Solving geophysical inverse problems with scientific machine learning

Ziyi Yin

CSE PhD dissertation defense
June 25, 2024

Committee members

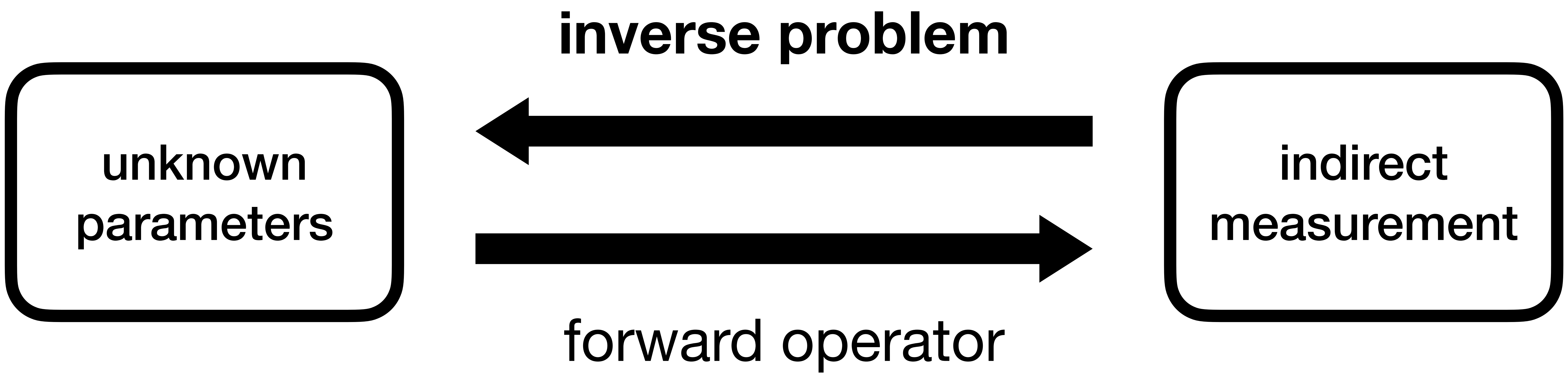
Dr. Felix J. Herrmann, advisor, School of CSE, ECE, EAS
Dr. Nisha Chandramoorthy, School of CSE
Dr. Peng Chen, School of CSE
Dr. J. Carlos Santamarina, School of CEE
Dr. Lars Ruthotto, Emory University
Dr. Olav Møyner, SINTEF Digital

SLIM
Georgia Institute of Technology

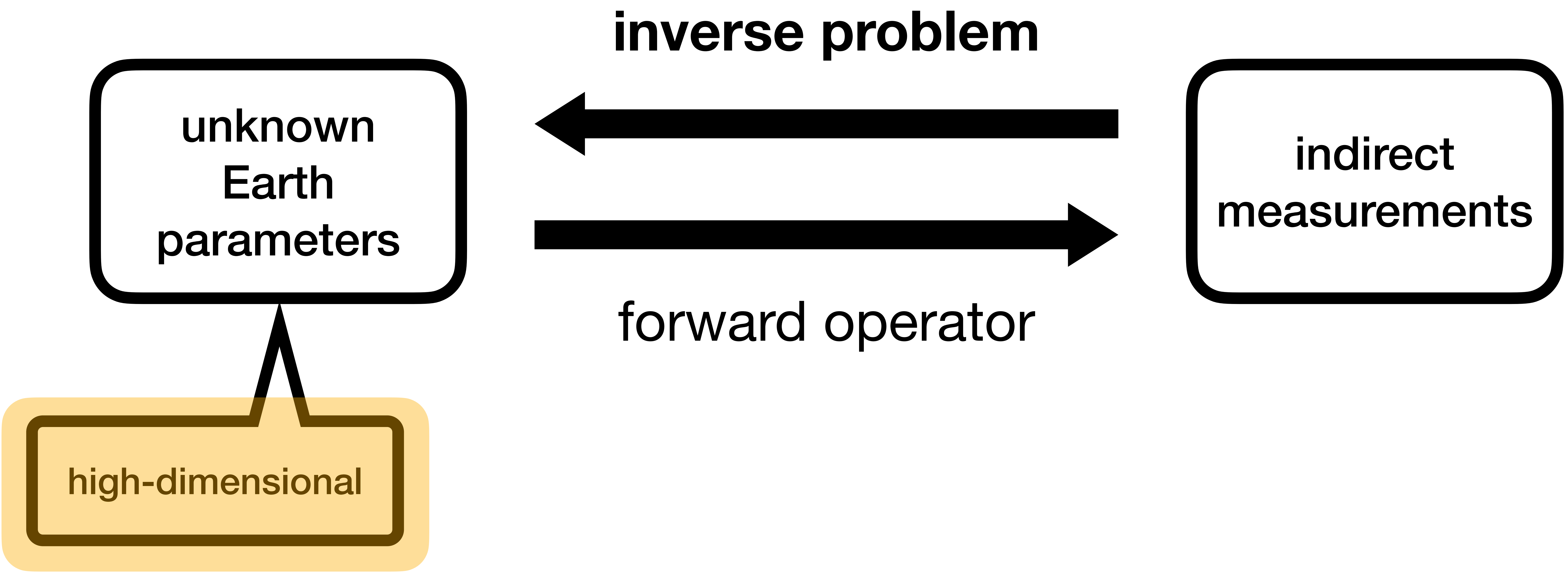
ML4Seismic

Georgia Tech

Inverse problems

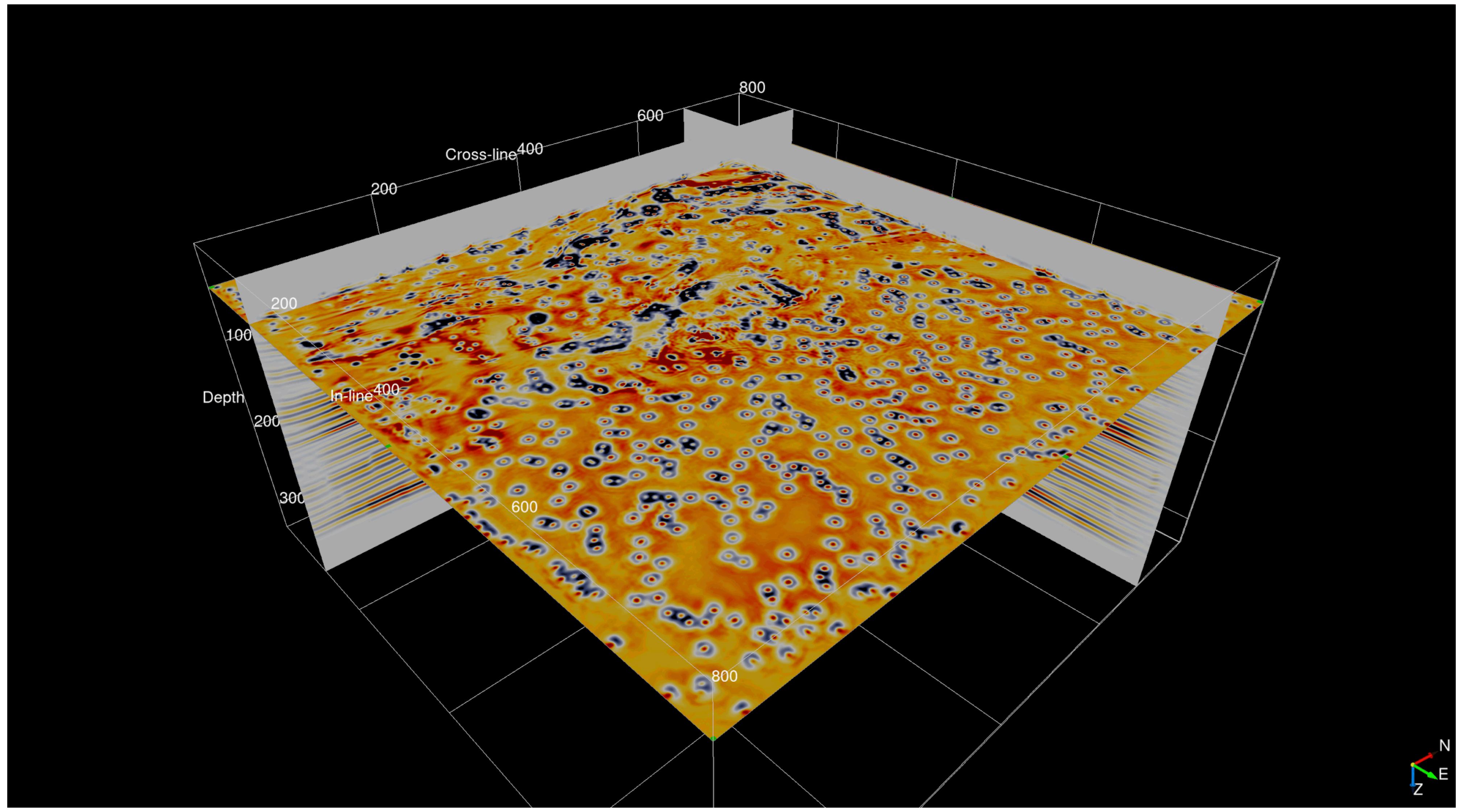


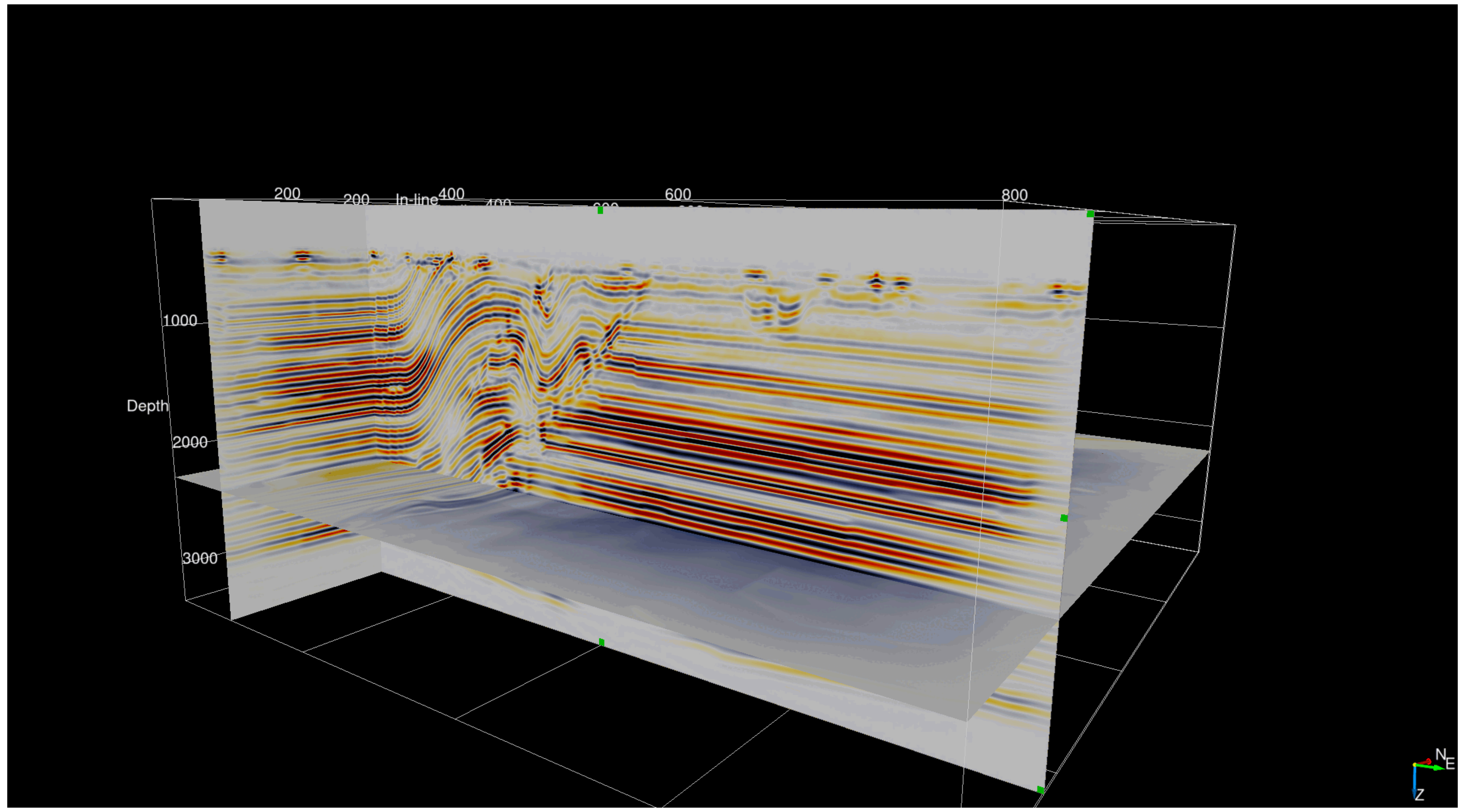
Geophysical inverse problems



Movies adapted from P. A. Witte, M. Louboutin, H. Modzelewski, C. Jones, J. Selvage and F. J. Herrmann, "An Event-Driven Approach to Serverless Seismic Imaging in the Cloud," in IEEE Transactions on Parallel and Distributed Systems, 2020.

High-dimensional parameter estimation

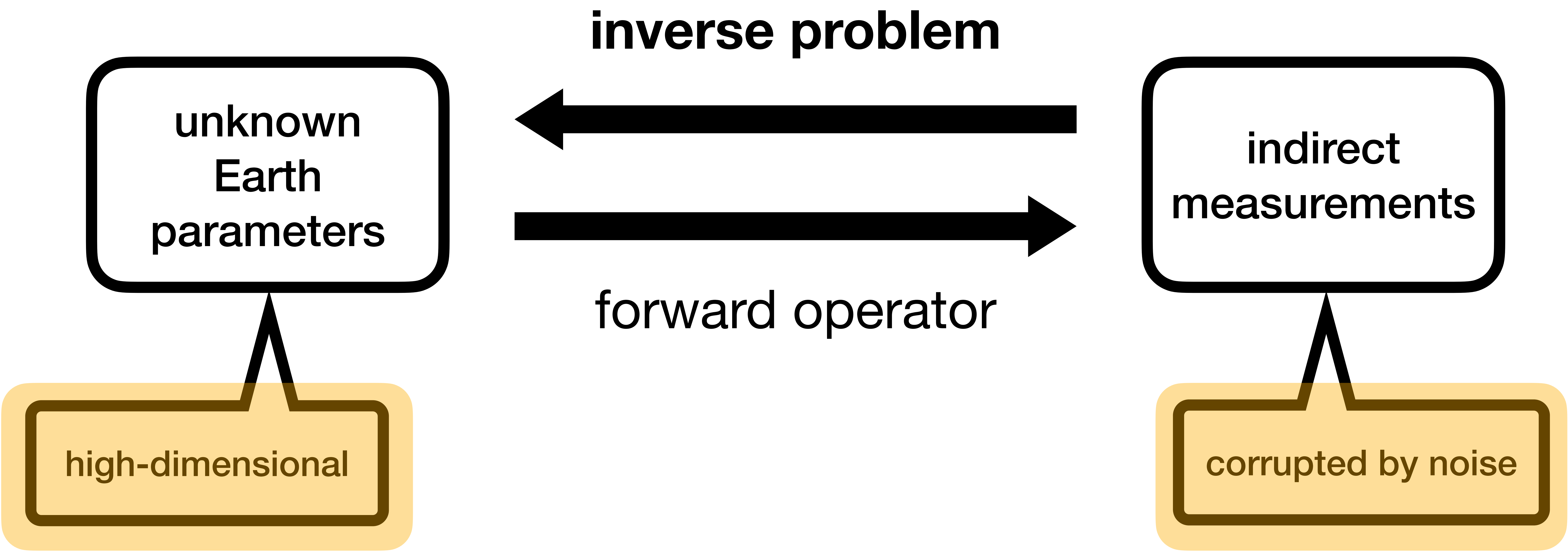




Geophysical exploration and monitoring

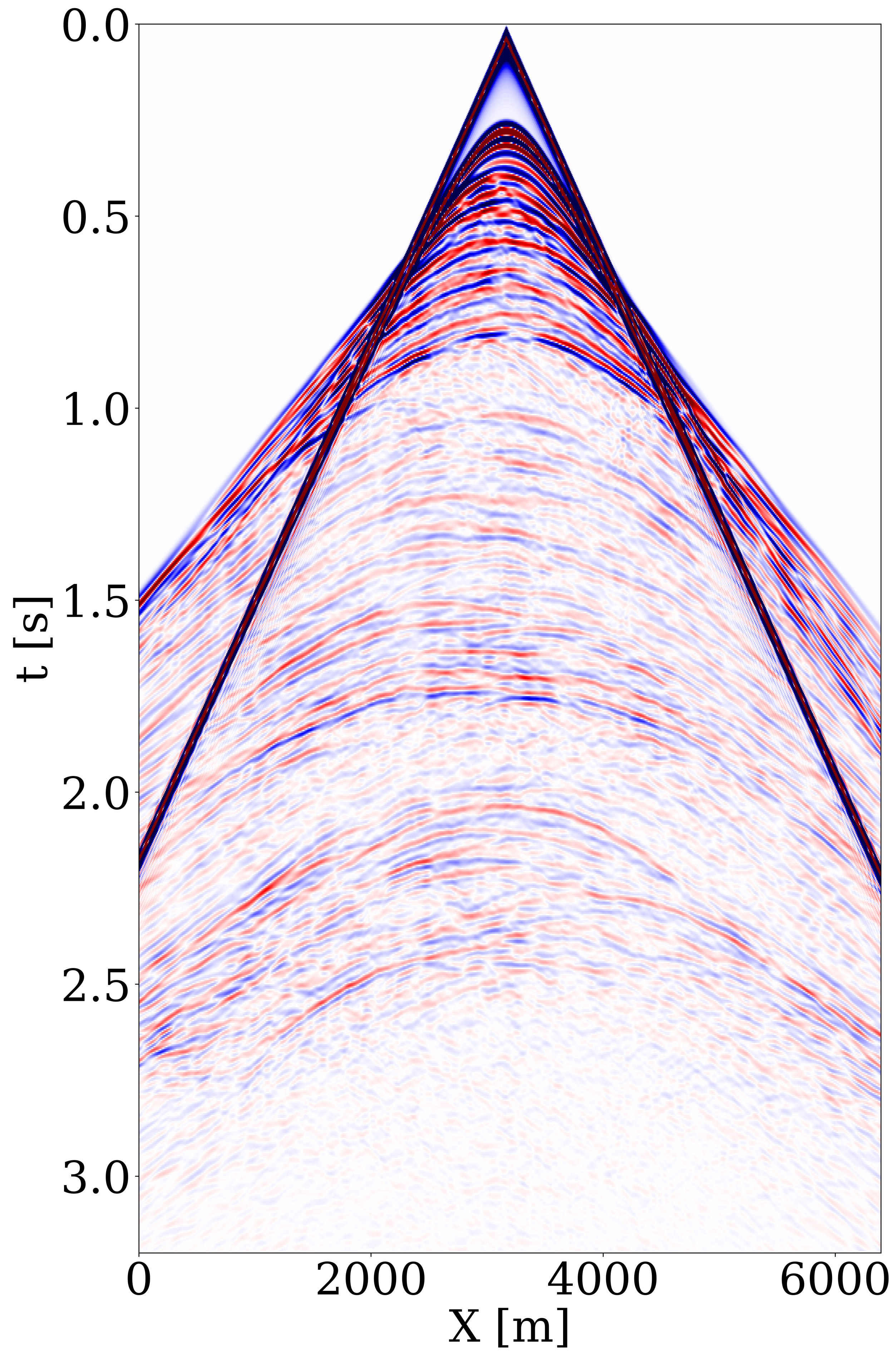
- over large subsurface areas
- require high-resolution Earth imaging
- $nx \times ny \times nz \sim \mathrm{O}(10^3 \times 10^3 \times 10^3)$ in realistic settings

Geophysical inverse problems

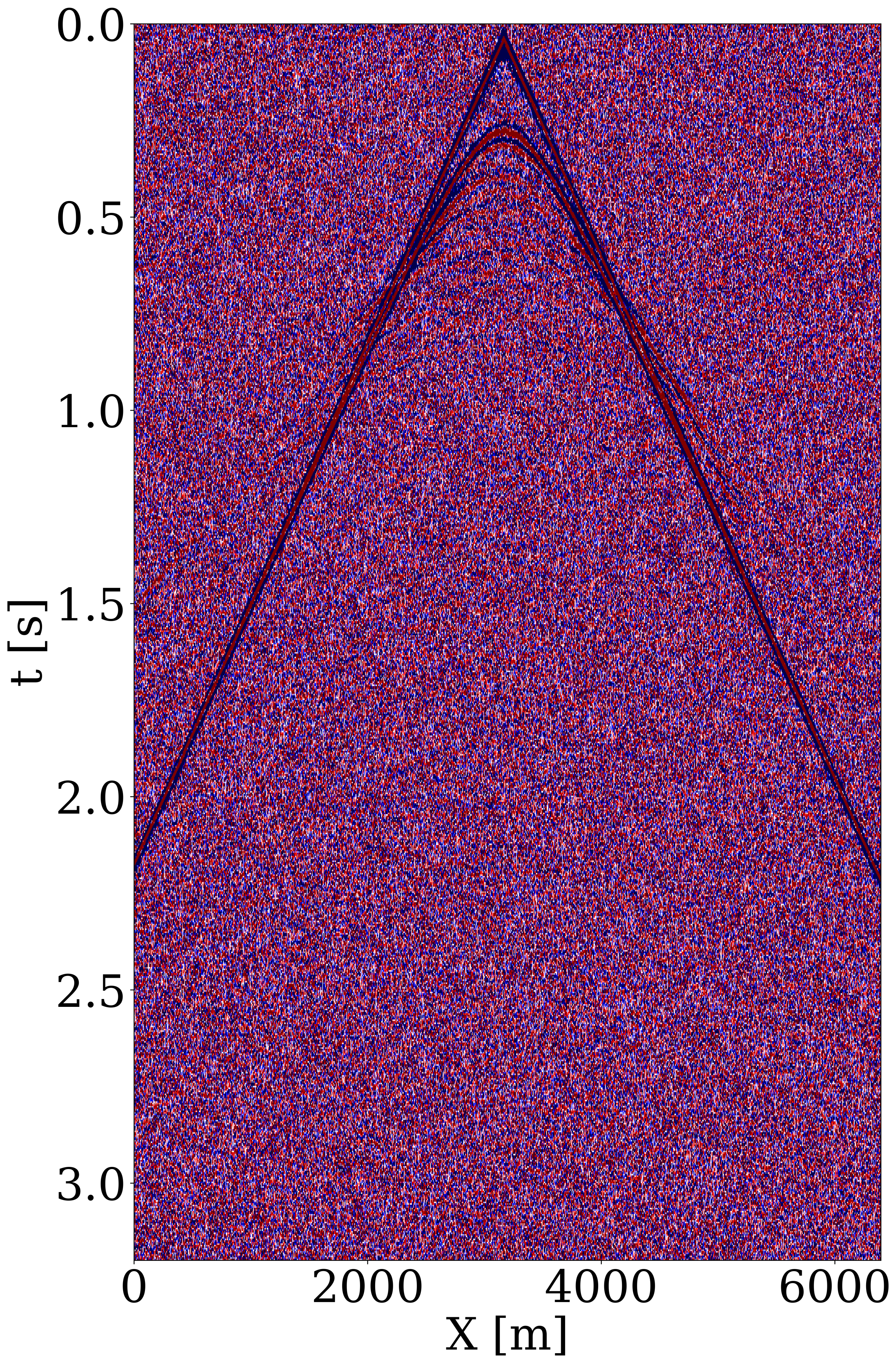


Noisy geophysical observations

noise-free data



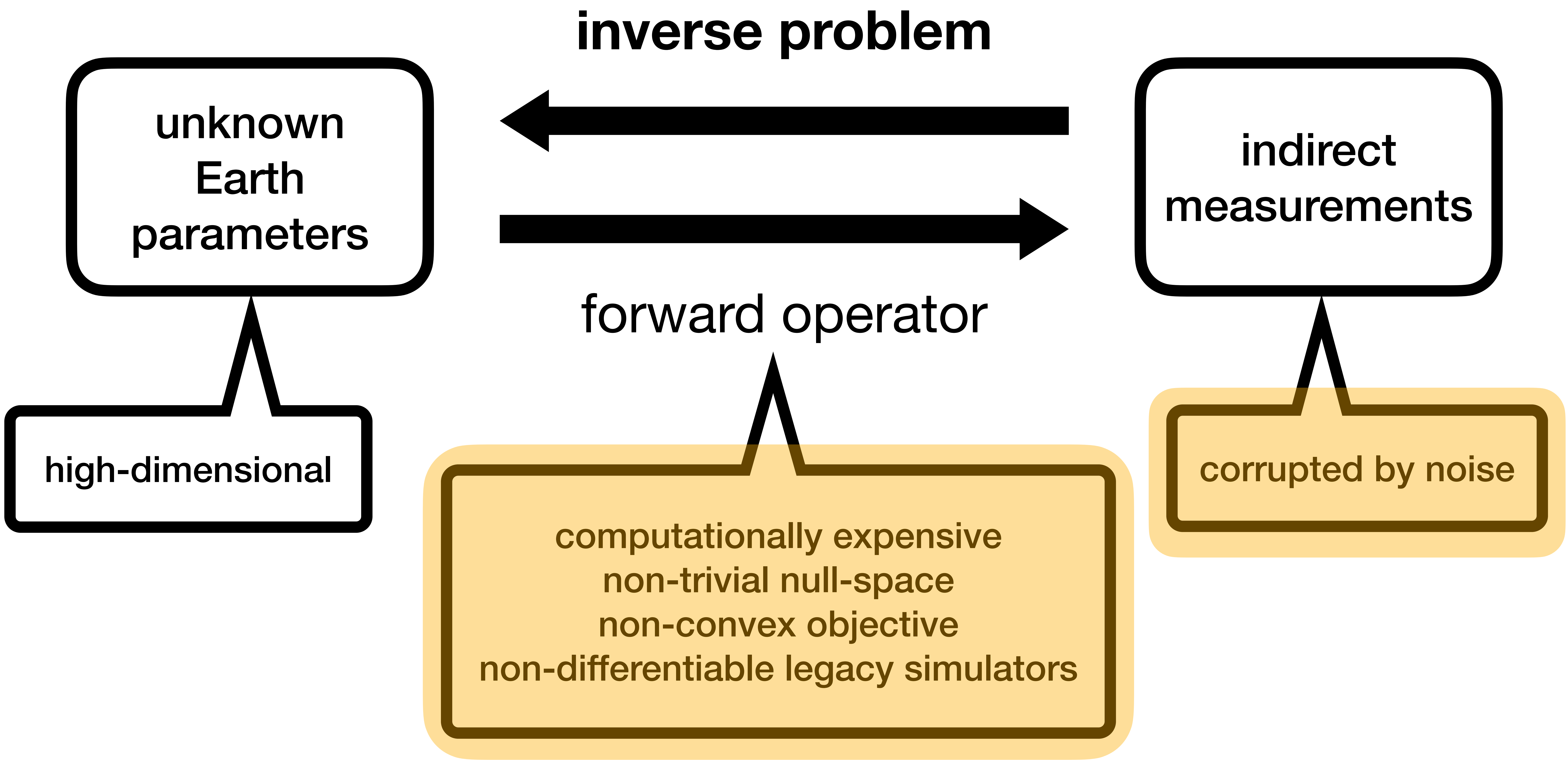
noisy data



Weak seismic signals are often corrupted by strong observational noise

Often lead to imaging artifacts

Geophysical inverse problems



Louboutin, Mathias, et al. "Devito (v3. 1.0): an embedded domain-specific language for finite differences and geophysical exploration." *Geoscientific Model Development* 12.3 (2019)
Luporini, Fabio, et al. "Architecture and performance of Devito, a system for automated stencil computation." *ACM Transactions on Mathematical Software (TOMS)* 46.1 (2020)
Rasmussen, Atgeirr Flø, et al. "The open porous media flow reservoir simulator." *Computers & Mathematics with Applications* 81 (2021)
Virieux, Jean, and Stéphane Operto. "An overview of full-waveform inversion in exploration geophysics." *Geophysics* 74.6 (2009): WCC1-WCC26.
Grady, Thomas J., et al. "Model-parallel Fourier neural operators as learned surrogates for large-scale parametric PDEs." *Computers & Geosciences* 178 (2023): 105402.

Forward modeling operators

numerical simulators

Computationally expensive

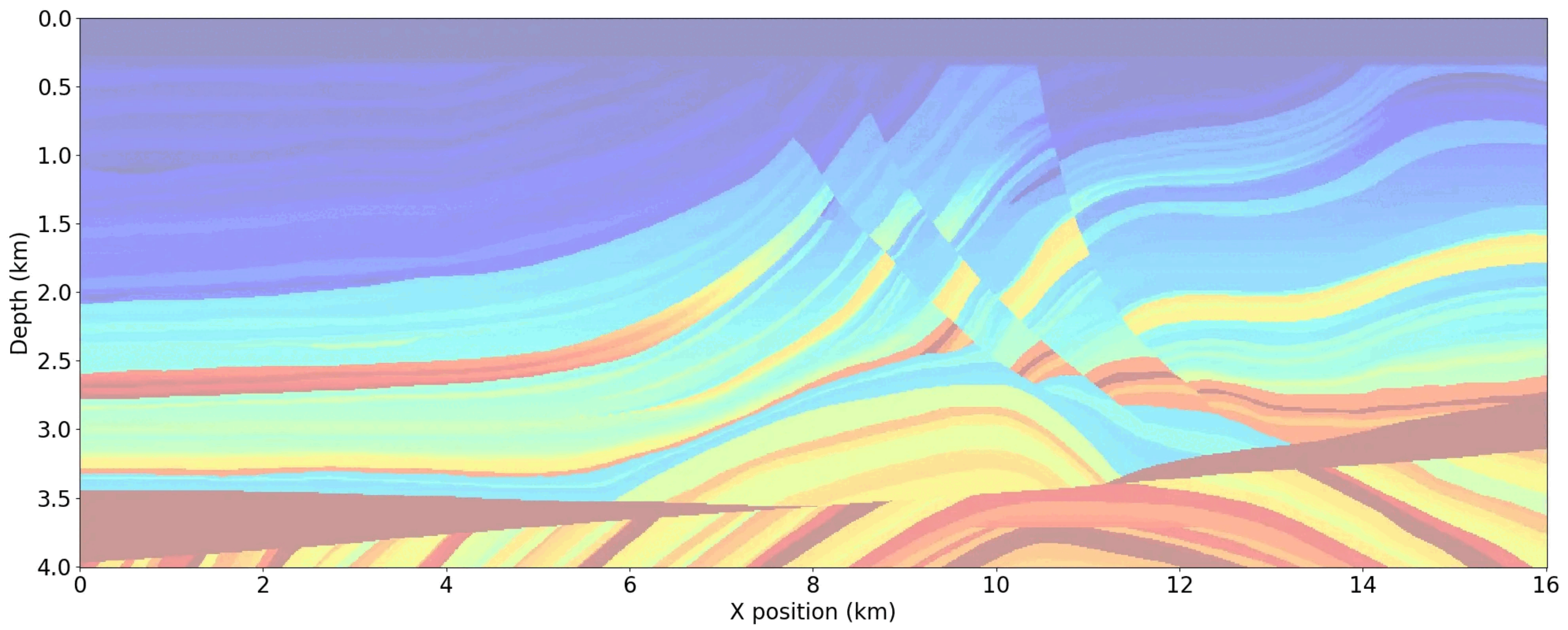
- physics-based simulation
- require solving PDEs

Legacy solvers

- lack interoperability
- difficult to derive sensitivities w.r.t. model parameters

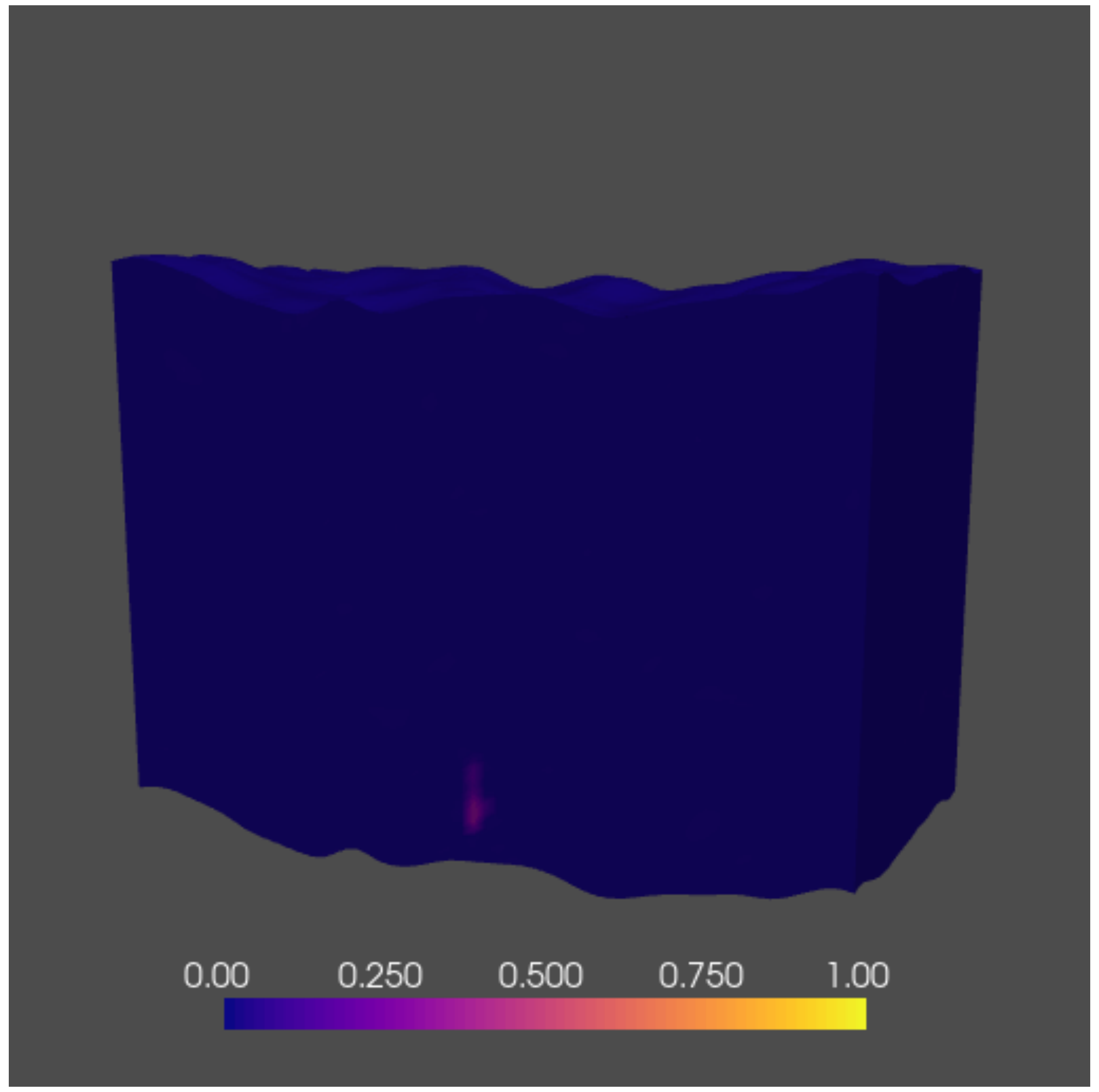
Mathematically challenging

- non-convex objective
- non-trivial null-space



acoustic wave

multiphase flow in porous media



Objectives of my dissertation

Develop scientific machine learning (SciML) methods at scale

- scalable, interoperable, differentiable programming frameworks
- achieve more accurate solutions
- accelerate the inversion process
- provide reliable & affordable inversion
- computationally feasible uncertainty quantification (UQ)

SLIM ML4Seismic

Mathias Louboutin*, Ziyi Yin*, Rafael Orozco, Thomas J. Grady II, Ali Siahkoohi, Gabrio Rizzuti, Philipp A. Witte, Olav Møyner, Gerard J. Gorman, and Felix J. Herrmann. “Learned multiphysics inversion with differentiable programming and machine learning.” The Leading Edge, 2023.
Ziyi Yin, Ali Siahkoohi, Mathias Louboutin, and Felix J. Herrmann. "Learned coupled inversion for carbon sequestration monitoring and forecasting with Fourier neural operators.” International Meeting for Applied Geoscience and Energy Expanded Abstracts, 2022. **(Best student paper honorable mention)**
Mathias Louboutin, Philipp A. Witte, Ali Siahkoohi, Gabrio Rizzuti, Ziyi Yin, Rafael Orozco, and Felix J. Herrmann. “Accelerating innovation with software abstractions for scalable computational geophysics.” International Meeting for Applied Geoscience and Energy Expanded Abstracts, 2022.

Chapter 2

Learned multiphysics inversion with differentiable programming and machine learning

Motivation

multiphysics inversion

Legacy software

- performant, optimized by domain experts — decades of efforts
- lack portability & interoperability
- difficult to maintain or add new features
- (some) lack differentiability & sensitivity calculation

⟷ **hinder R & D**

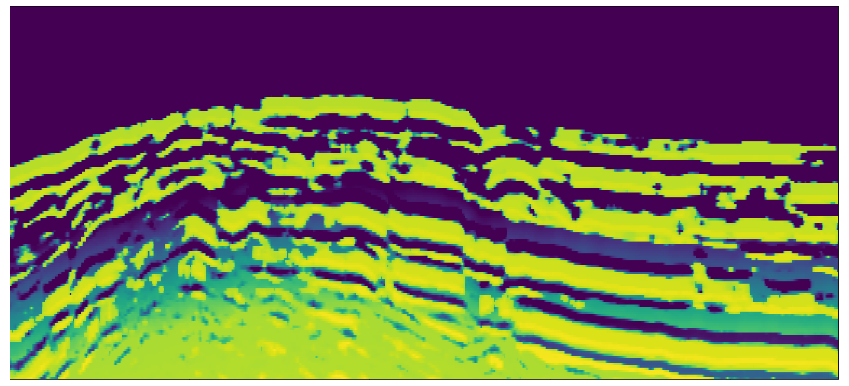
Time-lapse seismic monitoring of geological carbon storage (GCS)

- involves coupling of multiphysics modeling & inversion
- requires scalable, interoperable & differentiable software stack

Multiphysics modeling
GCS monitoring

permeability

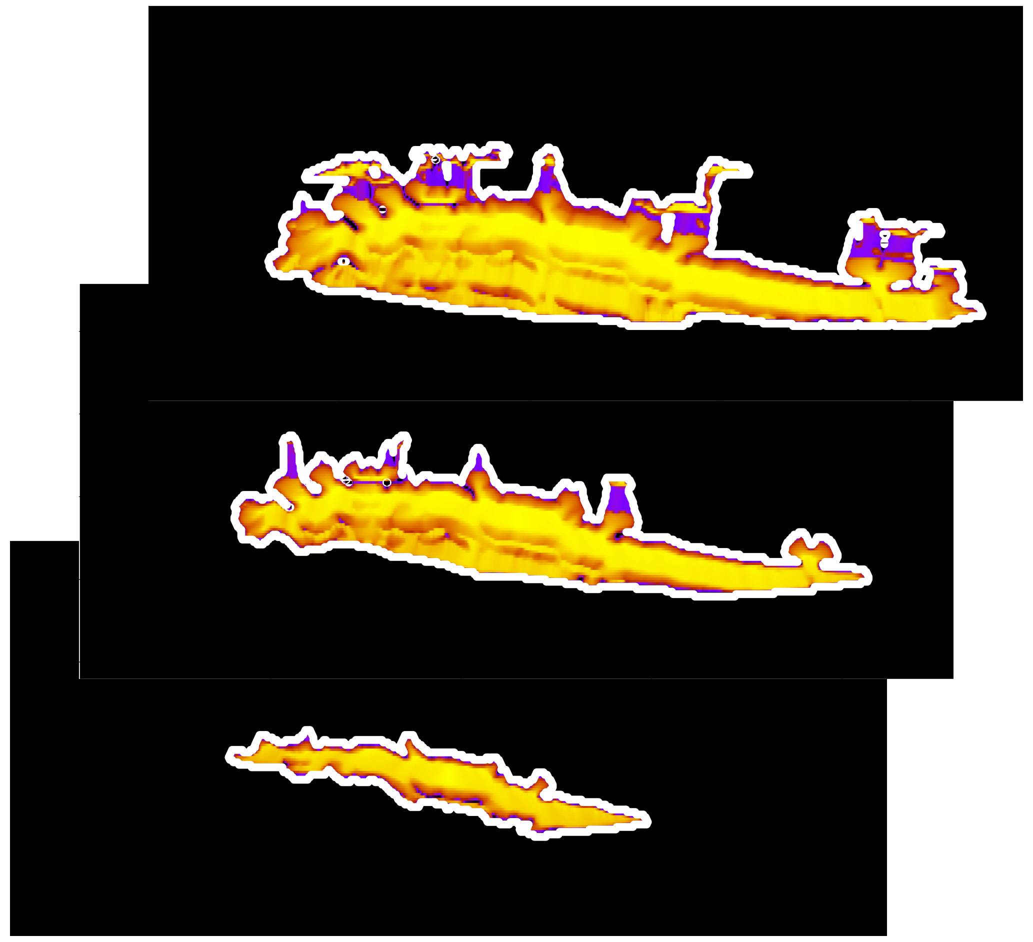
$$\mathbf{K}$$



$\mathscr{S}$ $\rightarrow$ **fluid-flow physics**

CO_2 saturation

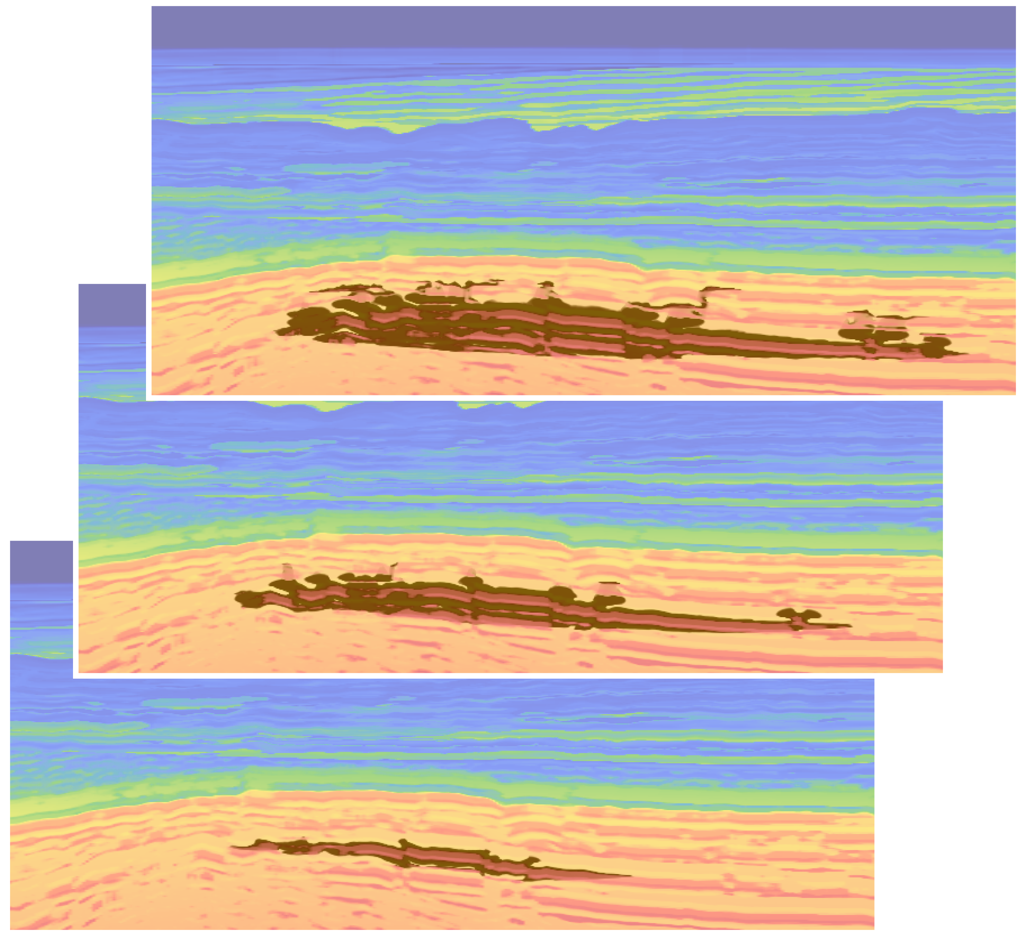
$$\mathbf{c} = \{\mathbf{c}_k\}_{k=1}^{n_k}$$



$\mathscr{R}$ $\rightarrow$ **rock physics**

seismic velocity

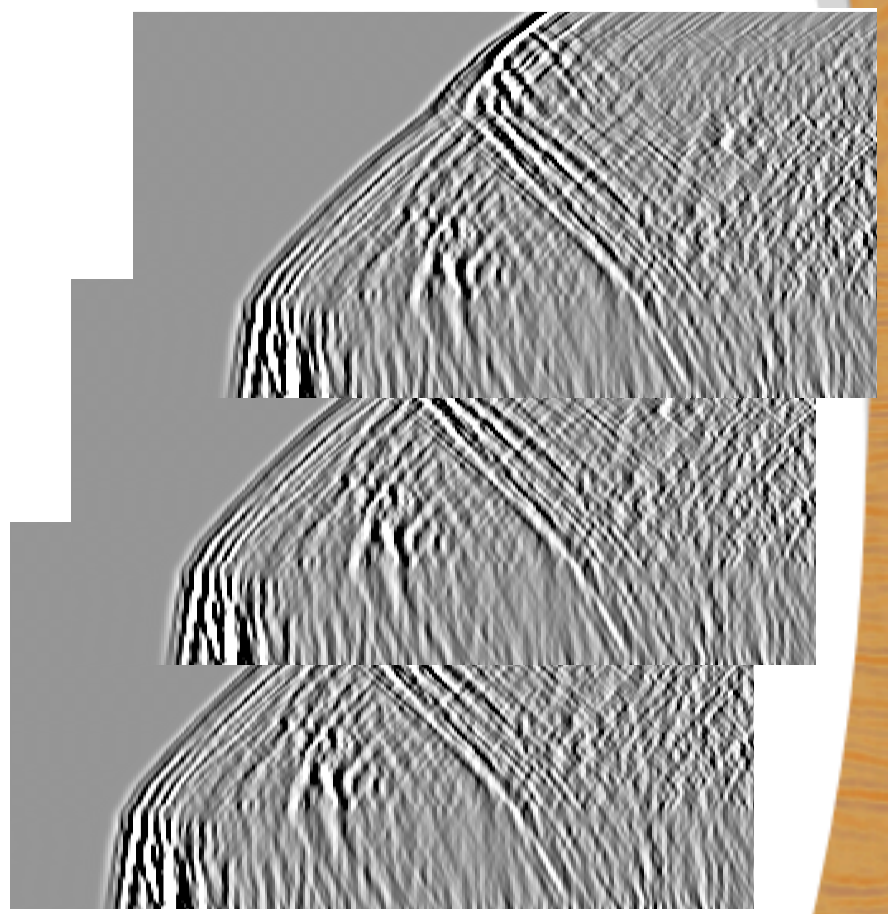
$$\mathbf{v} = \{\mathbf{v}_k\}_{k=1}^{n_k}$$



$\mathscr{F}$ $\rightarrow$ **wave physics**

time-lapse seismic data

$$\mathbf{d} = \{\mathbf{d}_k\}_{k=1}^{n_k}$$



Contributions
Chapter 2

Differentiable programming framework via math-inspired software abstractions

- *customized* automatic differentiation (AD) via integration with ChainRules.jl
- coupling of *disjoint* software libraries is feasible and scalable
- easily support *deep learning integration* (e.g., surrogate-assisted inversion)

Case study

- permeability inversion during GCS monitoring

End-to-end inversion framework
multiphysics coupling

permeability
$\mathbf{K}$

CO_2 saturation
$\mathbf{c} = \{\mathbf{c}_k\}_{k=1}^{n_k}$

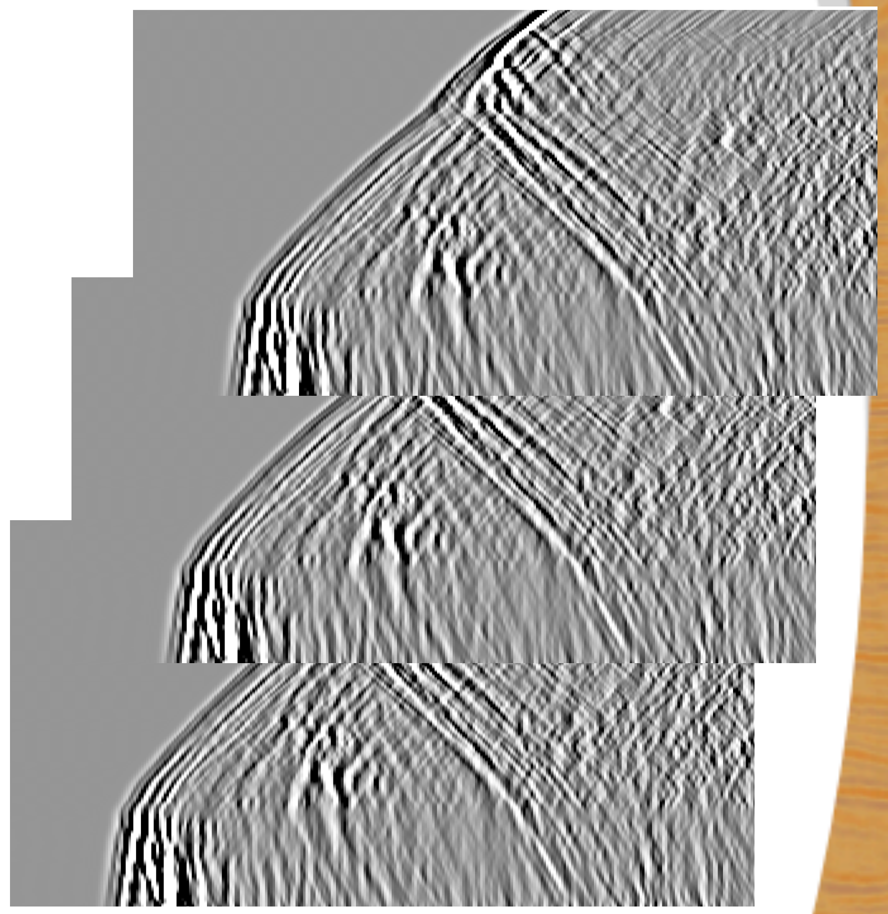
seismic velocity
$\mathbf{v} = \{\mathbf{v}_k\}_{k=1}^{n_k}$

time-lapse seismic data
$\mathbf{d} = \{\mathbf{d}_k\}_{k=1}^{n_k}$

?

$$\xrightarrow{\mathcal{F} \circ \mathcal{R} \circ \mathcal{S}}$$

coupled physics



$$\underset{\mathbf{K}}{\text{minimize}} \quad \|\mathcal{F} \circ \mathcal{R} \circ \mathcal{S}(\mathbf{K}) - \mathbf{d}\|_2^2$$

Li D, Xu K, Harris JM, Darve E. Coupled time-lapse full-waveform inversion for subsurface flow problems using intrusive automatic differentiation. Water Resources Research. 2020 Aug;56(8):e2019WR027032.

End-to-end inversion framework
physics-based

permeability
$\mathbf{K}$

CO_2 saturation
$\mathbf{c} = \{\mathbf{c}_k\}_{k=1}^{n_k}$

seismic velocity
$\mathbf{v} = \{\mathbf{v}_k\}_{k=1}^{n_k}$

time-lapse seismic data
$\mathbf{d} = \{\mathbf{d}_k\}_{k=1}^{n_k}$

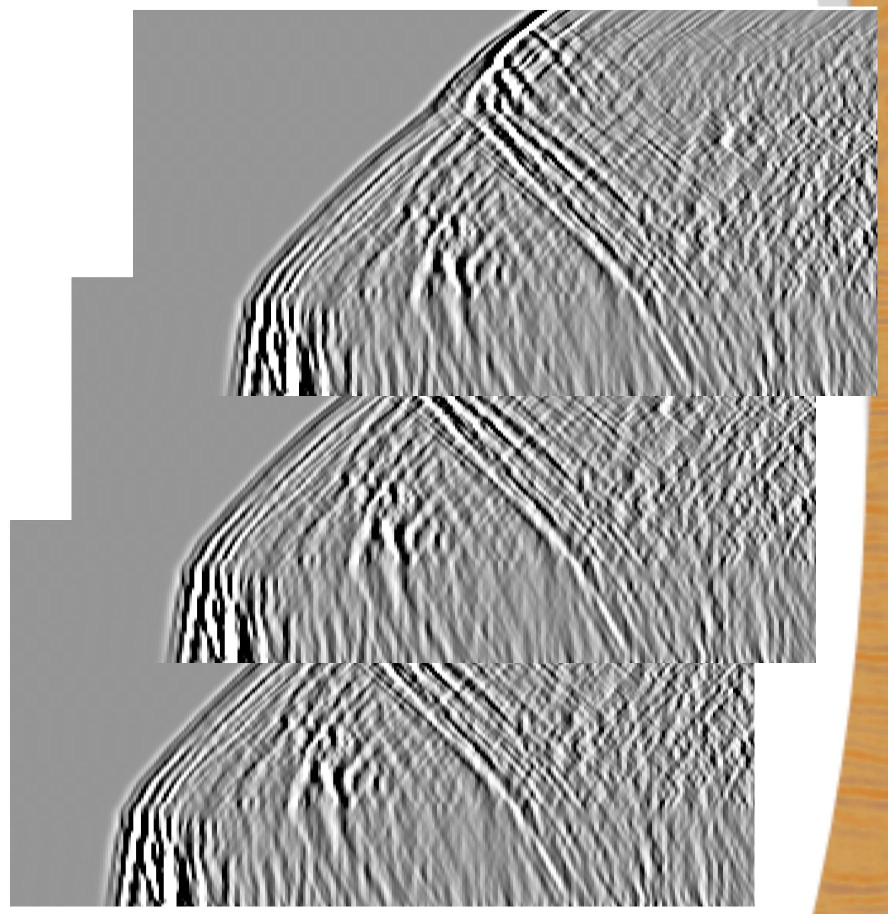
Devito / JUDI.jl

JutulDarcy.jl

$\mathscr{F} \circ \mathscr{R} \circ \mathscr{S}$

?

coupled physics



customized/hand-written

Julia native AD

Louboutin, Mathias, et al. "Devito (v3. 1.0): an embedded domain-specific language for finite differences and geophysical exploration." *Geoscientific Model Development* 12.3 (2019)
Luporini, Fabio, et al. "Architecture and performance of Devito, a system for automated stencil computation." *ACM Transactions on Mathematical Software (TOMS)* 46.1 (2020)
Witte PA, Louboutin M, Kukreja N, Luporini F, Lange M, Gorman GJ, Herrmann FJ. A large-scale framework for symbolic implementations of seismic inversion algorithms in Julia. Geophysics. 2019
Møyner, Olav, Grant Bruer, and Ziyi Yin. "Sintefmath/JutulDarcy. jl: V0. 2.3 (version v0. 2.3). Zenodo." (2023).
Yin, Ziyi, Grant Bruer, and Mathias Louboutin. "Slimgroup/JutulDarcyRules. jl: V0. 2.5 (version v0. 2.5). Zenodo." (2023).

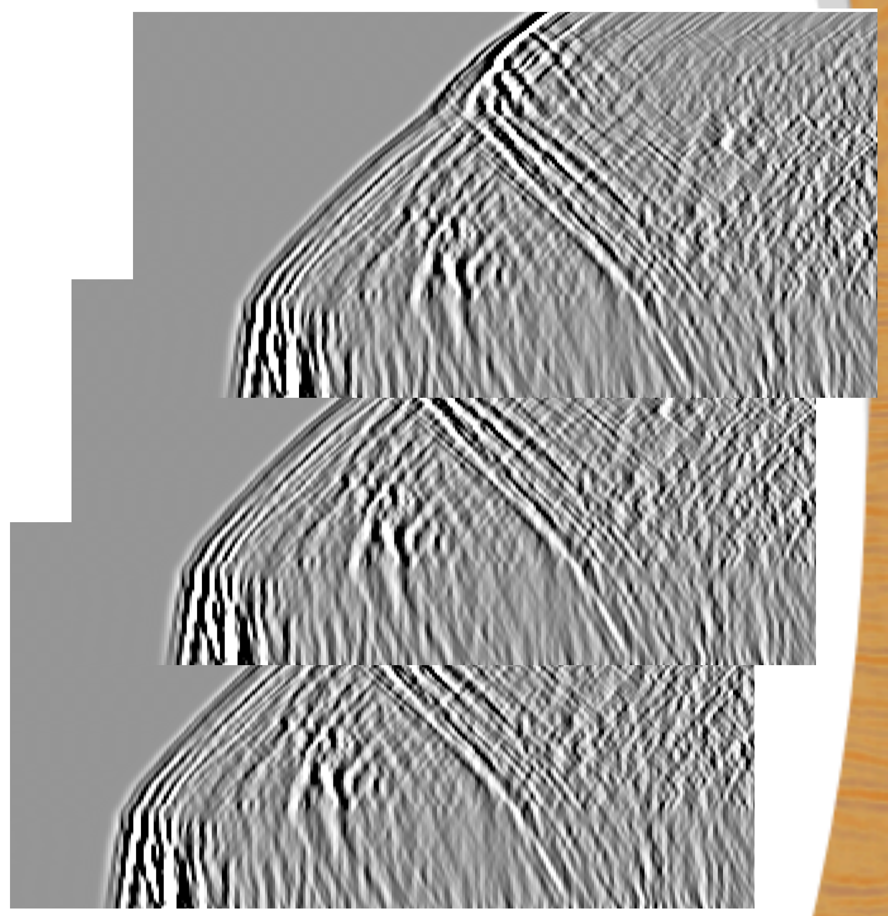
End-to-end inversion framework
surrogate-assisted

permeability
$\mathbf{K}$

CO_2 saturation
$\mathbf{c} = \{\mathbf{c}_k\}_{k=1}^{n_k}$

seismic velocity
$\mathbf{v} = \{\mathbf{v}_k\}_{k=1}^{n_k}$

time-lapse seismic data
$\mathbf{d} = \{\mathbf{d}_k\}_{k=1}^{n_k}$



?

$\mathcal{F} \circ \mathcal{R} \circ \mathcal{S}_{\boldsymbol{\theta}*}$

coupled physics

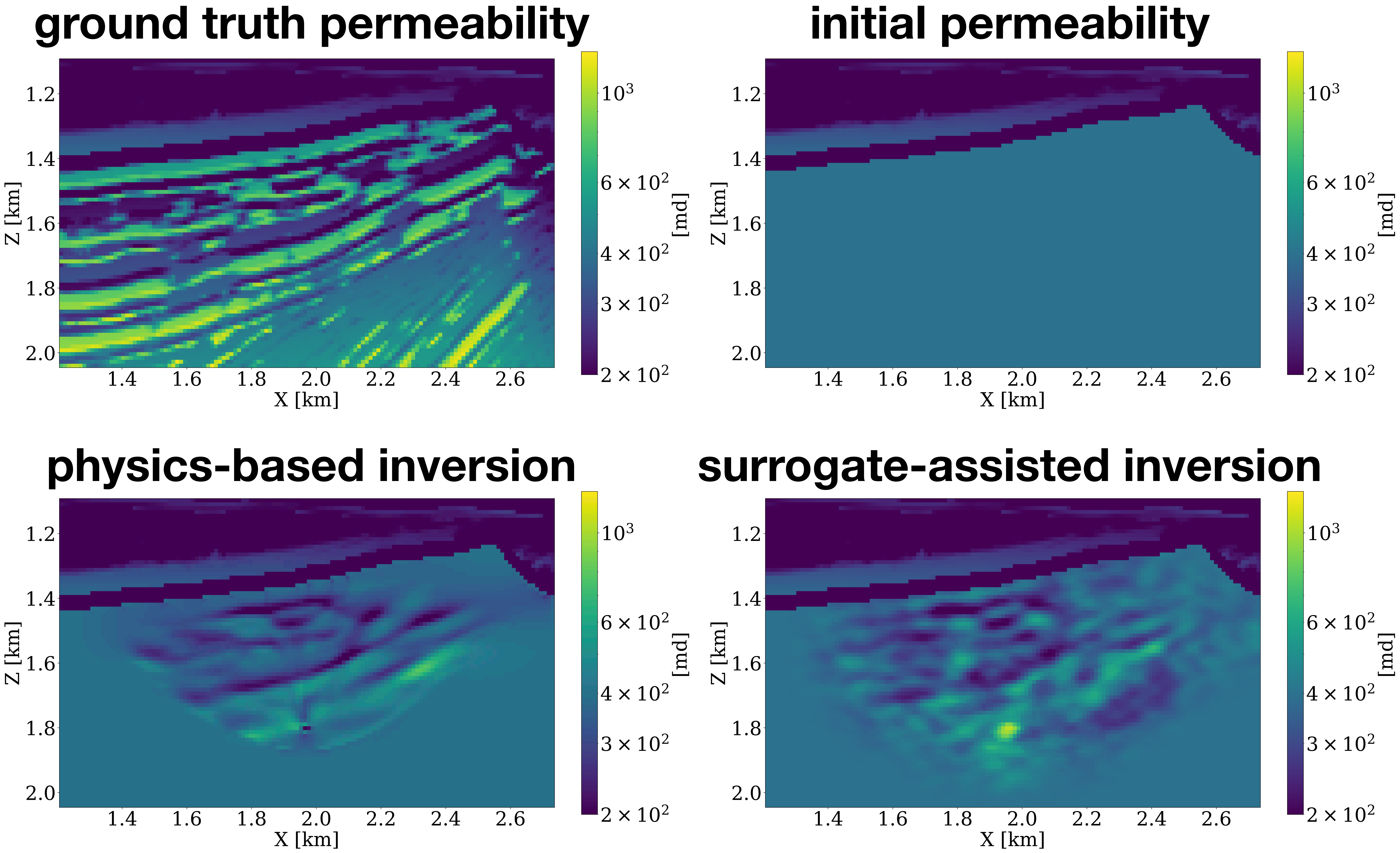
customized/hand-written

Julia native AD

$\mathcal{S}_{\boldsymbol{\theta}^*}$ **trained Fourier neural operators (FNOs)**

Li, Zongyi, et al. "Fourier neural operator for parametric partial differential equations." *arXiv preprint arXiv:2010.08895* (2020).

Case study on the Compass model



CE, JA Edgar, JI Selvage, and H Crook. 2012. “Building Complex Synthetic Models to Evaluate Acquisition Geometries and Velocity Inversion Technologies.” In *74th EAGE Conference and Exhibition Incorporating EUROPEC 2012*, cp–293. European Association of Geoscientists & Engineers.

SLIM ML4Seismic

Ziyi Yin, Mathias Louboutin, Olav Møyner, and Felix J. Herrmann. “Time-lapse full-waveform permeability inversion: a feasibility study”. The Leading Edge, 2024.

Chapter 3

Time-lapse full-waveform permeability inversion: a feasibility study

Contributions

Chapter 3

Examine the sensitivities of the permeability inversion framework w.r.t.

- initial model parameters
- modeling errors
- crosstalk during multiparameter inversion

Inversion leads to downstream tasks

- forecast CO_2 plume in the future w/o any observation

SLIM ML4Seismic

Ziyi Yin, Rafael Orozco, Mathias Louboutin, and Felix J. Herrmann. “Solving multiphysics-based inverse problems with learned surrogates and constraints.” Advanced Modeling and Simulation in Engineering Sciences, 2023.

Chapter 4

Solving multiphysics-based inverse problems with learned surrogates and constraints

Problem formulation

Solve inverse problem: $\mathbf{d} = \mathscr{H} \circ \mathscr{S}(\mathbf{K}) + \boldsymbol{\epsilon}$

- $\mathbf{d}$ observed data with noise $\boldsymbol{\epsilon}$
- $\mathbf{K}$ unknown parameter of interest
- $\mathscr{S}$ modeling operator
- $\mathscr{H}$ measurement operator

Motivation

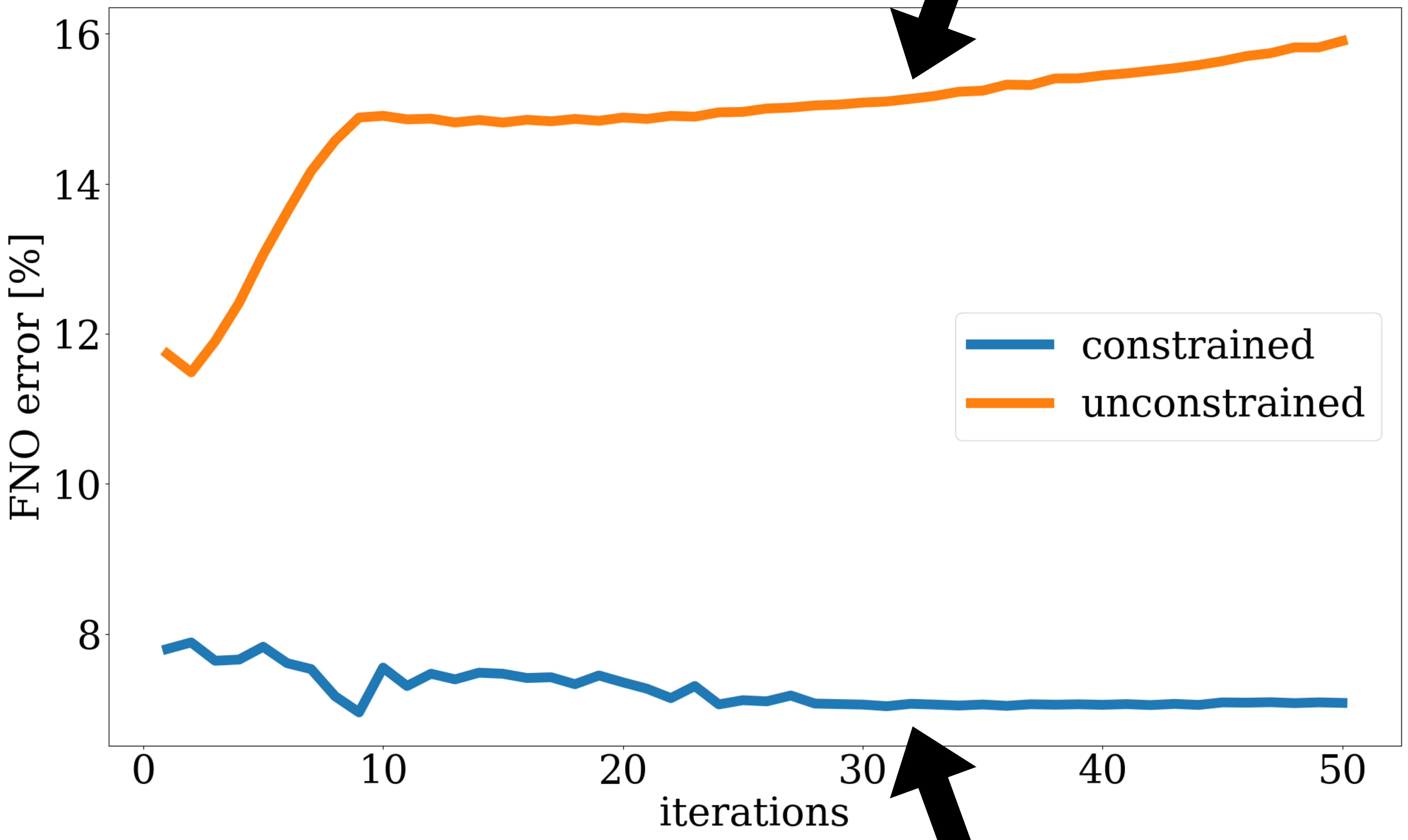
surrogate-assisted inversion

$$\underset{\mathbf{K}}{\text{minimize}} \quad \|\mathbf{d} - \mathscr{H} \circ \mathscr{S}_{\boldsymbol{\theta}*}(\mathbf{K})\|_2^2$$

Replace numerical simulator $\mathscr{S}$ by trained FNO $\mathscr{S}_{\boldsymbol{\theta}*}$

- orders of magnitude faster
- auto-differentiable
- intermediate $\mathbf{K}$ might go out-of-distribution (OOD)
- FNO prediction is less accurate $\mathscr{S}(\mathbf{K}) \neq \mathscr{S}_{\boldsymbol{\theta}*}(\mathbf{K})$

FNO error keeps increasing



Goal: “flatten the curve”

Objective

Chapter 4

Propose a learned inversion algorithm

- reap computational benefit of FNO surrogates - **fast**
- **constrain the FNO input to be always in-distribution - accurate**
- still bring down the data misfit via iterative optimization

Adapted from Rafael Orozco, Mathias Louboutin, Peng Chen, Felix J. Herrmann. “Amortized Bayesian full-waveform inversion and experimental design with normalizing flows”. International Meeting for Applied Geoscience & Energy, 2023.

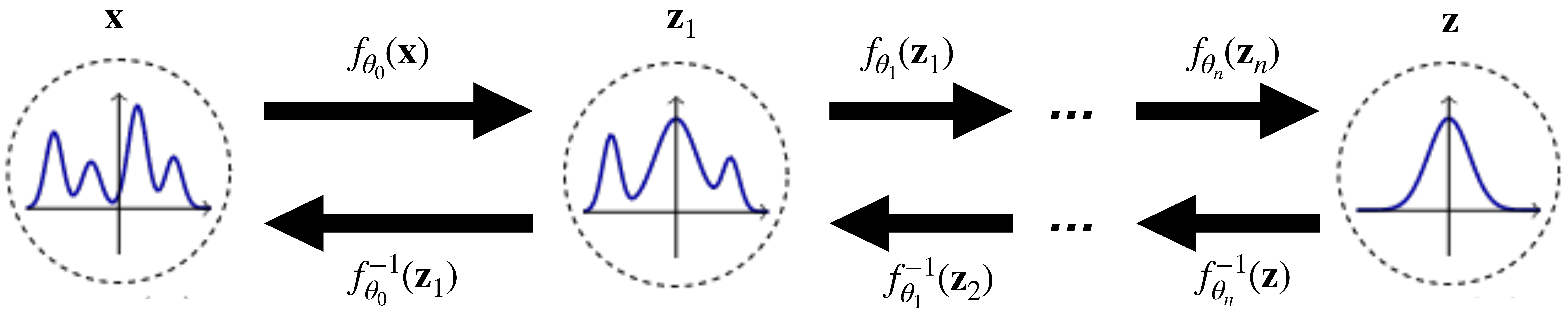
SLIM ML4Seismic

Normalizing flows (NFs)

transport maps

Learn distribution by mapping samples to Gaussian distribution

Mapping by design is **differentiable** and **invertible**



Kobyzev, Ivan, Simon Prince, and Marcus Brubaker. "Normalizing flows: An introduction and review of current methods." IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).

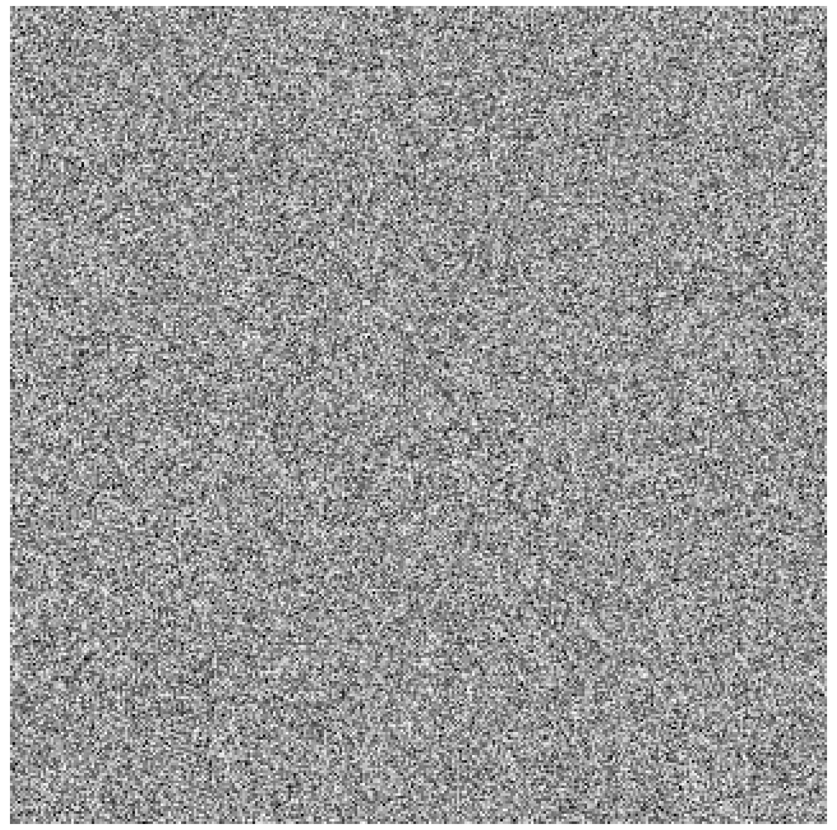
Normalizing flow (for cat)

model space | **latent space**

training:

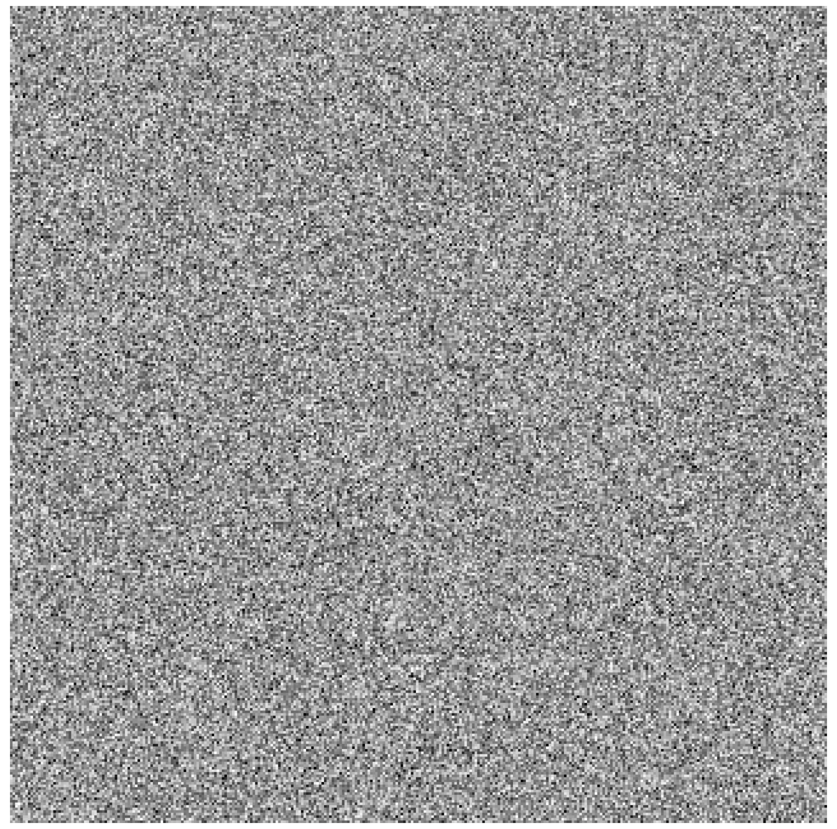
$\mathbf{x} \sim p_X(\mathbf{x})$

$$\xrightarrow{\mathscr{G}_{\mathbf{w}}^{-1}(\mathbf{x})}$$



$\mathbf{z} \sim p_Z(\mathbf{z})$

sampling:



$\mathbf{z} \sim p_Z(\mathbf{z})$

$$\xrightarrow{\mathscr{G}_{\mathbf{w}}(\mathbf{z})}$$

$\mathbf{x} \sim p_X(\mathbf{x})$

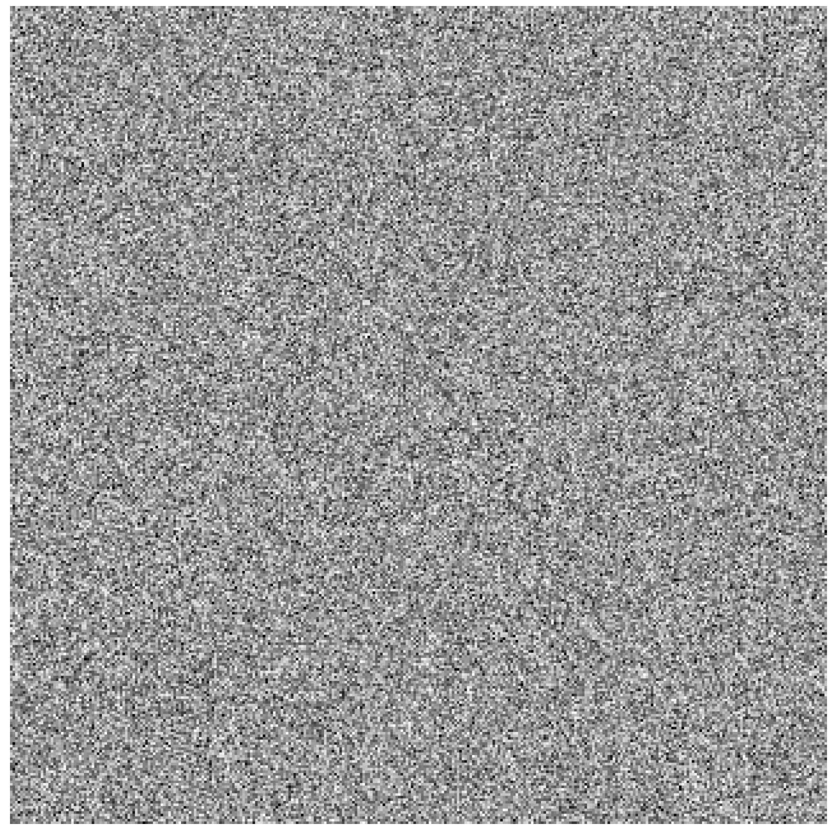
Normalizing flow (for Earth)

SLIM ML4Seismic

model space | **latent space**

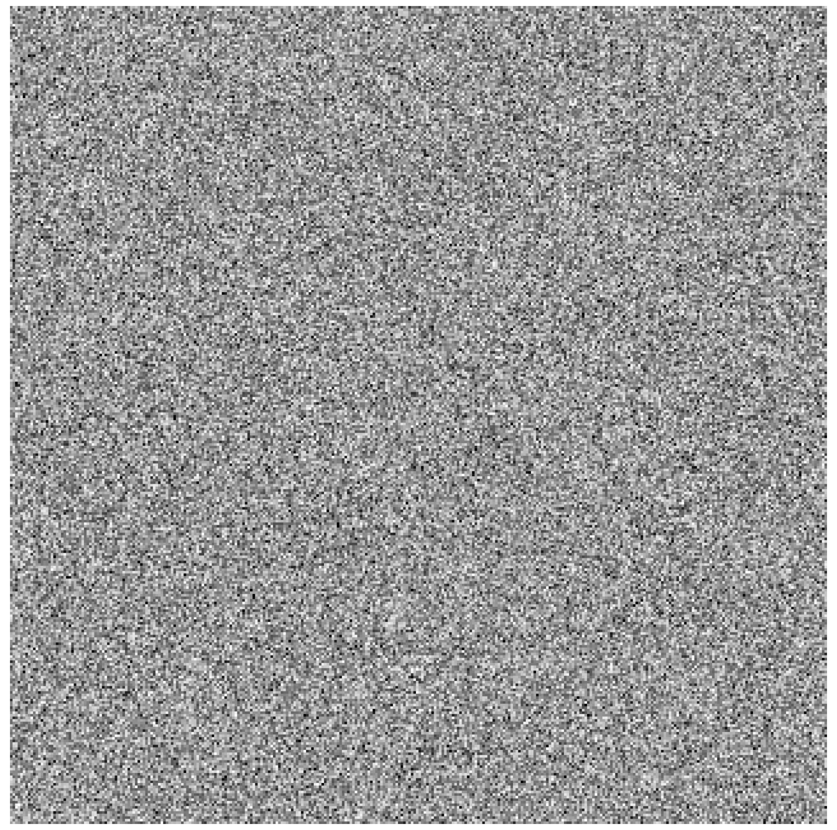
training:

$$\mathscr{G}_{\mathbf{w}}^{-1}(\mathbf{K}) \longrightarrow$$



$\mathbf{K} \sim p_K(\mathbf{K})$ | $\mathbf{z} \sim p_Z(\mathbf{z})$

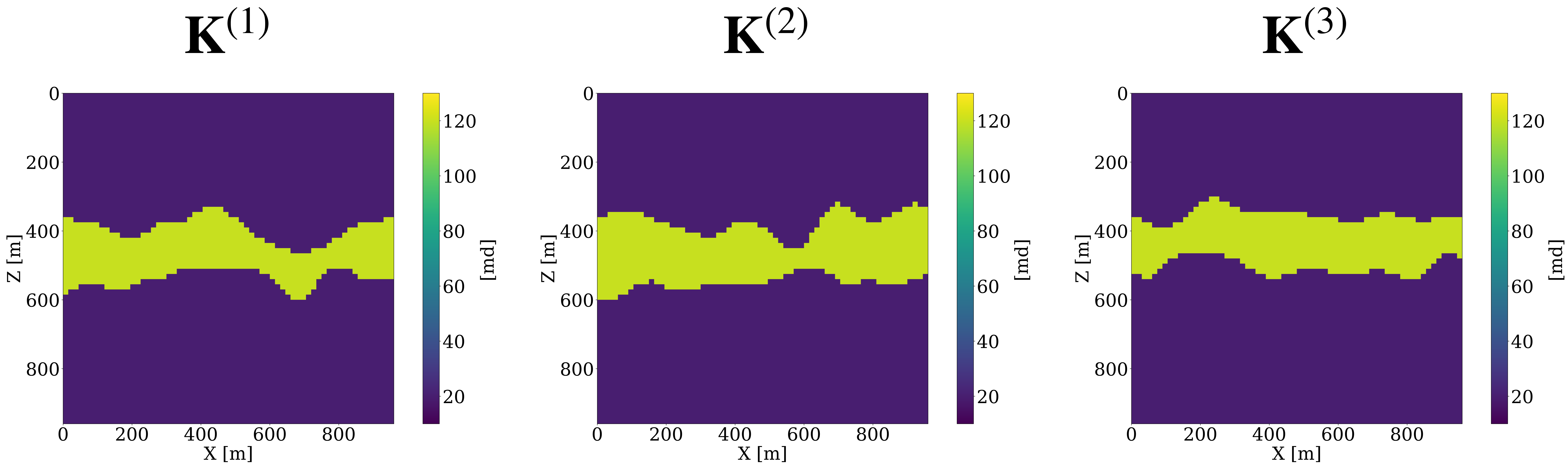
sampling:



$$\mathscr{G}_{\mathbf{w}}(\mathbf{z}) \longrightarrow$$

$\mathbf{z} \sim p_Z(\mathbf{z})$ | $\mathbf{K} \sim p_K(\mathbf{K})$

Prior distribution of the Earth
shared by FNO & NF training



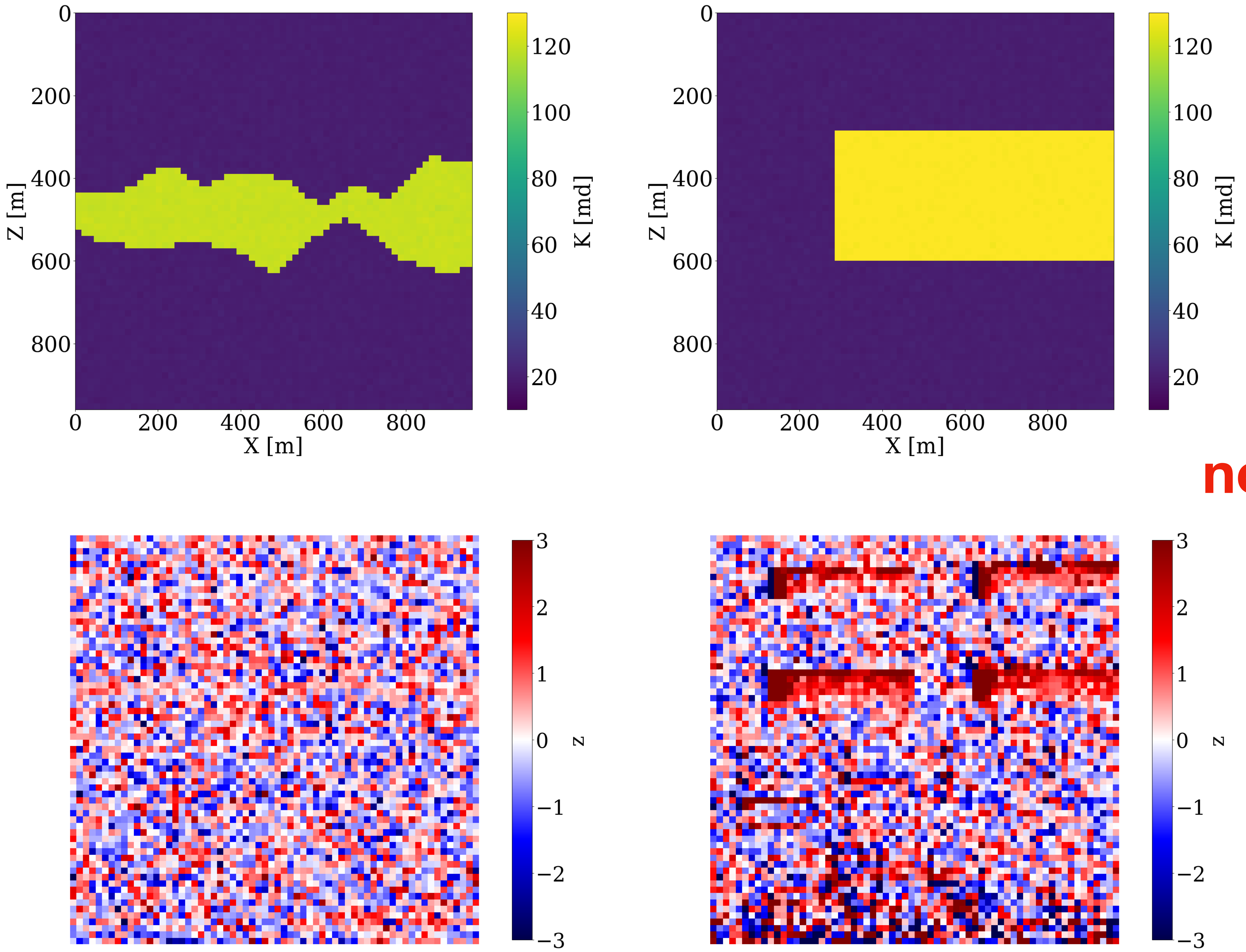
Invertibility of NF
enables probabilistic density evaluation

model space

$$\mathscr{G}^{-1}_{\mathbf{w}^*} \downarrow \quad \uparrow \mathscr{G}_{\mathbf{w}^*}$$

latent space

in-distribution | **out-of-distribution**



non-Gaussian for OOD sample

Shrinkage in the latent space of NFs

ℓ_2 norm ball shrinkage

SLIM ML4Seismic

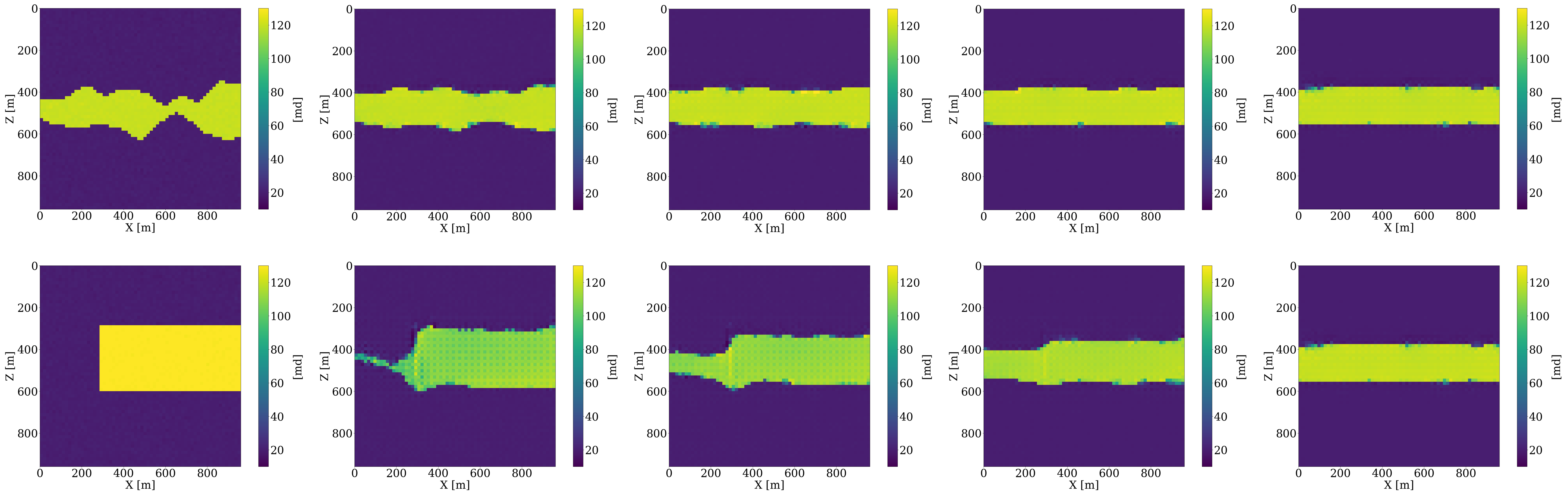
sequence $\widetilde{\mathbf{K}} = \mathcal{G}_{\mathbf{w}^*}(\alpha \mathbf{z})$ where $\mathbf{z} = \mathcal{G}^{-1}_{\mathbf{w}^*}(\mathbf{K})$ and $0 \leq \alpha \leq 1$

"shrink the latent code and observe the change in the model space"

$\alpha = 1$ | $\alpha = 0.4$ | $\alpha = 0.2$ | $\alpha = 0.1$ | $\alpha = 0$

in-distribution

out-of-distribution

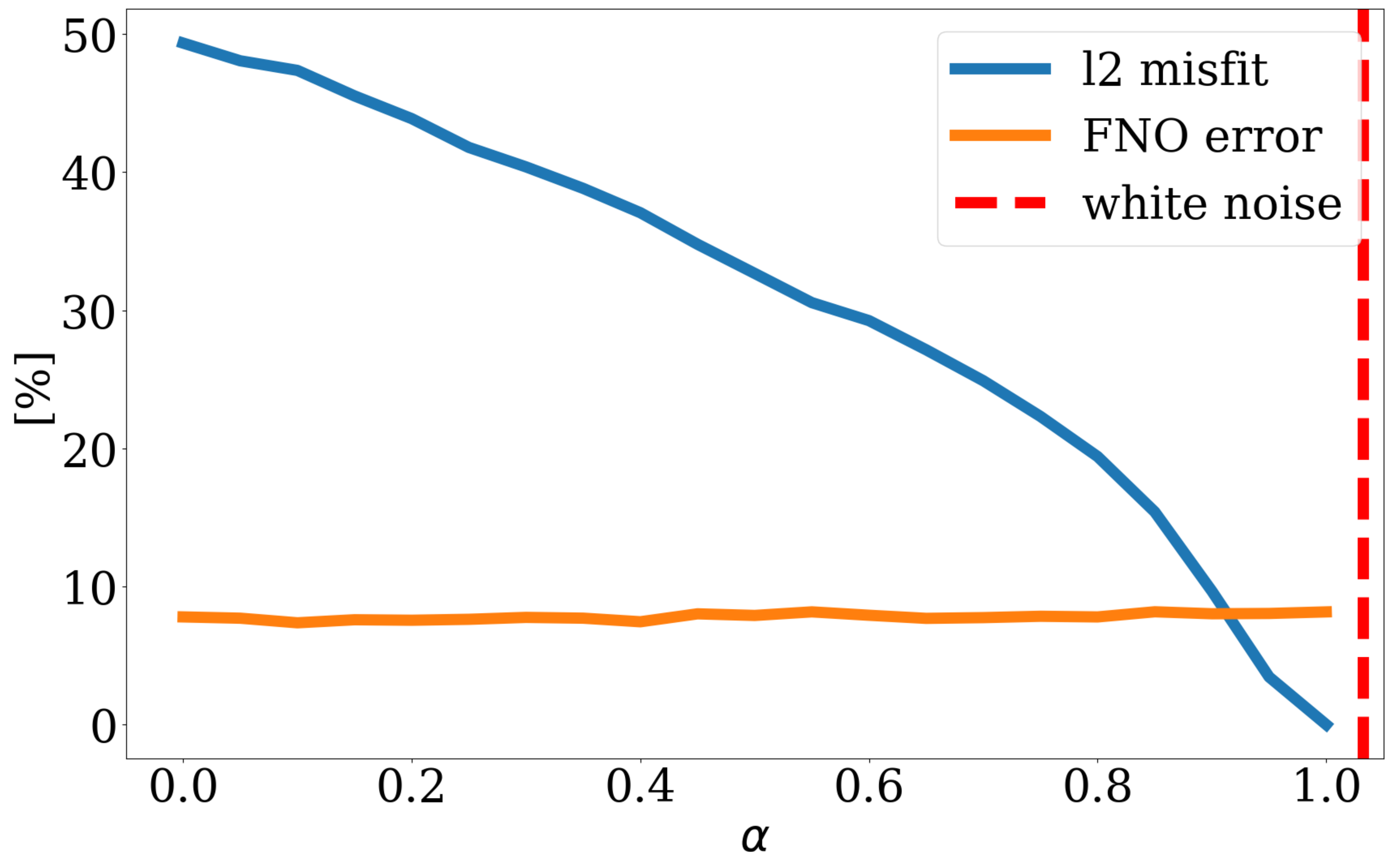


latent space shrinkage transitions from out-of-distribution to in-distribution

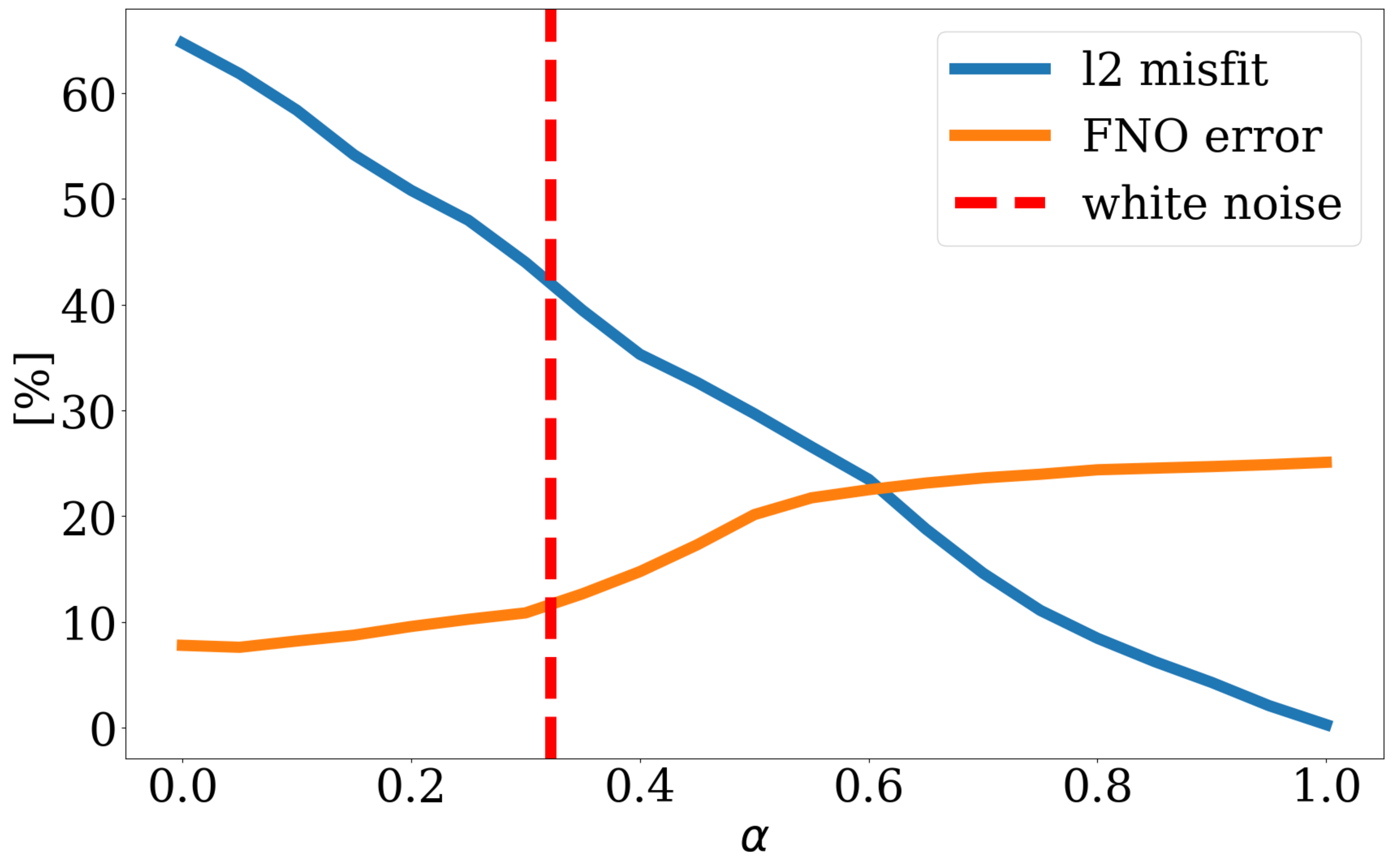
FNO errors during the shrinkage

transitioning from out-of-distribution to in-distribution reduces FNO error

in-distribution



out-of-distribution



Learned inversion algorithm

with learned surrogates (FNOs) and constraints (NFs)

$$\underset{\mathbf{z}}{\text{minimize}} \quad \|\mathbf{d} - \mathcal{H} \circ \mathcal{S}_{\boldsymbol{\theta}*} \circ \mathcal{G}_{\mathbf{w}*}(\mathbf{z})\|_2^2 \quad \text{subject to} \quad \|\mathbf{z}\|_2 \leq \tau$$

Trained FNO $\mathcal{S}_{\boldsymbol{\theta}*}$ replaces numerical simulator $\mathcal{S}$

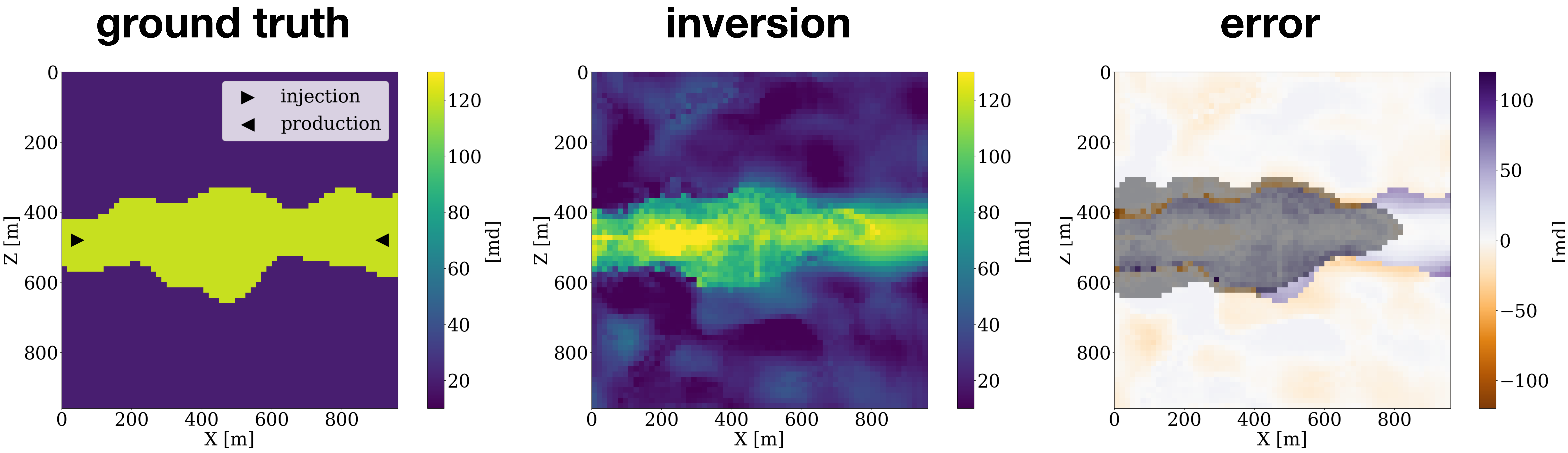
Reparameterize the unknown by trained NF $\mathcal{G}_{\mathbf{w}*}(\mathbf{z})$

τ controls size of the *iteratively relaxed* constraint set

- small τ at the beginning ensures to be in-distribution
- gradually increasing τ brings down the objective

Permeability inversion results
unconstrained inversion with FNO surrogates

$$\underset{\mathbf{K}}{\text{minimize}} \quad \|\mathbf{d} - \mathscr{H} \circ \mathscr{S}_{\boldsymbol{\theta}^*}(\mathbf{K})\|_2^2$$

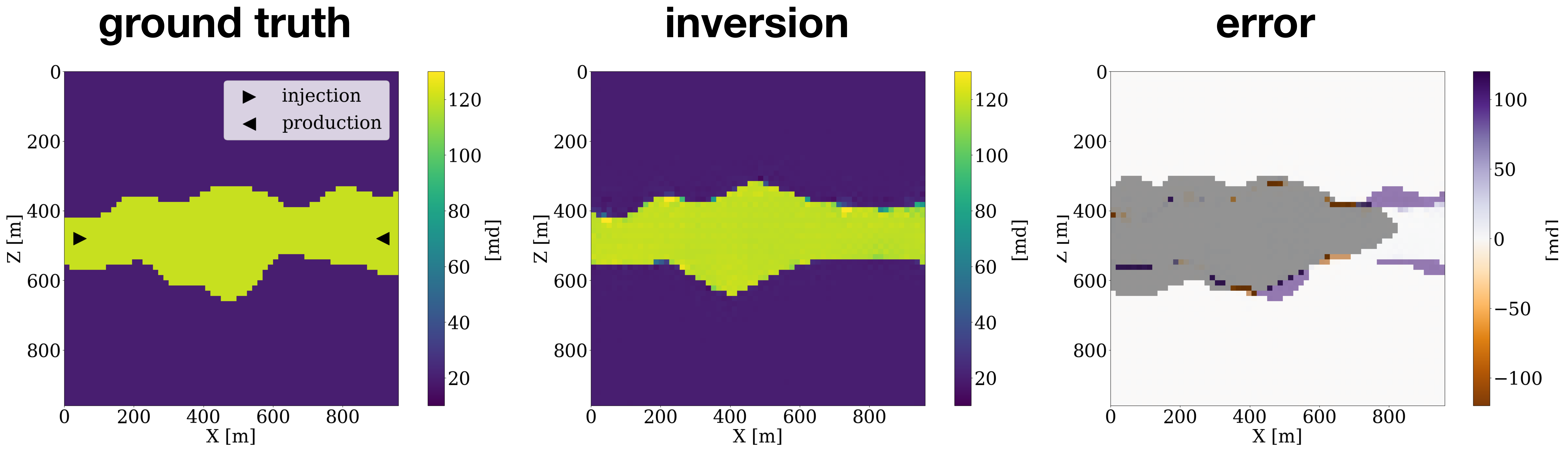


visible artifacts in the recovery

Permeability inversion results
constrained inversion with FNO surrogates

SLIM ML4Seismic

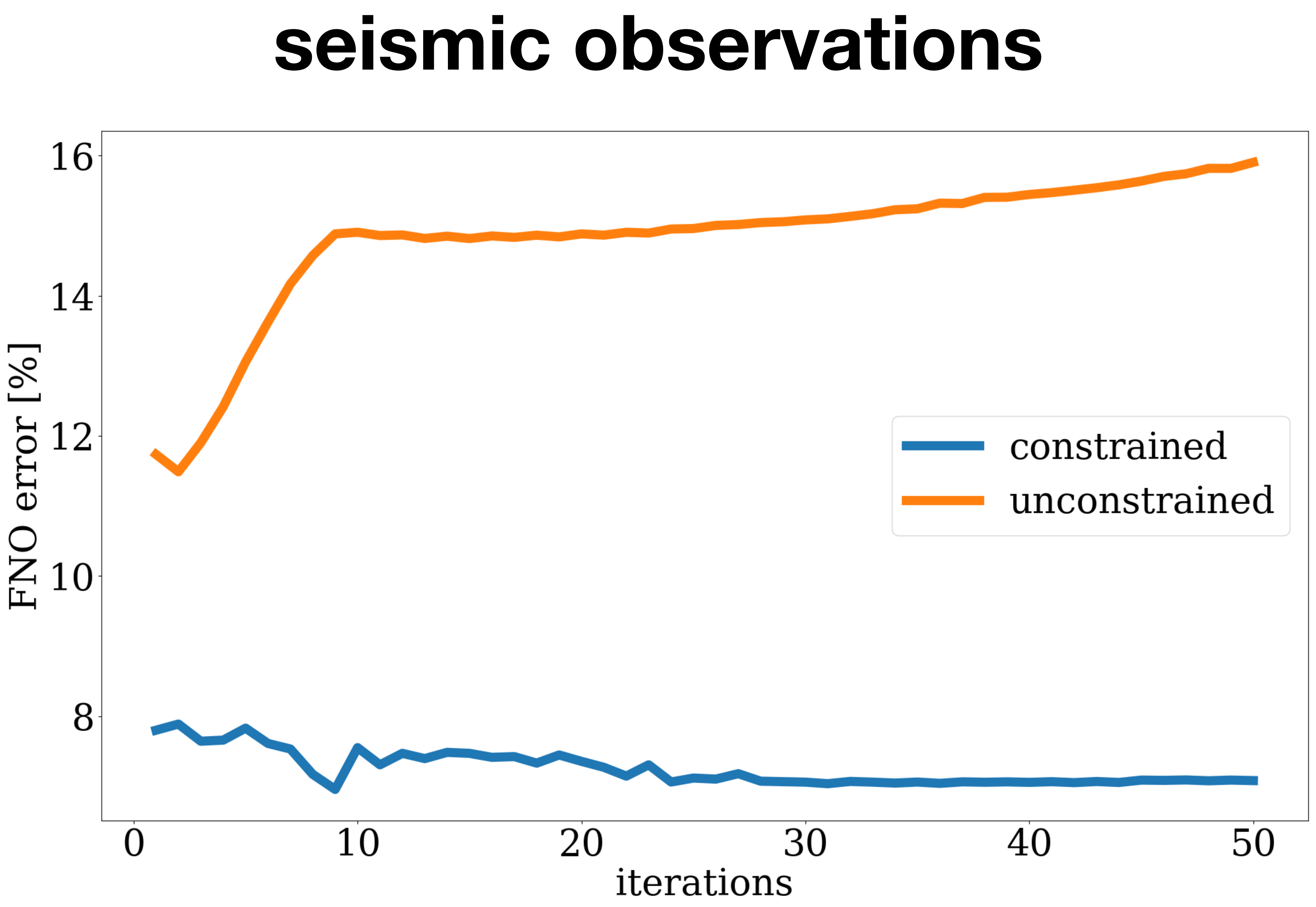
$$\underset{\mathbf{z}}{\text{minimize}} \quad \|\mathbf{d} - \mathcal{H} \circ \mathcal{S}_{\boldsymbol{\theta}^*} \circ \mathcal{G}_{\mathbf{w}^*}(\mathbf{z})\|_2^2 \quad \text{subject to} \quad \|\mathbf{z}\|_2 \leq \tau$$

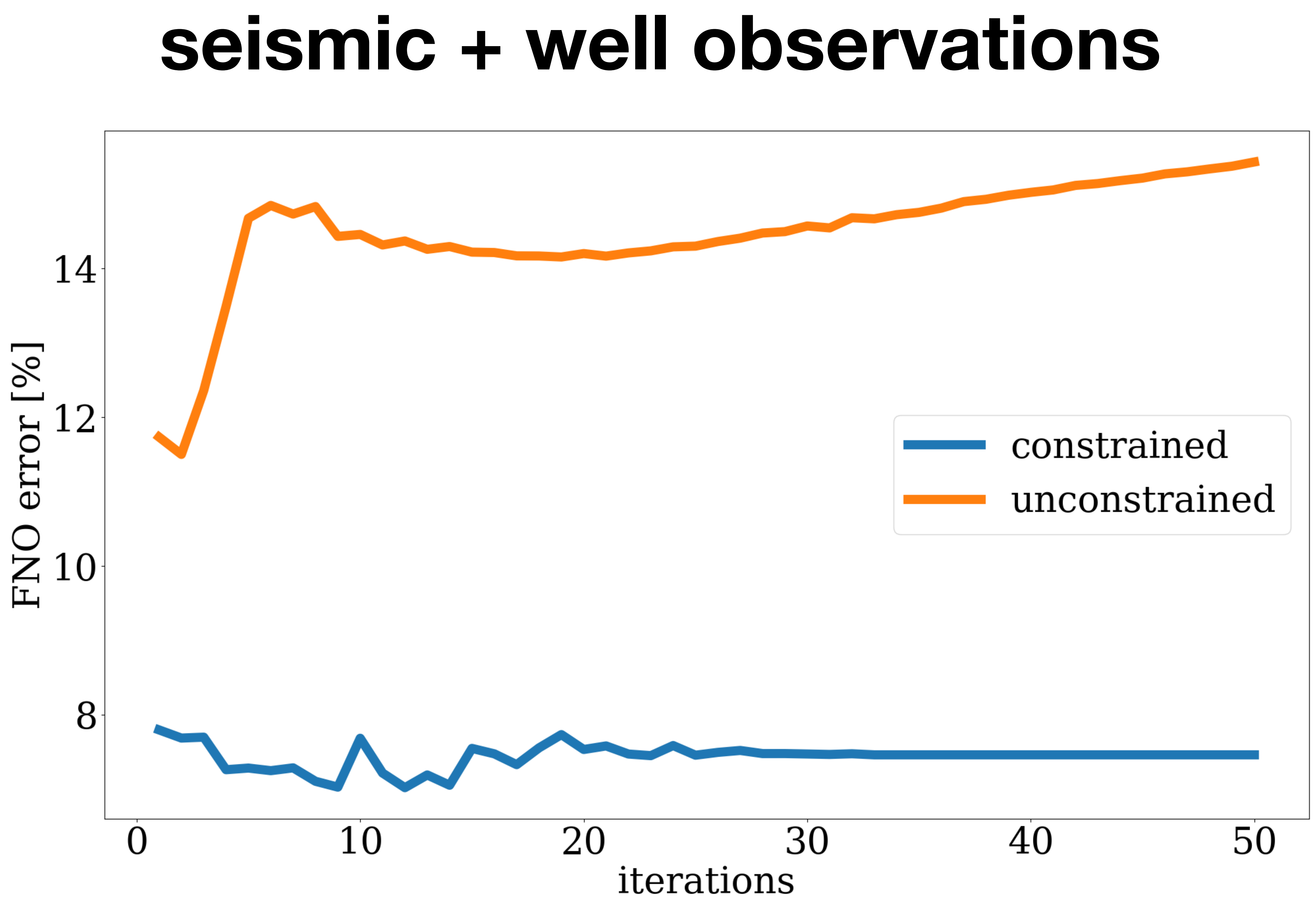


NF constraint greatly improves inversion

FNO error along iterations

constrained vs unconstrained inversion





FNO error remains relatively flatline during constrained inversion

Conclusions & Contributions

Chapter 4

After training FNO & NF on the same samples

- FNO error can be controlled by latent space shrinkage of NF

Propose learned inversion algorithm with FNO & NF

- NF reparameterization forms an efficient continuation scheme / homotopy
- iteratively relaxed constraint
  - safeguard FNO accuracy
  - bring down objective

Proof-of-concept permeability inversion from time-lapse seismic + well data

Ziyi Yin, Huseyin Tuna Erdinc, Abhinav Prakash Gahlot, Mathias Louboutin and F. Herrmann. “Derisking geological carbon storage from high resolution time-lapse seismic to explainable leakage detection.” The Leading Edge, 2023.
Huseyin Tuna Erdinc, Abhinav Prakash Gahlot, Ziyi Yin, Mathias Louboutin and F. Herrmann. “De-risking Carbon Storage and Sequestration with Explainable CO2 Leakage Detection in Time-lapse Seismic Monitoring Images.” AAAI fall symposium, 2022.
Ziyi Yin, Mathias Louboutin, and Felix J. Herrmann. "Compressive time-lapse seismic monitoring of carbon storage and sequestration with the joint recovery model." International Meeting for Applied Geoscience & Energy Expanded Abstracts, 2021.

SLIM
ML4Seismic

Chapter 5

Derisking geologic carbon storage from high-resolution time-lapse seismic to explainable leakage detection

Contributions
Chapter 5

Propose low-cost time-lapse seismic acquisition & imaging

Monitor CO_2 dynamics when it *fails to follow* multiphase flow equations

Deploy the joint recovery model (JRM)

- exploit *shared information* to enhance imaging quality
- reduce reliance on *replicating* source & receiver positions across surveys

Train deep neural classifiers

- automatic leakage detection from time-lapse seismic images
- explainable saliency maps

SLIM
ML4Seismic

SLIM ML4Seismic

Ziyi Yin*, Rafael Orozco*, Mathias Louboutin, and Felix J. Herrmann. “WISE: Full-waveform variational inference via subsurface extensions.” Geophysics, 2024. **(Featured in Geophysics Bright Spot in The Leading Edge)**

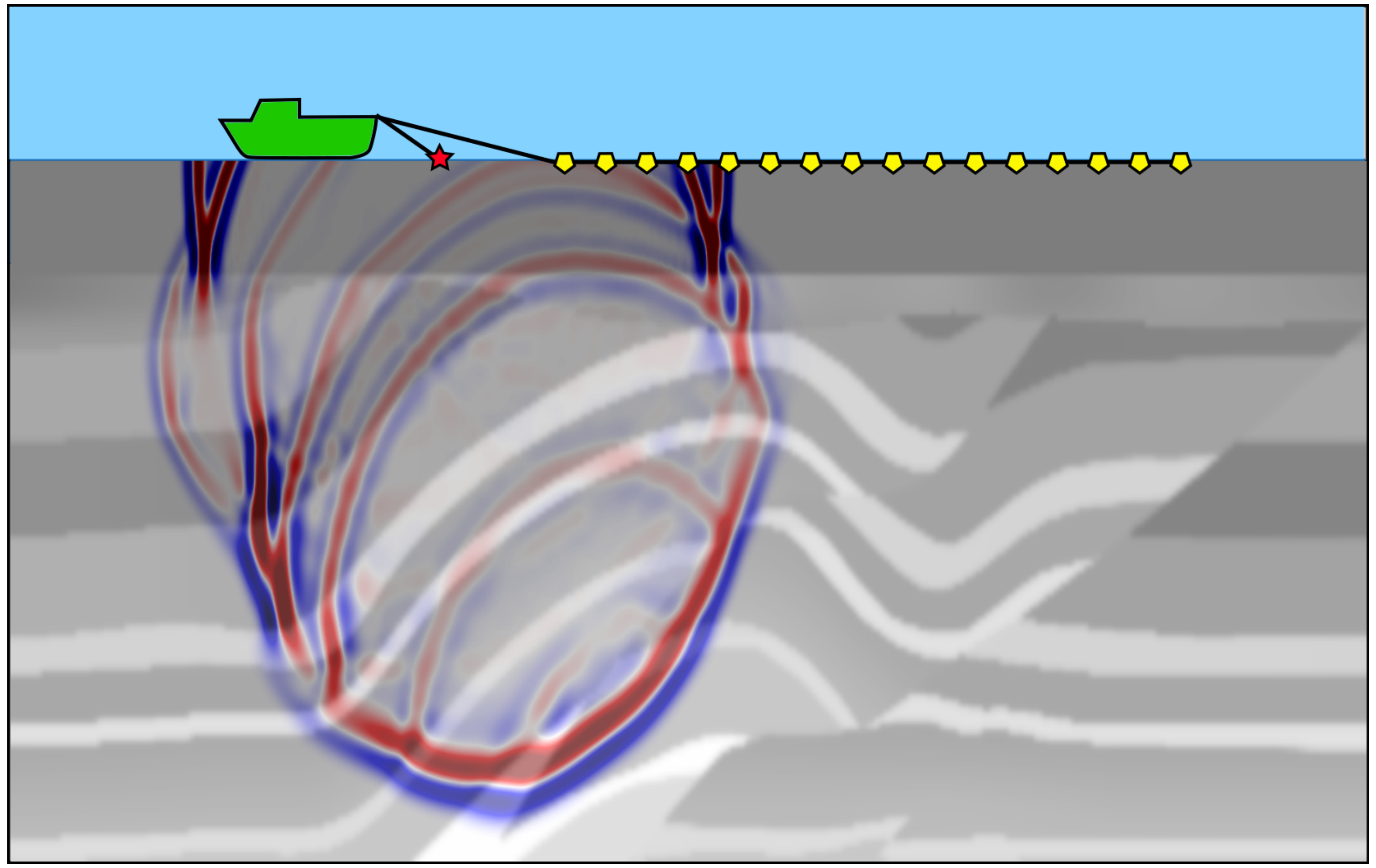
Chapter 6

WISE: full-Waveform variational Inference via Subsurface Extensions

Geophysical exploration

velocity

$\mathbf{x}$



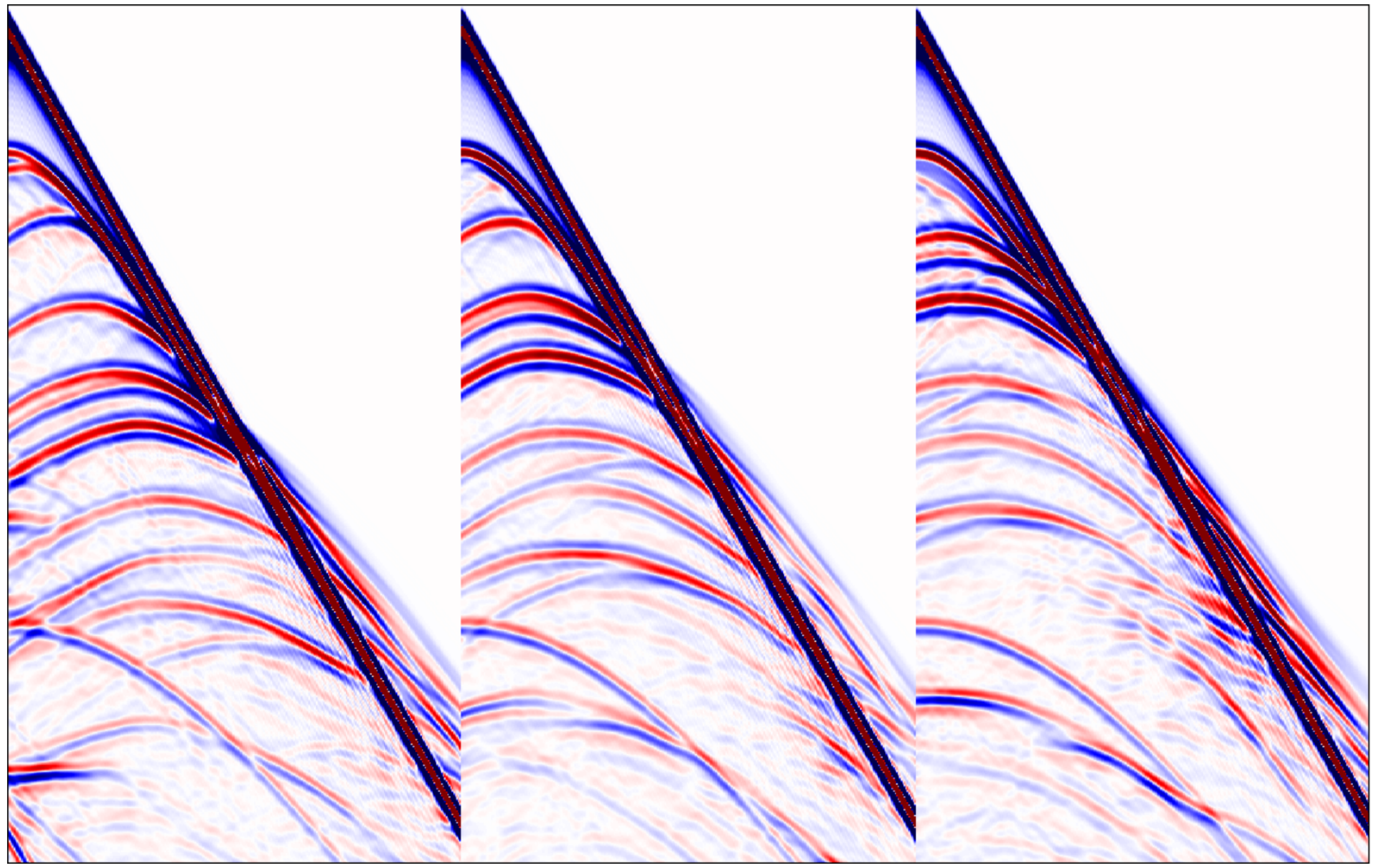
$\mathscr{F}$

$\longrightarrow$

wave physics

seismic data

$\mathbf{y}$



Full-waveform inversion (FWI)

Inverse problems related to PDE parameter estimation

velocity

$\mathbf{x}$

?

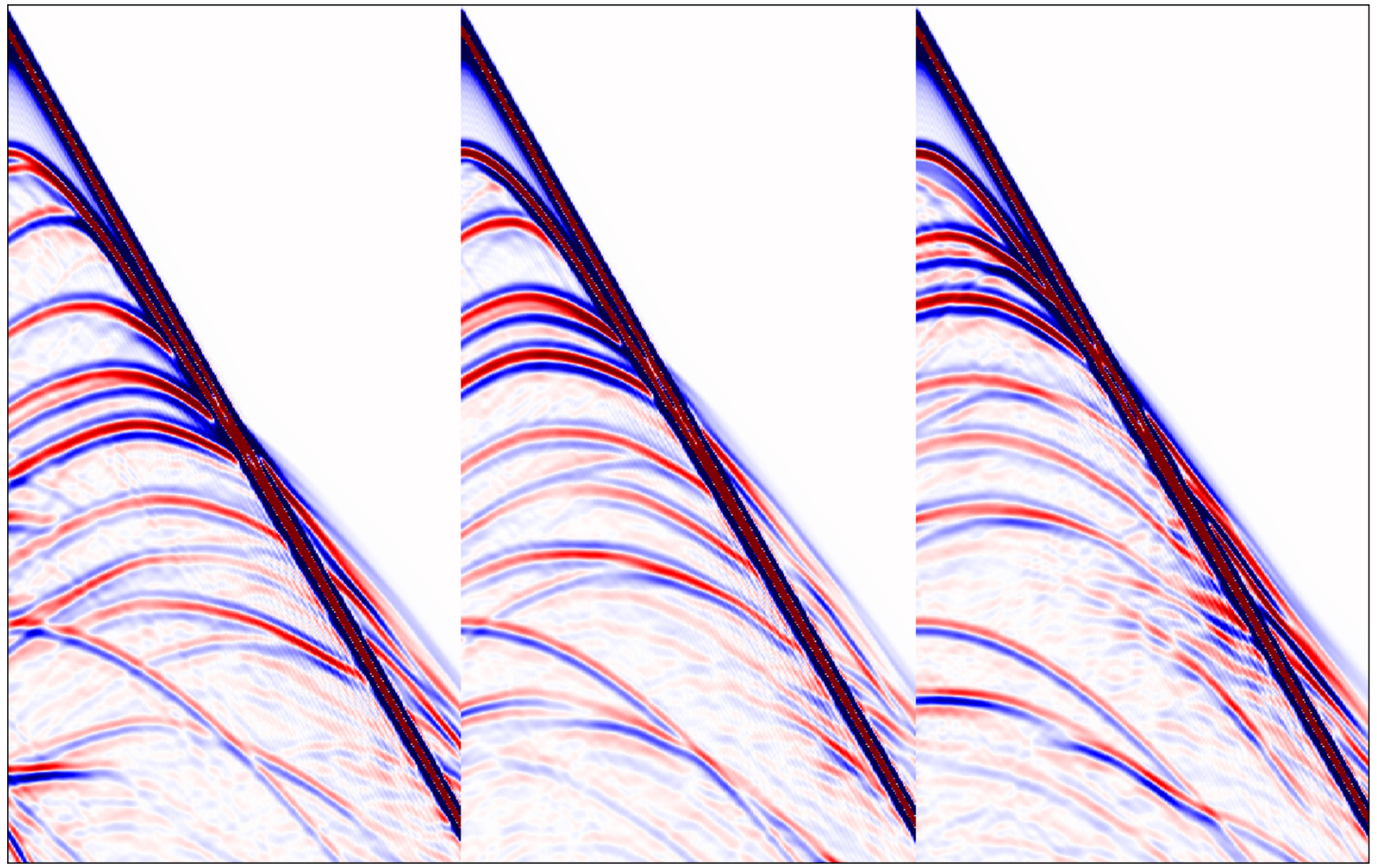
$\mathscr{F}$

$\longrightarrow$

wave physics

seismic data

$\mathbf{y}$



FWI cont'd

$$\mathbf{y} = \mathscr{F}(\mathbf{x}) + \boldsymbol{\epsilon}$$

- $\mathbf{x}$ acoustic velocity (unknown parameter of interest)
- $\mathscr{F}$ nonlinear forward modeling operator
- $\mathbf{y}$ observed seismic data
- $\boldsymbol{\epsilon}$ noise

Bayesian inference
posterior

SLIM
ML4Seismic

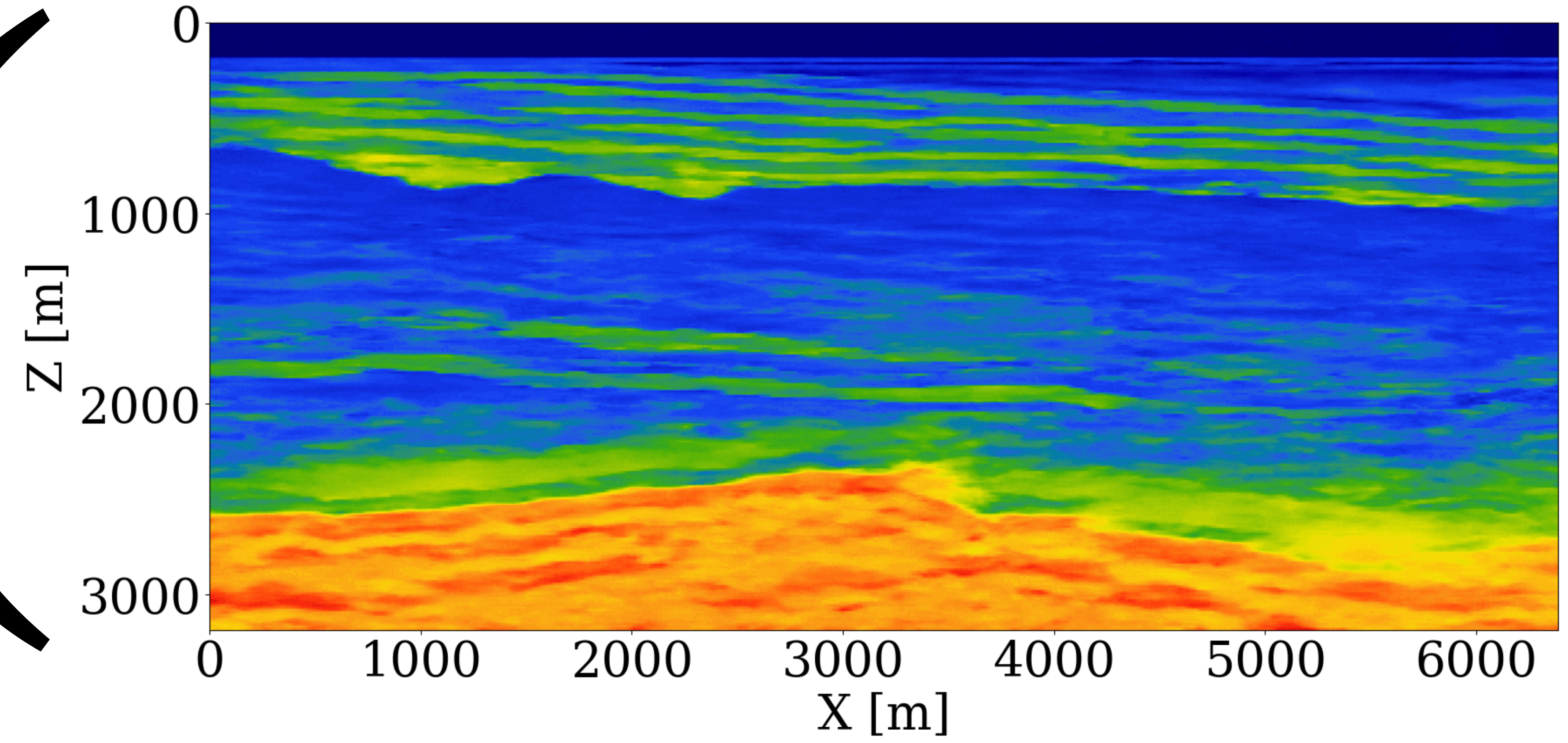
unknown parameter **observed data**

$$p(\ \mathbf{x}\ |\ \mathbf{y}\ )$$

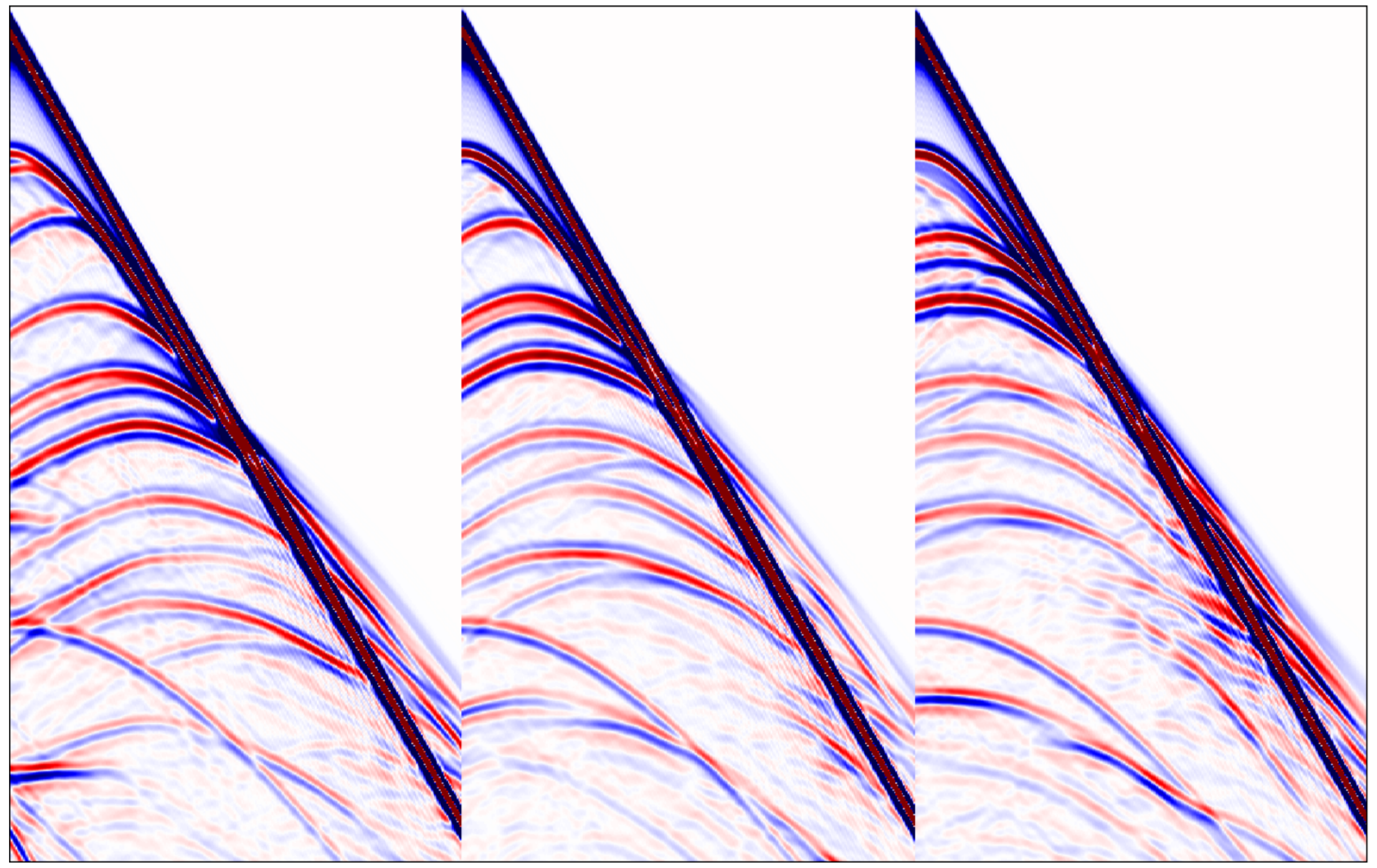
Full-waveform inversion & inference
posterior

velocity model $\mathbf{x}$ | **observed data $\mathbf{y}$**

$p\Big($



$\Big|$



$\Big)$

Radev, Stefan T., Ulf K. Mertens, Andreas Voss, Lynton Ardizzone, and Ullrich Köthe. "BayesFlow: Learning complex stochastic models with invertible neural networks." *IEEE transactions on neural networks and learning systems*, 2020.

SLIM ML4Seismic

Amortized variational inference (VI)

Learn $q_{\boldsymbol{\theta}}(\mathbf{x} \mid \mathbf{y}) \approx p(\mathbf{x} \mid \mathbf{y})$ via sample pairs $\{\mathbf{x}^{(i)}, \mathbf{y}^{(i)}\}_{i=1}^{N}$

Train conditional normalizing flows (CNFs)

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \quad \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{2}\|f_{\boldsymbol{\theta}}\left(\mathbf{x}^{(i)}; \mathbf{y}^{(i)}\right)\|_2^2 - \log\left|\det \mathbf{J}_{f_{\boldsymbol{\theta}}}\right|\right)$$

- p unknown target posterior distribution
- $q_{\boldsymbol{\theta}}$ approximated posterior distribution via CNFs $f_{\boldsymbol{\theta}}$
- expensive offline training
- cheap online inference

Radev, Stefan T., Ulf K. Mertens, Andreas Voss, Lynton Ardizzone, and Ullrich Köthe. "BayesFlow: Learning complex stochastic models with invertible neural networks." *IEEE transactions on neural networks and learning systems*, 2020.
Rafael Orozco, Ali Siahkoohi, Gabrio Rizzuti, Tristan van Leeuwen, and Felix J. Herrmann. "Adjoint operators enable fast and amortized machine learning based Bayesian uncertainty quantification." In *Medical Imaging 2023: Image Processing*, vol. 12464, pp. 357-367. SPIE, 2023.

Challenges

VI w/ CNFs

Practical challenges of training CNFs to approximate $p(\mathbf{x} \mid \mathbf{y})$

- need to *retrain* for new configurations (e.g., source/receiver positions)
- mapping between image $\mathbf{x}$ and data $\mathbf{y}$ is very *difficult to learn*
- does not incorporate any *physics* during training & inference

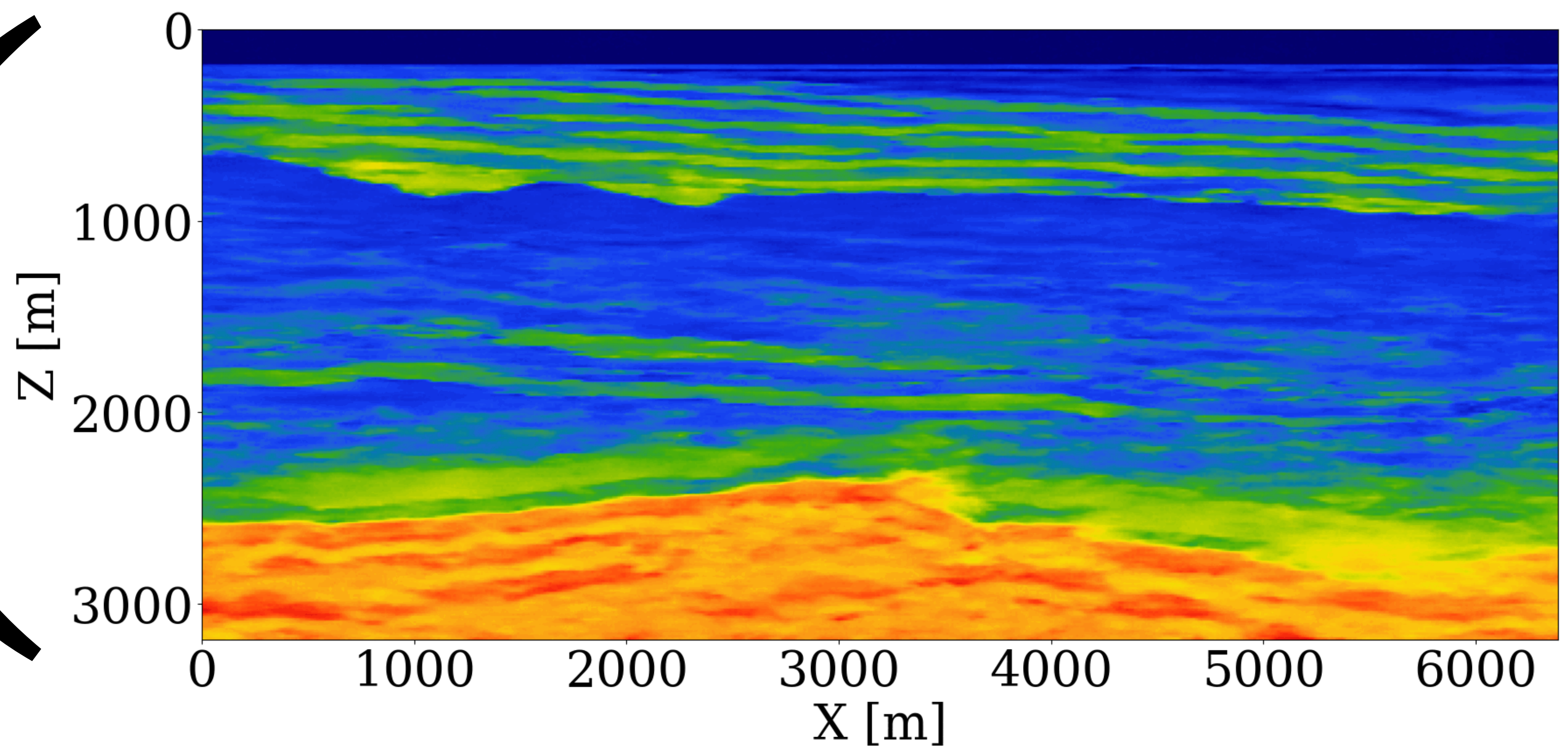
Current literature suggests

- **physics-informed summary statistics**
- partially *decode* the wave physics

Full-waveform inference
approximated posterior

SLIM ML4Seismic

velocity model $\mathbf{x}$ summary statistics $\overline{\mathbf{y}}$

$p($ 

$|$? $)$

Summary statistics need to
- preserve all information in data
- decode the complicated wave physics

Jie Hou, and William W. Symes. "An approximate inverse to the extended Born modeling operator." *Geophysics, 2015.*
Jie Hou, and William W. Symes. "Accelerating extended least-squares migration with weighted conjugate gradient iteration." *Geophysics*, 2016.
Jie Hou, and William W. Symes. "Inversion velocity analysis in the subsurface-offset domain." *Geophysics, 2018.*

Motivation

model extension & extended gradients

Orozco et al proved for linear inverse problems

- $\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}, \mathbf{I})$
- $p(\mathbf{x}\,|\,\mathbf{y}) \equiv p(\mathbf{x}\,|\,\overline{\mathbf{y}})$ where $\overline{\mathbf{y}} = \mathbf{A}^\top\mathbf{y}$

Linearize FWI problem at velocity $\mathbf{x}_0$

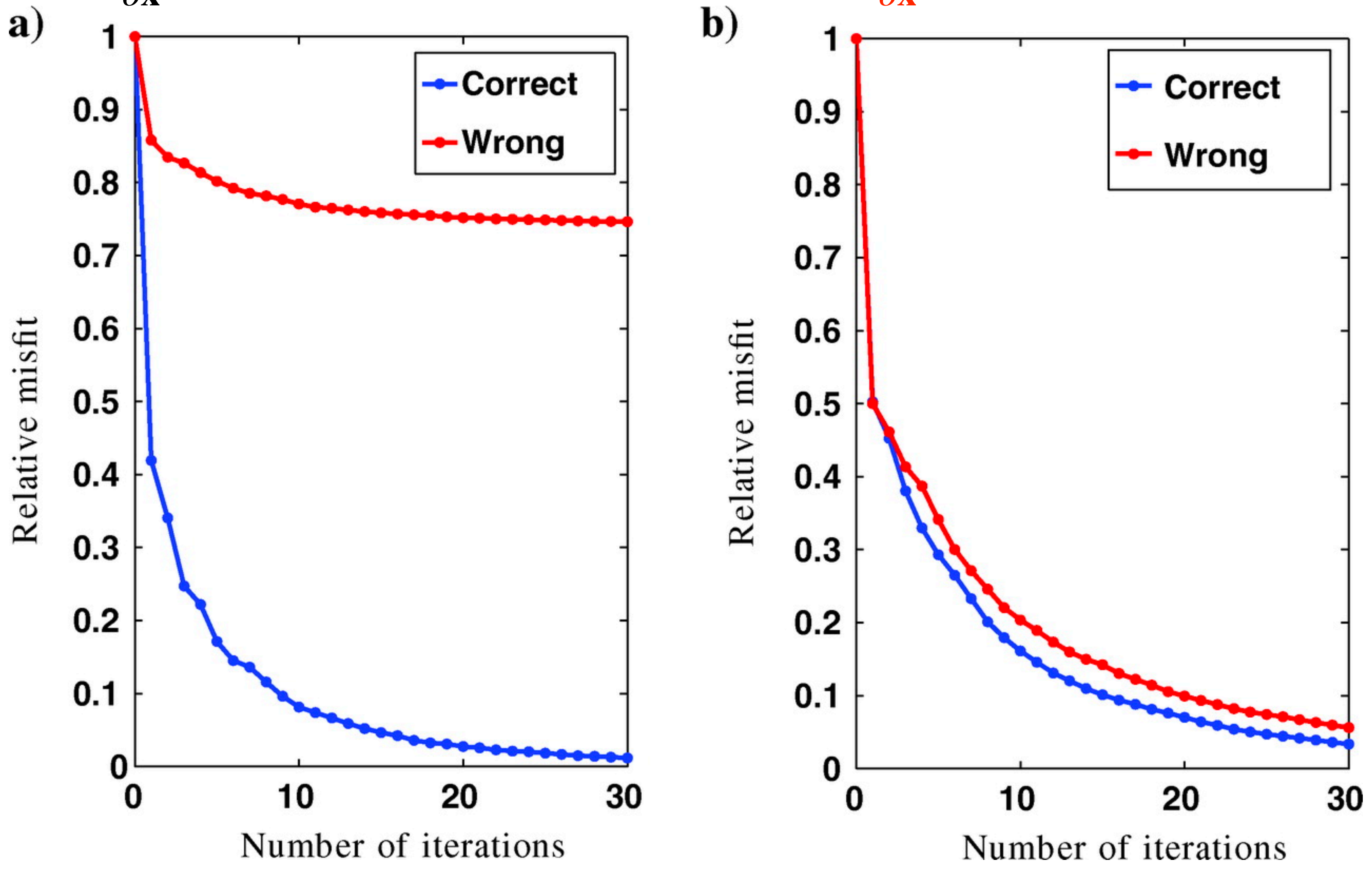
- $\mathscr{F}(\mathbf{x}) \approx \mathscr{F}(\mathbf{x}_0) + \nabla\mathscr{F}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$

Consider Gauss-Newton update at a bad linearization point $\mathbf{x}_0$

- standard Jacobian can't drive residual to 0, "*information is lost*"
- extended Jacobian can preserve information

standard $\quad\quad$ **extended**

$$\min_{\delta\mathbf{x}} \|\delta\mathbf{y} - \nabla\mathscr{F}(\mathbf{x}_0)\delta\mathbf{x}\|_2^2 \qquad \min_{\overline{\delta\mathbf{x}}} \|\delta\mathbf{y} - \overline{\nabla\mathscr{F}}(\mathbf{x}_0)\overline{\delta\mathbf{x}}\|_2^2$$



Correct: $\mathbf{x}_0$ is close to $\mathbf{x}$

Wrong: $\mathbf{x}_0$ is far from $\mathbf{x}$

Louboutin, Mathias, and Felix J. Herrmann. "Wave-based inversion at scale on graphical processing units with randomized trace estimation." Geophysical Prospecting 72.2 (2024)
Kroode, F. ten, 2023, An omnidirectional seismic image extension: Inverse Problems, 39

Model extension cont'd

extended gradient / common-image gathers

Standard gradient

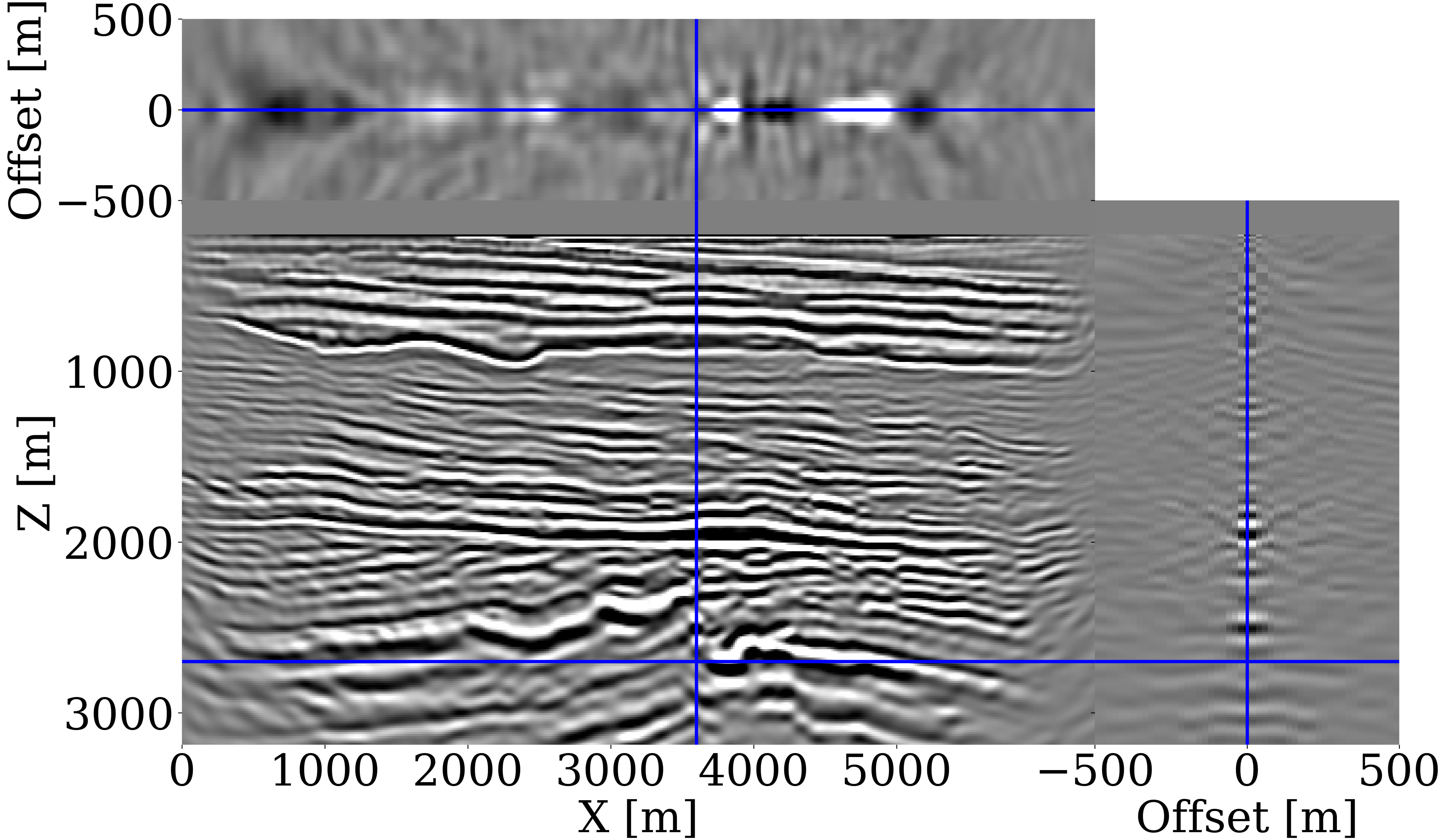
- $\mathbf{g}[\vec{x}] = \nabla \mathscr{F}(\mathbf{x}_0)^\top \delta \mathbf{y} = \sum_{i=1}^{n_s} \sum_{t=1}^{n_t} \ddot{\mathbf{u}}_i[\vec{x}, t] \odot \mathbf{v}_i[\vec{x}, t]$
- $\ddot{\mathbf{u}}[\vec{x}, t]$ second-time derivative solution of wave equation: $\mathbf{A}(\mathbf{x}_0)\mathbf{u}_i = \mathbf{q}_i$
- $\mathbf{v}[\vec{x}, t]$ solution of adjoint wave equation: $\mathbf{A}(\mathbf{x}_0)^\top \mathbf{v}_i = \mathbf{P}_r^\top \delta \mathbf{y}_i$

Extended gradient (with an extra *subsurface-offset* dimension)

- $\overline{\mathbf{g}}[\vec{x}, \vec{h}] = \overline{\nabla \mathscr{F}}(\mathbf{x}_0)^\top \delta \mathbf{y} = \sum_{i=1}^{n_s} \sum_{t=1}^{n_t} \ddot{\mathbf{u}}_i[\vec{x} + \vec{h}, t] \odot \mathbf{v}_i[\vec{x} - \vec{h}, t]$
- note: $\mathbf{g}[\vec{x}] = \overline{\mathbf{g}}[\vec{x}, \vec{0}]$
- near *isometry* & acts as an *embedding*

Extended gradient

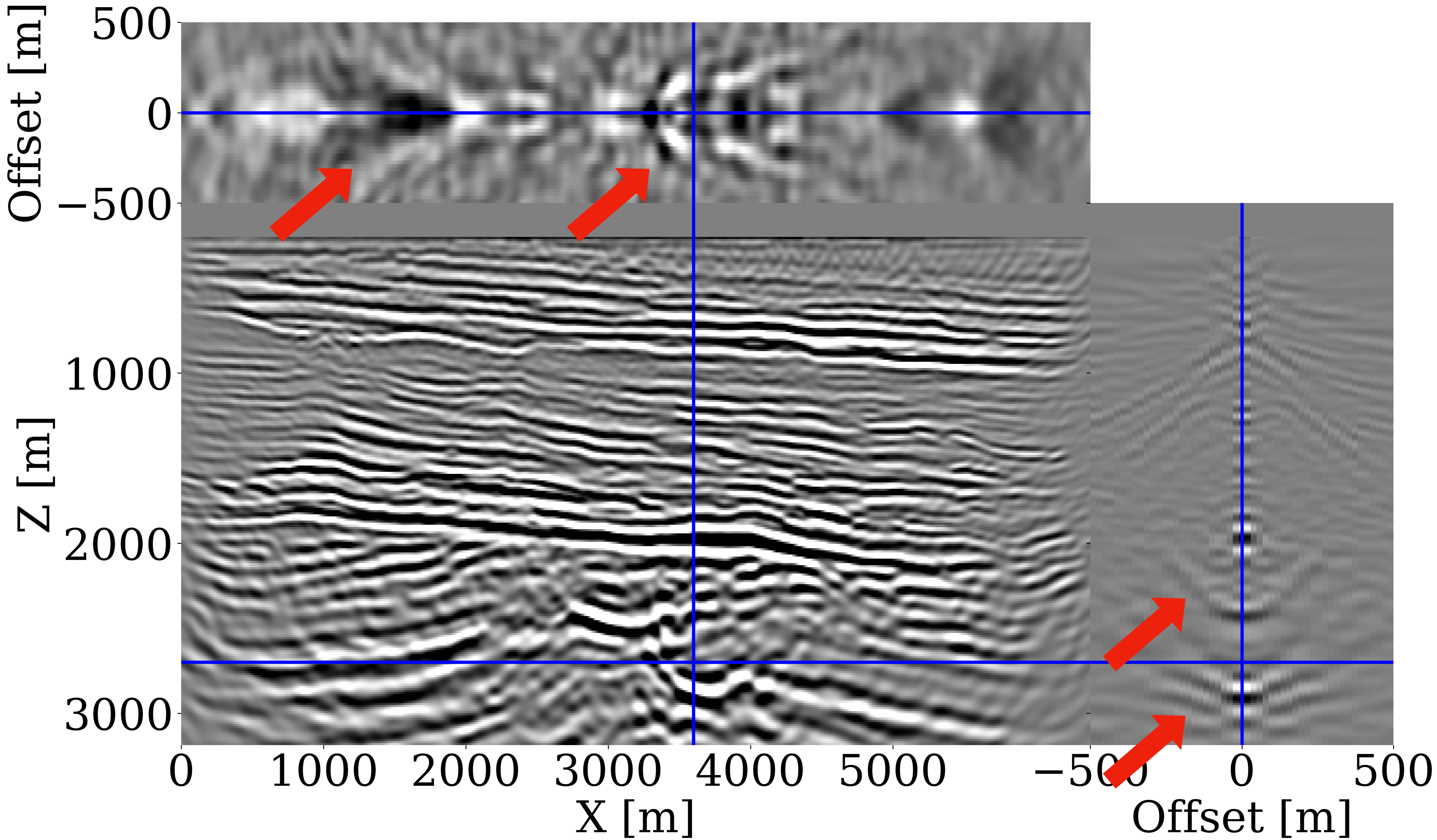
good $\mathbf{x}_0$



Extended gradient

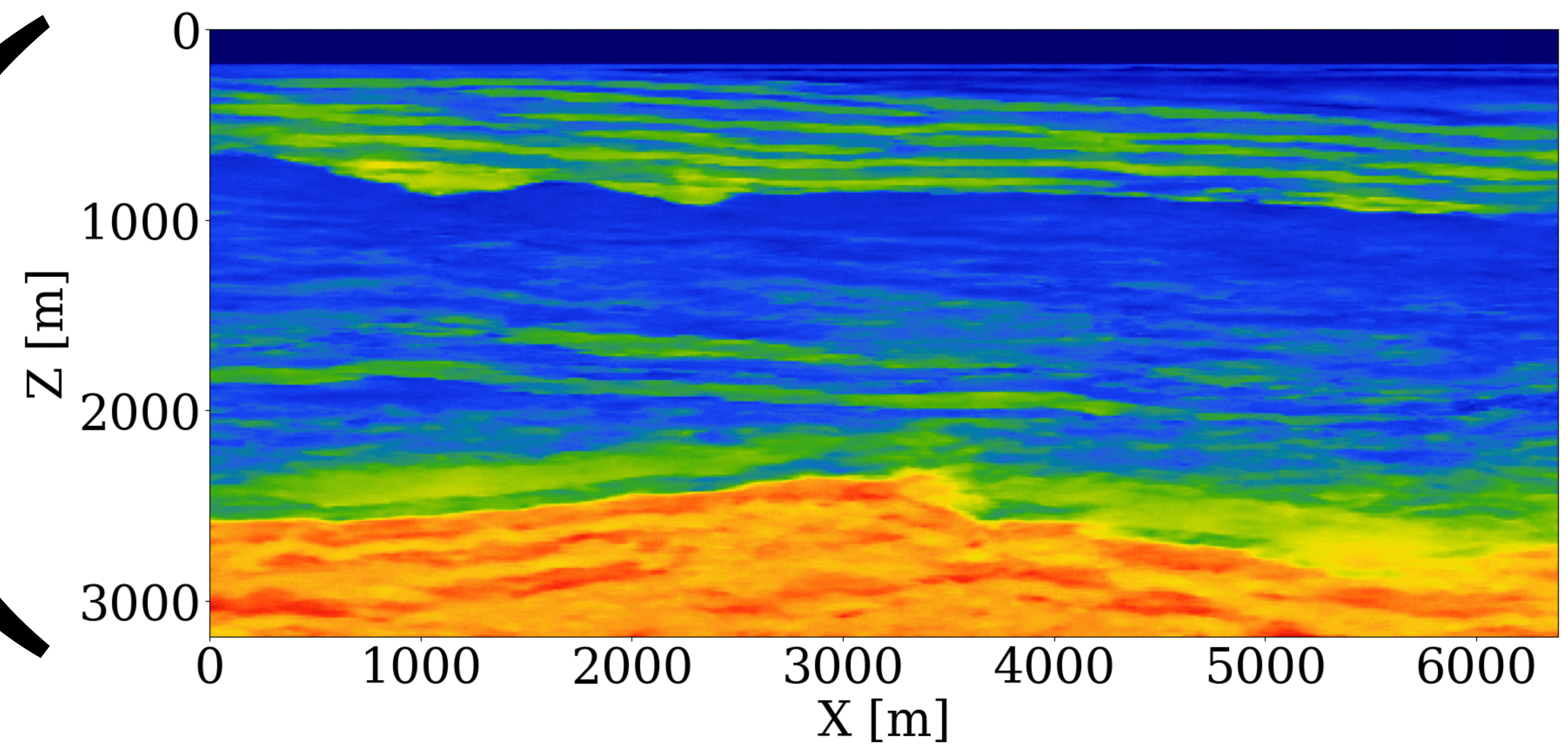
poor $\mathbf{x}_0$

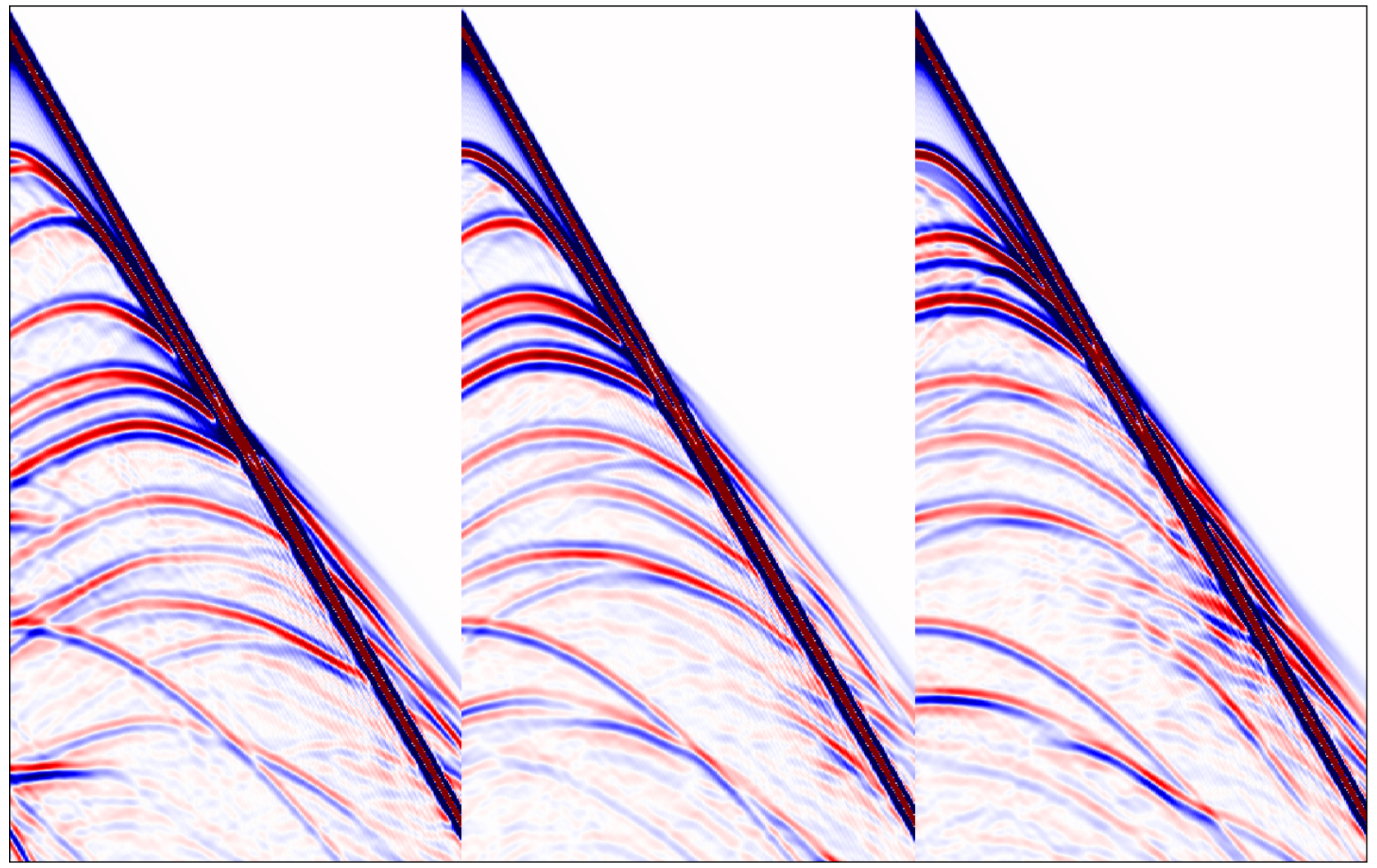
preserve information at non-zero offsets



Full-waveform inference
posterior

velocity model $\mathbf{x}$ | **observed data $\mathbf{y}$**

$p($ 

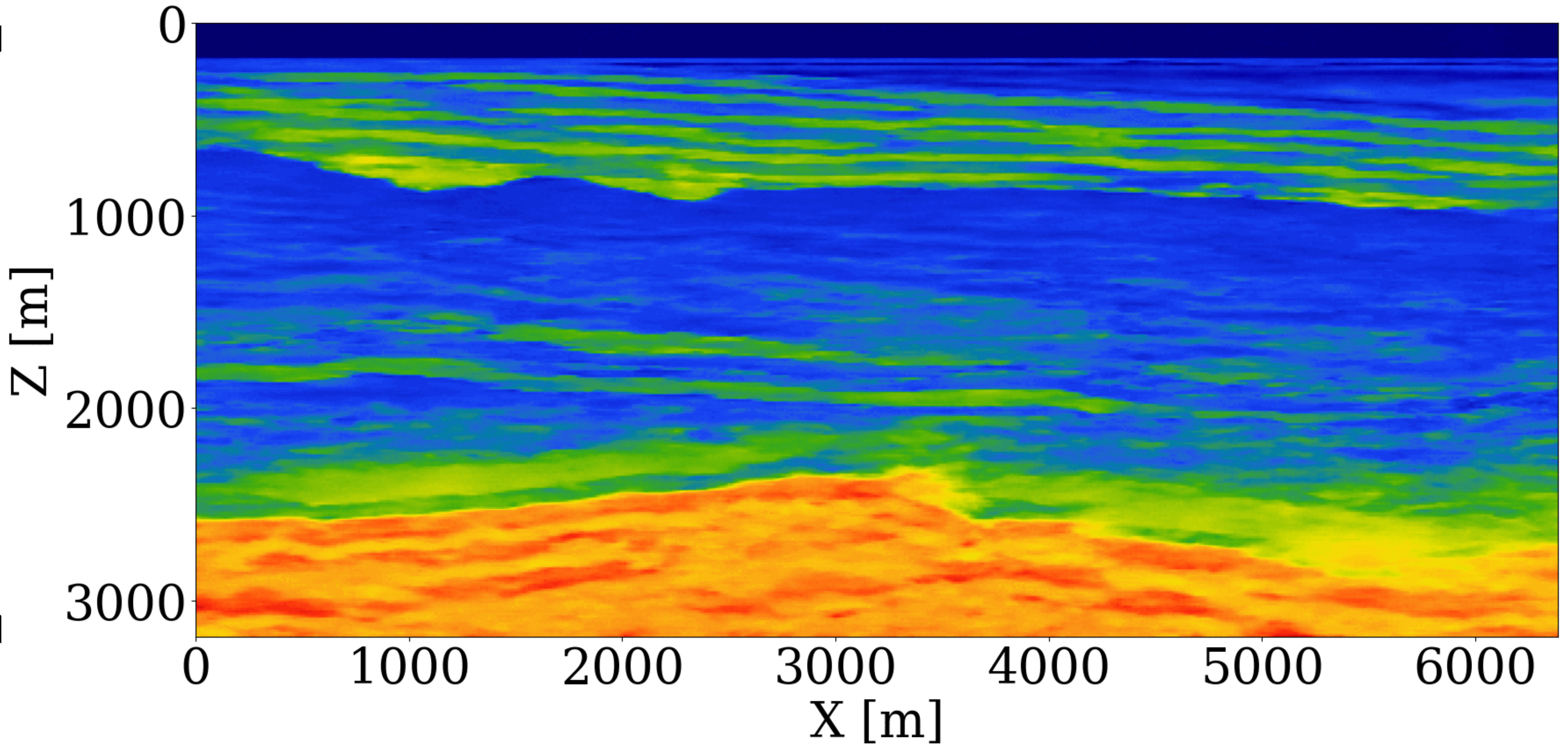
$|$  $)$

Mapping between velocity model and data is very difficult to learn

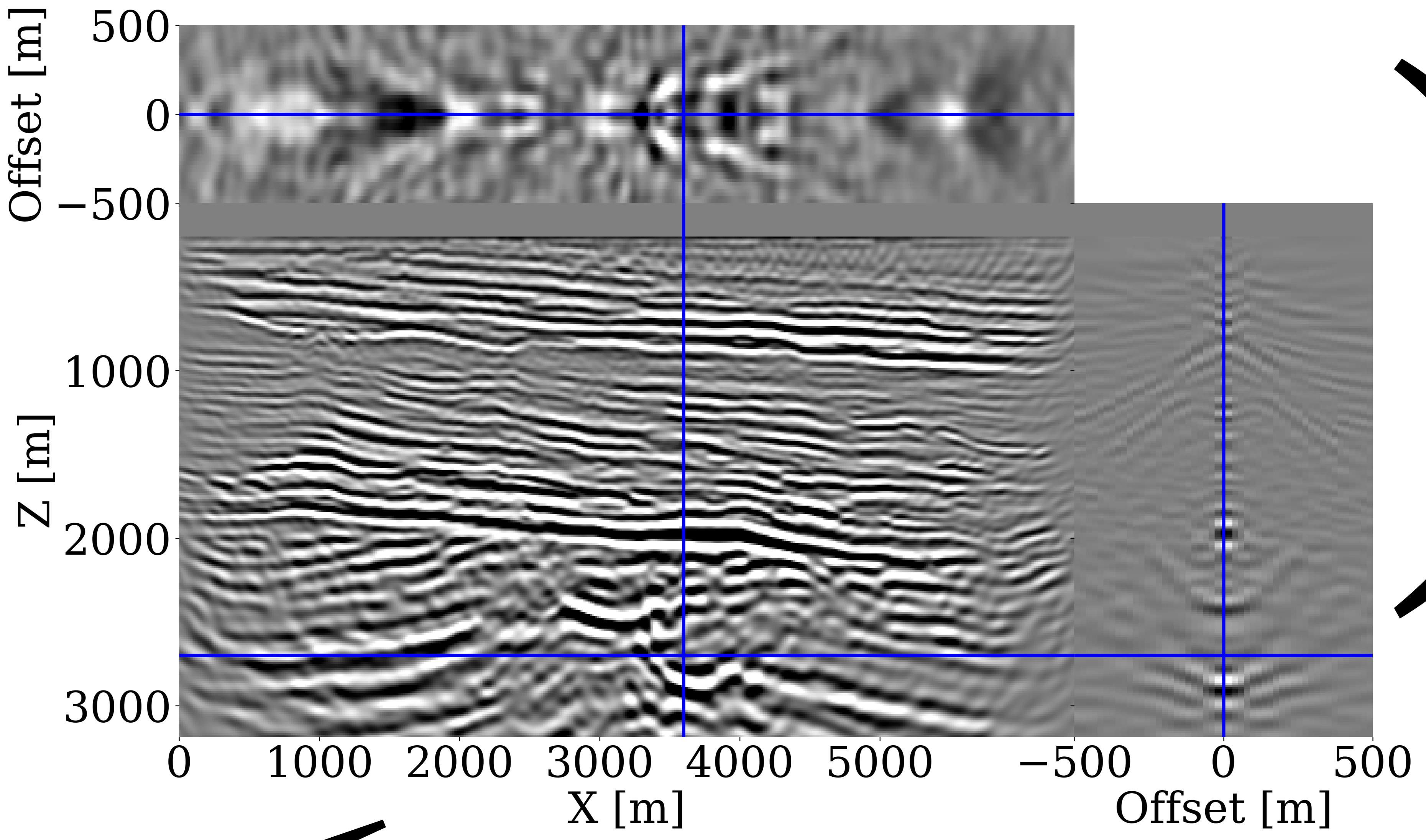
Full-waveform inference
summary statistics = extended gradient

velocity model $\mathbf{x}$ **extended gradient $\overline{\mathbf{y}}$**

$p($



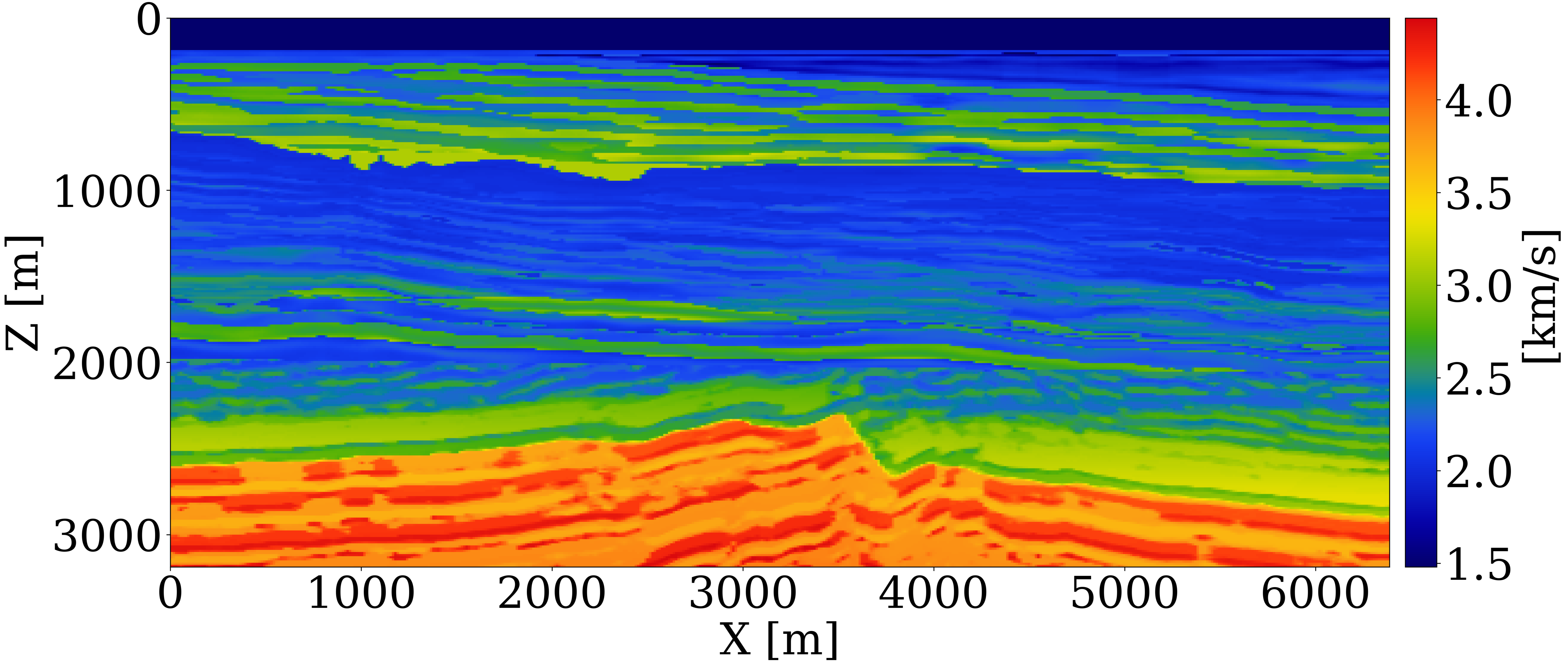
$|$



$)$

decode wave physics &
preserve information

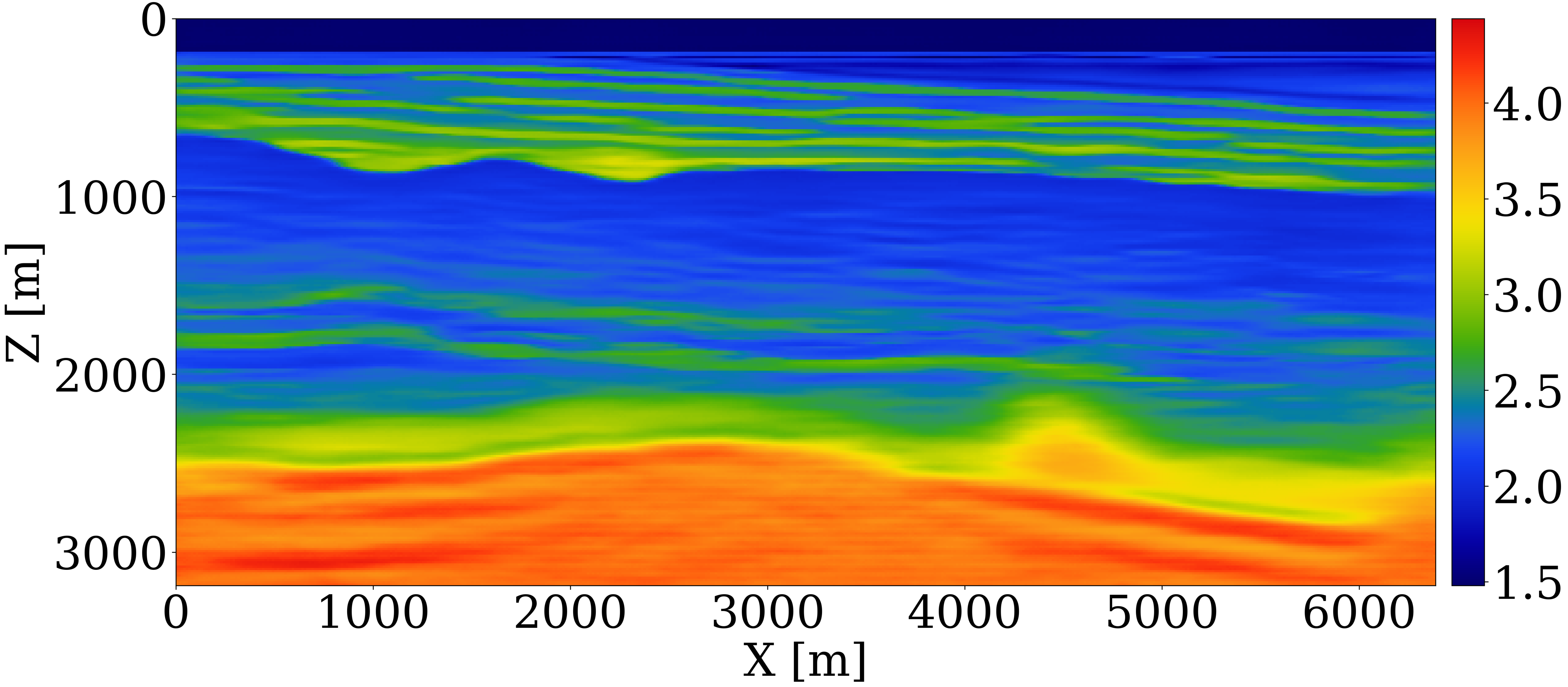
Unseen ground truth velocity



SLIM
ML4Seismic

Conditional mean estimate

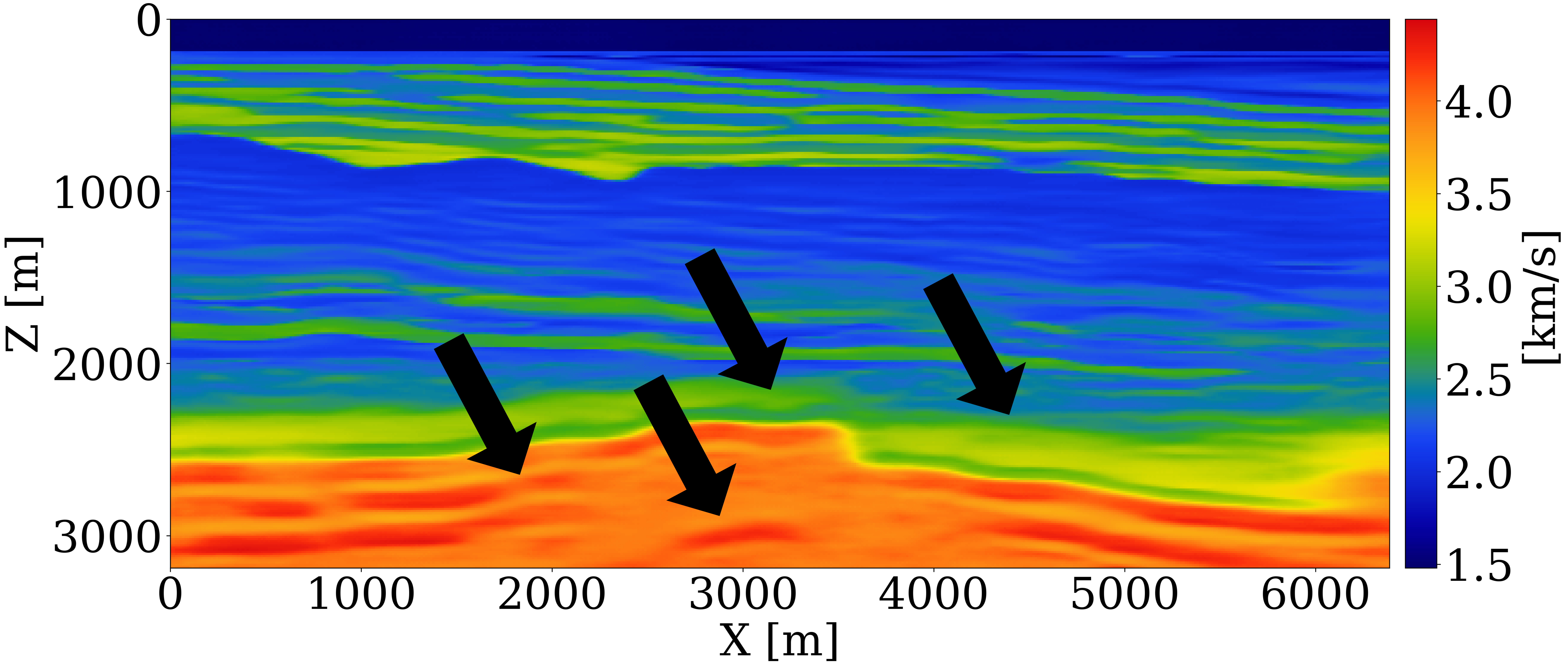
summary statistics = standard gradient



SLIM
ML4Seismic

Conditional mean estimate

summary statistics = extended gradient



Downstream task
high-frequency imaging

Uncertainty in imaged reflectivities entails important information to make business decisions

Amplitude variations
point-wise standard deviation

SLIM
ML4Seismic

Positioning variations
maximum vertical shift via cross correlation

Contributions
Chapter 6

Propose physics-informed summary statistics for *nonlinear FWI* problem

- based on *model extension* and geophysical knowledge
- reduce reliance on *accurate* initial model
- preserve information
- enhance CNF training

Perform *forward UQ* for downstream imaging tasks

Ziyi Yin, Rafael Orozco, and Felix J. Herrmann. “WISER: multimodal variational inference for full-waveform inversion without dimensionality reduction” ArXiv, 2024.

SLIM
ML4Seismic

Chapter 7

WISER: multimodal variational inference for full-waveform inversion without dimensionality reduction

Marino, Joe, Yisong Yue, and Stephan Mandt. 2018. "Iterative Amortized Inference." In *International Conference on Machine Learning*, 3403–12. PMLR.

WISER = WISE + Refinements

based on wave physics

Challenges

- *amortization gap*
  - network works well for a family of observations
  - but does not provide very accurate prediction for a single observation
- out of distribution at inference

Solution

- fine-tune network via a few physics-based iterations

Ali Siahkoohi, Gabrio Rizzuti, Rafael Orozco, and Felix J. Herrmann. "Reliable amortized variational inference with physics-based latent distribution correction." *Geophysics* 88, no. 3 (2023): R297-R322.

SLIM ML4Seismic

Physics-based latent space correction

constrained formulation

$$\underset{\boldsymbol{\phi}}{\text{minimize}} \quad \mathbb{KL}\left(p\left(h_{\boldsymbol{\phi}}(\mathbf{z})\right) || p_{\text{post}}(\mathbf{z} \mid \overline{\mathbf{y}}_{\text{obs}})\right)$$

$$= \mathbb{E}_{\mathbf{z}\sim\mathrm{N}(\mathbf{0},\mathbf{I})}\left[\frac{1}{2\sigma^2}\|\mathscr{F}\circ f_{\boldsymbol{\theta}^*}^{-1}\left(h_{\boldsymbol{\phi}}(\mathbf{z});\overline{\mathbf{y}}_{\text{obs}}\right)-\mathbf{y}_{\text{obs}}\|_2^2+\frac{1}{2}\|h_{\boldsymbol{\phi}}(\mathbf{z})\|_2^2-\log\left|\det \mathbf{J}_{h_{\boldsymbol{\phi}}}\right|\right].$$

likelihood **prior**

$f_{\boldsymbol{\theta}}$ trained *amortized* CNF from WISE, $h_{\boldsymbol{\phi}}$ refined *non-amortized* NF

Challenge: physics $\mathscr{F}$ (expensive) and networks $f_{\boldsymbol{\theta}^*}$, $h_{\boldsymbol{\phi}}$ are always coupled

Solution: decouple them via *weak* formulation

Siahkoohi, Ali, Gabrio Rizzuti, and Felix J. Herrmann. "Weak deep priors for seismic imaging." *SEG Technical Program Expanded Abstracts 2020*. Society of Exploration Geophysicists, 2020. 2998-3002.

SLIM ML4Seismic

Proposed WISER objective

weak formulation

$$\underset{\mathbf{x}_{1:M},\boldsymbol{\phi}}{\text{minimize}} \quad \frac{1}{M}\sum_{i=1}^{M}\left[\frac{1}{2\sigma^2}\|\mathcal{F}(\mathbf{x}_i)-\mathbf{y}_{\text{obs}}\|_2^2+\frac{1}{2\gamma^2}\|\mathbf{x}_i-f_{\boldsymbol{\theta}^*}^{-1}\left(h_{\boldsymbol{\phi}}\left(\mathbf{z}_i\right);\bar{\mathbf{y}}_{\text{obs}}\right)\|_2^2+\frac{1}{2}\|h_{\boldsymbol{\phi}}\left(\mathbf{z}_i\right)\|_2^2-\log\left|\det\mathbf{J}_{h_{\boldsymbol{\phi}}}\right|\right]$$

likelihood **weak prior** **prior**

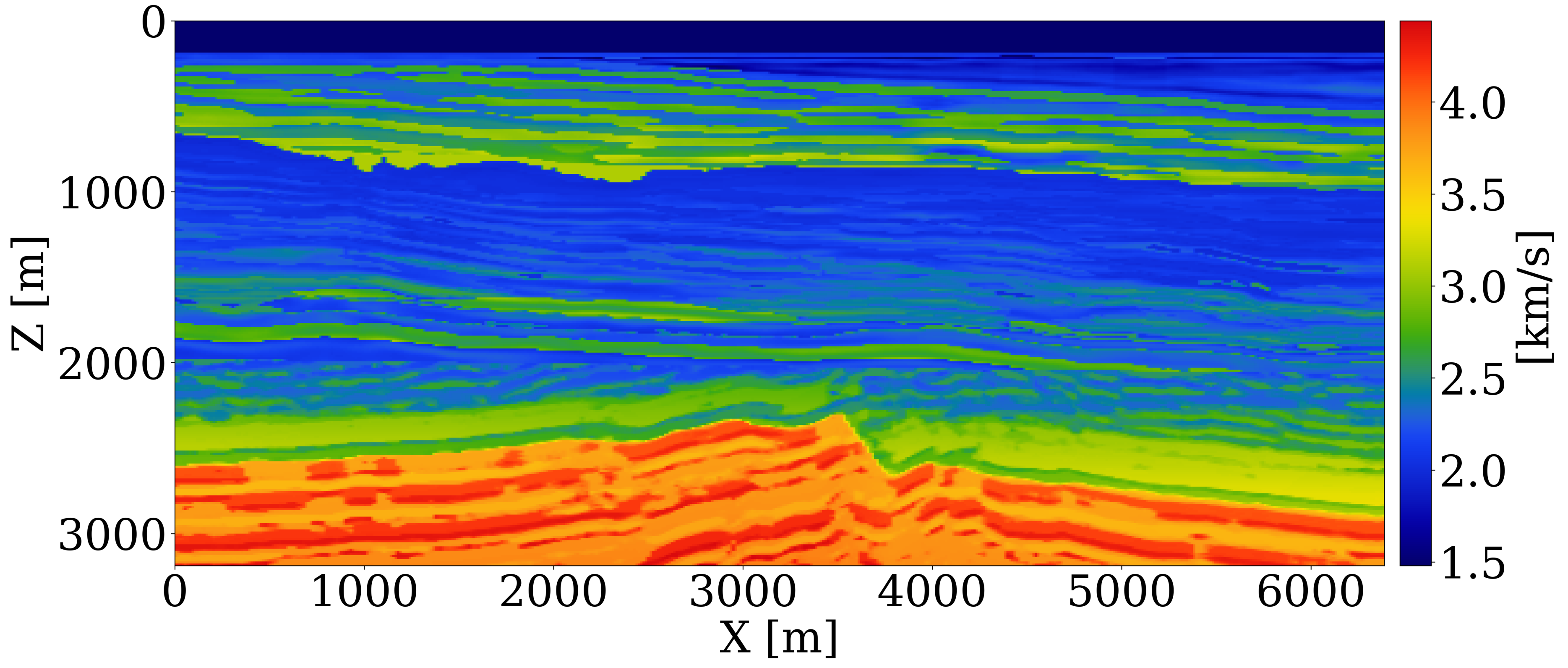
When $\gamma \to 0$, weak formulation → constrained formulation

Outer loop: update $\mathbf{x}_{1:M}$ using expensive physics $\mathcal{F}$ — a few times

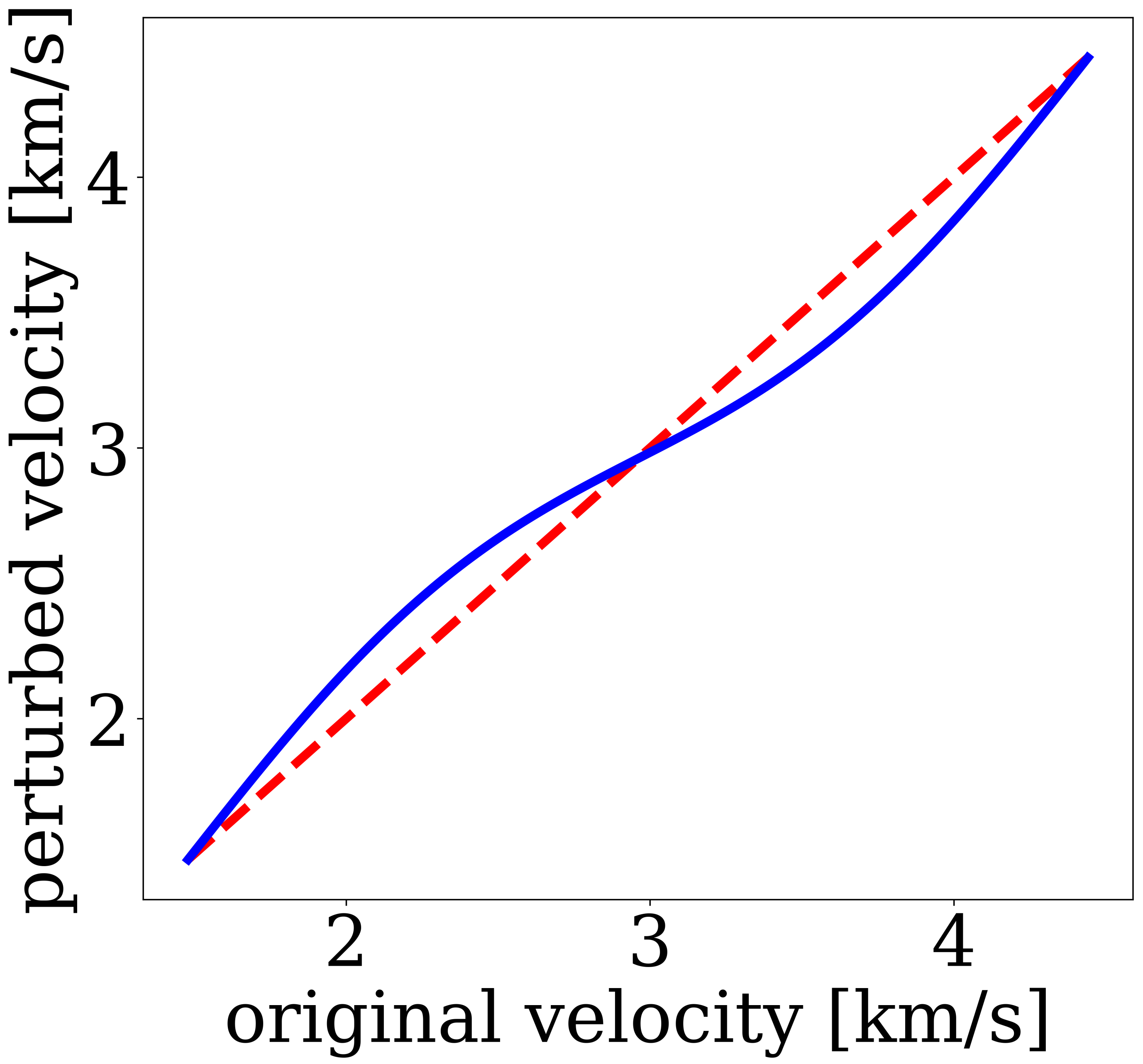
Inner loop: update $\boldsymbol{\phi}$ using only networks — many times

Distribution shift at inference

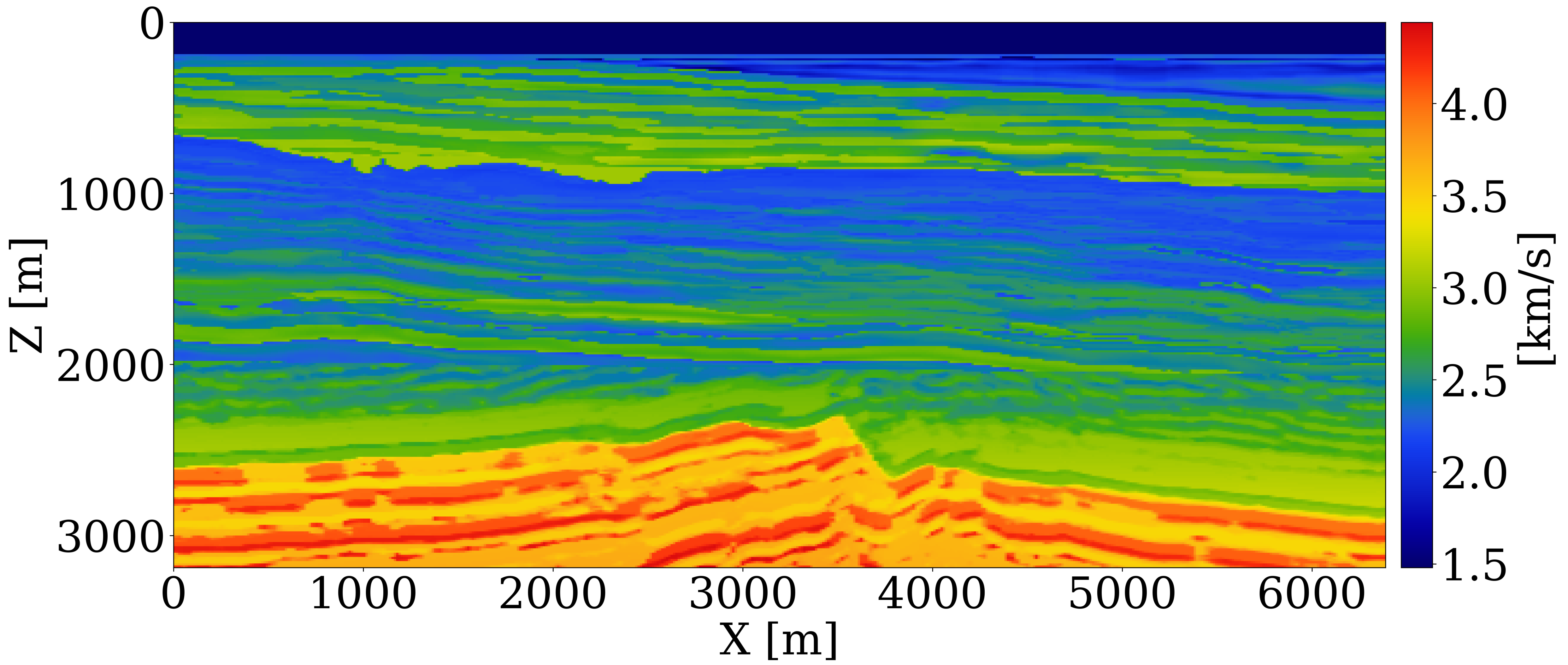
in distribution



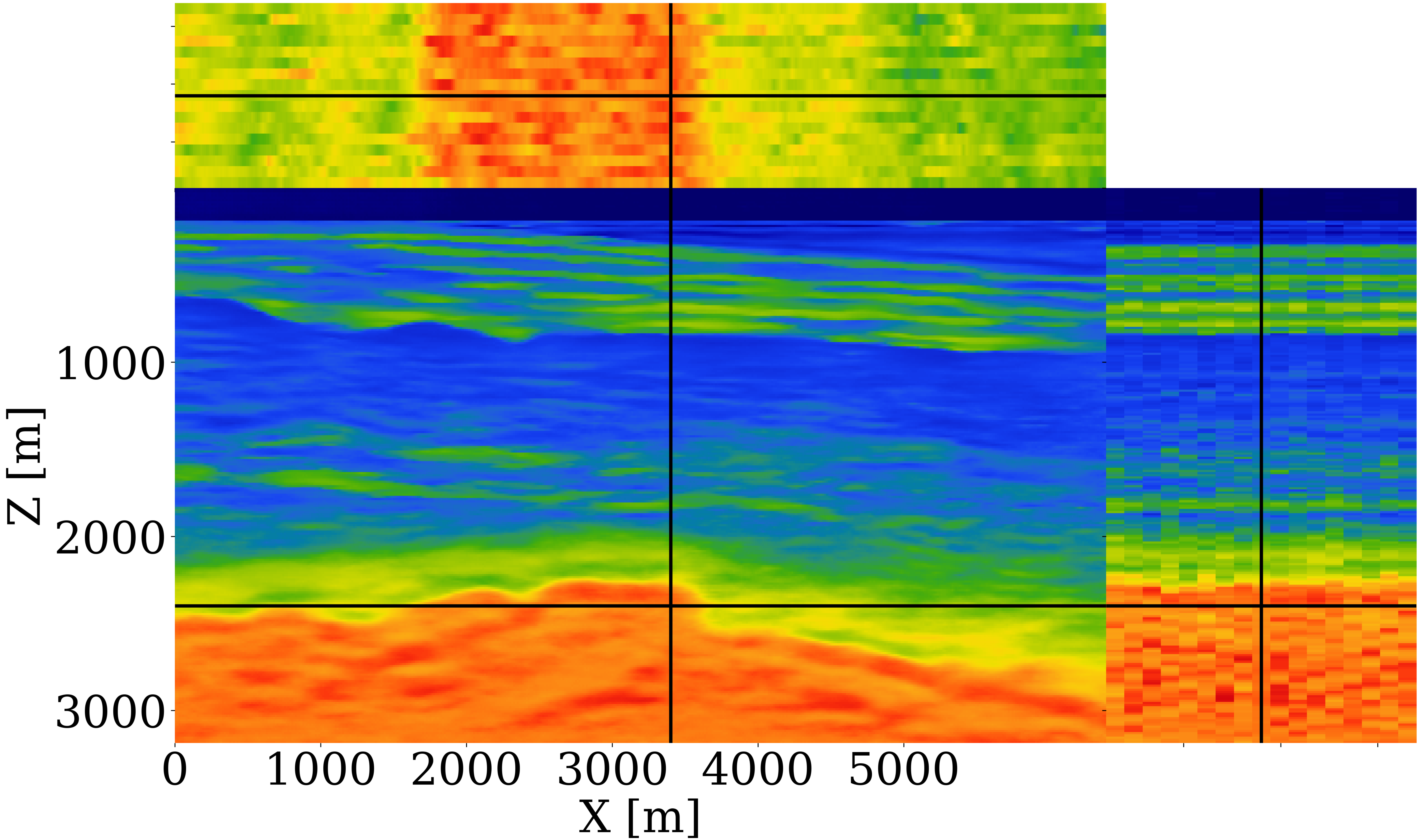
element-wise perturbation



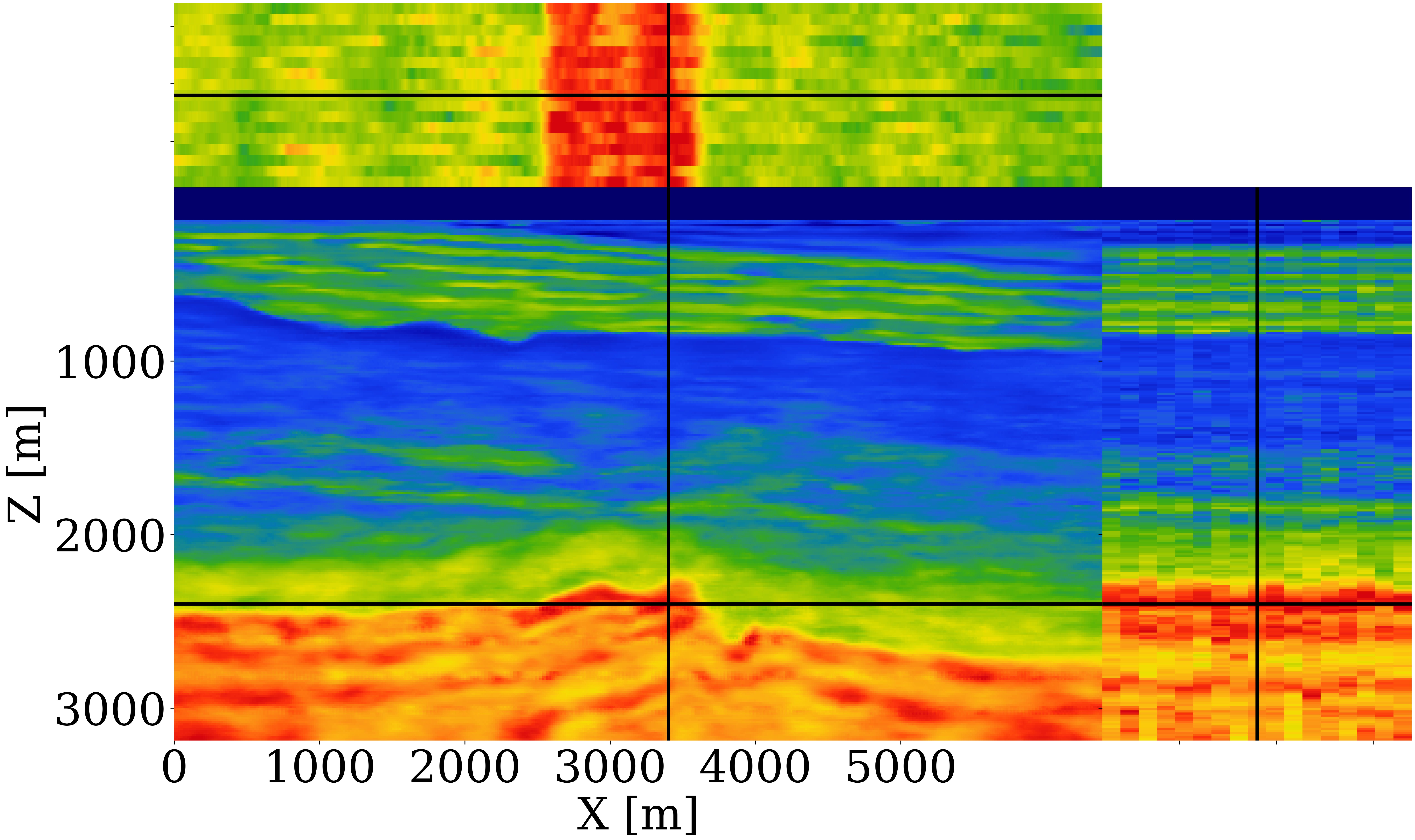
out of distribution



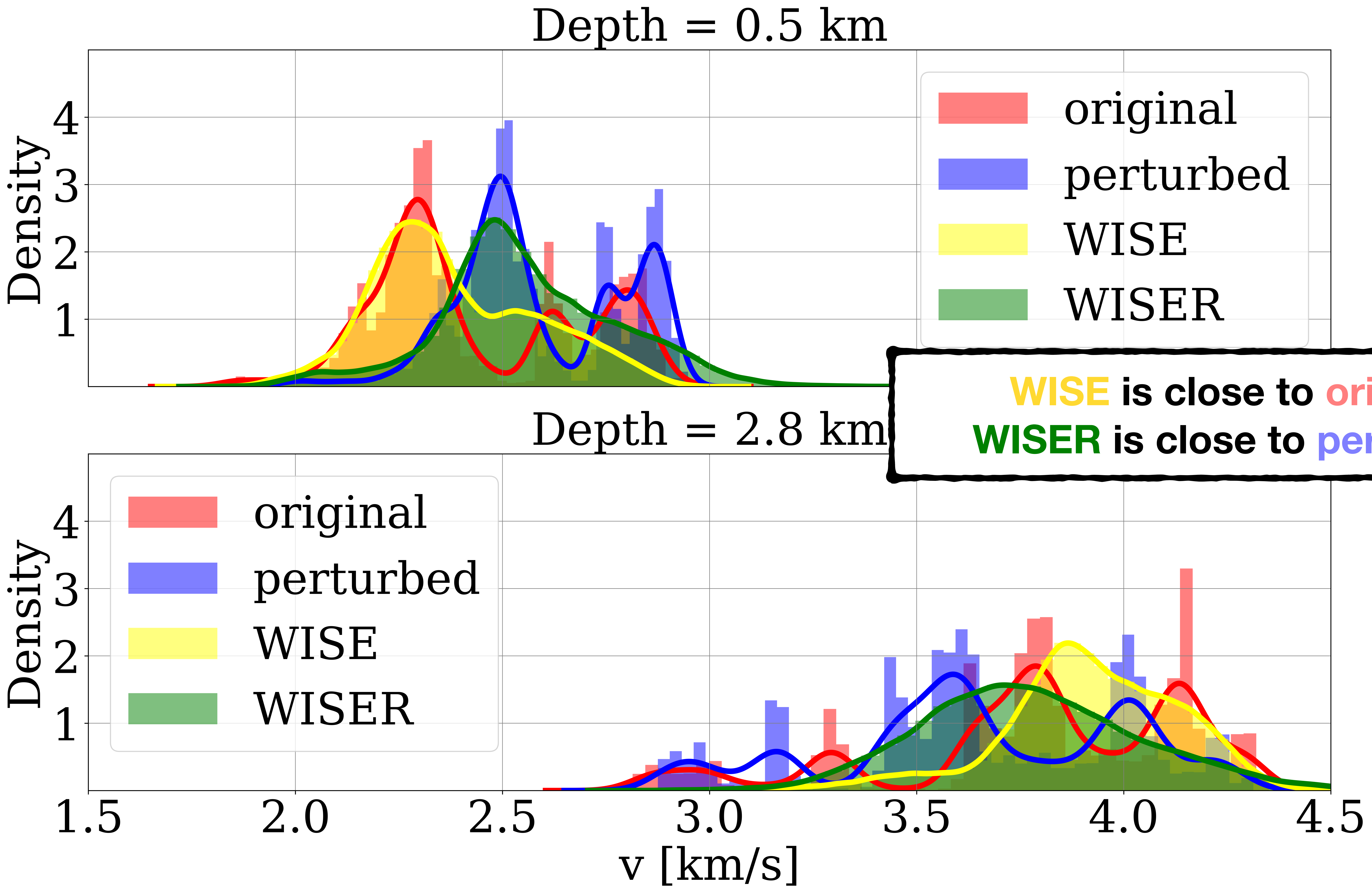
Predicted velocity models - WISE



Predicted velocity models - WISER



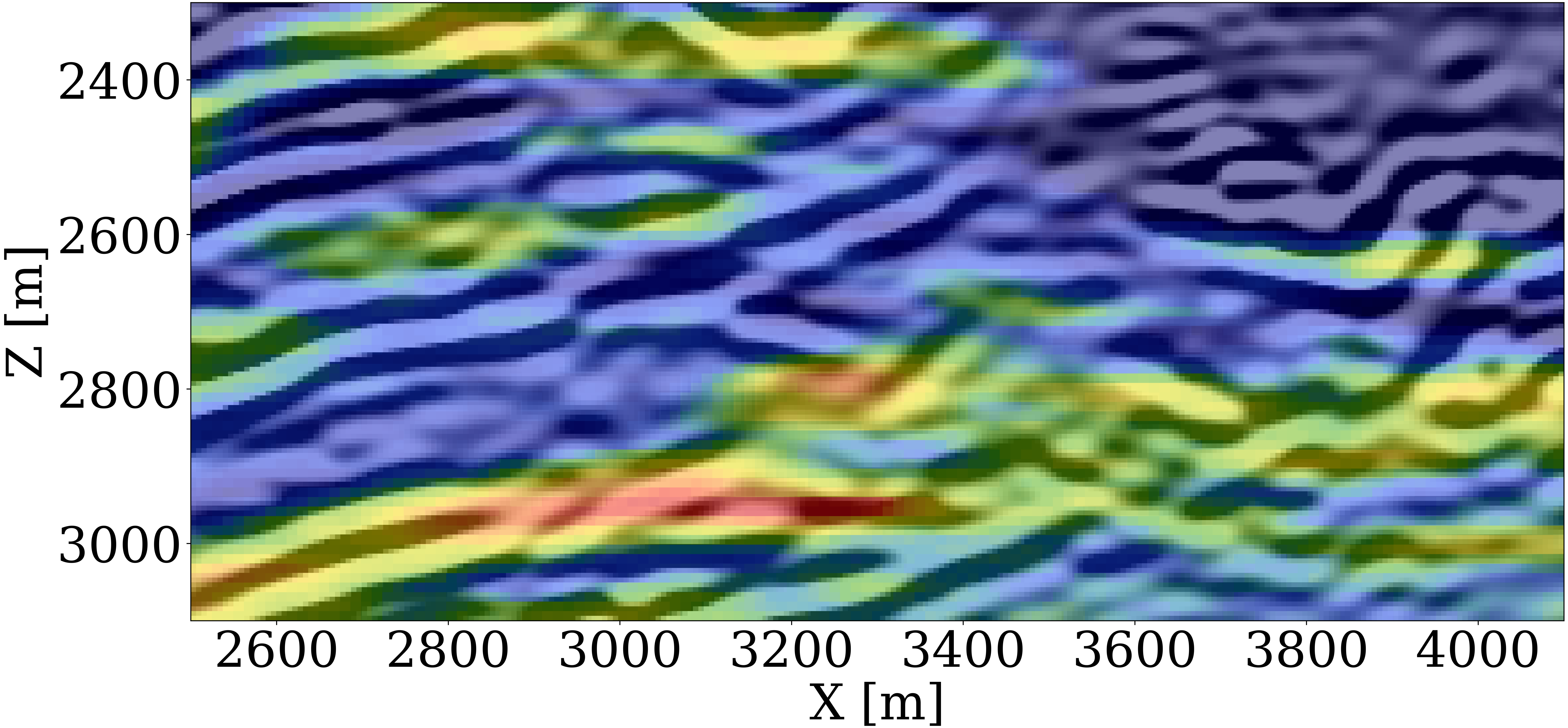
Histograms



WISE is close to original
WISER is close to perturbed

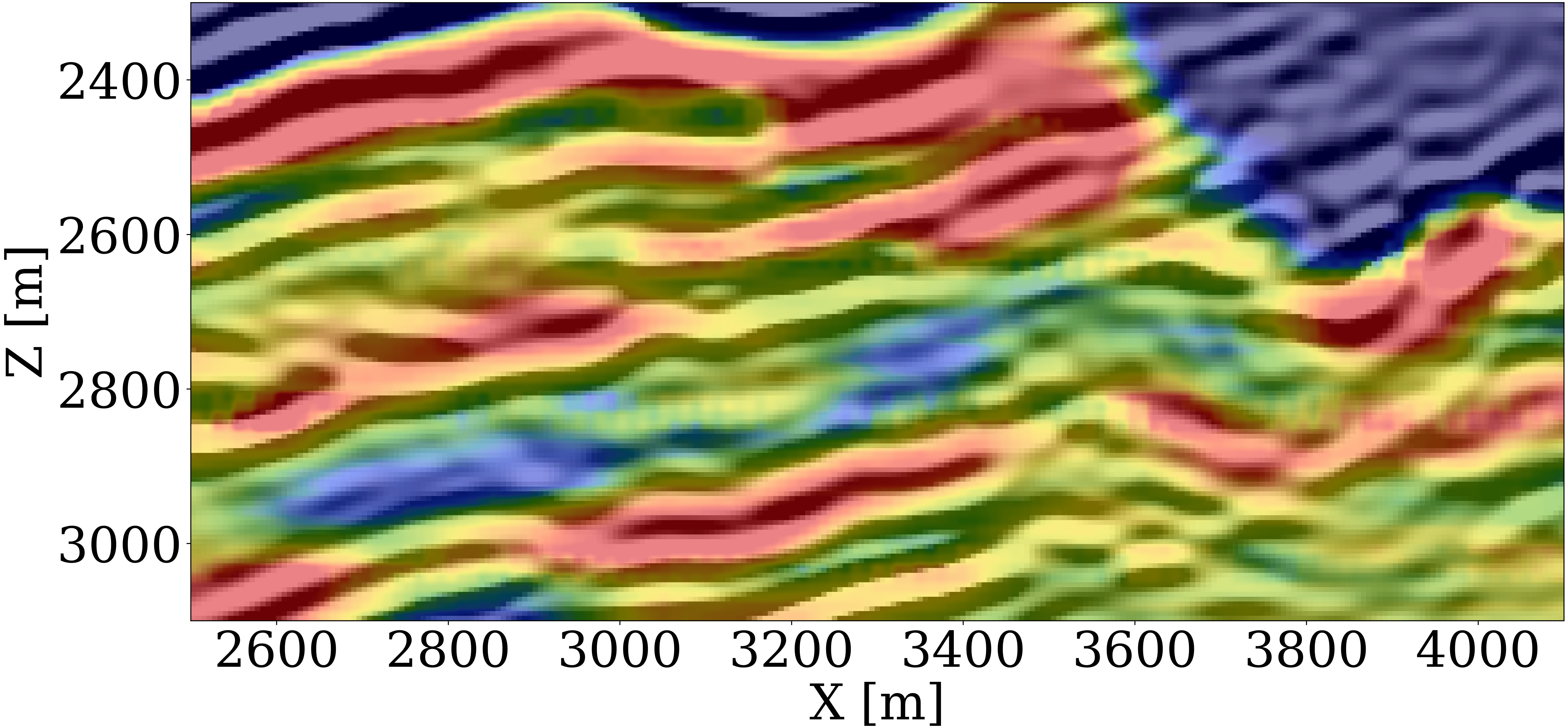
Imaged reflectivities
before correction - WISE

layers are disconnected



Imaged reflectivities
after correction - WISER

layers are connected and aligned with the velocity model



Contributions

Chapter 7

Propose physics-based refinement approach to improve WISE

- *frugal* usage of wave modeling and gradient
- robust w.r.t. *OOD scenarios* at inference

WISER leads to a novel **semi-amortized VI** paradigm

- computationally *affordable & scalable*
- physics-based & *reliable*
- not local, but *global* optimization & UQ
- w/o dimensionality reduction

Summary of contributions

Design *interoperable* and *differentiable* programming framework to support learned multiphysics inversion at scale

Explore deep neural networks as surrogate models to learn

- forward map
  - safeguard the accuracy of surrogate simulators during inversion via *learned constraints*
- (nonunique) inverse map
  - *physics-informed* & *information-preserving* summary statistics based on extension of wave physics
  - mitigate the amortization gap via *affordable physics-based refinements*

Employ the proposed SciML algorithm to solve inverse problems that are

- *high-dimensional*
- with *computationally expensive* forward operators

Including

- full-waveform inversion
- geological carbon storage monitoring

Future directions

Surrogate-assisted inversion with learned constraints

- examine different parameterizations
- derivative-informed surrogate-assisted inversion

Semi-amortized VI w/ WISE & WISER

- theoretically explore the family of model-extension-based summary statistics
- choice of initial model / fiducial point
  - experimental configuration in Bayesian optimal experimental design
  - nuisance parameter in simulation-based inference
- more challenging distribution of model parameters (salt bodies) & OOD scenarios

* denotes equal contributions

SLIM
ML4Seismic

Journal papers

Ziyi Yin, Mathias Louboutin, Olav Møyner, and Felix J. Herrmann. “Time-lapse full-waveform permeability inversion: a feasibility study”. Aug 2024. The Leading Edge. DOI: 10.48550/arXiv.2403.04083.

Ziyi Yin, Rafael Orozco, and Felix J. Herrmann. “WISER: multimodal variational inference for full-waveform inversion without dimensionality reduction”. May 2024. To be submitted. DOI: arXiv.2405.10327.

Ziyi Yin*, Rafael Orozco*, Mathias Louboutin, and Felix J. Herrmann. “WISE: Full-waveform variational inference via subsurface extensions”. Apr 2024. In: Geophysics. DOI: 10.1190/geo2023-0744.1.

Ziyi Yin, Rafael Orozco, Mathias Louboutin, and Felix J. Herrmann. “Solving multiphysics-based inverse problems with learned surrogates and constraints”. Oct 2023. In: Advanced Modeling and Simulation in Engineering Sciences. DOI: 10.1186/s40323-023-00252-0.

Mathias Louboutin*, **Ziyi Yin***, Rafael Orozco, Thomas J. Grady II, Ali Siahkoohi, Gabrio Rizzuti, Philipp A. Witte, Olav Møyner, Gerard J. Gorman, and Felix J. Herrmann. “Learned multiphysics inversion with differentiable programming and machine learning”. Jul 2023. In: The Leading Edge. DOI: 10.1190/tle42070474.1.

Thomas J. Grady II, Rishi Khan, Mathias Louboutin, **Ziyi Yin**, Philipp A. Witte, Ranveer Chandra, Russell J. Hewett, and Felix J. Herrmann. “Model-Parallel Fourier Neural Operators as Learned Surrogates for Large-Scale Parametric PDEs”. Jun 2023. In: Computers & Geosciences. DOI: 10.101 6/j.cageo.2023.105402.

Yijun Zhang, **Ziyi Yin**, Oscar Lopez, Ali Siahkoohi, Mathias Louboutin, Rajiv Kumar, and Felix J. Herrmann. “Optimized time-lapse acquisition design via spectral gap ratio minimization”. Apr 2023. In: Geophysics. DOI: 10.1190/geo2023-0024.1.

Ziyi Yin, Huseyin Tuna Erdinc, Abhinav Prakash Gahlot, Mathias Louboutin, and Felix J. Herrmann. “Derisking geologic carbon storage from high-resolution time-lapse seismic to explainable leakage detection”. Jan 2023. In: The Leading Edge. DOI: 10.1190/tle42010069.1.

* denotes equal contributions

SLIM ML4Seismic

Conference papers

Abhinav Prakash Gahlot, Huseyin Tuna Erdinc, Rafael Orozco, **Ziyi Yin**, Felix J. Herrmann. “Inference of CO2 flow patterns – a feasibility study”. Oct 2023. In: NeurIPS 2023 Workshop - Tackling Climate Change with Machine Learning. DOI: 10.48550/arXiv.2311.00290.

Yijun Zhang*, **Ziyi Yin***, Oscar Lopez, Ali Siahkoohi, Mathias Louboutin, and Felix J. Herrmann. “3D seismic survey design by maximizing the spectral gap”. Aug 2023. In: Third International Meeting for Applied Geoscience & Energy Expanded Abstracts. DOI: 10.1190/image2023-3895546.1.

Huseyin Tuna Erdinc*, Abhinav Prakash Gahlot*, **Ziyi Yin**, Mathias Louboutin, and Felix J. Herrmann. “De-risking Carbon Capture and Sequestration with Explainable CO2 Leakage Detection in Time-lapse Seismic Monitoring Images”. Nov 2022. In: AAAI 2022 Fall Symposium - The Role of AI in Responding to Climate Challenges. DOI: 10.48550/arXiv.2212.08596.

Ziyi Yin, Ali Siahkoohi, Mathias Louboutin, and Felix J. Herrmann. “Learned coupled inversion for carbon sequestration monitoring and forecasting with Fourier neural operators”. Aug 2022. In: Second * denotes equal contribution. International Meeting for Applied Geoscience & Energy Expanded Abstracts. DOI: 10.1190/image202 2-3722848.1. Student oral paper honorable mention.

Mathias Louboutin, Philipp A. Witte, Ali Siahkoohi, Gabrio Rizzuti, **Ziyi Yin**, Rafael Orozco, and Felix J. Herrmann. “Accelerating innovation with software abstractions for scalable computational geophysics”. Aug 2022. In: Second International Meeting for Applied Geoscience & Energy Expanded Abstracts. DOI: 10.1190/image2022-3750561.1.

Yijun Zhang, Mathias Louboutin, Ali Siahkoohi, **Ziyi Yin**, Rajiv Kumar and Felix J. Herrmann. “A simulation-free seismic survey design by maximizing the spectral gap”. Aug 2022. In: Second International Meeting for Applied Geoscience & Energy Expanded Abstracts. DOI: 10.1190/image202 2-3751690.1.

Ziyi Yin, Mathias Louboutin, Felix J. Herrmann. “Compressive time-lapse seismic monitoring of carbon storage and sequestration with the joint recovery model”. Sep 2021. In: First International Meeting for Applied Geoscience & Energy Expanded Abstracts. DOI: 10.1190/segam2021-3569087.1.

Ziyi Yin, Rafael Orozco, Philipp A. Witte, Mathias Louboutin, Gabrio Rizzuti, and Felix J. Herrmann. “Extended source imaging, a unifying framework for seismic & medical imaging”. Sep 2020. In: SEG Technical Program Expanded Abstracts 2020. DOI: 10.1190/segam2020-3426999.1.

Many thanks to

Felix J. Herrmann

PhD proposal & dissertation defense committee members

Colleagues at SLIM

Olav Møyner at SINTEF

Rishi Khan at Extreme Scale Solution

Internship colleagues at Chevron

NSF, DOE, ML4SEISMIC center, Georgia Research Alliance

Family and friends