SINBAD Consortium 2010

Randomized full-waveform inversion with compressive sensing

Xiang Li and Felix J. Herrmann

SLIM
University of British Columbia

SLIM

Full-waveform inversion

objective function:

$$\min_{\mathbf{m}} \ \Phi(\mathbf{m}) := \left\{ \frac{1}{2K} \sum_{i=1}^{K} \|\mathbf{d}_i - \mathcal{F}_i[\mathbf{m}]\mathbf{q}_i\|_2^2 = \frac{1}{2}\|\mathbf{D} - \boldsymbol{\mathcal{F}}[\mathbf{m}]\mathbf{Q}\|_F^2 \right\}$$

$$\begin{aligned} \mathbf{d}_i &= \text{monochromatic shot records} \\ \mathbf{q}_i &= \text{monochromatic sources} \\ \mathcal{F}_i[\mathbf{m}, \mathbf{q}_i] &= \text{monochromatic nonlinear forward operators} \end{aligned}$$

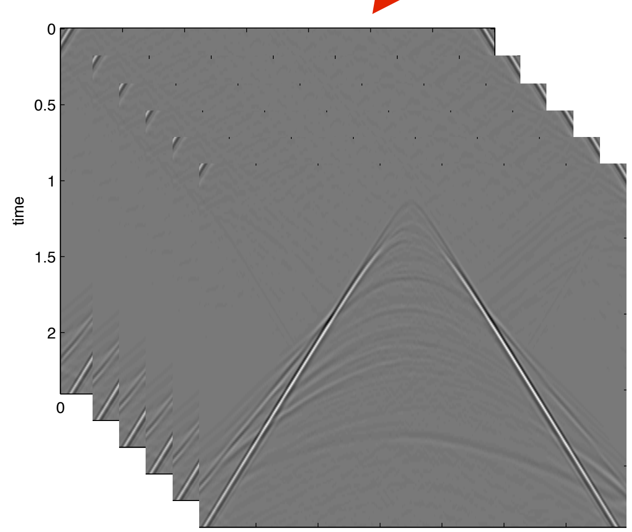
Gauss-Newton

GN subproblem:

$$\min_{\mathbf{m}} \quad \underline{\Phi}(\mathbf{m}) := \frac{1}{2} \| \underbrace{\mathbf{D} - \mathcal{F}[\mathbf{m};\mathbf{Q}]}_{\mathbf{b}} - \underbrace{\nabla \mathcal{F}[\mathbf{m};\mathbf{Q}]}_{\mathbf{A}} \underbrace{\delta \mathbf{m}}_{\mathbf{x}} \|_F^2 .$$

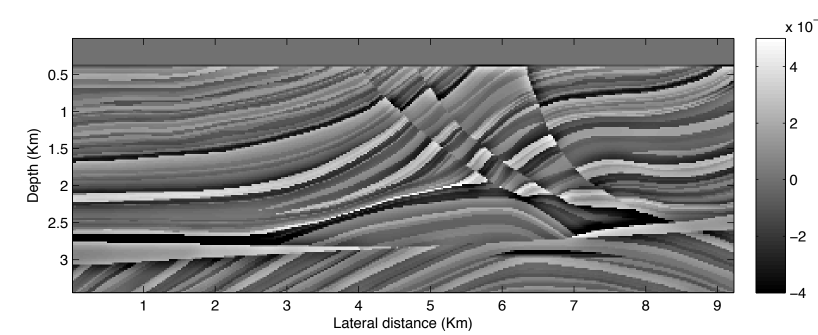
This is a least-squares problem:

$$\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2$$



$n_s \times n_f \times n_r$

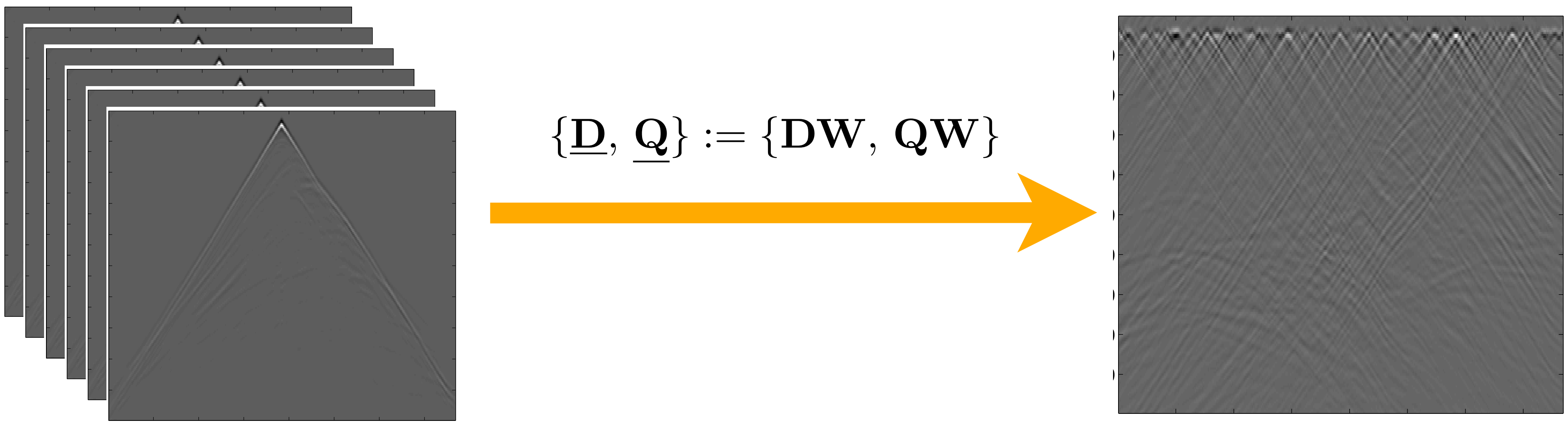
large over-determined linear operator



$n_x \times n_z$

Dimensionality reduction

$$\min_{\mathbf{m}} \ \underline{\Phi}(\mathbf{m}) := \left\{ \frac{1}{2K'} \sum_{i=1}^{K'} \|\mathbf{D}\mathbf{w}_i - \mathcal{F}_i[\mathbf{m}]\mathbf{w}_i\|_2^2 = \frac{1}{2}\|\underline{\mathbf{D}} - \boldsymbol{\mathcal{F}}[\mathbf{m};\underline{\mathbf{Q}}]\|_F^2 \right\}$$



GN subproblem:

$$\min_{\mathbf{m}} \quad \underline{\Phi}(\mathbf{m}) := \frac{1}{2}\|\underline{\mathbf{D}} - \boldsymbol{\mathcal{F}}[\mathbf{m};\underline{\mathbf{Q}}] - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m};\underline{\mathbf{Q}}]\delta\mathbf{m}\|_F^2.$$

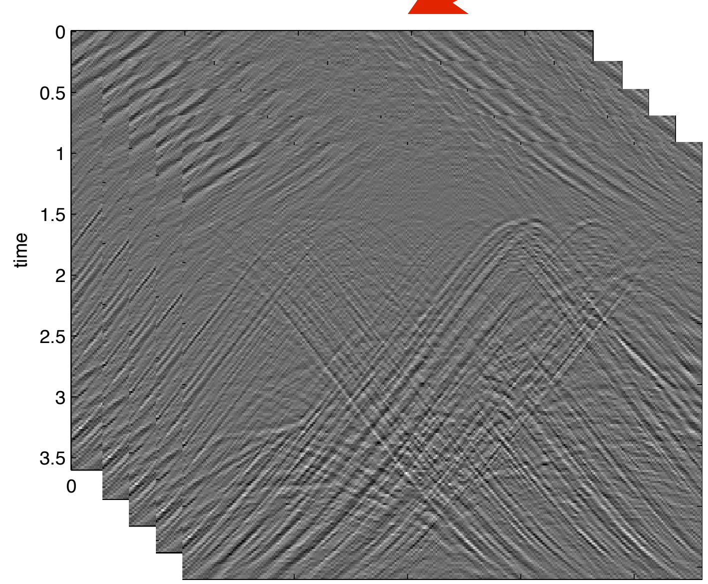
Gauss-Newton

GN subproblem:

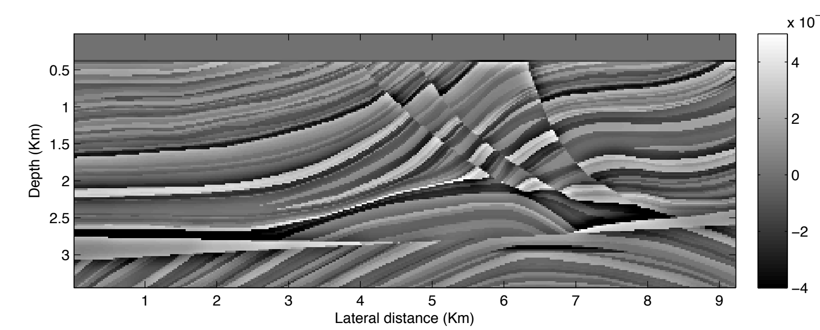
$$\min_{\mathbf{m}} \quad \underline{\Phi}(\mathbf{m}) := \frac{1}{2}\|\underbrace{\underline{\mathbf{D}} - \mathcal{F}[\mathbf{m};\underline{\mathbf{Q}}]}_{\underline{\mathbf{b}}} - \underbrace{\nabla\mathcal{F}[\mathbf{m};\underline{\mathbf{Q}}]}_{\underline{\mathbf{A}}}\underbrace{\delta\mathbf{m}}_{\mathbf{x}}\|_F^2.$$

This is a least-squares problem:

$$\min_{\mathbf{x}} \|\underline{\mathbf{b}} \quad - \quad \underline{\mathbf{A}} \quad \mathbf{x}\|_2^2$$



underdetermined
linear operator



$ns' \times nf' \times nr$

$nx \times nz$

Sparsity promotion

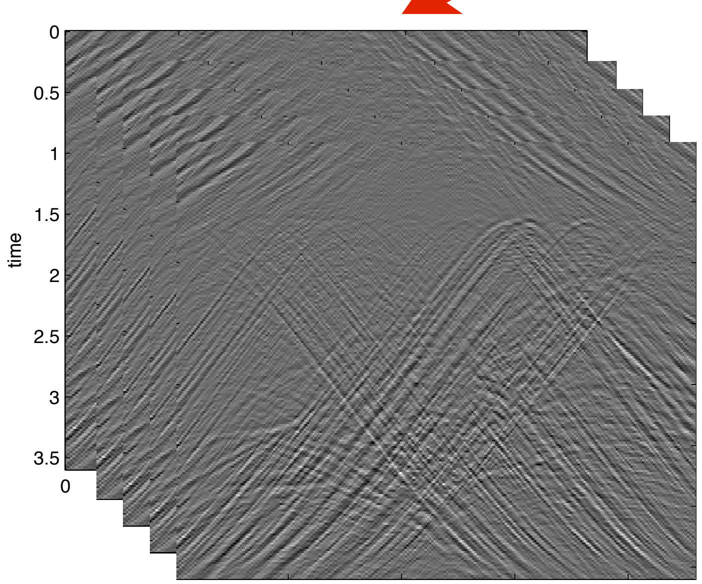
Modified GN subproblem (LASSO):

$$\min_{\mathbf{x}} \frac{1}{2}\|\underline{\boldsymbol{\delta}\mathbf{D}} - \nabla\mathcal{F}[\mathbf{m};\underline{\mathbf{Q}}]\mathbf{S}^H\mathbf{x}\|_F^2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \leq \tau$$

$$\underline{\mathbf{b}} \qquad \underline{\mathbf{A}} \quad \mathbf{S}^H \; \mathbf{x}$$

S is a sparsifying transform.

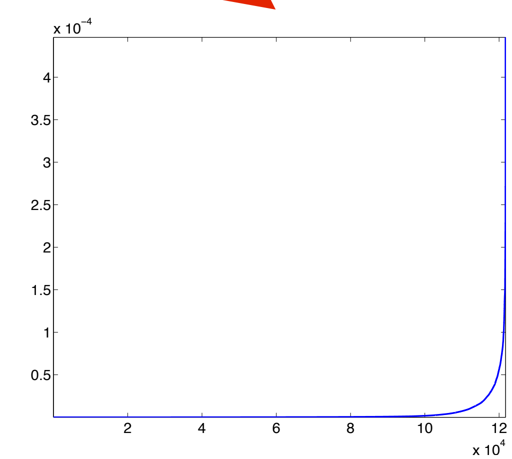
$$\min_{\mathbf{x}} \|\underline{\mathbf{b}} - \underline{\mathbf{A}} \; \mathbf{S}^H \; \mathbf{x}\|_2^2$$

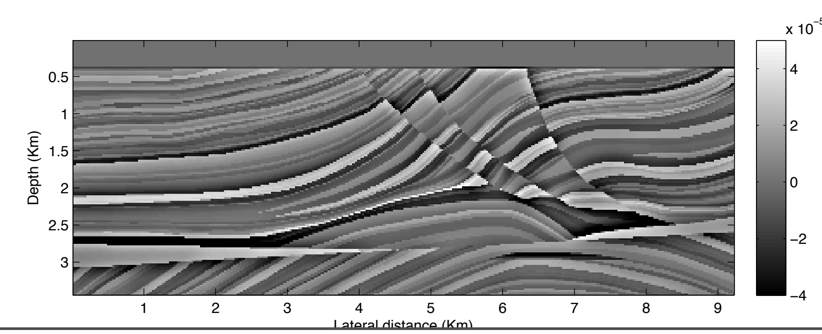


$ns' \times nf' \times nr$

under-determined linear operator

$\mathbf{S}^H$





$nx \times nz$

Sparse promotion

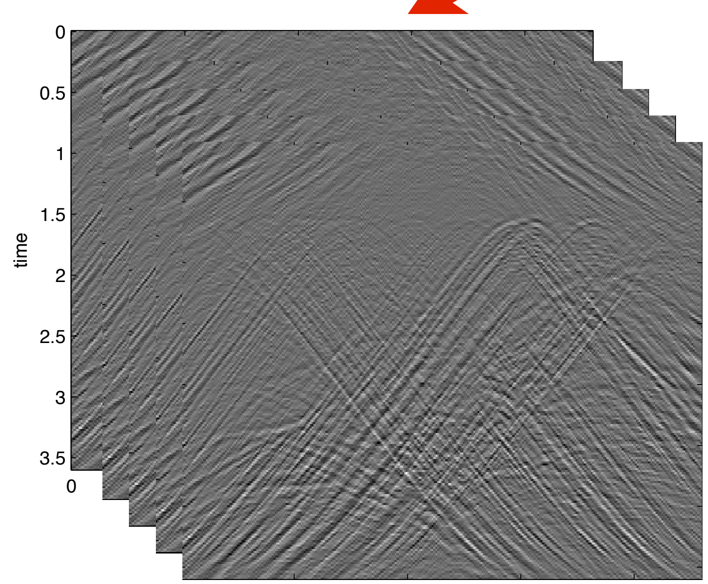
Modified GN subproblem (LASSO)

$$\min_{\mathbf{x}} \frac{1}{2}\|\underline{\boldsymbol{\delta}\mathbf{D}} - \nabla\mathcal{F}[\mathbf{m};\underline{\mathbf{Q}}]\mathbf{S}^H\mathbf{x}\|_F^2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \leq \tau$$

$$\underline{\mathbf{b}} \qquad \underline{\mathbf{A}} \quad \mathbf{S}^H \; \mathbf{x}$$

This is a least-squares problem:

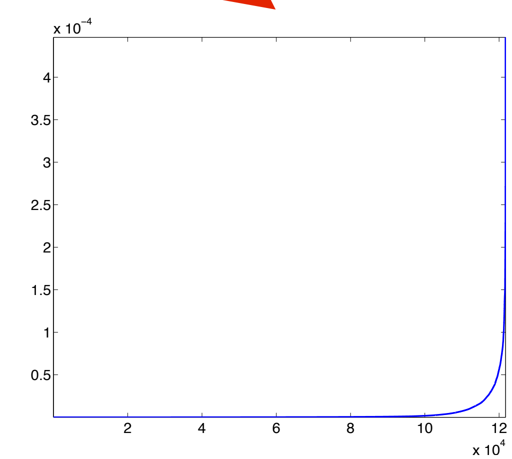
$$\min_{\mathbf{x}} \|\underline{\mathbf{b}} - \underline{\mathbf{A}}\,\mathbf{S}^H\,\mathbf{x}\|_2^2$$

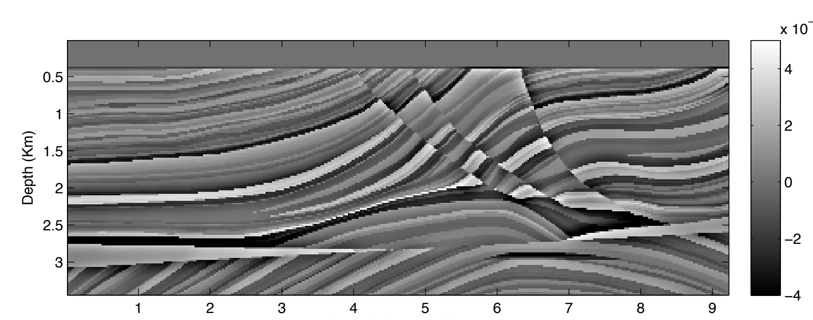


$ns' \times nf' \times nr$

under-determined linear operator

$\mathbf{S}^H$

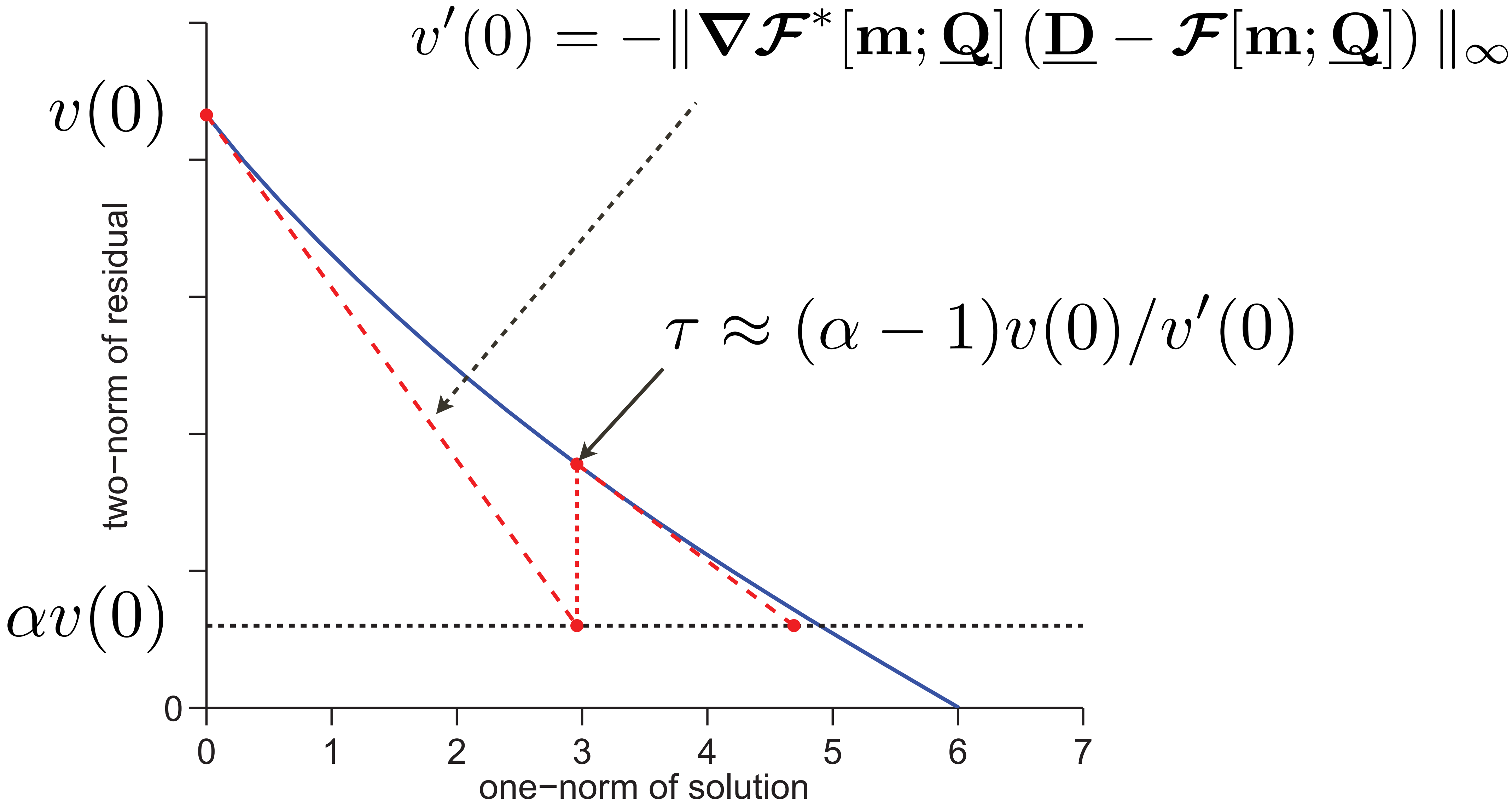


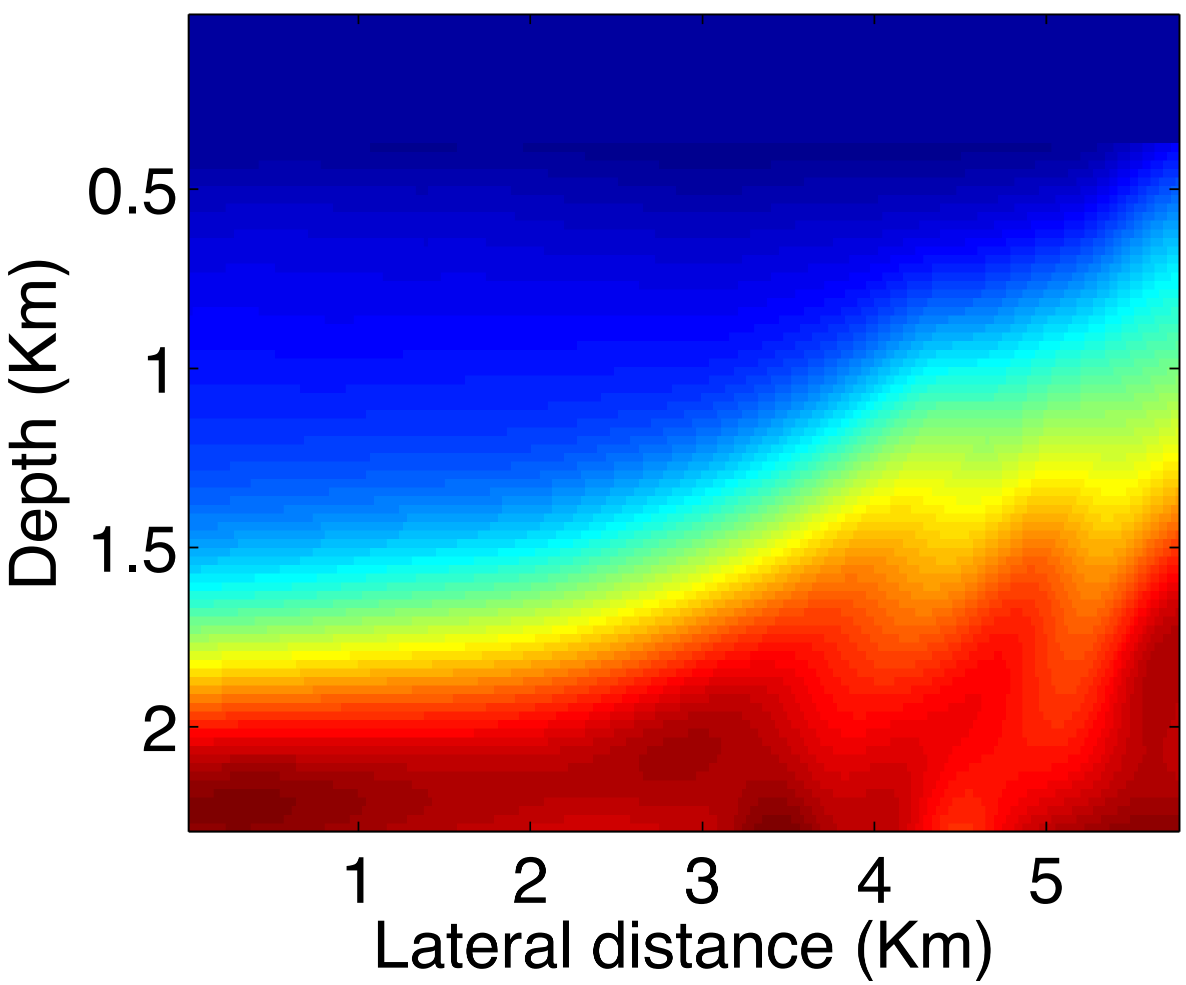


$nx \times nz$

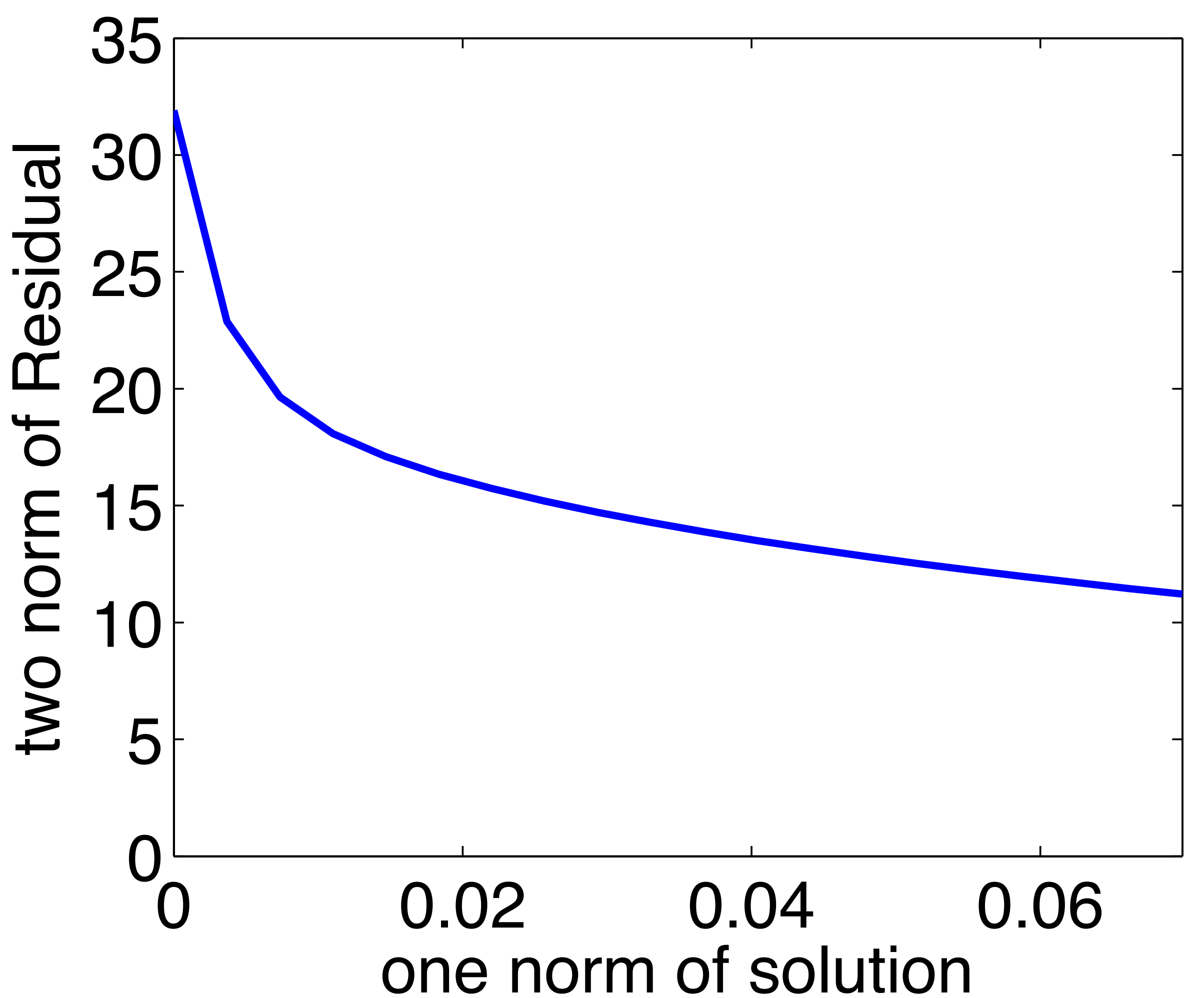
Pareto curve

Compute τ using Pareto curve [van den Berg & Friendlander, '08]

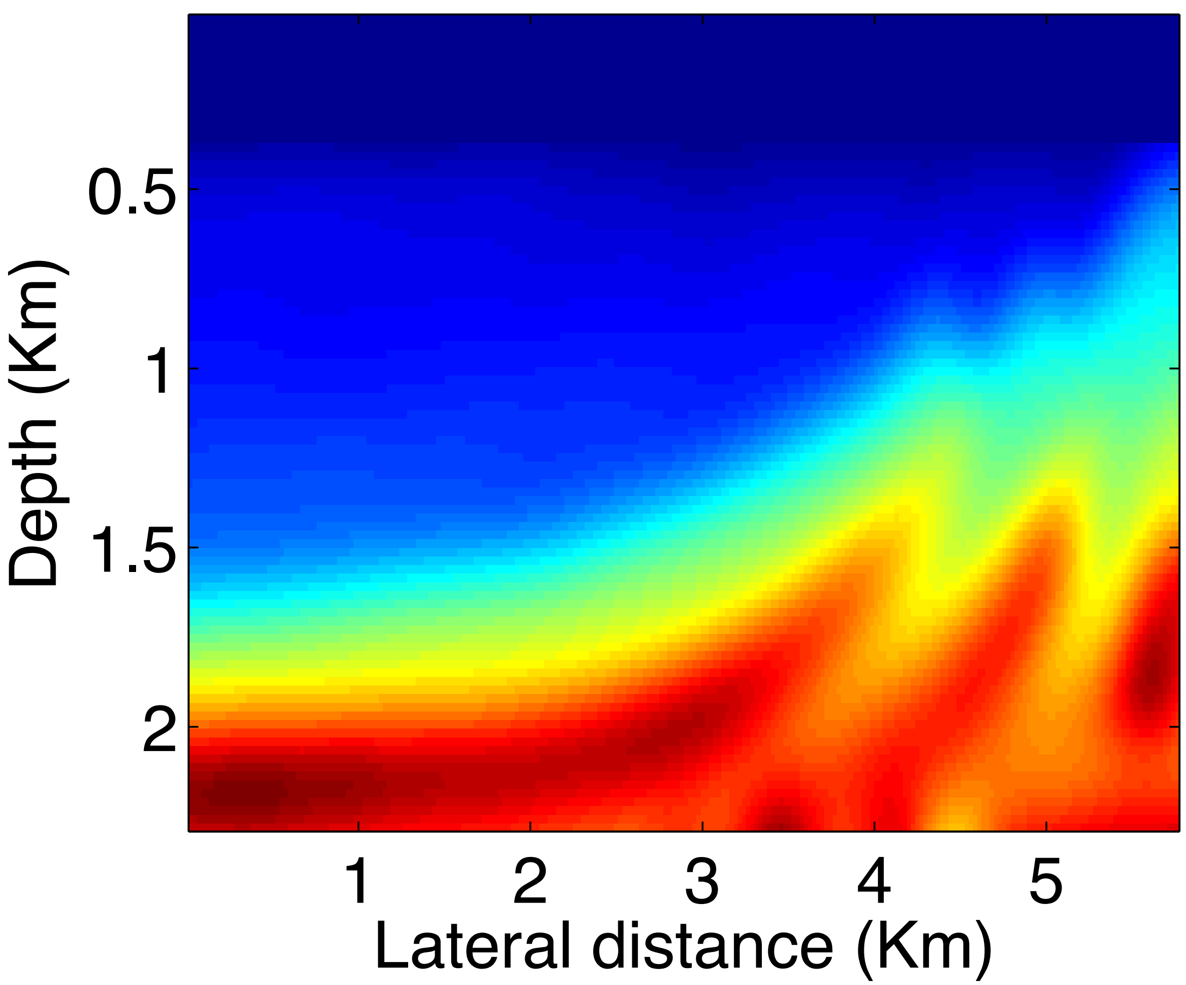


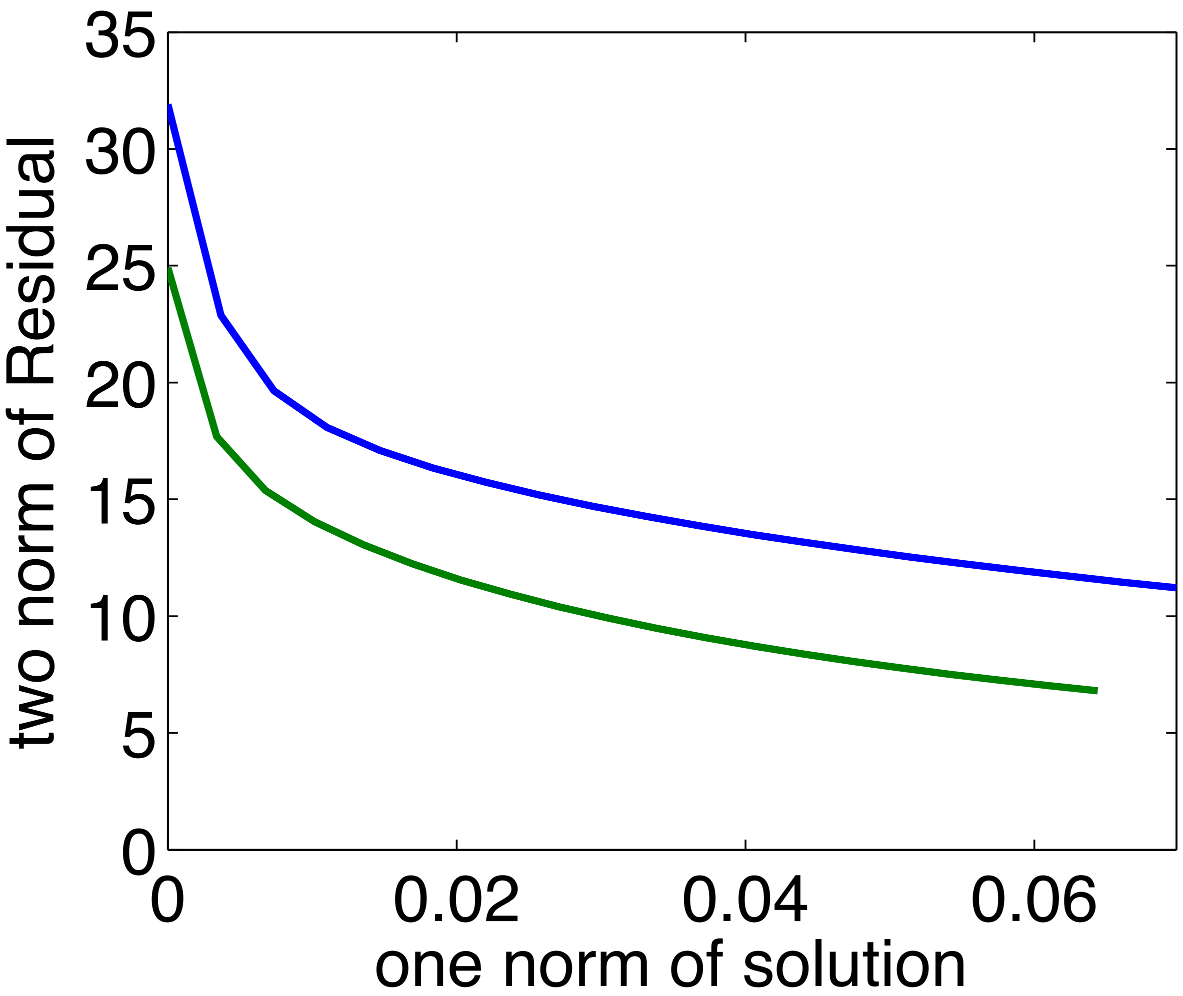


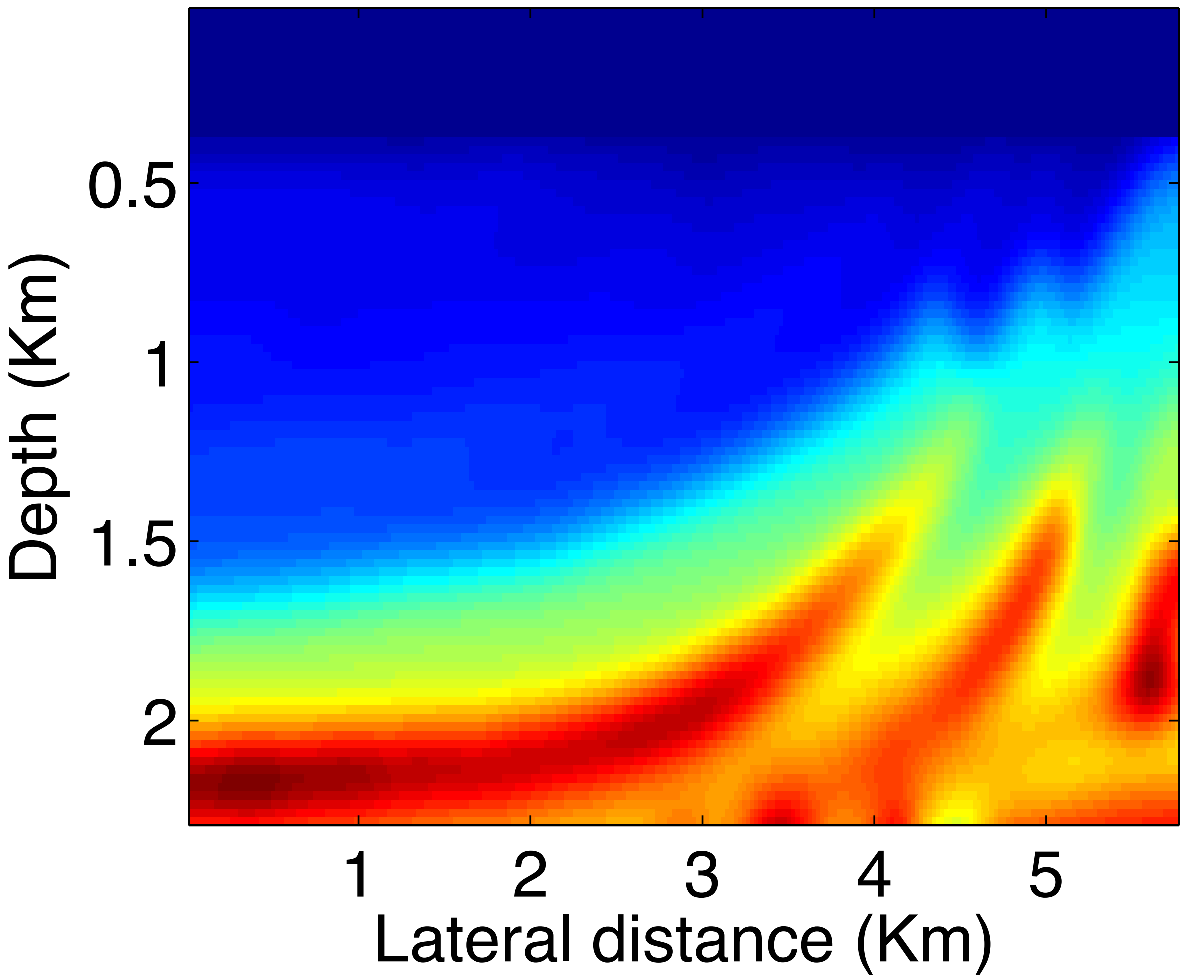
0.5
1
1.5
2
Depth (Km)
1
2
3
4
5
Lateral distance (Km)

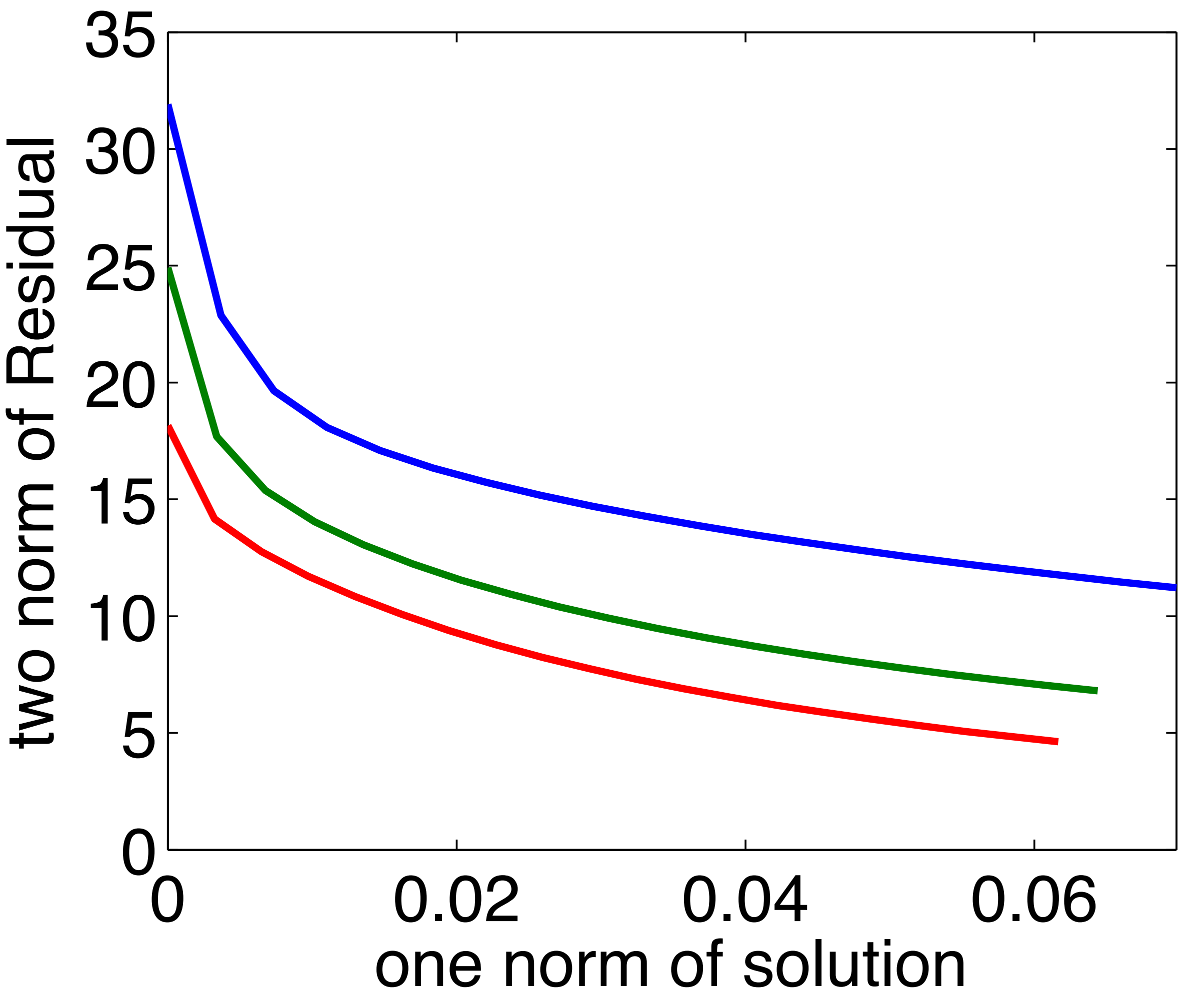


35
30
25
20
15
10
5
0
two norm of Residual
0
0.02
0.04
0.06
one norm of solution

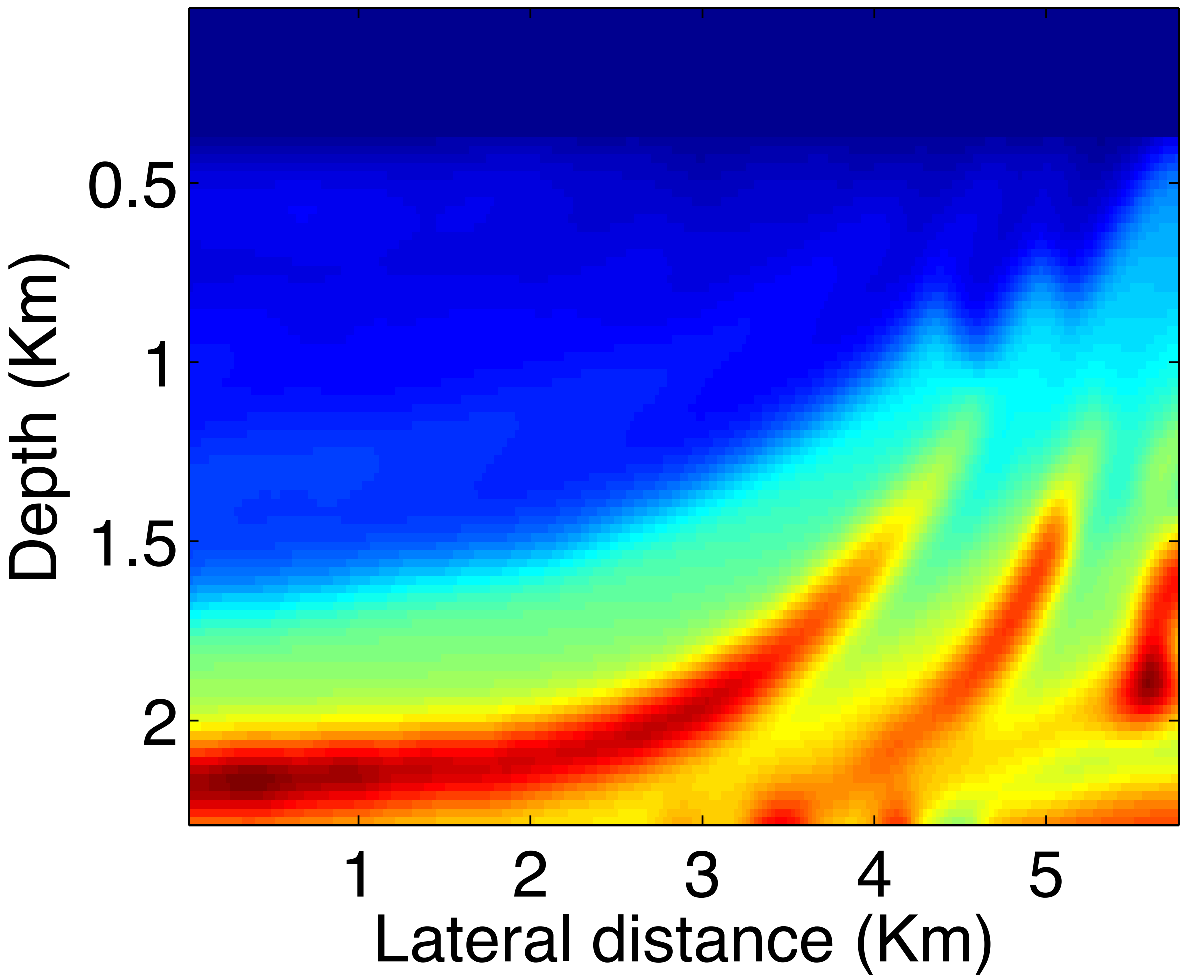

0.5
1
1.5
2
Depth (Km)
1 2 3 4 5
Lateral distance (Km)


35
30
25
20
15
10
5
0
two norm of Residual
0 0.02 0.04 0.06
one norm of solution

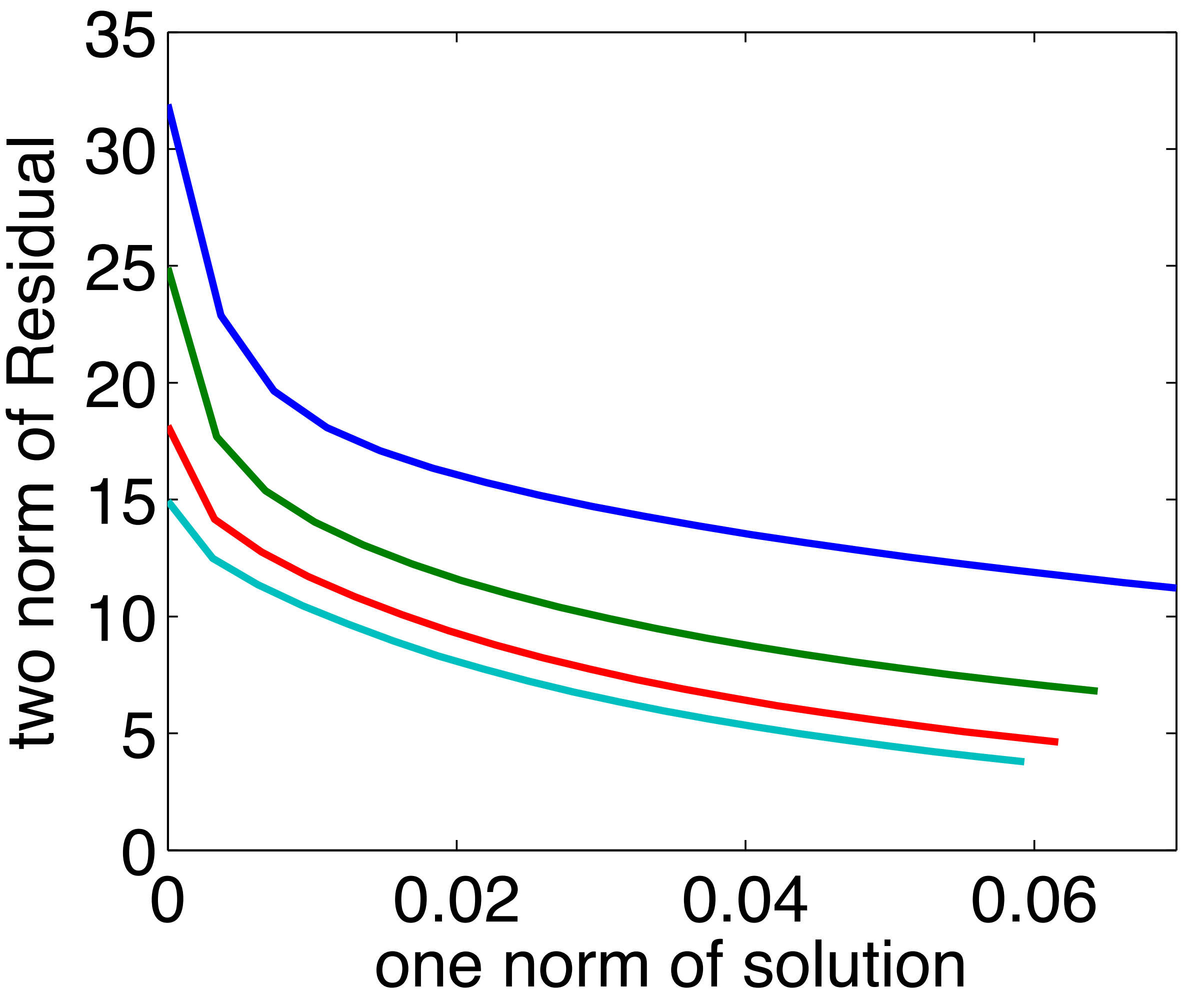

0.5
1
1.5
2
Depth (Km)
1
2
3
4
5
Lateral distance (Km)


35
30
25
20
15
10
5
0
two norm of Residual
0
0.02
0.04
0.06
one norm of solution

SLIM

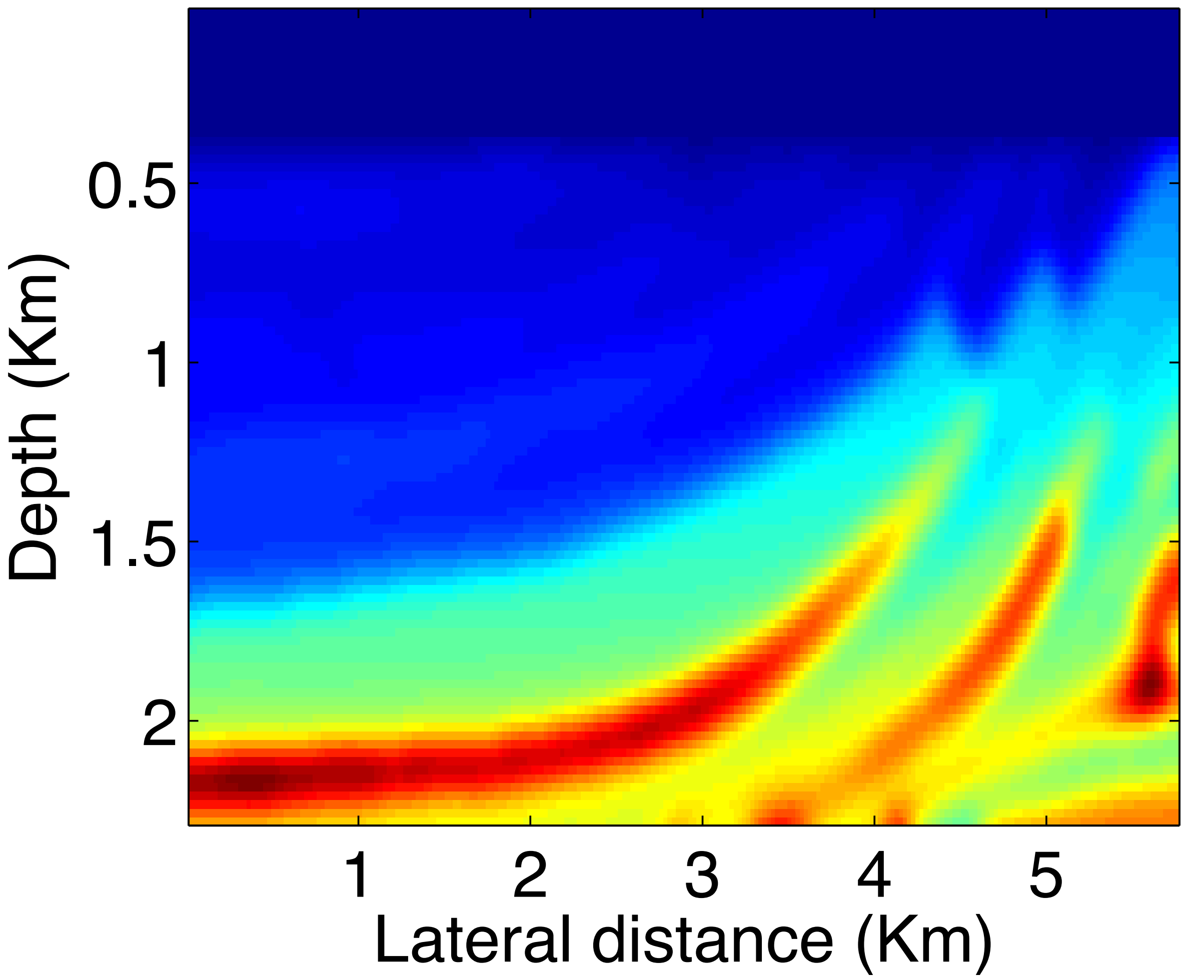


Depth (Km)
0.5
1
1.5
2
1
2
3
4
5
Lateral distance (Km)

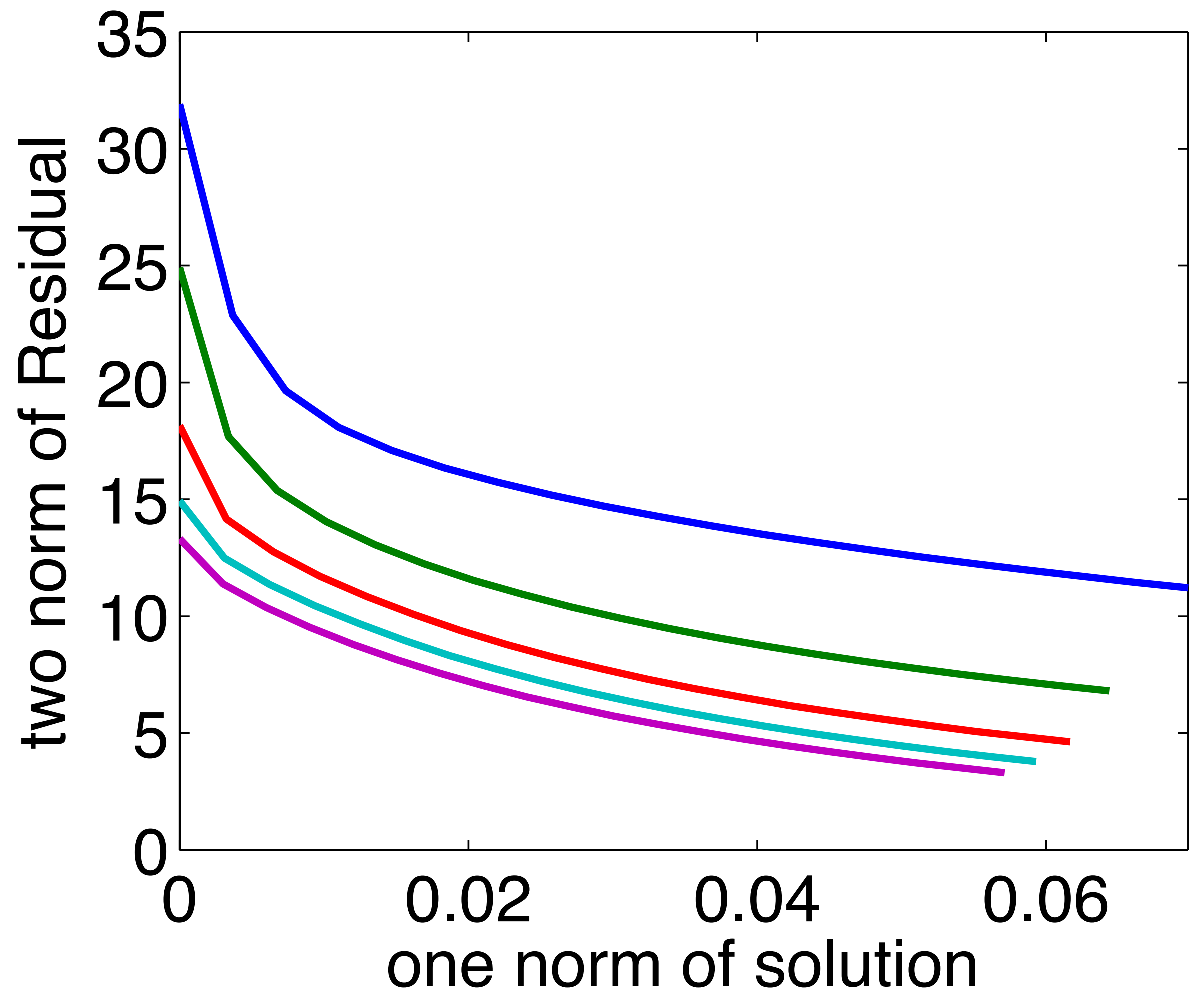


two norm of Residual
35
30
25
20
15
10
5
0
0
0.02
0.04
0.06
one norm of solution

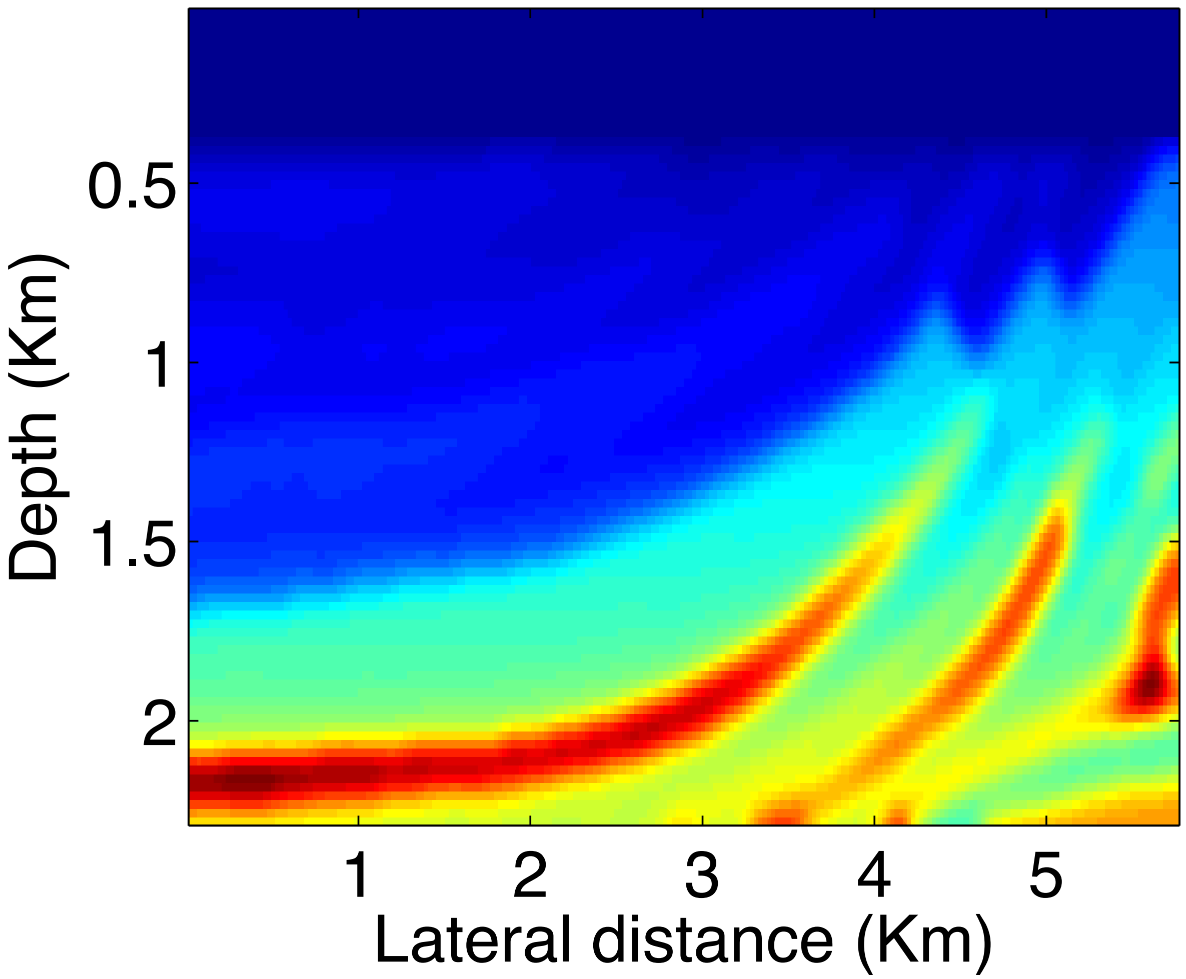
SLIM

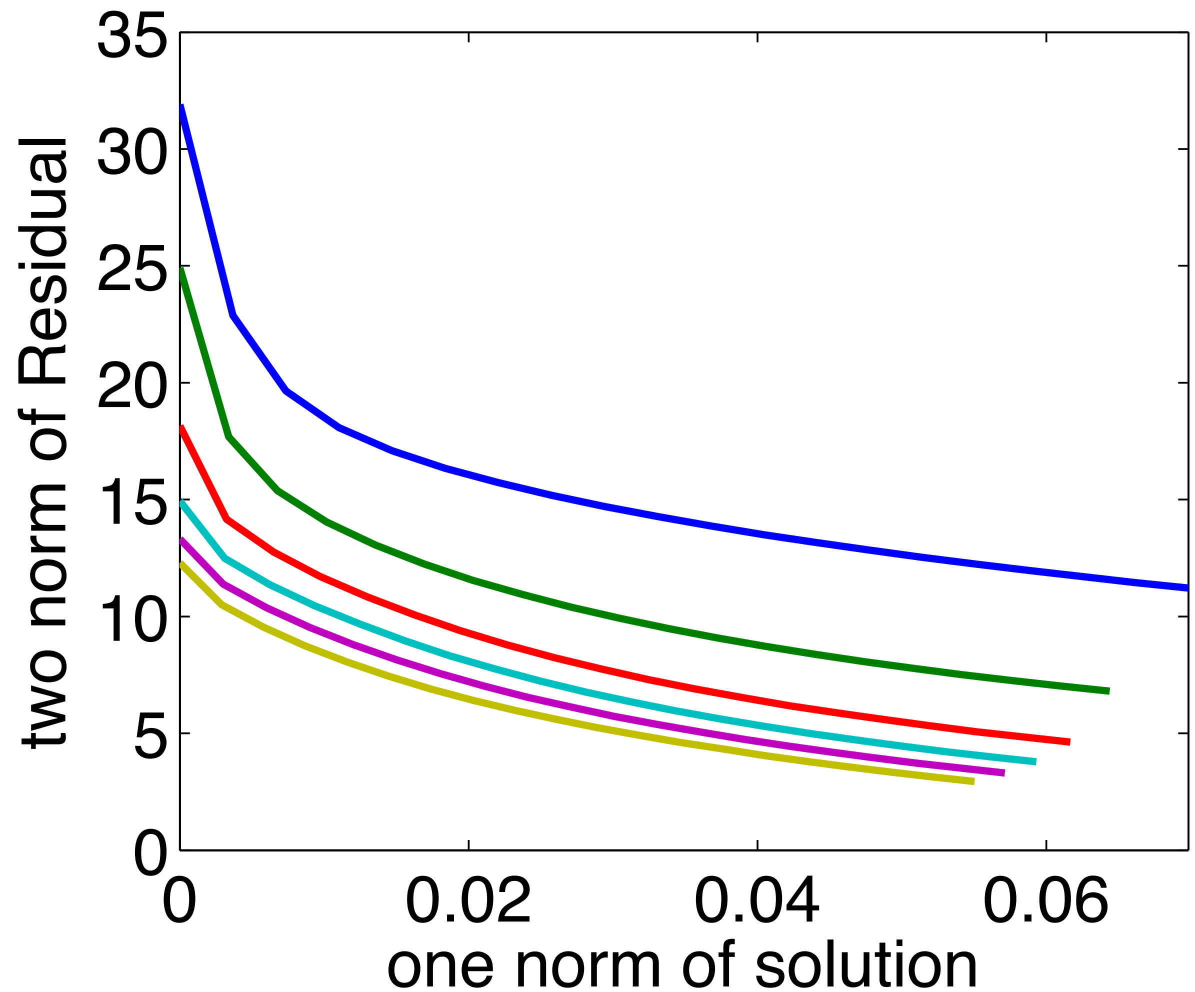


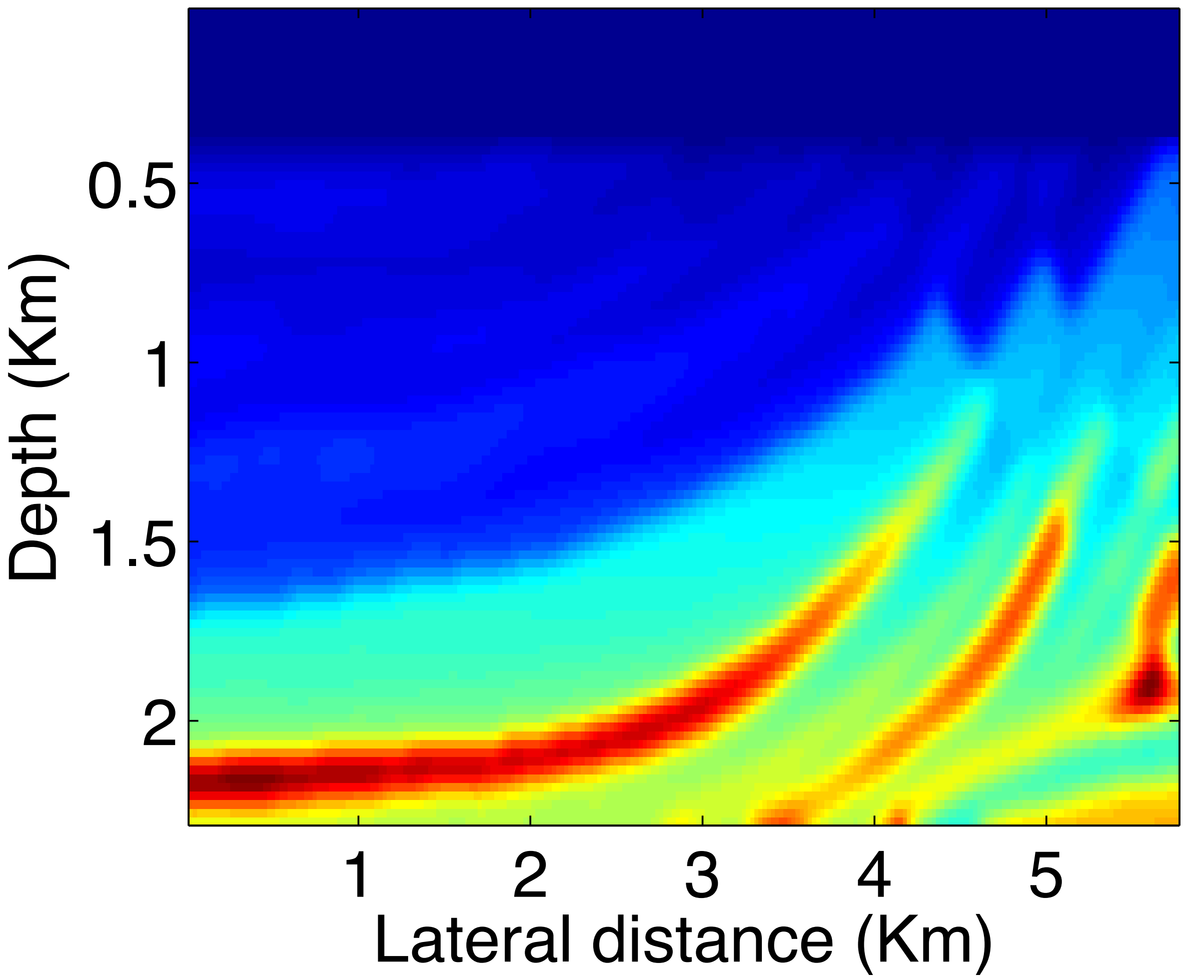
Depth (Km)
0.5
1
1.5
2
1
2
3
4
5
Lateral distance (Km)



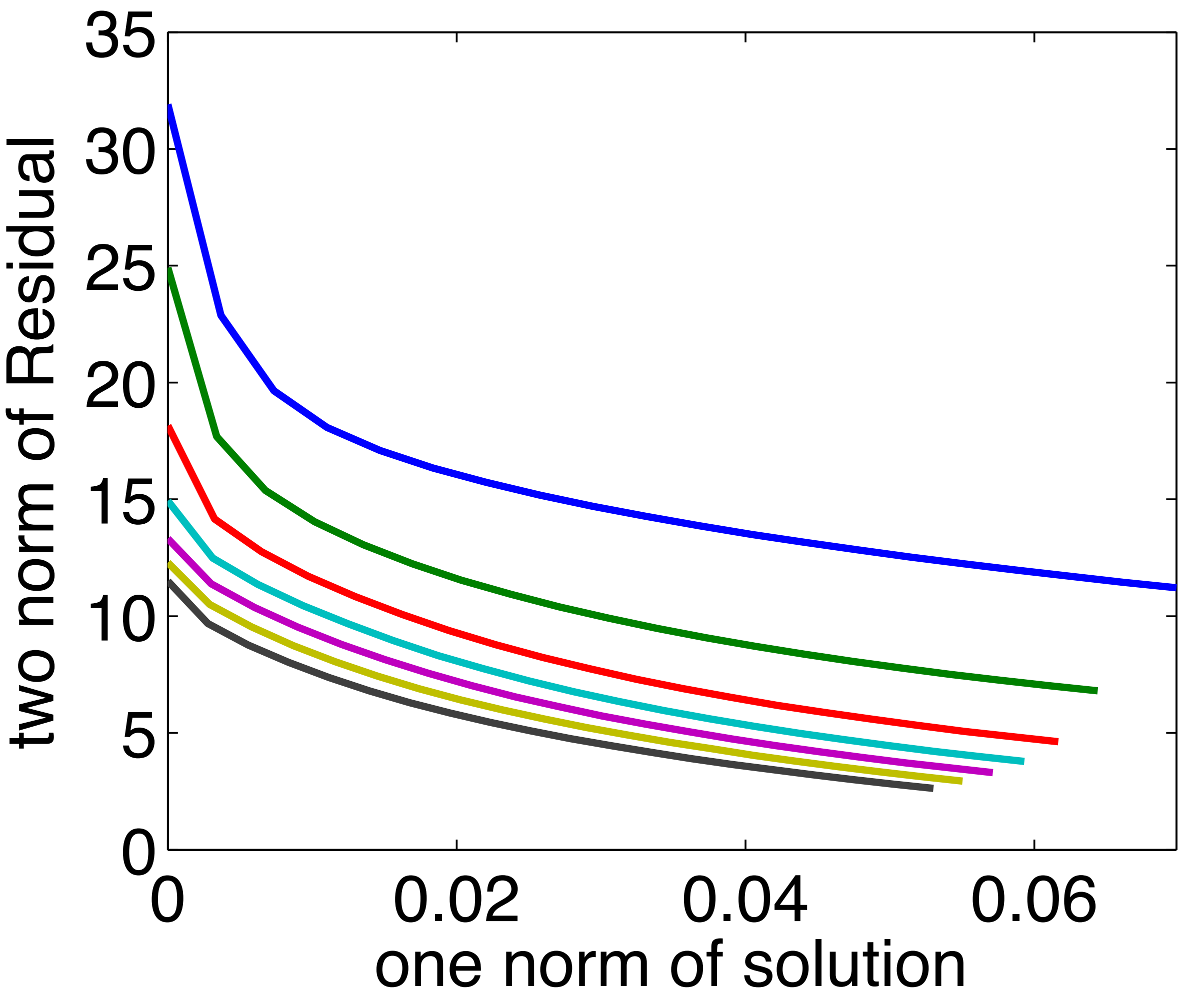
two norm of Residual
35
30
25
20
15
10
5
0
0
0.02
0.04
0.06
one norm of solution


Depth (Km)
0.5
1
1.5
2
1
2
3
4
5
Lateral distance (Km)

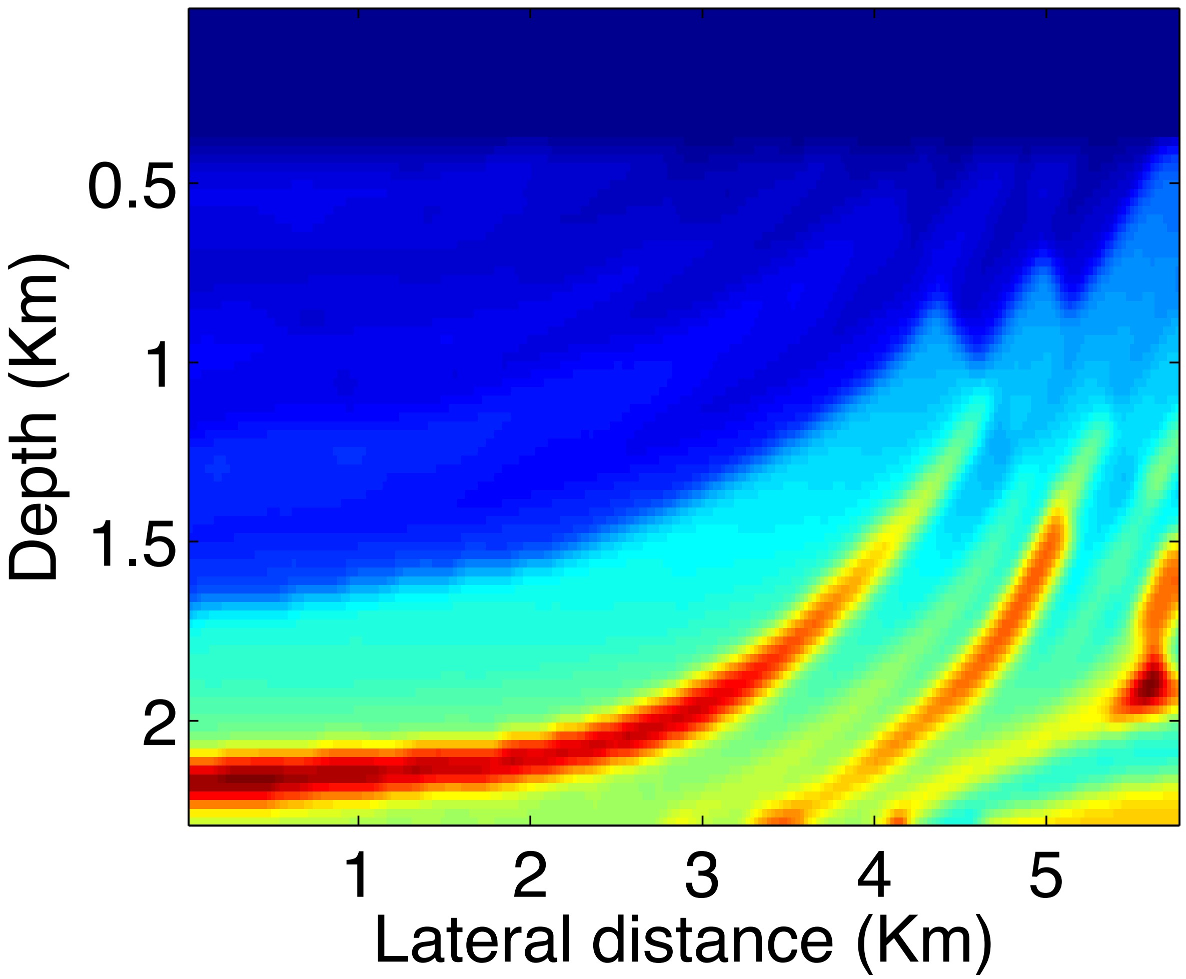

two norm of Residual
35
30
25
20
15
10
5
0
0
0.02
0.04
0.06
one norm of solution

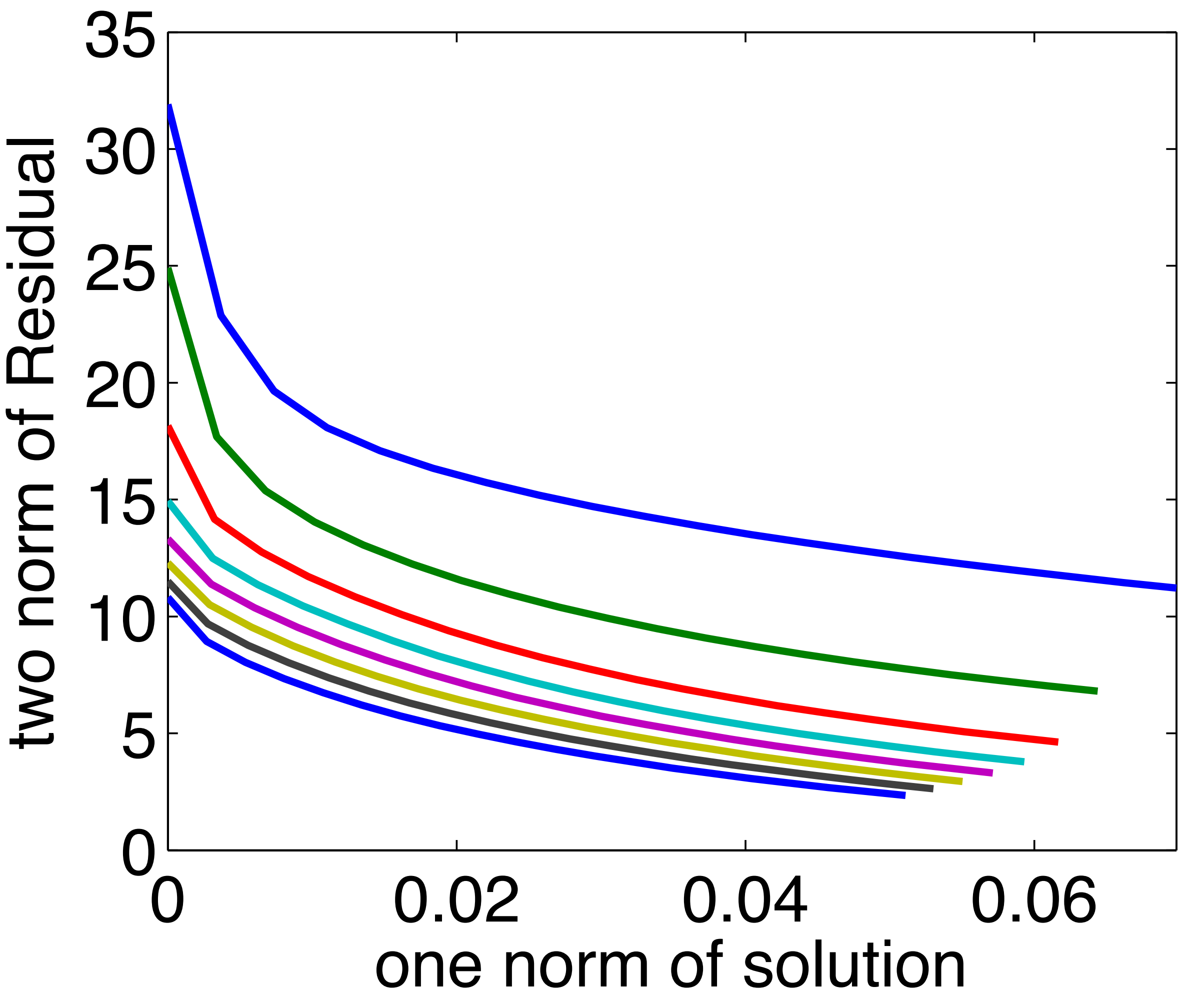


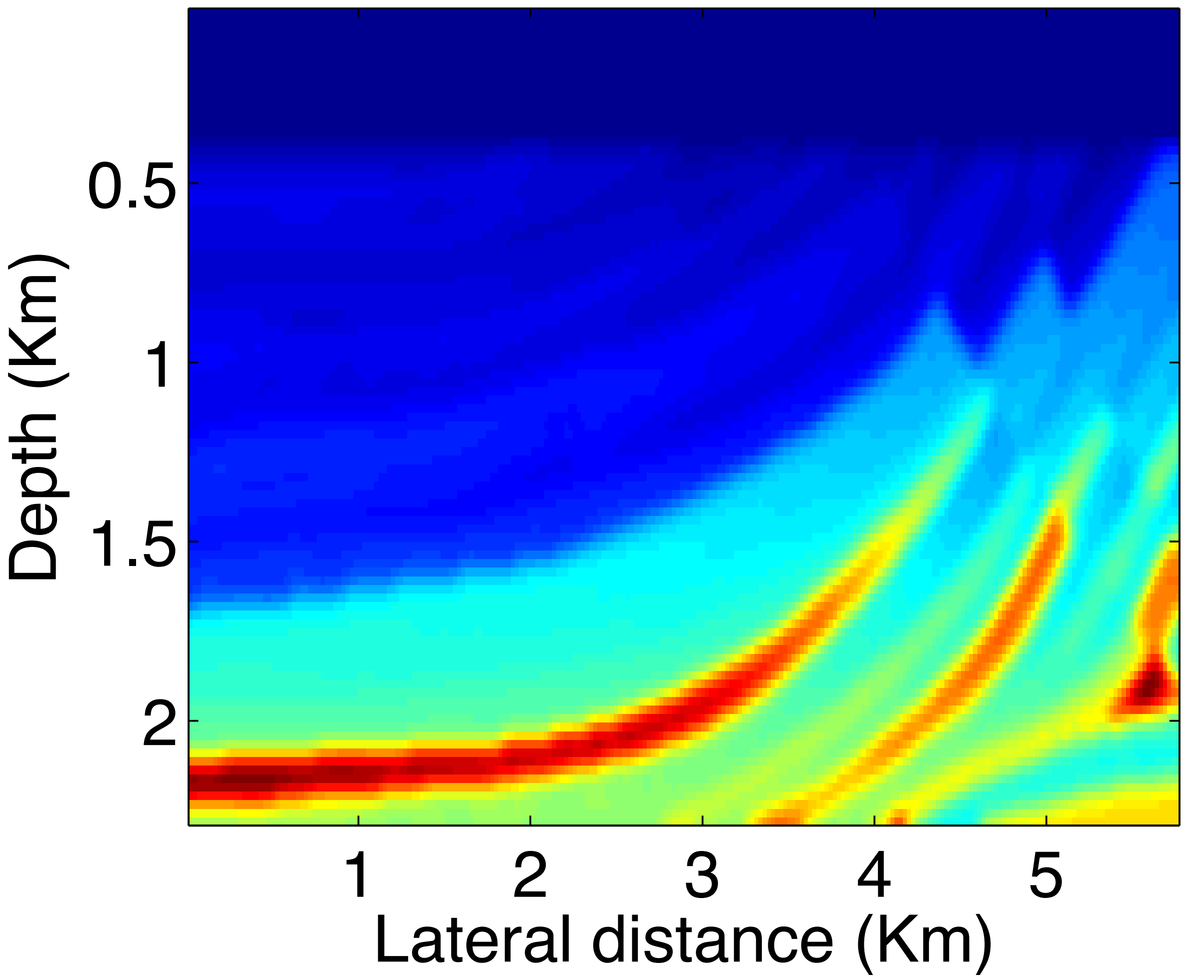
0.5
1
1.5
2
Depth (Km)
1
2
3
4
5
Lateral distance (Km)

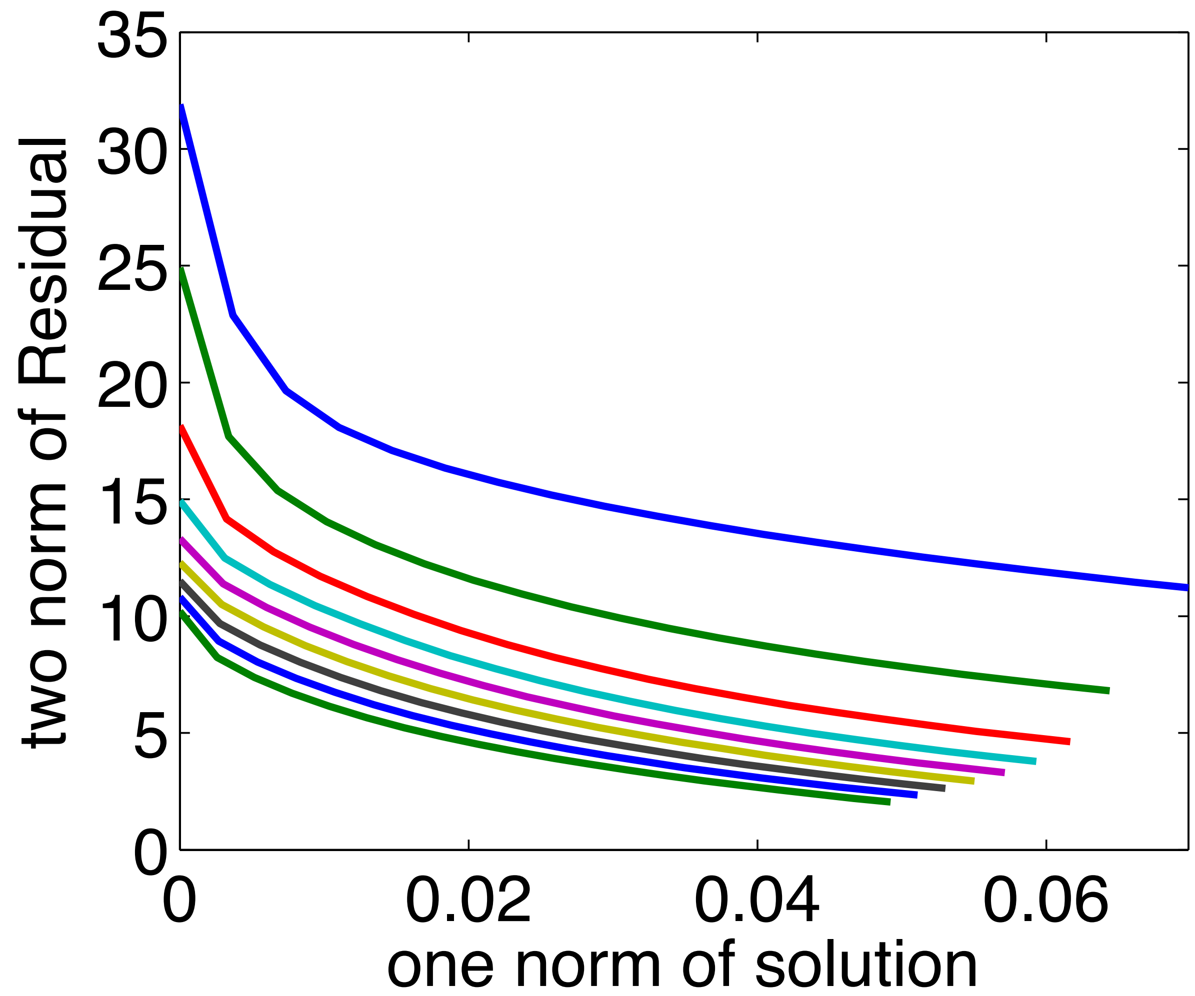


35
30
25
20
15
10
5
0
two norm of Residual
0
0.02
0.04
0.06
one norm of solution

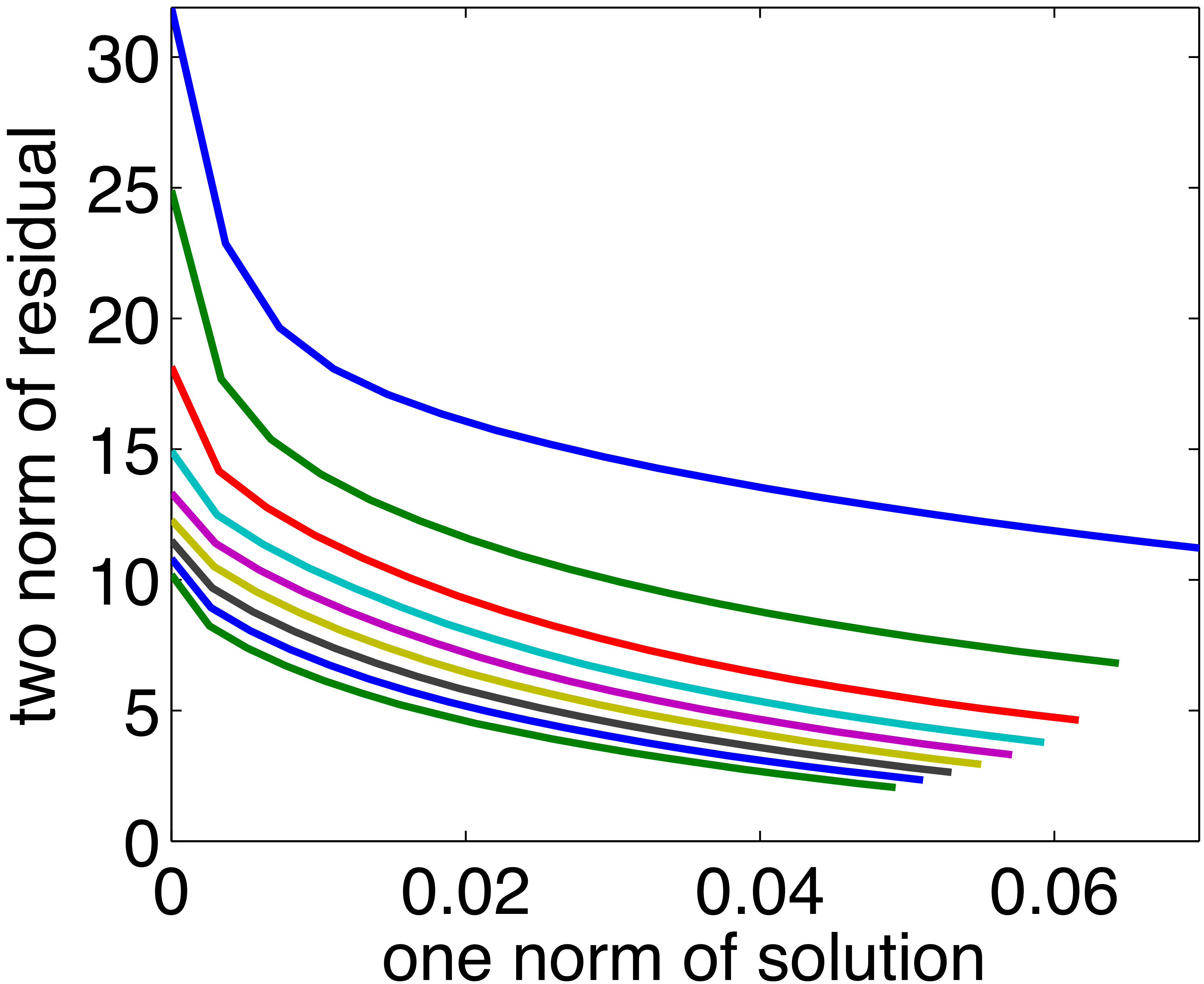

Depth (Km)
0.5
1
1.5
2
1
2
3
4
5
Lateral distance (Km)


two norm of Residual
35
30
25
20
15
10
5
0
0
0.02
0.04
0.06
one norm of solution


0.5
1
1.5
2
Depth (Km)
1
2
3
4
5
Lateral distance (Km)


0
5
10
15
20
25
30
35
two norm of Residual
0
0.02
0.04
0.06
one norm of solution

Pareto curves



Modified Gauss Newton

Instead of solving the *full* GN subproblem, we solve a series of *dimensionality*-reduced LASSO problems.

For each LASSO subproblem, we *redraw* a *new* random *subset* of (simultaneous) source experiments.

LASSO problem solved by Spectral Projected Gradient.

Jacobian & Hessian are *never* formed *explicitly*.

We use curvelets to *sparsely* represent *geophysical* model updates.

Modified Gauss-Newton

Algorithm 1: Dimensionality-reduced Gauss Newton with sparsity

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

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

while not converged **do**

$\{\underline{\mathbf{D}}^k, \underline{\mathbf{Q}}^k\} \longleftarrow \{\mathbf{DW}^k, \mathbf{QW}^k\}$ with $\mathbf{W}^k \in N(0,1)$; `// indep. draw.`

$\underline{\delta\mathbf{D}}^k \longleftarrow \underline{\mathbf{D}}^k - \boldsymbol{\mathcal{F}}[\mathbf{m}^k; \underline{\mathbf{Q}}^k]$; `// residual`

$\tau^k \longleftarrow (1-\alpha)\, \|\underline{\delta\mathbf{D}}^k\|_F \,/\, \|\nabla\boldsymbol{\mathcal{F}}^*[\mathbf{m}^k; \underline{\mathbf{Q}}^k]\underline{\delta\mathbf{D}}^k\|_\infty$; `// one-norm LASSO`

$$\delta\mathbf{x} \longleftarrow \begin{cases} \arg\min_{\delta\mathbf{x}} \frac{1}{2}\|\underline{\delta\mathbf{D}}^k - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}^k; \underline{\mathbf{Q}}^k]\delta\mathbf{x}\|_F^2 \\ \text{subject to} \quad \|\delta\mathbf{x}\|_1 \leq \tau^k \end{cases}$$

$\mathbf{m}^{k+1} \longleftarrow \mathbf{m}^k + \gamma^k \mathbf{C}^* \delta\mathbf{x}$; `// update with linesearch`

$k \longleftarrow k+1$;

end

FWI results

Time-harmonic Helmholtz:

- 205 X 701 with mesh size of 10m
- 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

FWI 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 12Hz
- recording time for each shot is 3.6s

FWI results

FWI:

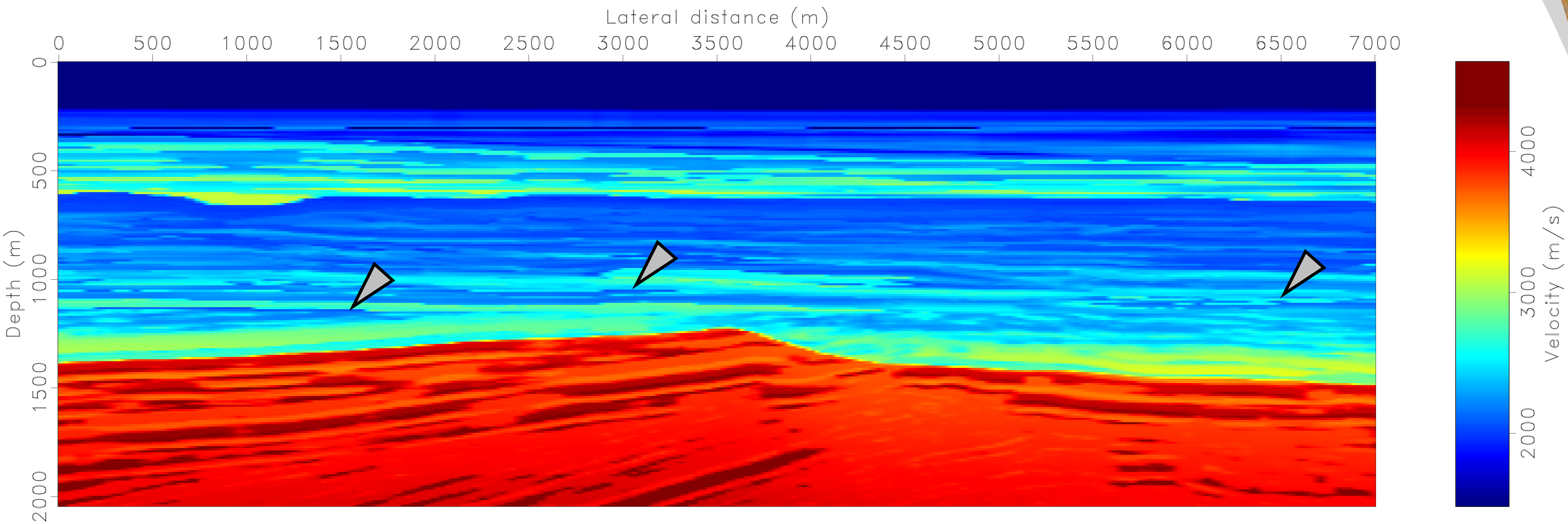
- 10 overlapping frequency bands with 10 frequencies (2.9Hz-25Hz)
- 10 Gauss-Newton steps for each frequency band (solved with max 20 spectral-projected gradient iterations)

[Bunks '95; Pratt '98]

SLIM

Results

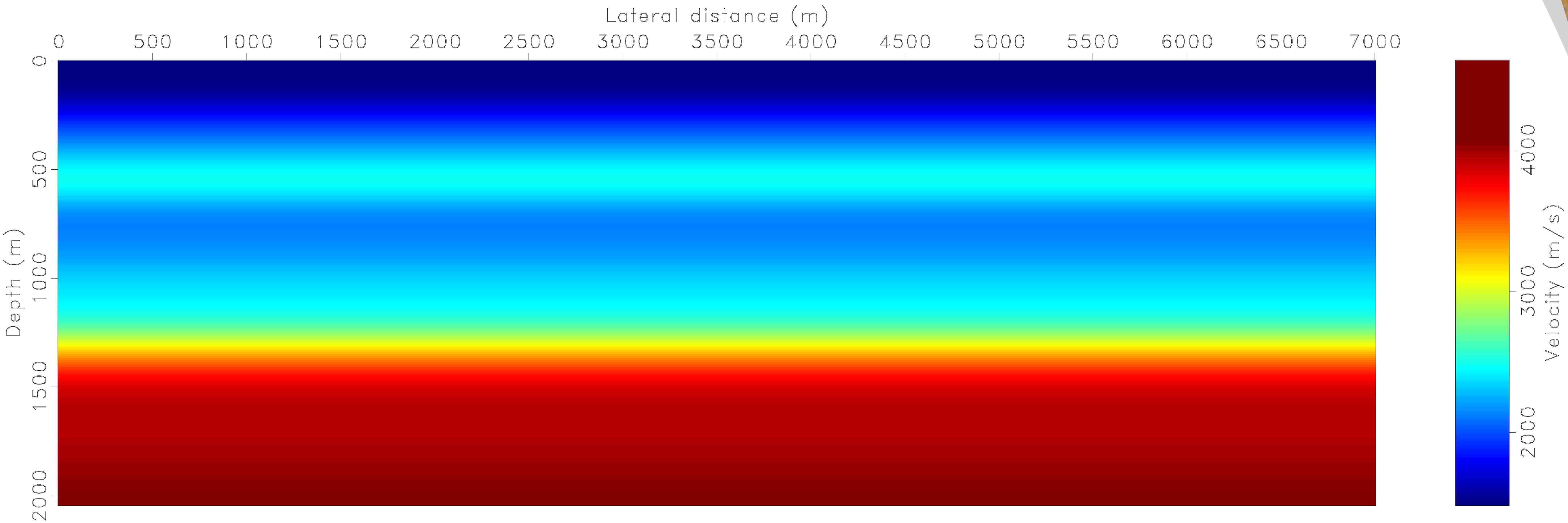
True model



SLIM

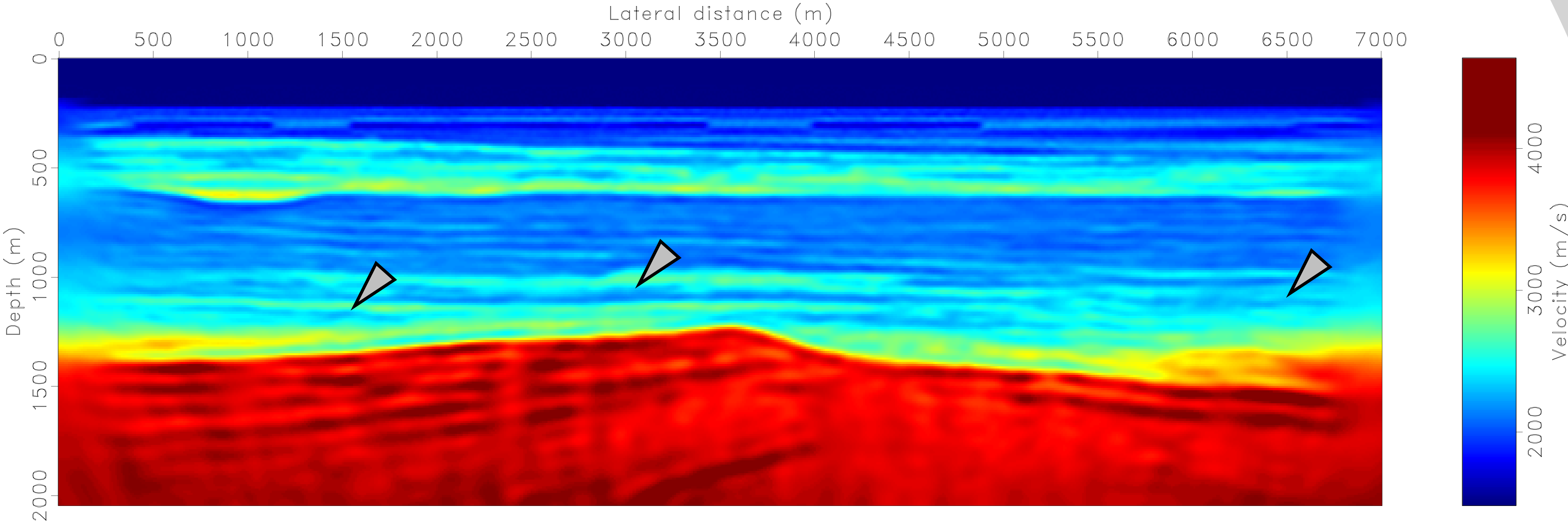
Results

Initial model



Results

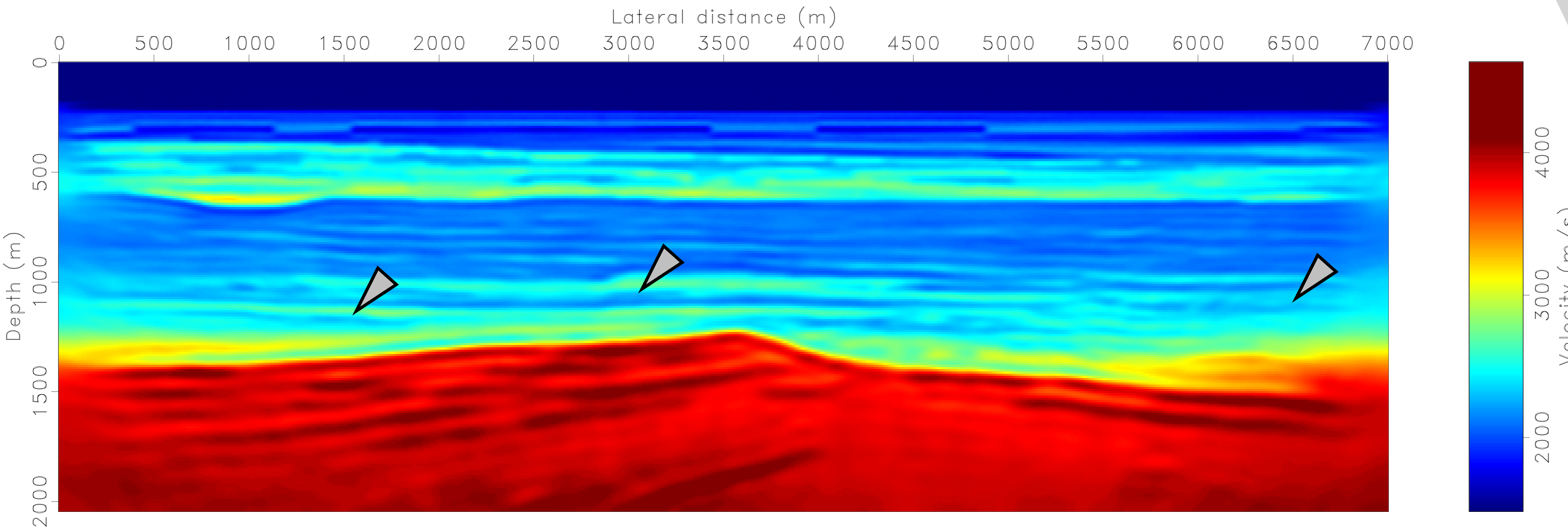
Modified GN 7 sim. shots *without renewals*



25 times speedup compared to full GN

Results

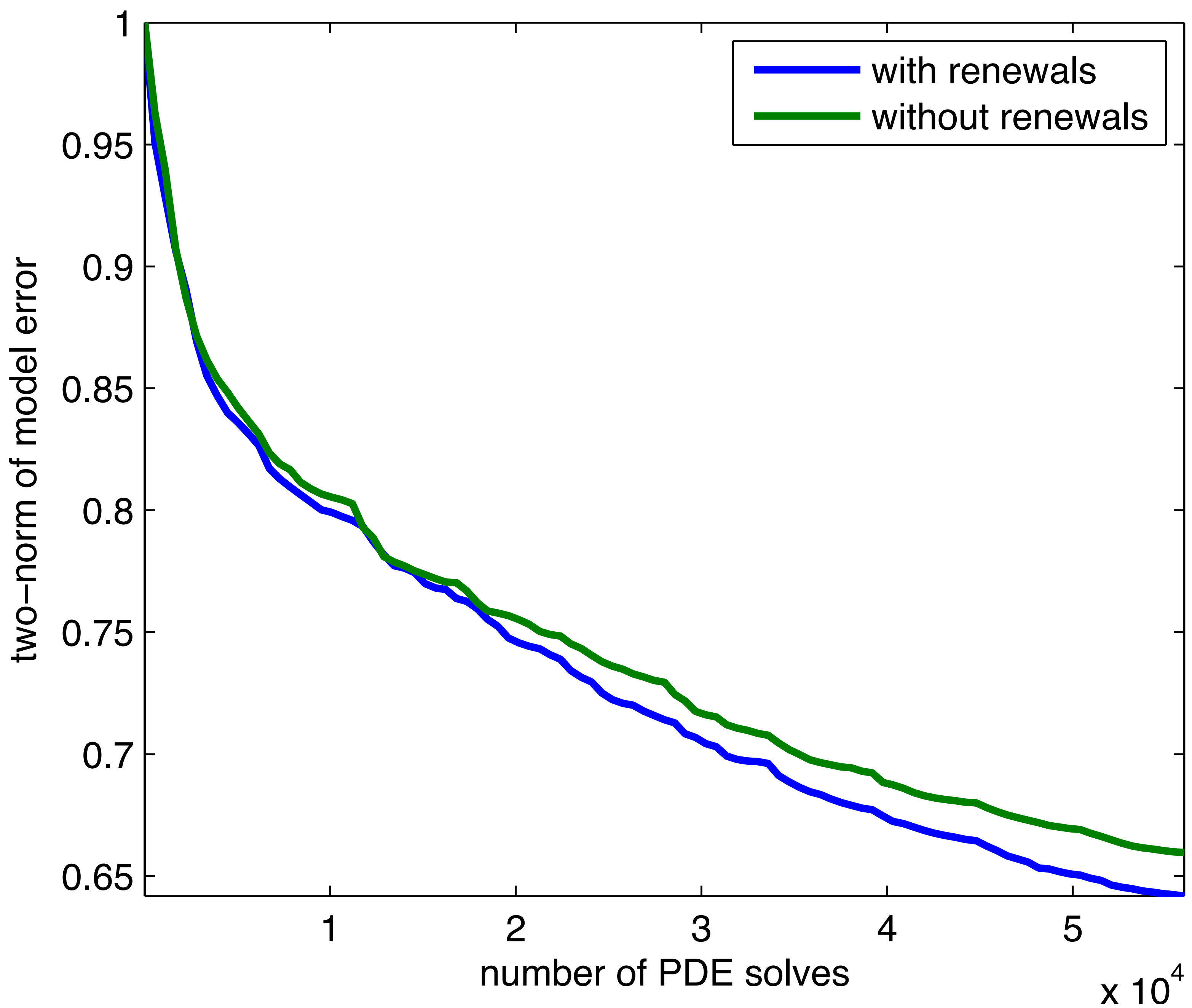
Modified GN 7 sim. shots *with renewals*



25 times speedup compared to full GN

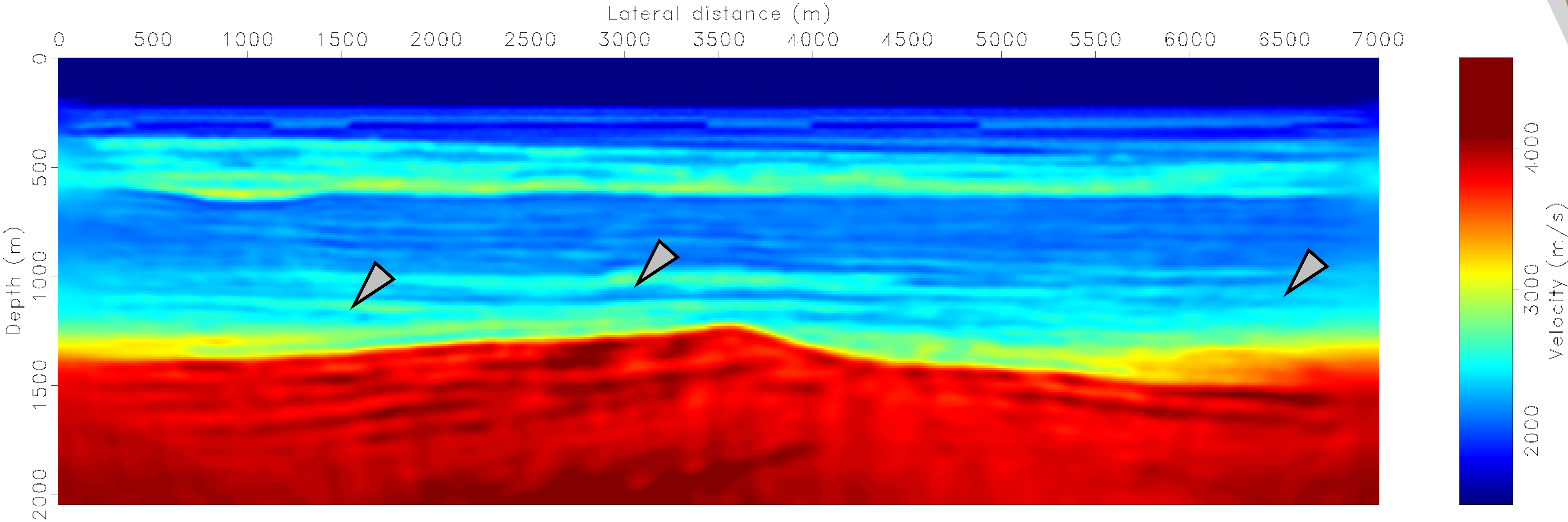
Results

Model error



Results

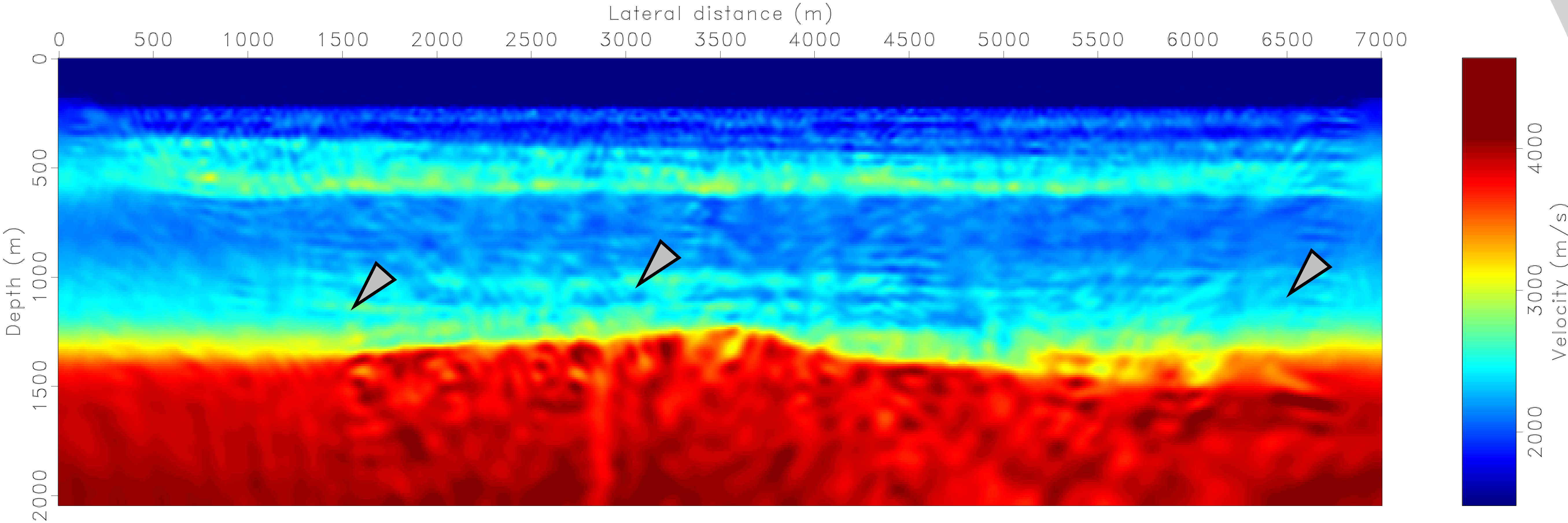
Modified GN 1 sim. shots *with renewals*



175 times speedup compared to full GN

Results

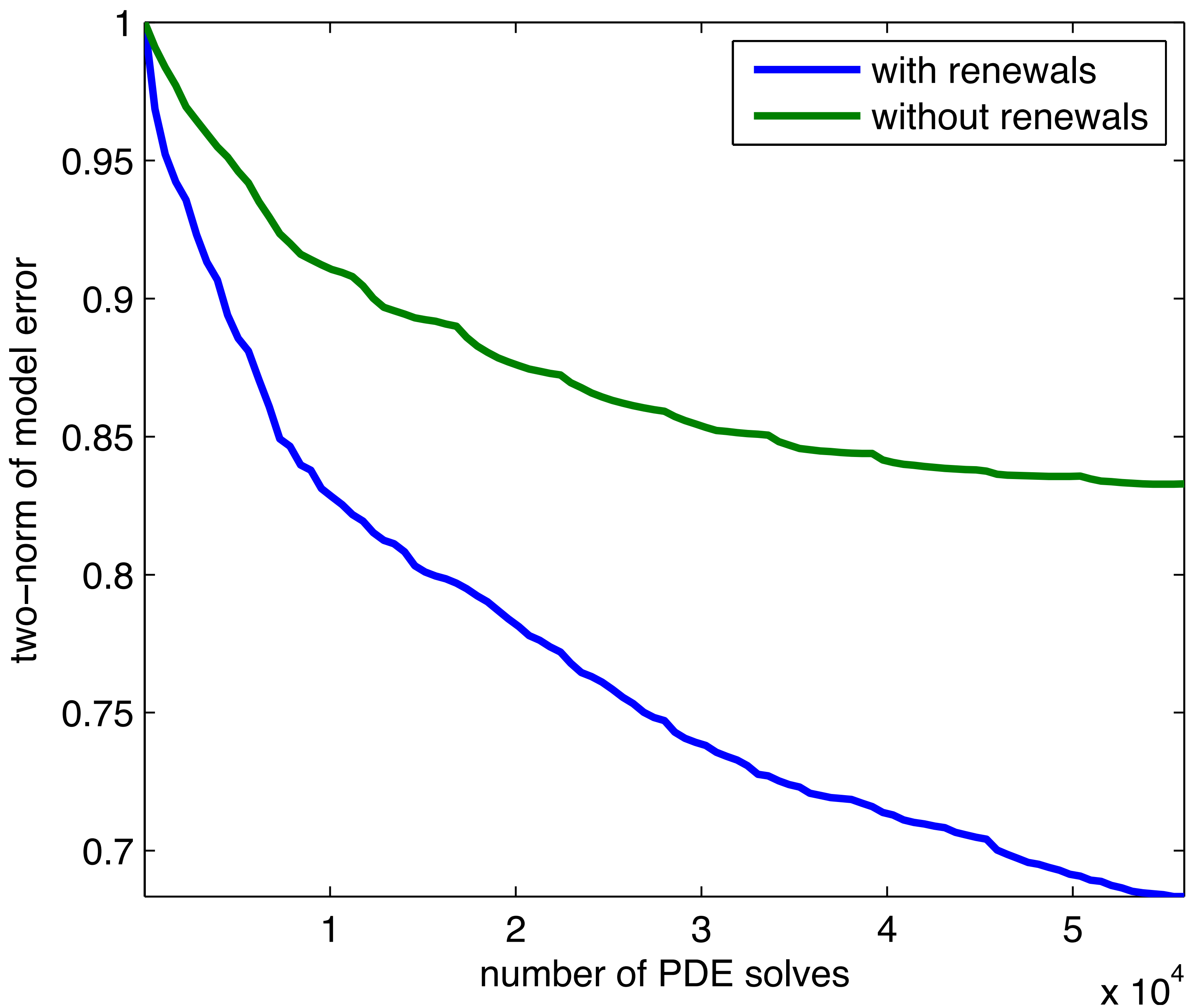
Modified GN 1 sim. shots *without renewals*



175 times speedup compared to full GN

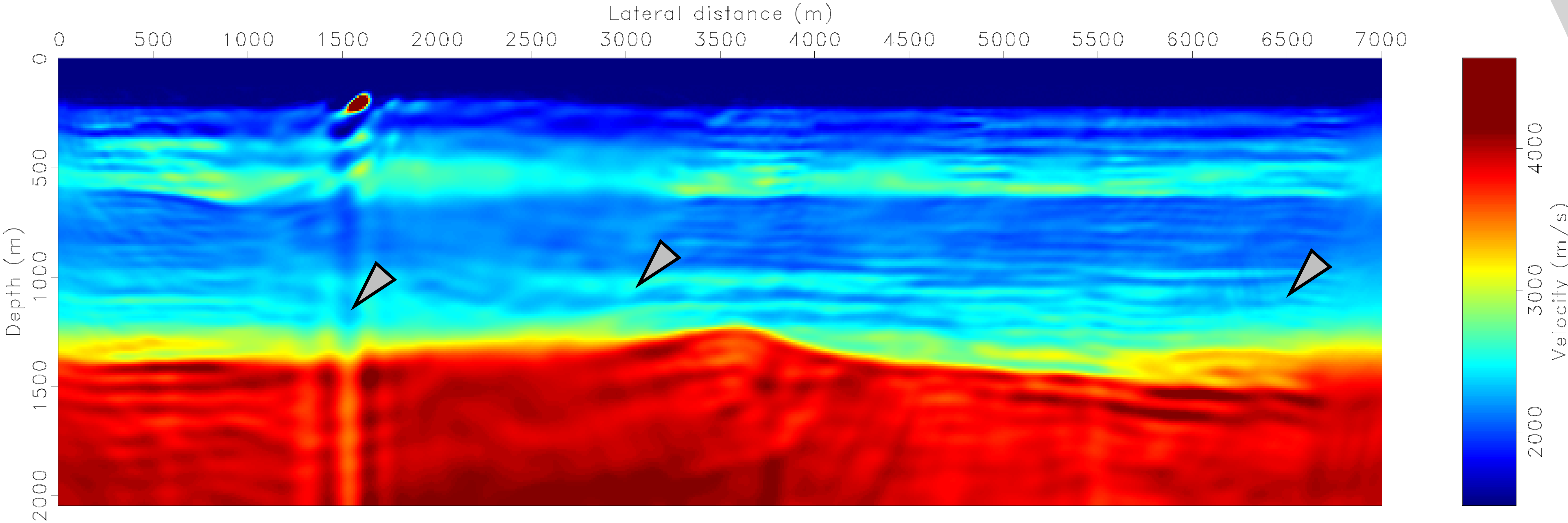
Results

Model error



Results

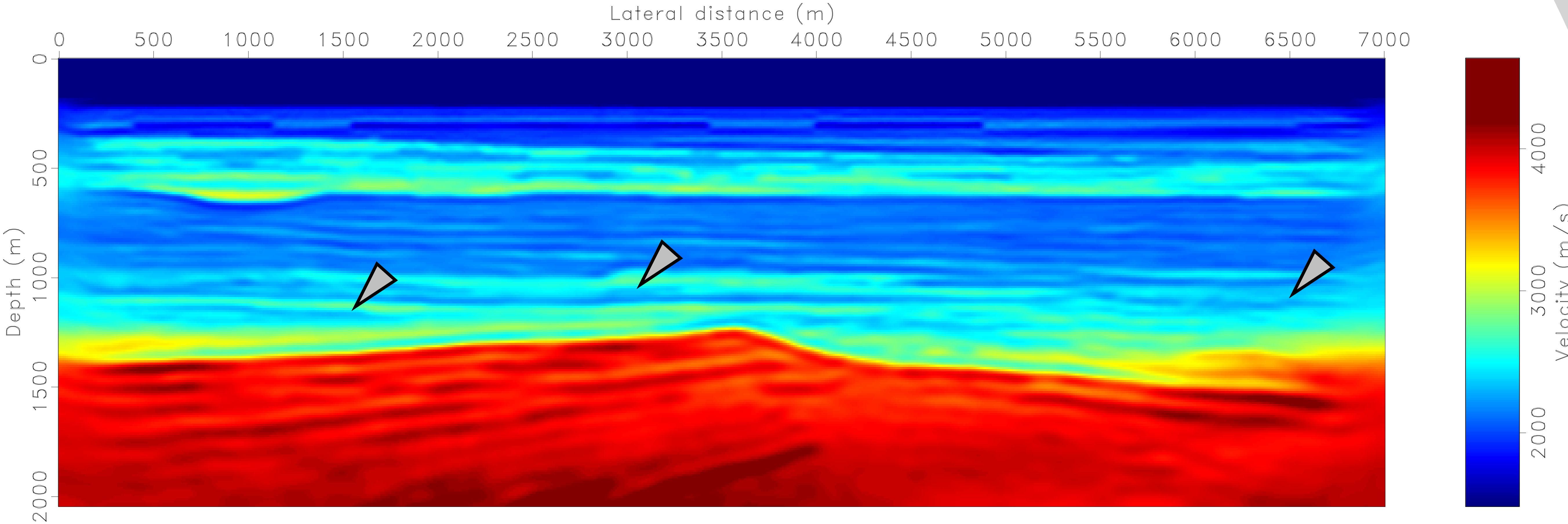
Modified GN 7 seq. shots *without renewals*



25 times speedup compared to full GN

Results

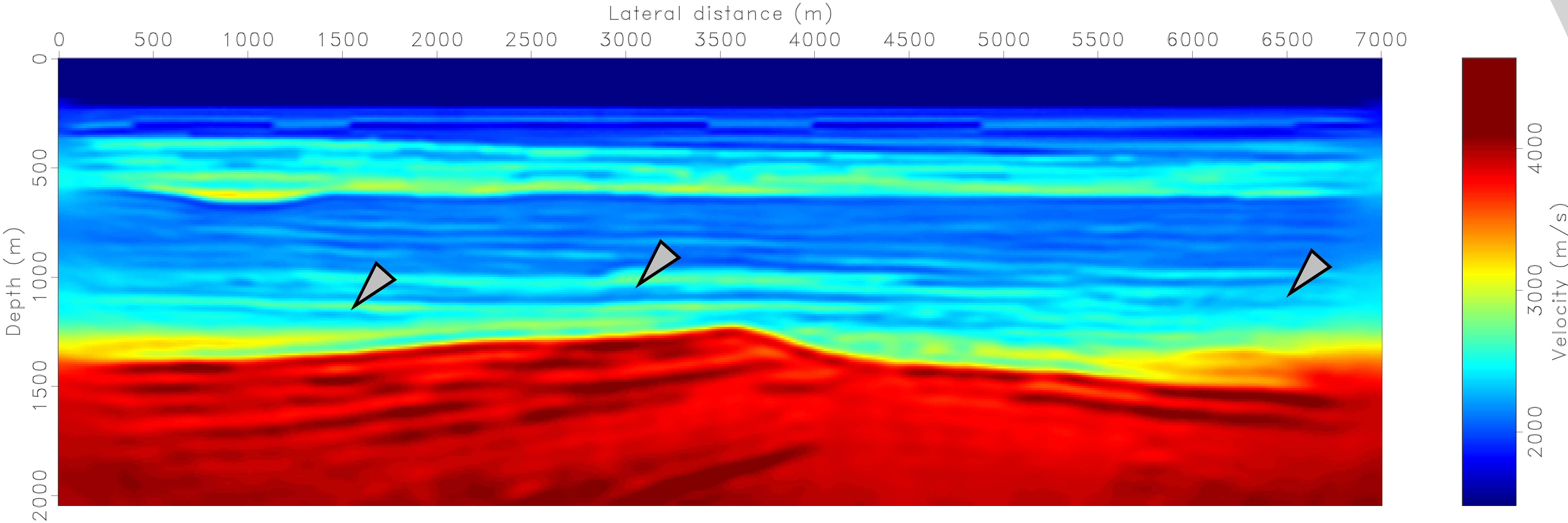
Modified GN 7 seq. shots *with renewals*



25 times speedup compared to full GN

Results

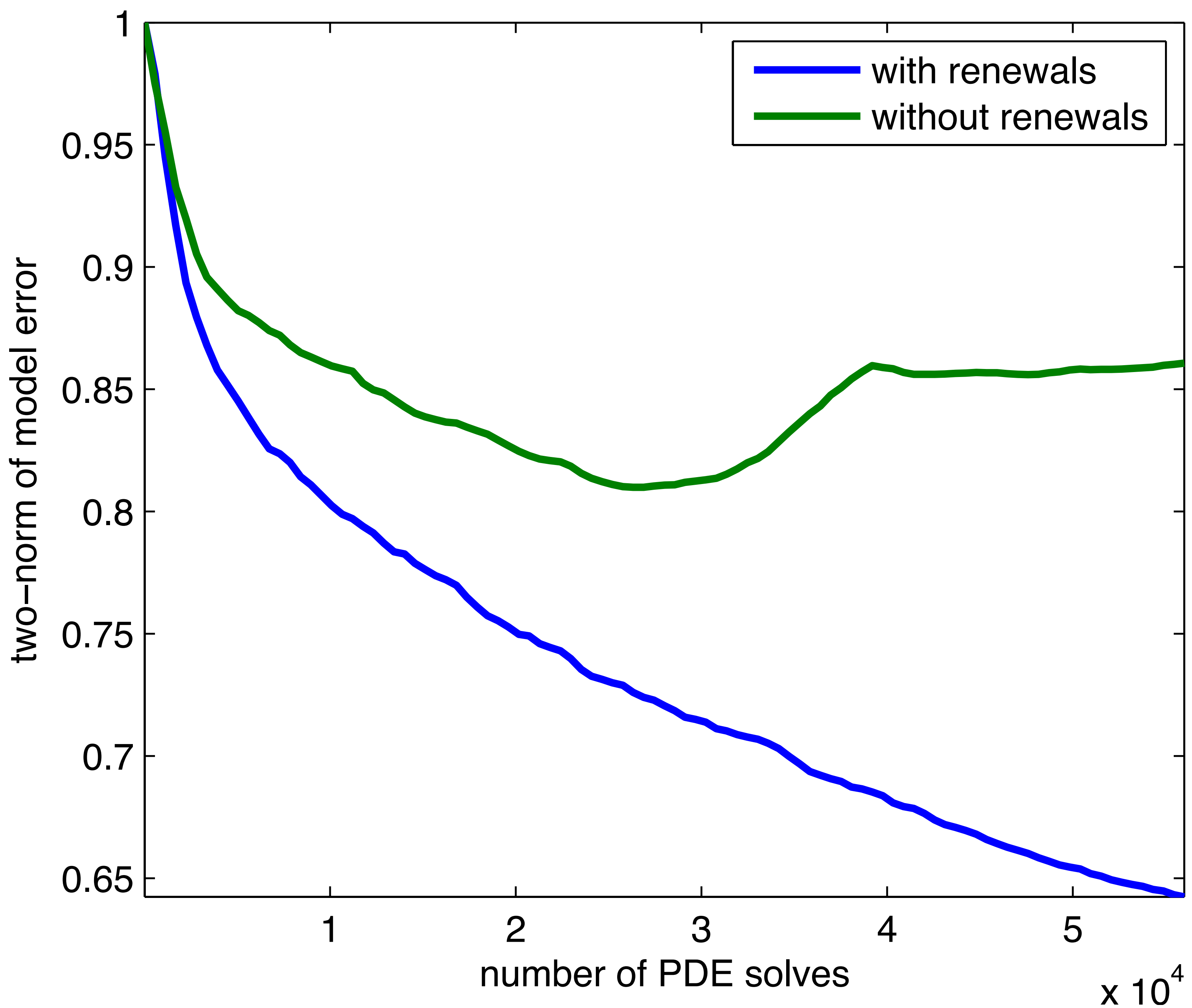
Modified GN 7 sim. shots *with renewals*



25 times speedup compared to full GN

Results

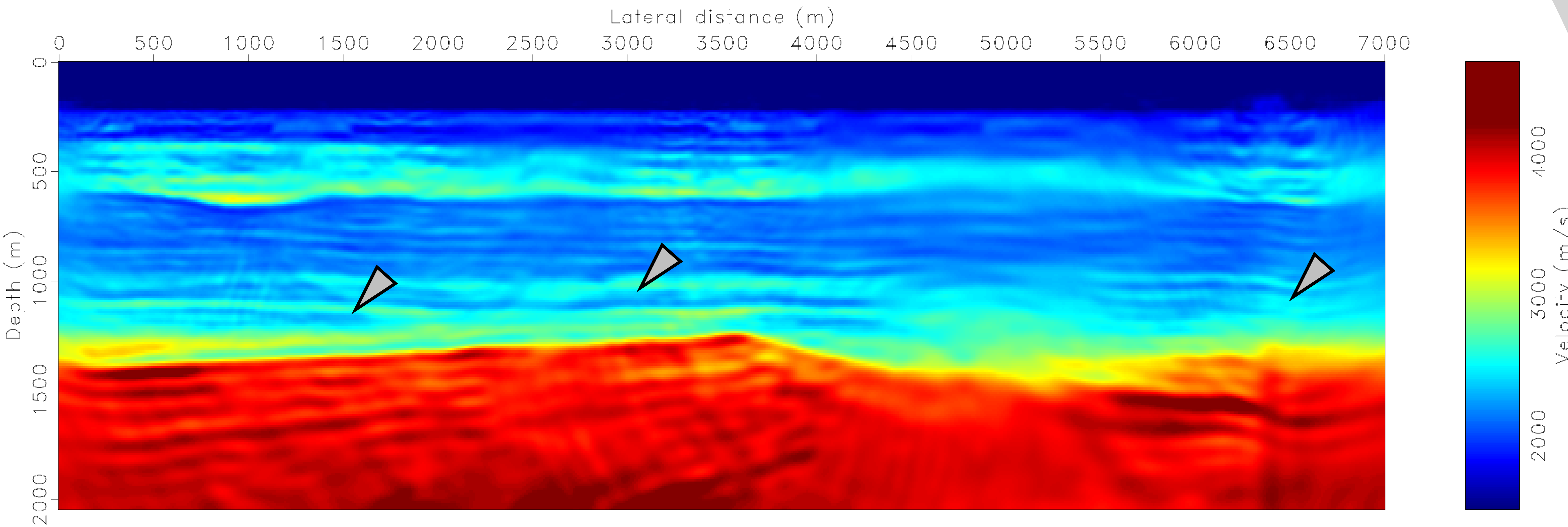
Model error



SLIM

Results

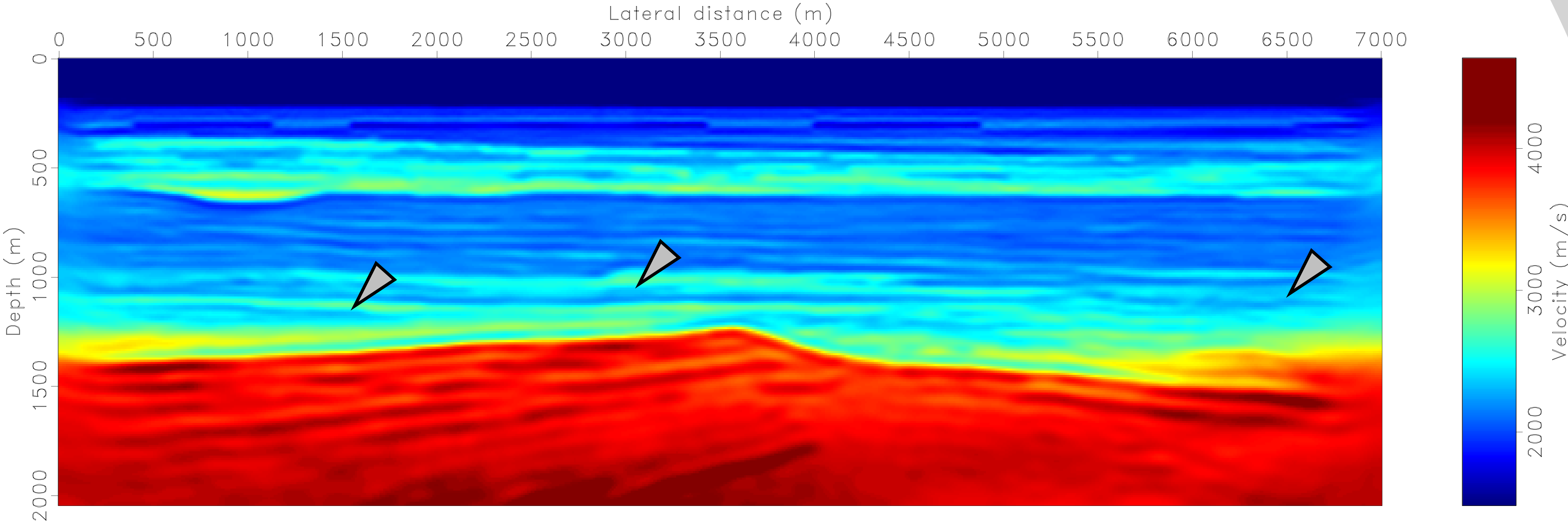
Modified GN 7 seq. shots *without renewals*



marine acquisition: min offset 100m; max offset 3km

Results

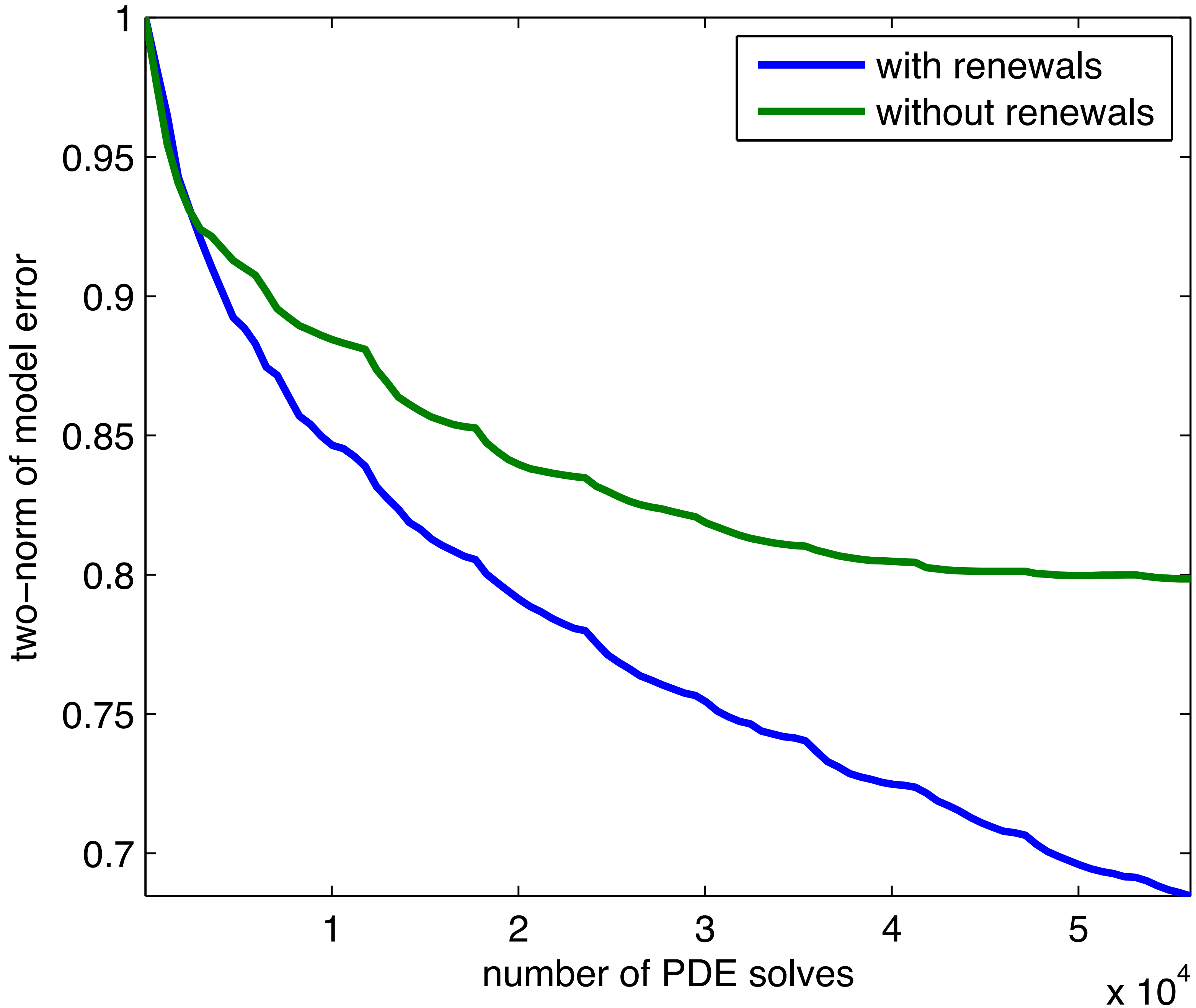
Modified GN 7 seq. shots *with renewals*



marine acquisition: min offset 100m; max offset 3km

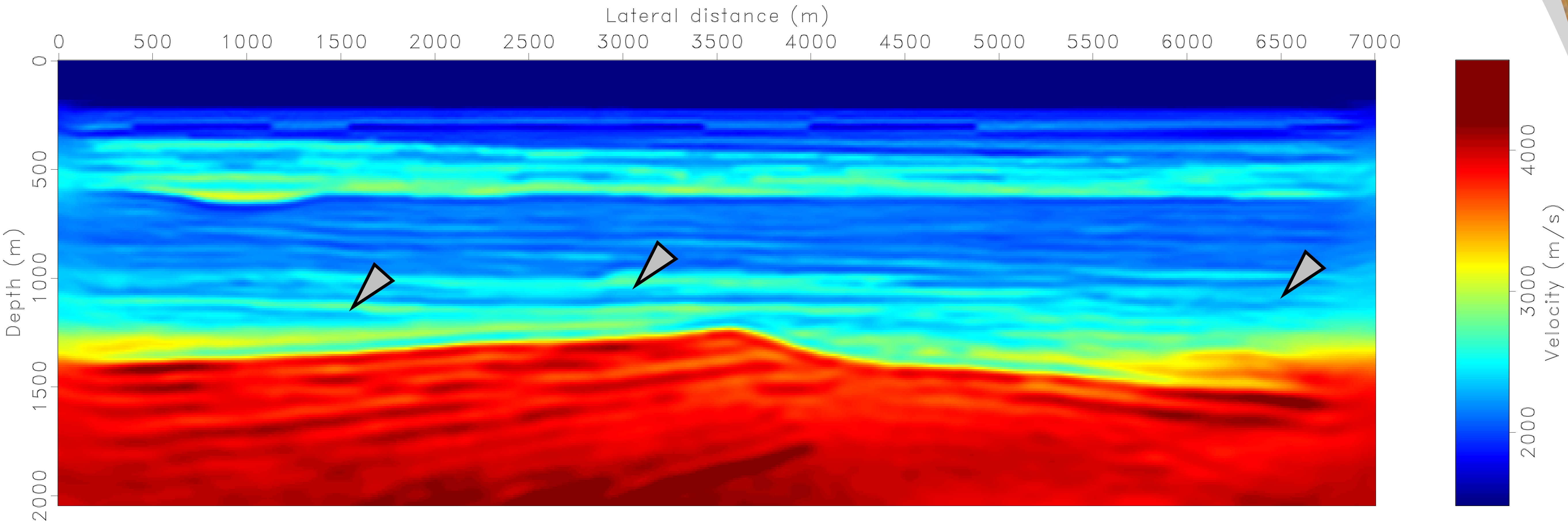
Results

Model error



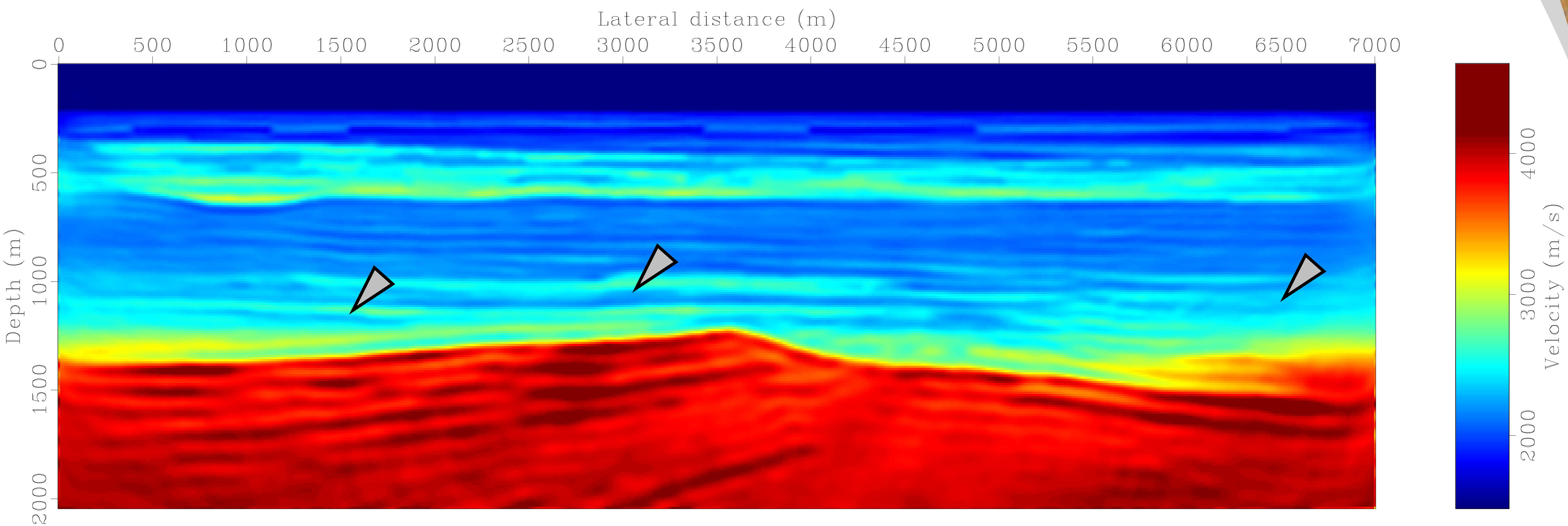
Results

Modified GN 7 seq. shots *with renewals*



Results

Batching LBFGS seq. shots [vanLeeuwen '11]



Conclusions

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

High-resolution *images* can be attainable by using *sparsity* promotion.

LASSO problems maintain fast decrease in residual while preserving *sparsity* of the *solution.*

Simultaneous shots produce *less* source cross-talk *artifacts*.

Renewals lead to significant improvements-especially for *random* subsets of *sequential* shots.

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

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