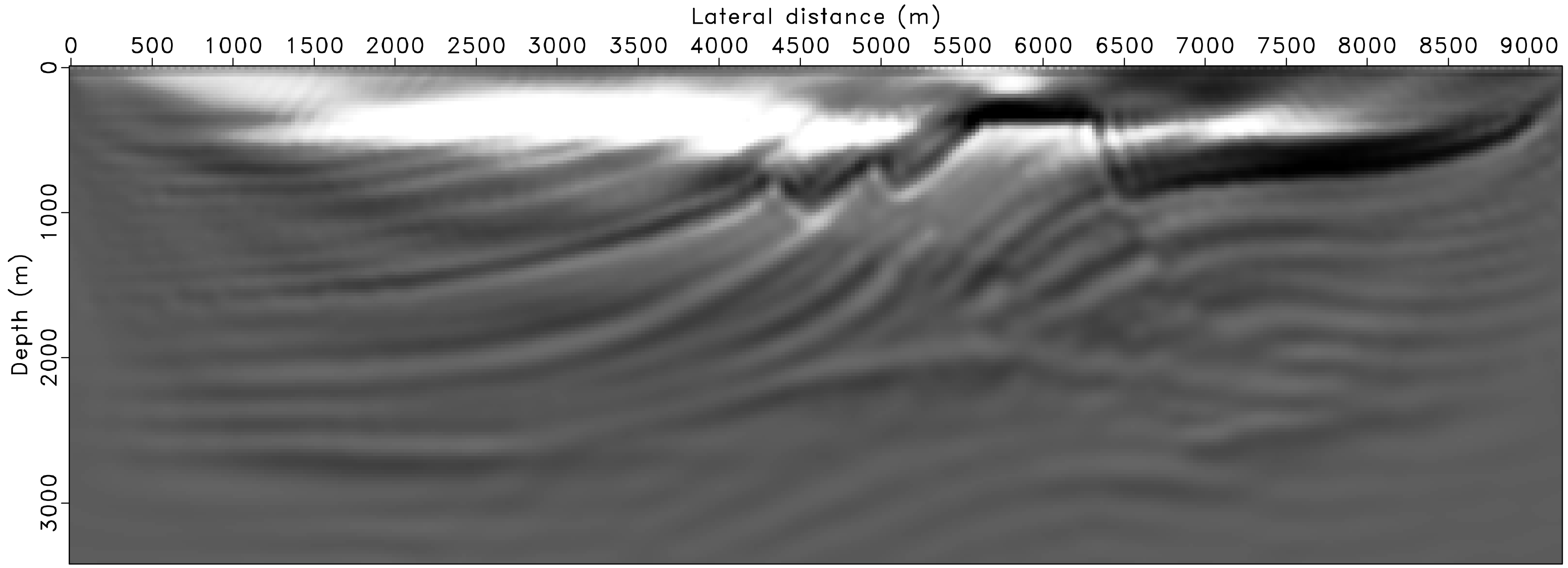
SINBAD Consortium 2010

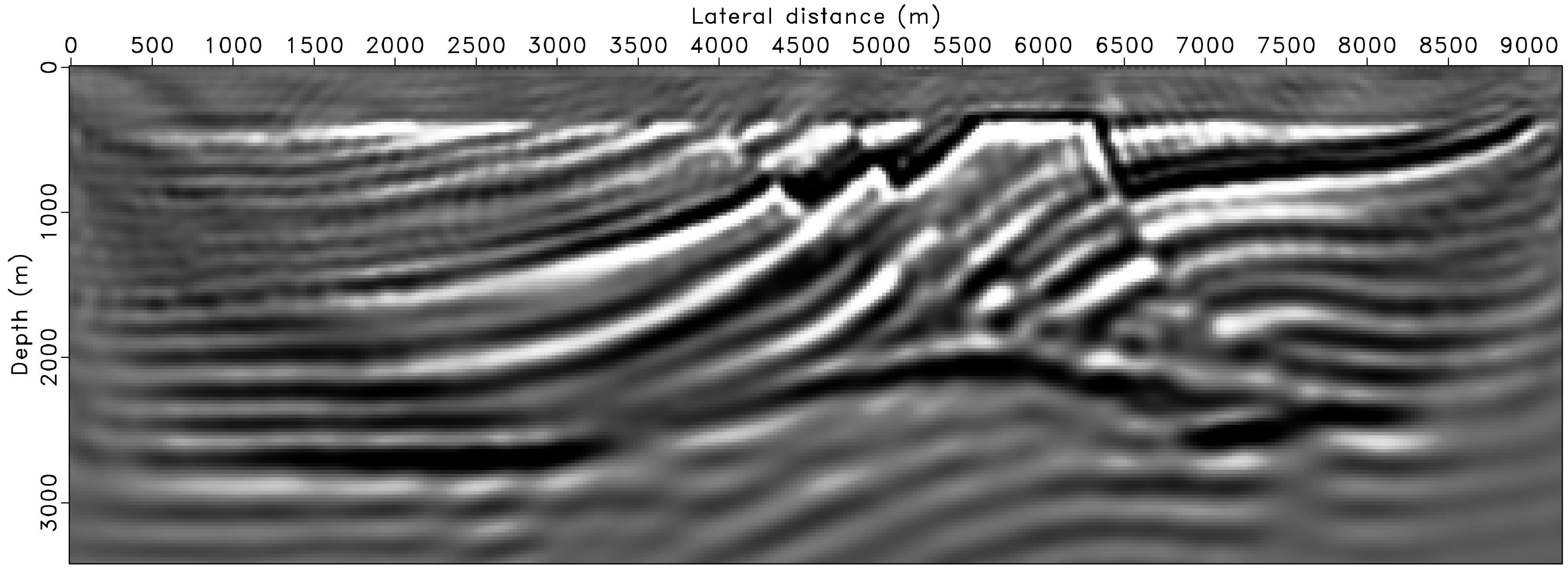
Efficient least-squares imaging with *sparsity* promotion and *compressive sensing*

Xiang Li and Felix J. Herrmann

SLIM
University of British Columbia

Migration vs *linearized* inversion





[Li & FJH et. al. '10-]

Compressive imaging

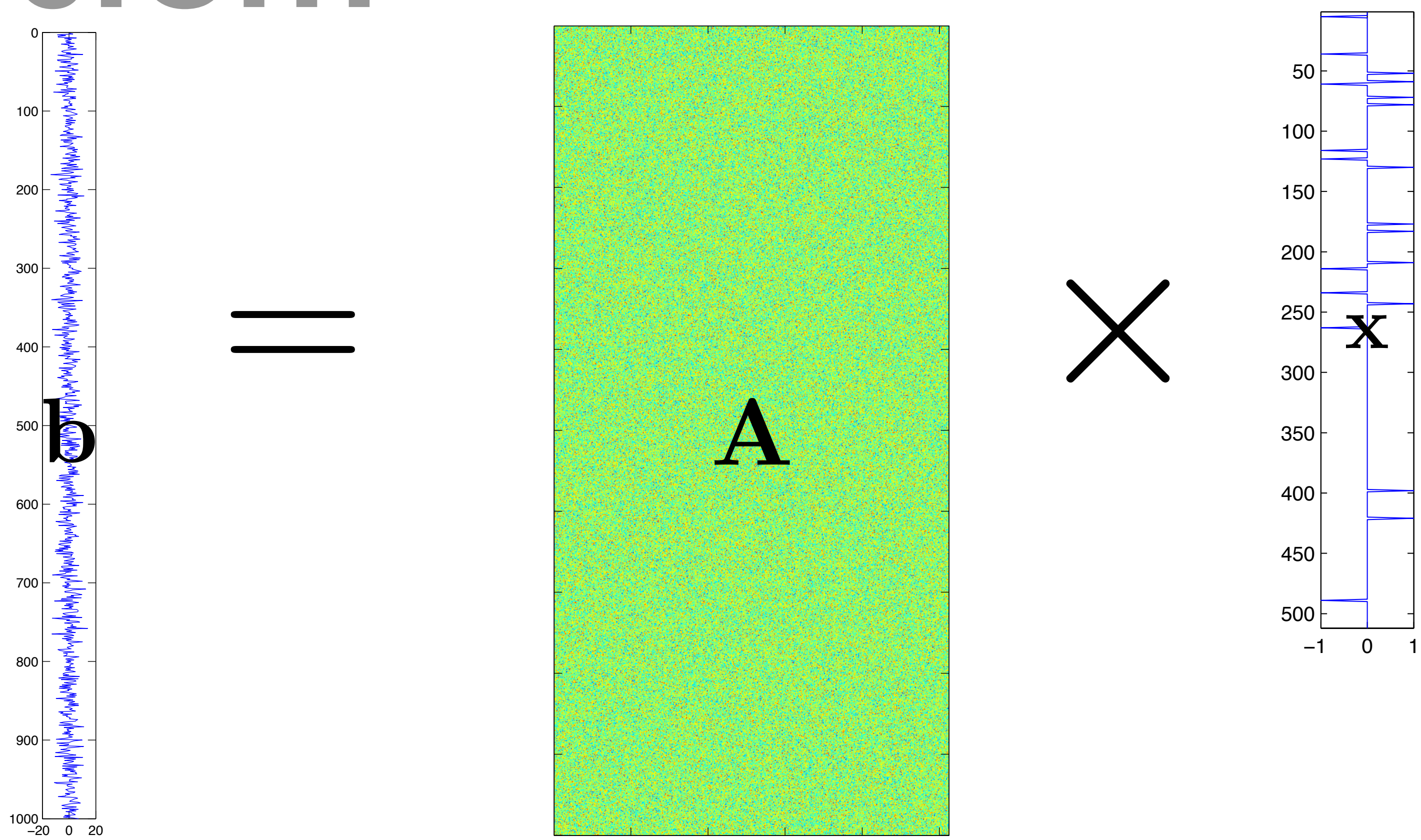
Challenge:

Least-squares migration requires *multiple passes & PDE solves*

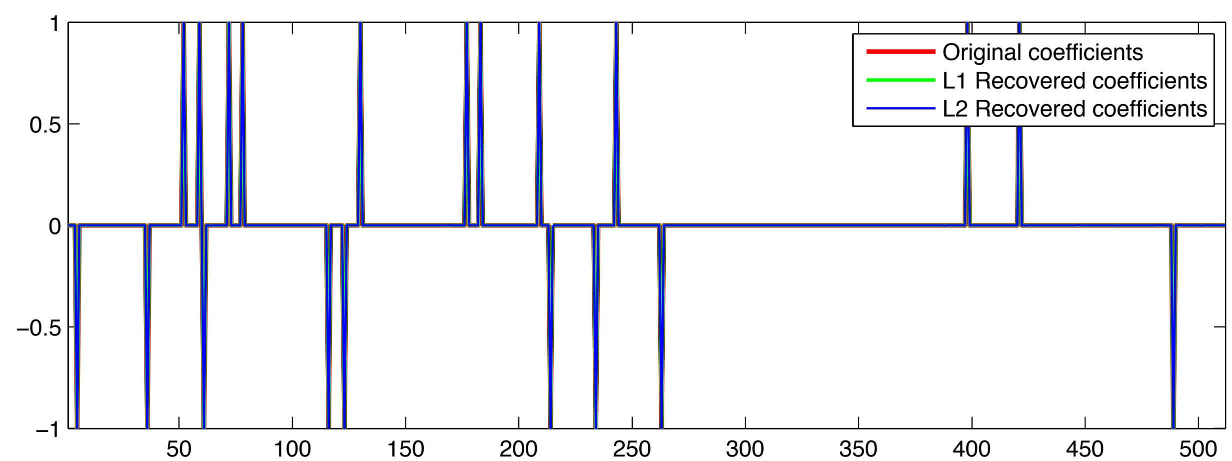
Key idea:

- combine *Compressive Sensing* & '*Phase encoding*'
- turn "*overdetermined*" imaging problem into *underdetermined* problem with *randomized* supershots
- use *curvelet*-based *sparse* recovery to remove *crosstalk*

Overdetermined system

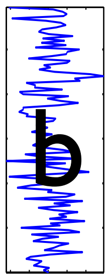


L1/L2 recovery

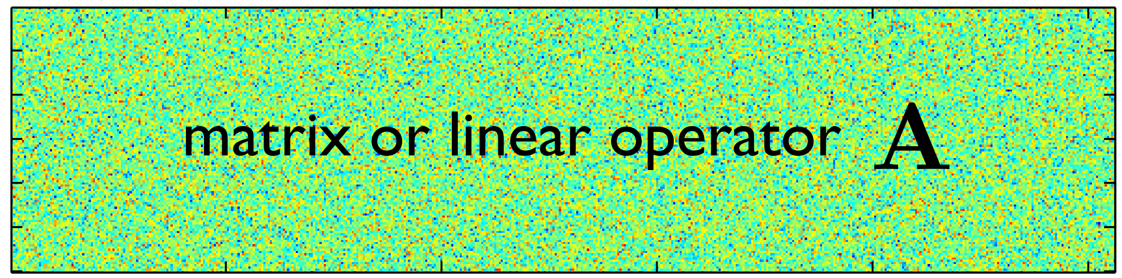


$$\mathbf{x} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1 or \ell 2} \quad \text{subject to} \quad \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \sigma$$

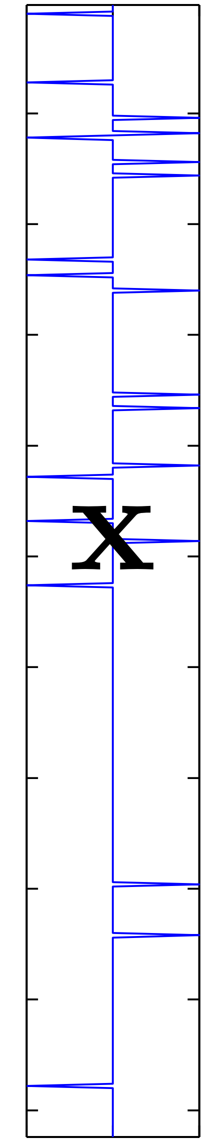
Underdetermined system



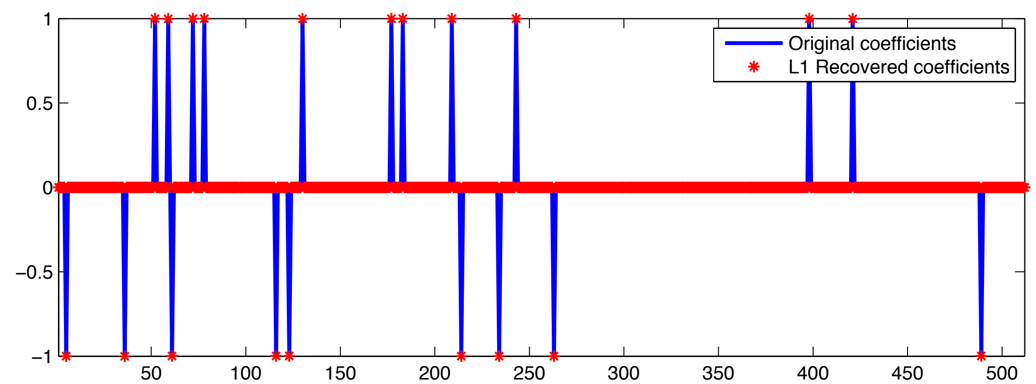
$=$



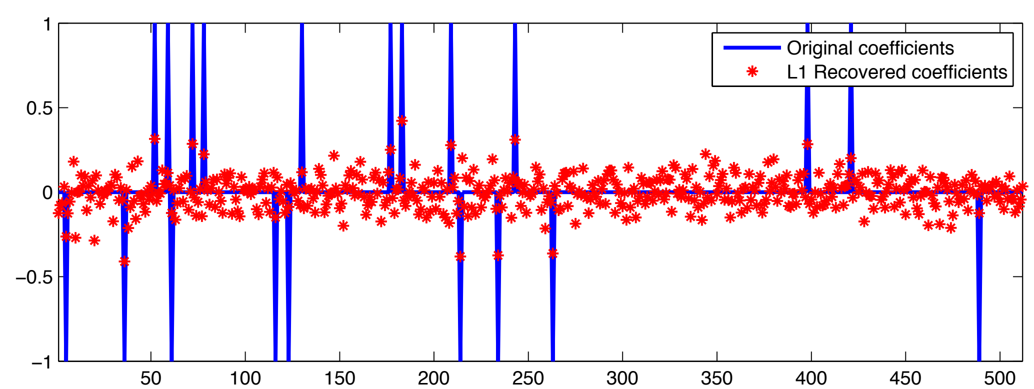
$\times$



L1 recovery



L2 recovery

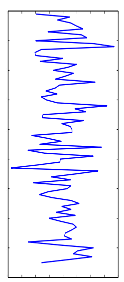


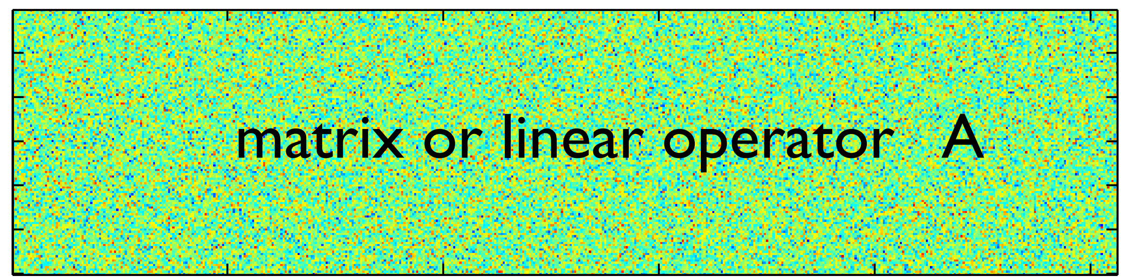
$$\mathbf{x} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1 or \ell 2} \quad \text{subject to} \quad \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \leq \sigma$$

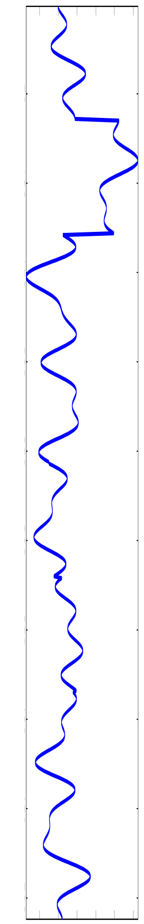
[van den Berg & Friedlander, ’08]

Compressible signals

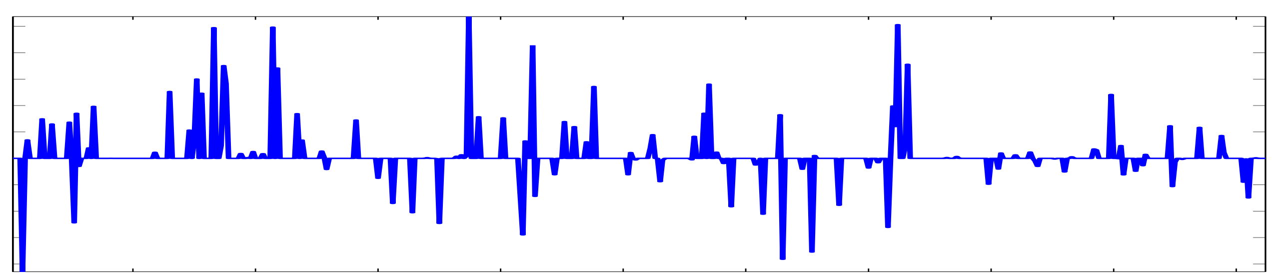
What if the signal is not sparse?







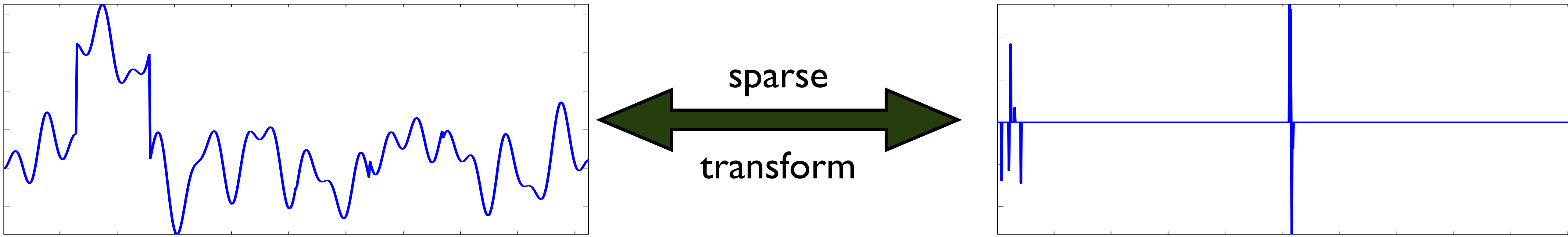
L1 recovery



$$\mathbf{x} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|b - A\mathbf{x}\|_2 \leq \sigma$$

Sparsifying domains

wavelet, curvelet, shearlet, wave atom...

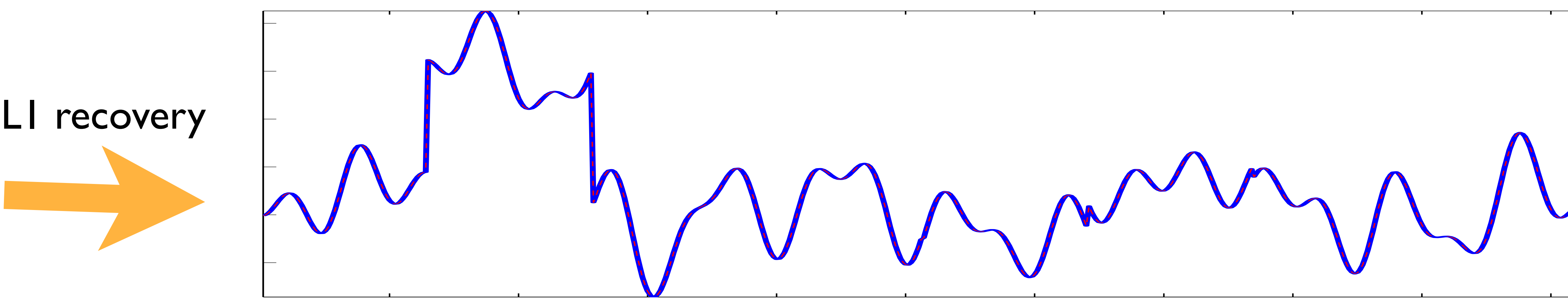


physical domain

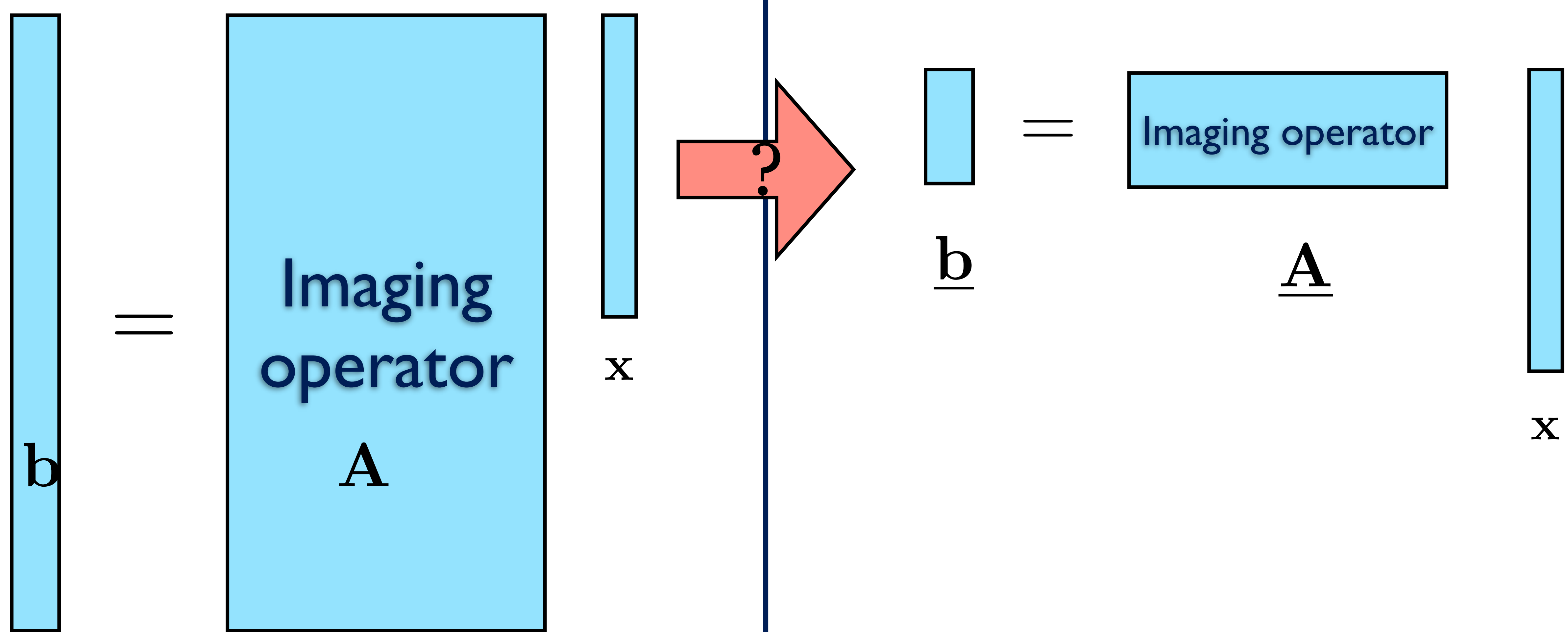
Sparse domain

$$x = \mathbf{S}^* \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|b - A\mathbf{S}^*\delta\mathbf{x}\|_2 \leq \sigma$$

L1 recovery



Dimensionality-reduced imaging



SLIM

[Nemeth et. al. '99]

Linearized inversion

Least-squares migration:

$$\delta\widetilde{\mathbf{m}} = \arg\min_{\delta\mathbf{m}} \frac{1}{2}\|\delta\mathbf{d} - \nabla\mathcal{F}[\mathbf{m}_0;\mathbf{Q}]\delta\mathbf{m}\|_2^2$$

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

[Nemeth et. al. '99]

Linearized inversion

Least-squares migration:

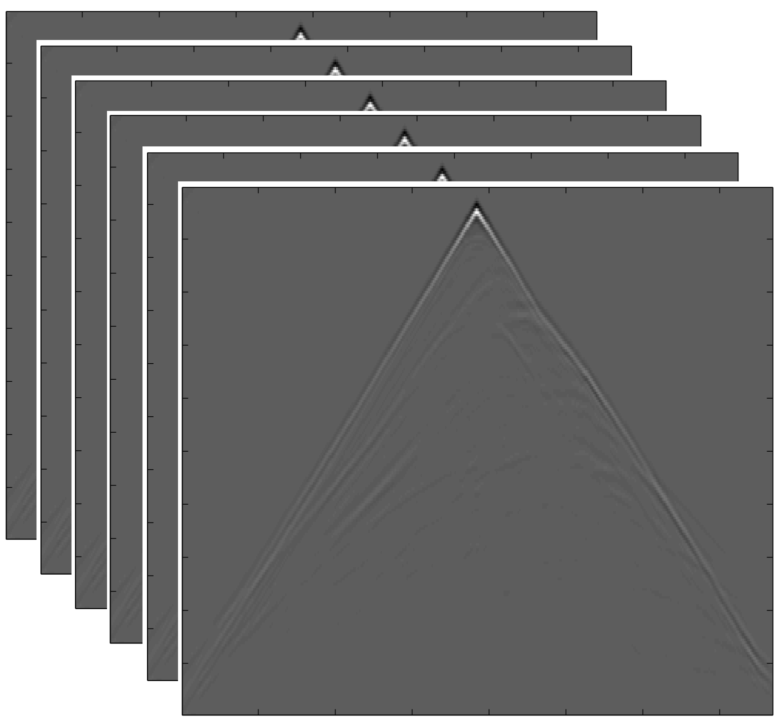
$$\delta\widetilde{\mathbf{m}} = \arg\min_{\delta\mathbf{m}} \frac{1}{2}\|\underbrace{\delta\mathbf{d}}_{\mathbf{b}} - \underbrace{\nabla\boldsymbol{\mathcal{F}}[\mathbf{m}_0;\mathbf{Q}]}_{\mathbf{A}}\underbrace{\delta\mathbf{m}}_{\mathbf{x}}\|_2^2$$

- overdetermined system, ($n_f \times n_s \times n_r$, $n_x \times n_z$)
- multiple passes trough all data

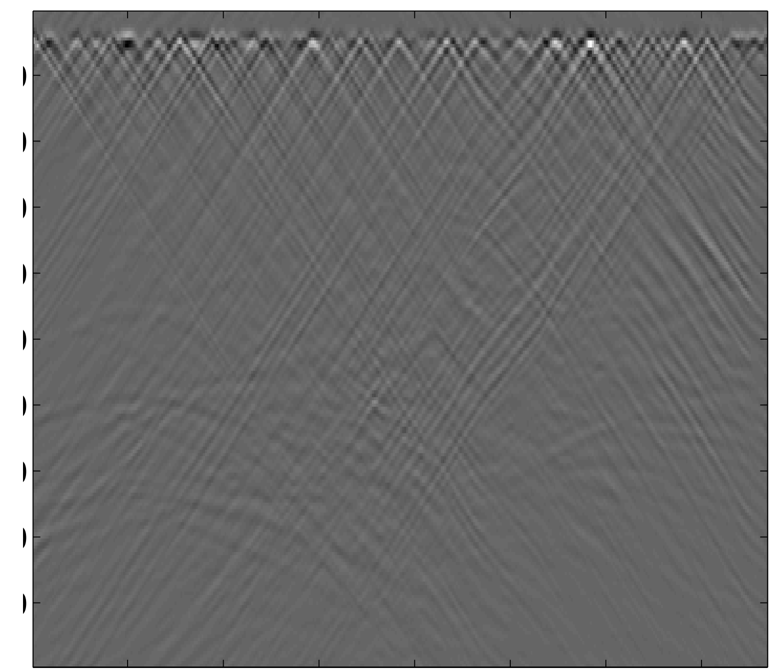
[Herrmann et. al. ’08-’10]

Dimensionality reduction

adapted from Herrmann et. al. ,09



$\mathbf{Q}$

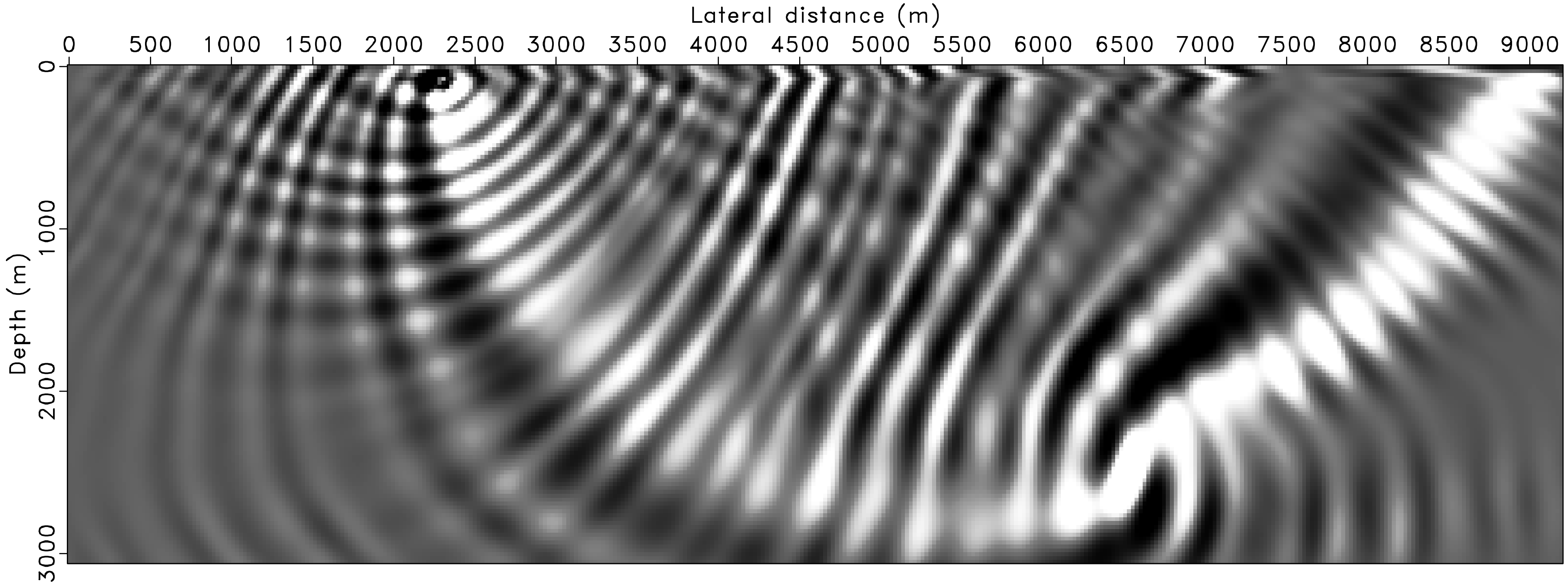


$\underline{\mathbf{Q}} = \mathbf{Q}\mathbf{W}$

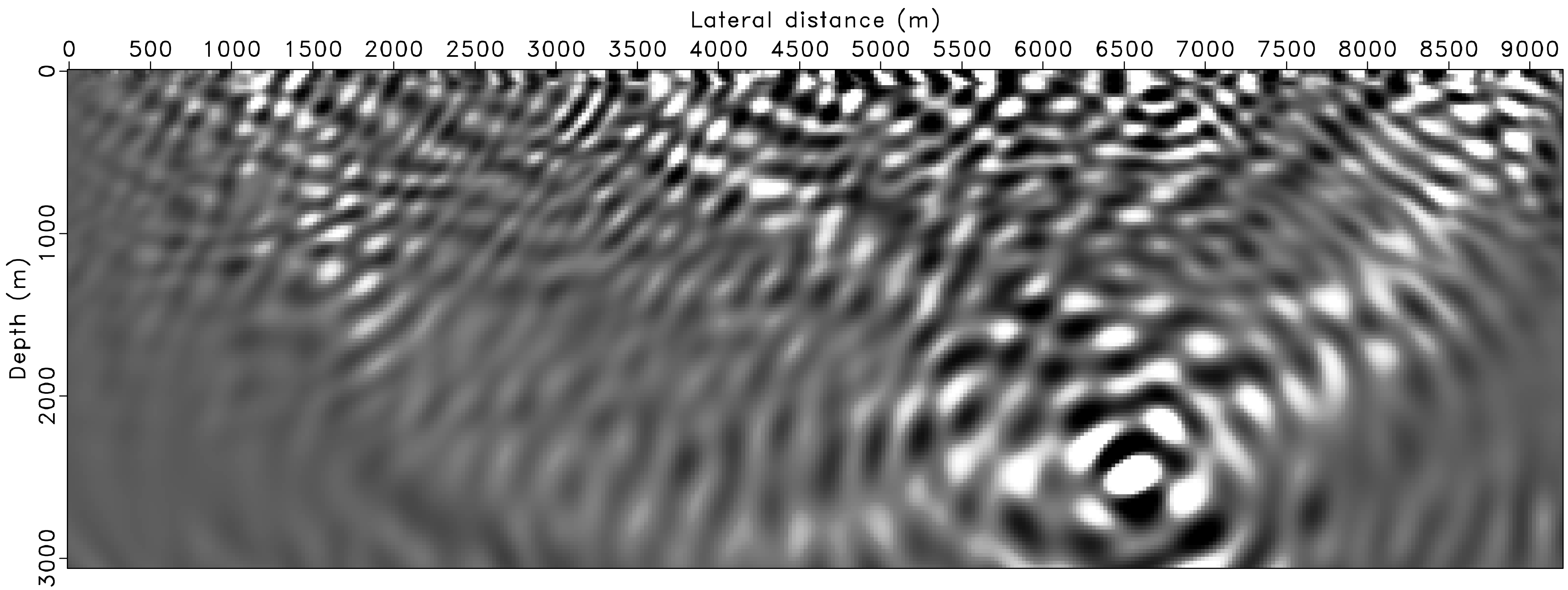
Collection of *K simultaneous-source* experiments with batch size $K \ll n_f \times n_s$

Single-shot image

one shot



Sequential shot image



Simultaneous shot image

Phase encoding

Least-squares migration:

$$\delta\widetilde{\mathbf{m}} = \arg\min_{\delta\mathbf{m}} \frac{1}{2}\|\delta\underline{\mathbf{d}} - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}_0;\underline{\mathbf{Q}}]\delta\mathbf{m}\|_2^2$$

$$\begin{aligned}\delta\underline{\mathbf{d}} &= \delta\mathbf{d}\mathbf{W} \ (\text{Simultaneous-source data residue})\\ \underline{\mathbf{Q}} &= \mathbf{Q}\mathbf{W} \ (\text{Simultaneous sources})\end{aligned}$$

[Wang & Sacchi, '07]

Sparse recovery

Least-squares migration with *sparsity* promotion

$$\min_{\delta\mathbf{x}} \frac{1}{2}\|\delta\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|\delta\underline{\mathbf{d}} - \nabla\mathcal{F}[\mathbf{m}_0;\underline{\mathbf{Q}}]\mathbf{S}^*\delta\mathbf{x}\|_2 \leq \sigma$$

$$\delta\widetilde{\mathbf{m}} = \mathbf{S}^*\delta\mathbf{x}$$

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

$\mathbf{S}^*$ = Curvelet synthesis

Migration results

Time-harmonic Helmholtz:

- 409 X 1401 with mesh size of 5m
- 9 point stencil [C. Jo et. al., ’96]
- absorbing boundary condition with damping layer with thickness proportional to wavelength
- solve wavefields on the fly with direct solver

SLIM

Migration results

Split-spread surface-free 'land' acquisition:

- 350 sources with sampling interval 20m
- 701 receivers with sampling interval 10m
- maximal offset 7km (3.5 X depth of model)
- Ricker wavelet with central frequency of 30Hz
- recording time for each shot is 3.6s

Migration results

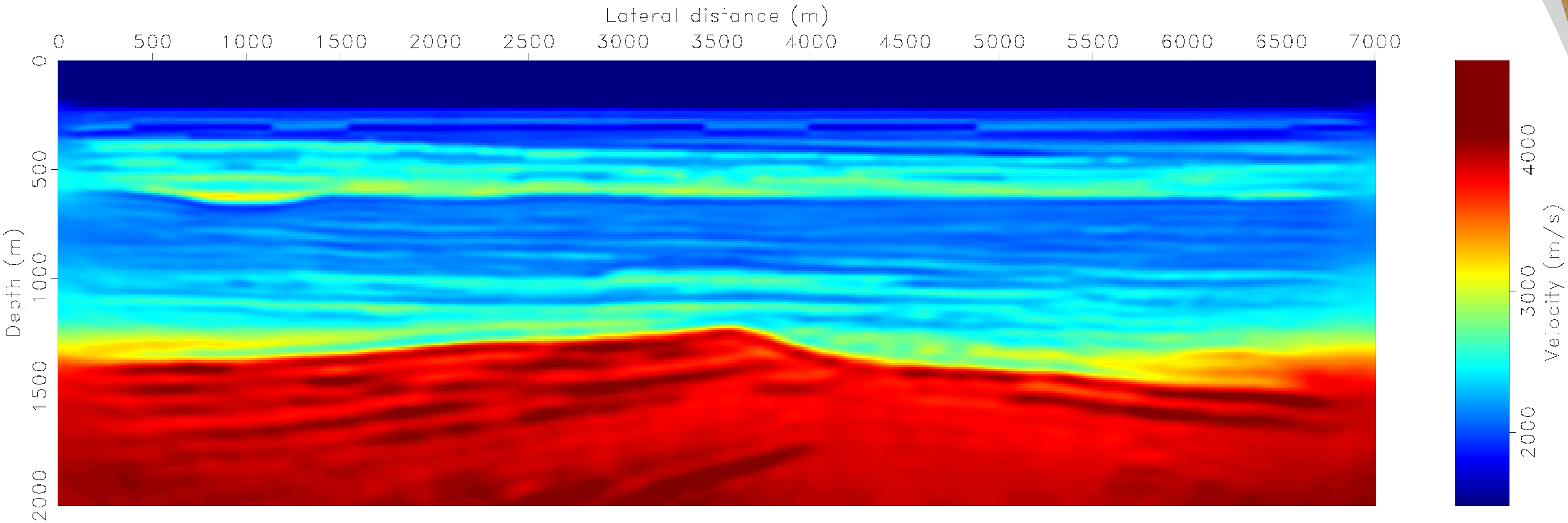
Migration:

- 10 random frequencies (20Hz-50Hz)
- 17 *simultaneous* shots (*versus 350 sequential shots*)
- LASSO problems determined by SPGL1

[Herrmann & Xi, 2011]

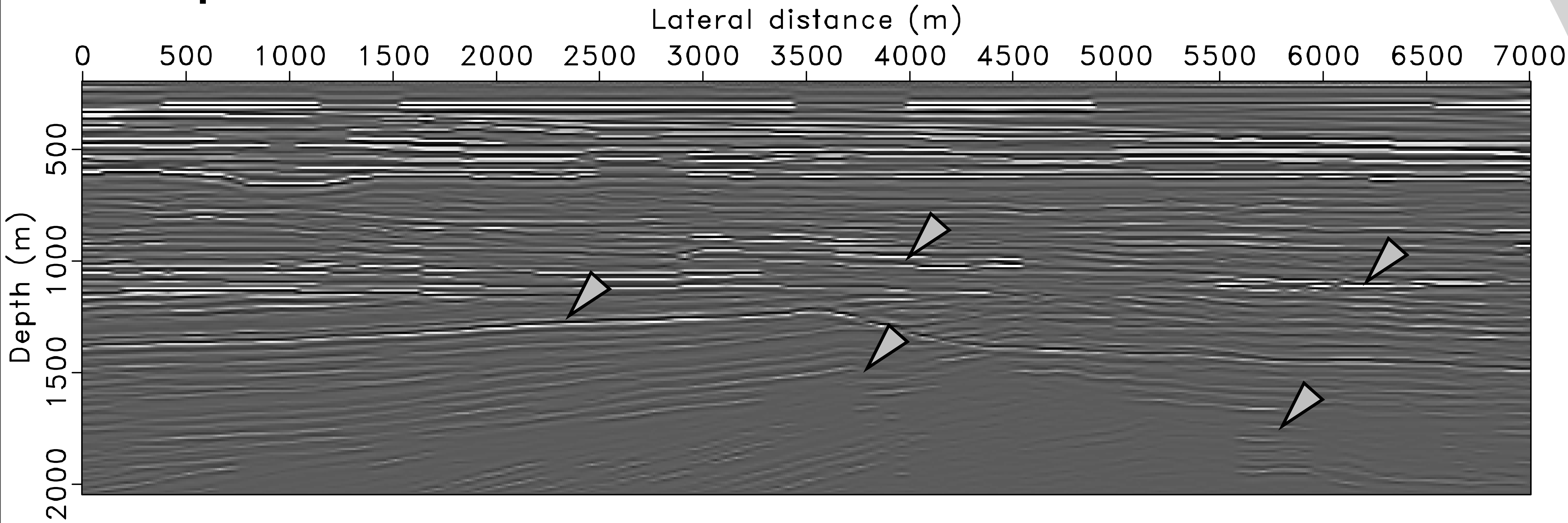
Migration results

true perturbation



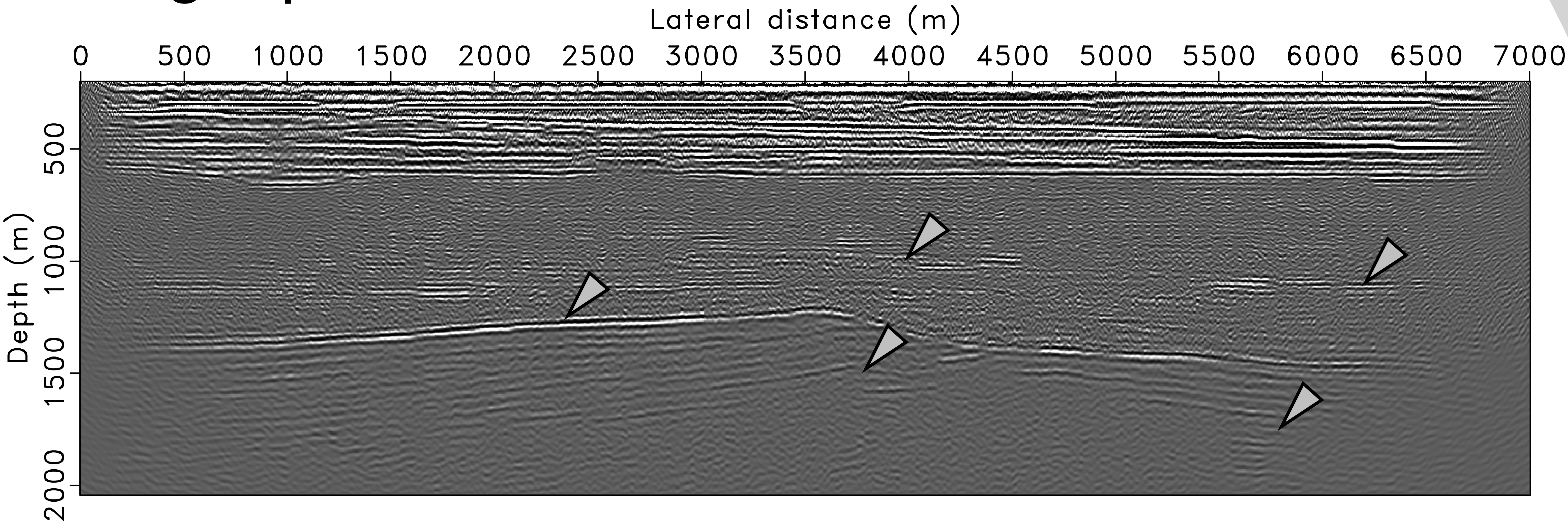
Migration results

true perturbation



Migration results
overdetermined

imaged perturbation with L2

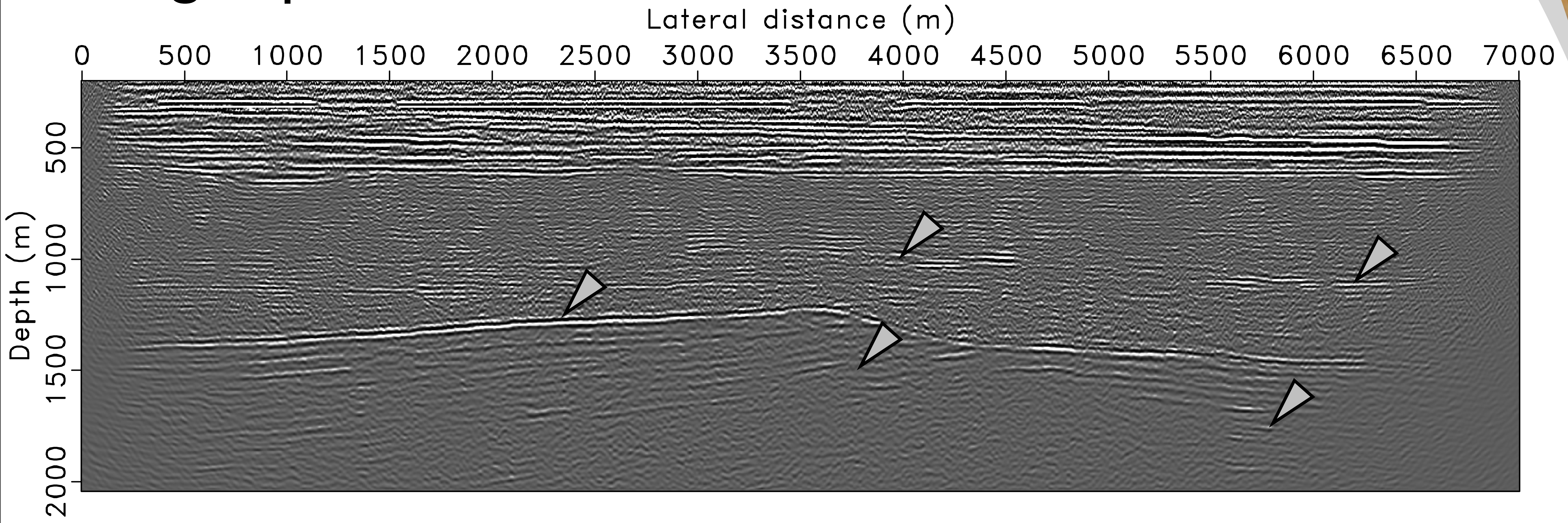


71 *sequential* shots

# of PDEs: 56800

Migration results
underdetermined

imaged perturbation with L1

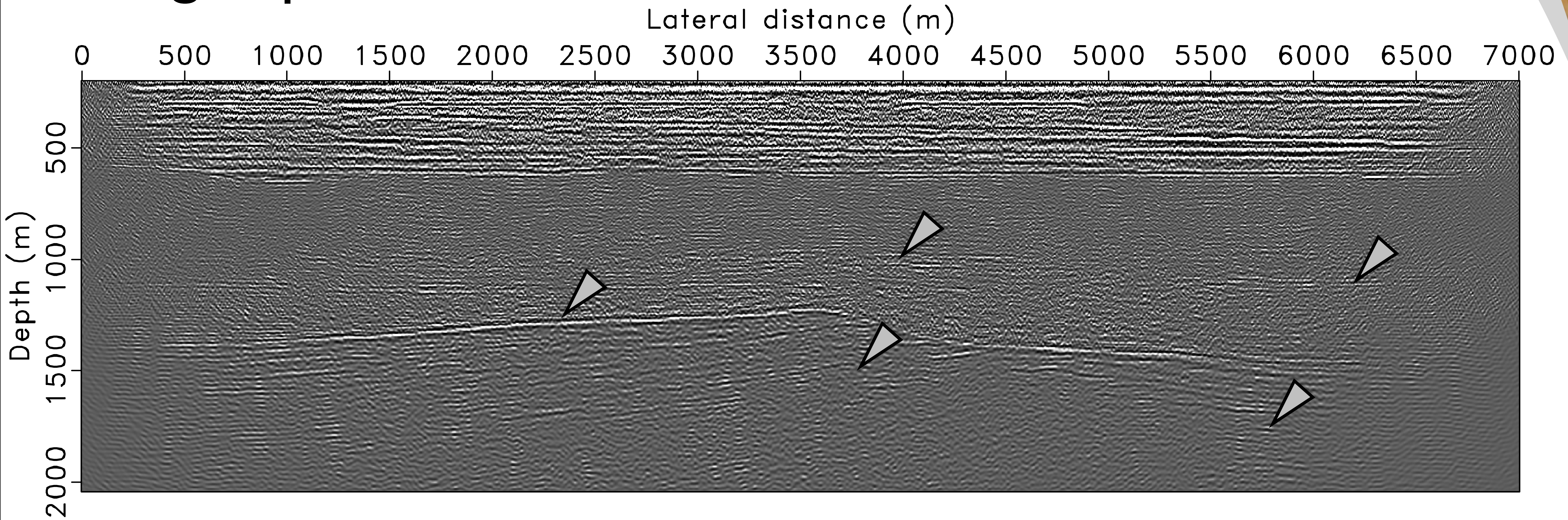


with 17 *simultaneous* shots

# of PDEs: 34000

Migration results

imaged perturbation with L2

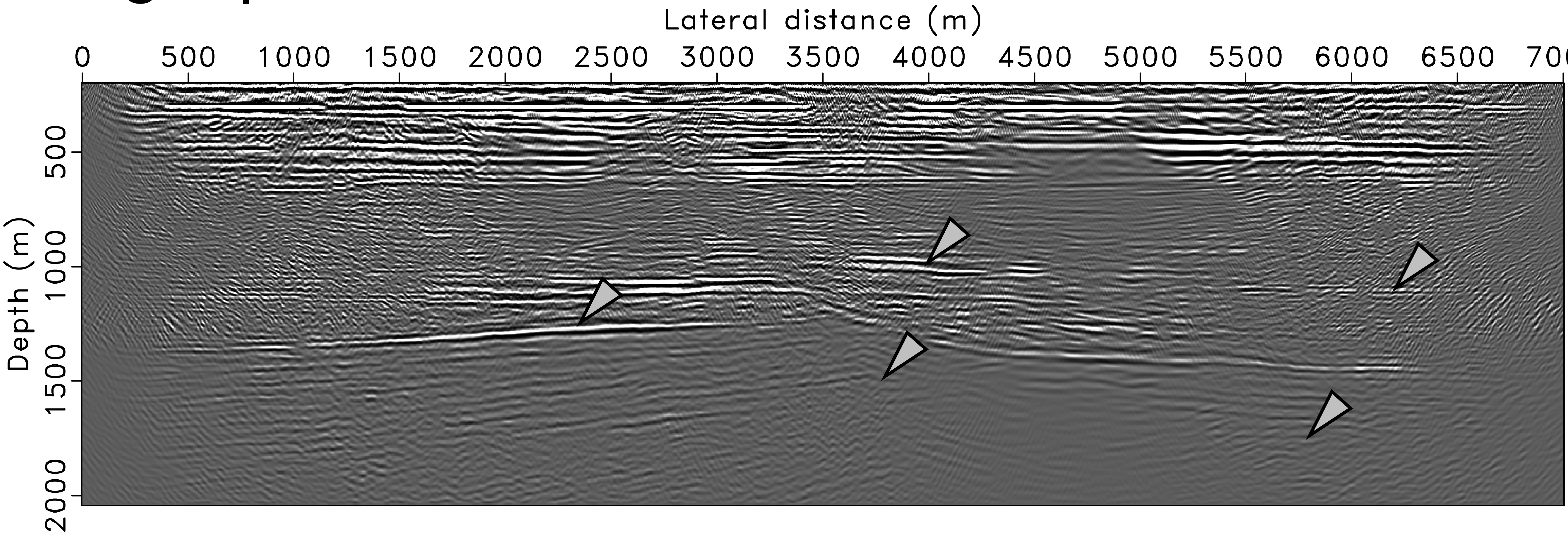


with 17 simultaneous shots

# of PDEs: 34000

Migration results
underdetermined

imaged perturbation with L1



with 17 *sequential* shots

# of PDEs: 34000

Observations

- $\ell 1$ regularization works better than $\ell 2$ for *underdetermined* systems
- Since $\ell 1$ relies on the *sparsity*, it is more *efficient* in removing Gaussian type *noise*
- Migration artifact can be reduced by using *simultaneous* shots instead of *sequential* shots
- By turning *overdetermined* system into *underdetermined* system, we can save on demans for computational resources

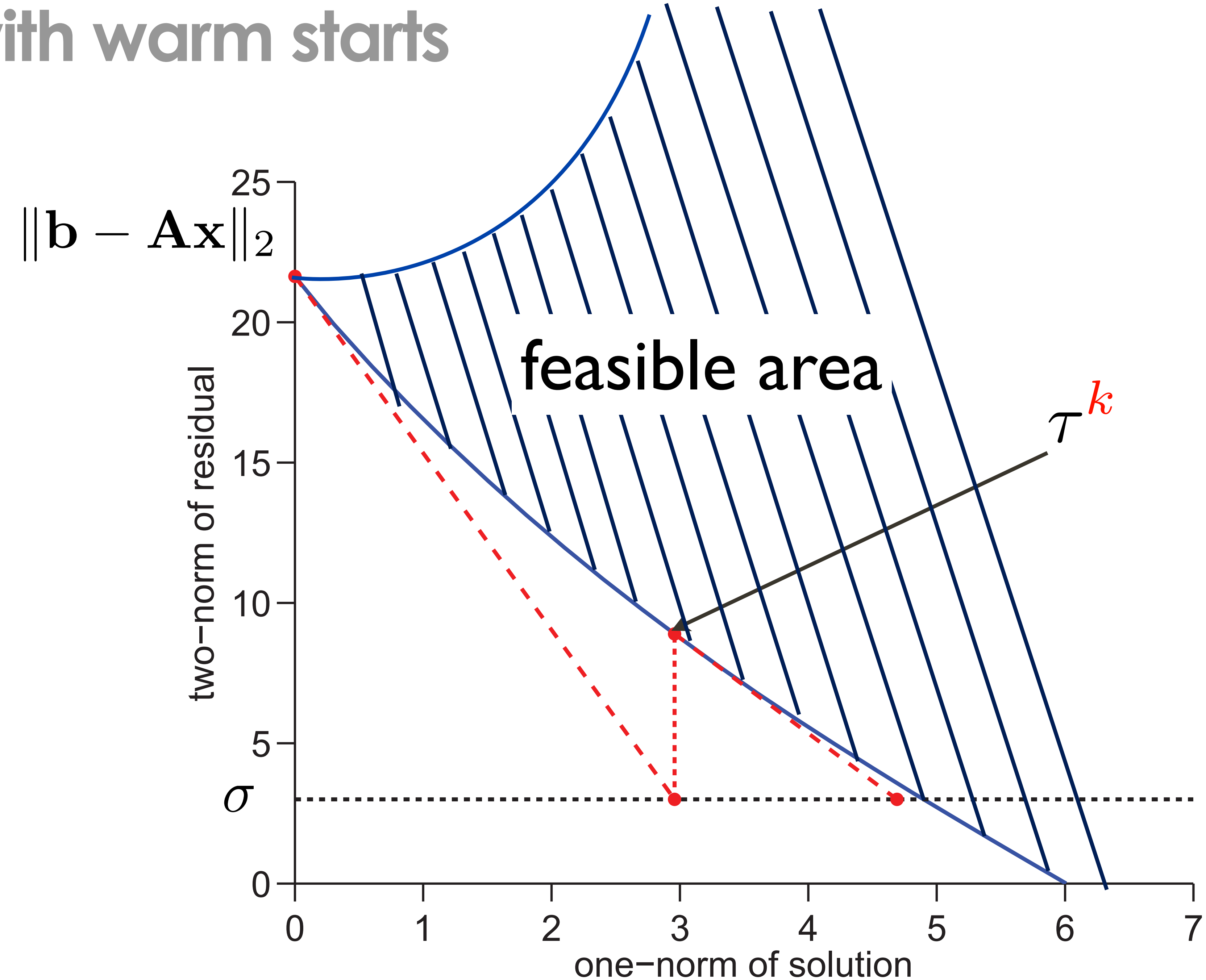
Is this all we can do?

Continuation

Large-scale *sparsity*-promoting solvers *limit* the number of *matrix-vector* multiplies by

- solving an *intelligent* series of LASSO *subproblems* for *decreasing* sparsity levels
- *exploring* properties of the *Pareto* trade-off curve
- slowly allowing *components* to *enter* into the *solution*

Picking Lasso parameter
with warm starts

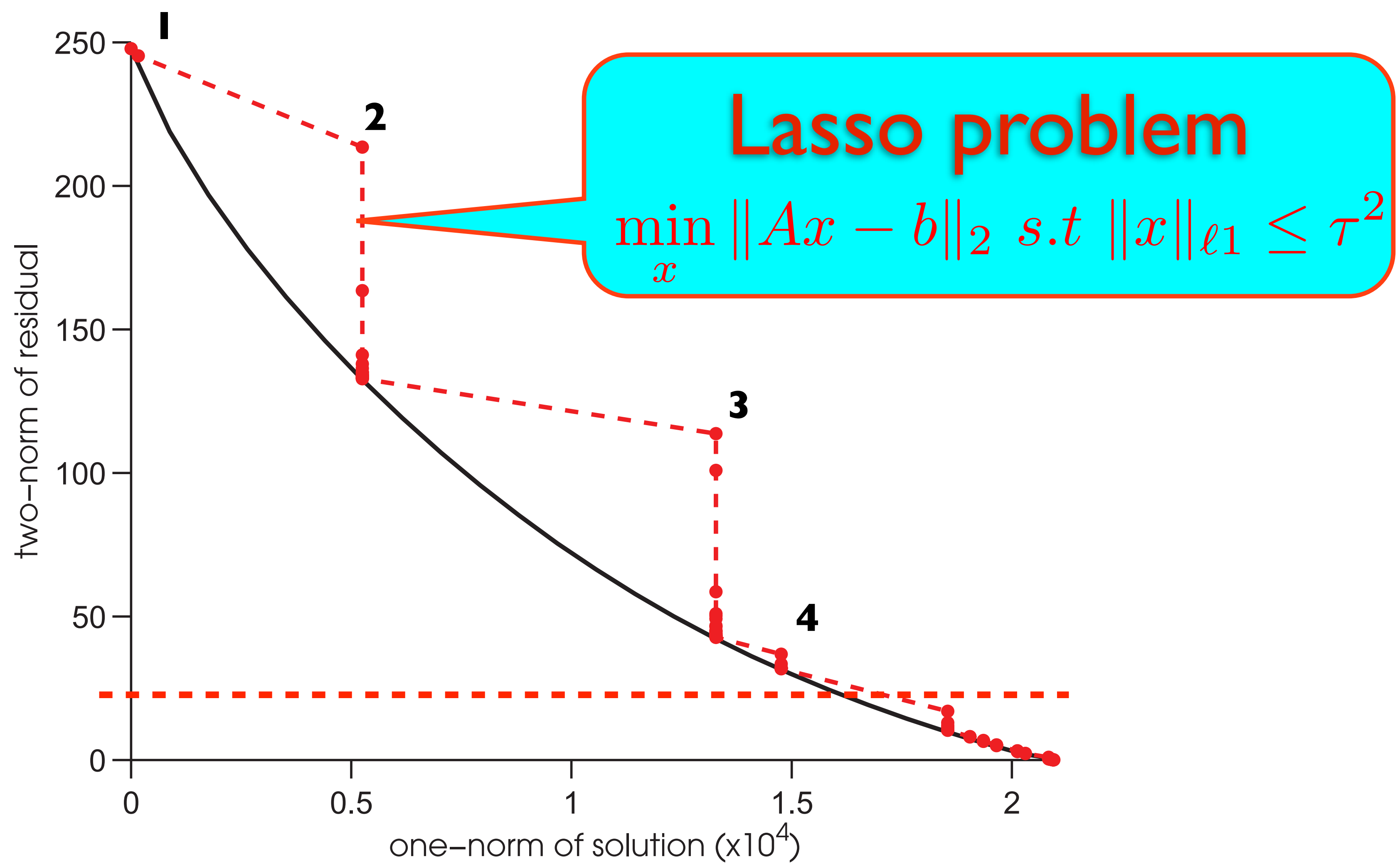


Root finding

$$\text{basis pursuit denoise:} \quad \min \|\mathbf{x}\|_1 \quad s.t \quad \|\mathbf{b}-\mathbf{Ax}\|_2 \leq \sigma$$

Pareto curve

subproblems



[Wang & Sacchi, '07]

Sparsity recovery

This is a BPDN problem:

$$\min_{\delta\mathbf{x}} \tfrac{1}{2}\|\delta\mathbf{x}\|_{\ell_1} \quad \text{subject to} \quad \|\delta\underline{\mathbf{d}} - \nabla\mathcal{F}[\mathbf{m}_0;\underline{\mathbf{Q}}]\mathbf{S}^*\delta\mathbf{x}\|_2 \leq \sigma$$

$$\delta\widetilde{\mathbf{m}} = \mathbf{S}^*\delta\mathbf{x}$$

BPDN problem is a *series* of LASSO *subproblems:*

$$\min_{\delta\mathbf{x}} \tfrac{1}{2}\|\delta\underline{\mathbf{d}} - \nabla\mathcal{F}[\mathbf{m}_0;\underline{\mathbf{Q}}]\mathbf{S}^*\delta\mathbf{x}\|_2 \leq \sigma \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \leq \tau^k$$

Renewals & warm starts

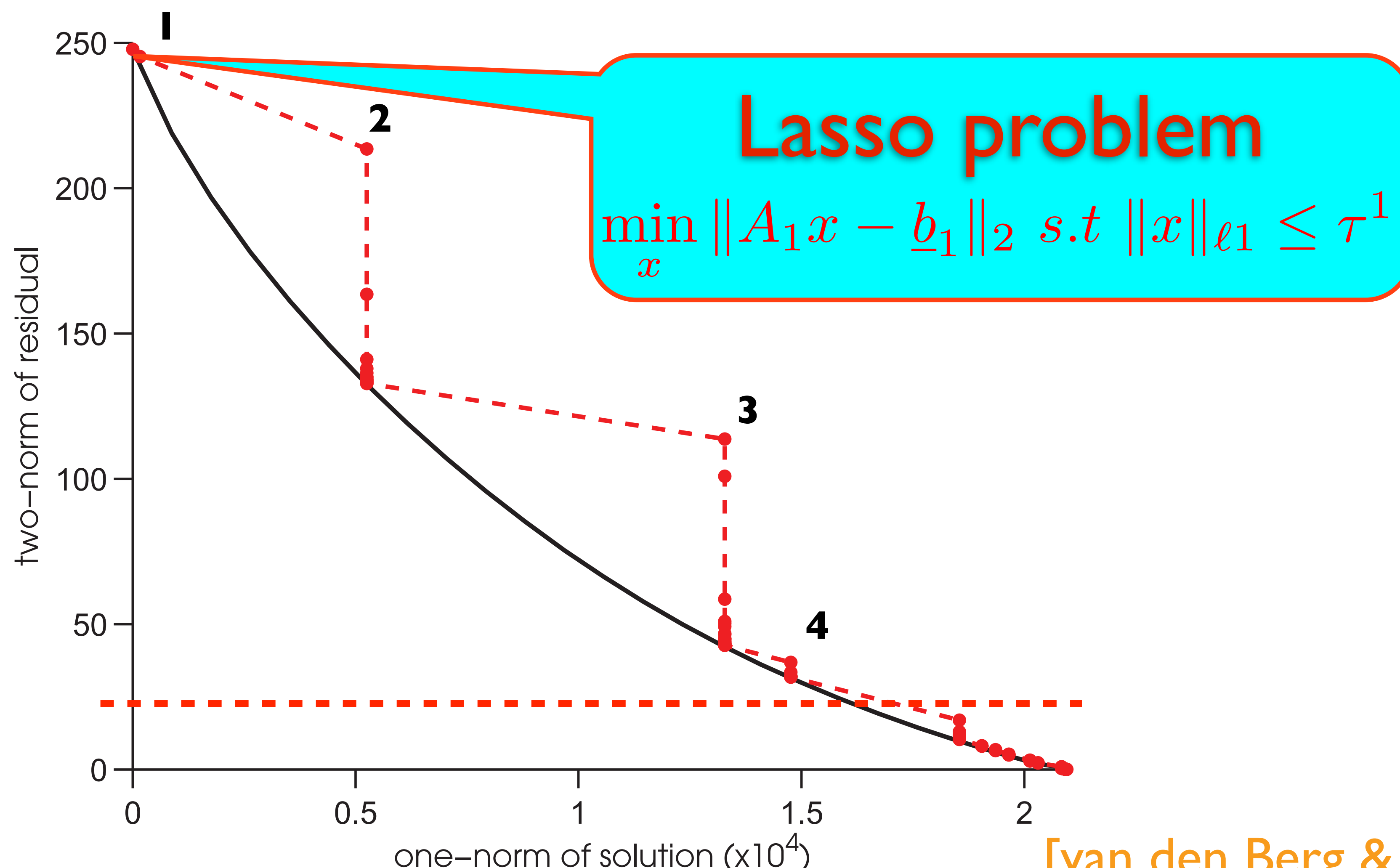
For *each* subproblem:

$$\min_{\delta \mathbf{x}} \frac{1}{2}\|\delta\underline{\mathbf{d}}^{k} - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}_0; \underline{\mathbf{Q}}^{k}]\mathbf{S}^*\delta\mathbf{x}\|_2 \leq \sigma \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \leq \tau^{k}$$

- *redraw* a *new* set of *randomized* source *experiments*
- use $\delta\mathbf{x}^{k-1}$ as *warm* start for next *subproblem*

Pareto curve

subproblems



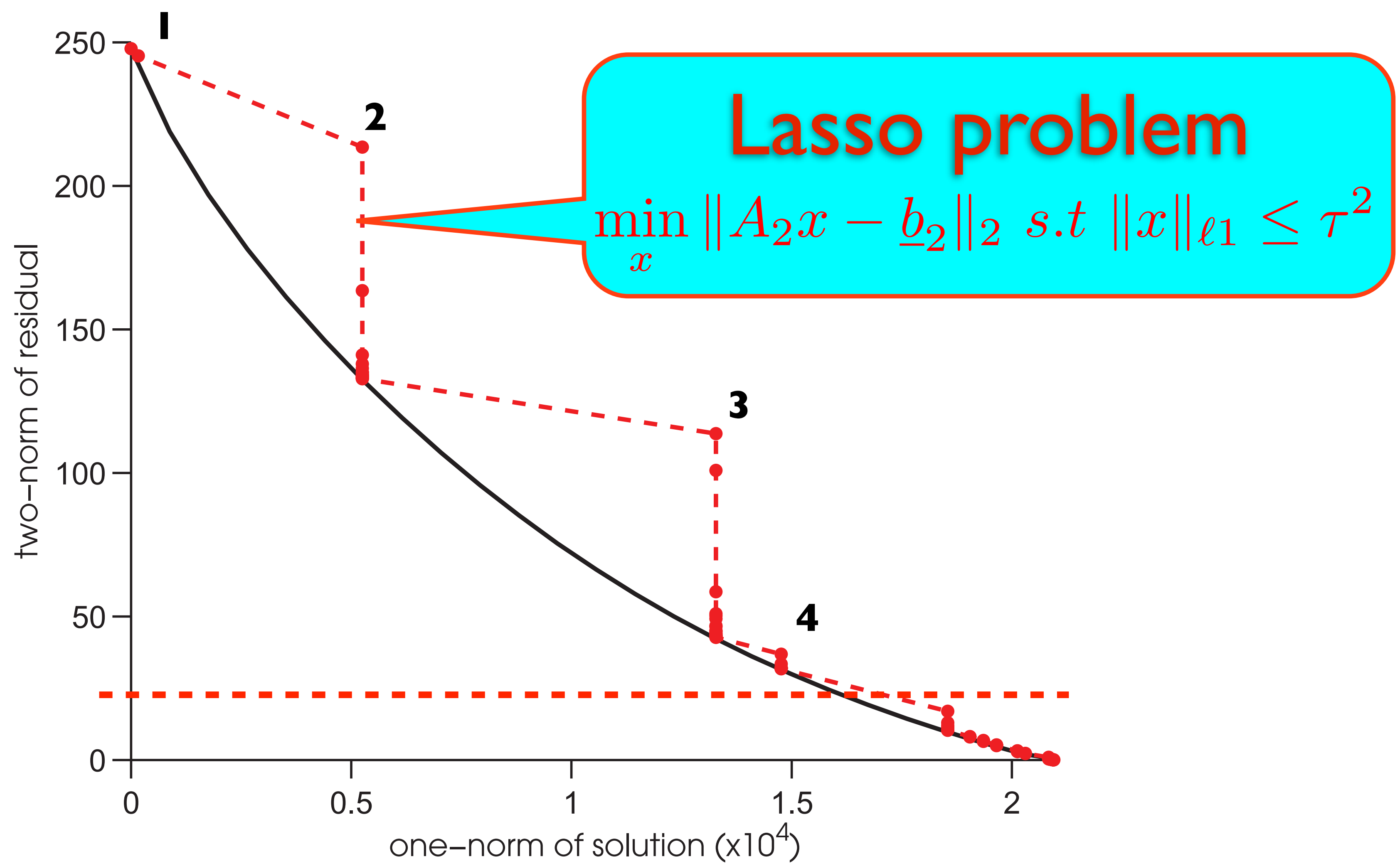
[van den Berg & Friedlander, '08]

[Hennefent et. al., '08]

[Lin & FJH, '09-]

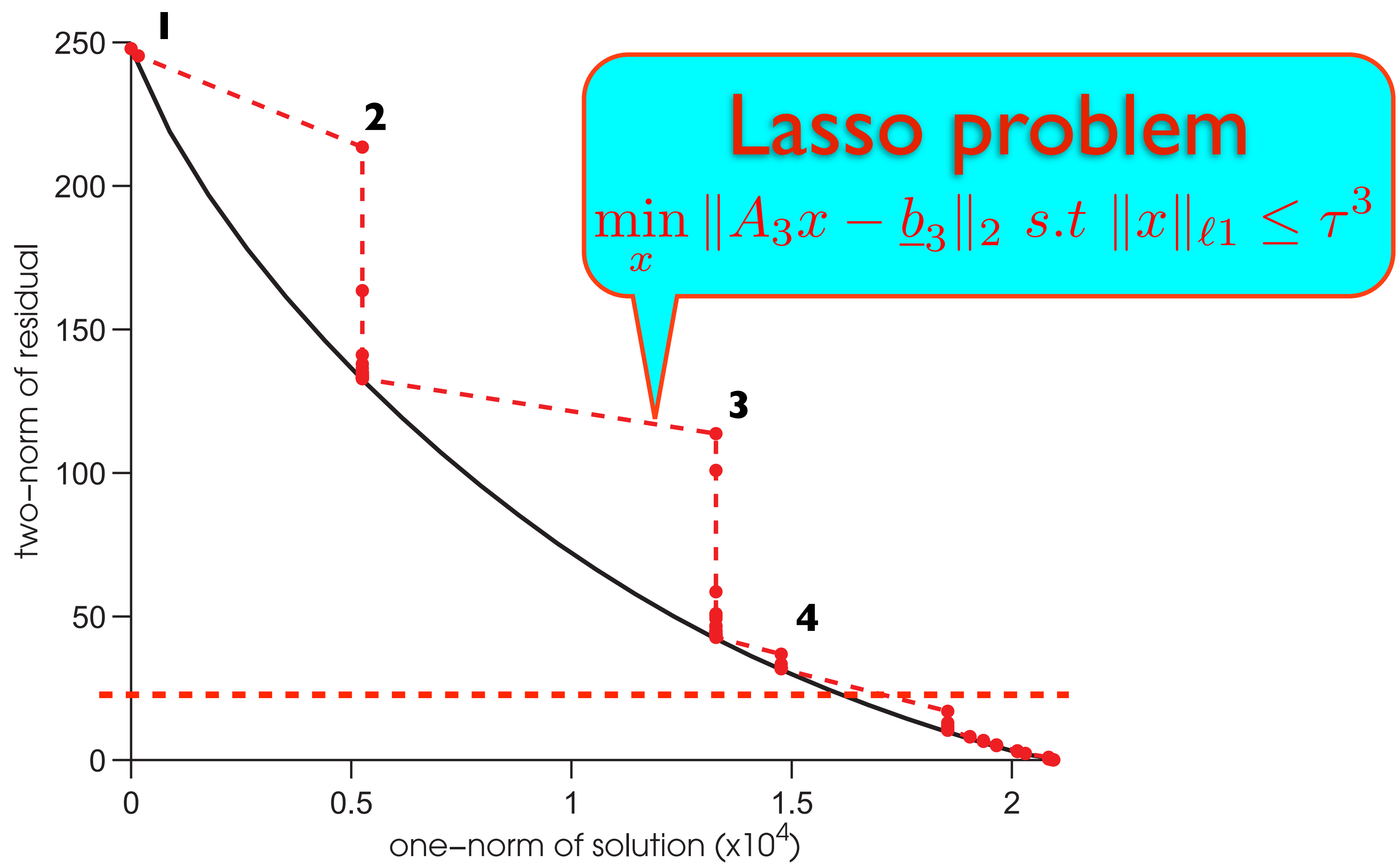
Pareto curve

subproblems



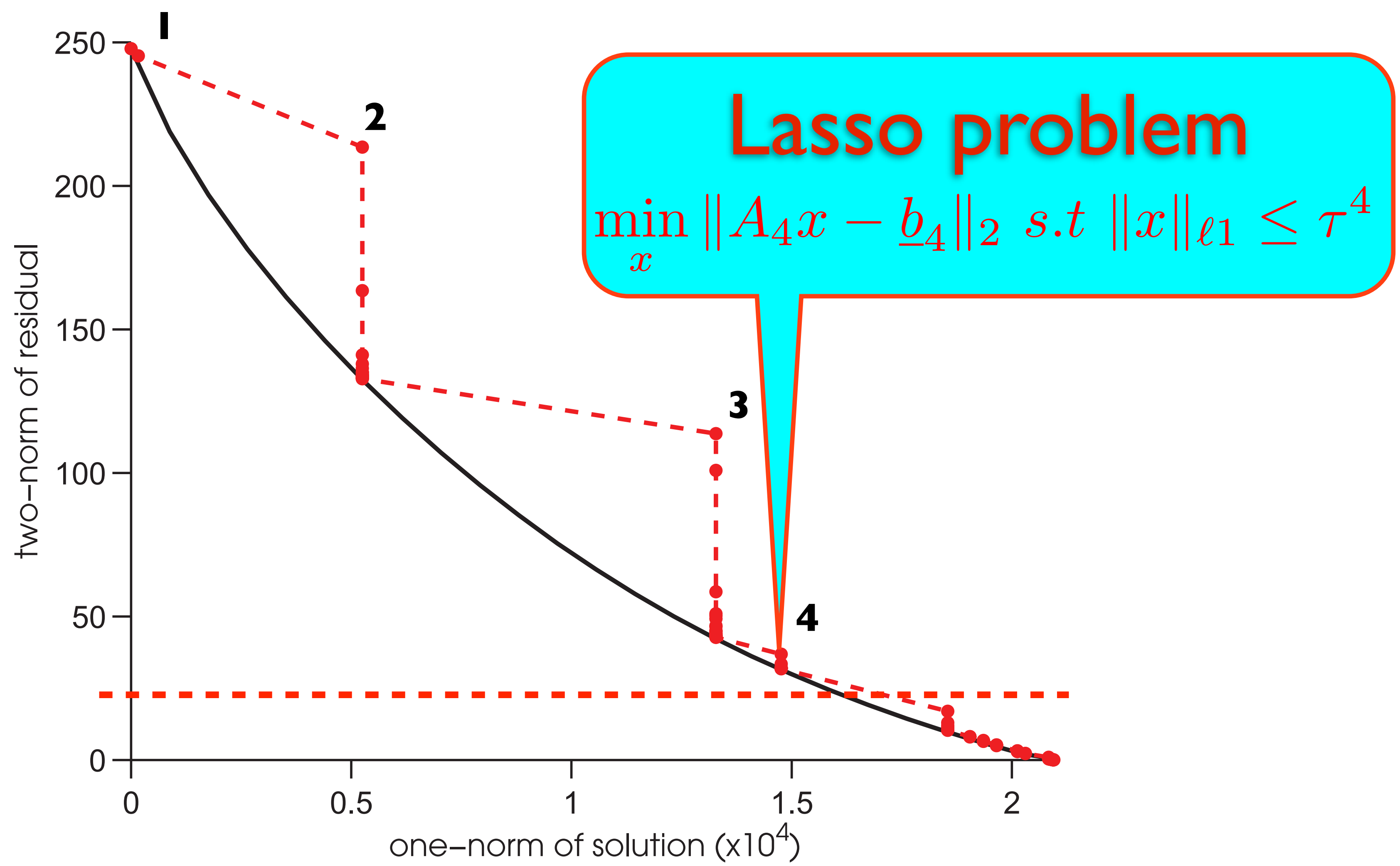
Pareto curve

subproblems

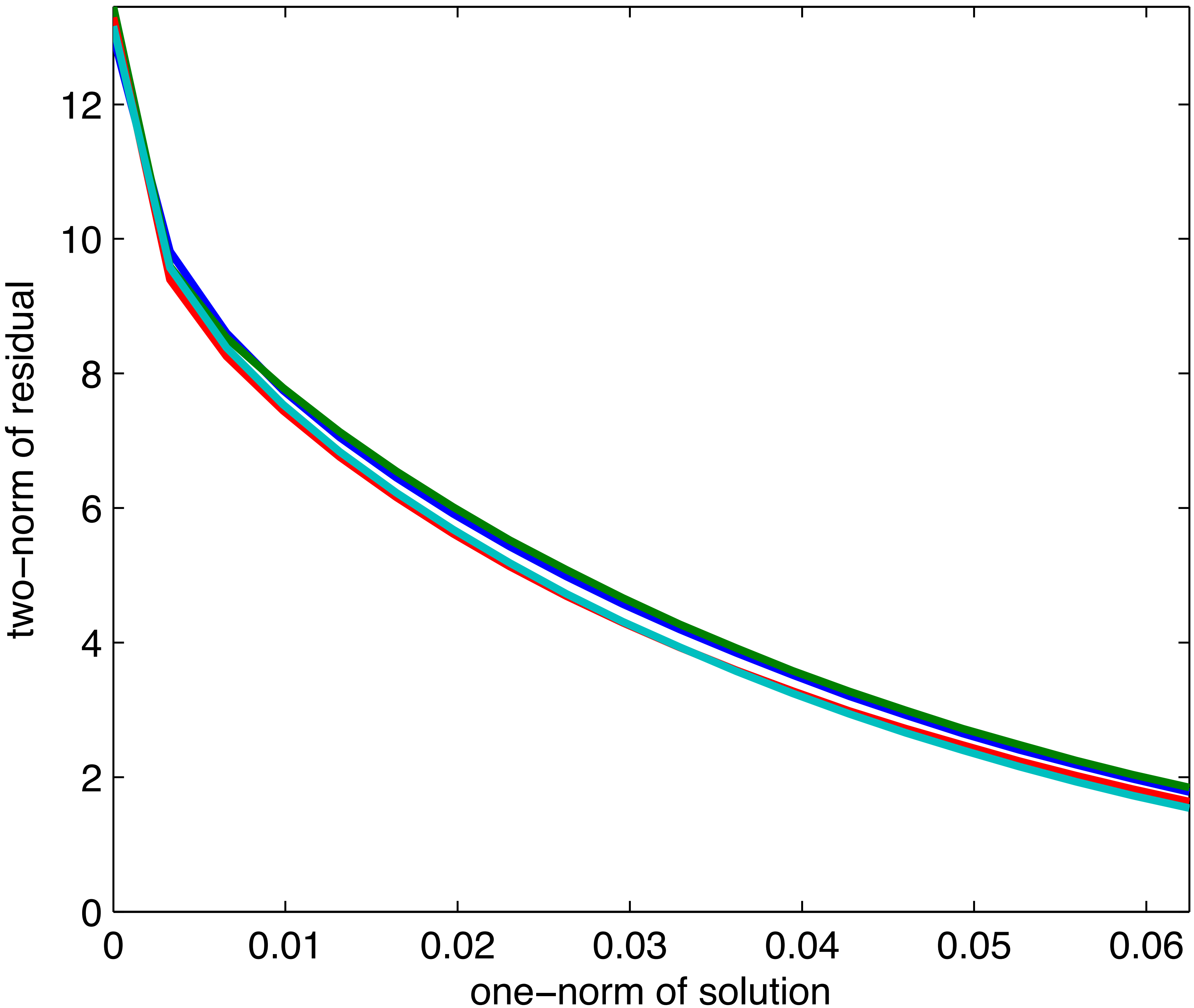


Pareto curve

subproblems



Pareto curves



Continuation methods & *renewals*

Underlying assumption is that Pareto curves are similar

- for large enough batch sizes

In that case the *warm starts* are effective

Renewals remove biases

Migration results

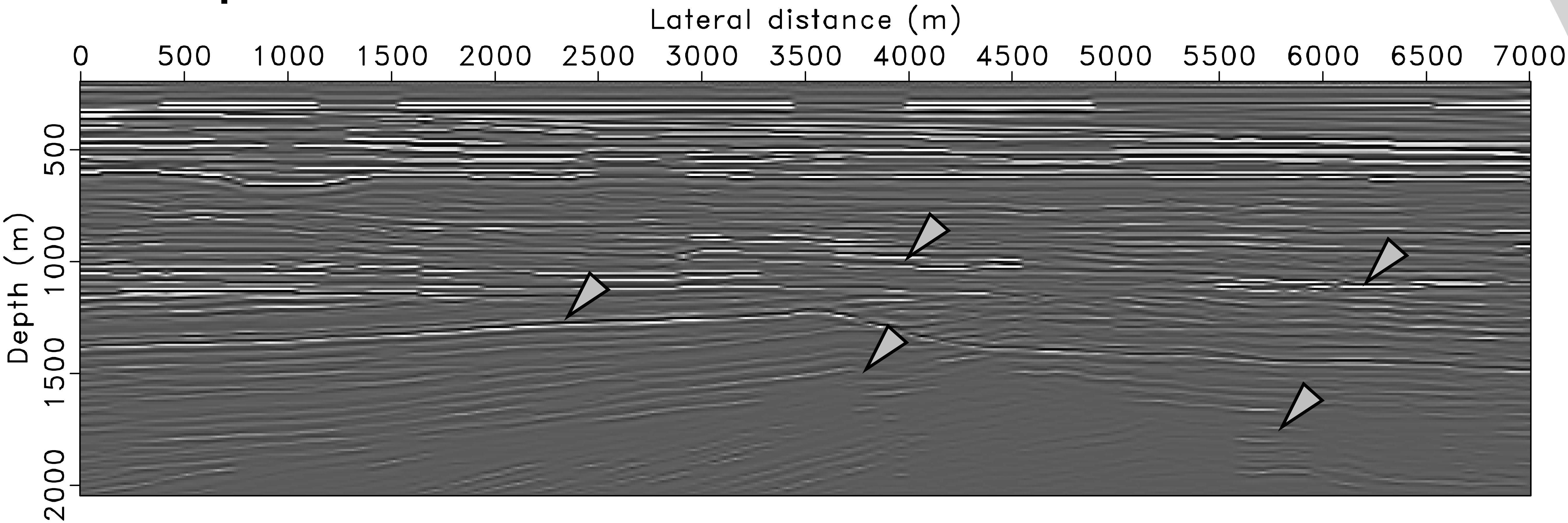
Migration:

- 10 random frequencies (20Hz-50Hz)
- 17 *simultaneous* shots (*versus 350 sequential shots*)
- LASSO problems determined by SPGL1

[Herrmann & Xi, 2011]

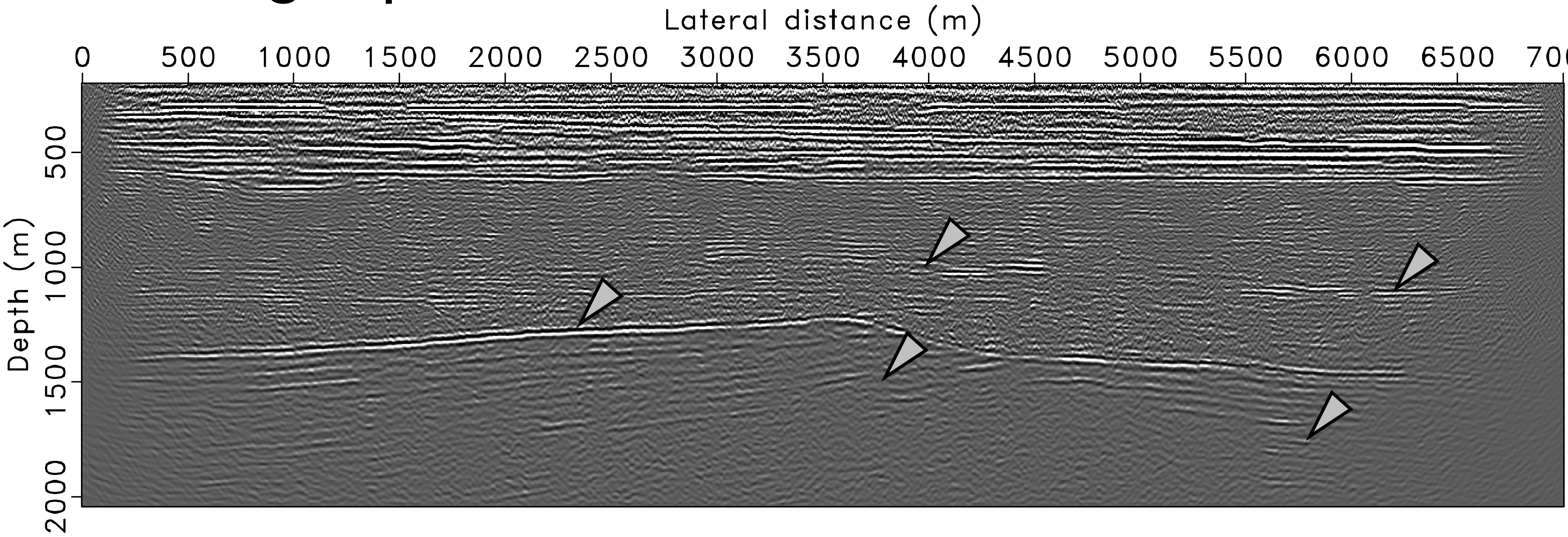
Migration results

true perturbation



Migration results

L1 imaged perturbation *without* renewals

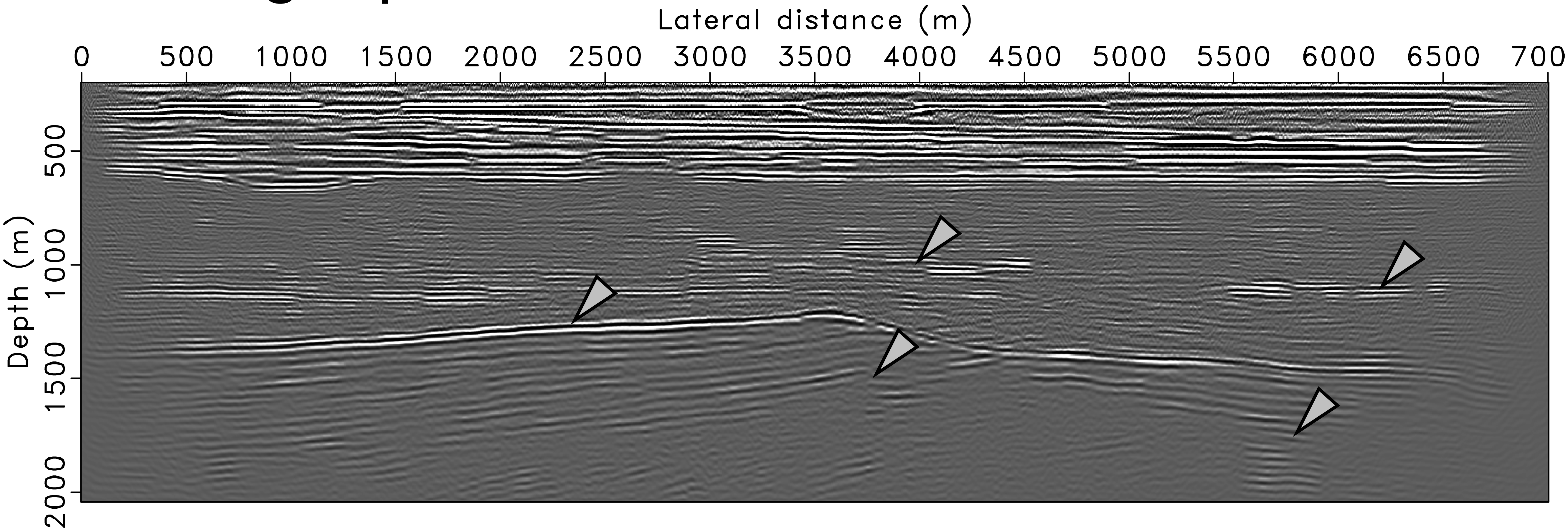


with 17 simultaneous shots

SLIM

Migration results

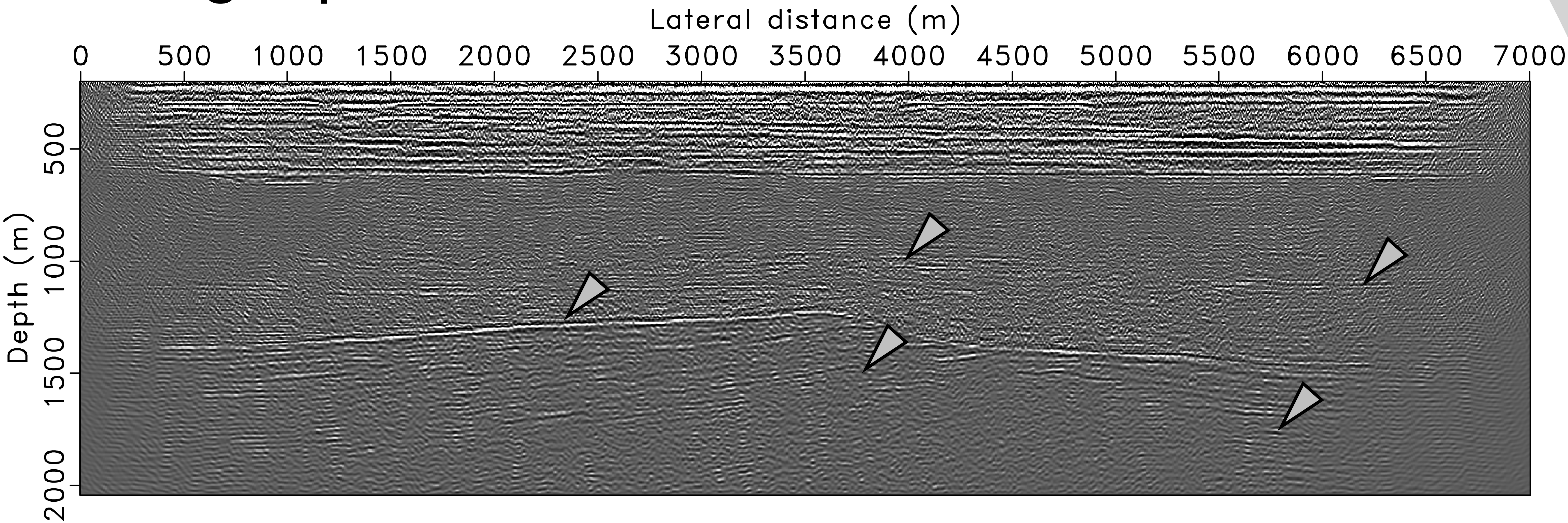
L1 imaged perturbation *with* renewals



with 17 *simultaneous* shots

Migration results

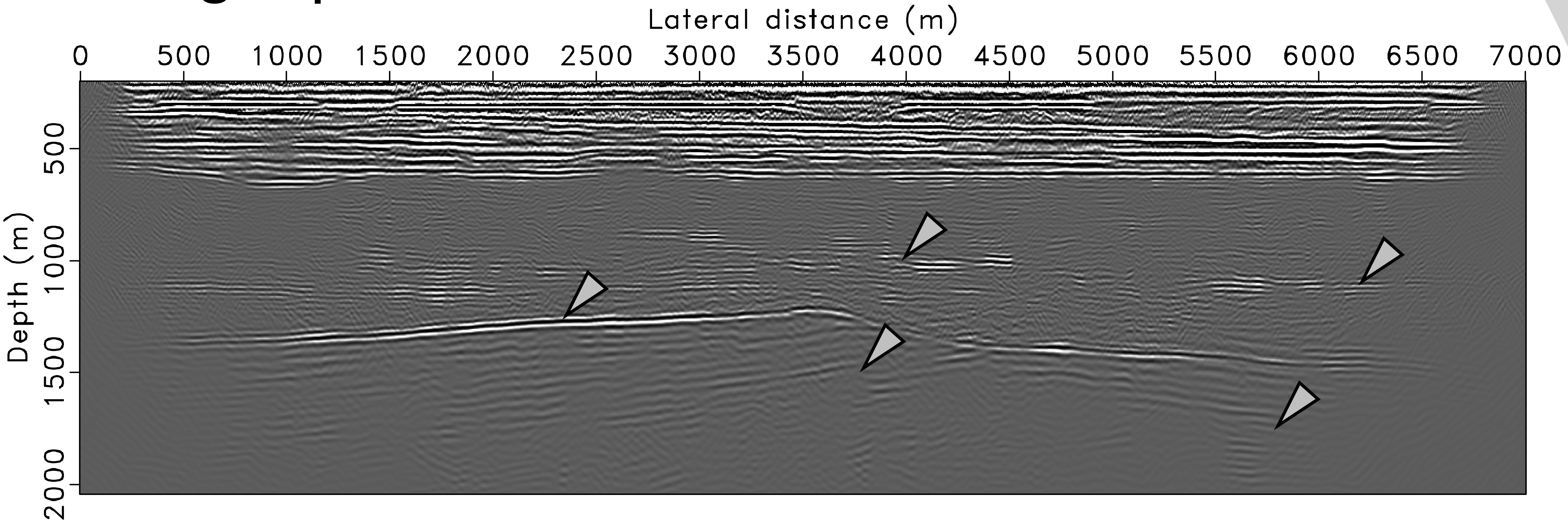
imaged perturbation with L2 *without* renewals



with 17 *simultaneous* shots

Migration results

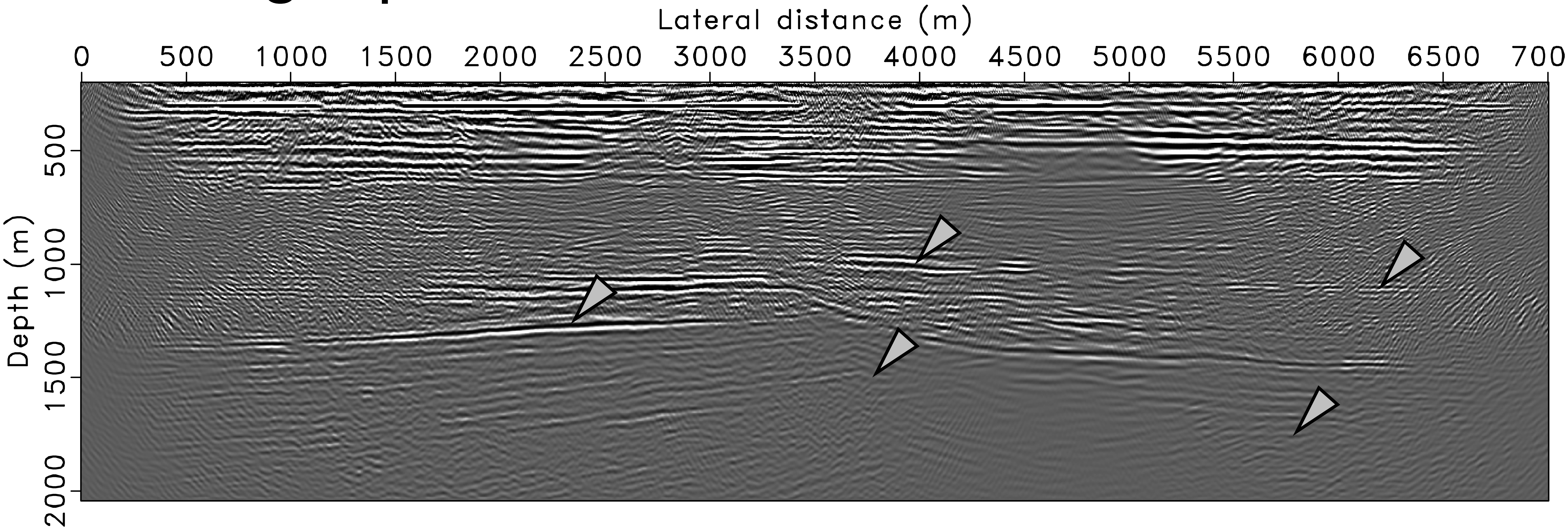
imaged perturbation with L2 *with* renewals



with 17 *simultaneous* shots

Migration results

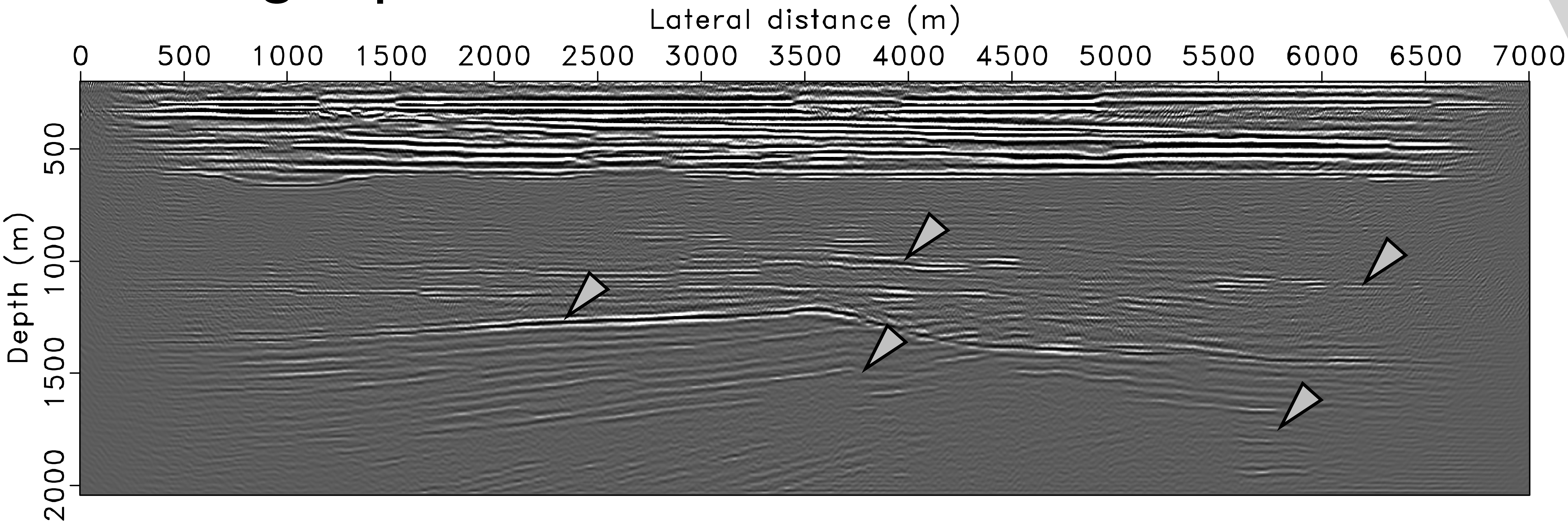
L1 imaged perturbation *without* renewals



with same # of *randomly* selected *sequential* shots

Migration results

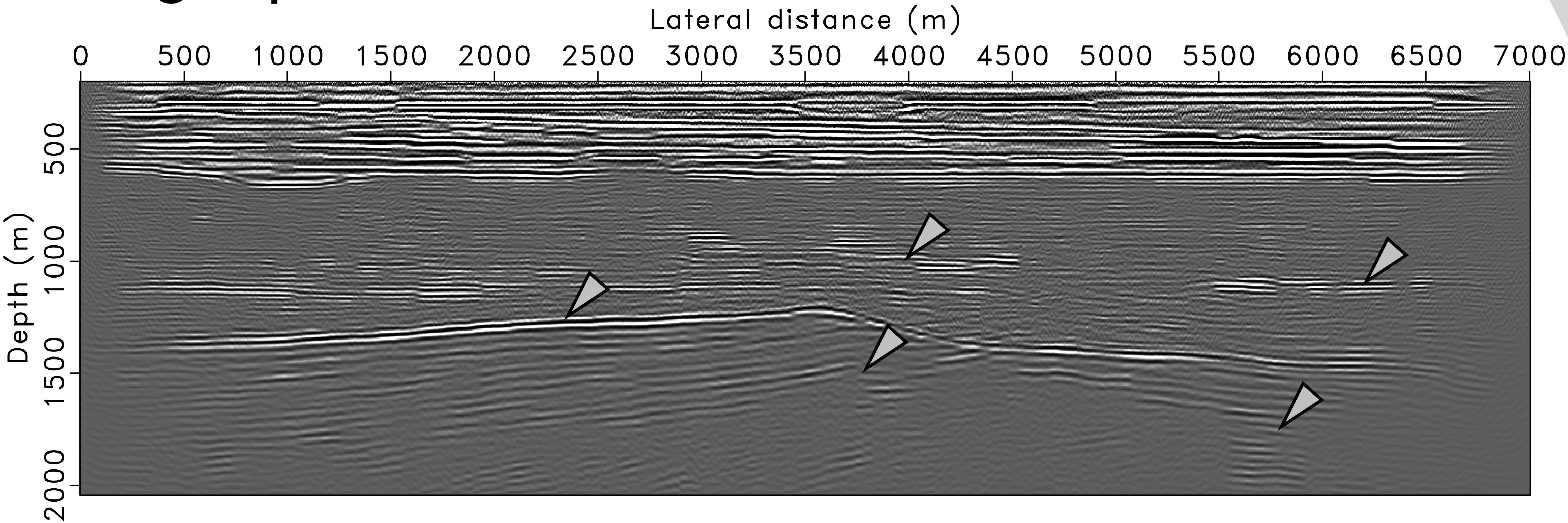
L1 imaged perturbation *with* renewals



with same # of *randomly* selected *sequential* shots

Migration results

imaged perturbation *with renewals*



with 17 *simultaneous* shots

Conclusions

Computational cost can be *reduced* significantly by using *randomized dimensionality reduction*

Underdetermined system can be solved by *sparsity* promotion in a *sparsifying* (e.g curvelet) domain

Within the same *computational* cost, simultaneous shots produce *less* migration *artifacts*

Source cross-talk bias can be *removed* by *renewals* & *warm starts*.

Imaged reflectors in GN updates are *compressible* in the *curvelet* domain

Acknowledgments

We would like to thank Charles Jones from BG for providing us with the BG Compass model. 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, BP, BGP, Chevron, ConocoPhillips, Petrobras, PGS, Total SA, and WesternGeco.

NSERC CRSNG

Thank you

slim.eos.ubc.ca