Full-Waveform Inversion with Gauss-Newton-Krylov Method

Yogi A. Erlangga and Felix J. Herrmann
{yerlangga,fherrmann}@eos.ubc.ca

Seismic Laboratory for Imaging and Modeling
The University of British Columbia (UBC), Vancouver

The 79th SEG Meeting: SI3 Methods
Houston, October 27, 2009

SLIM
Seismic Laboratory for Imaging and Modeling

Full-Waveform Inversion (FWI)

Given experiment data $\mathbf{P}$.

With the misfit functional: $E[\mathbf{m}] = \frac{1}{2}\|\mathbf{P} - \mathbf{F}[\mathbf{m}]\|_2^2$

Optimization Problem: Find

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}\in\mathcal{M}} E[\mathbf{m}] \quad \text{subject to} \quad \mathbf{F}[\mathbf{m}] = \mathbf{D}\mathbf{U}[\mathbf{m}]$$

the (forward) modeling wavefields $\mathbf{U}$ restricted to the receivers by $\mathbf{D}$.

- Lailly, 1983
- Tarantola, 1984, 1986, 1987
- Pratt and co-authors, 1996, 1998, 1999, 2003

Frequency domain FWI

Forward model: Helmholtz equation

$$\mathbf{H}[\omega, \mathbf{m}]\mathbf{U} = \mathbf{Q}, \quad \mathbf{m} = (m_1 \ \ldots \ m_M)^T$$

- $\mathbf{H}$: the Helmholtz matrix, function of angular freq ω
- $\mathbf{Q} = [\mathbf{q}_1 \ \ldots \ \mathbf{q}_{n_s}]$: the source matrix, with n_s shots
- $\mathbf{U} = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_{n_s}]$: the wavefield matrix

Impediments

- Fast, scalable solver for the forward and adjoint systems
  - iterative method with
    - Preconditioning with shifted Laplacian [E. et al. (2006), Riyanti et al., (2006)]
    - Multilevel Krylov method [E. & Nabben (2009), E. & Herrmann (2008)]

- Multidimensional experiments (shots, frequencies): more data than model
  - Data reduction via frequency subsampling [Sirgue & Pratt (2004), Mulder & Plessix (2004)]
  - Compressive Sampling (CS) framework : data reduction via *shot* and *frequency* subsampling
    - compressive wavefield computation [Lin, Herrmann (2007), Herrmann, E. & Lin (2009)]
    - extension to compressive imaging

- Fast minimization solver (GN-type: Hessian)
  - Gauss-Newton method with implicit computation of Hessian

Our solution

- Gauss-Newton with implicit Hessian (Gauss-Newton-Krylov, GNK)

- Dimensionality reduction
  - [Herrmann, E. & Lin (2009)]
  - [Tim Lin: Compressive simultaneous full-waveform simulation, this meeting, SM1]

$$\begin{cases} \mathbf{Q} = \mathbf{D}^* \underbrace{\mathbf{s}}_{\text{single shots}} \\ \mathbf{HU} = \mathbf{Q} \\ \mathbf{y} = \mathbf{RMDU} \end{cases} \iff \begin{cases} \underline{\mathbf{Q}} = \underline{\mathbf{D}}^* \underbrace{\mathbf{RMs}}_{\text{simul. shots}} \\ \underline{\mathbf{HU}} = \underline{\mathbf{Q}} \\ \underline{\mathbf{y}} = \underline{\mathbf{DU}} \end{cases}$$

- FWI with CS

SLIM Seismic Laboratory for Imaging and Modeling

FWI with CS

The misfit functional: $$\underline{E}[\mathbf{m}] = \frac{1}{2}\|\mathbf{RM}(\mathbf{P} - \mathbf{F}[\mathbf{m}])\|_2^2$$

with $\mathbf{RM}$ a CS-sampling matrix (reduces data size).

Optimization Problem: Find

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}\in\mathcal{M}} \underline{E}[\mathbf{m}] \quad \text{subject to} \quad \mathbf{F}[\mathbf{m}] = \mathbf{DU}[\mathbf{m}]$$

Main contribution:

[Hermann et al. (2009), EAGE]

See also: Krebs et al. (2009), this meeting

$$\underline{E}[\mathbf{m}] = \frac{1}{2}\|\underline{\mathbf{P}} - \underline{\mathbf{F}}[\mathbf{m}]\|_2^2$$

In line with this:

Sampling of overdetermined systems [Drineas, Mahoney & Muthukhrisnan (2006)]

- $\min \underline{E} \neq \min E$ but is a bounded approximation.

SLIM

Seismic Laboratory for Imaging and Modeling

Outline

- Newton method: Hessian
- Implicit computation of the GN Hessian
- Extension to CS framework
  - Reduced numbers of shots and frequencies
- Examples

- related work: in time domain [Akcelic, Biros & Ghattas (2002)]
- PDE-constrained optimization: KKT sytems, reduced systems, etc [Heinkenschloss (1991), Biros & Ghattas (2005), ...]

Newton Method

$$E[\mathbf{m} + \delta\mathbf{m}] = E[\mathbf{m}] + \mathbf{g}^T\delta\mathbf{m} + \frac{1}{2}\delta\mathbf{m}^T\mathcal{H}\delta\mathbf{m}$$

Initial model $\mathbf{m}_0$;

Update until convergence:

$$\delta\mathbf{m} = -\mathcal{H}_{k-1}^{-1}\mathbf{g}_{k-1};$$
$$\mathbf{m}_k = \mathbf{m}_{k-1} + \gamma_{k-1}\delta\mathbf{m};$$

with

- $\mathbf{g}_{k-1} \equiv \mathbf{g}[\mathbf{m}_{k-1}]$: the gradient,
- $\mathcal{H}_{k-1} \equiv \mathcal{H}[\mathbf{m}_{k-1}]$: the Hessian,
- γ_{k-1} : the step length.

SLIM Seismic Laboratory for Imaging and Modeling

Hessian: $\mathcal{H} = [h_{i,j}]$ with

$$\begin{aligned} h_{i,j} &= \frac{\partial}{\partial m_i}\left(\frac{\partial E}{\partial m_j}\right) \\ &= \mathrm{rowsum}\left(\frac{\partial^2 \mathbf{U}}{\partial m_i \partial m_j} - \frac{\partial \mathbf{U}}{\partial m_i}\frac{\partial \mathbf{F}}{\partial m_j}\right). \end{aligned}$$

- Negative sign: not necessarily SP(S)D
- Fast/quadratic convergence only if close to the minimizer
- From the adjoint system: $\dfrac{\partial \mathbf{U}}{\partial m_i} \longrightarrow \mathbf{V}$ (back-propagated)

SLIM Seismic Laboratory for Imaging and Modeling

Gauss-Newton Method

- Simplify the Hessian by setting $\frac{\partial \mathbf{U}}{\partial m_i}\frac{\partial \mathbf{F}}{\partial m_j} = 0$
  - nonlinear wave phenomena (e.g. multiples)
- Giving

$$h_{i,j}^{GN} = \text{rowsum}\left(\frac{\partial^2 \mathbf{U}}{\partial m_i \partial m_j}\right).$$

- This is associated with setting the back-propagated wavefield $\mathbf{V} = 0$ in the Hessian
- $\mathcal{H}^{GN} = [h_{i,j}^{GN}]$ is SP(S)D.

SLIM
Seismic Laboratory for Imaging and Modeling

Inverting the Hessian: Krylov

$$\mathcal{H}^{GN}_{k-1}\delta\mathbf{m} = -\mathbf{g}_{k-1}$$

SP(S)D Hessian: compute $\delta\mathbf{m}$ with Conjugate Gradient (CG).

Four important steps in CG:

- compute: $\mathbf{w} := \mathcal{H}^{GN}_{k-1}\mathbf{p}$
- solution update: $\delta\mathbf{m} \leftarrow \delta\mathbf{m} + \alpha\mathbf{p}$
- residual update: $\mathbf{r} \leftarrow \mathbf{r} - \alpha\mathbf{w}$
- search dir. update: $\mathbf{p} \leftarrow \mathbf{r} + \beta\mathbf{w}$

α, β :
CG step lengths, satisfying orthogonal projection

$\mathbf{p} \sim \delta^2\mathbf{m}$: second variation (of the Lagrangian) of $\mathbf{m}$.

Second variation system of GN

- derived from second variations of the Lagrange minimization functional
- detailed treatment in weak (bilinear) formulation, see the abstract.

Forward model: $\mathbf{H}[\mathbf{m}]\widetilde{\mathbf{U}} = -\omega^2 \mathrm{diag}(\mathbf{p})\mathbf{U}$

- $\widetilde{\mathbf{U}}$: second variation of $\mathbf{U}$

Adjoint system/back propagation: $\mathbf{H}[\mathbf{m}]\widetilde{\mathbf{V}} = \mathbf{D}^*\mathbf{D}\widetilde{\mathbf{U}}$

- $\widetilde{\mathbf{V}}$: the second variation of $\mathbf{V}$

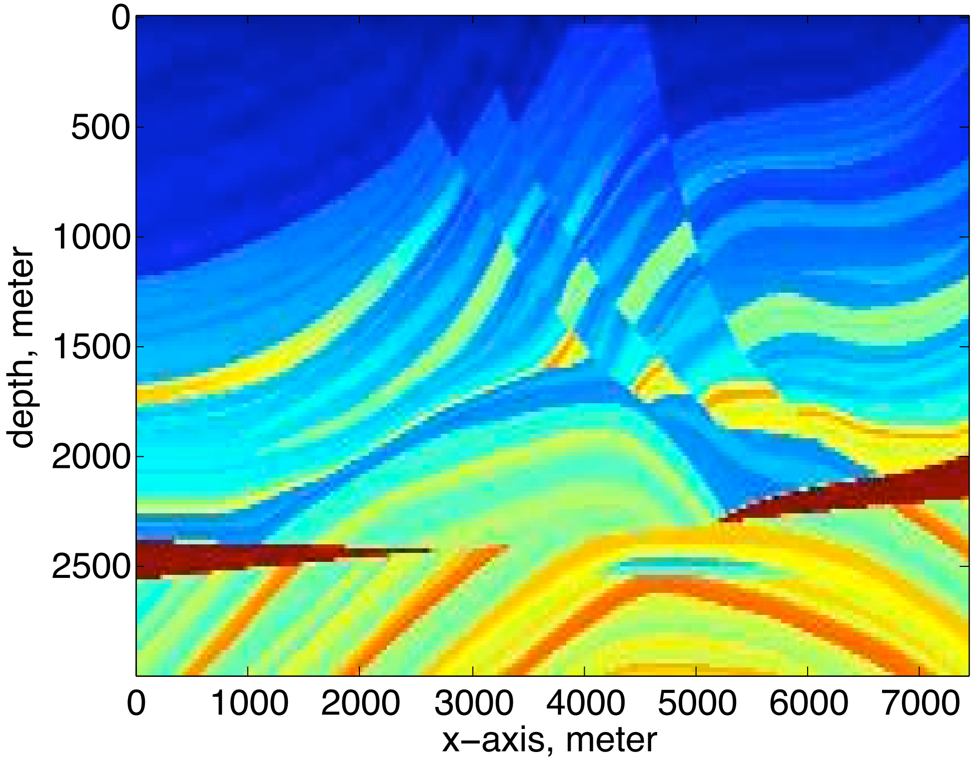
The action of Hessian on $\mathbf{p}$

$$\mathcal{H}^{GN}\mathbf{p} := -\omega^2 \mathrm{rowsum}(\mathbf{U} \odot \widetilde{\mathbf{V}})$$

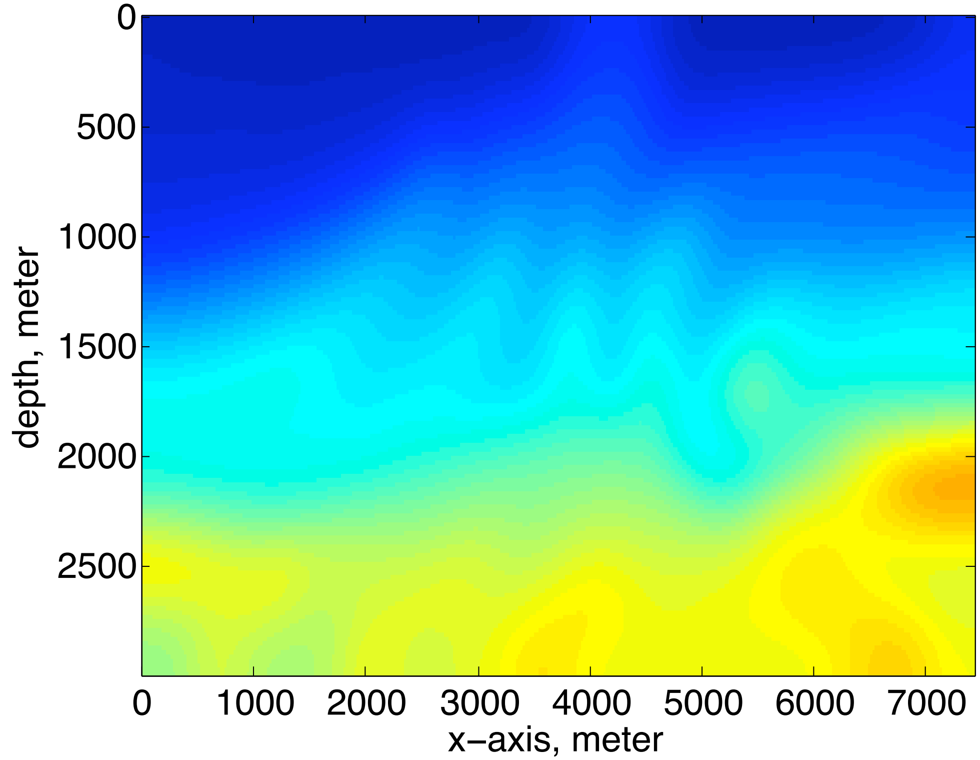
SLIM
Seismic Laboratory for Imaging and Modeling

FWI: Examples

Marmousi model: 7420 x 2980 m, 372 x 150 gridpoints, 370 shots. Freqs: 3, 5, 9 Hz. 10 CG iters for the Hessian.



Hard model

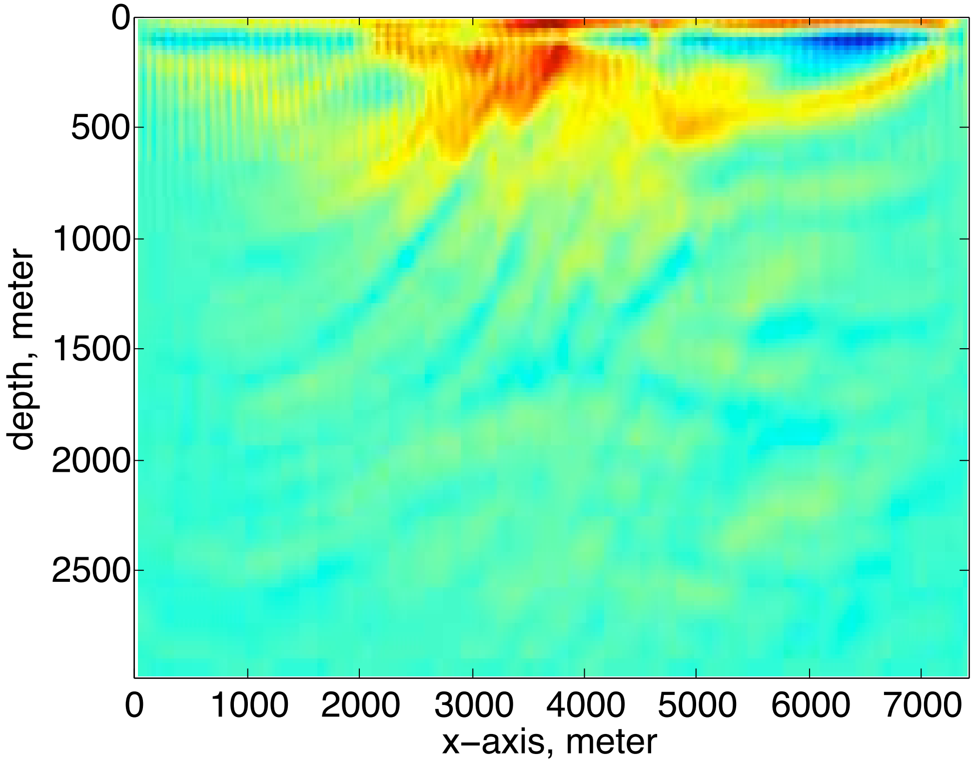


Smooth model

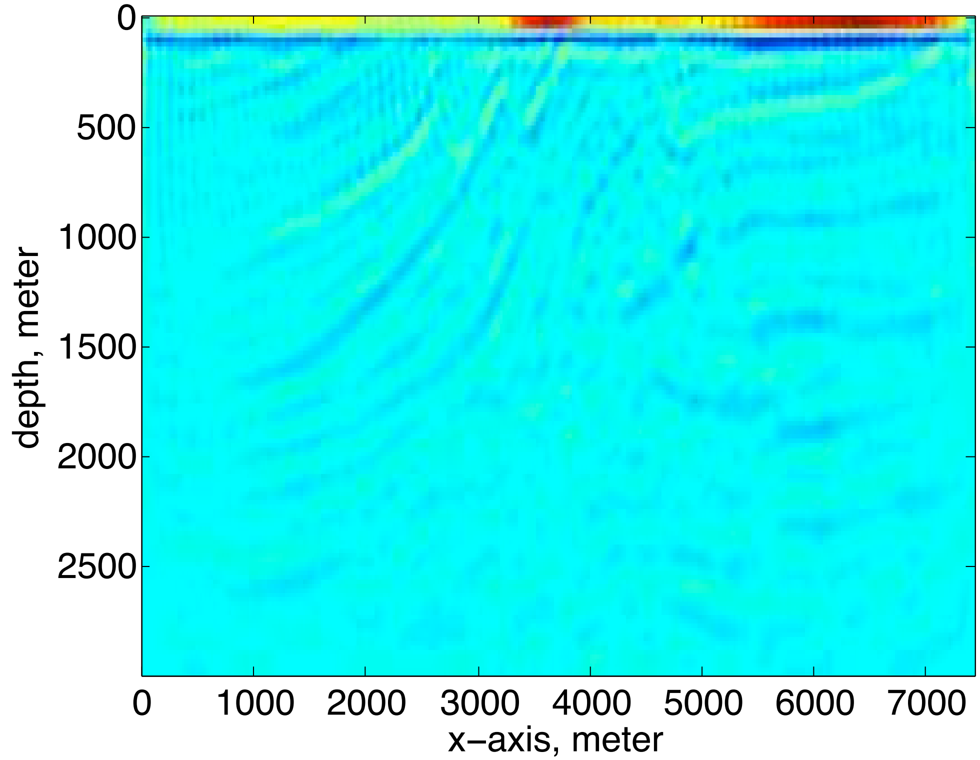
SLIM Seismic Laboratory for Imaging and Modeling

First Update (in $\delta\mathbf{m}$)

Gradient Method



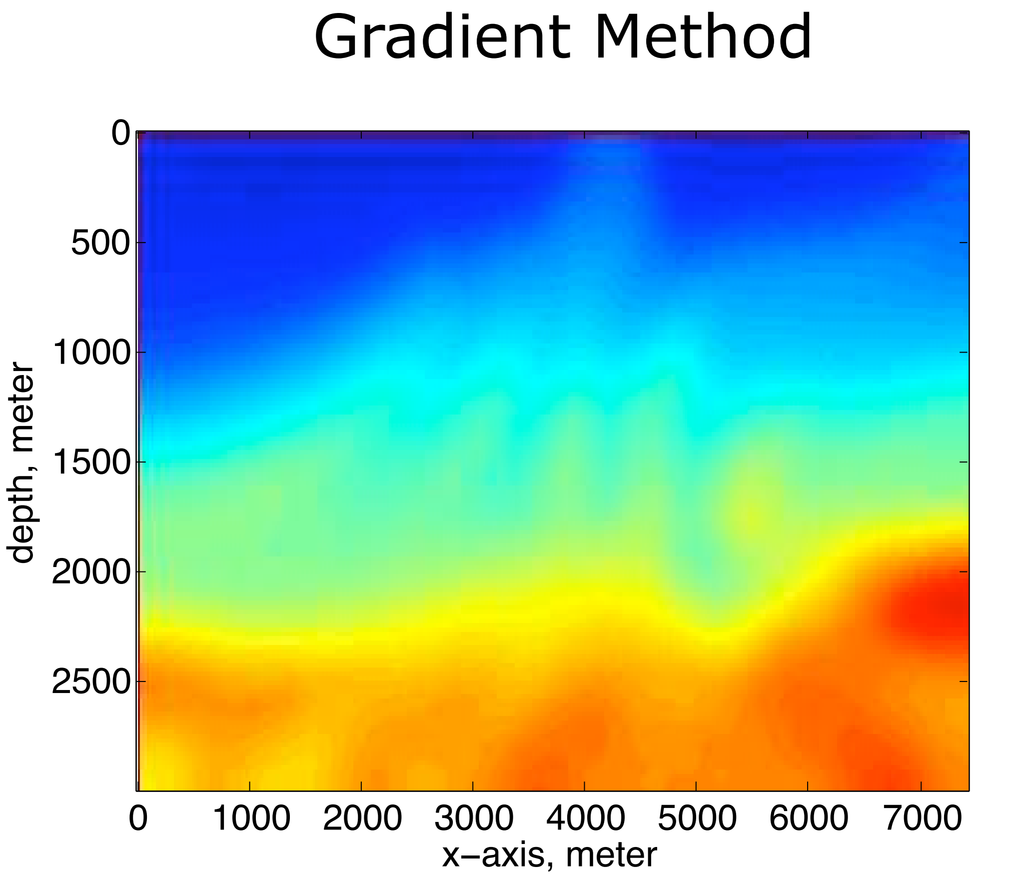
GNK



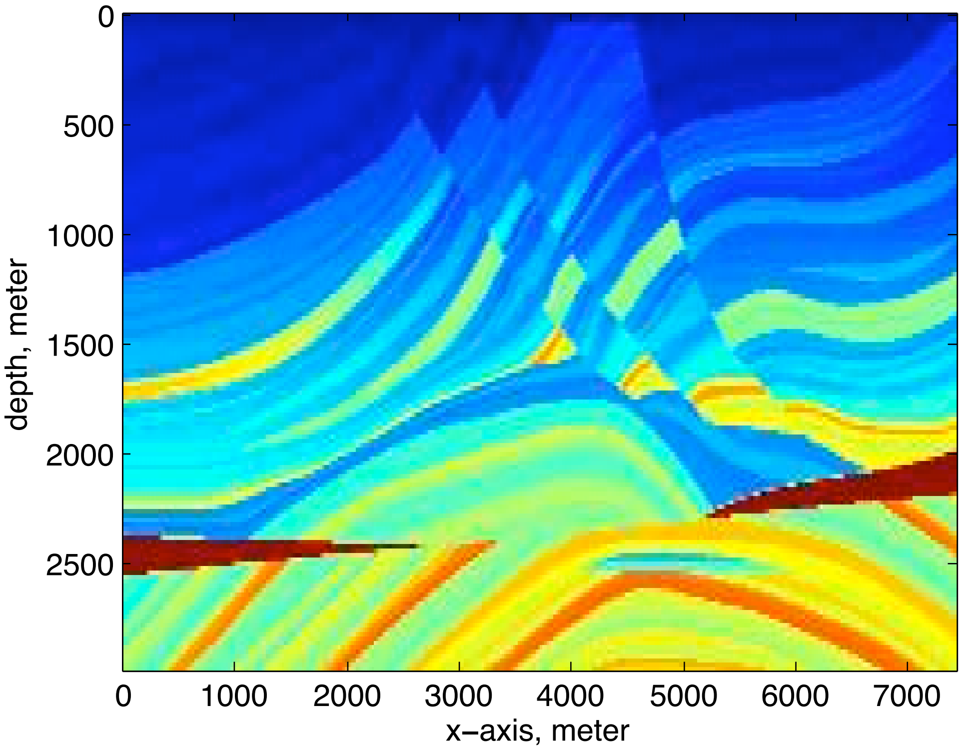
Note: different scale (by 10^3)

SLIM
Seismic Laboratory for Imaging and Modeling

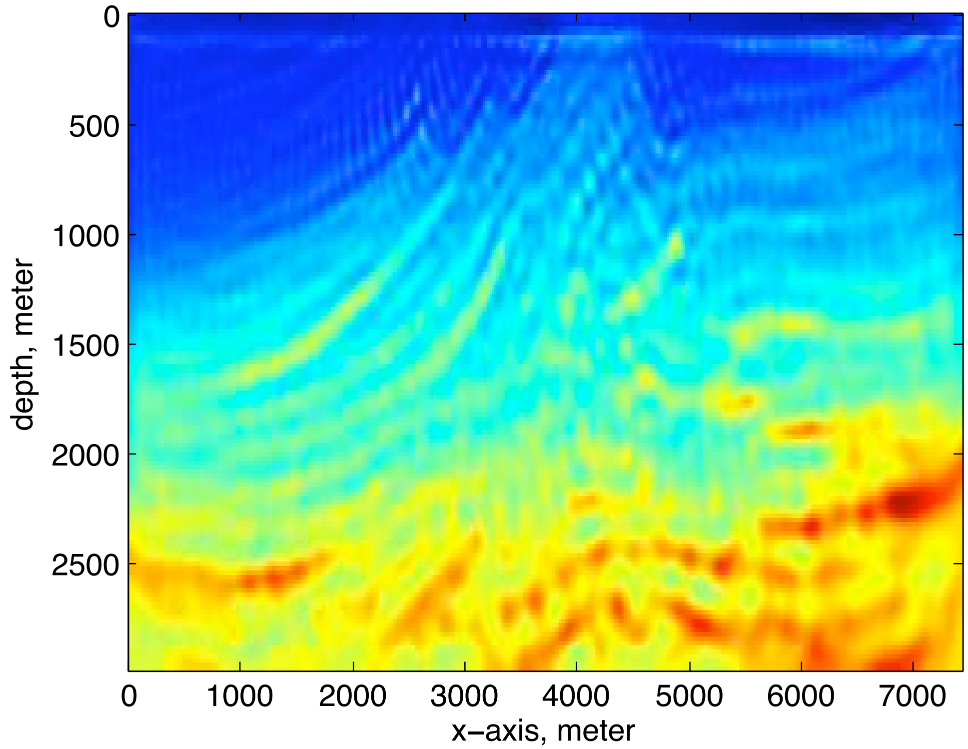
Velocity after the first update



After 5 iterations



Hard model



Inverted Result

SLIM
Seismic Laboratory for Imaging and Modeling

FWI with compressive simultaneous source (CFWI)

Minimization problem:

$$\hat{\mathbf{m}}_R = \arg \min_{\mathbf{m} \in \mathcal{M}} \frac{1}{2} \|\mathbf{RM}(\mathbf{P} - \mathbf{F})\|_2^2$$

$\mathbf{RM}$:

- CS-sampling matrix
- turns single shots into randomized simultaneous shots
- subsamples the shots (fewer shots) and frequencies

Simultaneous shots:

Beasley, Chambers & Jiang (1998), Beasley (2008)
Berkhout (2008)
Neelamani, Krohn, Krebs, Deffenbaugh & Romberg (2008)
Herrmann, E. & Lin (2009)

SLIM Seismic Laboratory for Imaging and Modeling

Gradient method of CFWI

Minimize functional:

$$\begin{aligned} E &= \frac{1}{2}\|\mathbf{RM}(\mathbf{P}-\mathbf{DU})\|_2^2 \\ &= \frac{1}{2}(\mathbf{RM}(\mathbf{P}-\mathbf{DU}))^T\overline{\mathbf{RM}(\mathbf{P}-\mathbf{DU})} \end{aligned}$$

Gradient update: $\mathbf{g}_R = \text{rowsum}\left(\mathbf{J}^T\overline{(\mathbf{RM}(\mathbf{P}-\mathbf{DU}))}\right)$

with the Jacobian $\mathbf{J} \equiv \mathbf{J}(\mathbf{RMDU})$.

$$\mathbf{g}_R = -\text{rowsum}\left(\mathbf{J} \odot \overline{(\mathbf{RM}(\mathbf{P}-\mathbf{DU}))}\right)$$

SLIM
Seismic Laboratory for Imaging and Modeling

Using wavefield-source equivalence [Herrmann, E., & Lin, 2009]

Gradient update $$\delta\mathbf{m} = \mathbf{g} = \underline{\mathbf{J}}^T \overline{(\underline{\mathbf{P}} - \underline{\mathbf{D}\mathbf{U}})}$$

with

- $\underline{\mathbf{J}} \equiv \underline{\mathbf{J}}(\underline{\mathbf{D}\mathbf{U}})$: the (compressed) Jacobian w.r.t. to the compressed simultaneous sources
- $\underline{\mathbf{P}}$: data obtained with simultaneous shot

Computing the Jacobian

Compressed forward model: $\underline{\mathbf{H}}\underline{\mathbf{U}} = \underline{\mathbf{Q}}$

$$\Rightarrow \frac{\partial \underline{\mathbf{U}}}{\partial m_i} = -\underline{\mathbf{H}}^{-1} \frac{\partial \underline{\mathbf{H}}}{\partial m_i} \underline{\mathbf{U}}.$$

$$\delta \mathbf{m} = -\mathrm{rowsum}\left(\underline{\mathbf{U}}^T \left[\frac{\partial \underline{\mathbf{H}}^T}{\partial m_1} \ldots \frac{\partial \underline{\mathbf{H}}^T}{\partial m_M} \right] \overbrace{\underline{\mathbf{H}}^{-T}(\underline{\mathbf{P}} - \underline{\mathbf{D}}\underline{\mathbf{U}})}^{\underline{\mathbf{V}}} \right)$$

$\underline{\mathbf{V}}$: backpropagated wavefield ass. with $\underline{\mathbf{Q}} = \mathbf{RMQ}.$

The GN Hessian can be derived in the similar way!

Complexity Analysis

Gauss-Newton-Krylov (GNK)

- Gradient : forward + back-propagation $n_f n_s n^2 \log n$
- Hessian : forward + back-propagation per CG iteration $n_{CG} n_f n_s n^2 \log n$
- Overall : $n_{CG} n_f n_s n^2 \log n$

Compressive FWI with GNK: $n_{CG} n'_f n'_s n^2 \log n$

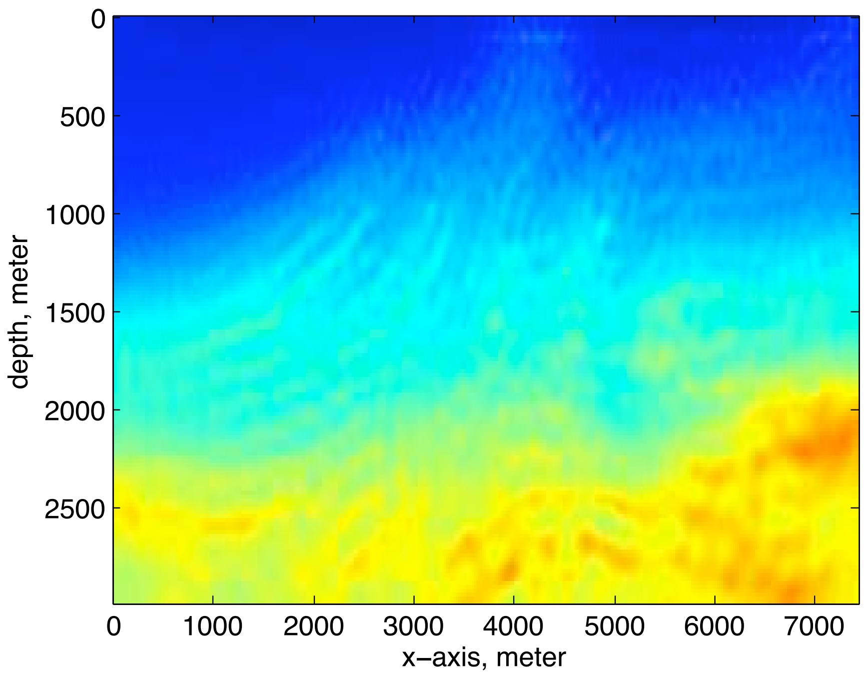
$$n'_f \ll n_f, \quad n'_s \ll n_s$$

Construction of $\mathbf{RM}$ negligible compared to FWI

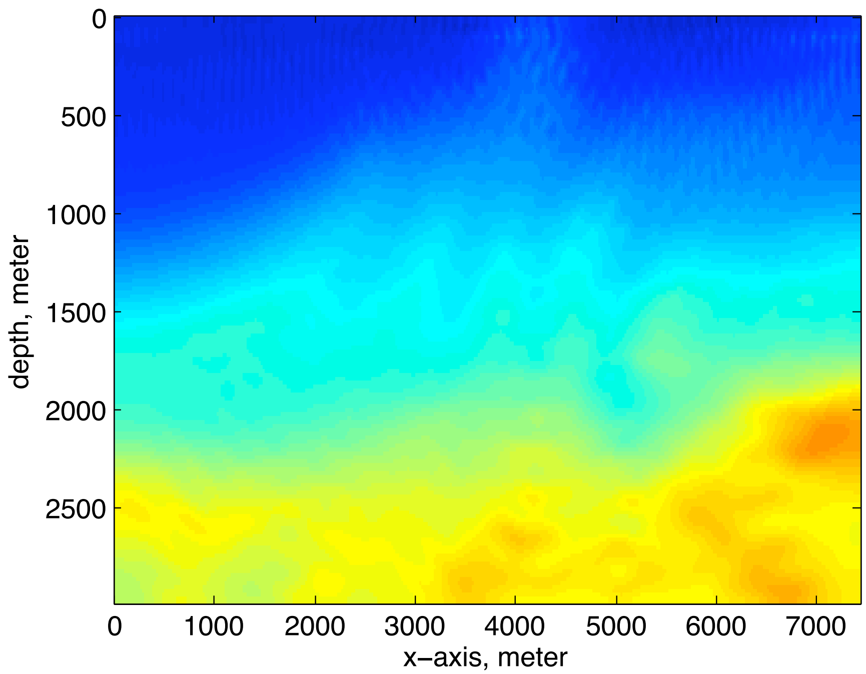
CFWI: Examples, GNK Iter #1

90% subsampled

37 randomized simul. shots



37 periodic shots

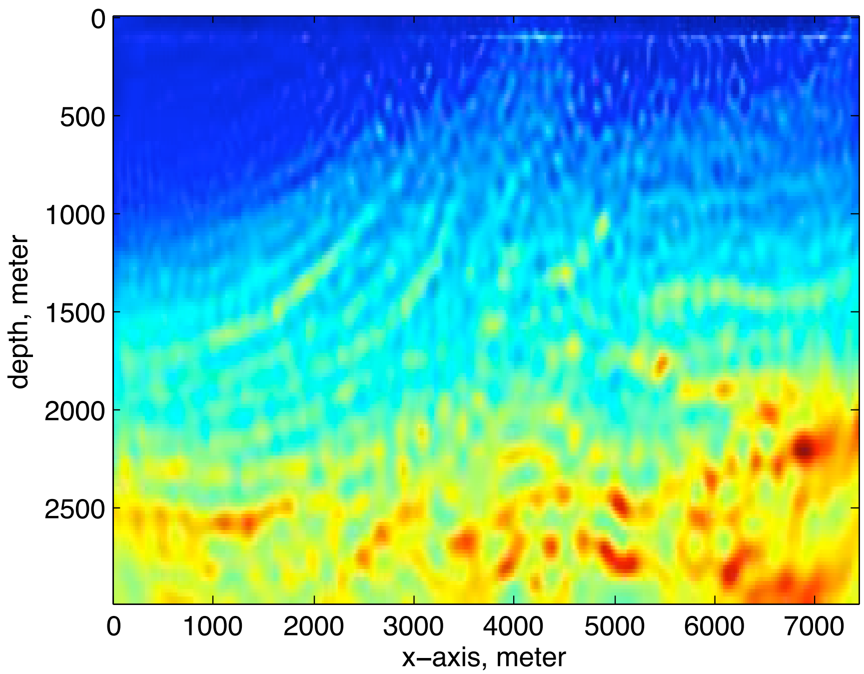


Noisy image --> recover the image via sparsity promoting

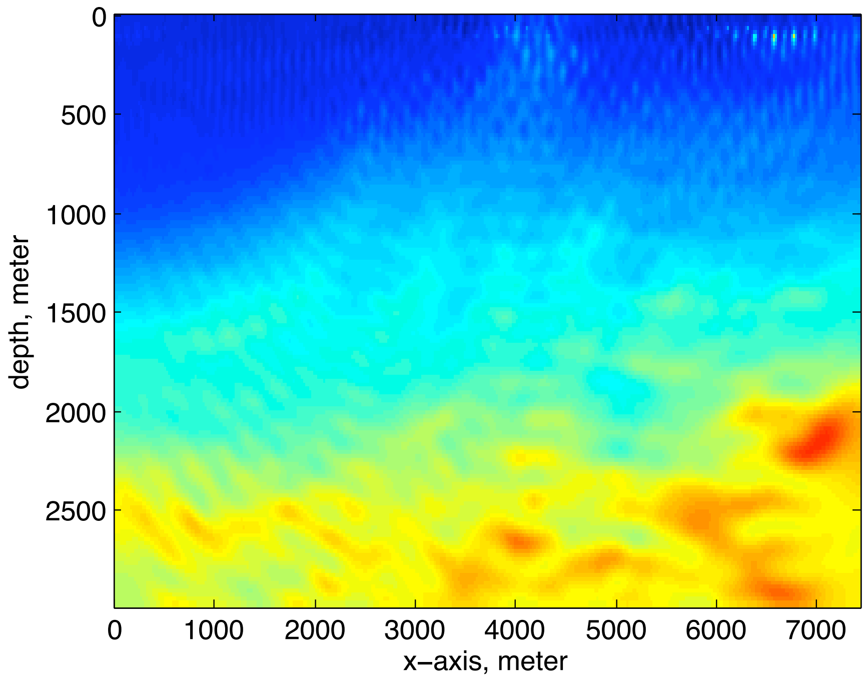
CFWI: Examples, GNK Iter #5

90% subsampled

37 randomized simul. shots



37 periodic shots

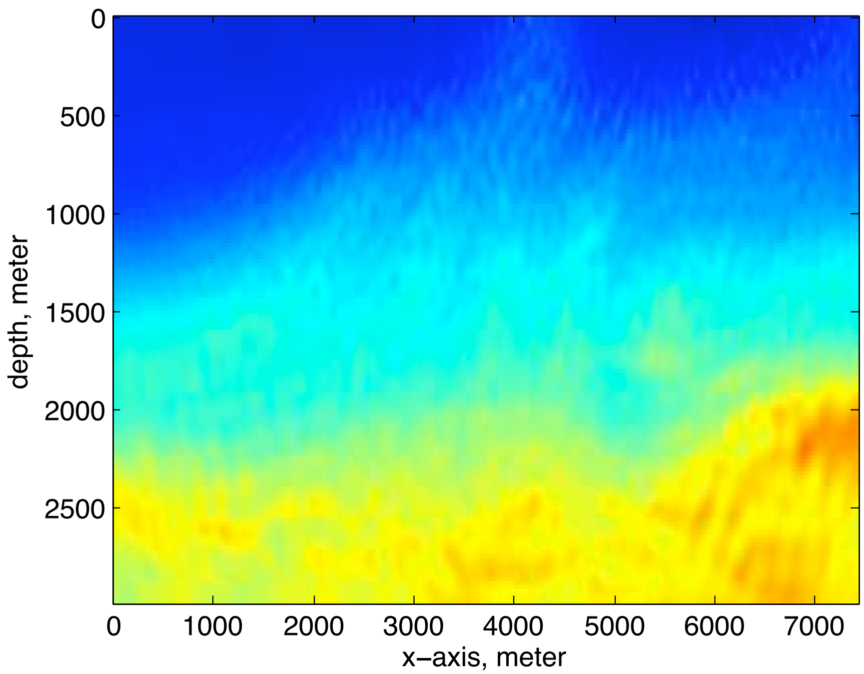


SLIM
Seismic Laboratory for Imaging and Modeling

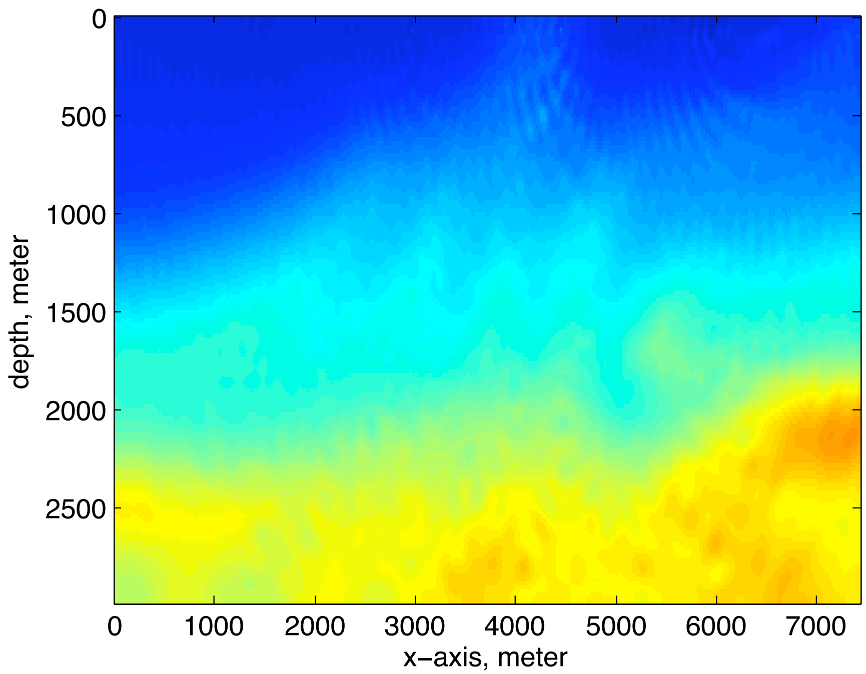
CFWI: Examples, GNK Iter #1

99% subsampled

4 randomized simul. shots



4 periodic shots

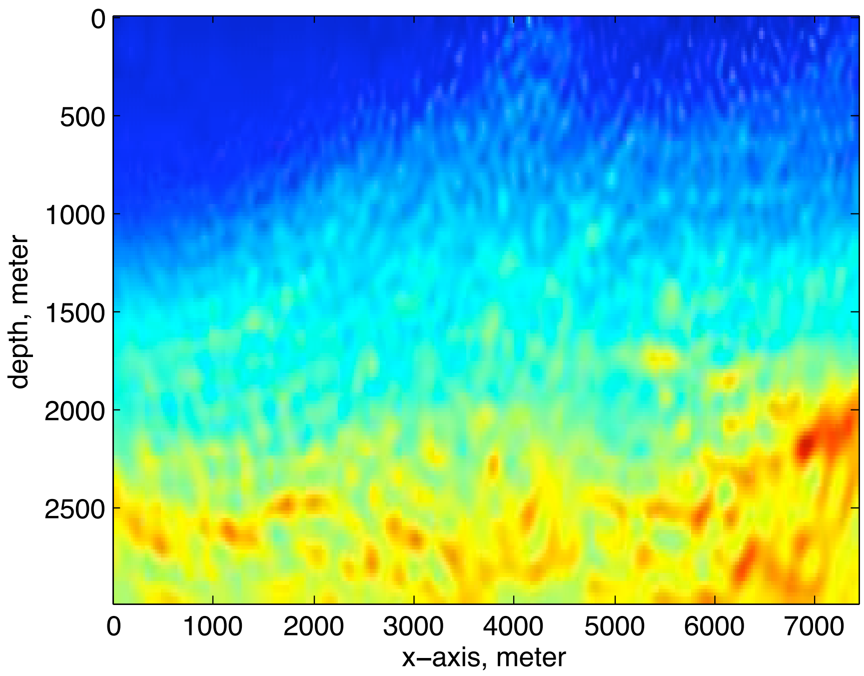


SLIM
Seismic Laboratory for Imaging and Modeling

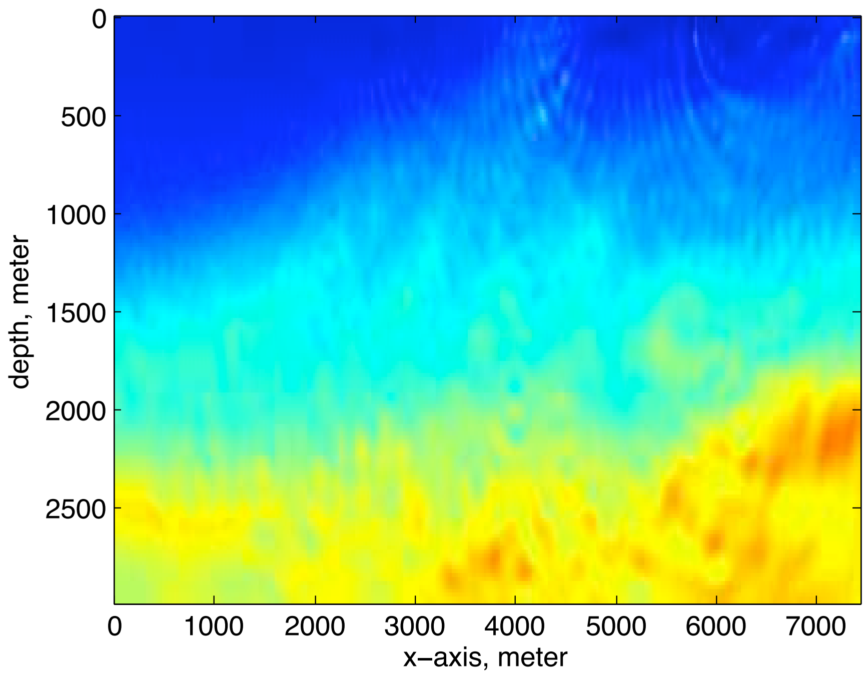
CFWI: Examples, GNK Iter #5

99% subsampled

4 randomized simul. shots



4 periodic shots



SLIM
Seismic Laboratory for Imaging and Modeling

Conclusion

- Viable inversion of GN Hessian with Krylov method
  - Accuracy of the inversion of Hessian depends on the number of iterations --> better FWI result
  - Faster convergence of CG by preconditioners
    - Implicit BFGS-type preconditioner
    - Curvelet-based preconditioner [Herrmann, Brown, E. & Moghaddam (2009)]
  - Memory-friendly algorithm (gradient and Hessian can be computed on the fly)
  - With scalable implicit solver for forward and adjoint systems, matrix-free algorithm [E., Oosterlee & Vuik (2006), E. & Nabben (2009), E. & Herrmann (2008)]
- Natural extension to compressive FWI
  - Similar results but less computational work
  - In the CS framework: l1 inversion

SLIM
Seismic Laboratory for Imaging and Modeling

Acknowledgments

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*** *(334810-05)* of *Felix J. Herrmann.*

This research was carried out as part of ***the SINBAD II*** *project* with support from the following organizations: ***BG Group***, ***BP***, ***Petrobras***, and ***Schlumberger***.

Further information: slim.eos.ubc.ca

SLIM
Seismic Laboratory for Imaging and Modeling