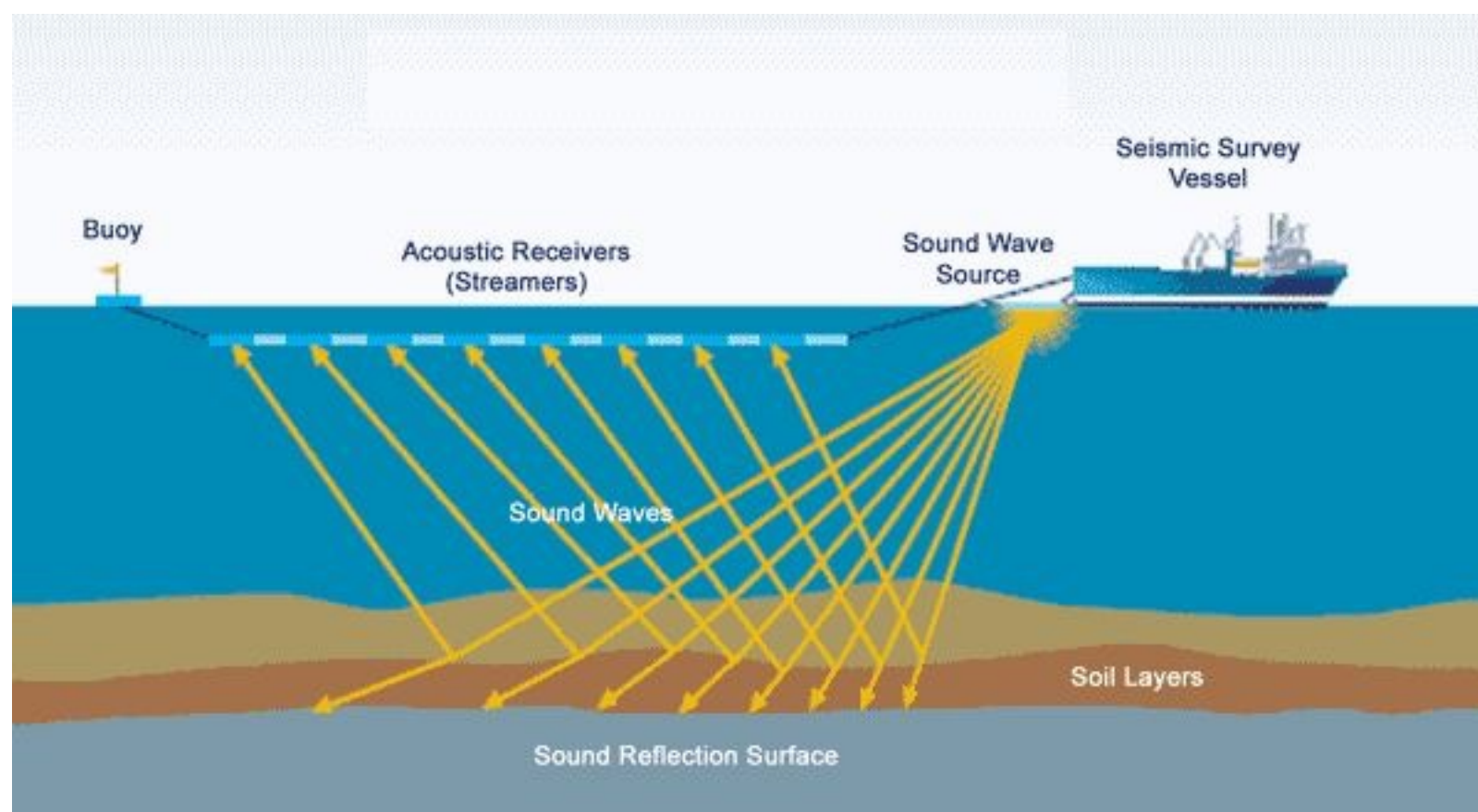


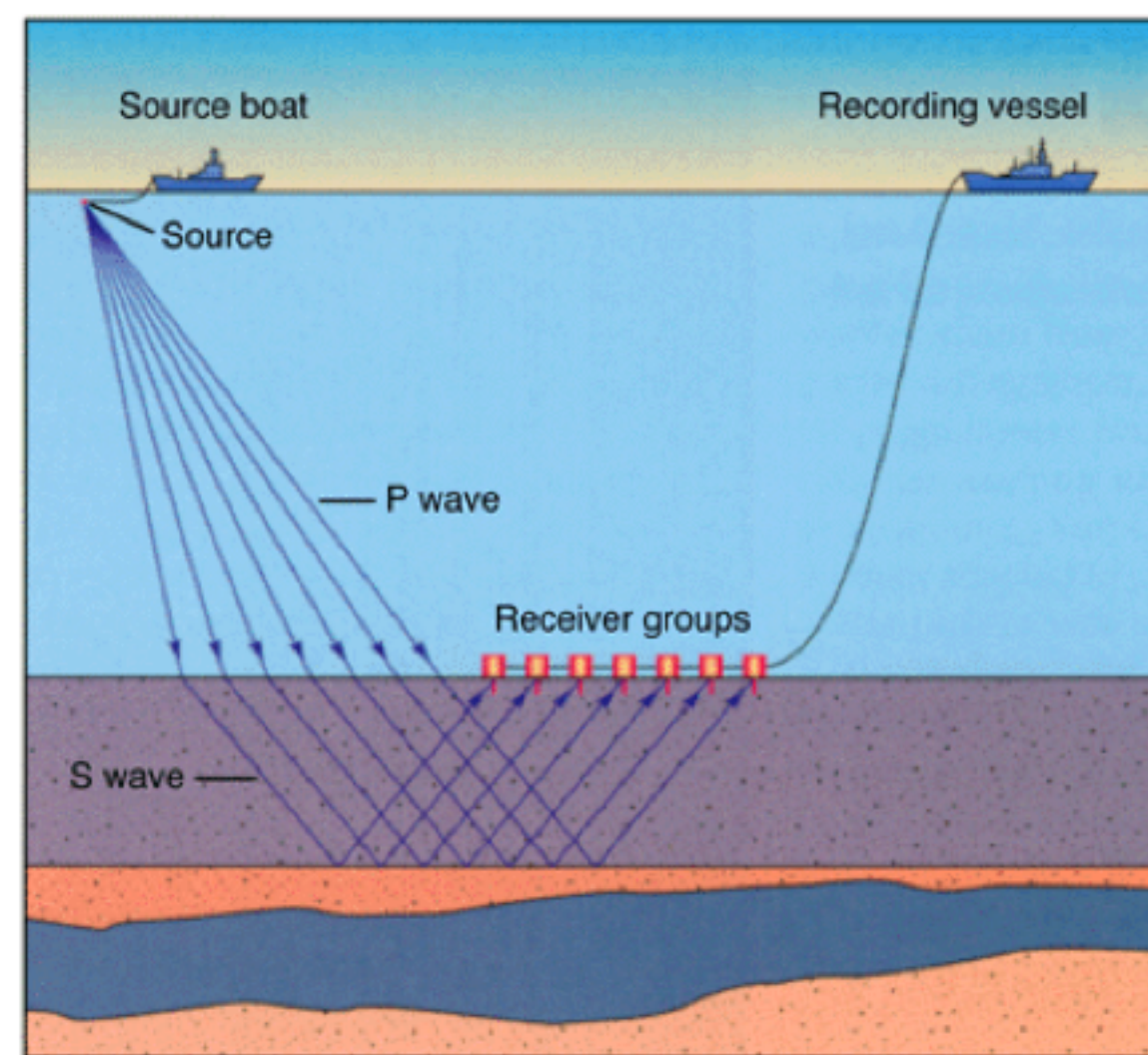
Seismic Imaging with Extended Image Volumes and Source Estimation

Mengmeng Yang
PhD Defense of Dissertation - March 11, 2020
Supervised by Dr. Felix J. Herrmann

Marine acquisition



using Hydrophone



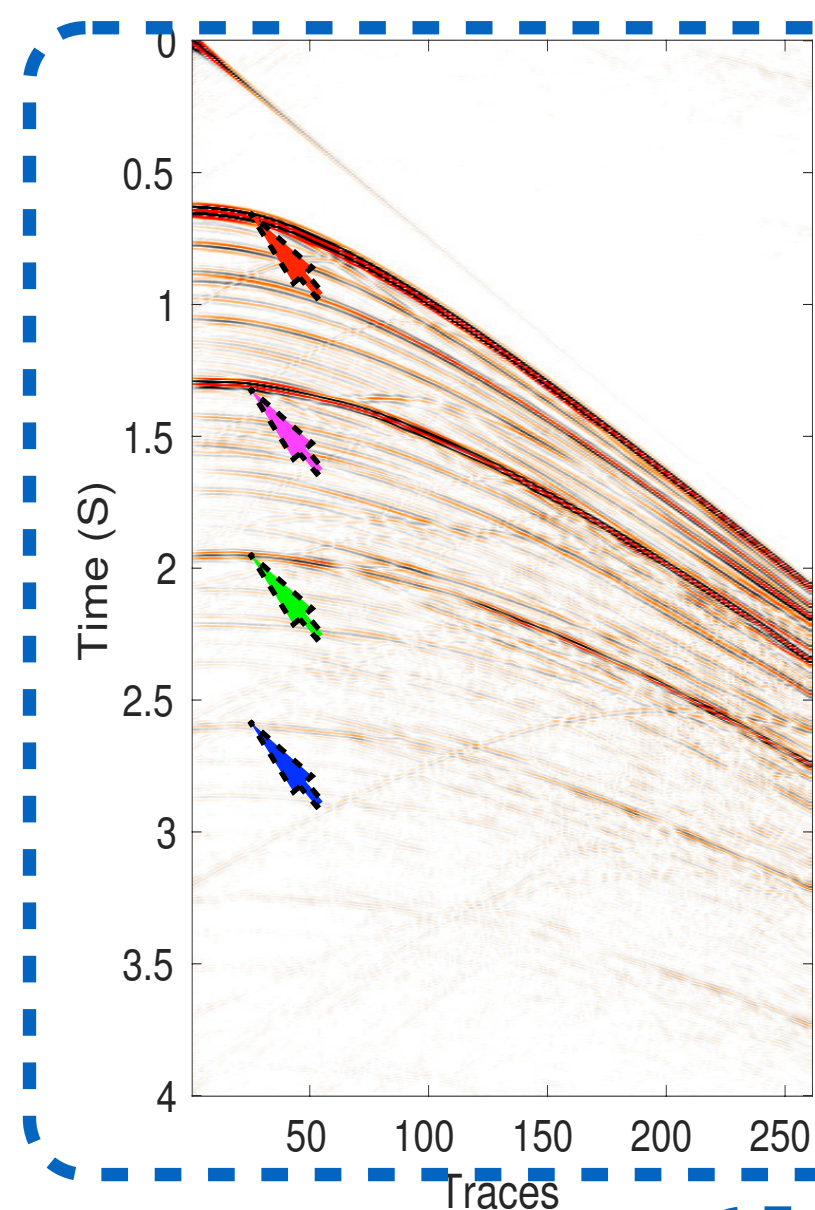
using ocean bottom node

cover 10^2 km^2

subsurface gridpoint $10^6 - 10^9$

10^8 traces

Workflow



Demultiple

Deghost

Starting background model

Processed data

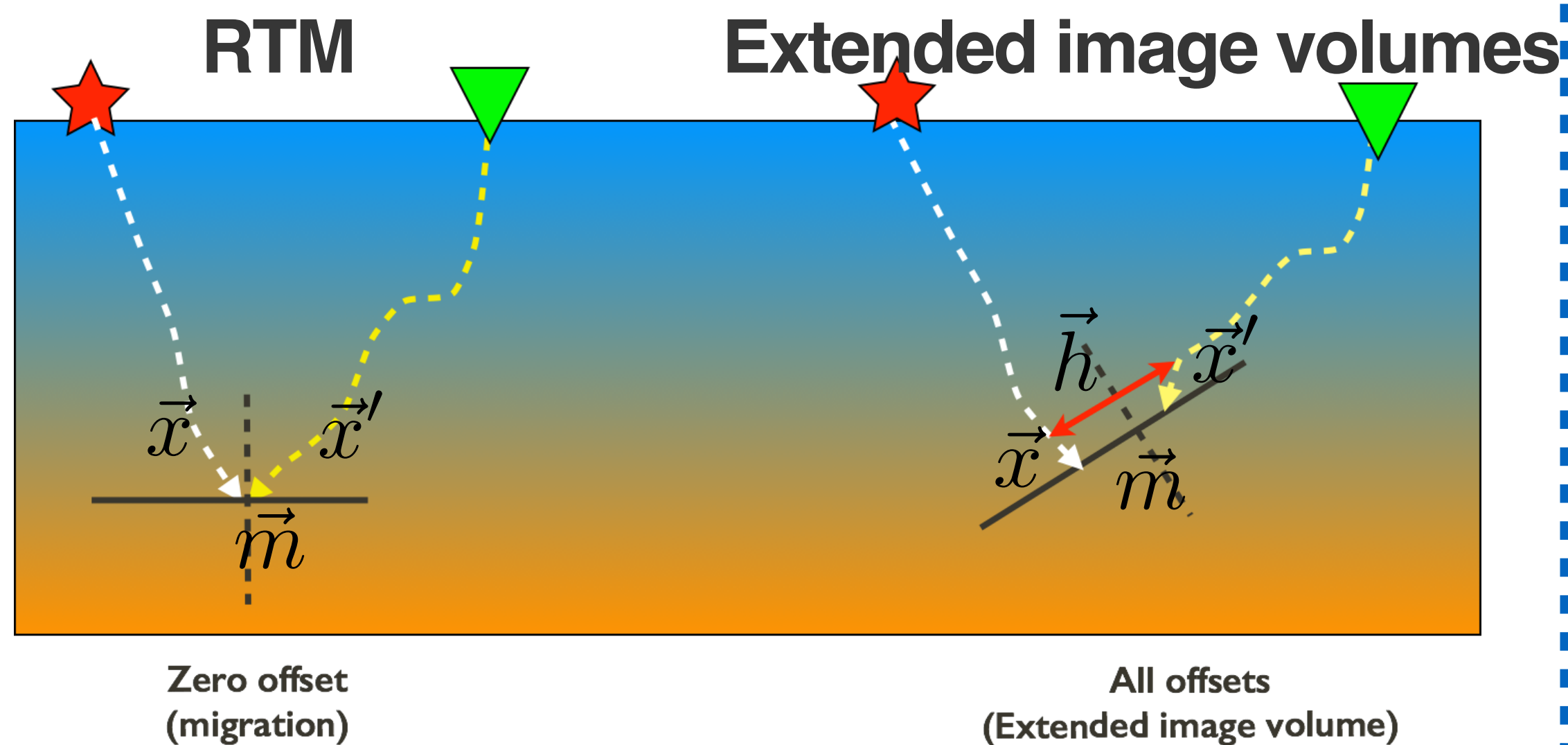
Migration

Wavelet estimation

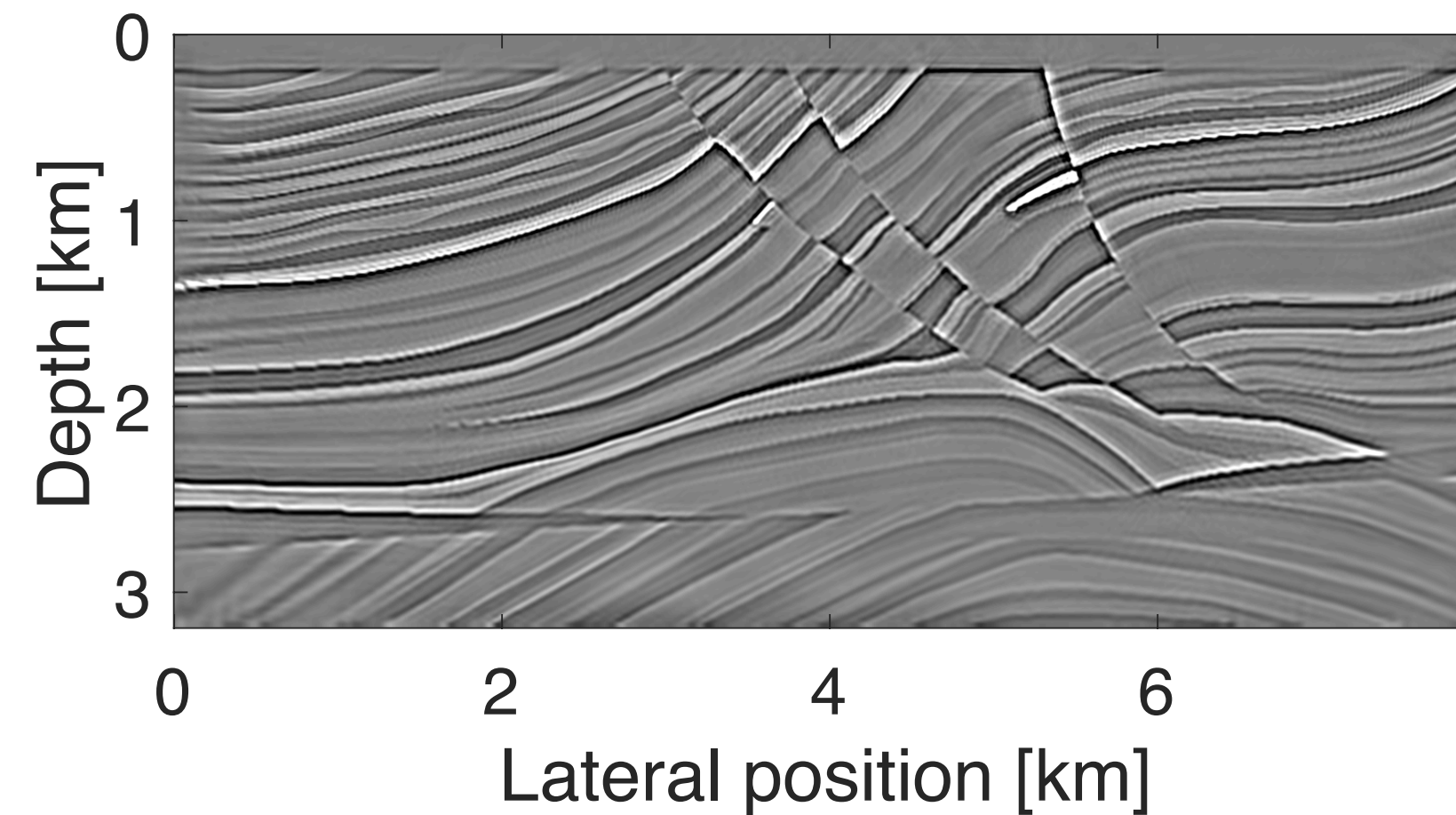
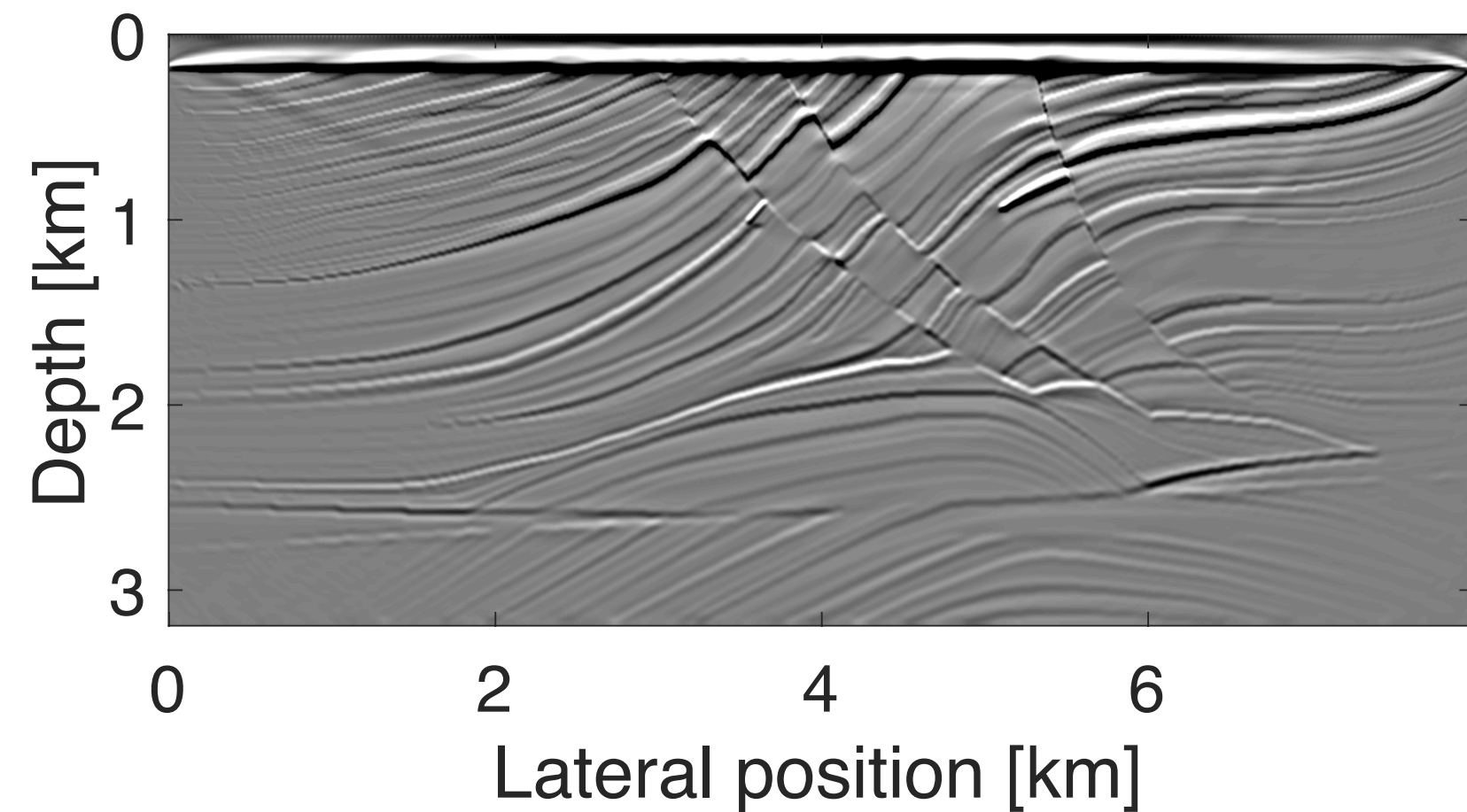
Tomographic inversion, FWI

Updated background model

Final background model



Imaging inversion & wavelet influence

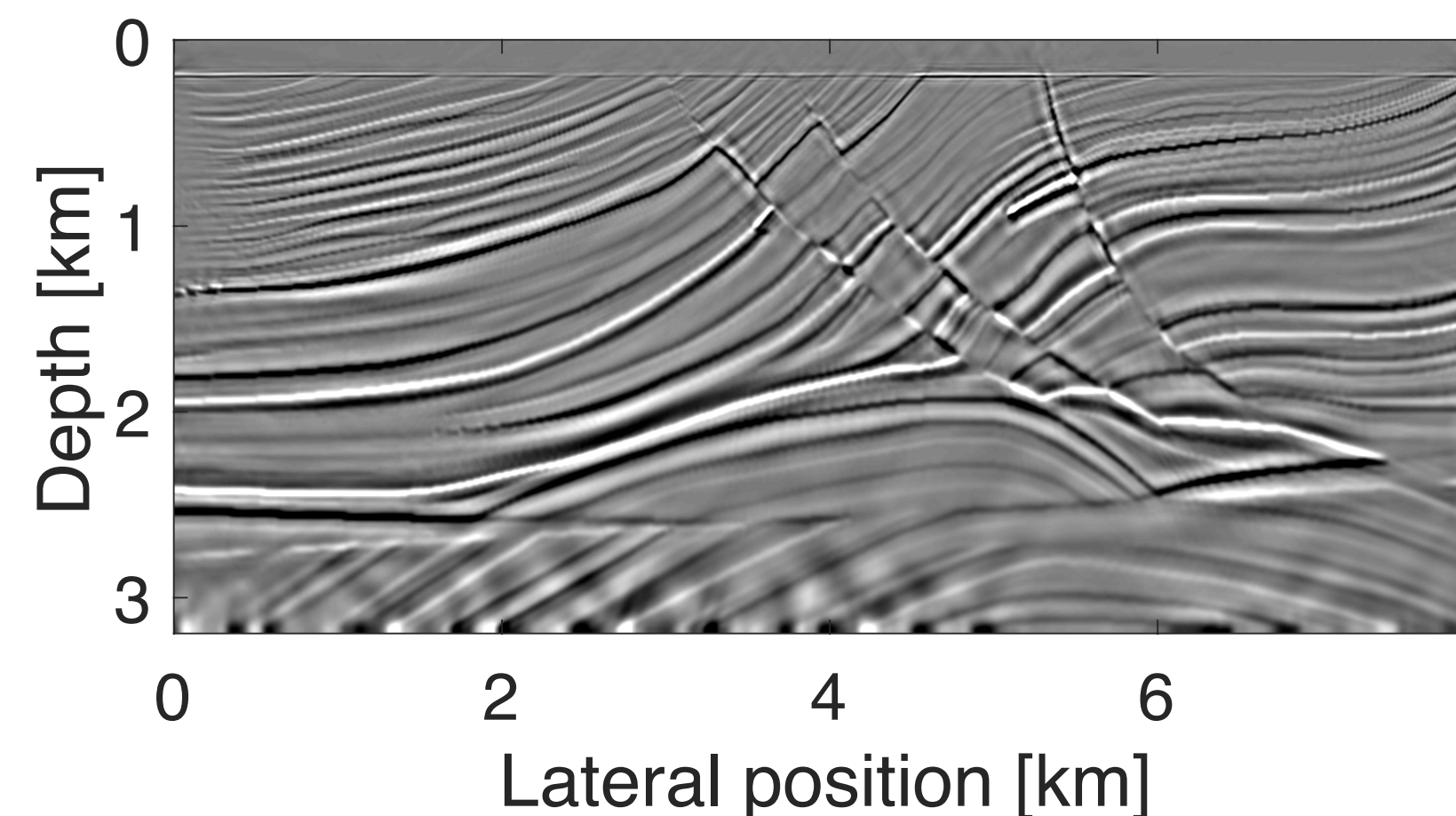


From imaging to inversion

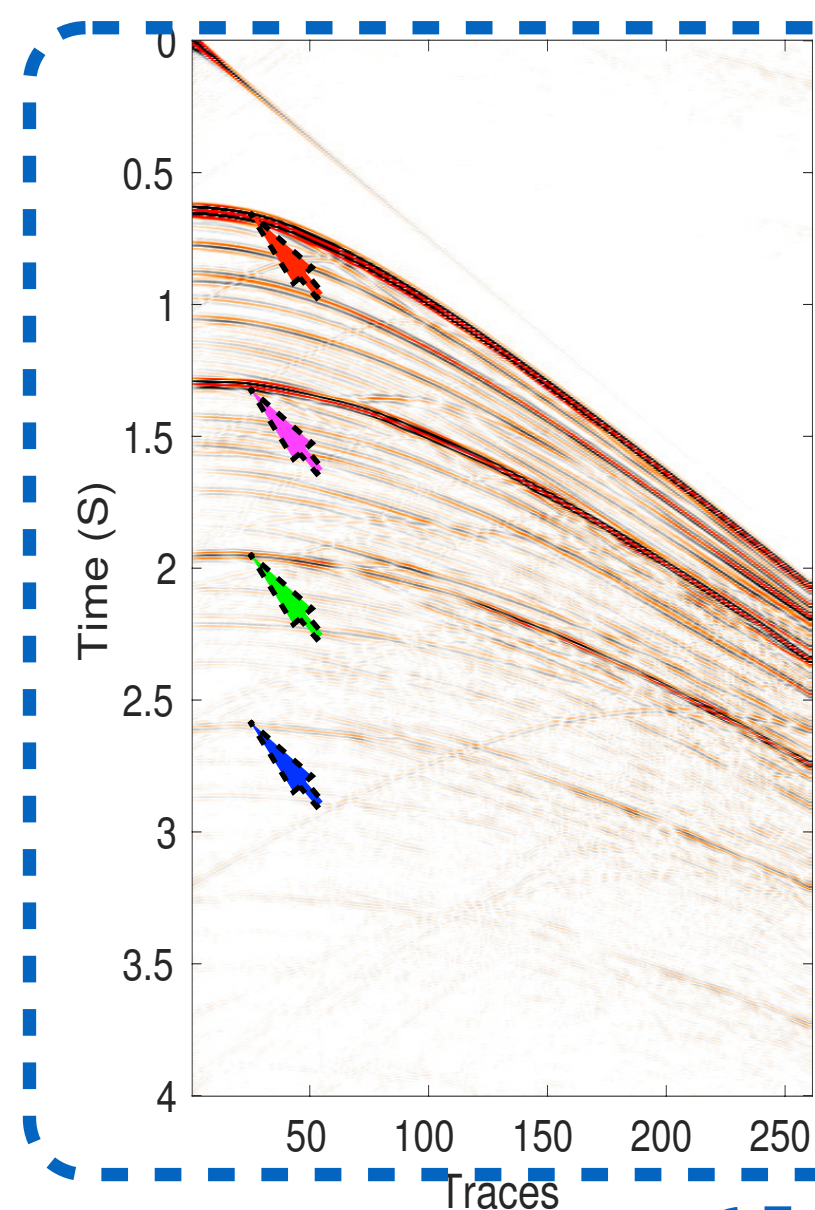
- remove source and receiver imprint
- remove band-limited effect from source

However

- expensive cost in iterations
- prior knowledge of source



Workflow



Demultiple

Deghost

Starting background model

Processed data

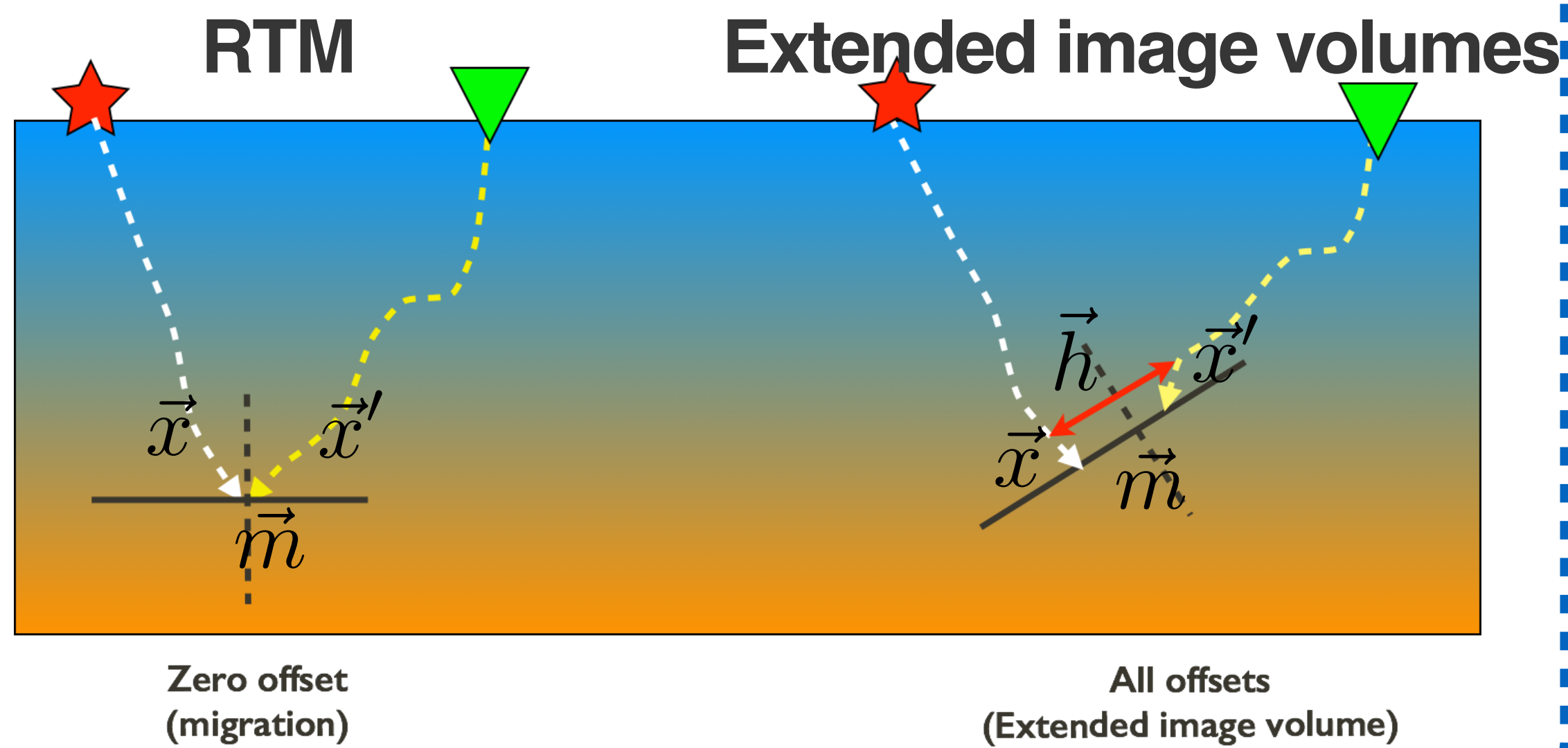
Migration

Wavelet estimation

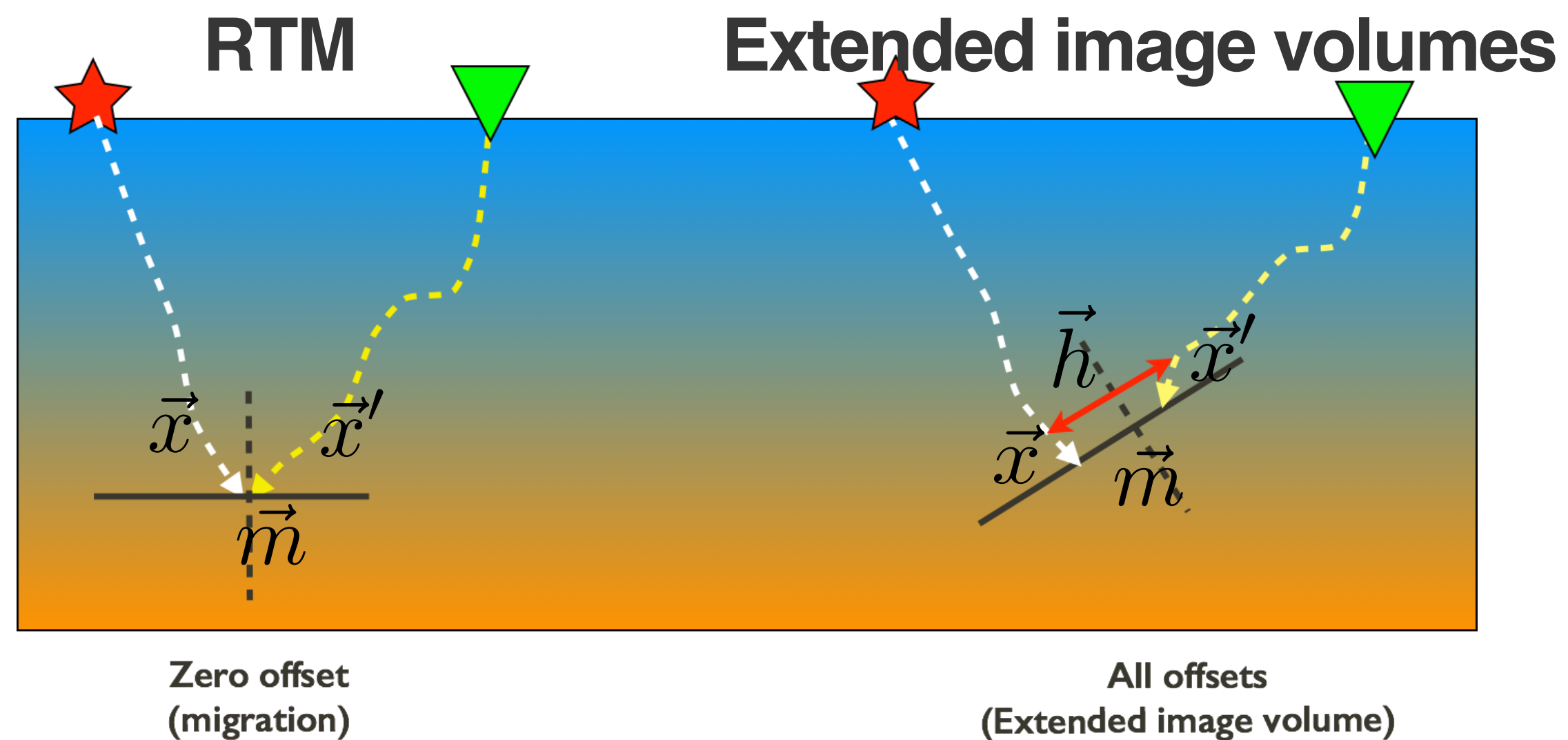
Tomographic inversion, FWI

Updated background model

Final background model



Rich but expensive Extended image volumes



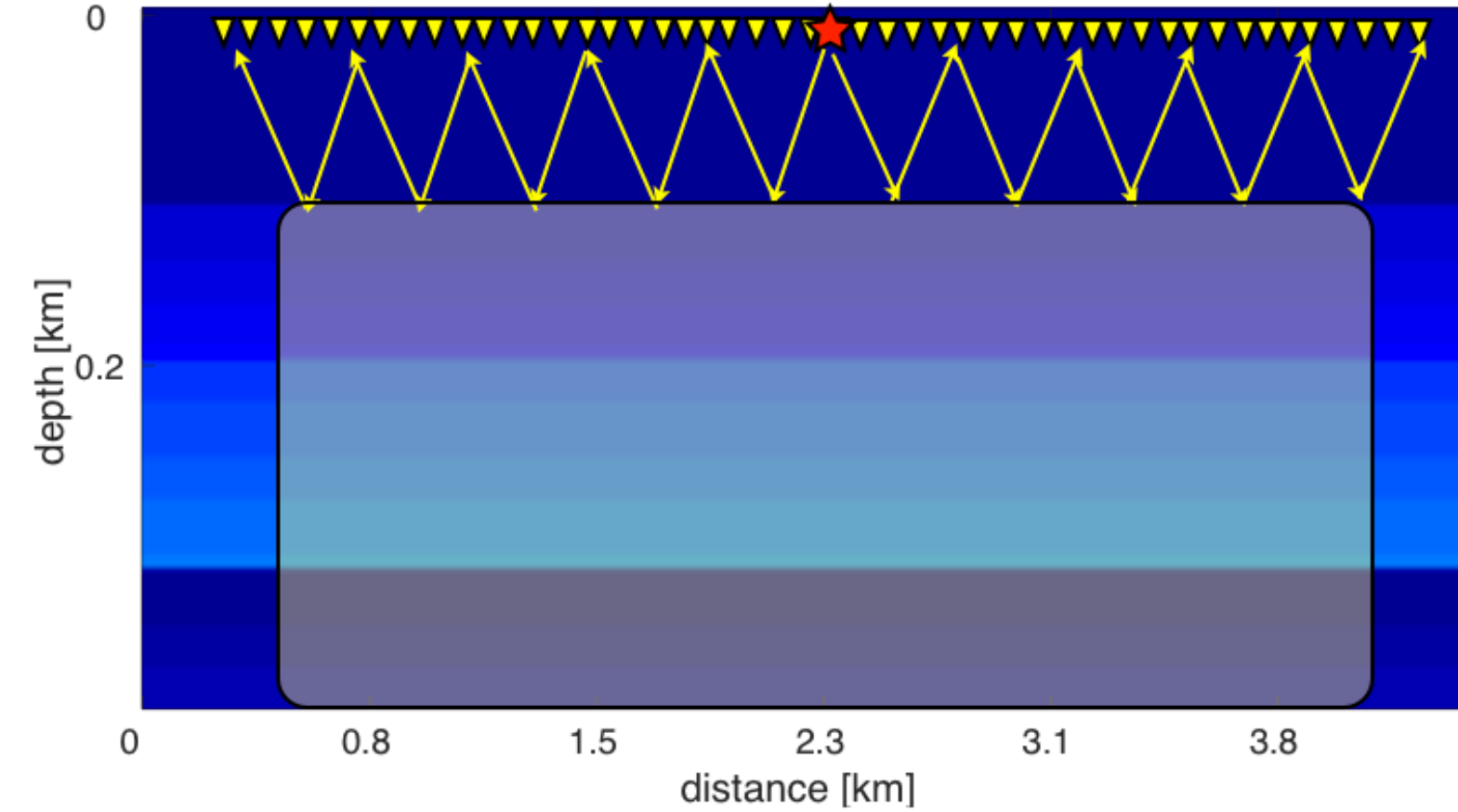
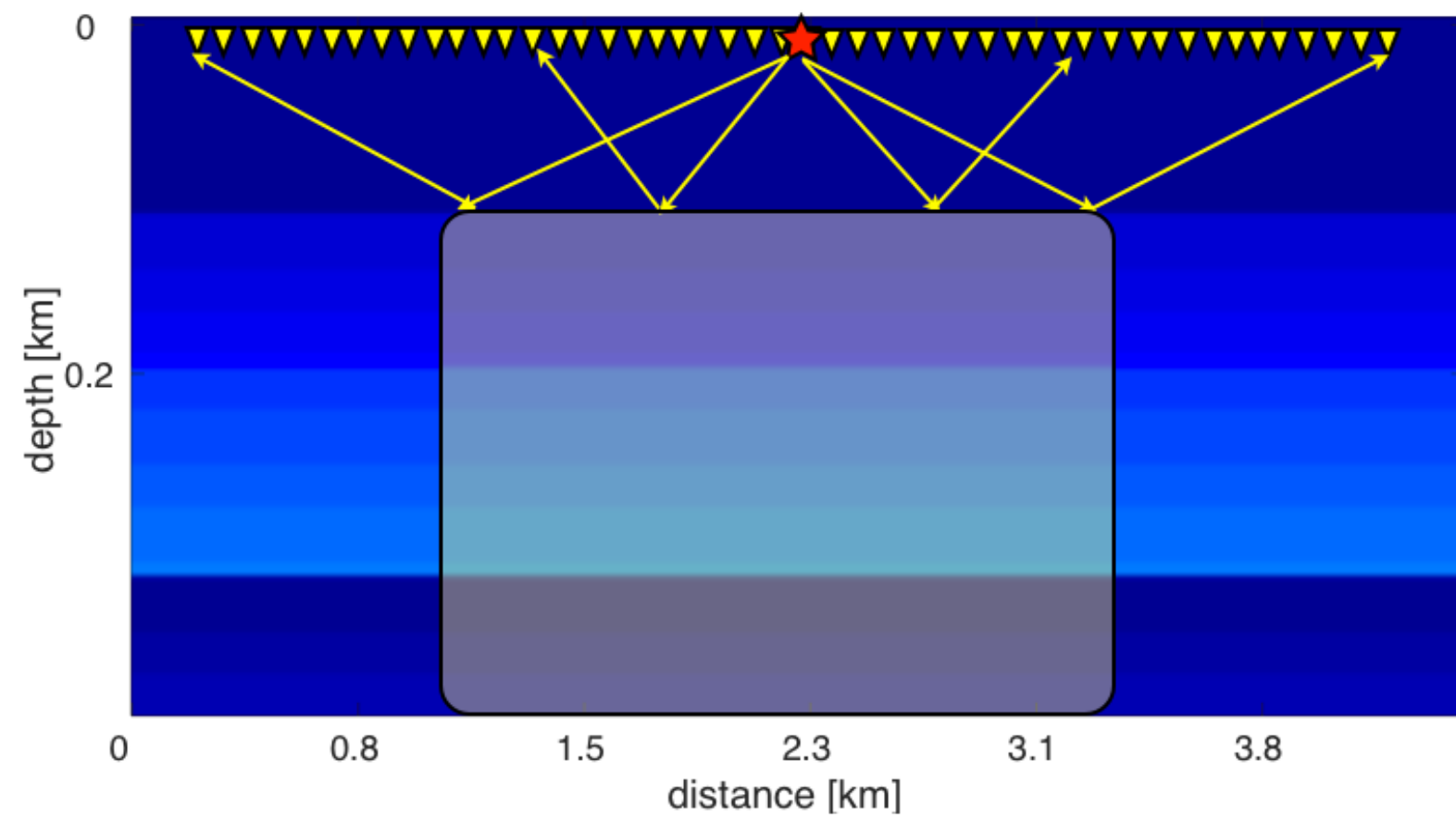
Information can be used to

- create images
- infer rock properties
- QC background velocity model

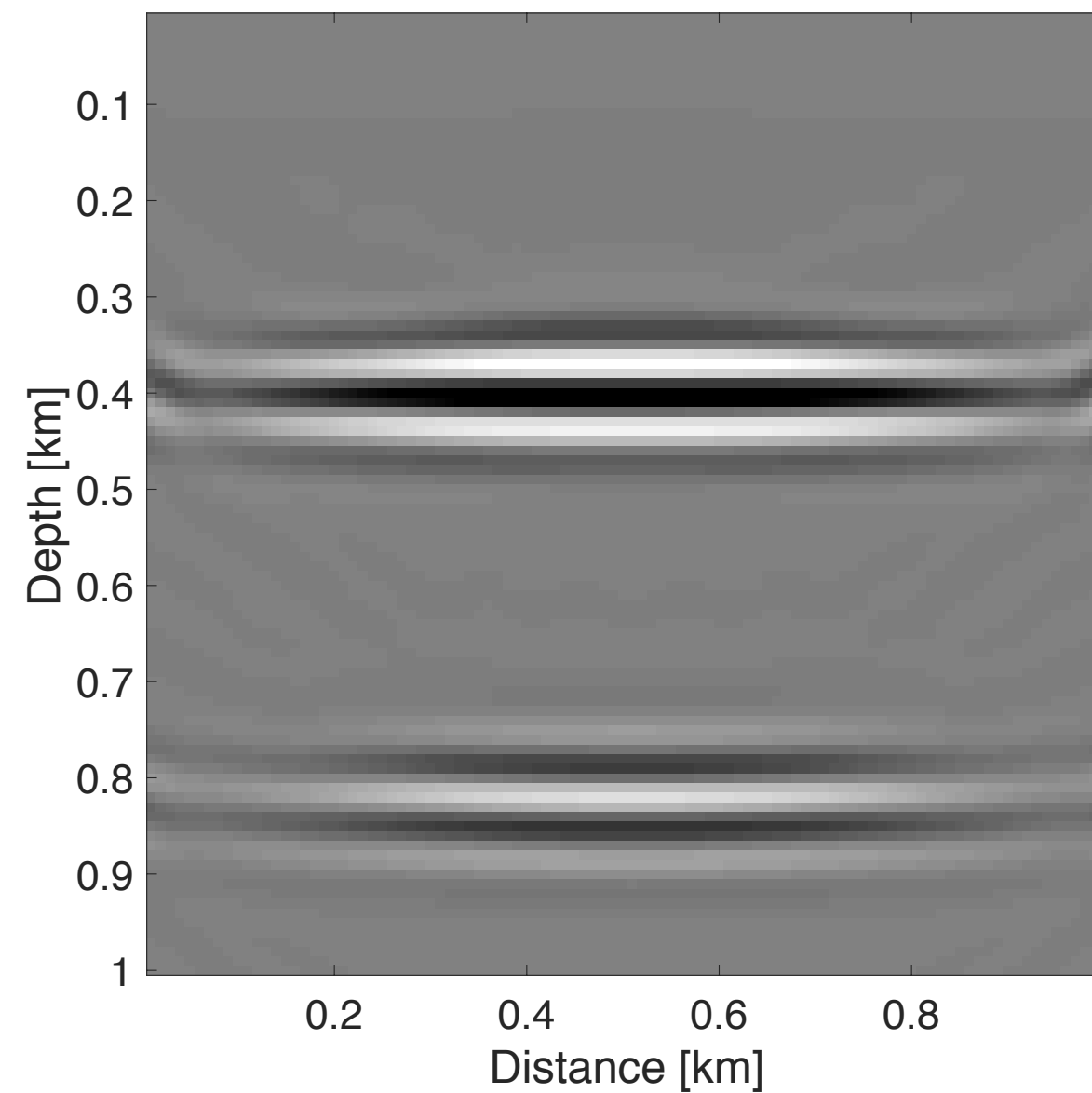
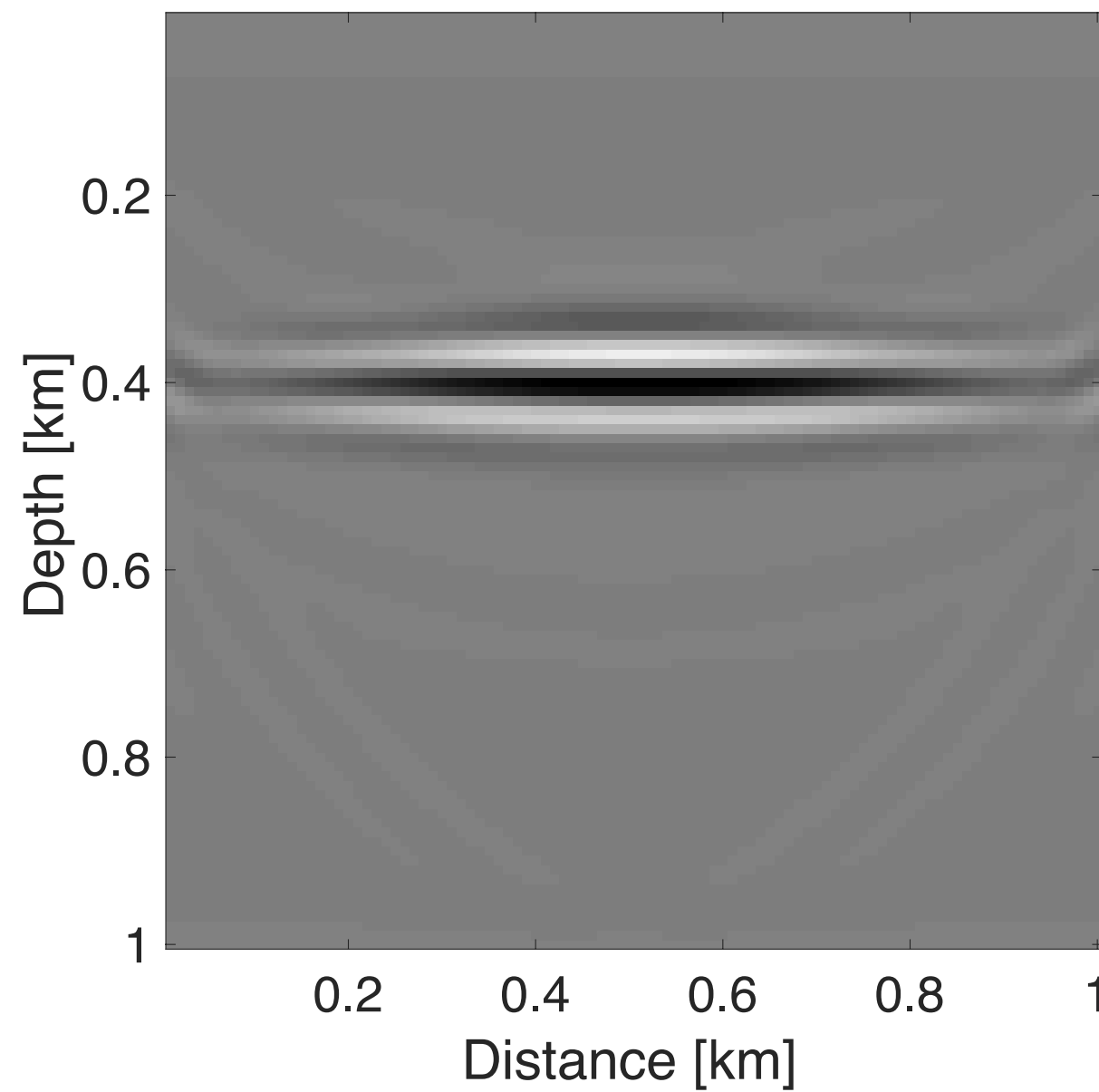
Expensive in memory

Expensive computations involving wave-equation solves that scale with $2n_s$

Imaging w/ multiples



- Multiples elimination removes illumination
- Naive usage of multiples in imaging introduces cross-talk



Purpose of thesis

Outline:

- **Chapter 2-3:** Low-rank recovery for subsurface extended image volumes based on time-stepping propagator and power schemes, velocity continuation via invariance relationship
- **Chapter 4:** Source estimation for time-domain sparsity promoting least-squares reverse-time migration (LS-RTM)
- **Chapter 5-6:** Sparsity promoting least-squares reverse-time migration with multiples, and removing density effects in least square reverse-time migration with matched low-rank filter

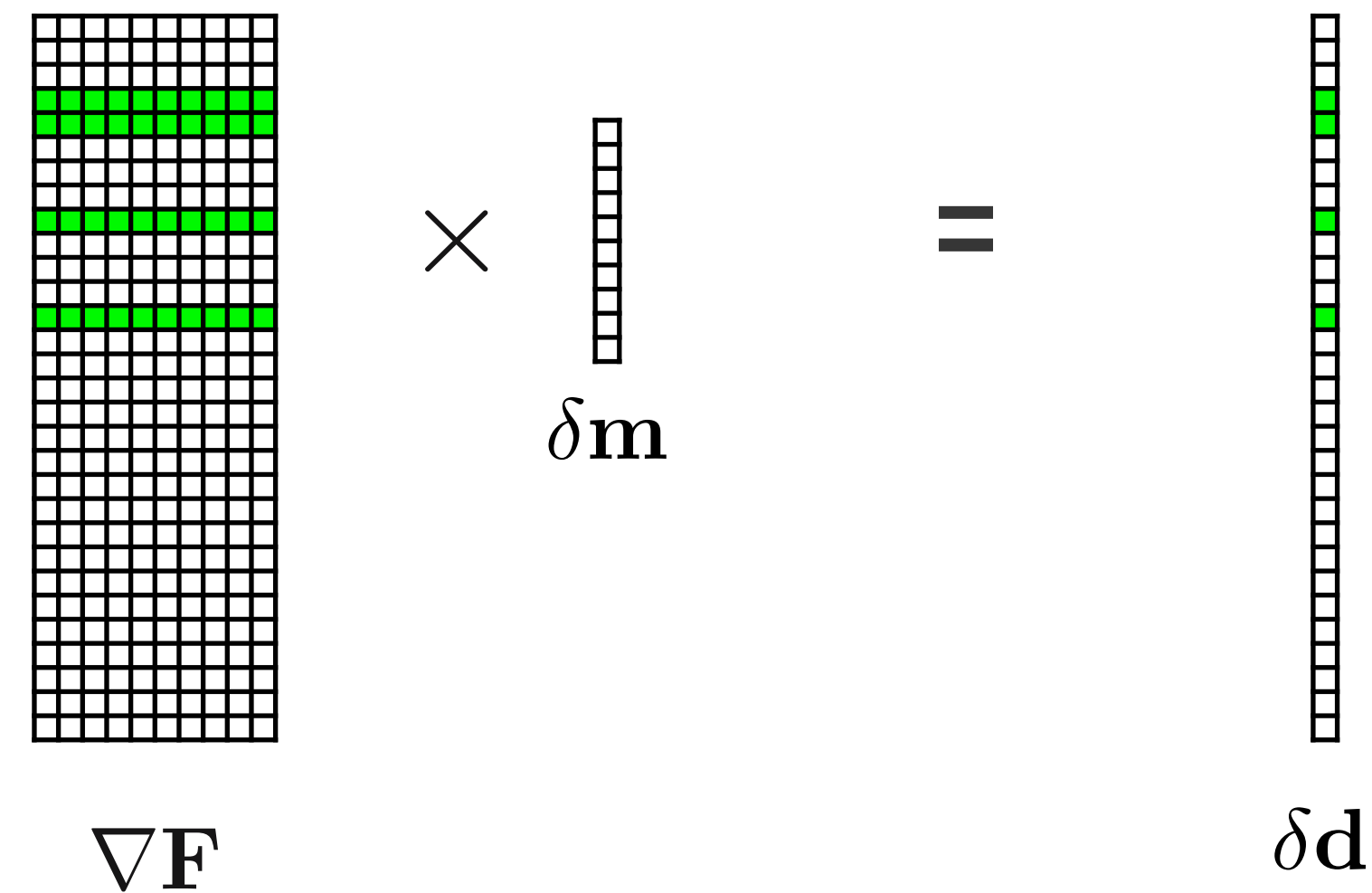
Chapter 4 Source estimation for time-domain sparsity promoting least-squares reverse-time migration (LS-RTM)

From LS-RTM to SPLS-RTM

- In practice, we have to solve the following optimization problem: (LS-RTM)

$$\min_{\delta \mathbf{m}} \frac{1}{2} \sum_{i=1}^{n_s} \|\nabla \mathbf{F}_i(\mathbf{m}_0, \mathbf{q}) \delta \mathbf{m} - \delta \mathbf{d}_i\|^2.$$

- Very large overdetermined system
- Computationally expensive in each iteration



From LS-RTM to SPLS-RTM

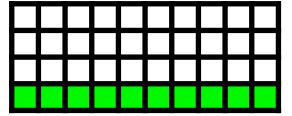
- **Dimensional reduction + sparsity promoting optimization:**

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

subject to $\sum_{i=1}^{n_s} \|\nabla \mathbf{F}_i(\mathbf{m}_0, \mathbf{q}) \mathbf{C}^T \mathbf{x} - \delta \mathbf{d}_i\|_2 \leq \sigma$


Tolerance for noise

Curvelet transform




$\nabla \mathbf{F}_k \mathbf{C}^T$

×

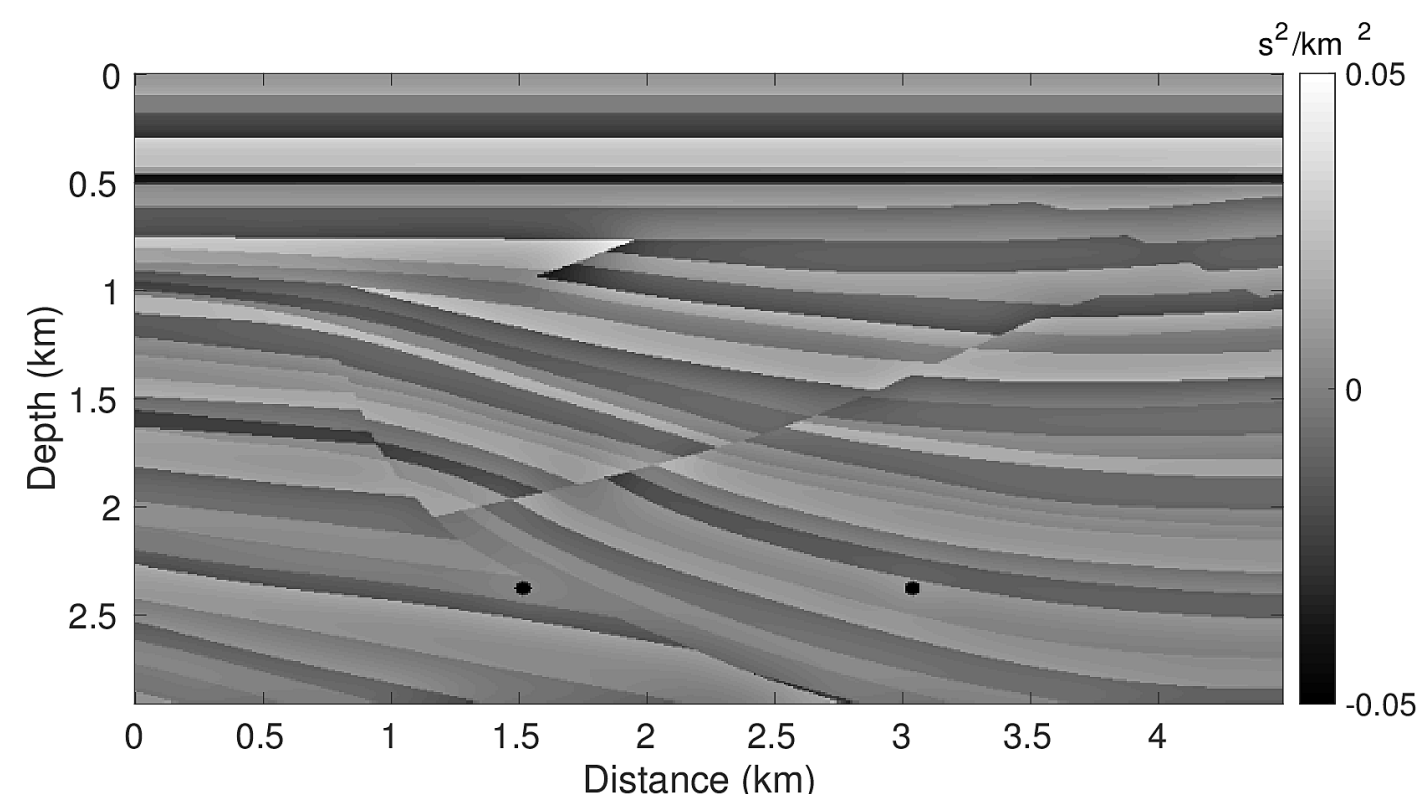


\mathbf{x}

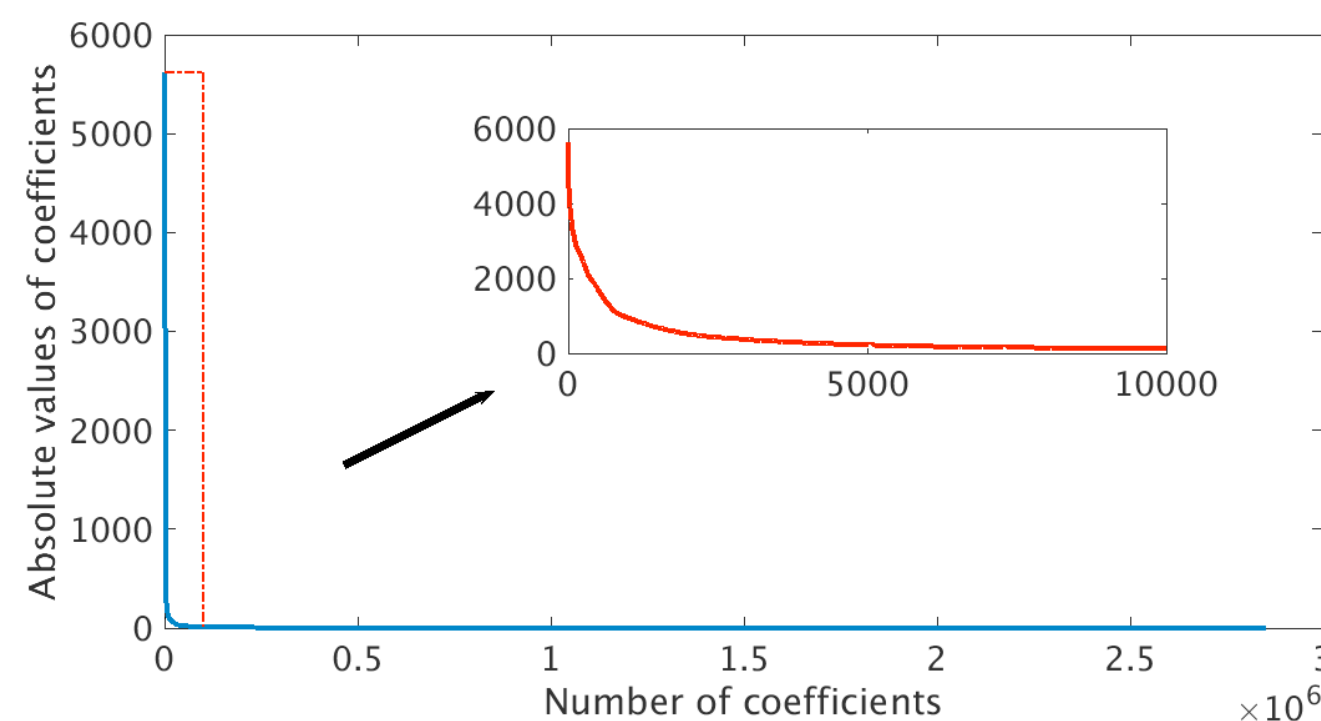
=



$\delta \mathbf{d}_k$



Model perturbation

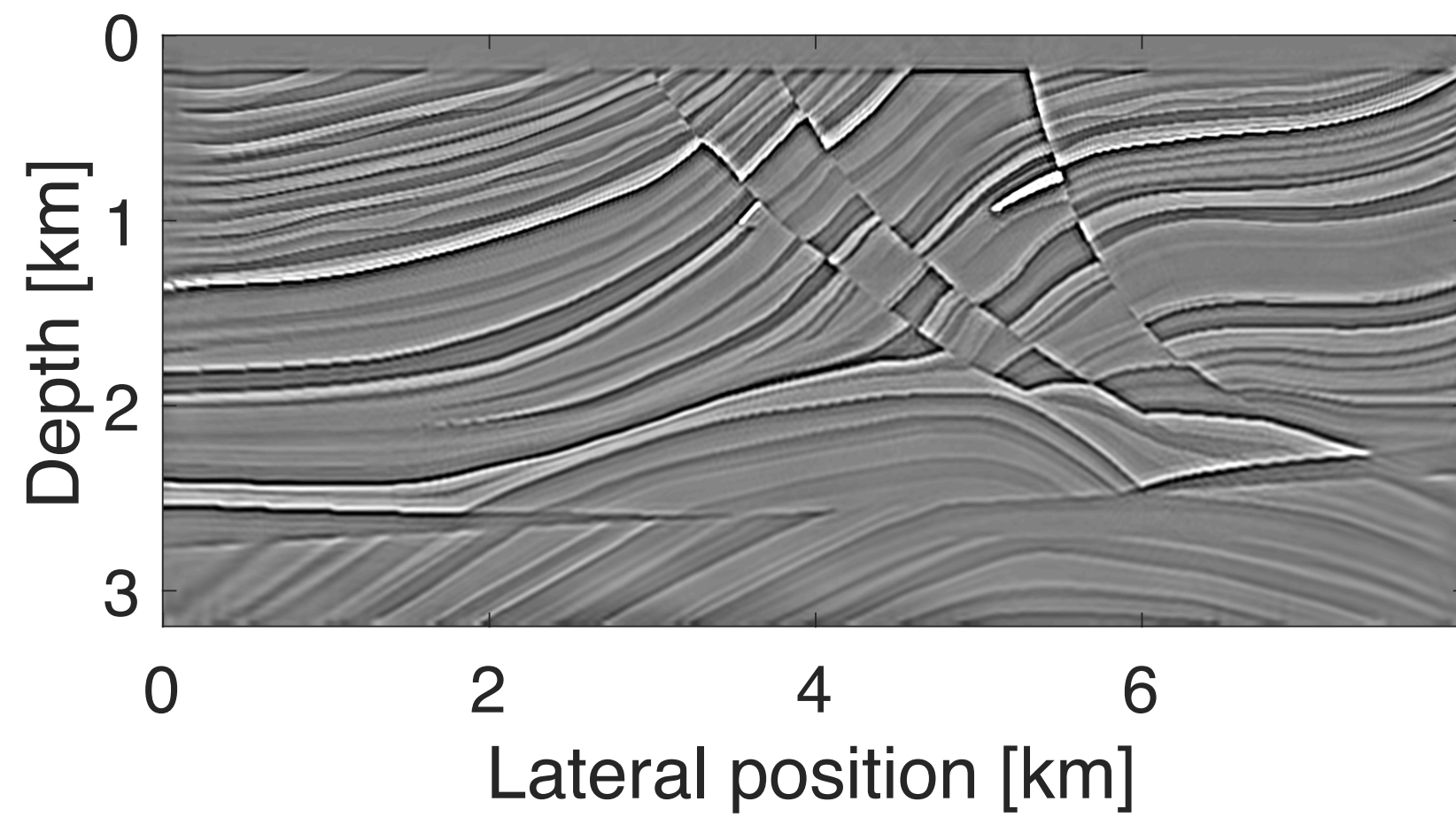


Sorted Curvelet

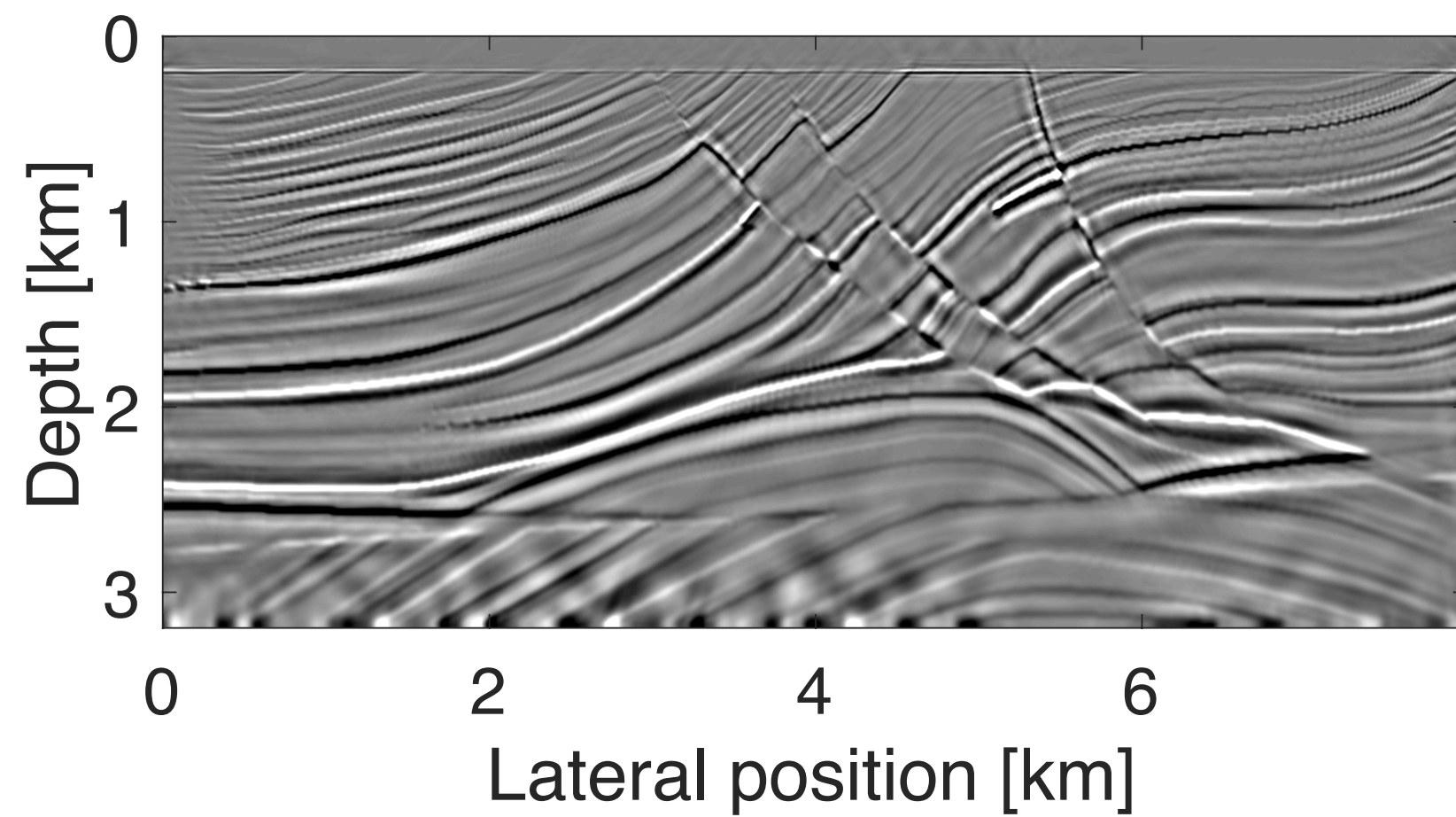
Still need prior information of source function

Lorenz D. et al, "The linearized bregman method via split feasibility problems: Analysis and generalizations", SIAM, 2014
 Witte P. et al, "Compressive least-squares migration with on-the-fly fourier transforms", Geophysics, 2019

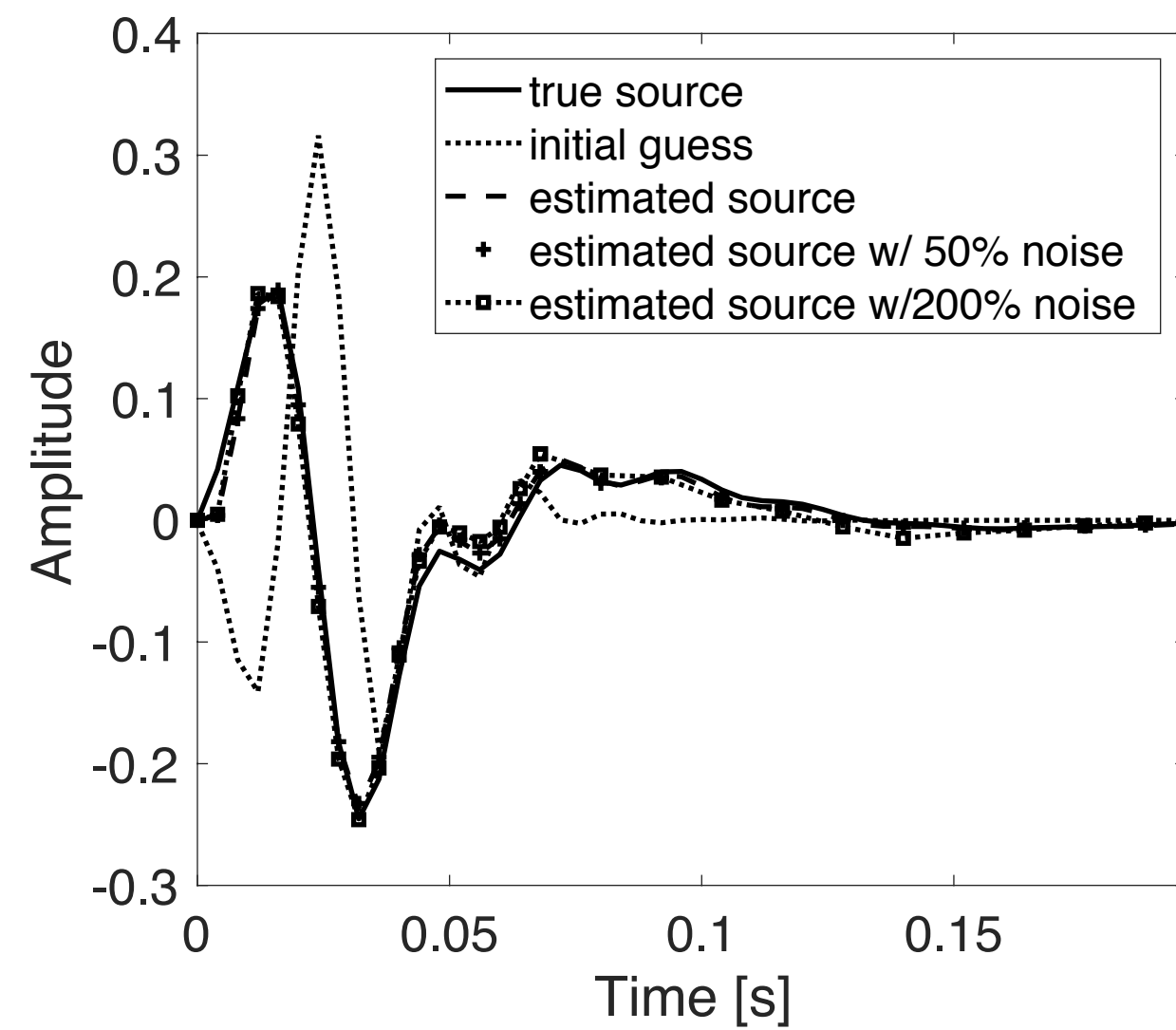
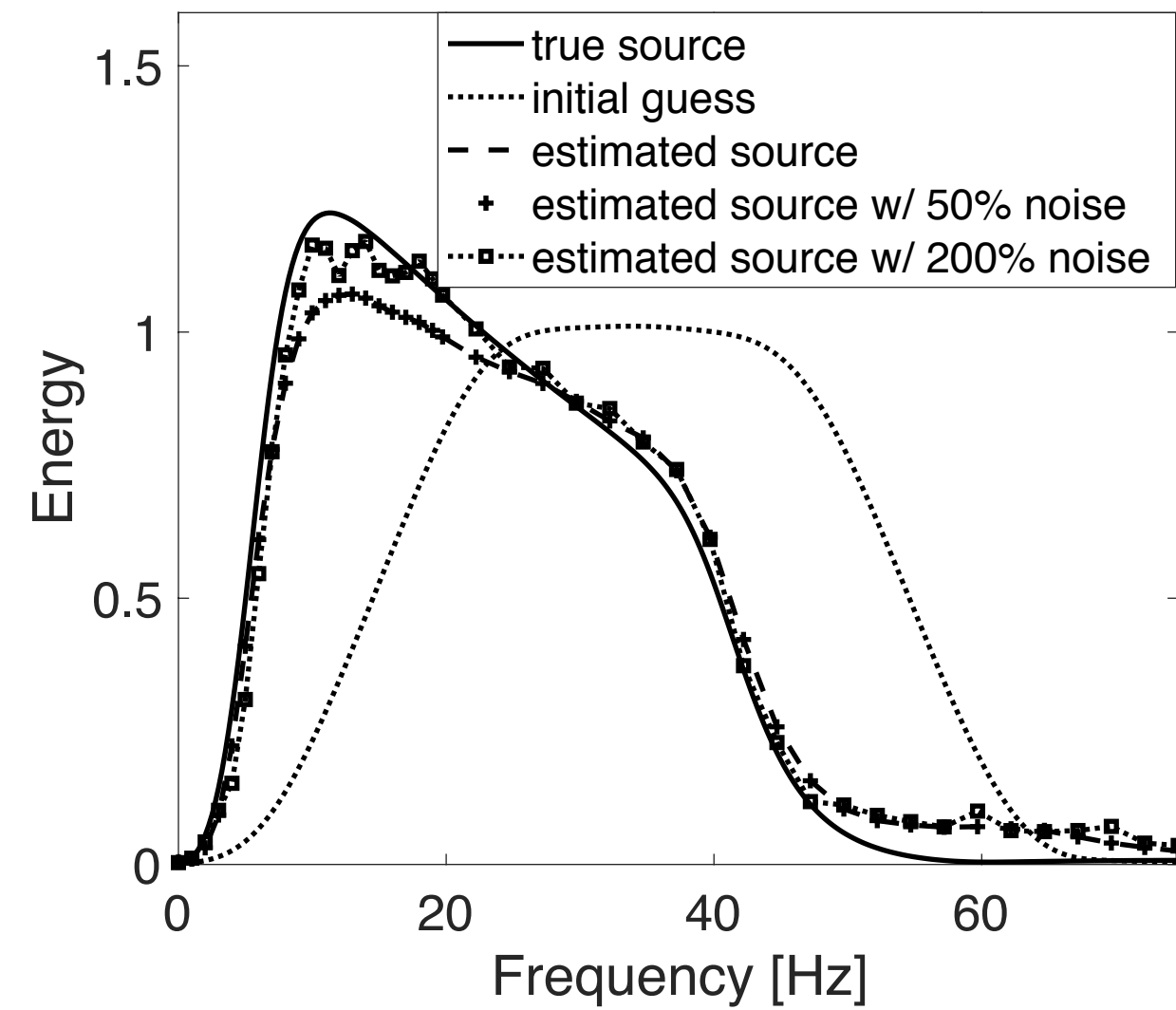
Influence of wavelet - Marmousi model, linear data



one data pass w/ true source



one data pass w/ wrong source



Source estimation

Solutions: deconvolution + penalization

Let: $q = w * q_0$, q_0 is well defined.

Property 1:

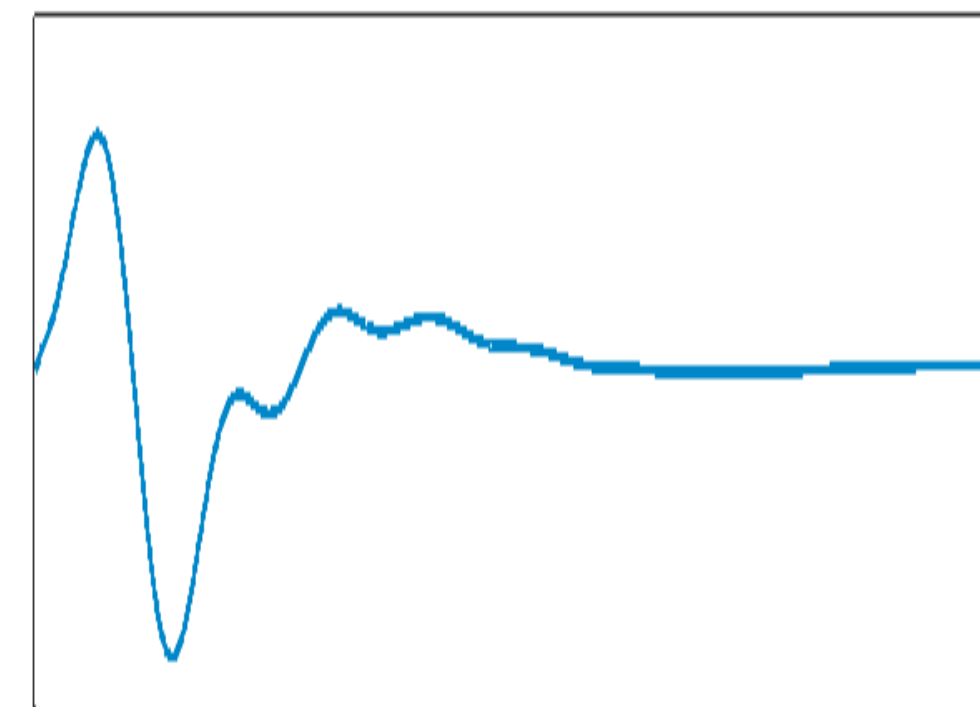
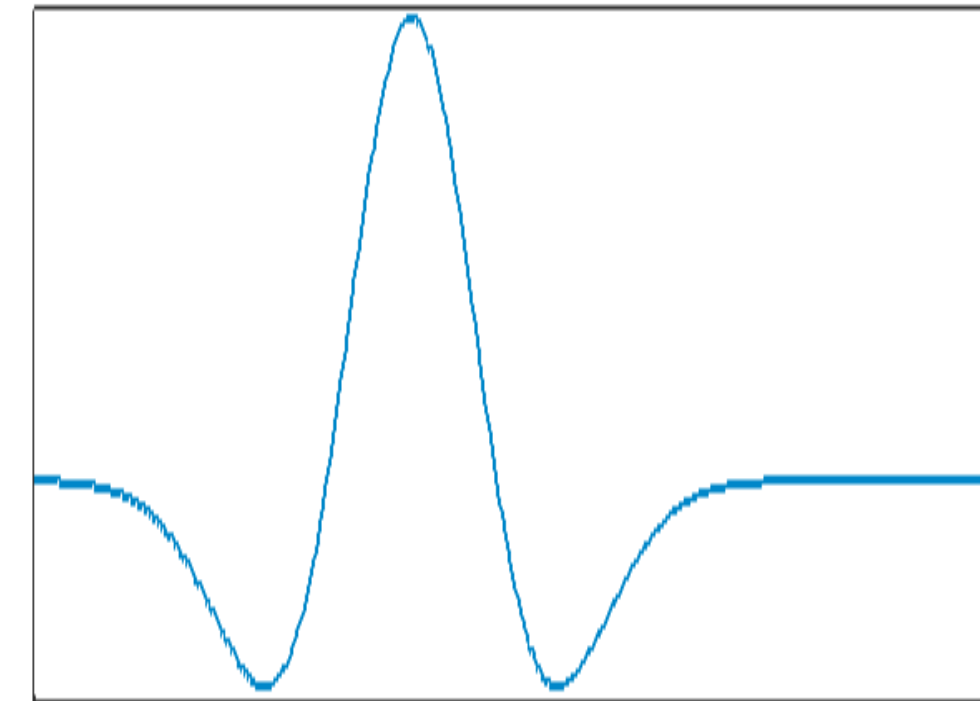
$$\nabla F_i(\mathbf{m}_0, w * q_0) \mathbf{C}^T \mathbf{x} = w * \nabla F_i(\mathbf{m}_0, q_0) \mathbf{C}^T \mathbf{x}$$

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

Time-domain SPLS-RTM w/ on-the-fly source estimation

New subproblem:

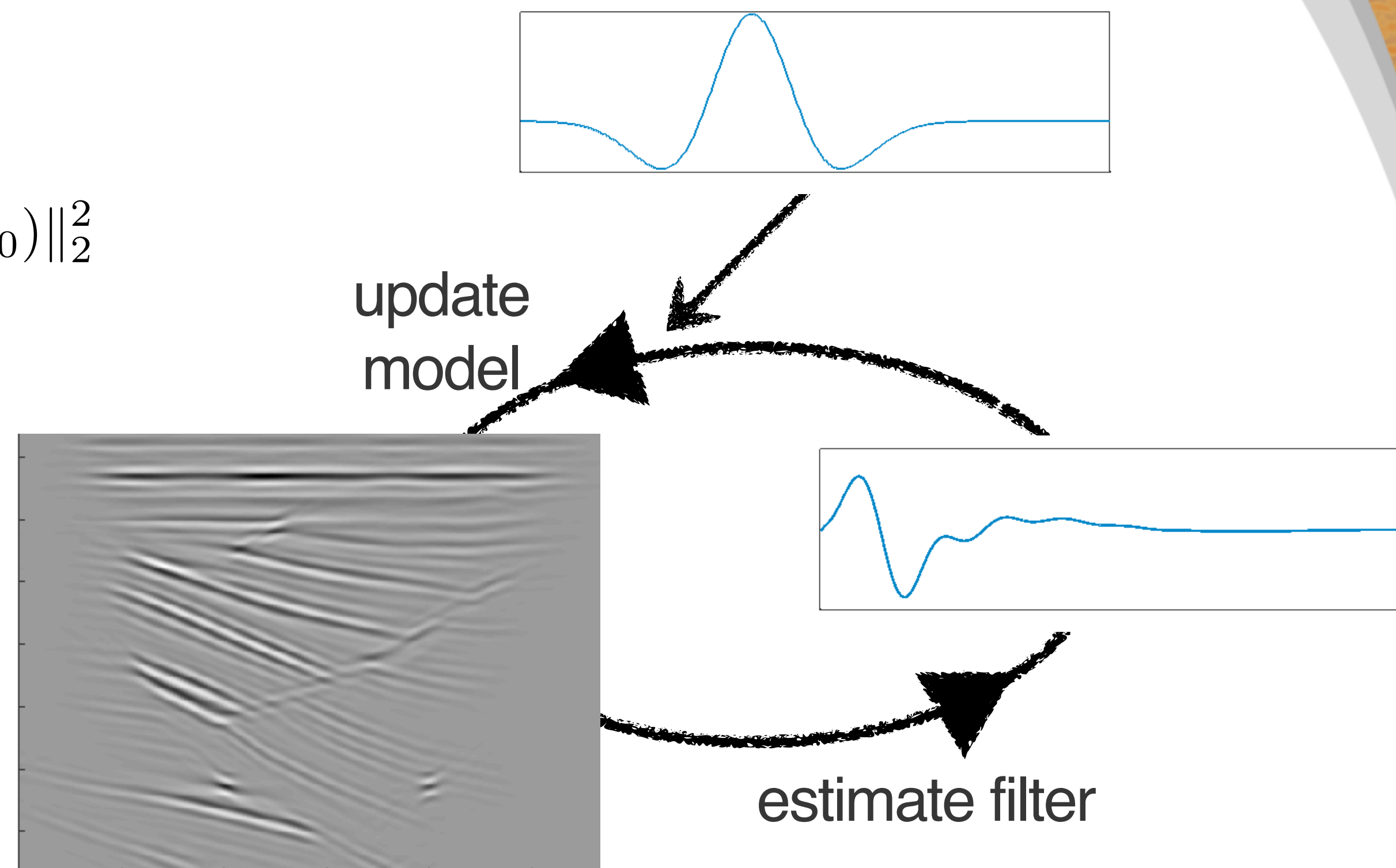
$$\min_{\mathbf{w}} \sum_{i \in \mathcal{I}_k} \|\mathbf{w} * \nabla \mathbf{F}_i(\mathbf{m}_0, \mathbf{q}_0) \mathbf{C}^T \mathbf{x} - \delta \mathbf{d}_i\|_2^2 + \|\mathbf{r} \odot (\mathbf{w} * \mathbf{q}_0)\|_2^2$$

with weights

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

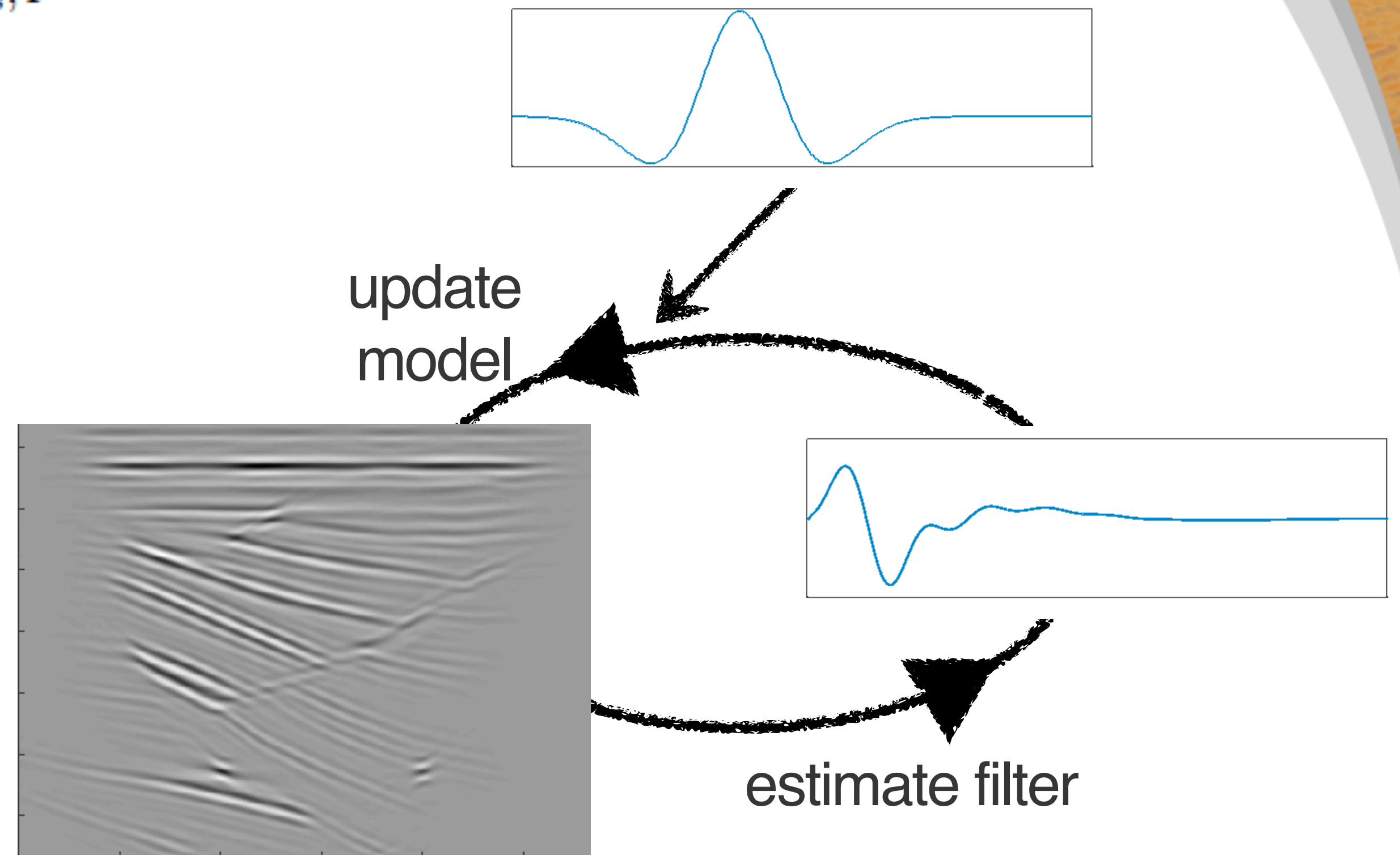
Property 3

Property 2

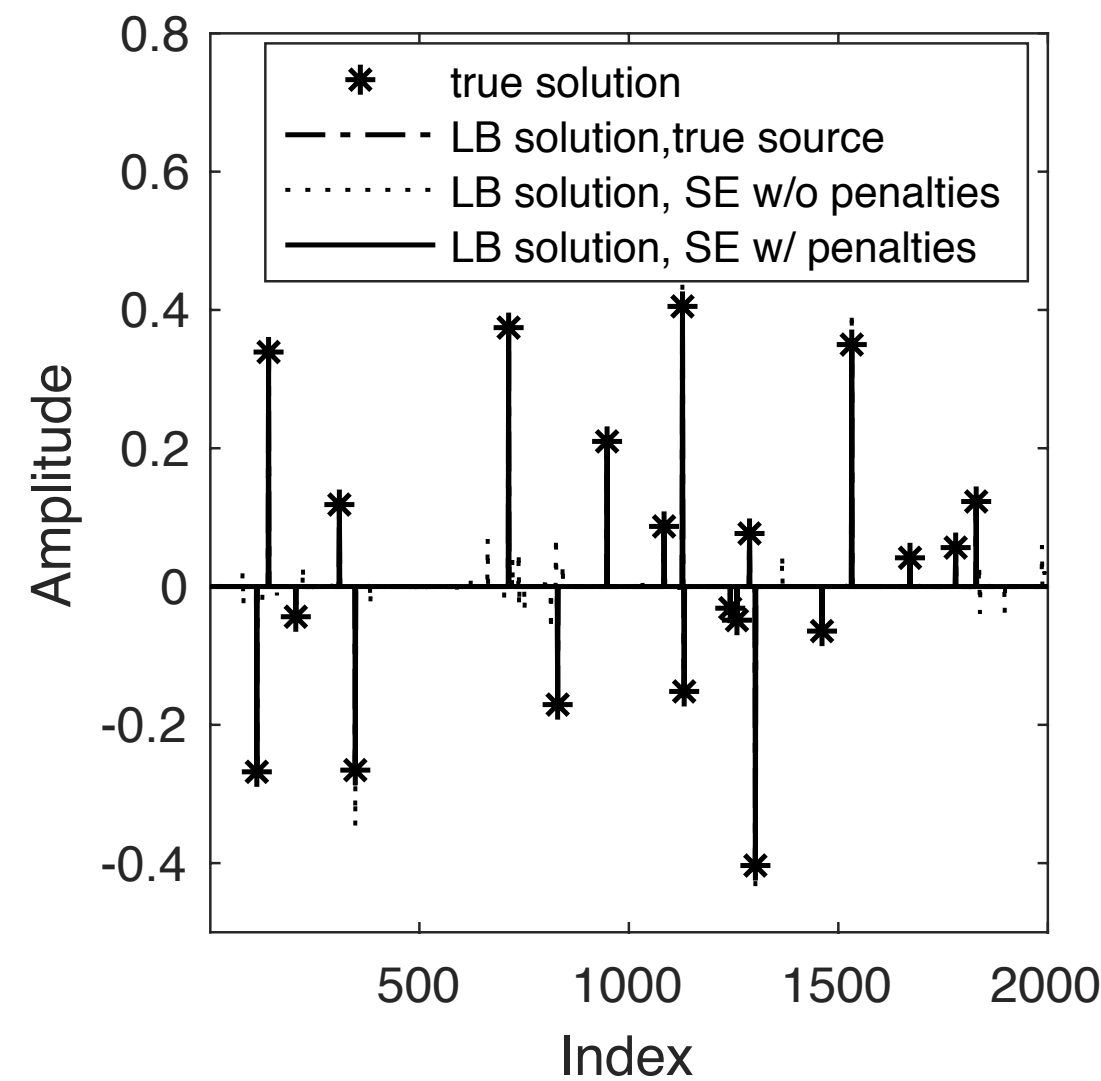


Workflow - LB + on-the-fly source estimation

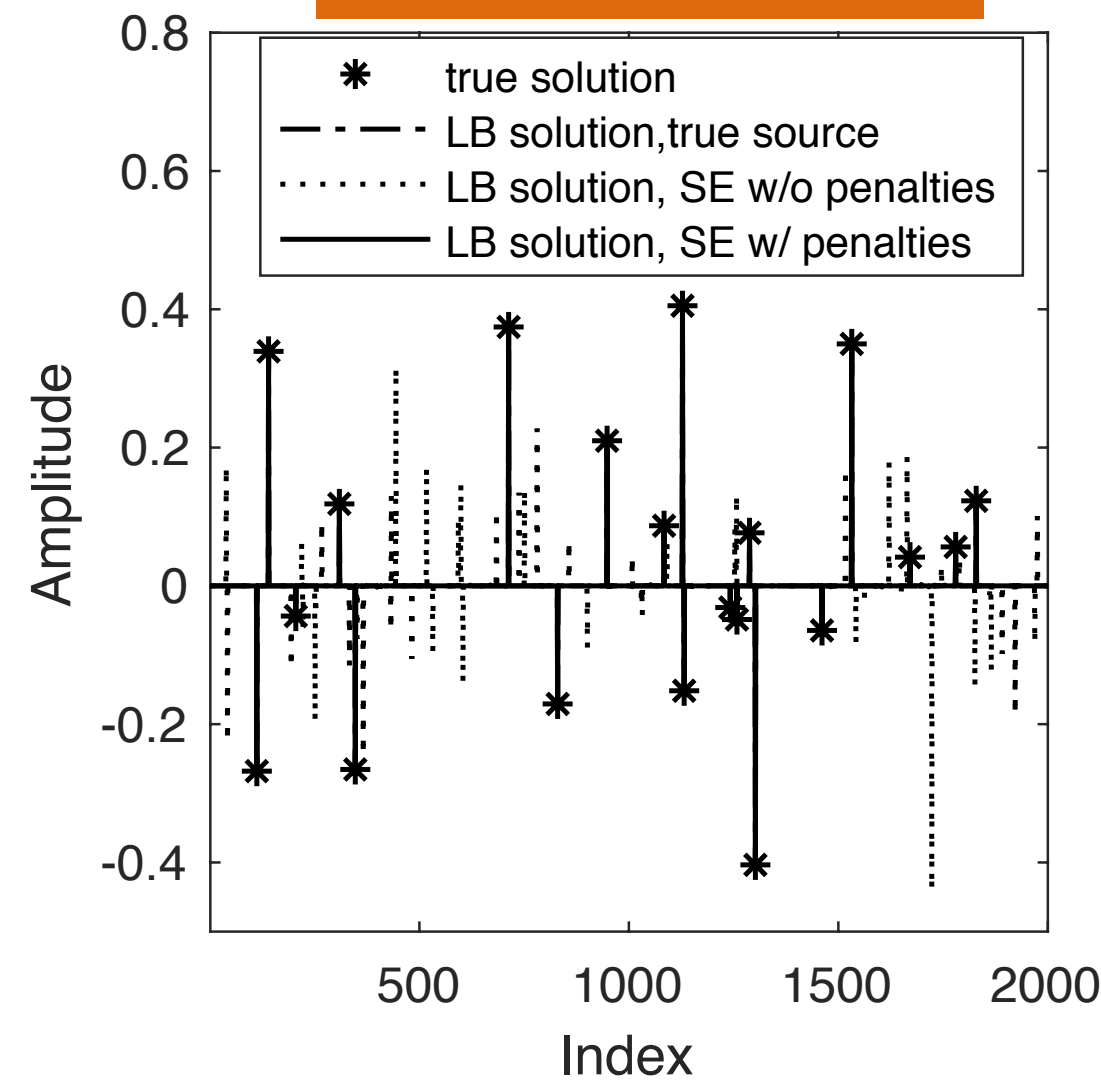
- 1: Initialize $\mathbf{x}_0 = \mathbf{0}, \mathbf{z}_0 = \mathbf{0}, \mathbf{q}_0, \lambda_1, \mathbf{w}_0 = \delta, \nu$, batch size $n'_s \ll n_s, \mathbf{r}$
- 2: **for** $k = 0, 1, \dots$ **do**
- 3: Randomly choose shot subsets $\mathcal{I} \subset [1 \dots n_s], |\mathcal{I}| = n'_s$
- 4: $\mathbf{A}_k = \{\nabla F_i(\mathbf{m}_0, \mathbf{q}_0) \mathbf{C}^\top\}_{i \in \mathcal{I}}$
- 5: $\mathbf{b}_k = \{\delta \mathbf{d}_i\}_{i \in \mathcal{I}}$
- 6: $\tilde{\mathbf{b}}_k = \mathbf{A}_k \mathbf{x}_k$
- 7: $t_k = \|\tilde{\mathbf{b}}_k - \mathbf{b}_k\|_2^2 / \|\mathbf{A}_k^\top (\tilde{\mathbf{b}}_k - \mathbf{b}_k)\|_2^2$
- 8: $\mathbf{z}_{k+1} = \mathbf{z}_k - t_k \mathbf{A}_k^\top (\mathbf{w}_k * \mathcal{P}_\sigma(\mathbf{w}_k * \tilde{\mathbf{b}}_k - \mathbf{b}_k))$
- 9: $\mathbf{x}_{k+1} = S_{\lambda_1}(\mathbf{z}_{k+1})$
- 10: $\mathbf{w}_{k+1} = \arg \min_{\mathbf{w}} \|\mathbf{w} * \tilde{\mathbf{d}}_k - \mathbf{b}_k\|_2^2 + \|\text{diag}(\mathbf{r})(\mathbf{w} * \mathbf{q}_0)\|_2^2$
- 11: **end for**
- 12: **Output:** $\mathbf{q} = \hat{\mathbf{w}}_{k+1} * \mathbf{q}_0$, and $\delta \hat{\mathbf{m}} = \mathbf{C}^\top \mathbf{x}_{k+1}$



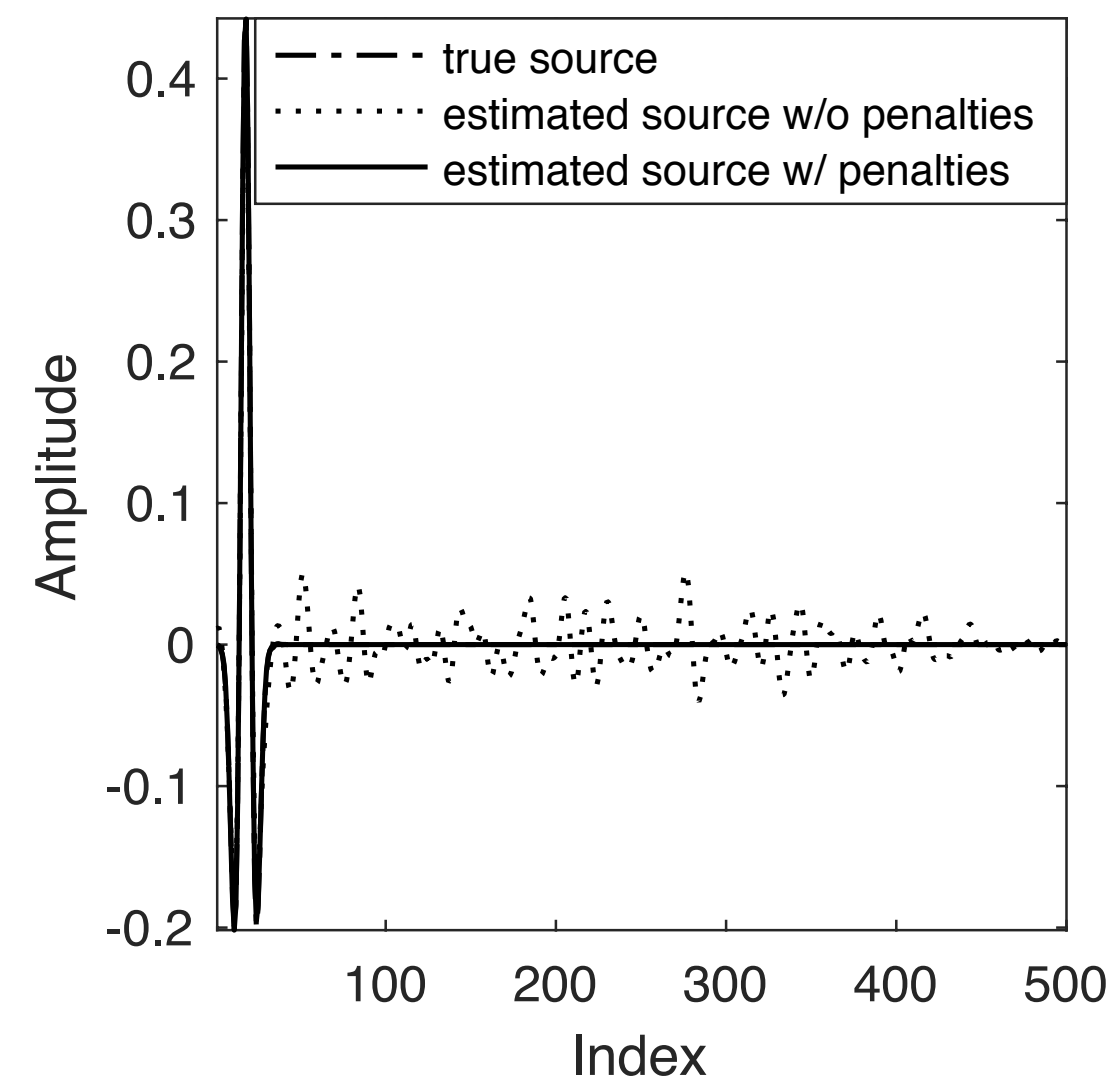
Stylized example - penalization avoiding overfitting



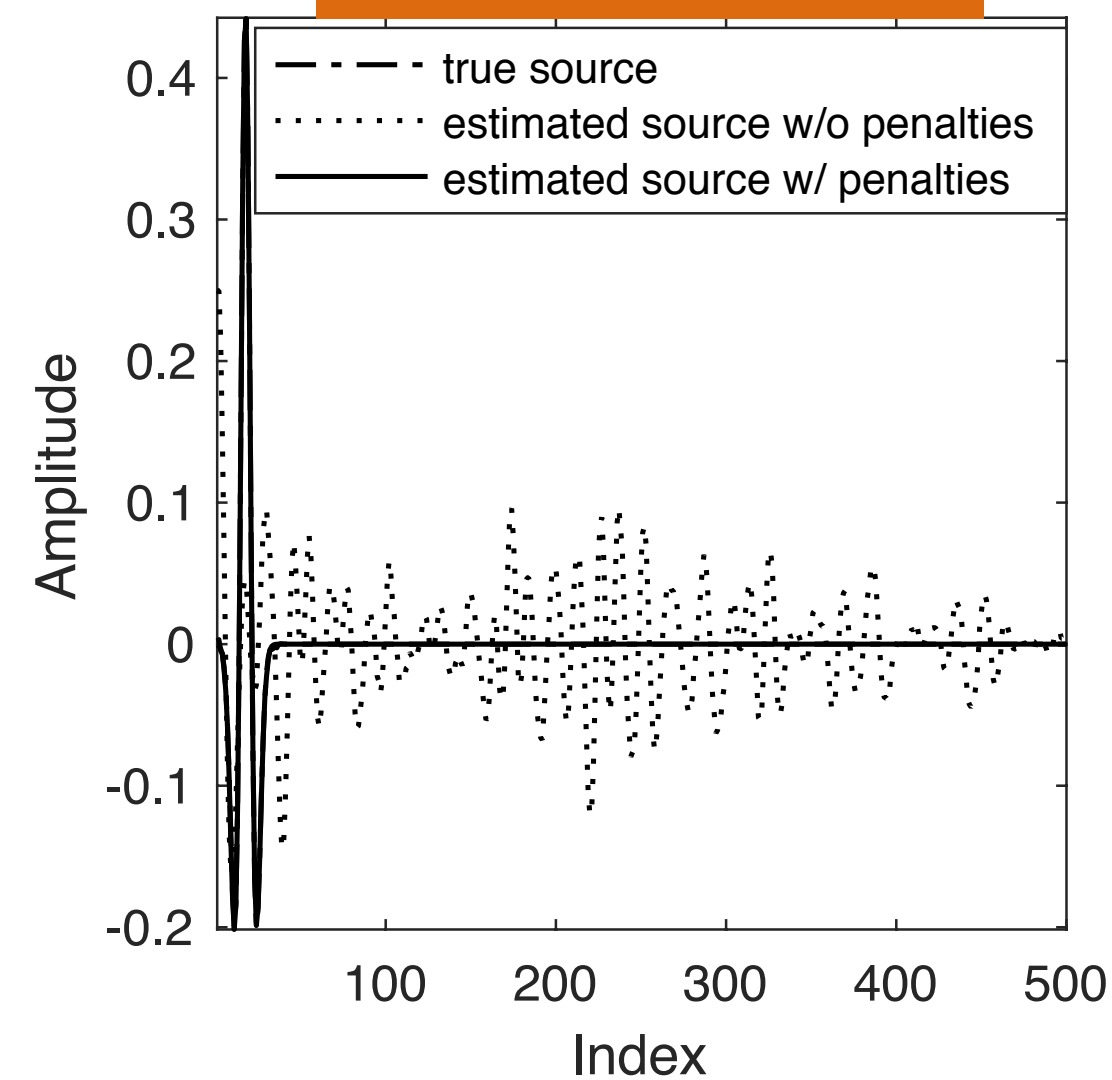
data w/o noise



data w/ noise



data w/o noise



data w/ noise

$$\mathbf{W}\mathbf{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{A} \in \mathcal{R}^{20000 \times 10000}$$

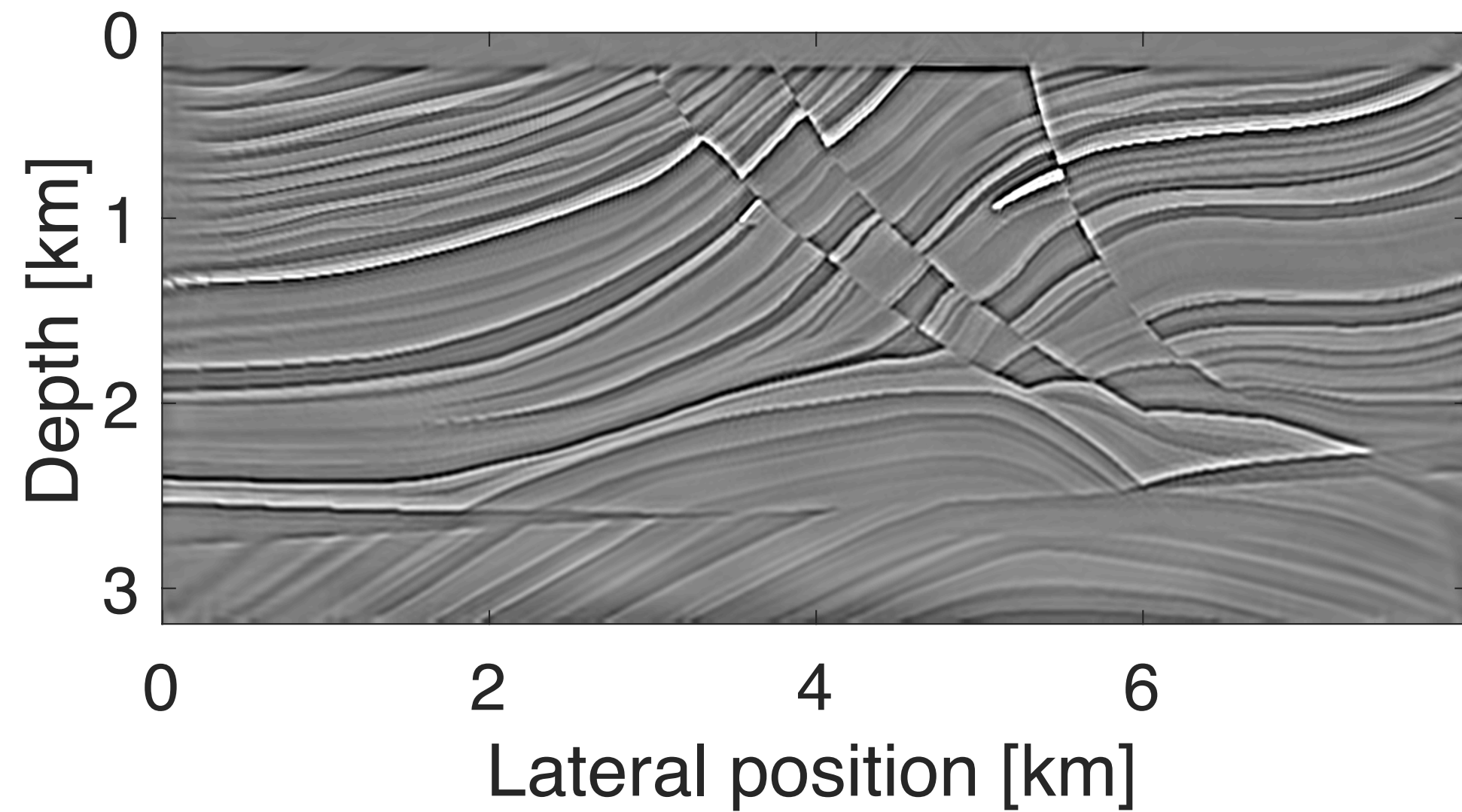
$$\text{rank}(\mathbf{A}) = 500$$

$\mathbf{x} \in \mathcal{R}^{10000 \times 1}$ has 20 non-zeros

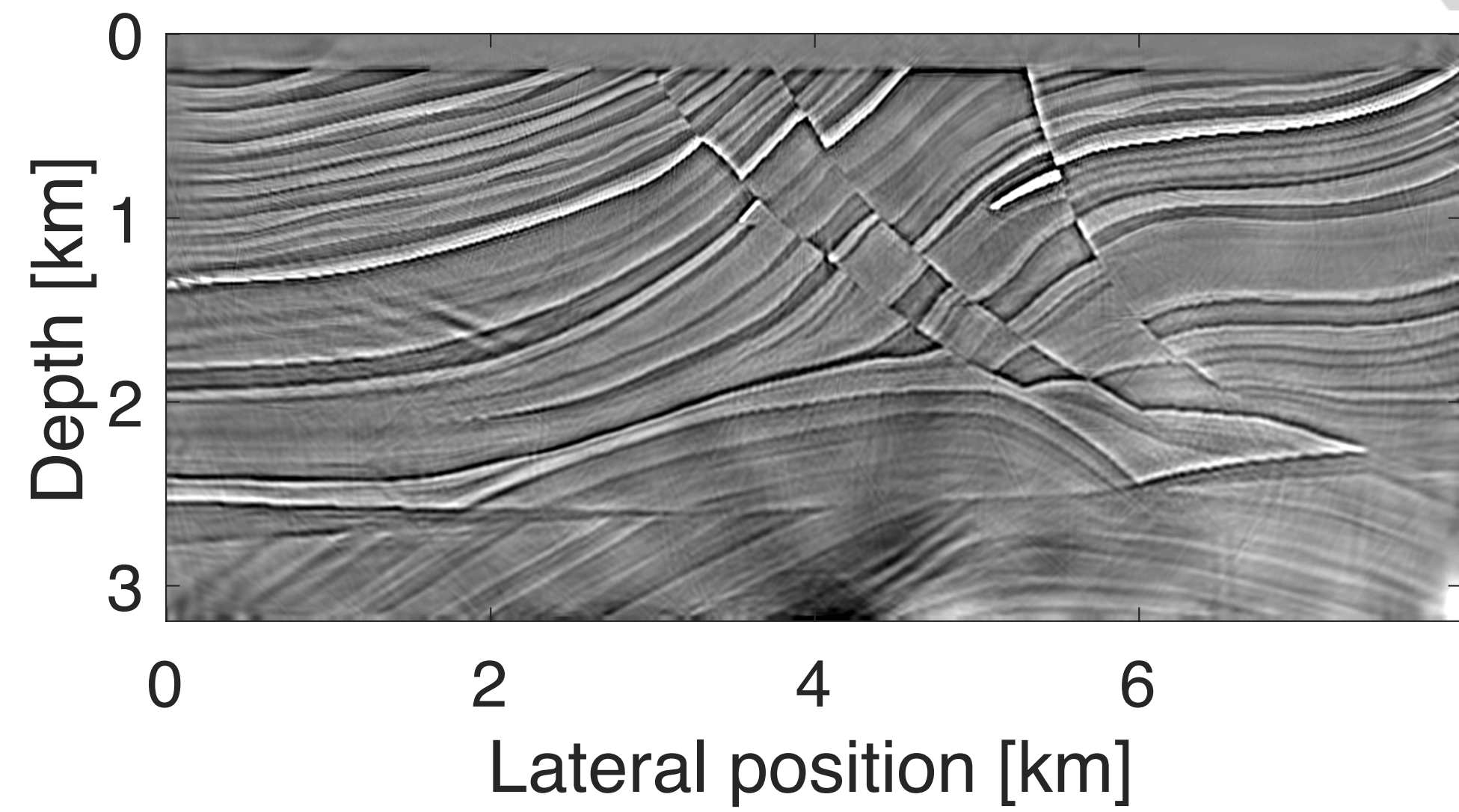
$$\mathbf{A}_i \in \mathcal{R}^{500 \times 10000}, i \in [1 \dots 40]$$

w filter

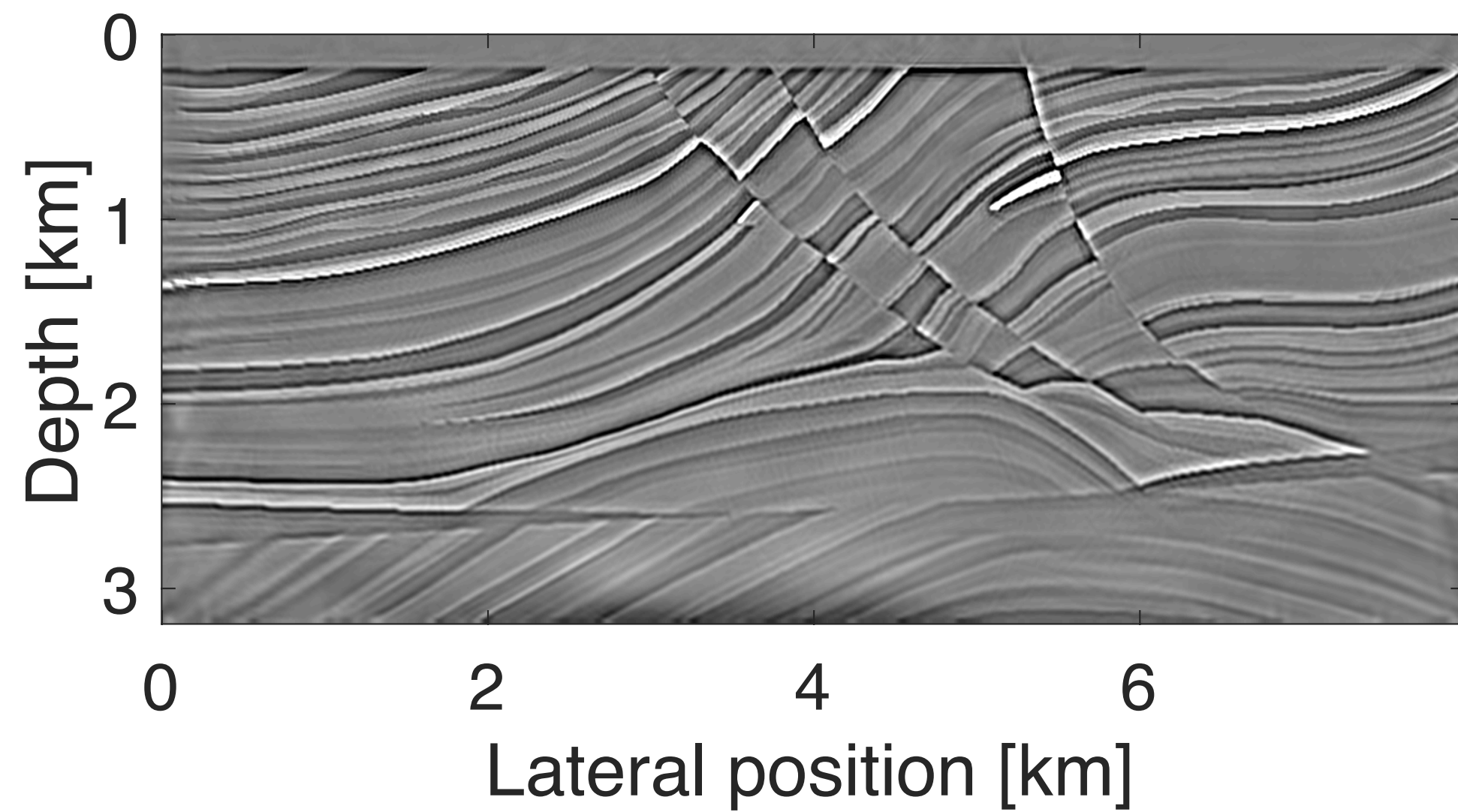
Robustness over noise



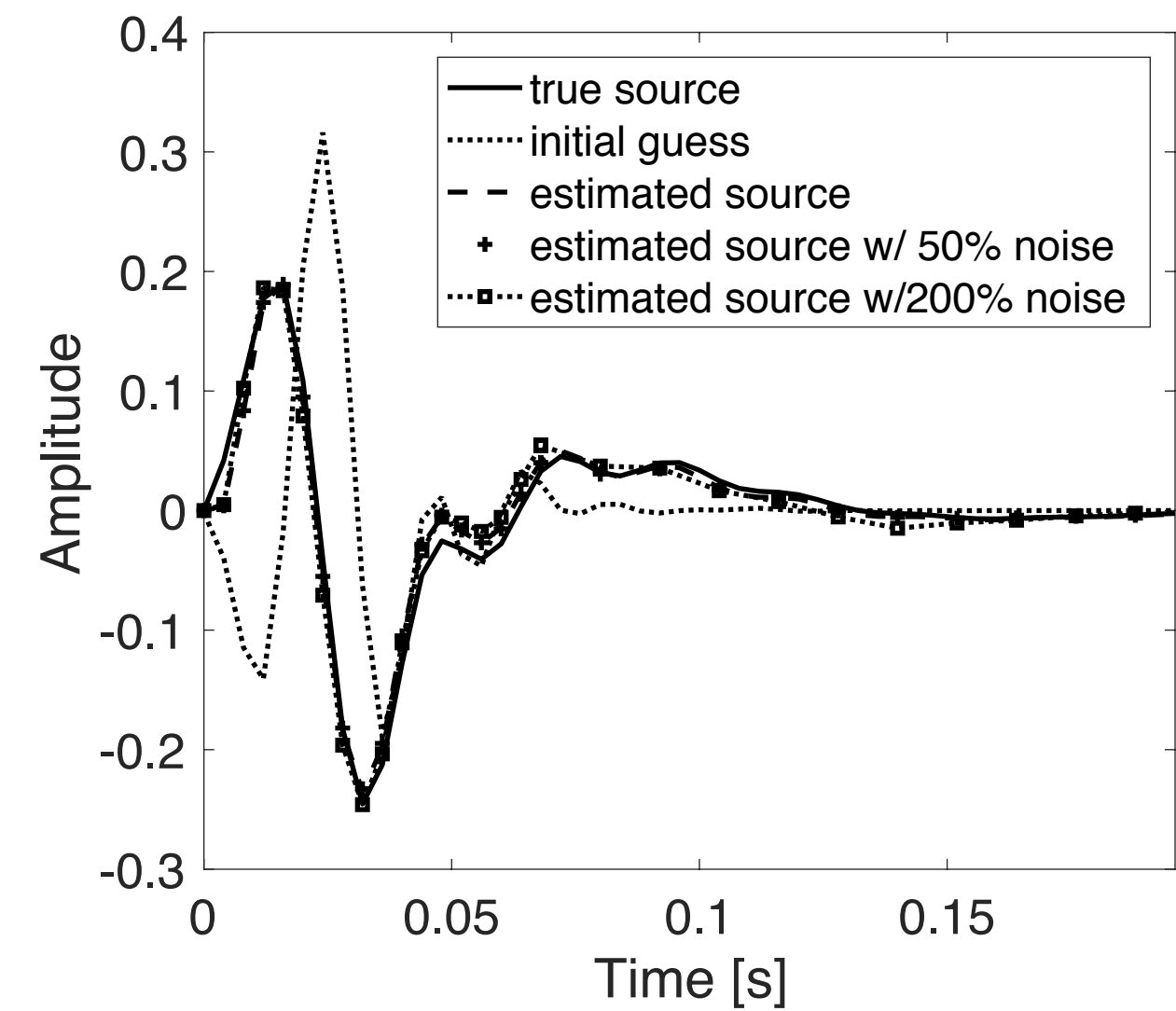
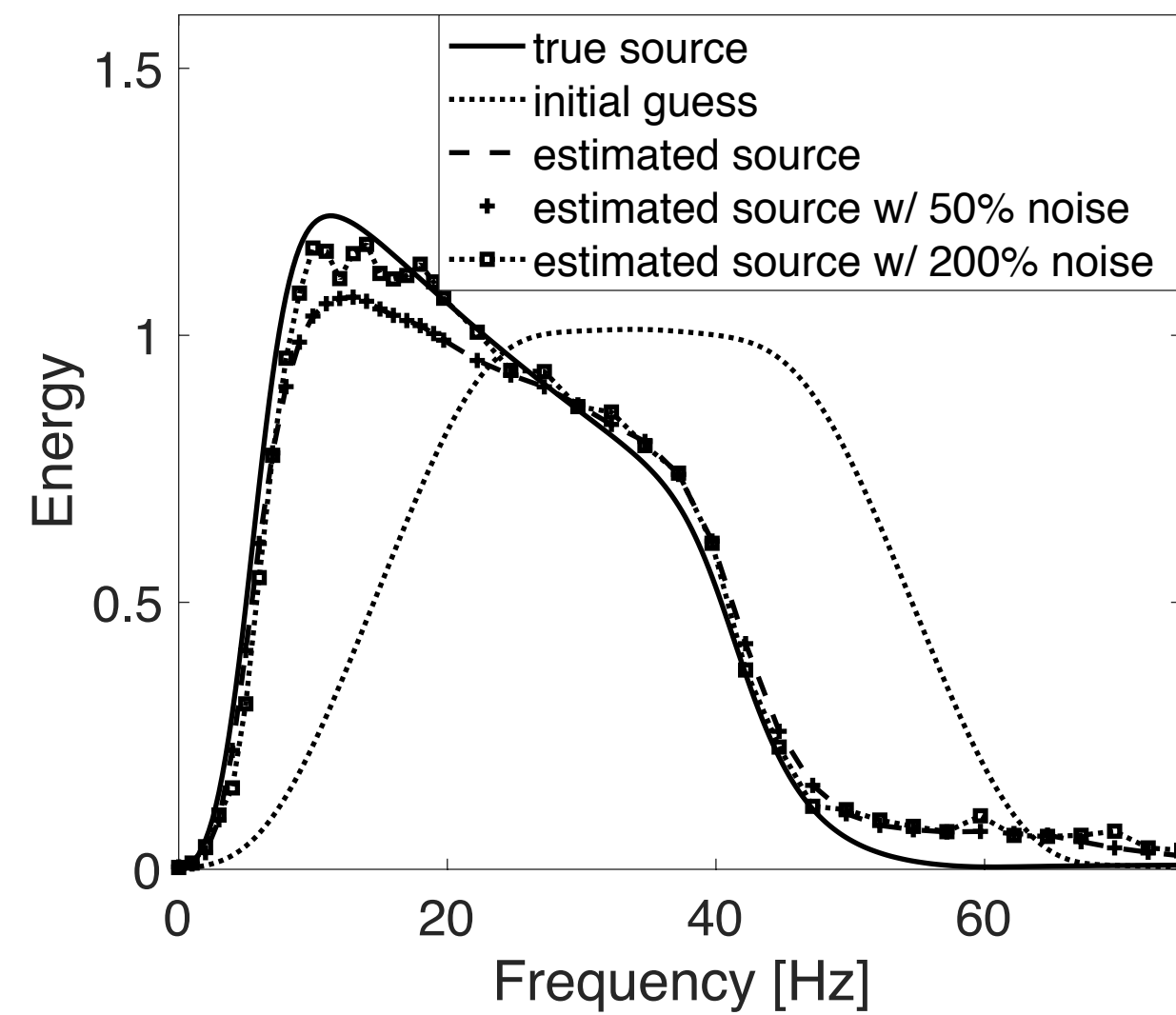
No noise



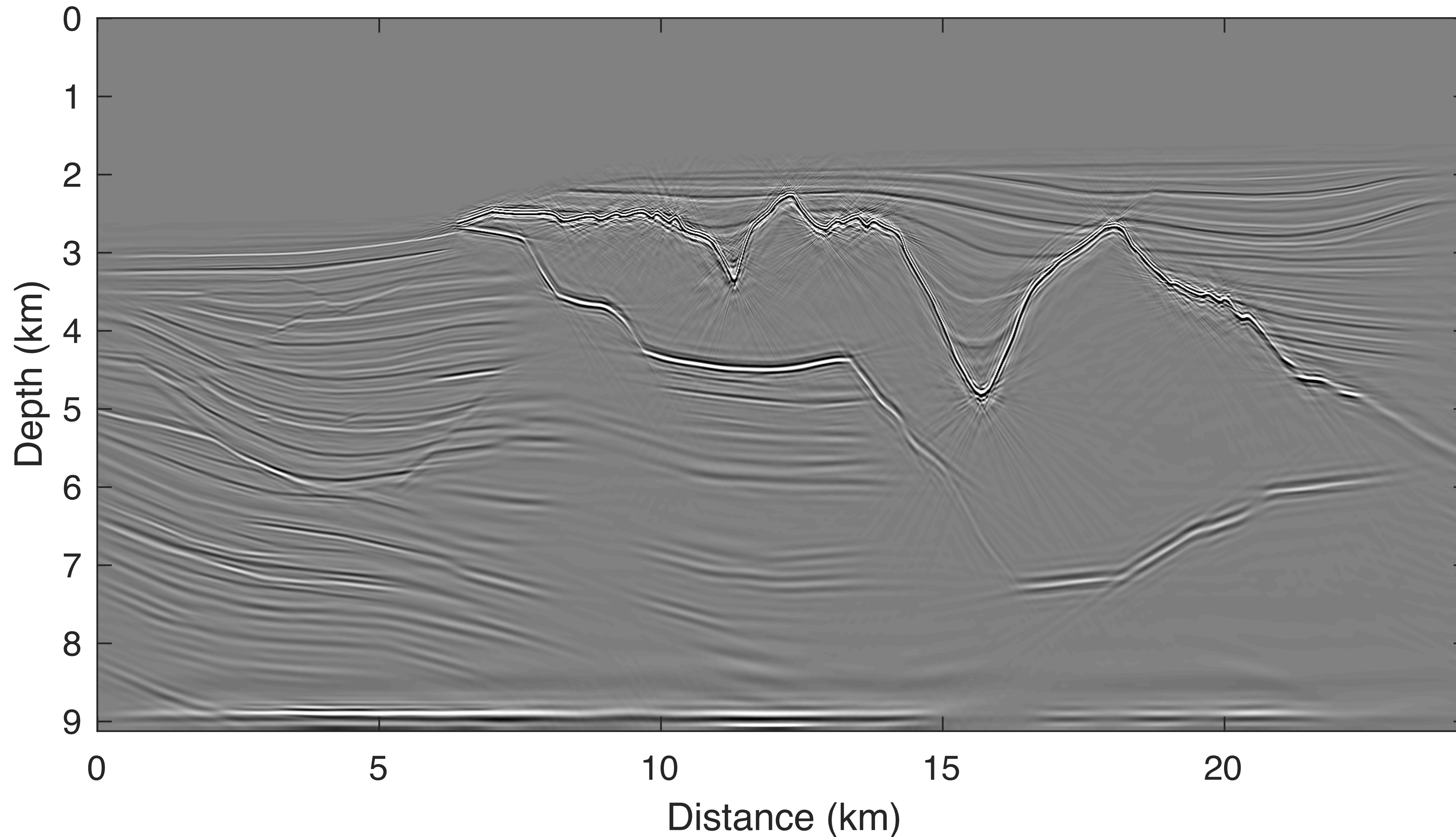
200% noise



50% noise



Challenge of salt model

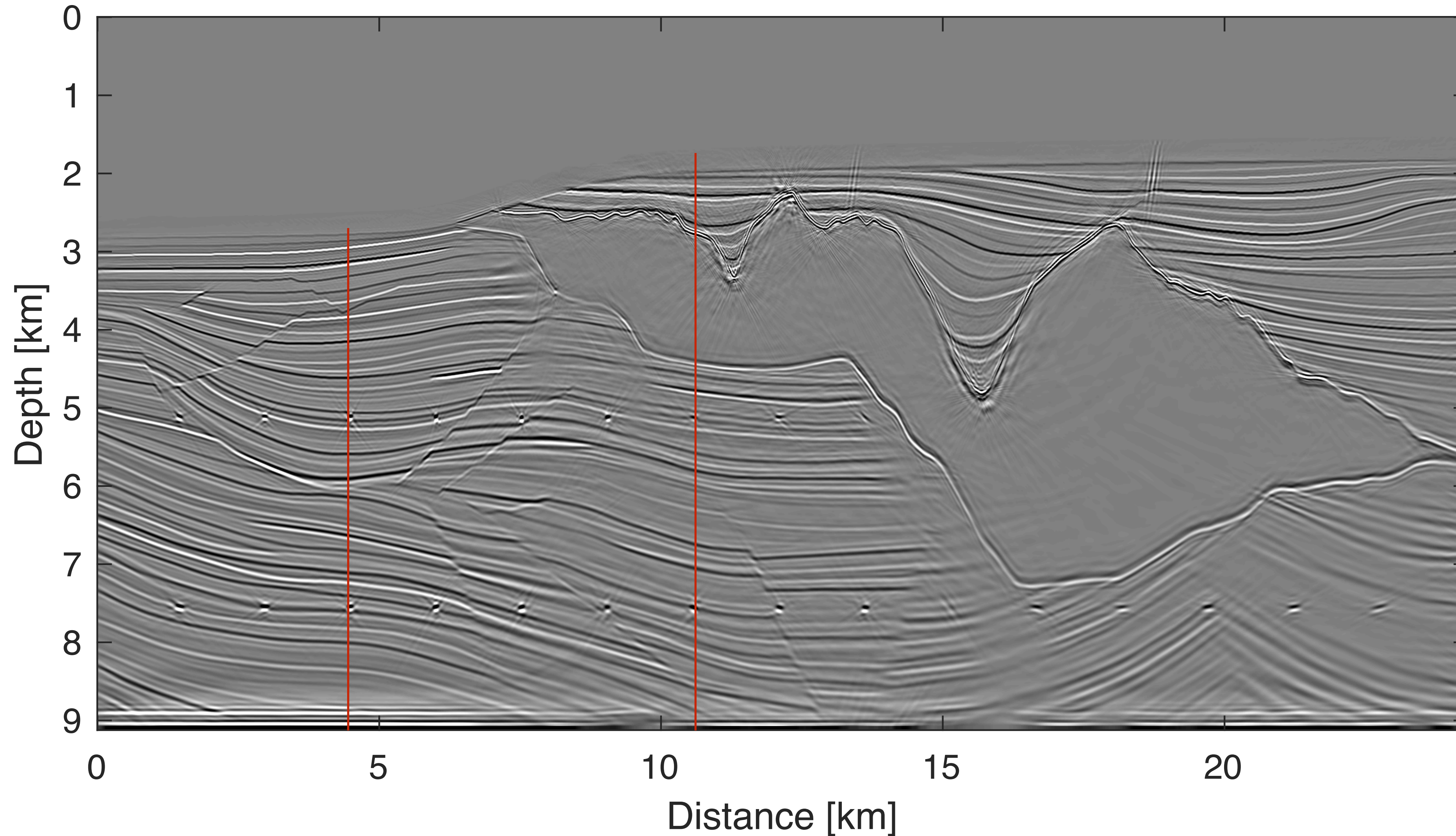


Data :

- Marine acquisition
- nonlinear
- 960 source & receivers
- 8km offset
- S & R interval 25m
- recording 10s

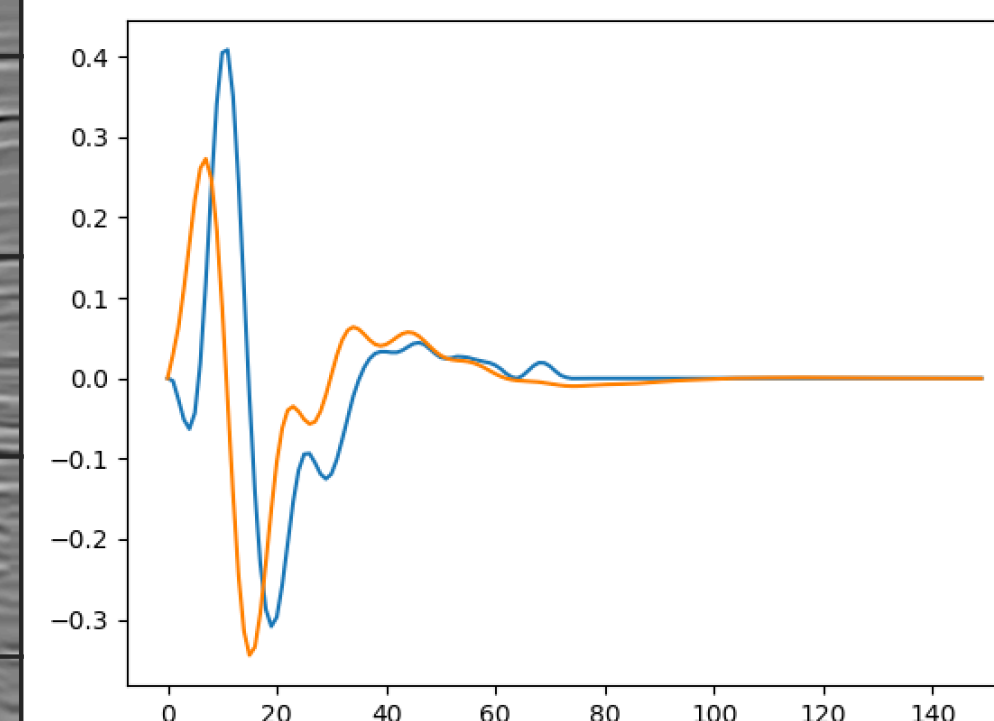
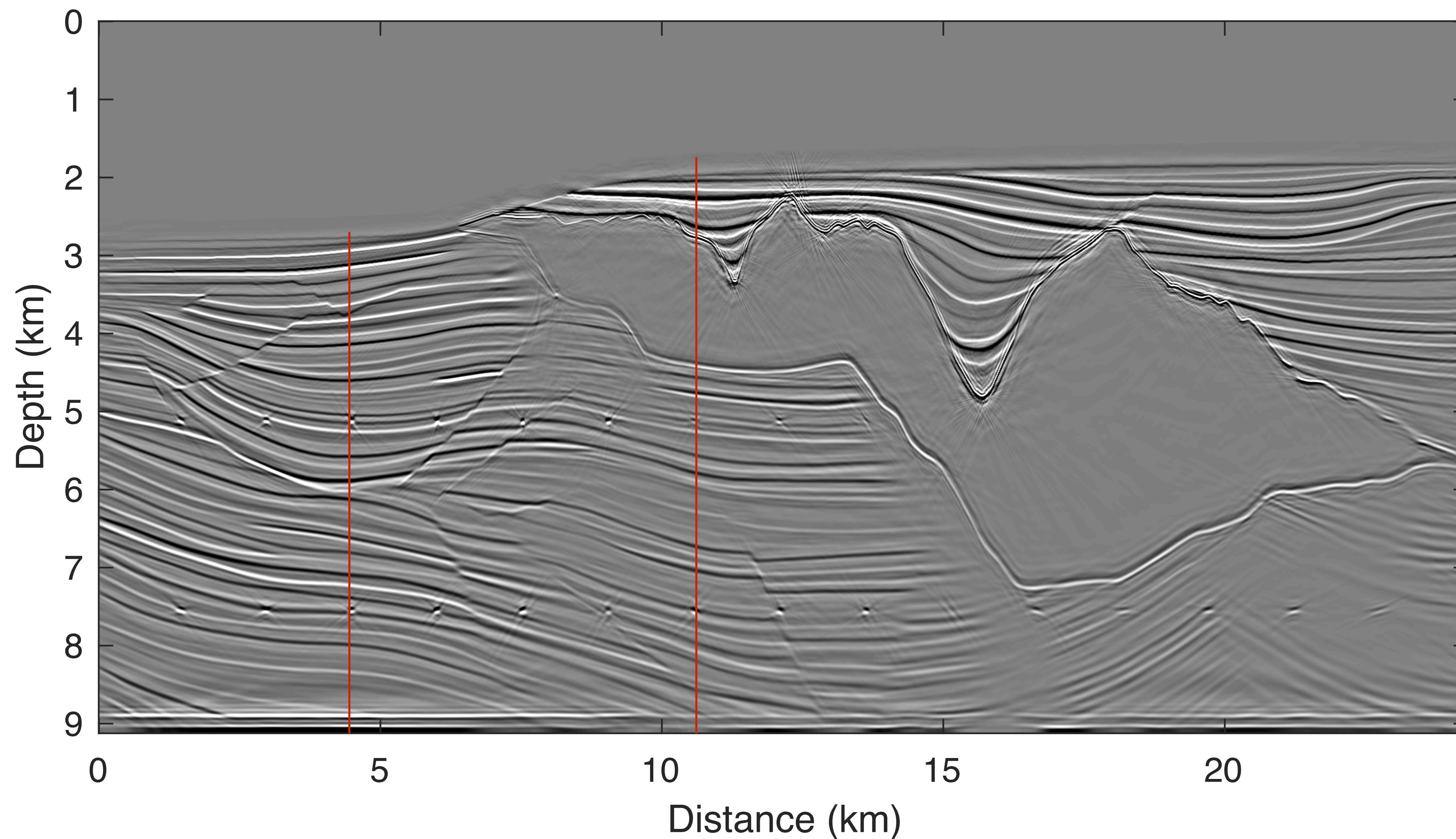
Two data pass w/ true source

inversion w/ inverse scattering image condition



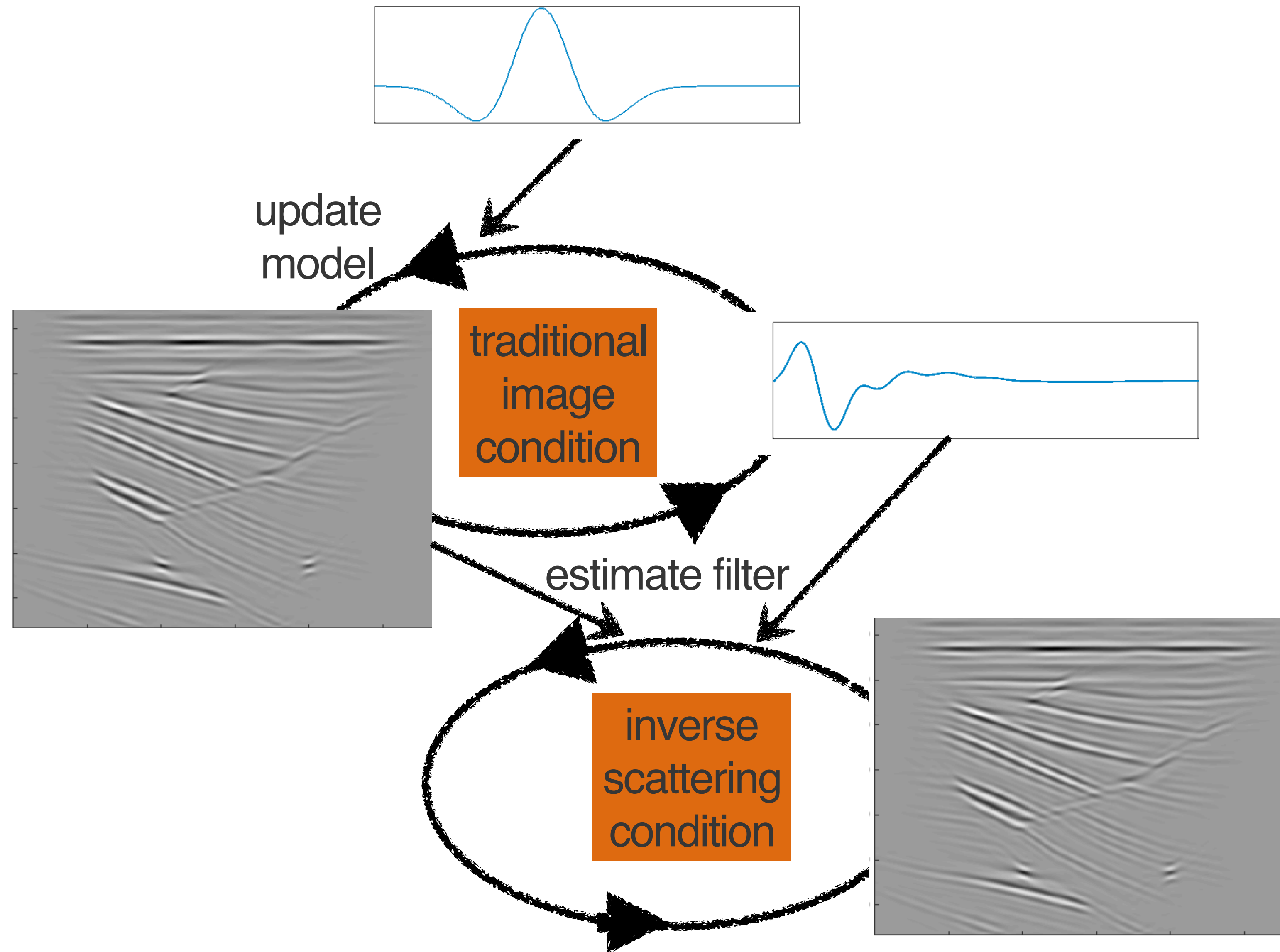
Two data pass w/ true source

inversion w/ inverse scattering image condition

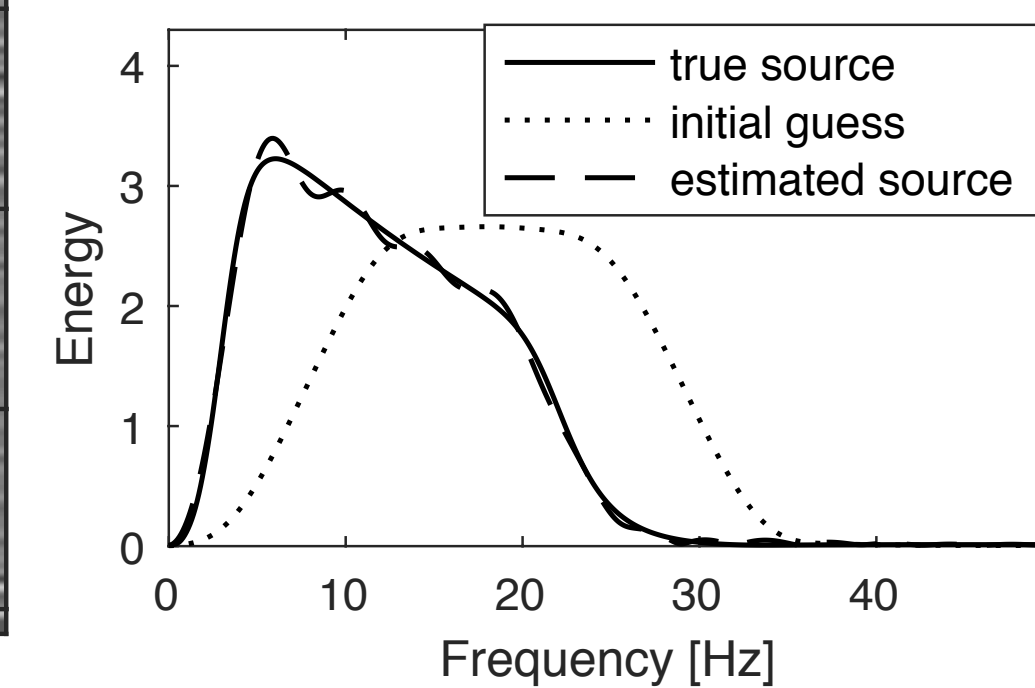
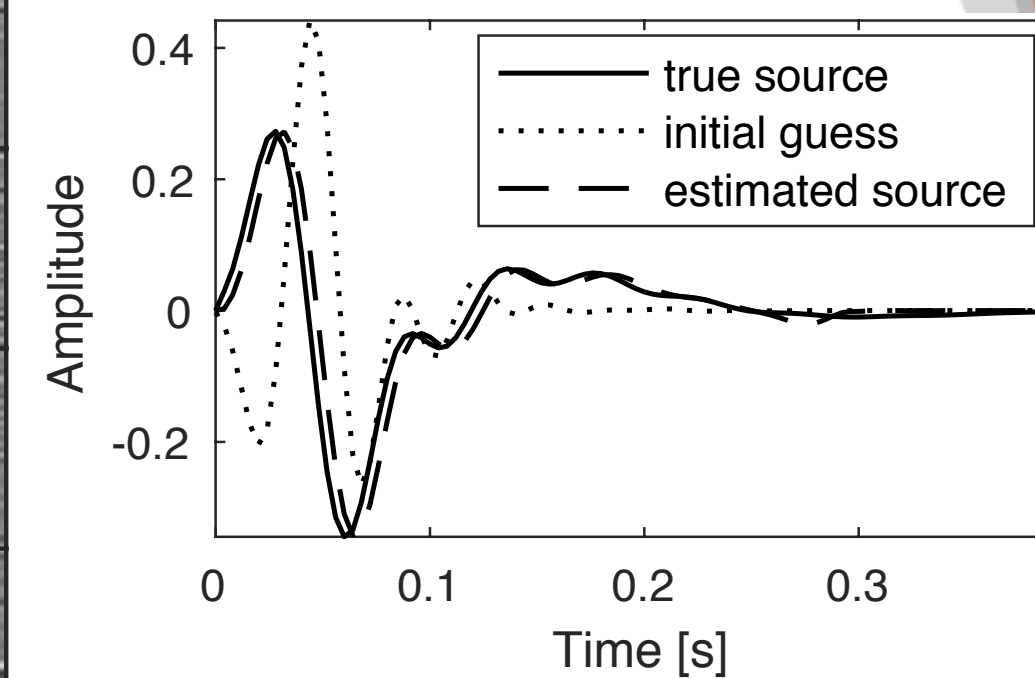
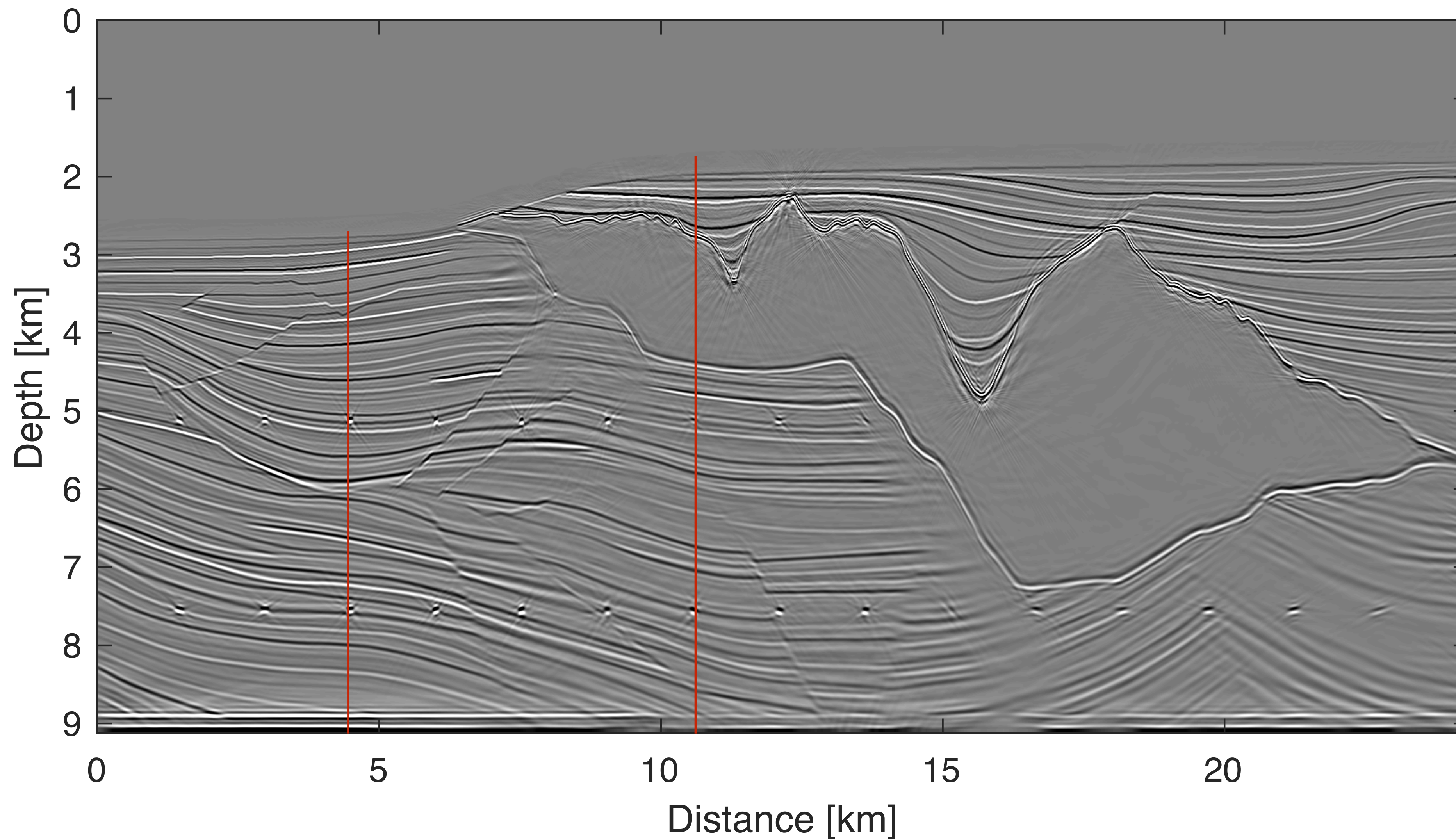


Two data pass w/ source estimation

Hybrid workflow for salt model

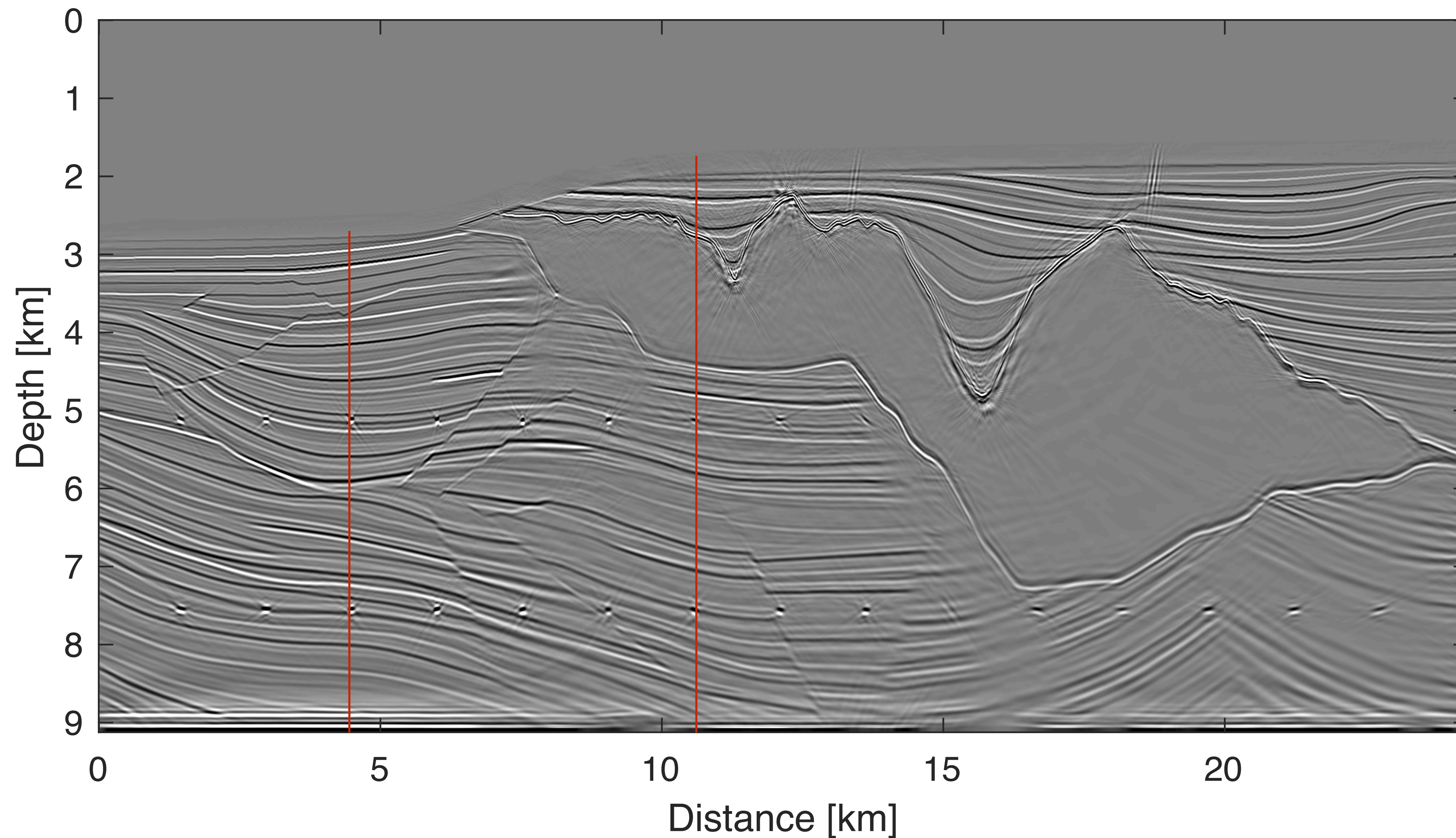


Hybrid workflow for salt model



Two data pass w/ hybrid workflow

inversion w/ inverse scattering image condition



Two data pass w/ true source

Conclusion

- LB combined w/ on-the-fly source estimation is capable of generating high-fidelity true-amplitude images at the cost of 1~2 RTM.
- The penalized sub-problem solved during variable projection can avoid overfitting to noise.
- Hybrid framework of LB w/ on-the-fly source estimation and LB based on inverse-scattering condition generates artifact-free images.

Chapter 2 Low-rank recovery for subsurface extended image volumes based on time-stepping propagator and power schemes

Chapter 3 mapping for velocity variation scenarios via invariance relationship

Extended Image Volumes

Extended image volume for single frequency

$$\mathbf{E}_i = -\omega_i^2 \mathbf{V}_i \mathbf{U}_i^* = \mathbf{H}_i^{-*} \mathbf{P}_r^\top \dot{\mathbf{D}}_i \dot{\mathbf{Q}}_i^* \mathbf{P}_s \mathbf{H}_i^{-*}$$

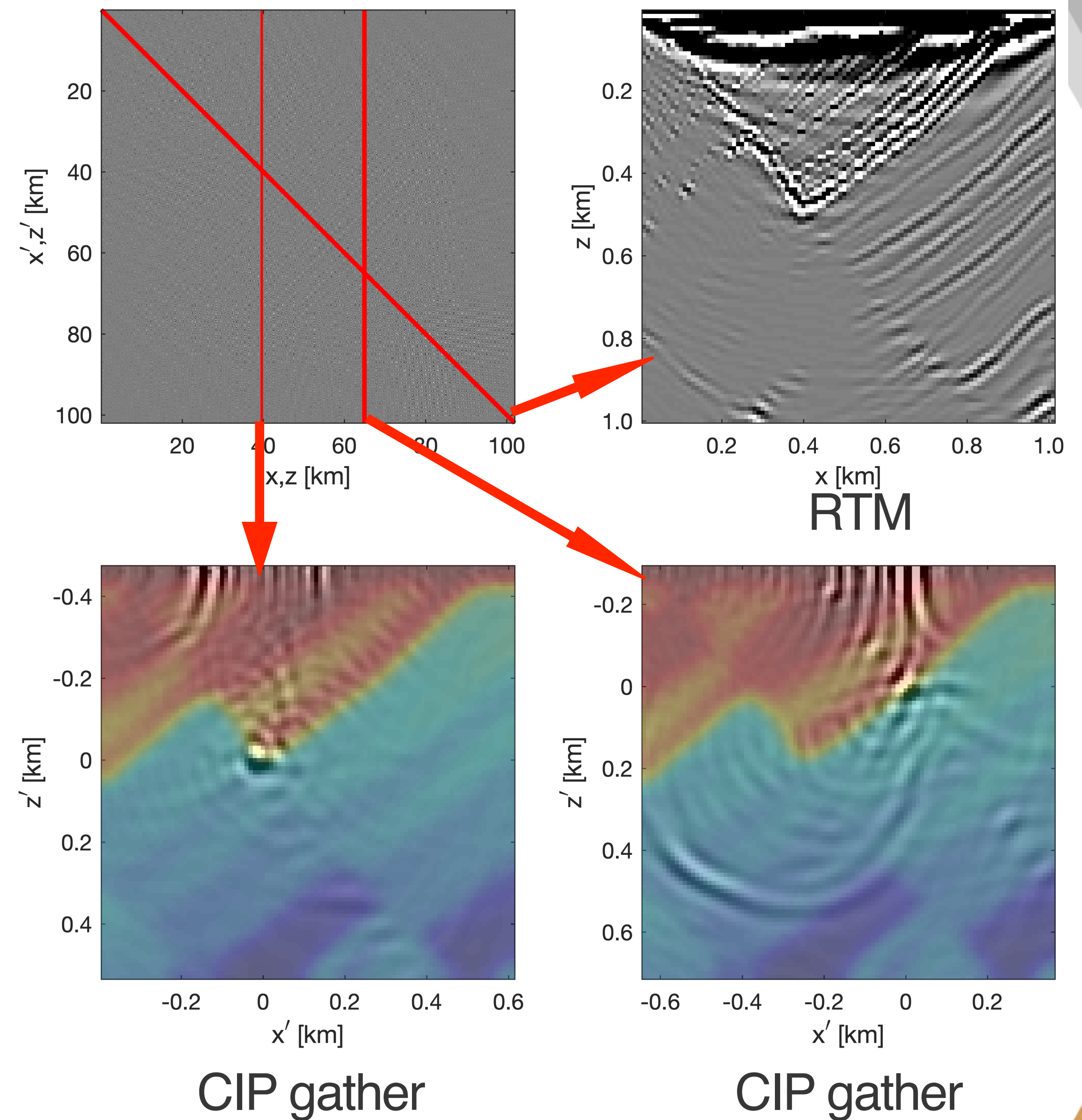
ω_i : angular frequency

$\mathbf{H}_i(\mathbf{m})$: discretization of the Helmholtz operator

\mathbf{Q}_i : source

\mathbf{D}_i : data matrix

$\mathbf{P}_s, \mathbf{P}_r$: projection operators that restrict the wavefields



Extended Image Volumes

Express image volume for single frequency

$$\mathbf{E}_i = -\omega_i^2 \mathbf{V}_i \mathbf{U}_i^* = \mathbf{H}_i^{-*} \mathbf{P}_r^\top \mathbf{D}_i \mathbf{Q}_i^* \mathbf{P}_s \mathbf{H}_i^{-*}$$

ω_i : angular frequency

$\mathbf{H}_i(\mathbf{m})$: discretization of the Helmholtz operator

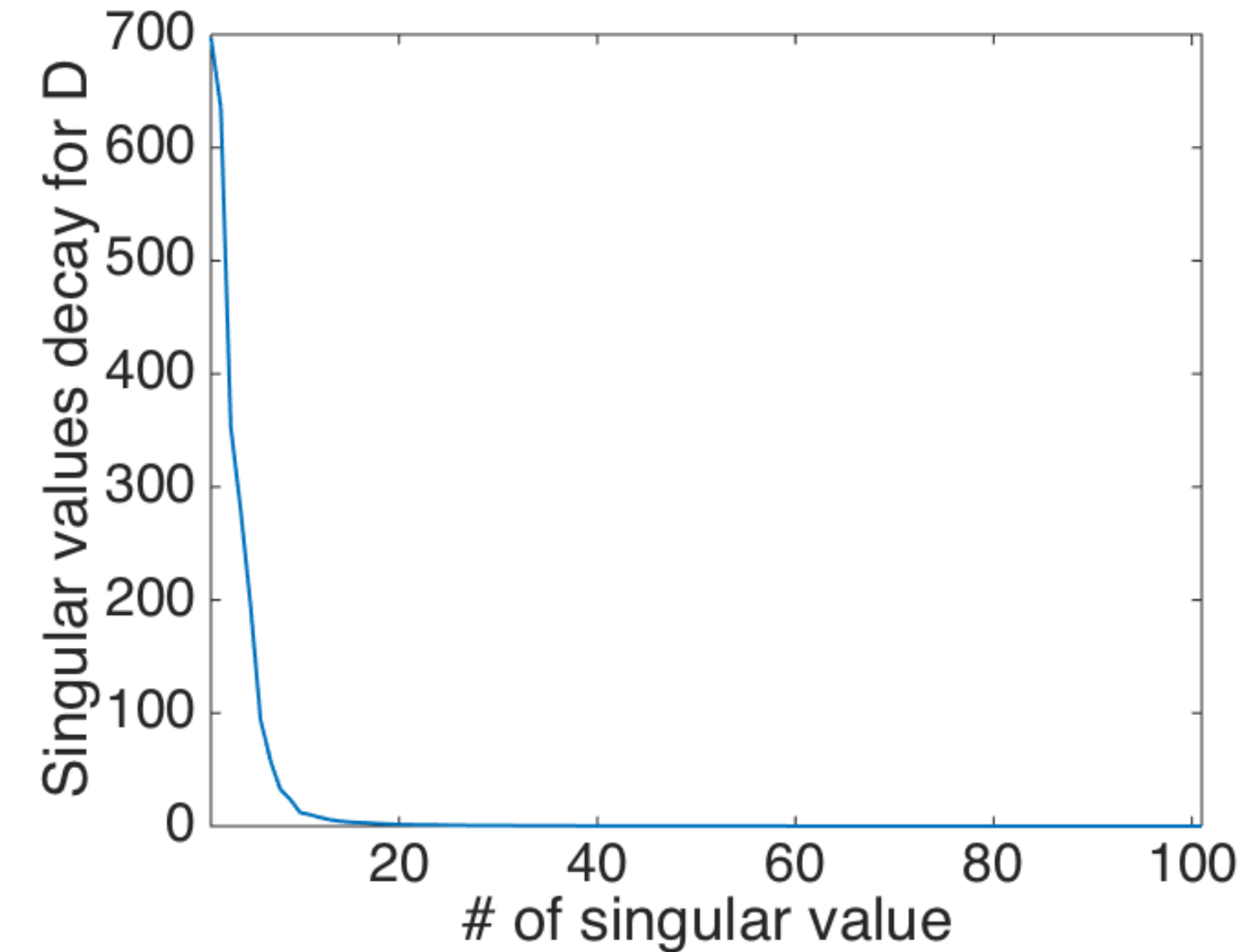
\mathbf{Q}_i : source

\mathbf{D}_i : data matrix

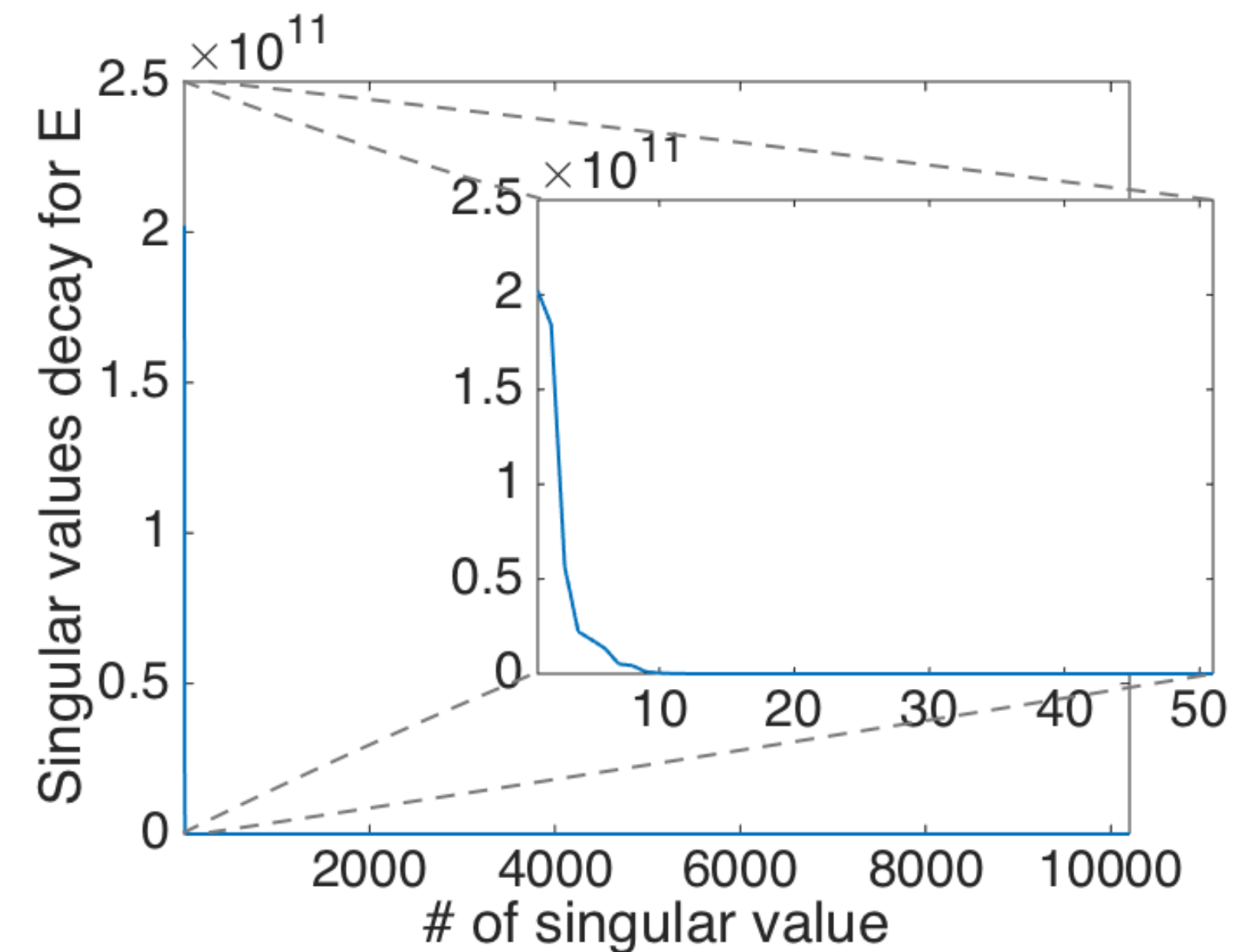
$\mathbf{P}_s, \mathbf{P}_r$: projection operators that restrict the wavefields

Too expensive to compute or store

Rank limited by Data's Rank



Monochromatic data at 5Hz



Monochromatic EIV at 5Hz

Monochromatic randomized SVD algorithm

0. **Input:** q and n_p random Gaussian vectors $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_{n_p}]$
1. $\mathbf{Y} := \mathbf{E}\mathbf{W}$, $\mathbf{Y} \in \mathbb{C}^{N \times n_p}$
2. $[\mathbf{Q}, \sim] = \text{qr}(\mathbf{K})$, $\mathbf{Q} \in \mathbb{C}^{N \times n_p}$
3. $\mathbf{Z} = \mathbf{E}^* \mathbf{Q}$, $\mathbf{Z} \in \mathbb{C}^{N \times n_p}$
4. $[\Phi, \Sigma, \Psi] = \text{svd}(\mathbf{Z}^*)$, svd computes the top n_p singular vectors
5. set $\Phi \leftarrow \mathbf{Q}\Phi$
6. $\mathbf{L} = \Phi \sqrt{\Sigma}$, $\mathbf{R} = \Psi \sqrt{\Sigma}$
7. **Output:** factors $\{\mathbf{L}, \mathbf{R}\}$ yielding $\mathbf{E} \approx \mathbf{L}\mathbf{R}^*$

Pros : only $4n_p$ PDEs

: access to each element (various image gathers, e.g. RTM)

Low rank recovery for Extended Image Volumes

Monochromatic randomized SVD algorithm

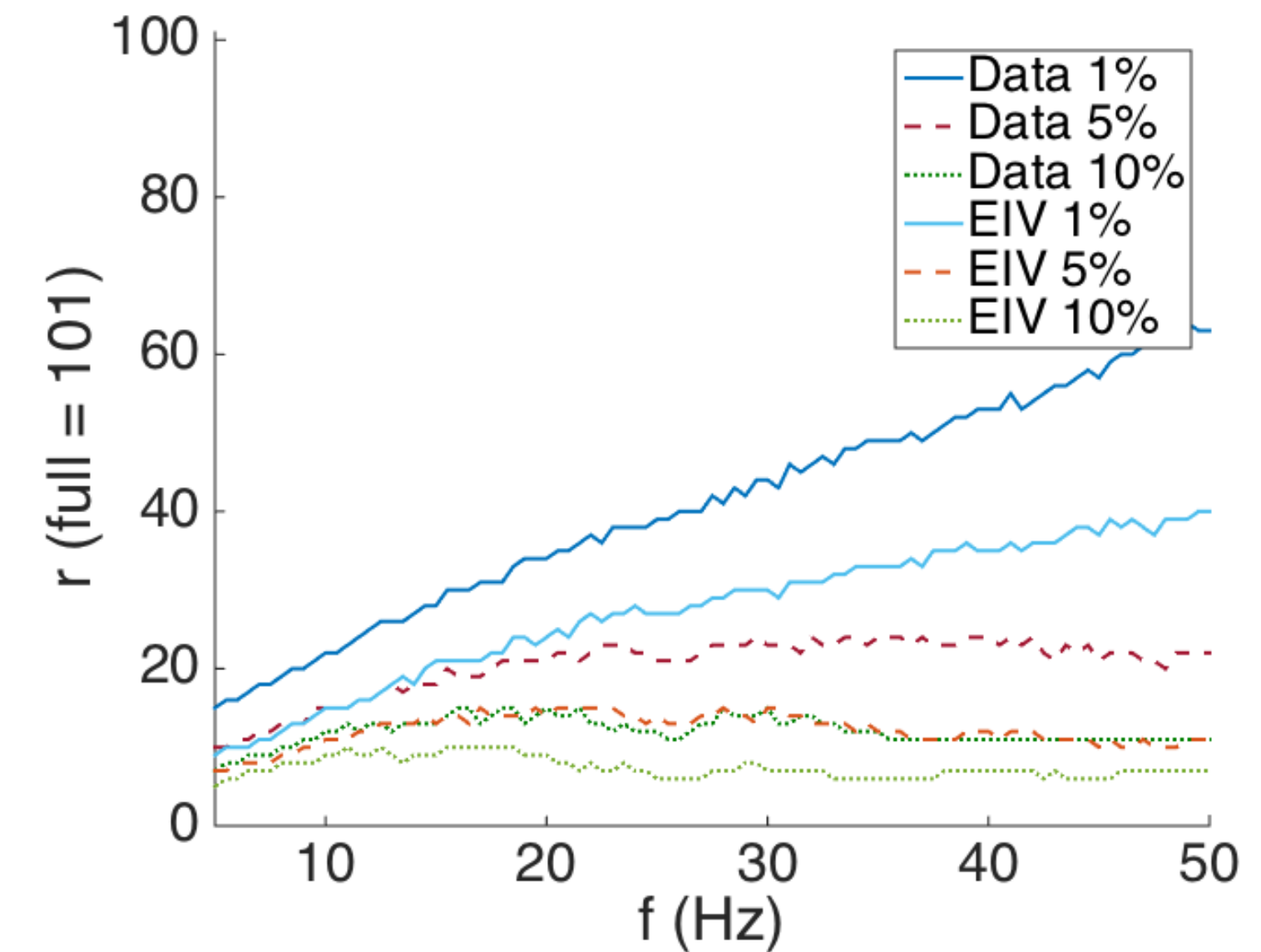
Pros : only $4n_p$ PDEs

: access to each element (various image gathers)

Cons : Rank increasing along frequency

: Inefficient PDE solves for large 2D or 3D problem

: Investment in QR and SVD factorizations



$n_p : 9 \sim 40$

Threshold according 1%,5%,10% of the maximum of singular values

Power scheme combined with rSVD : Simultaneous Iterations vs Block Krylov Iterations

| | SI | BKI |
|-------------|-----------------------|-------------------------------------|
| Cost | $2n_p + 4qn_p$ | $2n_p + 4qn_p$ wave-equation solves |
| | $\mathcal{O}(Nn_p^2)$ | $\mathcal{O}(N((q+1)n_p)^2)$ flops |
| | $2n_p$ | $2(q+1)n_p$ wave-equation solves |
| | $\mathcal{O}(Nn_p^2)$ | $\mathcal{O}(N(q+1)n_p^2)$ flops |

$$\mathbf{K} := (\mathbf{E}\mathbf{E}^*)^q \mathbf{E}\mathbf{W}, \mathbf{K} \in \mathbb{C}^{N \times n_p} \quad \text{vs} \quad \mathbf{K} := [\mathbf{E}\mathbf{W}, (\mathbf{E}\mathbf{E}^*)\mathbf{E}\mathbf{W}, \dots, (\mathbf{E}\mathbf{E}^*)^q \mathbf{E}\mathbf{W}], \mathbf{K} \in \mathbb{C}^{N \times (q+1)n_p}$$

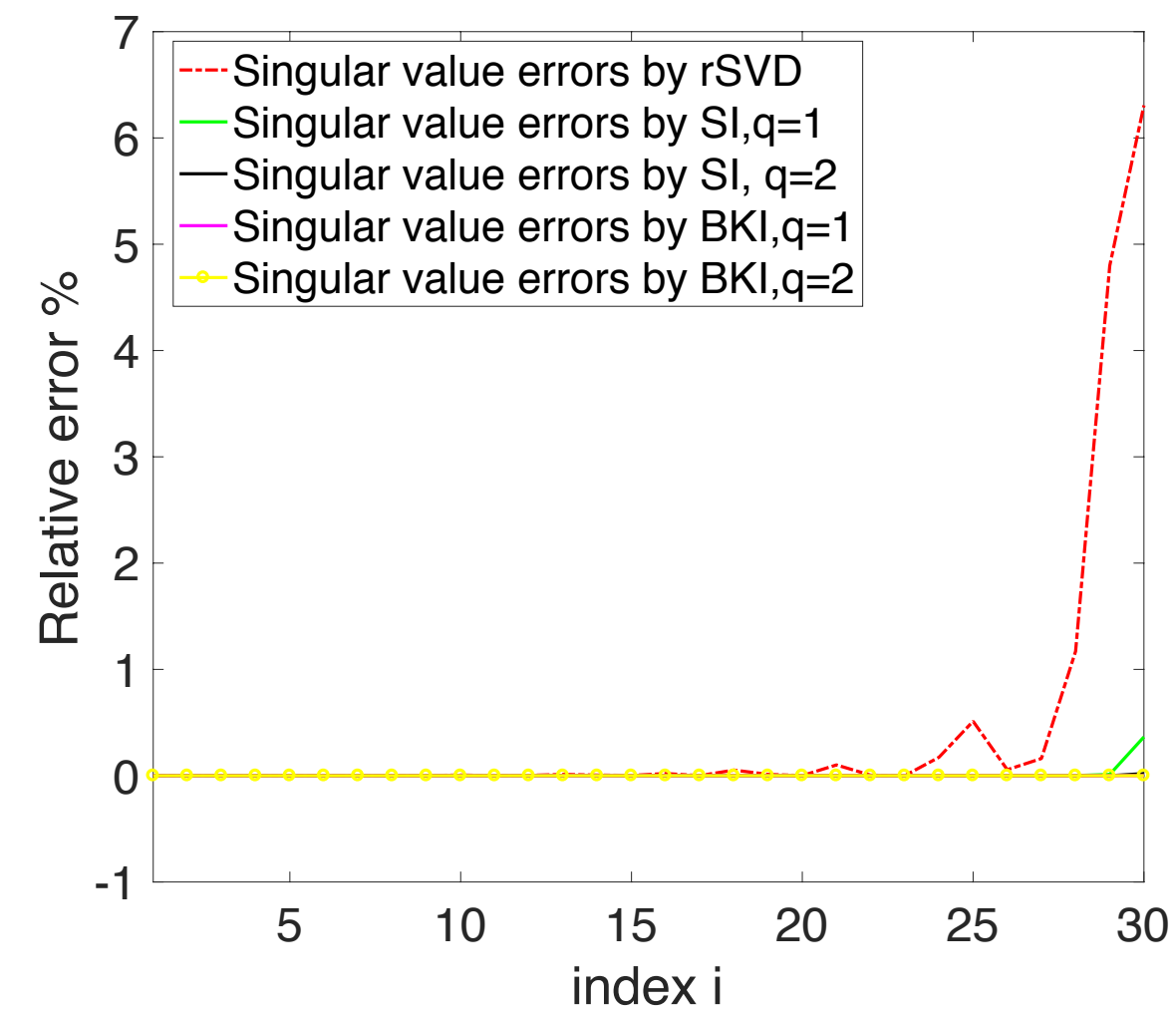
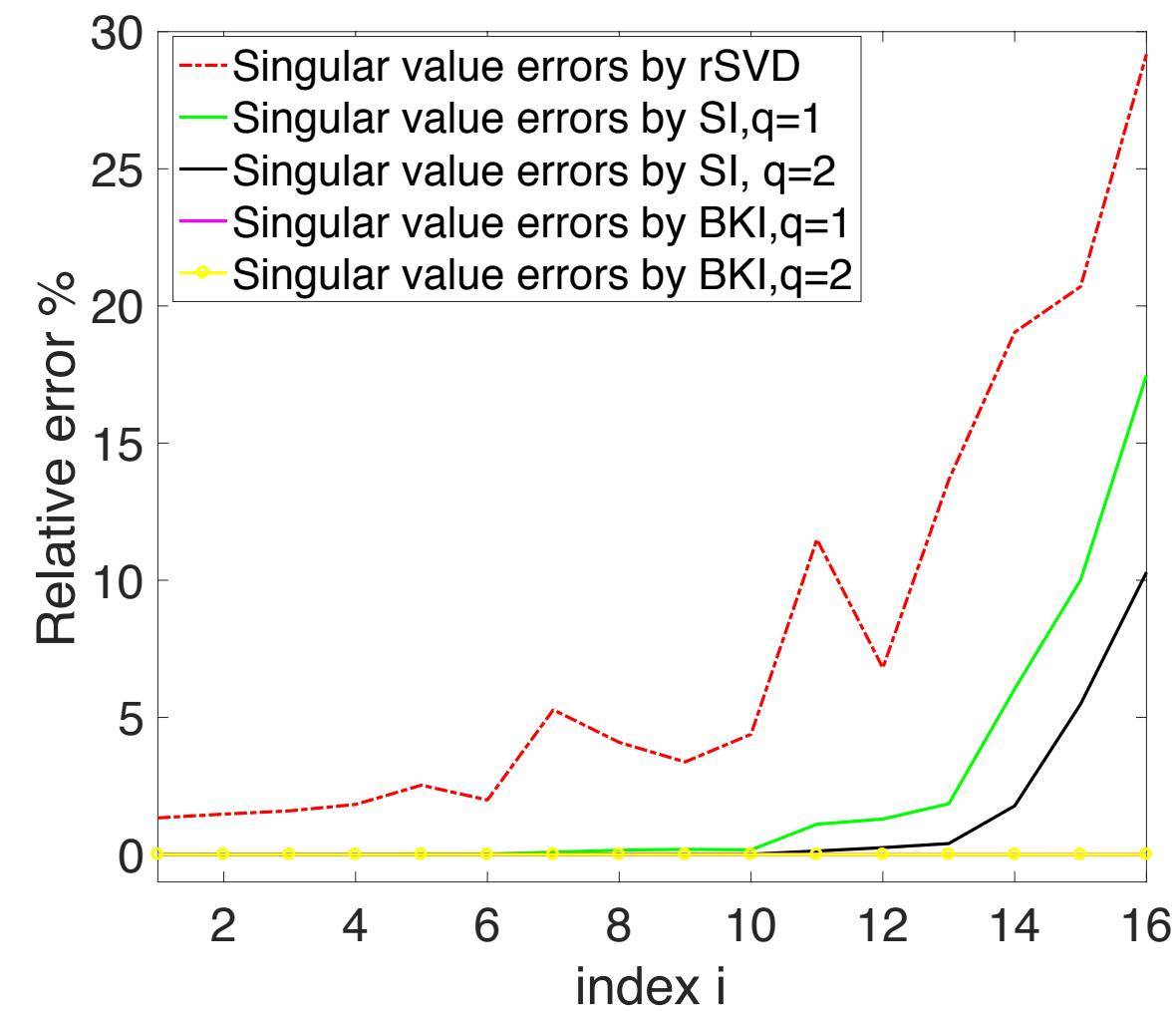
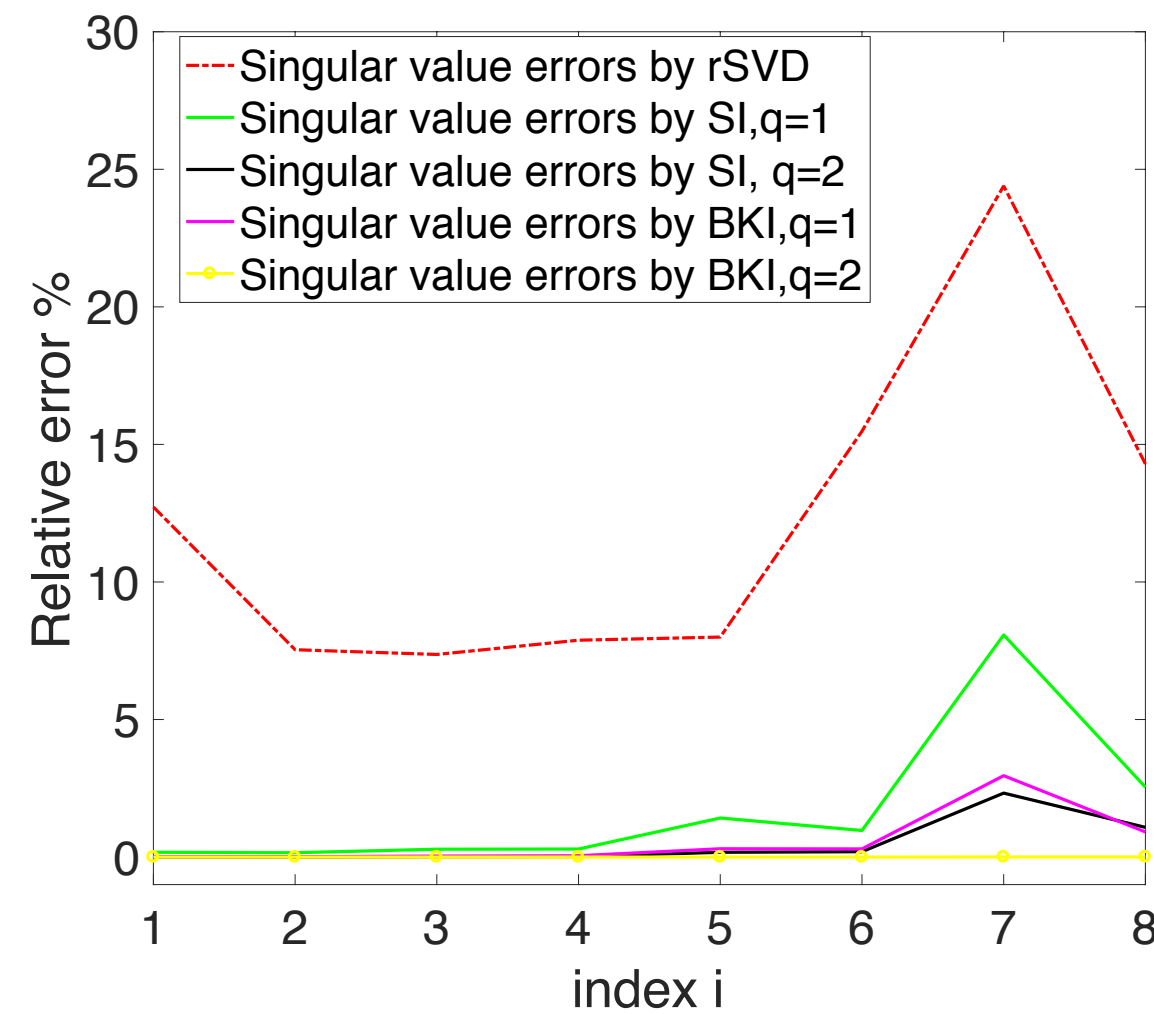
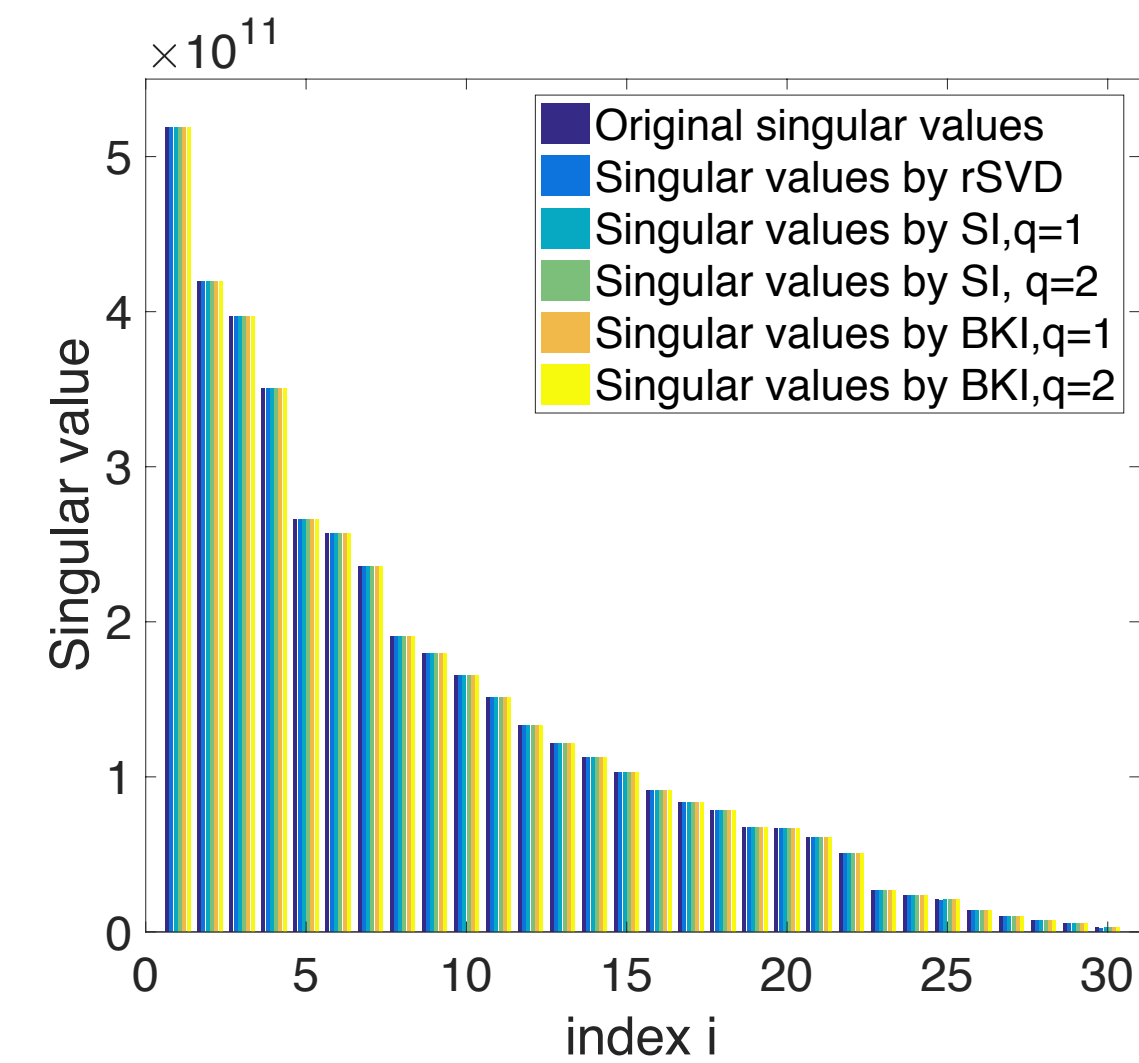
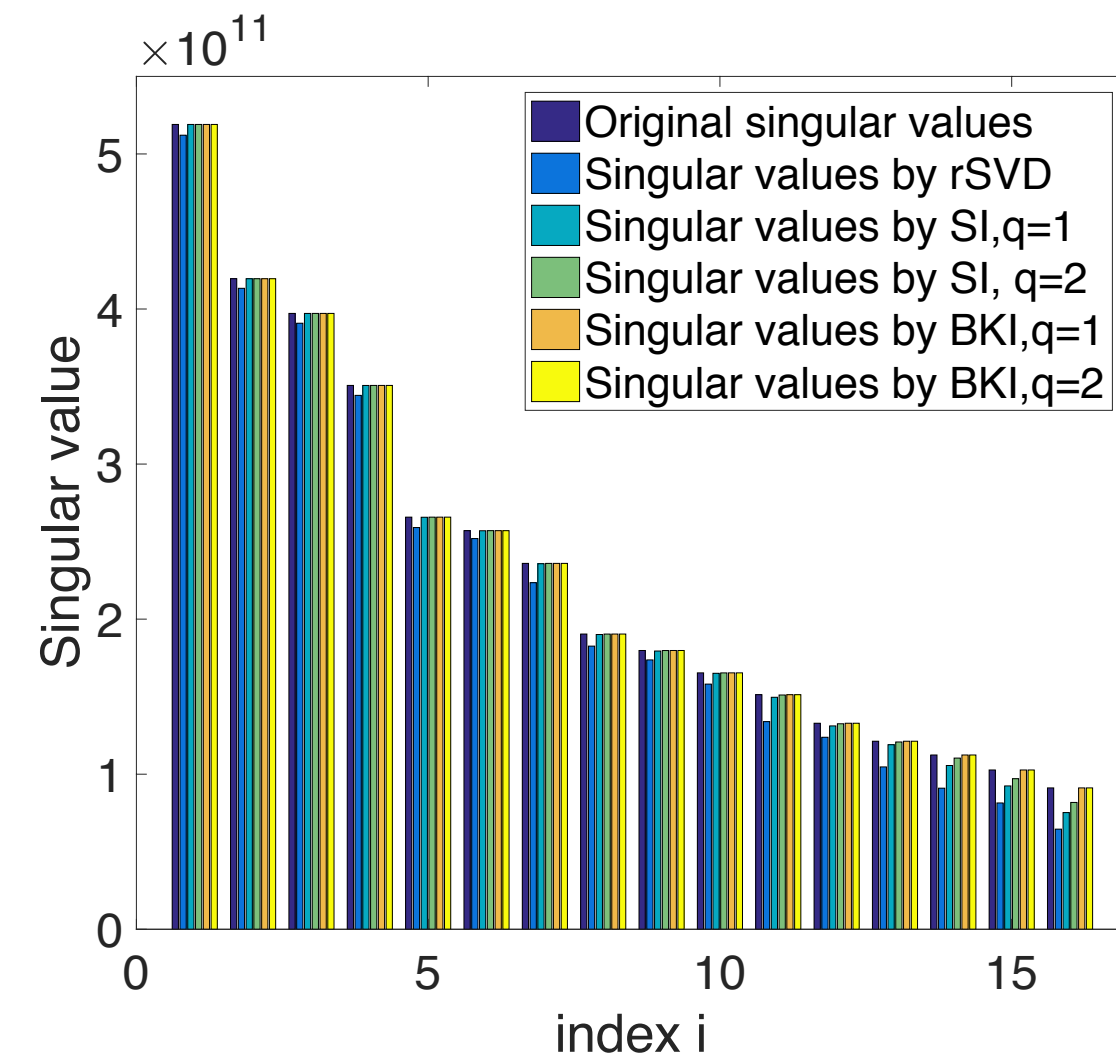
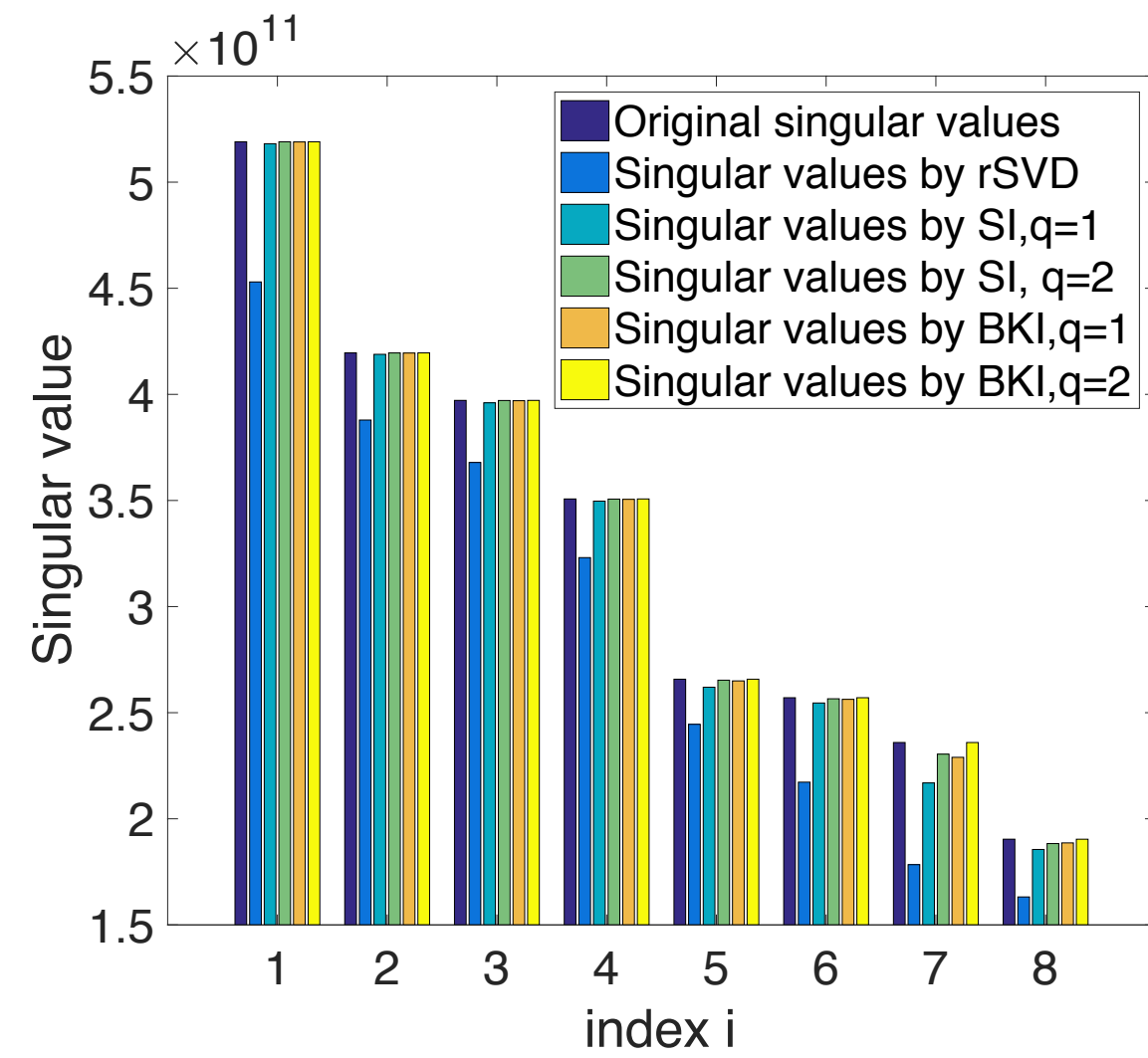
Pros : increasing accuracy w/o increasing probing size

Cons : more WE scale w/ power

: more flops in qr, svd

Musco C et al, "Randomized block krylov methods for stronger and faster approximate singular value decomposition", Advances in Neural Information Processing systems, 2015

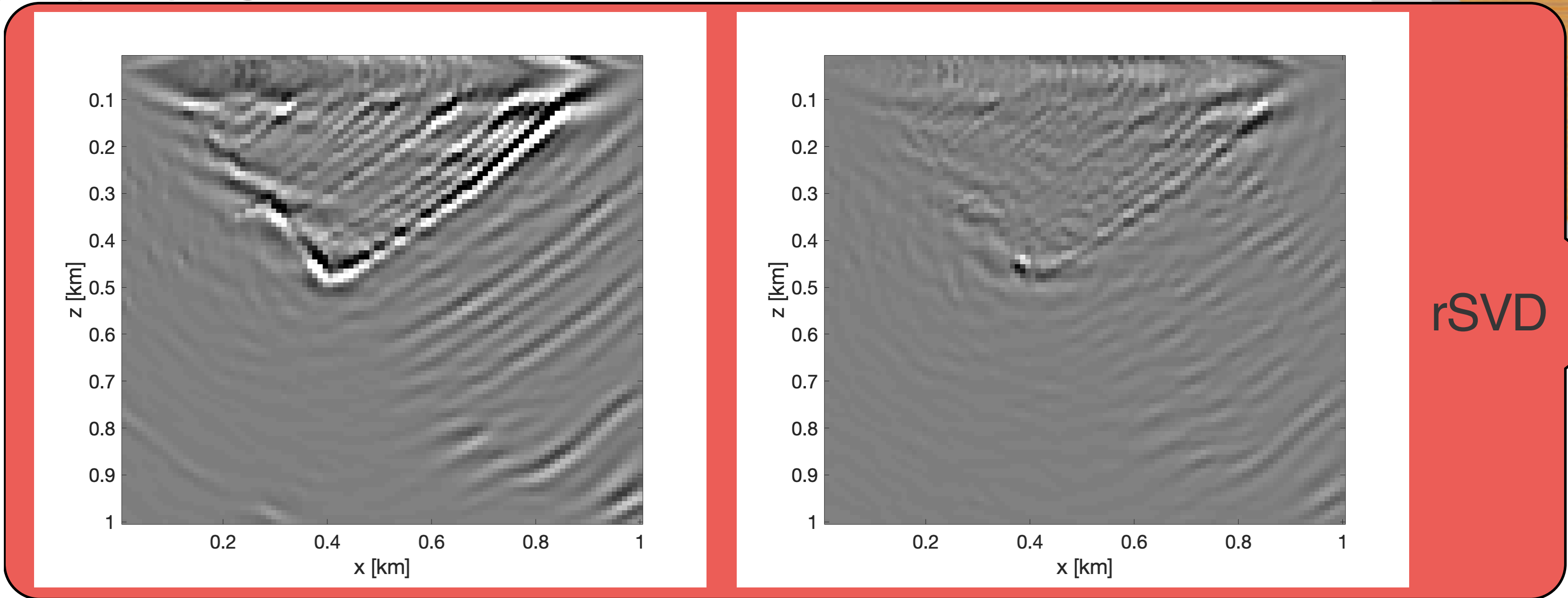
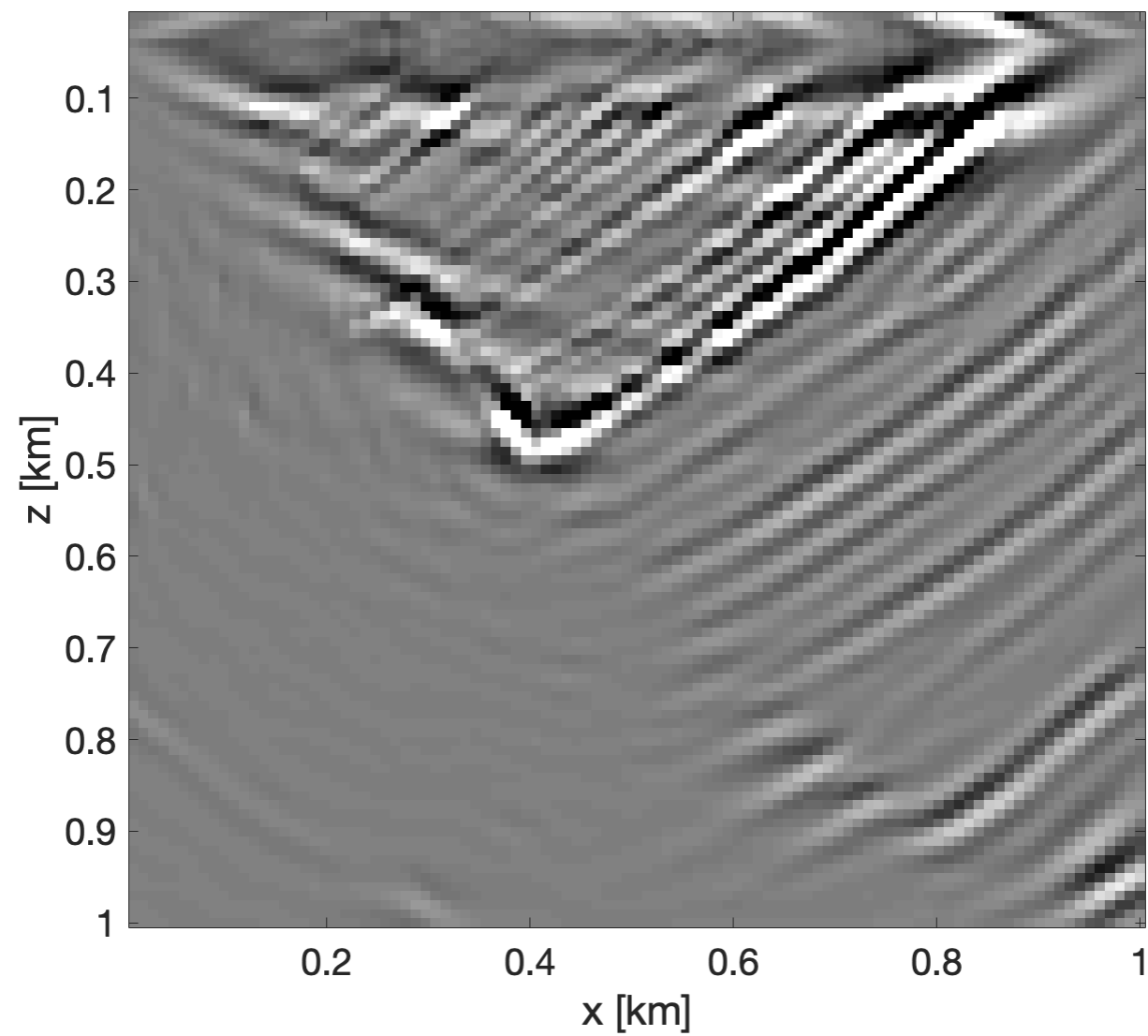
Simultaneous Iterations vs Block Krylov Iterations



Monochromatic EIV
at 25 Hz

$$n_s = 100$$

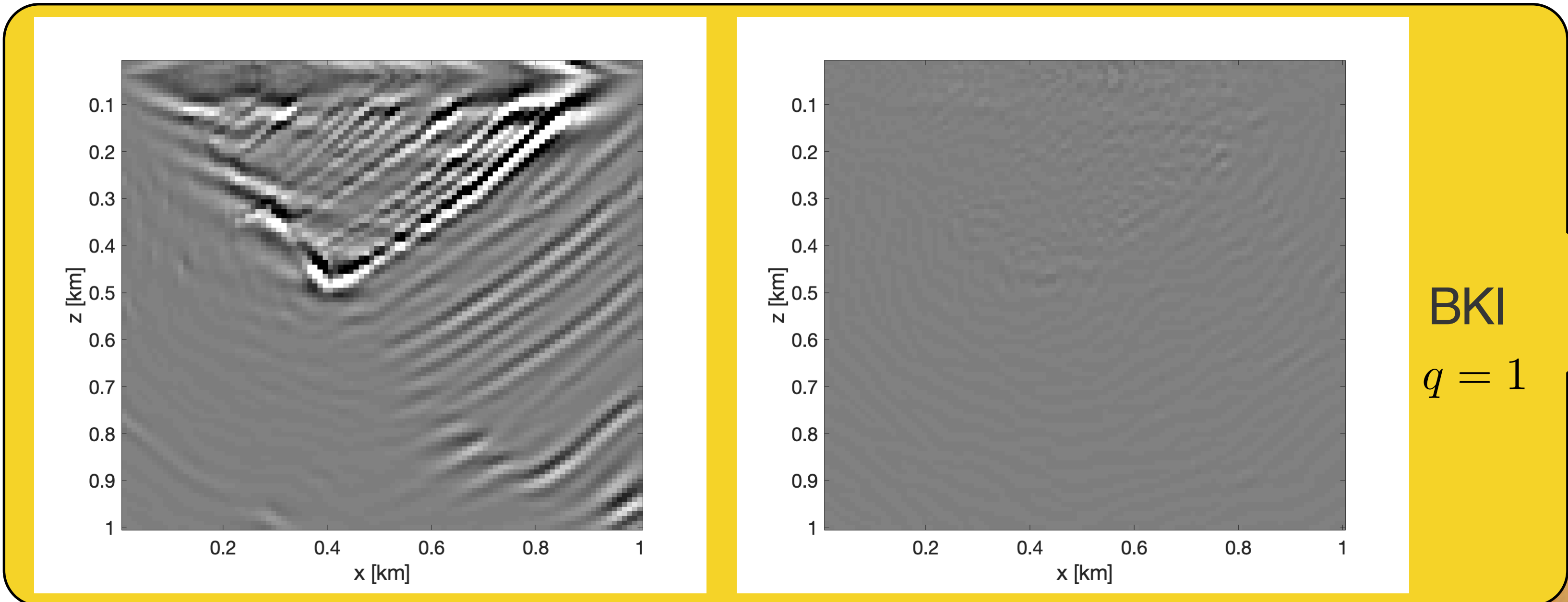
Block Krylov Iterations vs rSVD



Standard, $k=8$

Pros : higher accuracy w/o increasing n_p

Cons : additional WEs scale w/ q
: more flops in qr, svd



Full extended image volumes

$$\mathbf{E}_i = -\omega_i^2 \mathbf{V}_i \mathbf{U}_i^* = \mathbf{H}_i^{-*} \mathbf{P}_r^\top \dot{\mathbf{D}}_i \dot{\mathbf{Q}}_i^* \mathbf{P}_s \mathbf{H}_i^{-*}.$$

$$\begin{aligned} \mathbf{E} &= \overbrace{\mathcal{F} \circ \mathcal{A}^{-\top} \circ \mathcal{P}_r^\top [\dot{\mathbf{D}}]}^{\dot{\mathbf{V}}} \cdot \overbrace{(\mathcal{F} \circ \mathcal{A}^{-1} \circ \mathcal{P}_s^\top [\dot{\mathbf{Q}}])^*}^{\dot{\mathbf{U}}^*} \\ &= \mathcal{F} \circ \mathcal{A}^{-\top} \circ \mathcal{P}_r^\top \circ \mathcal{F}^\top [\dot{\mathbf{D}} \cdot \dot{\mathbf{Q}}^* \cdot (\mathcal{F} \circ \mathcal{P}_s \circ \mathcal{A}^{-\top} \circ \mathcal{F}^\top [\mathbf{I}])]. \end{aligned}$$

Probing

$$\mathbf{Y}_i = \mathbf{E}_i \mathbf{W}_i = \mathbf{H}_i^{-*} \mathbf{P}_r^\top \dot{\mathbf{D}}_i \dot{\mathbf{Q}}_i^* \mathbf{P}_s \mathbf{H}_i^{-*} \mathbf{W}_i$$

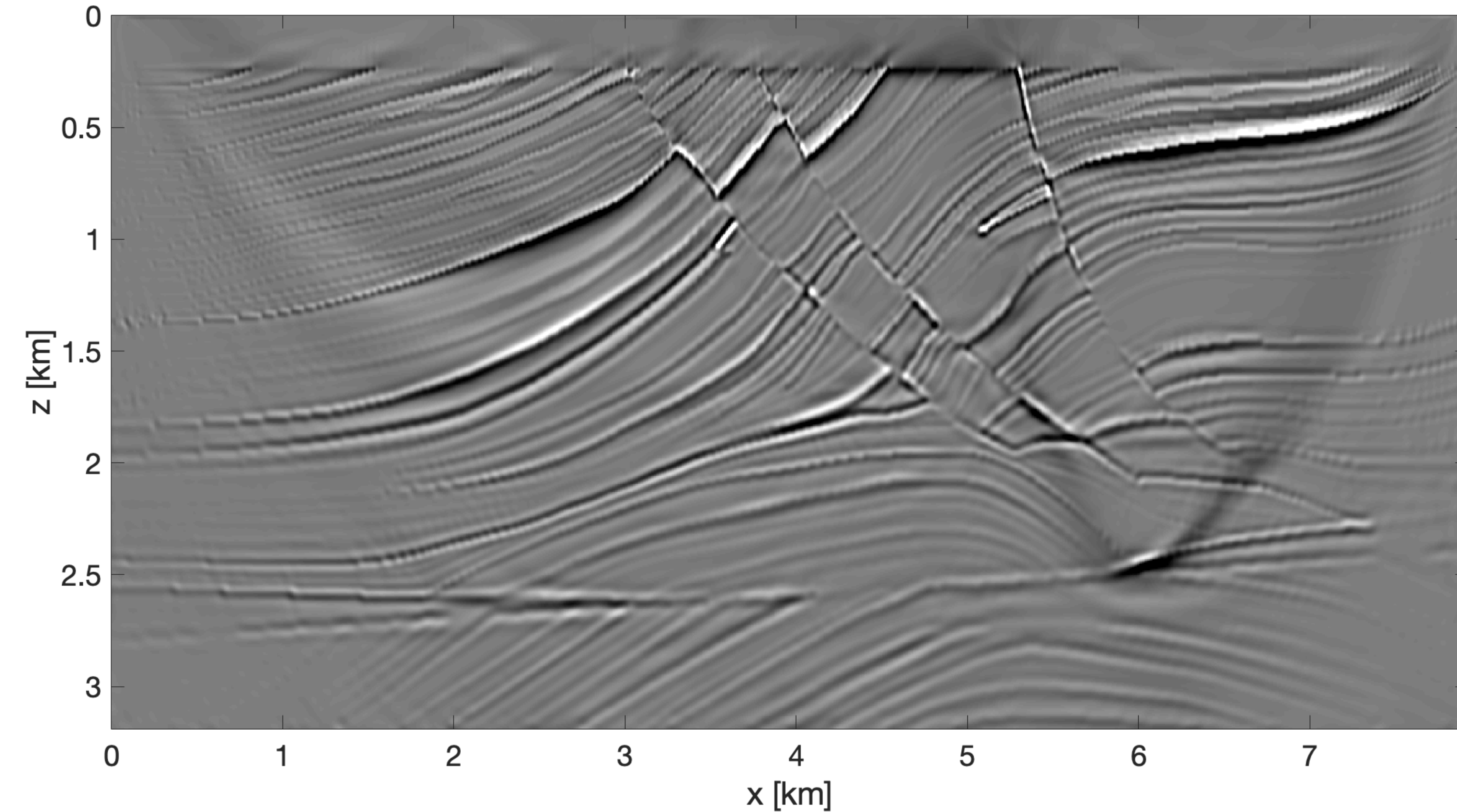
$$\begin{aligned} \mathbf{Y} &= \mathcal{F} \circ \mathcal{A}^{-\top} \circ \mathcal{P}_r^\top [\dot{\mathbf{D}} *_t (\mathcal{P}_s \circ \mathcal{A}^{-\top} [\mathbf{W}])] \\ &= \mathcal{F} \circ \mathcal{A}^{-\top} \circ \mathcal{P}_r^\top \circ \mathcal{F}^\top [\dot{\mathbf{D}} \cdot (\mathcal{F} \circ \mathcal{P}_s \circ \mathcal{A}^{-\top} [\mathbf{W}])] \end{aligned}$$

Fast Fourier transform

Gaussian noise

Band-limited Gaussian noise

Traditional RTM vs extracted RTM

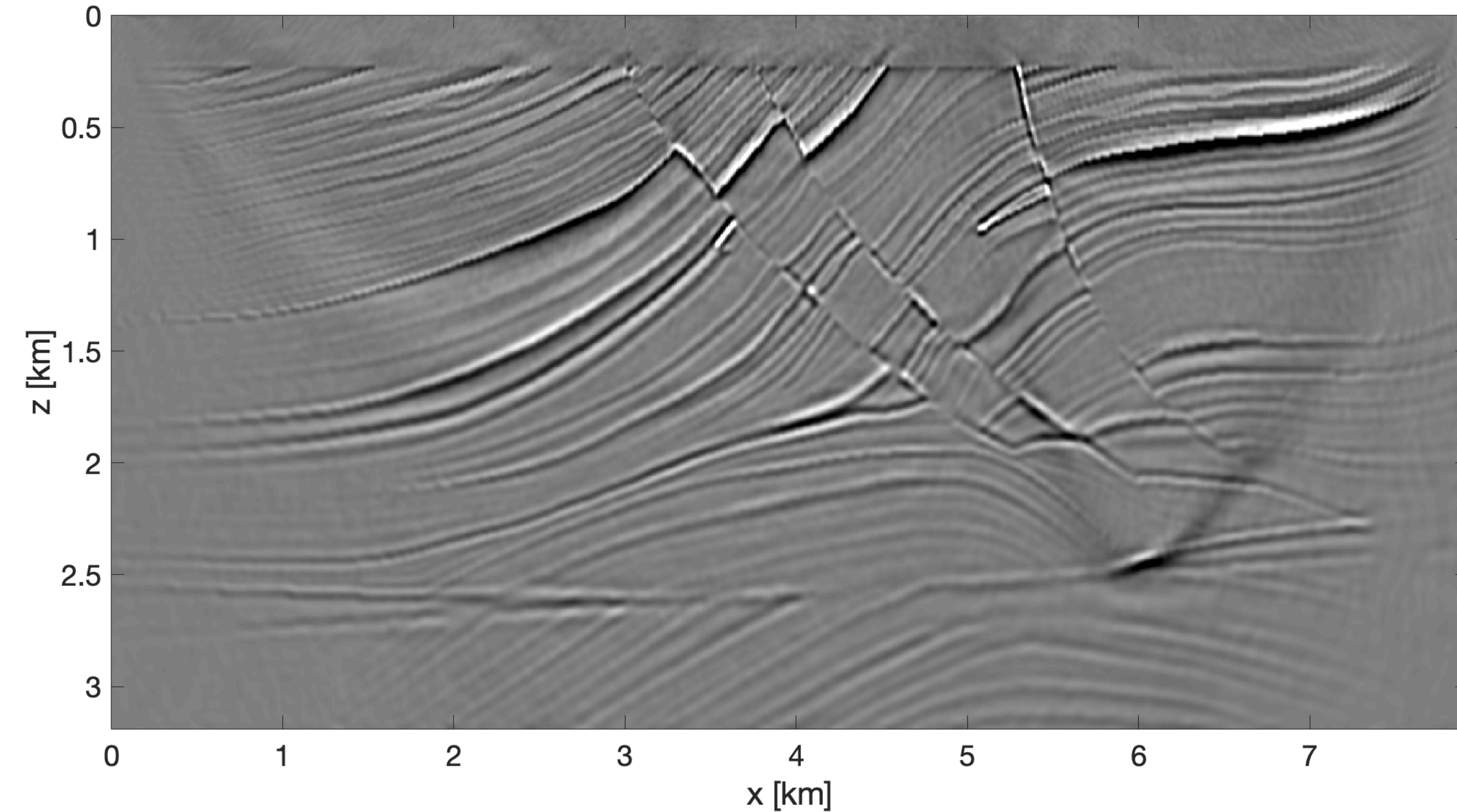


Experiment:

- 650 co-located sources and receivers
- Ricker wavelet centered at 23 Hz
- Direct wave removed

$$2n_s = 1300$$

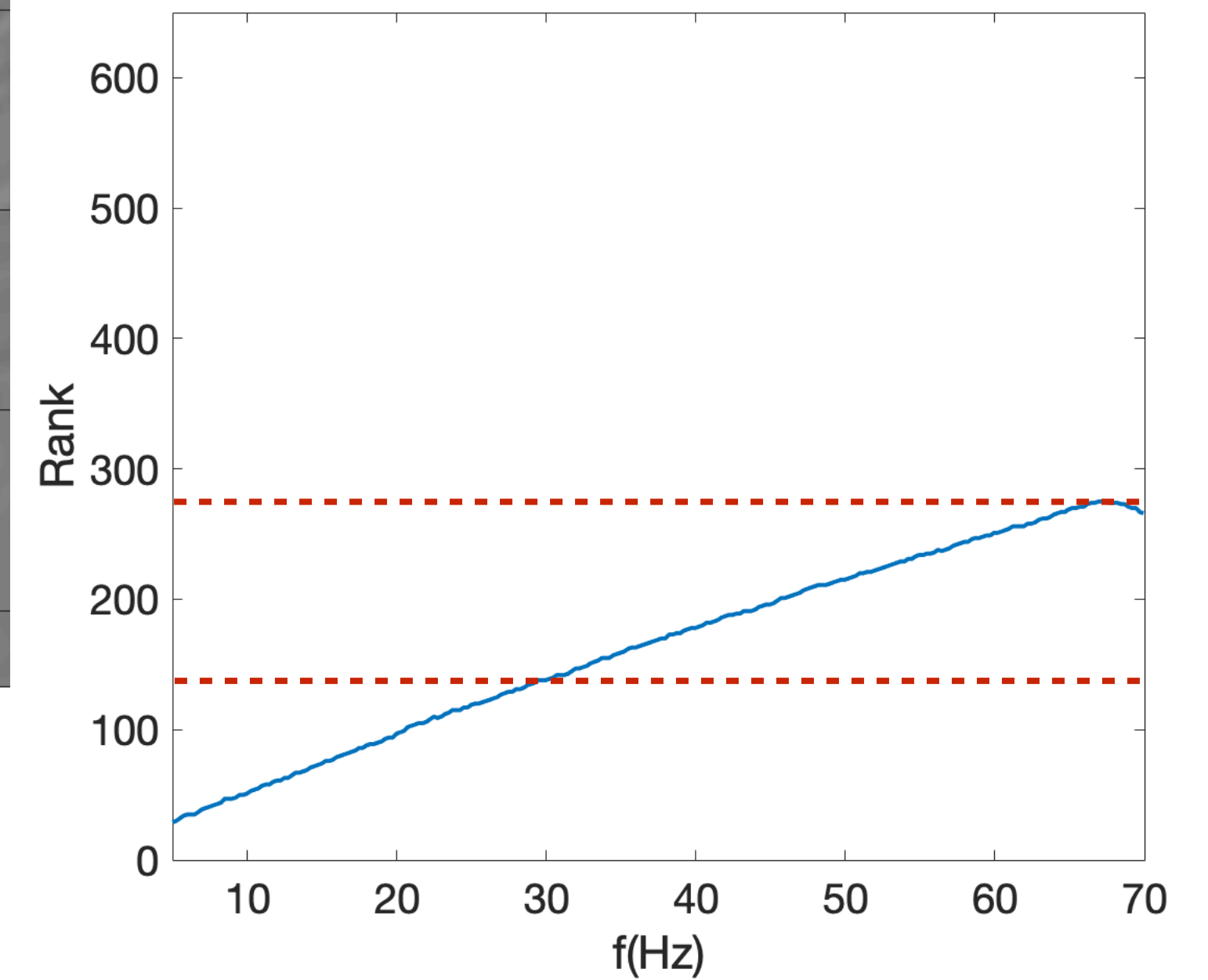
Traditional RTM vs extracted RTM



$$10n_p = 1300$$

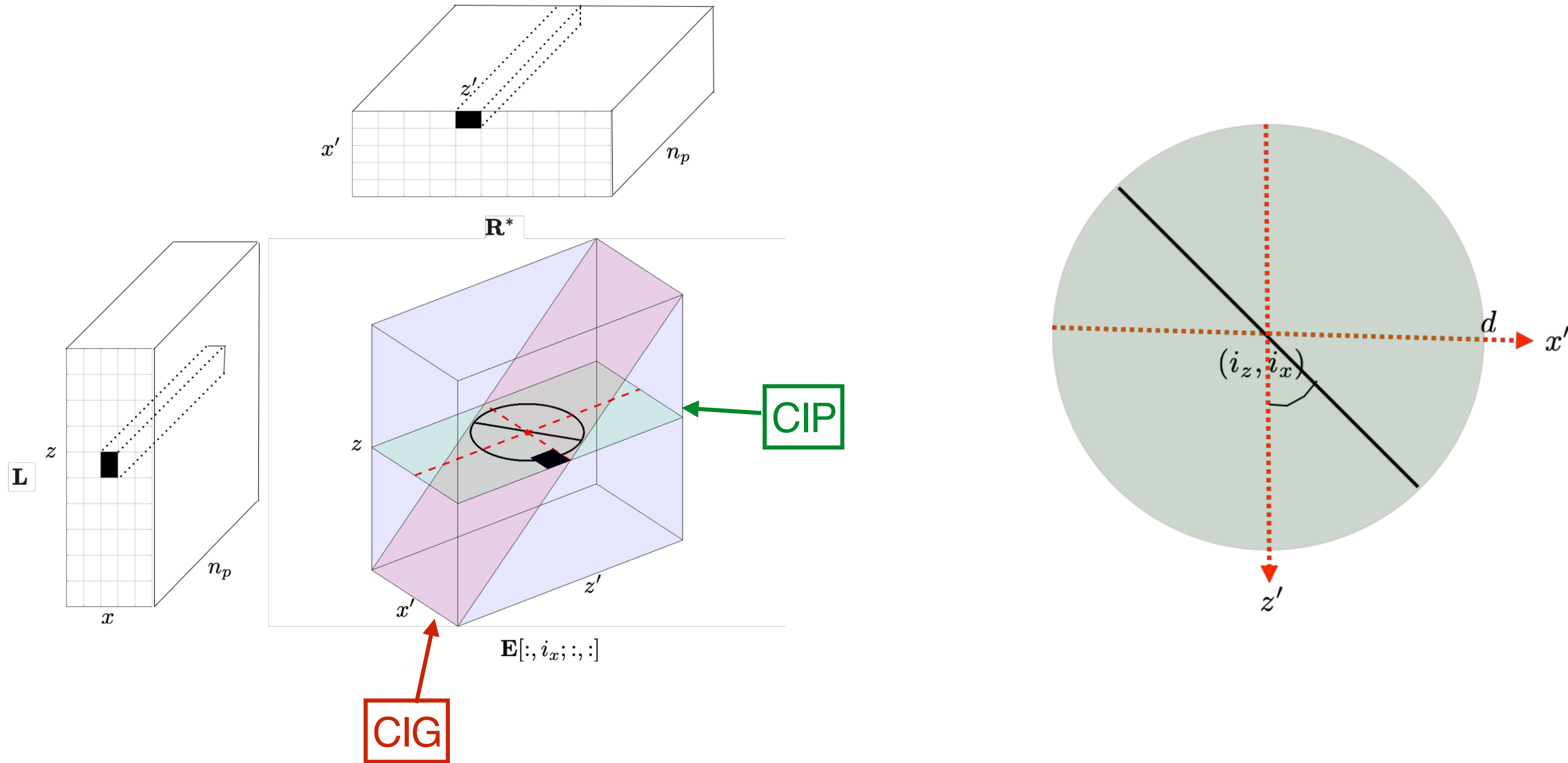
Experiment:

- BKI, $q = 1$
- Probing size $n_p = 130$

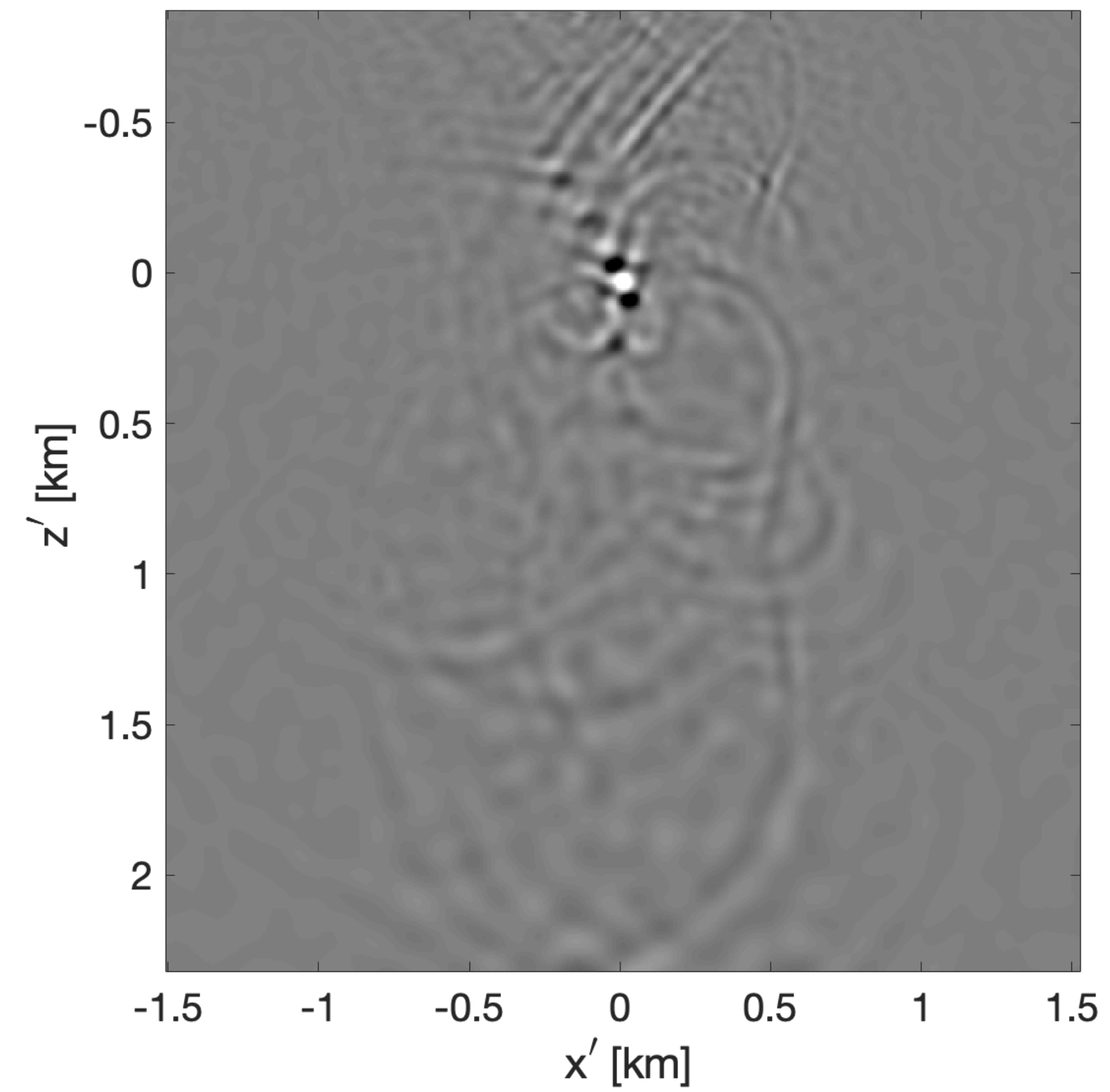
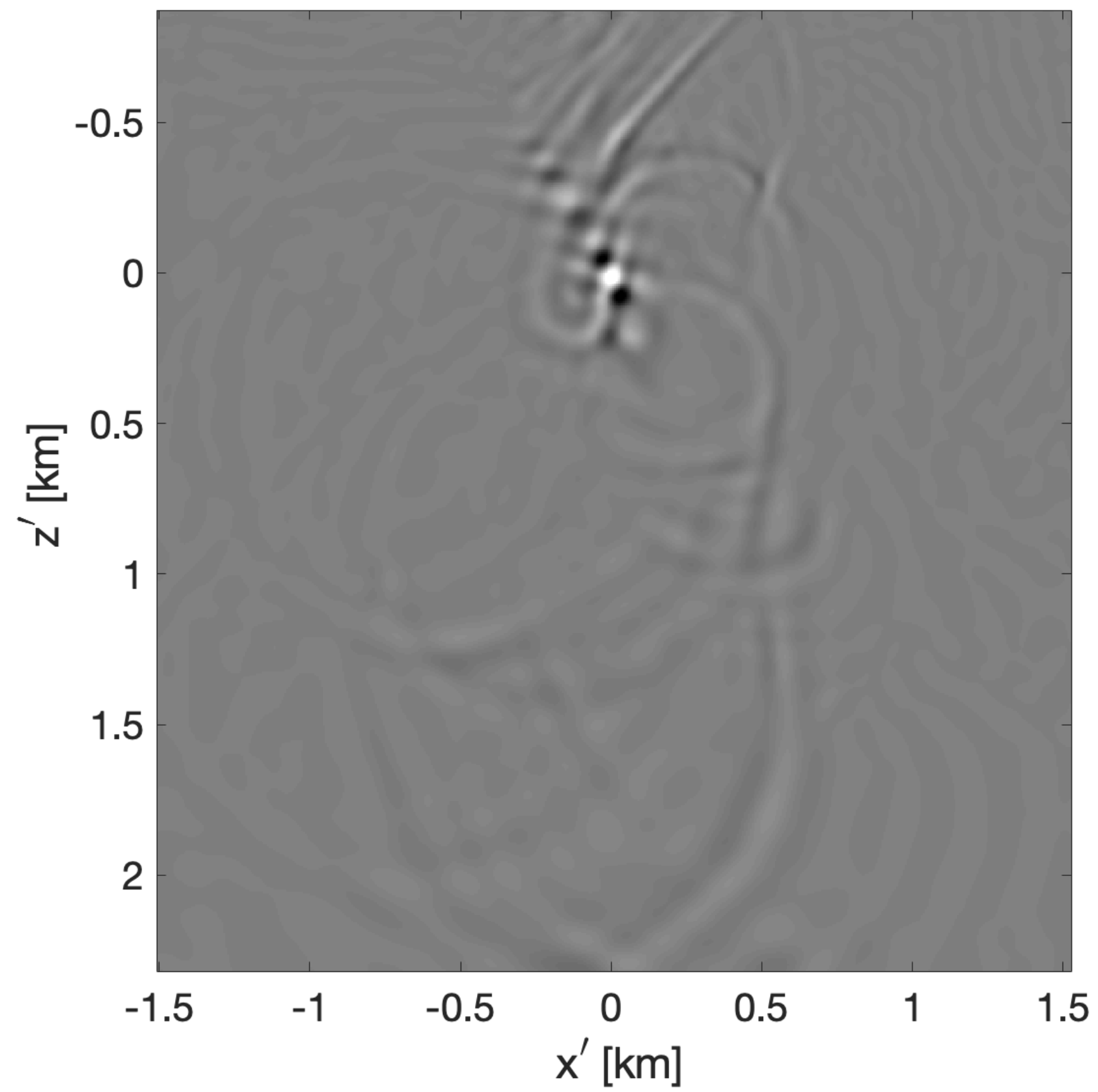


Selected rank to capture 95% of data's energy

Traditional image gathers vs extracted image gathers

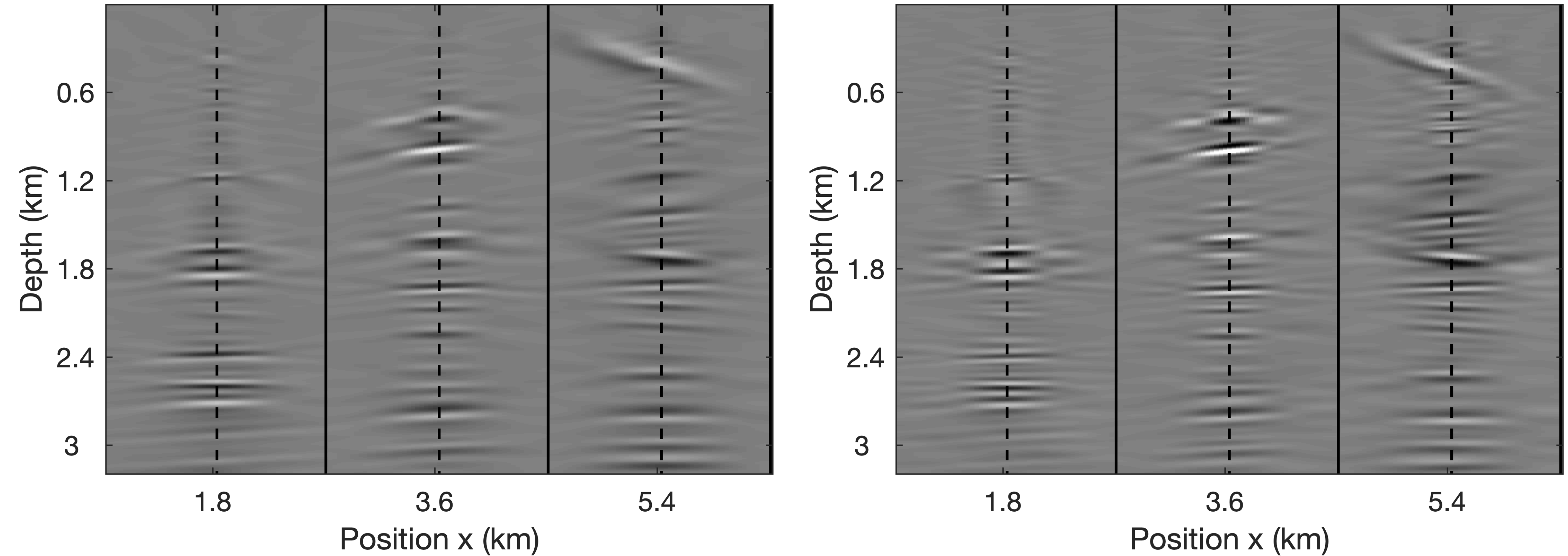


Traditional CIP vs extracted CIP



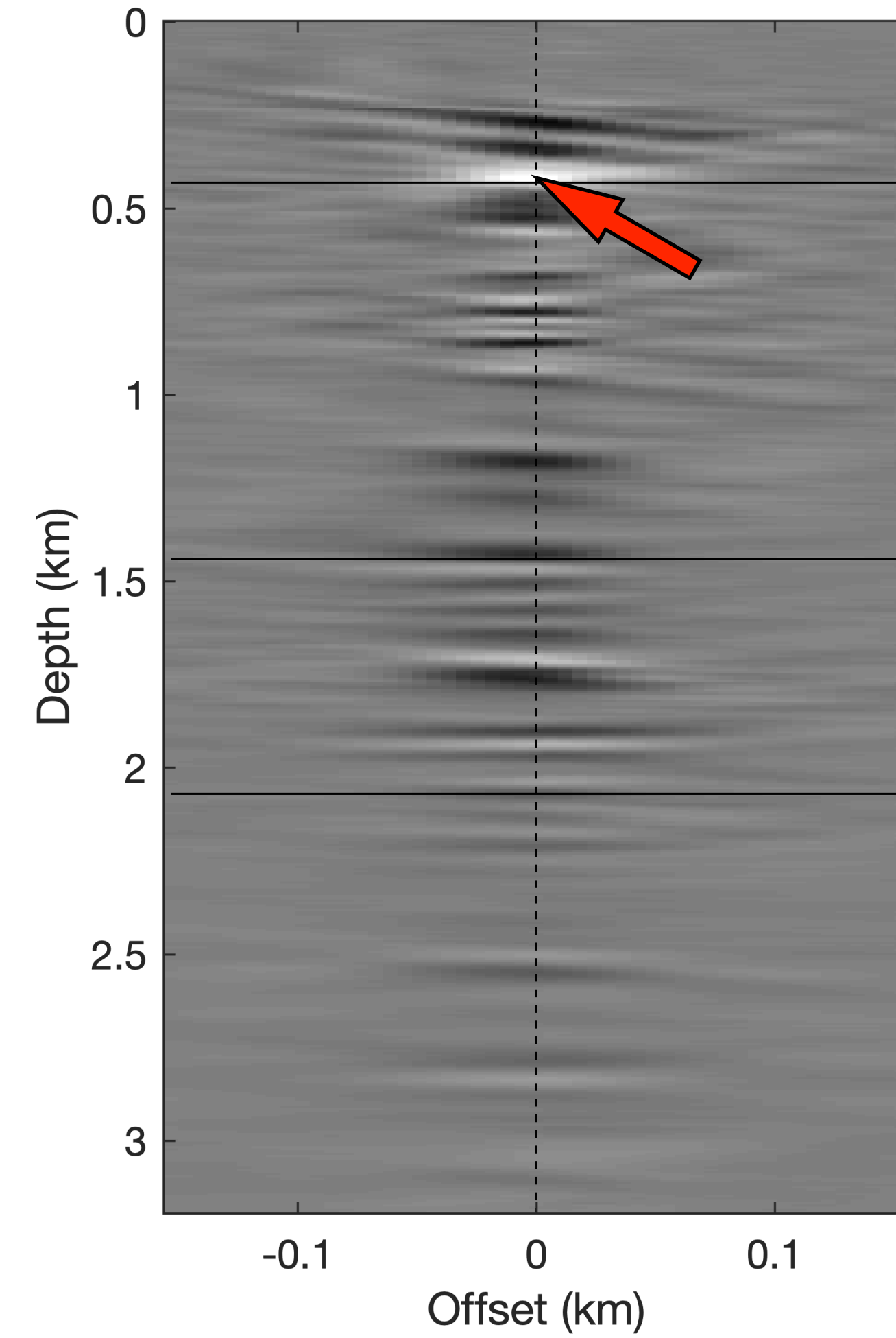
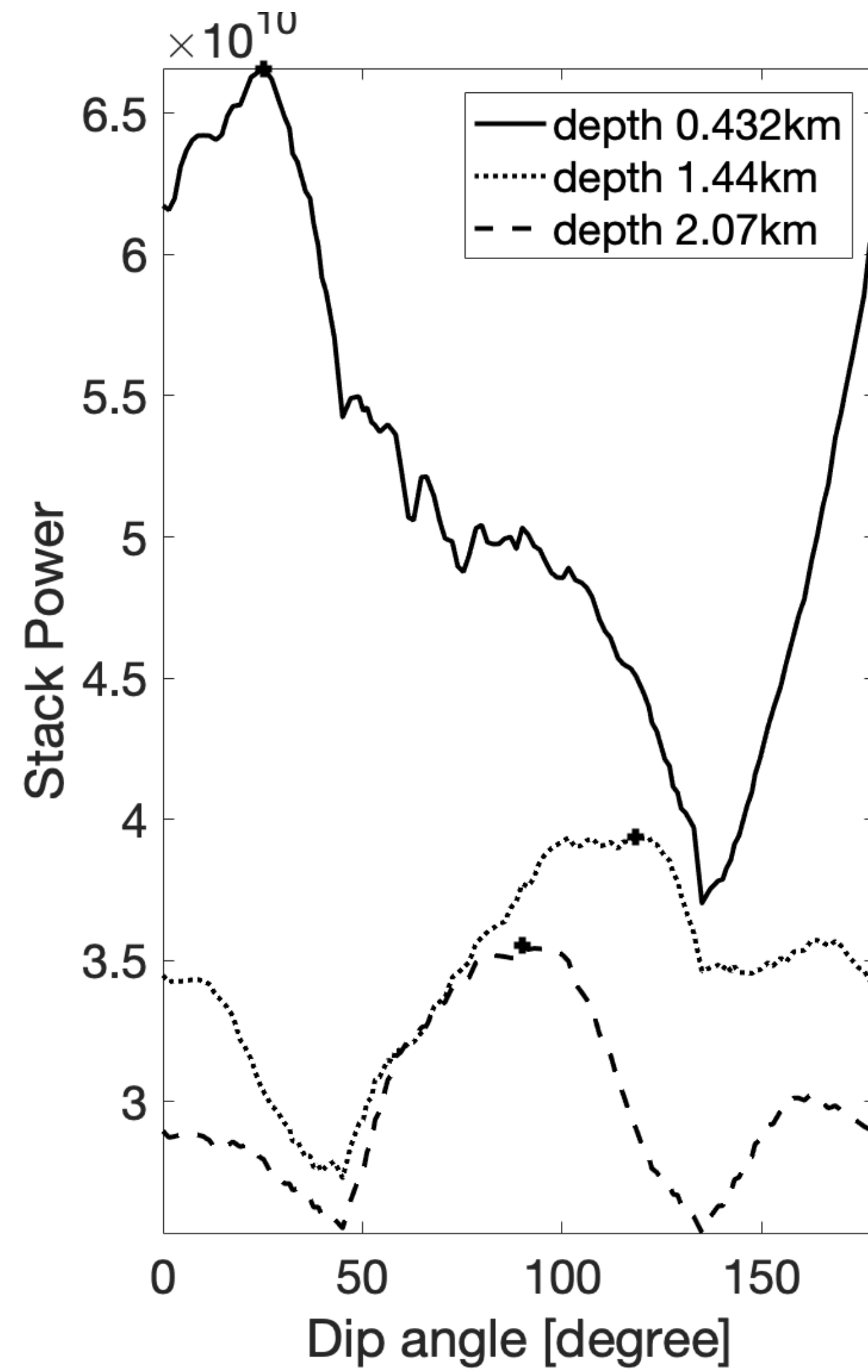
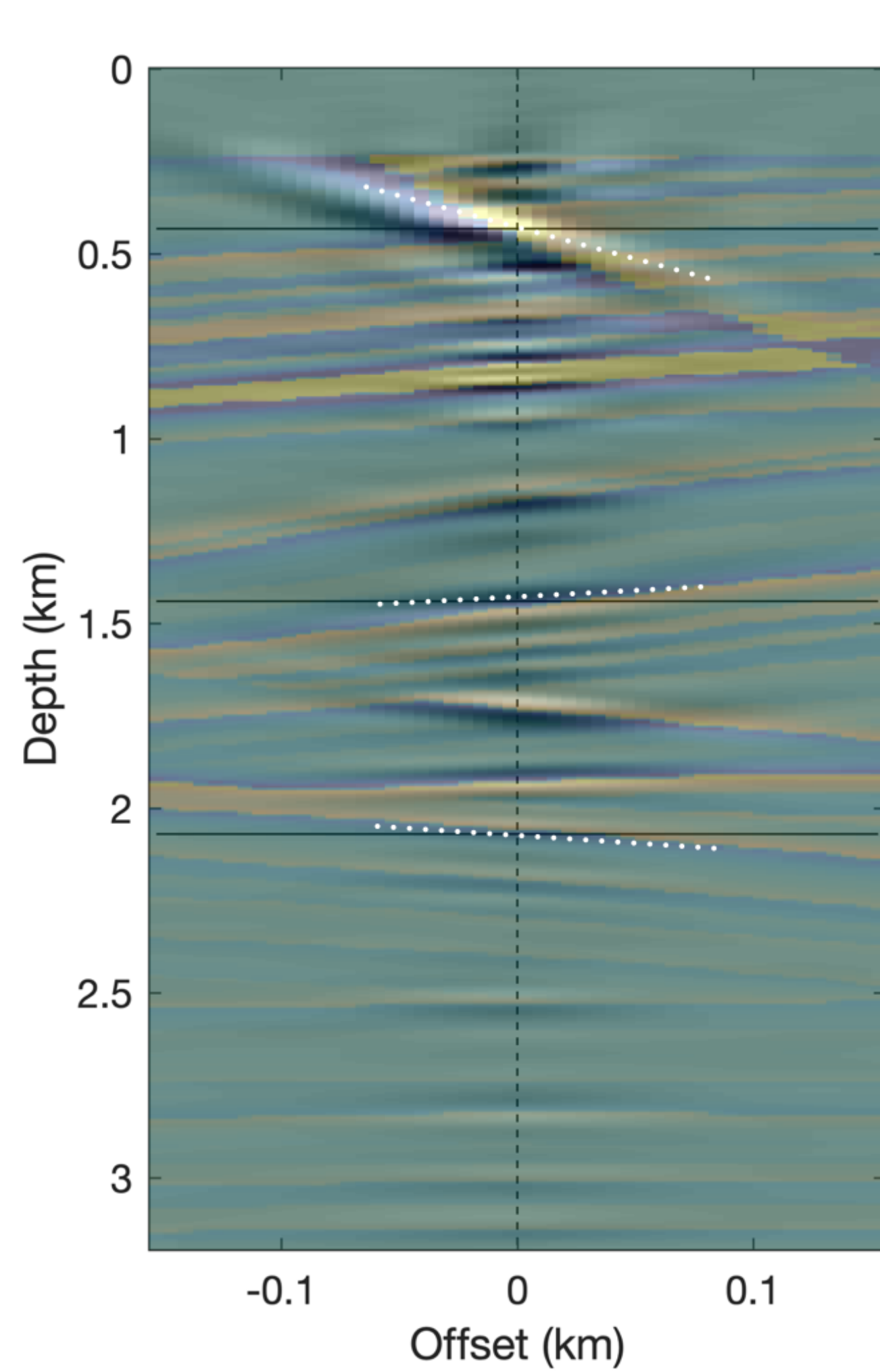
$2n_s$ VS 0

Traditional CIG vs extracted CIG

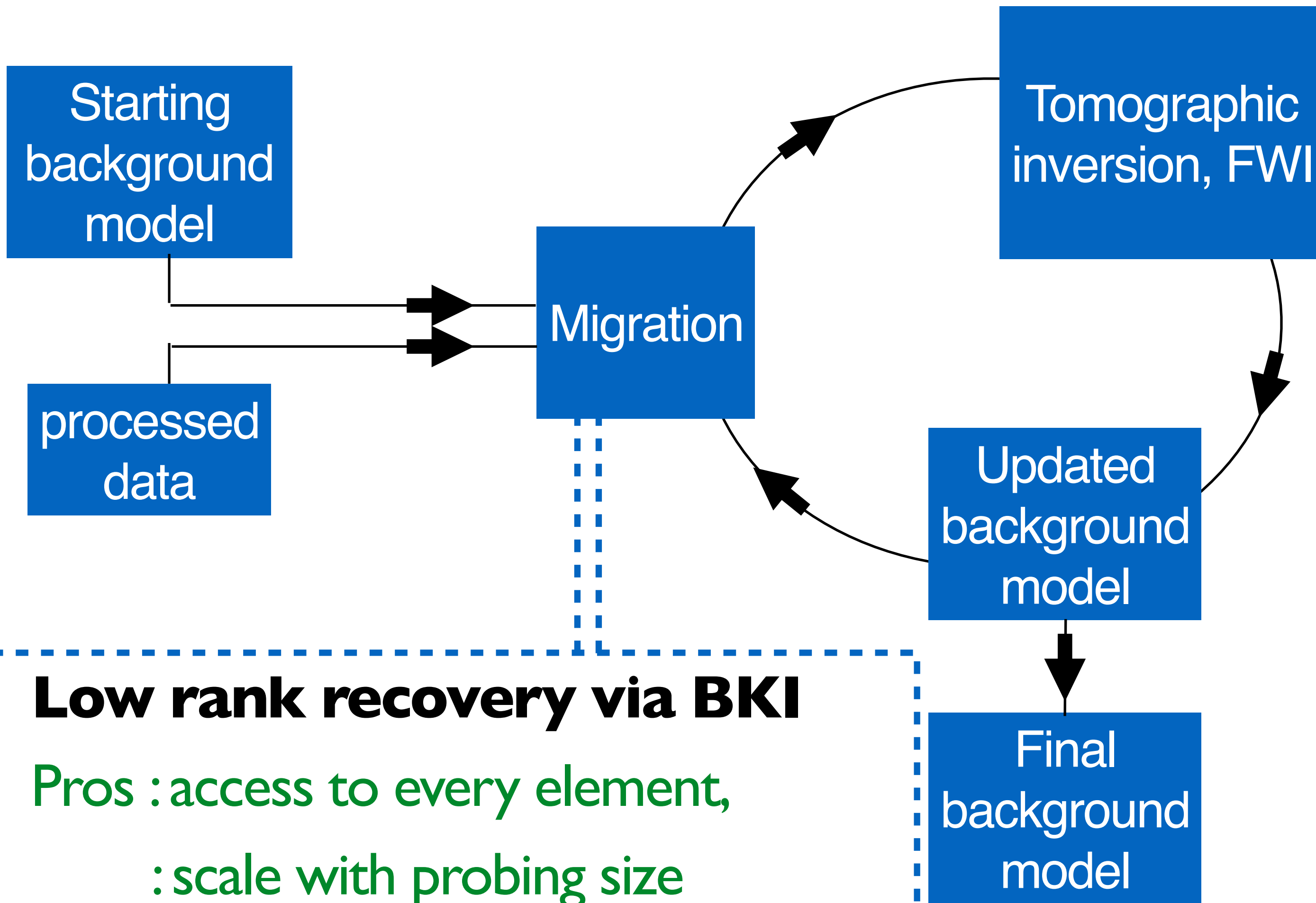


$2n_s$ VS 0

Geologic dip-corrected CIG



Mapping via Invariance relationship



Low rank recovery via BKI

Pros : access to every element,
: scale with probing size

~~Cons : additional WEs scale w/ power~~
~~: more flops in qr, svd~~

Scenario 1
BKI w/
model 1
 $\{\mathbf{L}_1, \mathbf{R}_1\}$

qr, svd
WEs scale w/
power

Scenario 2
BKI w/
model 2
 $\{\mathbf{L}_2, \mathbf{R}_2\}$

Scenario 3
mapping
 $\{\mathbf{L}_1, \mathbf{R}_1\} \rightarrow \{\mathbf{L}_2, \mathbf{R}_2\}$

WEs scale w/
only probing
size

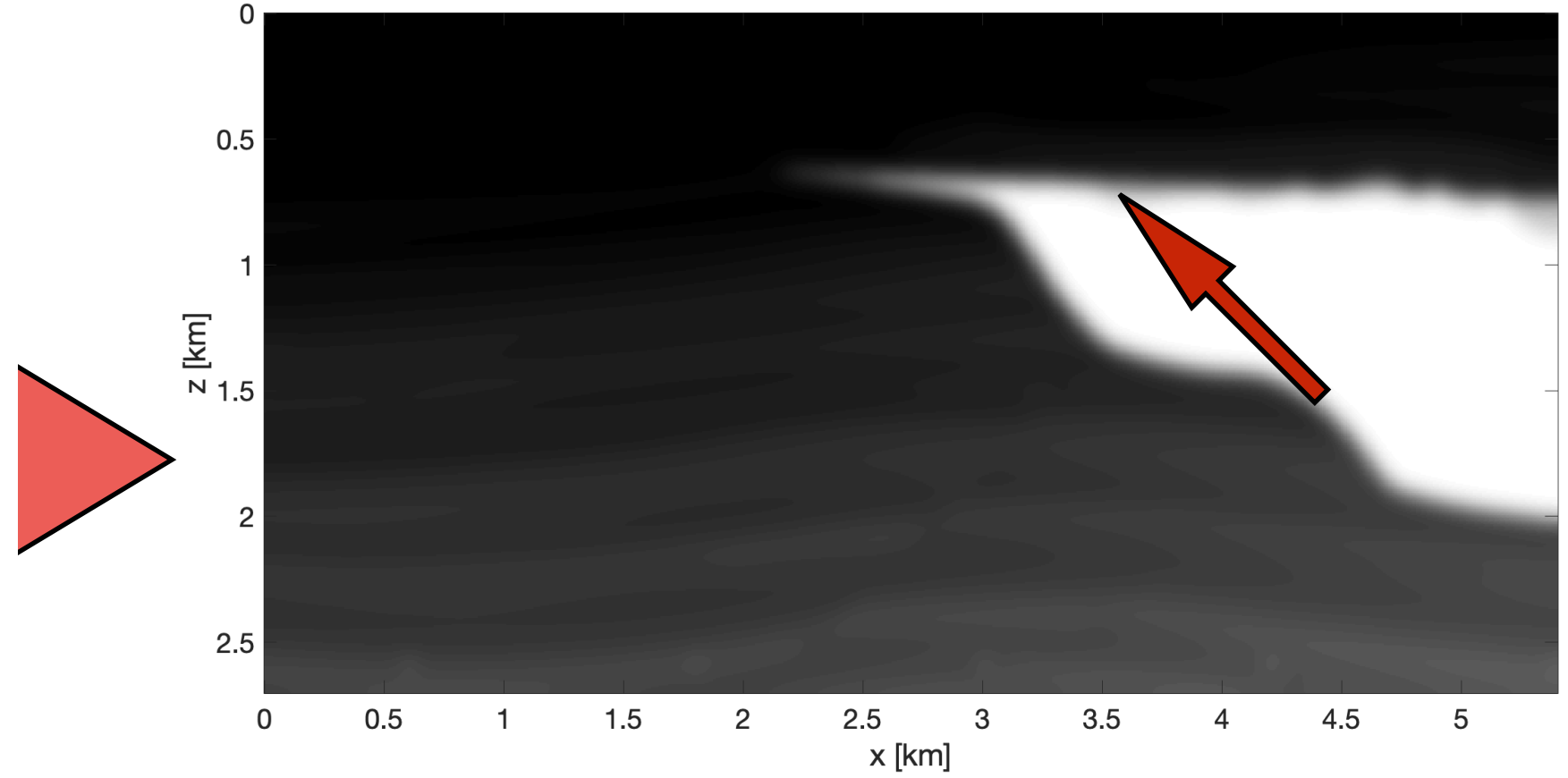
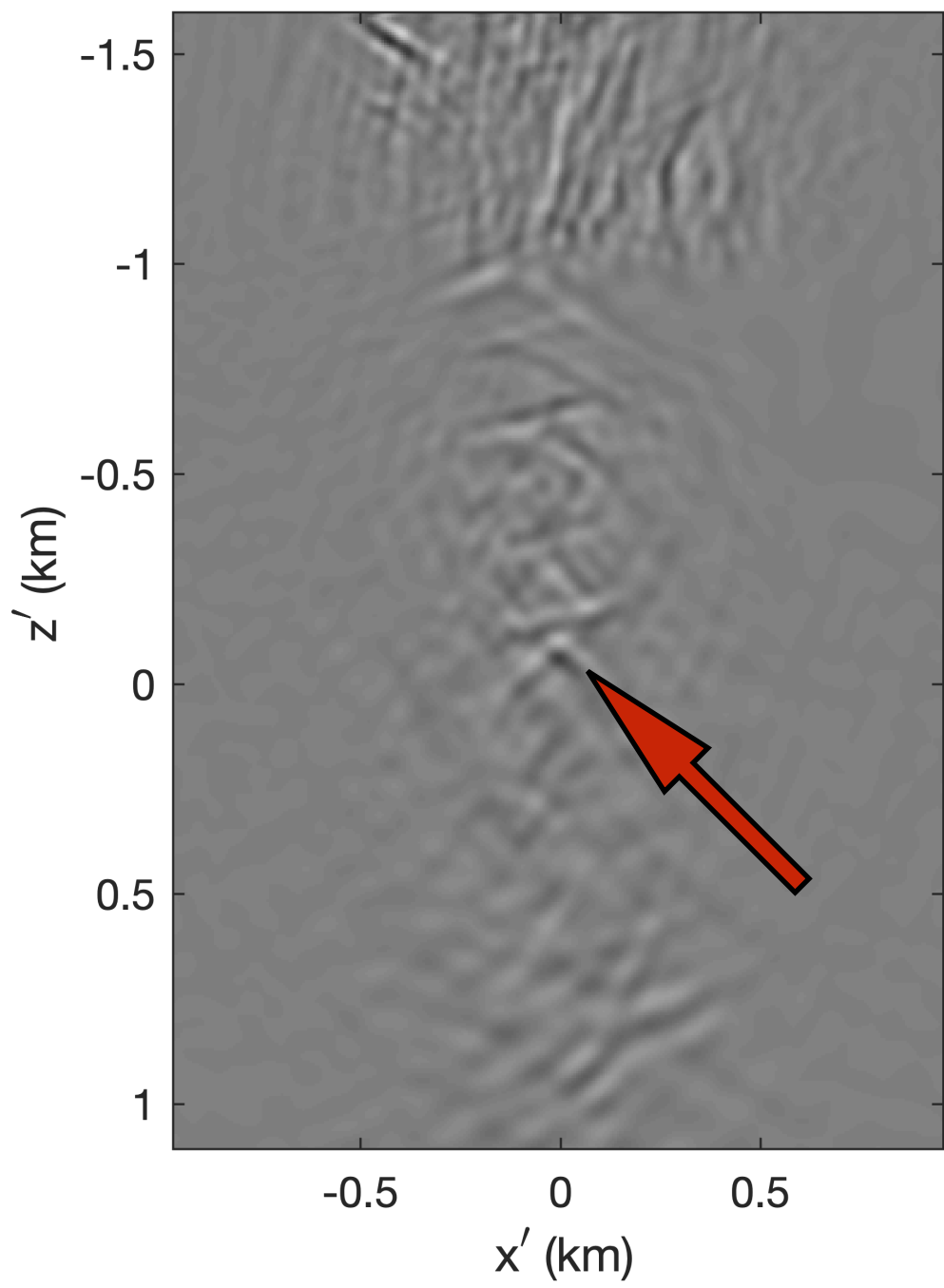
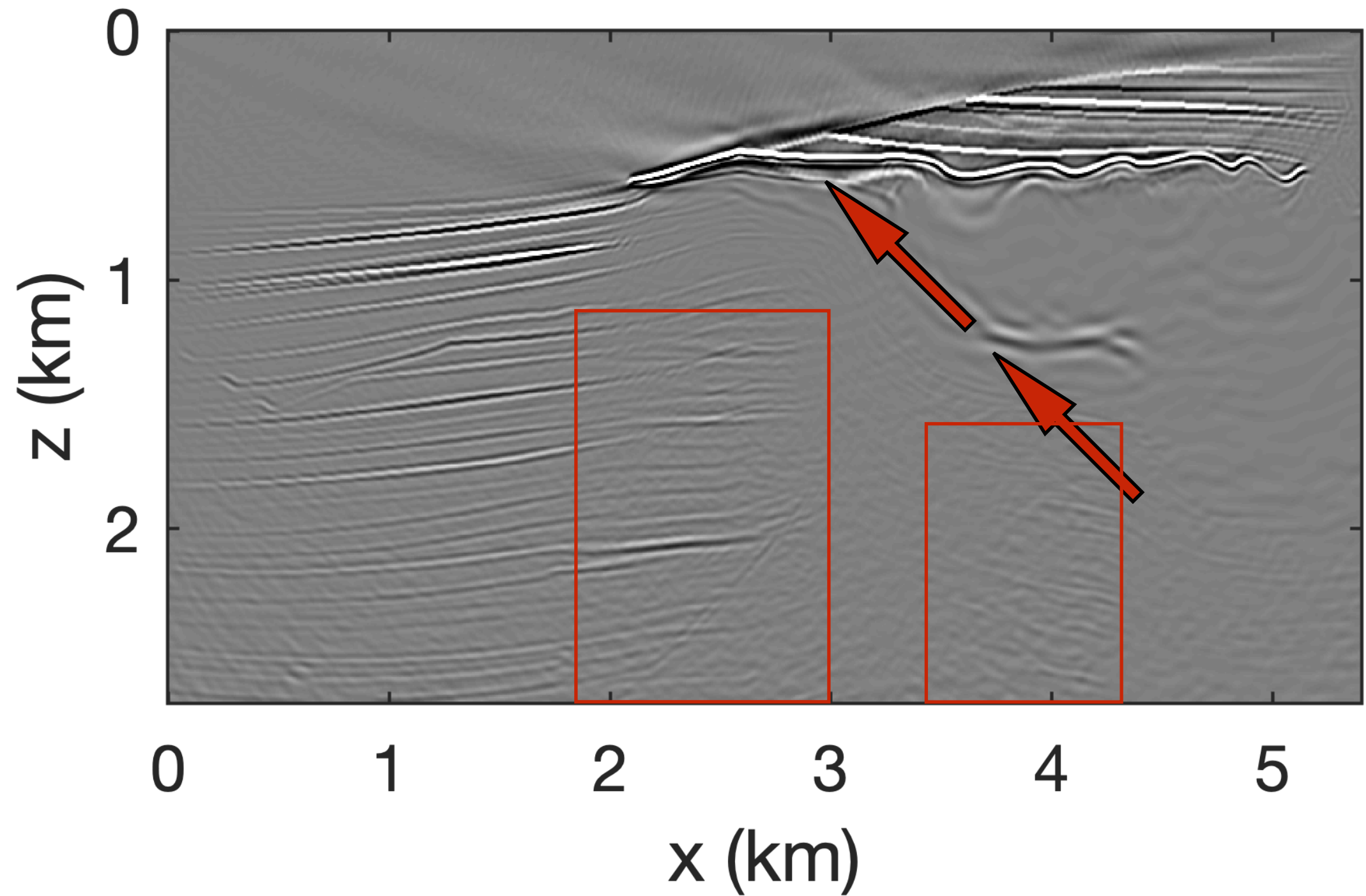
Invariant

$$\mathbf{E}_i = \mathbf{H}_i^{-*} \mathbf{P}_r^\top \mathbf{D}_i \mathbf{Q}_i^* \mathbf{P}_s \mathbf{H}_i^{-*}.$$

$$\mathbf{L}_2 = \mathcal{F} \circ \mathcal{A}_2^{-\top} \circ \mathcal{A}_1^\top [\mathbf{L}_1] \text{ with } \mathbf{L}_1 = \mathcal{F}^\top [\mathbf{L}_1],$$

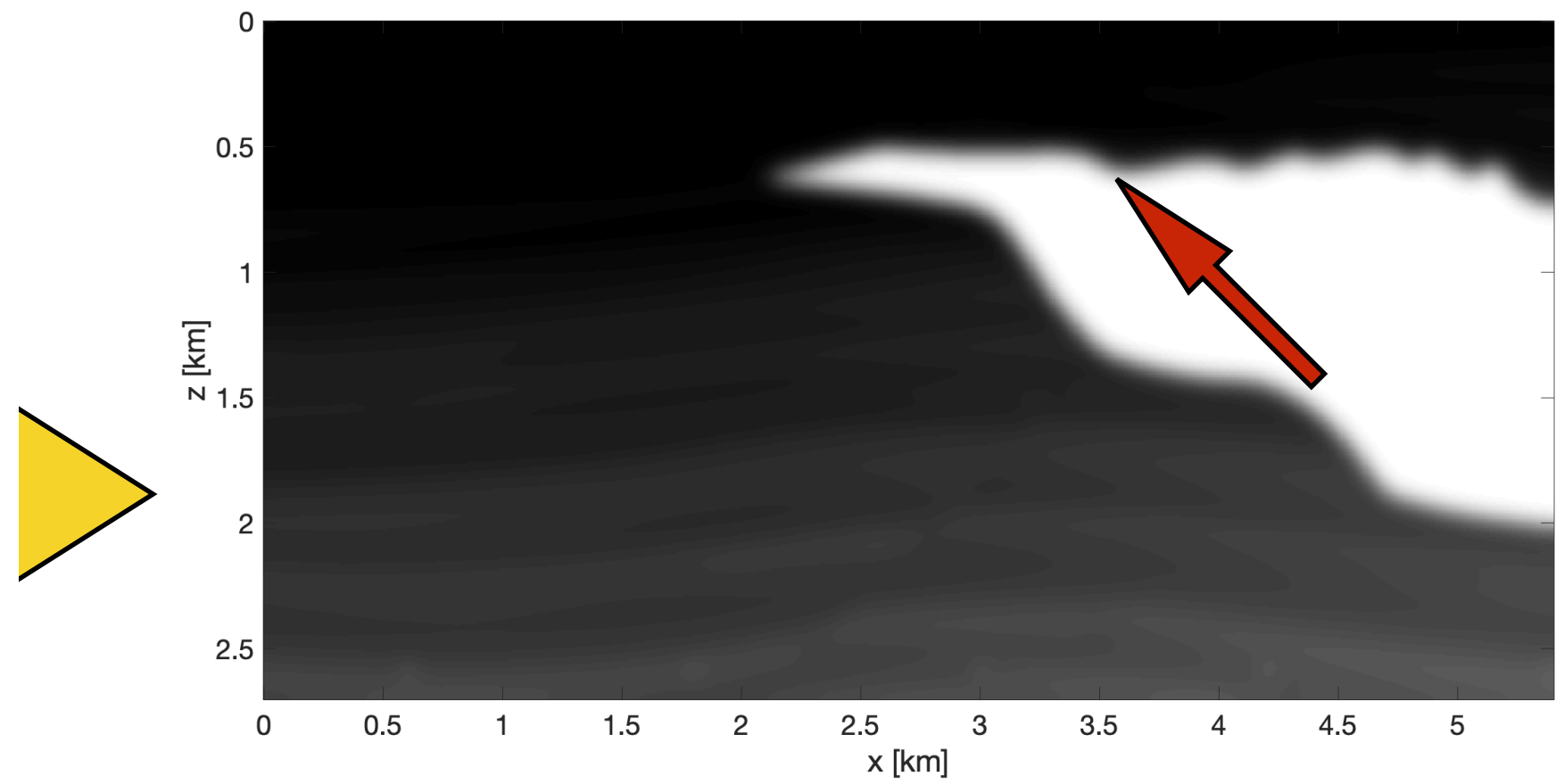
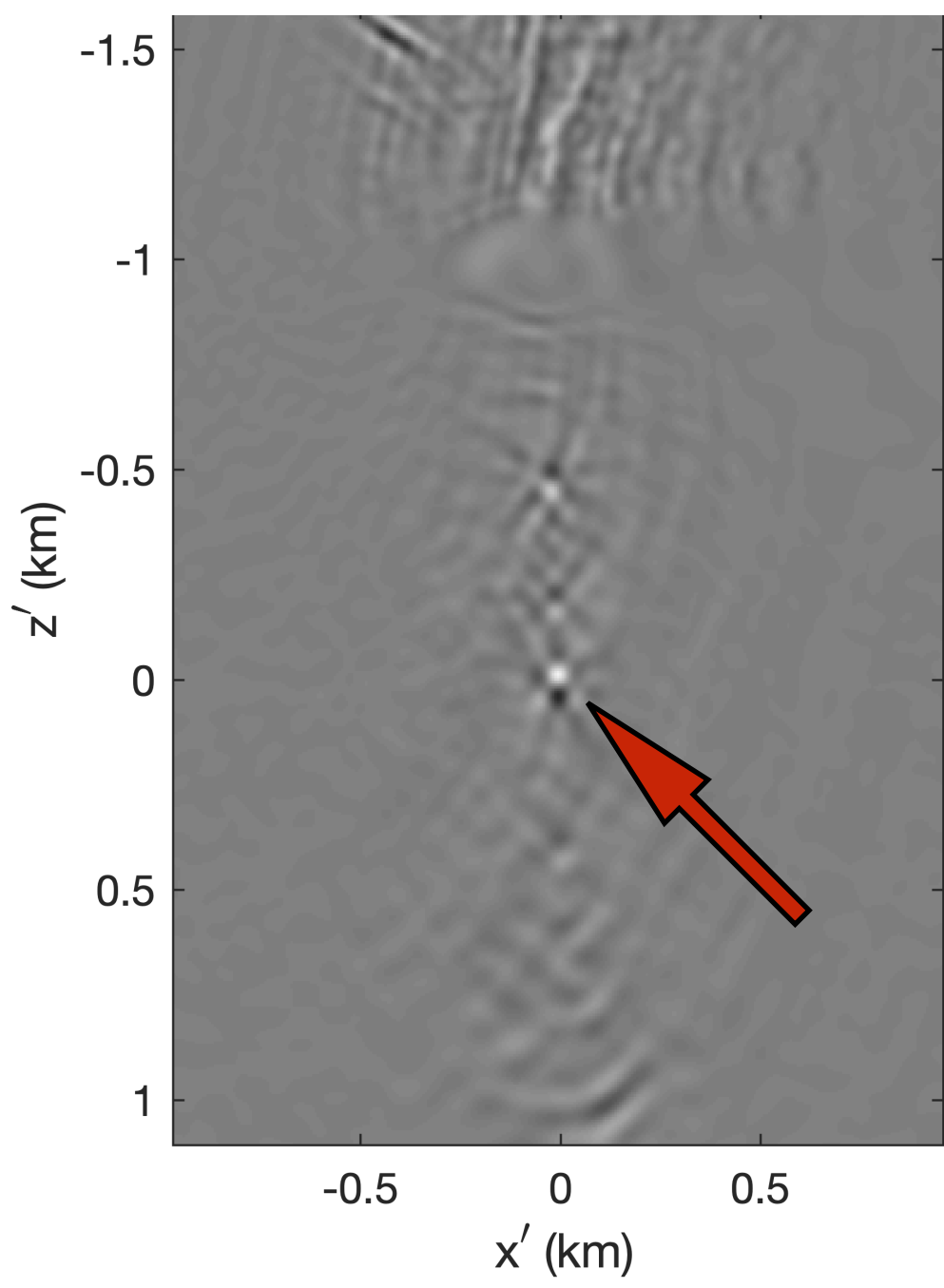
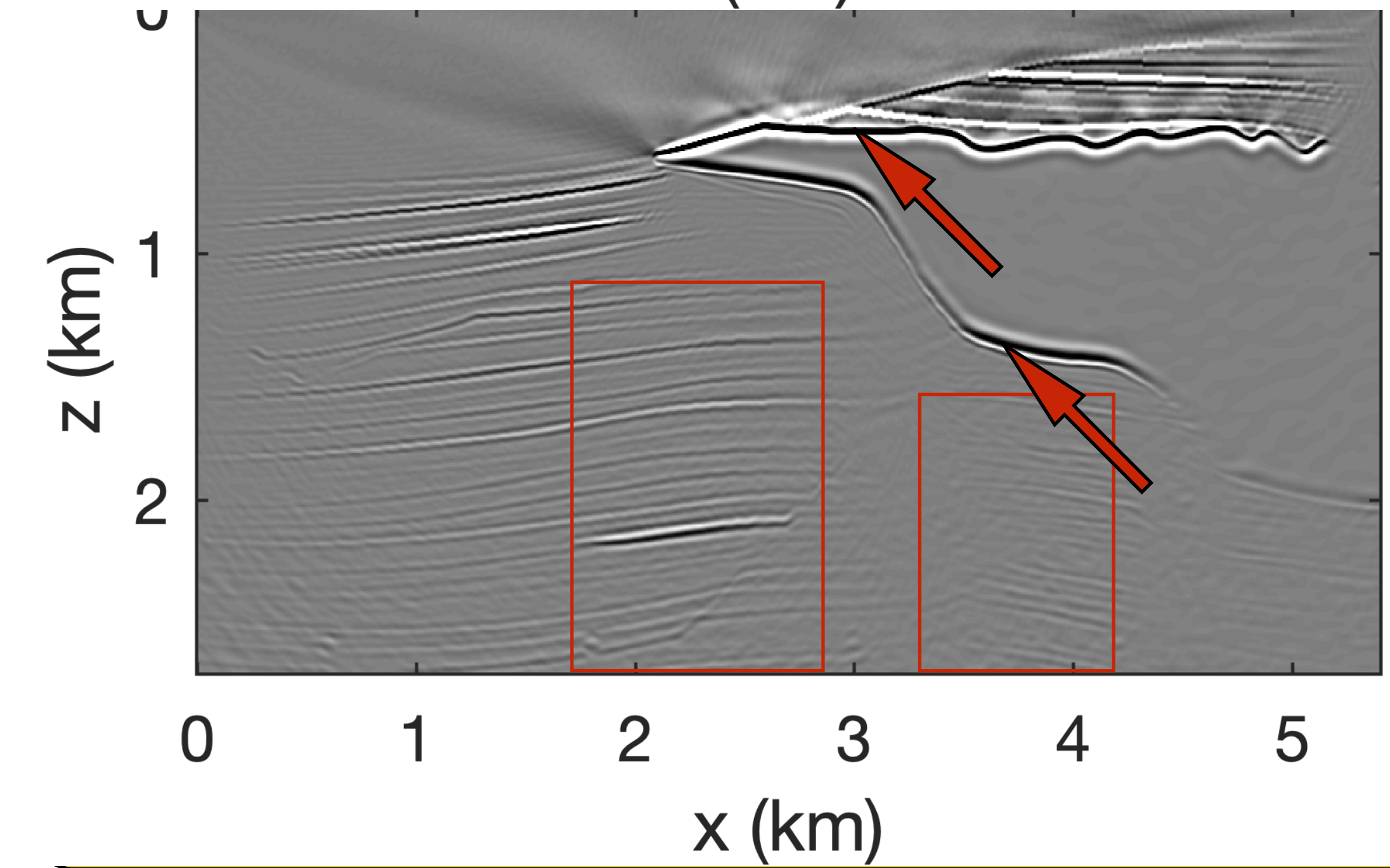
$$\mathbf{R}_2 = \mathcal{F} \circ \mathcal{A}_2^{-1} \circ \mathcal{A}_1 [\mathbf{R}_1] \text{ with } \mathbf{R}_1 = \mathcal{F}^\top [\mathbf{R}_1].$$

Scenario 1

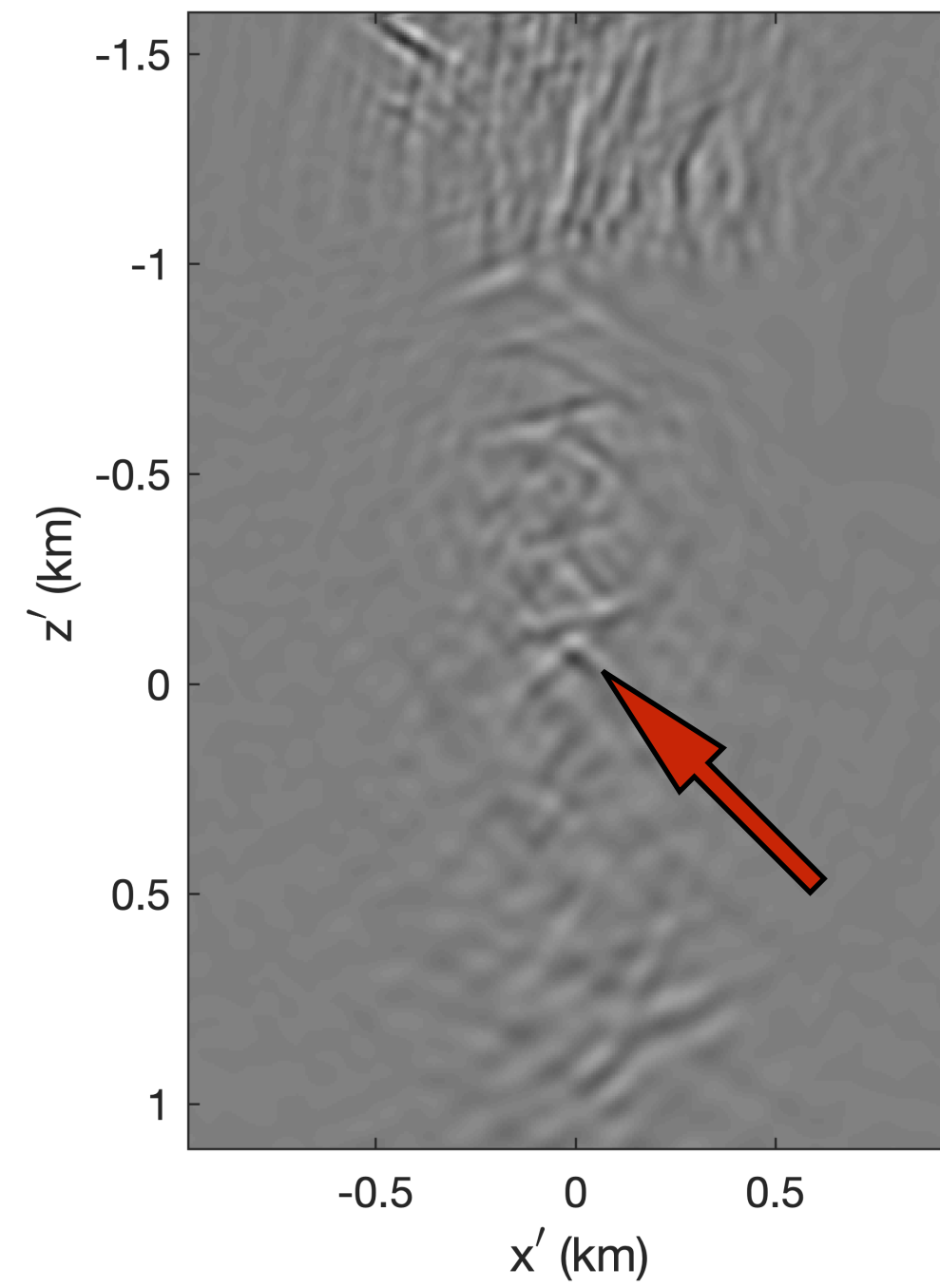
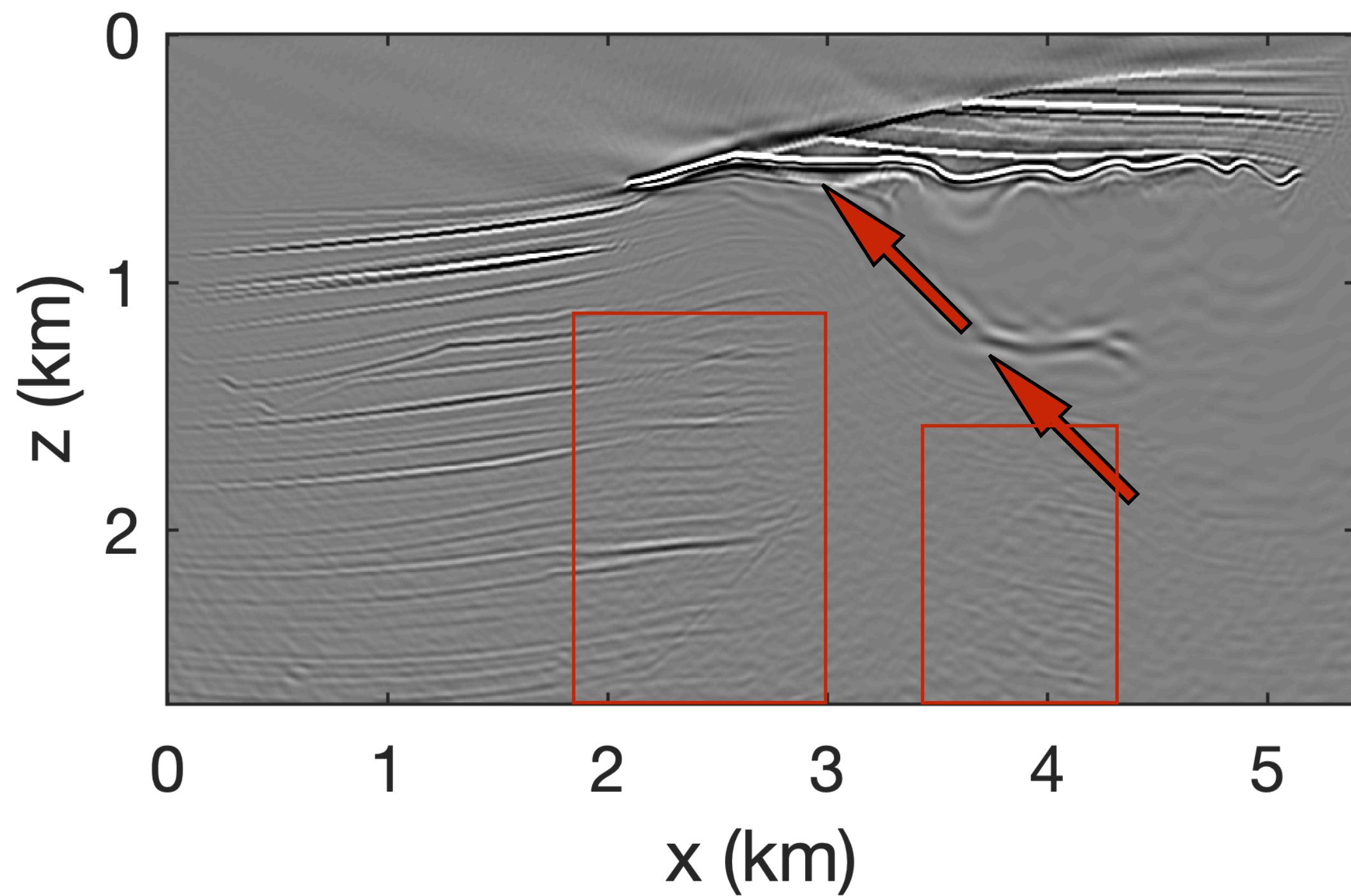


$10n_p = 1000$, QR+SVD

Scenario 2

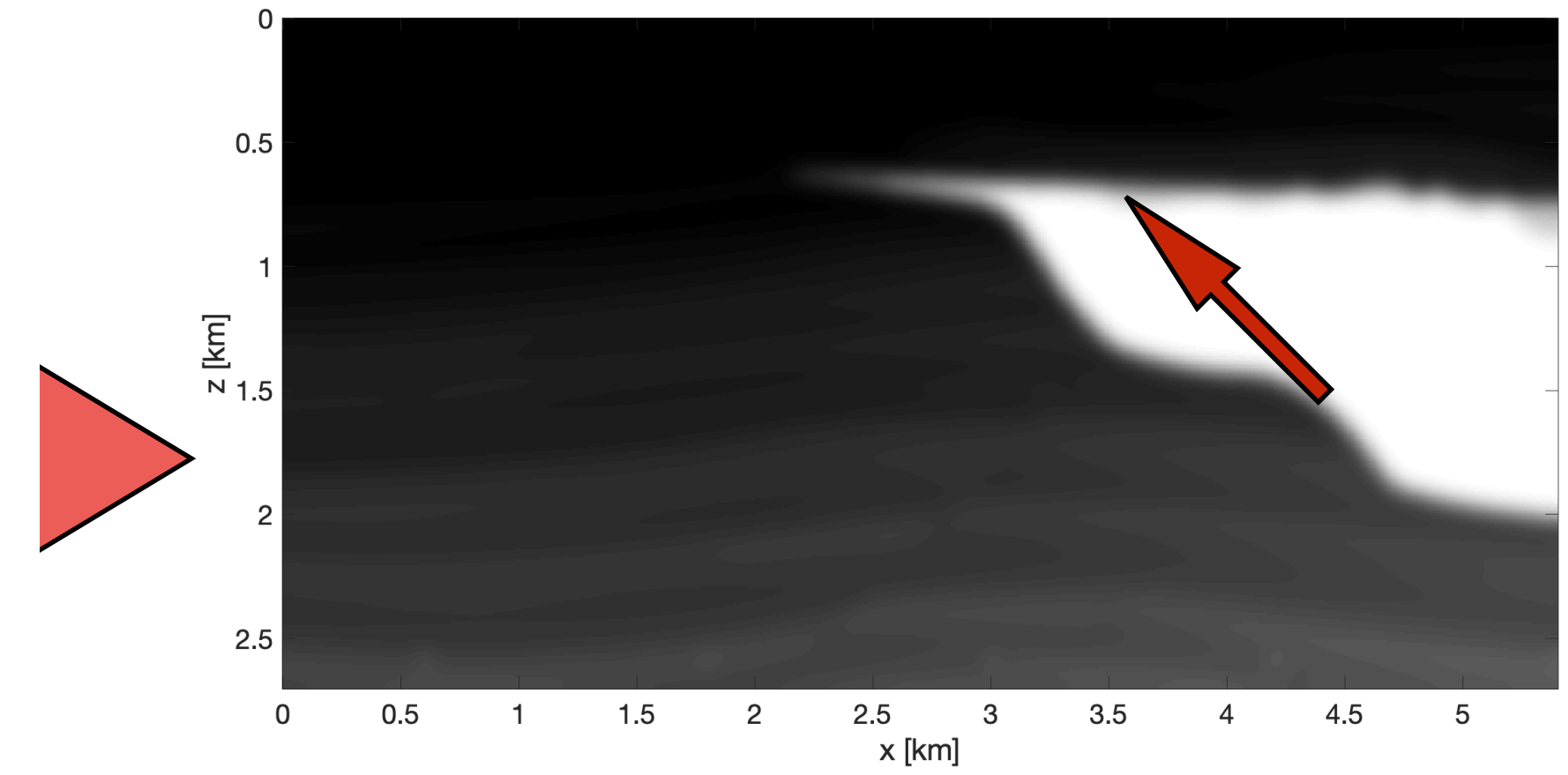


$10n_p = 1000$, QR+SVD

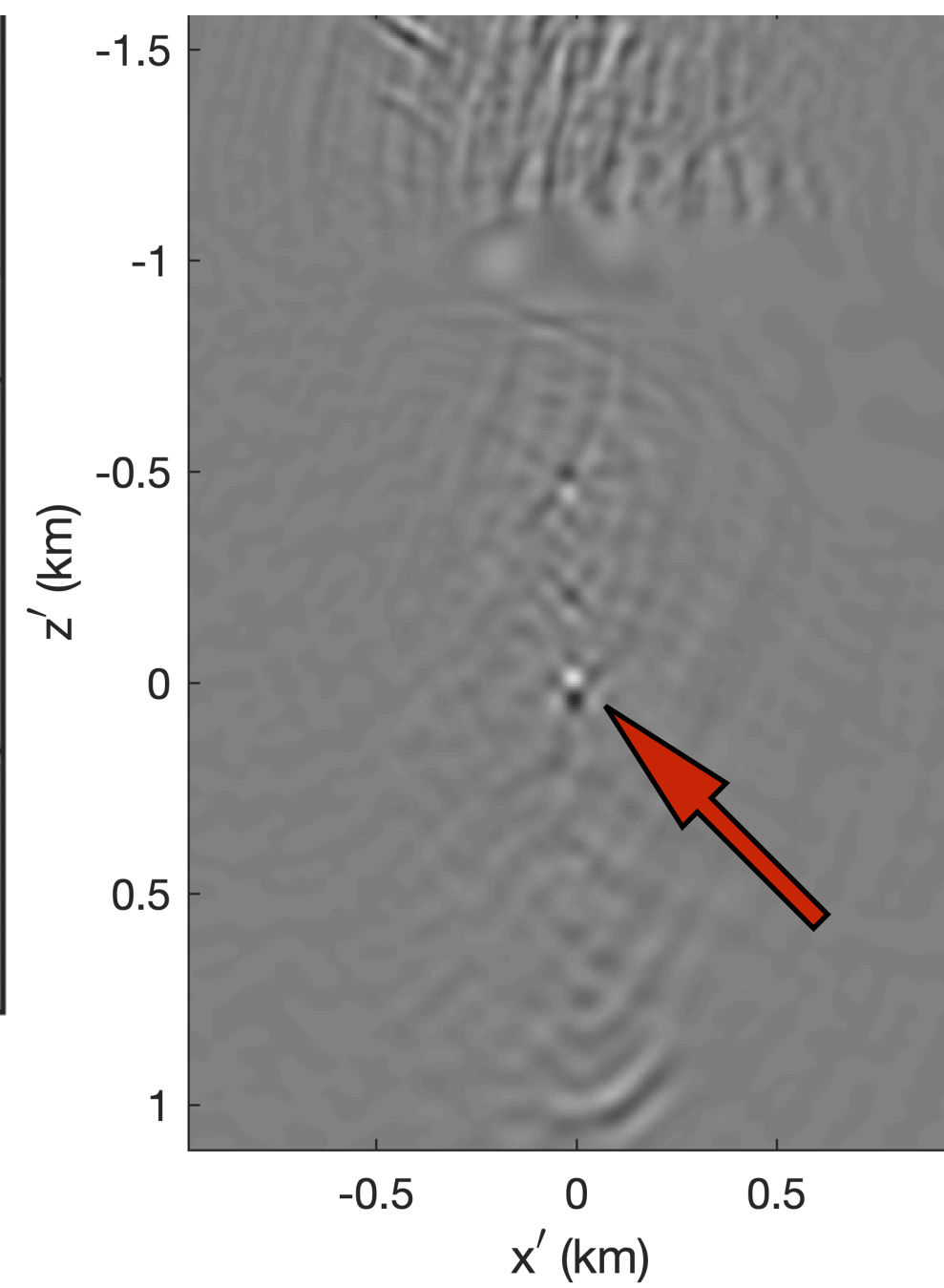
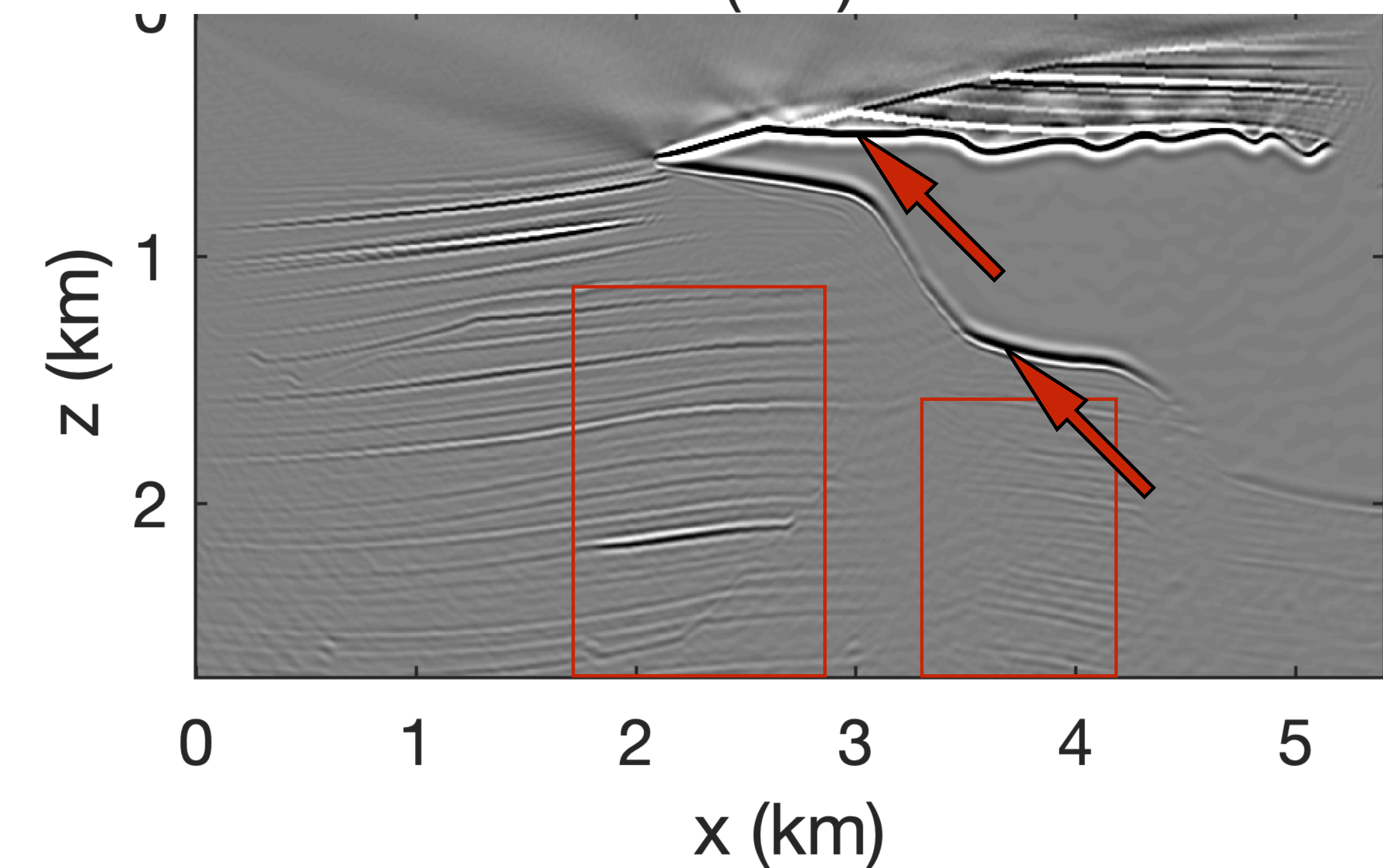


Scenario 1

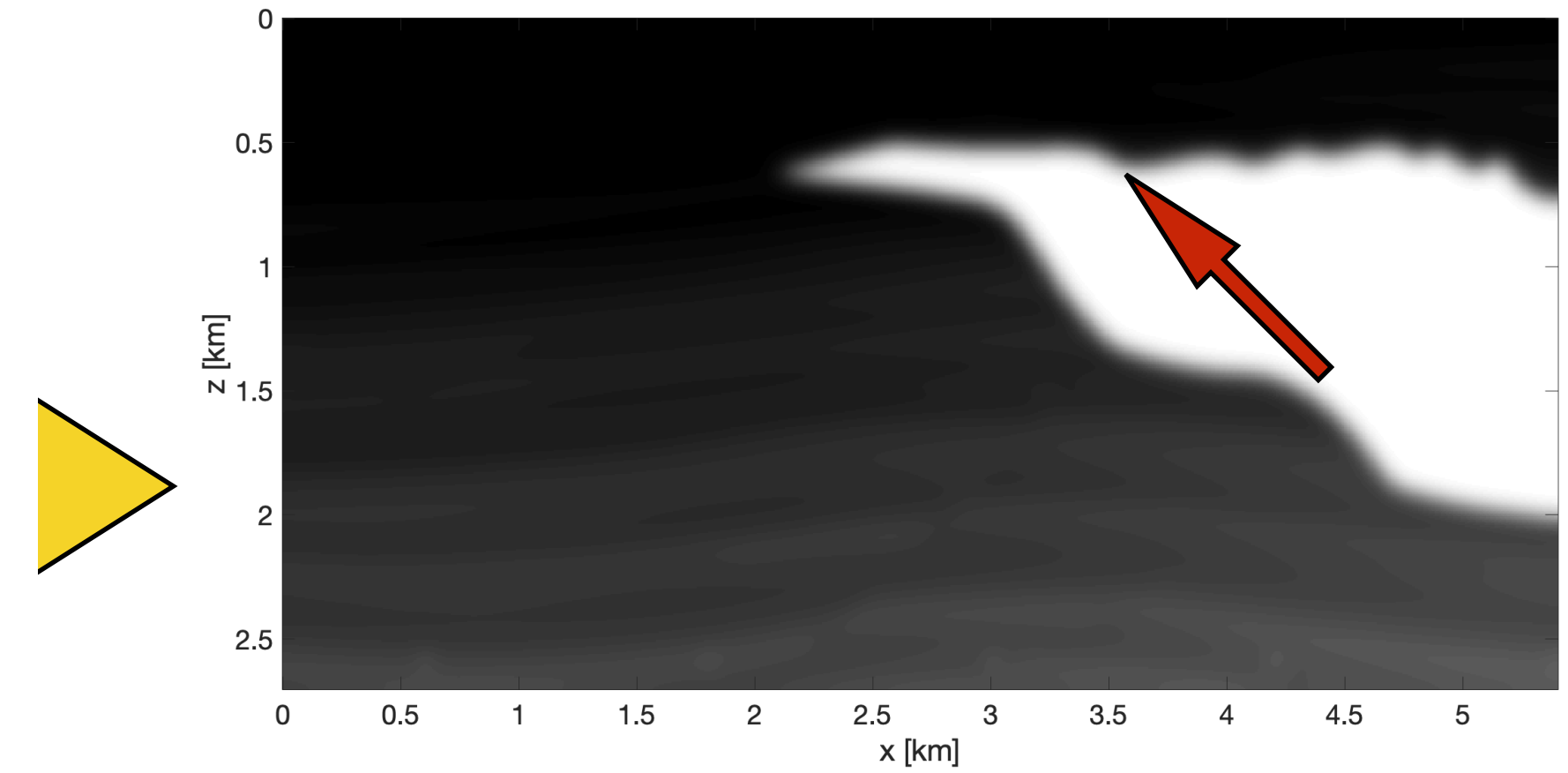
SLIM



$10n_p = 1000$, QR+SVD

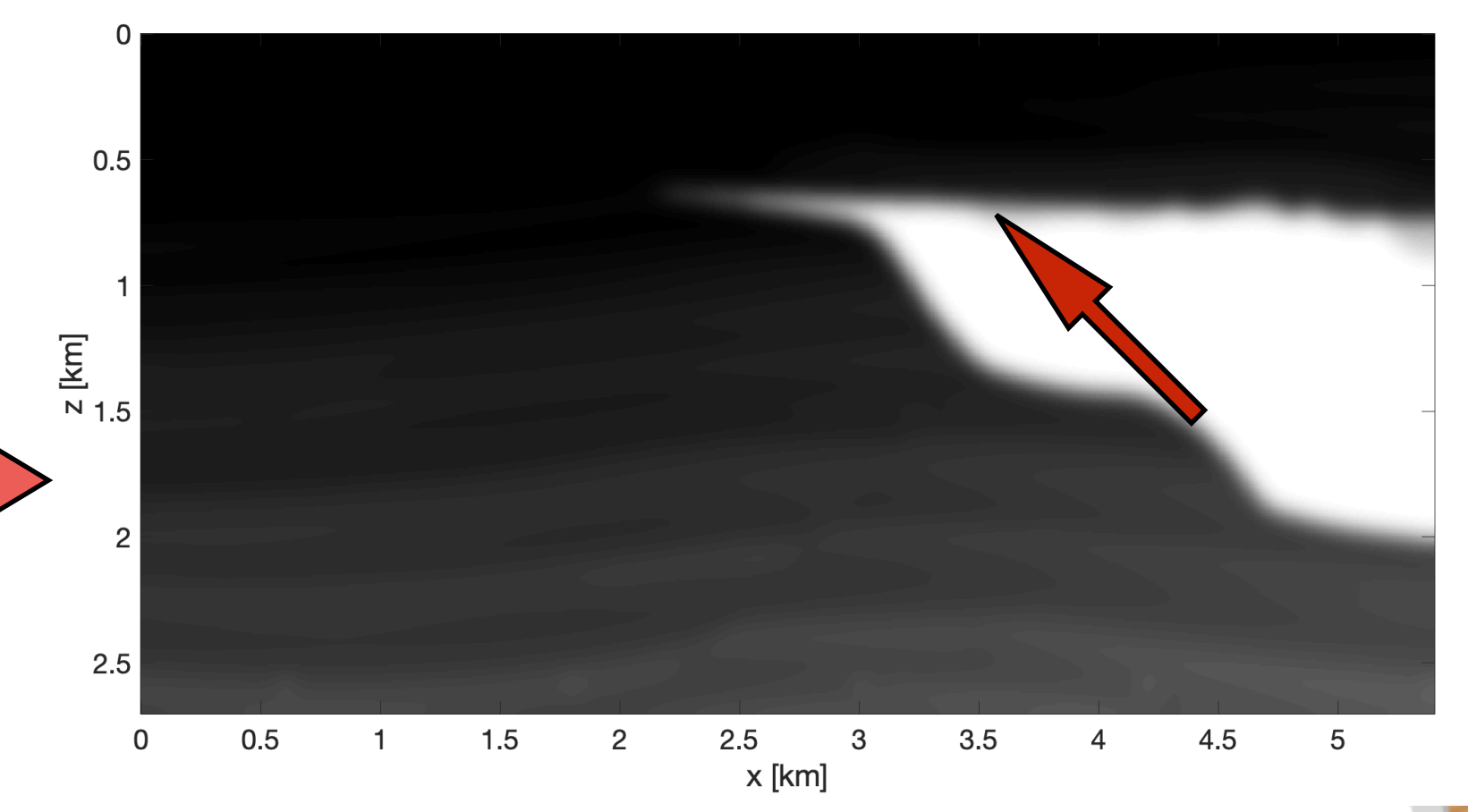


Scenario 3



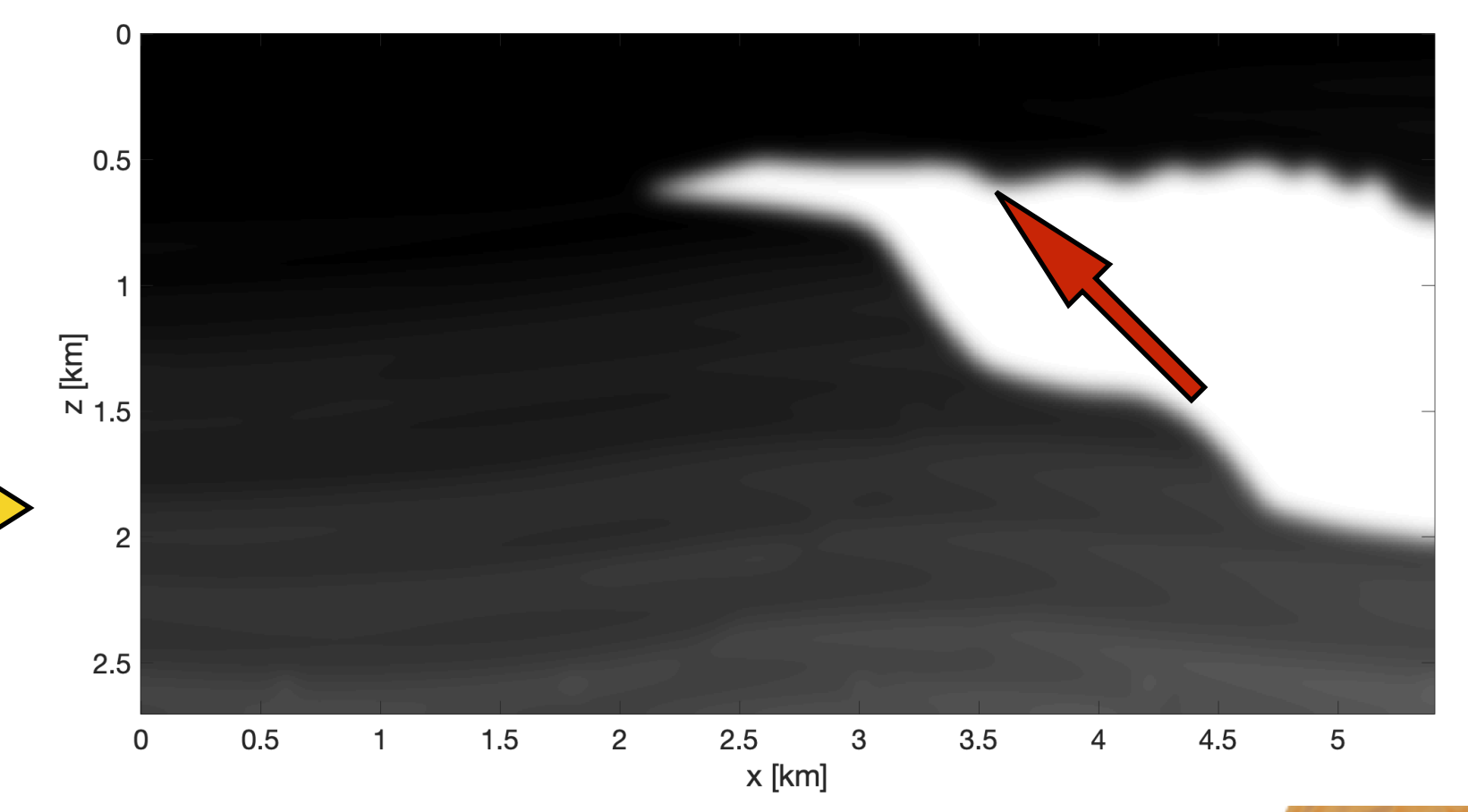
$4n_p = 400$

Scenario 1

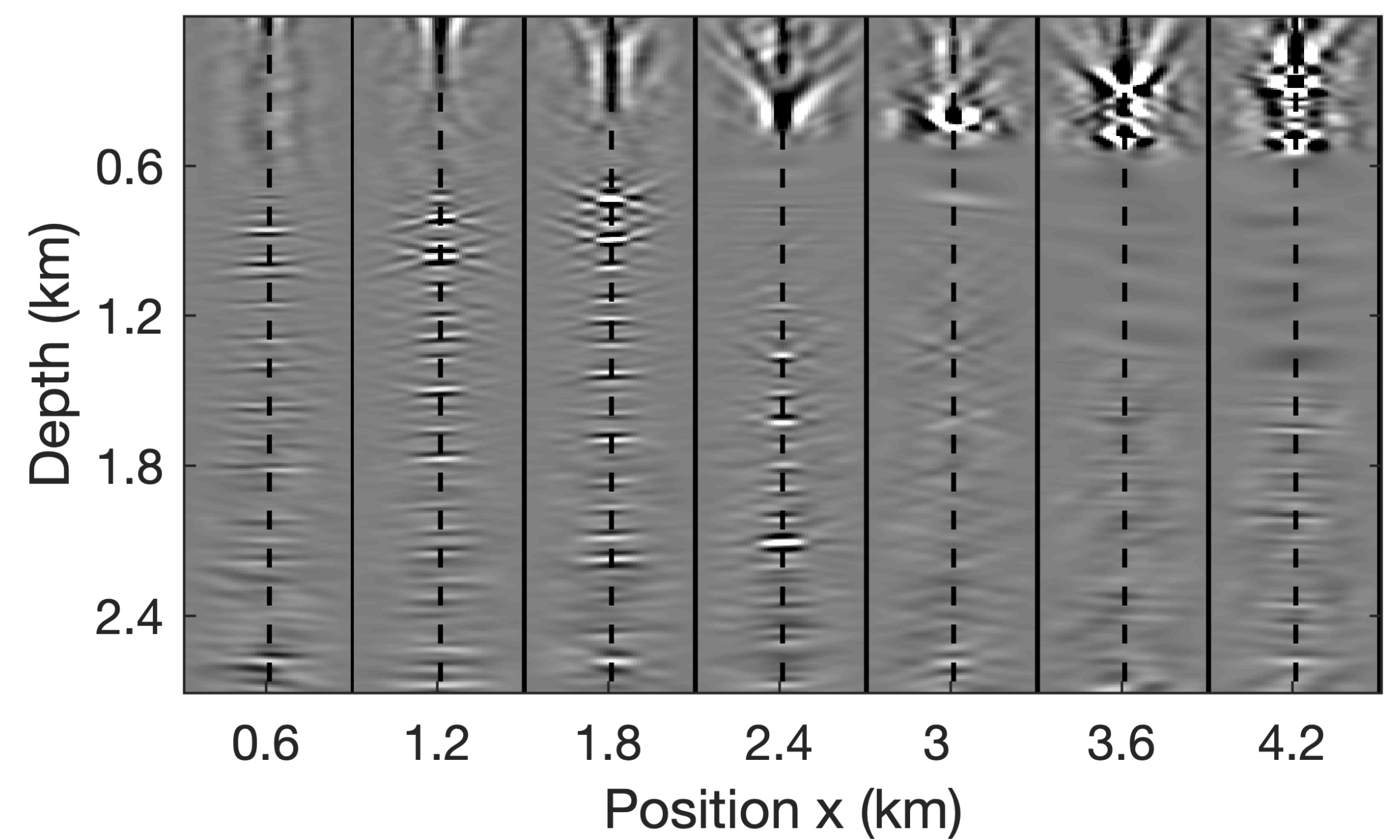
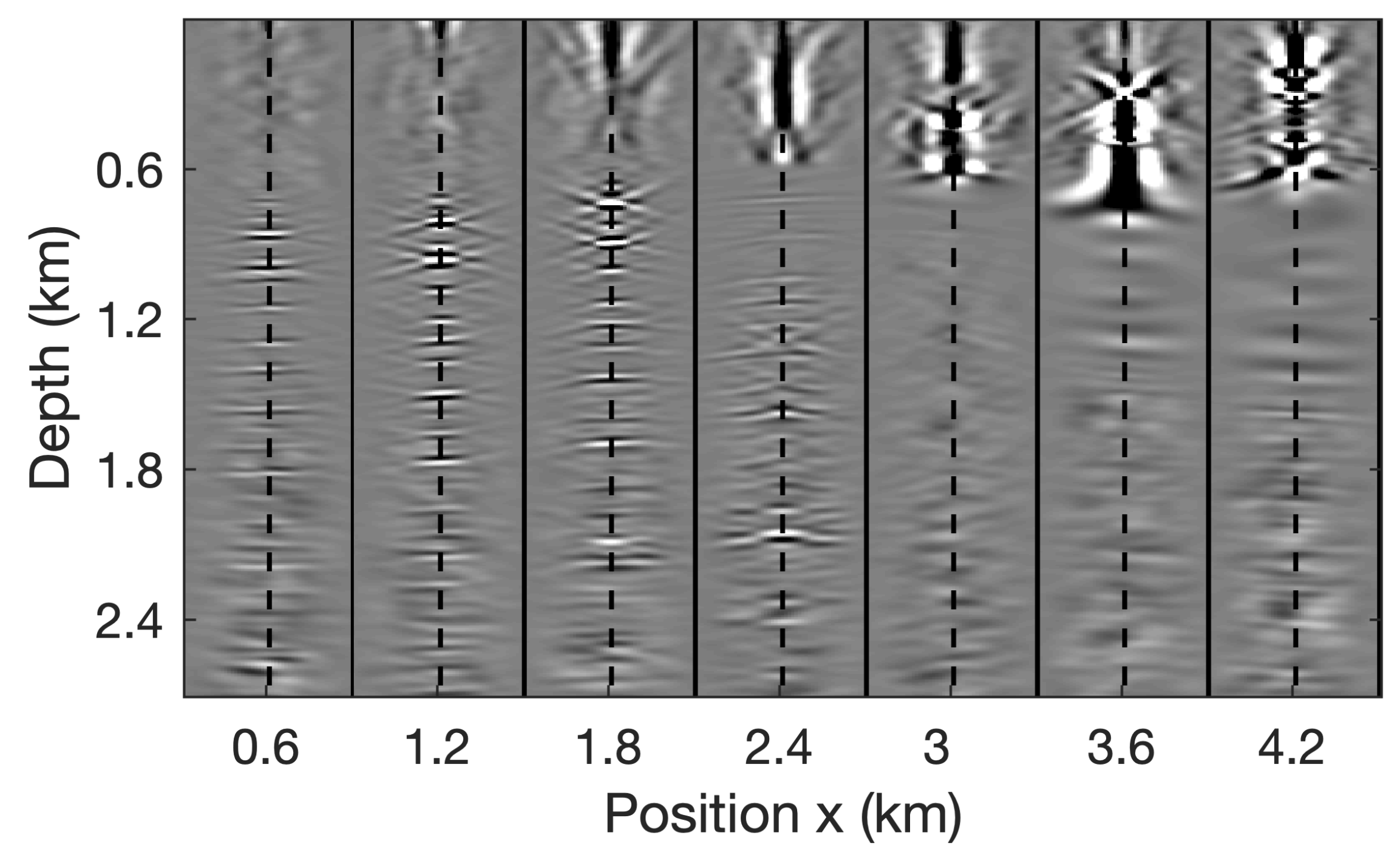


$10n_p = 1000, \text{QR+SVD}$

Scenario 2

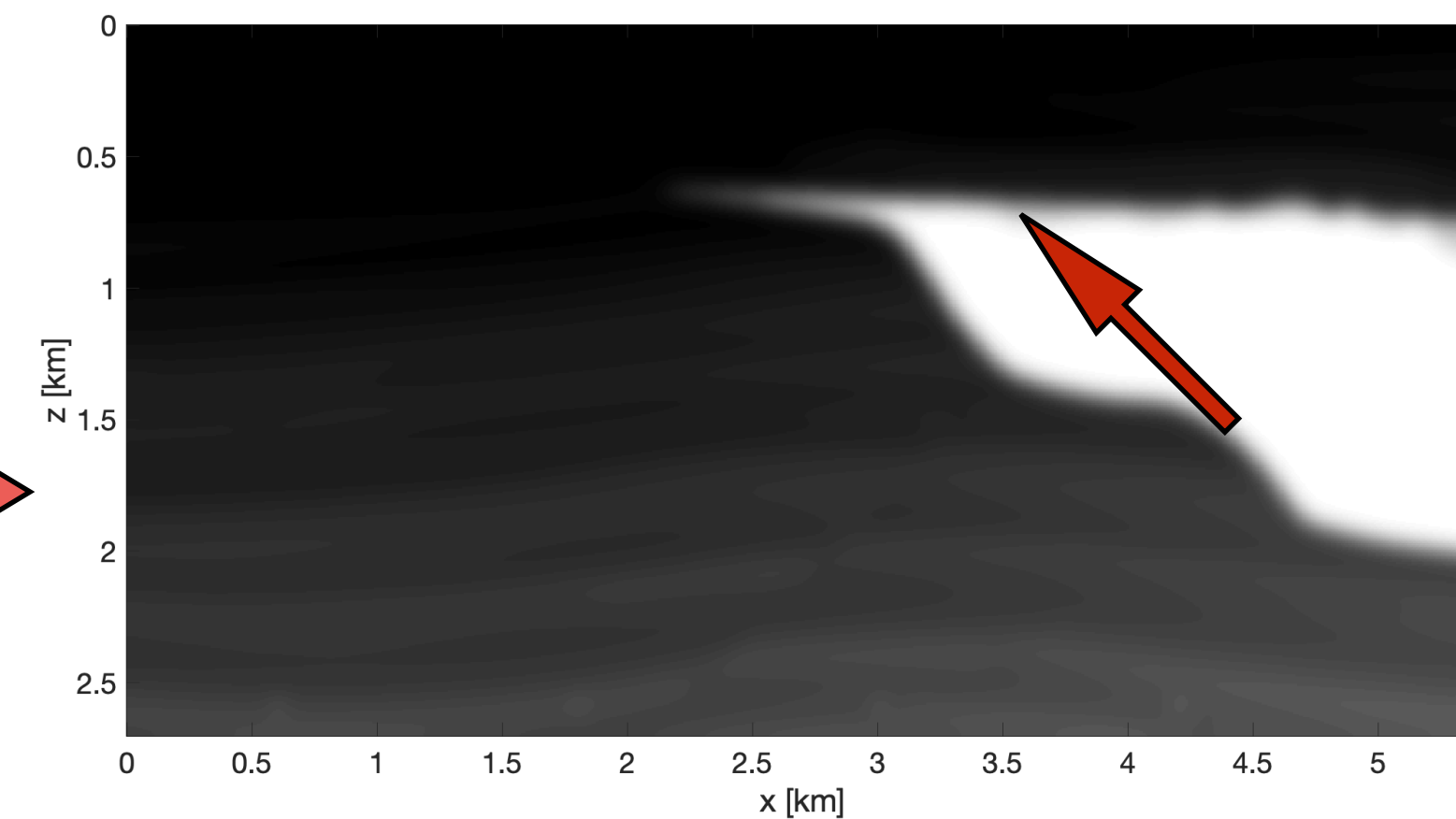


$10n_p = 1000, \text{QR+SVD}$



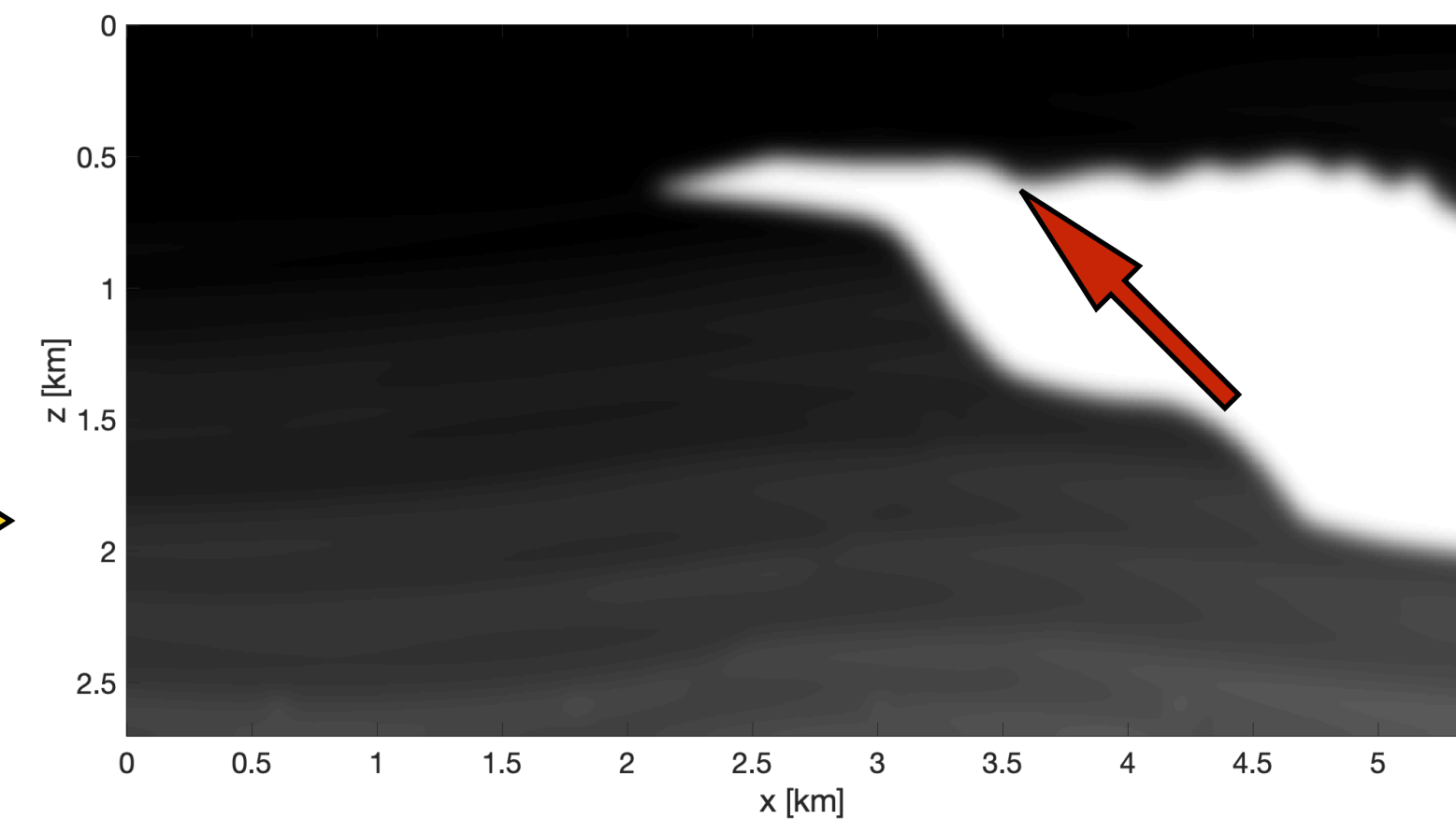
Scenario 1

SLIM 

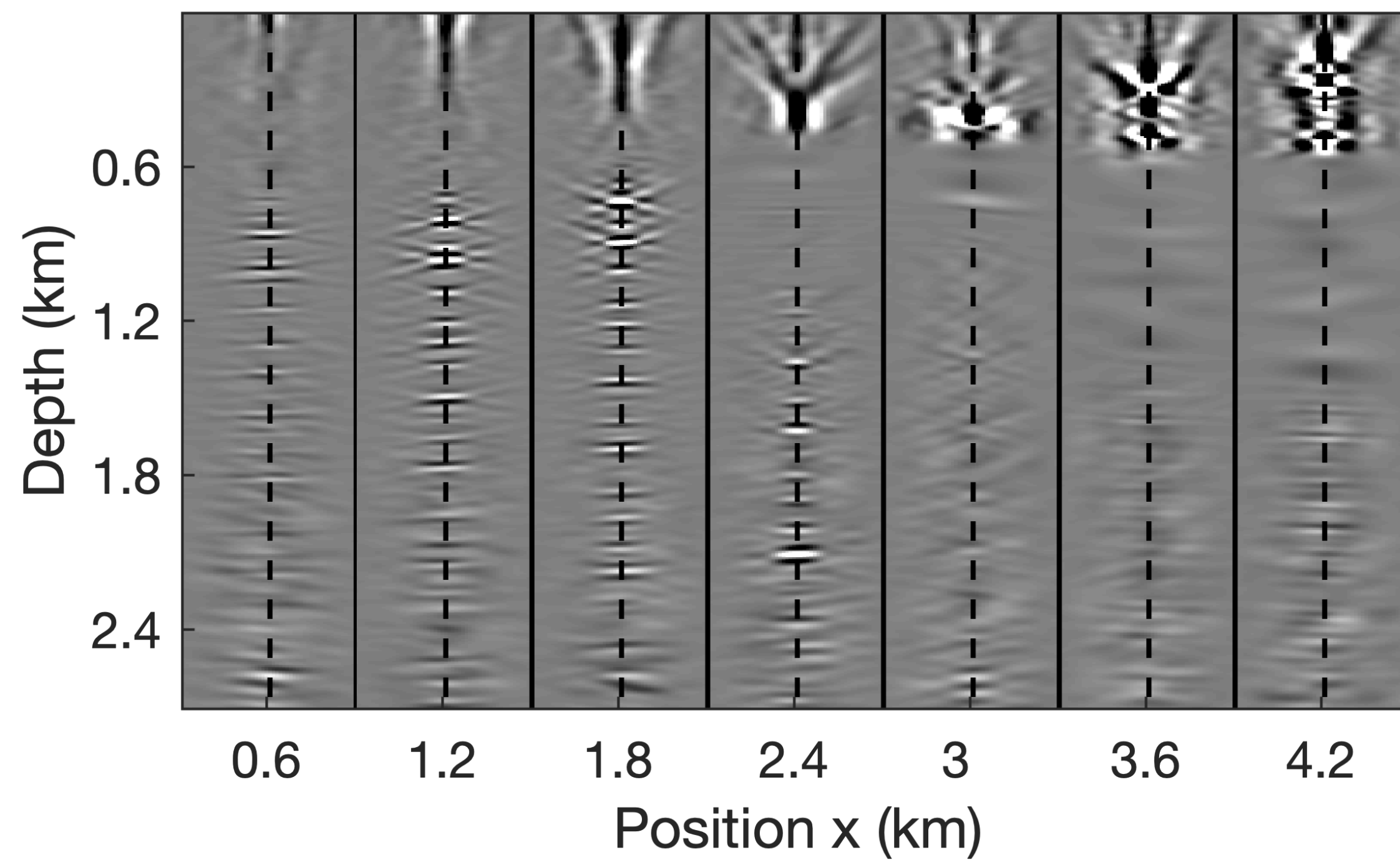
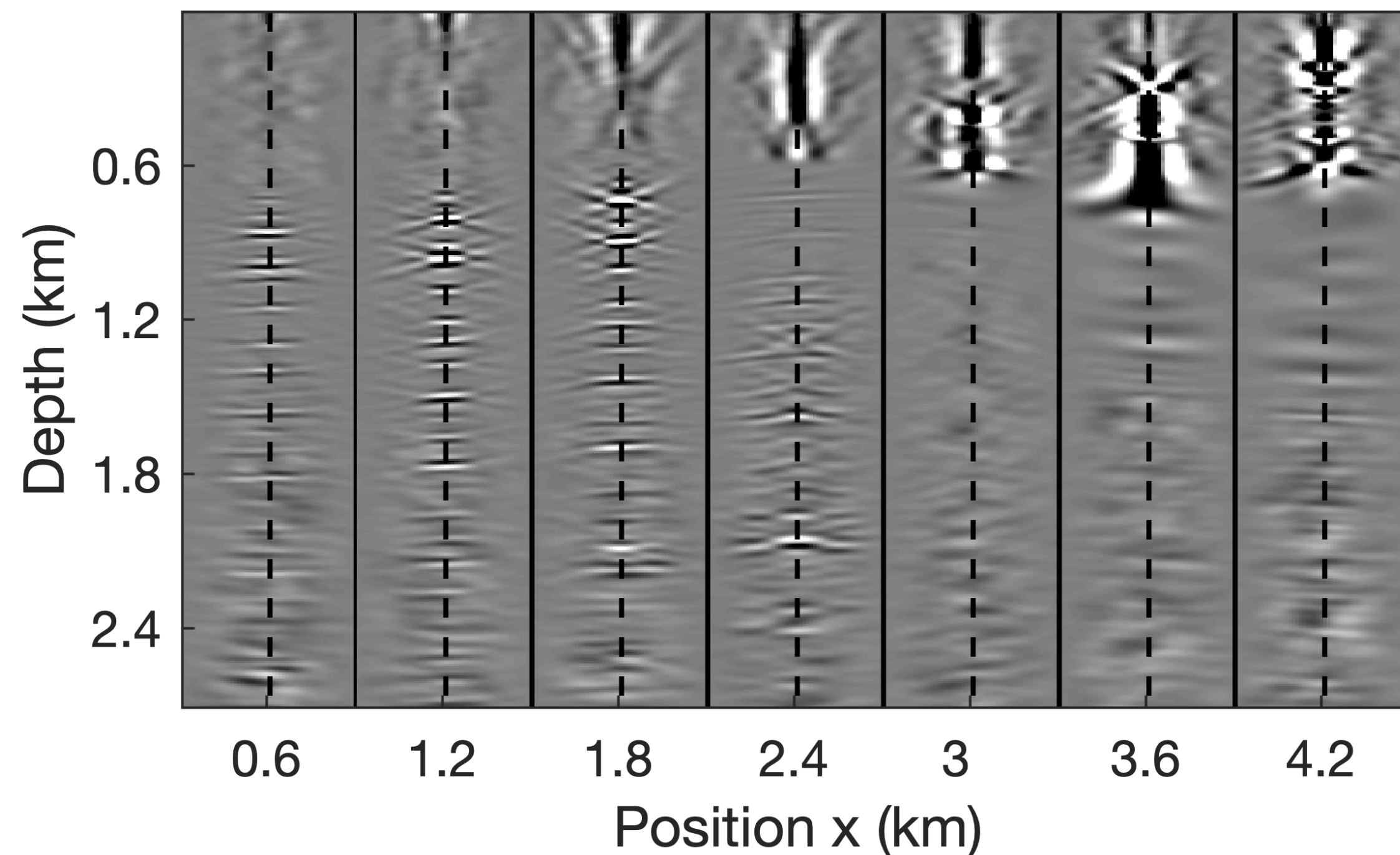


$$10n_p = 1000, \text{ QR+SVD}$$

Scenario 3



$$4n_p = 400$$



Conclusion

By combining EIV probing w/ double two-way wave equation w/ randomized linear algebra

- ▶ EIVs in highly compressed form & manipulations on factors

High-resolution high-accuracy imaging via

- ▶ probing w/ time-domain propagators & Block-Krylov rSVD

The low-rank factors provide

- ▶ access RTM, CIPs, CIGs, geologic dip-corrected CIGs
- ▶ w/o additional wave-equation solves

Velocity continuation

- ▶ via direct mapping factors from one background model to another
- ▶ wave-equation solves scale w/ only probing size

**Chapter 5 Sparsity promoting least-squares
reverse-time migration with multiples**

**Chapter 6 removing density effects in least
square reverse time migration with matched
low-rank filter**

Contribution

- Computationally efficient method for recovering the low-rank representations of the full subsurface extended image volumes, based on time-stepping propagator.
- Examination of power schemes combined with basic randomized linear algebra.
- SVD-free approach to mapping the low-rank factors for velocity variation scenarios.
- On-the-fly source estimation for time-domain sparsity-promoting least-squares reverse-time migration avoiding overfitting.
- Design of a low-rank filter that matches the destiny effect from the strong density variations at the ocean bottom in least-squares reverse time migration w/ only velocity-related Born modelling.

Future work

- Limit large memory usage in 3D by applying on-the-fly Fourier transform with time-stepping, and choose optimal probing size per frequency.
- Design preconditioner to mitigating the ill-conditioning of the subsurface extended image volumes, during the low-rank recovery.
- Design preconditioner for specific image gathers w/ low-rank factors
- Design other optimization method to avoid the SVD factorizations in the estimation of the matched low-rank filter and combine with joint inversion of primaries and multiples.

Publication

- [1] Mengmeng Yang, Zhilong Fang, Philipp Witte and Felix J. Herrmann, “Time-domain sparsity promoting least-squares reverse time migration with source estimation”, submitted to Geophysical Prospecting.
- [2] Mengmeng Yang, Marie Graff, Rajiv Kumar and Felix J. Herrmann, “Low-rank representation of omnidirectional subsurface extended image volumes”, submitted to Geophysics.
- [3] Mengmeng Yang, Marie Graff, Rajiv Kumar and Felix J. Herrmann, “Low-rank representation of subsurface extended image volumes with power iterations”, SEG Technical Program Expanded Abstracts 2019.
- [4] Mengmeng Yang, Rajiv Kumar, Rongrong Wang and Felix J. Herrmann, “Removing density effects in LS-RTM with low-rank matched filter”, SEG Technical Program Expanded Abstracts 2018.
- [5] Mengmeng Yang, Emmanouil Daskalakis and Felix J. Herrmann, “Fast sparsity-promoting least-squares migration with multiples in the time domain”, SEG Technical Program Expanded Abstracts 2017.
- [6] Philipp Witte, Mengmeng Yang and Felix J. Herrmann, “Sparsity-promoting least-squares migration with the linearized inverse scattering imaging condition”, 79th EAGE Conference and Exhibition 2017.
- [7] Mengmeng Yang, Philipp Witte, Zhilong Fang and Felix J. Herrmann, “Time-domain sparsity-promoting least-squares migration with source estimation”, SEG Technical Program Expanded Abstracts 2016.
- [8] Xintao Chai, Mengmeng Yang, Philipp Witte, Rongrong Wang, Zhilong Fang and Felix J. Herrmann, “A linearized Bregman method for compressive waveform inversion”, SEG Technical Program Expanded Abstracts 2016.

Thank you !

PhD Advisory Committee members

PhD Exam Committee members

SLIM Group Members

SINBAD sponsors & NSERC

Georgia Research Alliance

Family and Friends