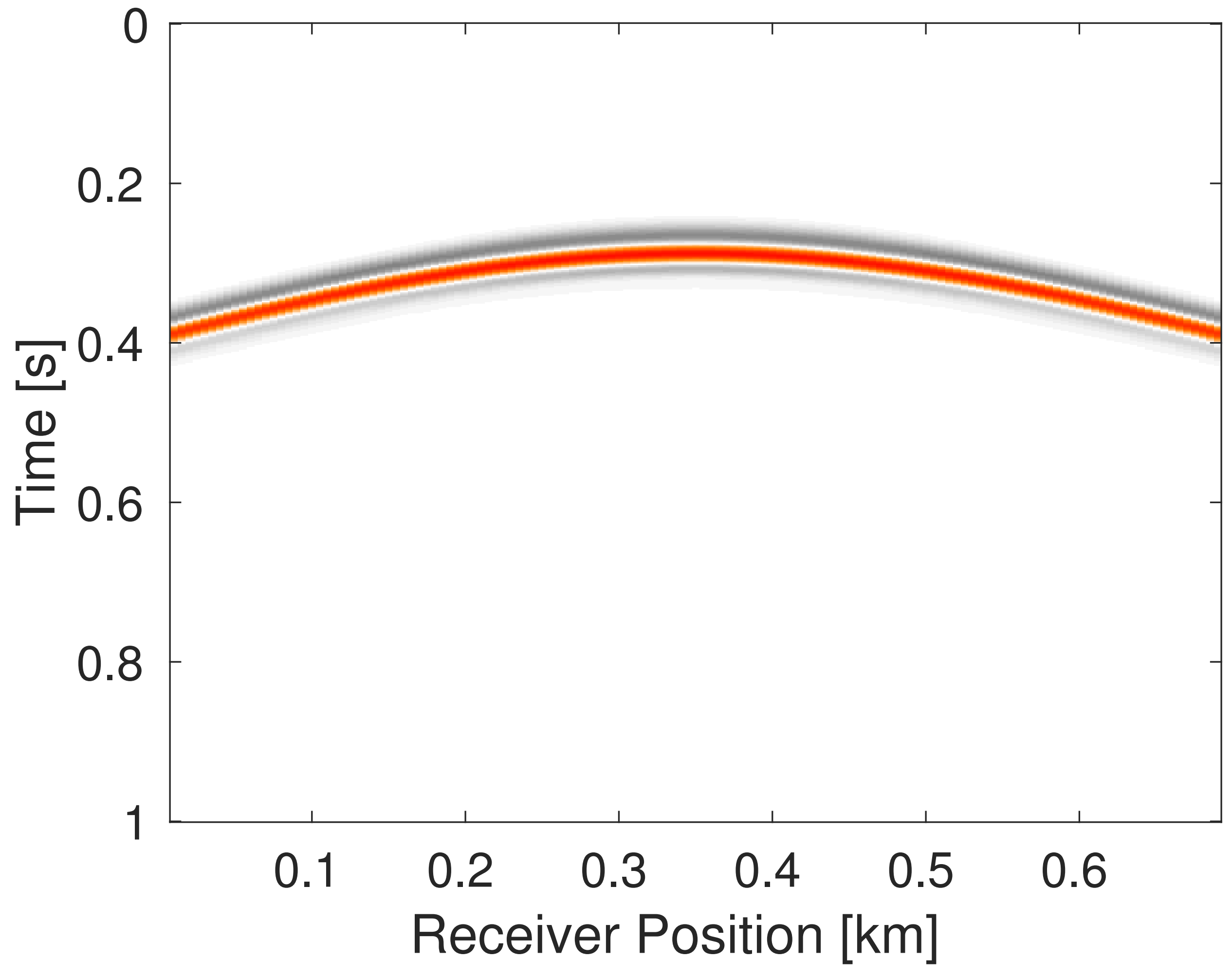
Adjoint solve



$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

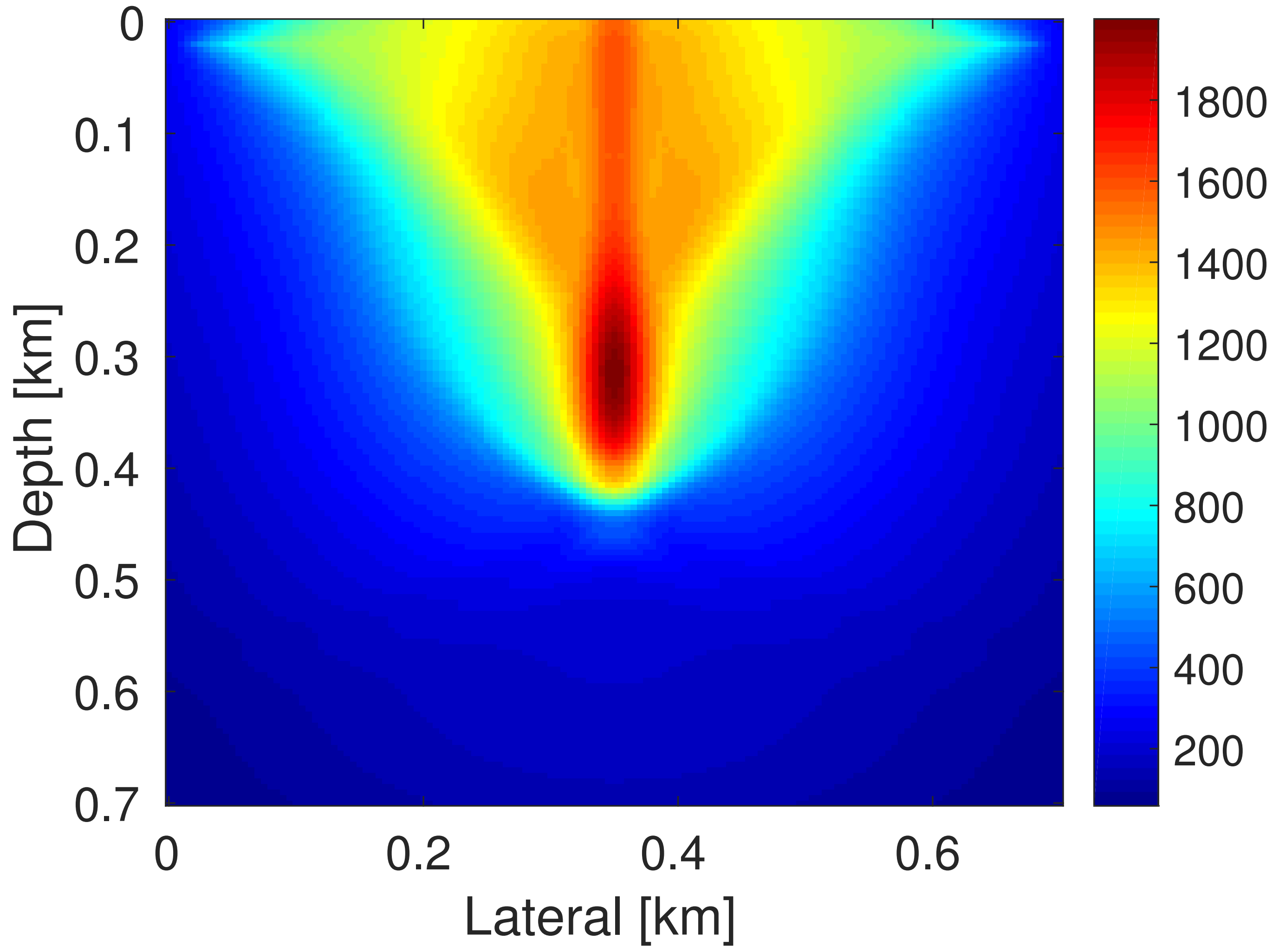
Adjoint solve





$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

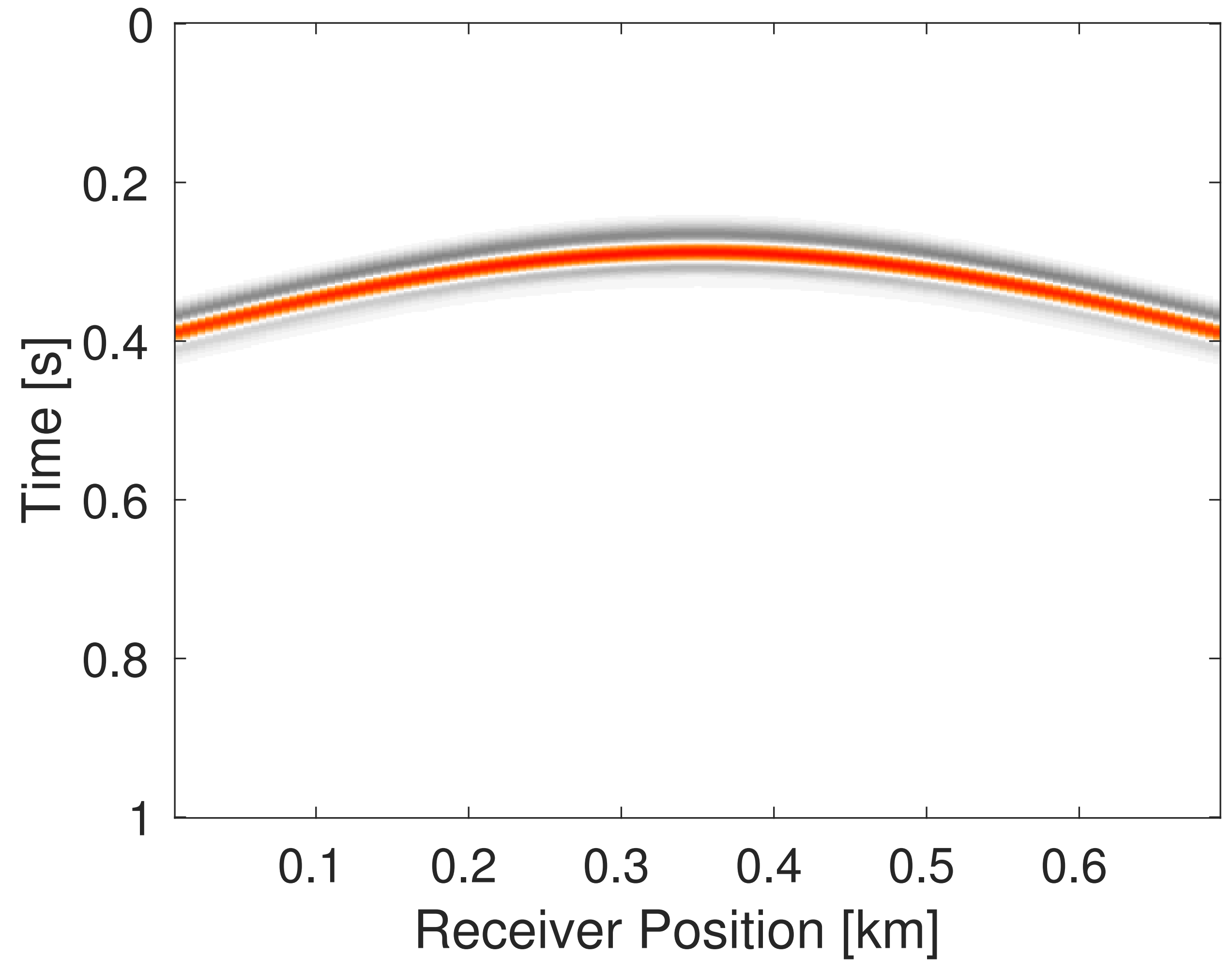
Adjoint solve



Auxiliary variable update

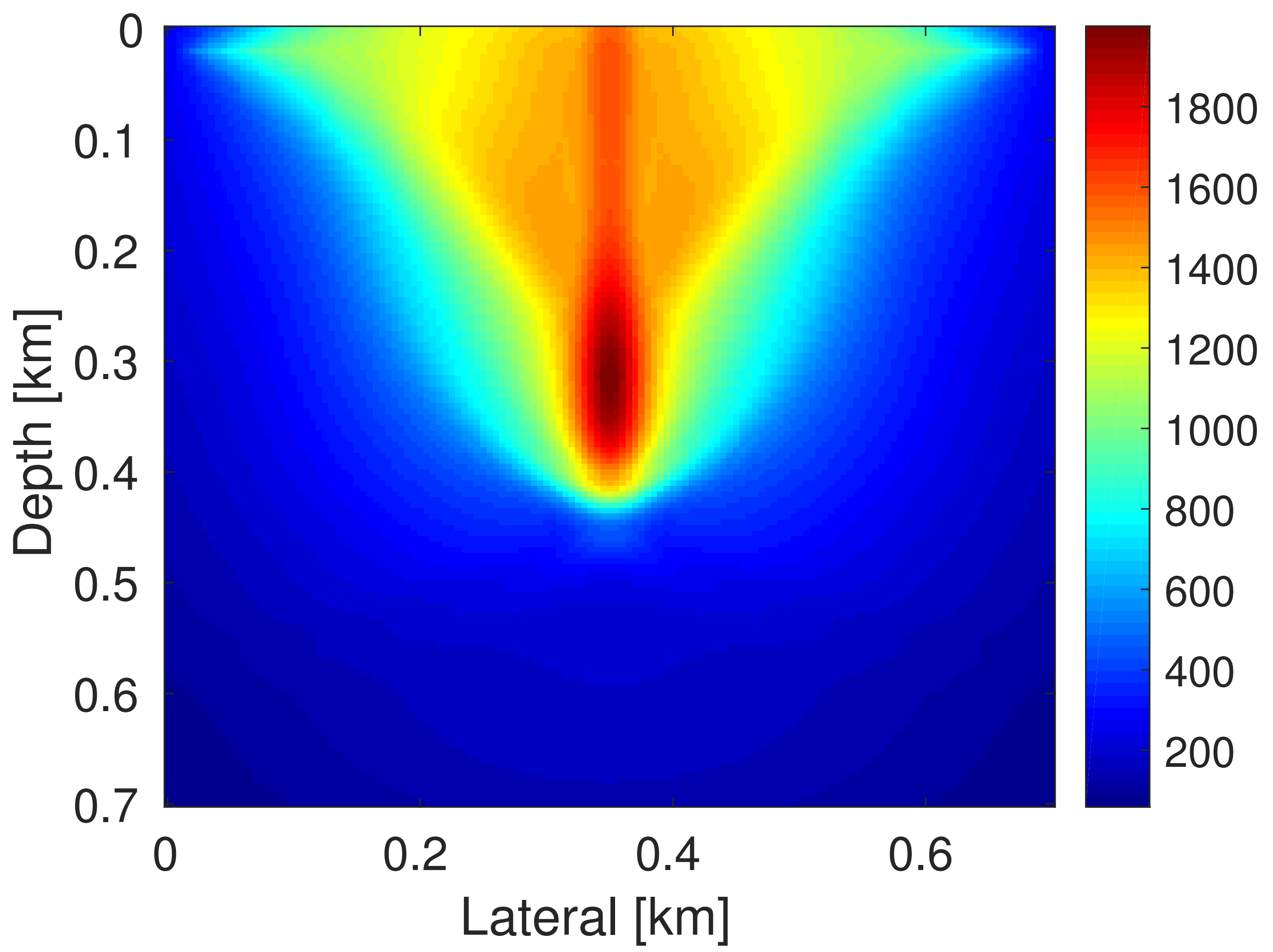
$$\mathbf{Z}_1 = \mathbf{Z}_0 - t_1 \mathbf{V}_1$$

SLIM



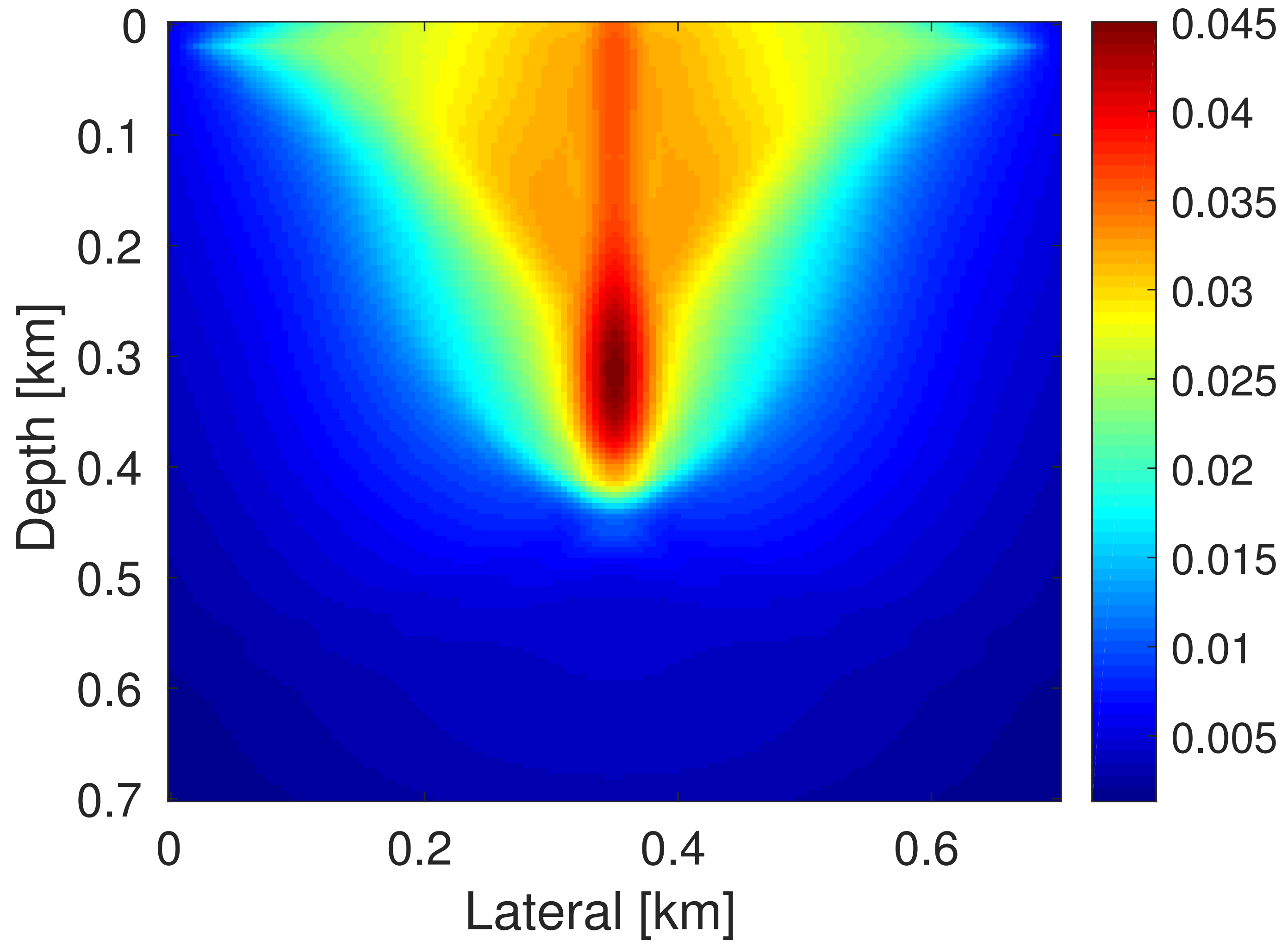
$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

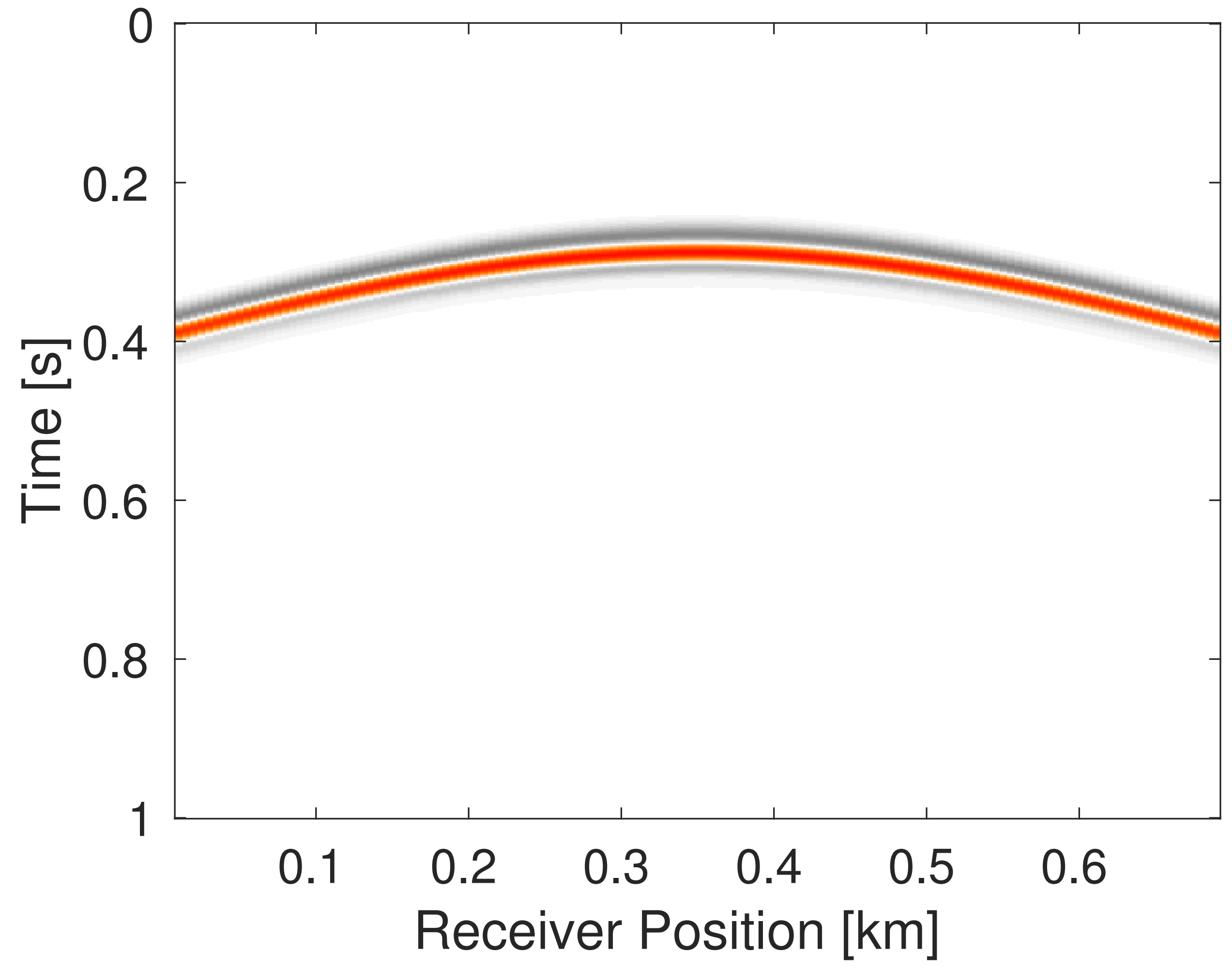
Adjoint solve



Auxiliary variable update

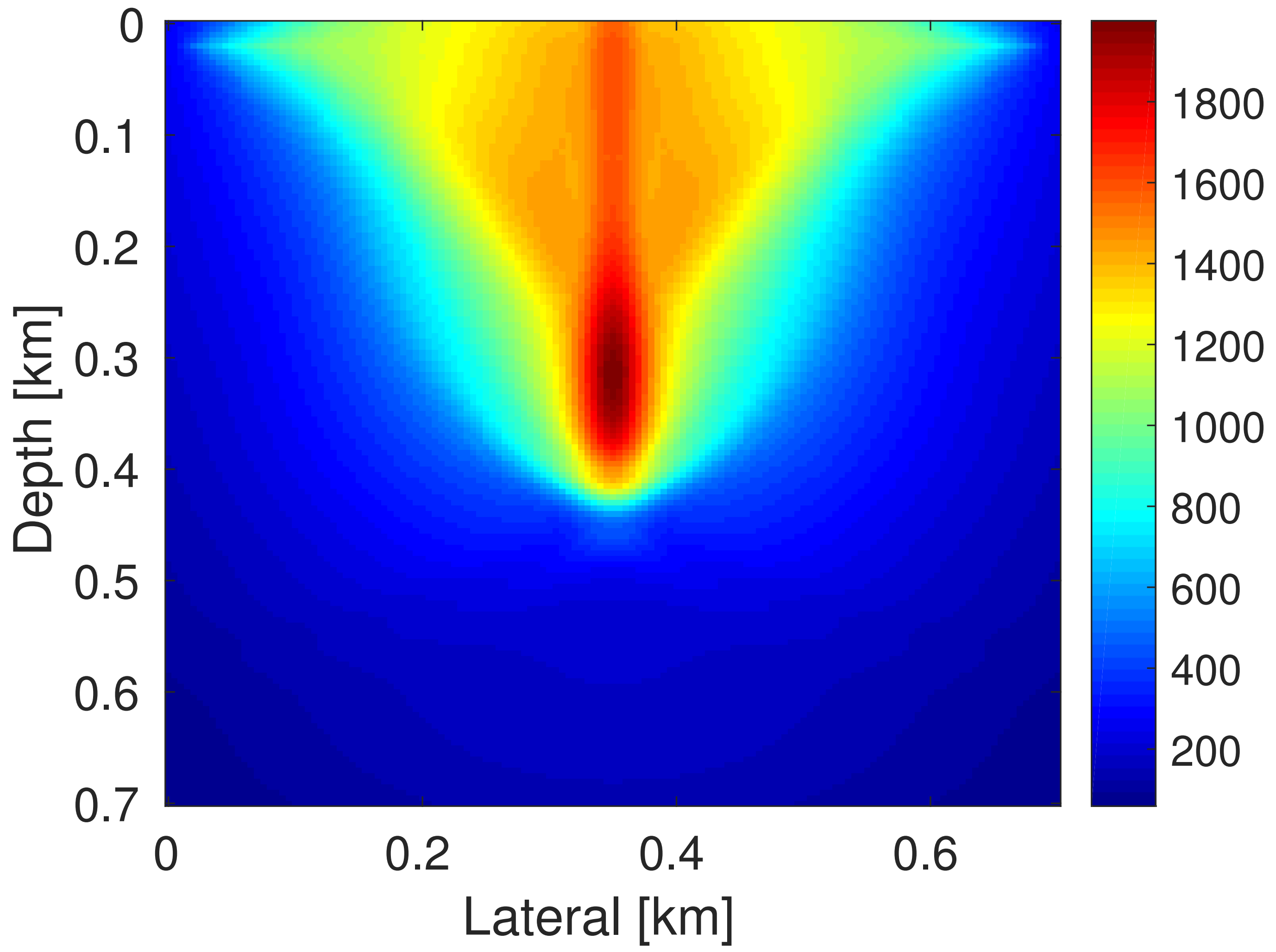
$$\mathbf{Z}_1 = \mathbf{Z}_0 - t_1\mathbf{V}_1$$





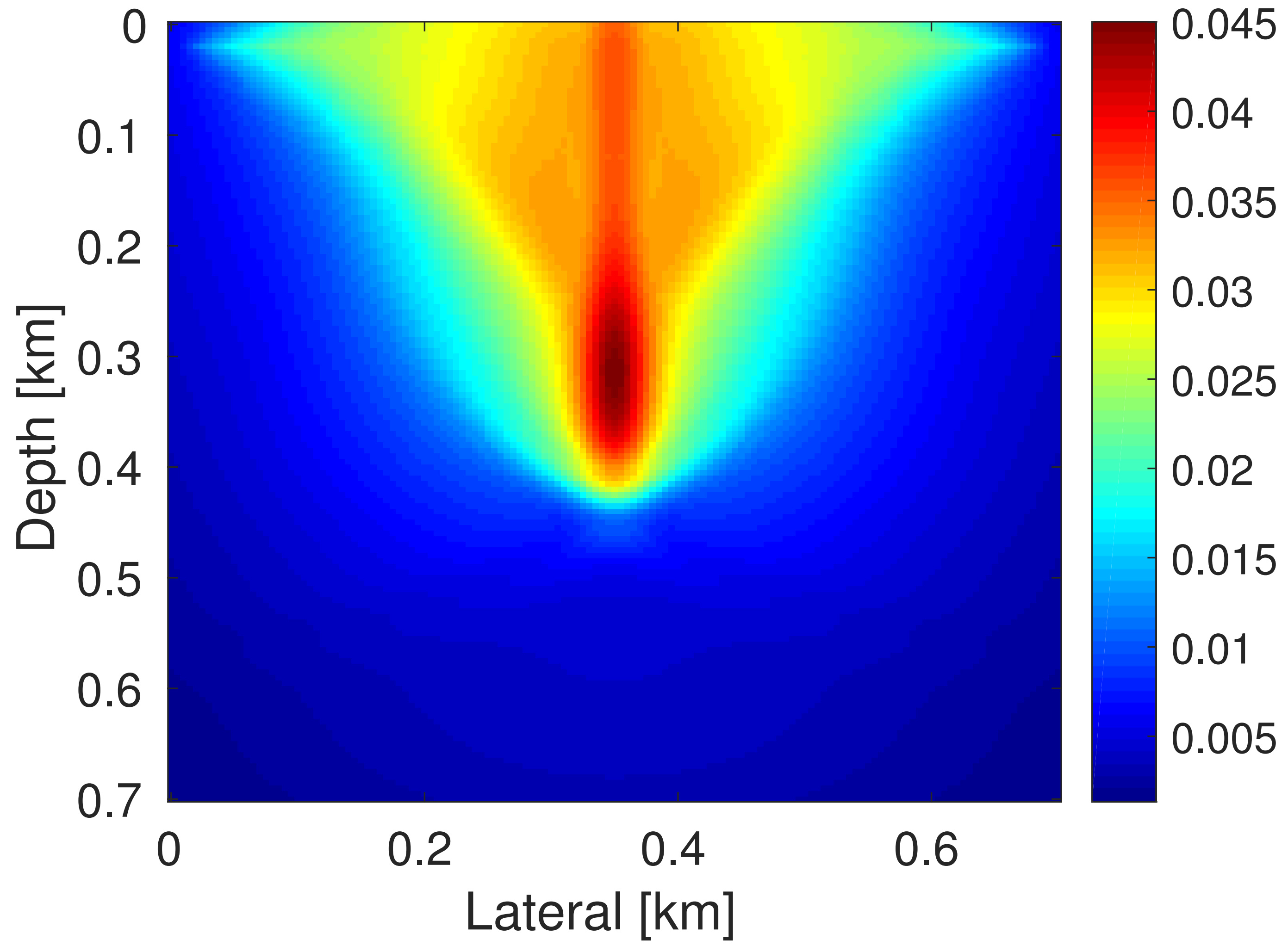
$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

Adjoint solve



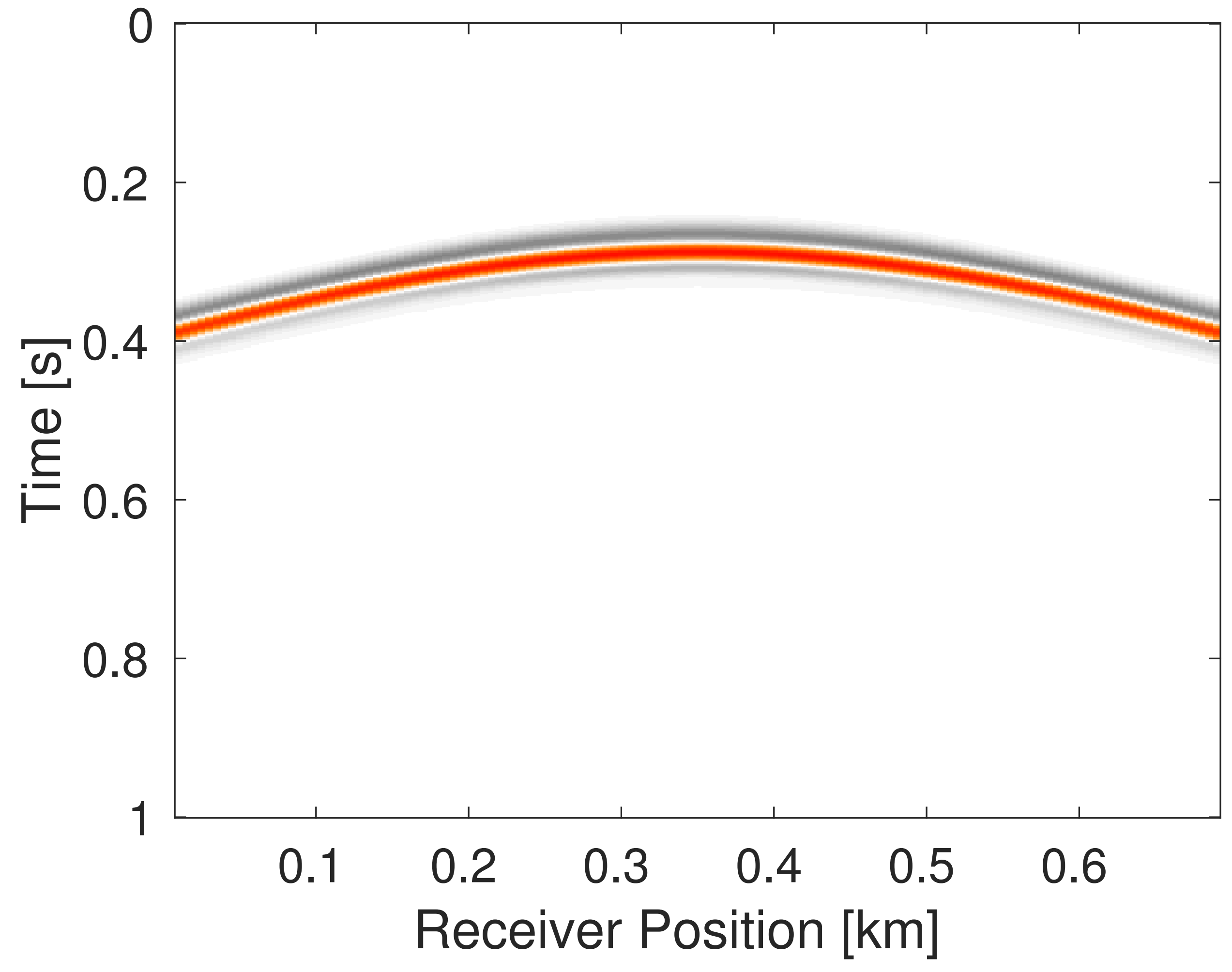
Auxiliary variable update

$$\mathbf{Z}_1 = \mathbf{Z}_0 - t_1 \mathbf{V}_1$$



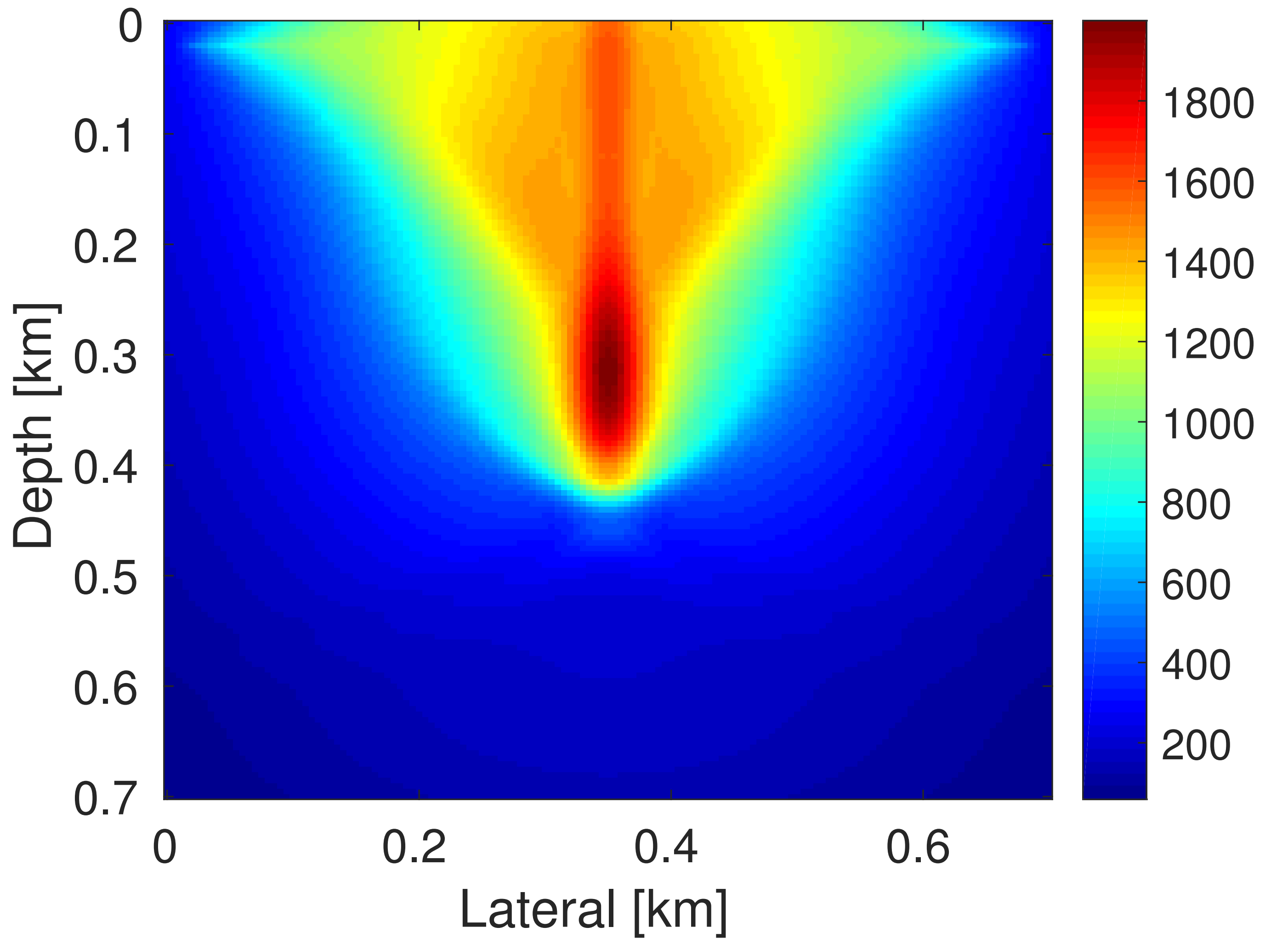
$$\mathbf{Q}_1 = \mathrm{Prox}_{\mu \ell_{2,1}}(\mathbf{Z}_1)$$

Sparsity promotion



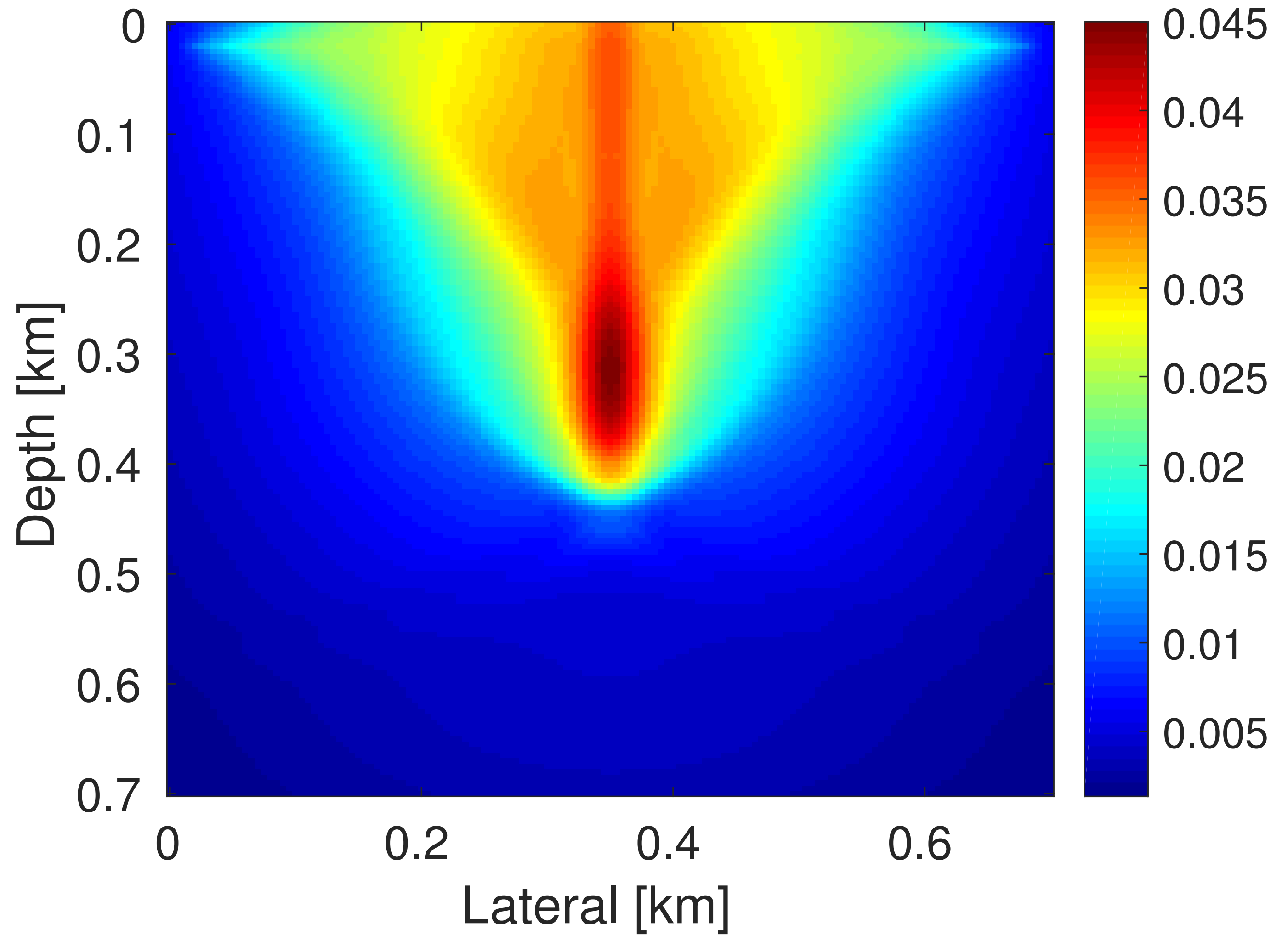
$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

Adjoint solve



Auxiliary variable update

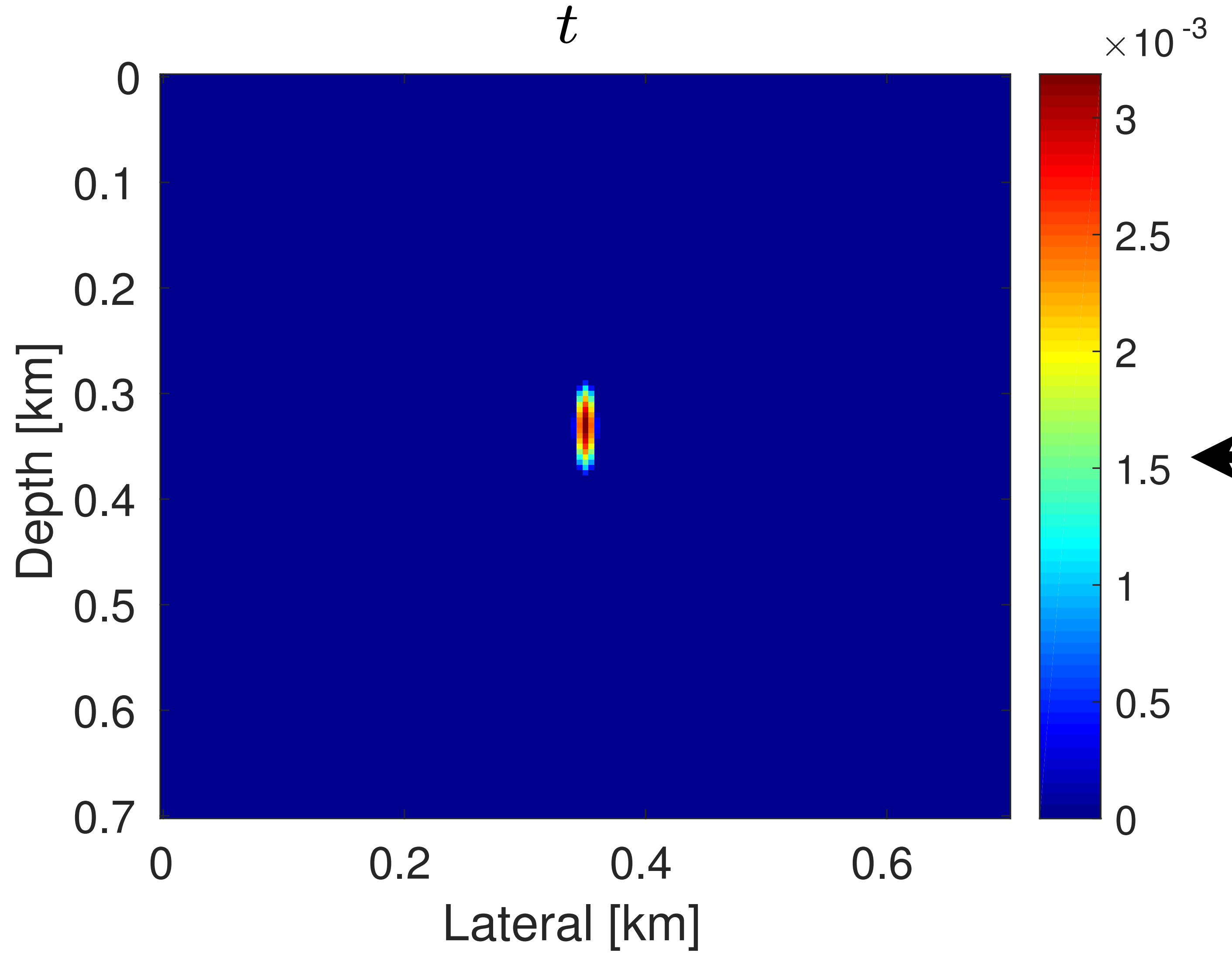
$$\mathbf{Z}_1 = \mathbf{Z}_0 - t_1\mathbf{V}_1$$

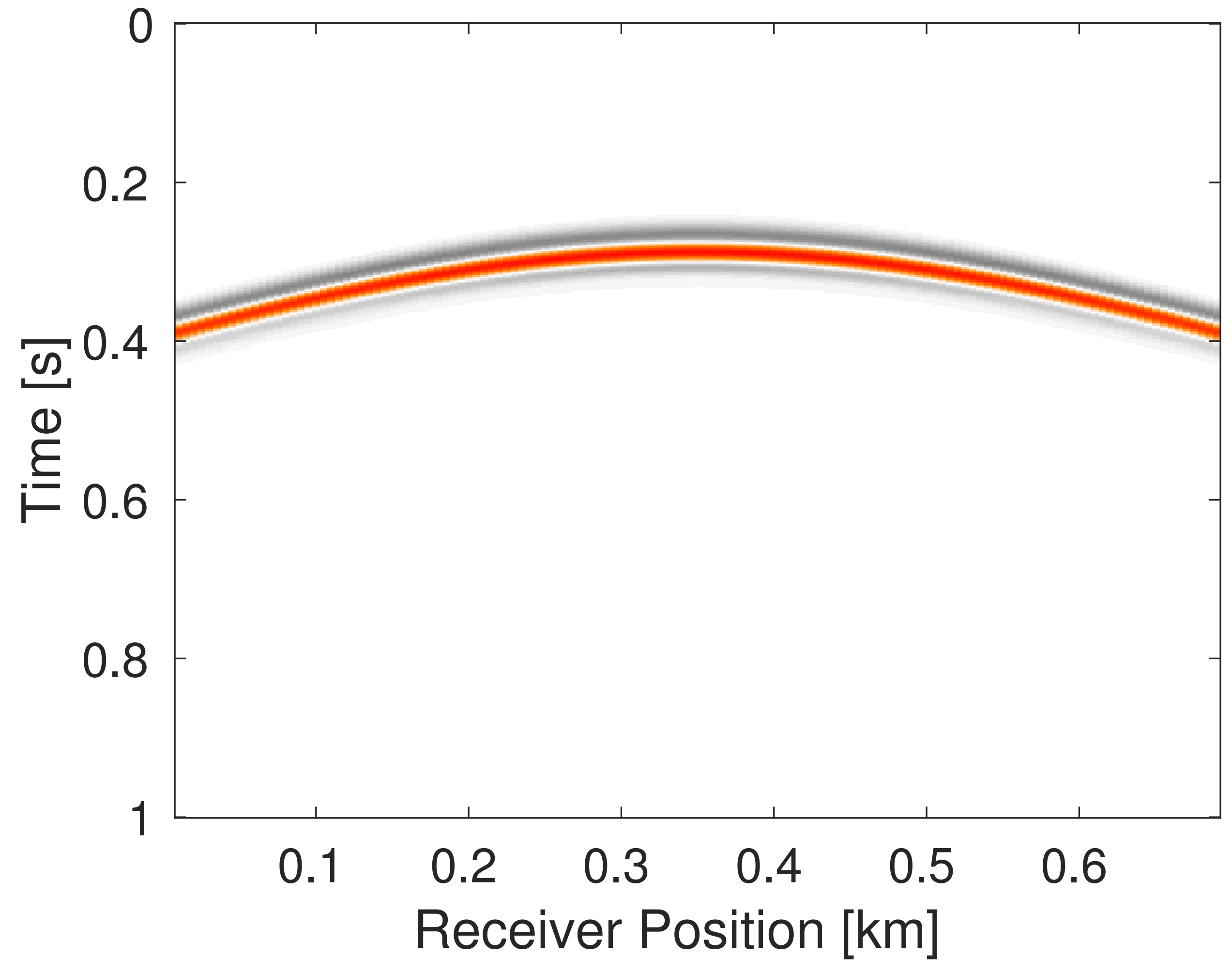


$$\mathbf{Q}_1 = \mathrm{Prox}_{\mu\ell_{2,1}}(\mathbf{Z}_1)$$

Sparsity promotion

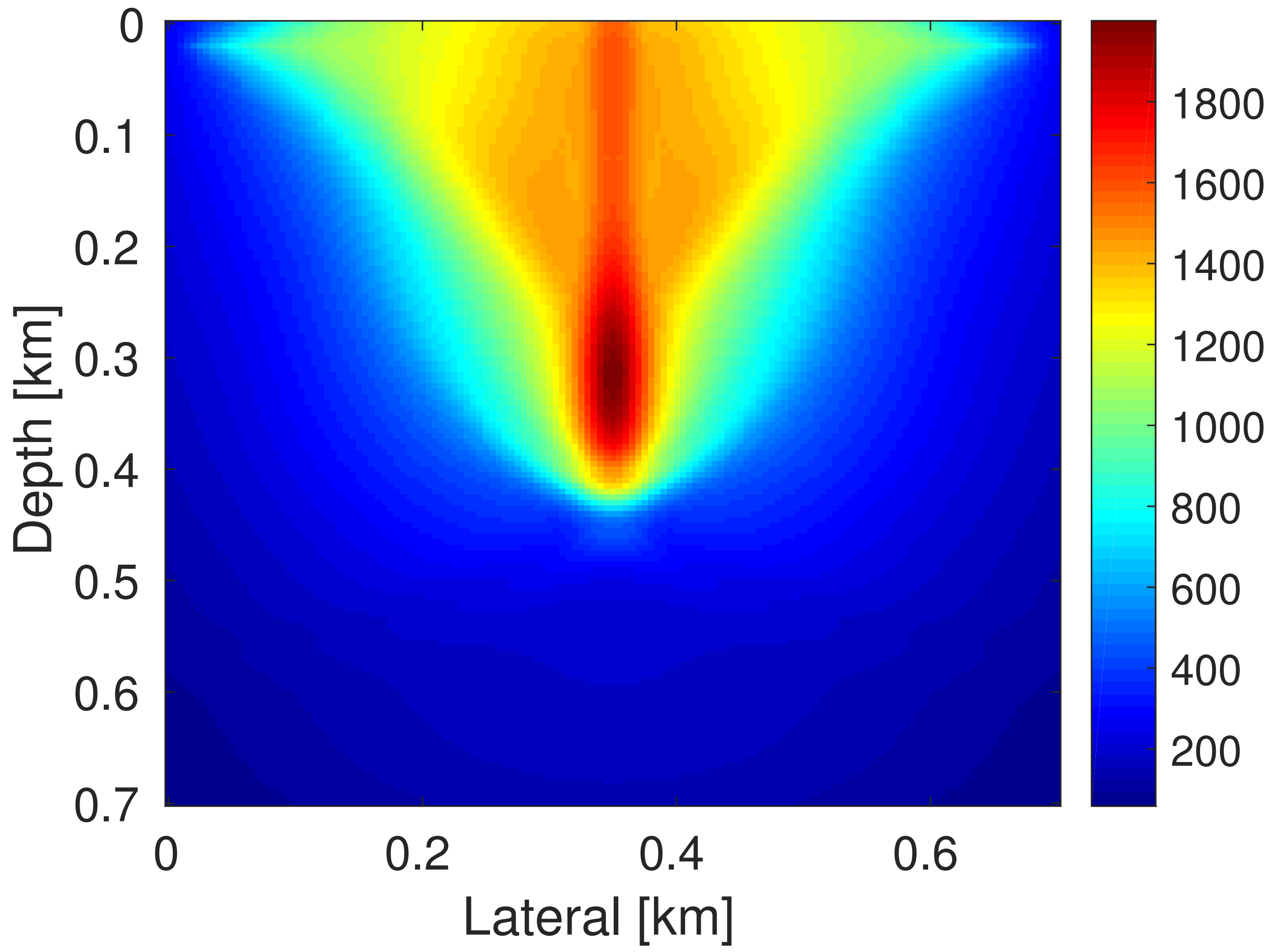
$$\mathbf{I}(\mathbf{x}) = \sum_t \mid \mathbf{Q}_1(\mathbf{x}, t) \mid$$





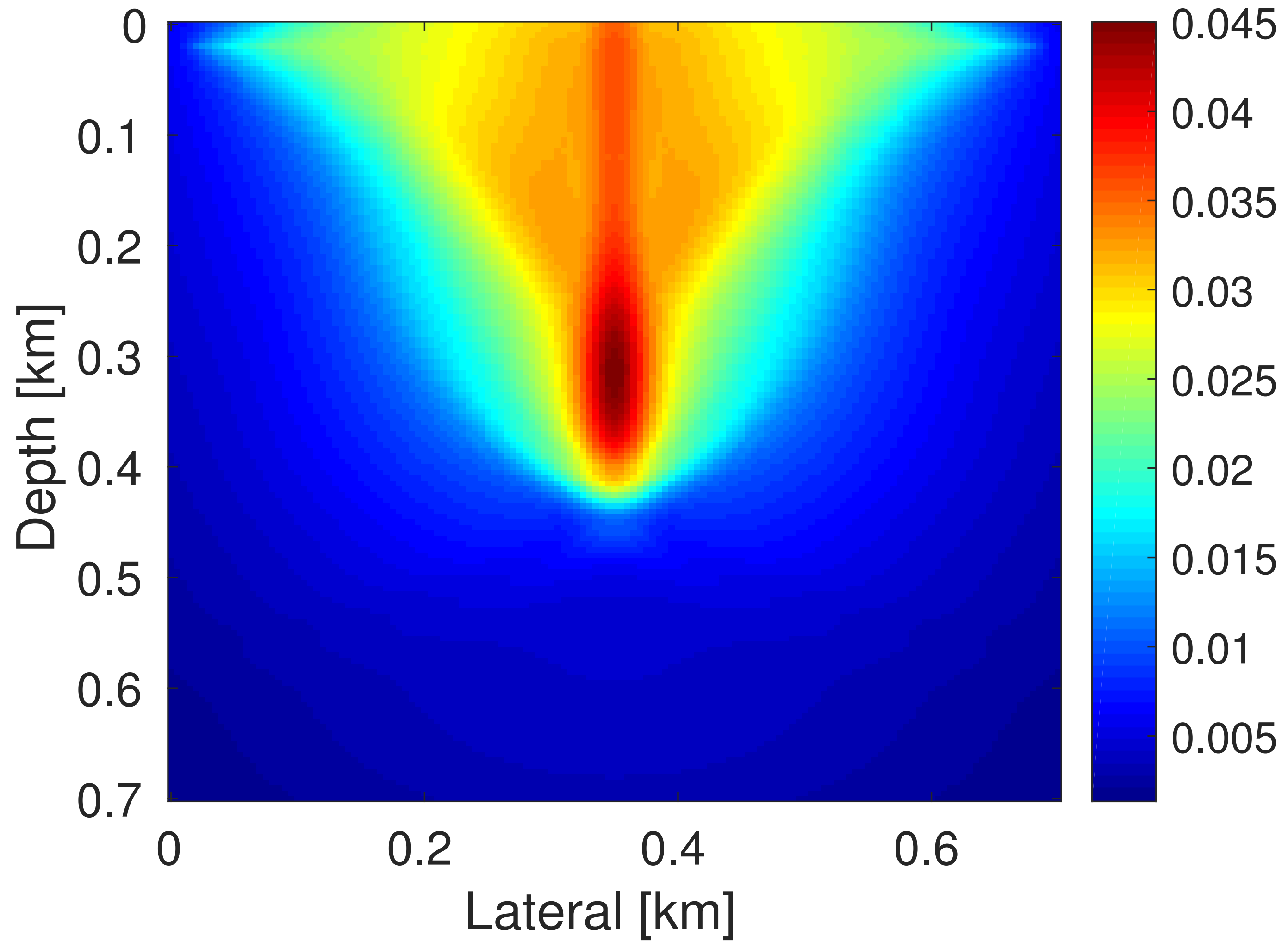
$$\mathbf{V}_1 = \mathcal{F}^\top[\mathbf{m}](\Pi_\epsilon(\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d}))$$

Adjoint solve



Auxiliary variable update

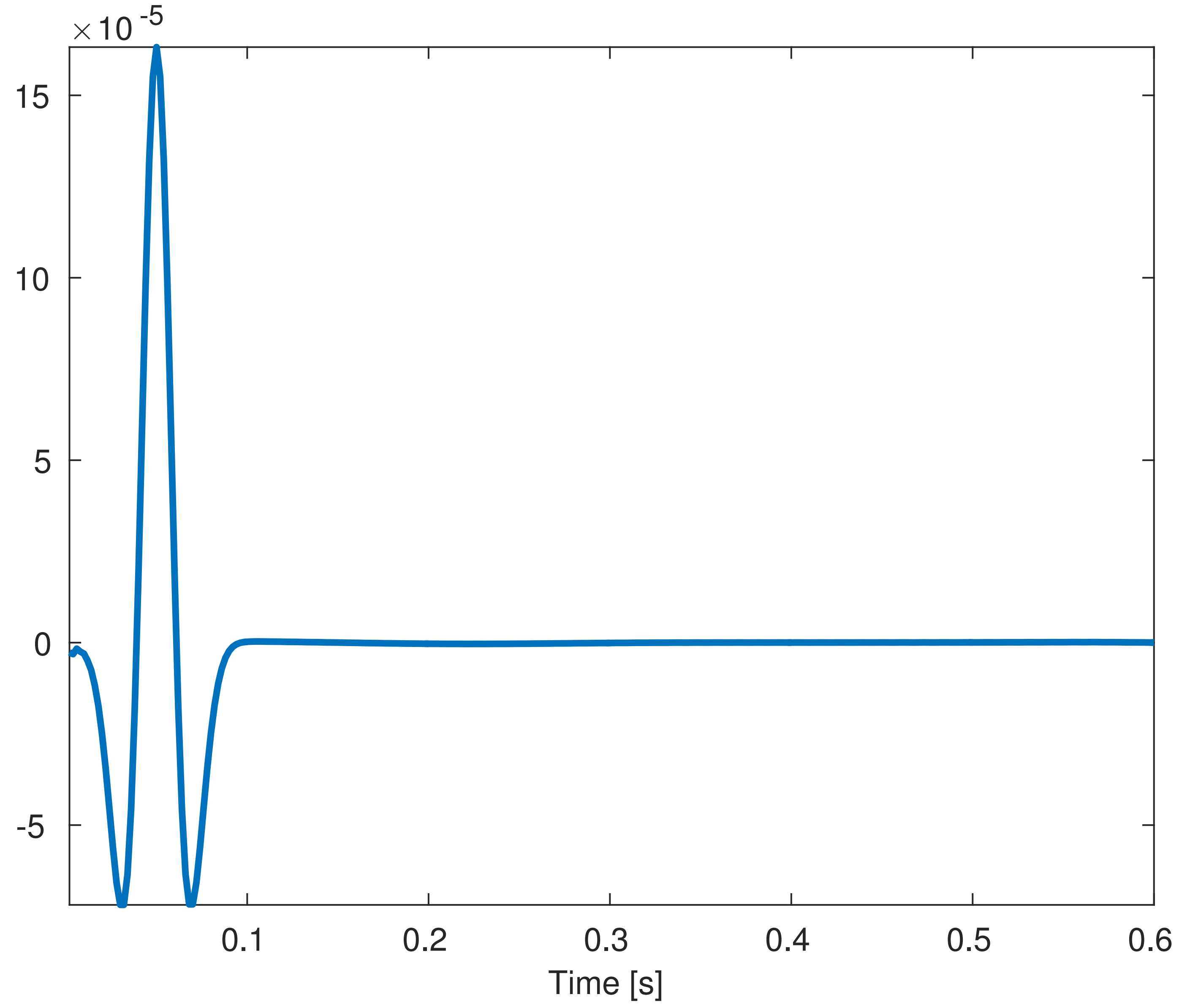
$$\mathbf{Z}_1 = \mathbf{Z}_0 - t_1\mathbf{V}_1$$



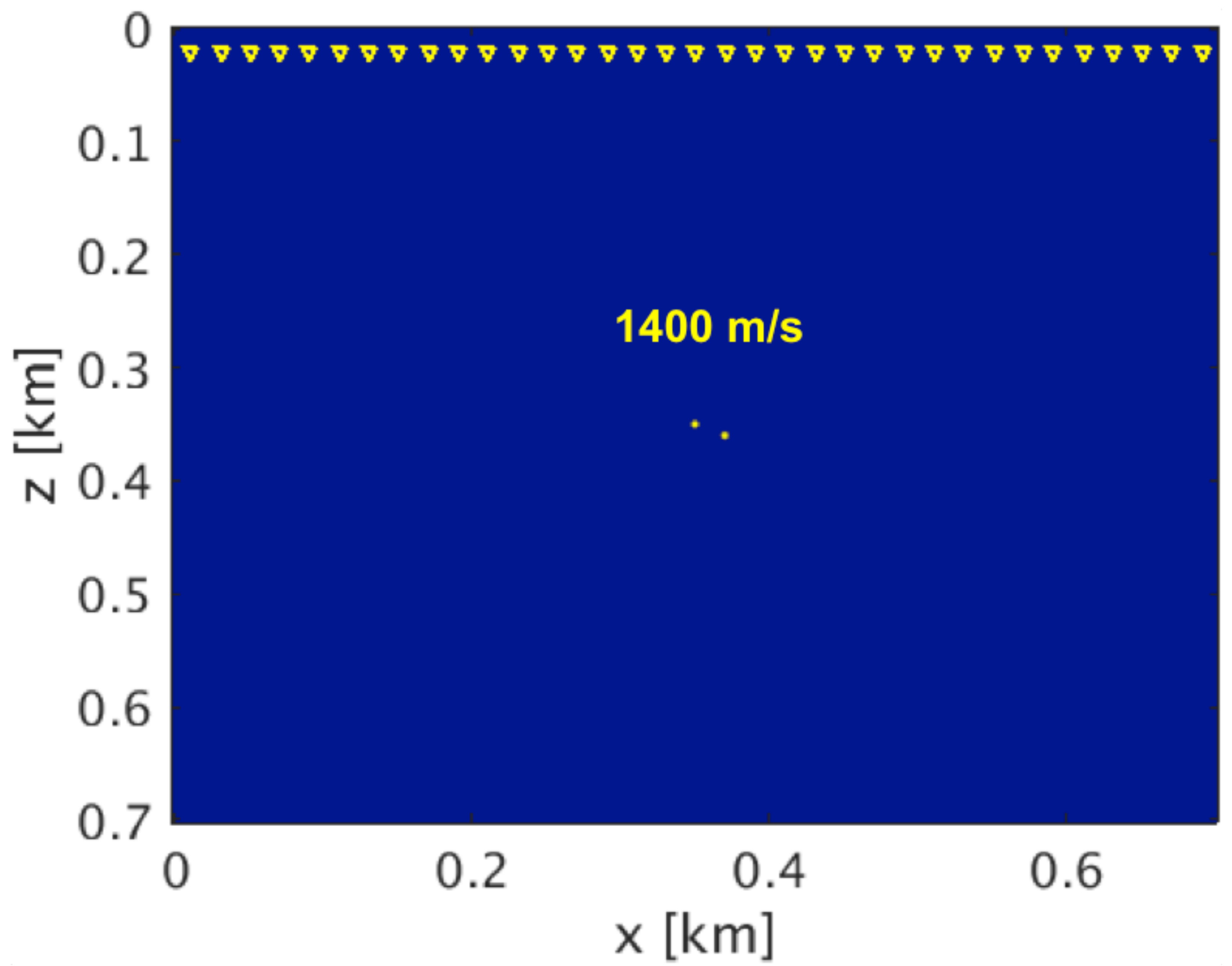
$$\mathbf{Q}_1 = \mathrm{Prox}_{\mu\ell_{2,1}}(\mathbf{Z}_1)$$

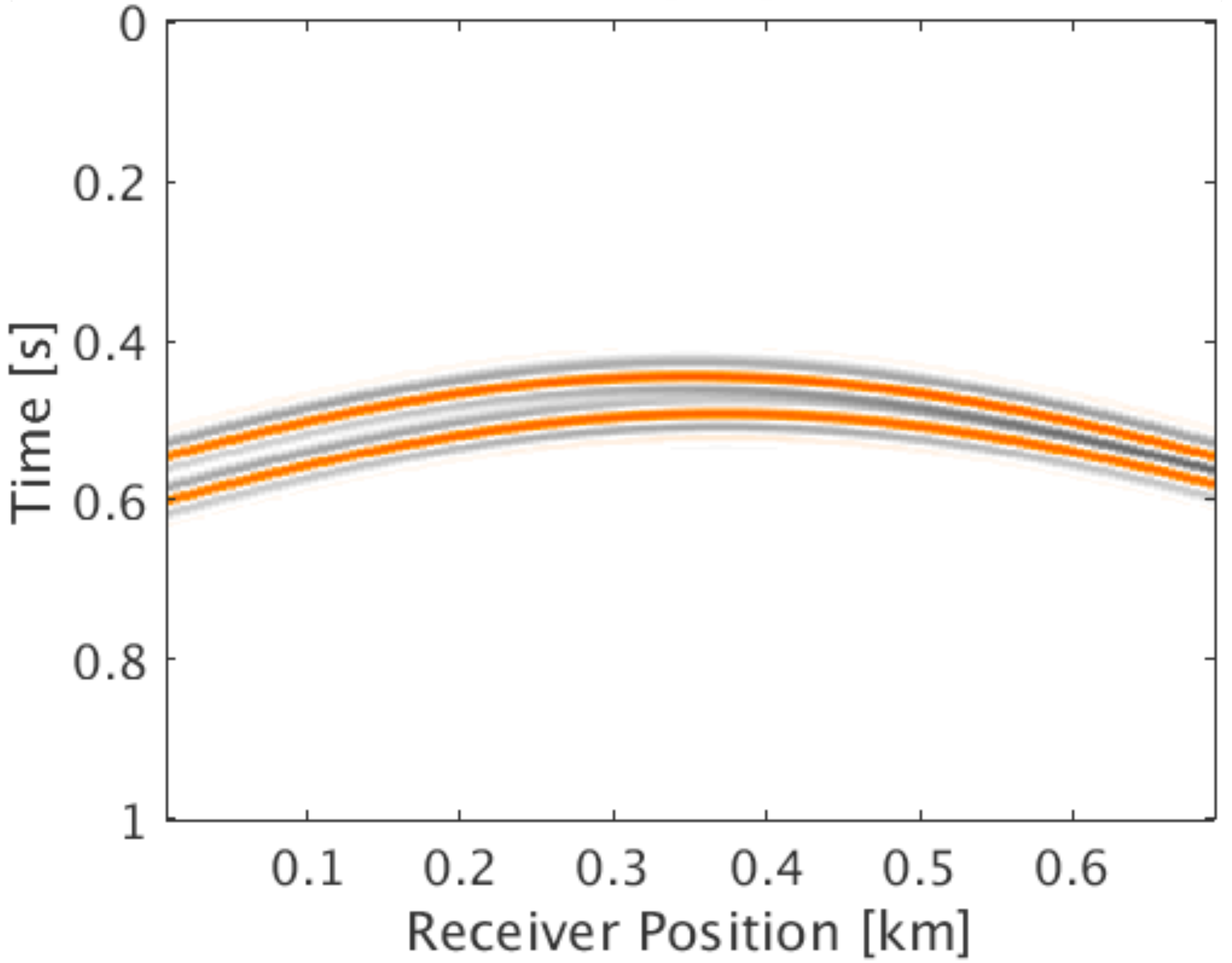
Sparsity promotion

Source-time function



Case study: two nearby sources

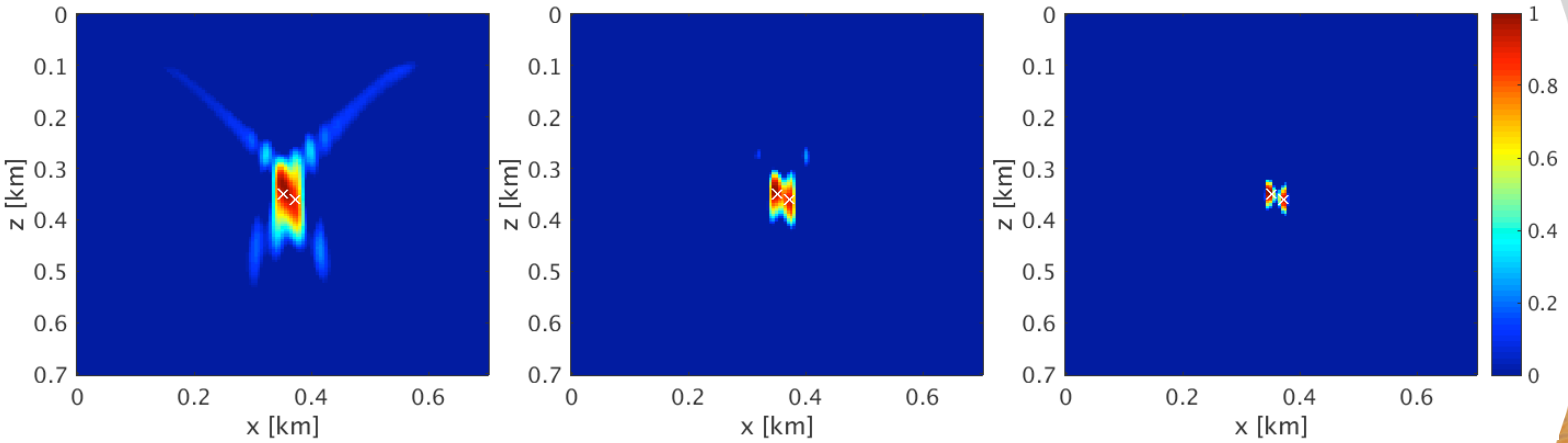




Modeling information:

Model size: 0.7 km x 0.7 km
Grid spacing: 5m
Receiver spacing: 10m
Receiver depth: 20m
Fixed spread: 0.69km
Sampling interval: 2 ms
Recording length: 1s
Peak frequency : 30 Hz
Dominant wavelength: 46 m
Source separation: 22 m

Results w/ different threshold parameters



μ= 8e-5, 200 iterations

μ= 8e-4, 600 iterations

μ = 8e-3, 4900 iterations

SLIM

Acceleration with quasi-Newton: Algorithm

1. **Data d, slowness square m, number of iterations** k //Input
2. **Initialize dual variable** $\mathbf{y} = 10^{-3}\mathbf{d}$
3. $\hat{\mathbf{y}} = \text{L-BFGS}(f(\mathbf{y}), g(\mathbf{y}), \mathbf{y}, k)$ //Dual solution

where $f(\mathbf{y}) = \Psi(\mathbf{y}) - \epsilon\|\mathbf{y}\|_2$ //L-BFGS objective

and $g(\mathbf{y}) = \Psi'(\mathbf{y}) - \epsilon\mathbf{y}/\|\mathbf{y}\|_2$ //L-BFGS gradient

4. $\hat{\mathbf{Q}} = \text{Prox}_{\mu\ell_{2,1}}(\mu\mathcal{F}[\mathbf{m}]^\top(\hat{\mathbf{y}}))$ //Primal solution
5. $\mathbf{I}(\mathbf{x}) = \sum_t |\hat{\mathbf{Q}}(\mathbf{x}, t)|$ //Intensity plot

*where $\Psi(\mathbf{y}) = \min_{\mathbf{Q}} \quad \|\mathbf{Q}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{Q}\|_F - \mathbf{y}^\top(\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathbf{d})$

* $\Psi'(\mathbf{y}) = \mathbf{d} - \mathcal{F}[\mathbf{m}](\text{Prox}_{\mu\ell_{2,1}}(\mu\mathcal{F}[\mathbf{m}]^\top(\mathbf{y})))$ is the gradient of $\Psi(\mathbf{y})$

[Huang et al.,'13; Yin ,'10]

SLIM

Acceleration with quasi-Newton: Algorithm

1. **Data d, slowness square m, number of iterations** k //Input
2. **Initialize dual variable y** $= 10^{-3}\mathbf{d}$
3. $\hat{\mathbf{y}} = \text{L-BFGS}(f(\mathbf{y}), g(\mathbf{y}), \mathbf{y}, k)$ //Dual solution

where $f(\mathbf{y}) = \Psi(\mathbf{y}) - \epsilon\|\mathbf{y}\|_2$ //L-BFGS objective

and $g(\mathbf{y}) = \Psi'(\mathbf{y}) - \epsilon\mathbf{y}/\|\mathbf{y}\|_2$ //L-BFGS gradient

4. $\hat{\mathbf{Q}} = \text{Prox}_{\mu\ell_{2,1}}(\mu\mathcal{F}[\mathbf{m}]^\top(\hat{\mathbf{y}}))$ //Primal solution
5. $\mathbf{I}(\mathbf{x}) = \sum_t |\ \hat{\mathbf{Q}}(\mathbf{x}, t)\ |$ //Intensity plot

lives in much smaller space

- dimensions equals that of observed data

*where $\Psi(\mathbf{y}) = \min_{\mathbf{Q}} \quad \|\mathbf{Q}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{Q}\|_F - \mathbf{y}^\top(\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathbf{d})$

* $\Psi'(\mathbf{y}) = \mathbf{d} - \mathcal{F}[\mathbf{m}](\text{Prox}_{\mu\ell_{2,1}}(\mu\mathcal{F}[\mathbf{m}]^\top(\mathbf{y})))$ is the gradient of $\Psi(\mathbf{y})$

[Herrmann et al.,’09; Song et al. ,’95]

SLIM

Further acceleration w/ 2D Preconditioning

$$\underset{\mathbf{Q}}{\text{minimize}} \;\; \|\mathbf{Q}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{Q}\|_F^2$$

$$\text{subject to } \|\mathcal{M}_L\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathcal{M}_L\mathbf{d}\|_2 \leq \gamma$$

*with $\mathcal{M}_L := \partial_{|t|}^{1/2}$ is the half differentiation correction

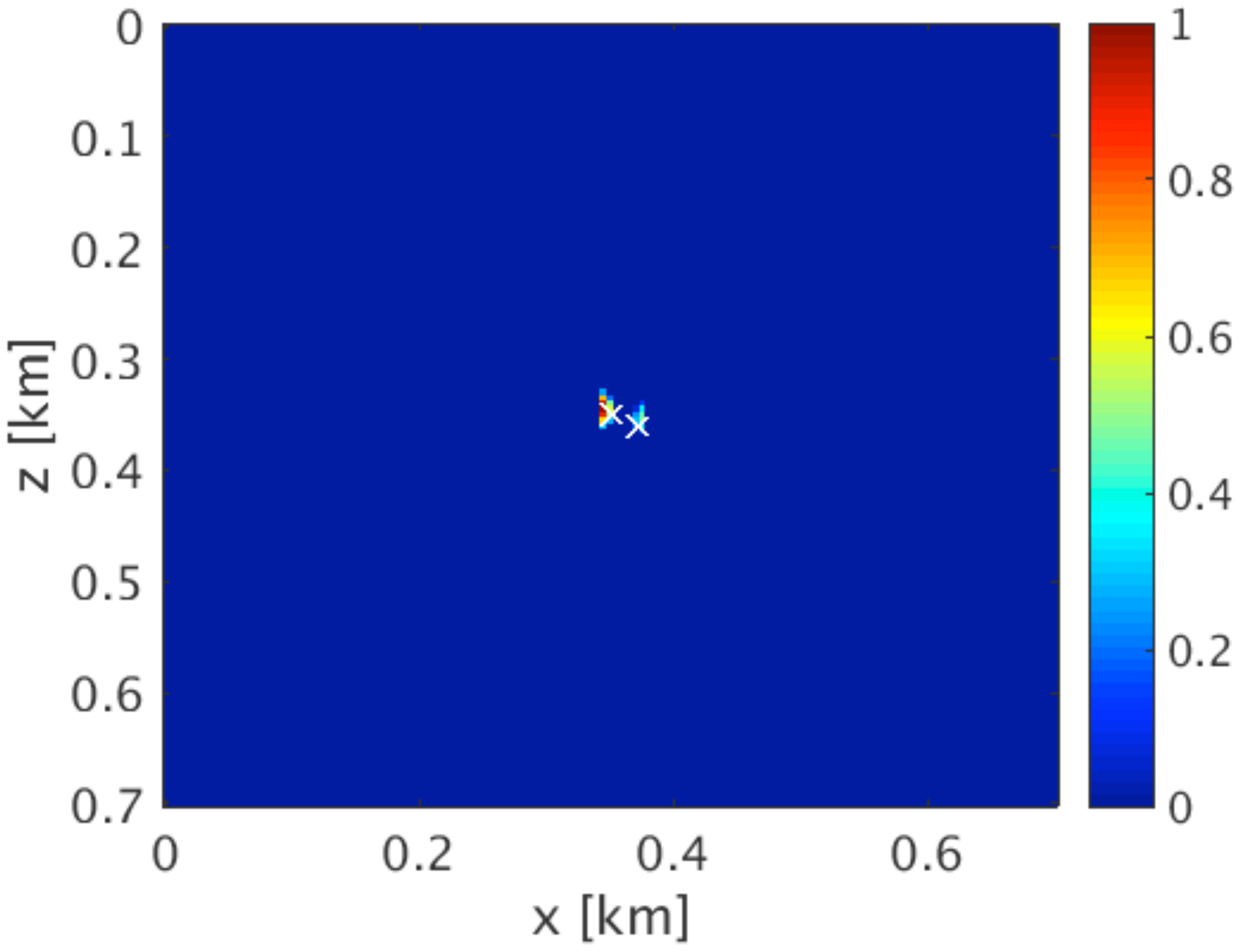
*where $\partial_{|t|}^{1/2} = \mathbf{F}^{-1}|\omega|^{1/2}\mathbf{F}$

*$\mathbf{F}$ is the Fourier transform and ω is the frequency

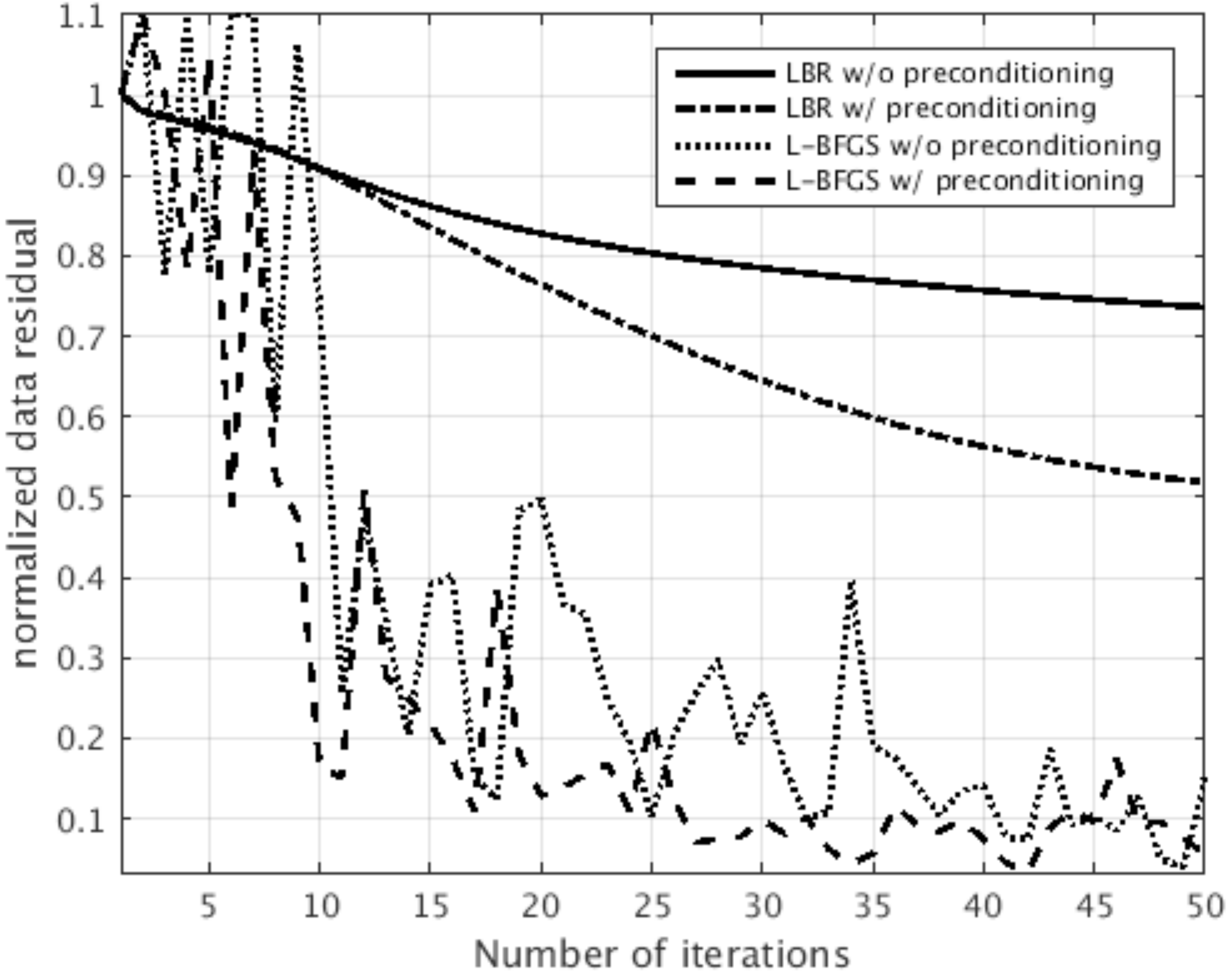
*γ is the noise level

Estimated location

w/μ = 8e-2 and 10 iterations



Convergence comparison: LBR vs L-BFGS



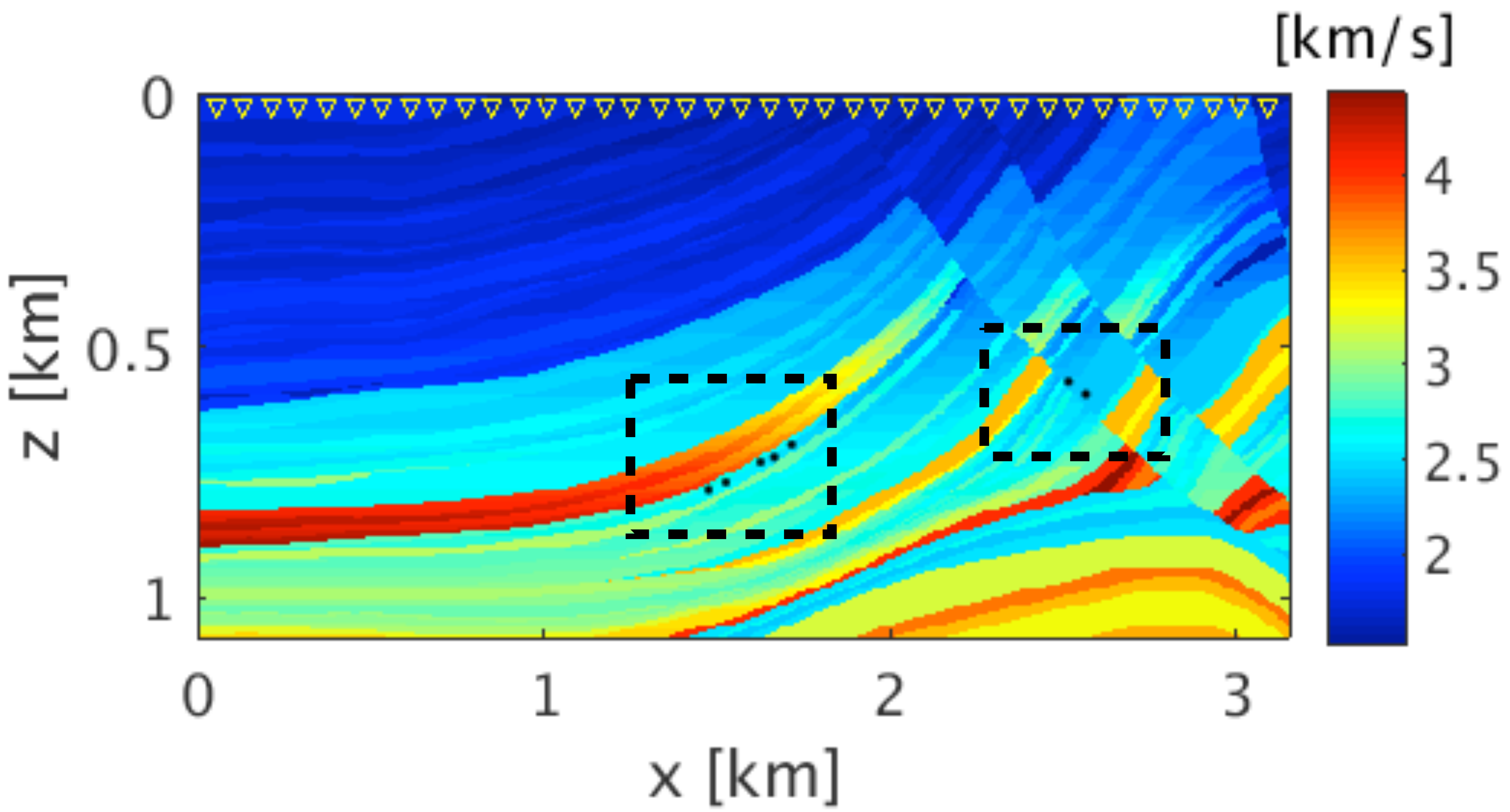
Convergence comparison

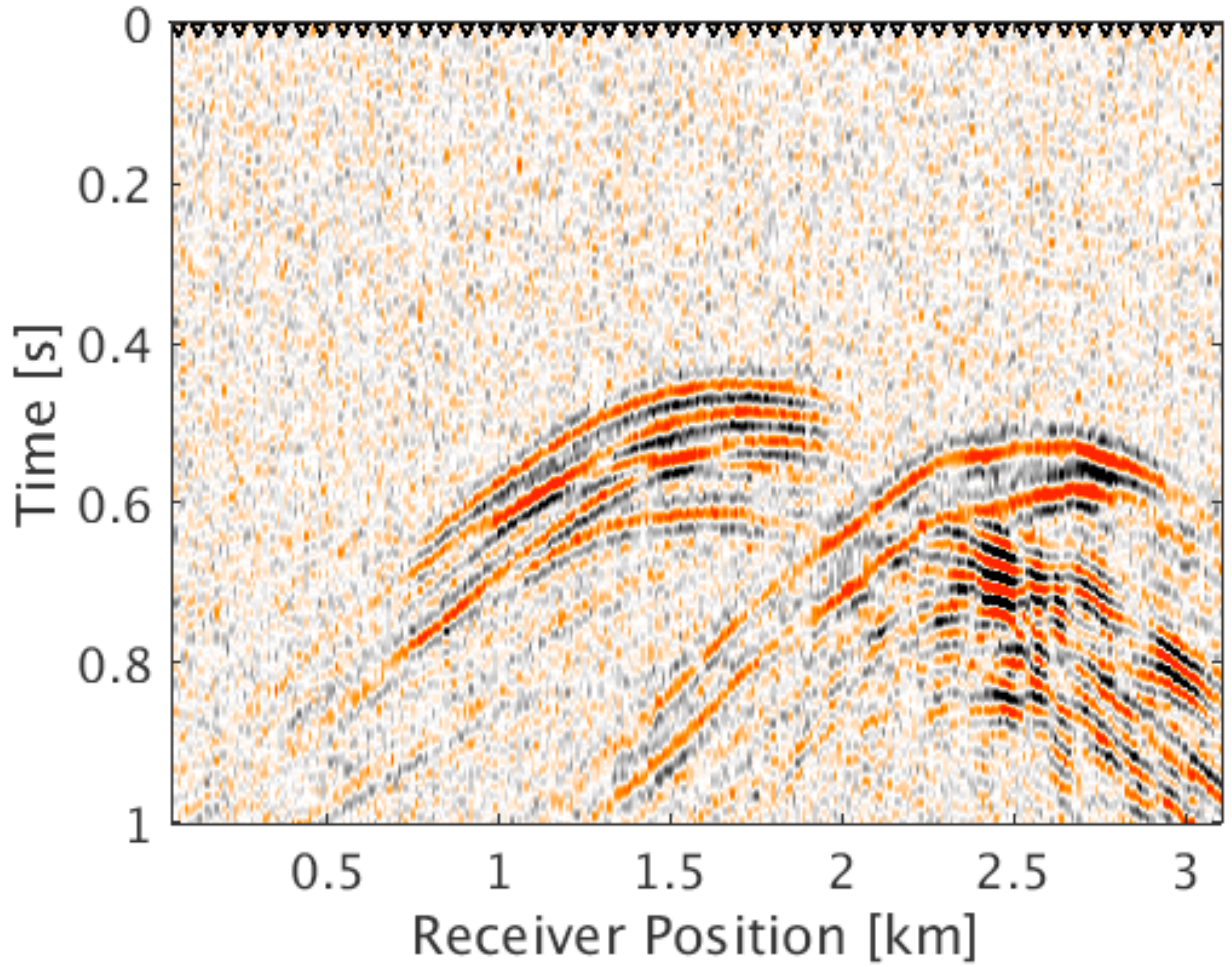
- Using same value of μ

Improvement in convergence with

- Dual formulation and
- 2D Preconditioning

Multiple source cluster experiment in Marmousi model

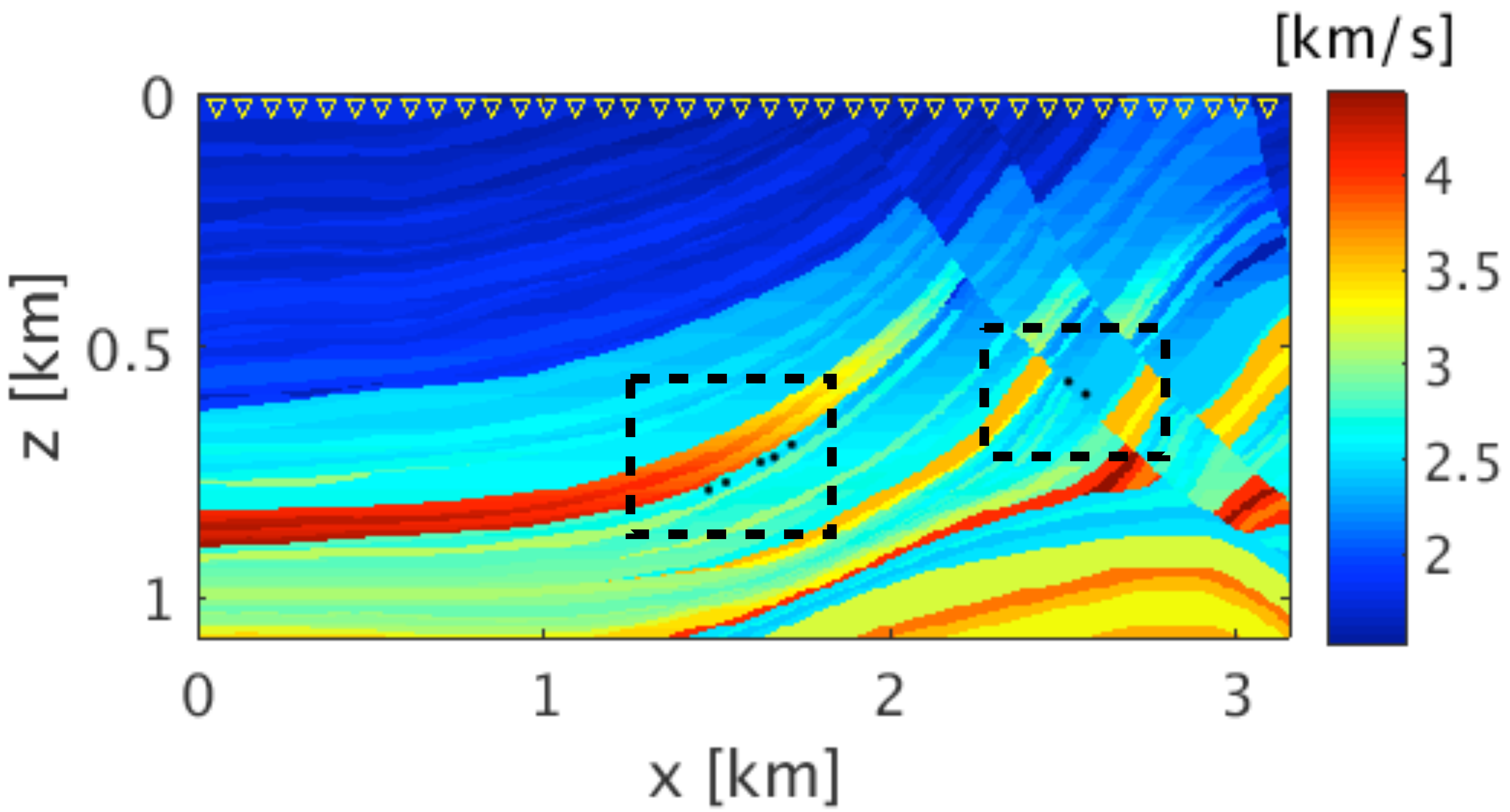


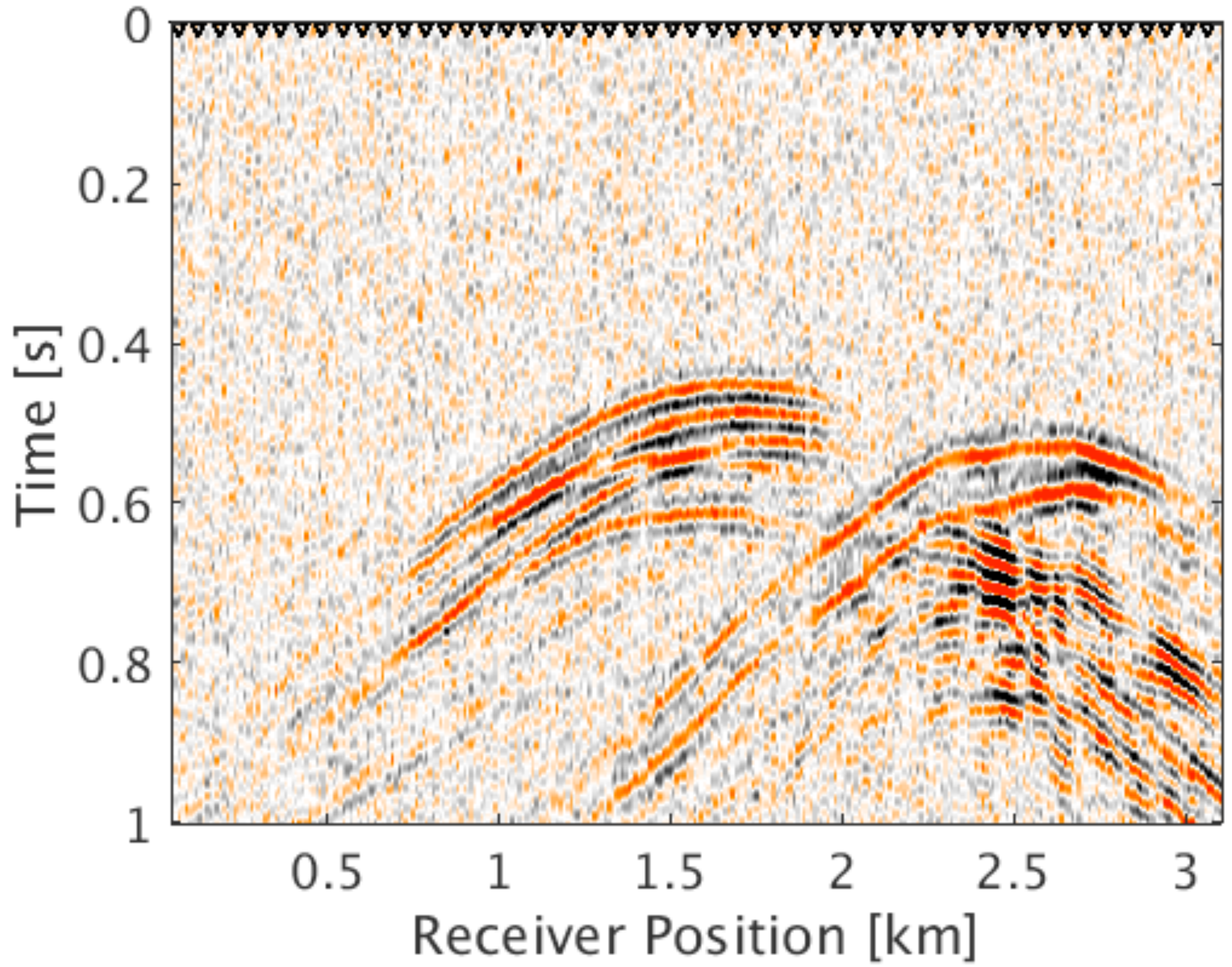


Modeling information:

Model size: 3.15 km x 1.08 km
Grid spacing: 5 m
Total number of sources: 7
Peak frequency : 22 Hz, 25 Hz & 30 Hz
Receiver spacing: 10m
Receiver depth: 20m
Sampling interval: 0.5 ms
Recording length: 1 s
Free surface: No

Multiple source cluster experiment in Marmousi model



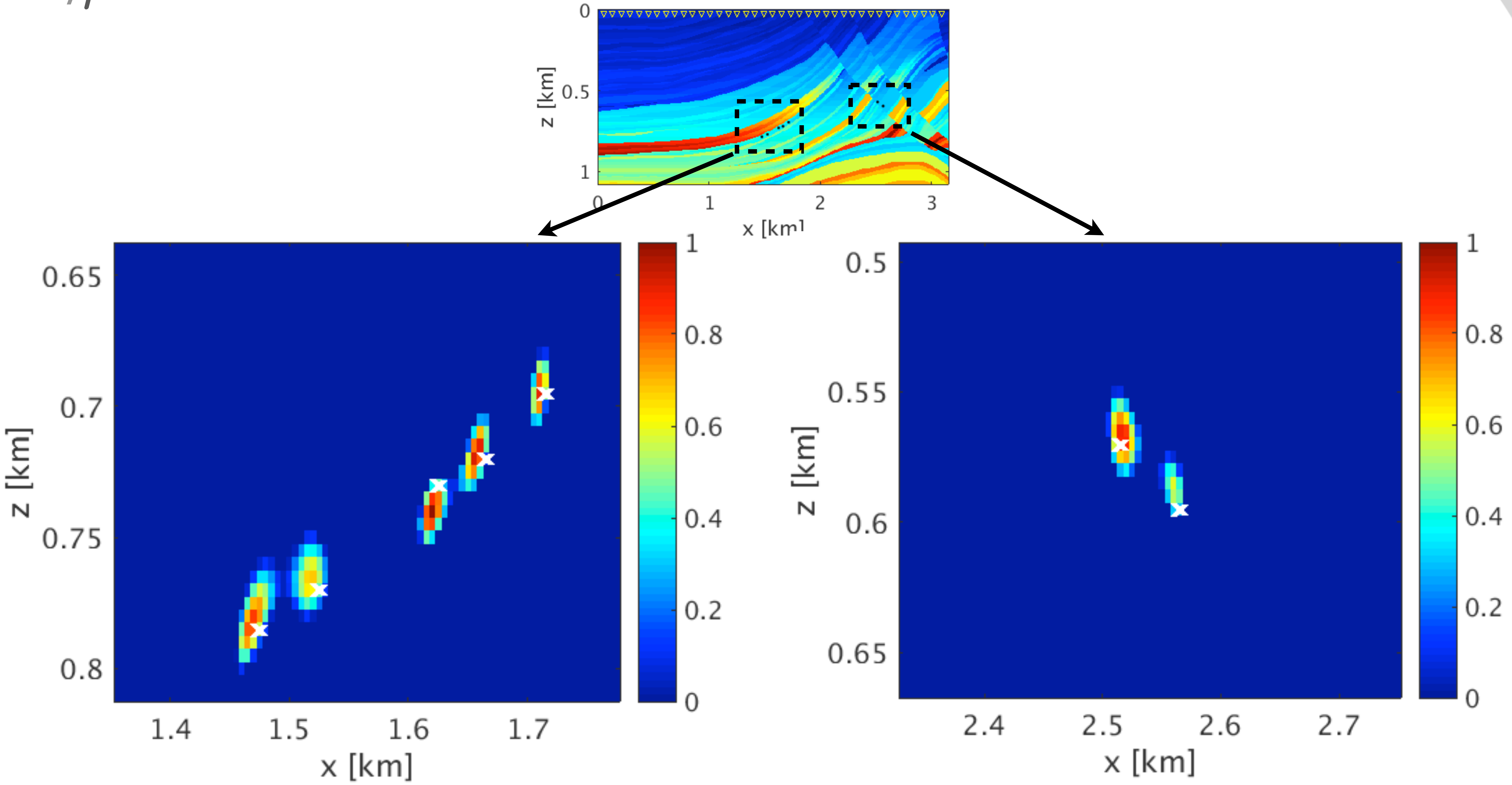


- Contaminated with 5 to 45 Hz random noise
- SNR = 3.5 dB

Modeling information:

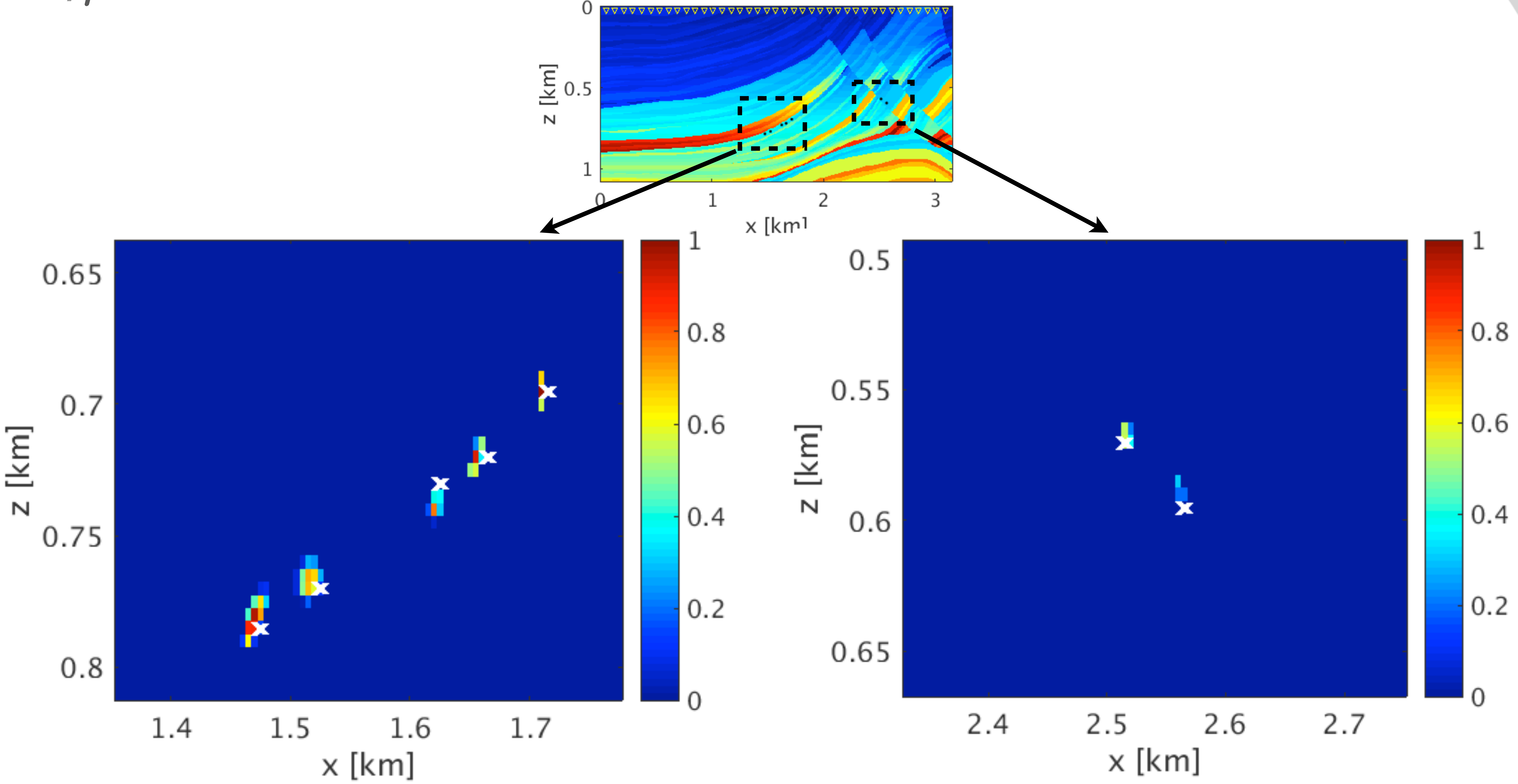
Model size: 3.15 km x 1.08 km
Grid spacing: 5 m
Total number of sources: 7
Peak frequency : 22 Hz, 25 Hz & 30 Hz
Receiver spacing: 10m
Receiver depth: 20m
Sampling interval: 0.5 ms
Recording length: 1 s
Free surface: No

Estimated location
w/μ = 9e-4 and 10 iterations

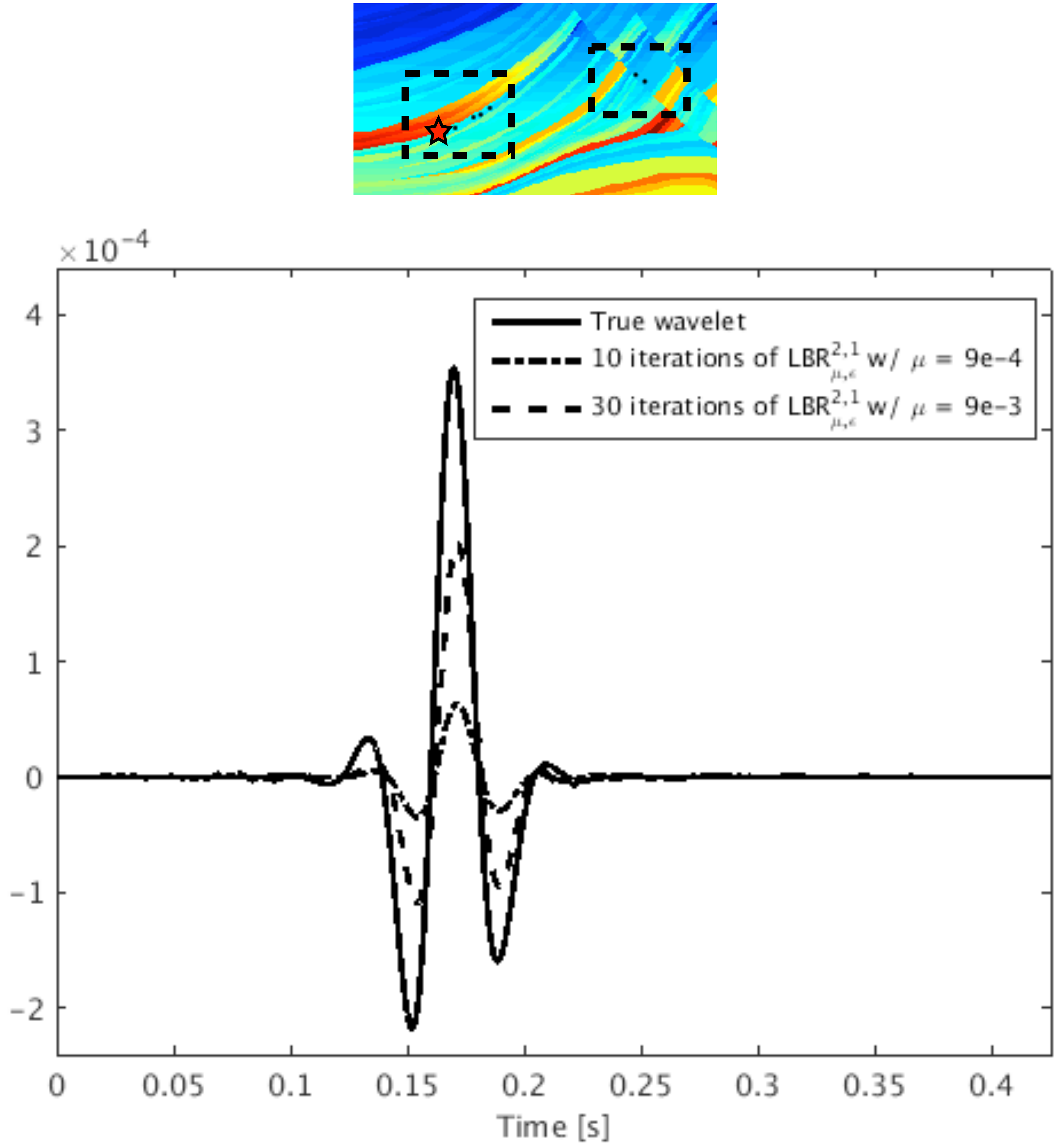


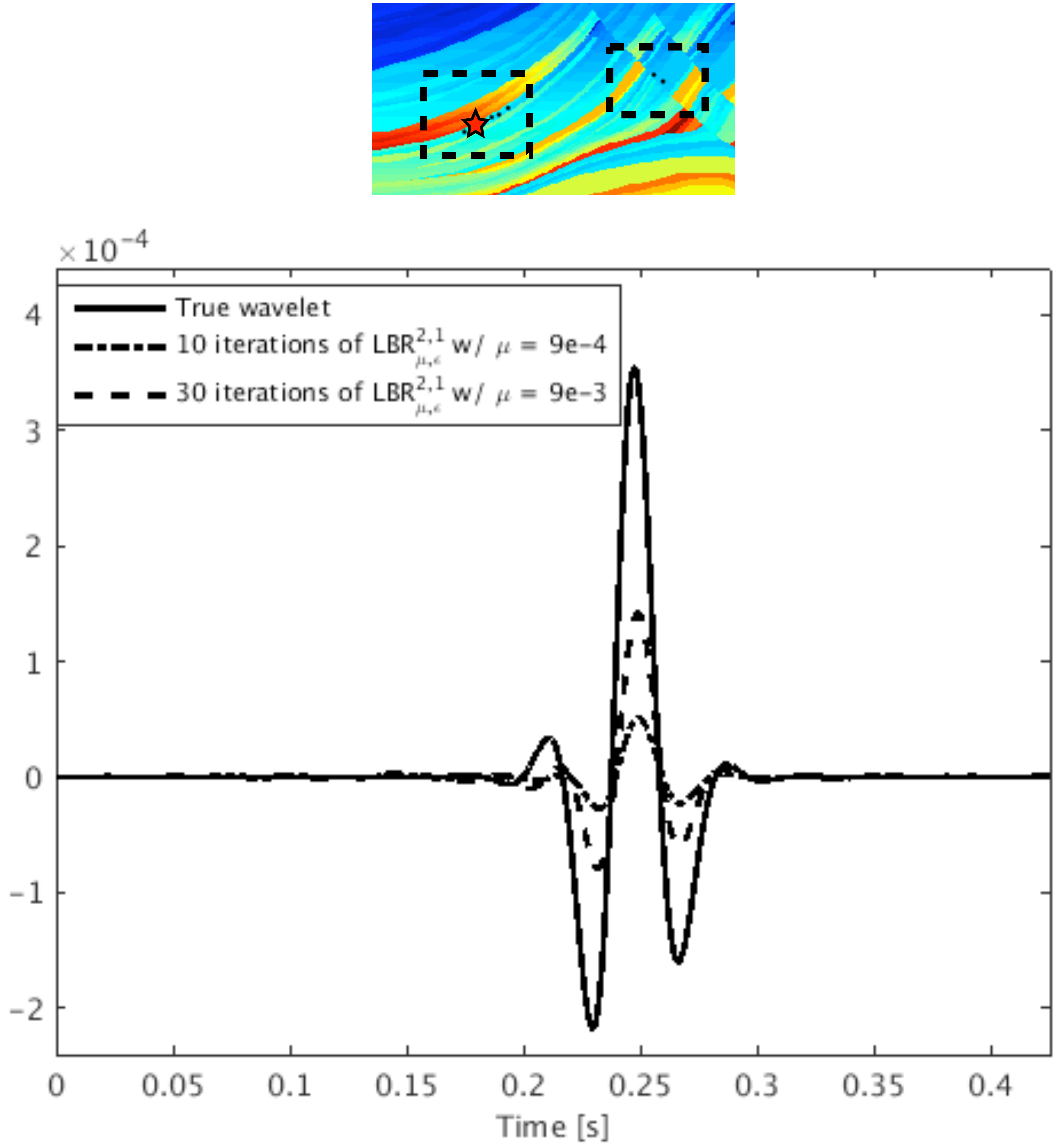
Estimated location

w/μ = 9e-3 and 30 iterations

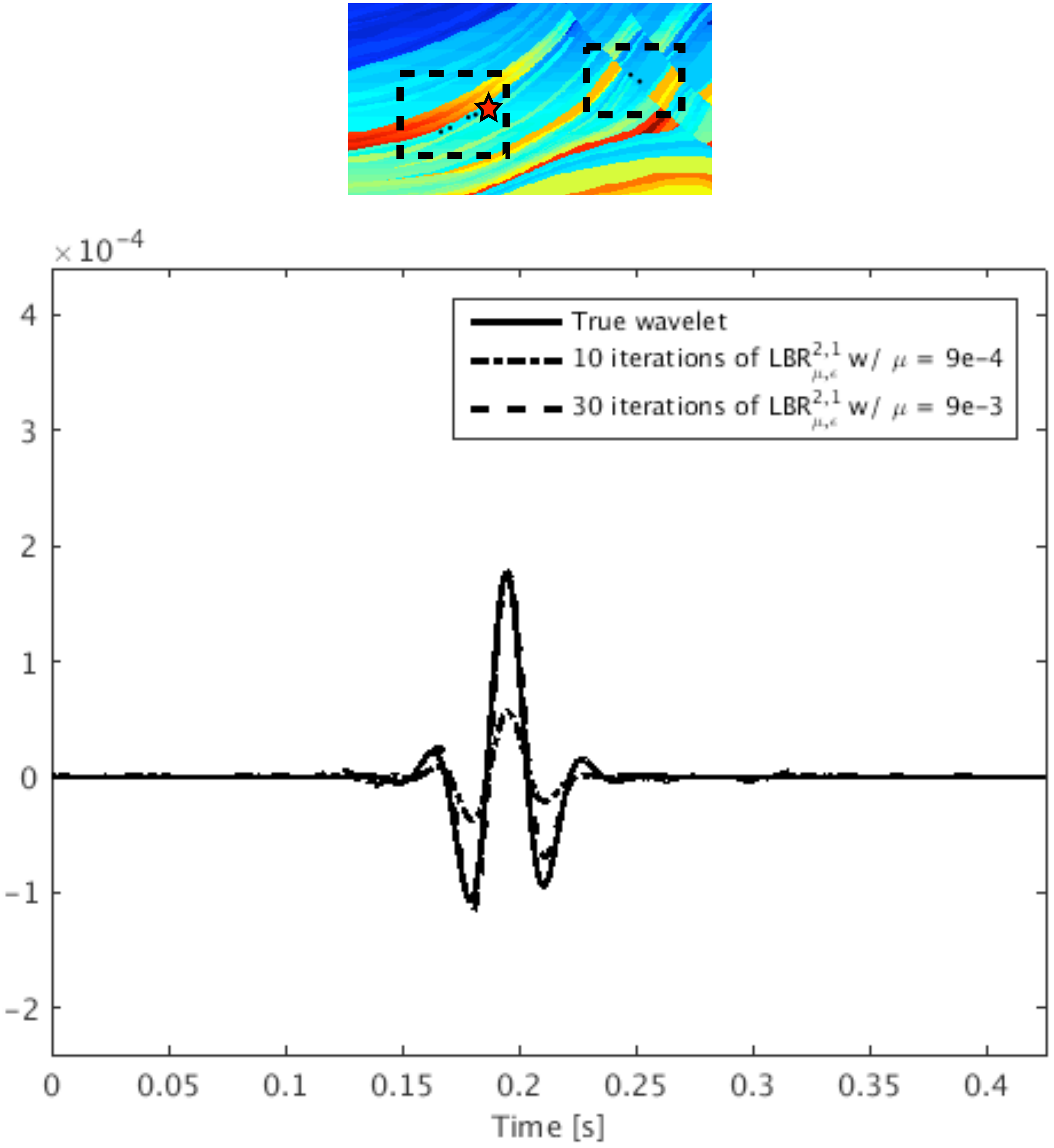


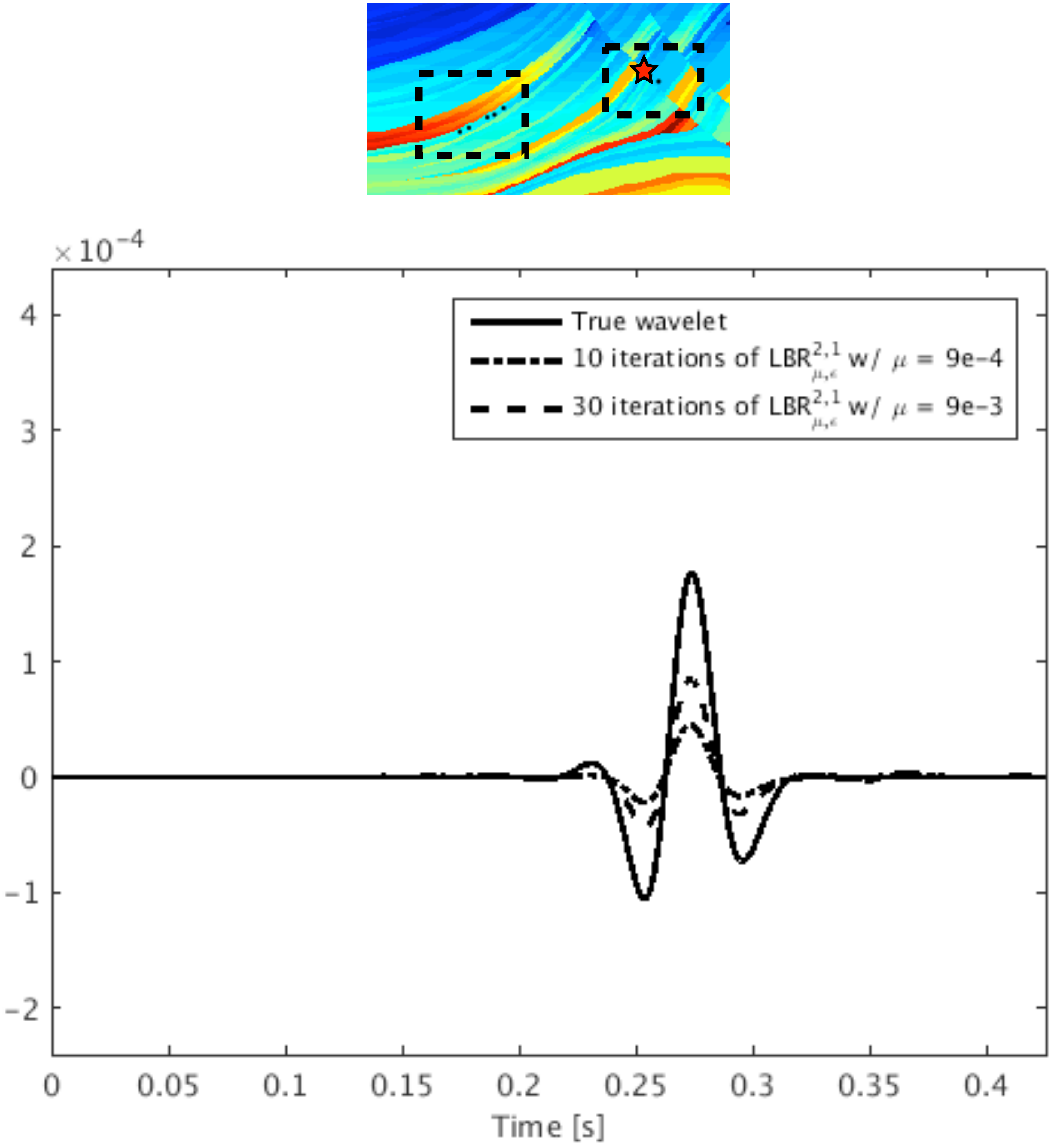
Wavelet comparison



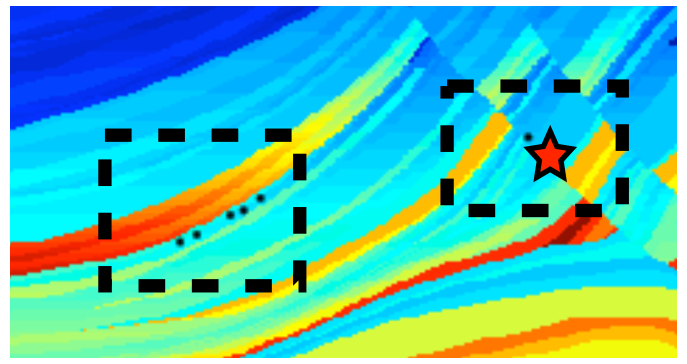


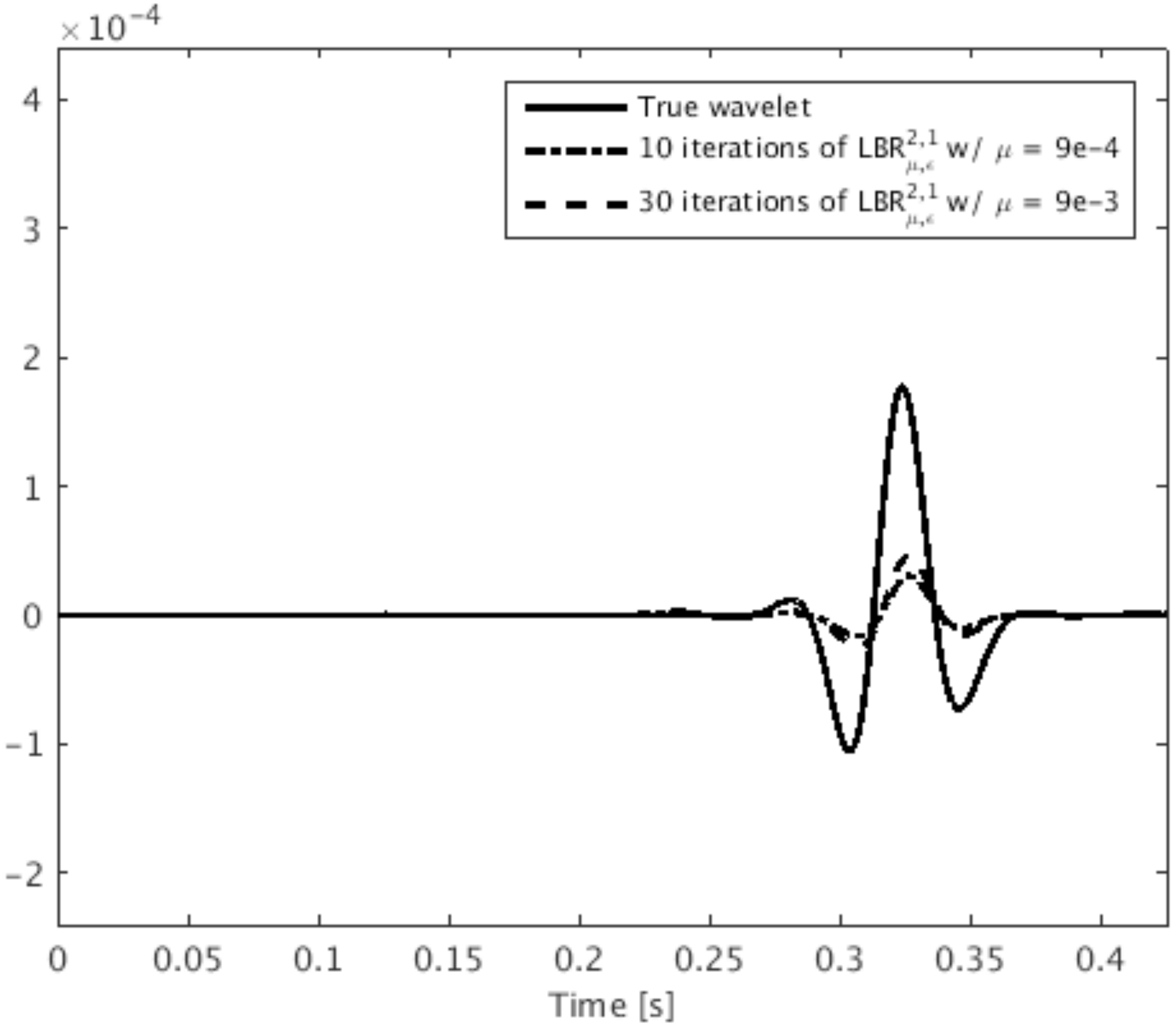
Wavelet comparison



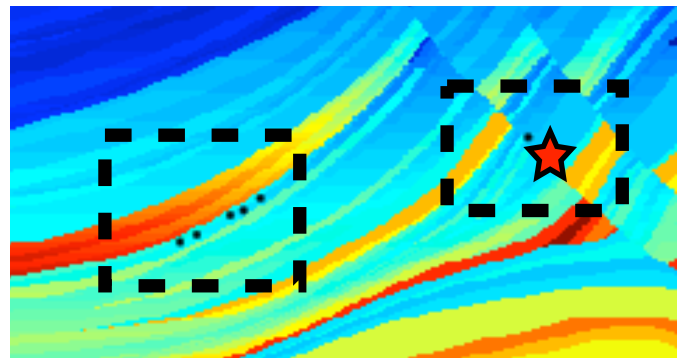


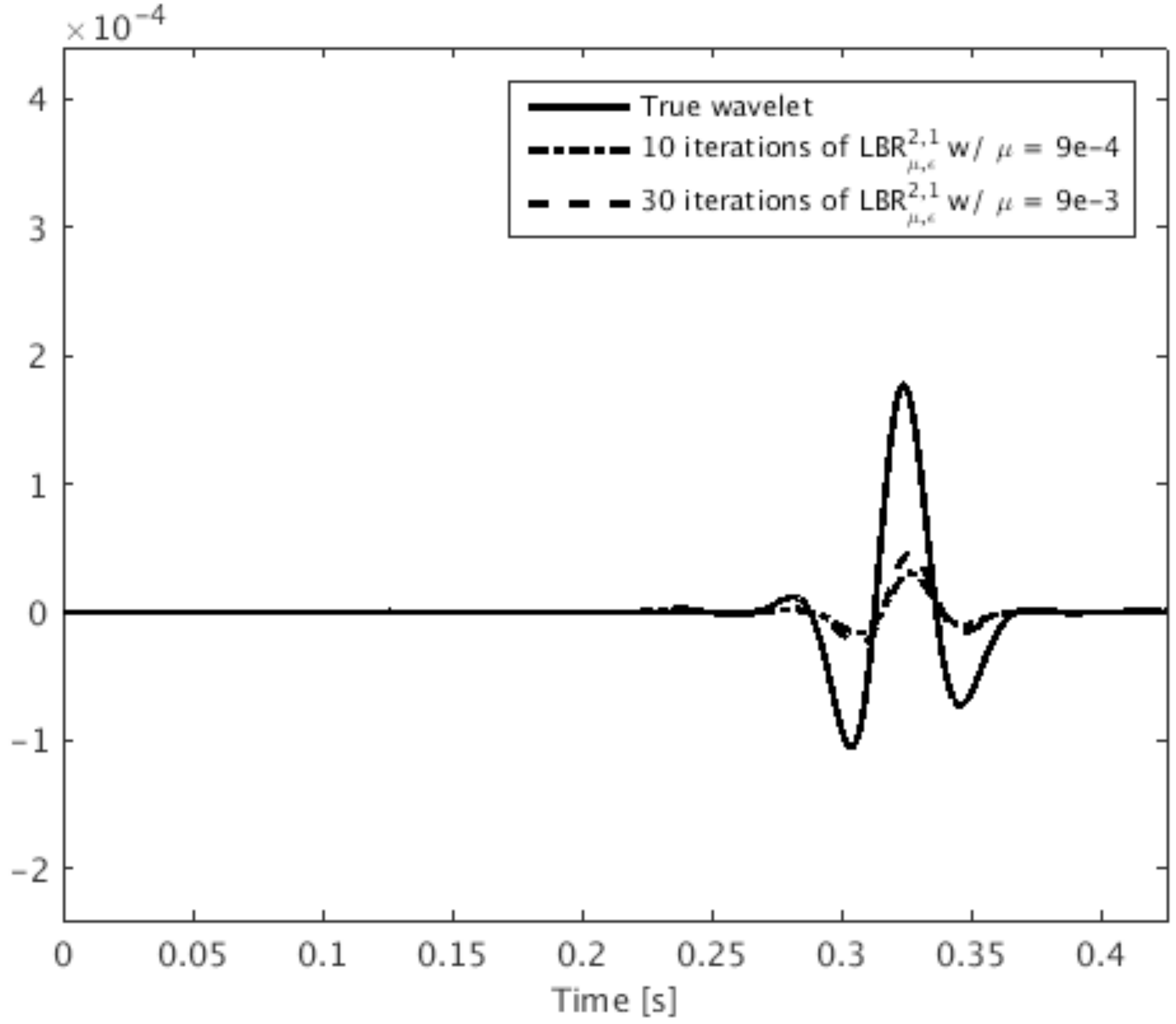
Wavelet comparison





Wavelet comparison





In chapter 5

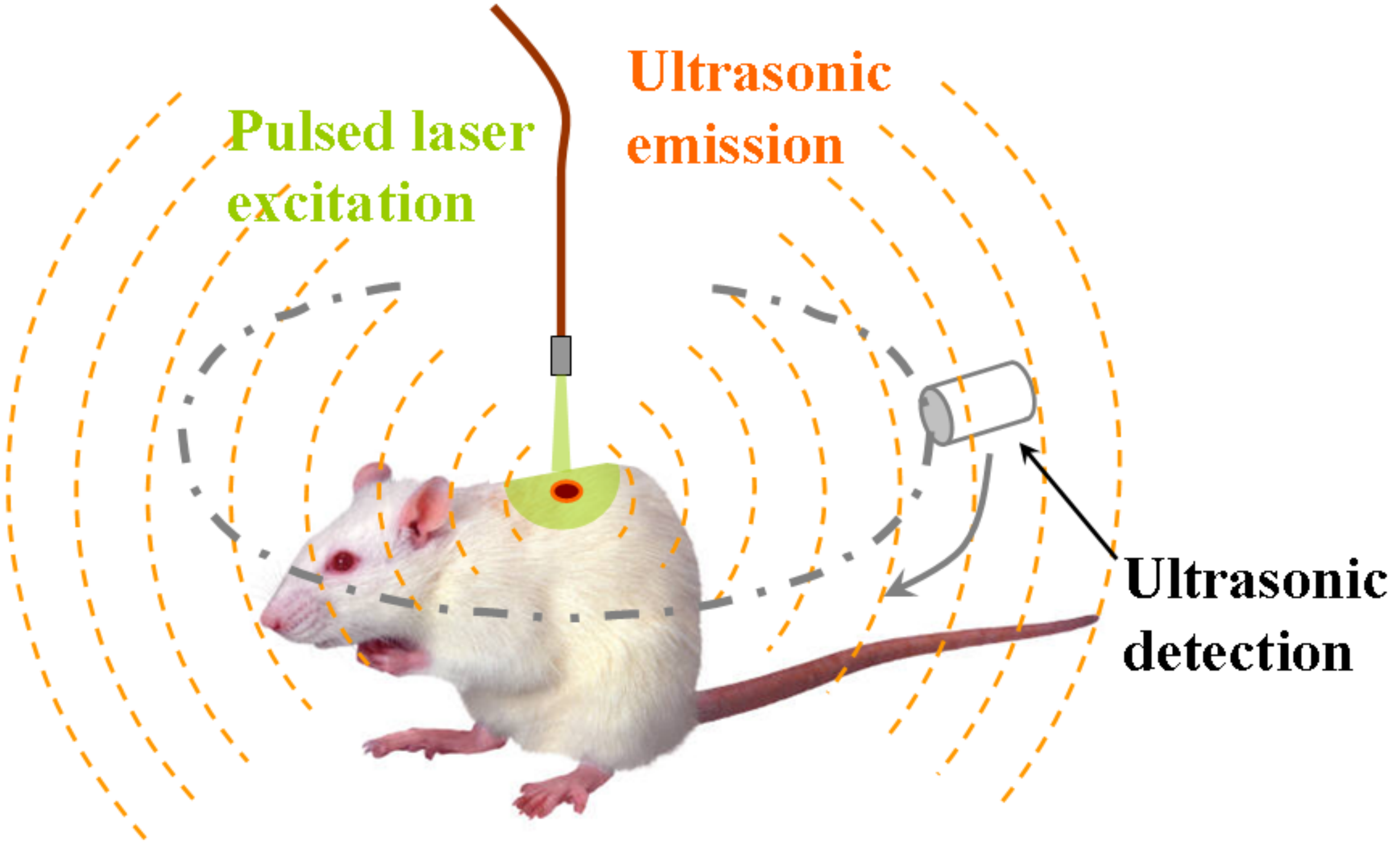
- Debiasing step to correct amplitude and
- Detection of microseismic sources from noisy data

Key Contributions: Chapter 6

Sparsity-promoting photo acoustic imaging

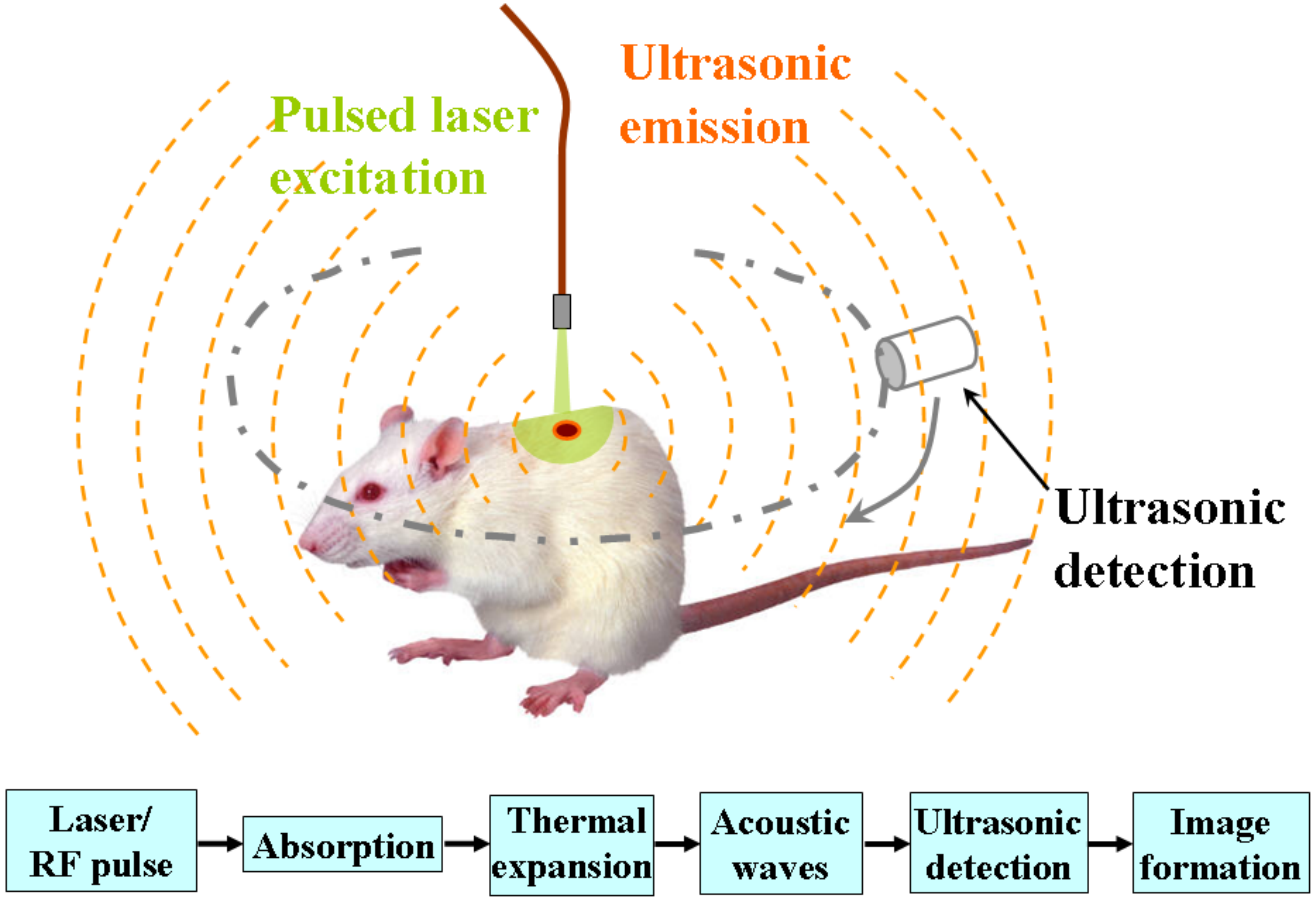
- Simultaneous imaging of absorption map and source estimation
- With reduced number of transducers and with smooth velocity model

Photoacoustic Imaging



Laser/ RF pulse → Absorption → Thermal expansion → Acoustic waves → Ultrasonic detection → Image formation

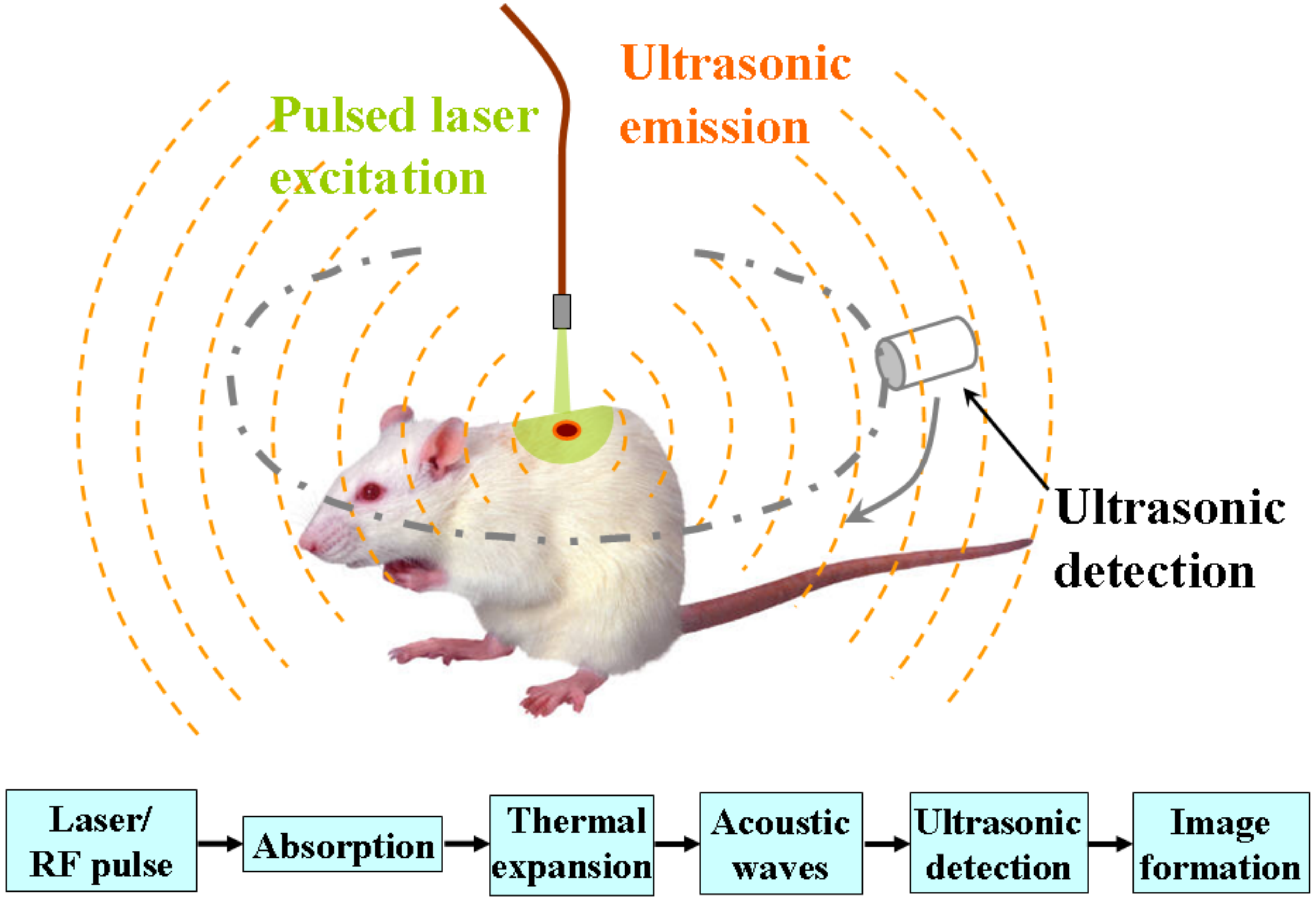
Photoacoustic Imaging



Objectives

- detection of photoabsorbers
- estimation of associated source-time function

Photoacoustic Imaging



Objectives

- detection of photoabsorbers
- estimation of associated source-time function

Challenges

- Time-reversal methods do not estimate source-time function
- Require dense transducer coverage

SLIM

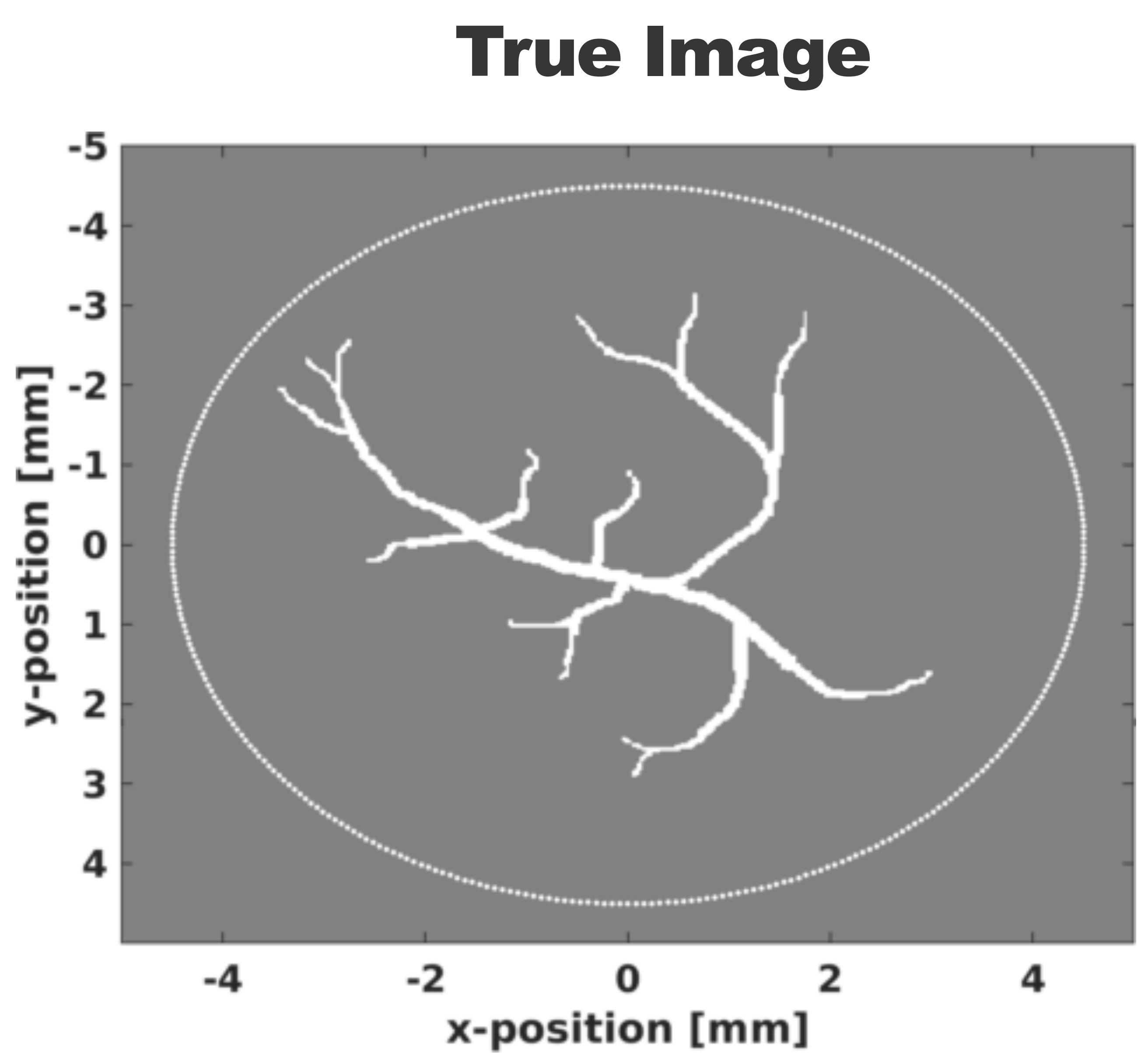
Solving w/ Linearized Bregman

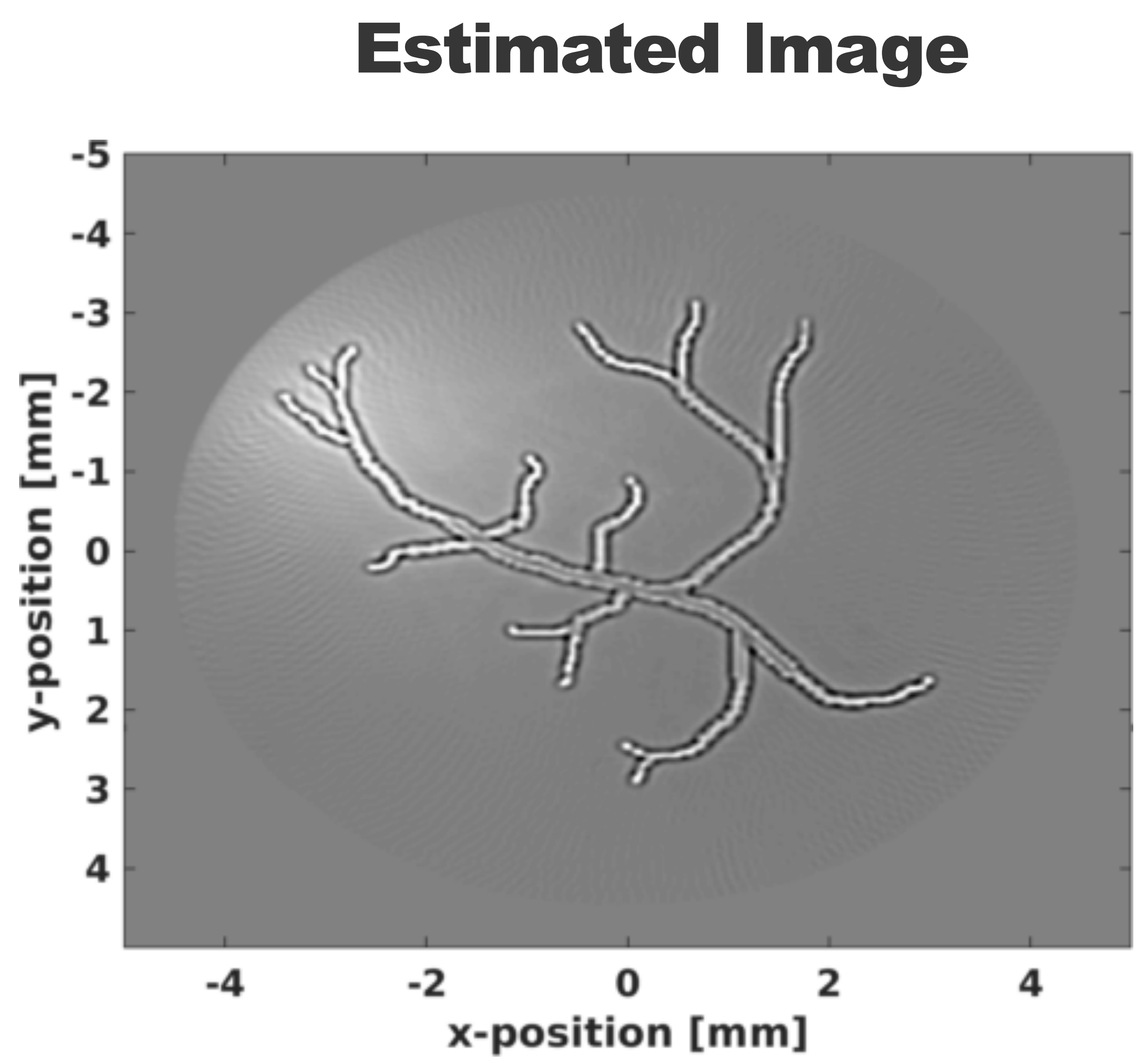
$$\begin{aligned} \underset{\mathbf{Q}\in\mathcal{T}_\tau}{\text{minimize}} \quad & \|\mathbf{Q}\|_{2,1} + \frac{1}{2\mu}\|\mathbf{Q}\|_F^2 \\ \text{subject to} \quad & \|\mathcal{F}[\mathbf{m}](\mathbf{Q}) - \mathbf{d}\|_2 \leq \epsilon \end{aligned}$$

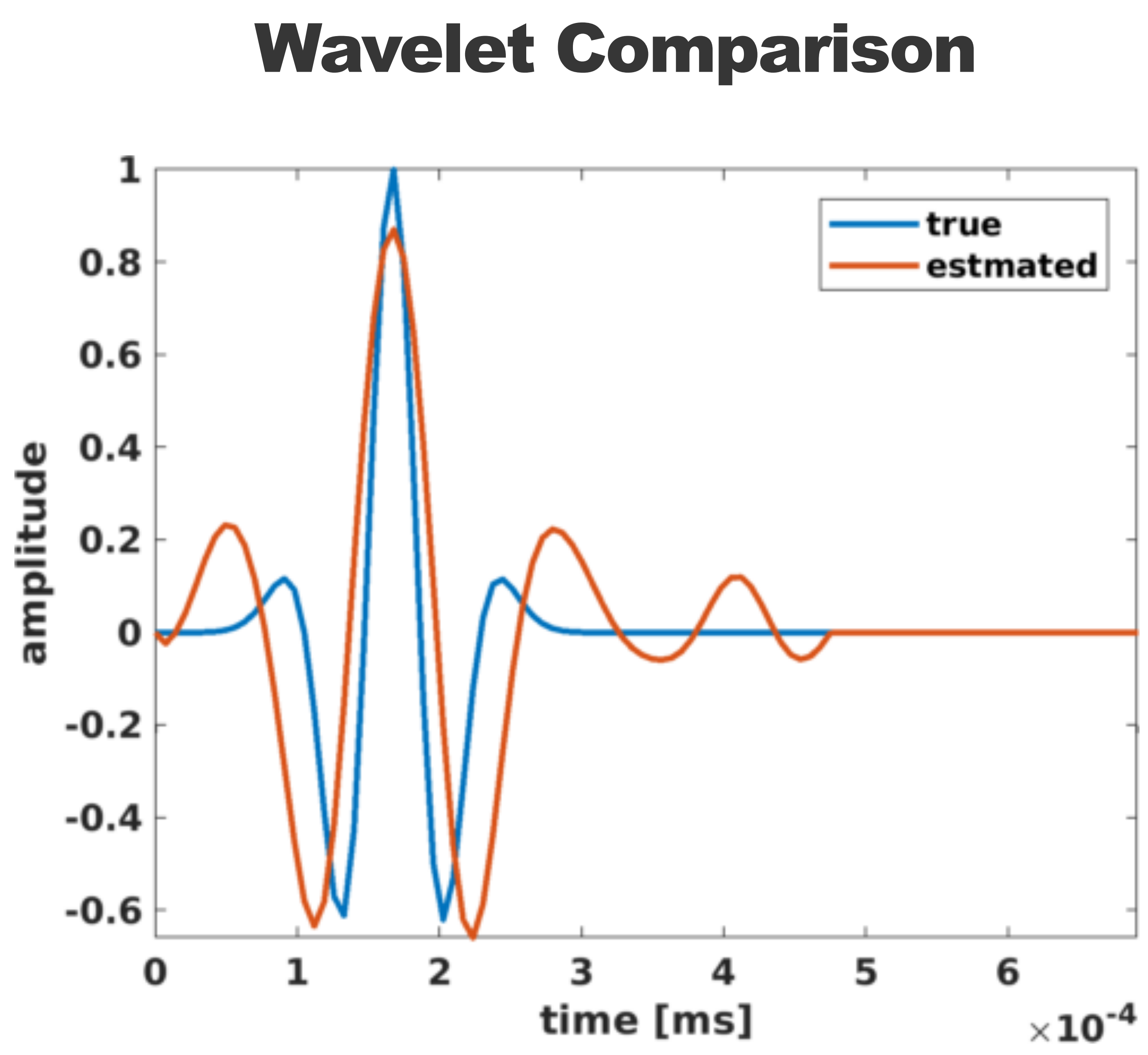
*where $\mathcal{T}_\tau = \{\mathbf{Q}|\mathbf{Q}(\cdot, t) = 0,\ t > \tau\}$

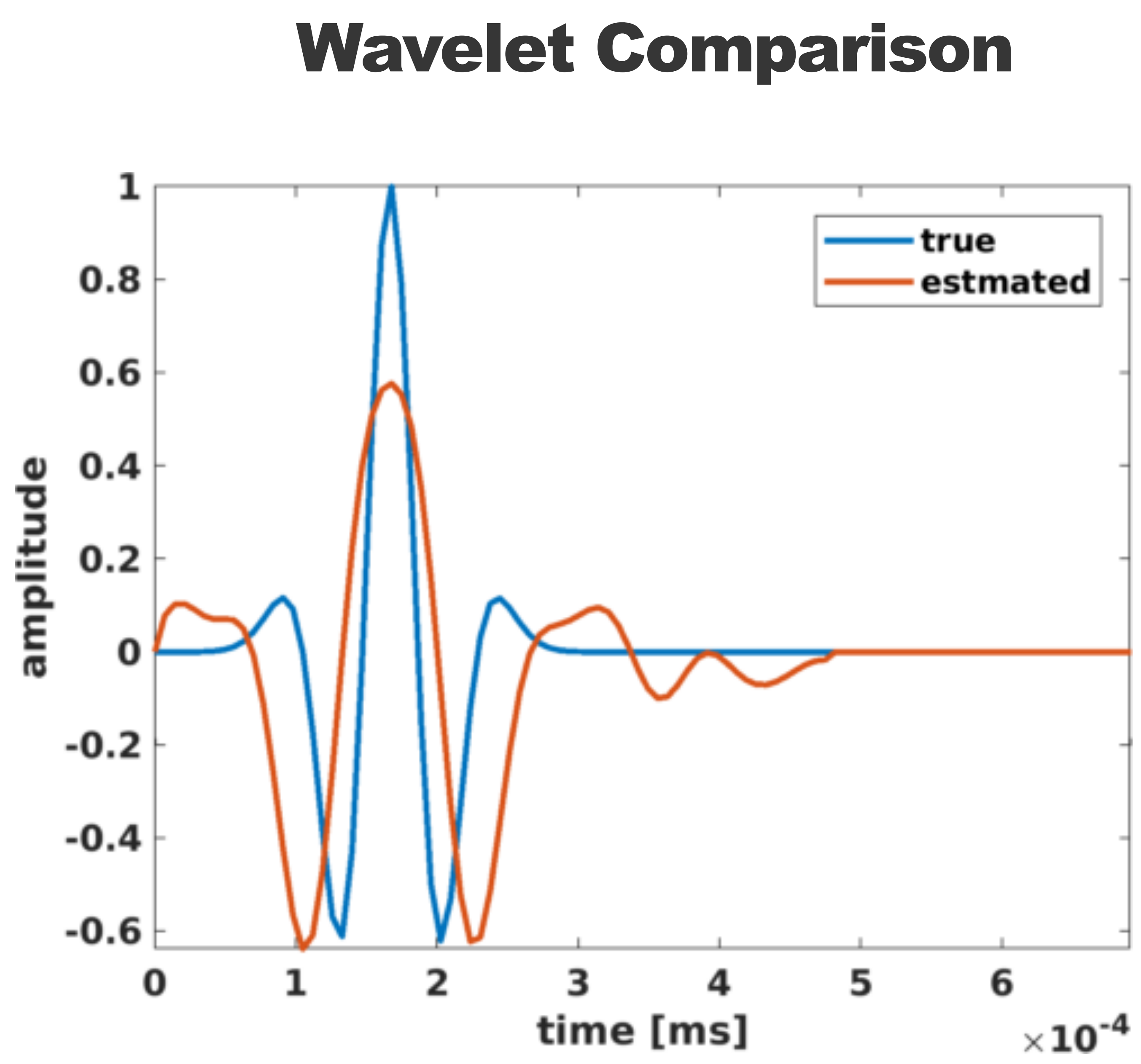
*where τ is the user defined duration

Case Study: Blood vessel phantom









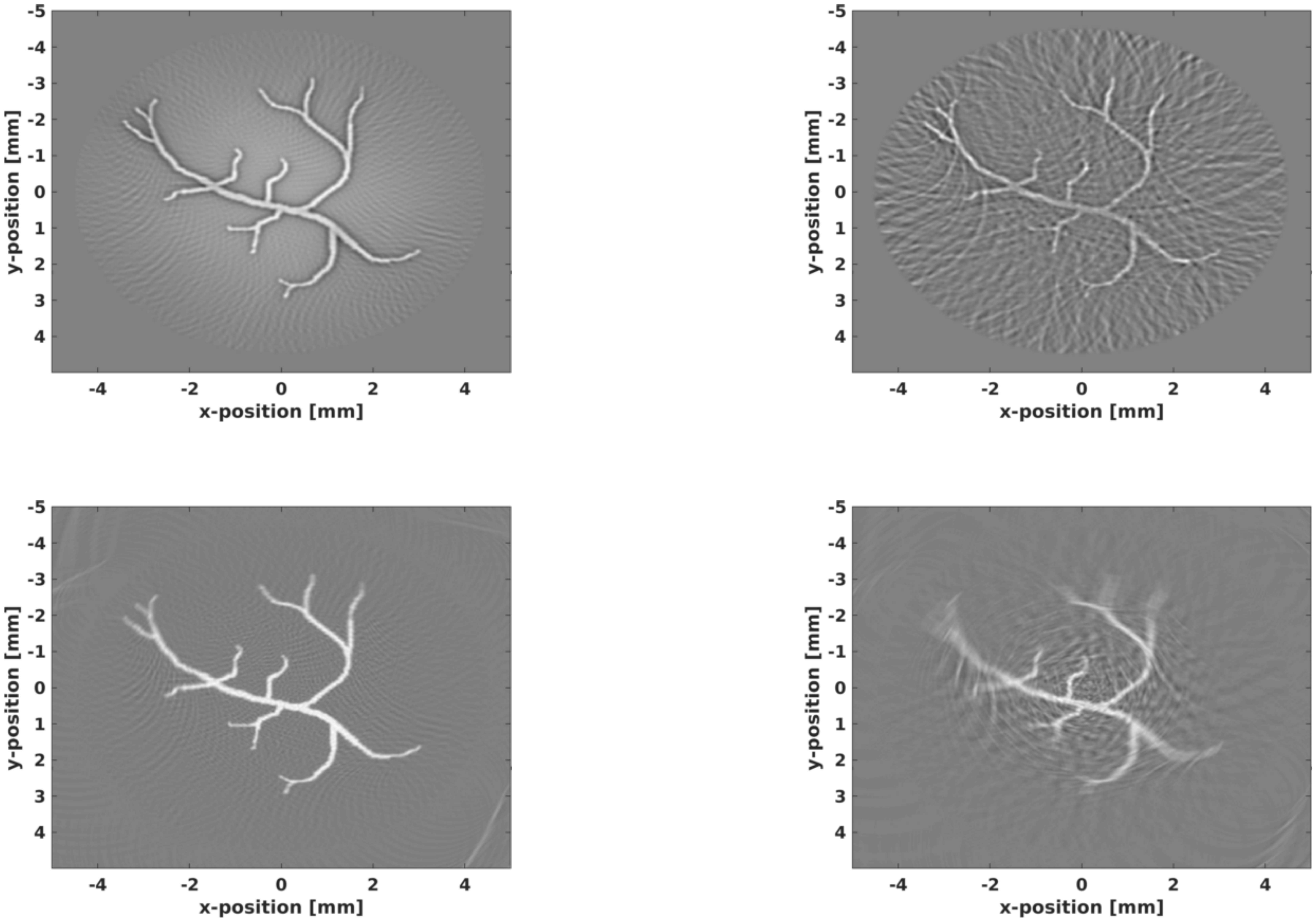
Effect of transducer sampling

Transducers every 2 degrees

Transducers every 6 degrees

Sparsity-Promoting

Time reversal w/ K-wave



Conclusions

Sparsity promotion based method

- can simultaneously estimate multiple source locations & source-time functions
- can provide locations of fractures by resolving microseismic sources within half a wavelength
- works w/ sources of different frequencies & origin times

SLIM

Conclusions

Sparsity promotion based method

- can simultaneously estimate multiple source locations & source-time functions
- can provide locations of fractures by resolving microseismic sources within half a wavelength
- works w/ sources of different frequencies & origin times

Dual formulation provides a computationally efficient scheme

[Sharan et al.,'18]

SLIM

Conclusions

Sparsity promotion based method

- can simultaneously estimate multiple source locations & source-time functions
- can provide locations of fractures by resolving microseismic sources within half a wavelength
- works w/ sources of different frequencies & origin times

Dual formulation provides a computationally efficient scheme

Wavelet scaling can be corrected using debiasing method

Future work

Checkpointing strategy to avoid storage of complete source-wavefield

Inclusion of TV-norm instead of ℓ_1 norm in space

- for sources along a plane
- with sharp boundaries

Publications

1. S. Sharan, Y. Zhang, O. Lopez and F.J. Herrmann, 2020, *Large scale high-frequency wavefield reconstruction with recursively weighted matrix factorizations*, Submitted to Geophysics

2. Y. Zang, S. Sharan, O. Lopez and F.J. Herrmann, 2020, *Wavefield recovery with limited-subspace weighted matrix factorizations*, SEG Technical Program Expanded Abstracts

3. S. Sharan, R. Wang and F.J. Herrmann, 2019, *Fast sparsity-promoting microseismic source estimation*, Geophysical Journal International, vol. 216, pp. 164-181

4. Y. Zhang, S. Sharan and F.J. Herrmann, 2019, *High-frequency wavefield recovery with weighted matrix factorizations*, SEG Technical Program Expanded Abstracts

5. R. Kumar, S. Sharan, N. Moldoveanu and F.J. Herrmann, 2018, *Compressed sensing based land simultaneous acquisition using encoded sweeps*, EAGE Annual Conference Proceedings

6. S. Sharan, R. Kumar, R. Wang and F.J. Herrmann, 2018, *A debiasing approach to microseismic*, SEG Technical Program Expanded Abstracts

7. S. Sharan, R. Kumar, D.S. Dumani, M. Louboutin, R. Wang, S. Emelianov and F.J. Herrmann, 2018, *Sparsity-promoting photoacoustic imaging with source estimation*, IEEE International Ultrasonics Symposium

8. S. Sharan, R. Wang and F.J. Herrmann, 2017, *High resolution fast microseismic source collocation and source time function estimation*, SEG Technical Program Expanded Abstracts

9. R. Kumar, H. Wason, S. Sharan and F.J. Herrmann, 2017, *Highly repeatable 3D compressive full-azimuth towed-streamer time-lapse acquisition — a numerical feasibility study at scale*, The Leading Edge, vol. 36, pp. 677-687

10. R. Kumar, S. Sharan, H. Wason and F.J. Herrmann, 2016, *Efficient large-scale 5D seismic data acquisition and processing using rank-minimization*, SEG Technical Program Expanded Abstracts

Acknowledgement

Many thanks to

- my advisor Prof Felix J. Herrmann
- PhD advisory committee at Georgia Tech and UBC
- PhD examination committee
- My colleagues at SLIM Lab
- Georgia Research Alliance, NSERC, SINBAD consortium
- Nick Moldoveanu from Schlumberger
- Developers of open source software packages (Devito, JUDI)