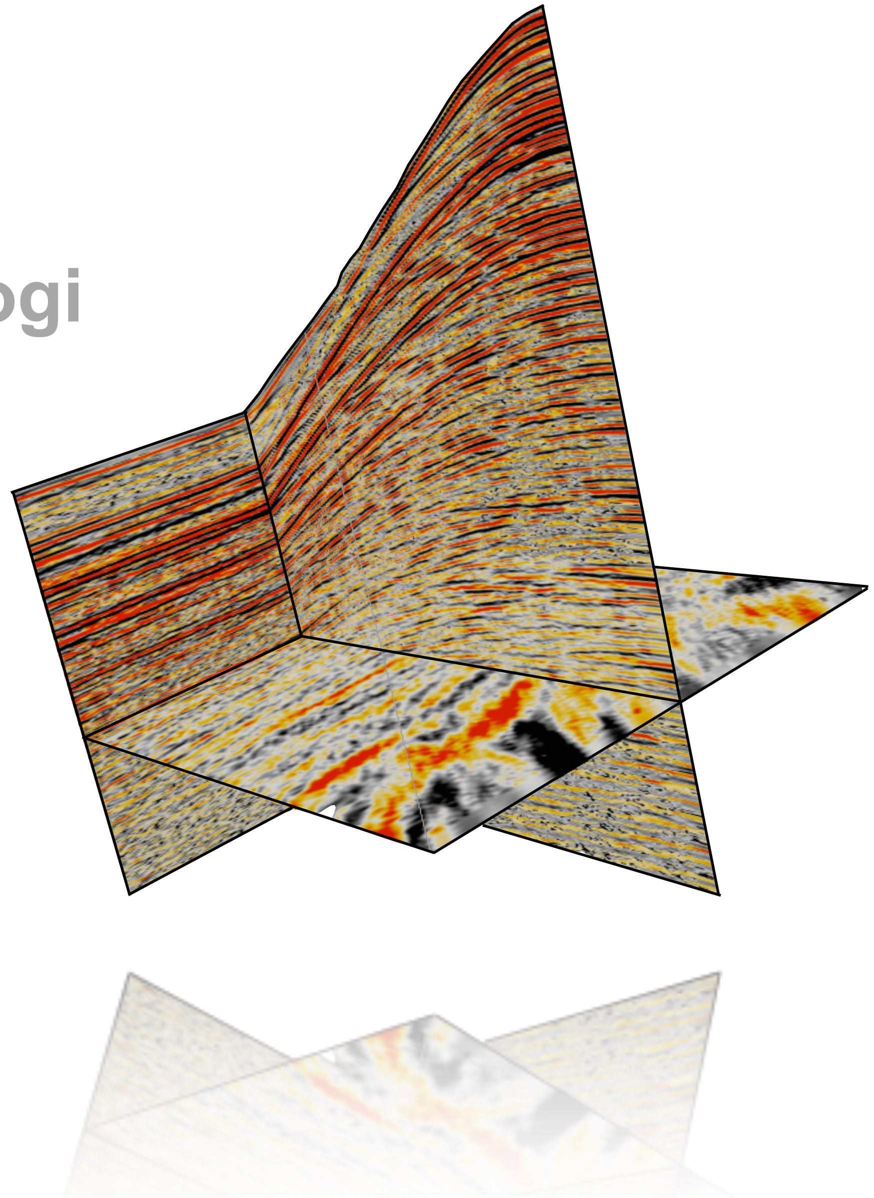
Compressive sampling meets seismic imaging

Felix J. Herrmann
fherrmann@eos.ubc.ca
http://slim.eos.ubc.ca

joint work with Tim Lin and Yogi Erlangga

Seismic Laboratory for Imaging & Modeling
Department of Earth & Ocean Sciences
The University of British Columbia



SIAM Conference on Imaging Science
San Diego, July 2008

Motivation

Seismic data processing, modeling & imaging

- firmly rooted in Nyquist's paradigm
  - sampling (e.g. of wavefields)
  - sampling of solutions (e.g. of PDEs)
- acquisition, modeling & inversion **costs** are proportional to the **size** of *data* and *model*

New paradigm of *compressive sensing* (CS)

- Nyquist is too *pessimistic* for signals with *structure*
  - existence of some sparsifying transform (e.g. wavelets)
  - existence of some low-dimensional structure (smooth manifolds)
- allows for recovery from sample **rates ≈ computational cost** *proportional* to the **complexity** of *data* and *model*

SLIM
Seismic Laboratory for Imaging and Modeling

Main ingredients

New **preconditioner** for the *Helmholtz* operator

[Erlanga & Nabben, '06-'08, Elangga, Lin, F.J.H., '08]

Current advent of **simultaneous & continuous** source acquisition and modeling

[Romero et. al., '00; Neelamani & C.E. Krohn, '08]

Sparsity-promoting **recovery** using results from **CS**

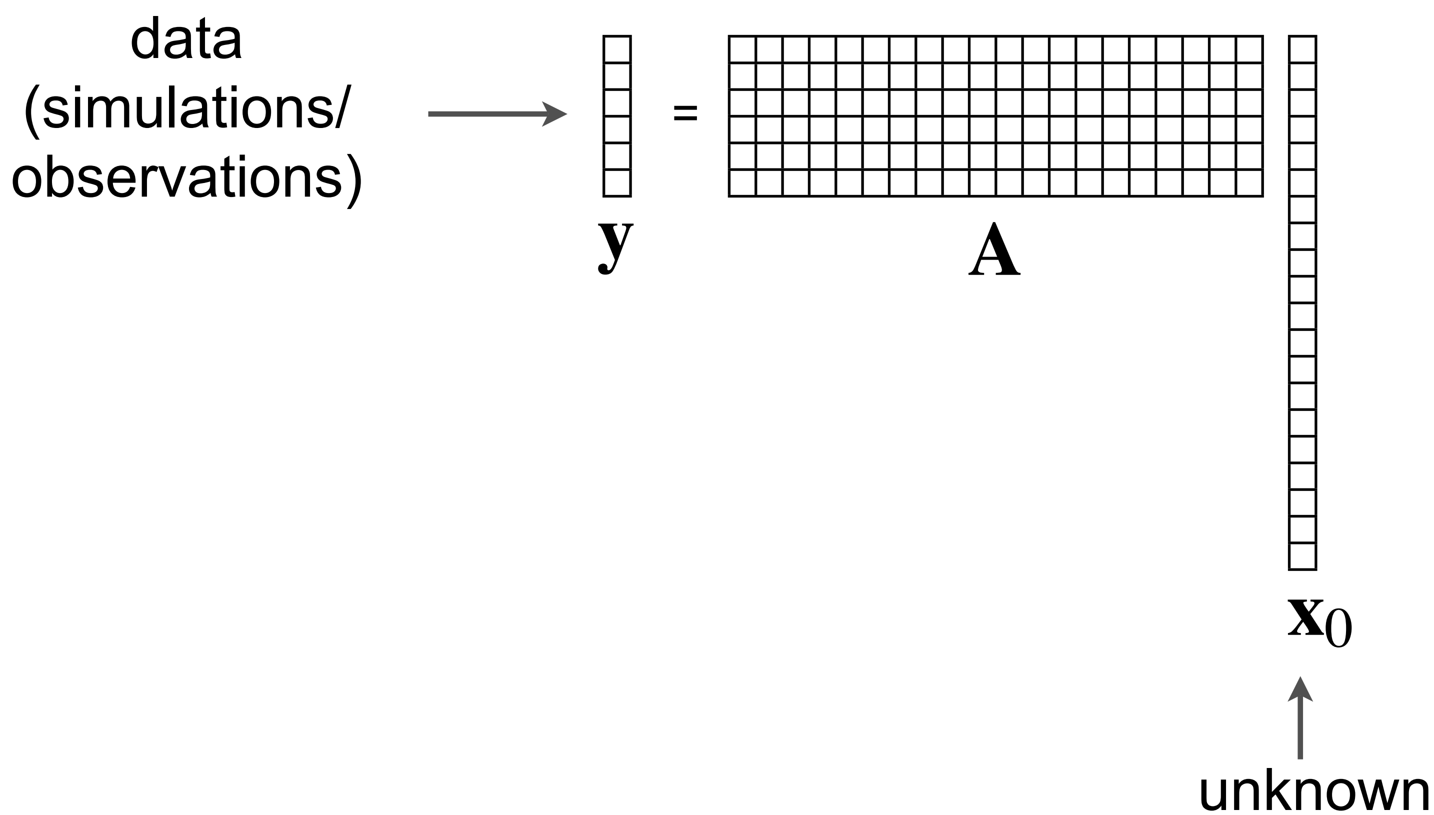
[Donoho, '06; Candes et al., '06; Candes and Tao, '06]

SLIM Seismic Laboratory for Imaging and Modeling

CS

problem statement

Consider the following (severely) underdetermined system of linear equations

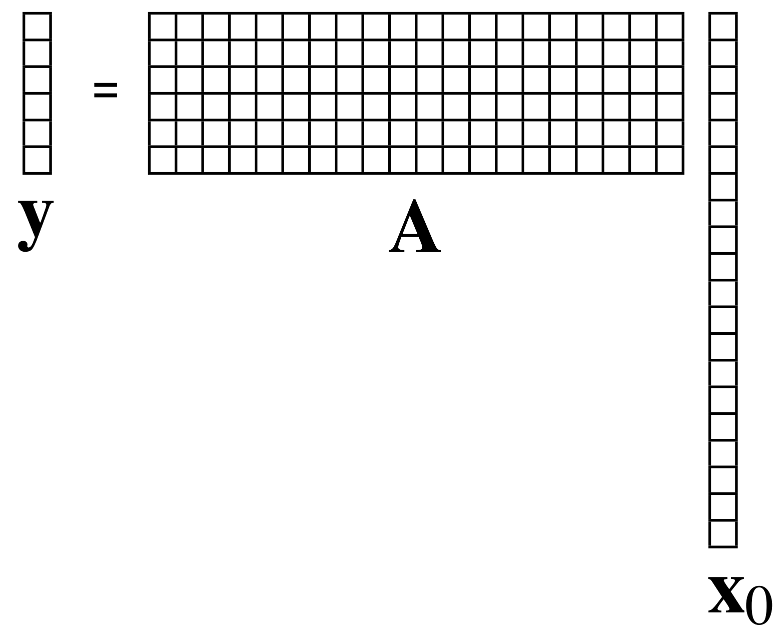


Is it possible to recover $\mathbf{x}_0$ accurately from $\mathbf{y}$?

[Candès et al.‘06]
[Donoho‘06]

CS

perfect recovery



conditions:

- **A** obeys the *uniform uncertainty principle*
- $\mathbf{x}_0$ is *sufficiently sparse*

procedure:

$$\underbrace{\min_{\mathbf{x}} \|\mathbf{x}\|_1}_{\text{sparsity}} \quad \text{s.t.} \quad \underbrace{\mathbf{A}\mathbf{x} = \mathbf{y}}_{\text{perfect reconstruction}}$$

performance:

- *S*-sparse vectors recovered from roughly on the order of *S* measurements (to within constant and *log* factors)

SLIM
Seismic Laboratory for Imaging and Modeling

Adjoint state method

[Lailly, Pratt]
[Plessix, Mulder]

Unconstrained nonlinear LS problem

$$\min_{\mathbf{m}\in\mathcal{M}} \frac{1}{2}\|\mathbf{b}-\mathbf{F}[\mathbf{m}]\|_2^2$$

with

$$\mathbf{F}[\mathbf{m}] = \mathbf{D}\mathbf{A}^{-1}[\mathbf{m}]\mathbf{f}$$

and the ***gradient*** = - ***migrated image***,

$$\left[\nabla J(\mathbf{m})\right]_i = -\Re\left(\sum_{\omega}\sum_{s}\langle(\frac{\partial\mathbf{A}}{\partial\mathbf{m}}\mathbf{e}_i)\mathbf{u}_s,\ \mathbf{v}_s\rangle\right)$$

involves for each **monochromatic shot** the solution of

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

with

$$\mathbf{r} = \mathbf{D}^H(\mathbf{b}-\mathbf{F}[\mathbf{m}])$$

SLIM
Seismic Laboratory for Imaging and Modeling

Forward modeling

Current paradigm: *time-domain finite differences*

Pro: relatively simple, implicit and fast

Con:

- discretization criteria for numerical stability
- storage requirements for
  - model (domain decompositions)
  - imaging conditions (check pointing)

New 'paradigm': *implicit preconditioned Helmholtz solvers*

Pro:

- matrix free, favorable criteria for numerical stability
- embarrassing parallelization over angular frequency

Con:

- slow or no convergence of indirect Krylov methods

Solution: preconditioner

SLIM
Seismic Laboratory for Imaging and Modeling

Forward modeling

Discretize frequency-domain acoustic wave equation

$$\mathcal{H}u(\omega, x_s; x) := -\left(\nabla \cdot \nabla - \frac{\omega^2}{c(x)^2}\right) u(\omega, x_s; x) = b$$

Monochromatic linear system

$$\mathbf{A}_\omega[\mathbf{c}]\mathbf{u}^s = \mathbf{b}^s$$

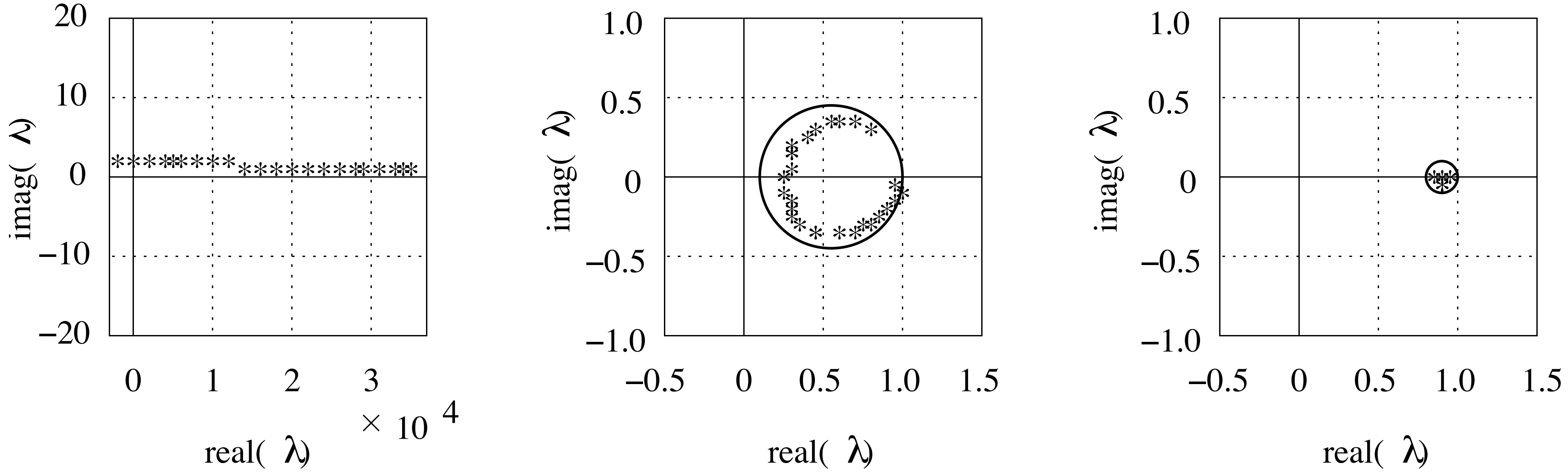
Preconditioned system

$$\mathbf{AM}^{-1}\hat{\mathbf{u}} = \mathbf{b}, \qquad \mathbf{u} = \mathbf{M}^{-1}\hat{\mathbf{u}},$$

derived from shifted Laplacian

$$\mathcal{M} := -\nabla \cdot \nabla - \frac{\omega^2}{c(x)^2}(1 - \beta i) \quad \text{with} \quad i = \sqrt{-1},\ \beta > 0$$

Forward modeling cont'd



Preconditioning [Erlangga & Nabben, '06-'08]**:**

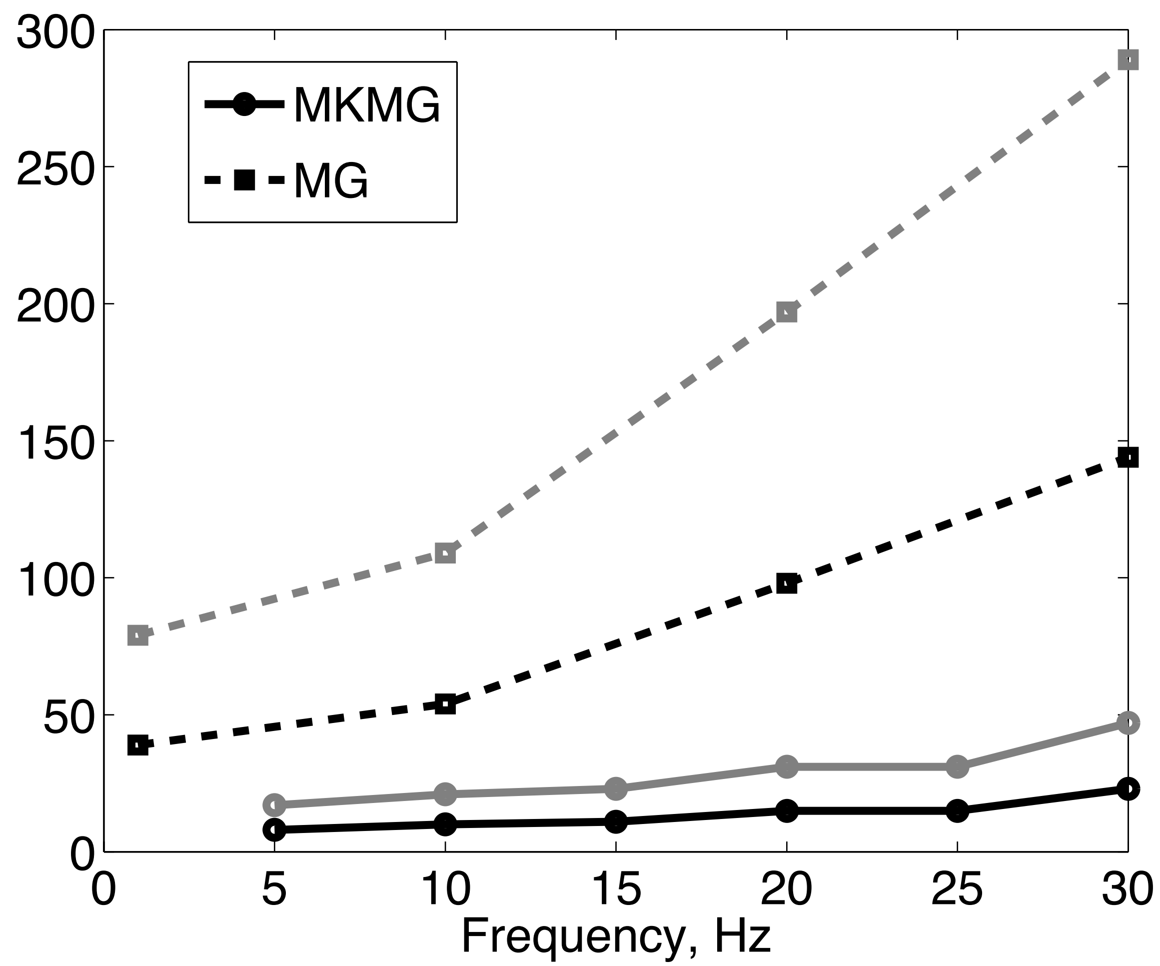
- moves eigenvalues to circle in complex plane
- is inverted using multigrid (no longer elliptic)

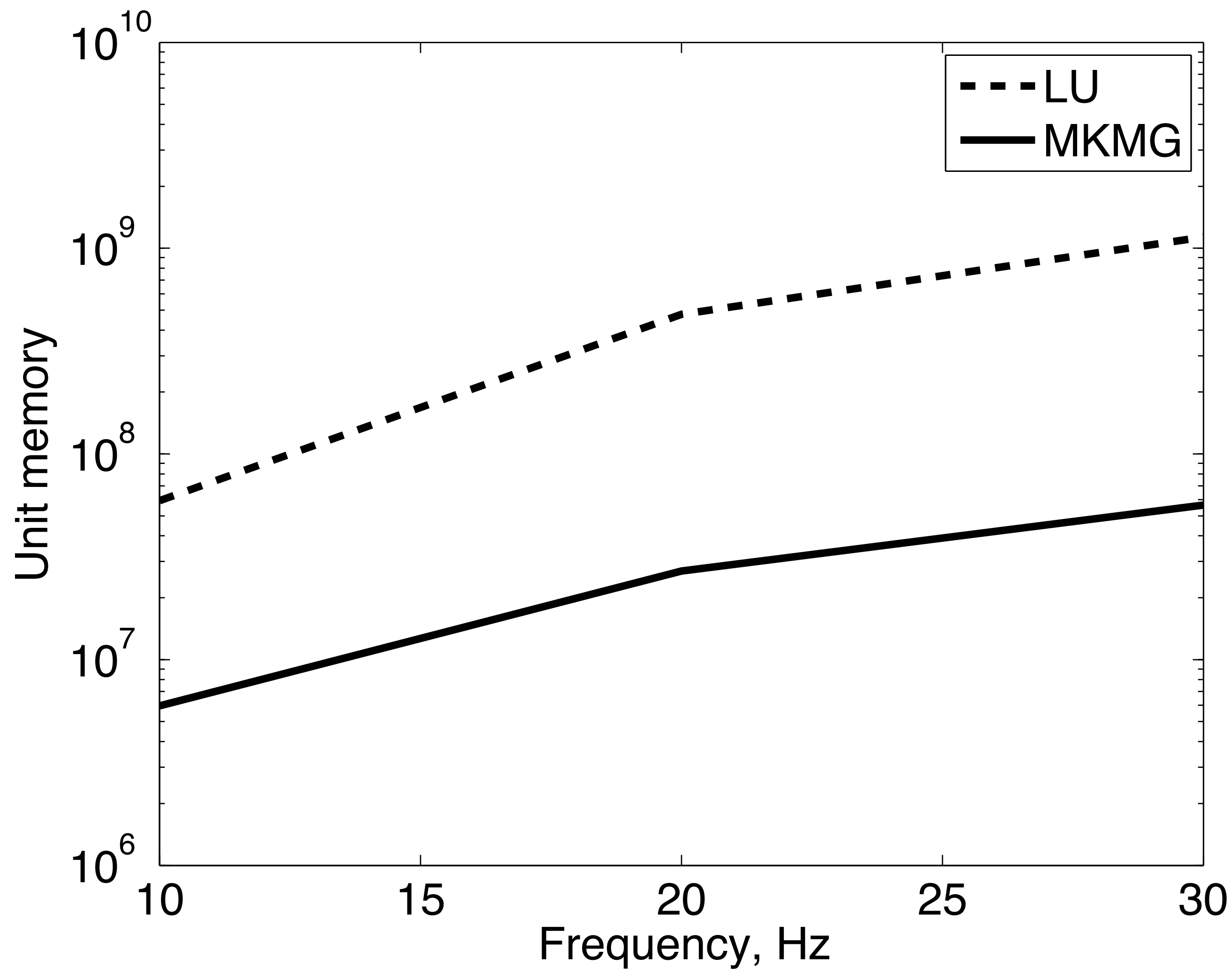
Additional multi-level Krylov projection:

$$\mathbf{A}\mathbf{M}^{-1}\mathbf{Q}\hat{\mathbf{u}} = \mathbf{b}, \quad \mathbf{u} = \mathbf{M}^{-1}\mathbf{Q}\hat{\mathbf{u}},$$

- moves eigenvalues to real axis near 1
- improves condition number

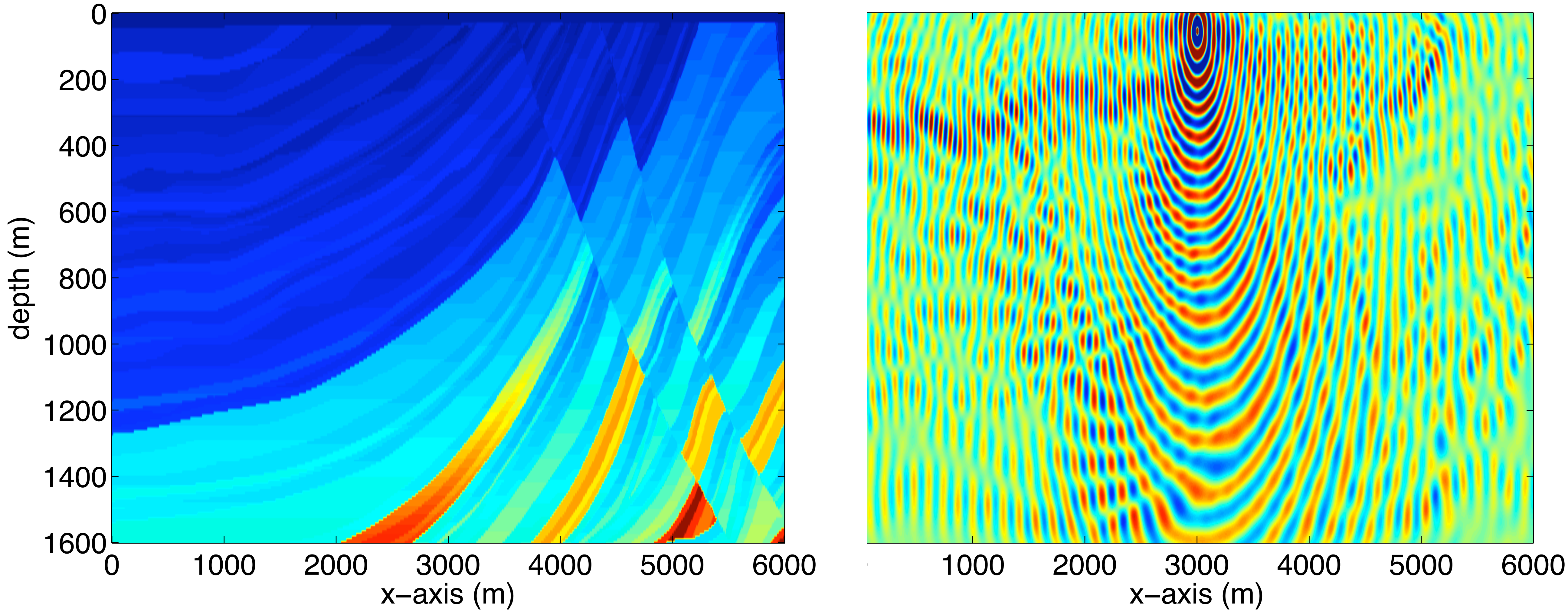
Forward modeling cont'd





- implicit solvers converge
- number of iterations flat in grid size & frequency
- opens perspective to large-scale parallel solver for 3-D models

Forward modeling cont'd



Despite significant improvement by Helmholtz preconditioner

- redundancy <=> extreme large size seismic data volumes
- multiple frequencies & multiple right-hand sides
- expensive modeling, imaging & inversion costs

Leverage new paradigm of CS ...

SLIM
Seismic Laboratory for Imaging and Modeling

Relation to existing work

Simultaneous & continuous acquisition:

- *Simultaneous Sourcing without Compromise* by R. Neelamani & C.E. Krohn, ’08.

Simultaneous simulations & migration:

- *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.

Imaging:

- *How to choose a subset of frequencies in frequency-domain finite-difference migration* by Mulder & Plessix, ’04.
- *Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies* by Sirque and Pratt, ’04.

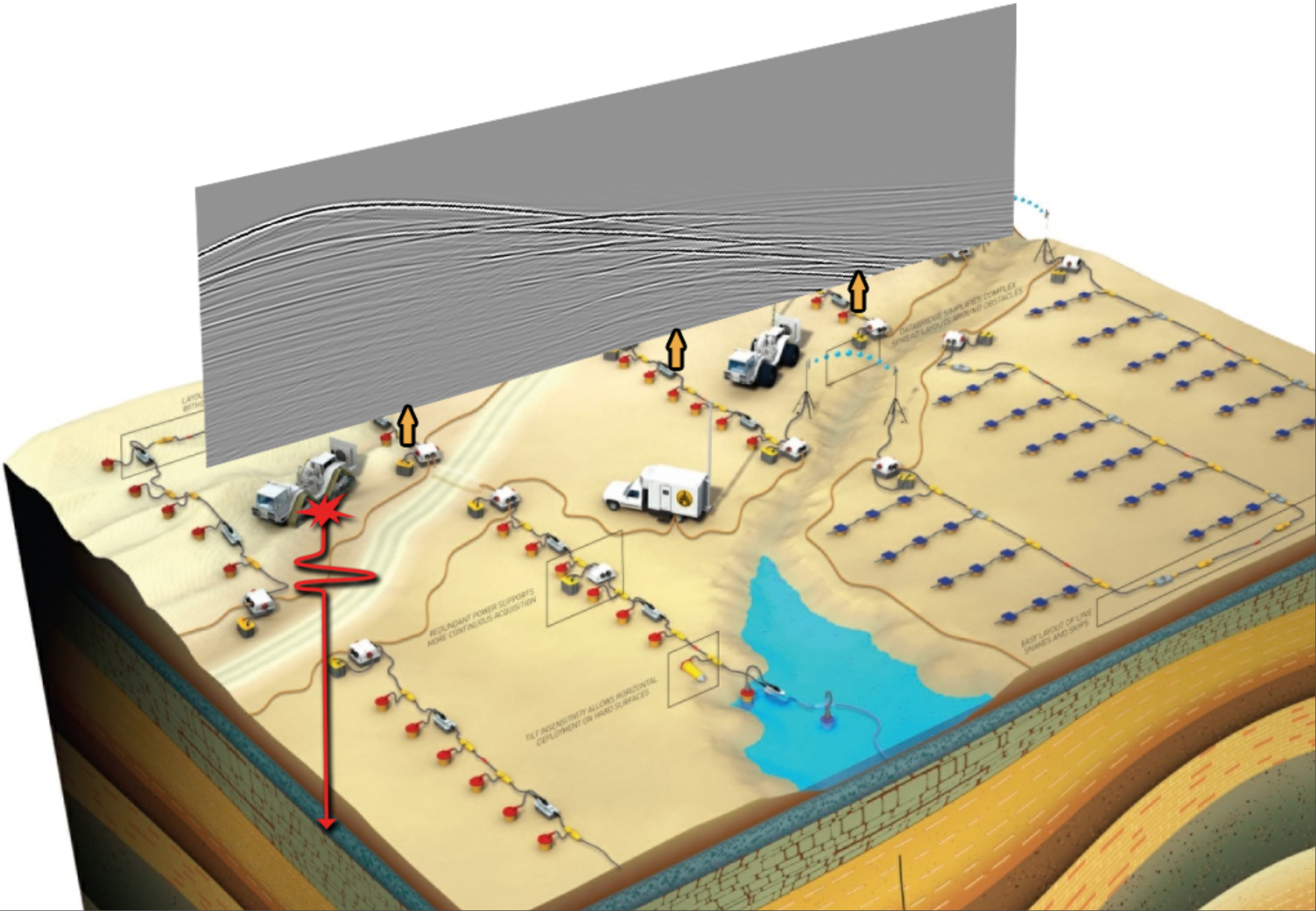
Wavefield extrapolation:

- *Compressed wavefield extrapolation* by T. Lin and F.J.H, ’07
- *Compressive wave computations* by L. Demanet in MS79

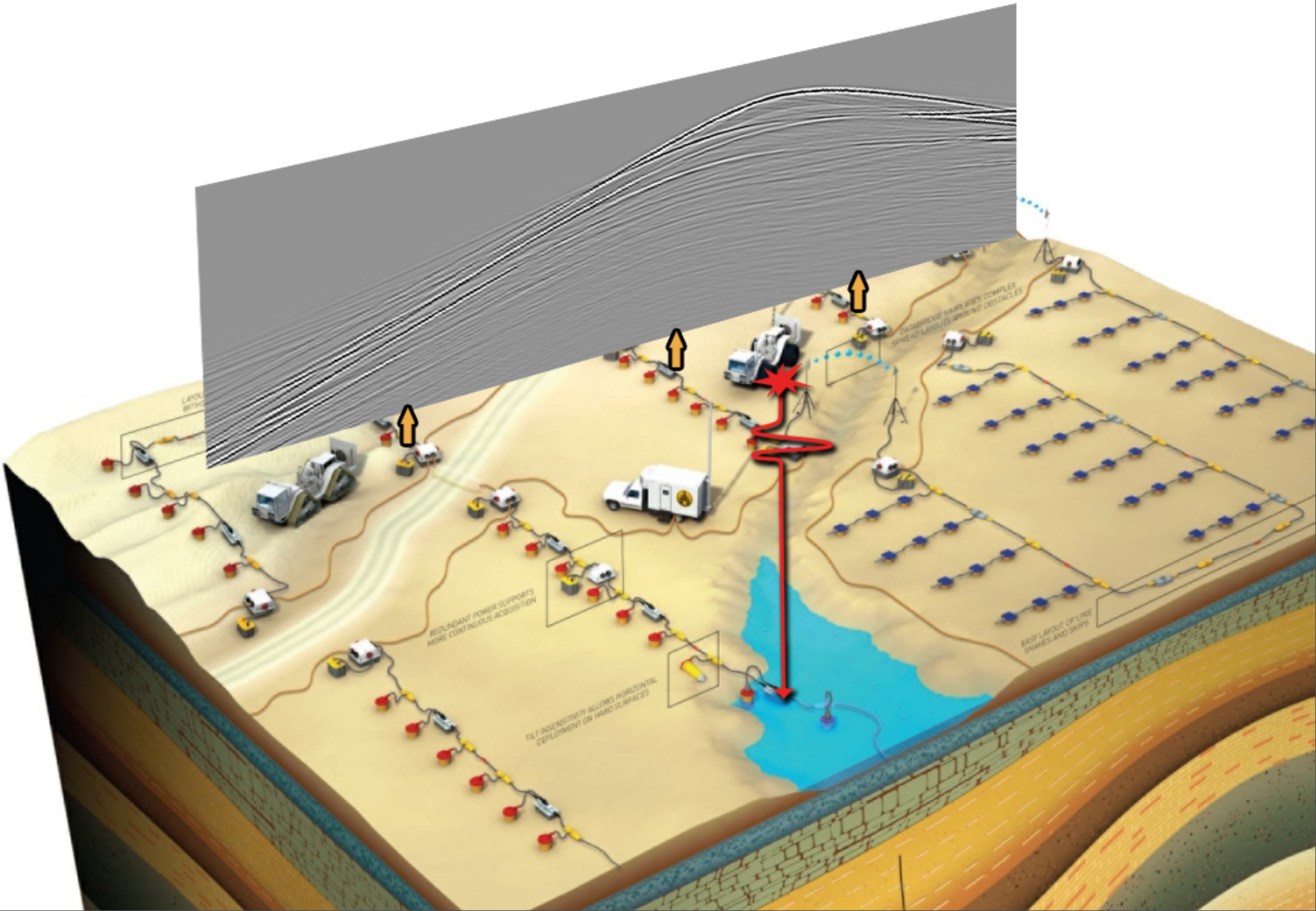
SLIM
Seismic Laboratory for Imaging and Modeling

Acquisition image courtesy ION (www.iongeo.com)

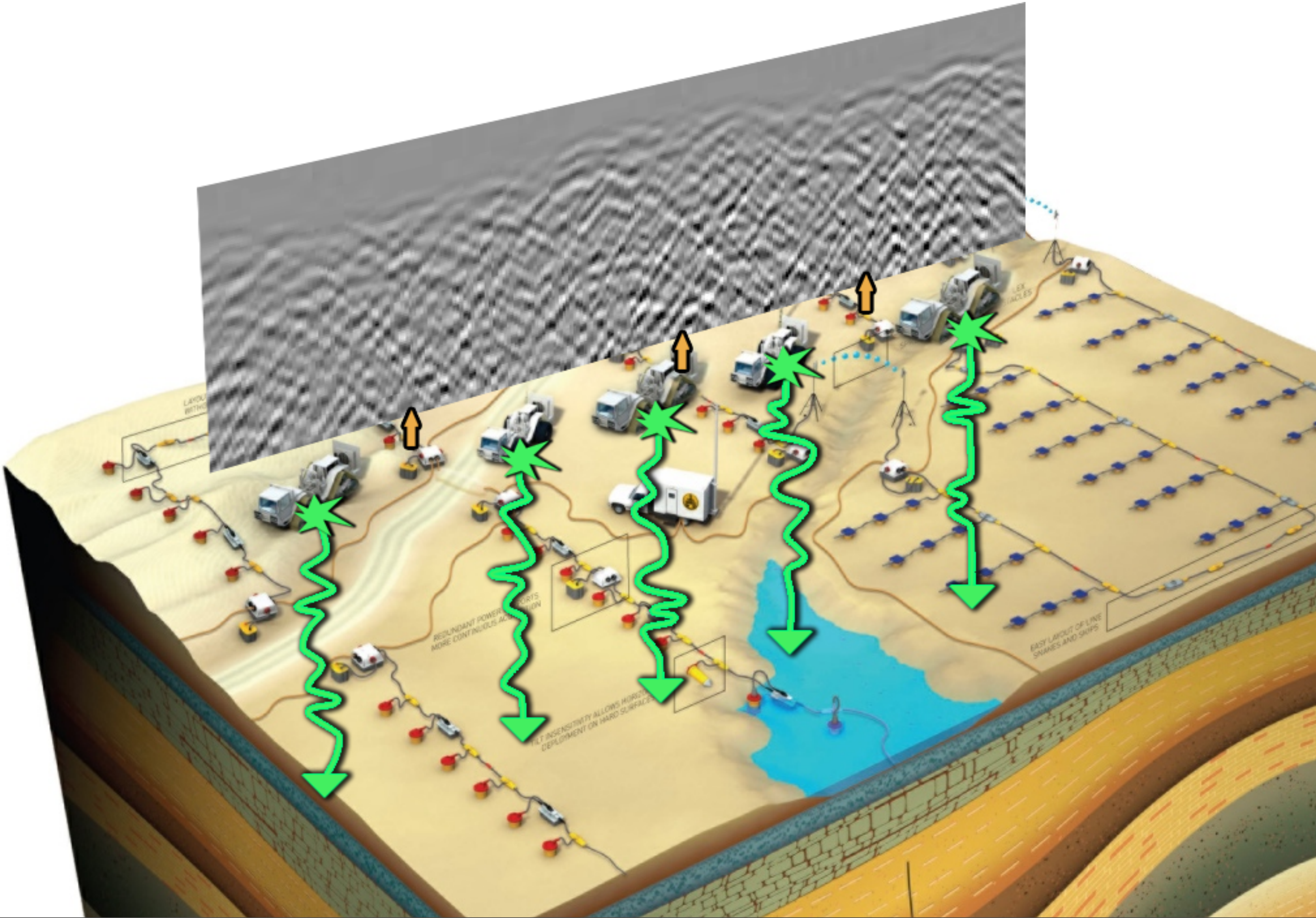
Individual shots



Individual shots



Simultaneous & continuous shots



Simultaneous

modeling & acquisition

Current paradigm:

- ***separate*** single-source experiments in the field
- ***separate*** single-shot simulations in the computer
- **Con:** expensive

New paradigm:

- ***simultaneous*** & ***continuous*** source experiments in the field
- ***simultaneous*** (continuous) simulations in the computer
- **continuous** simultaneous simulations are equivalent to **multiple** simultaneous experiments
- **Con:** *postprocessing necessary to separate into individual shots*

Key observation: this is *really* CS ...

SLIM
Seismic Laboratory for Imaging and Modeling

Forward modeling

multishot

$$
\begin{bmatrix}
\mathbf{A}_{\omega_1} & 0 & & \\
0 & \mathbf{A}_{\omega_2} & \ddots & \\
 & \ddots & \ddots & 0 \\
 & & 0 & \mathbf{A}_{\omega_{n_f}}
\end{bmatrix}
\begin{bmatrix}
\overbrace{[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{n_s}]_{\omega_1}}^{\mathbf{U}_{\omega_1}} \\
\vdots \\
\vdots \\
\underbrace{[\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_{n_s}]_{\omega_{n_f}}}_{\mathbf{U}_{n_f}}
\end{bmatrix}
=
\begin{bmatrix}
\overbrace{[\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_{n_s}]_{\omega_1}}^{\mathbf{B}_{\omega_1}} \\
\vdots \\
\vdots \\
\underbrace{[\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_{n_s}]_{\omega_{n_f}}}_{\mathbf{B}_{n_f}}
\end{bmatrix}
$$

Helmholtz equation solved for:

- individual angular frequencies, i.e.,

$$\omega_i = 2\pi i \cdot \Delta f,\ i = 1 \cdots n_f,$$

  - n_f number of frequencies
  - Δf the sample interval in frequency

- individual shots, i.e., $\mathbf{b}_i = \mathbf{e}_i$ for $i = 1, \cdots, n_s$

SLIM
Seismic Laboratory for Imaging and Modeling

Forward modeling
multishot

Rewrite into

$$\underbrace{\begin{bmatrix} \mathbf{A}_{\omega_1} & 0 & & \\ 0 & \mathbf{A}_{\omega_2} & \ddots & \\ & \ddots & \ddots & 0 \\ & & 0 & \mathbf{A}_{\omega_{n_f}} \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} \mathbf{U}_{\omega_1} \\ \vdots \\ \vdots \\ \mathbf{U}_{n_f} \end{bmatrix}}_{\mathbf{U}} = \underbrace{\begin{bmatrix} \mathbf{B}_{\omega_1} \\ \vdots \\ \vdots \\ \mathbf{B}_{n_f} \end{bmatrix}}_{\mathbf{B}}$$

or

$$\mathbf{L}\mathbf{U} = \mathbf{B}.$$

Modeling involves the inversion of the matrix

$$\mathbf{L} \in \mathbb{C}^{n_d \times n_d} \text{ with } n_d = 2 n_f n_s n_r$$

Equivalence

Show equivalence between

- CS sampling of ***full*** solution for separate single-source (sweep) experiments
- Solution of ***reduced*** *system* after CS sampling the collective single-shot source wavefield => **simultaneous source experiments**

$$\begin{cases} \mathbf{B} = \mathbf{D}^* \underbrace{\mathbf{s}}_{\text{single shots}} \\ \mathbf{L}\mathbf{U} = \mathbf{B} \\ \mathbf{y} = \mathbf{R}\mathbf{M}\mathbf{D}\mathbf{U} \end{cases} \iff \begin{cases} \underline{\mathbf{B}} = \underline{\mathbf{D}}^* \underbrace{\mathbf{R}\mathbf{M}\mathbf{s}}_{\text{simul. shots}} \\ \underline{\mathbf{L}\mathbf{U}} = \underline{\mathbf{B}} \\ \underline{\mathbf{y}} = \underline{\mathbf{D}\mathbf{U}} \end{cases}$$

Show that $\mathbf{y} = \underline{\mathbf{y}}$ for which it is sufficient to show that

$$\mathbf{R}_\Omega \overbrace{\mathbf{L}^{-1}\mathbf{B}}^{\text{full system}} \mathbf{R}_\Sigma^* = \underline{\underline{\mathbf{U}}} \iff \underline{\mathbf{U}} = \overbrace{\left(\mathbf{R}_\omega \mathbf{L} \mathbf{R}_\Sigma^*\right)^{-1} \underline{\mathbf{B}}}^{\text{reduced system}}$$

SLIM Seismic Laboratory for Imaging and Modeling

Equivalence cont'd

Fourier restriction:

$\mathbf{R}_\Omega : n'_f \times n_f$ block matrix, $\quad n'_f = \#\{\Omega\}, \Omega \subset \{\omega_i\}, i = 1, \ldots, n_f, n_f \gg n'_f$

$$[\mathbf{R}_\Omega]_{J,I} = \begin{cases} \mathbf{I}_{n_x \times n_z}, & I \in \mathcal{I} \\ \mathbf{O}_{n_x \times n_z}, & I \notin \mathcal{I}, \end{cases}$$

with $\mathcal{I}$ the index set of Ω, and $J = 1, \ldots, n'_f$.

Identity: $\mathbf{R}_\Omega \mathbf{L} = \underline{\mathbf{L}} \mathbf{R}_\Omega$, where

$$\underline{\mathbf{L}} = \mathrm{diag}\left(\mathbf{A}_{\omega_I}\right), \qquad I \in \mathcal{I}.$$

This implies: $\mathbf{R}_\Omega \mathbf{L}^{-1} = \underline{\mathbf{L}}^{-1} \mathbf{R}_\Omega$.

SLIM Seismic Laboratory for Imaging and Modeling

Equivalence cont'd

Shot restriction:

$\mathbf{R}_\Sigma : n_s' \times n_s$ rectangular matrix, $\quad n_s' = \#\{\mathcal{N}_s'\}, \mathcal{N}_s' \subset \mathcal{N}_s,$ with $\mathcal{N}_s$ the index set of $\mathbf{b}_i$.

$$[\mathbf{R}_\Sigma]_{j,i} = \begin{cases} 1, & i \in \mathcal{N}_s', \\ 0, & i \notin \mathcal{N}_s', \end{cases}$$

for $j = 1, \ldots, n_s'$ with $n_s' \ll n_s$.

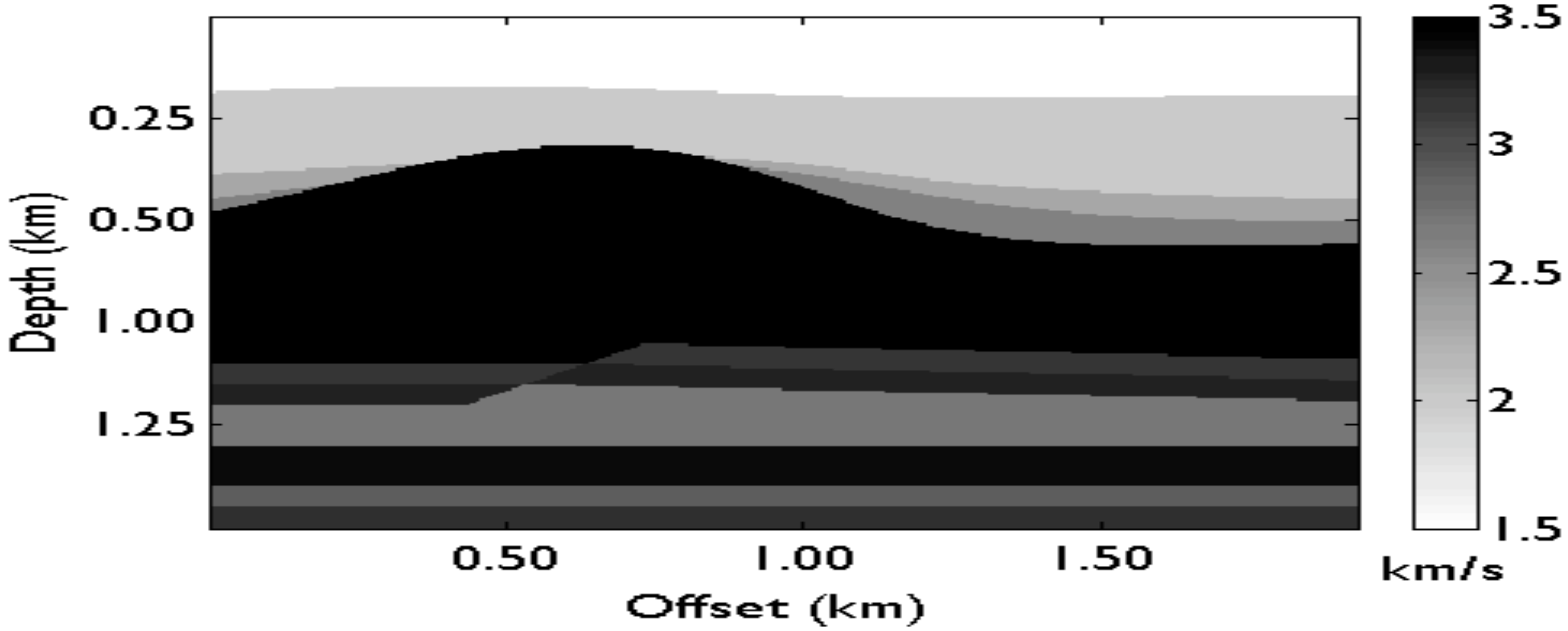
So we have,

$$\mathbf{R}_\Omega \mathbf{L}^{-1} = \underline{\mathbf{L}}^{-1} \mathbf{R}_\Omega \Rightarrow \underbrace{\mathbf{R}_\Omega \overbrace{\mathbf{L}^{-1} \mathbf{B}}^{\mathbf{U}} \mathbf{R}_\Sigma^*}_{\underline{\underline{\mathbf{U}}}} = \underline{\mathbf{L}}^{-1} \underbrace{\mathbf{R}_\Omega \mathbf{B} \mathbf{R}_\Sigma^*}_{\underline{\mathbf{B}}} = \underline{\mathbf{U}}.$$

implying

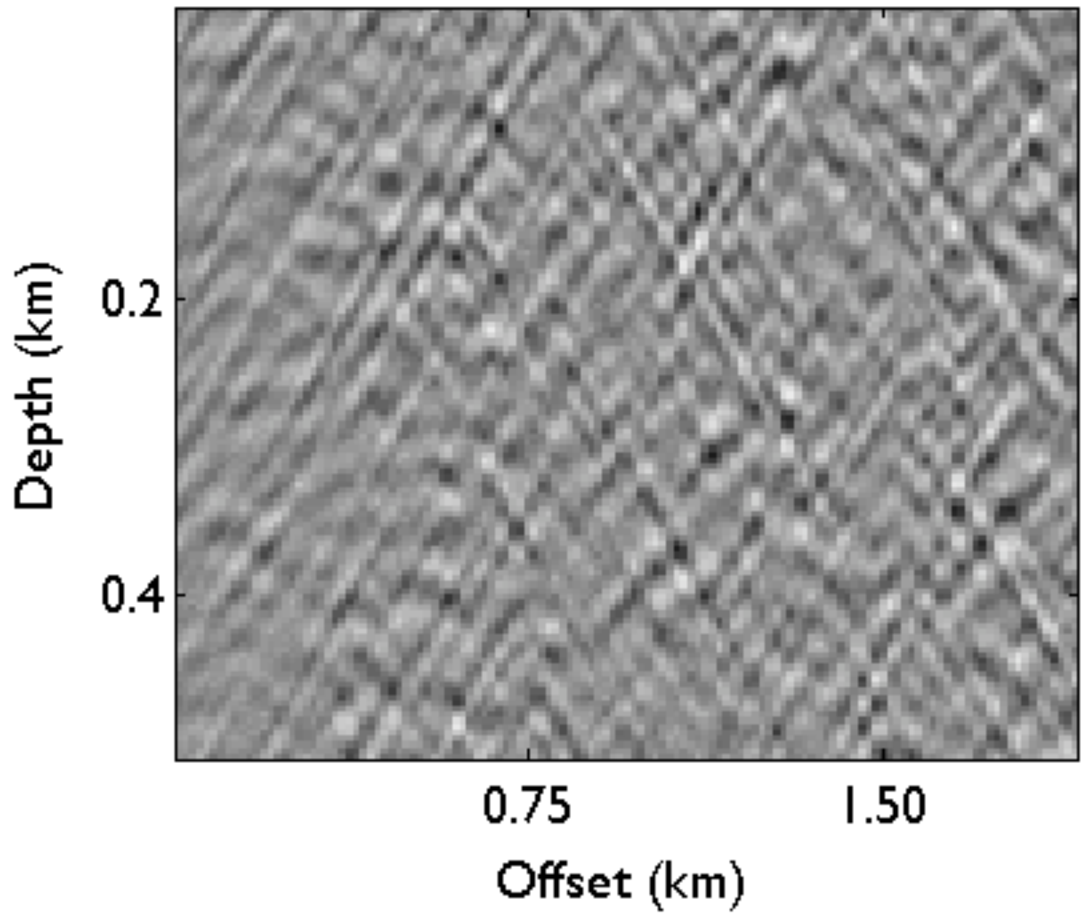
$$\mathbf{y} = \underline{\mathbf{y}}$$

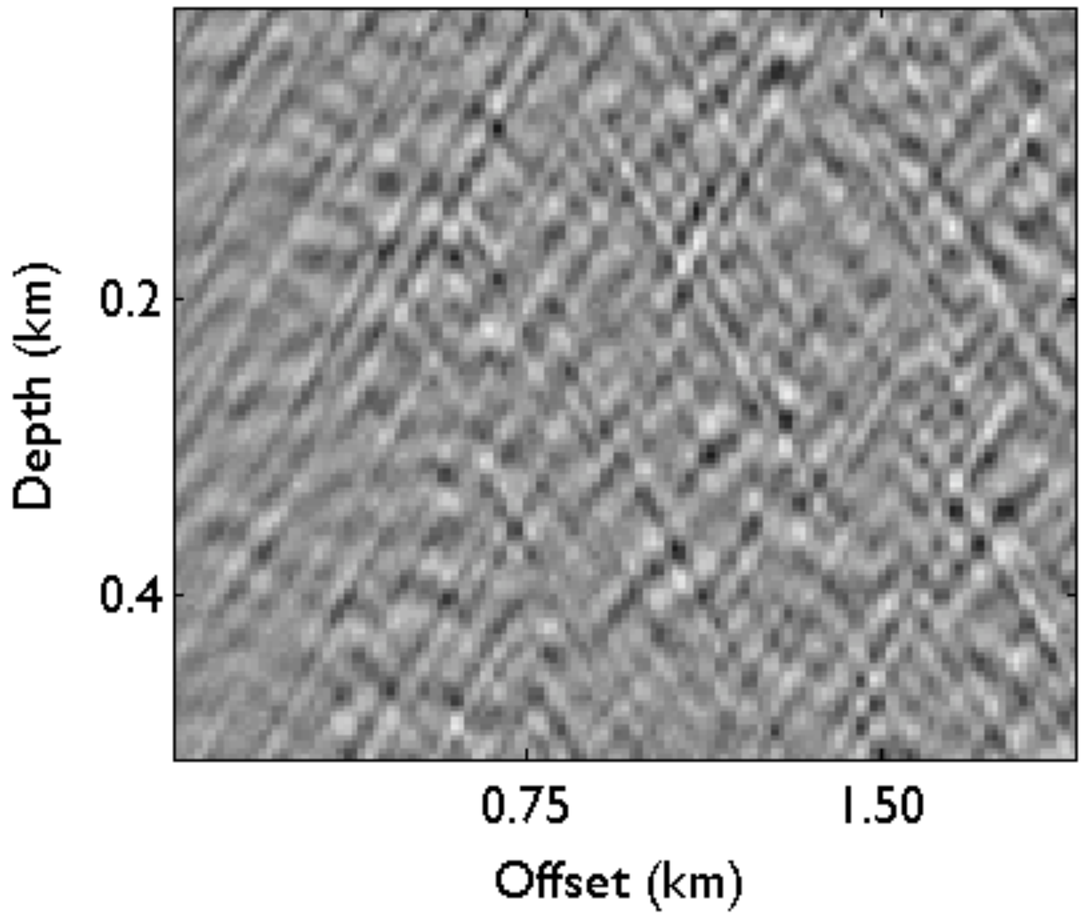
Experiment



$\mathbf{y} = \mathbf{RMDU}$

$\underline{\mathbf{y}} = \underline{\mathbf{DU}}$





Seismic Laboratory for Imaging and Modeling

CS

Current paradigm: *Nyquist sampling*

Pro:

- linear
- signal independent (aside from Nyquist frequency)

Con:

- cost dependent on the Nyquist frequency and model size
- overly pessimistic for signals with *structure*

New paradigm: *Compressive sensing*

Pro:

- cost dependent on signal's *complexity*

Con:

- solve a nonlinear recovery problem

Can lead to reduced cost when recovery cost < reduced simulation costs ...

SLIM
Seismic Laboratory for Imaging and Modeling

CS

$$
\mathbf{P_1}: \quad \begin{cases} \mathbf{y} & = \mathbf{RMf} \\ \tilde{\mathbf{x}} & = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{Ax} = \mathbf{y} \\ \mathbf{A} & = \mathbf{RMS}^* \\ \tilde{\mathbf{f}} & = \mathbf{S}^* \tilde{\mathbf{x}} \end{cases}
$$

CS provides conditions under which P1 recovers f:

- selection of CS-matrix (Measurement & Restriction matrices)
- selection of sparsifying transform

Additional complications

- large-to-extremely large problem size
- projected gradient with root finding method ($\mathrm{SPG}\ell_1$, Friedlander & van den Berg, '07-'08)
- CS matrix has to lead to *physically* ***realizable*** source wavefield for modeling & acquisition

CS

Selection of the CS-matrix

- natural restriction in Fourier (**F**) with *importance* sampling in the temporal direction
- CS with Gaussian (**N**) matrix along shots => simultaneous sources
- assures *incoherence* with sparsifying transform

For each **simultaneous** shot, define different restrictions

$$\mathbf{RM} = \begin{bmatrix} \mathbf{R}_1^{\Sigma} \otimes \mathbf{R}_1^{\Omega} \\ \vdots \\ \mathbf{R}_{n_{s'}}^{\Sigma} \otimes \mathbf{R}_{n_{s'}}^{\Omega} \end{bmatrix} \otimes \left(\mathbf{N} \otimes \mathbf{F} \right)$$

yielding the reduced simulated data

$$\mathbf{y} = \underline{\mathbf{y}} = \mathbf{RMd},\ \mathbf{y} \in \mathbb{C}^{n'_d}$$

$$\text{with } n'_d = n'_f n'_s n_r \ll n_d = 2 n_f n_s n_r$$

CS

Selection of the sparsifying transform:

- wavelet transform is known to compresses seismic data [Donoho '99]
- successfully applied in MRI (reconstructions from incomplete Fourier data)[Lustig et. al. '07]

Define

$$\mathbf{S} = \mathbf{W} \otimes \mathbf{W} \otimes \mathbf{W}$$

Bottom line:

Computational gain of CS proportional to undersampling ratio

$$\frac{n'_d}{n_d} \text{ with } n'_d \approx 5 \times \#\{\mathcal{N}_\Omega \circ \mathcal{N}'_s\}$$

at the expense of solving a CS problem.

SLIM
Seismic Laboratory for Imaging and Modeling

Complexity analysis

Assume discretization size in each dimension is *n*, and

$$n_s = n_t = n_f = \mathcal{O}(n)$$

Time-domain finite differences:

- $\mathcal{O}(n^4)$ in 2-D
- large constants

Preconditioned Helmholtz (Riyanti '06):

- $\mathcal{O}(n^5) = n_f n_s n_{it} \mathcal{O}(n^2)$ with $n_{it} = \mathcal{O}(n)$ asymptotically
- small constants

Multilevel-Krylov preconditioned (Erlangga and Nabben 08')

- $\mathcal{O}(n^4) = n_f n_s n_{it} \mathcal{O}(n^2)$ with $n_{it} = \mathcal{O}(1)$
- small constants

SLIM
Seismic Laboratory for Imaging and Modeling

Complexity analysis cont'd

Cost sparsity promoting optimization problem dominated by matrix-vector products

- 3-D wavelets are $\mathcal{O}(n^3)$
- Gaussian projection $\mathcal{O}(n^3)$ per frequency
- **Cost** $\mathcal{O}(n^4)$, which does not lead to asymptotic improvement

Use fast transforms instead (e.g. Noiselets by Coifman '01)

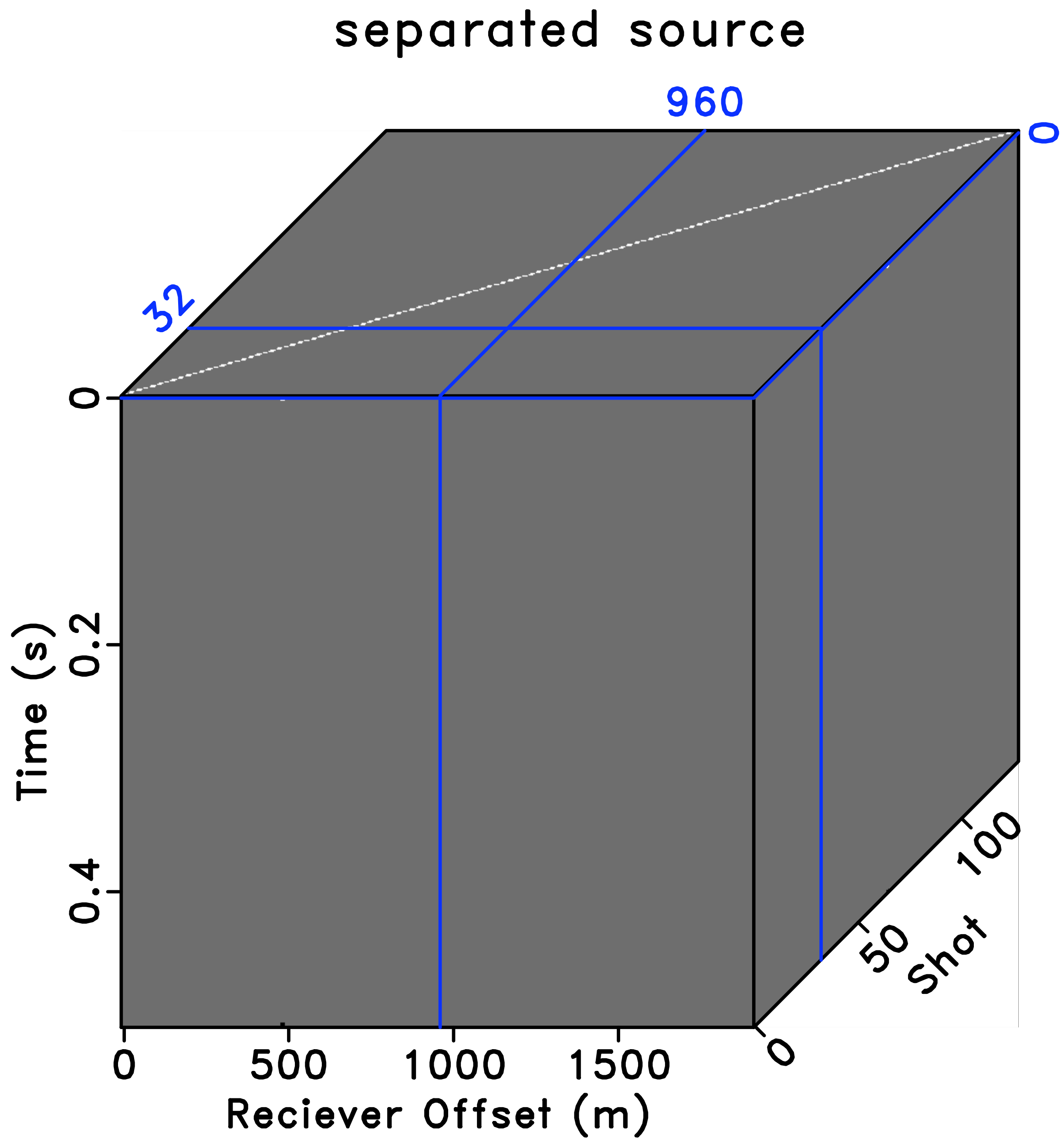
- fast projection in time & shot directions: $\mathcal{O}(n \log n)$
- **Cost** $\mathcal{O}(n^3 \log n)$ instead of $\mathcal{O}(n^4)$

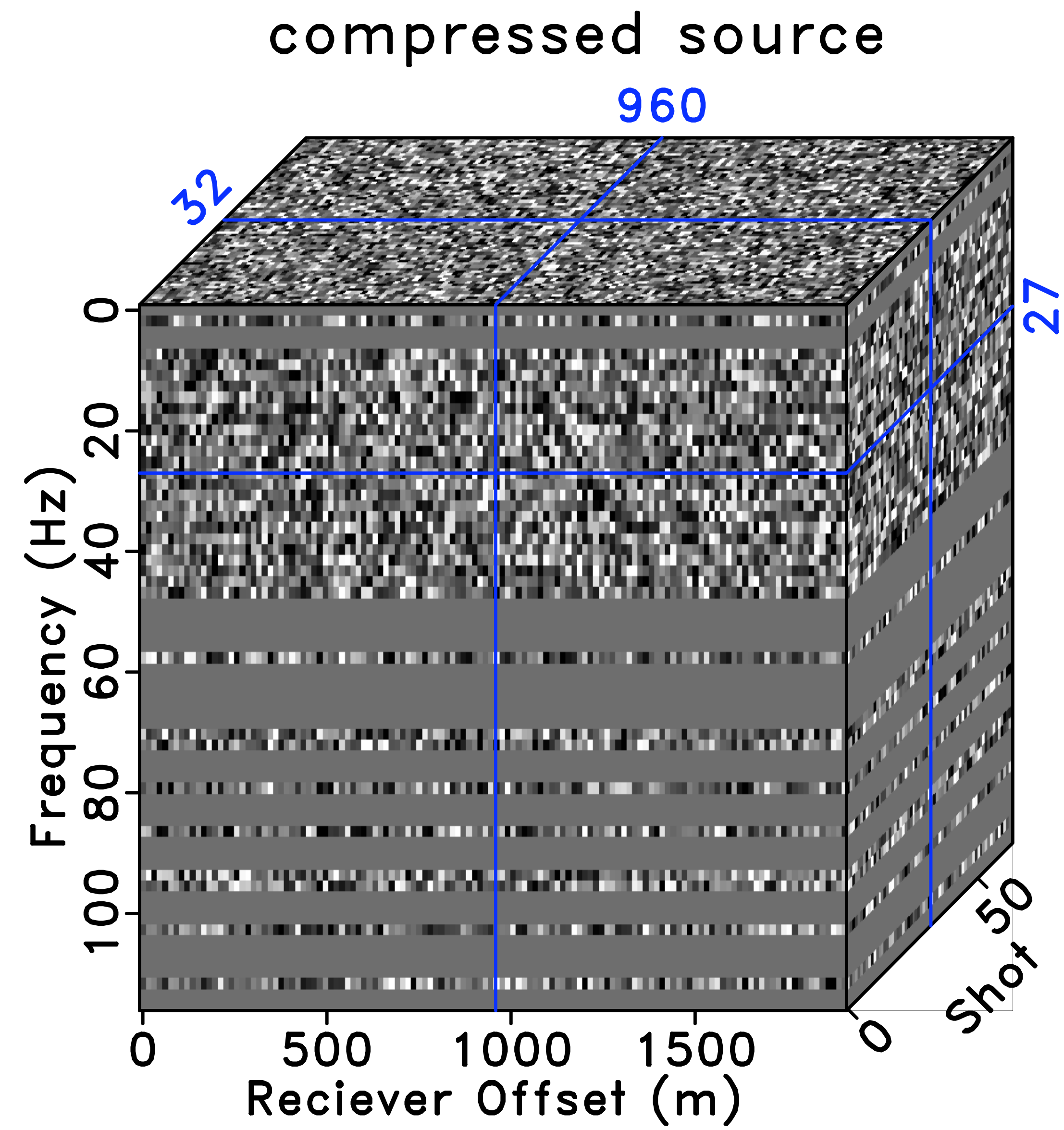
Bottom line: Computational cost for the ℓ_1-solver is less ($\mathcal{O}(n^3 \log n)$ vs. $\mathcal{O}(n^4)$) than the cost for solving Helmholtz

- smaller memory imprint
- smaller data volume requirement
- cost reduction dependent on complexity

SLIM
Seismic Laboratory for Imaging and Modeling

Source wavefields

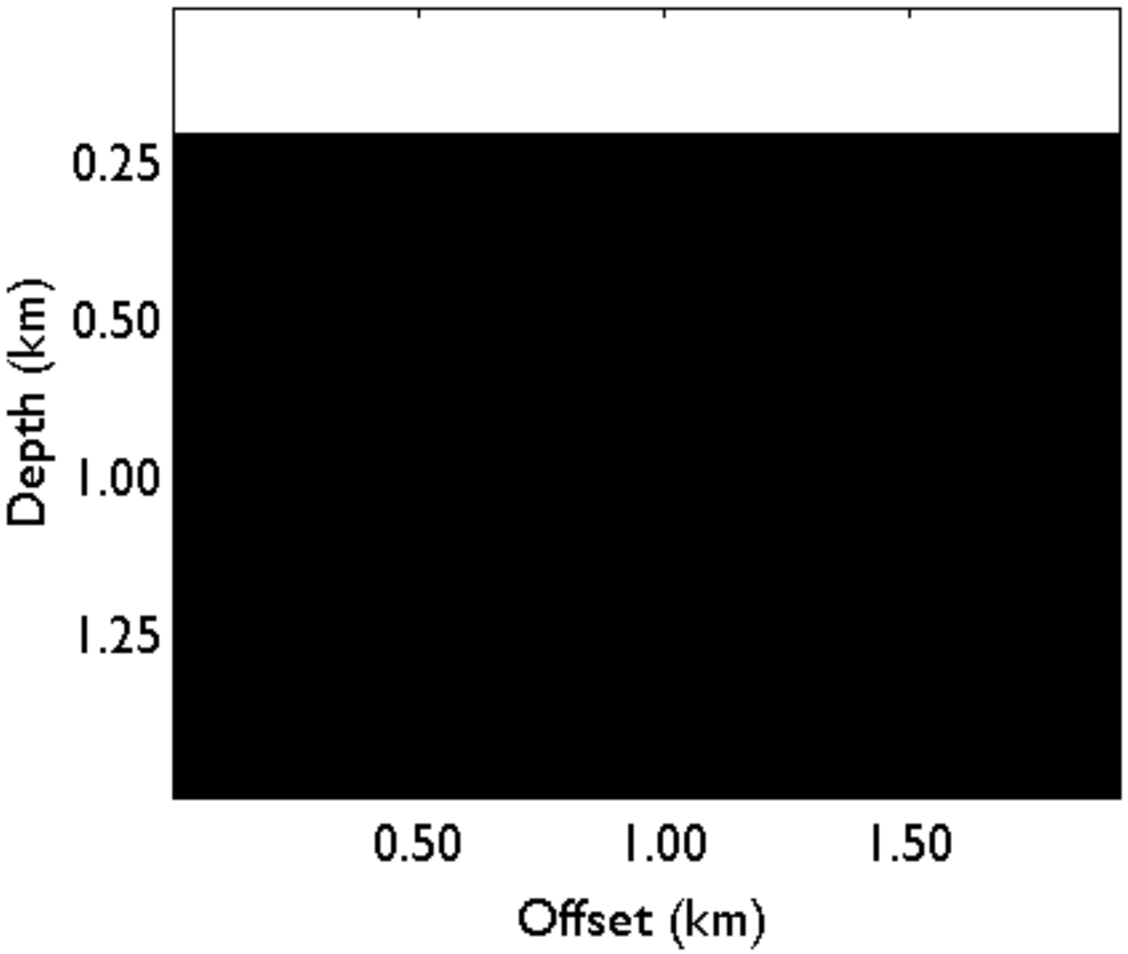




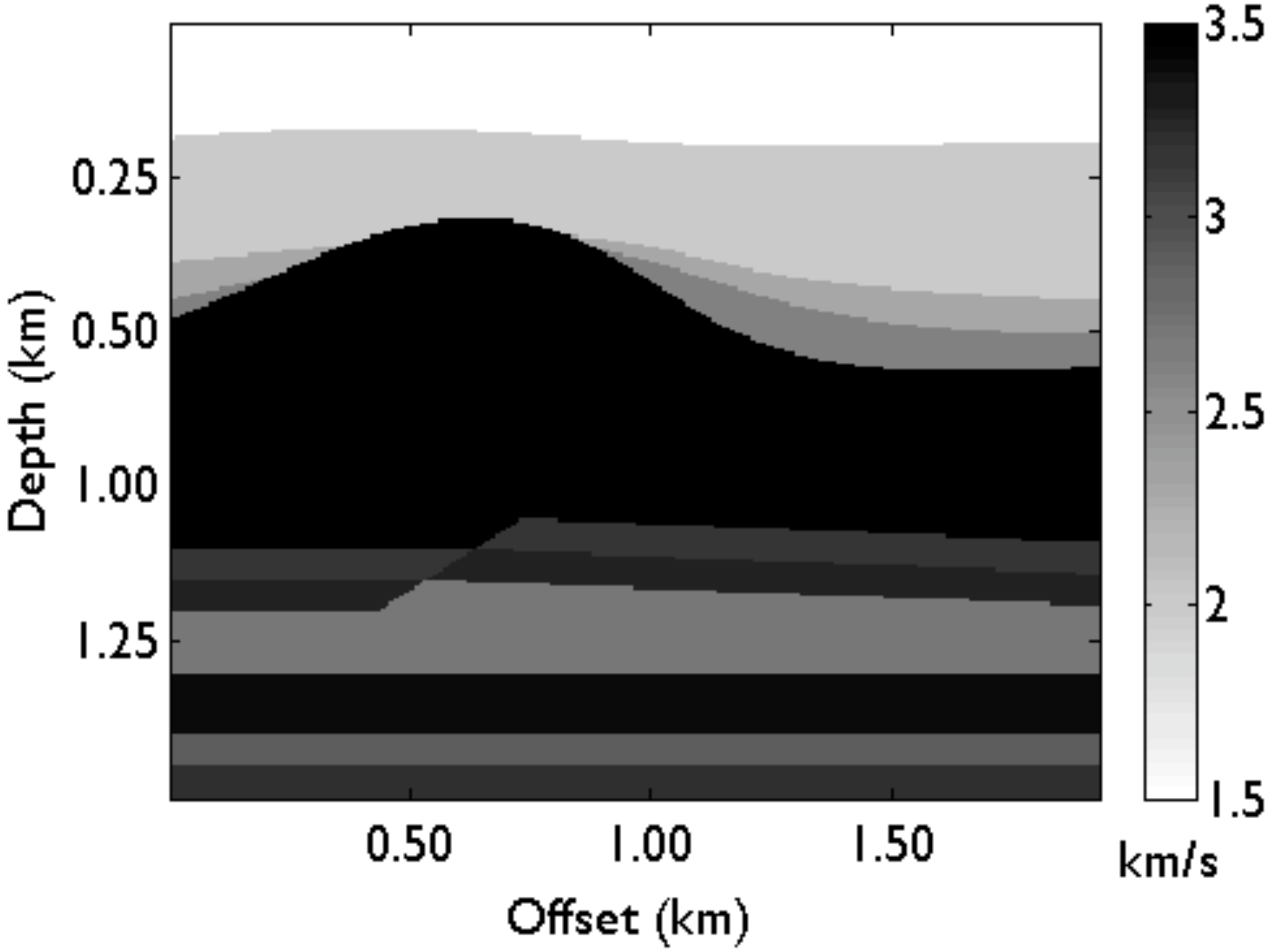
Freq sample 50%
Shot sample 50%
total sample 25%

Velocity models

simple model

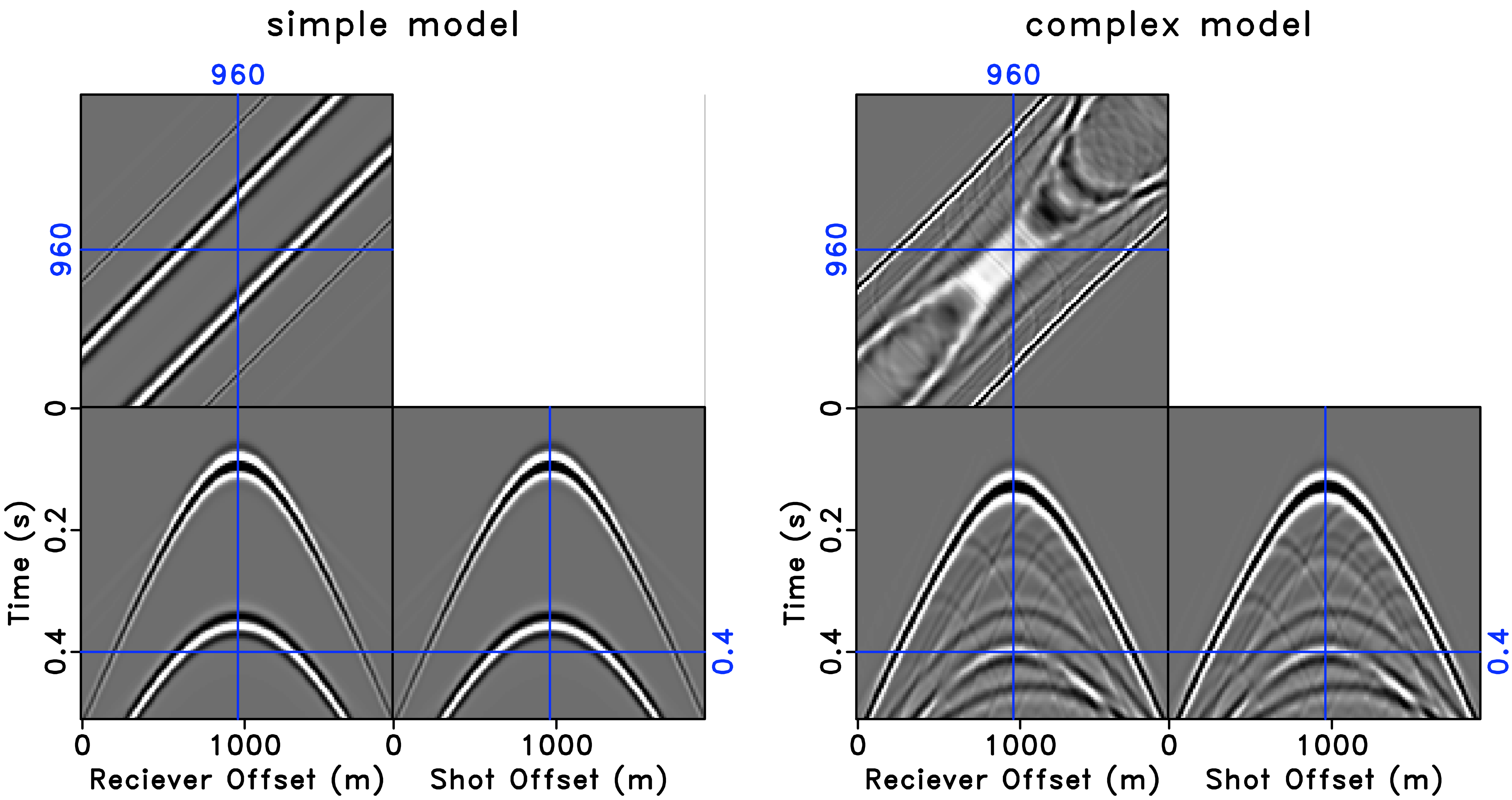


complex model

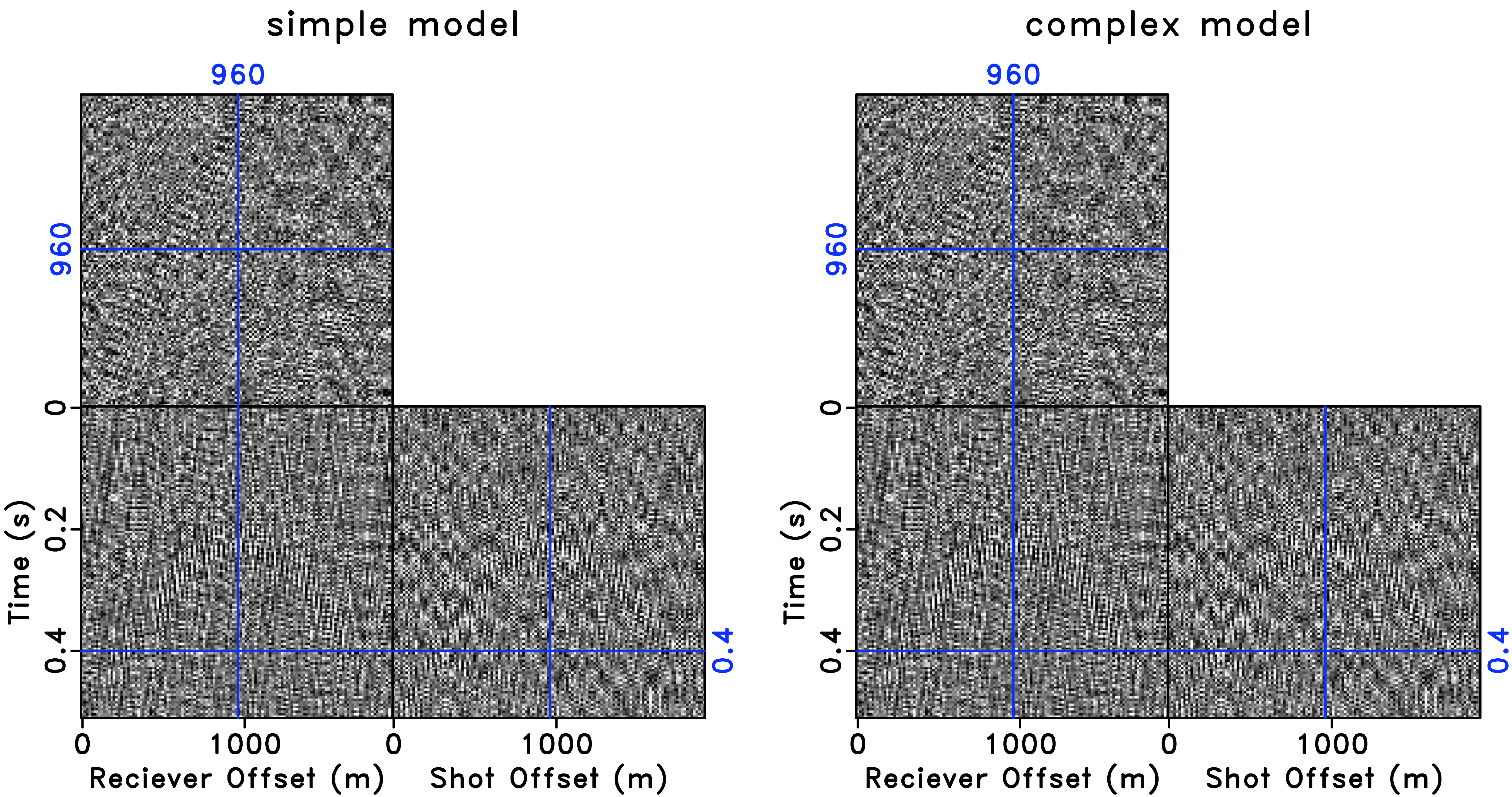


SLIM
Seismic Laboratory for Imaging and Modeling

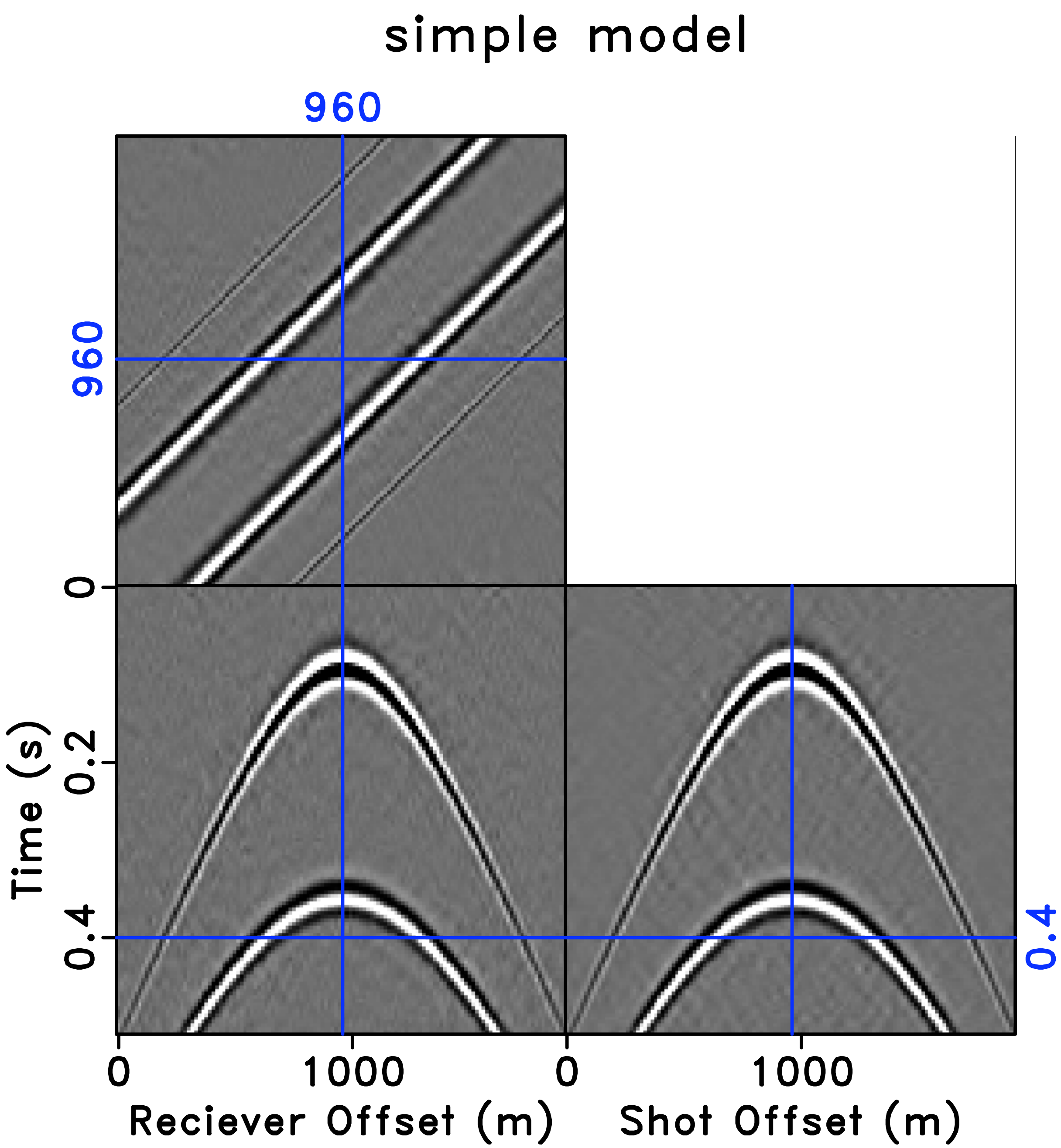
Green's functions



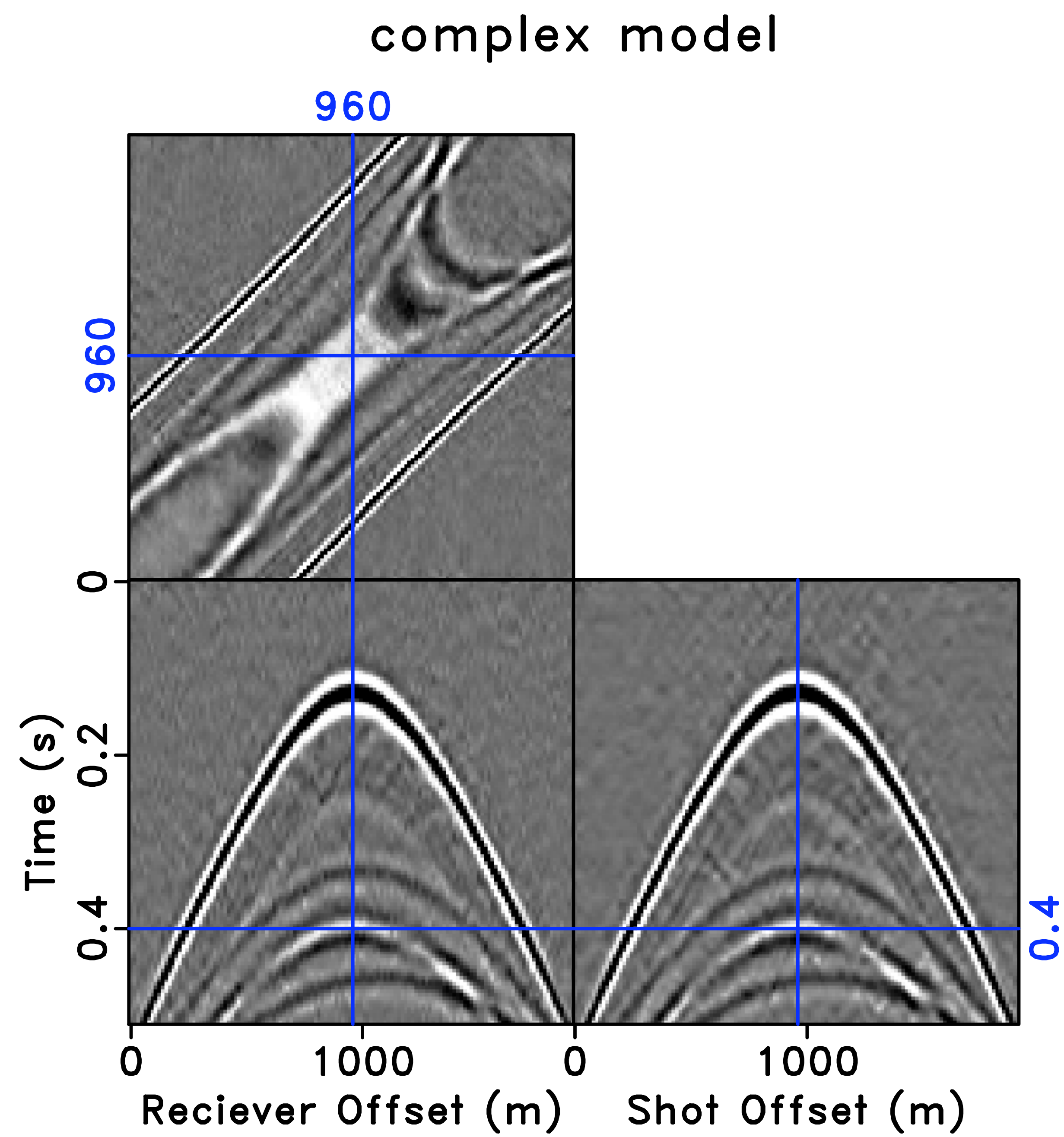
Matched filter



Recovered data



16.9dB

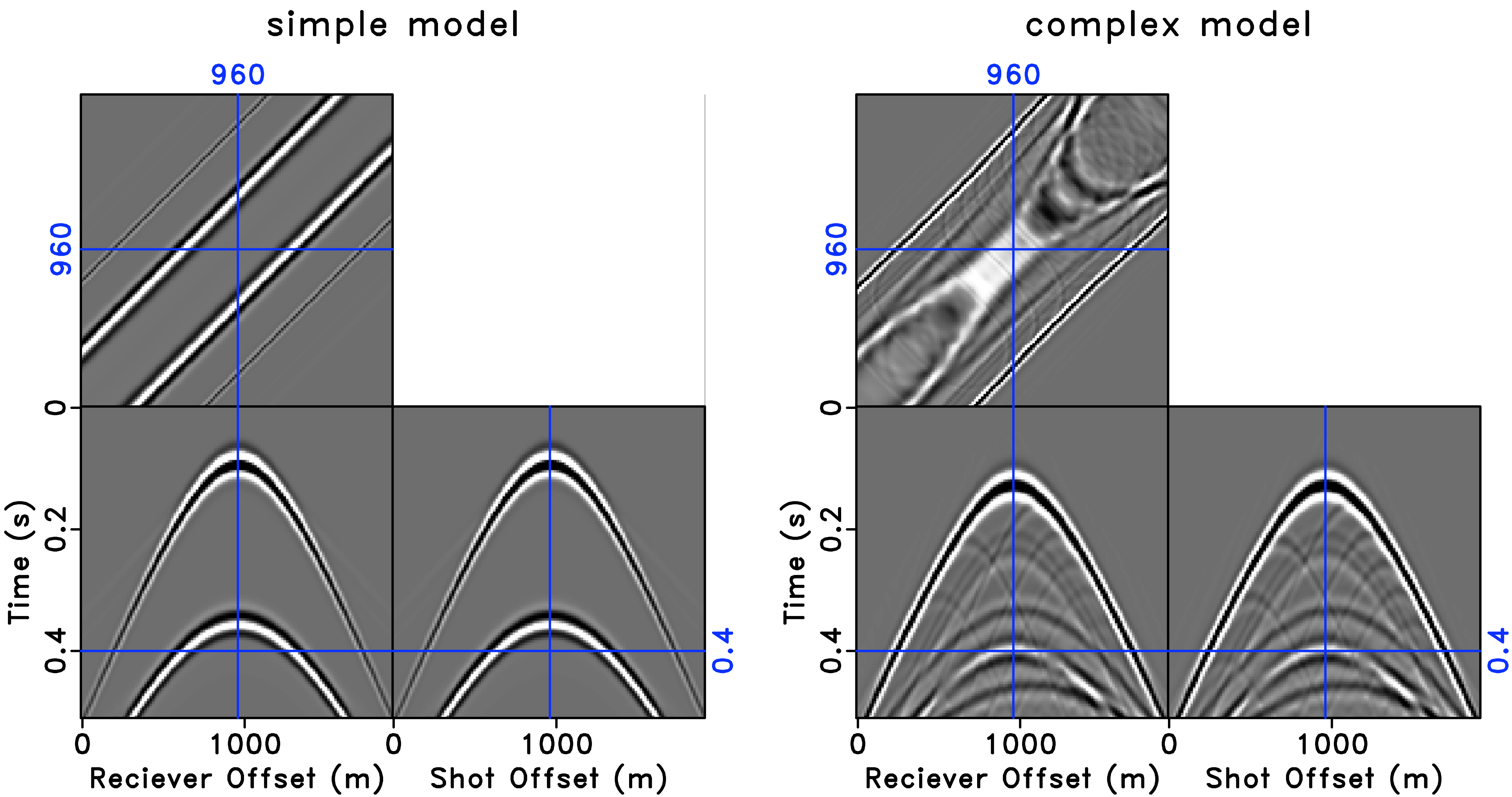


14.3dB

BP Solution ~3000 SPGL1 iteration

SLIM
Seismic Laboratory for Imaging and Modeling

Green's functions



SLIM
Seismic Laboratory for Imaging and Modeling

Sample ratio SNR (dB)

problem size 2^{21}

Total computed data fraction

# Frequencies / # Shots	0.25	0.15	0.07
2	9.3	7.0	4.3
1	13.7	9.2	3.7
0.5	11.6	7.4	3.4

$$\mathrm{SNR} = -20\log\frac{\|\mathbf{d}-\tilde{\mathbf{d}}\|_2}{\|\mathbf{d}\|_2}$$

Discussion & extensions

Compressive samplings are ***cumulative***

- **more** simultaneous experiments **improve** recovery
- equivalent to longer simultaneous & continuous acquisition

and allow for **design** of beneficial insonifying **waveforms.**

Add sparsity-promoting prior to PDE constrained optimization problem:

$$\min_{\underline{\mathbf{U}}\in\underline{\boldsymbol{\mathcal{U}}},\,\mathbf{x}\in\boldsymbol{\mathcal{X}}} \frac{1}{2}\|\mathbf{y}-\underline{\mathbf{DU}}\|_2^2+\lambda\|\mathbf{x}\|_1 \quad \text{subject to} \quad \underline{\mathbf{L}}[\mathbf{S}^H\mathbf{x}]\underline{\mathbf{U}}=\underline{\mathbf{B}}$$

Unconstrained optimization problem:

$$\min_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \frac{1}{2}\|\mathbf{y}-\underline{\mathbf{F}}[\mathbf{x}]\|_2^2+\lambda\|\mathbf{x}\|_1 \;\; \text{with} \quad \underline{\mathbf{F}}[\mathbf{x}]=\underline{\mathbf{DL}}^{-1}[\mathbf{S}^H\mathbf{x}]\underline{\mathbf{B}}$$

Requires extension of projected gradient ℓ_1-solver to nonlinear forward map ...

SLIM
Seismic Laboratory for Imaging and Modeling

Conclusions

Confluence of Compressive sensing, Simultaneous acquisition/modeling, and Helmholtz preconditioners leads to a formulation where cost to compute/acquire Green's functions are

- no longer dependent on the problem **size** but on the **complexity** (=sparsity) of the wavefield
- computed/acquired with a **gain** in speed *proportional* to the **compression** rate of the wavefield & behavior CS matrix
- obtained with an **overhead** for the recovery problem that becomes *negligible* for **large** problem sizes.

Extends to other forward modeling operators.

Room for analyses.

Interesting

- link with simultaneous acquisition and source design
- outlook towards complexity-driven solutions to inversion problems.

Acknowledgments

E. van den Berg and M. P. Friedlander for *SPGL1* (www.cs.ubc.ca/labs/scl/spgl1) & *Sparco* (www.cs.ubc.ca/labs/scl/sparco)

Sergey Fomel for Madagascar (rsf.sf.net)

This work was carried out as part of the SINBAD project with financial support from the collaborative research & development (CRD) grant DNOISE (334810-05) funded by the Natural Science and Engineering Research Council (NSERC) and matching contributions from BG, BP, Chevron, ExxonMobil and Shell.

Thank you!

SLIM
Seismic Laboratory for Imaging and Modeling