Extended source imaging - a unifying framework for seismic & medical imaging

Ziyi Yin, Rafael Orozco, Philipp Witte[2], Mathias Louboutin, Gabrio Rizzuti, Ali Siahkoohi, and Felix J. Herrmann

[2] Microsoft Azure

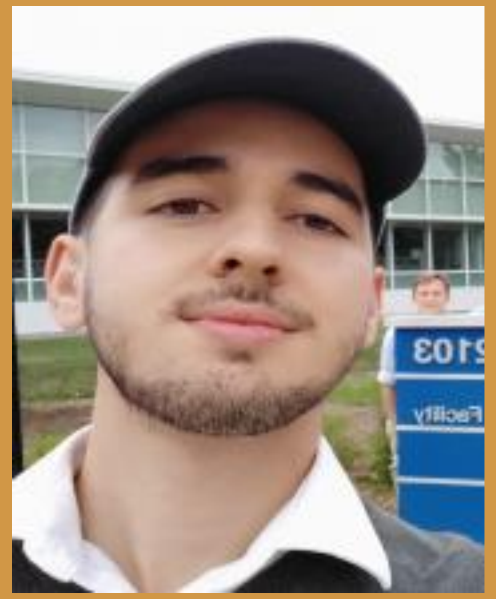
SOCIETY OF EXPLORATION GEOPHYSICISTS

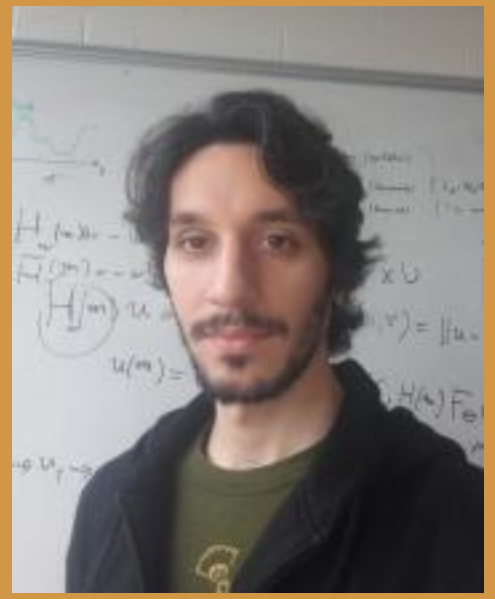
SLIM
Georgia Institute of Technology

Georgia Tech

Extended source imaging - a unifying framework for seismic & medical imaging

Ziyi Yin, Rafael Orozco, Philipp Witte, Mathias Louboutin, Gabrio Rizzuti, Ali Siahkoohi, and Felix J. Herrmann

SLIM
Georgia Institute of Technology

Georgia Tech

Outline

Overview

- contrasts & synergies between seismic & medical imaging
- benefits separation of concerns, abstractions & automatic code generation

Examples *extended volume source imaging*

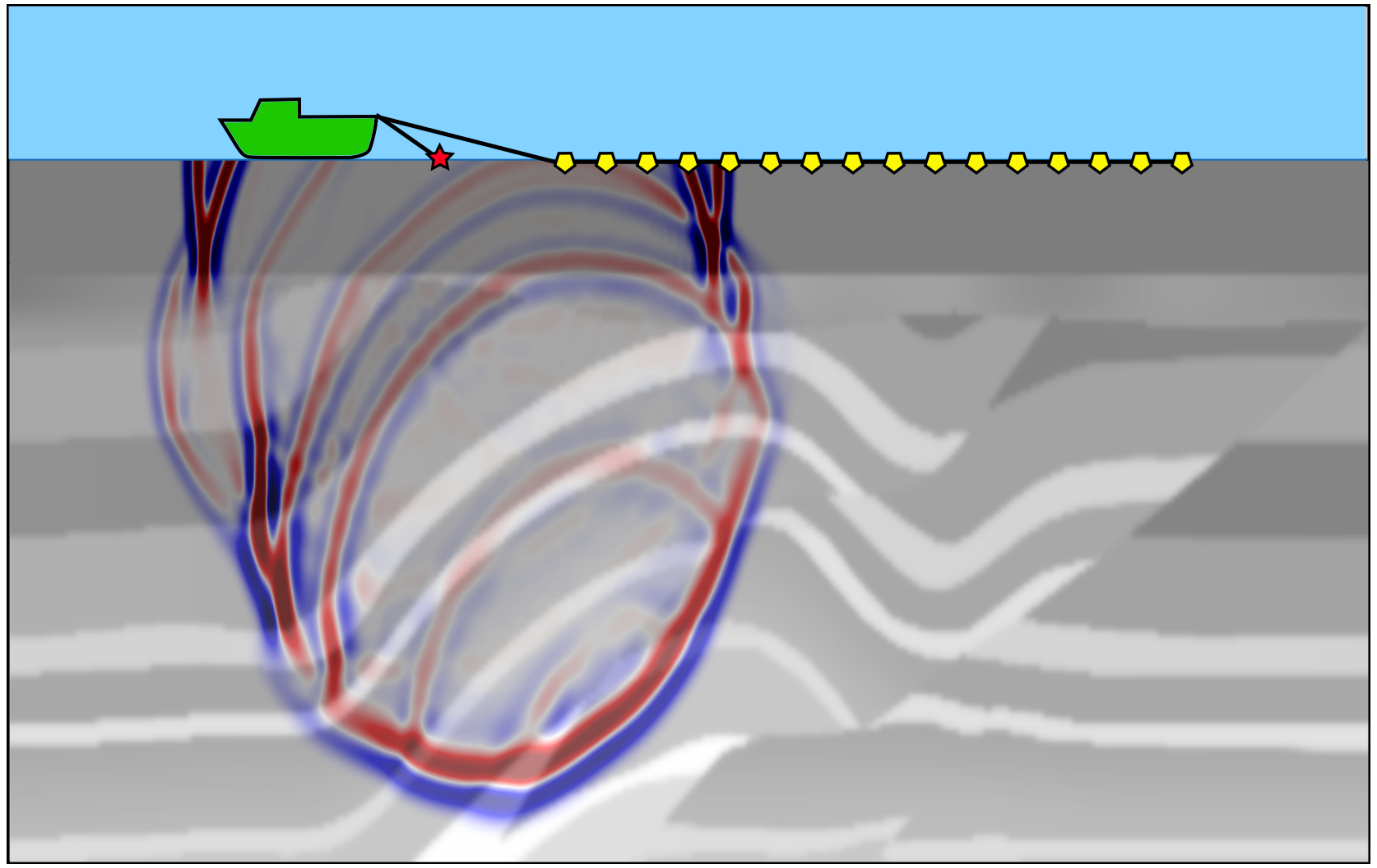
Discuss steps next steps & opportunities

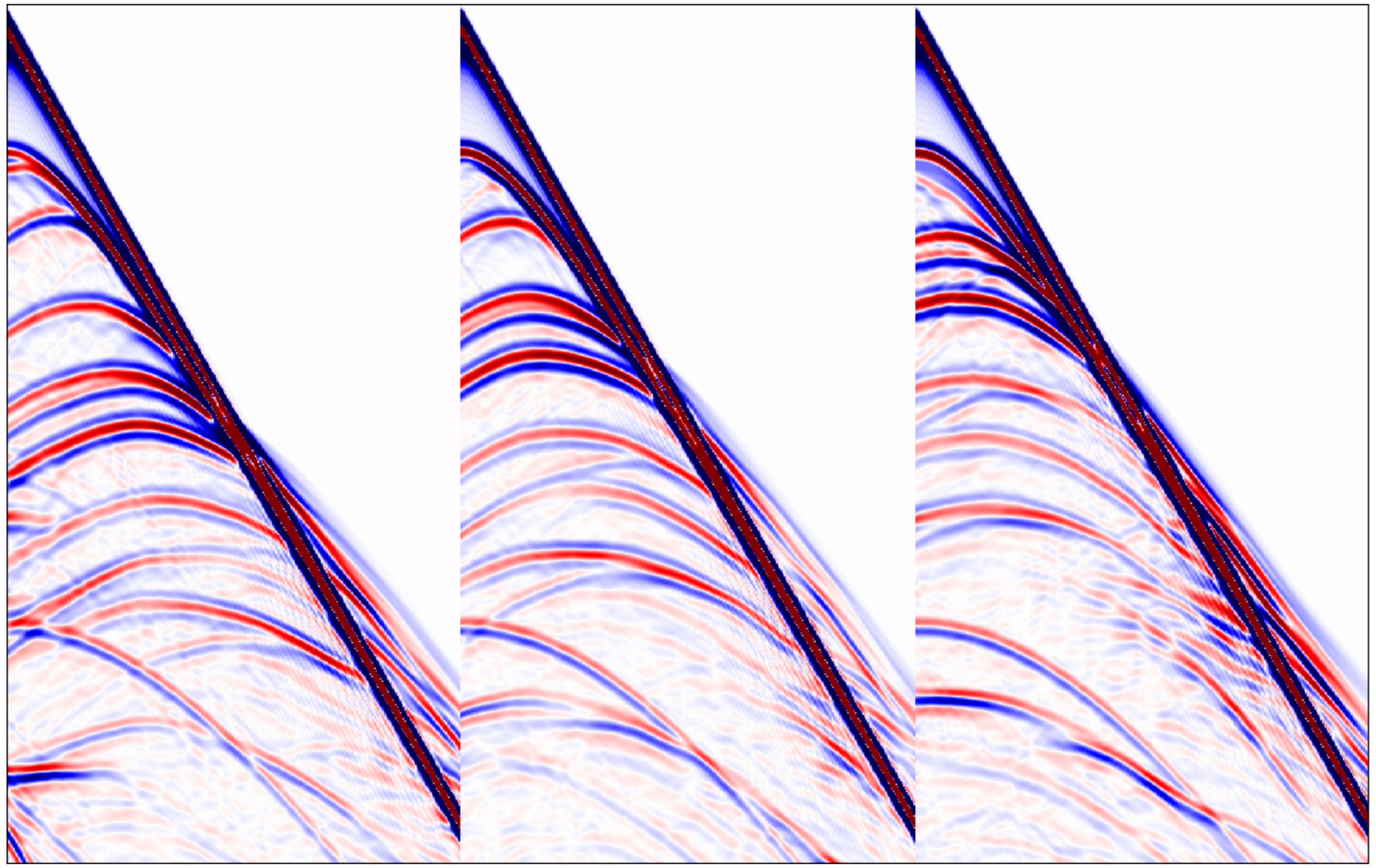
- scale up
- quantitative assessment of risk

Seismic inverse problems

Estimate unknown subsurface properties from seismic (surface) data:

- image geological structures/discontinuities
- estimate physical rock properties (wavespeed, density, etc.)
- turn around times order 1 year



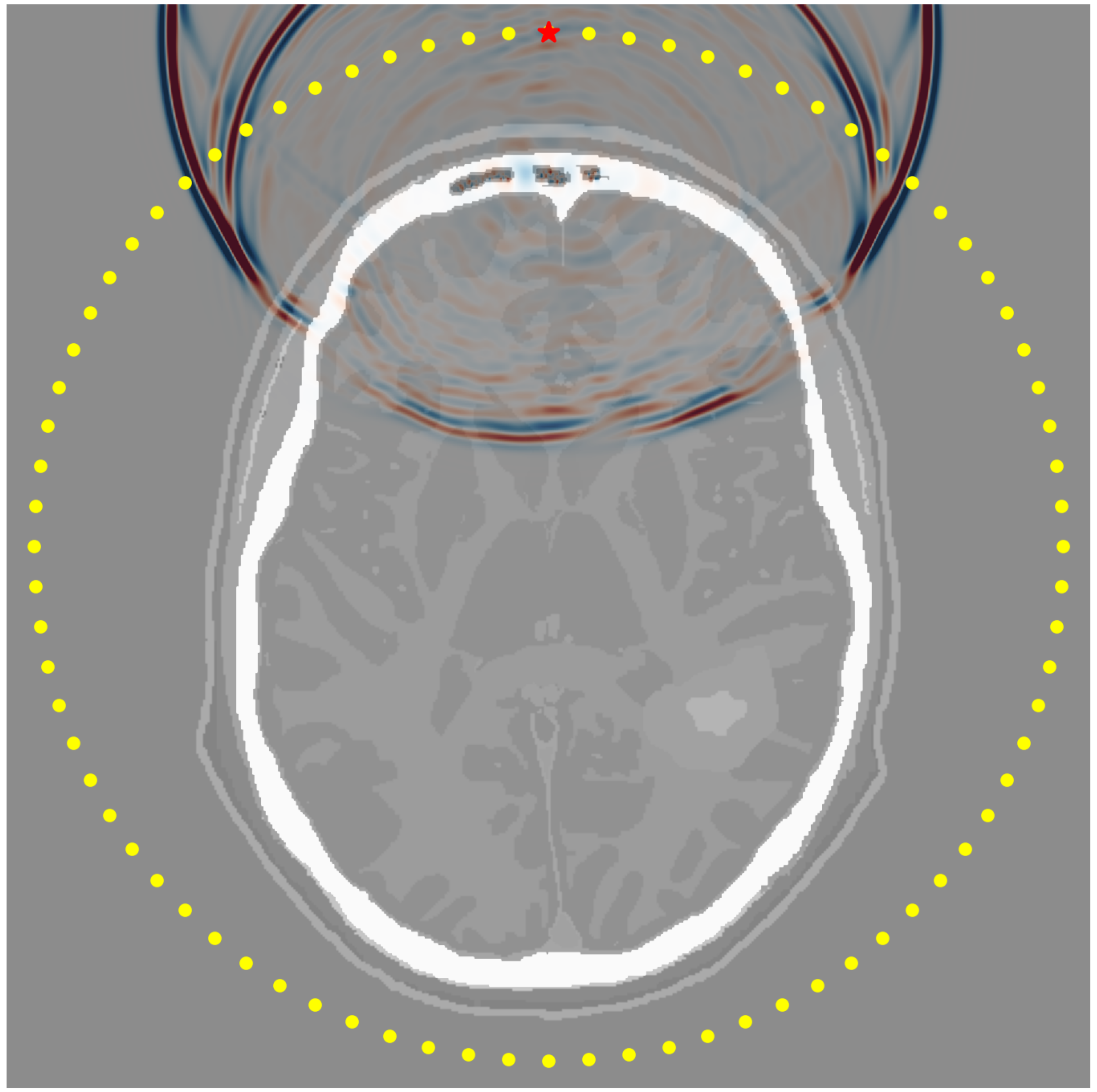


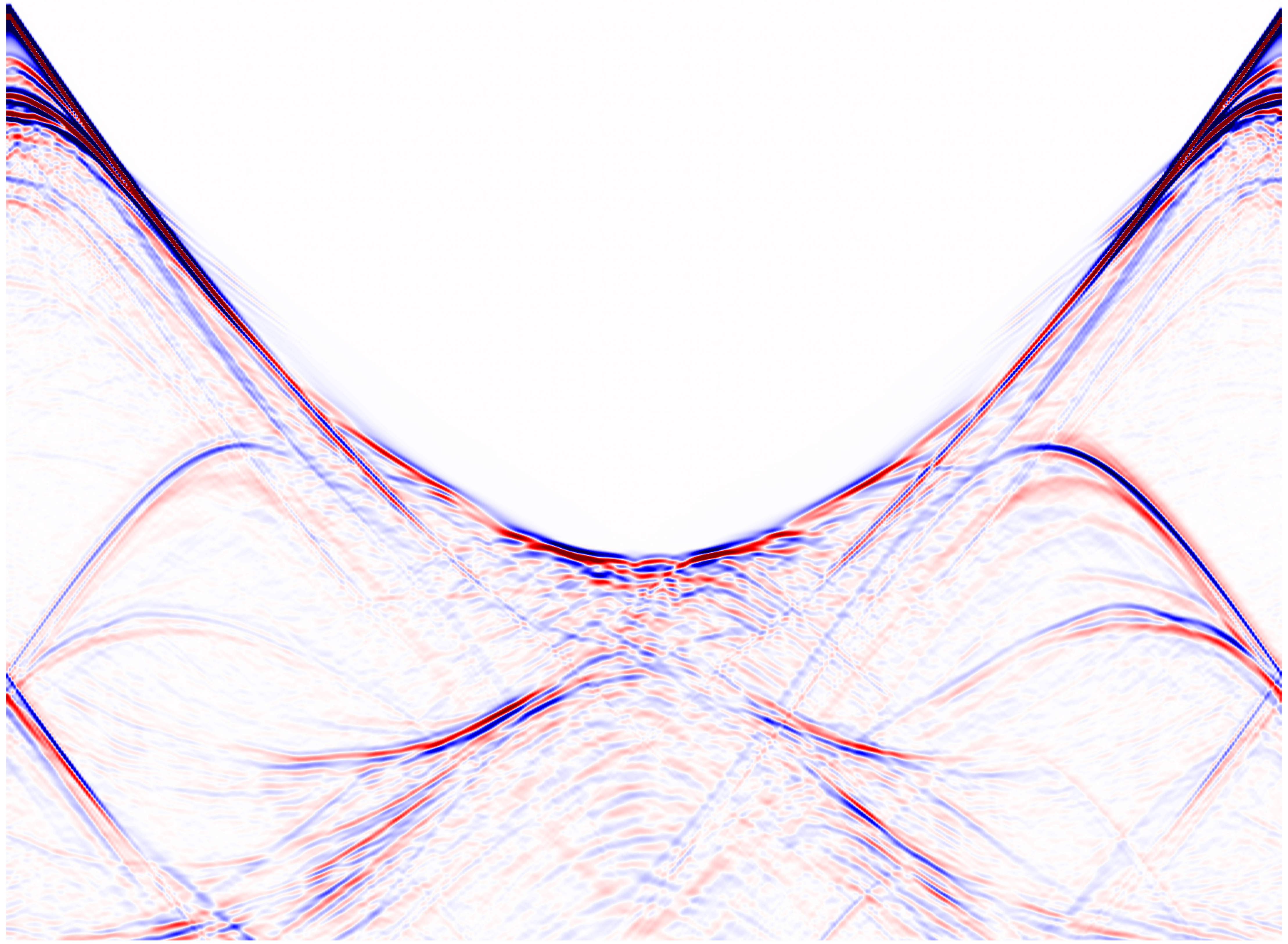
[1] Iacono, Maria Ida, et al. "MIDA: a multimodal imaging-based detailed anatomical model of the human head and neck." *PloS one* 10.4 (2015)
[2] Guasch, Lluís, et al. "Full-waveform inversion imaging of the human brain." *NPJ digital medicine* 3.1 (2020): 1-12.

Medical inverse problems

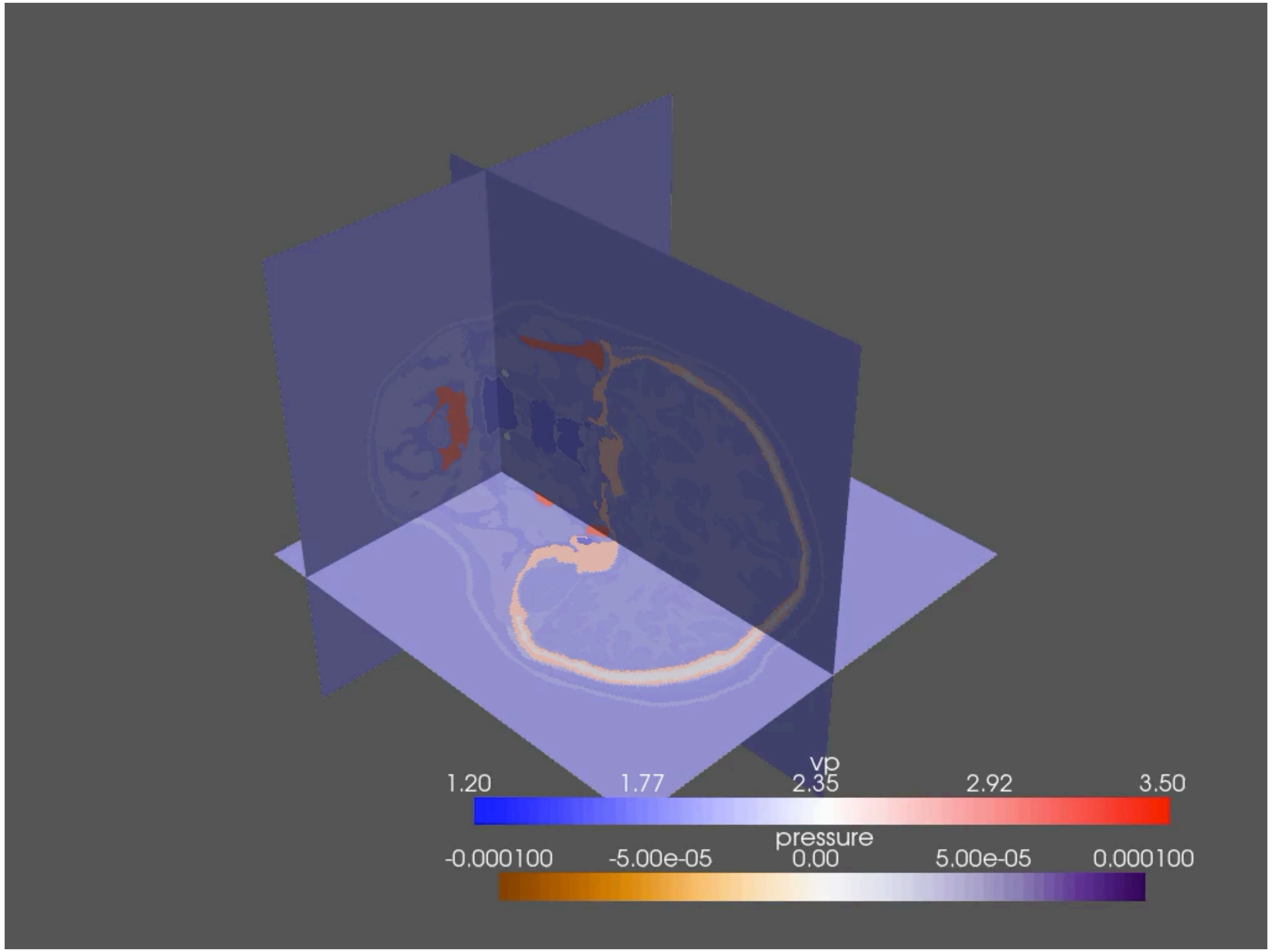
Estimate unknown tissue properties from (omnidirectional) ultrasound data:

- image healthy & malignant tissue
- estimate physical tissue properties (wavespeed, density, etc.)
- turn around times seconds to minutes





Challenge

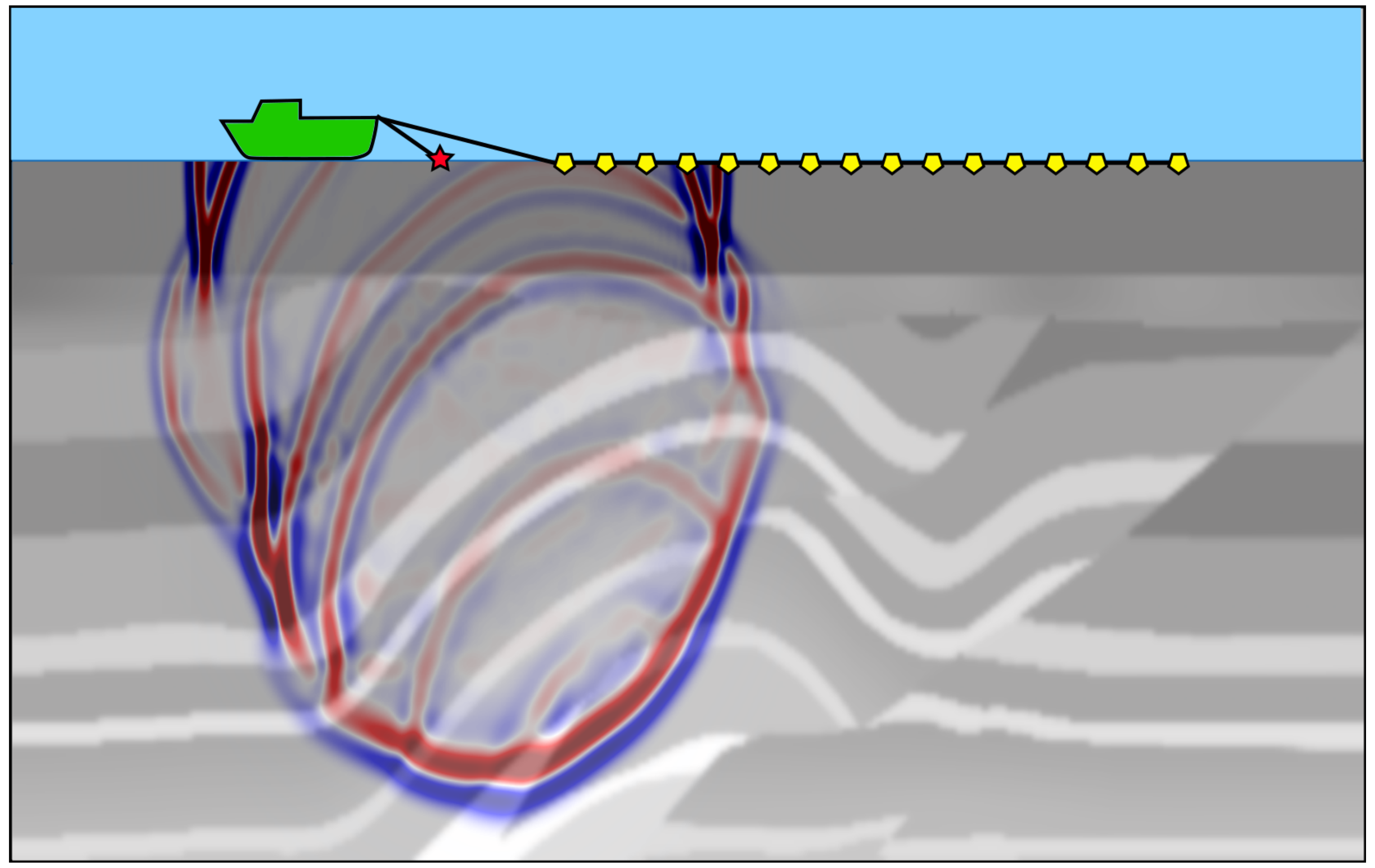

vp
1.20
1.77
2.35
2.92
3.50
pressure
-0.000100
-5.00e-05
0.00
5.00e-05
0.000100

PDE-constrained optimization

Problem formulation:

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}\sum_{i=1}^{n_s} \|\mathbf{P}_r\mathbf{u}_i - \mathbf{d}_i\|_2^2$$

$$\text{subject to} \quad \mathbf{A}[\mathbf{m}]\mathbf{u}_i = \mathbf{P}_s^\top \mathbf{q}_i$$



$\mathbf{u}_i$	state variables (wavefields)
$\mathbf{A}[\mathbf{m}]$	discretized PDE parametrized by $\mathbf{m}$
$\mathbf{q}_i$	spatio-temporal source distribution
$\mathbf{P}_r$	receiver sampling operator
$\mathbf{P}_s^\top$	source injection operator
$\mathbf{d}_i$	observed data

Unconstrained reduced formulation

Eliminate PDE constraints:

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}\sum_{i=1}^{n_s} \|\mathbf{P}_r \mathbf{u}_i - \mathbf{d}_i\|_2^2 \quad \text{subject to} \quad \mathbf{A}[\mathbf{m}]\mathbf{u}_i = \mathbf{P}_s^\top \mathbf{q}_i \iff \min_{\mathbf{m}} \frac{1}{2}\sum_{i=1}^{n_s} \|\underbrace{\mathbf{P}_r \mathbf{A}[\mathbf{m}]^{-1}\mathbf{P}_s^\top}_{\mathbf{F}[\mathbf{m}]} \mathbf{q}_i - \mathbf{d}_i\|_2^2$$

Application areas:

- Geophysics (seismic imaging, global seismology, monitoring)
- Medical imaging (ultrasound/photoacoustic imaging)
- Mathematical finance (volatility from option prices)
- Aerodynamics (aerodynamic shape optimization)
- Neural networks (CNNs, neural ODEs/PDEs)

Forward problem

The *forward* problem:

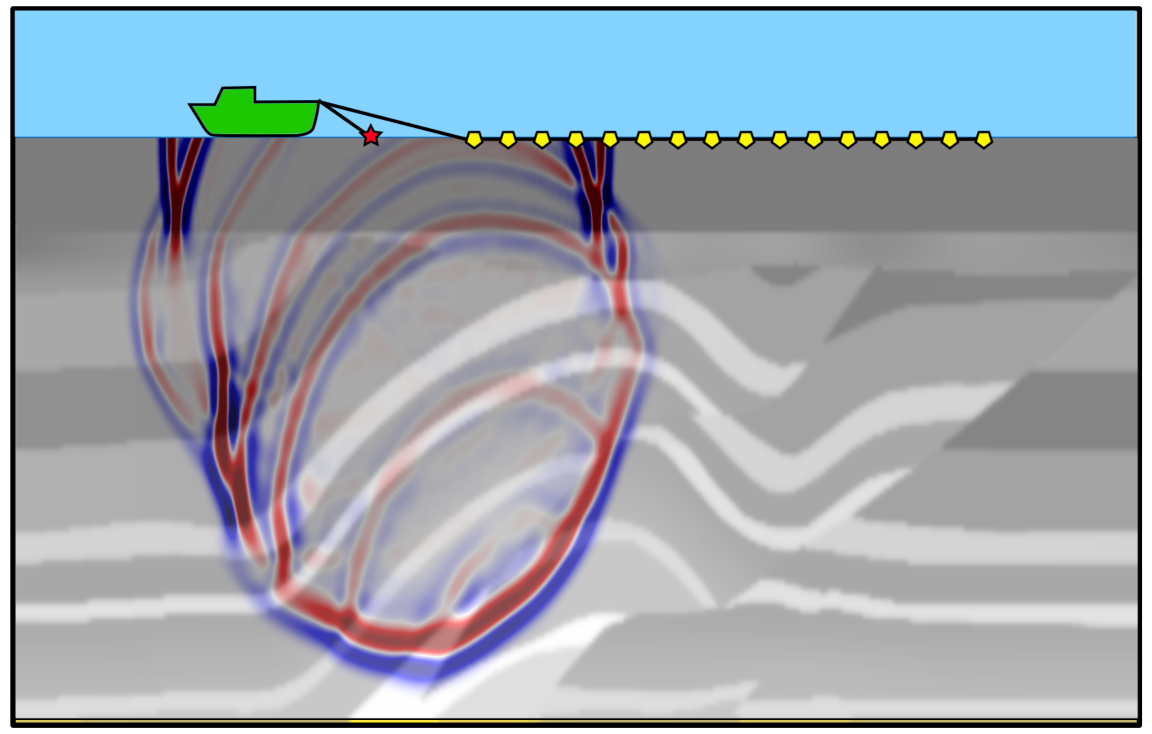
$$\mathbf{F}[\mathbf{m}] \cdot \mathbf{q} = \mathbf{P}_r \underbrace{\mathbf{A}[\mathbf{m}]^{-1}\mathbf{P}_s^\top \mathbf{q}}_{\mathbf{u}}$$

Discretized acoustic wave equation:

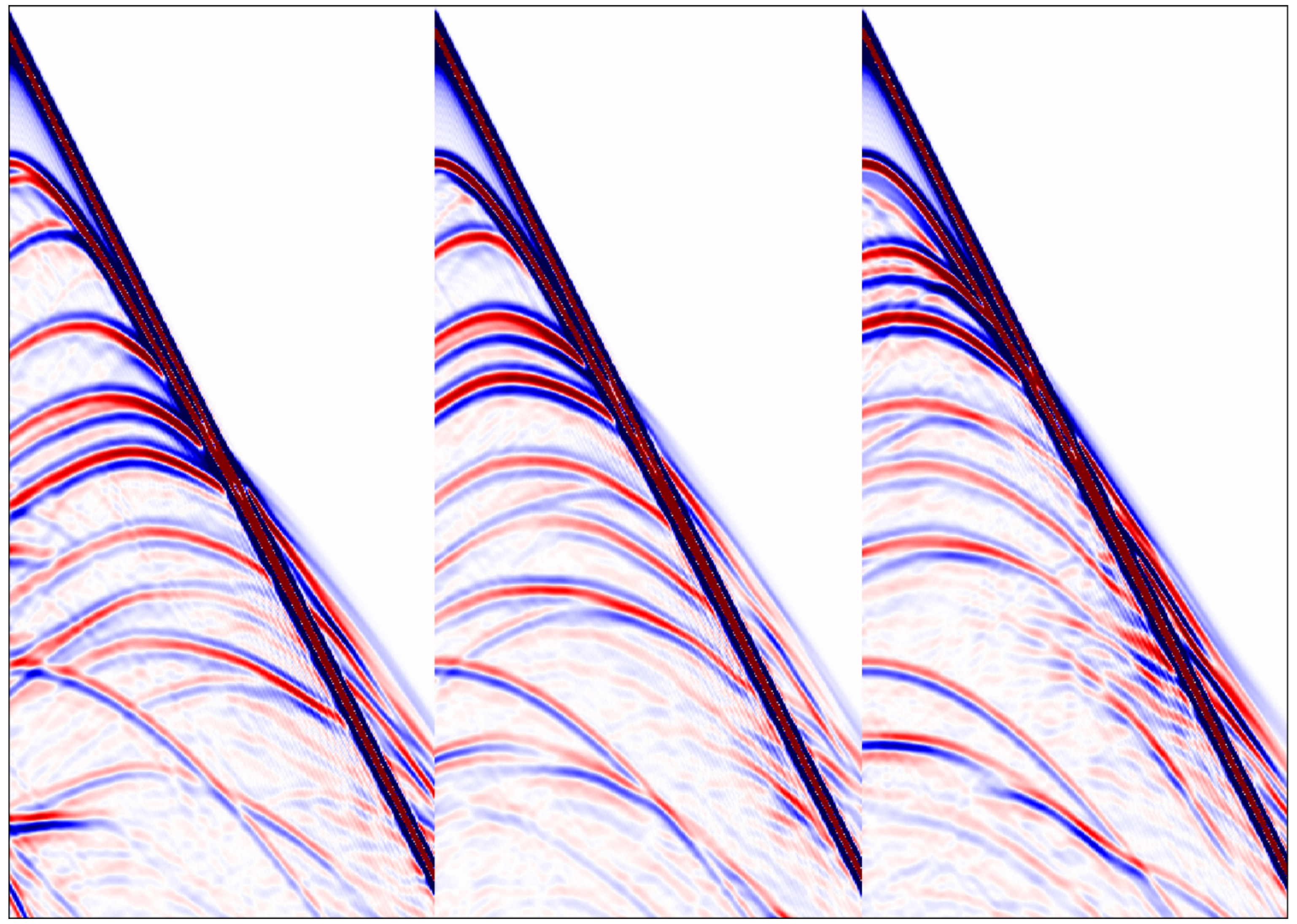
$$\mathbf{A}[\mathbf{m}] = \mathbf{m} \odot \frac{\partial^2}{\partial t^2} - \nabla^2$$

Solve via finite-difference time-stepping:

$$\mathbf{u}^{n+1} = [2 + \frac{\Delta t^2}{\mathbf{m}} \odot \mathbf{L}]\mathbf{u}^n - \mathbf{u}^{n-1} + \frac{\Delta t^2}{\mathbf{m}} \odot \mathbf{P}_s^\top \mathbf{q}^{n+1}$$



$\mathbf{F}[\mathbf{m}] \cdot \mathbf{q}$



Inverse problem

The *inverse* problem:

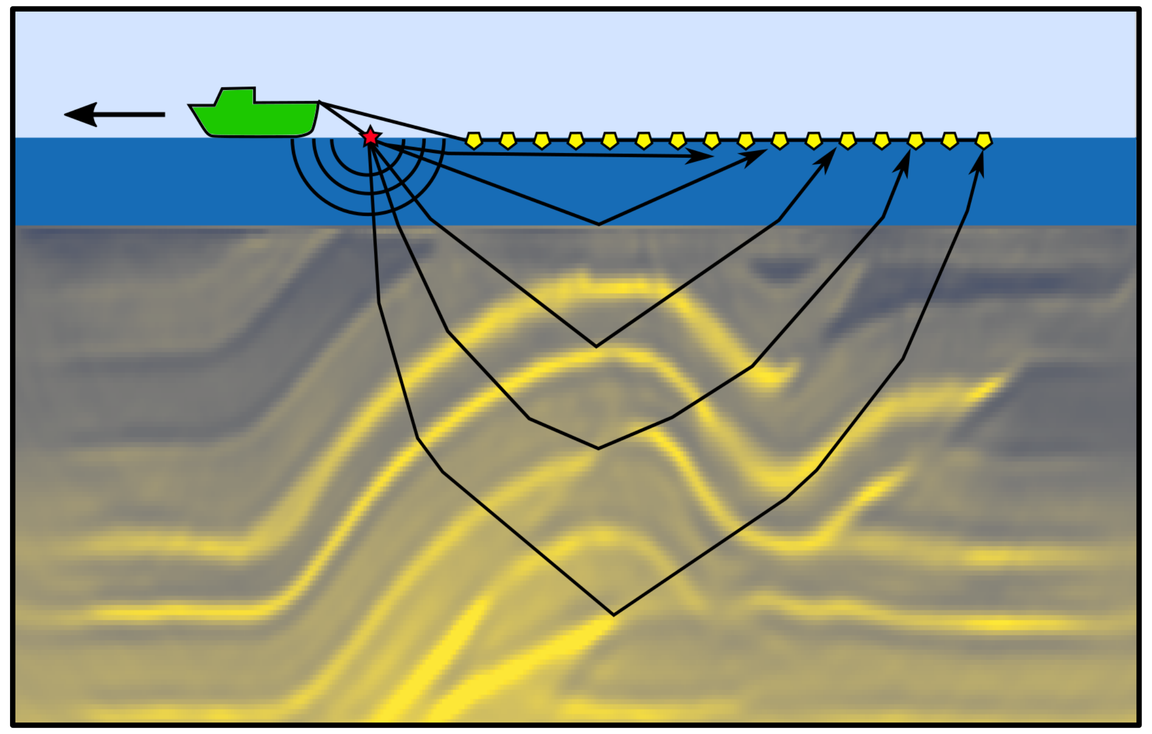
$$\min_{\mathbf{m}} \Phi(\mathbf{m}) = \frac{1}{2}\sum_{i=1}^{n_s} \|\mathbf{d}_i - \mathbf{F}[\mathbf{m}]\mathbf{q}_i\|_2^2$$

Sensitivities w.r.t. model parameters: $\dfrac{\partial \mathbf{F}[\mathbf{m}]\mathbf{q}}{\partial \mathbf{m}}$

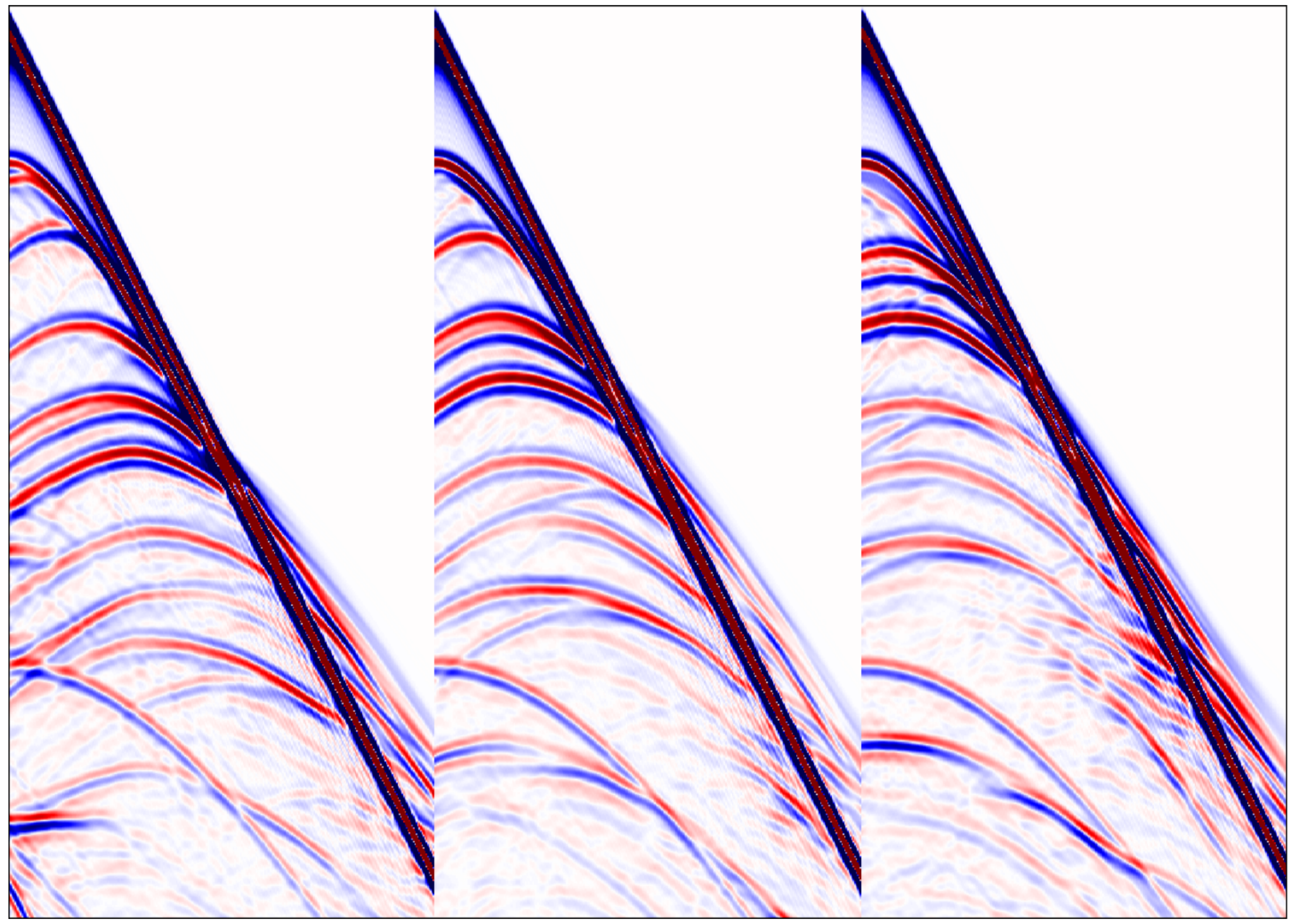
$$\mathbf{J} = -\mathbf{P}_r \mathbf{A}[\mathbf{m}]^{-1} \mathrm{diag}\left(\frac{\partial \mathbf{A}[\mathbf{m}]}{\partial \mathbf{m}} \mathbf{A}[\mathbf{m}]^{-1} \mathbf{P}_s^\top \mathbf{q}\right)$$

Gradient of objective function via backpropagation:

$$\mathbf{g} = \sum_{i=1}^{n_s} \mathbf{J}^\top (\mathbf{F}[\mathbf{m}]\mathbf{q}_i - \mathbf{d}_i)$$



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



SLIM

Mengmeng Yang, Zhilong Fang, Philipp A. Witte, and Felix J. Herrmann, **"Time-domain sparsity promoting least-squares reverse time migration with source estimation"**

Active source imaging

Nonlinear least squares optimization problem:

$$\min_{\mathbf{m},\mathbf{s}} f(\mathbf{m},\mathbf{s}) = \frac{1}{2\sigma^2}\sum_{i=1}^{n_s} \|\mathbf{d}_i - \mathbf{F}[\mathbf{m}]\mathbf{q}_i^{\delta}[\mathbf{s}]\|_2^2 + \lambda R_t(\mathbf{s})$$

$\mathbf{q}_i^{\delta}[\mathbf{s}]$ sources w/ known positions & unknown source-time function $\mathbf{s}$

$\mathbf{F}[\mathbf{m}]$ receiver restricted Green's function

$\lambda R_t(\mathbf{s})$ source-time function regularization via weighted penalty

Sources considered discretizations of $\delta(\mathbf{x} - \mathbf{x}_i^s)s(t)$

Variable projection

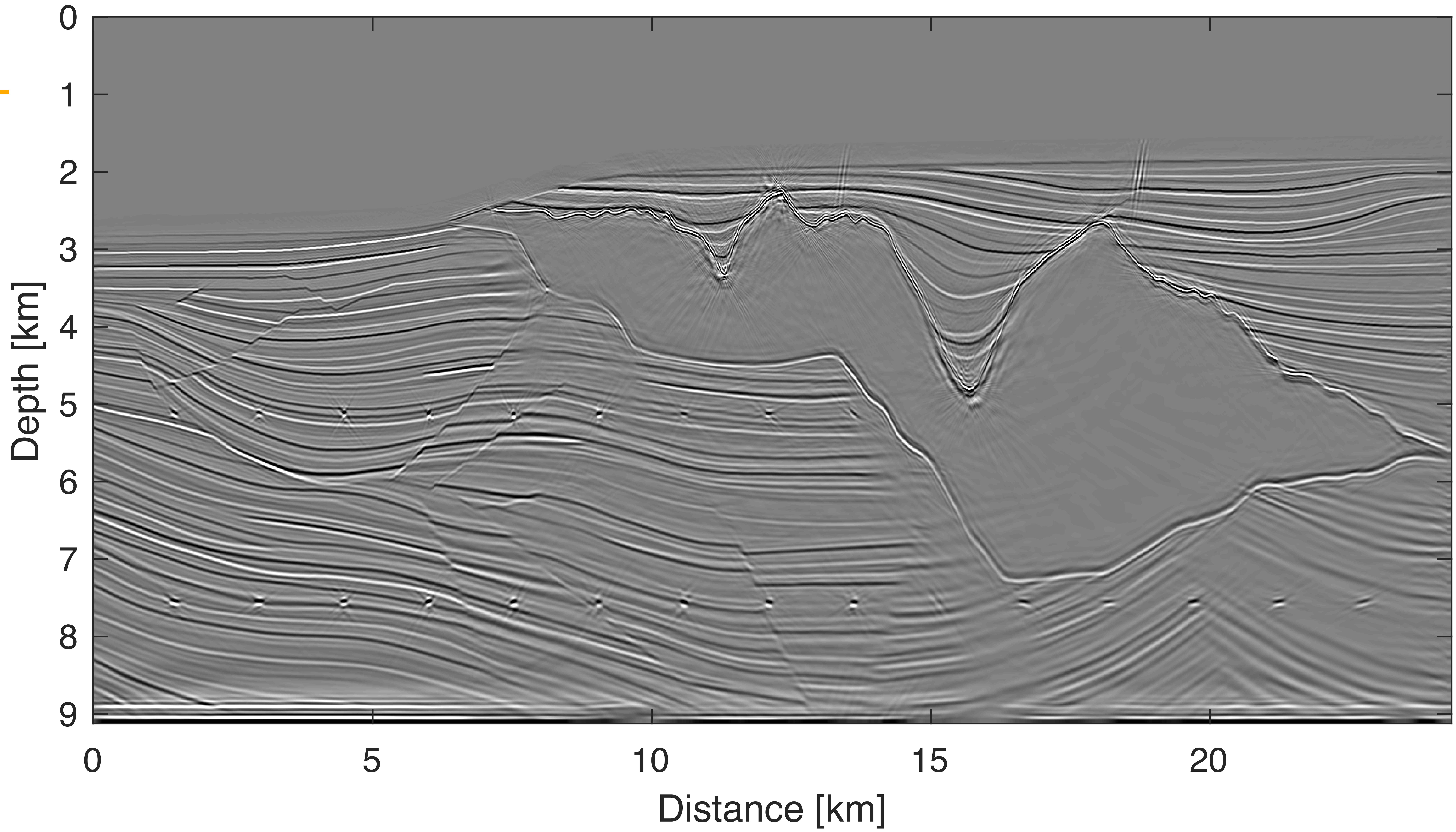
Reduced objective:

$$\min_{\mathbf{m}} \bar{f}(\mathbf{m}) = f(\mathbf{m}, \bar{\mathbf{s}}(\mathbf{m})) \quad \text{where} \quad \bar{\mathbf{s}}(\mathbf{m}) = \arg\min_{\mathbf{s}} f(\mathbf{m}, \mathbf{s})$$

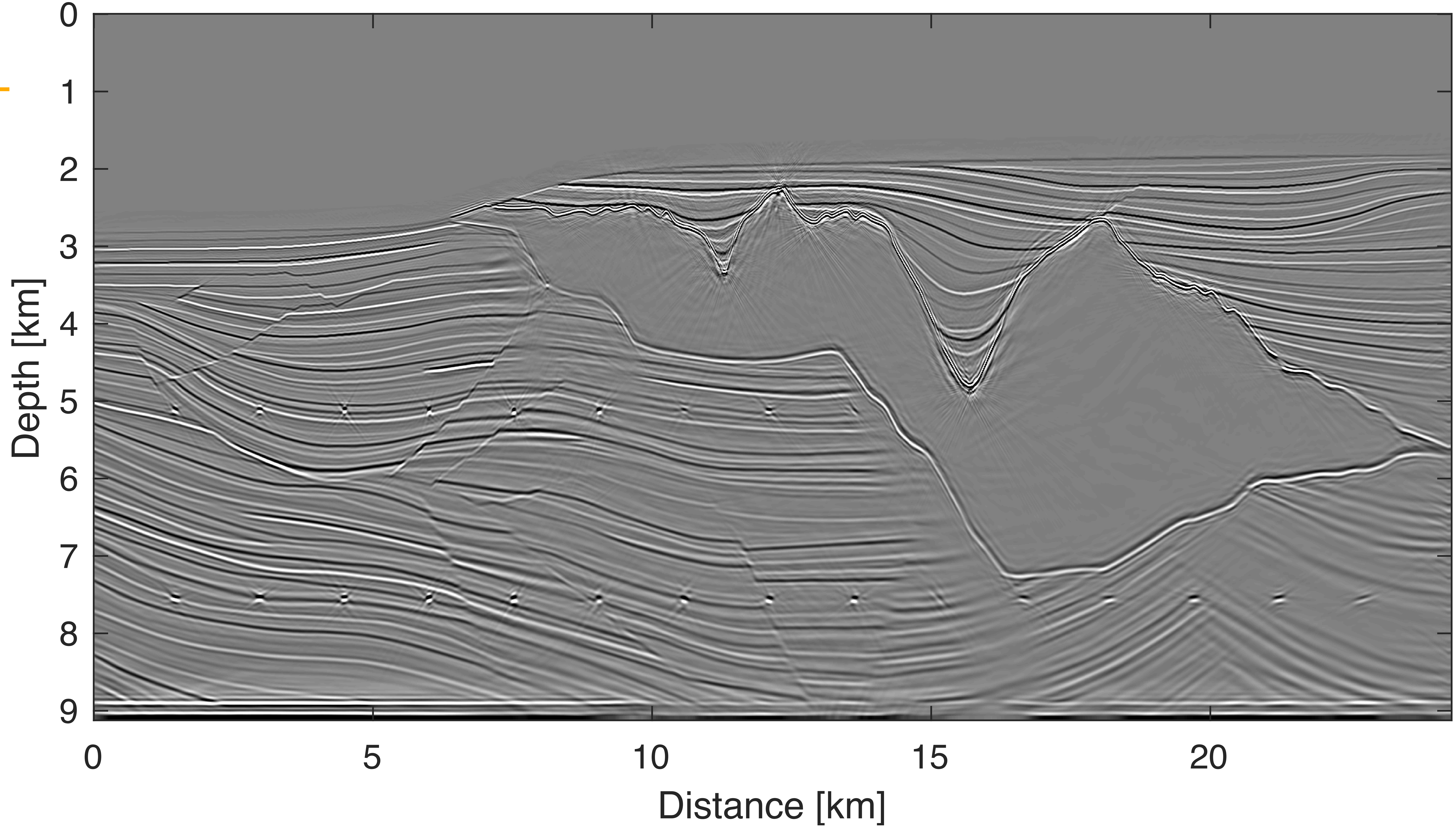
Gradient w.r.t. $\mathbf{m}$ for all n_s sources:

$$\mathbf{g} = \nabla \bar{f}(\mathbf{m}) = \frac{1}{2\sigma^2} \sum_{i=1}^{n_s} \nabla \mathbf{F}_i^\top \left(\mathbf{F}[\mathbf{m}]\mathbf{q}^\delta[\bar{\mathbf{s}}] - \mathbf{d}_i \right)$$

Compressive imaging w/ known source



Compressive imaging w/ source estimation



Observations

Seismic imaging w/ on-the-fly source estimation

- works as well or better
- extension of the search space (a la AWI/WRI)

Q: H*ow can we leverage active source seismic technology to medical imaging?*

- unifying theory
- computational framework

Observations

Seismic imaging w/ on-the-fly source estimation

- works as well or better
- extension of the search space (a la AWI/WRI)

Q: *how can we leverage active source seismic technology to medical imaging?*

- **unifying theory**
- computational framework

Extended volume source imaging

Nonlinear least squares optimization problem:

$$\min_{\mathbf{m},\mathbf{u}} f(\mathbf{m},\mathbf{u}) = \frac{1}{2\sigma^2}\sum_{i=1}^{n_s} \|\mathbf{d}_i - \mathbf{F}[\mathbf{m}]\mathbf{q}^e[\mathbf{u}_i]\|_2^2 + \lambda R_x(\mathbf{u_i})$$

$\mathbf{q}^e[\mathbf{u}_i]$ extended source w/ known wavelet & unknown location

$\lambda R_x(\mathbf{u}_i)$ spatial penalty term

Sources considered discretizations of $u_i(\mathbf{x})s(t)$

u_i is unknown function of space only

$s(t)$ is known causal source-time function

SLIM

Aleksandr Y. Aravkin and Tristan van Leeuwen, **“Estimating nuisance parameters in inverse problems”**
Huang, G., Nammour, R., and Symes, W. W., **“Volume source-based extended waveform inversion”**

Variable projection

Reduced objective:

$$\min_{\mathbf{m}} \bar{f}(\mathbf{m}) = f(\mathbf{m}, \bar{\mathbf{u}}(\mathbf{m})) \quad \text{where} \quad \bar{\mathbf{u}}(\mathbf{m}) = \arg\min_{\mathbf{s}} f(\mathbf{m}, \mathbf{u})$$

Gradient w.r.t. $\mathbf{m}$ for all n_s sources:

$$\mathbf{g} = \nabla \bar{f}(\mathbf{m}) = \frac{1}{2\sigma^2} \sum_{i=1}^{n_s} \nabla \mathbf{F}_i^\top \left(\mathbf{F}[\mathbf{m}]\mathbf{q}[\bar{\mathbf{u}}_i] - \mathbf{d}_i \right)$$

Comparison active/extended source imaging

Active source imaging:

$$\mathbf{q}^{\delta}[\mathbf{u}_i] = \mathbf{e}_i \mathbf{s}^{\top}$$

- known source locations
- unknown source-time function
- **modality:** seismic exploration

Extended source imaging:

$$\mathbf{q}^{e}[\mathbf{u}_i] = \mathbf{u}_i \mathbf{s}^{\top}$$

- unknown spatial source distribution
- known source-time function
- **modality:** photoacoustic image w/ opto/thermal-induced sources & seismic exploration

Examples

Case I — Photoacoustic imaging w/ constraints

- leverage constraints vs penalties
- image through mouse's skull w/ correct wave physics

Case II — Seismic imaging w/ extended volume sources

- connection photoacoustic imaging & volume source imaging
- demonstrate ability to image w/o knowledge source locations

Case III — Photoextended imaging

- demonstration new photoacoustic imaging modality w/ extended volume source imaging

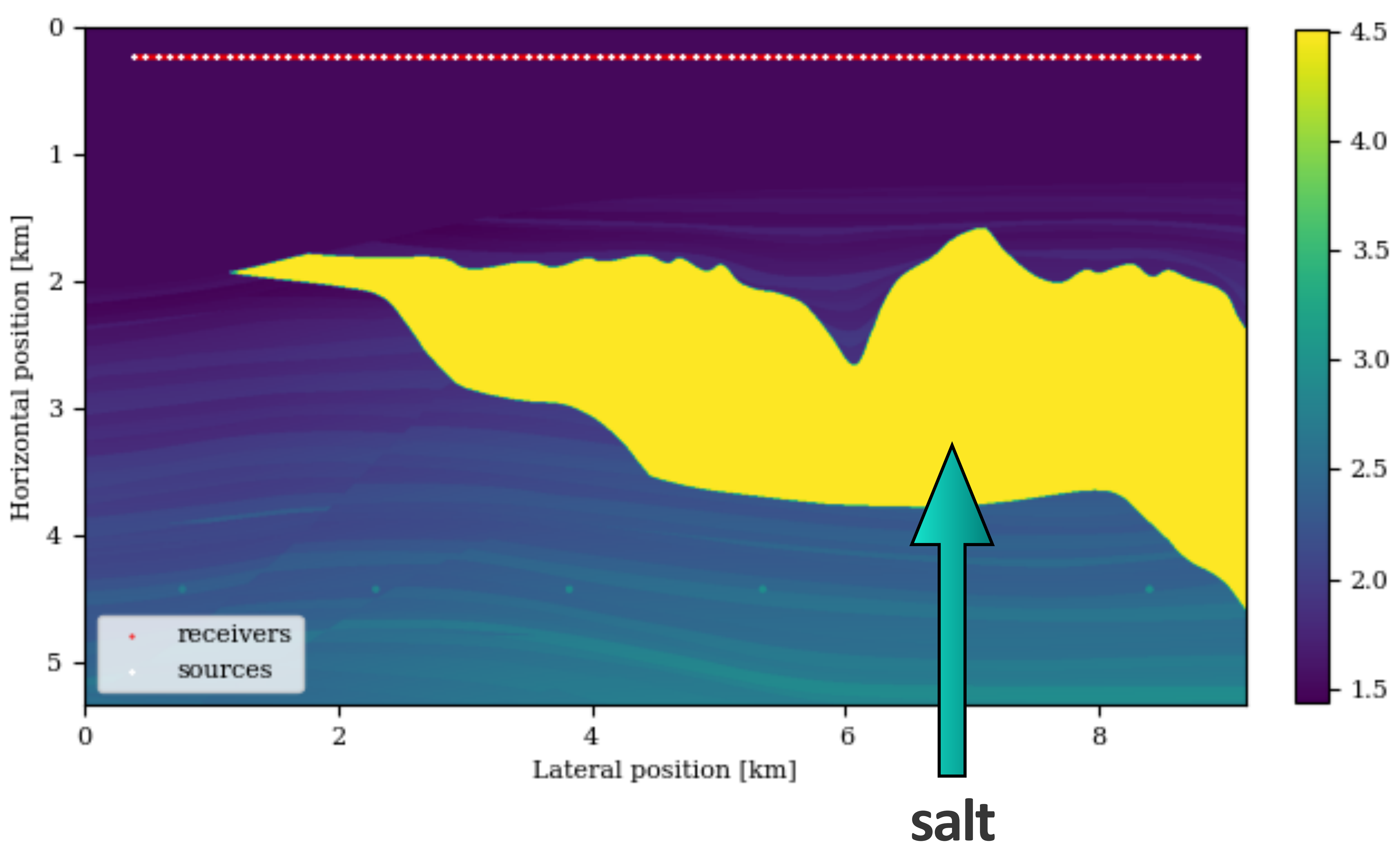
Compare seismic & photoacoustic imaging

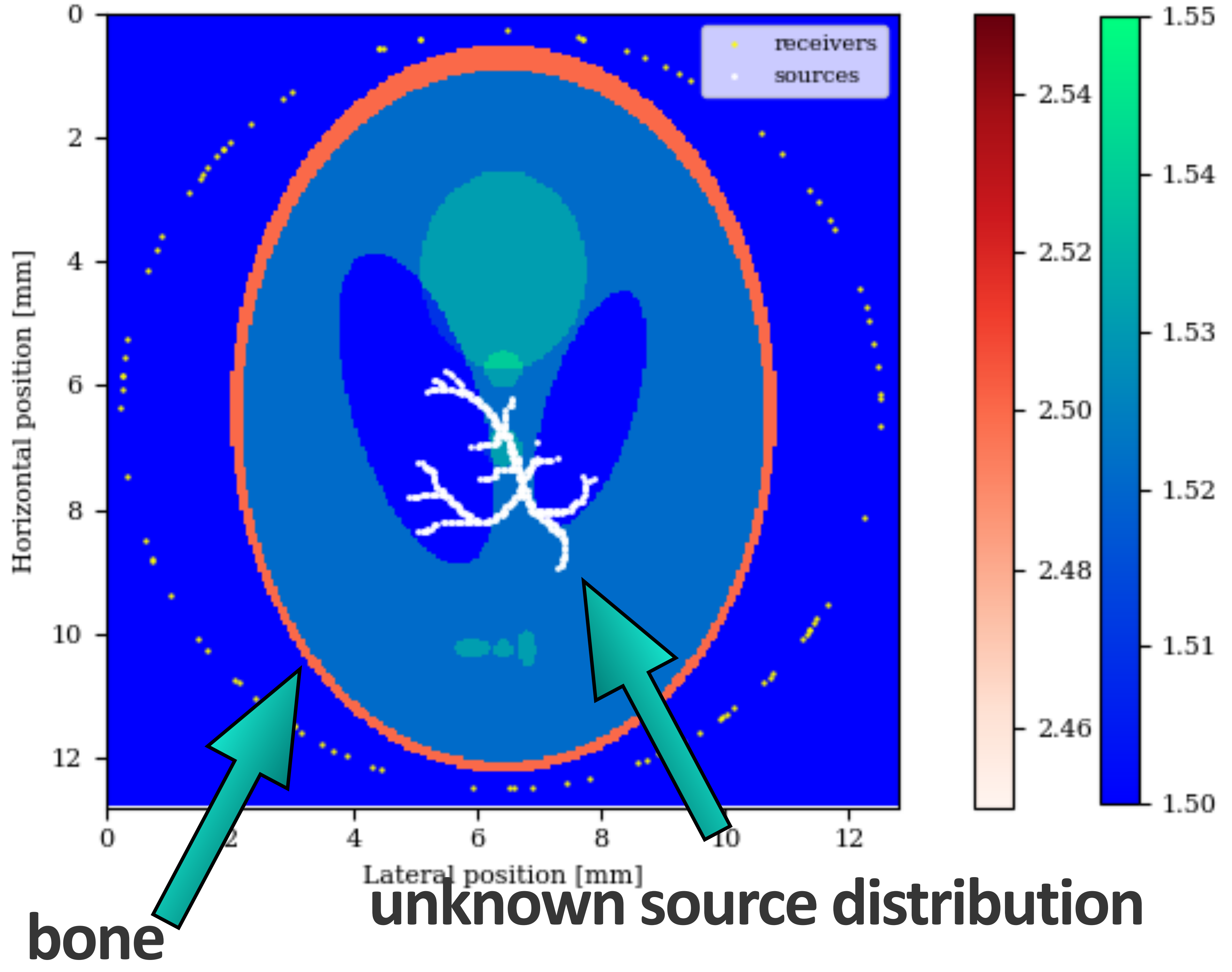
Seismic imaging:

- known sources
- acquisition only on the top
- locate oil & gas
- image through salt

Photoacoustic imaging:

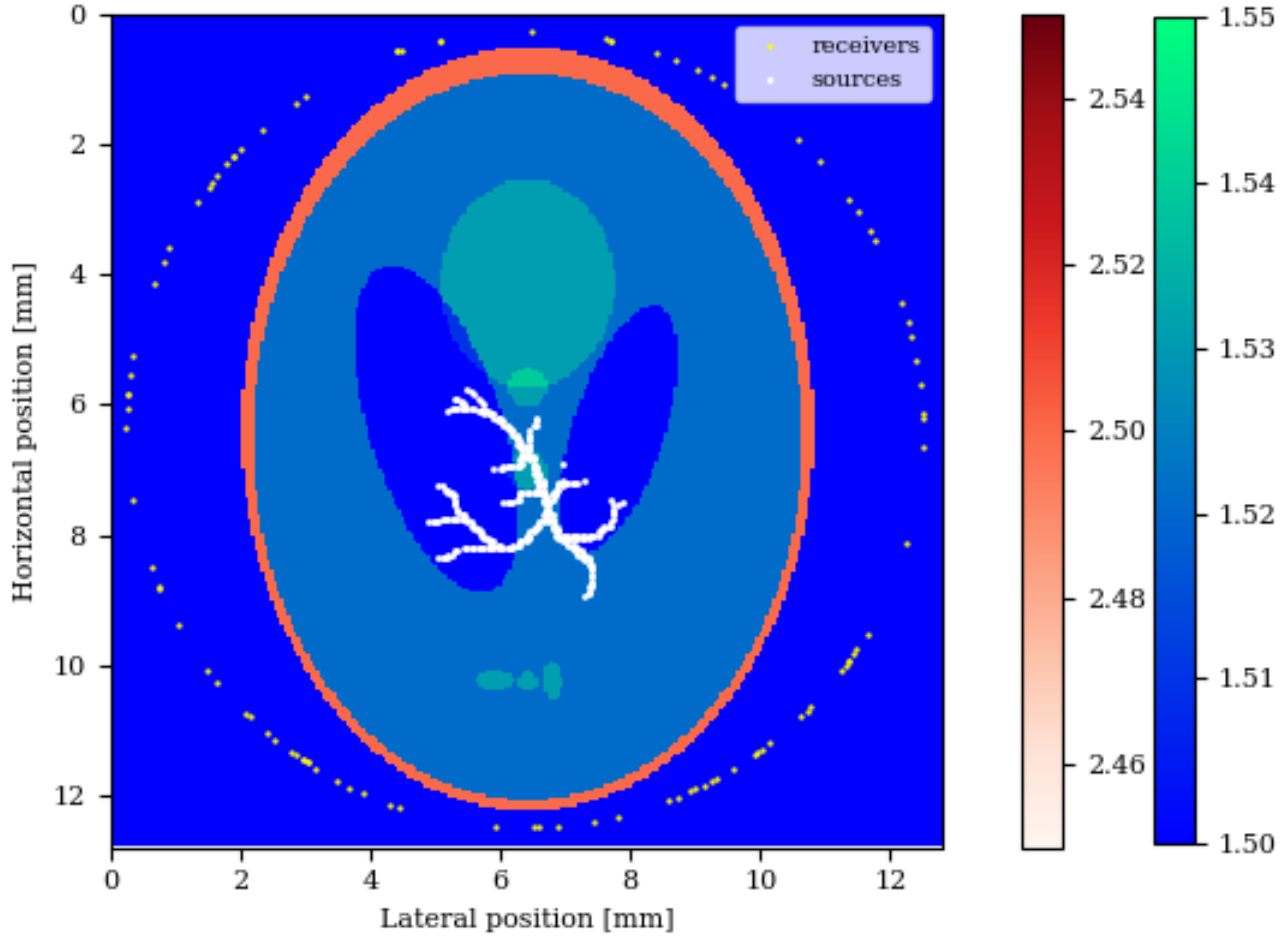
- unknown source distribution
- acquisition all around
- detect anomalous tissue
- image through skull



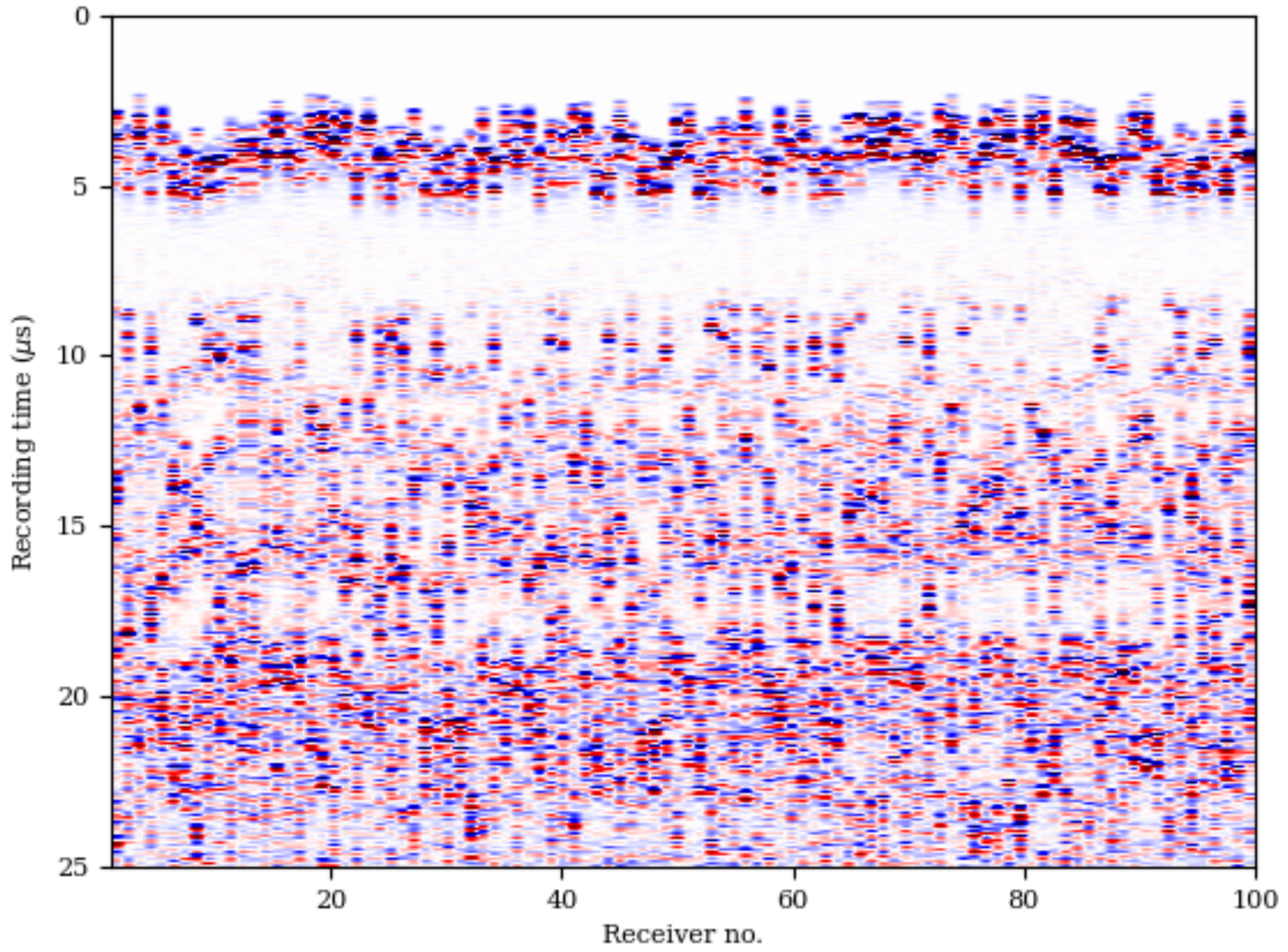


Photoacoustic imaging w/ TV constraints

$$\min_{\mathbf{u}} \frac{1}{2\sigma^2}\|\mathbf{d}-\mathbf{F}[\mathbf{m}]\mathbf{q}^e[\mathbf{u}]\|_2^2 \quad \text{subject to} \quad \|\mathbf{u}\|_{TV} \leq \tau$$



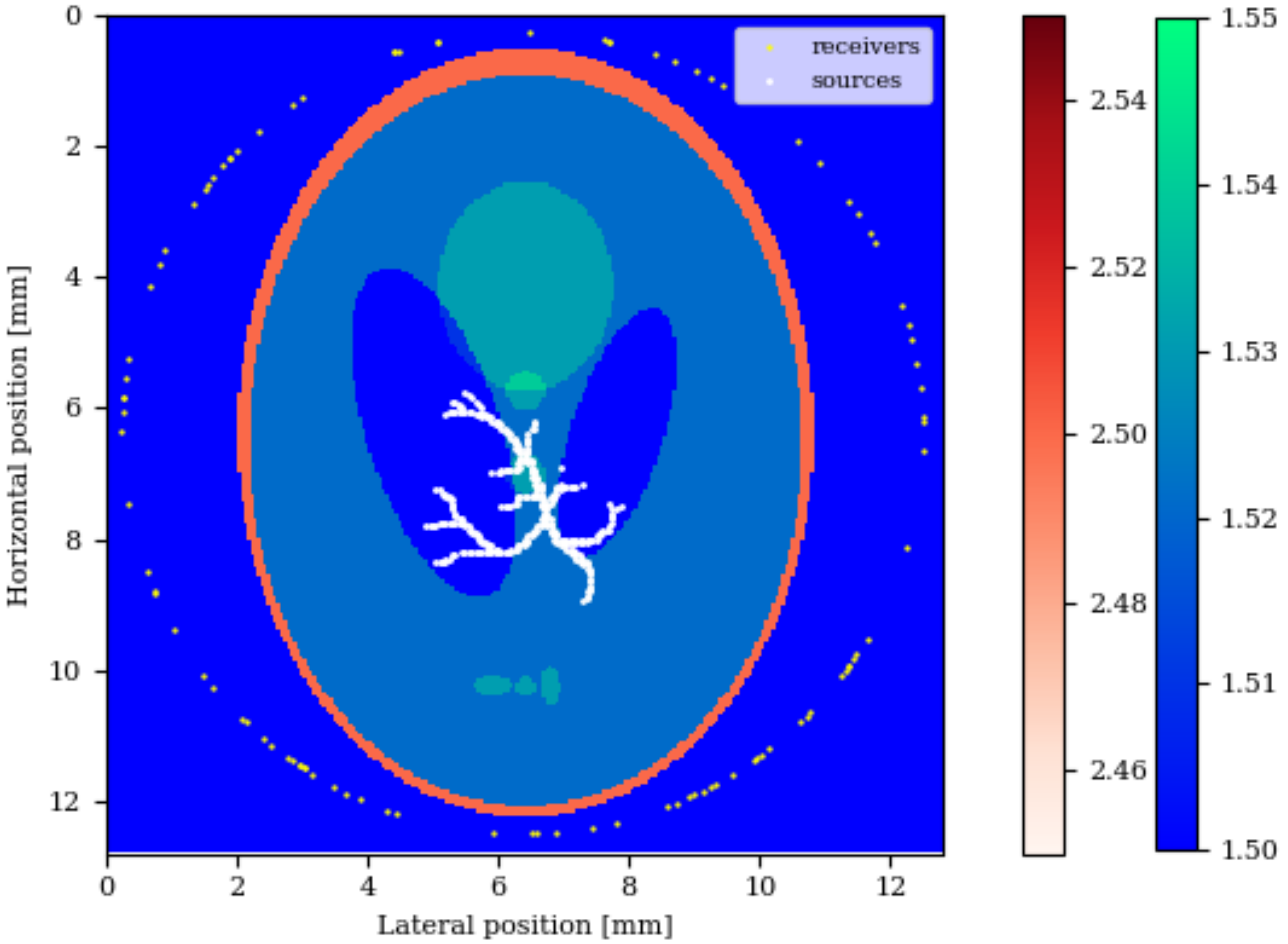
Acquisition

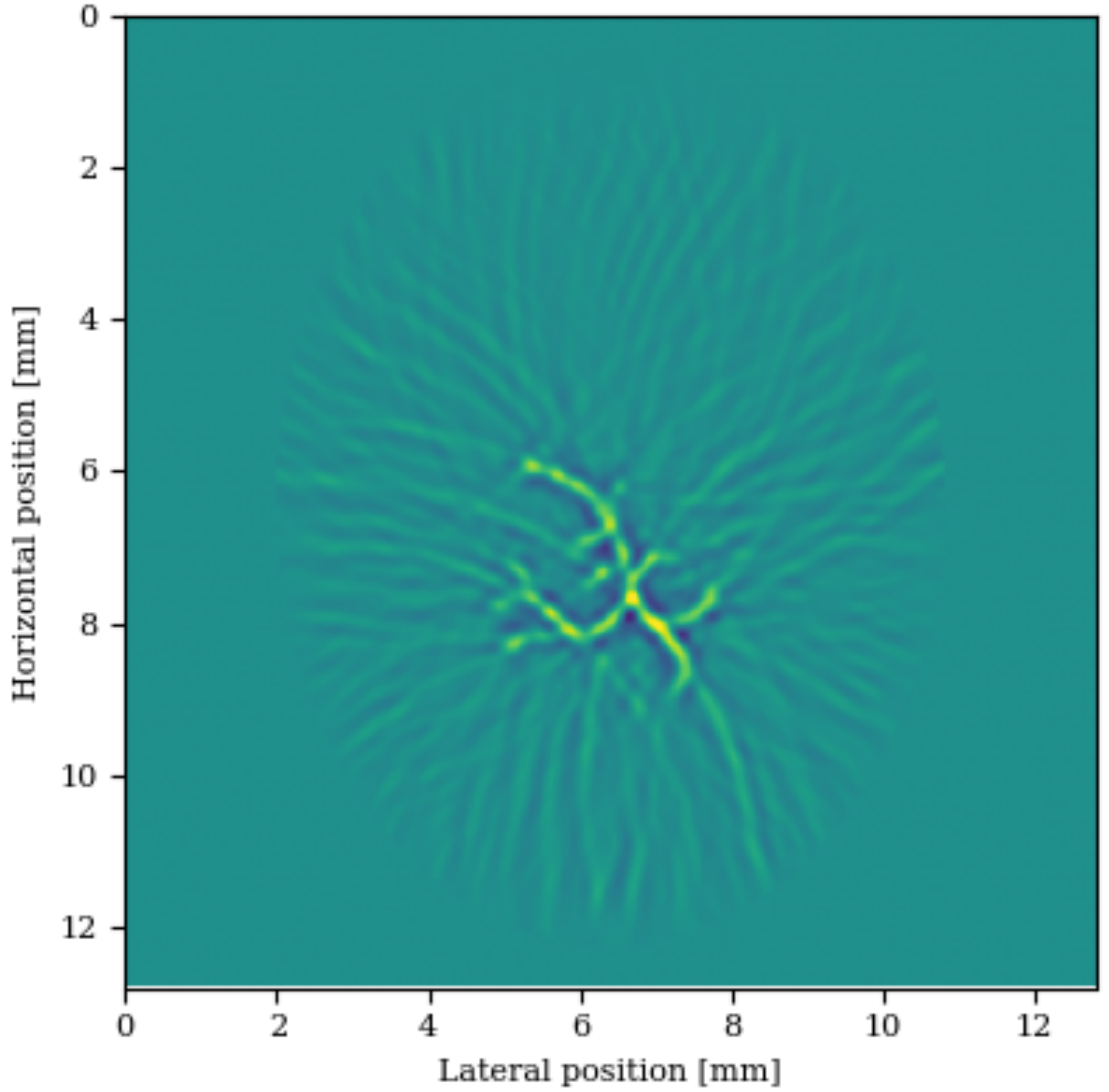


Data

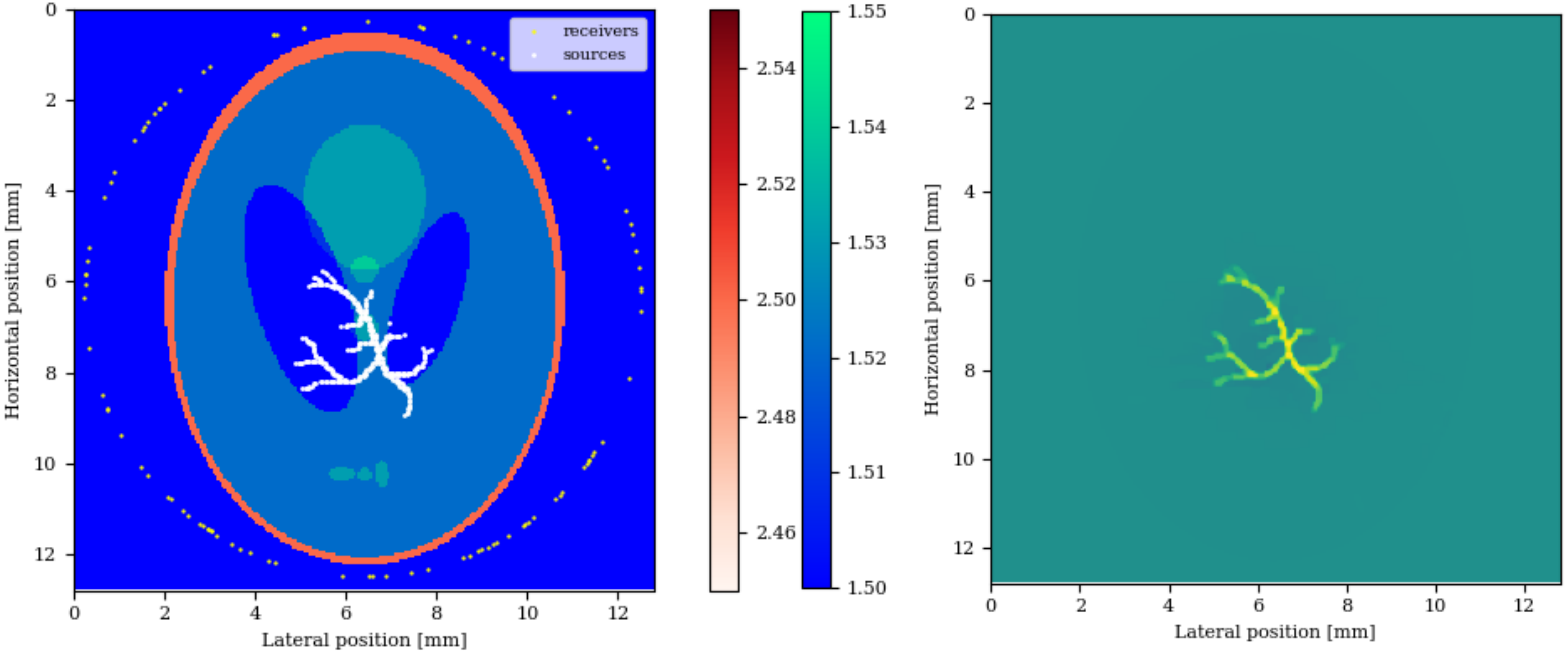
Ernie Esser, Lluís Guasch, Tristan van Leeuwen, Aleksandr Y. Aravkin, and Felix J. Herrmann, “Total-variation regularization strategies in full-waveform inversion”
Bas Peters, Brendan R. Smithyman, and Felix J. Herrmann, “Projection methods and applications for seismic nonlinear inverse problems with multiple constraints”

Back propagation

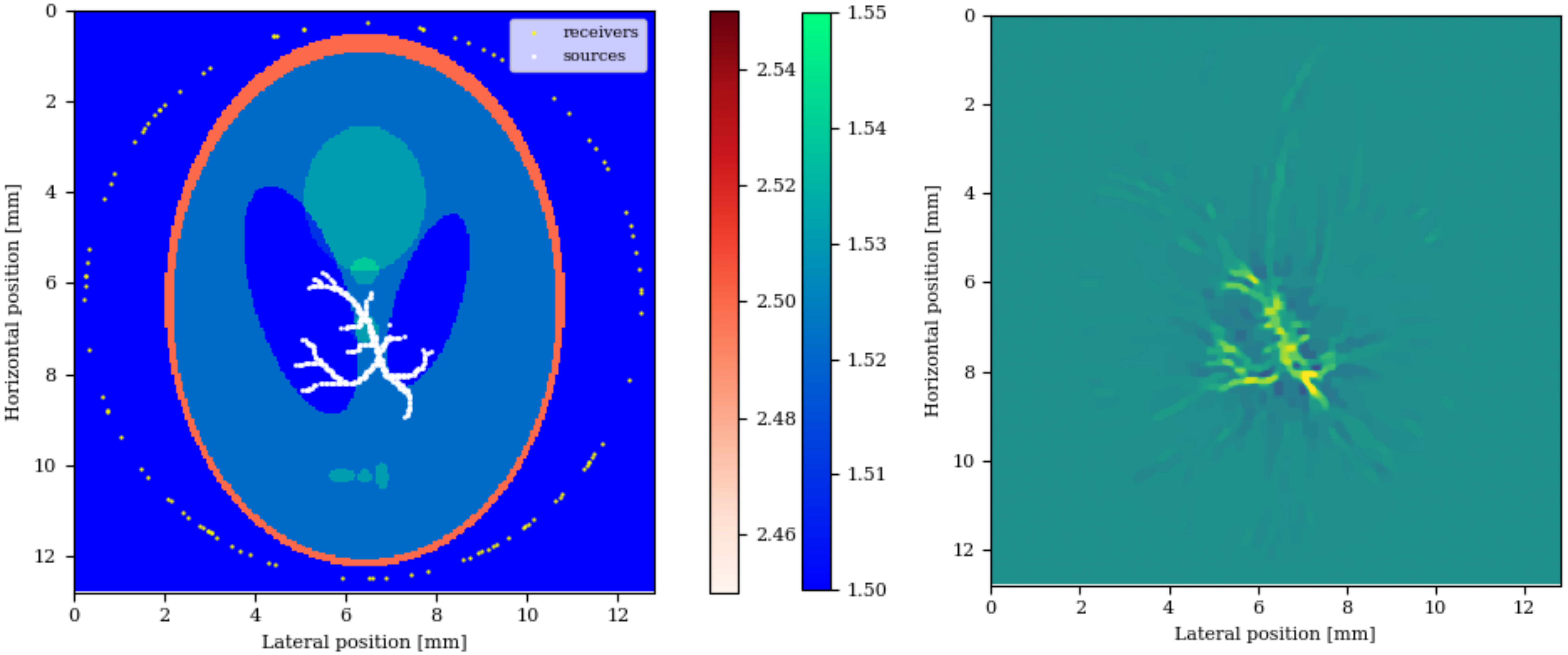




Inversion w/ TV constraints + correcto velocity



Inversion w/ TV constraints w/ wrong water velocity



Examples

Case I — Photoacoustic imaging w/ constraints

- leverage constraints vs penalties
- image through mouse skull w/ correct wave physics

Case II — Seismic imaging w/ extended volume sources

- connection photoacoustic imaging & volume source imaging
- demonstrate ability to image w/o knowledge source locations

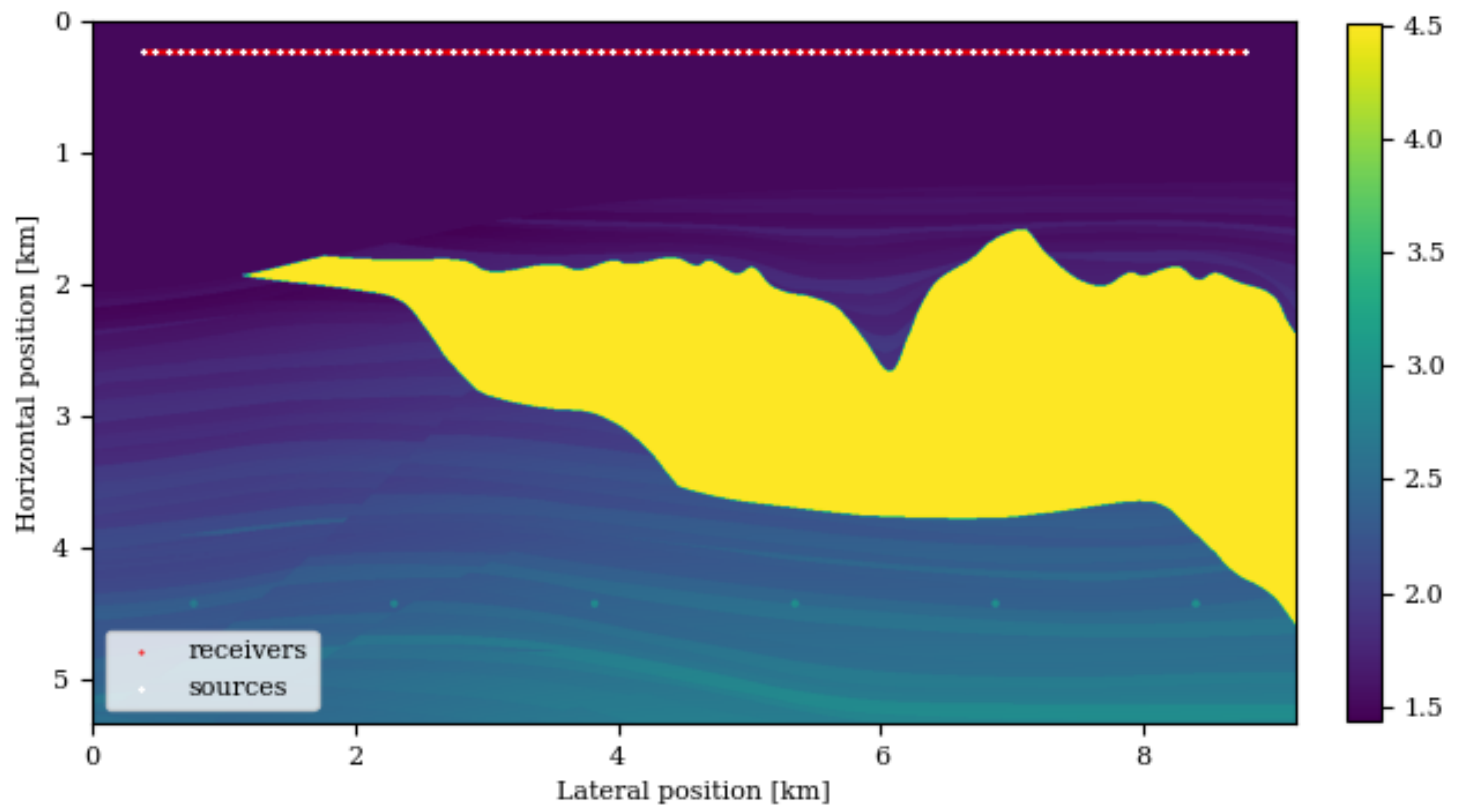
Case III — Photoextended imaging

- demonstration new photoacoustic imaging modality w/ extended volume source imaging

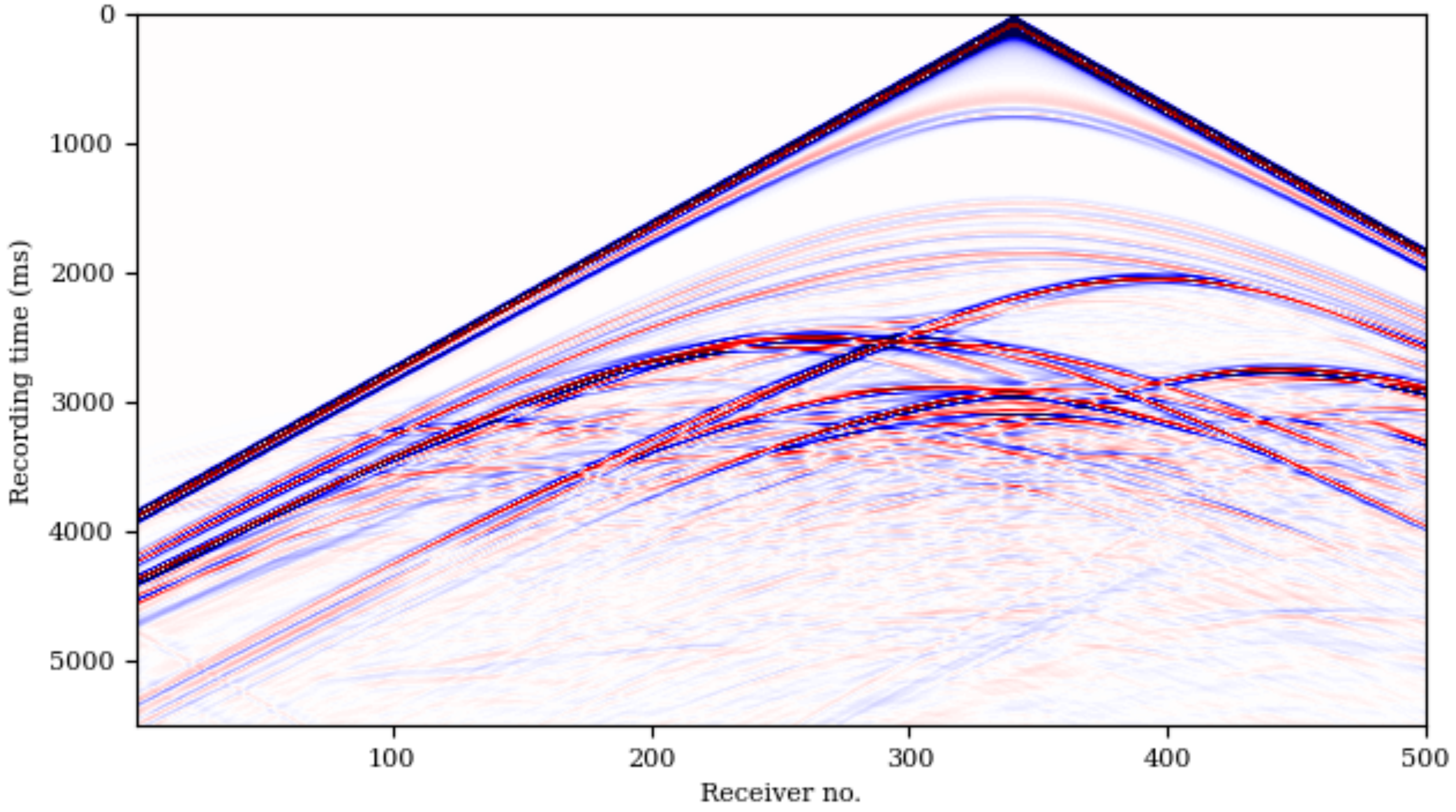
SLIM

Seismic imaging w/ focused sources

$$R(\mathbf{u}) = \|\mathbf{W}\mathbf{u}\|_2^2 \quad \text{where} \quad \mathbf{W}\mathbf{u} = |\mathbf{x} - \mathbf{x}_s|\mathbf{u}$$

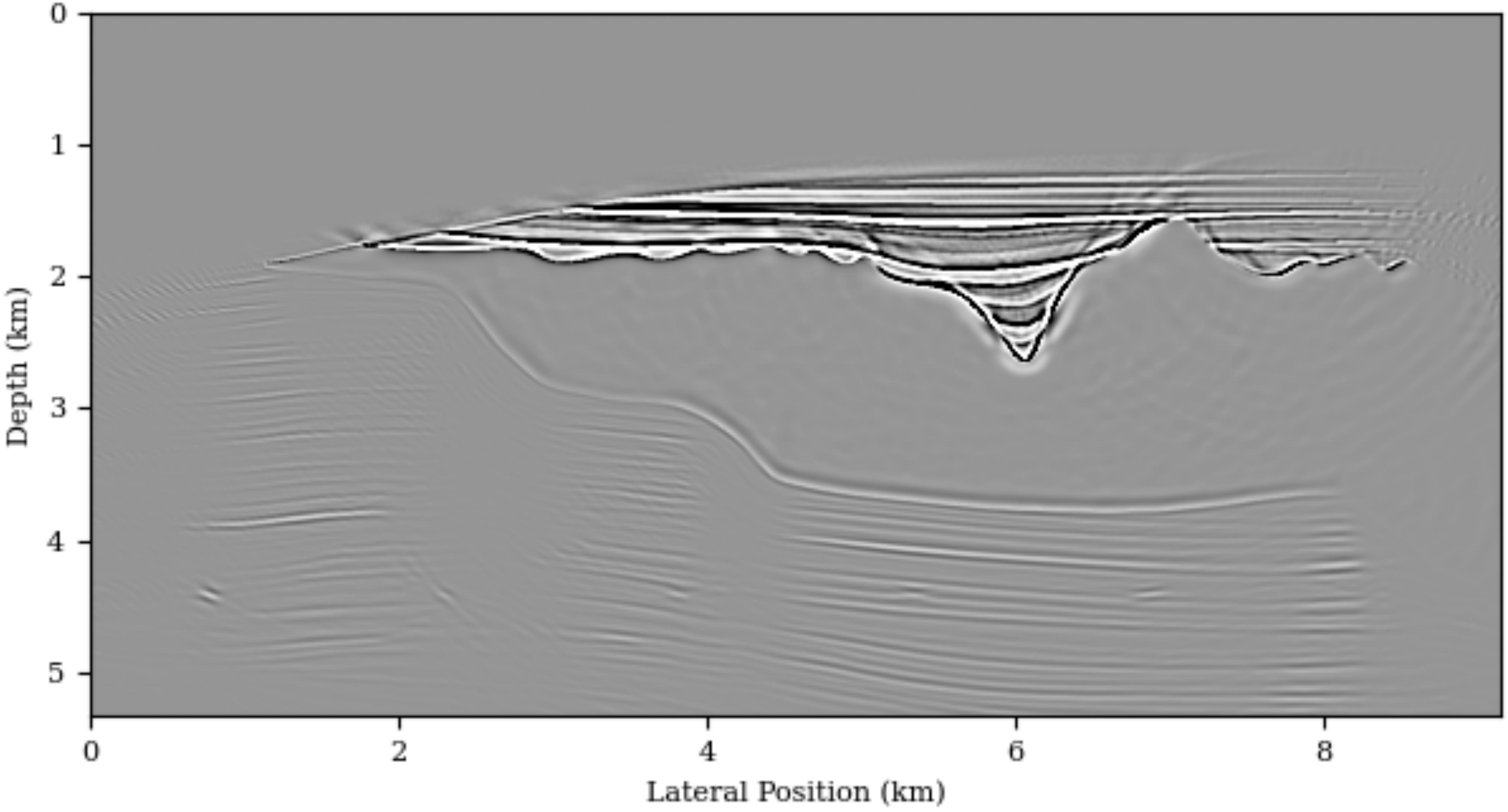


Acquisition

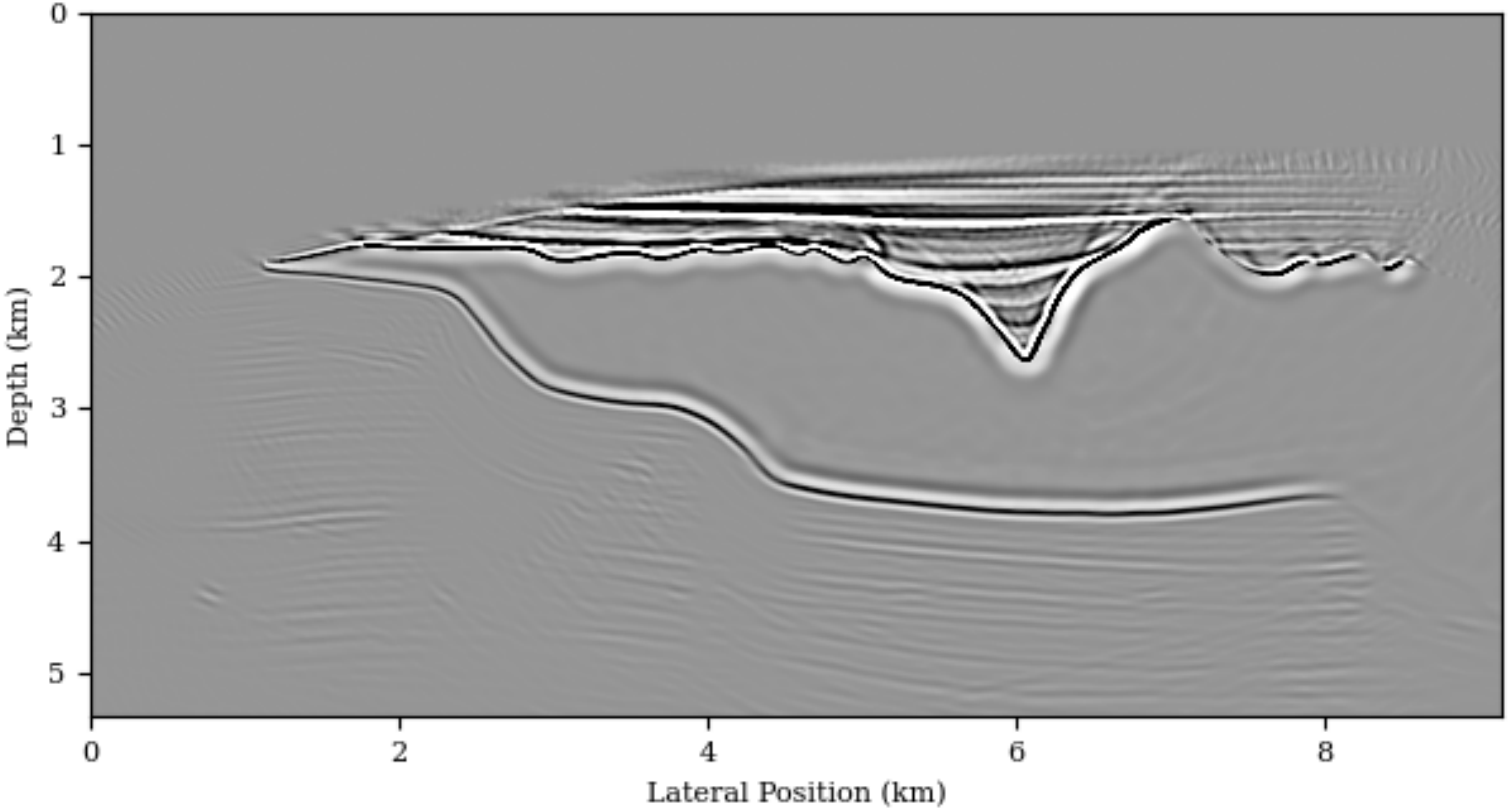


Shot record

RTM imaging



Extended source imaging



Examples

Case I — Photoacoustic imaging w/ constraints

- leverage constraints vs penalties
- image through mouse skull w/ correct wave physics

Case II — Seismic imaging w/ extended volume sources

- connection photoacoustic imaging & volume source imaging
- demonstrate ability to image w/o knowledge source locations

Case III — Photoextended imaging

- demonstration new photoacoustic imaging modality w/ extended volume source imaging

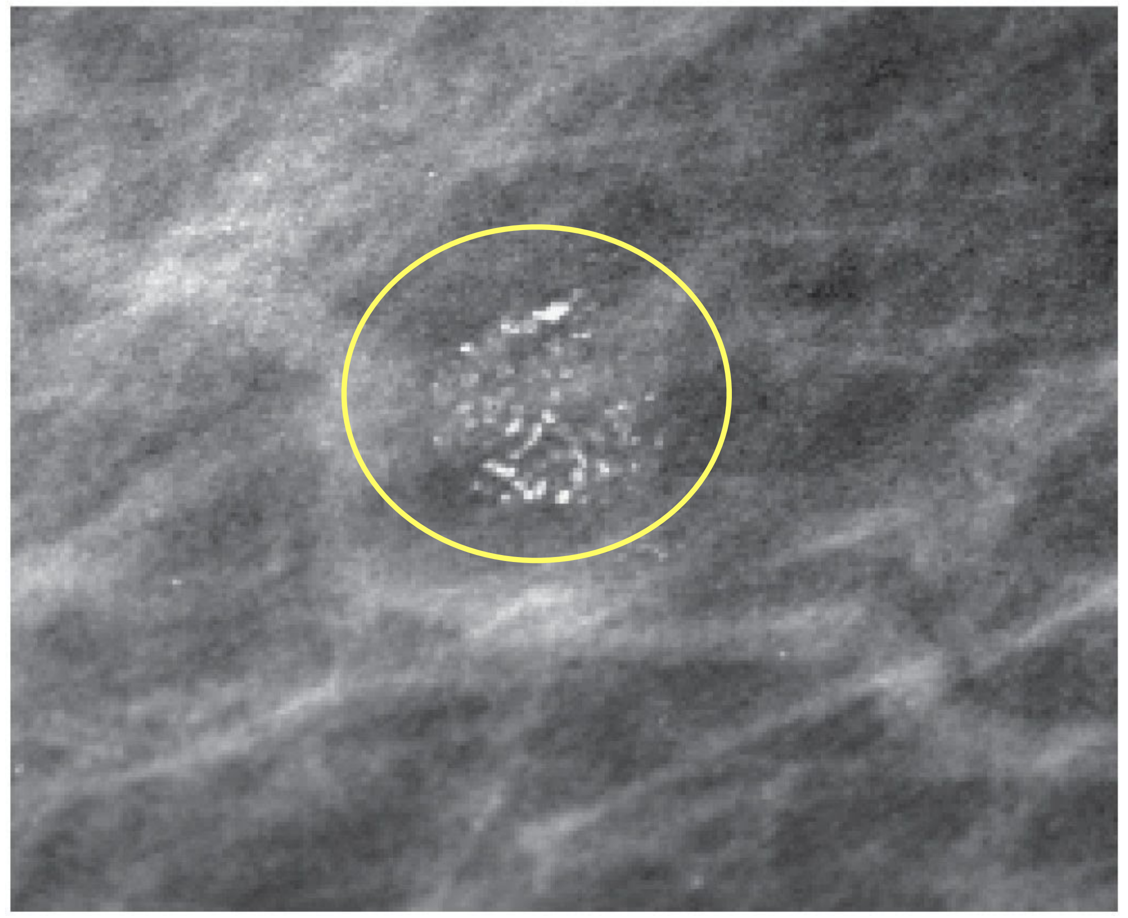
[1] Gregory, Adriana, et al. "Effect of calcifications on breast ultrasound shear wave elastography: an investigational study." PloS one 10.9 (2015): e0137898.

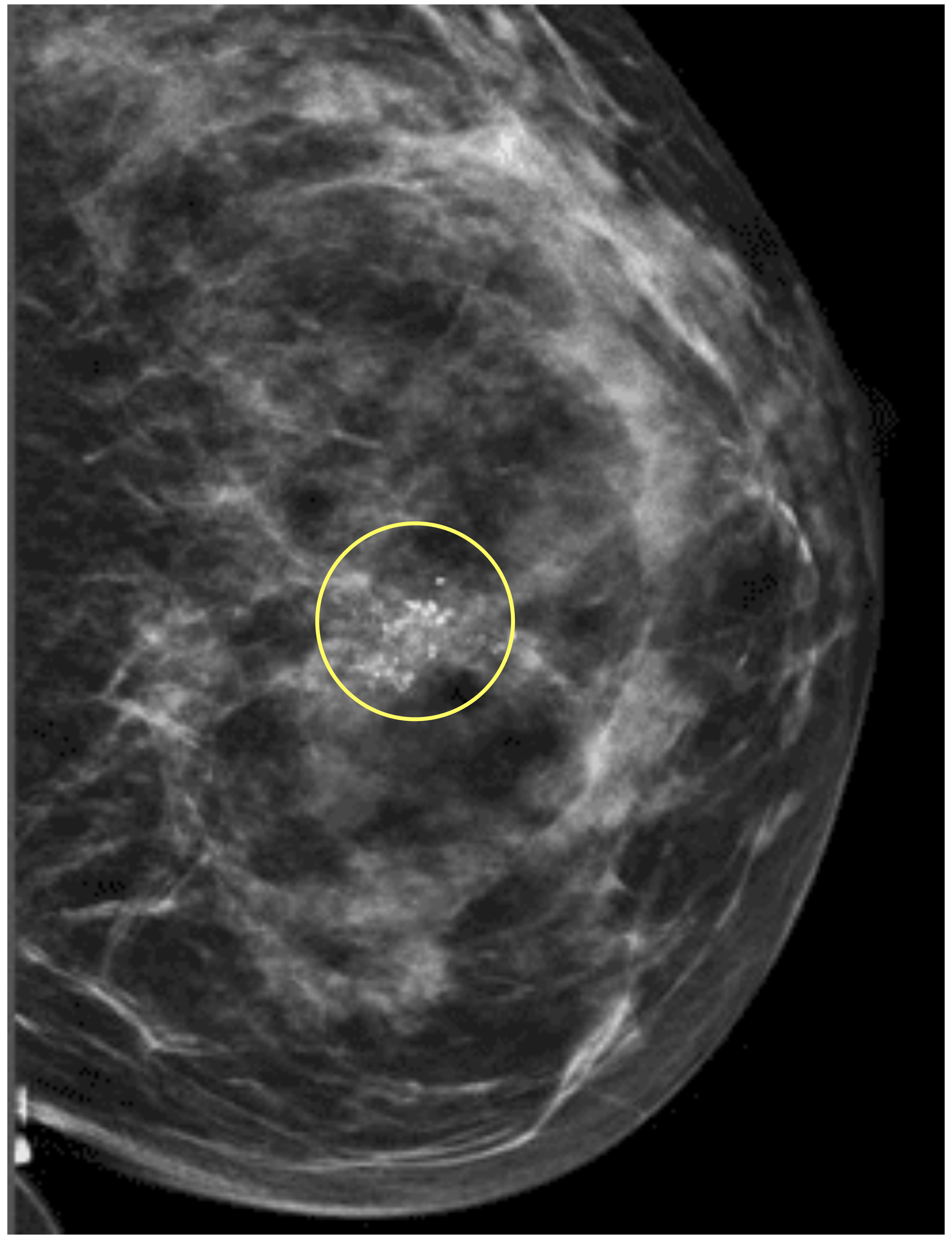
[2] Images adapted from https://info.iowaradiology.com/

Breast microcalcifications

Typically composed of:

- calcium oxalate
- calcium hydroxyapatite





Can be precursors to cancer

Highly dense compounds so high speed of sound contrast

[1] Lou, Yang, et al. "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging." Journal of Biomedical Optics 22.4 (2017): 041015.

[2] Breast model downloaded from https://anastasiolab.wustl.edu/downloadable-content/oa-breast-database/

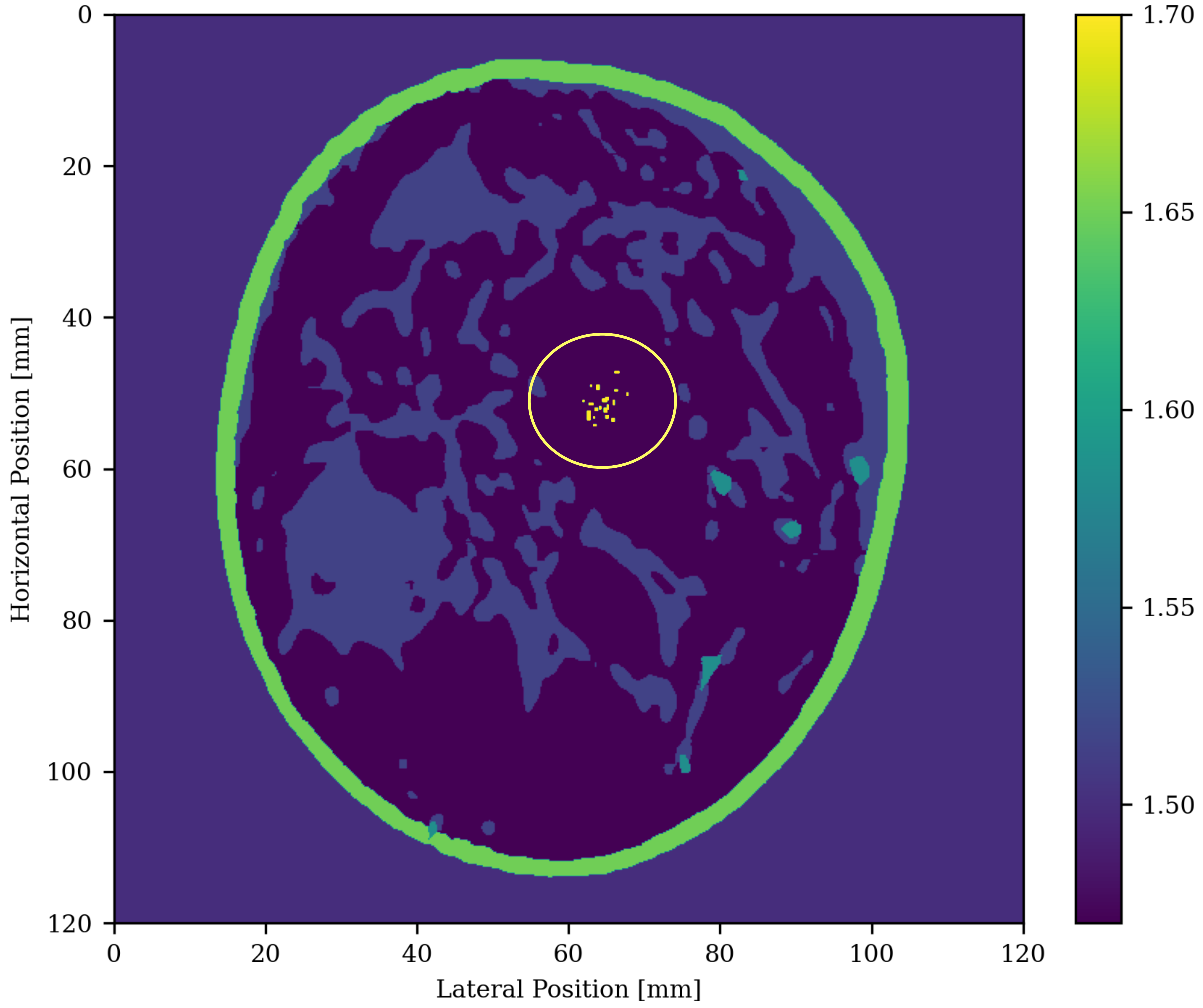
[3] Deganello, bS. "The structure of whewellite, CaC2O4. H2O at 328 K." Acta Crystallographica Section B: Structural Crystallography and Crystal Chemistry 37.4 (1981): 826-829.

[4] Cleverley, J. R., A. R. Jackson, and A. C. Bateman. "Pre-operative localization of breast microcalcification using high-frequency ultrasound." Clinical radiology 52.12 (1997): 924-926.

Breast model

Starting from breast model[1], we simulate calcifications:

- number of specks ≈ 20
- size $0.2 - 0.6\,\mathrm{mm}$
- speed of sound $3432\,\mathrm{m/s}$



[1] Nyayapathi, Nikhila, and Jun Xia. "Photoacoustic imaging of breast cancer: a mini review of system design and image features." Journal of biomedical optics 24.12 (2019): 121911.

SLIM

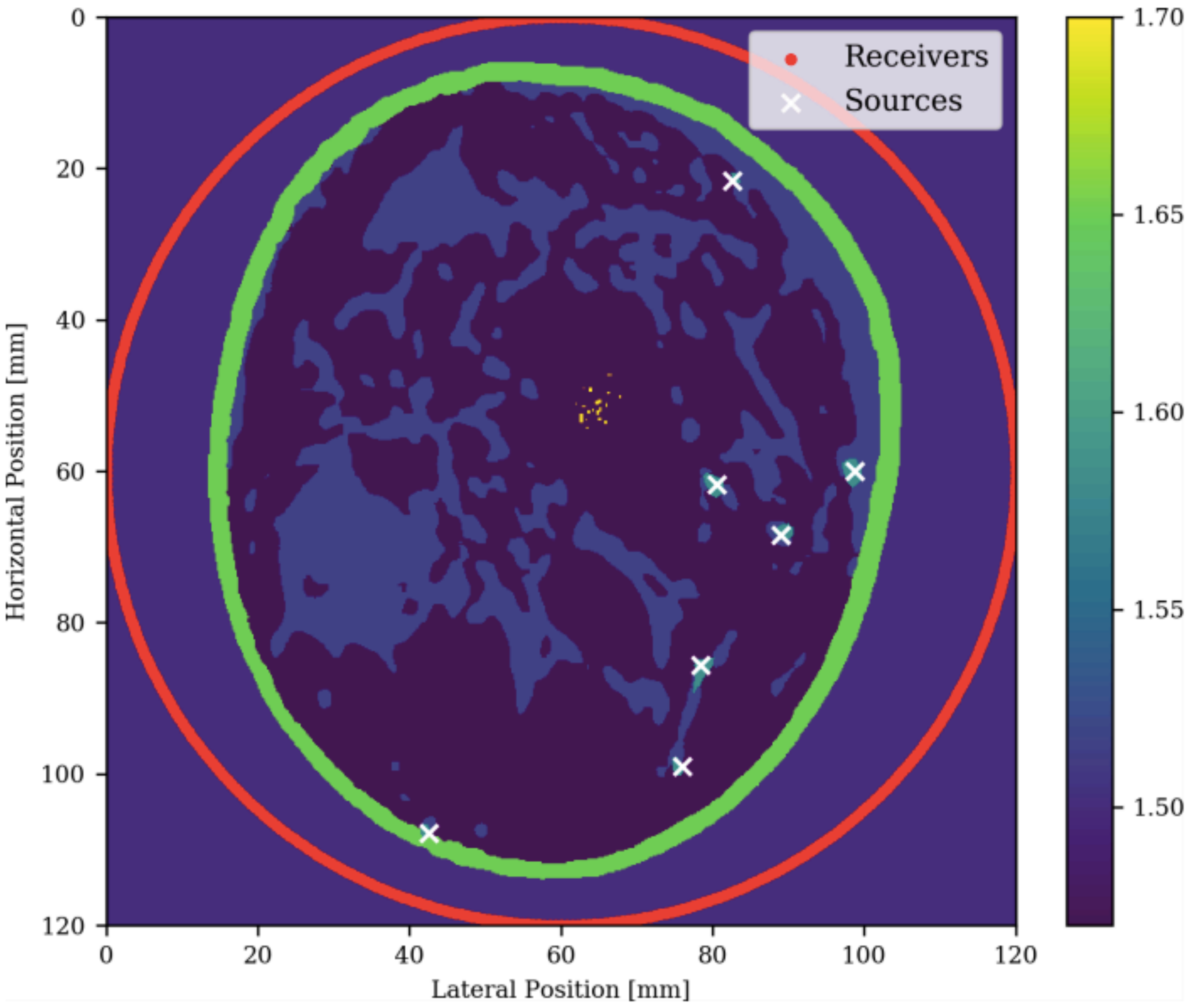
Breast model

Acquisition:

- 512 receivers in 12cm diameter ring
- laser based acoustic stimulus

Blood vessels act as secondary photoacoustic sources

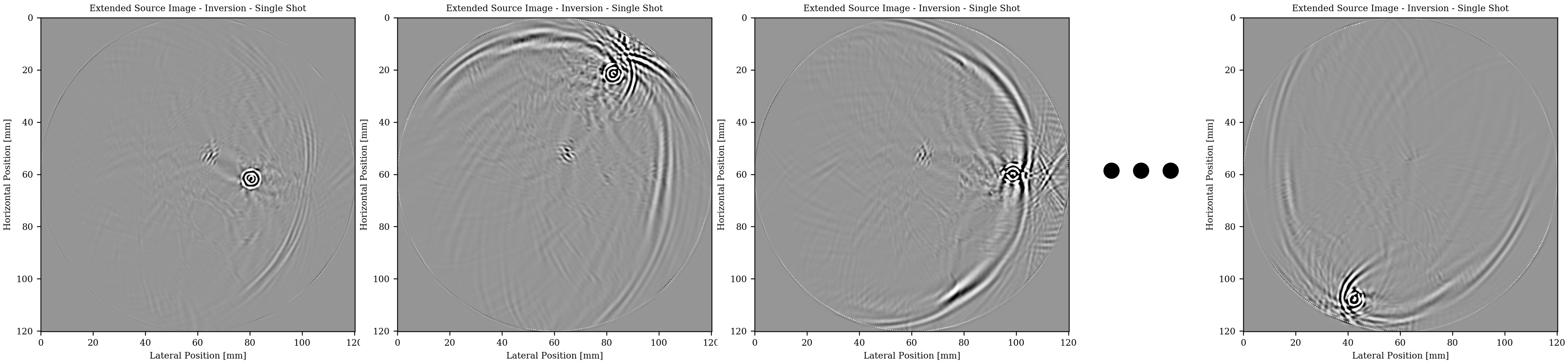
Source-time function modeled by 6MHz Ricker wavelet



Breast Imaging

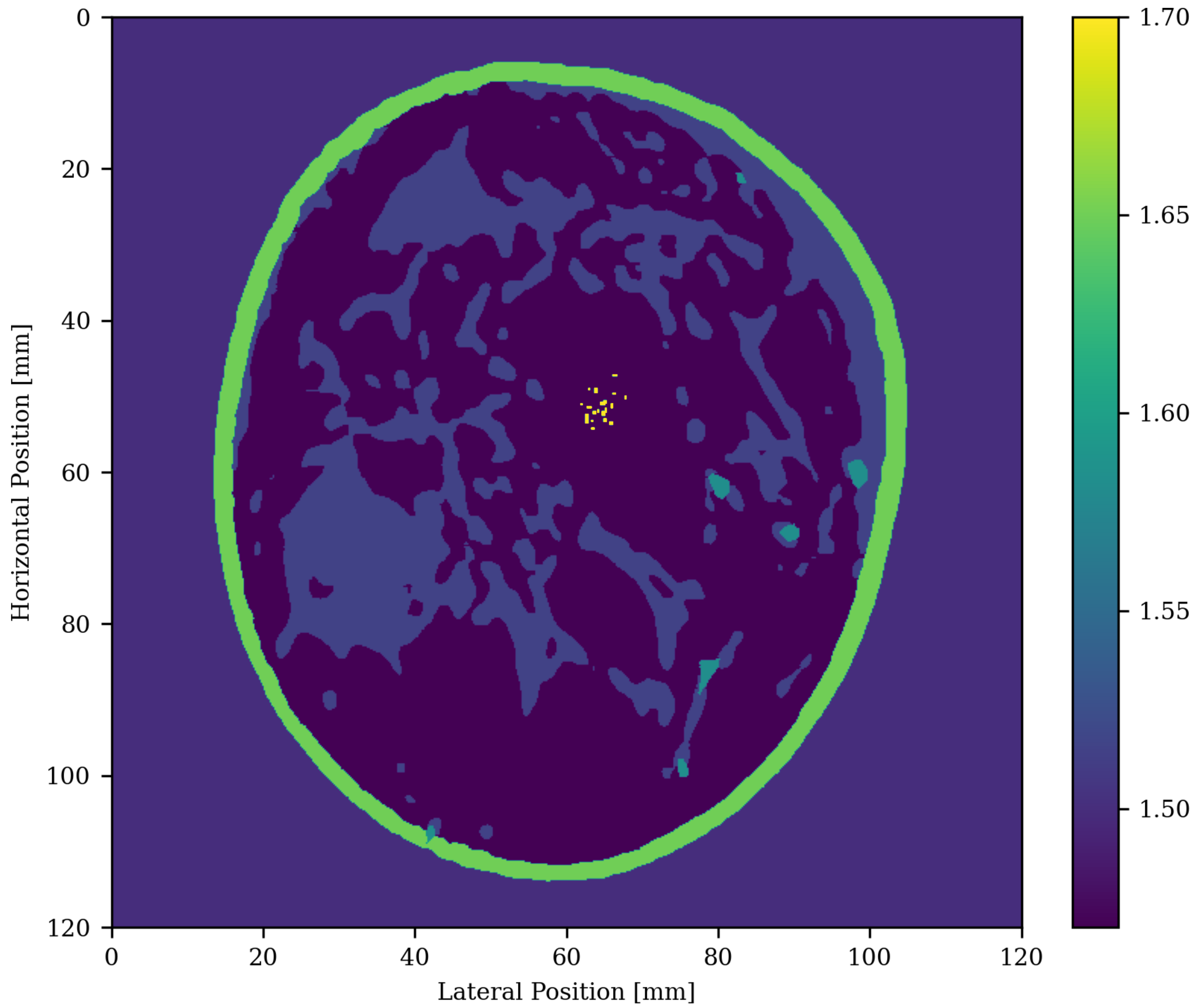
Extended source imaging procedure as in seismic

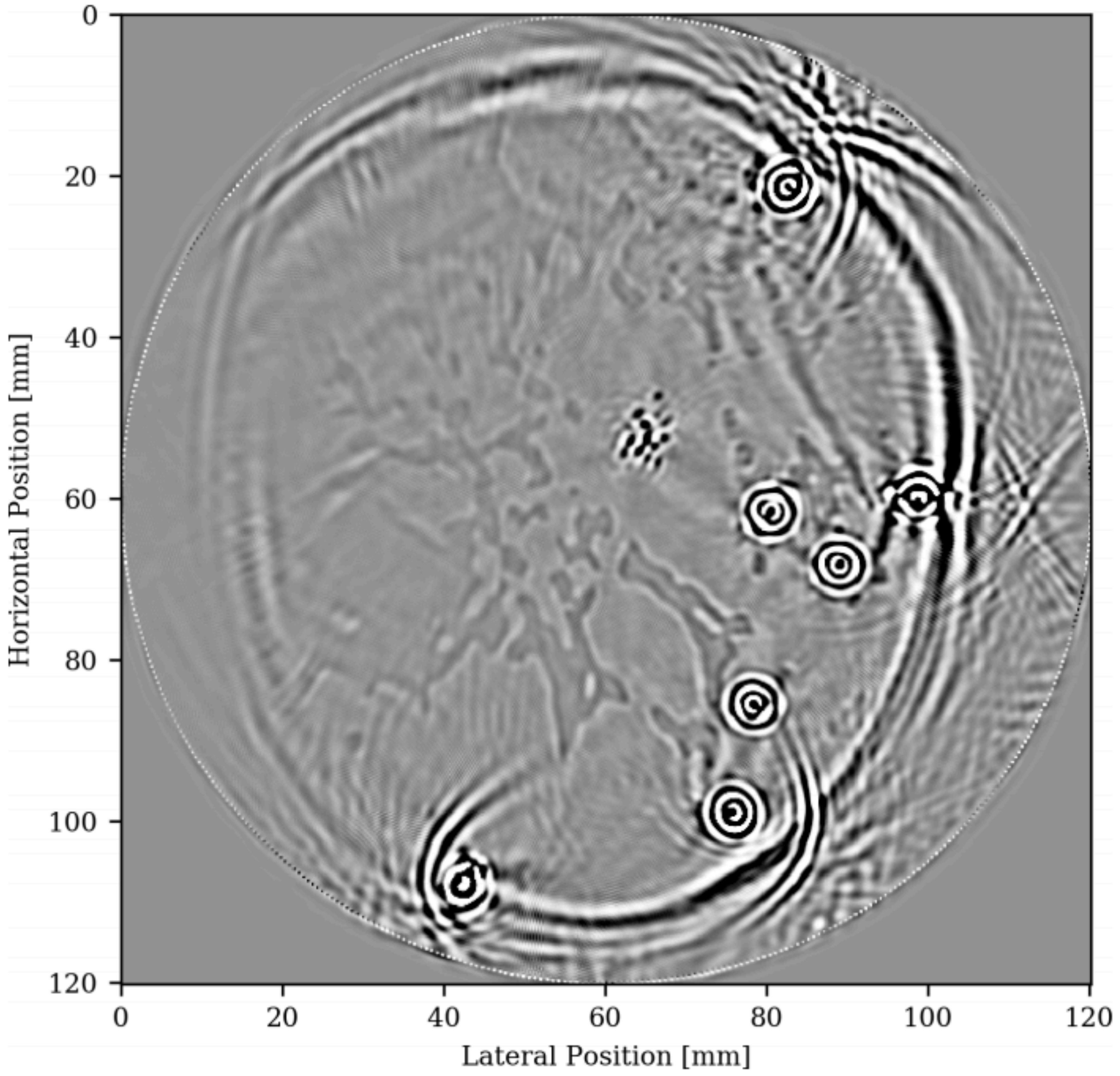
Each blood vessel illuminated separately



Breast Imaging

Stacked gradients for final image





Observations

Seismic imaging w/ extended volume source estimation

- connects seismic imaging to photoacoustic imaging
- importance including correct wave physics
- opens potentially new imaging modalities

Q: *how can we leverage active source seismic technology in medical imaging?*

- unifying theory
- computational framework

Observations

Seismic imaging w/ on-the-fly extended volume source estimation

- connects seismic imaging to photoacoustic imaging
- importance including correct wave physics
- opens potentially new imaging modalities

Q: *how can we leverage active source seismic technology in medical imaging?*

- unifying theory
- **computational framework**

The Julia Devito Inversion framework (JUDI)

Julia

linear opeartors, data containers, IO

↓ parallel modeling function

parallelization: distribute sources, data

↓ serial modeling function

interface to Devito (Python)

↓

Python

Devito: symbolic definition of PDE

↓ automatic code generation and JIT compilation

C

solve PDE w/ OpenMP parallelism

```
g = J'*(Pr*A_inv*Ps'*q - d_obs)
```

```
# Symbolic wavefield
u = TimeFunction(name='u', space_order=16)

# Acoustic wave equation
pde = model.m * u.dt2 - u.laplace
```

```
float (*rec_coords)[2] = (Float (*[2]) rec_coords_vec;
{
  struct timeval start_kernel, endkernel;
    gettimeofday(&start_kernel, NULL);
    int t0;
    int t1;
    int t2;
    ;
    for (int i3 = 0; i3 < 3; i3+=1)
    {
      flops->kernel += 2.000000;
      {
        {
          t0 = (i3)%(3);
```

[1] P. A. Witte, M. Louboutin, N. Kukreja, F. Luporini, M. Lange, G. G. Gorman and Felix J. Herrmann, 2019, *A large-scale framework for symbolic implementations of seismic inversion algorithms in Julia*, Geophysics, 84, 3.
[2] M. Louboutin, M. Lange, F. Luporini, N. Kukreja, P. A. Witte, F. J. Herrmann, P. Velesko and G. J. Gorman, 2019, Devito 3.1.0: an embedded domain-specific language for finite differences and geophysical exploration: Geoscientific model development, 12, 3.
[3] F. Luporini, M. Lange, M. Louboutin, N. Kukreja, J. Hückelheim, , C. Yount, P. A. Witte, P. H. J. Kelly, F. J. Herrmann and G. J. Gorman, 2019, Architecture and performance of Devito, a system for automated stencil computation, arXiv preprints.

The Julia Devito Inversion framework (JUDI)

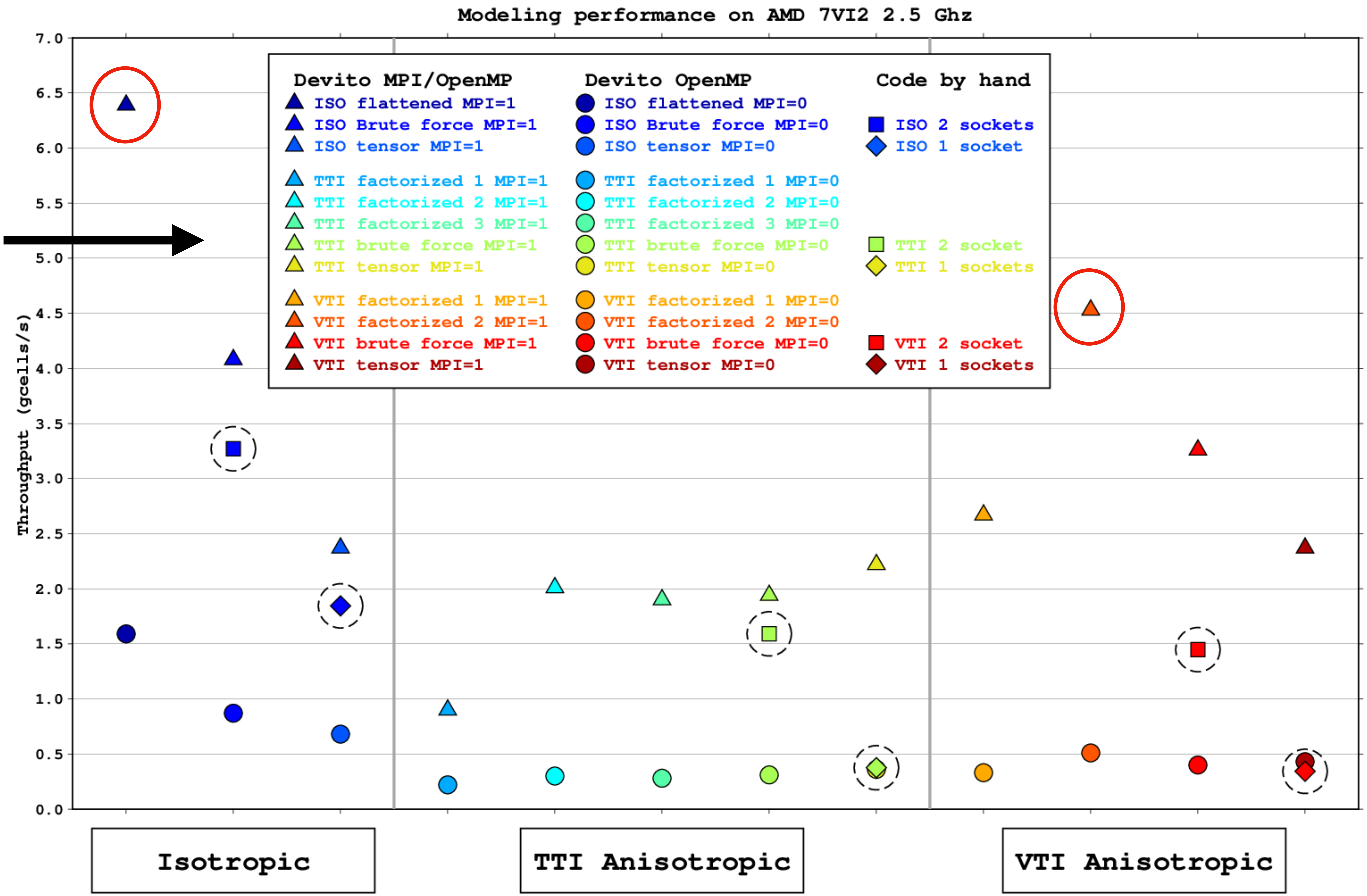
JUDI: [1]

- matrix-free linear operators and abstract data containers
- parallel I/O based on look-up tables
- data parallelism
- interface to Julia's machine learning package Flux
- *https://github.com/slimgroup/JUDI.jl/* (MIT license)

Devito: [2][3]

- a domain-specific language for finite-differences in Python
- code generation with automated performance optimizations
- model parallelism (MPI, multithreading)
- *https://github.com/devitocodes/devito* (MIT license)

State of the art performance



[1] P. A. Witte, M. Louboutin, F. Luporini, G. G. Gorman and Felix J. Herrmann, 2019, *Compressive least-squares migration with on-the-fly Fourier transforms* , Geophysics, 84, 5.

SLIM

Example 1: Compressive seismic imaging

Least-squares migration as an elastic net:[1]

$$\begin{aligned}
&\underset{\delta\mathbf{m}}{\text{minimize}} \quad \lambda \|\mathbf{C}\,\delta\mathbf{m}\|_1 + \frac{1}{2}\|\mathbf{C}\,\delta\mathbf{m}\|_2^2 \\
&\text{subject to:} \quad \sum_{i=1}^{n_s}\sum_{j=1}^{n_f} \left\|\mathbf{M}_l^{-1}\mathbf{J}\mathbf{M}_r^{-1}\,\delta\mathbf{m} - \mathbf{M}_l^{-1}\bar{\mathbf{d}}_{ij}\right\| \leq \sigma
\end{aligned}$$

Solve via the linearized Bregman method:

- at each iteration: random subsets of sources + frequencies
- memory per source: $\mathcal{O}(\bar{n}_f)$
- compressive sensing: no. of samples $\sim$ no. of grid points, non-zero entries

Example 1: Compressive seismic imaging

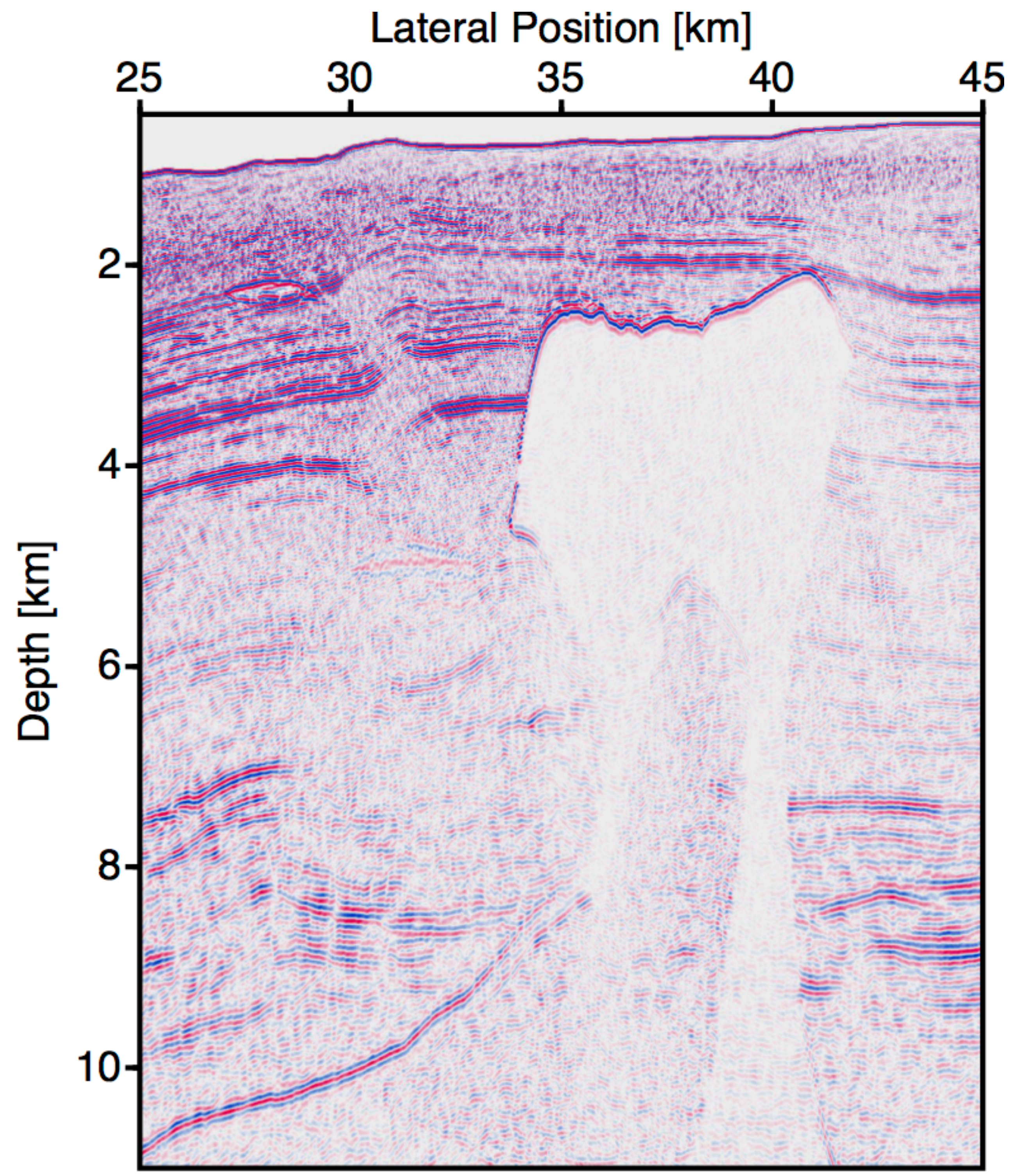
Linearized Bregman method with JUDI:

```
for j=1:maxiter

    # Compute residual and gradient
    i = randperm(d_obs.nsrc)[1:batchsize_source]
    select_frequencies!(J, batchsize_freq)
    r = Ml*J[i]*Mr*x - Ml*d_obs[i]
    g = Mr'*J[i]'*Ml'*proj_l2(r)

    # Residual and function value
    res[j] = norm(r, 2)
    fval[j] = λ*norm(C*z, 1) + .5f0*norm(C*z, 2)^2

    # Update variables
    global z -= α*g
    global x = C*soft_thresholding(C*z, λ)
end
```



Example 1: Compressive seismic imaging

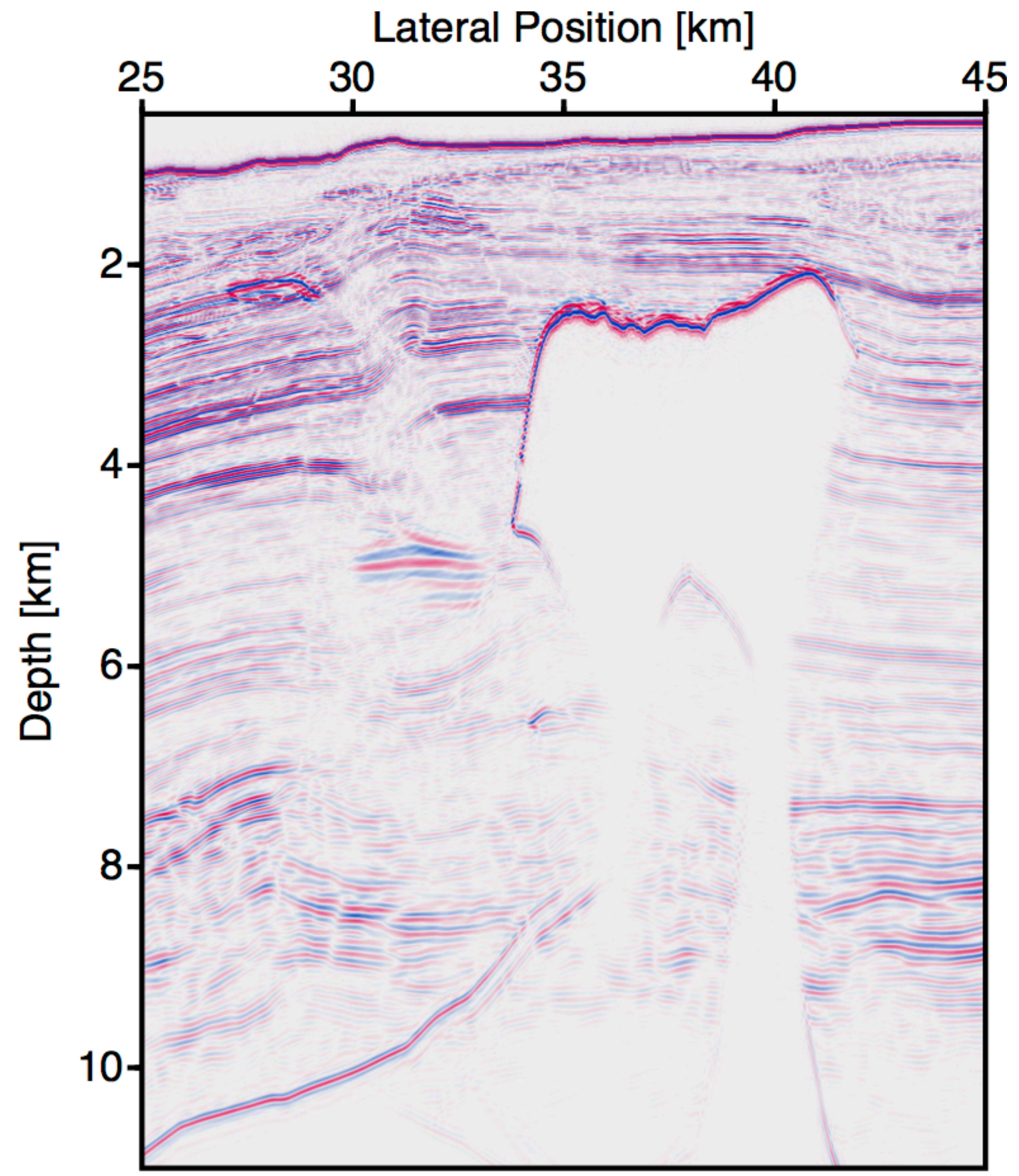
Linearized Bregman method with JUDI:

```
for j=1:maxiter

    # Compute residual and gradient
    i = randperm(d_obs.nsrc)[1:batchsize_source]
    select_frequencies!(J, batchsize_freq)
    r = Ml*J[i]*Mr*x - Ml*d_obs[i]
    g = Mr'*J[i]'*Ml'*proj_l2(r)

    # Residual and function value
    res[j] = norm(r, 2)
    fval[j] = λ*norm(C*z, 1) + .5f0*norm(C*z, 2)^2

    # Update variables
    global z -= α*g
    global x = C*soft_thresholding(C*z, λ)
end
```



Example 2: Deep Learning

JUDI operators with deep learning packages (e.g. *Flux.jl*):

```
network(x) = W2*(J*(W1*x + b1)) + b2
# Or:
network(x) = conv2(𝓕(conv1(x)))
loss(x, y) = Flux.mse(network(x), y)
```

Once again, abstractions pay off:

- No need to backpropagate through solvers using AD
- Backpropagation through JUDI operators

```
@adjoint *(J::judiJacobian, x) = *(J, x), Δ -> (nothing, transpose(J) * Δ)
```

Example 2: Deep Learning

Learned optimization schemes:[1]

- Unroll gradient descent loop
- Physics + data driven
- Parameterized by:

$$\theta = \{\mathbf{w}_1, \mathbf{b}_1, \mathbf{w}_2, \mathbf{b}_2, \mathbf{w}_3, \mathbf{b}_3\}$$

Training:

$$\underset{\theta}{\text{minimize}} \quad \sum_{i=1}^{n_{\text{train}}} \frac{1}{2} \|\Lambda_\theta(\mathbf{J}_i, \mathbf{d}_i) - \delta\mathbf{m}_i\|_2^2$$

function $\Lambda_\theta(\mathbf{J}, \mathbf{d})$
    $\mathbf{x}_0 = 0,\ \mathbf{s} = 0$
    for $j = 1, ..., n$
        $\mathbf{g} = \mathbf{J}^\top(\mathbf{J}\mathbf{x}_{j-1} - \mathbf{d})$
        $\mathbf{u} = [\mathbf{x}_{j-1}, \mathbf{g}, \mathbf{s}]$
        $\mathbf{u} = \mathrm{ReLU}\big(\mathbf{u} * \mathbf{w}_1 + \mathbf{b}_1\big)$
        $\mathbf{u} = \mathrm{ReLU}\big(\mathbf{u} * \mathbf{w}_2 + \mathbf{b}_2\big)$
        $\mathbf{u} = \mathbf{u} * \mathbf{w}_3 + \mathbf{b}_3$
        $\mathbf{s} = \mathrm{ReLU}\big(\mathbf{u}[:, 1\text{:end-}1, :, :]\big)$
        $\mathbf{\Delta x} = \mathbf{u}[:, \text{end}, :, :]$
        $\mathbf{x}_j = \mathbf{x}_{j-1} - \mathbf{\Delta x}$
    end
    return $\mathbf{x}_n$
end

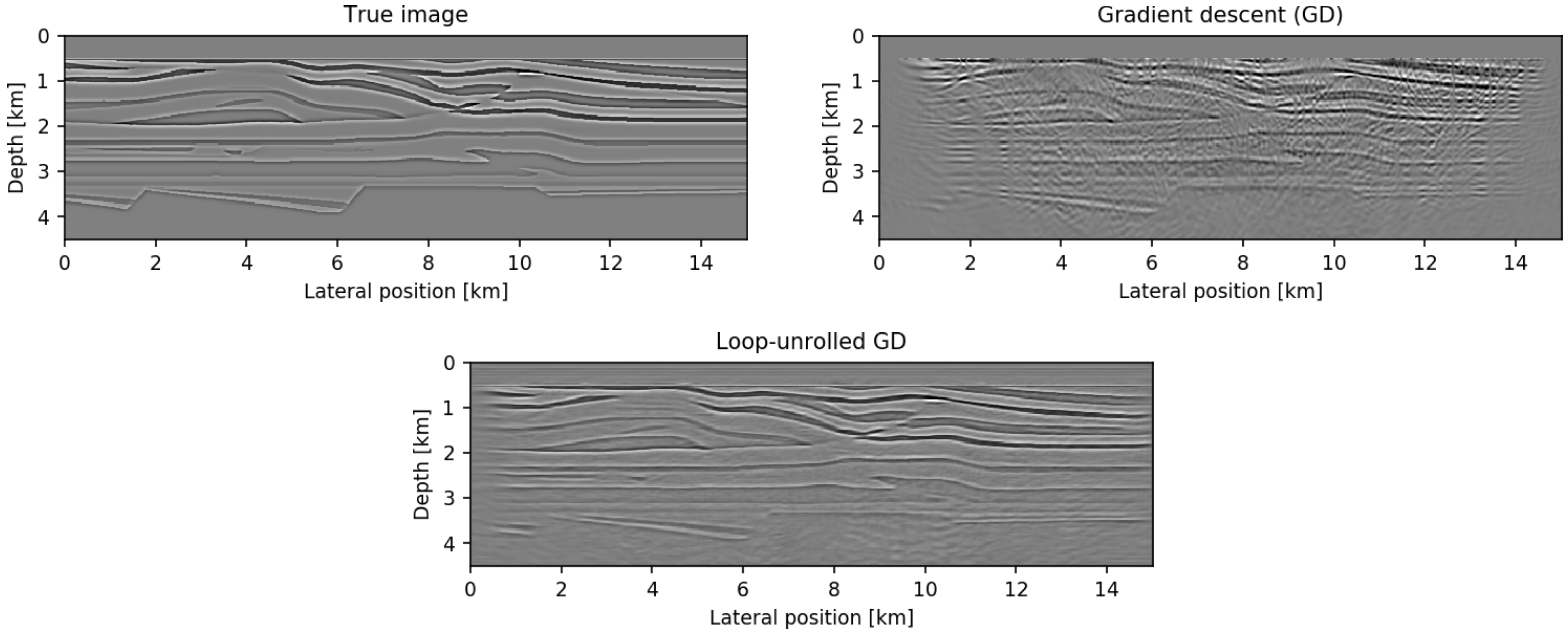
Loop unrolled GD

[1] Adler, Jonas, and Ozan Öktem, 2017, Solving ill-posed inverse problems using iterative deep neural networks, Inverse Problems 33, 12.

Example 2: Deep Learning

Training loop unrolled network:

- 1000 pairs of seismic data + true image
- Single epoch w/ ADAM



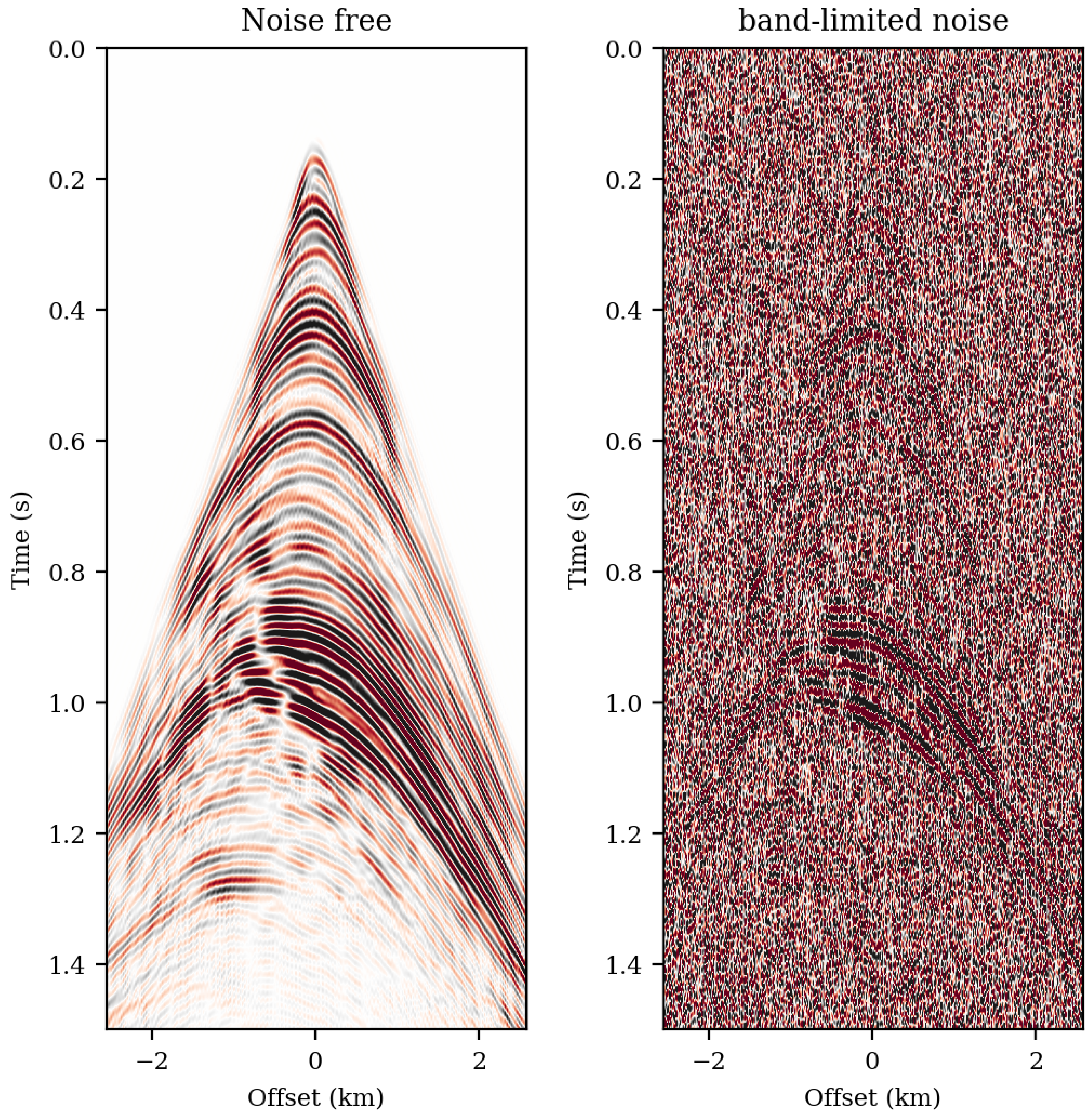
Observations

Abstracted software framework allows us to

- rapidly prototype new algorithms
- scale up
- integrate w/ optimization & ML

Q: H*ow can we assess risk in wave-equation based imaging/ inversion?*

- deep prior + MCMC
- conditional neural nets + normalizing flows



Single source noise-free (left) and noisy (right) linearized data

Linearized imaging

$$\min_{\delta\mathbf{m}} \quad \frac{1}{2\sigma^2}\sum_{i=1}^{n_s} \|\delta\mathbf{d}_i - \mathbf{J}(\mathbf{m}_0, \mathbf{q}_i)\delta\mathbf{m}\|_2^2$$

linearized Born operator, $\mathbf{J}(\mathbf{m}_0, \mathbf{q}_i)$

linearized data, $\delta\mathbf{d}_i$

noise variance, σ^2

Deep priors in seismic

imaging

full-waveform inversion

Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann. "A Deep-Learning Based Bayesian Approach to Seismic Imaging and Uncertainty Quantification". In: 2020.1 (2020), pp. 1–5. DOI: https://doi.org/10.3997/2214-4609.202010770. URL: https://arxiv.org/pdf/2001.04567.pdf.

Yulang Wu and George A McMechan. "Parametric convolutional neural network-domain full-waveform inversion". In: *GEOPHYSICS* 84.6 (2019), R881–R896. DOI: 10.1190/geo2018-0224.1.

Deep prior

$$\delta\mathbf{m} = g(\mathbf{z}, \mathbf{w}), \quad \mathbf{w} \sim \mathrm{N}(\mathbf{0}, \lambda^{-2}\mathbf{I})$$

untrained CNN, $g(\mathbf{z}, \mathbf{w})$

CNN weights, $\mathbf{w}$

fixed input, $\mathbf{z}$

V. Lempitsky, A. Vedaldi, and D. Ulyanov. "Deep Image Prior". In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 2018, pp. 9446–9454. DOI: 10.1109/CVPR.2018.00984.

Negative-log posterior

$$-\log \pi_{\mathbf{w}|\delta\mathbf{d}}\left(\mathbf{w} \mid \{\delta\mathbf{d}_i\}_{i=1}^{n_s}\right)$$

$$= \underbrace{\frac{1}{2\sigma^2}\sum_{i=1}^{n_s}\|\delta\mathbf{d}_i - \mathbf{J}(\mathbf{m}_0, \mathbf{q}_i)g(\mathbf{z}, \mathbf{w})\|_2^2}_{\text{negative-log likelihood}} + \underbrace{\frac{\lambda^2}{2}\|\mathbf{w}\|_2^2}_{\text{negative-log prior}} + \underbrace{\text{const}}_{\text{Ind. of } \mathbf{w}}$$

SGLD—an stochastic-approximation based MCMC sampling approach

$$
\begin{aligned}
\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\epsilon}{2}\mathbf{M}\nabla_{\mathbf{w}} \left[n_s \log \pi_{\text{like}} \left(\delta \mathbf{d}_i \mid \mathbf{w}_k \right) \right. \\
\left. + \log \pi_{\text{prior}} \left(\mathbf{w}_k \right) \right] + \boldsymbol{\eta}_k, \quad \boldsymbol{\eta}_k \sim \mathrm{N}(\mathbf{0}, \epsilon \mathbf{M}),
\end{aligned}
$$

Max Welling and Yee Whye Teh. "Bayesian Learning via Stochastic Gradient Langevin Dynamics". In: *Proceedings of the 28th International Conference on International Conference on Machine Learning*. ICML'11. 2011, pp. 681–688. DOI: 10.5555/3104482.3104568.

Chunyuan Li et al. "Preconditioned Stochastic Gradient Langevin Dynamics for Deep Neural Networks". In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. AAAI'16. 2016, pp. 1788–1794. DOI: 10.5555/3016100.3016149.

Posterior distribution, $\pi_{\mathbf{w}|\delta\mathbf{d}}$

MAP, conditional mean, pointwise standard deviation (UQ), ...

MAP estimate, $\widehat{\delta \mathbf{m}}_{\mathsf{MAP}} = g(\mathbf{z}, \widehat{\mathbf{w}}_{\mathsf{MAP}})$

$$\widehat{\mathbf{w}}_{\mathsf{MAP}} := \underset{\mathbf{w}}{\arg\max} \;\; \pi_{\mathbf{w}|\delta\mathbf{d}} \left(\mathbf{w} \mid \{\delta \mathbf{d}_i\}_{i=1}^{n_s}\right)$$

conditional (posterior) mean, $\widehat{\delta \mathbf{m}}_{\mathrm{CM}}$

$$\widehat{\delta \mathbf{m}}_{\mathrm{CM}} := \mathbb{E}_{\mathbf{w} \sim \pi_{\mathbf{w} \mid \delta \mathbf{d}}} \left[g(\mathbf{z}, \mathbf{w}) \right] \approx \frac{1}{n_{\mathrm{w}}} \sum_{j=1}^{n_{\mathrm{w}}} g(\mathbf{z}, \widehat{\mathbf{w}}_j),$$

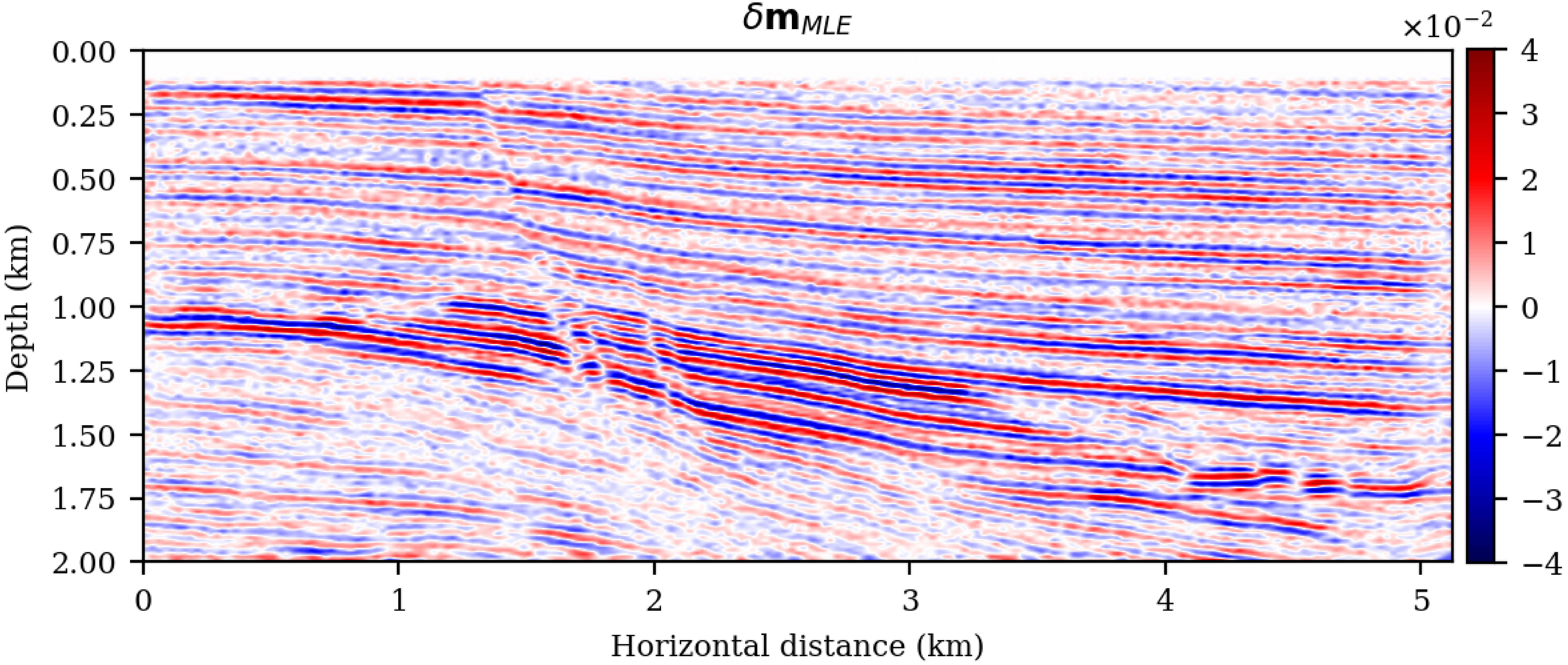
samples from posterior, $\left\{ \widehat{\mathbf{w}}_j \right\}_{j=1}^{n_{\mathrm{w}}} \sim \pi_{\mathbf{w} \mid \delta \mathbf{d}} \left(\mathbf{w} \mid \{ \delta \mathbf{d}_i \}_{i=1}^{n_s} \right)$

pointwise standard deviation (UQ)

$$\widehat{\boldsymbol{\sigma}}^2 := \frac{1}{n_{\mathrm{w}} - 1} \sum_{j=1}^{n_{\mathrm{w}}} (g(\mathbf{z}, \widehat{\mathbf{w}}_j) - \delta\widehat{\mathbf{m}}_{\mathsf{CM}}) \odot (g(\mathbf{z}, \widehat{\mathbf{w}}_j) - \delta\widehat{\mathbf{m}}_{\mathsf{CM}}),$$

samples from posterior, $\left\{\widehat{\mathbf{w}}_j\right\}_{j=1}^{n_{\mathrm{w}}} \sim \pi_{\mathbf{w}|\delta\mathbf{d}}(\mathbf{w} \mid \{\delta\mathbf{d}_i\}_{i=1}^{n_s})$

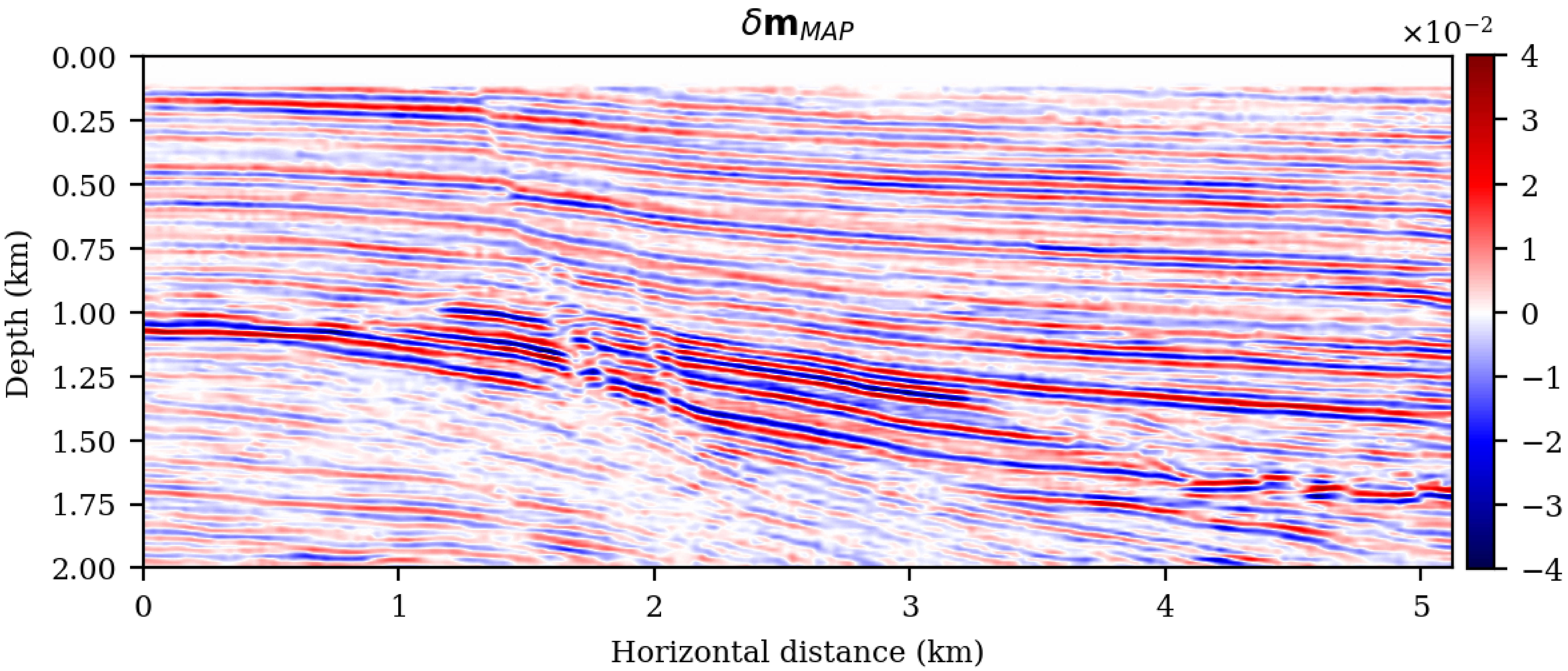
Maximum Likelihood estimate



MLE—no (deep) prior

Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann, "Weak deep priors for seismic imaging"

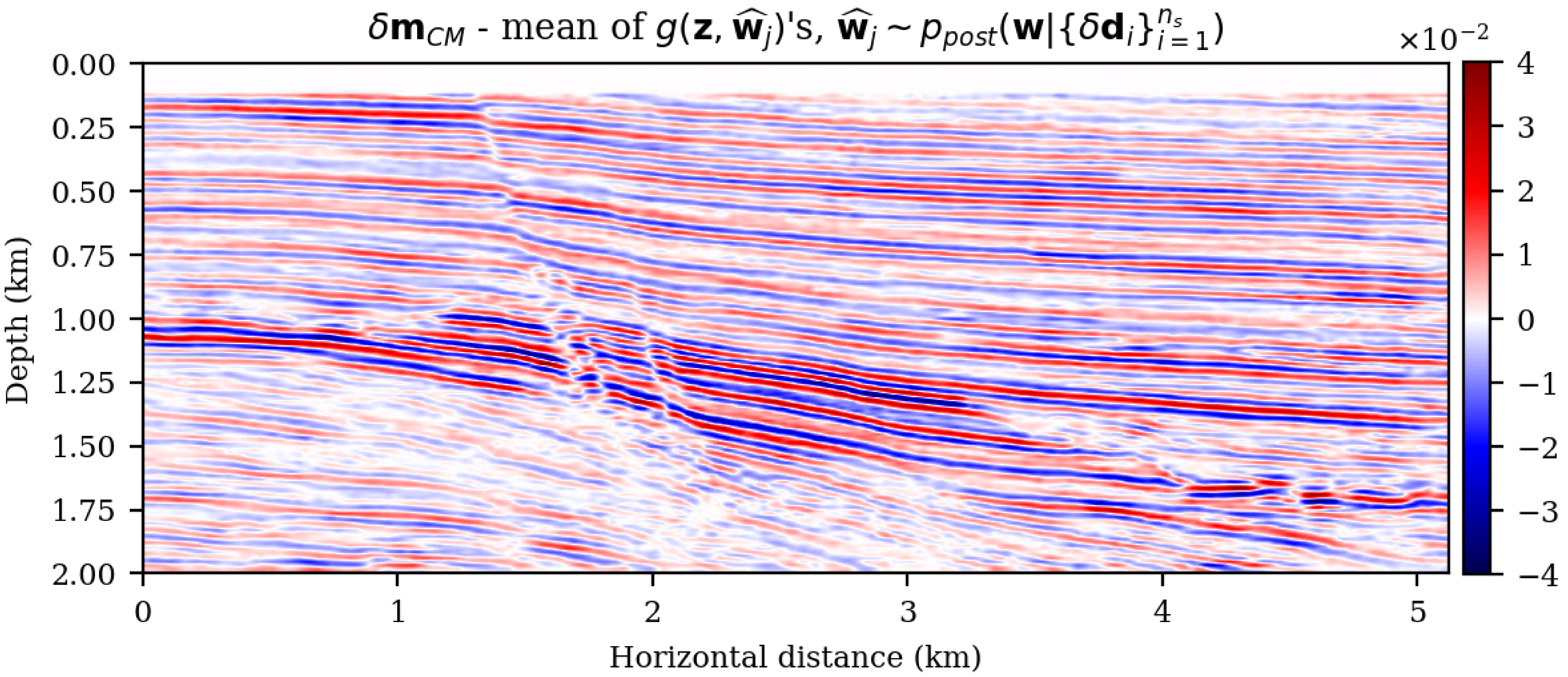
MAP estimate



MAP estimate—deep-prior based result with early stopping

Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann, "**Weak deep priors for seismic imaging**"

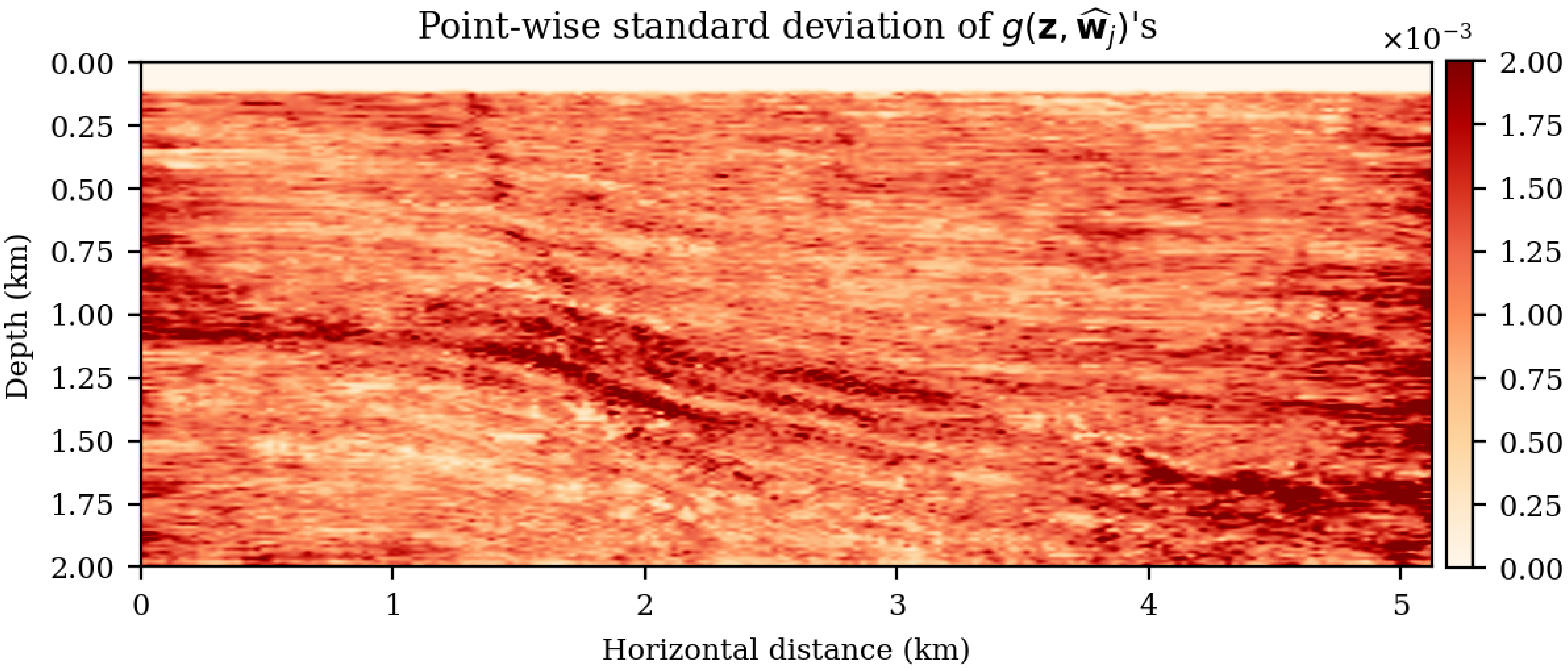
Conditional mean estimate



Conditional mean estimate — more robust to overfitting

Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann, "Weak deep priors for seismic imaging"

Standard deviation estimate

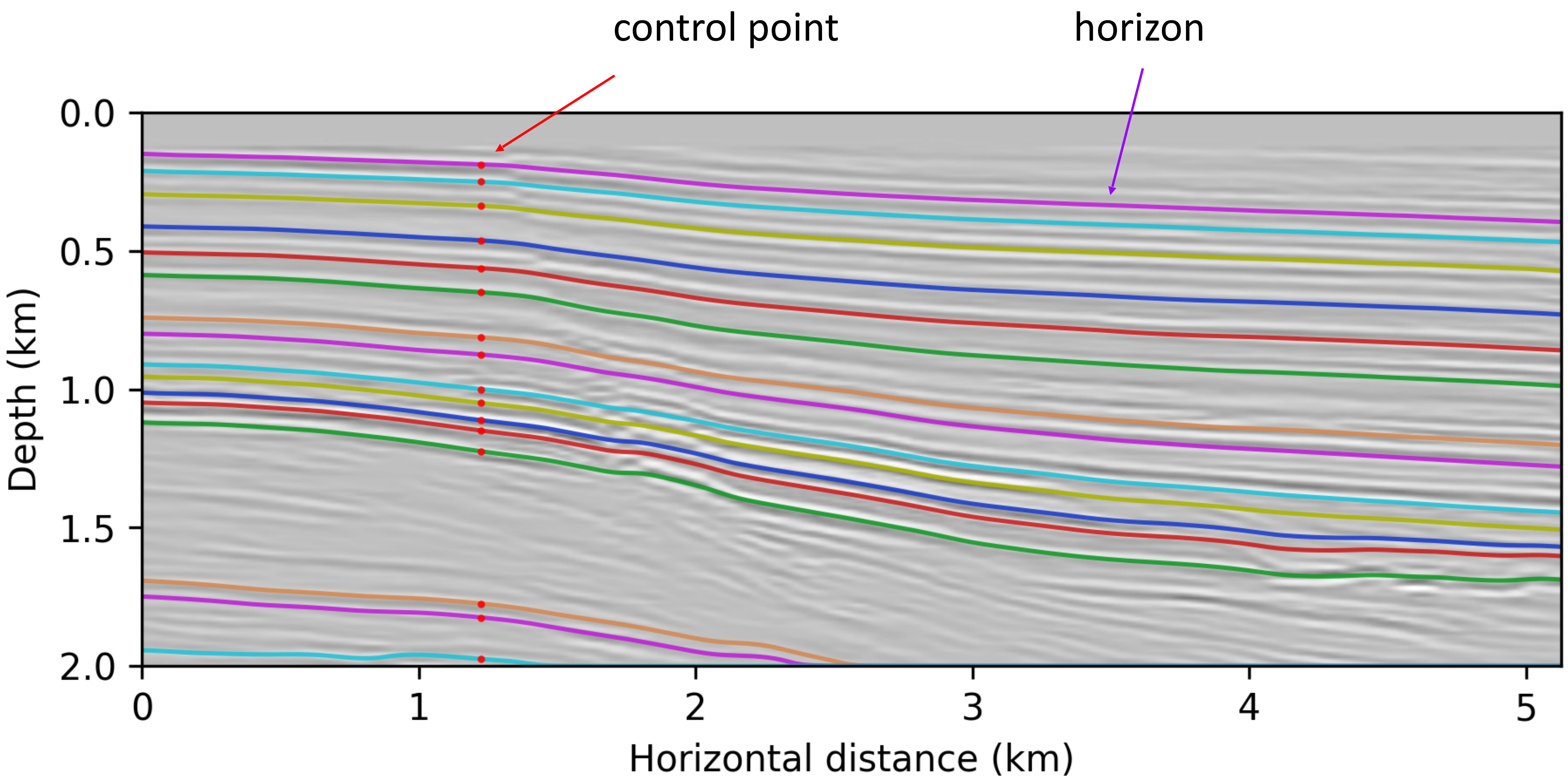


Imaging UQ

Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann, “Weak deep priors for seismic imaging”

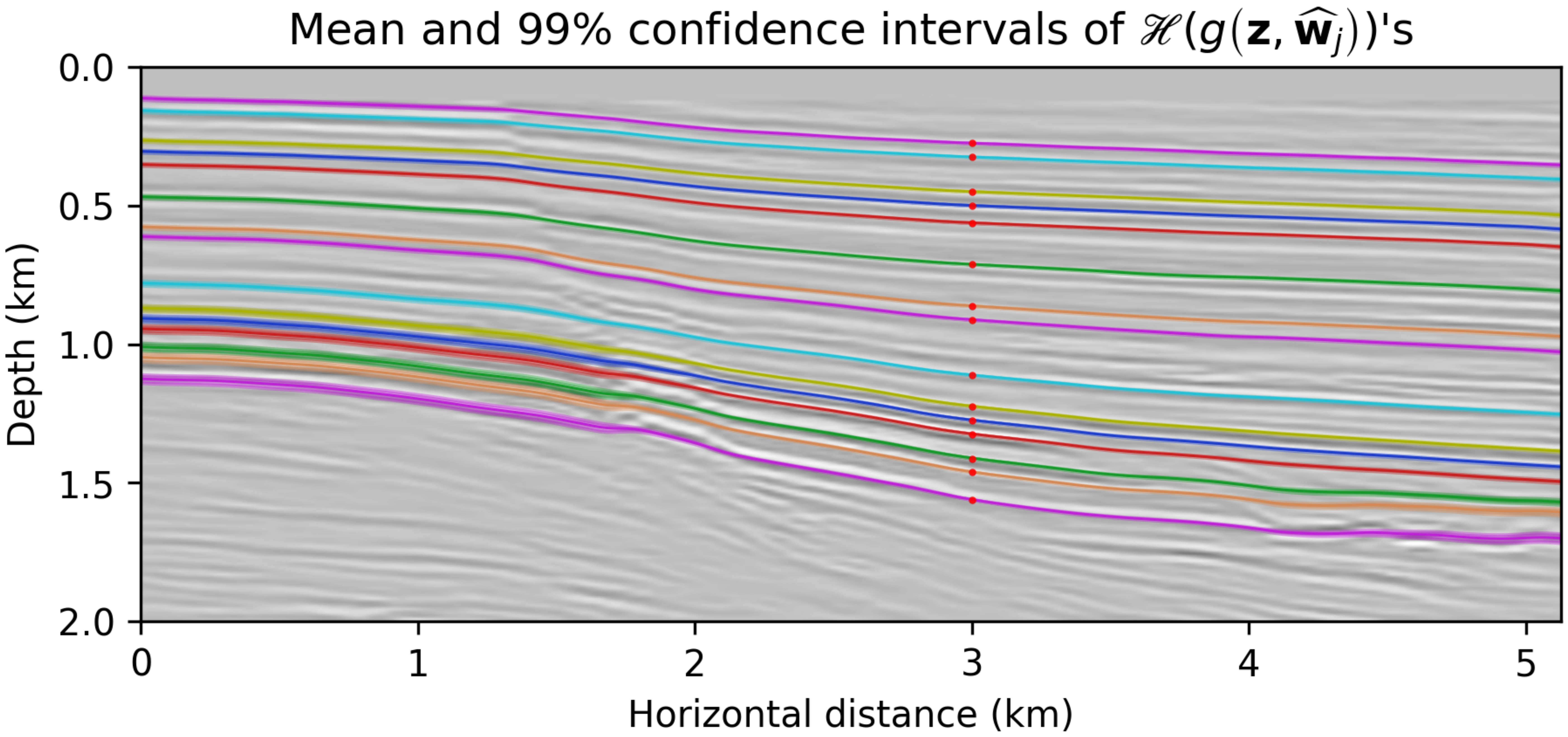
Horizon tracking

Wu, Xinming, and Sergey Fomel. "Least-squares horizons with local slopes and multigrid correlations." Geophysics 83.4 (2018): IM29-IM40.



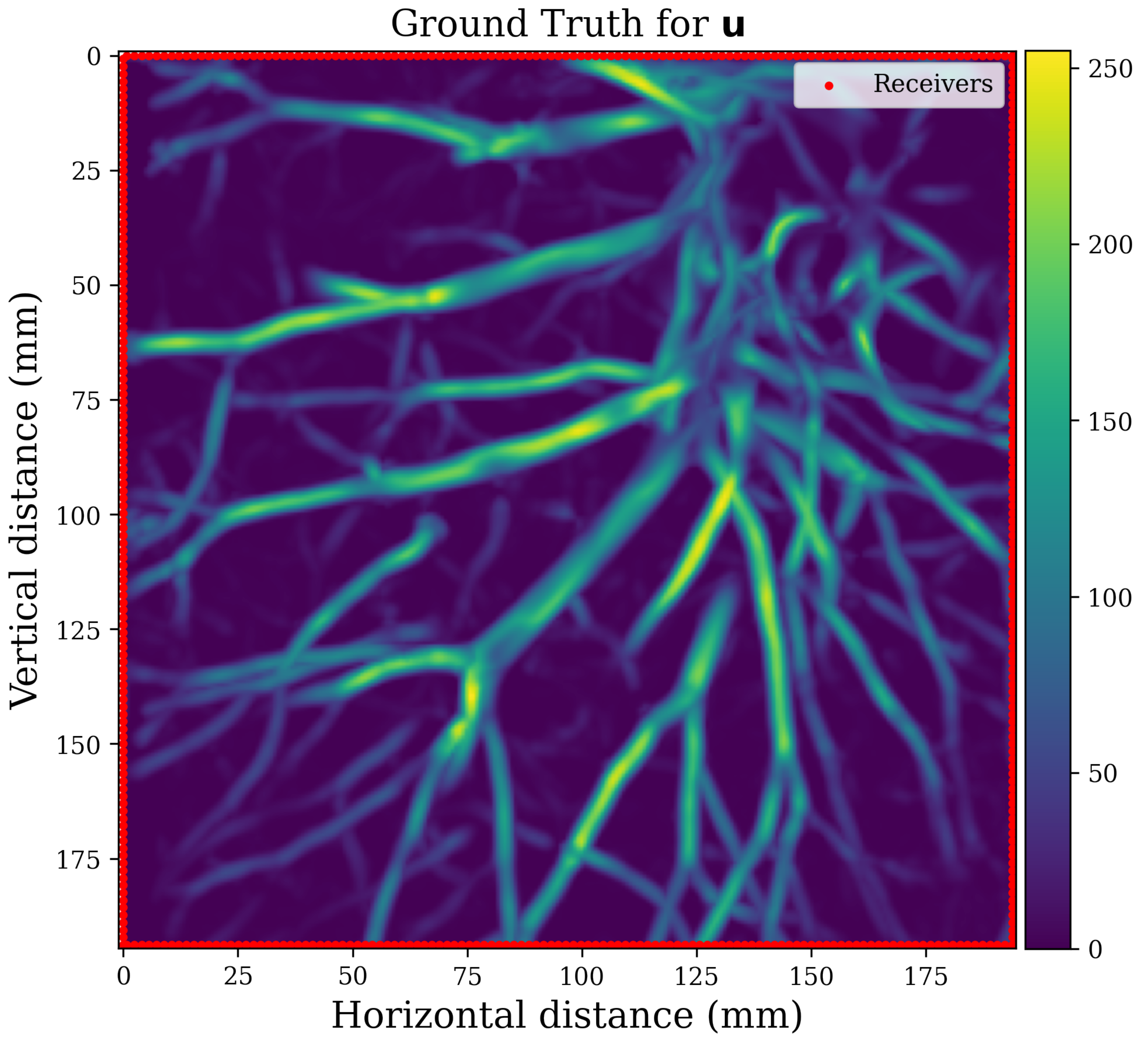
Horizon tracking w/ UQ

Ali Siahkoohi, Gabrio Rizzuti, and Felix J. Herrmann, "**Uncertainty quantification in imaging and automatic horizon tracking—a Bayesian deep-prior based approach**"

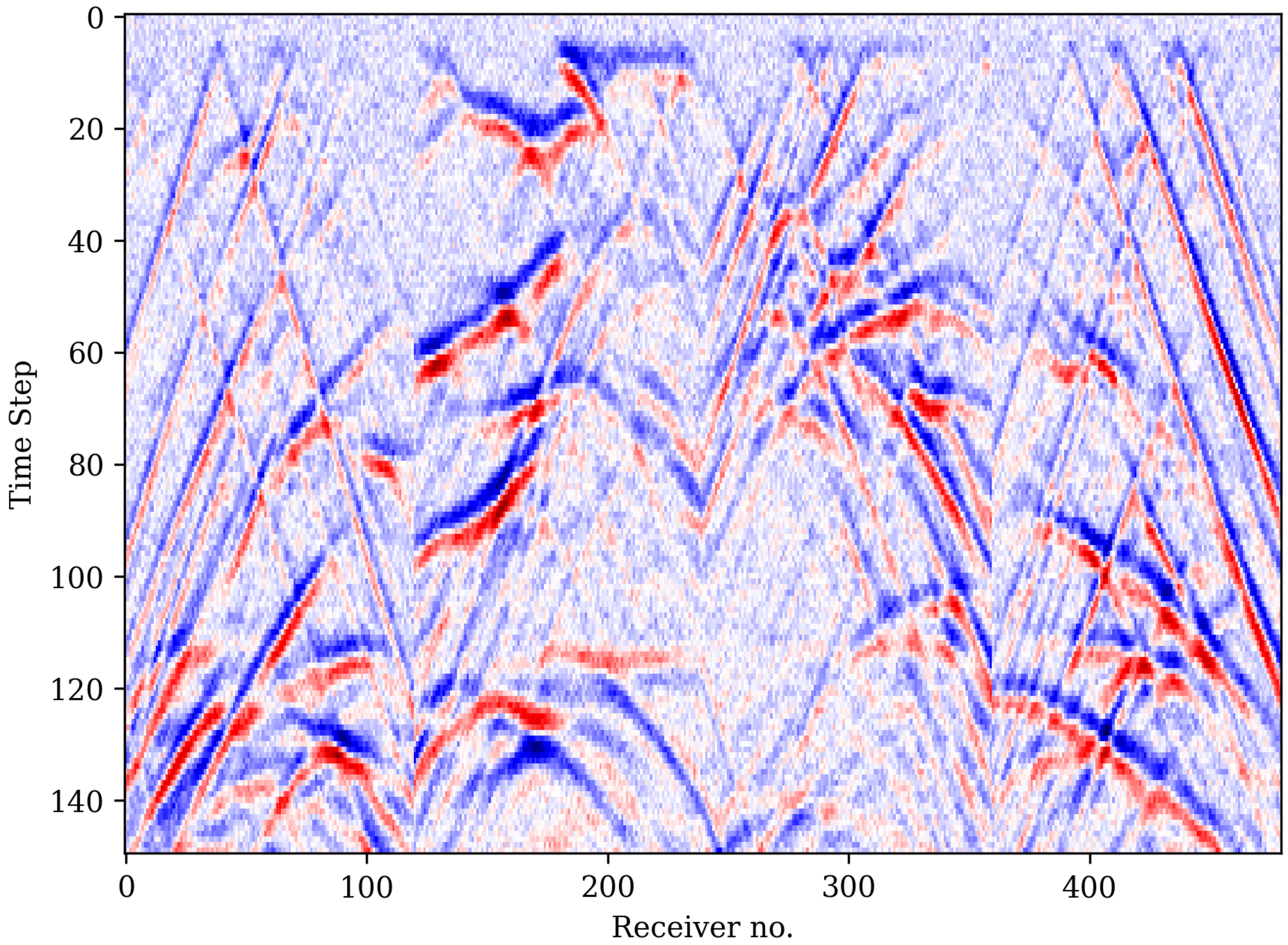


[1] Reeves, Anthony P., et al. "A public image database to support research in computer aided diagnosis." 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2009.

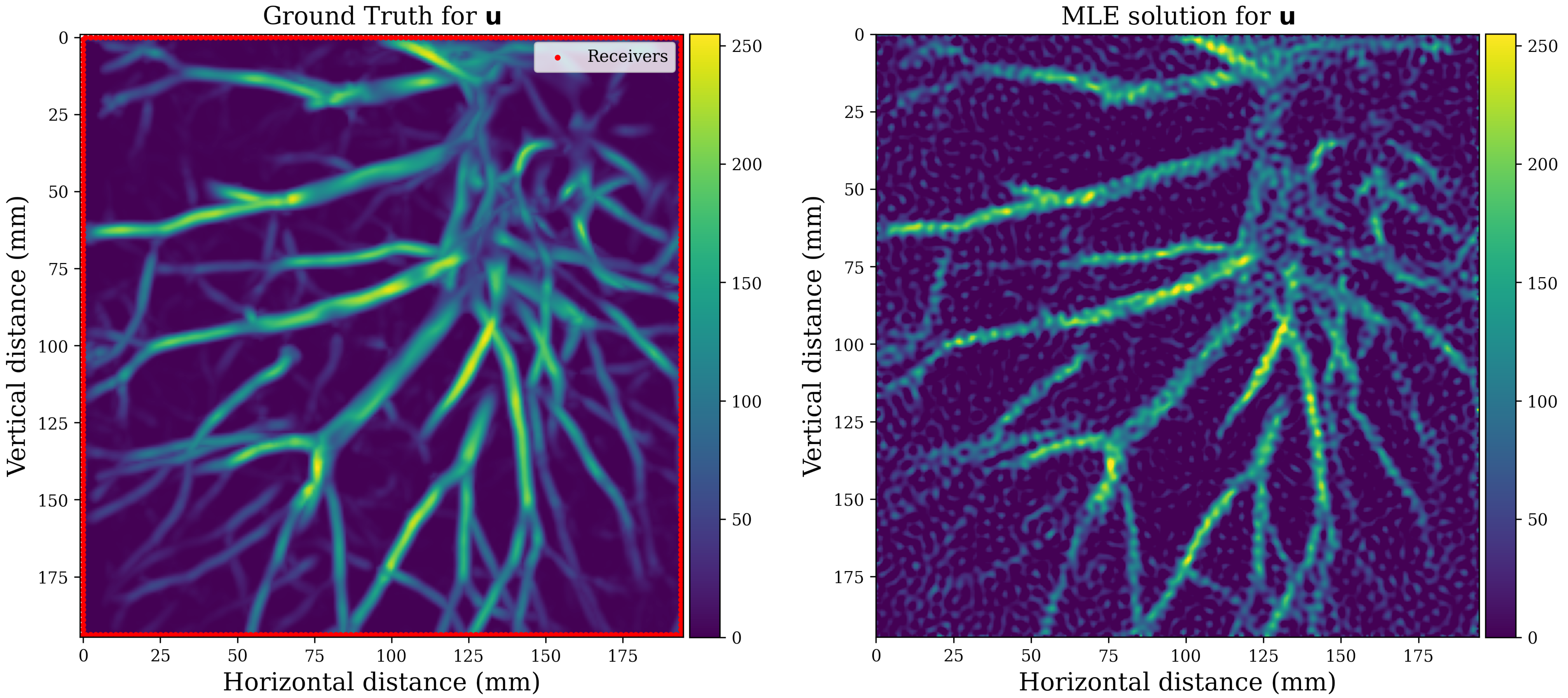
Photoacoustic imaging



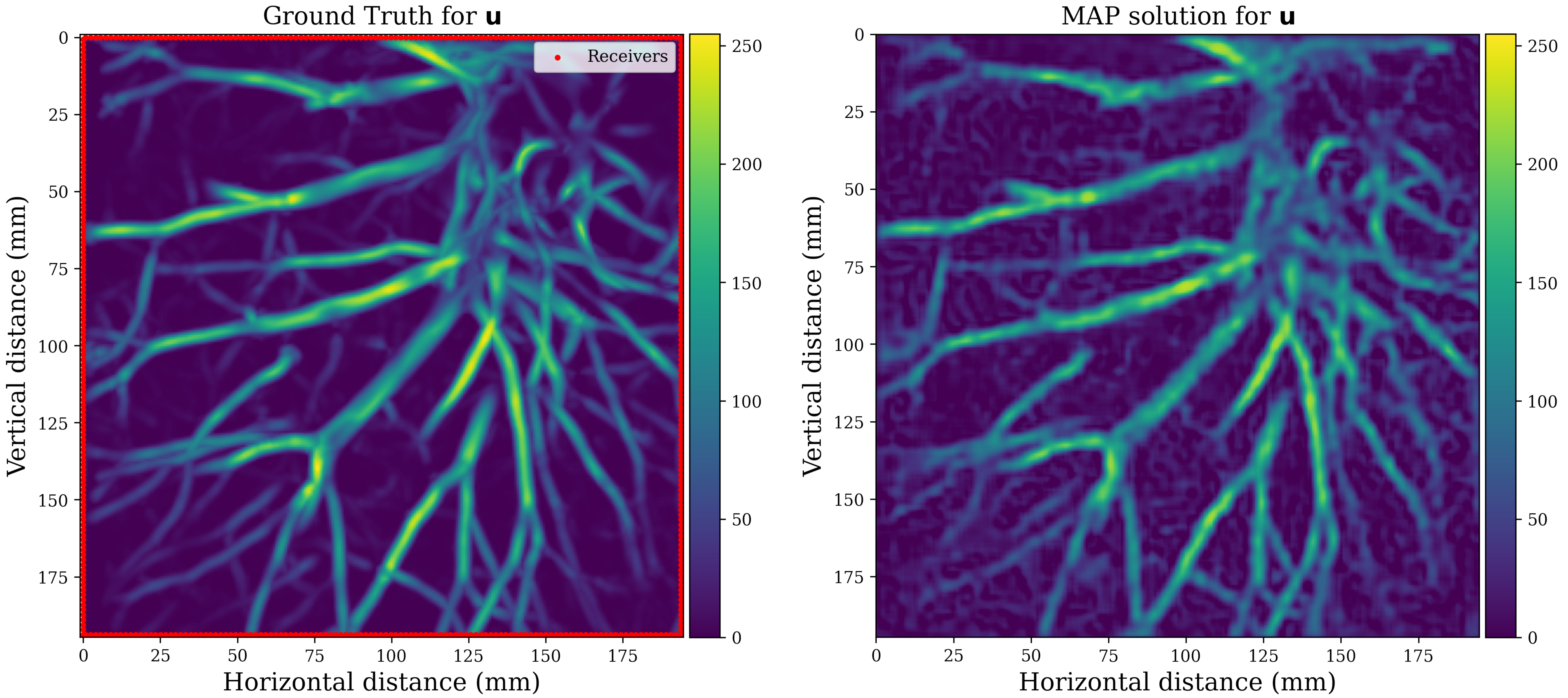
Noisy observed data



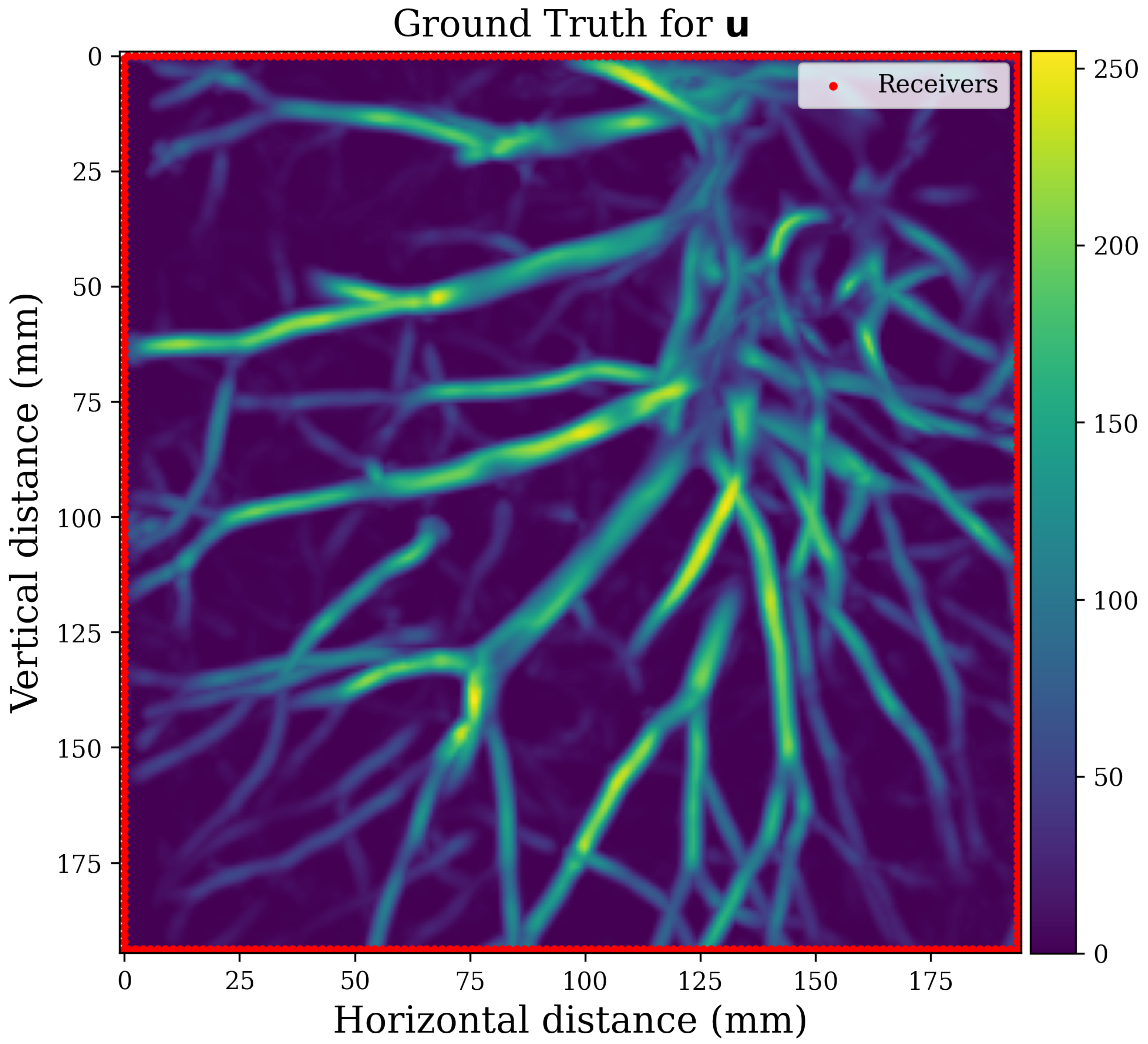
Maximum likelihood estimate

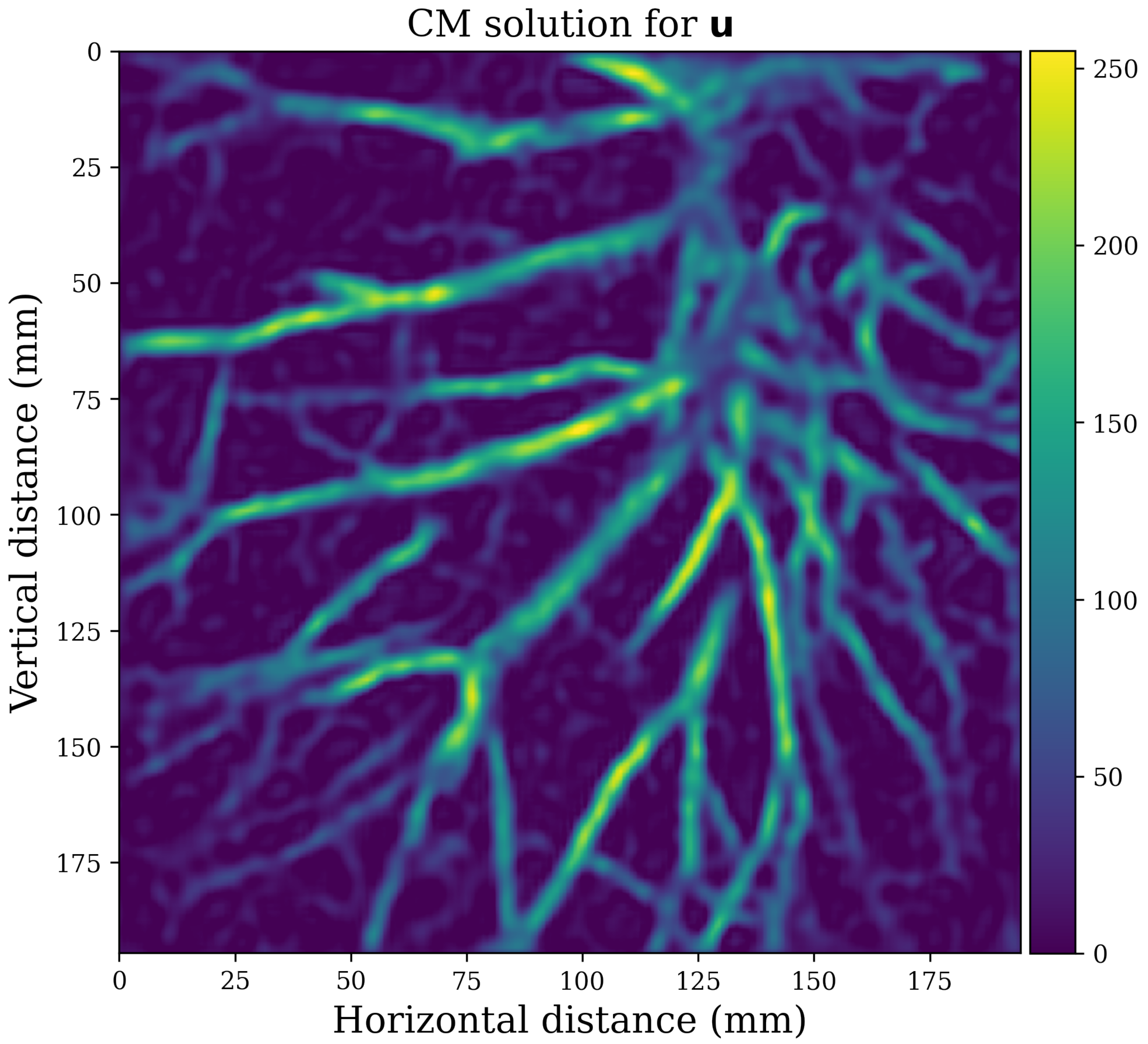


Map estimate

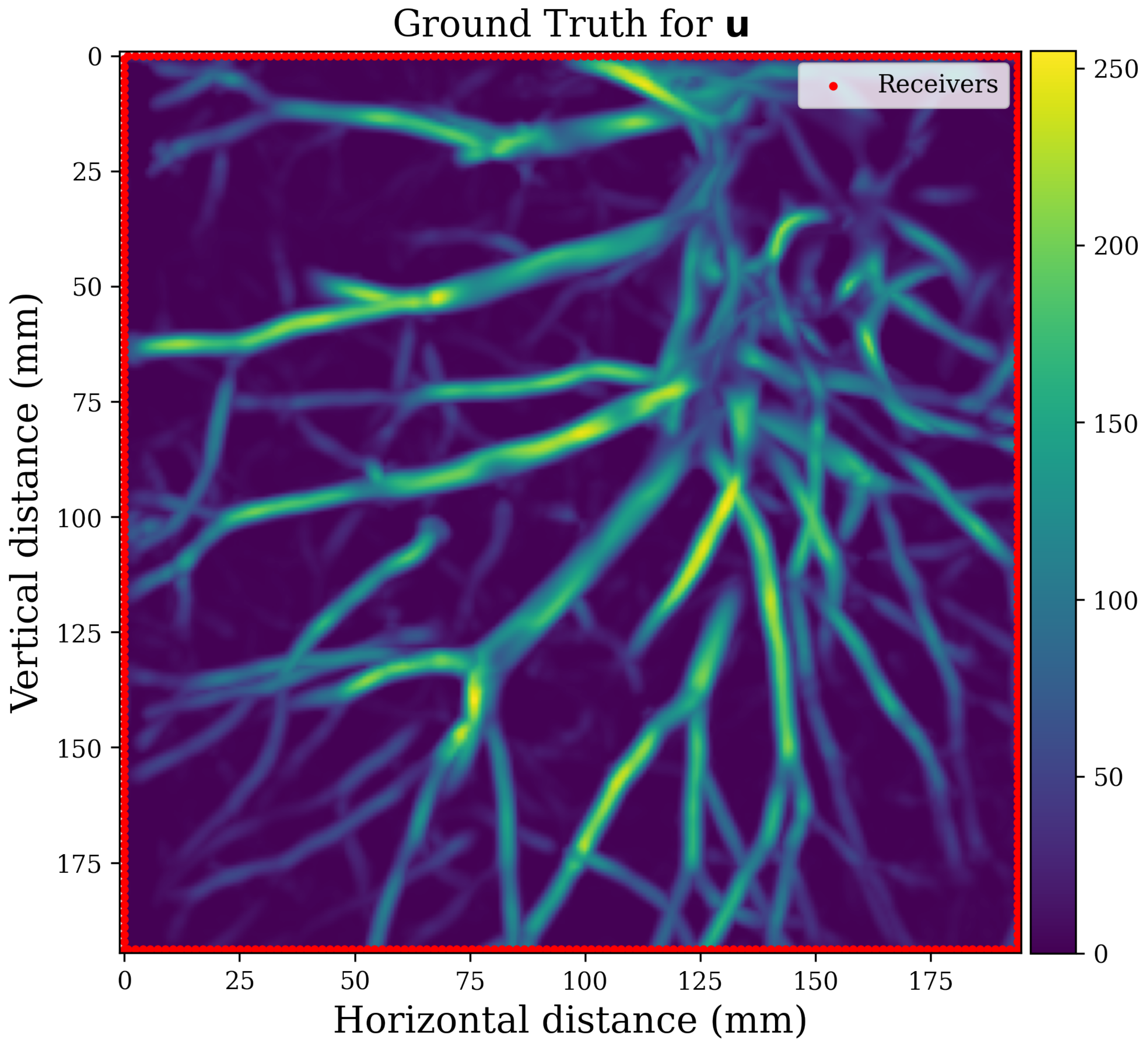


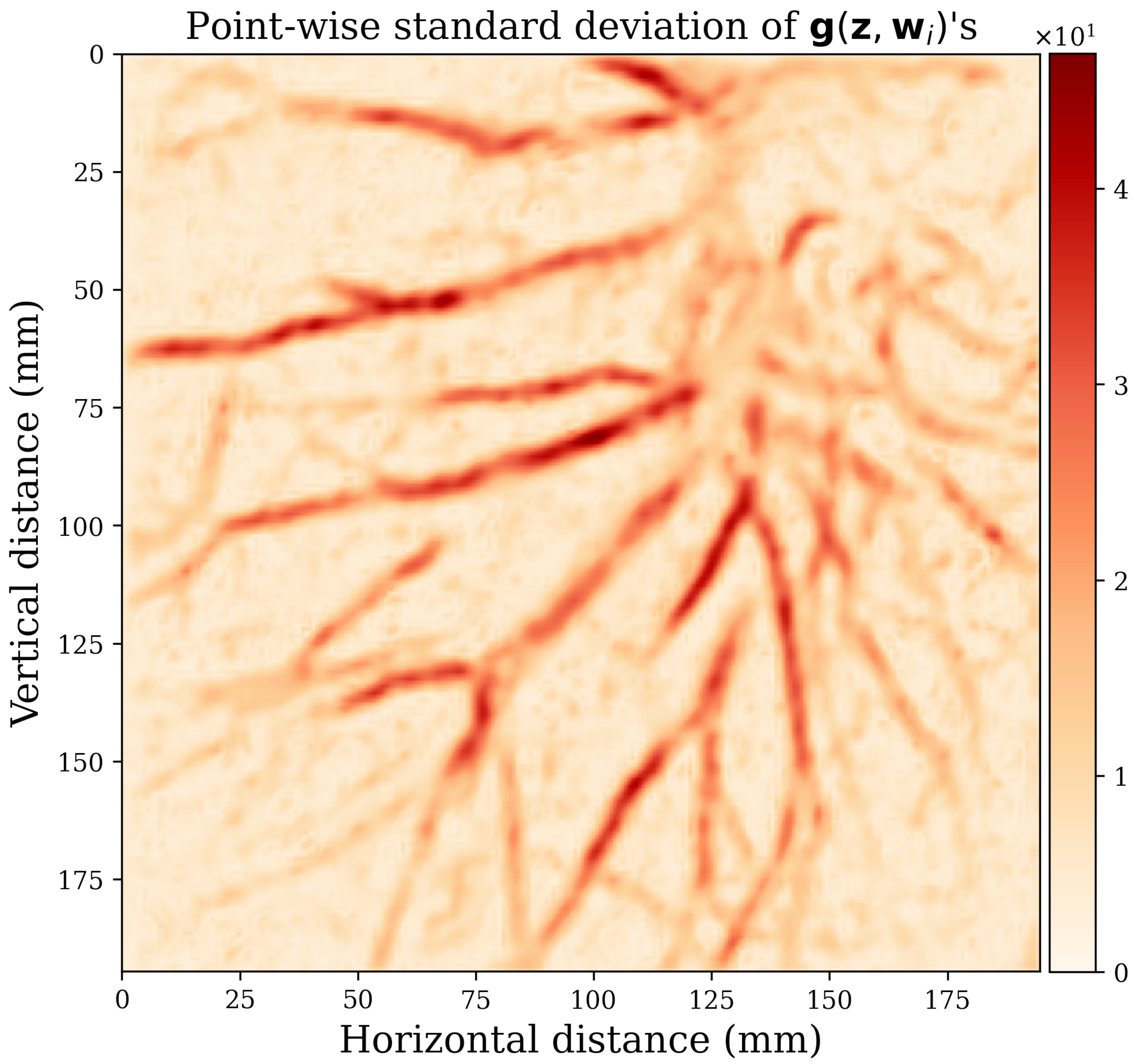
Conditional mean estimate





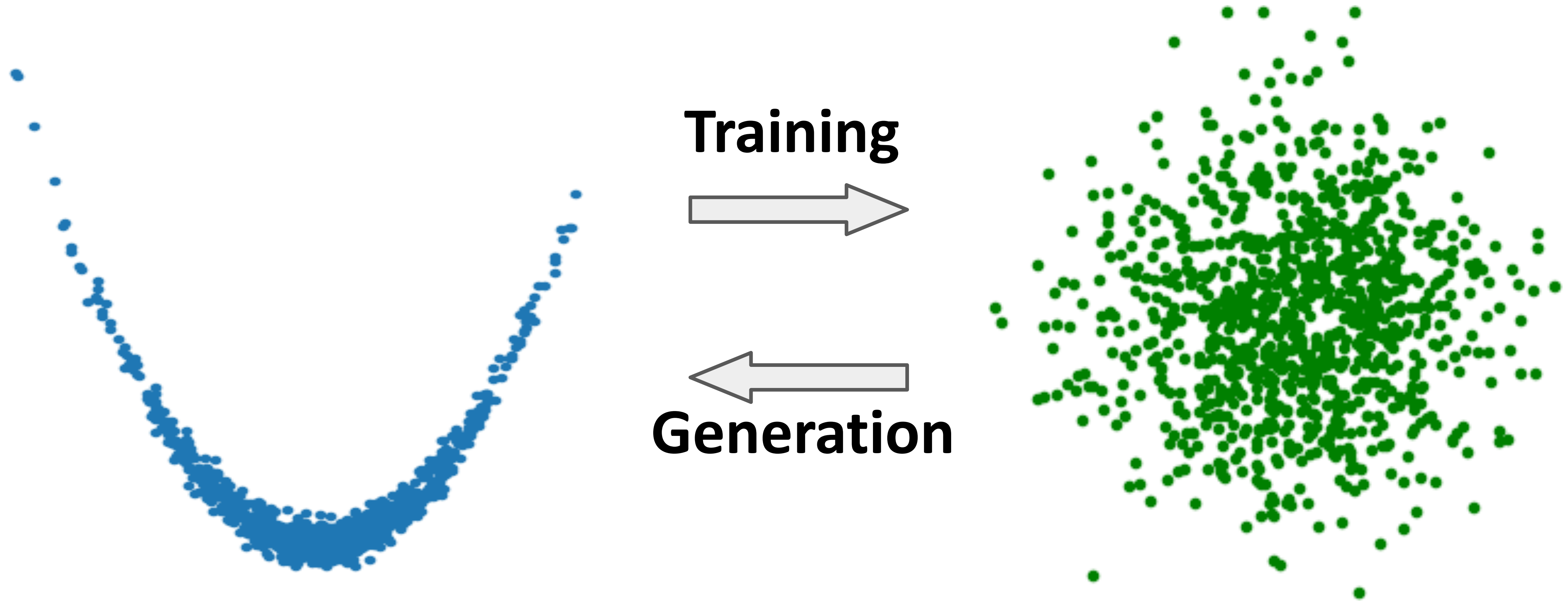
Standard deviation estimate





Ivan Kobyzev, Simon J.D. Prince, and Marcus A. Brubaker. “Normalizing Flows: An Introduction and Review of Current Methods”, arXiv:1908.09257
Gabrio Rizzuti, Ali Siahkoohi, Philipp A. Witte, and Felix J. Herrmann, “Parameterizing uncertainty by deep invertible networks, an application to reservoir characterization”
Ali Siahkoohi, Gabrio Rizzuti, Philipp A. Witte, and Felix J. Herrmann, “Faster Uncertainty Quantification for Inverse Problems with Conditional Normalizing Flows”

Normalizing flows



$\mathbf{x} \sim p_X(\mathbf{x})$ $\qquad$ $\mathbf{z} \sim p_Z(\mathbf{z})$

$$G_\theta : \mathcal{X} \to \mathcal{Z}$$

Jakob Kruse, Gianluca Detommaso, Robert Scheichl, and Ullrich Köthe. “HINT: Hierarchical Invertible Neural Transport for Density Estimation and Bayesian Inference”. arXiv:1905.10687

SLIM

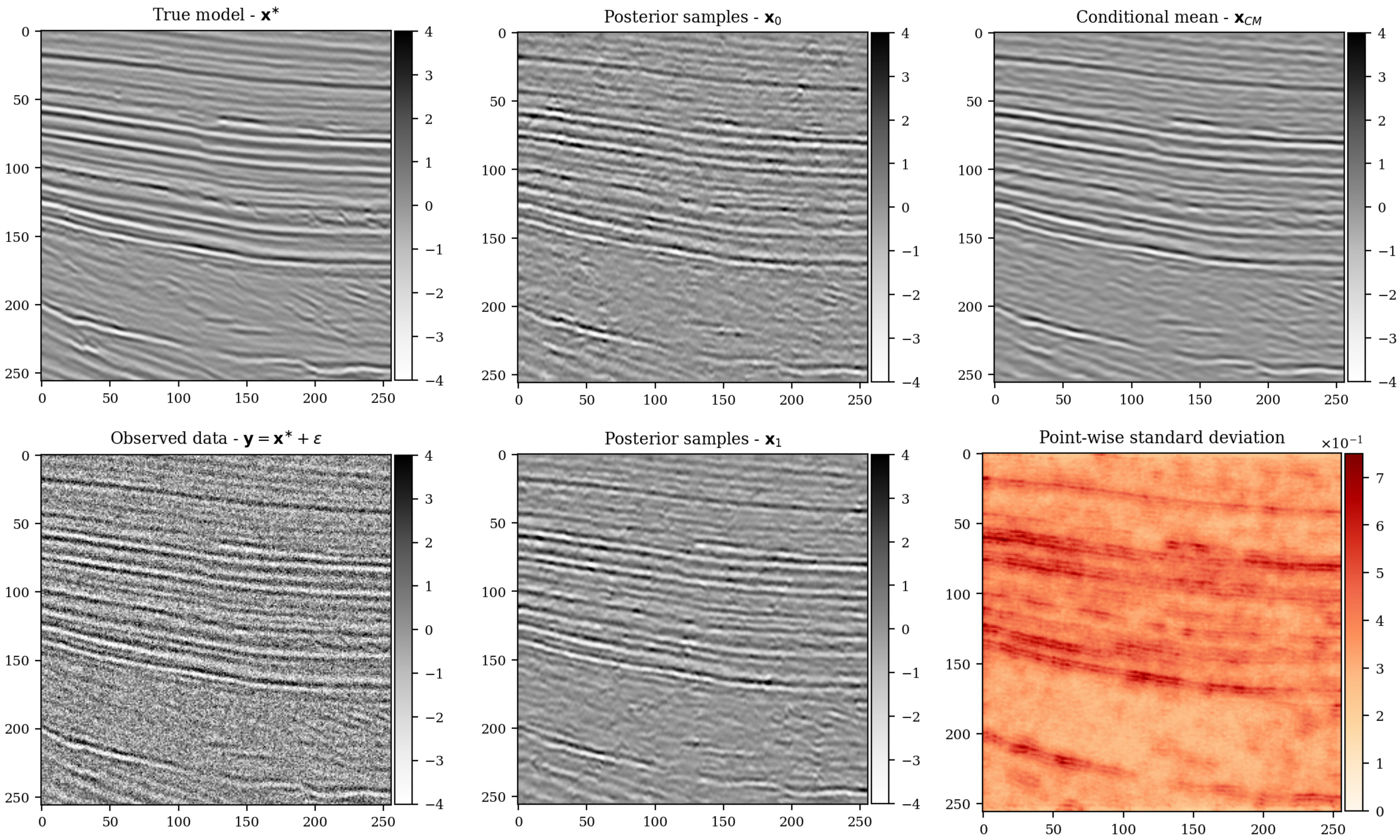
Supervised imaging and UQ

Given joint pairs (**m**, **d**),

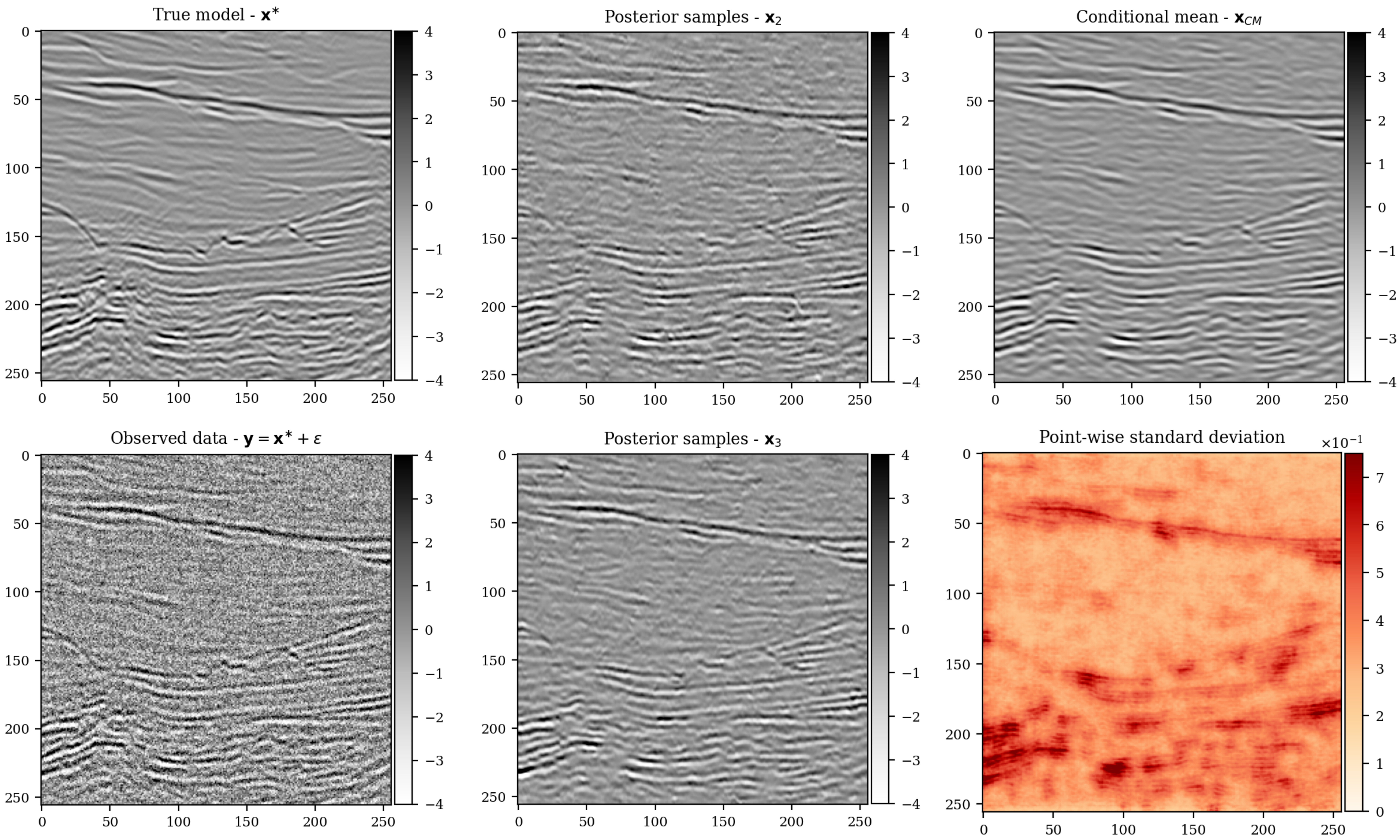
$$\begin{aligned}\min_\theta \quad & \mathrm{KL}(p_{\mathrm{joint}}(\mathbf{m},\mathbf{d})||p_\theta(\mathbf{m},\mathbf{d})) \\ & = \mathbb{E}_{\mathbf{m},\mathbf{d}\sim p_{\mathrm{joint}}(\mathbf{m},\mathbf{d})}\frac{1}{2}||G_\theta(\mathbf{m},\mathbf{d})||^2 - \log|\det J_{G_\theta}(\mathbf{m},\mathbf{d})|\end{aligned}$$

+ special **conditional structure** on G

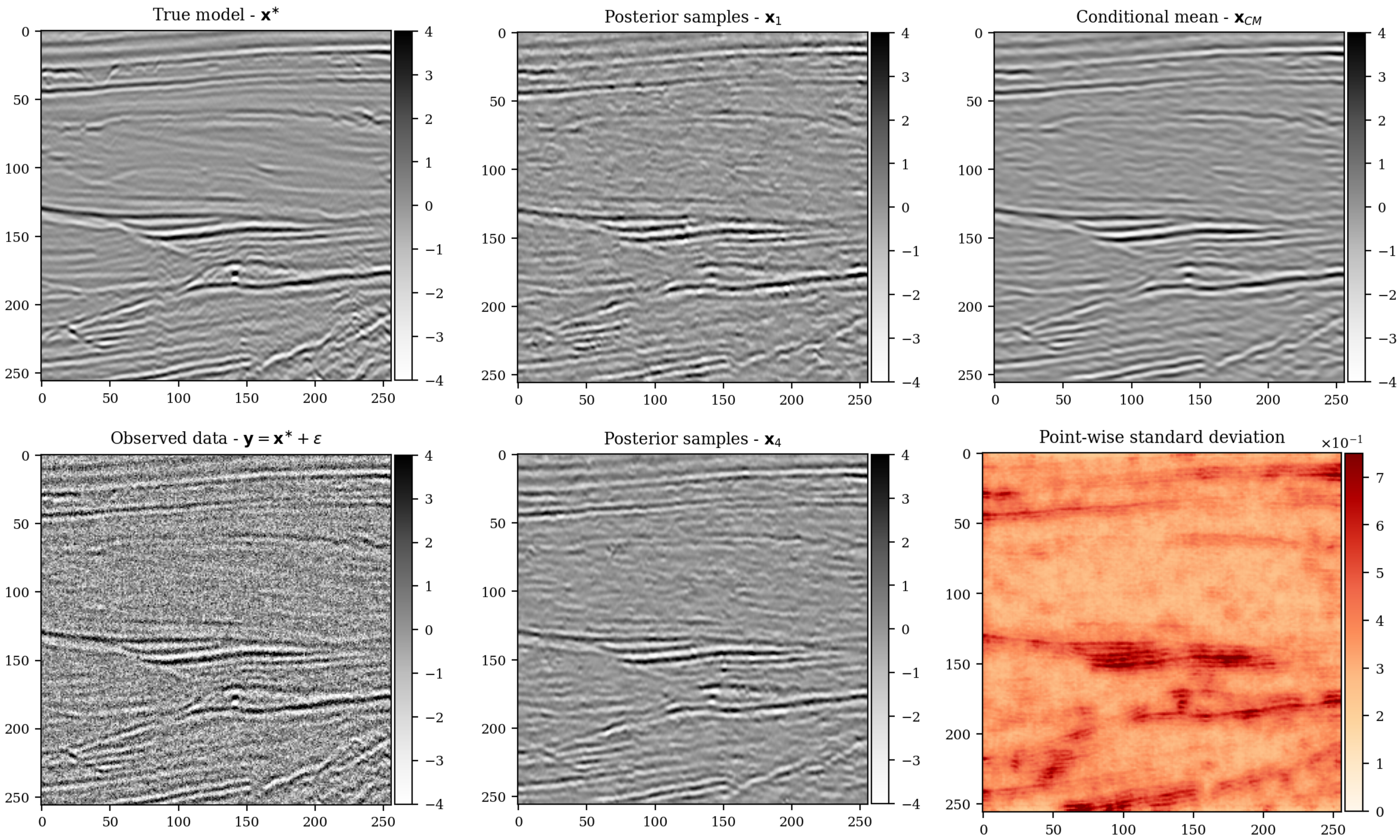
Realization 1



Realization 2



Realization 3



Observations

Medical & seismic imaging are challenged by

- correct wave physics (spatially varying acoustic/elastic properties)
- scale up to 3D (required compute + feasible return times)
- training (compute + access training data)

Use of unifying framework + abstractions

- facilitate rapid innovations
- manage complexity

Quantitative assessment of risk remains elusive…