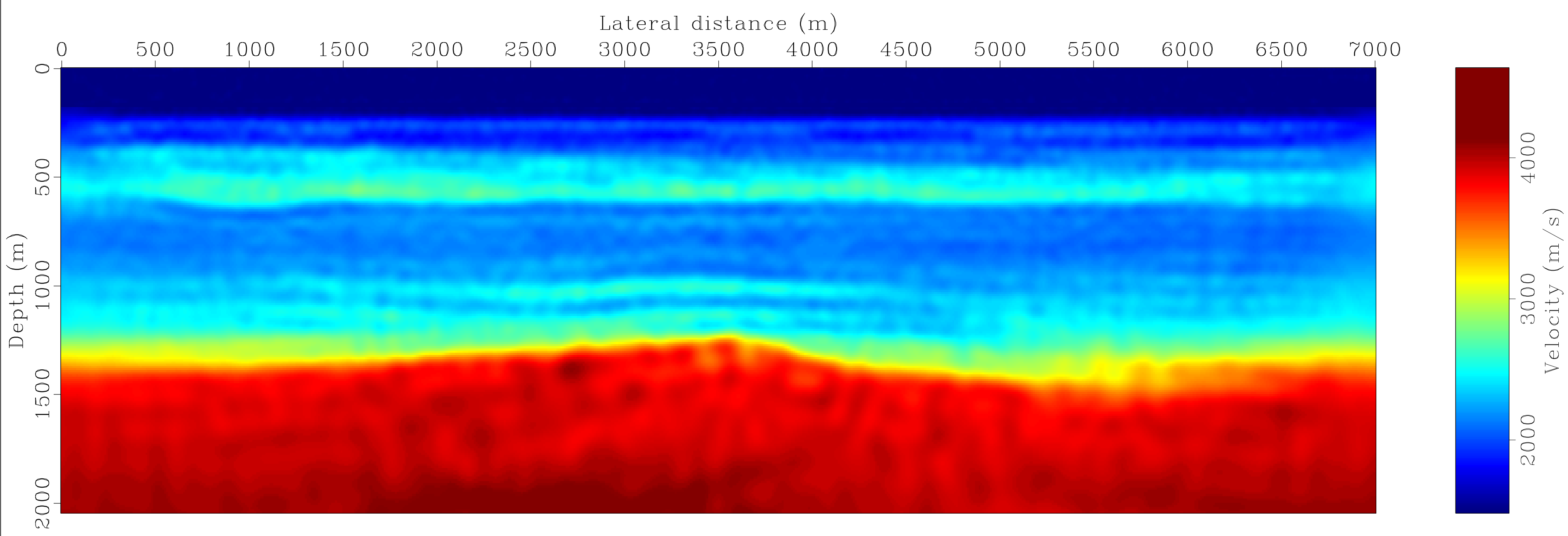
SINBAD Consortium 2012

Fast Gauss-Newton full-waveform inversion with sparsity regularization

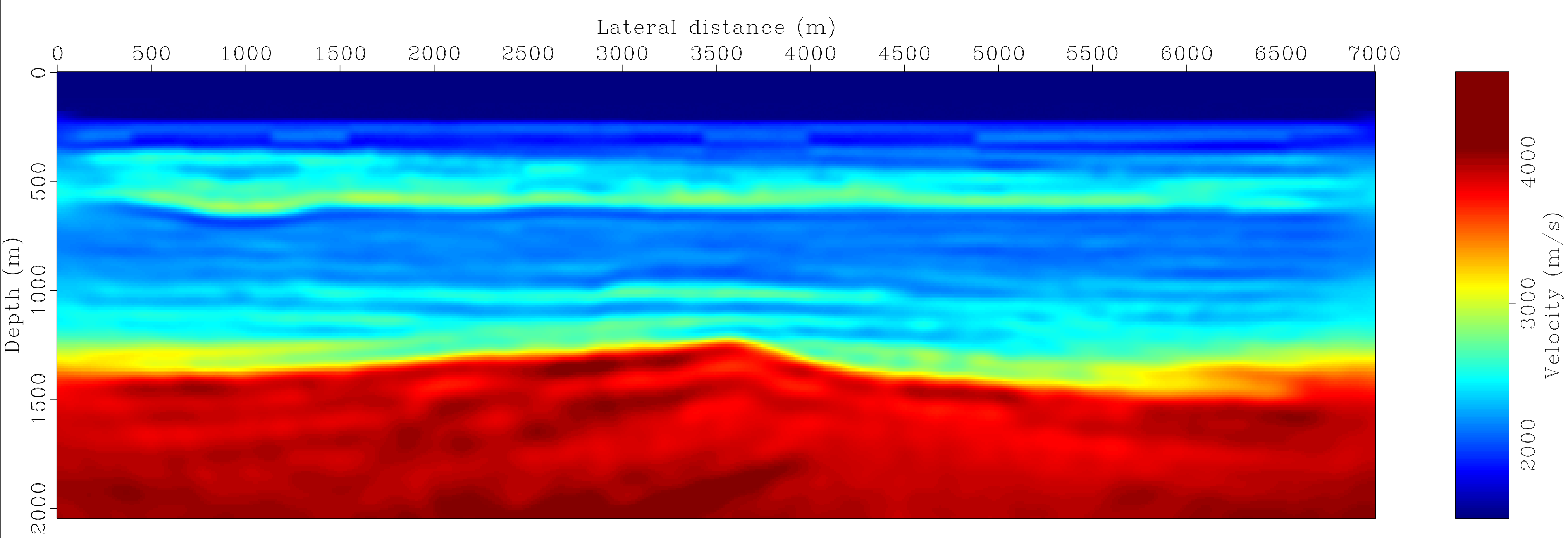
Xiang Li and Felix J. Herrmann

SLIM
University of British Columbia

Motivation

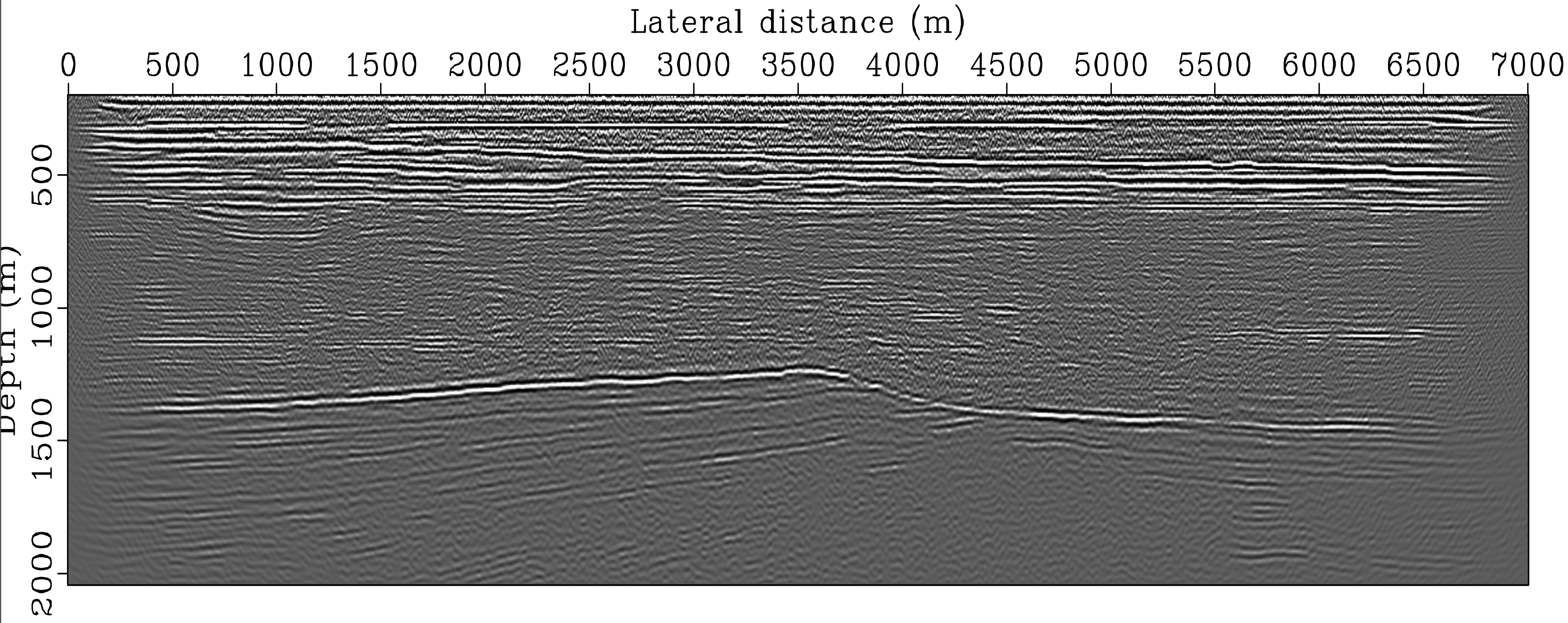


standard GN FWI

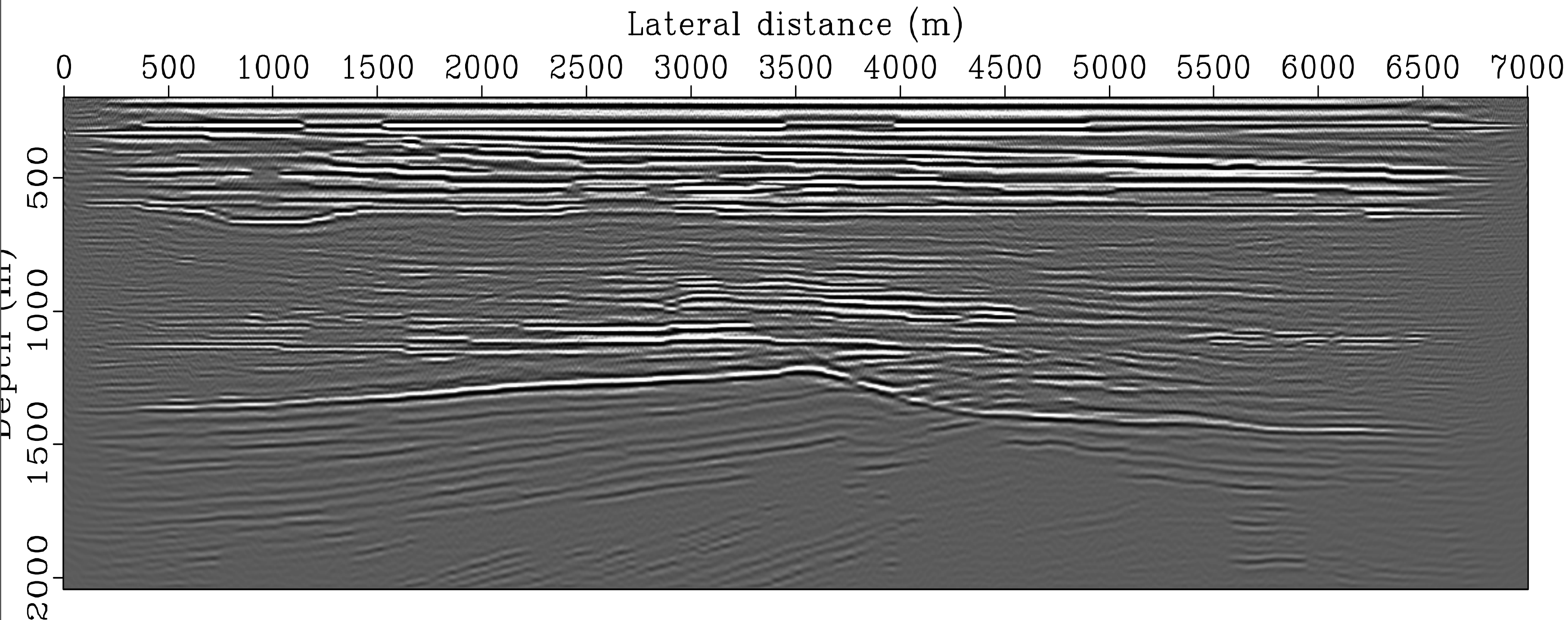


sparsity-promoting GN FWI

Motivation



standard least-squares migration



sparsity promoting least-squares migration

Outline

- sparsity promoting Gauss-Newton FWI to generate initial model
- sparsity-promoting imaging for migration

Making sparsity-promoting imaging computationally feasible...

Full-waveform inversion

Full waveform inversion

$$\underset{\mathbf{m},\alpha}{\text{minimize}}\ \frac{1}{2}\|\mathbf{D} - \alpha\boldsymbol{\mathcal{F}}[\mathbf{m}]\|_F^2$$

$\mathbf{D}$: observed data

$\boldsymbol{\mathcal{F}}$: forward modelling kernel

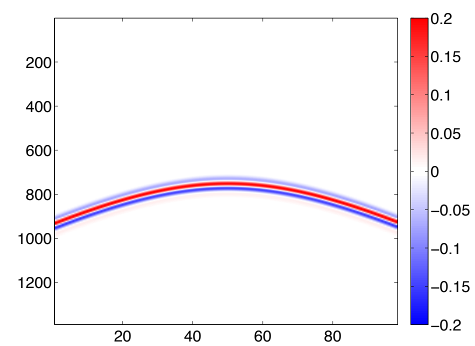
α : source wavelet

$\mathbf{m}$: model parameters

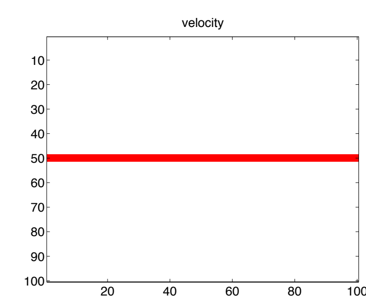
Gauss-Newton

Gauss-Newton subproblem:

$$\underset{\delta\mathbf{m}}{\text{minimize}}\ \frac{1}{2}\|\underbrace{\mathbf{D}-\alpha\mathcal{F}(\mathbf{m})}_{\mathbf{b}}-\underbrace{\alpha\nabla\mathcal{F}(\mathbf{m})}_{\mathbf{A}}\underbrace{\delta\mathbf{m}}_{\mathbf{x}}\|_F^2$$



Jacobian operator (born modeling operator)



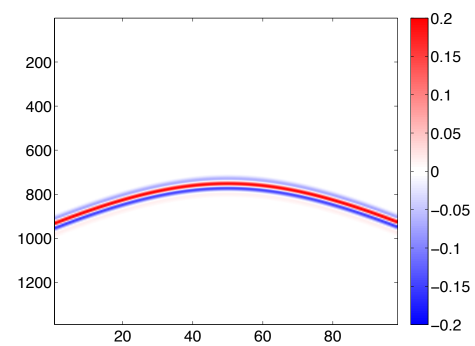
- least-squares inversion problem
- no explicit Jacobian required

Gauss-Newton

Gauss-Newton subproblem:

$$\underset{\delta\mathbf{m}}{\text{minimize}}\ \frac{1}{2}\|\underbrace{\mathbf{D}-\alpha\boldsymbol{\mathcal{F}}(\mathbf{m})}_{\mathbf{b}}-\underbrace{\alpha\nabla\boldsymbol{\mathcal{F}}(\mathbf{m})}_{\mathbf{A}}\underbrace{\delta\mathbf{m}}_{\mathbf{x}}\|_F^2$$

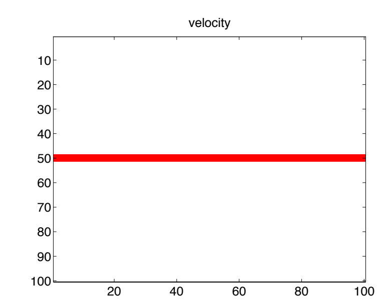
$\mathbf{b}$



$\mathbf{A}$

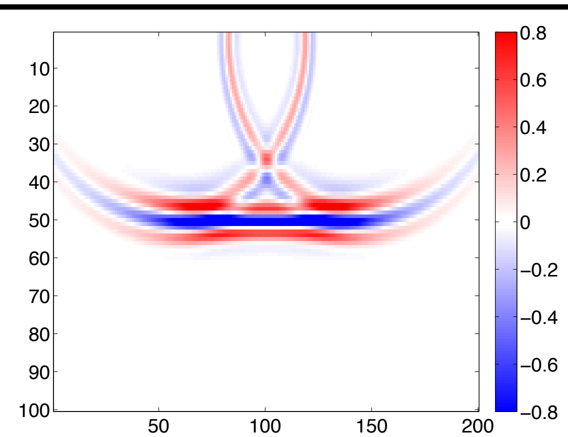
Jacobian operator (born modeling operator)

$\mathbf{x}$



Hint:

$$\mathbf{A}^T\mathbf{b} =$$

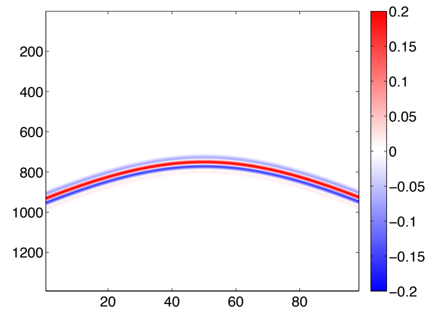


RTM and Born modeling needs to satisfy adjoint-state

Sparsity-promoting GN

Gauss-Newton subproblem:

$$\underset{\delta \mathbf{m}}{\text{minimize}} \frac{1}{2}\|\mathbf{D} - \alpha\boldsymbol{\mathcal{F}}(\mathbf{m}) - \alpha\nabla\boldsymbol{\mathcal{F}}(\mathbf{m})\mathbf{C}^T\delta\mathbf{x}\|_F^2 \quad \text{s.t} \quad \|\delta\mathbf{x}\|_1 < \tau$$



Jacobian operator (born modeling operator)

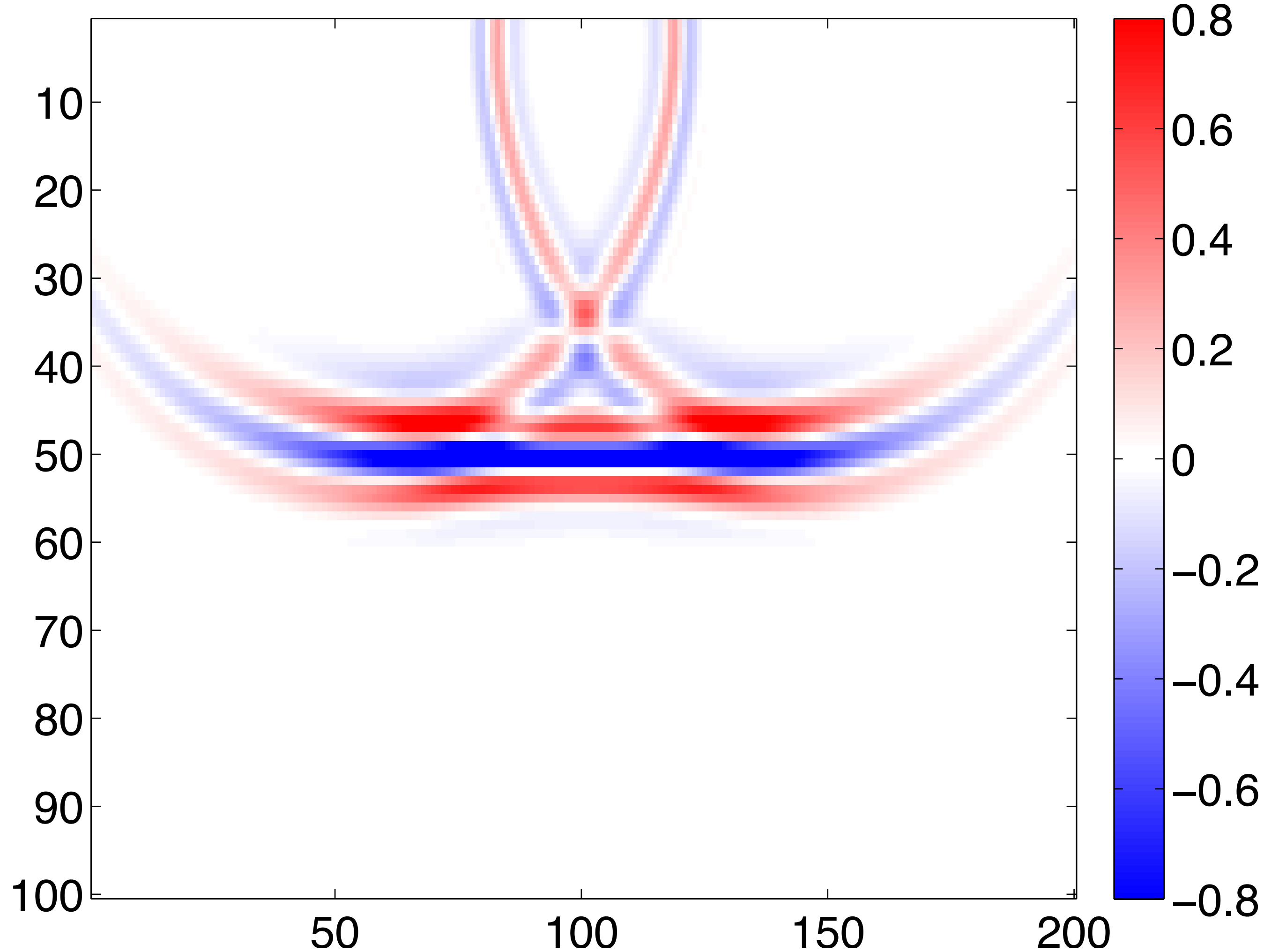
Sparsifying transform

$$\delta\mathbf{m} = \mathbf{C}^T\delta\mathbf{x}$$

- suppress incoherent noise/artifacts by one-norm constraint in a transform domain
- randomized subsets of shots to reduce computational costs

Curvelet regularization

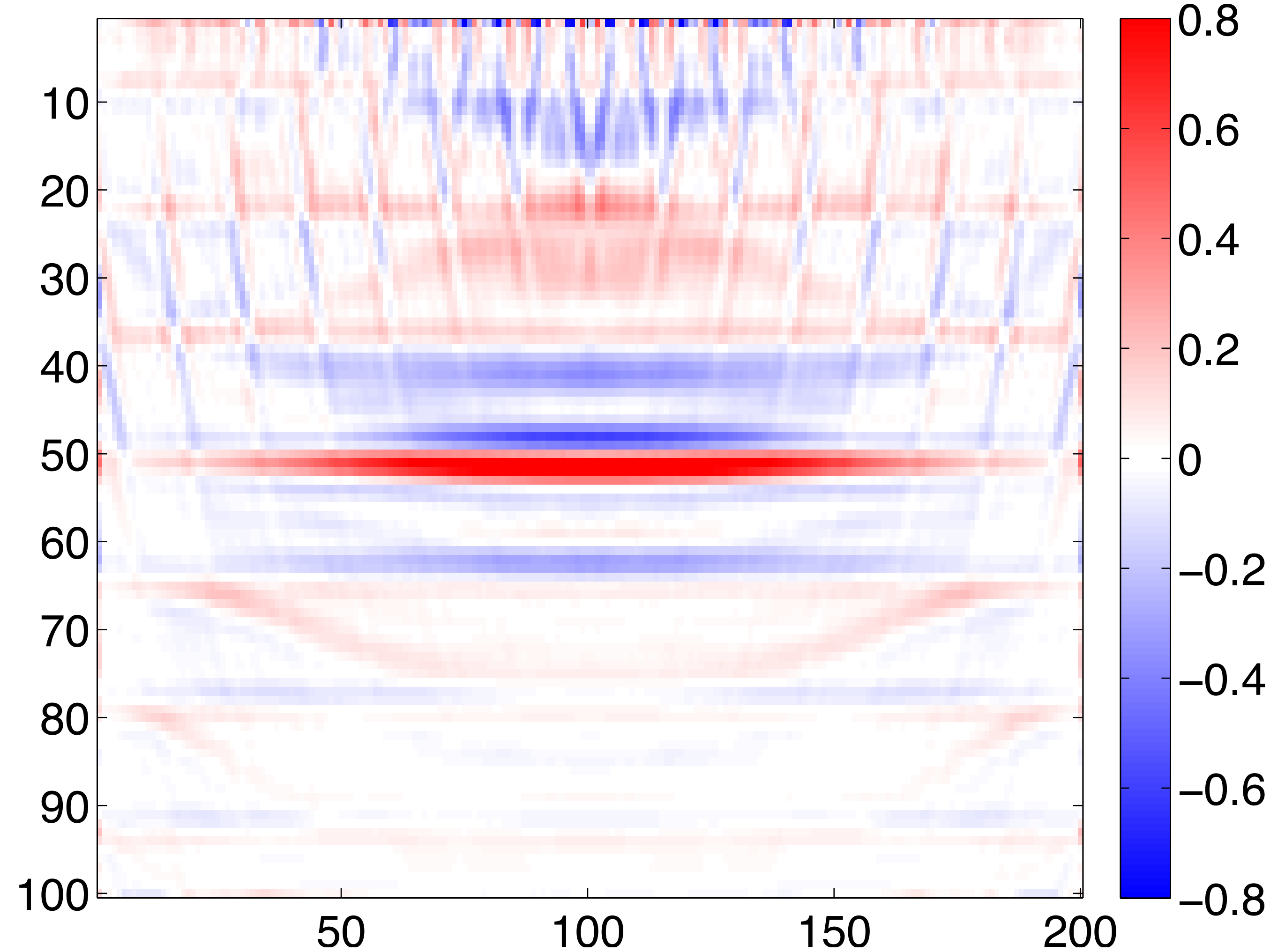
Gradient (RTM) of one shot for all frequencies



Curvelet regularization

Gradient of all shots for subsampled frequencies

- used as a gradient for FWI *(expensive)*
- subsampling frequencies causes periodic artifacts



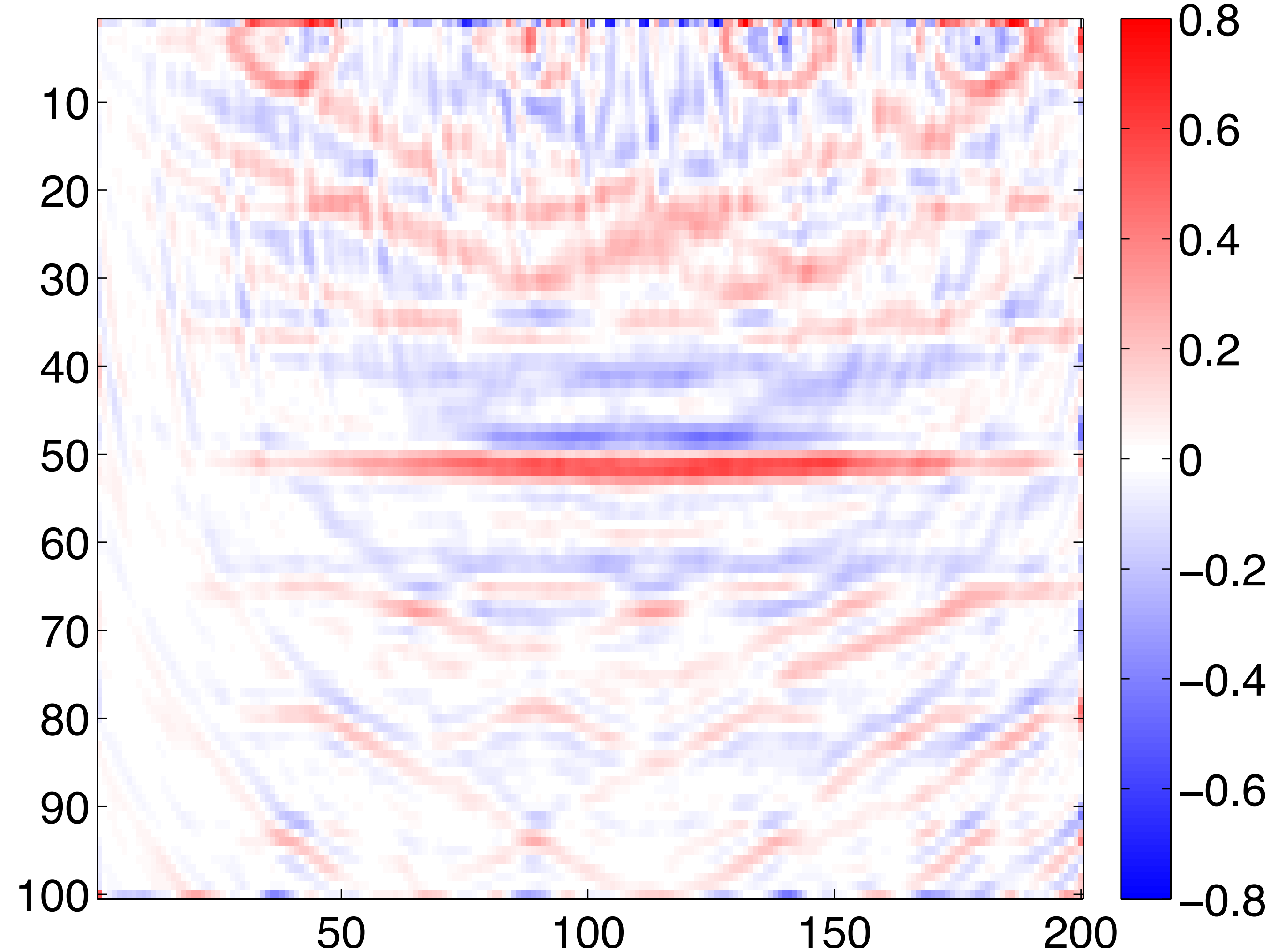
Curvelet regularization

gradient of subsampled shots and frequencies

- subsampling sources

  will introduce even

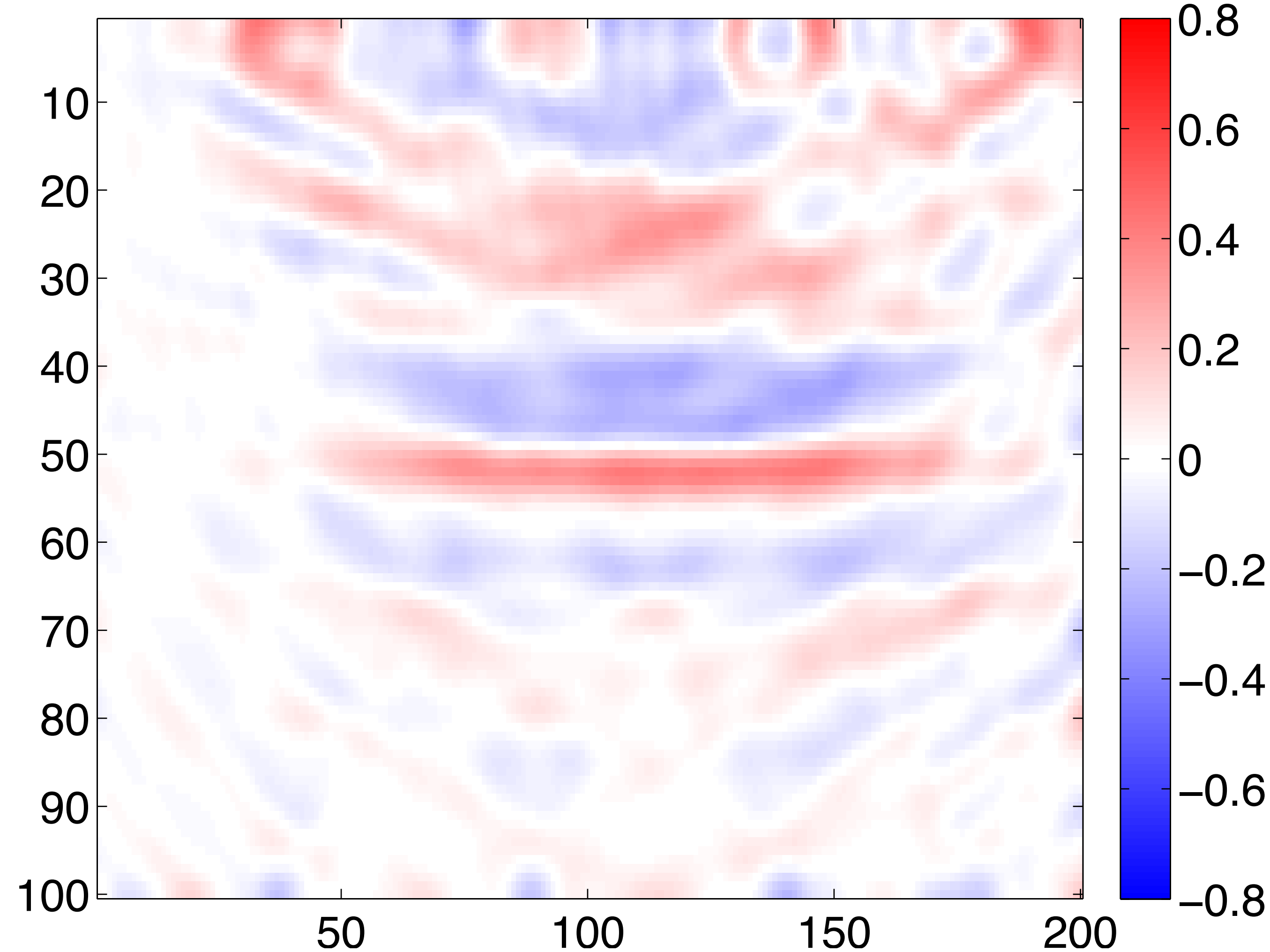
  more artifacts



Curvelet regularization

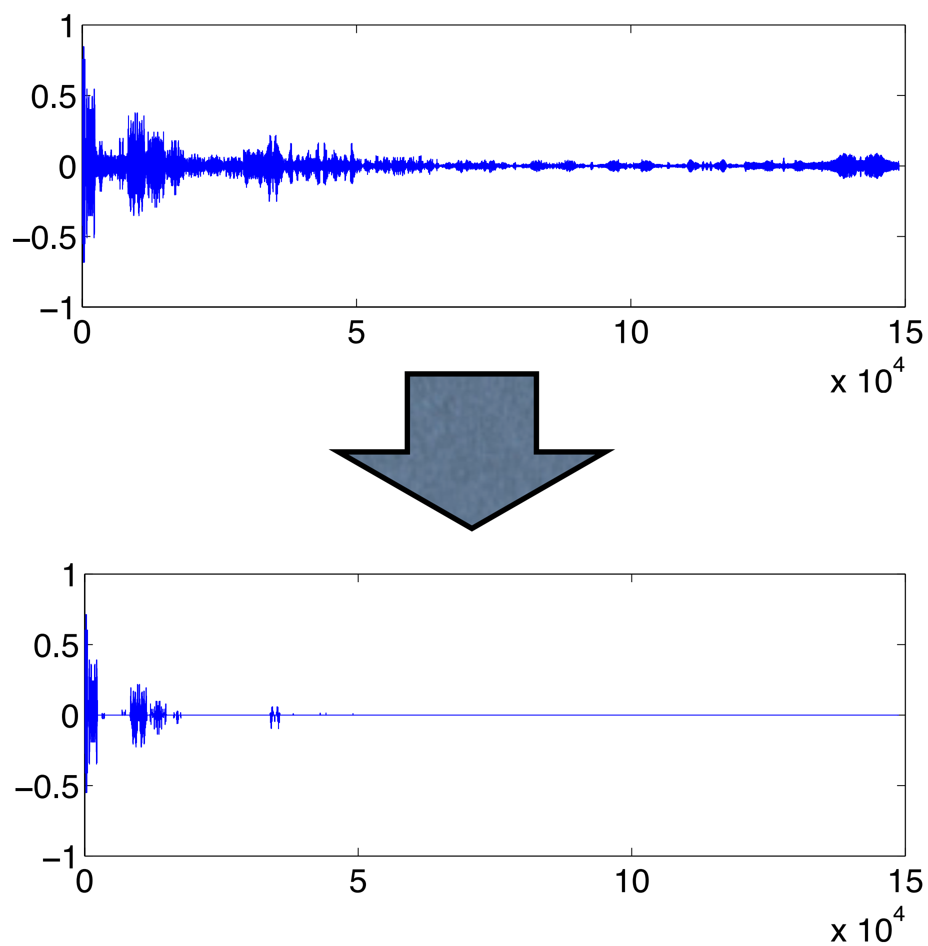
one solution to suppress artifacts is by smoothing

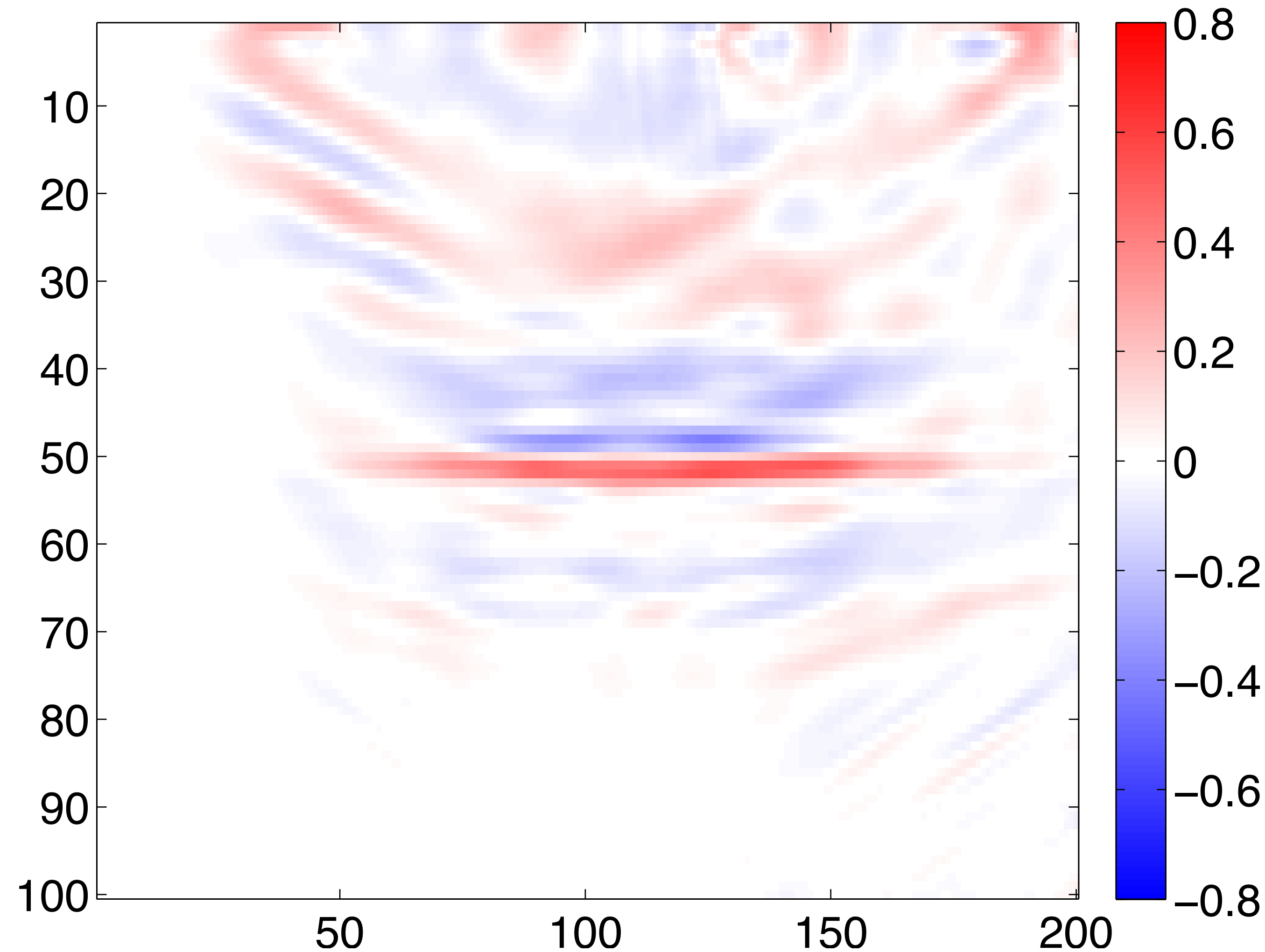
- loss high frequency information



Curvelet regularization

Sparsity regularization in Curvelet domain

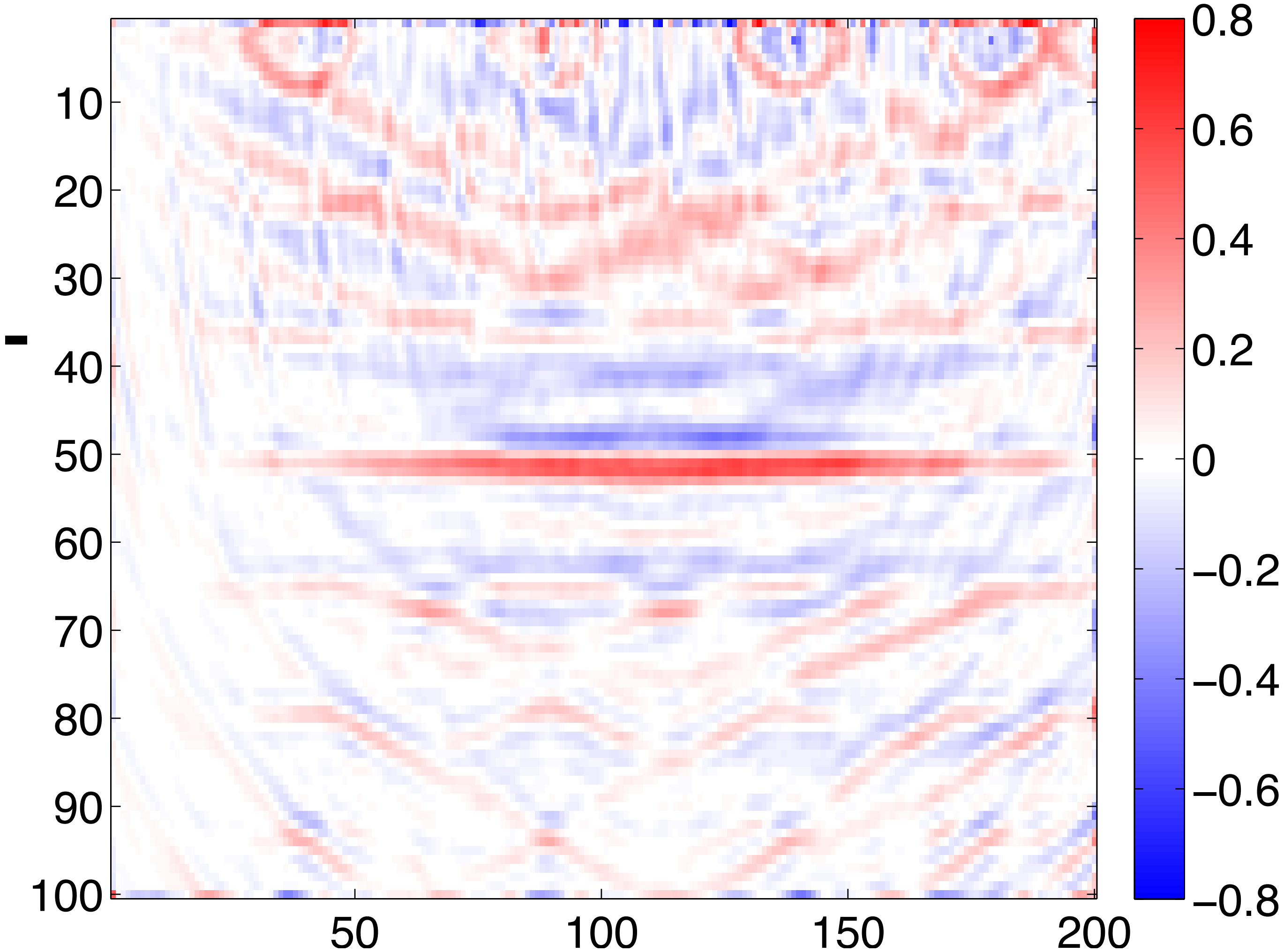




Curvelet regularization

Sparsity regularization in Curvelet domain

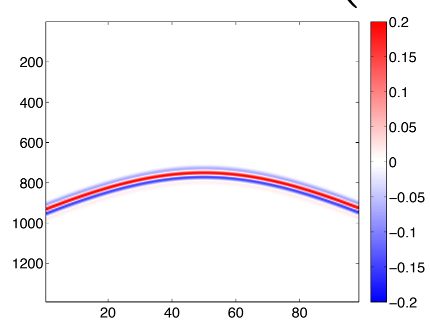
- efficiently suppresses artifacts
- maximally preserves geo-logical structures



Sparsity-promoting GN

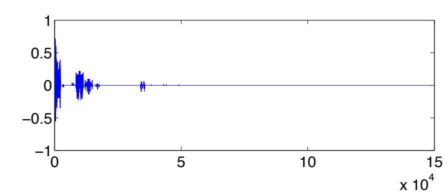
Gauss-Newton subproblem:

$$\underset{\delta \mathbf{m}}{\text{minimize}} \; \frac{1}{2}\|\mathbf{D} - \alpha\boldsymbol{\mathcal{F}}(\mathbf{m}) - \alpha\nabla\boldsymbol{\mathcal{F}}(\mathbf{m})\mathbf{C}^T\delta\mathbf{x}\|_F^2 \quad \text{s.t} \quad \|\delta\mathbf{x}\|_1 < \tau$$



Jacobian operator (born modeling operator)

Sparsifying transform



$$\delta\mathbf{m} = \mathbf{C}^T\delta\mathbf{x}$$

- suppress incoherent noise/artifacts by one-norm constraint in a transform domain
- randomized subsets of shots to reduce computational costs

[Pratt, '98]

Source estimation

Source estimation at the k^{th} iteration:

$$\alpha_k = \arg\min_{\alpha} \|\mathbf{D}_k - \alpha \boldsymbol{\mathcal{F}}(\mathbf{m}_k)\|_F^2$$

- source estimation handles amplitude problems
- cheap to evaluate

Sparsity-promoting GN FWI

Algorithm 1: Sparsity-promoting Gauss-Newton FWI

Result: Output estimate for the model $\mathbf{m}$

$\mathbf{m} \longleftarrow \mathbf{m}_0$; $k \longleftarrow 0$; `// initial model`

while not converged **do**

- $\alpha_k \longleftarrow \arg\min_\alpha \|\mathbf{D}_k - \alpha \boldsymbol{\mathcal{F}}(\mathbf{m}_k)\|_F^2$; `// source estimation`
- $\delta\mathbf{D}_k \longleftarrow \mathbf{D}_k - \alpha_k \boldsymbol{\mathcal{F}}(\mathbf{m}_k)$; `// wavefield residual`
- $\delta\mathbf{x}_k \longleftarrow \arg\min_{\delta\mathbf{x}} \frac{1}{2}\|\delta\mathbf{D}_k - \alpha\nabla\boldsymbol{\mathcal{F}}(\mathbf{m}_k)\mathbf{C}^T\delta\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \|\delta\mathbf{x}\|_1 \leq \tau_k$
- $\mathbf{m}^{k+1} \longleftarrow \mathbf{m}^k + \mathbf{C}^T\delta\mathbf{x}$; `// update model`
- $k \longleftarrow k + 1$;

end

Examples

Acquisition geometry:

- 350 shots with interval 20 m
- 701 receivers with interval 10 m

Observed data is generated by

- time domain finite difference modeling method with PML boundary
- 12 Hz Ricker wavelet

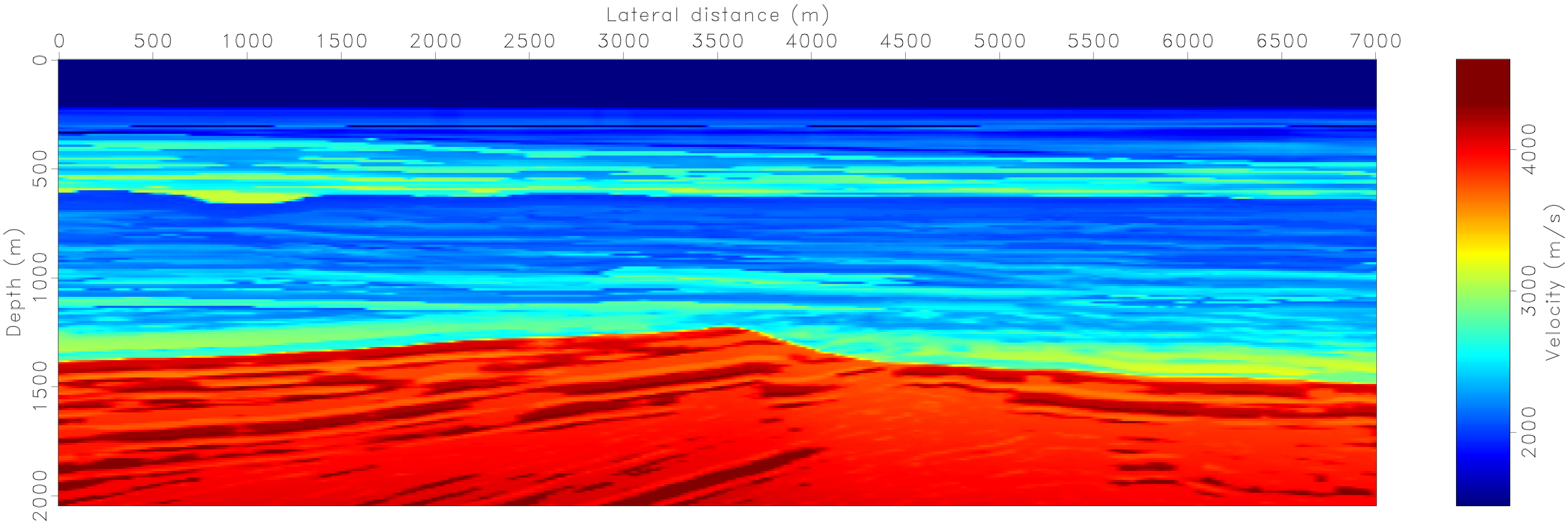
Examples

Inversion parameters:

- 10 frequency bands for 3-12 Hz.
- each band contain 4 frequencies
- inversion using frequency modeling kernel (Helmholtz)
- grid size is determined by minimal wavelength
- solve 5-10 GN subproblems for each frequency band
- < 10 iterations for each GN subproblem

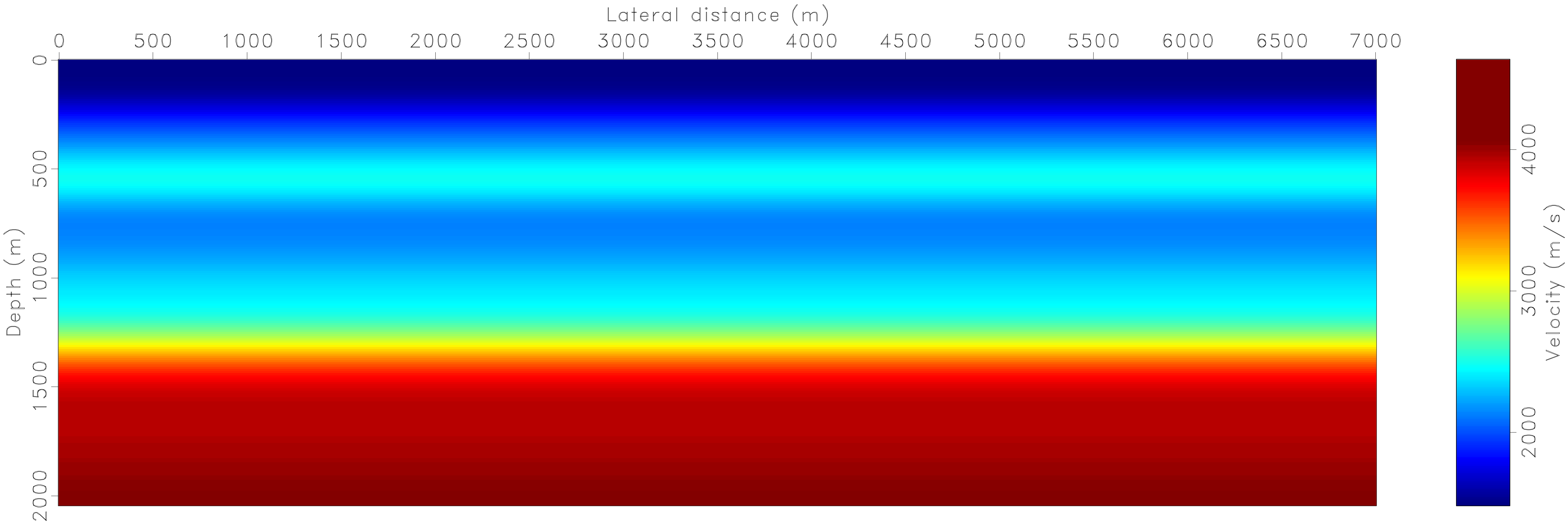
True model

BG compass



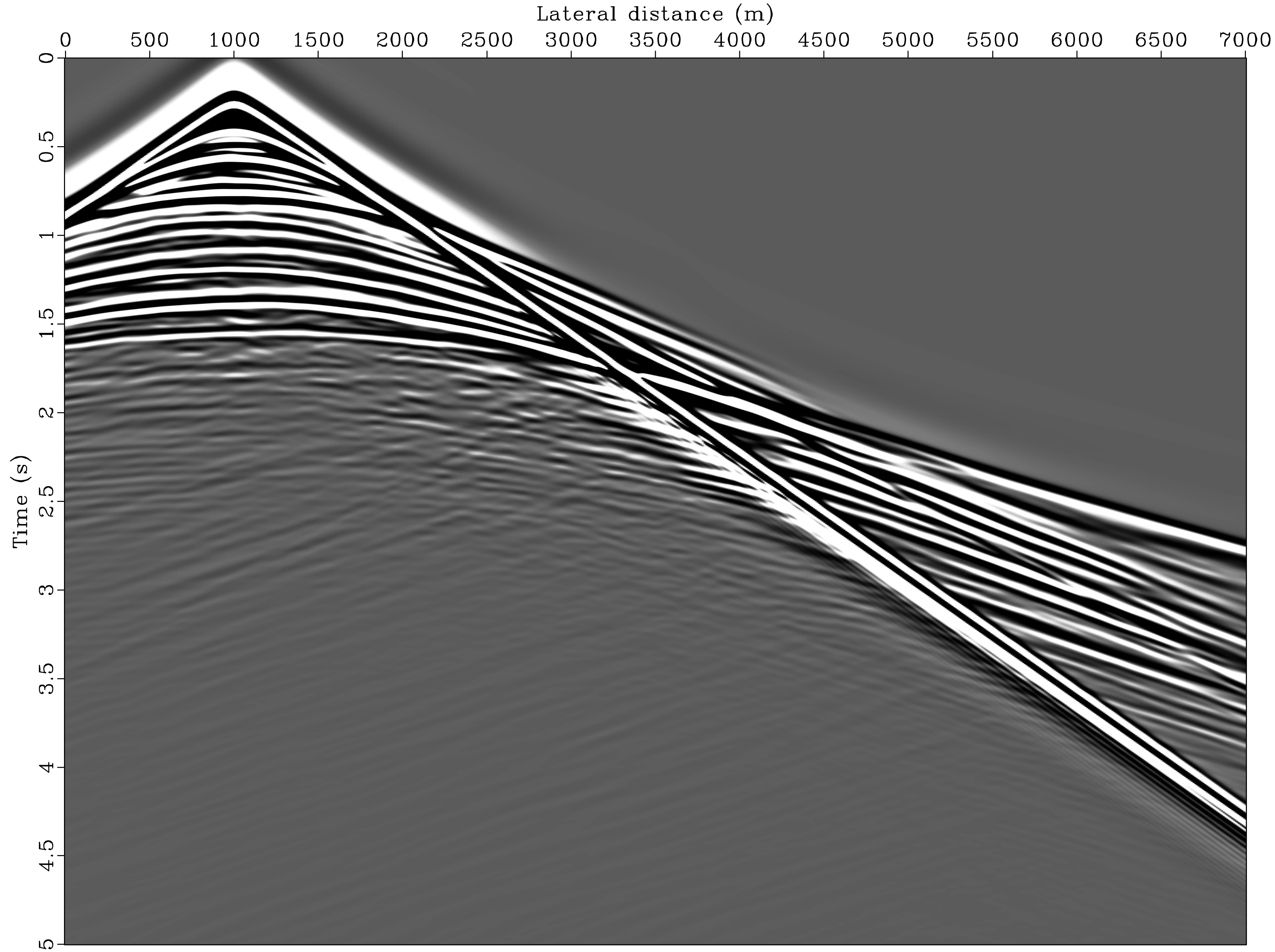
Initial model

Initial model



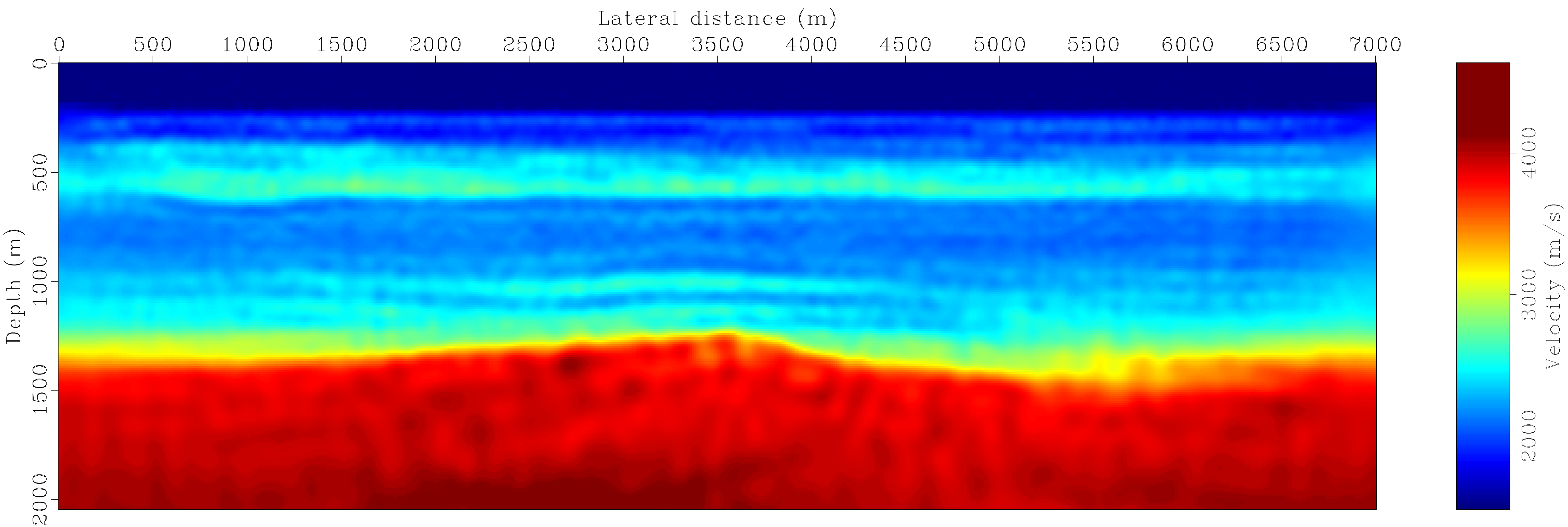
One shot record

Time-domain finite-differences (PML boundary)



Normal Gauss-Newton

Each GN subproblem uses 20 randomly selected sequential shots

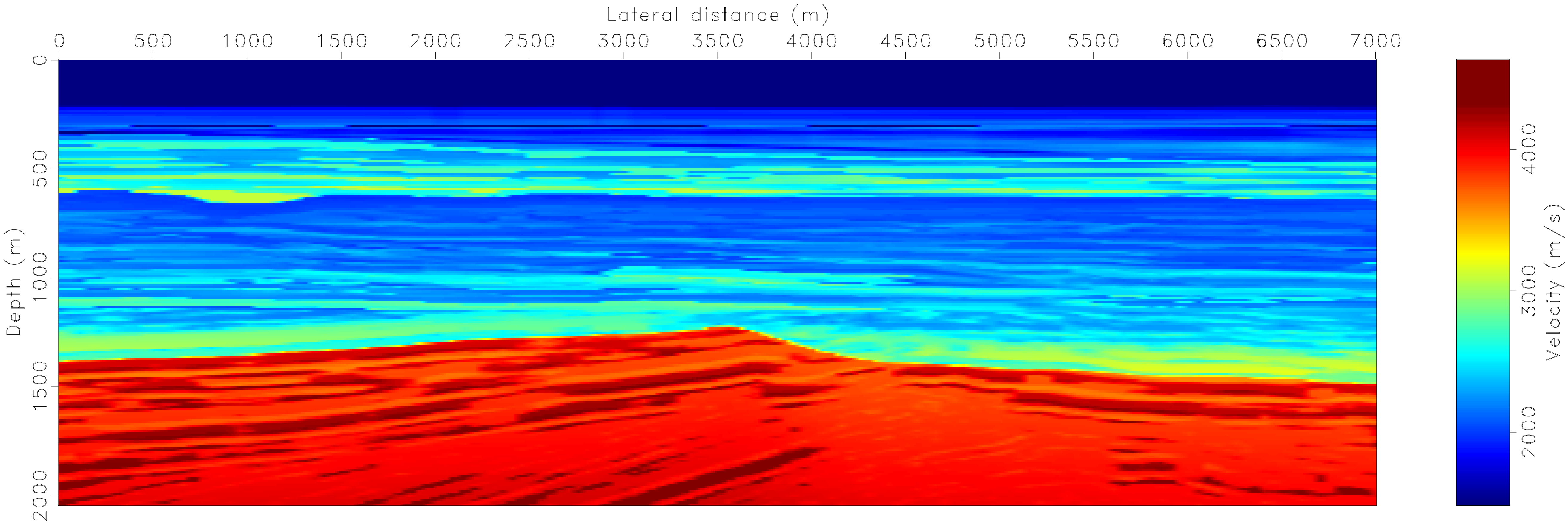


Each GN subproblem using 5 lsqr iterations

1x node, 4 x cpus, an hour

True model

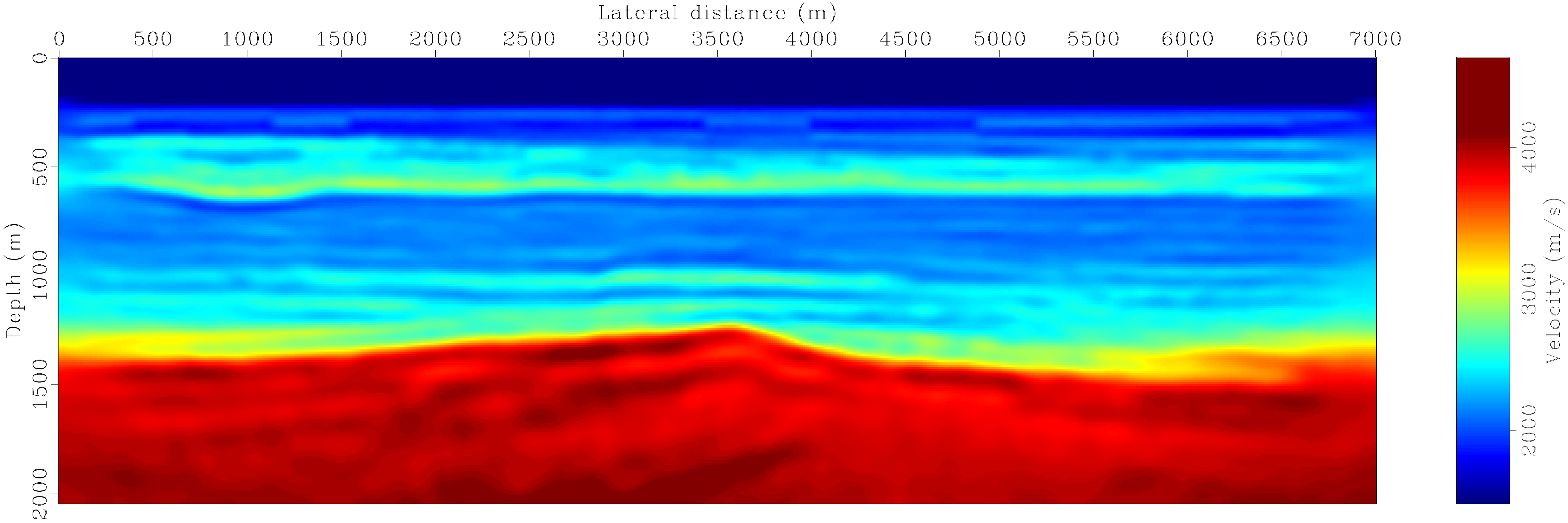
BG compass



Sparsity promoting Gauss-Newton

SLIM

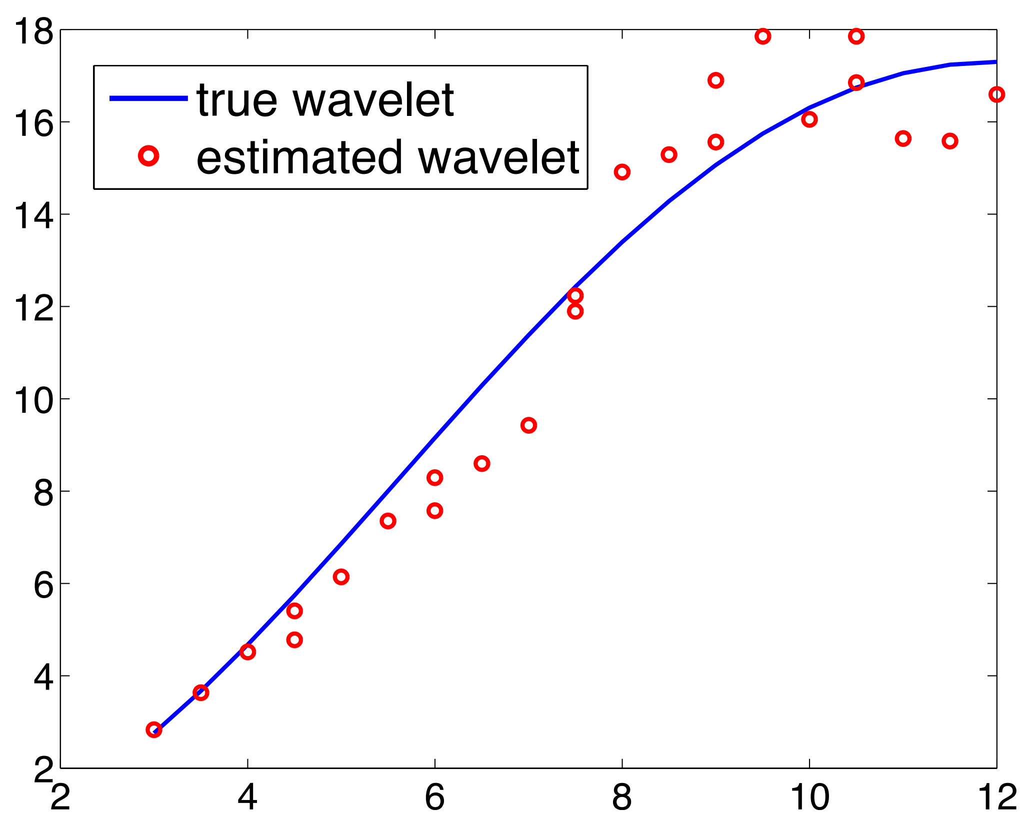
Each GN subproblem using 20 randomly selected shots



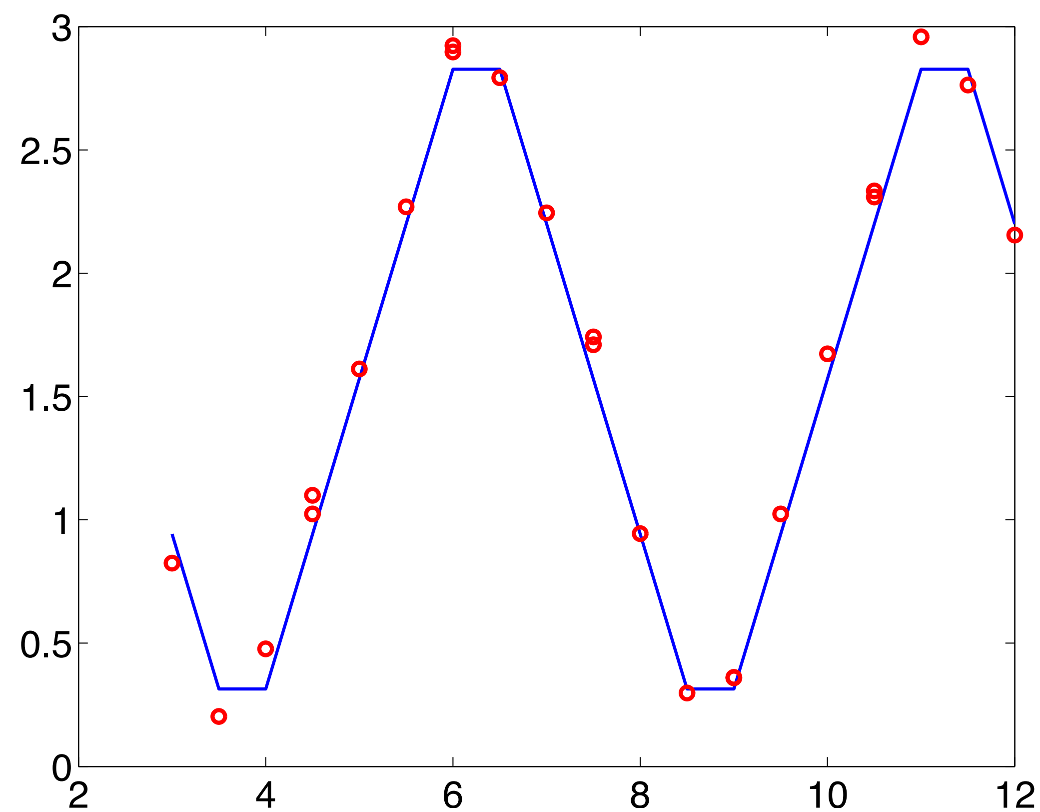
Uses <10 spectral-projected gradient iterations

Source wavelet

Estimates source function from 3 to 12 Hz



magnitude



phase

Observation

- sparsity recovery and *randomized* subsampling can lead to a significant speedup
- Curvelet transform is efficient in representing geological models
- sparse regularization in Curvelet domain can greatly suppress incoherent artifacts

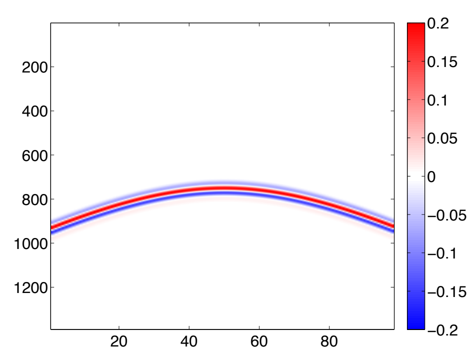
[Wang & Sacchi, '07]

Seismic imaging

Least-squares migration:

$$\delta\widetilde{\mathbf{m}} = \arg\min_{\delta\mathbf{m}} \frac{1}{2}\|\delta\mathbf{d} - \alpha\nabla\boldsymbol{\mathcal{F}}(\mathbf{m}_0)\delta\mathbf{m}\|_2^2$$

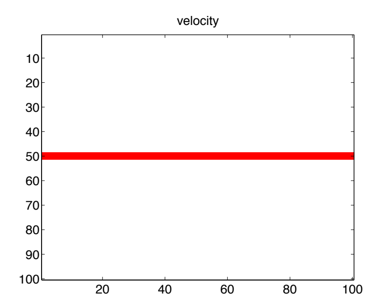
$\mathbf{b}$



$\mathbf{A}$

Jacobian operator (born modeling operator)

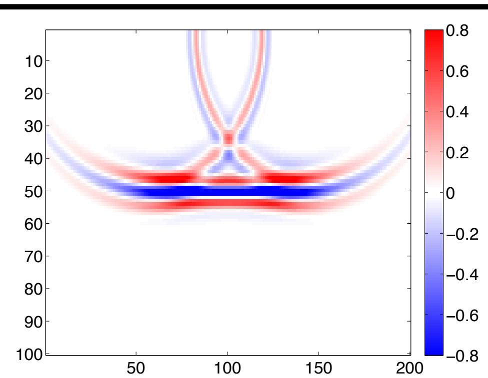
$\mathbf{x}$



$$\begin{aligned}
\delta\mathbf{d} &= \text{Multi-source multi-frequency data residue}\\
\nabla\boldsymbol{\mathcal{F}}(\mathbf{m}_0) &= \text{Linearized Born-scattering operator}\\
\mathbf{m}_0 &= \text{Background velocity model}\\
\mathbf{Q} &= \text{Sources}\\
\delta\tilde{\mathbf{m}} &= \text{image}
\end{aligned}$$

Hint:

$\mathbf{A}^T\mathbf{b} =$



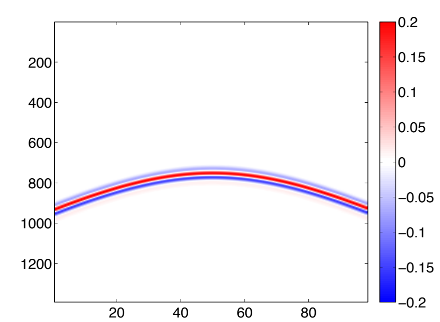
[Donoho Chen, '06; Li&Herrmann, '10]

Sparsity-promoting imaging

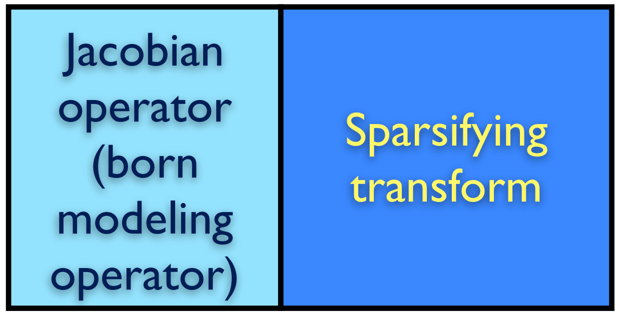
Basis pursuit denoising problem:

$$\delta\widetilde{\mathbf{m}} = \mathbf{C}^T \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|\delta\mathbf{d} - \alpha\nabla\boldsymbol{\mathcal{F}}[\mathbf{m}_0;\mathbf{Q}]\mathbf{C}^T\delta\mathbf{x}\|_2 \le \sigma$$

b



A

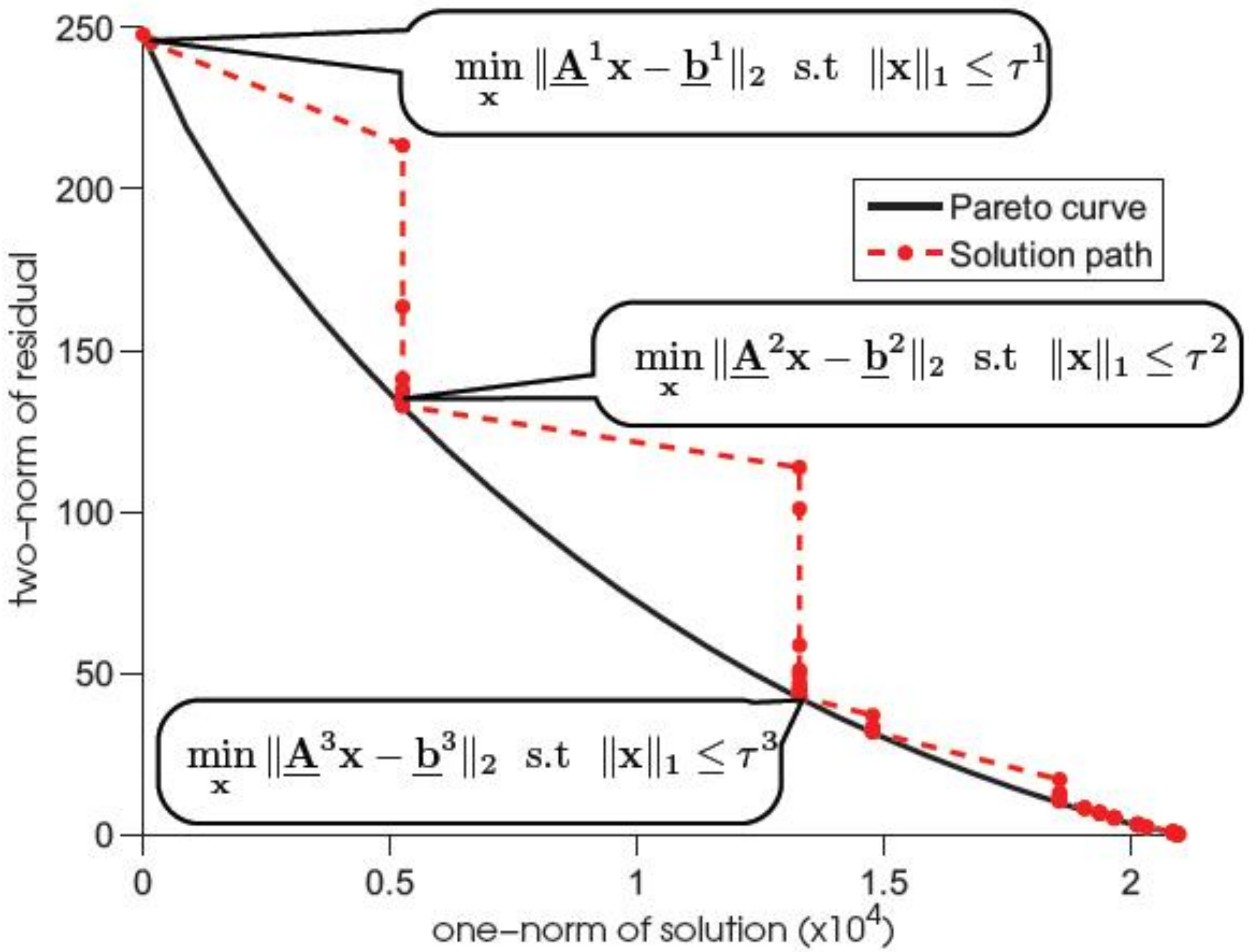


$$\delta\mathbf{x} = \text{Sparse curvelet-coefficient vector}$$

$$\mathbf{C}^T = \text{Curvelet synthesis}$$

Remarkable speedup of convergence can be obtained by *Message passing*

Pareto curve



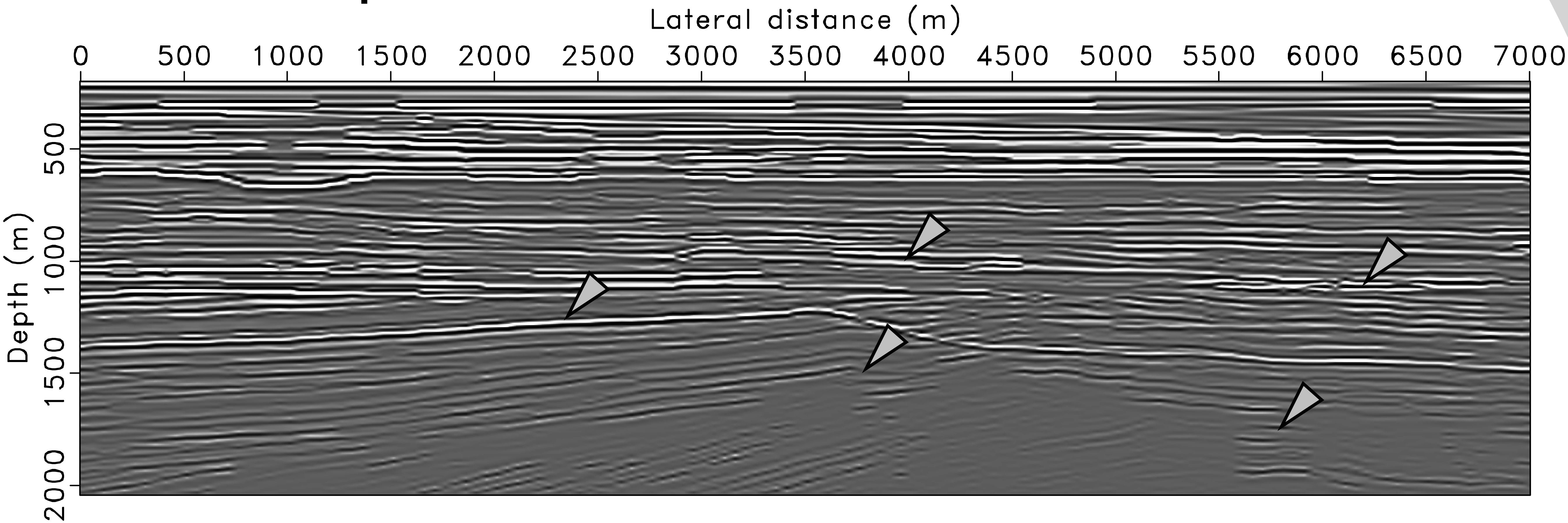
$$\delta\widetilde{\mathbf{m}} = \mathbf{C}^T \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|\delta\underline{\mathbf{d}} - \alpha\nabla\boldsymbol{\mathcal{F}}[\mathbf{m}_0;\underline{\mathbf{Q}}]\mathbf{C}^T\delta\mathbf{x}\|_2 \leq \sigma$$

Examples

- 409x1401 with mesh size of 5m
- 10 randomly selected frequencies (30-50Hz)
- 3 randomly combined simultaneous shots / 17 randomly selected sequential shots
- 60 iterations with 10 redraws

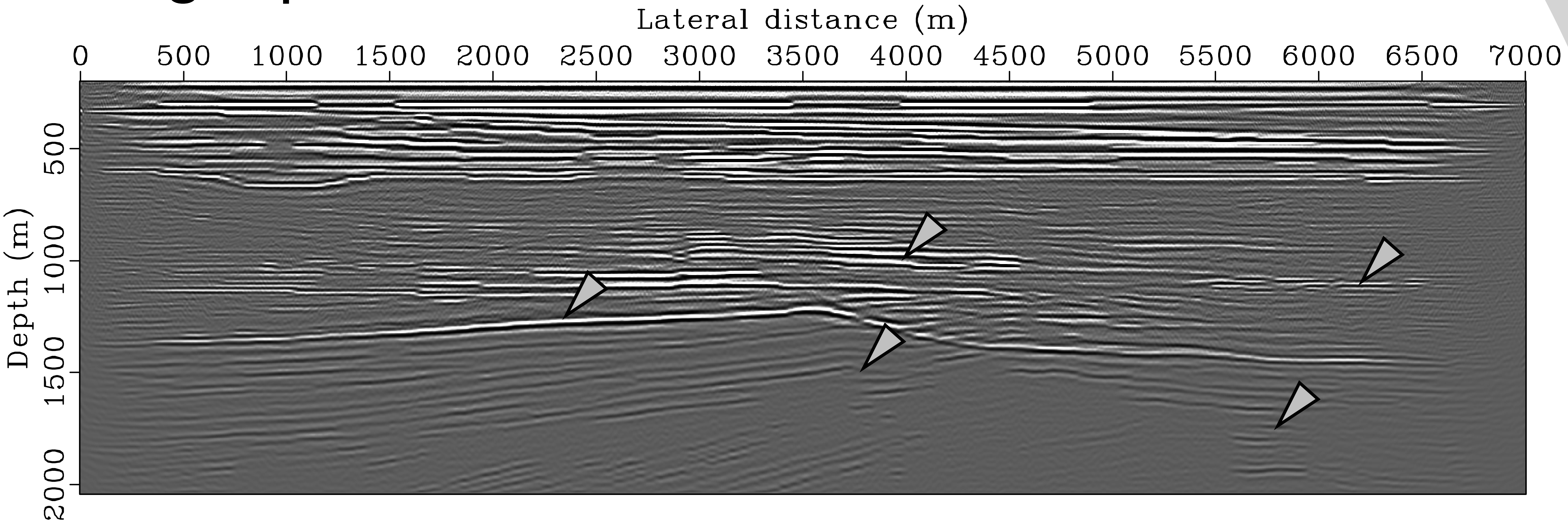
Migration results

True perturbation



Migration results
underdetermined

imaged perturbation with L1 with AMP

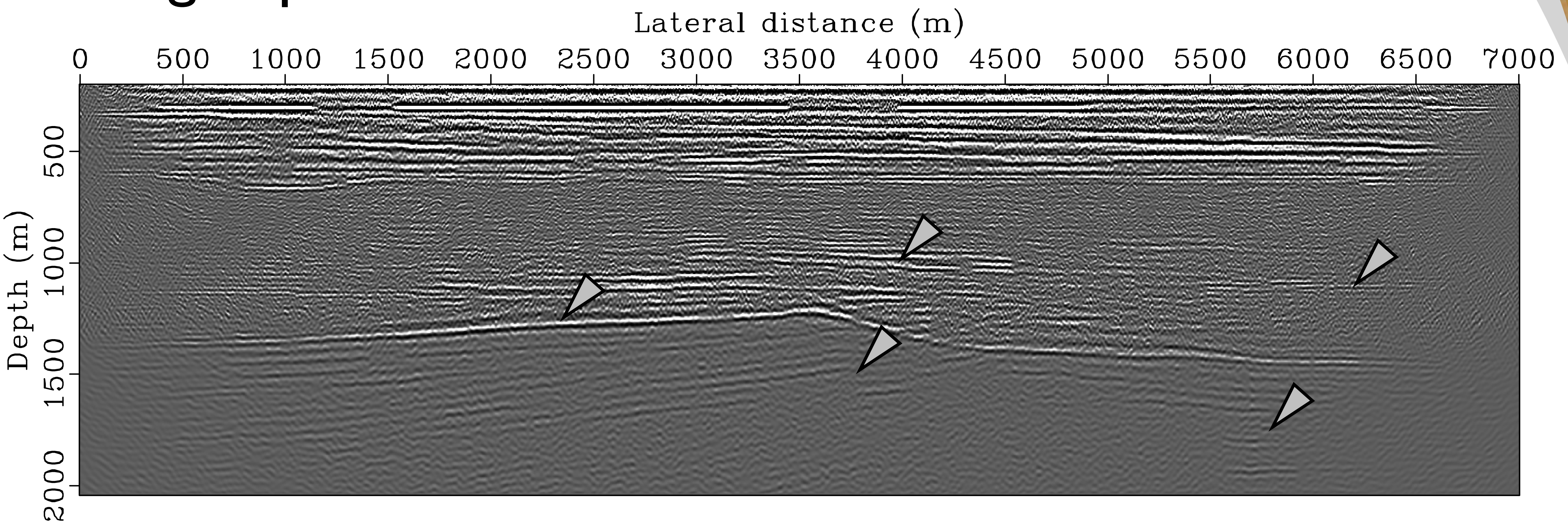


3 *simultaneous* shots

time: ~= 24 hours

Migration results
underdetermined

imaged perturbation with L2 with AMP



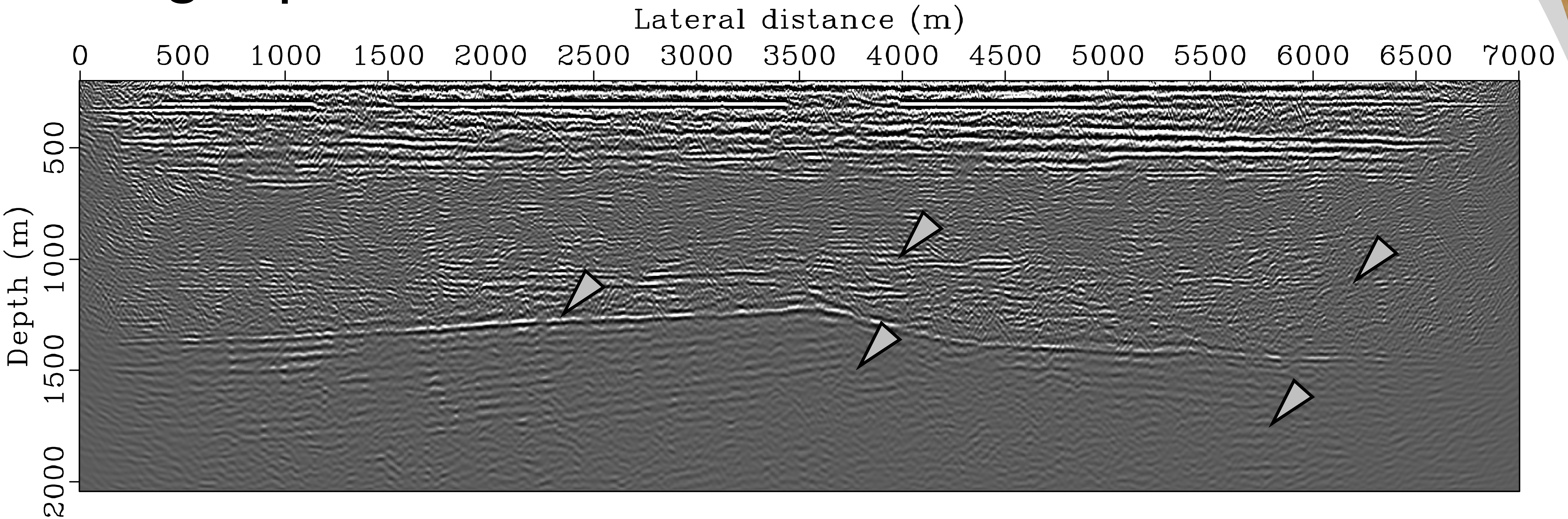
3 *simultaneous* shots

time: ~= 24 hours

SLIM

Migration results
underdetermined

imaged perturbation with L1 without AMP

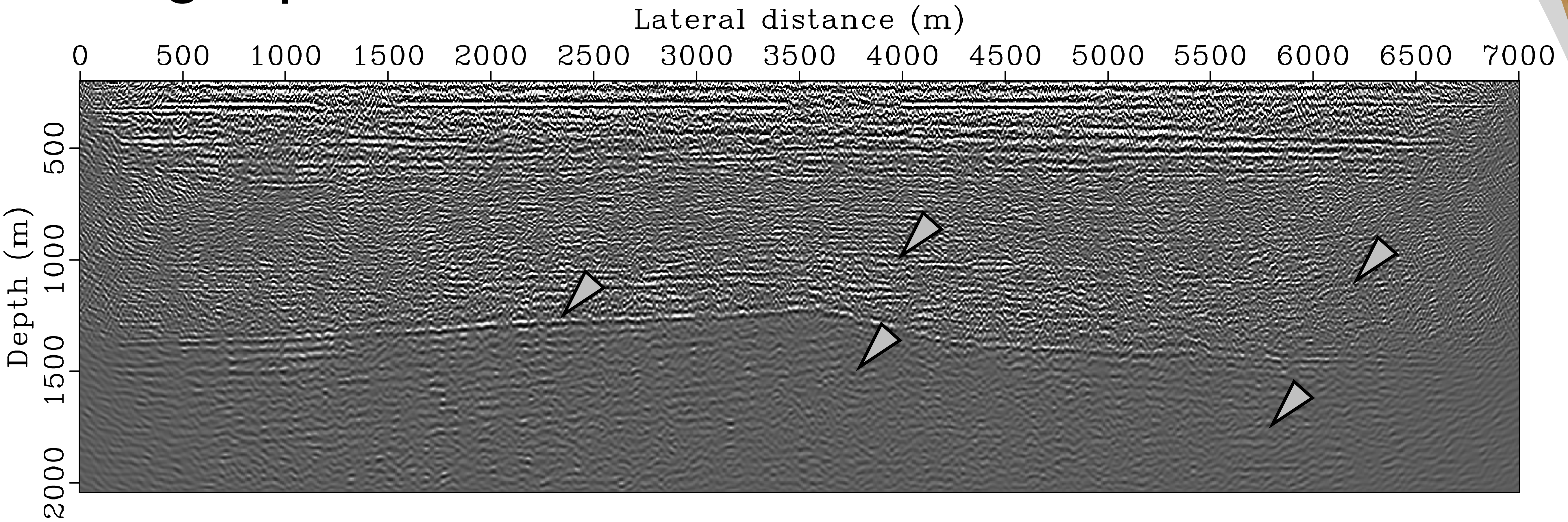


3 *simultaneous* shots

time: ~= 24 hours

Migration results
underdetermined

imaged perturbation with L2 without AMP

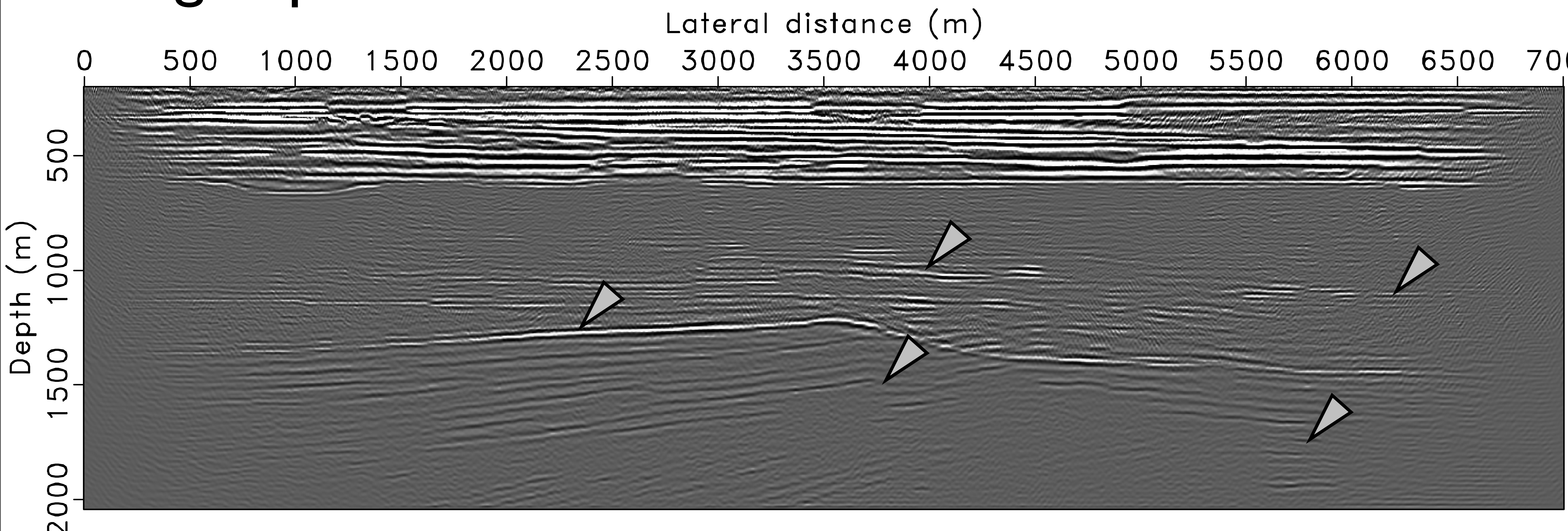


3 *simultaneous* shots

time: ~= 24 hours

Migration results
underdetermined

imaged perturbation with L1 with AMP

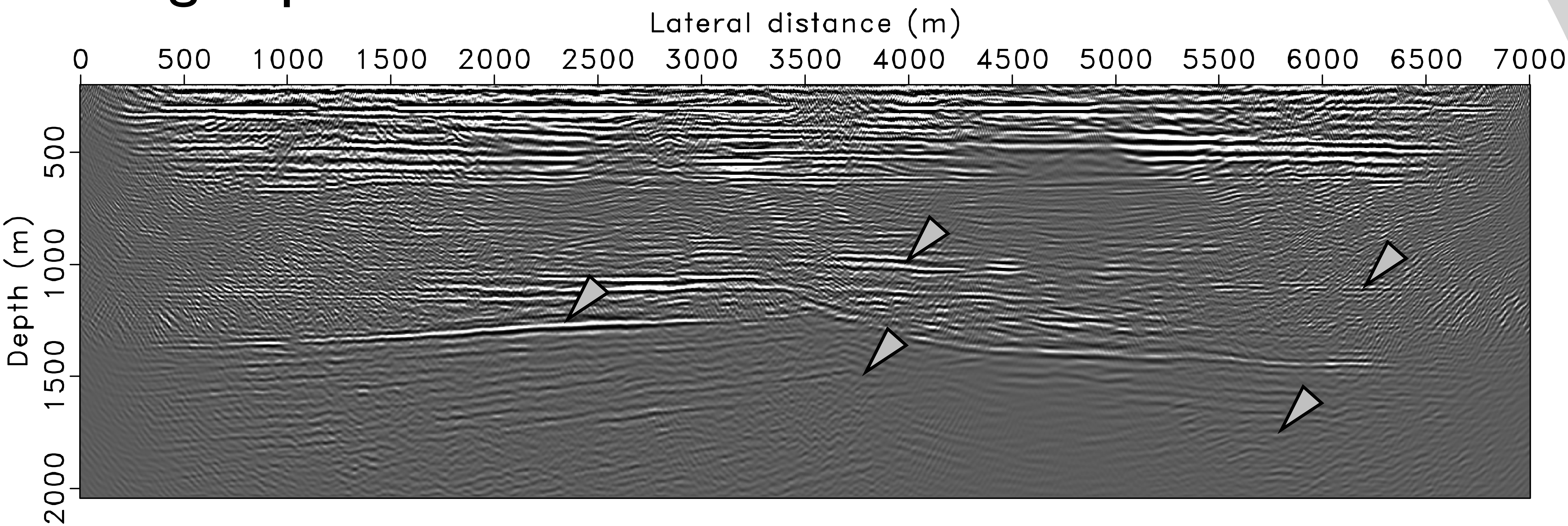


17 *sequential* shots with marine acquisition

time: ~= 40 hours

Migration results
underdetermined

imaged perturbation with L1 without AMP



17 *sequential* shots with marine acquisition

time: ~= 40 hours

Gulf of Mexico data

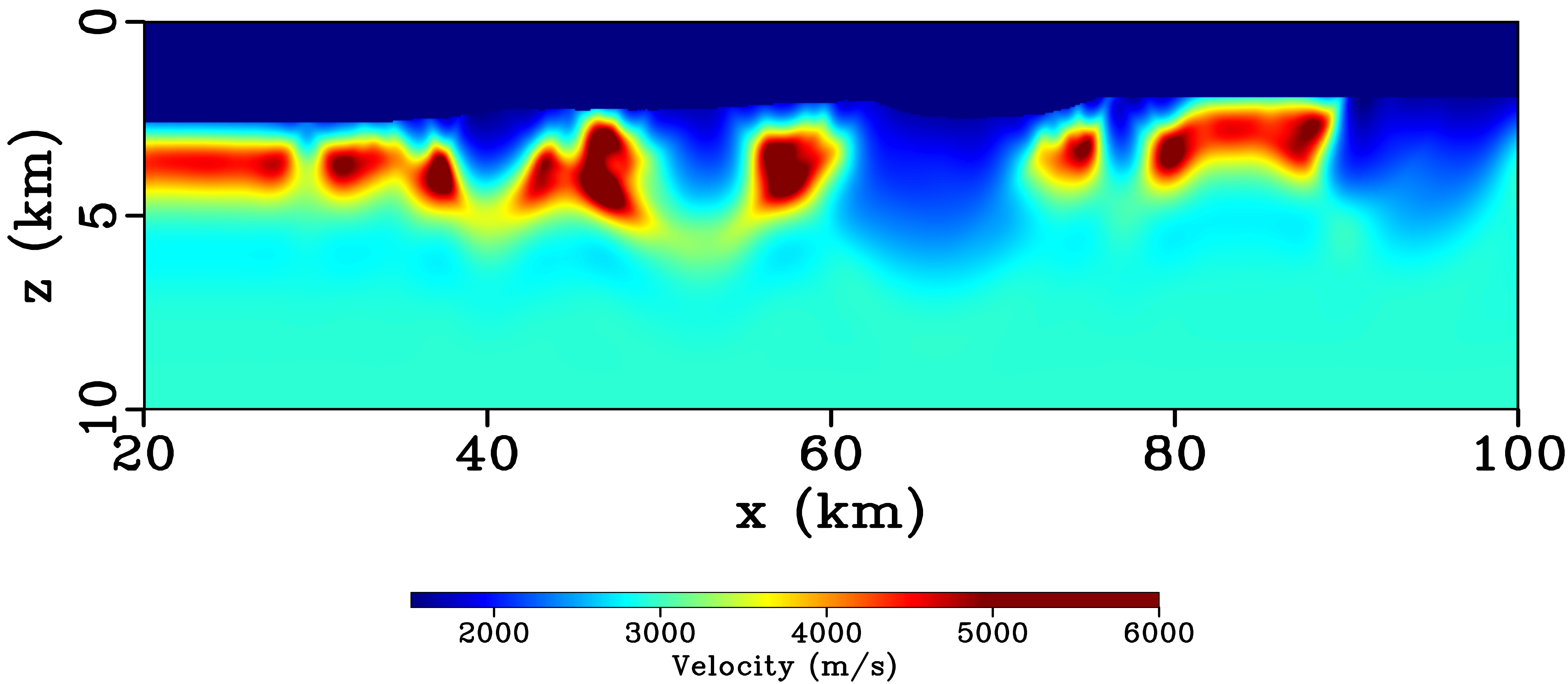
Chevron blind test

Modified Gauss-Newton

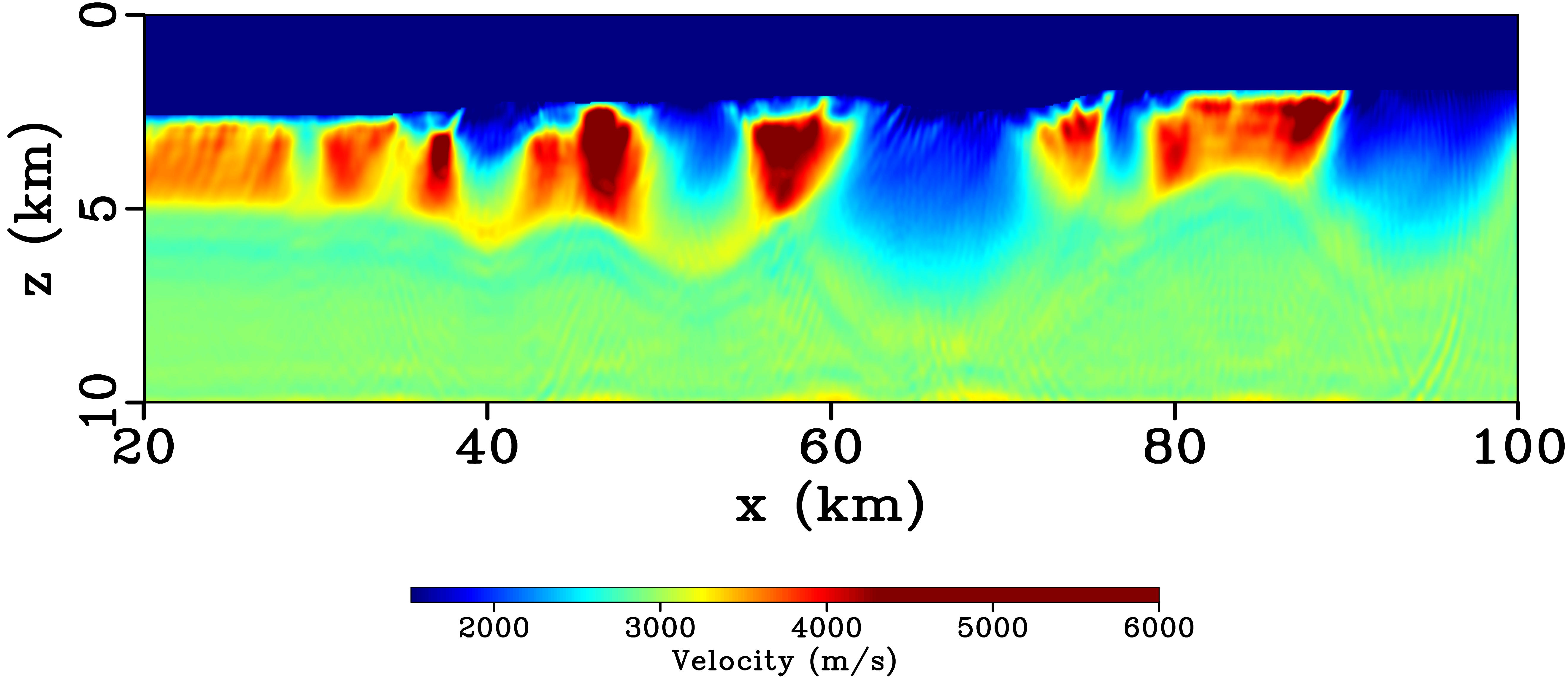
- 7 frequency bands (2-5 Hz), each contain 4 frequencies
- randomly selected 600 shots (totally 3201 shots)
- 6 Gauss-Newton iteration for each frequency band
- modeling uses Helmholtz
- depth weighting and water layer projection

Initial model

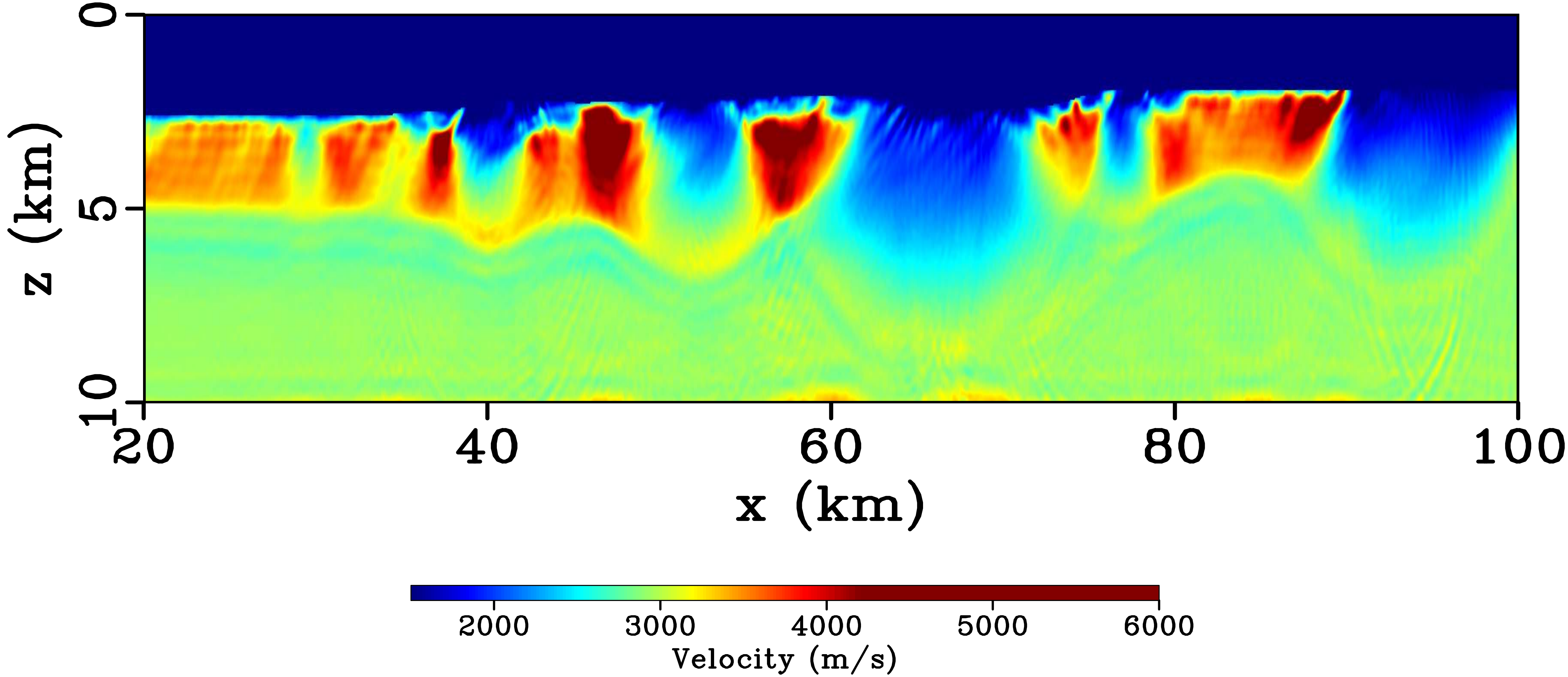
[ray-based tomography]



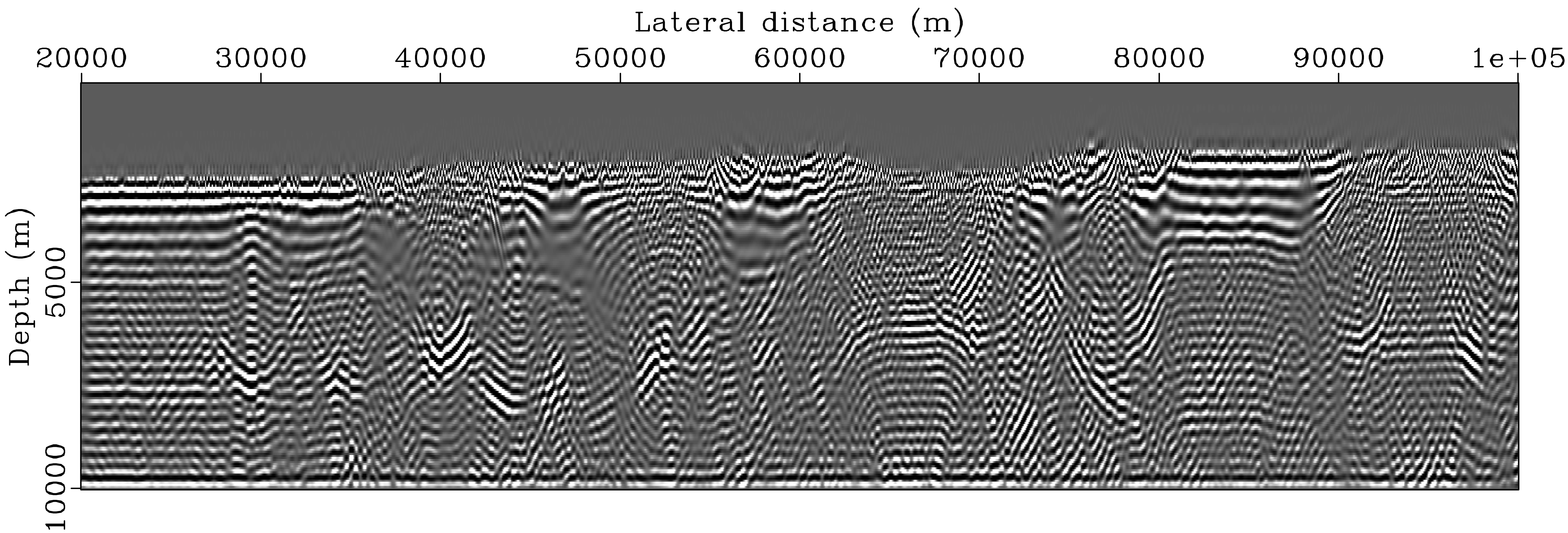
Inverted result with raw data



Inverted result with denoised data



Sparsity promoting migration



8 frequencies, 600 sequential shots

4 CPUs, < 7 days

Conclusion

- *Curvelet transform* efficiently represents geological models
- *Sparsity regularization in Curvelet domain* can significantly suppress model space artifacts
- High-resolution FWI and Imaging results is attainable through Sparsity promotion
- Ideas from message passing leads to a remarkable speedup of convergence

Acknowledgements

- Charles Jones for BG compass model
- Authors of CurveLab, SPGL1 and Spot
- My colleagues

SINBAD

NSERC CRSNG

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, BP, Chevron, ConocoPhillips, Petrobras, PGS, Total SA, and WesternGeco.

Thank you

https://www.slim.eos.ubc.ca