SINBAD Consortium 2010

Compressive Imaging

Xiang Li and Felix J. Herrmann

SLIM
University of British Columbia

Motivation

Curse of dimensionality for d>2

- *Exponentially* increasing *data volumes*
- *Helmholtz* requires *iterative* solvers to address *bandwidth*
- Computational complexity grows *linearly* with # RHS's
- Multi-dimensional correlations between source wavefield and residual wavefield
- High-resolution image

SLIM

Wish list

An *imaging* technology that

- is based on a *time-harmonic* PDE solver, which is easily *parallelizable*, and *scalable* to 3D
- does *not* require *multiple* iterations with *all* data
- removes the *linearly* increasing costs of *iterative* solvers for increasing numbers of frequencies & RHS's
- produces *high-resolution* least square migration result for full-waveform inversion

Key technologies

Simultaneous sources & phase encoding [Beasley, '98, Berkhout, '08] [Morton, '98, Romero, '00]

- supershots [Krebs et.al., '09, Operto et. al., '09, Herrmann et.al., '08-10']

Stochastic optimization & machine learning [Bertsekas, '96]

- stochastic gradient descent

Compressive sensing [Candès et.al, Donoho, '06]

- *sparse* recovery & *randomized* subsampling

[Nemeth et. al. '99]

Imaging

Least-squares migration:

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

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

[Pratt et. al., '98]
[Plessix '06]

SLIM

Adjoint state

Solves of Helmholtz system for each source $\mathbf{q}$

$$\mathbf{H}[\mathbf{m}]\mathbf{u} = \mathbf{q} \quad \text{and} \quad \mathbf{H}^*[\mathbf{m}]\mathbf{v} = \mathbf{r}$$

with

$$\mathbf{r} = \mathbf{D}^*(\mathbf{p} - \mathcal{F}[\mathbf{m}, \mathbf{q}])$$

and compute gradient by

$$\delta\mathbf{m} = \Re\left(\sum_{\omega} \omega^2 \sum_{s} (\bar{\mathbf{u}} \odot \mathbf{v})_{s,\omega}\right)$$

SLIM

[Plessix '06]

Multishot gradient

Post-stack migration:

$$\delta \mathbf{m} = \Re \left(\sum_{\omega} \omega^2 \sum_{s} \left(\bar{\mathbf{u}} \odot \mathbf{v} \right)_{s,\omega} \right) = \nabla \mathcal{F}^* [\mathbf{m}, \mathbf{Q}] \delta \mathbf{d}$$

with

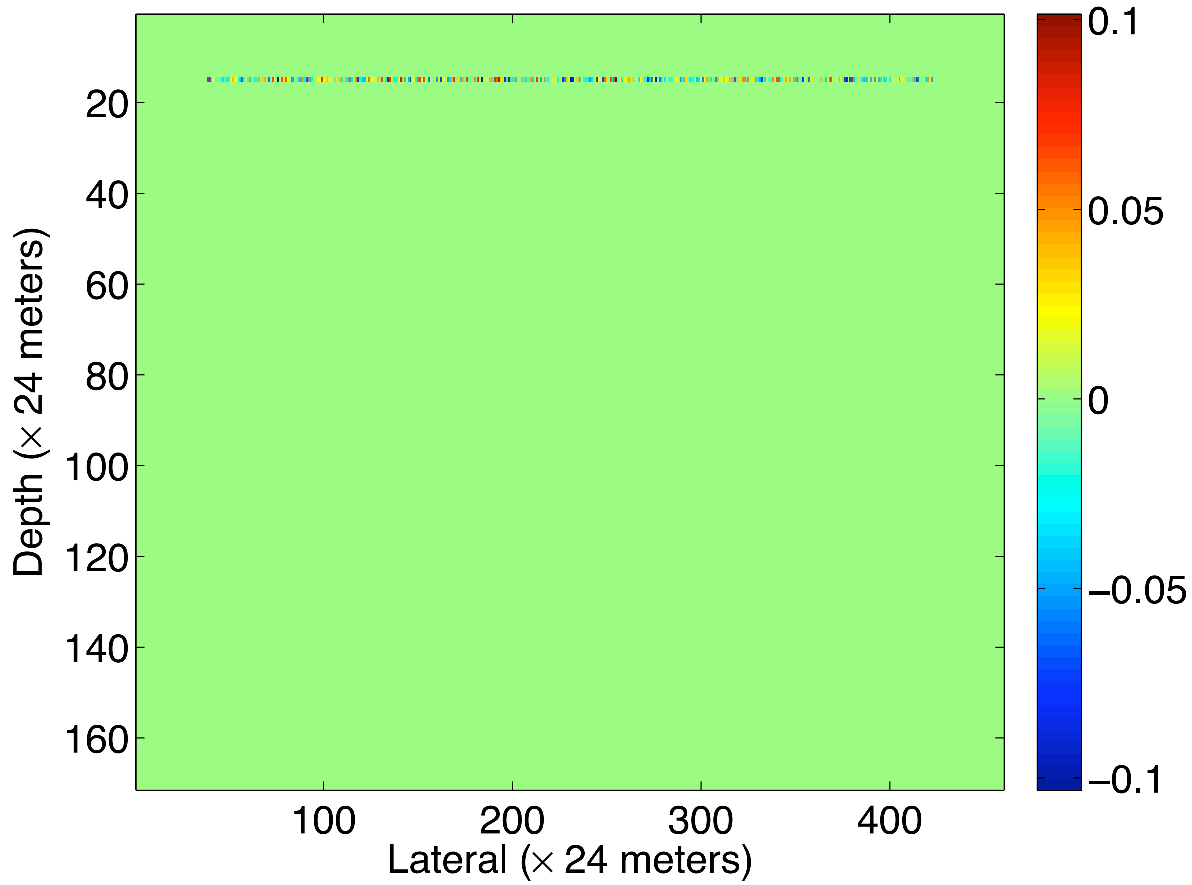
$$\boldsymbol{\delta} \mathbf{d} = \mathrm{vec}(\mathbf{P} - \mathcal{F}[\mathbf{m}, \mathbf{Q}])$$

and the multi-experiment

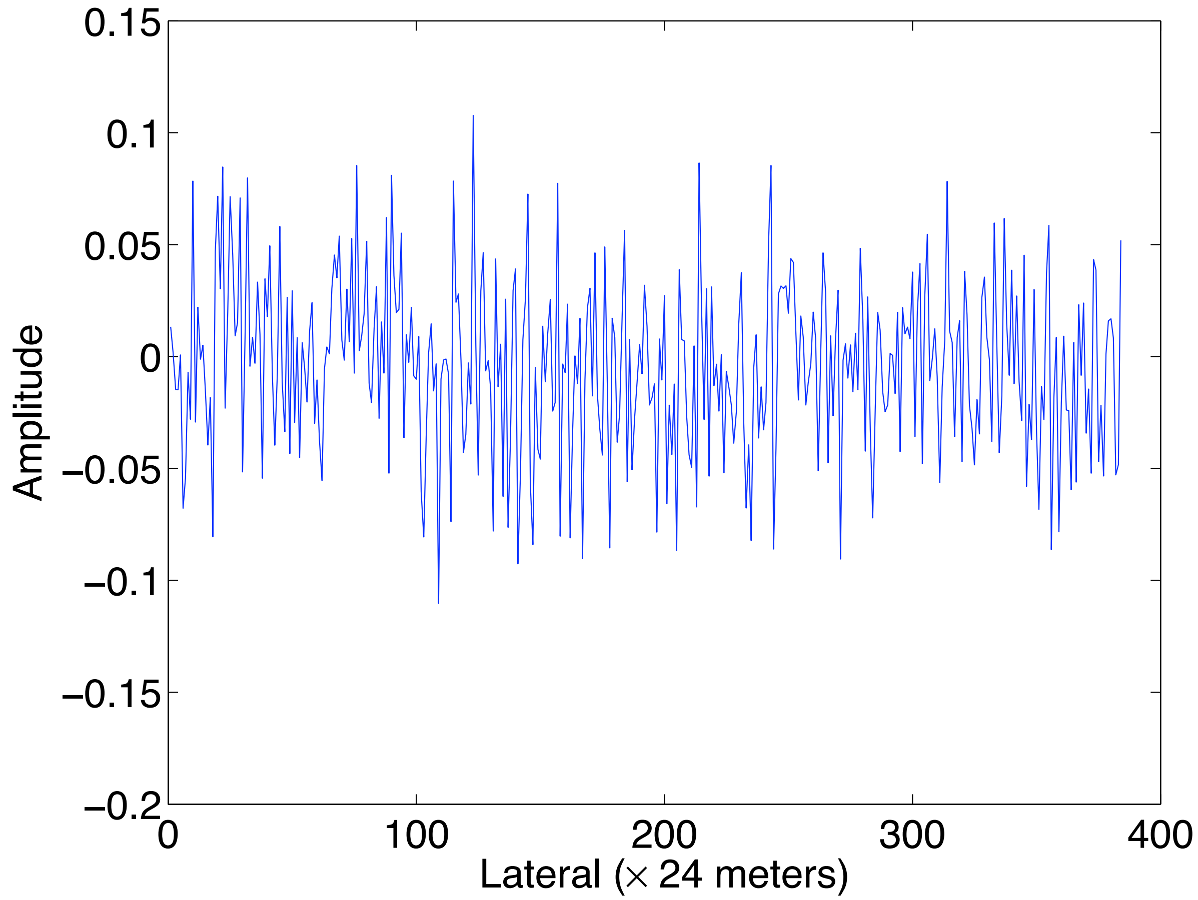
$$\mathcal{F}[\mathbf{m}, \mathbf{Q}] = \mathbf{D}\mathbf{H}^{-1}[\mathbf{m}]\mathbf{Q}$$

Phase encoding

Simultaneous source



Randomized amplitudes along the shot line

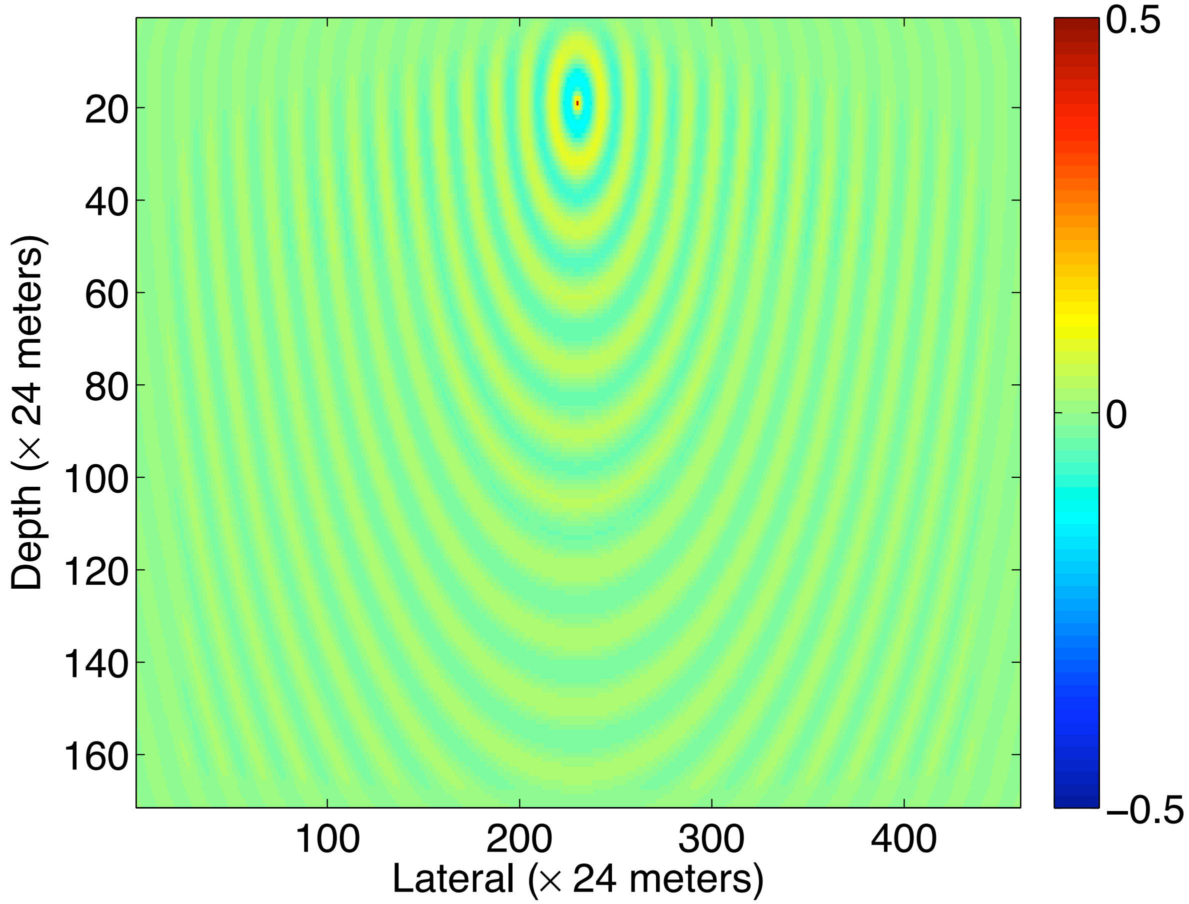


Create *supershot* via *superposition*

[Morton, '98, Romero, '00]

Simultaneous shot at 5 Hz

Sequential-source wavefield



Simultaneous-source wavefield

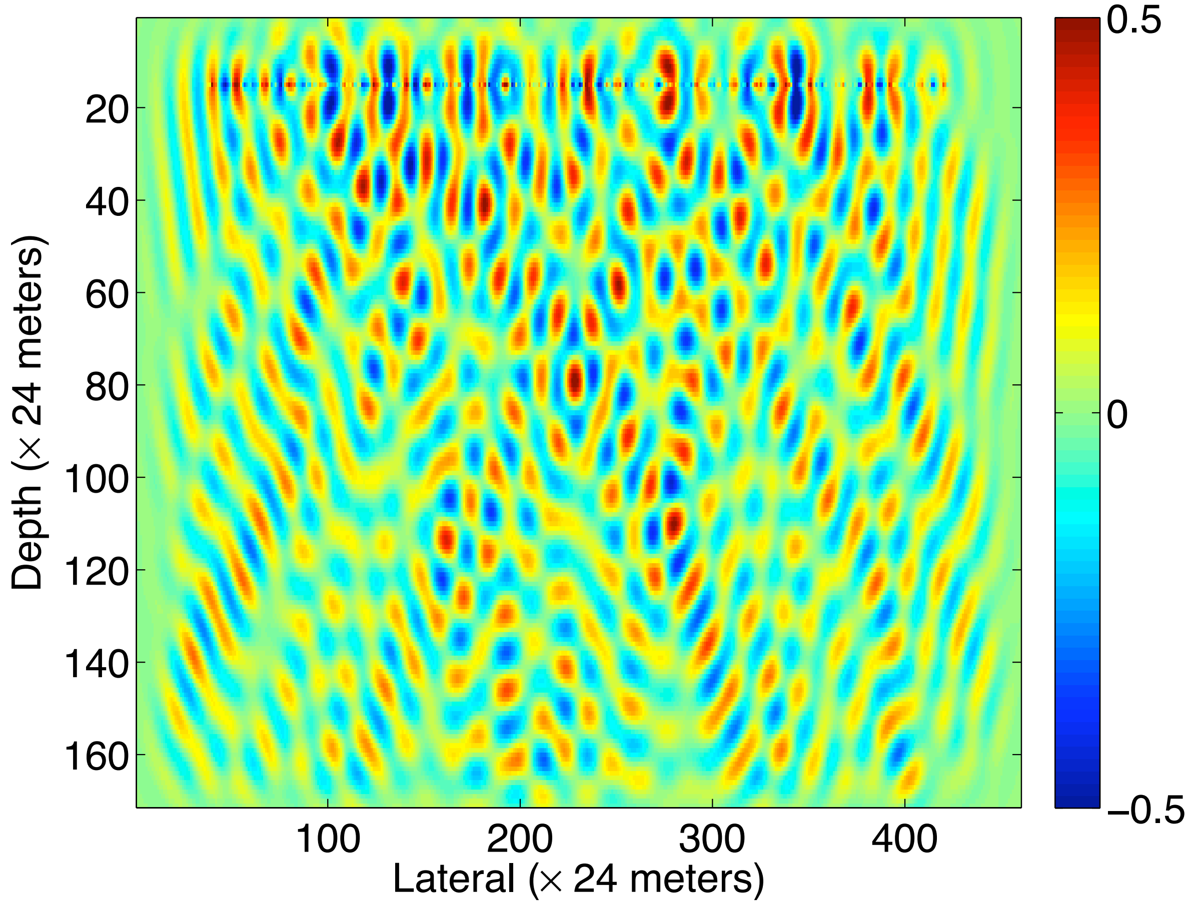
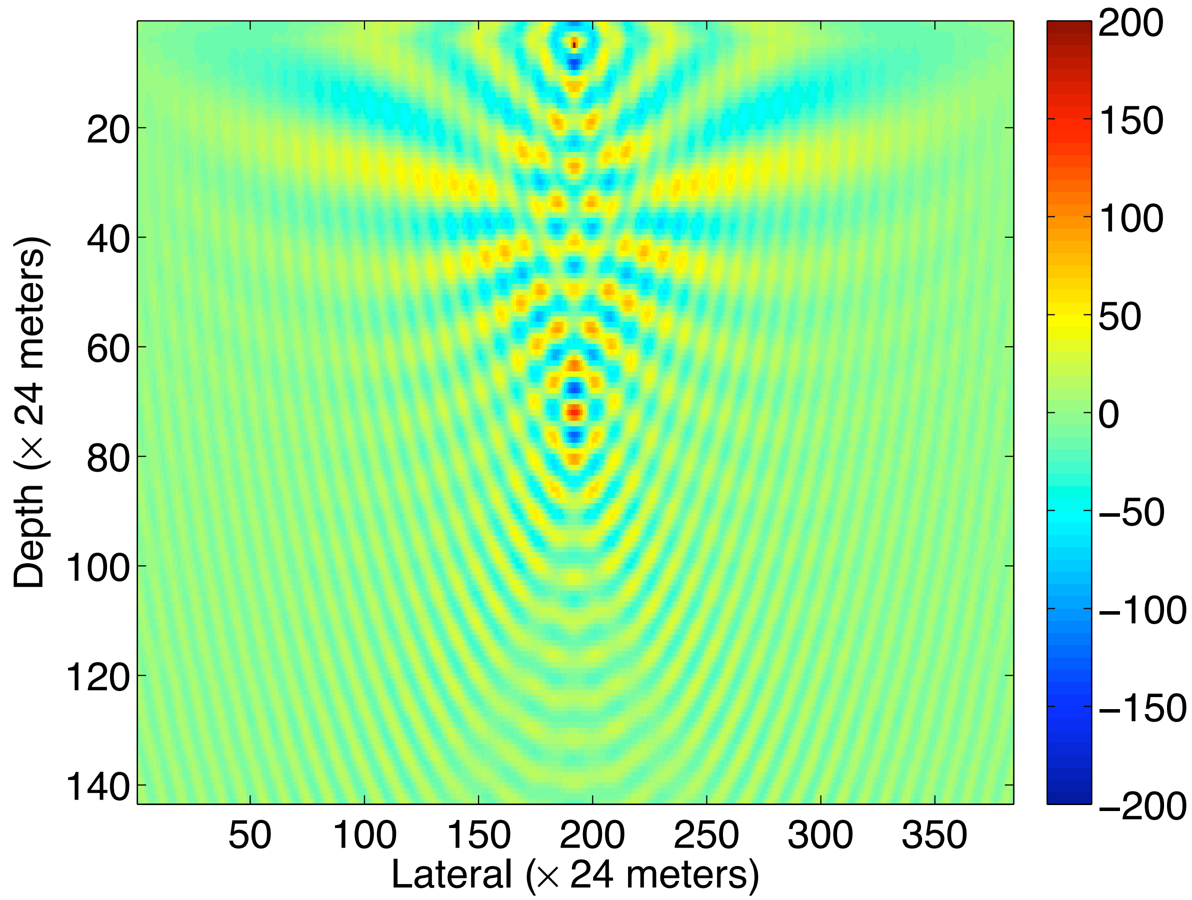
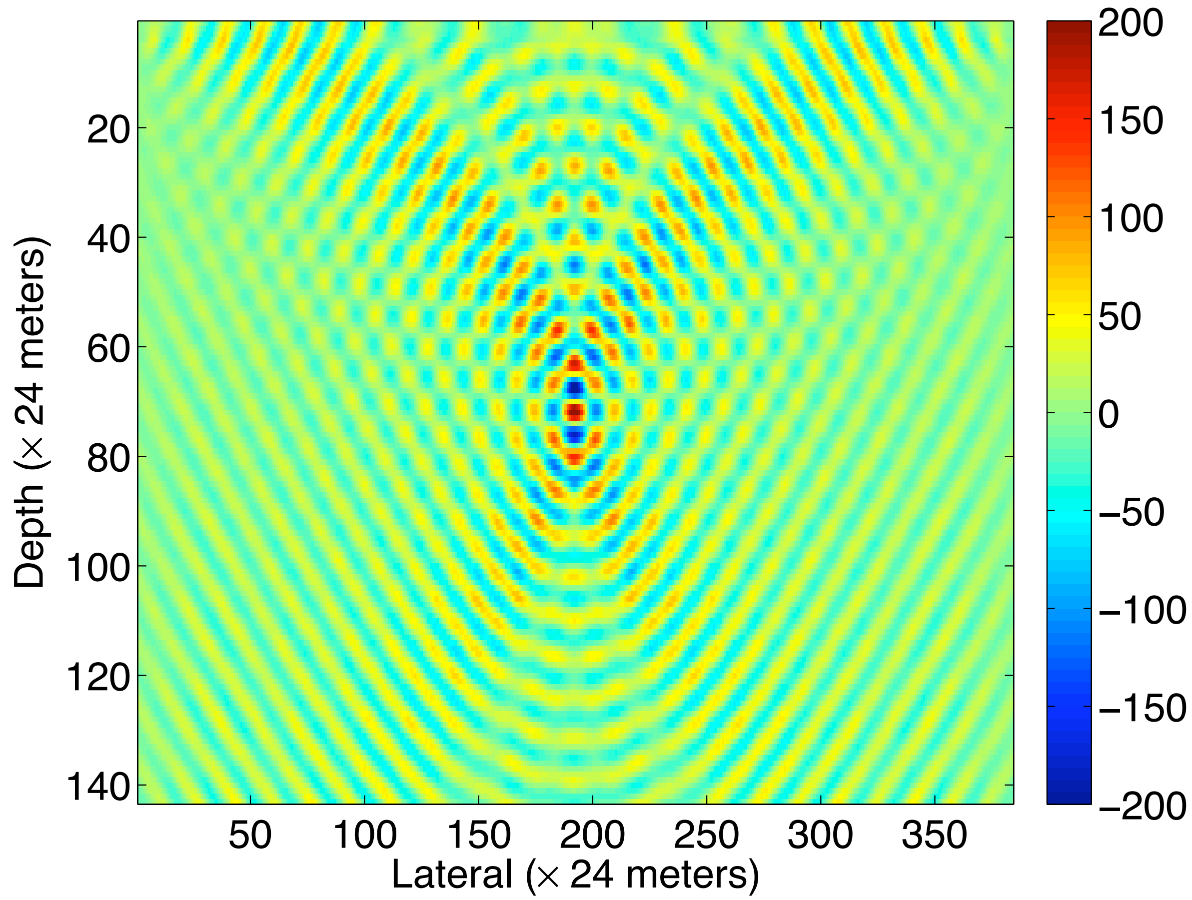


Image at 5 Hz

Sequential-source image



Simultaneous-source image

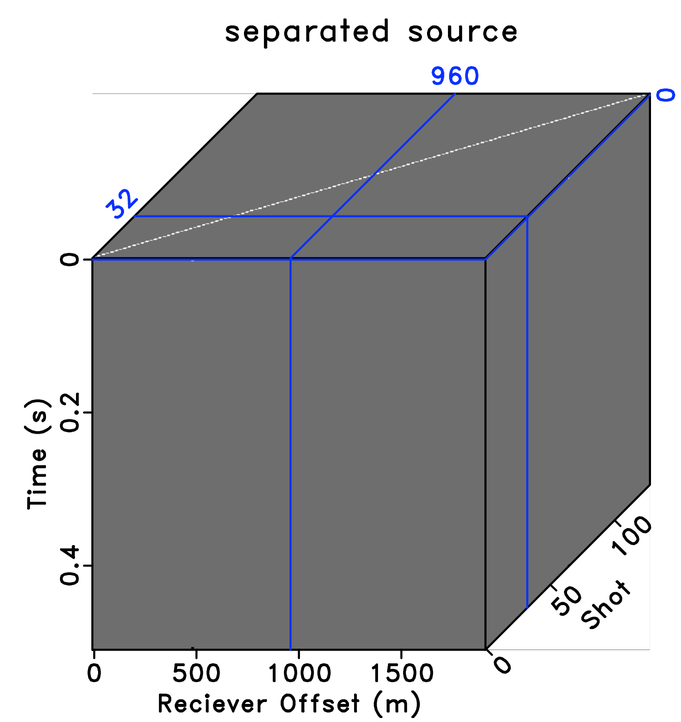


[Morton, '98, Romero, '00]

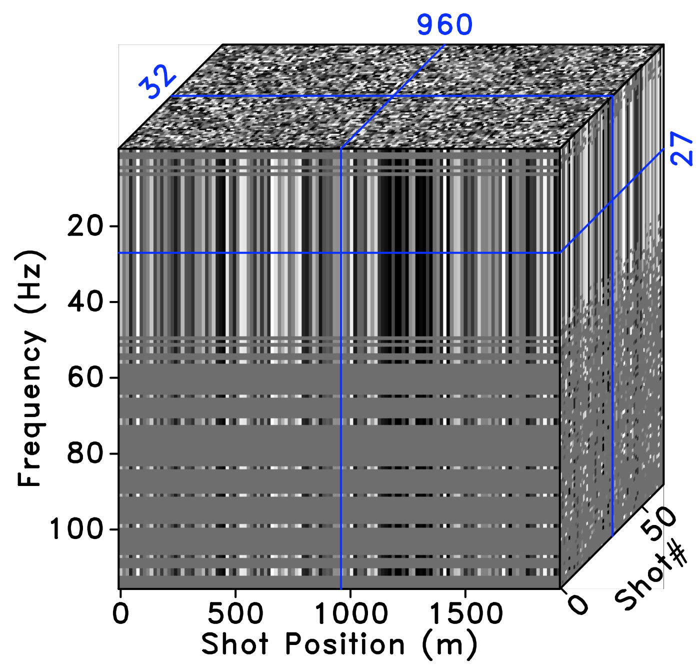
[Herrmann et. al. ’08-’10]

Supershot

adapted from Herrmann et. al. ,09



$$\mathbf{Q}$$



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

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

Phase encoding

Least-squares migration:

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

$$\delta\underline{\mathbf{d}} = \text{Simultaneous-source data residue}$$

$$\underline{\mathbf{Q}} = \text{Simultaneous sources}$$

SLIM

[Wang & Sacchi, '07]

Sparse recovery

Least-squares migration with *sparsity* promotion

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

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

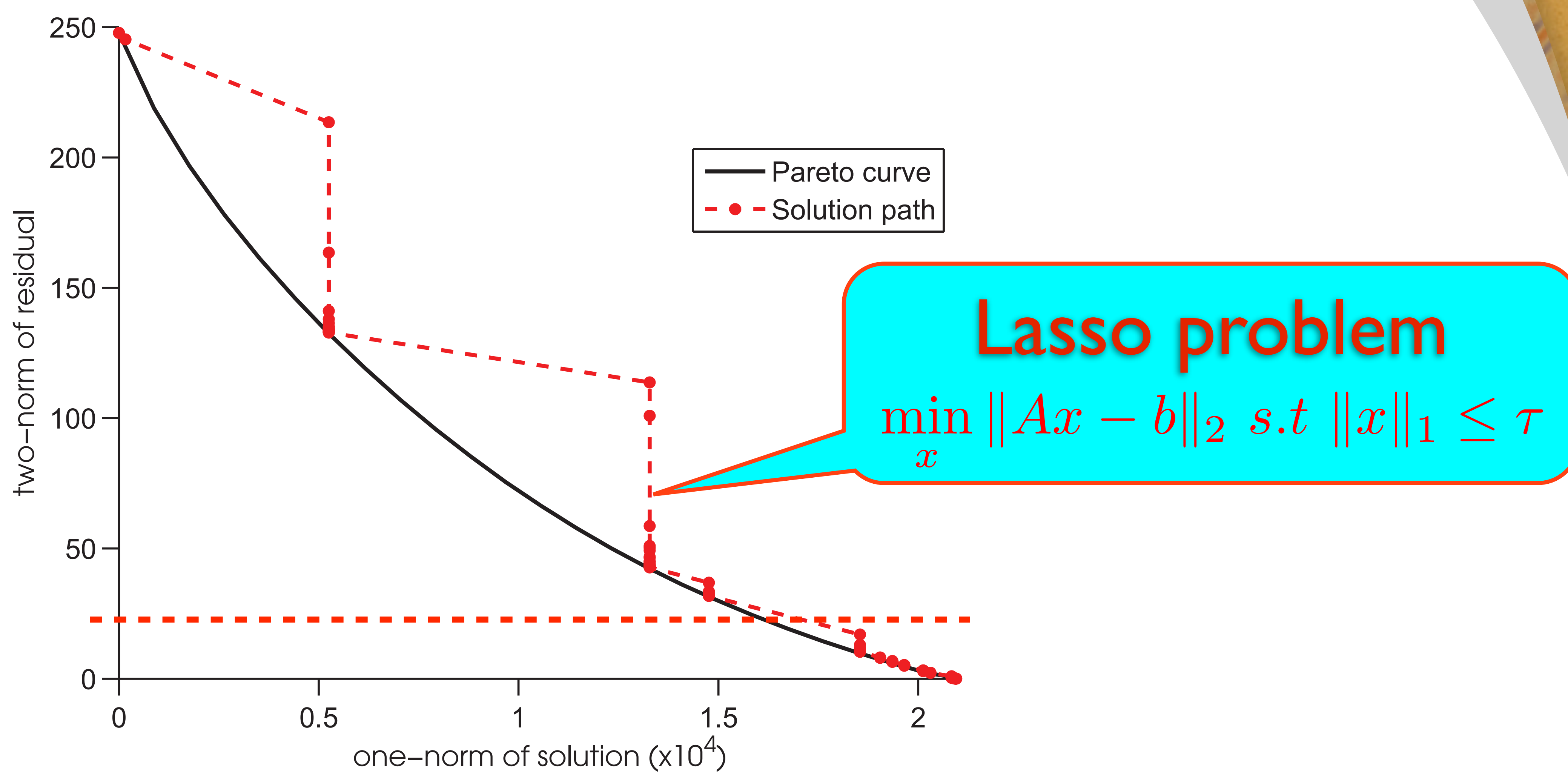
$$\mathbf{S}^* = \text{Curvelet synthesis}$$

leads to *significant* speedup as long as

$$n^{\ell_1}_{PDE} \times K \ll n^{\ell_2}_{PDE} \times n_f \times n_s$$

SLIM

Pareto Curve



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

Renewals

Redraw *different* simultaneous shots and frequencies when the pareto curve is reached, i.e.,

$$\underline{\mathbf{Q}} \mapsto \underline{\mathbf{Q}}^k$$

$$\underline{\mathbf{f}} \mapsto \underline{\mathbf{f}}^k$$

- does NOT increase the size of the problem
- gives "new" information

Experiment

Linearized *sparsity promoting* least-squares migration

- Marmousi model (128x384) with grid size 24 m
- 12 Hz ricker wavelet
- use different
  - # of simultaneous shots
  - # of frequencies

Comparison

compressive recovery versus ***L2 recovery***

Compressive recovery

- 3-10 sim-shots
- 8-20 freqs
- 200 iterations

Batch size roughly equals to 50

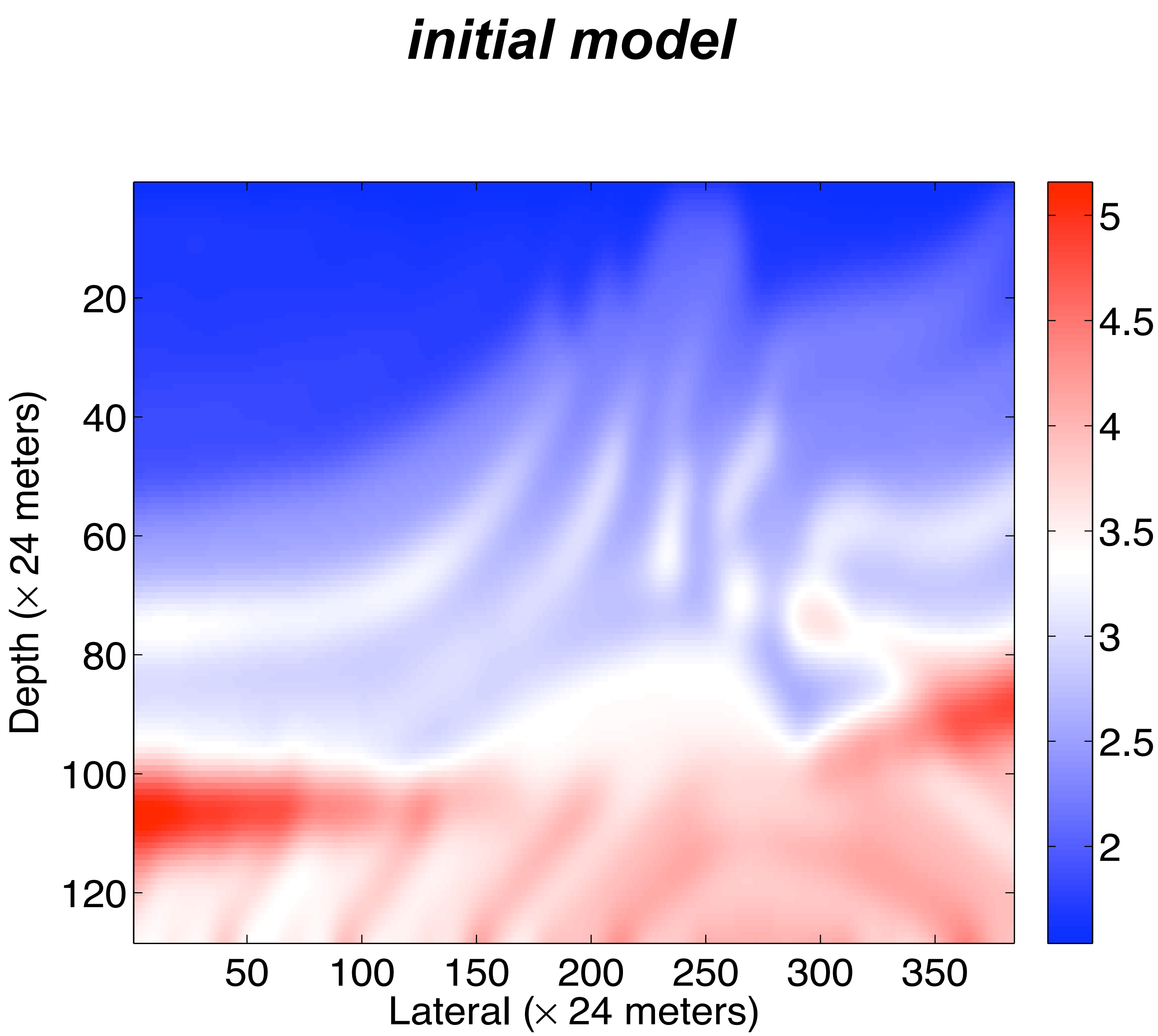
L2 recovery

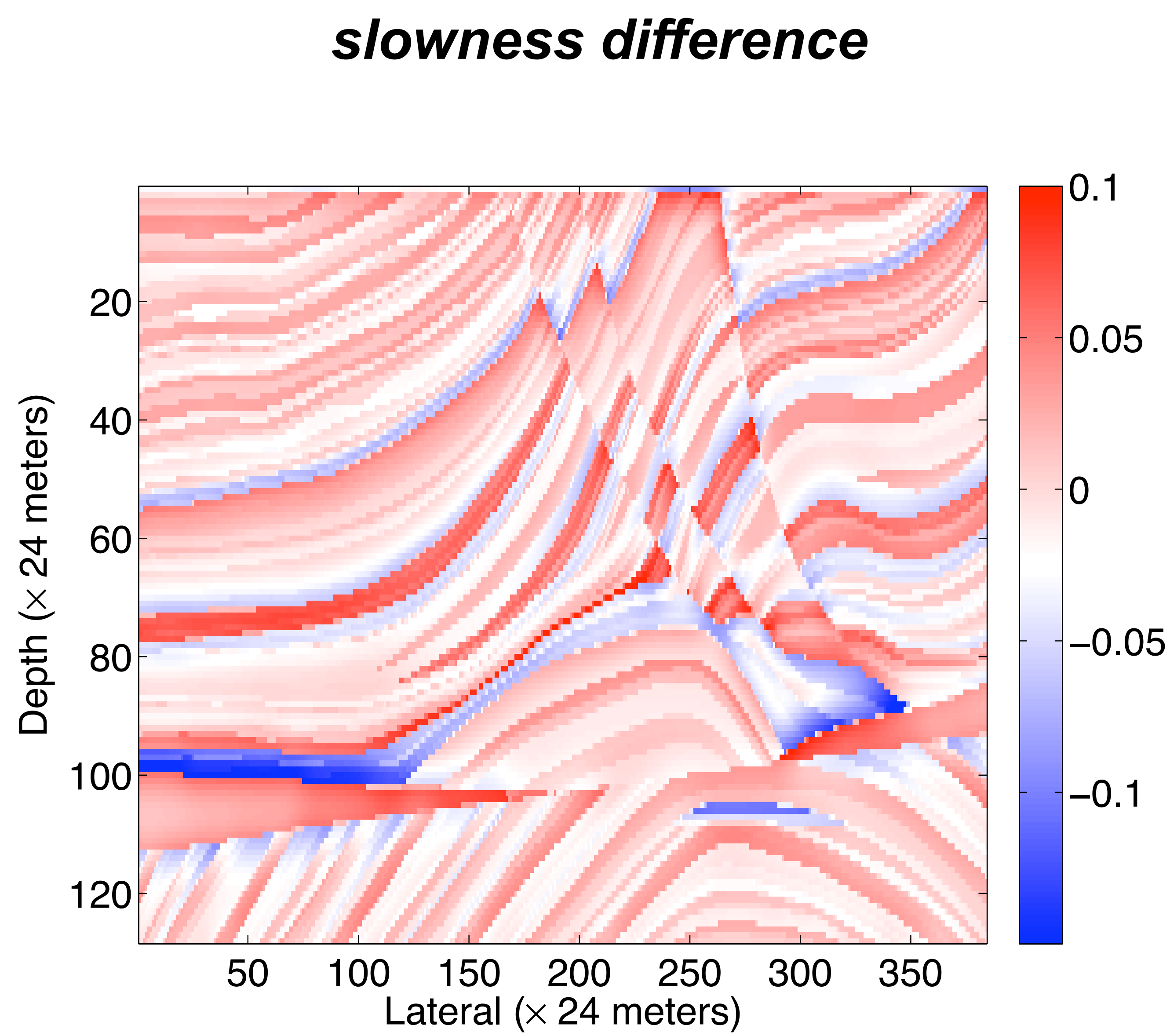
- all 192 shots
- all 50 freqs
- 10 iterations

Batch size is 8100

Initial model

Marmousi model experiment

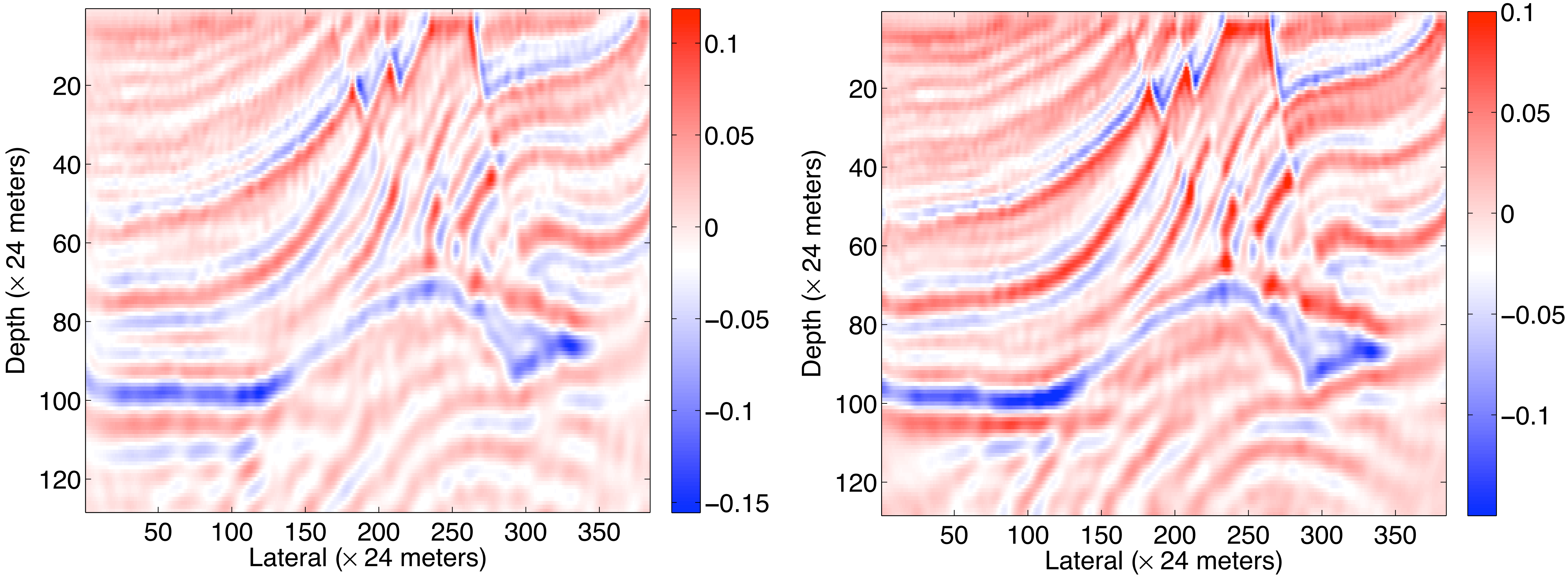




Linearized sparse inversion

14 simultaneous shots 7 random frequencies

L2 recovery with all data | ***sparse* recovery with curvelets**

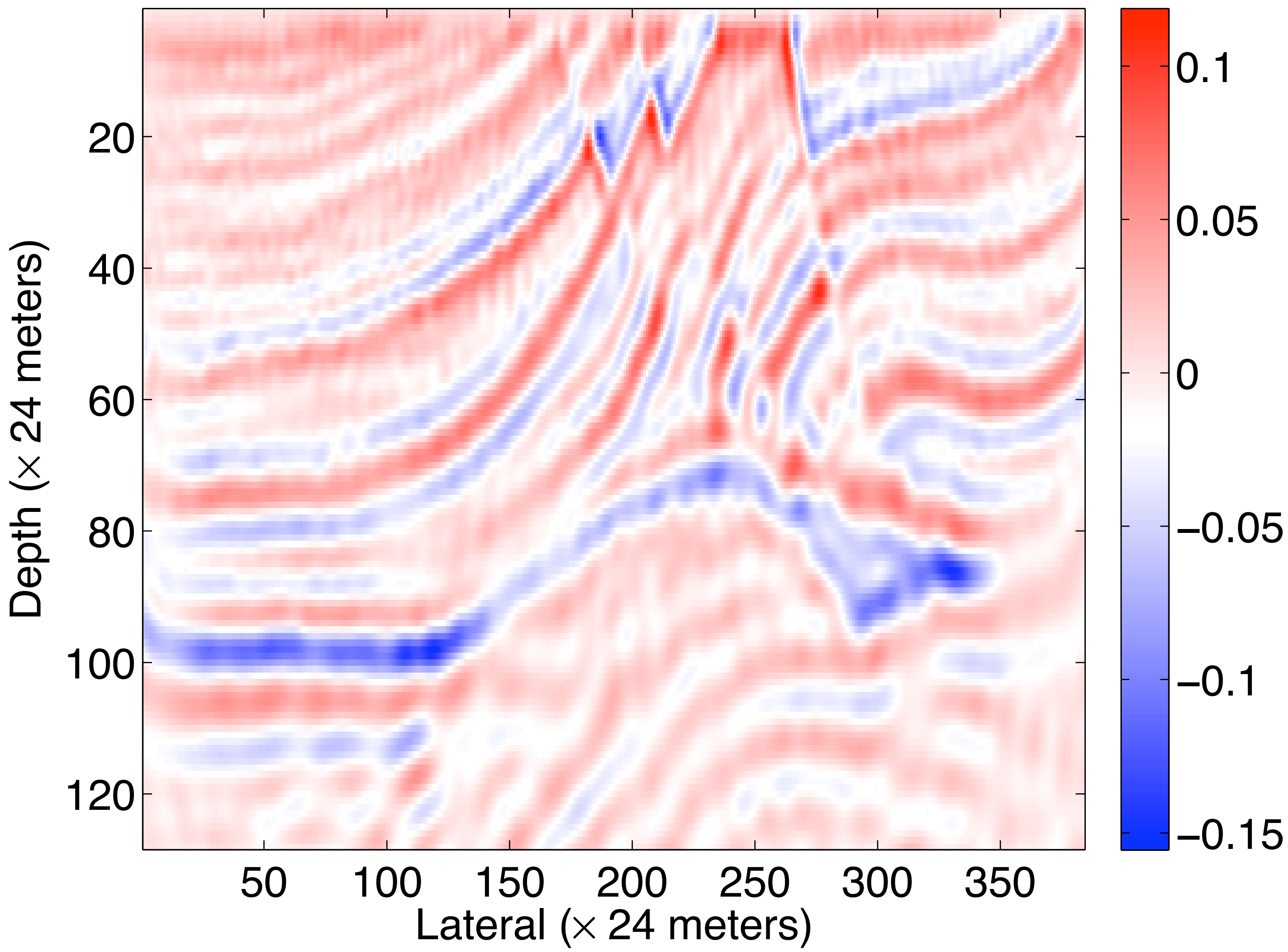


Speed up: x8.3

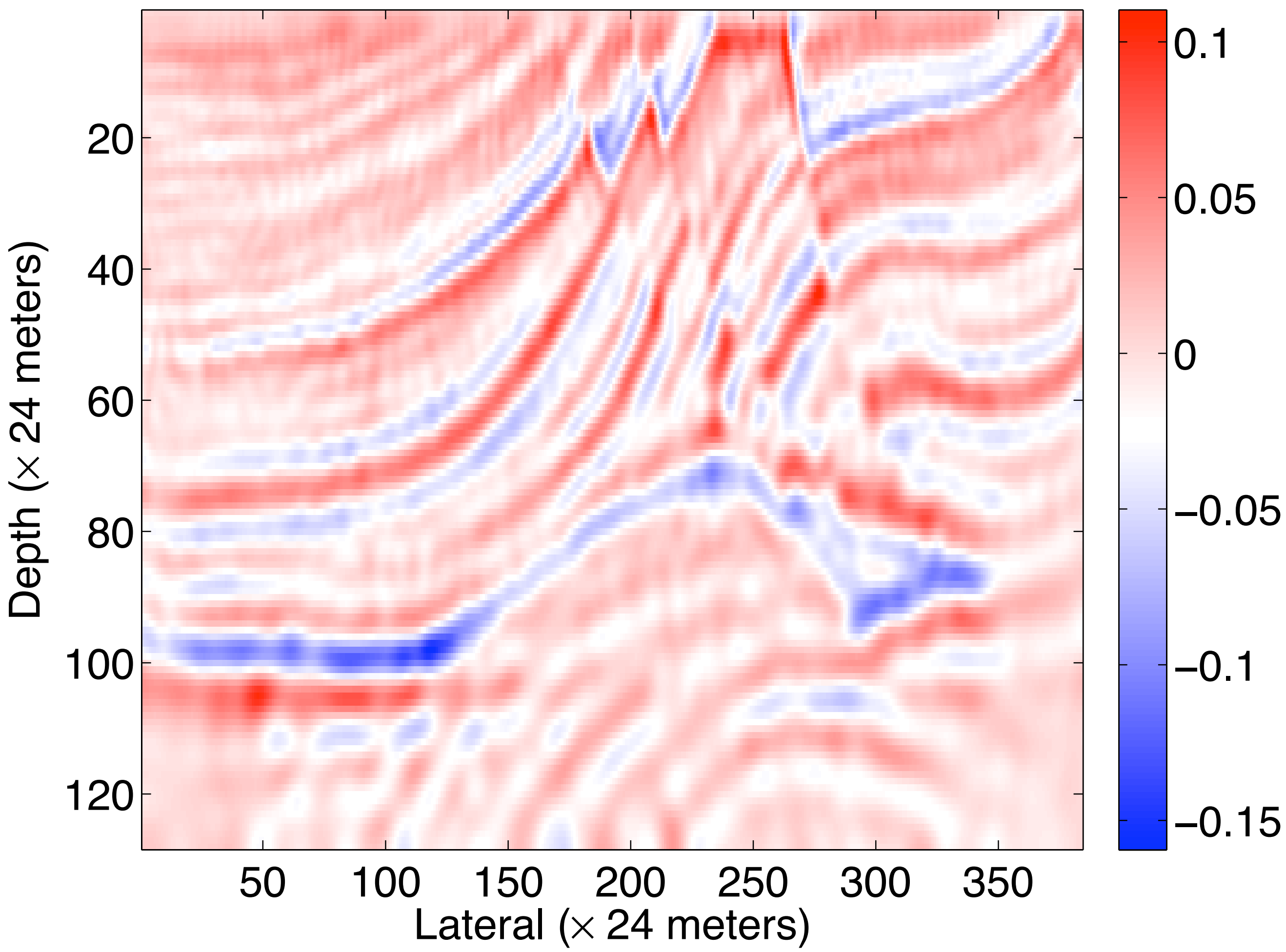
Linearized sparse inversion

8 simultaneous shots 3 random frequencies

L2 recovery with all data



***sparse* recovery with curvelets**

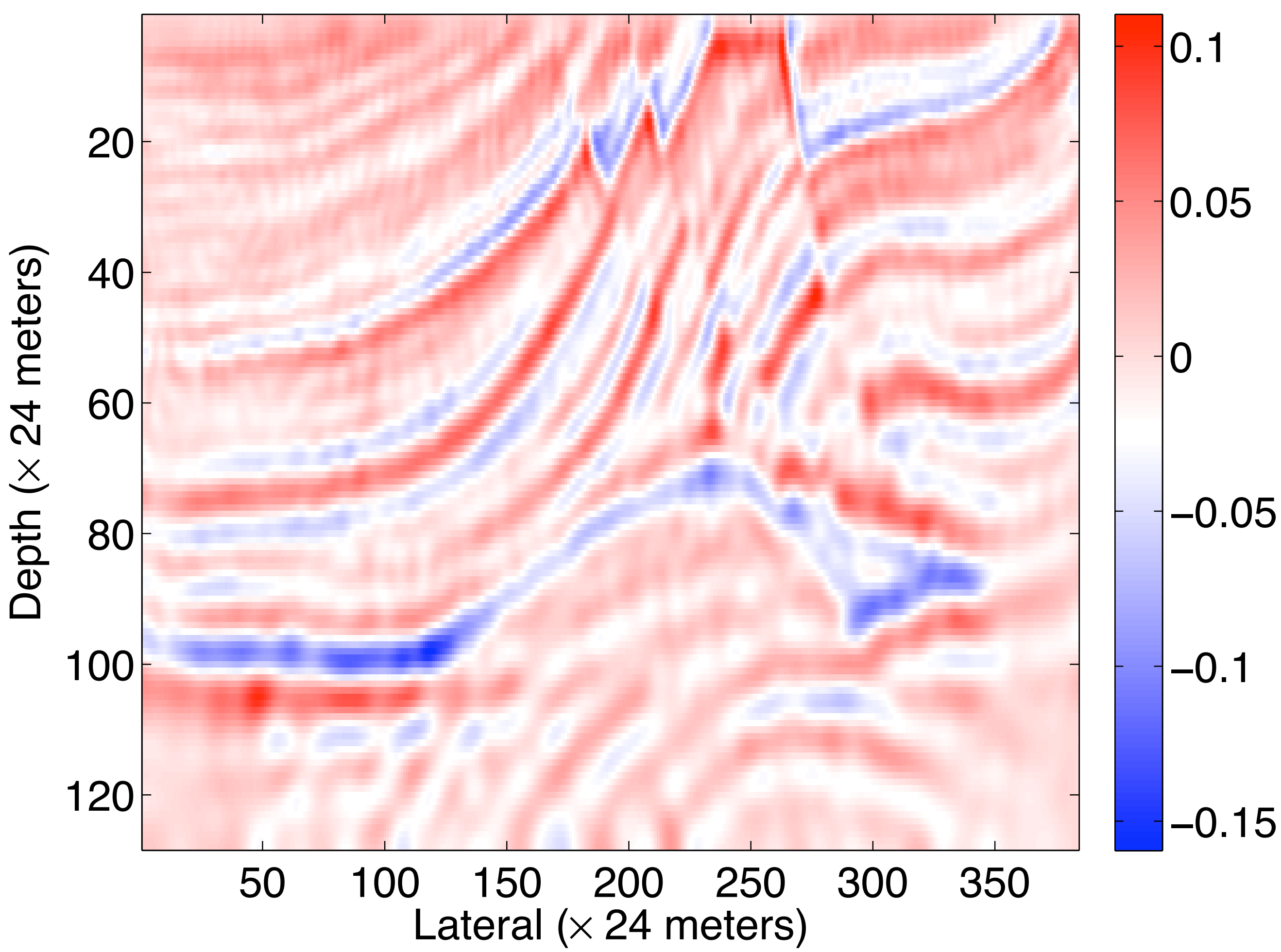


Speed up: x22

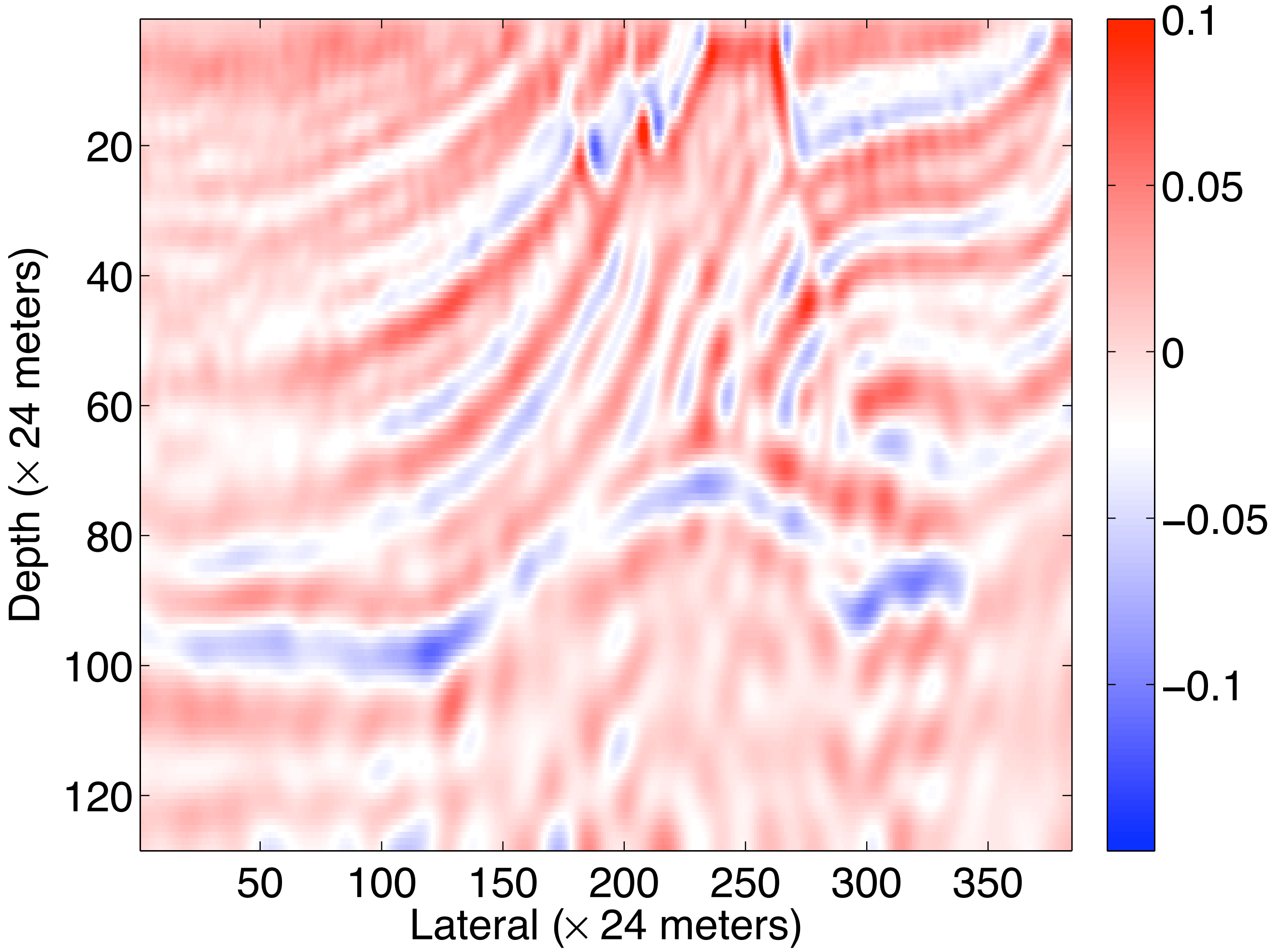
Linearized sparse inversion

8 simultaneous shots 3 random frequencies

sparse recovery with renewals



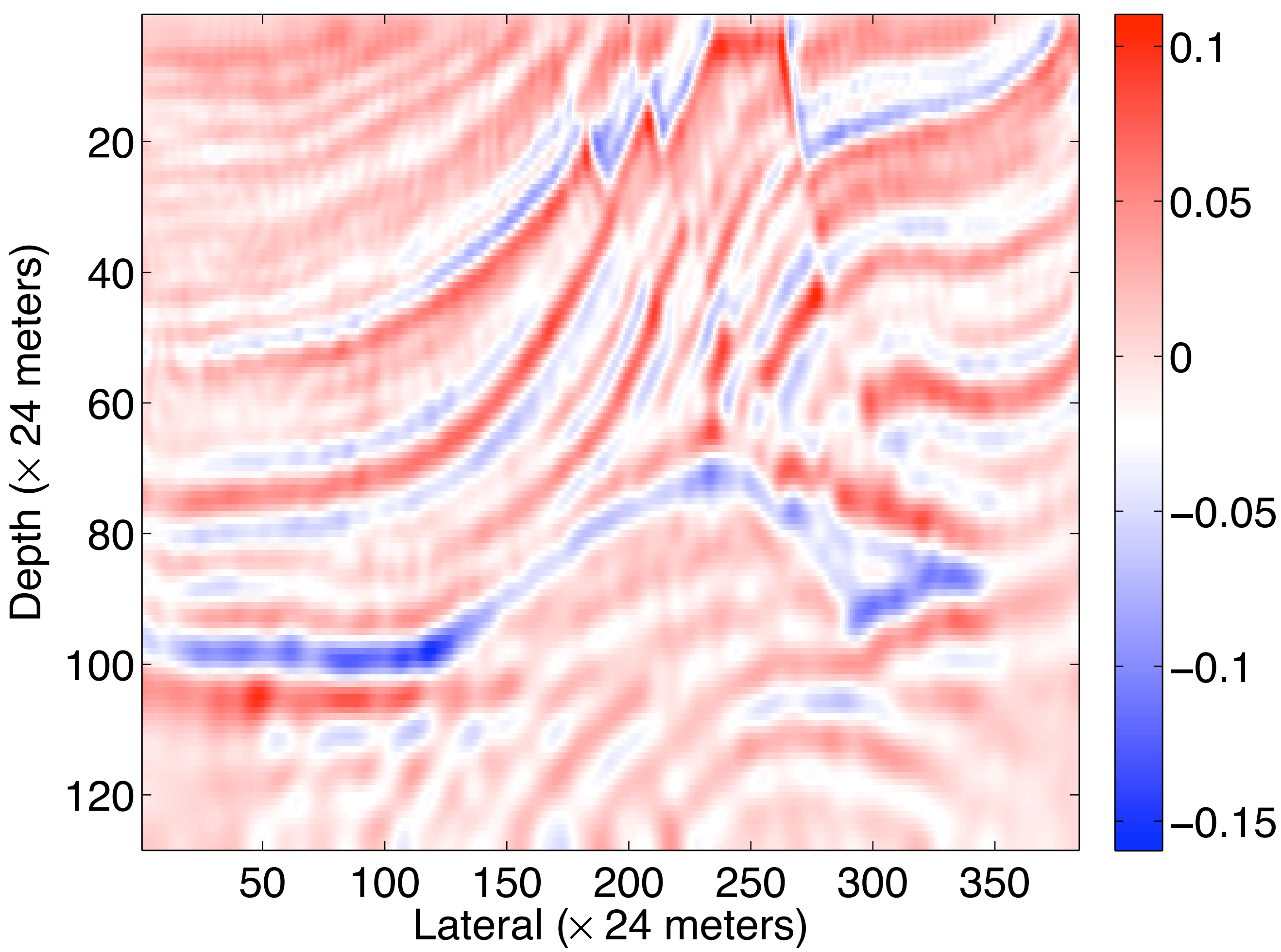
***sparse* recovery without renewals**



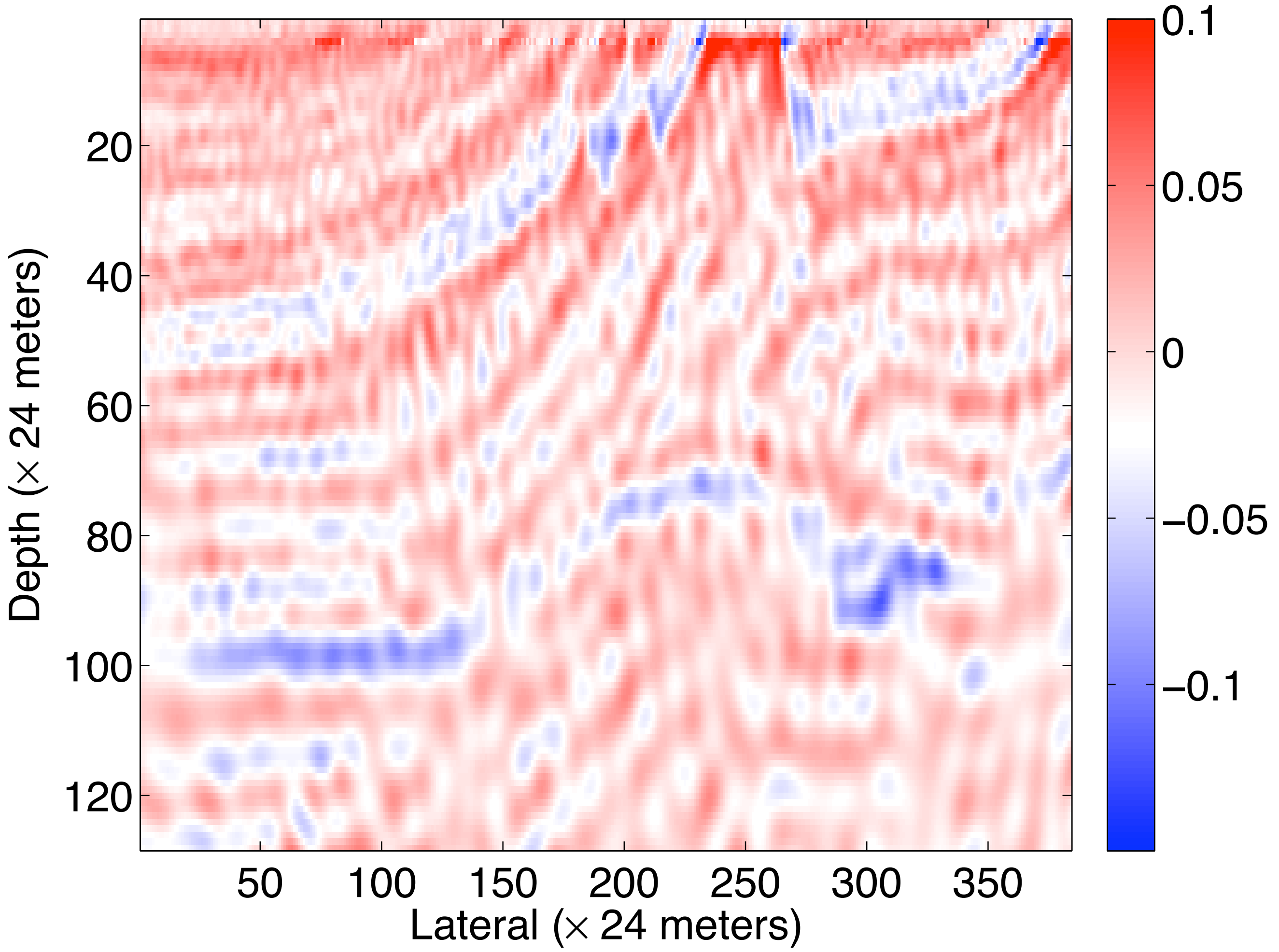
Linearized sparse inversion

8 simultaneous shots 3 random frequencies

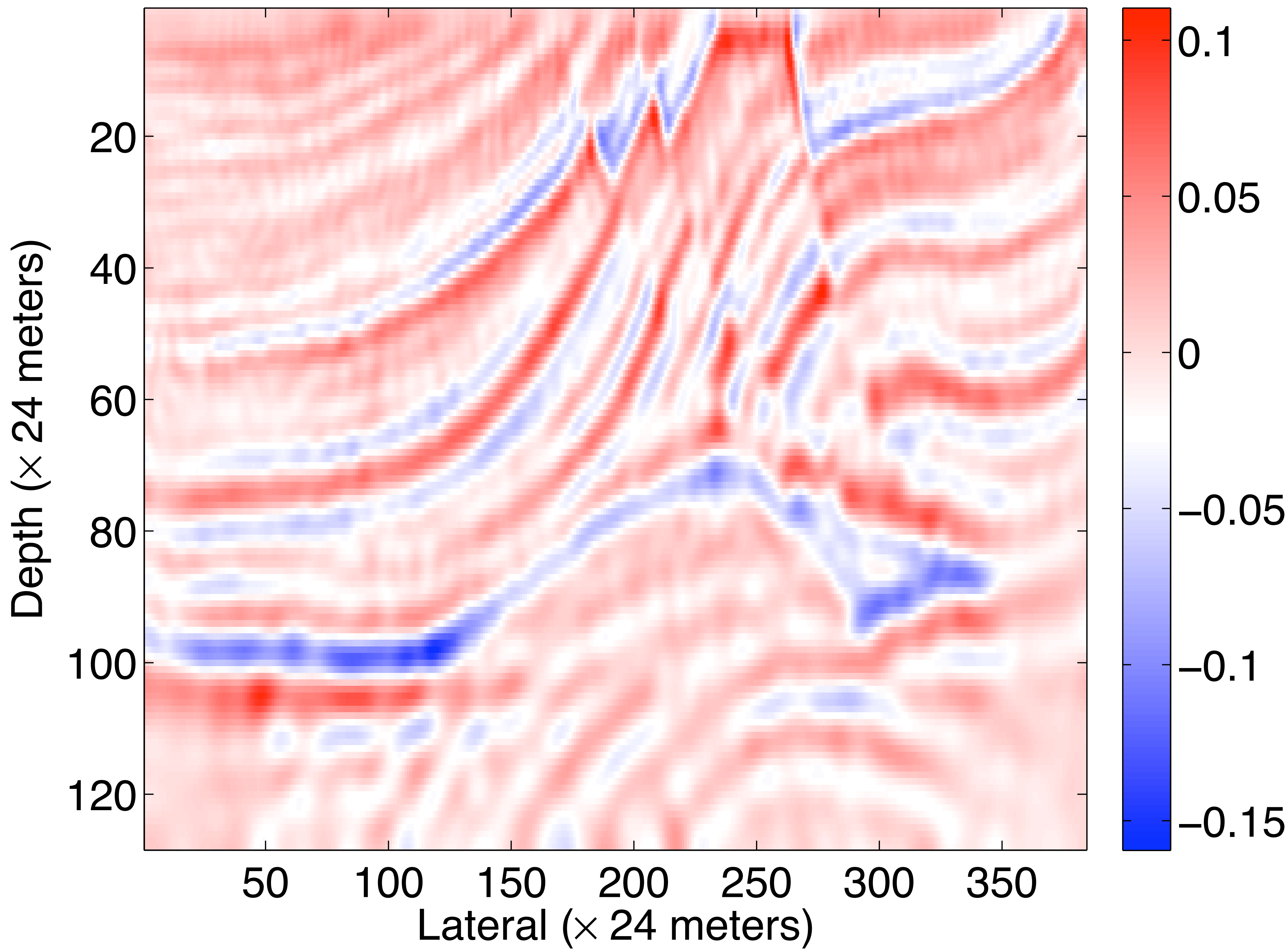
sparse recovery with L1 solver



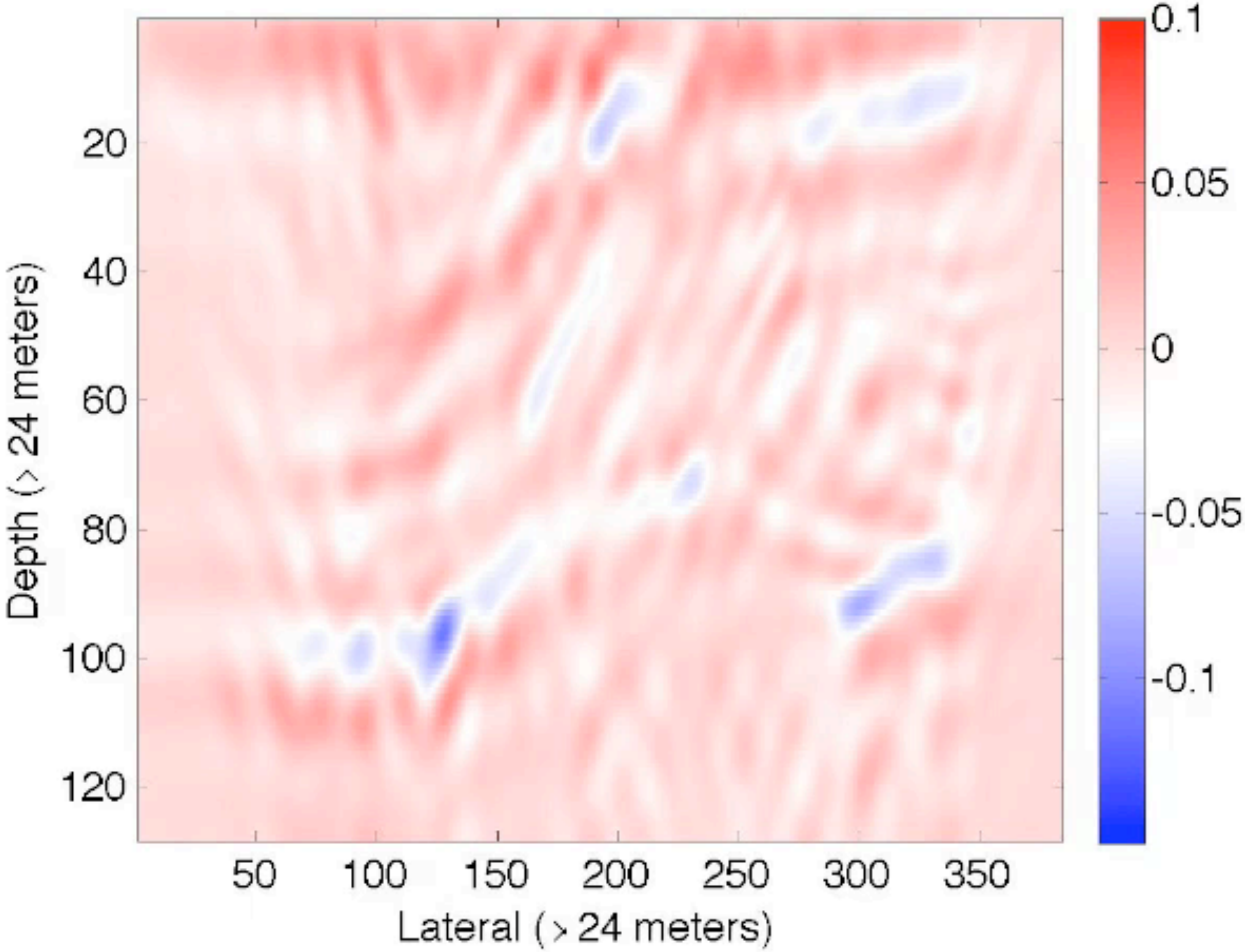
recovery with L2 solver



Linearized sparse inversion



Linearized sparse inversion



Linearized sparse inversion

Subsample ratio	0.0006	0.0013	0.0026	0.0033
n'_f/n'_s		Signal-noise ratio (dB)		
2	**3.1652** (1.4964)	**3.3452** (1.5326)	**3.4022** (1.5529)	**3.4243** (1.5572)
1	**3.2019** (1.5011)	**3.3832** (1.5377)	**3.4523** (1.5610)	**3.4865** (1.5915)
0.5	**3.2253** (1.5128)	**3.3864** (1.5964)	**3.4765** (1.5984)	**3.5063** (1.6245)
Speed up ($\times$)	1536	768	384	307

$$\mathbf{SNR} = 20 \times log_{10}\left(\frac{\|f\|_2}{\|f - \hat{f}\|_2}\right)$$

SNRs for migration without renewals in parentheses

Performance

Methods	L2	L1 with renewals	L1 without renewals
Number of freqs	**30**	**3**	**3**
Number of shots	**192**	**8**	**8**
Number of PDE	**10**	**107**	**107**
Number of Matrix Multipulication	**21**	**226**	**216**
Total cost	**120960**	**5424**	**5184**
Speed up ($\times$)	1	22	23

Observations

Reconstruct images

- from *randomized* subsamplings
- with correct amplitudes

Recovery quality depends on *degree* of *subsampling*

Significant speedups attainable...

Conclusions

A reduction in the # of PDE solves cost by virtue of the reduced system size

Sparse recoveries offset *random interferences* due to *source encoding*

Hight-quality & high-resolution migration images have been *achieved with significant accelerations*

Improvements come from *sparsity promotion & curvelets*

Indications that the *curse of dimensionality* can be removed...

Use this formulation to solve Gauss-Newton steps part of FWI (tomorrow's talk)

SLIM

Acknowledgements

- Authors of CurveLab.
- 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, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco.

SLIM

Thank you

slim.eos.ubc.ca

11:30-12:00 PM Xiang Li
Full-waveform inversion with randomized L1 recovery for the model updates

Further reading

Compressive sensing

- *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information by Candes, 06*.
- *Compressed Sensing* by D. Donoho, '06
- Curvelets and Wave Atoms for Mirror-Extended Images by L. Demanet, L. Ying, 07.

Simultaneous acquisition

- *A new look at simultaneous sources* by Beasley et. al., '98.
- *Changing the mindset in seismic data acquisition* by Berkhout '08.

Simultaneous simulations, imaging, and full-wave inversion:

- *Faster shot-record depth migrations using phase encoding* by Morton & Ober, '98.
- *Phase encoding of shot records in prestack migration* by Romero et. al., '00.
- *Efficient Seismic Forward Modeling using Simultaneous Random Sources and Sparsity* by N. Neelamani et. al., '08.
- *Compressive simultaneous full-waveform simulation* by FJH et. al., '09.
- *Fast full-wavefield seismic inversion using encoded sources* by Krebs et. al., '09
- *Randomized dimensionality reduction for full-waveform inversion* by FJH & X. Li, '10

Stochastic optimization and machine learning:

- *A Stochastic Approximation Method* by Robbins and Monro, 1951
- *Neuro-Dynamic Programming* by Bertsekas, '96
- *Robust stochastic approximation approach to stochastic programming* by Nemirovski et. al., '09
- *Stochastic Approximation and Recursive Algorithms and Applications* by Kushner and Lin
- *Stochastic Approximation approach to Stochastic Programming* by Nemirovski
- *An effective method for parameter estimation with PDE constraints with multiple right hand sides. by* Eldad Haber, Matthias Chung, and Felix J. Herrmann. '10