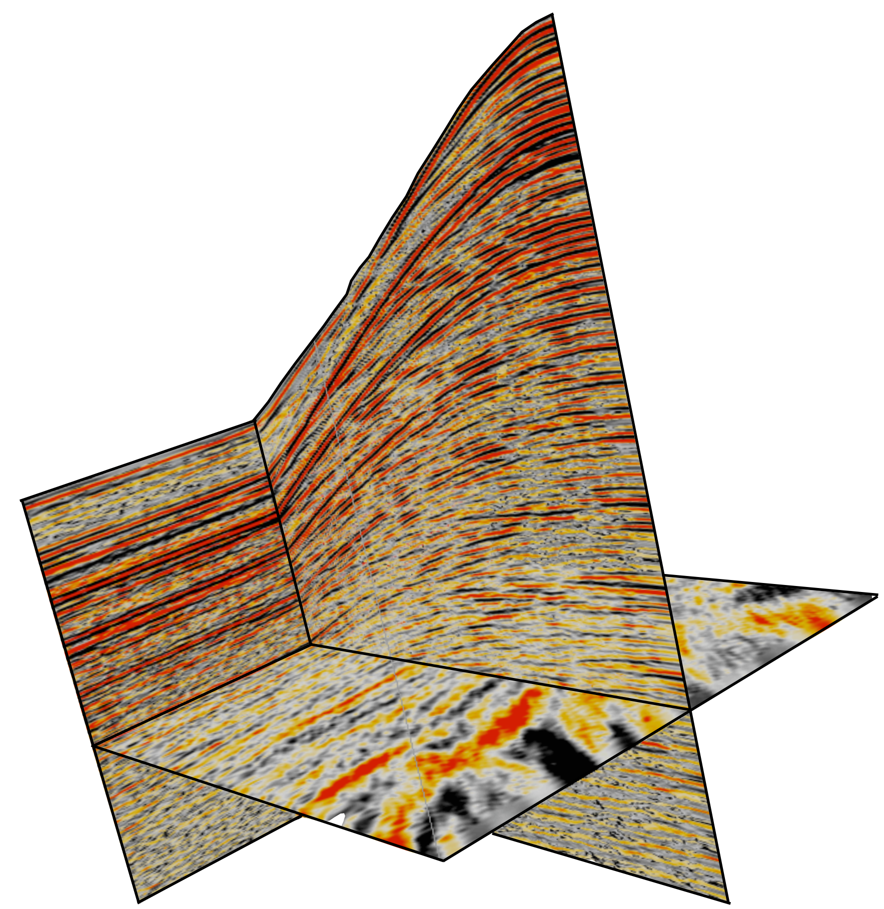
THE UNIVERSITY OF BRITISH COLUMBIA | VANCOUVER

Compressive-wavefield simulations

Felix J. Herrmann*
fherrmann@eos.ubc.ca

Joint work with Yogi Erlangga, and Tim Lin

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



Delft University of Technology
Delft, April 8, 2009

Seismic imaging & inversion

$$\min_{\mathbf{U}\in\mathcal{U},\,\mathbf{m}\in\mathcal{M}} \frac{1}{2}\|\mathbf{P}-\mathbf{D}\mathbf{U}\|_2^2 \quad \text{subject to} \quad \mathbf{H}[\mathbf{m}]\mathbf{U}=\mathbf{F}$$

$$\begin{aligned}
\mathbf{P} &= \text{Total multi-source and multifrequency data volume}\\
\mathbf{D} &= \text{Detection operator}\\
\mathbf{U} &= \text{Solution of the Helmholtz equation}\\
\mathbf{H} &= \text{Discretized multi-frequency Helmholtz system}\\
\mathbf{F} &= \text{Seismic sources}\\
\mathbf{m} &= \text{Unknown medium profile, e.g. } c^{-2}(x)
\end{aligned}$$

- **massive problem size**
- **non-uniqueness**

Adjoint state methods

For each ***separate*** source **p** solve the **unconstrained problem**

$$\min_{\mathbf{m}\in\mathcal{M}} \frac{1}{2}\|\mathbf{p}-\mathcal{F}[\mathbf{m}]\|_2^2 \qquad \text{with } \mathcal{F}[\mathbf{m}] = \mathbf{D}\mathbf{H}^{-1}[\mathbf{m}]\mathbf{f}$$

with ***model updates*** **= -** "***post-stack*" *image***,

$$\delta\mathbf{m} = \Re\left(\sum_{\omega}\sum_{s}\omega^2\mathbf{U}\odot\mathbf{V}^*\right)$$

involving solutions of the Helmholtz system

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

with

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

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* or *model*

- *New insights* in inversion using *extended* (prestack) image volumes
  - quadratic minimizers that promote ***focusing***
  - uniqueness result for migration velocity analysis using *differential semblance*

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
[Morton, S. A. and C. C. Ober, ‘98, 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]

- Focusing combining results from **migration-velocity analysis** joint **sparsity promotion**
[Shen, P. and W. W. Symes, ’08, van den Berg and Friedlander, ‘08]

Today's agenda

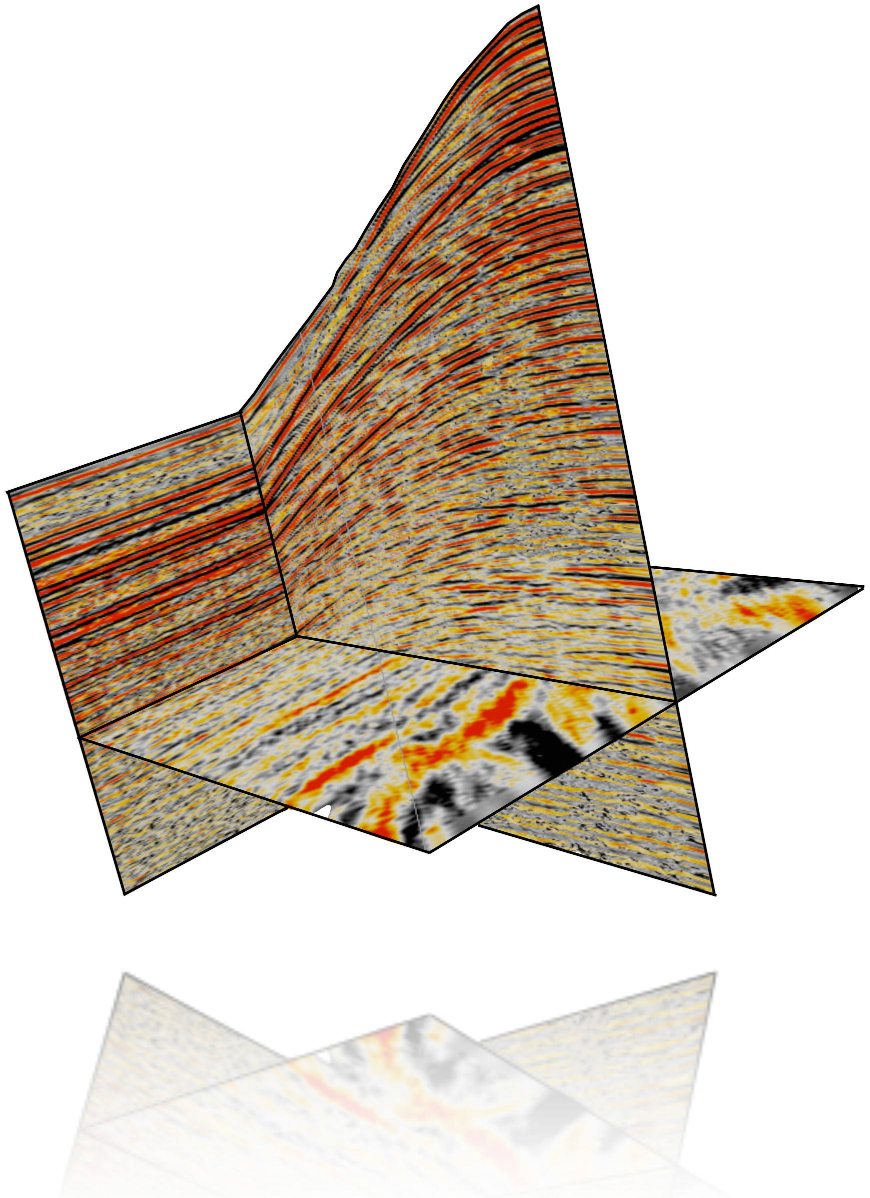
- Brief introduction of *compressive sensing*
  - sparsifying transforms
  - randomization
  - nonlinear recovery by sparsity promotion

- CS applied to ***explicit*** one-way wavefield computations
  - use of the ***modal*** domain as the ***sampling*** domain
  - reduction of the number of eigenvectors & frequencies

- CS applied to ***implicit*** simultaneous full-waveform simulation
  - use ***simultaneous sources*** as the ***sampling*** domain
  - reduction of the number of right-hand sides & frequencies

- CS applied to ***explicit*** prestack imaging
  - leverage focusing
  - reduction of the model-space wavefields

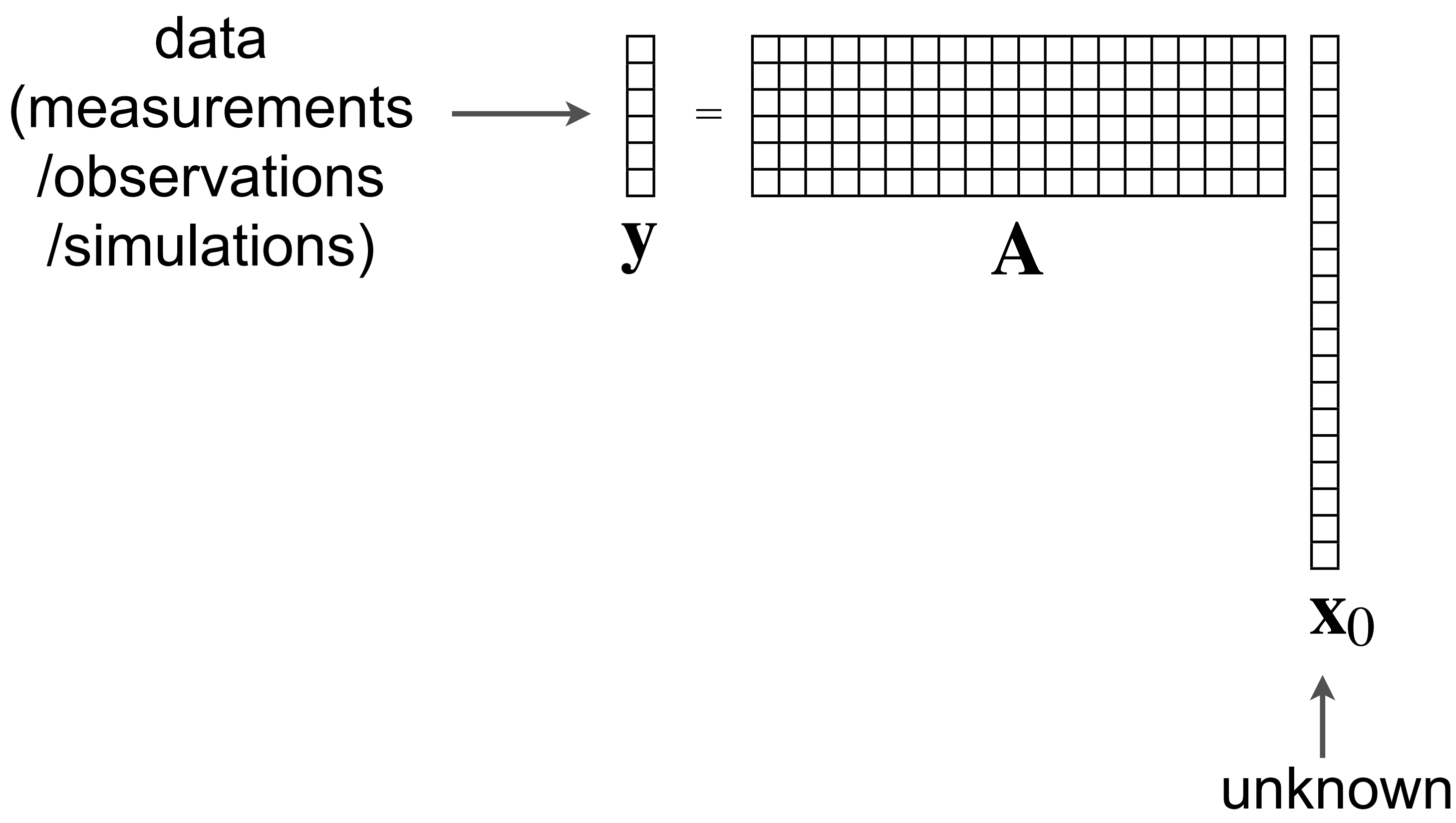
Compressive sensing



Delft University of Technology
Delft, April 8, 2009

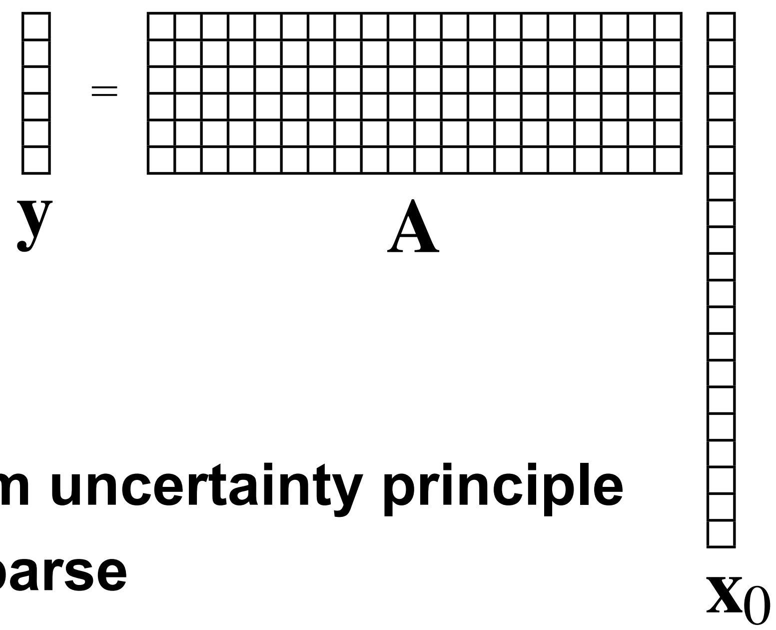
Problem statement

Consider the following (severely) underdetermined system of linear equations



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

Perfect recovery



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

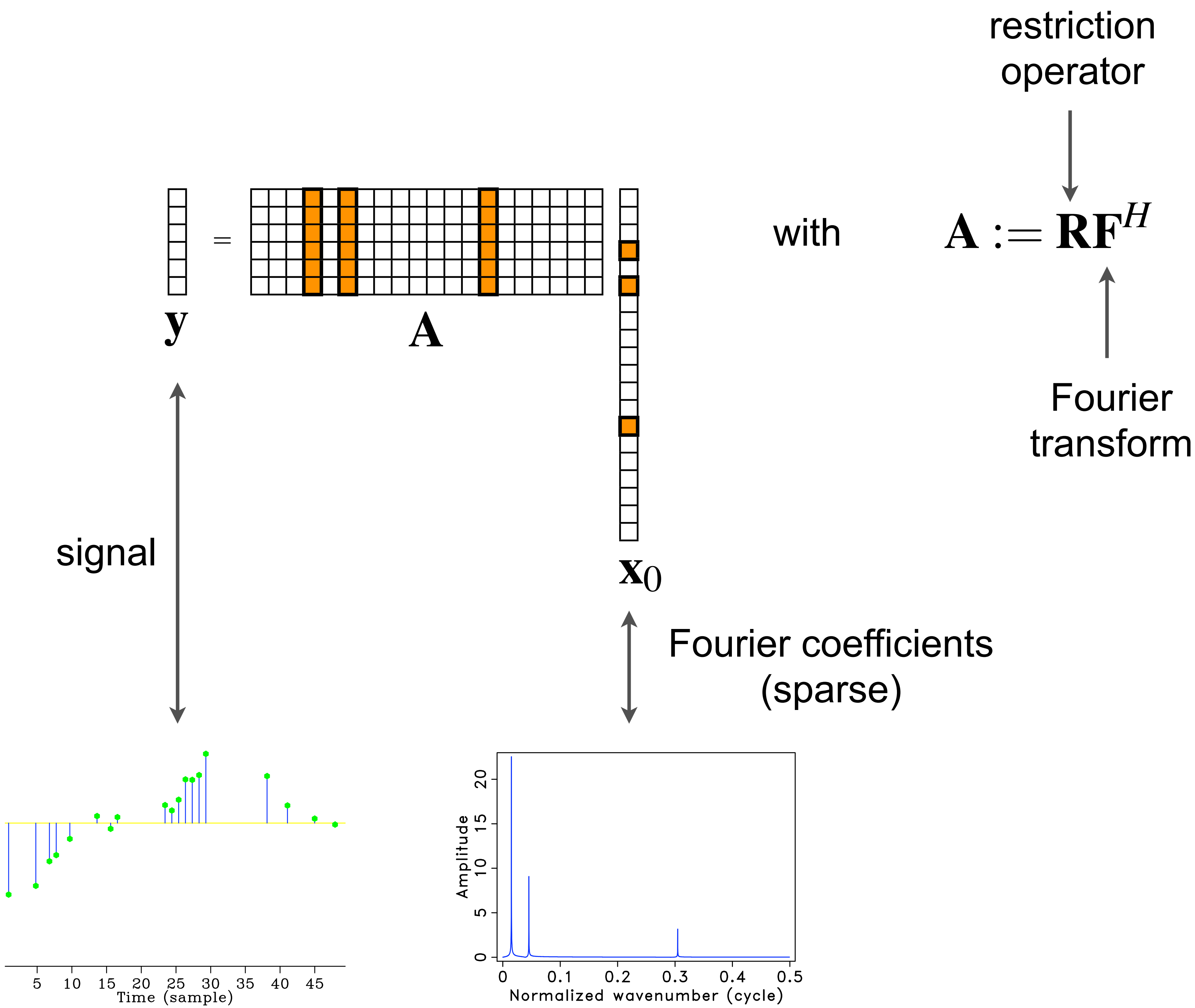
- procedure

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

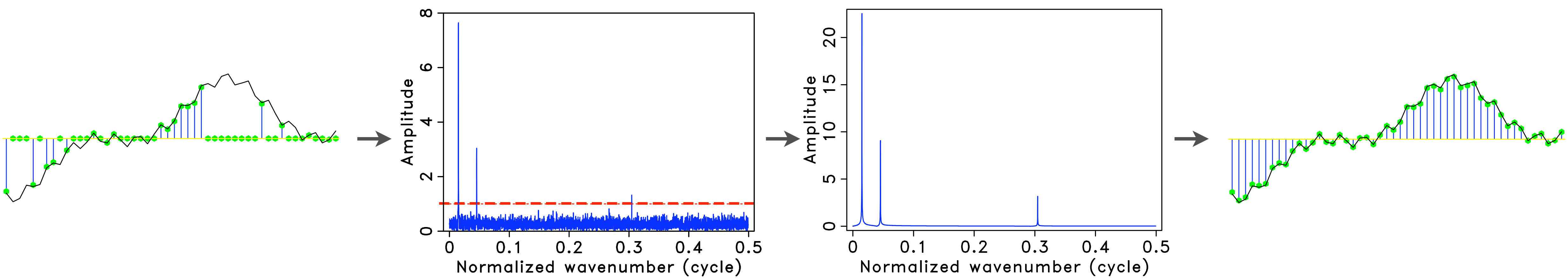
- performance
  - *S*-sparse vectors recovered from roughly on the order of *S* measurements (to within constant and *log* factors)

[Candès et al.'06]
[Donoho'06]

Simple example



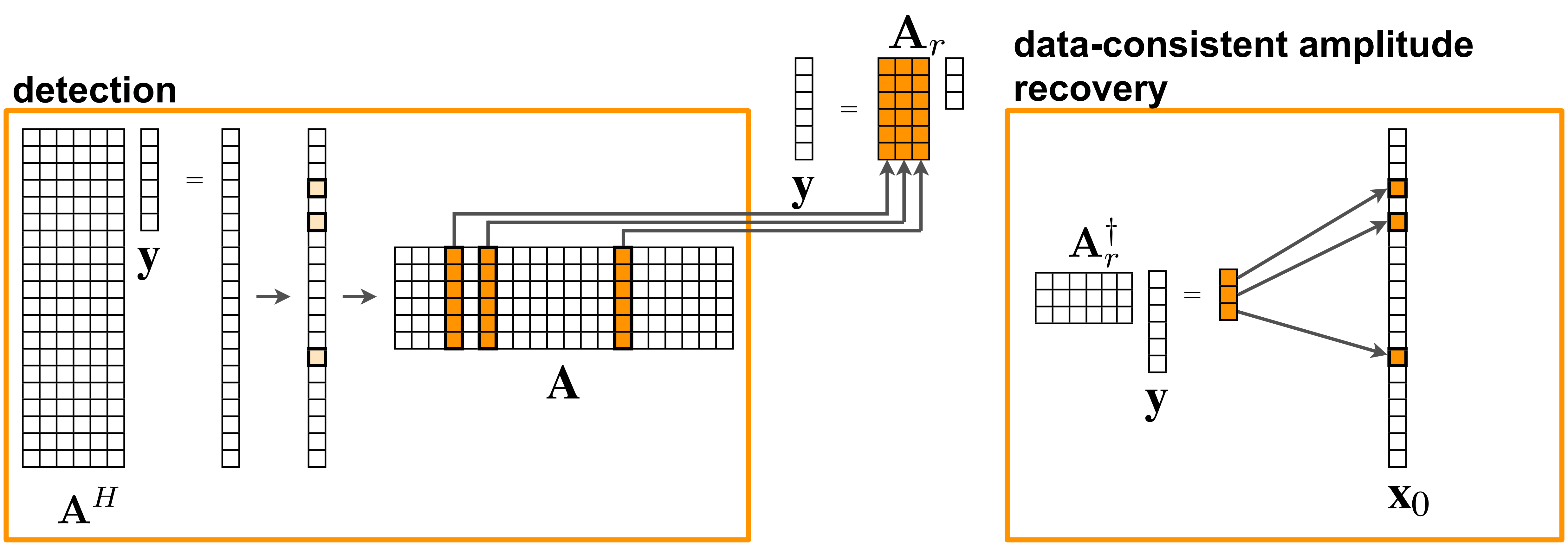
NAIVE sparsity-promoting recovery



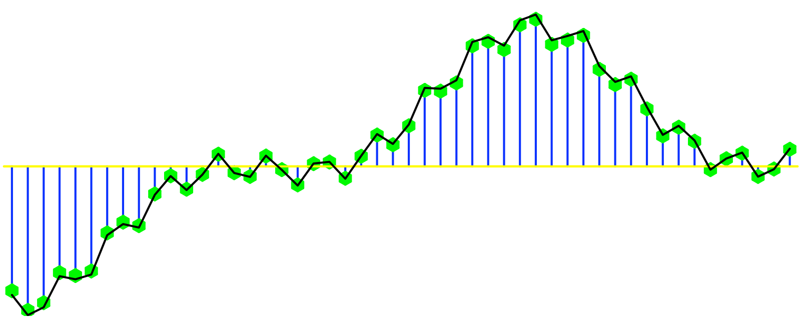
inverse Fourier transform

detection + data-consistent amplitude recovery

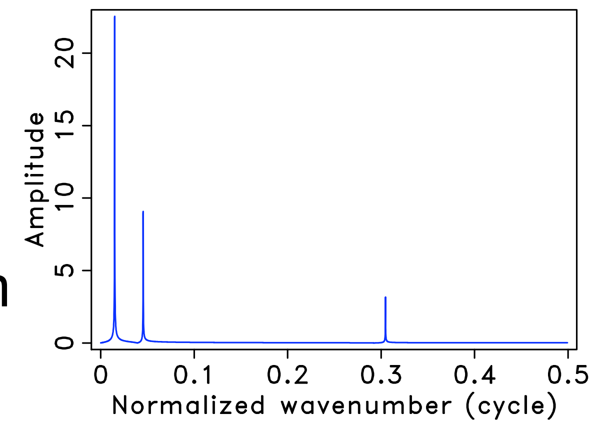
Fourier transform



Coarse sampling schemes

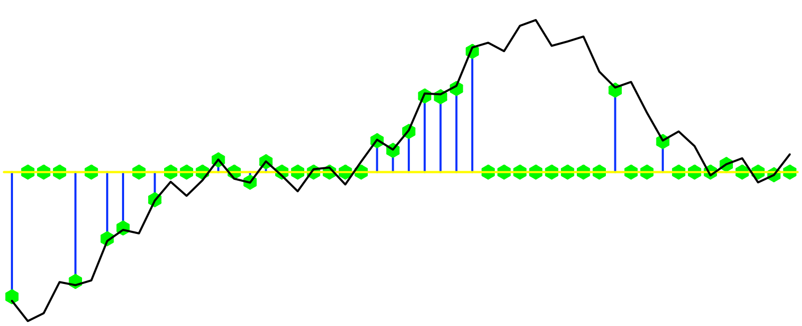


Fourier
→
transform

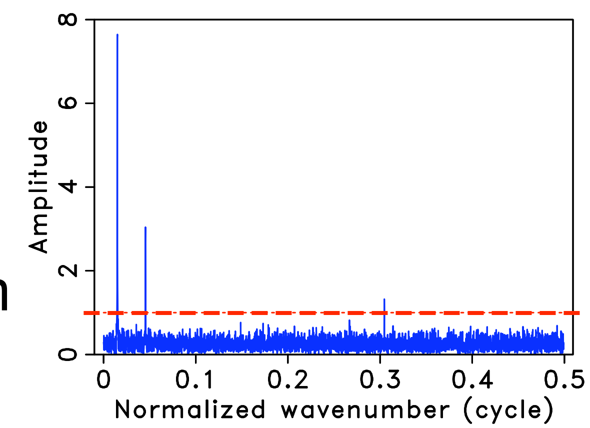


few significant coefficients

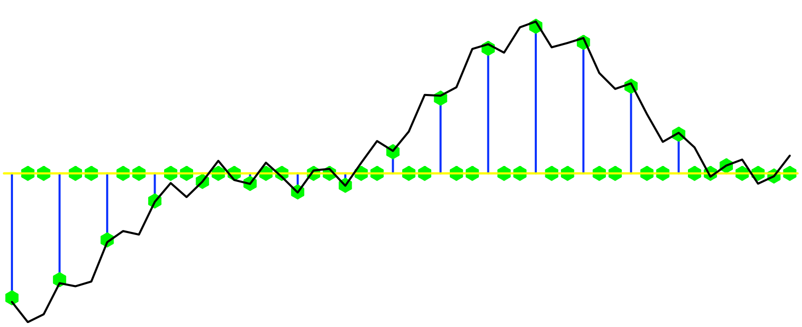
3-fold under-sampling



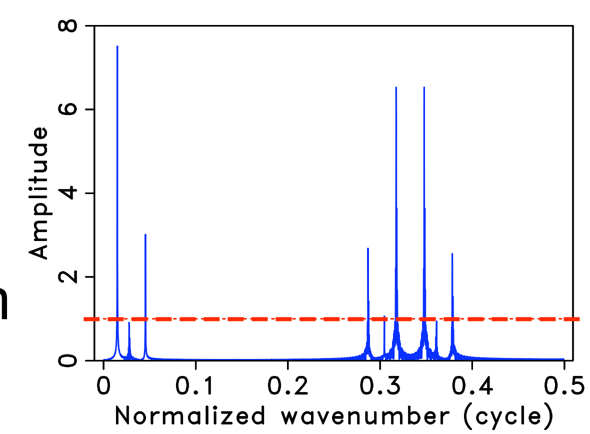
Fourier
→
transform



✓ significant coefficients detected



Fourier
→
transform



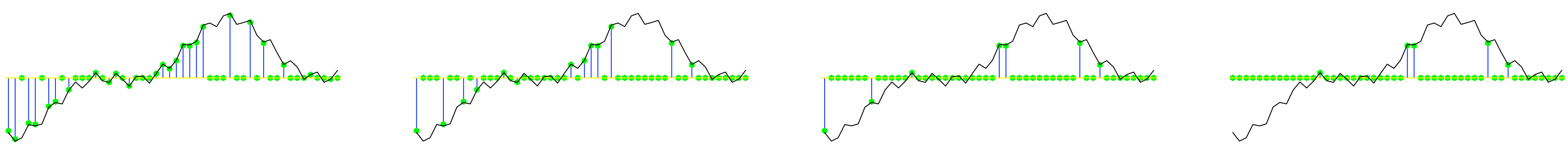
✗ ambiguity

[Hennenfent and Herrmann‘08]

Undersampling “noise”

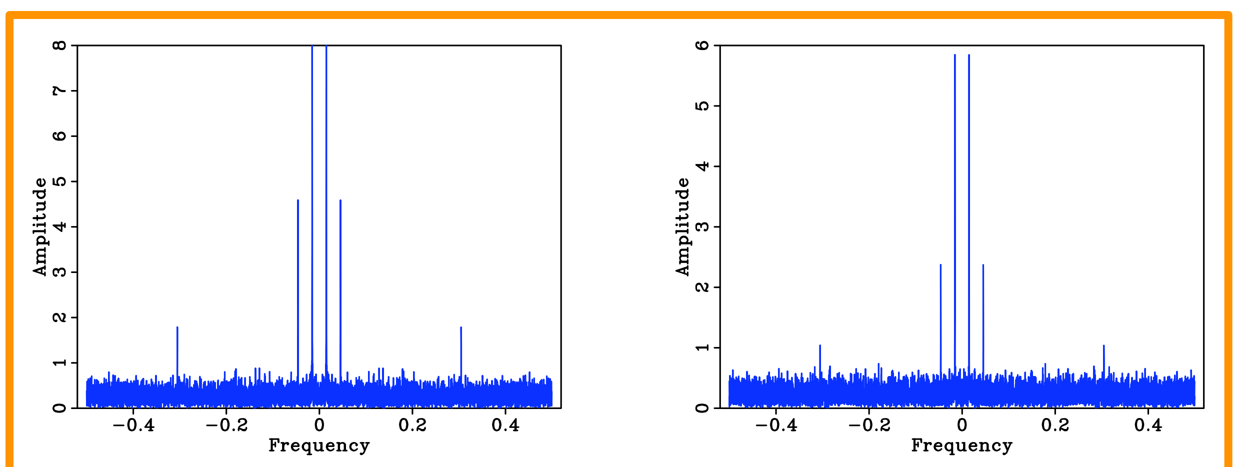
- “noise”
  - due to $\mathbf{A}^H\mathbf{A} \neq \mathbf{I}$
  - defined by $\mathbf{A}^H\mathbf{A}\mathbf{x}_0-\alpha\mathbf{x}_0 = \mathbf{A}^H\mathbf{y}-\alpha\mathbf{x}_0$

1 out of 2 | 1 out of 4 | 1 out of 6 | 1 out of 8

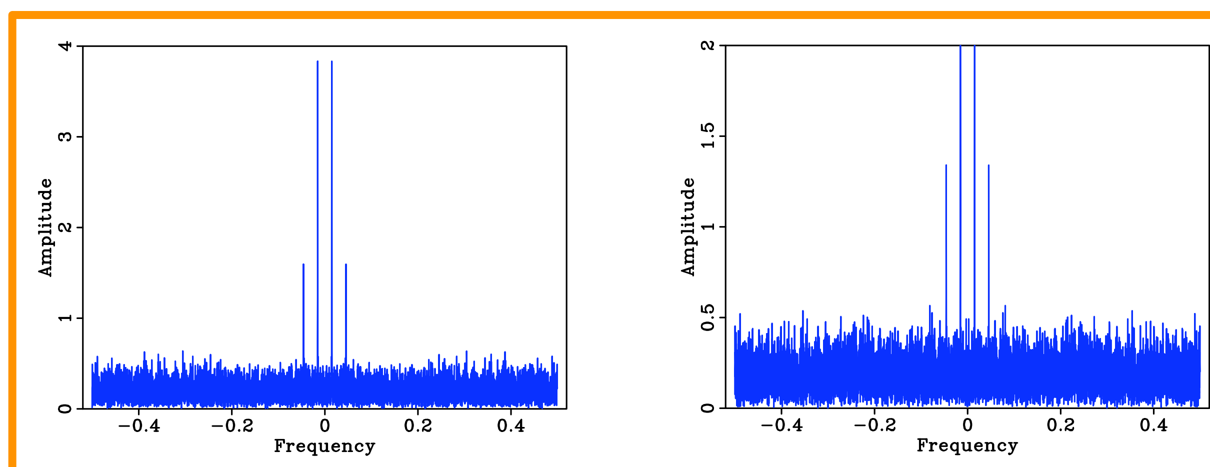


less acquired data

3 detectable Fourier modes



2 detectable Fourier modes



Extensions

Incomplete and noisy measurements:

$$\mathbf{y} = \overbrace{\mathbf{R}}^{\text{Restriction}} \underbrace{\mathbf{M}}_{\text{Measurement}} \mathbf{m} + \mathbf{n}$$

y *incomplete* (compressively sampled) and noisy data

m the *unknown* model

M "**arbitrary**" *measurement* matrix

- Fourier, eigenfunctions
- simultaneous sources

R *restriction* matrix

n Gaussian noise

Extensions

- Use CS principles to select appropriate
  - restriction matric **R**
  - measurement basis **M**
  - sparsifying transform **S**
  - driven by signal type, physics, and type of acquisition (e.g. fMRI vs seismic)

  } **subsampled randomized phase encoder**

- Sparse signal representation:

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0$$

with

$$\mathbf{A} = \mathbf{R}\mathbf{M}\mathbf{S}^H$$

restriction matrix **measurement matrix** **sparsity matrix**

Selection is aimed at turning *aliases/coherent subsampling artifacts*

Extensions

- According to CS theory (valid for orthonormal bases for **M** & **S**) recovery depends on ***restriction***, ***mutual coherence***, and ***sparsity***

- ***Mutual coherence*** between **M** & **S** (off-diagonals of Gramm matrix),

$$\mu(\mathbf{M},\mathbf{S}) = \sqrt{m} \max_{(k,l)\in[1\cdots m]\times[1\cdots m]} |\langle m_k,\ s_l\rangle|$$

  and appropriate subsampling
  - controls leakage
  - maps interferences into noise
  - importance sample in the band

- ***Compressibility***, i.e.,

$$|\mathbf{x}_{i\in I}| \leq Ci^{-r},\ r \geq 1 \quad \text{and} \quad x_{I(1)} \geq x_{I(2)} \geq \cdots \geq x_{I(m)}$$

Extensions [Donoho, Candes, Tao, and Romberg, '06]

- If the number of samples *n* of and unknown *S*-sparse m-length signal

$$n \propto \mu^2 \cdot S$$

then recovery according to

$$(P_1): \begin{cases} \min_{\mathbf{x}\in\mathbb{R}^m} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2 \leq \epsilon \\ \hat{\mathbf{m}} = \mathbf{S}^H\hat{\mathbf{x}} \end{cases}$$

yields

$$\|\hat{\mathbf{x}}-\mathbf{x}_0\|_2 \leq C_3 \cdot \epsilon + C_4 \cdot S^{-r+1/2}.$$

- Recovery
  - within the noise level ϵ
  - determined by the compression rate *r*, the higher the rate the better the recovery
  - for compressible signals it is as if CS recovers the first *S* largest entries
  - true miracle compressive sensing gives you a nonlinear approximation at a cost of a logarithmic oversampling

Strategy

- Leverage CS towards solutions of wave simulation & imaging problems

- Subsample solution deliberately, followed by CS recovery

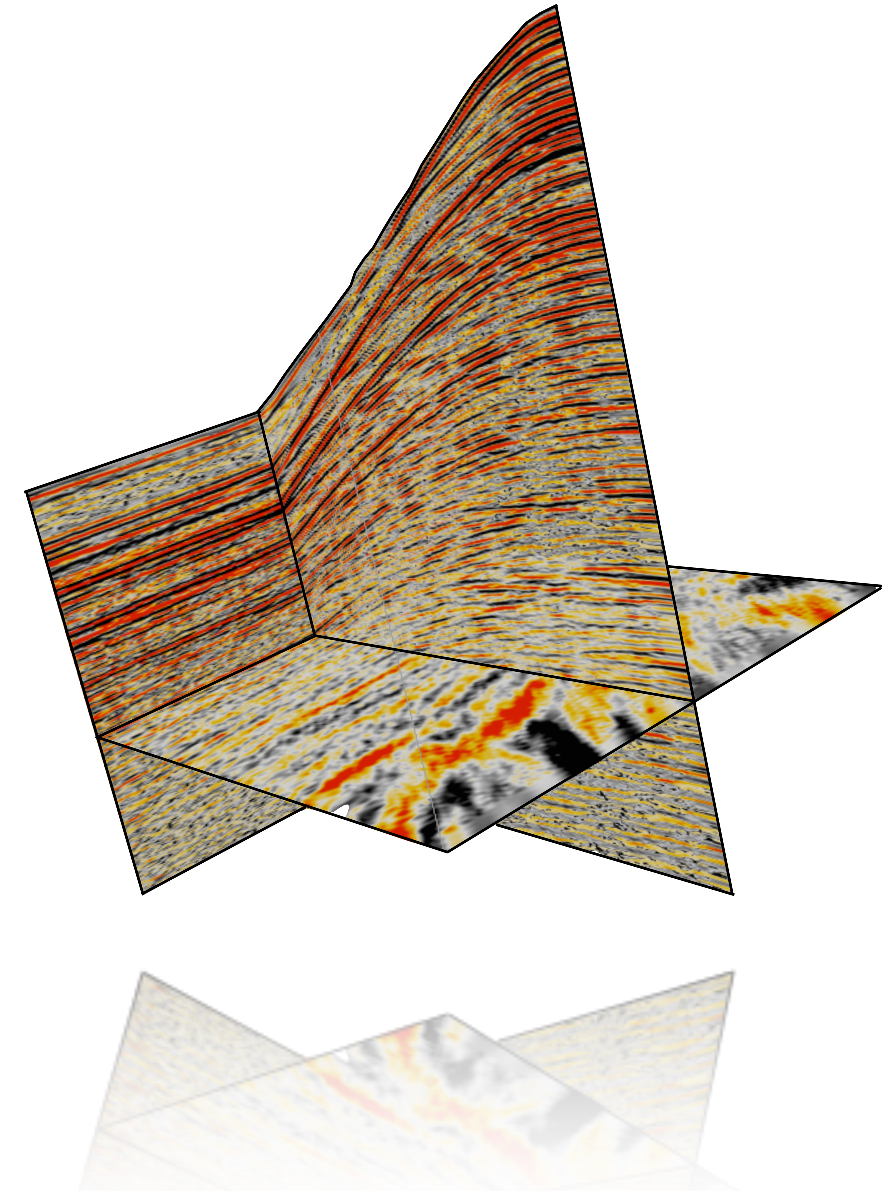
- Speedup if recovery costs < gain in reduced system size
  - computation
  - storage

- Examples
  - compressed explicit one-way wave propagation with CS sampling in the modal domain
  - compressed implicit forward modeling by CS sampling in the data space via simultaneous sources
  - compressed imaging by CS sampling in the model space

Compressive one-way wavefield extrapolation

T. T. Y. Lin and F. J. Herrmann. Compressed wavefield extrapolation. Geophysics, 72(5):SM77–SM93, 2007.

L. Demanet and G. Peyre. Compressive wave computation, 2008. Stanford. Submitted for publication.



Delft University of Technology
Delft, April, 2009

Motivation

- Syntheses of the discretized operators form bottle neck of imaging
- Operators have to be applied to multiple right-hand sides
- Explicit operators are feasible in 2-D and can lead to an order-of-magnitude performance increase in performance
- Extension towards 3-D problematic
  - storage of the explicit operators
  - convergence of implicit time-harmonic approaches
- First go at the problem using CS techniques to compress the operator by subsampling the spectral representation of the operator ...

Basic idea

- Compute one-way wavefields from an incomplete set of eigenfunctions of the Helmholtz equation
- Recover the solution by solving a CS recovery problem
- Based on an incoherence argument
  - between eigenfunctions and the representation in which the solution is sparse
- Formal proofs derived by Demanet and Peyre, ‘08 for 1-D media
  - interestingly medium needs to be BV

One-Way Wave Operator

- Solution of the one-way wave equation

$$\mathcal{W}(x_3; x_3') = \exp(-j(x_3 - x_3')\mathcal{H}_1)$$

- After discretization solve eigenproblem on $\mathbf{H}_2$

$$\mathbf{H}_2 = \begin{bmatrix} \left(\frac{\omega}{\bar{c}_1}\right)^2 & 0 & \cdots & 0 \\ 0 & \left(\frac{\omega}{\bar{c}_2}\right)^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\frac{\omega}{\bar{c}_{n_1}}\right)^2 \end{bmatrix} + \mathbf{D}_2$$

  - Helmholtz operator is Hermitian
  - monochromatic
  - velocity $\bar{c}$ varies laterally

(Claerbout, 1971; Wapenaar and Berkhout, 1989)

Modal transform

- Solve eigenproblem & take square root

$$\mathbf{H}_1 = \mathbf{L}\boldsymbol{\Lambda}^{1/2}\mathbf{L}^H$$

- $\mathbf{L}$ is orthonormal & defines the modal transform that diagonalizes one-way wavefield extrapolation
- Eigenvalues play role of vertical wavenumbers
- Extrapolation operator is diagonalized

$$\mathbf{W} = \boldsymbol{\mathcal{F}}^H \mathbf{L} e^{-j\boldsymbol{\Lambda}^{1/2}(x_3 - x_3')} \mathbf{L}^H \boldsymbol{\mathcal{F}}$$

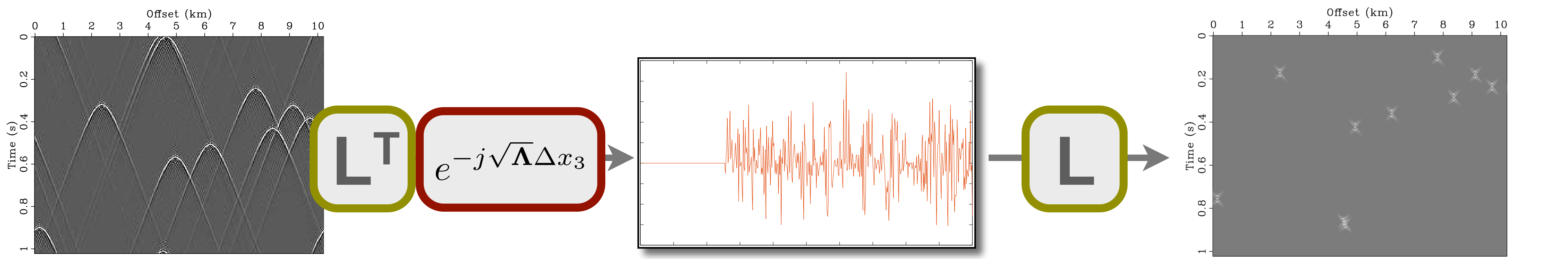
Compressed wavefield extrapolation

$$\begin{cases} \mathbf{y} & = \mathbf{R}\mathbf{L}^H\mathbf{u} \\ \mathbf{A} & = \mathbf{R}e^{j\mathbf{\Lambda}^{1/2}\Delta x_3}\mathbf{L}^H \\ \tilde{\mathbf{x}} & = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \\ \tilde{\mathbf{v}} & = \tilde{\mathbf{x}} \end{cases}$$

- Randomly subsample & phase rotate in Modal domain
- Recover by norm-one minimization
- Capitalize on
  - the incoherence modal functions and point scatterers
  - reduced explicit matrix size
  - constant velocity <=> Fourier recovery

SLIM Seismic Laboratory for Imaging and Modeling

Straightforward 1-Way inverse Wavefield Extrapolation

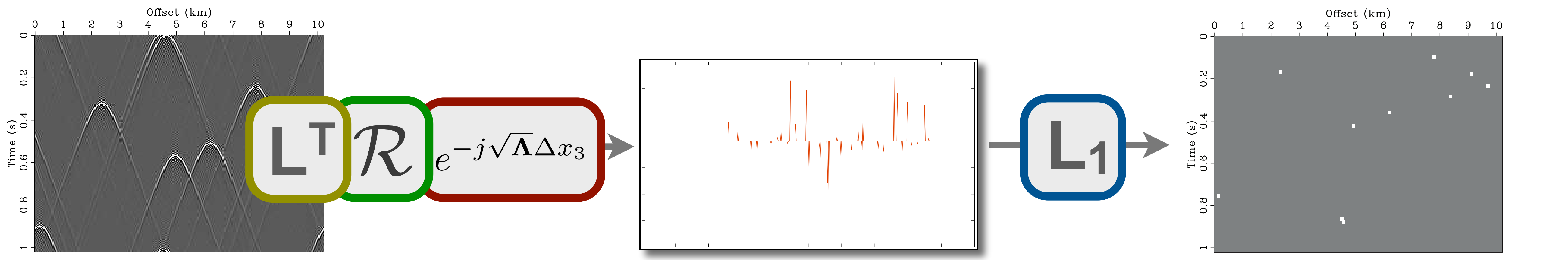


wavefield in space-time domain

back-extrapolated wavefield in H2 domain

back-extrapolated to impulse source in space-time domain

Compressed 1-Way Wavefield Extrapolation



wavefield in space-time domain

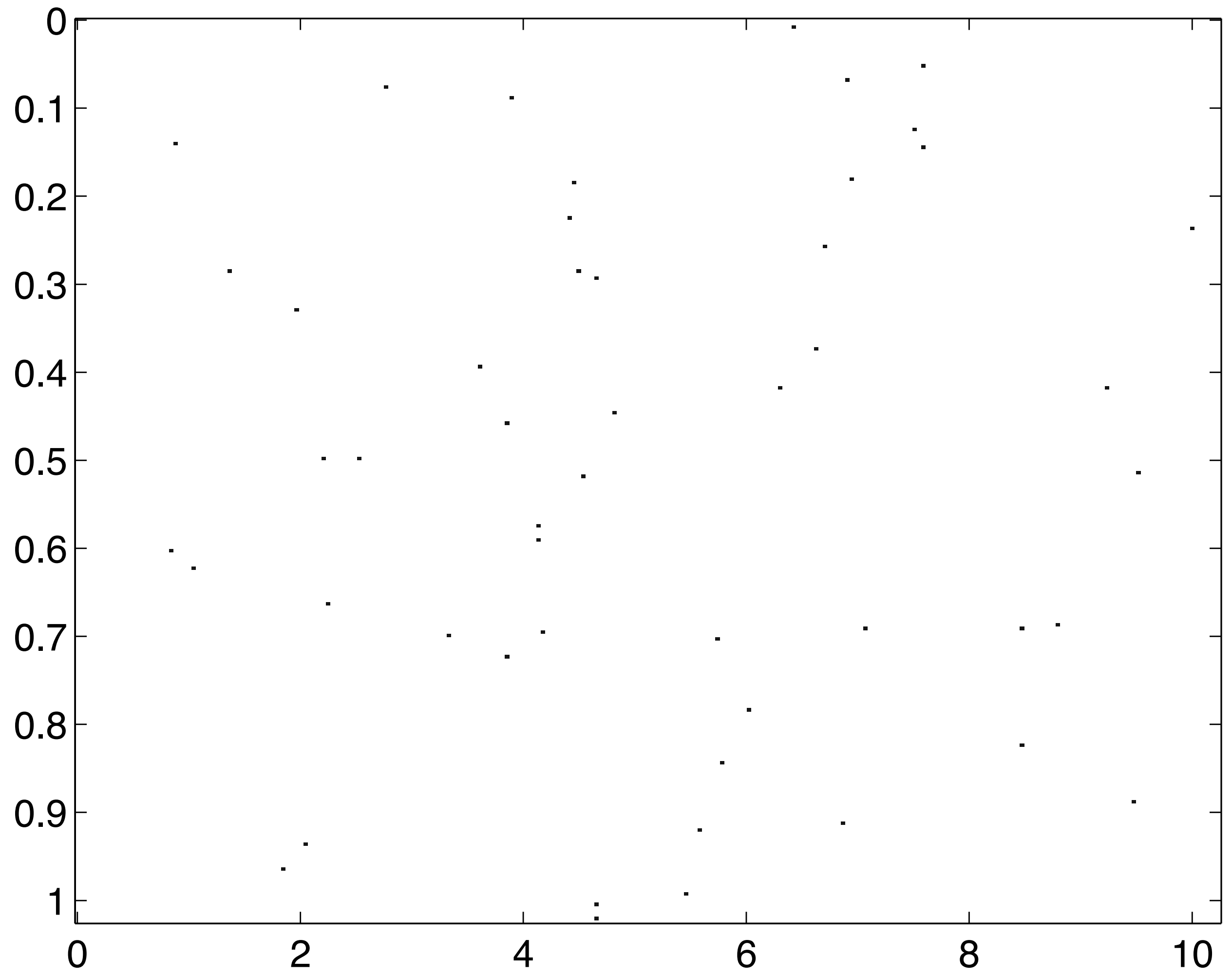
incomplete back-extrapolated wavefield in H2 domain

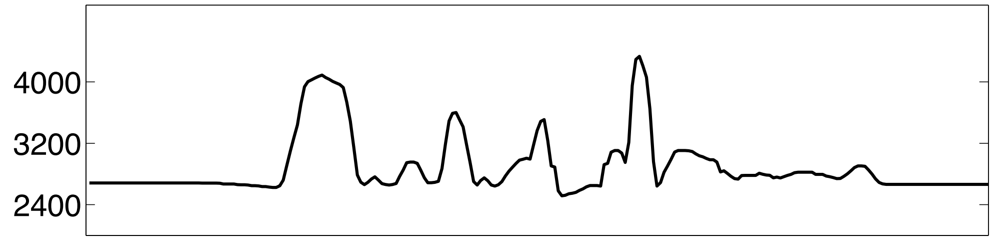
back-extrapolated to impulse source in space-time domain

SLIM Seismic Laboratory for Imaging and Modeling

Compressed wavefield extrapolation

simple 1-D space/time propagation example with point scatters

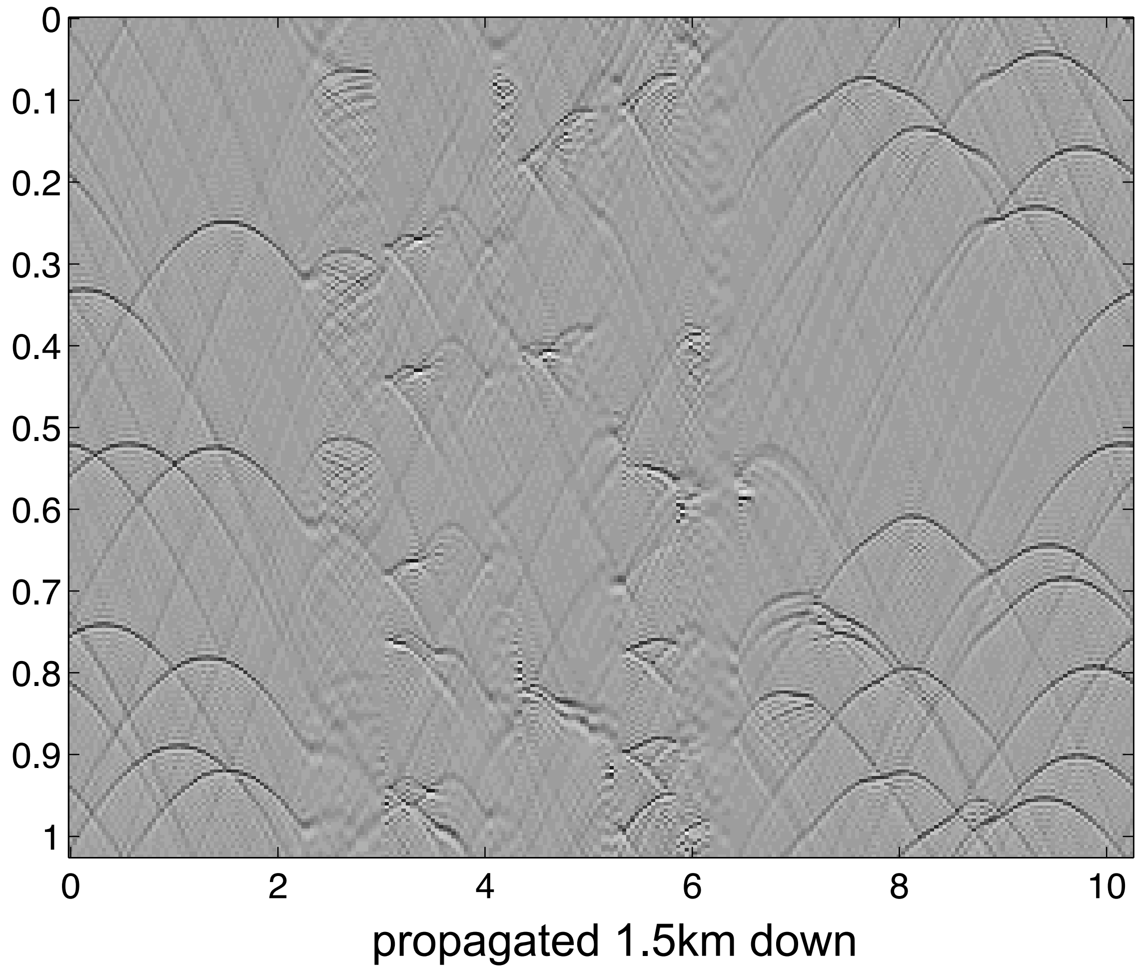


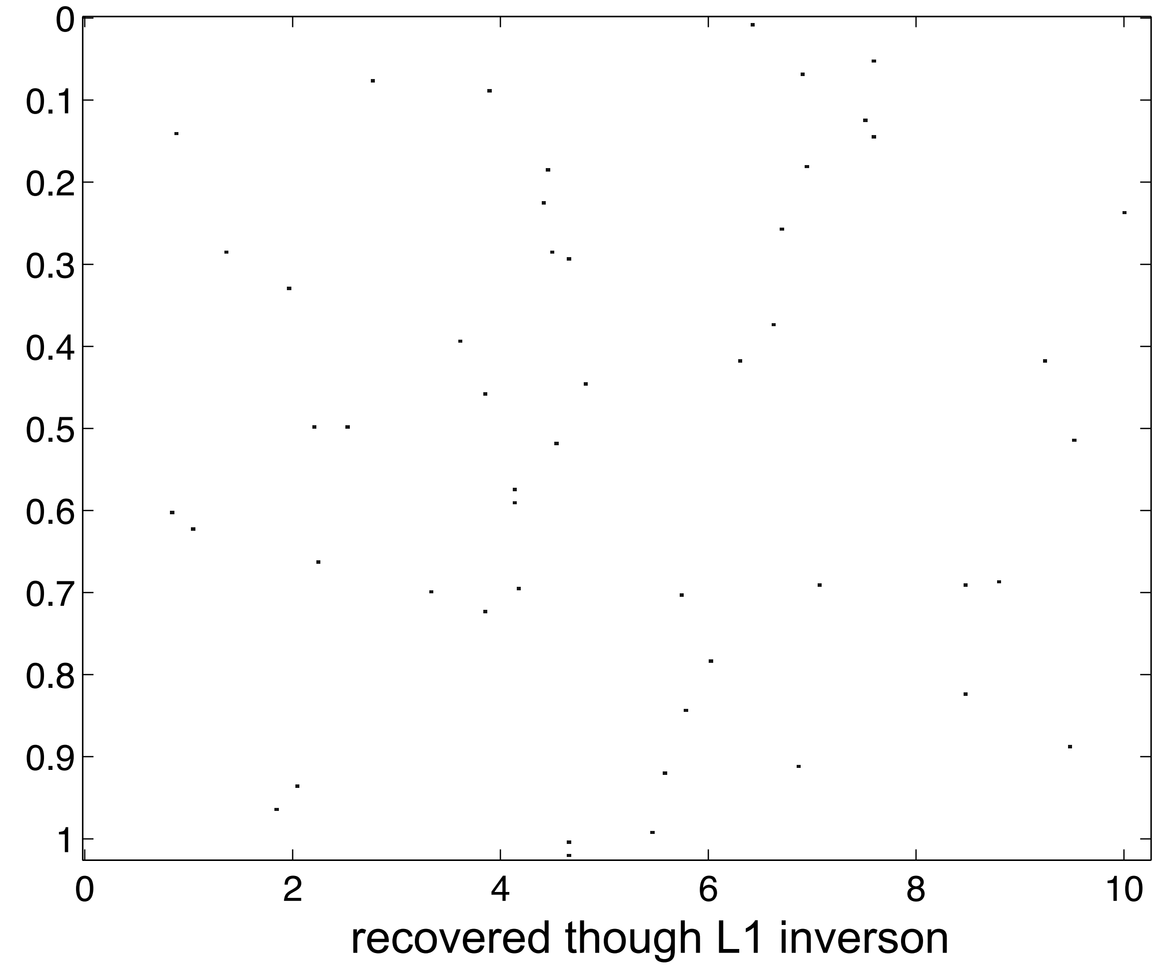


SLIM
Seismic Laboratory for Imaging and Modeling

Compressed wavefield extrapolation

simple 1-D space/time propagation example with point scatters





Restricted L transform to ~0.01 of original coefficients

SLIM
Seismic Laboratory for Imaging and Modeling

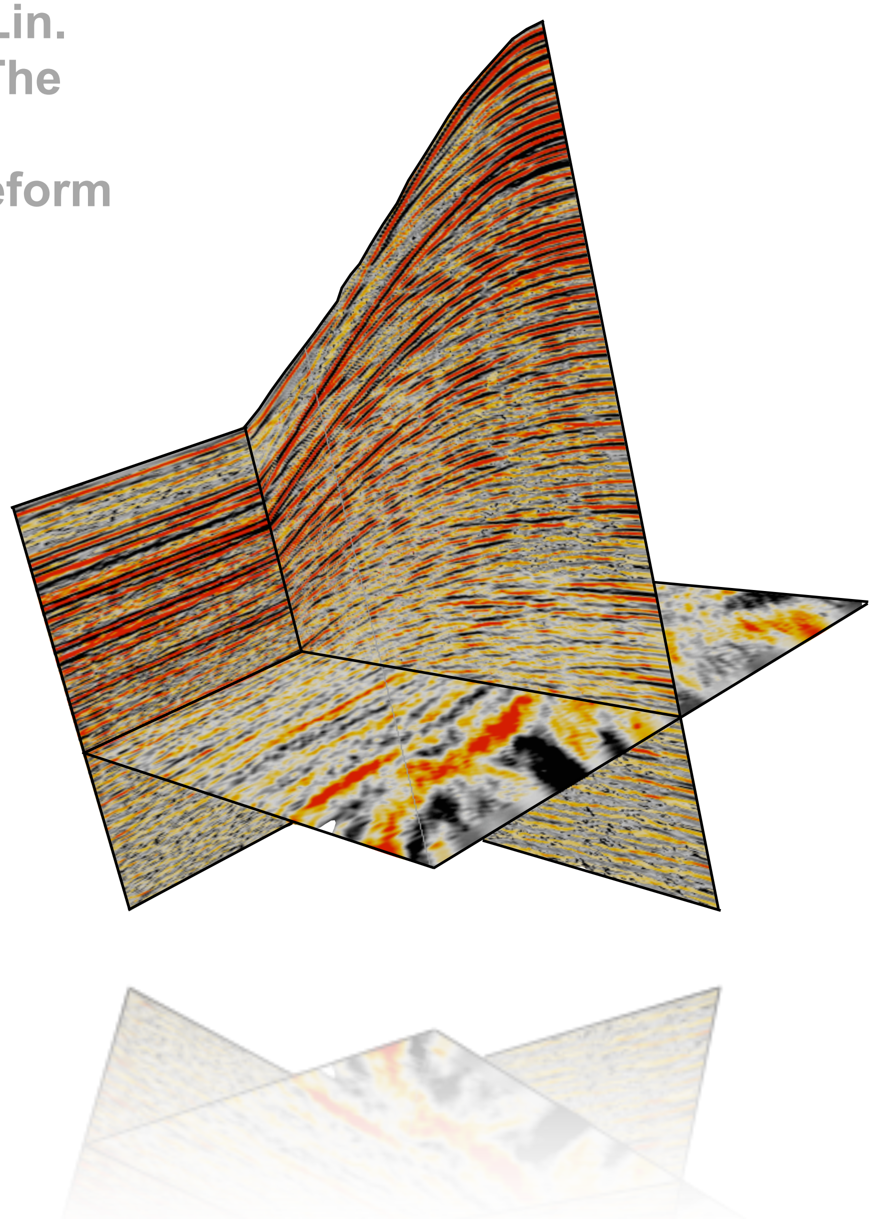
Observations

- Solution can be obtained from incomplete spectral representation
  - feasible if somebody comes up with a fast random eigensolver
- Performance depends on
  - transform-domain sparsity of the extrapolated wavefield
  - mutual incoherence between spectral representation of the operator and the transform that sparsifies the solution
- Mutual coherence and sparsity are at odds
  - curvelets are localized eigenfunction like

THE UNIVERSITY OF BRITISH COLUMBIA | VANCOUVER

Compressive simultaneous full-waveform simulation

Felix J. Herrmann, Yogi Erlangga, and Tim T. Y. Lin.
Seismic Laboratory for Imaging and Modeling. The university of British Columbia Technical Report.
TR-2008-3. Compressive simultaneous full-waveform simulation. To appear as a letter in Geophycis.



Delft University of Technology
Delft, April 8, 2009

Motivation

- New *implicit* solvers for the preconditioned time-harmonic Helmholtz
  - combination of multigrid and deflation
  - numerical *convergence* for increasing *frequencies* and decreasing *grid* sizes

- *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 ≈ acquisition & computational cost** *proportional* to the **complexity** of *data* and *model*

Wavefield Computation
acoustic case

Time domain: $U(x,t)$

$$m\frac{\partial^2 U}{\partial t^2} = \Delta U, \quad m = c^{-2}(x), \quad t = (0,T)$$

Numerical method

- Time marching: $U^{n_t+1} = U^{n_t} + W(U^{n_t})$, explicit method
- Courant-Friedrichs-Lewy (CFL) condition for stability

Frequency domain: $u(x,\omega)$

$$\mathcal{H}u = -\Delta u - \omega^2 m u = b$$

($\mathcal{H}$: Helmholtz operator)

Numerical method

- Implicit: $$\boldsymbol{\mathcal{H}}(\omega)\mathbf{u} = \mathbf{b} \rightarrow \mathbf{u} = \boldsymbol{\mathcal{H}}^{-1}(\omega)\mathbf{b} \qquad \forall\omega \in \Omega$$
- Nyquist sampling, aliases

SLIM
Seismic Laboratory for Imaging and Modeling

Wavefield computations

- For n_s shots and n_f frequencies, the linear (Helmholtz) systems are independent
- The multi-shot and multi-frequency problem is embarrassingly parallel

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

$$\mathcal{H}_{\omega_j} := \mathcal{H}(\omega_j), \quad \omega_j = 2\pi j \Delta f, \quad j = 1, \ldots, n_f$$

Δf frequency sample interval

Wavefield computations

$$\overbrace{\begin{bmatrix} \mathcal{H}_{\omega_1} & 0 & & \\ 0 & \mathcal{H}_{\omega_2} & \ddots & \\ & \ddots & \ddots & 0 \\ & & 0 & \mathcal{H}_{\omega_{n_f}} \end{bmatrix}}^{\mathbf{H}} \overbrace{\begin{bmatrix} \mathbf{U}_{\omega_1} \\ \mathbf{U}_{\omega_1} \\ \vdots \\ \mathbf{U}_{n_f} \end{bmatrix}}^{\mathbf{U}} = \overbrace{\begin{bmatrix} \mathbf{B}_{\omega_1} \\ \mathbf{B}_{\omega_2} \\ \vdots \\ \mathbf{B}_{n_f} \end{bmatrix}}^{\mathbf{B}}$$

$$\Downarrow$$

$$\mathbf{HU} = \mathbf{B}$$

- Natural for CS setting!

Iterative Helmholtz Solver

Old method: slow convergence

- The matrix $\mathcal{H}$ is indefinite and ill-conditioned

Recent progress:

- **Preconditioner** [Erlangga, Oosterlee, Vuik, 2006]

$$\mathcal{M} \triangleq \left(-\Delta - (1-\beta\hat{i})\omega^2 m\right)_h, \quad \beta = (0,1]$$

- **Deflation operator** [Erlangga, Nabben, 2008]

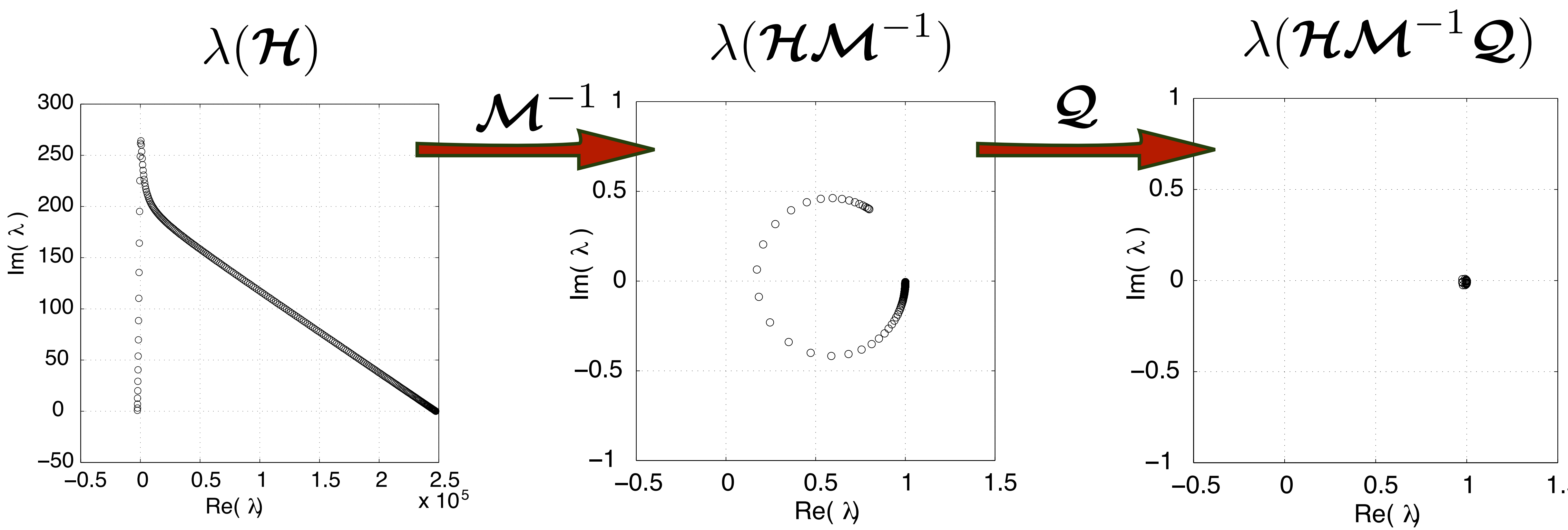
$$\mathcal{Q} := \mathbf{I} - \mathbf{Z}\mathbf{E}^{-1}\mathbf{Y}^\top\mathcal{H}\mathcal{M}^{-1} - \mathbf{Z}\mathbf{E}^{-1}\mathbf{Y}^\top$$

with: $\mathbf{E} = \mathbf{Y}^\top\mathcal{H}\mathcal{M}^{-1}\mathbf{Z}$

$\mathbf{Z}, \mathbf{Y}$ multigrid-type interpolation matrices

Iterative Helmholtz Solver cont'd

Eigenvalue λ of 1D Helmholtz equation



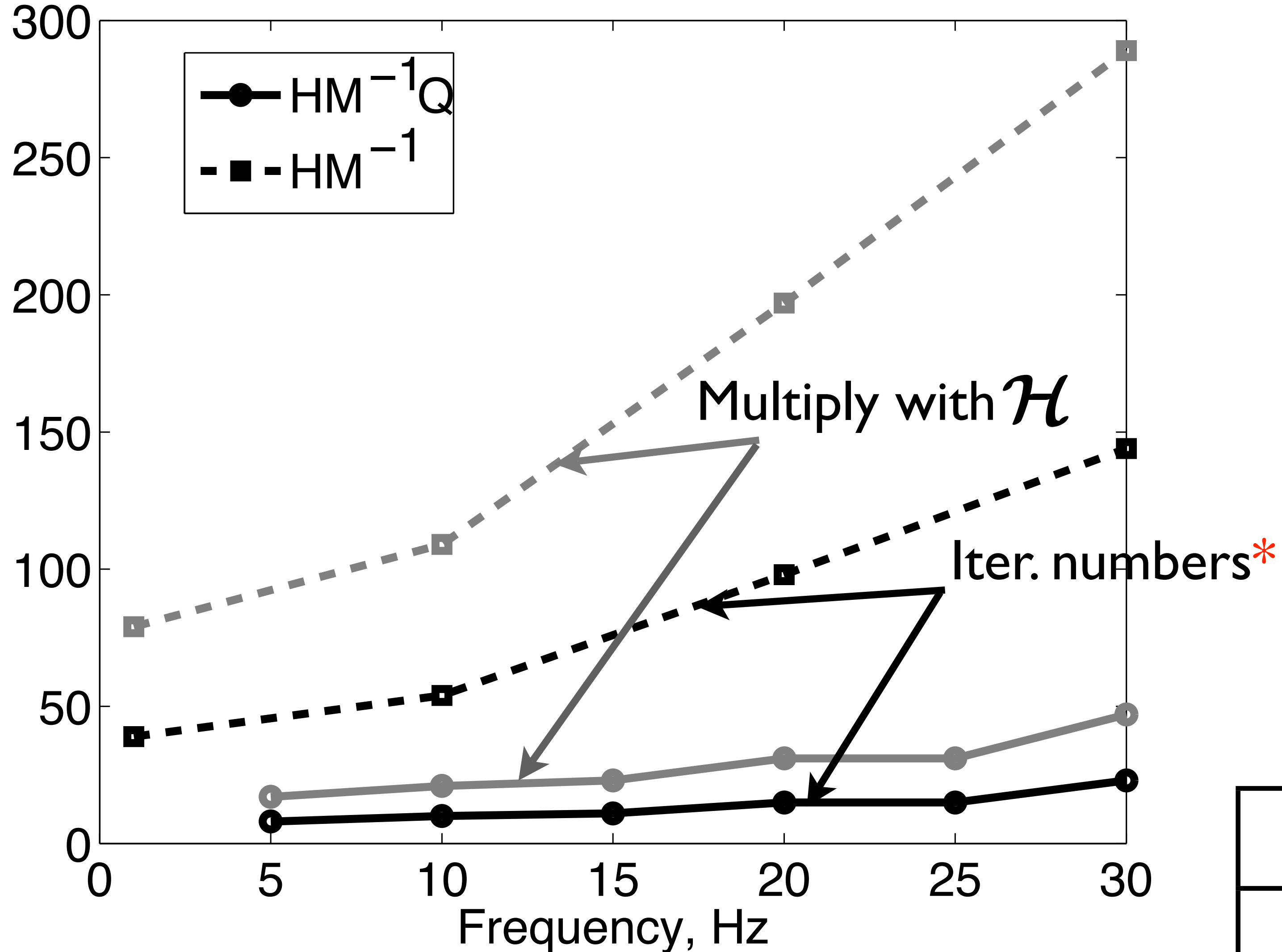
- $\mathcal{M}^{-1}$ shifts the eigenvalues to the positive half plane (solve indefiniteness)
- $\mathcal{Q}$ clusters the eigenvalue around one (solve ill-conditioning)

For iterative methods, fast convergence

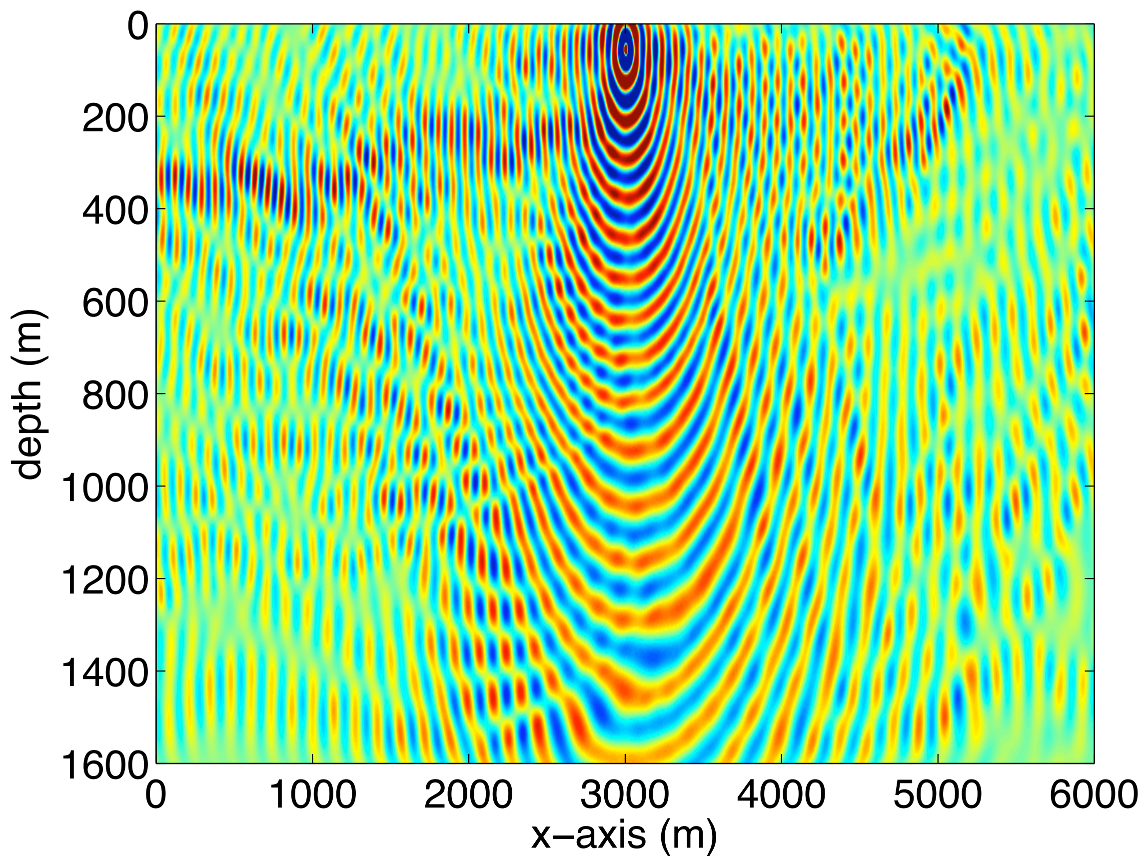
SLIM
Seismic Laboratory for Imaging and Modeling

Example: Marmousi

Forward Model

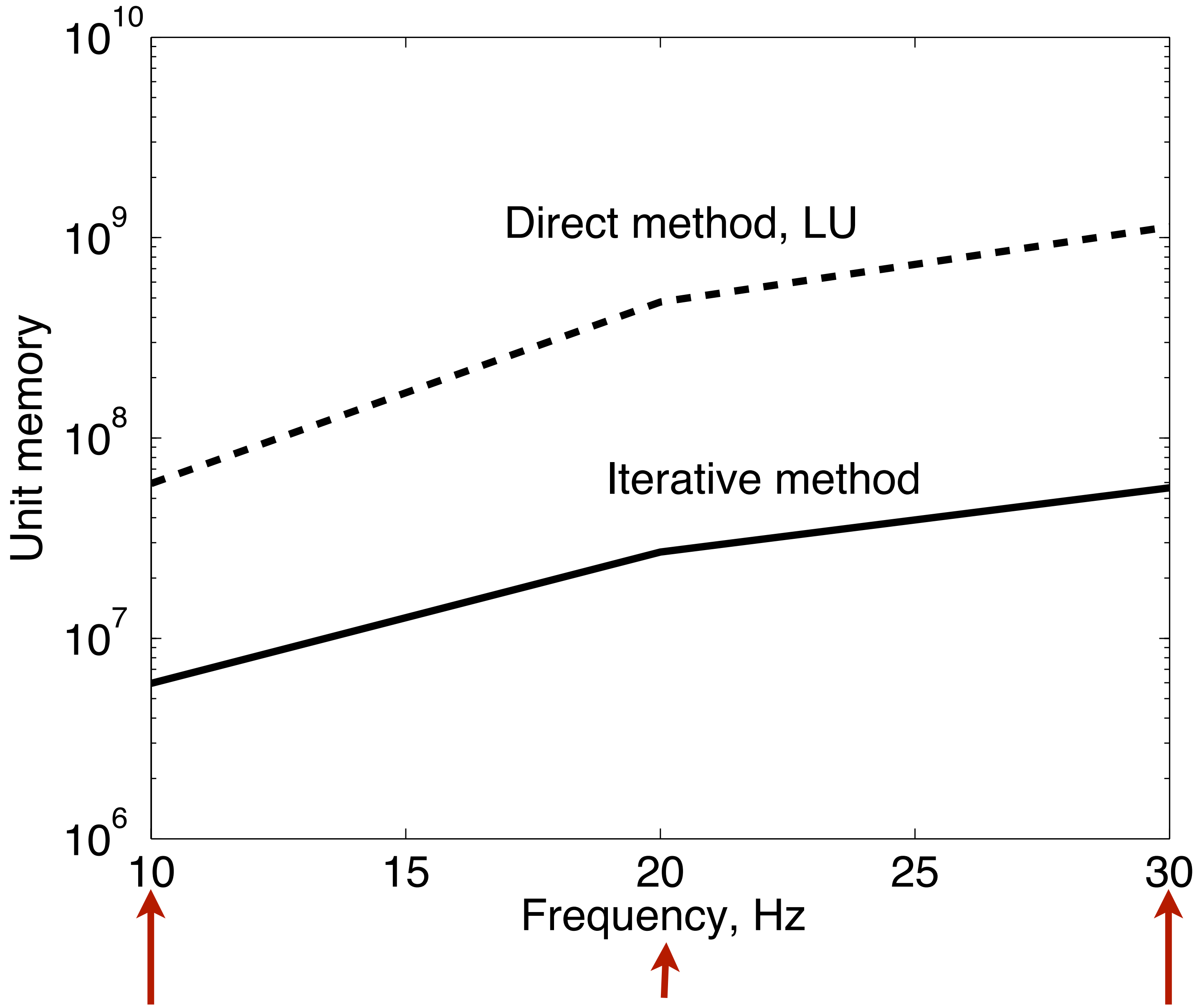


* Residual reduced to 10^{-6}



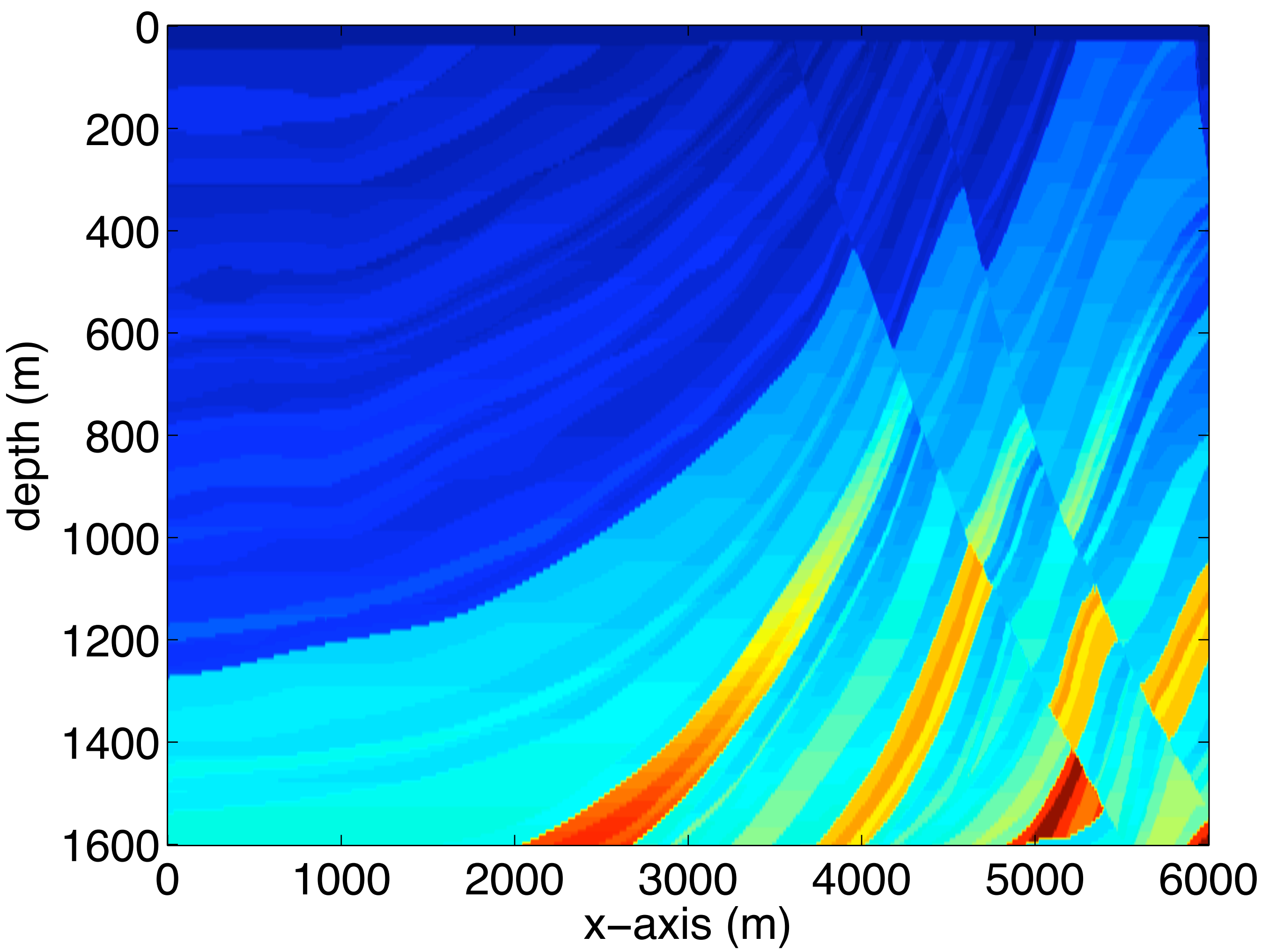
Frequency [Hz]	5	10	20	30
Forward	8	10	15	23
Back propagation	8	10	15	23

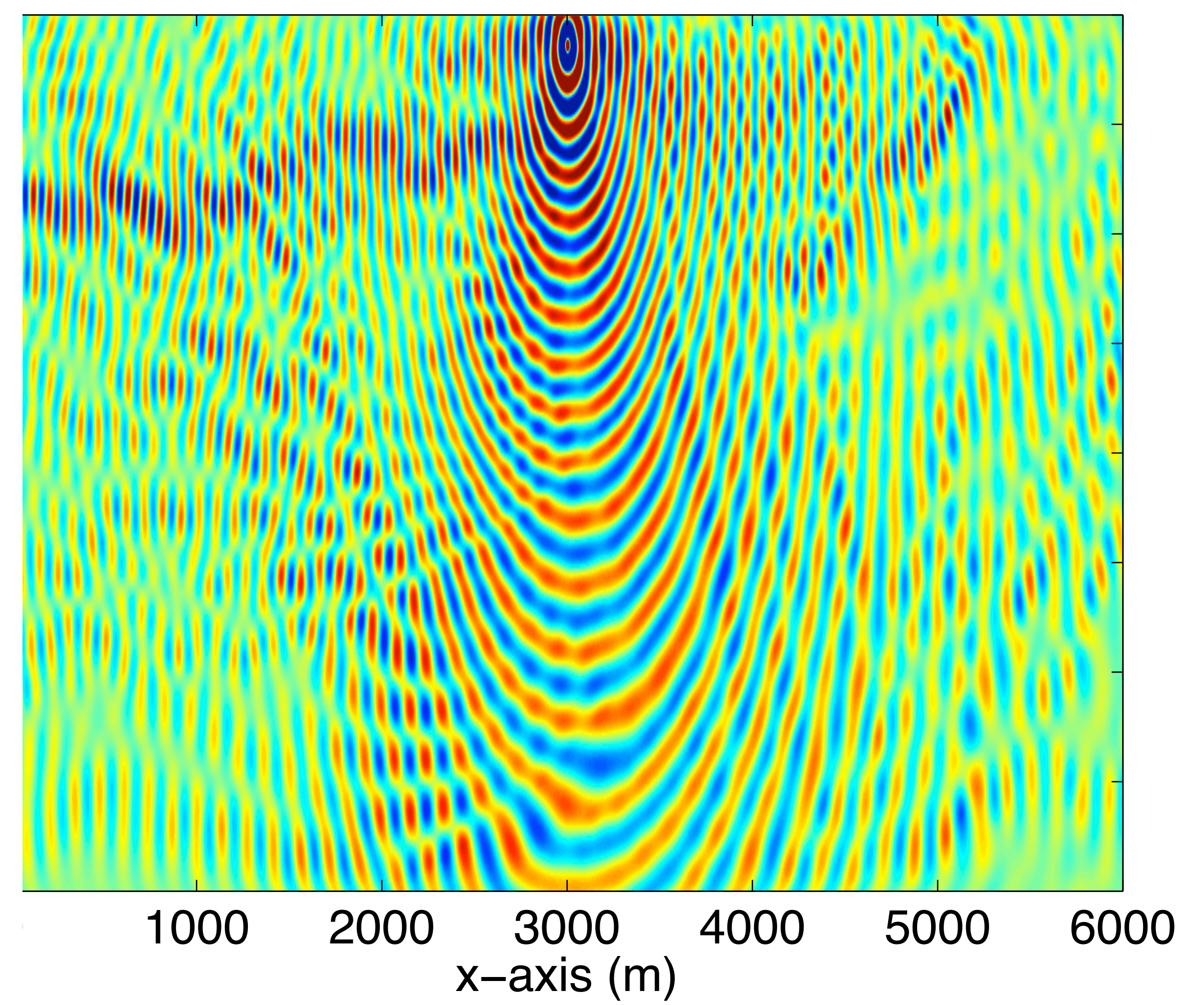
Example: Marmousi, cont'd



gridpoint : 751×201 1501×401 2001×534

Forward modeling cont’d





Despite significant improvement by Helmholtz preconditioner

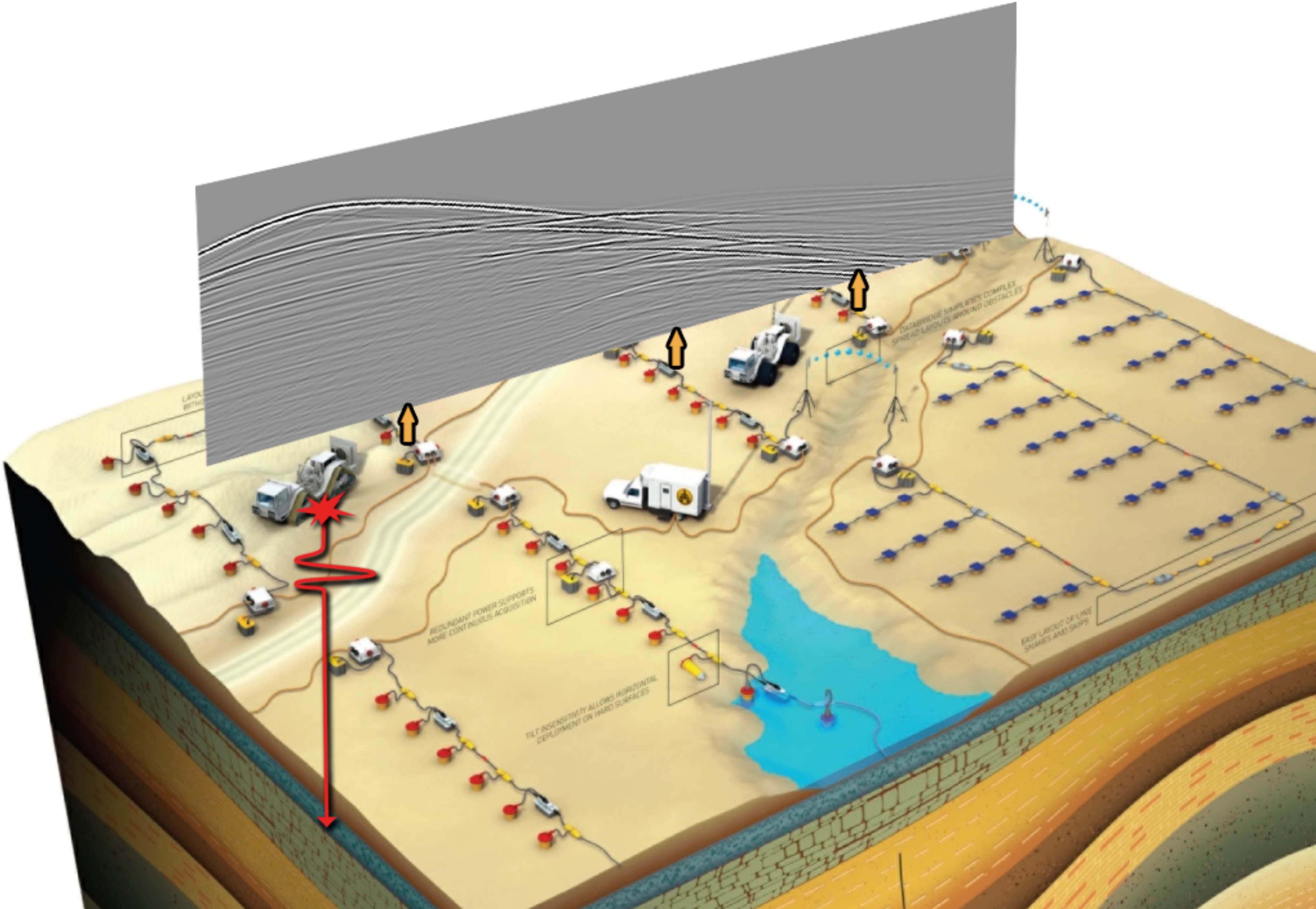
- redundancy <=> extreme large size seismic data volumes
- multiple frequencies & multiple right-hand sides
- expensive modeling, imaging & inversion costs

Leverage new paradigm of CS ...

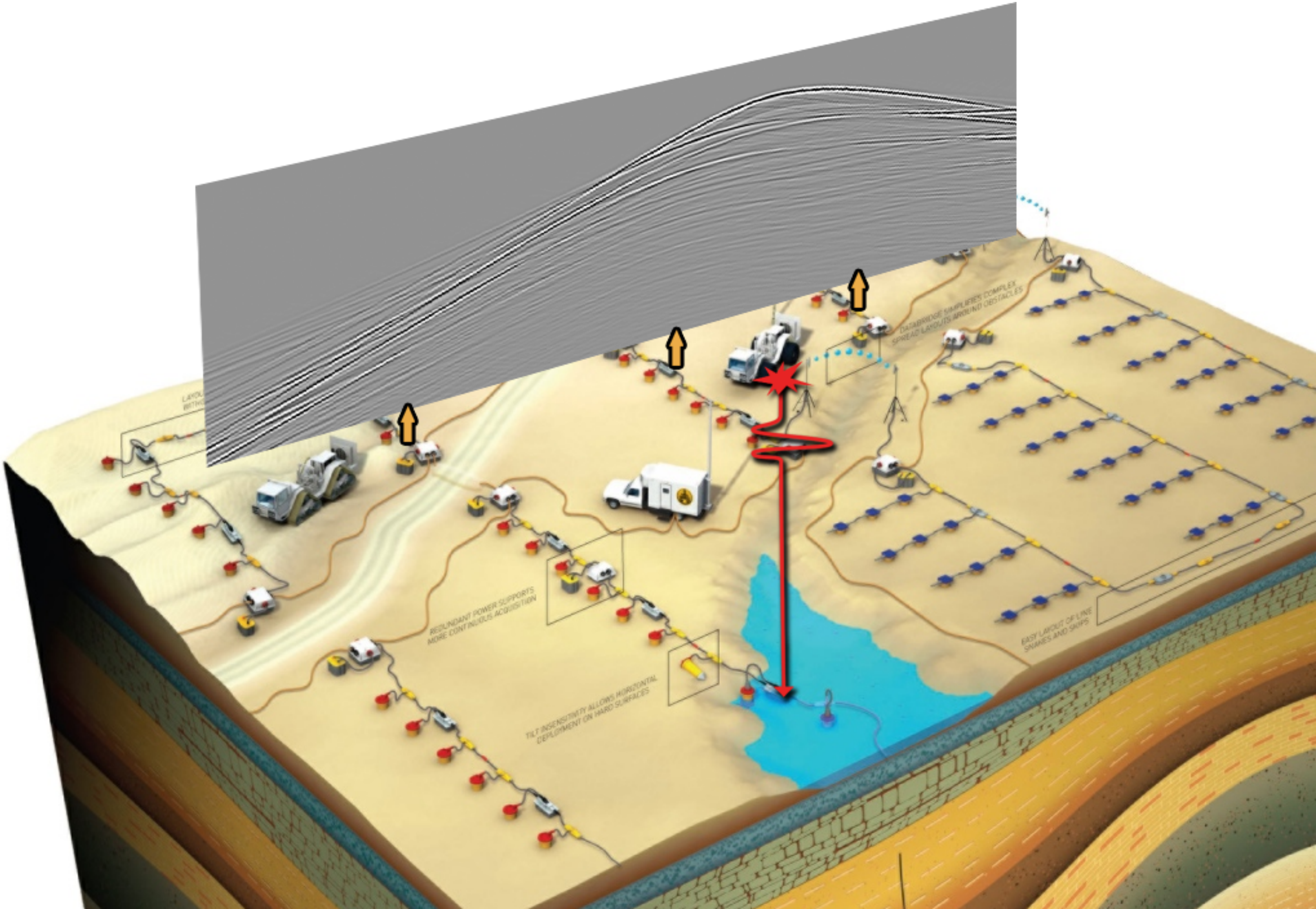
SLIM
Seismic Laboratory for Imaging and Modeling

Acquisition image courtesy ION (www.iongeo.com)

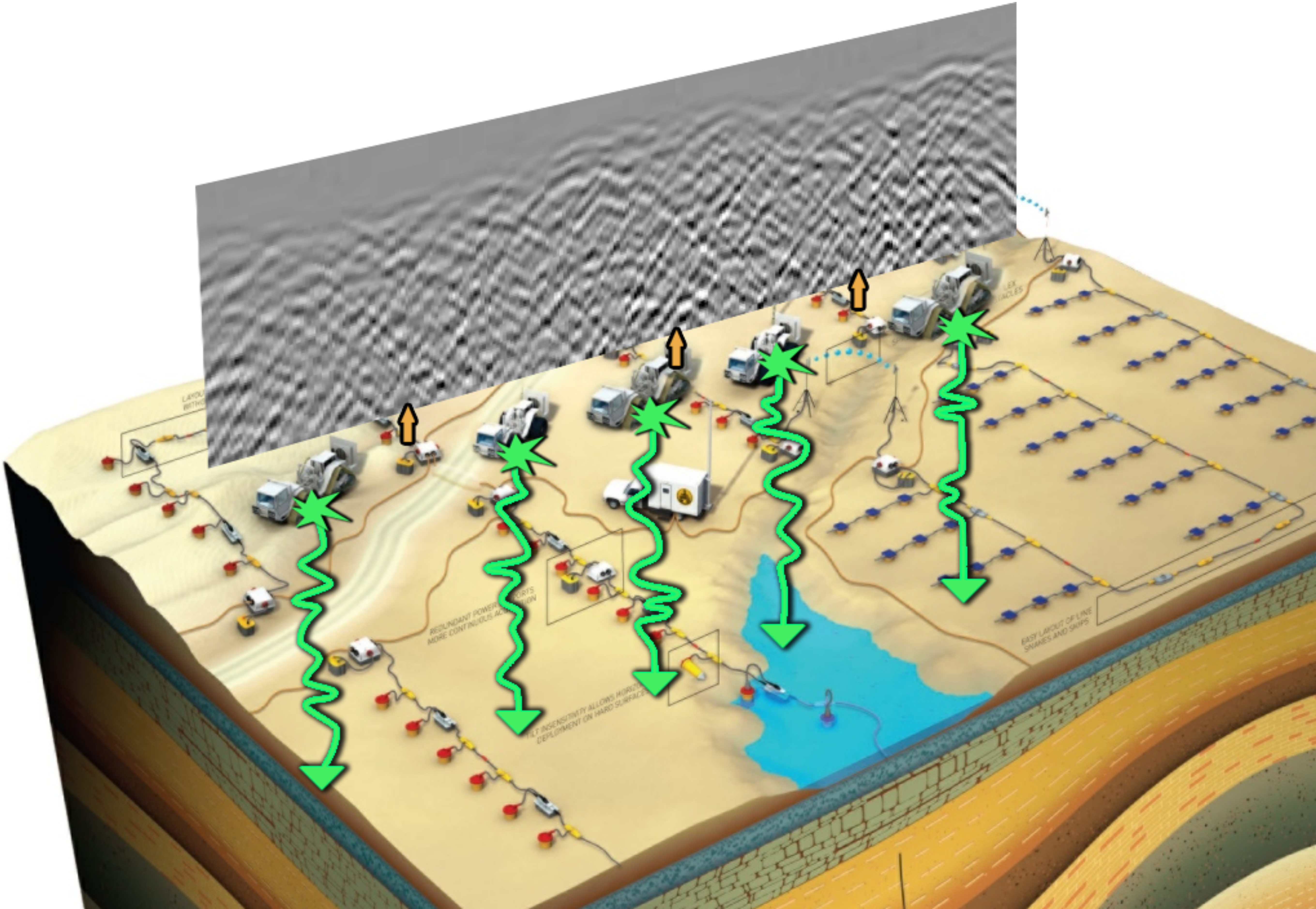
Individual shots



Individual shots



Simultaneous & continuous shots



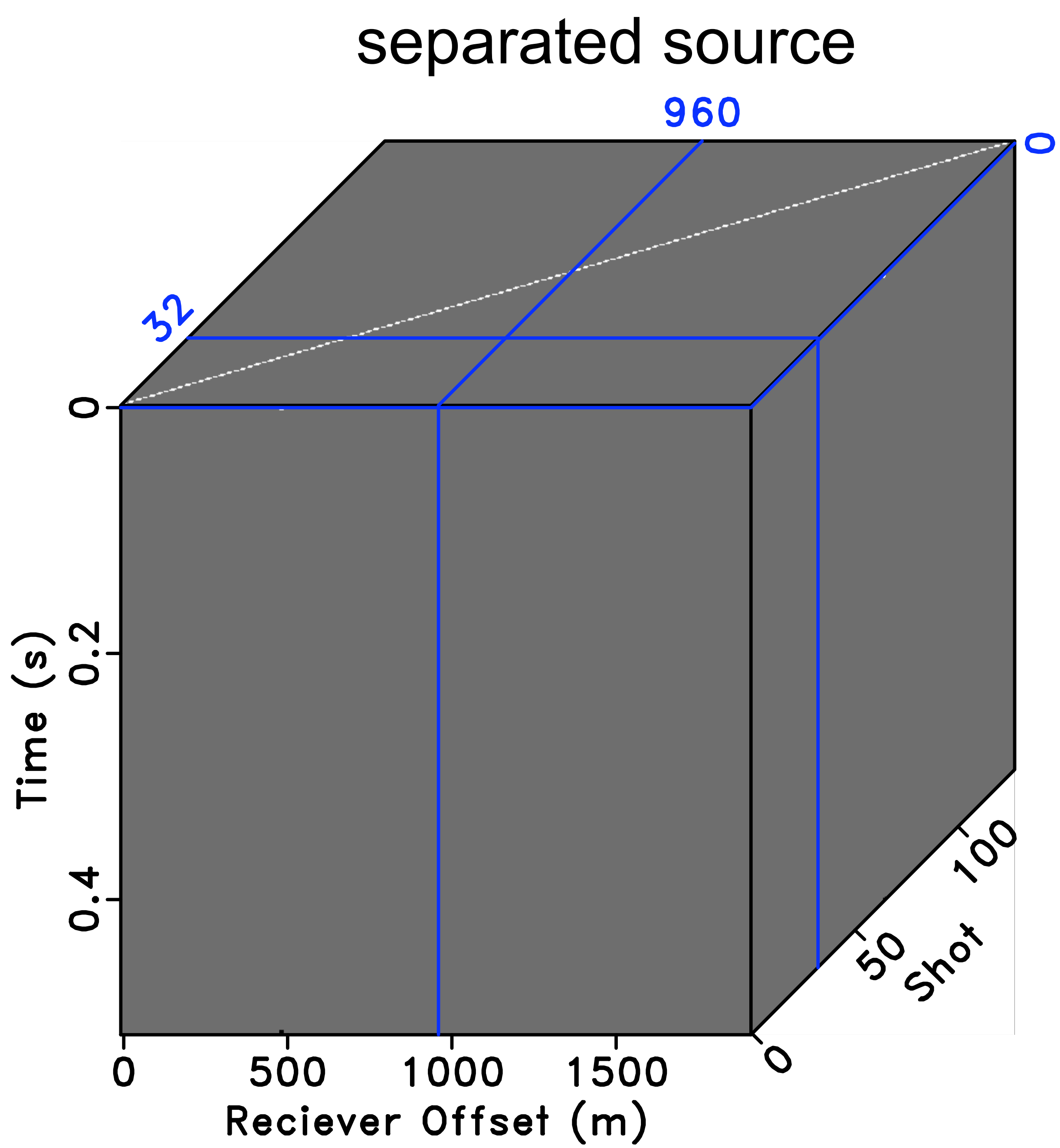
CS sampling of frequencies and shots (rhs)

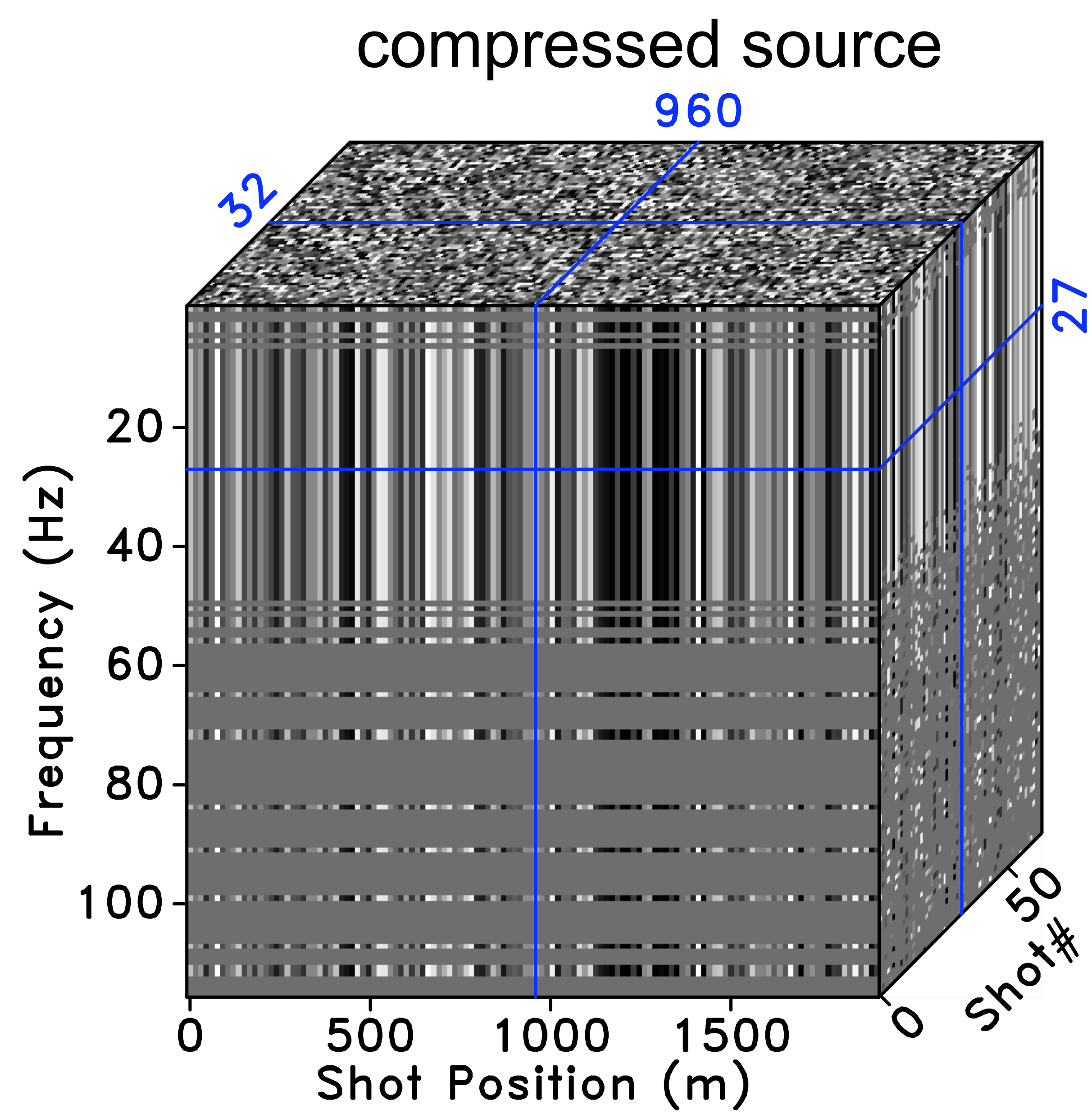
- **CS with Random Convolution (Romberg '08)**
  - Replace Gaussian matrix along shots with restricted **random convolution** over the whole seismic data
  - CS with 3D Fourier Transform $\mathbf{F}_3$ and multiply each coefficient with a unit-norm complex number of randomly determined phase, followed by inverse 2D Fourier on shot-receiver plane, then restrict in both temporal-frequency and shot coordinates

$$\mathbf{RM} = \overbrace{\begin{bmatrix} \mathbf{R}_1^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{R}_1^{\Omega} \\ \vdots \\ \mathbf{R}_{n_{s'}}^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{R}_{n_{s'}}^{\Omega} \end{bmatrix}}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_2^* \operatorname{diag}\left(e^{\hat{i}\boldsymbol{\theta}}\right) \otimes \mathbf{I}\right) \mathbf{F}_3}^{\text{random phase encoder}},$$

$$\theta_w = \text{Uniform}([0, 2\pi])$$

Applying to Shot Sources





SLIM
Seismic Laboratory for Imaging and Modeling

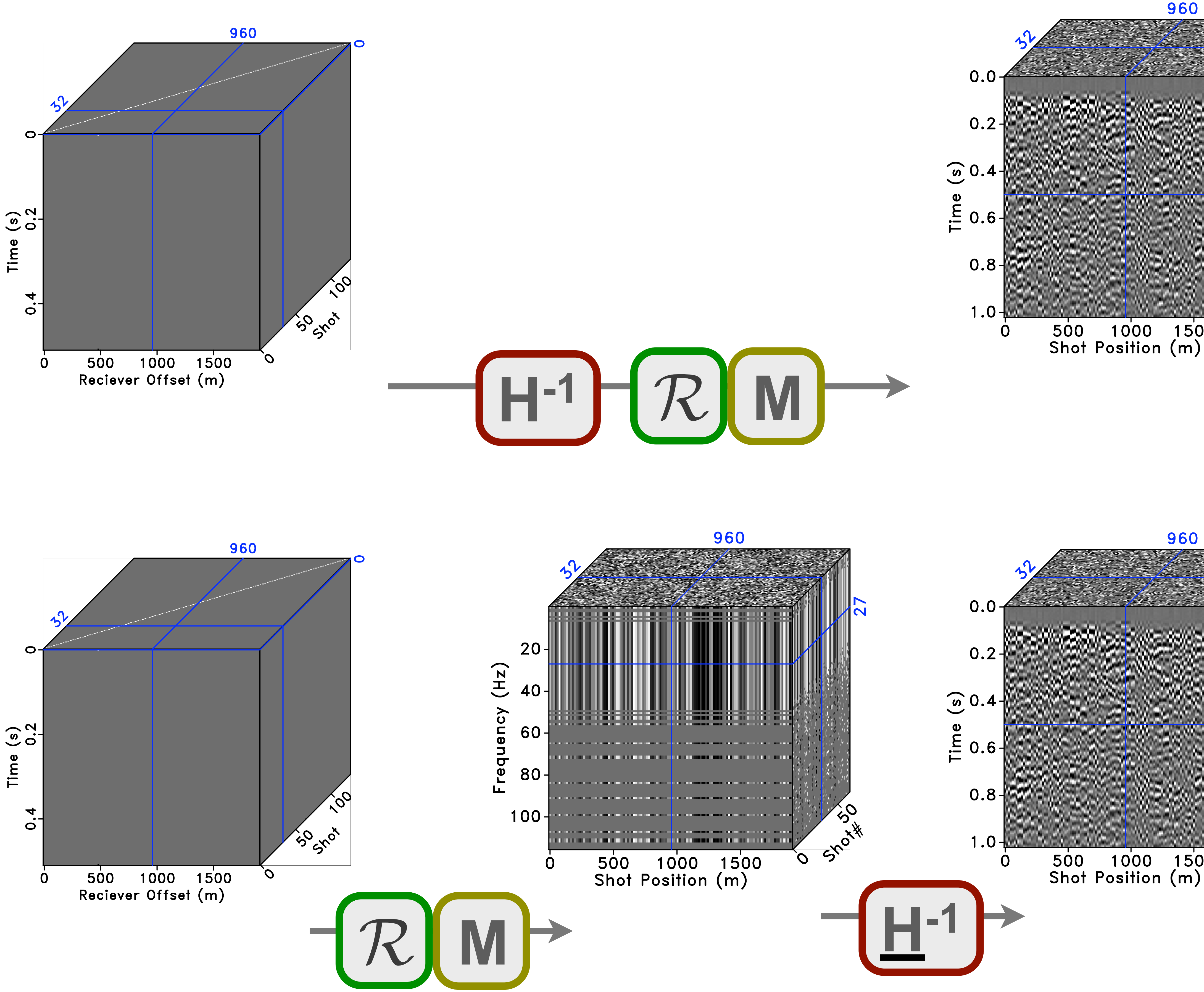
Equivalence

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**

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

Show that $\mathbf{y} = \underline{\mathbf{y}}$.



SLIM
Seismic Laboratory for Imaging and Modeling

CS

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

CS provides conditions under which P1 recovers **d**:

- 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

SLIM
Seismic Laboratory for Imaging and Modeling

Composite sparsity transform

Using Curvelet transform for shot and receiver coordinates

- **Frequency-domain** restrictions perform well under **Wavelet** transform for seismic data (Lin et. al. ’08)
- **Spatial-domain** restrictions perform well under **Curvelet** transform for seismic data (Hennefent et. al. ’07)

Combine both transforms in the coordinate they are most suited for

- Wavelet sparsity on temporal-frequency coordinate
- 2D Curvelet sparsity on shot and receiver plane

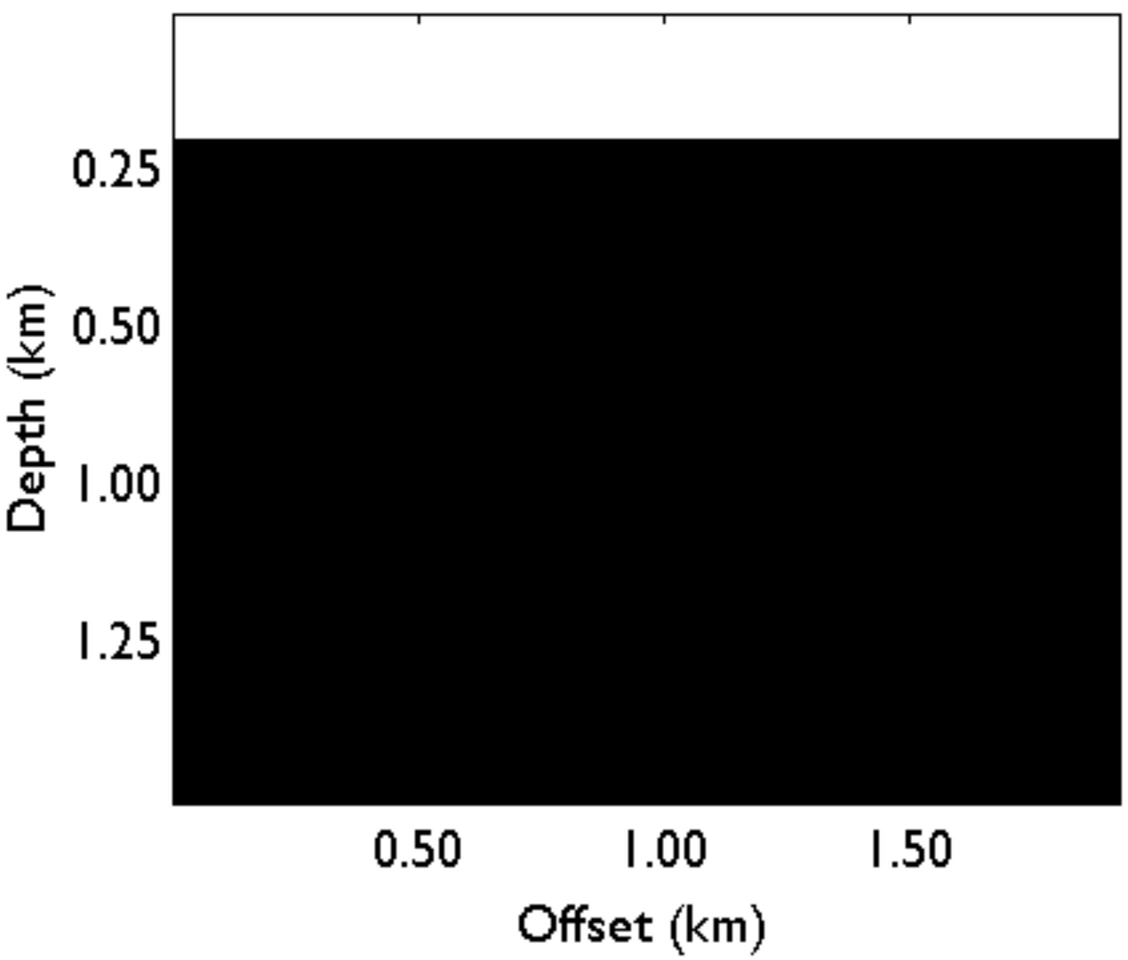
$$\mathbf{S} = \mathbf{C}_{2d} \otimes \mathbf{W}$$

Complexity $\mathcal{O}(n^3 \log n)$

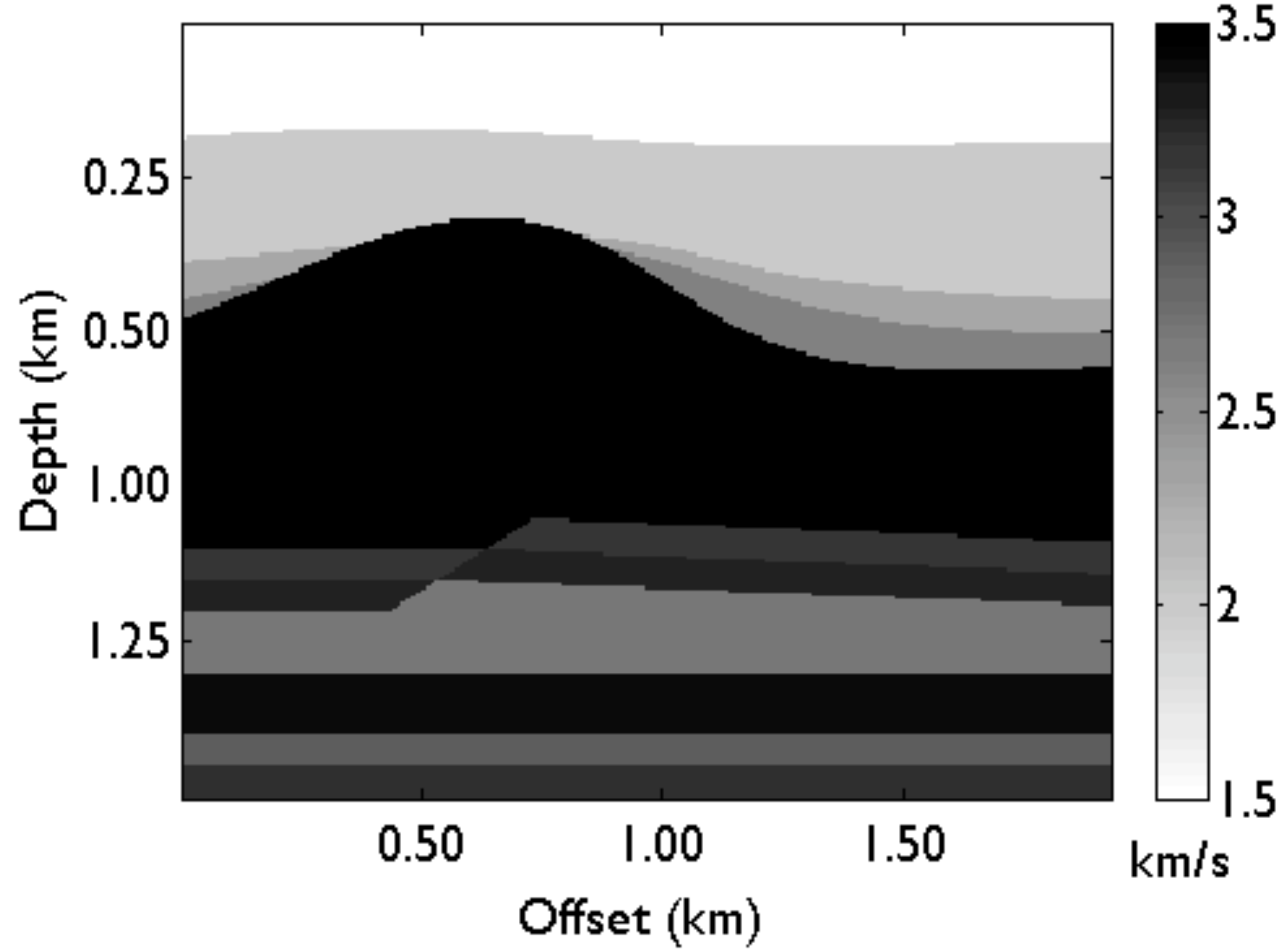
SLIM
Seismic Laboratory for Imaging and Modeling

Velocity models

simple model

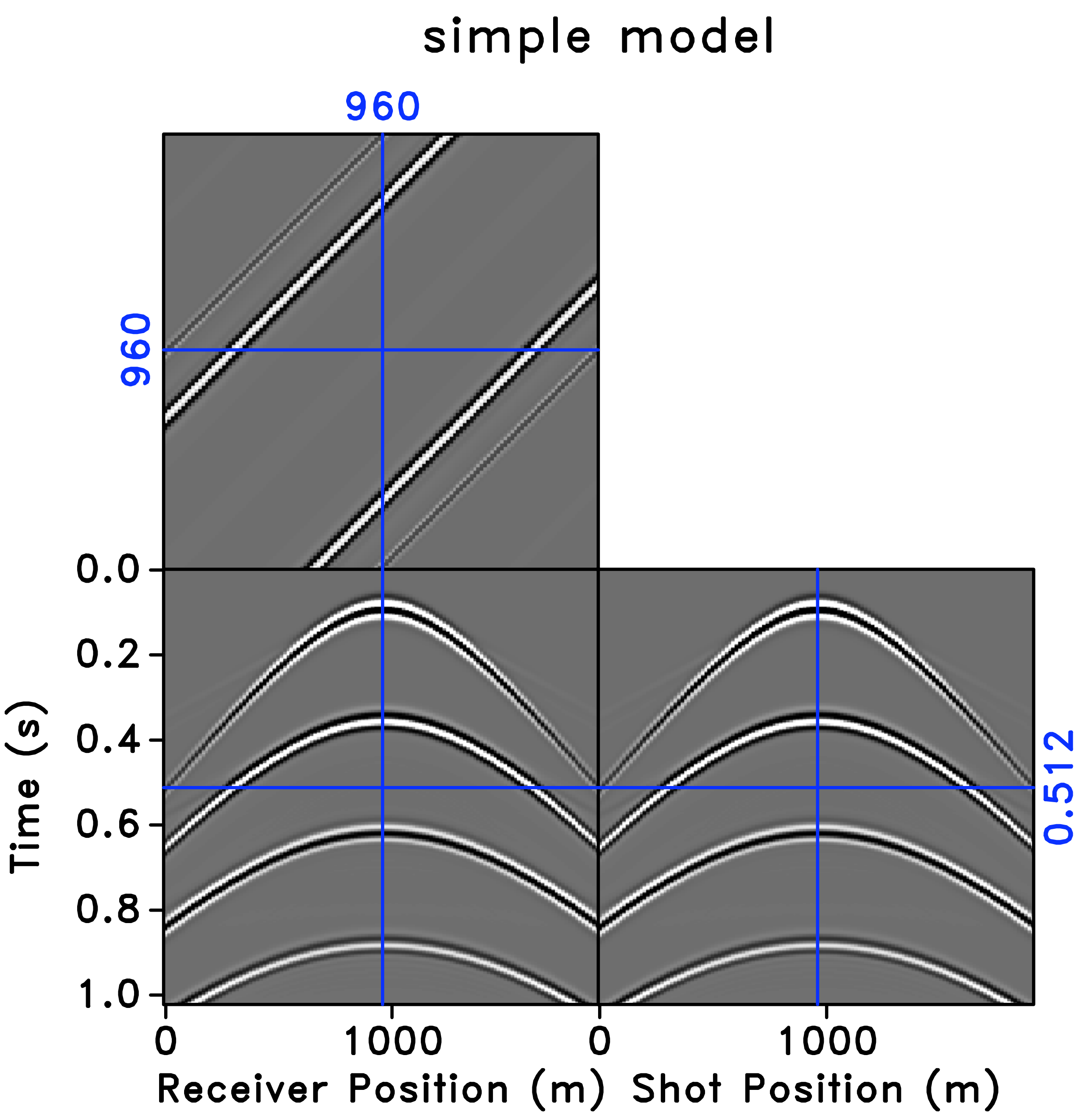


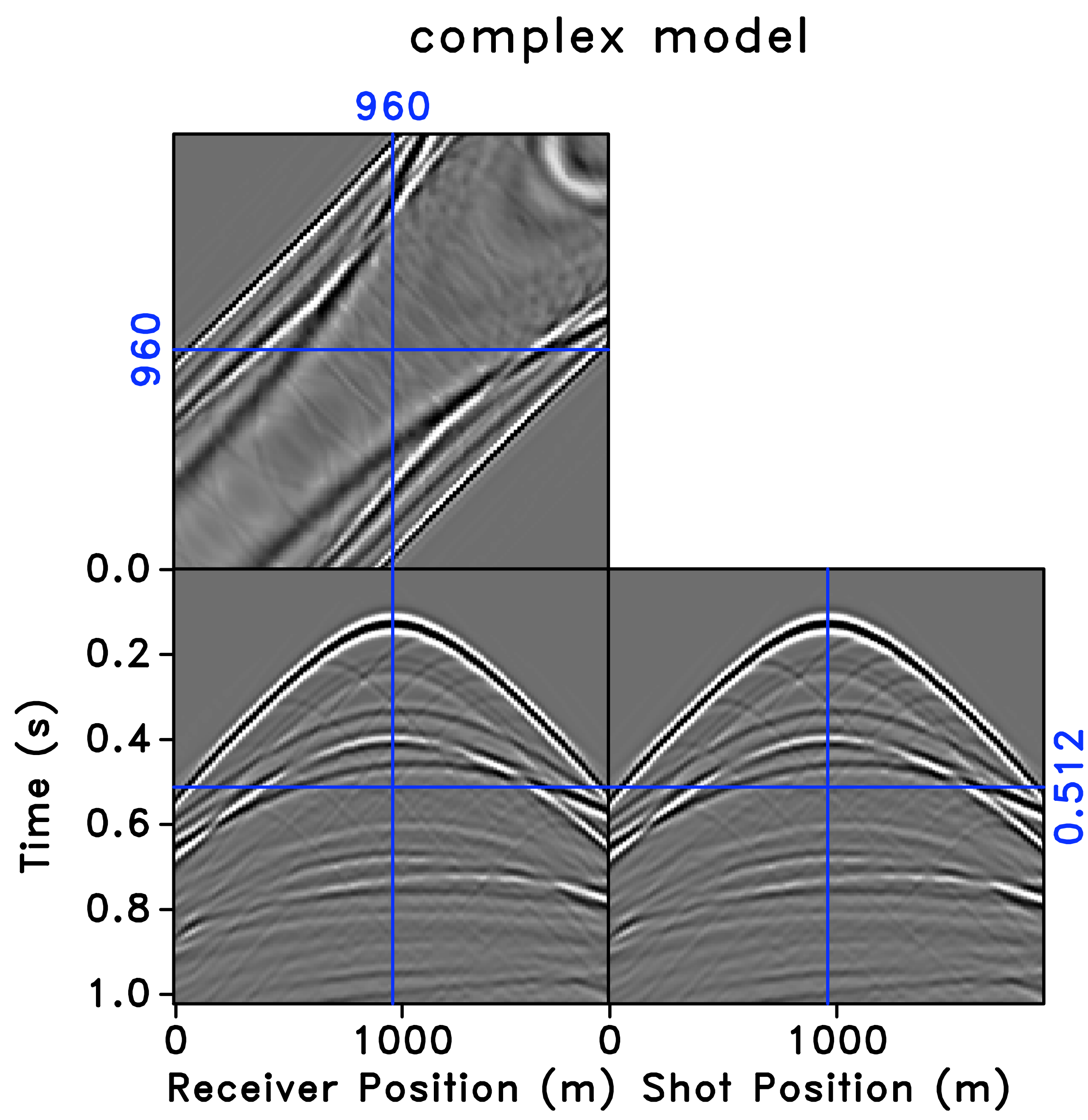
complex model



SLIM
Seismic Laboratory for Imaging and Modeling

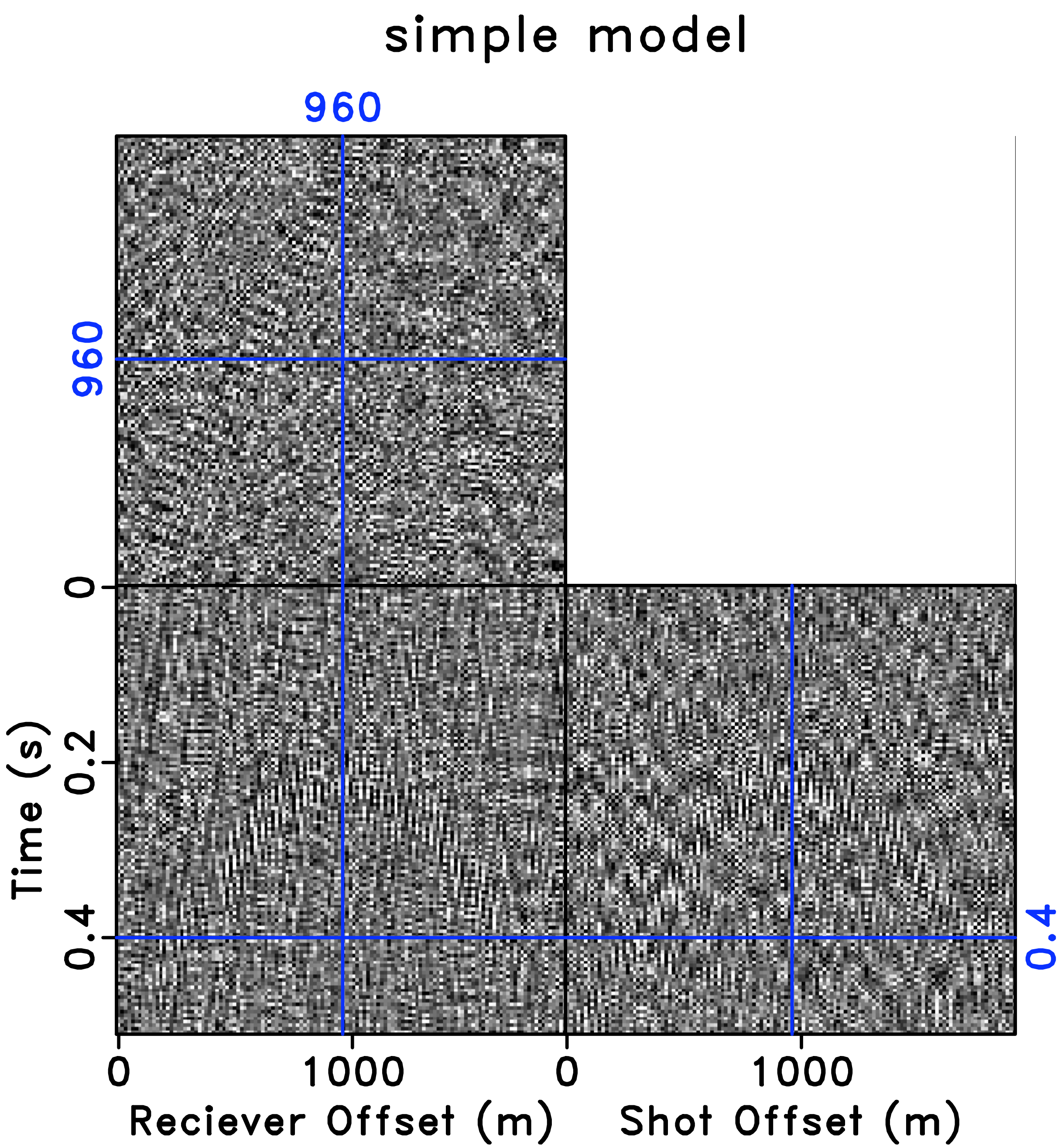
Green’s functions

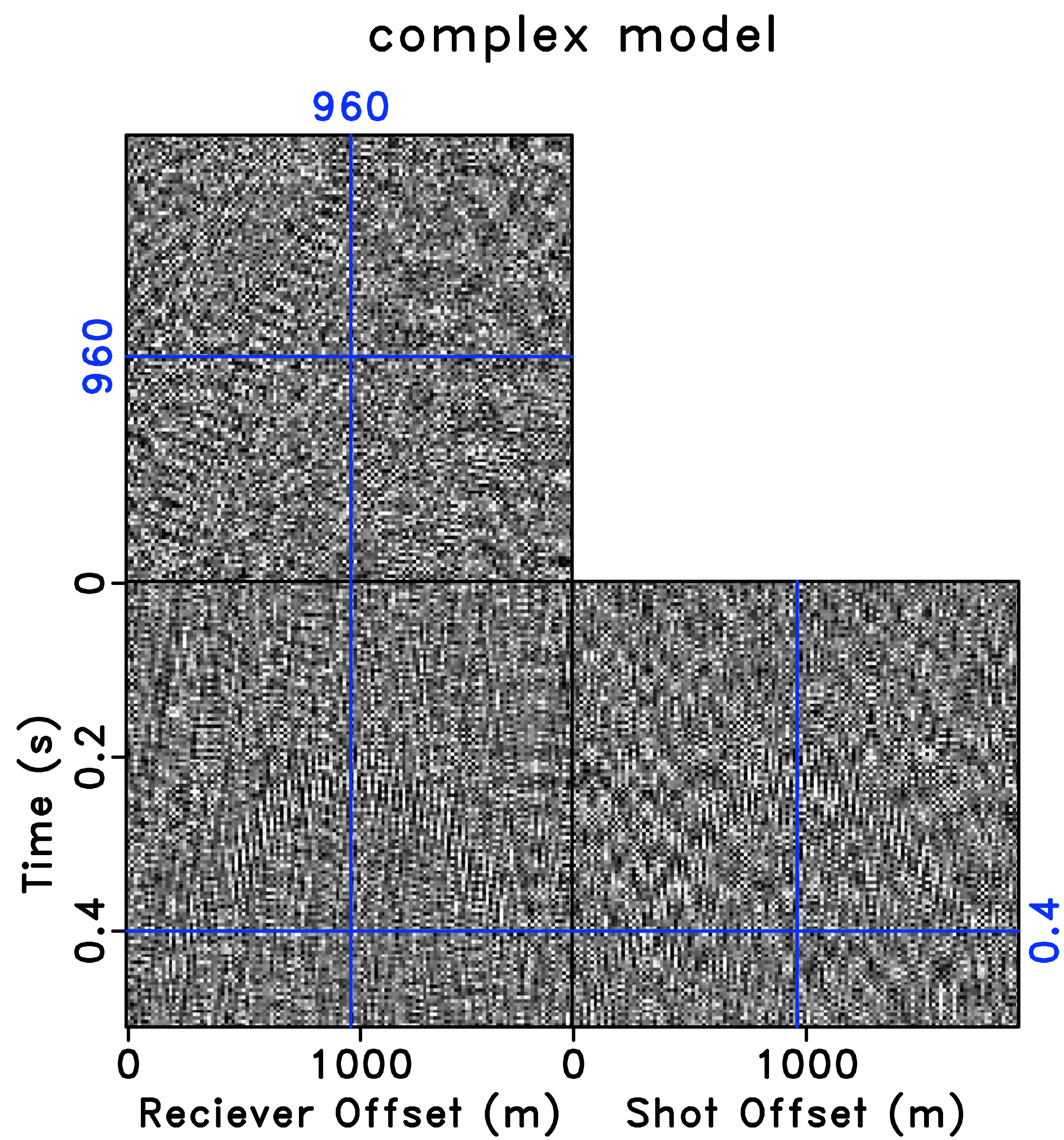




Seismic Laboratory for Imaging and Modeling

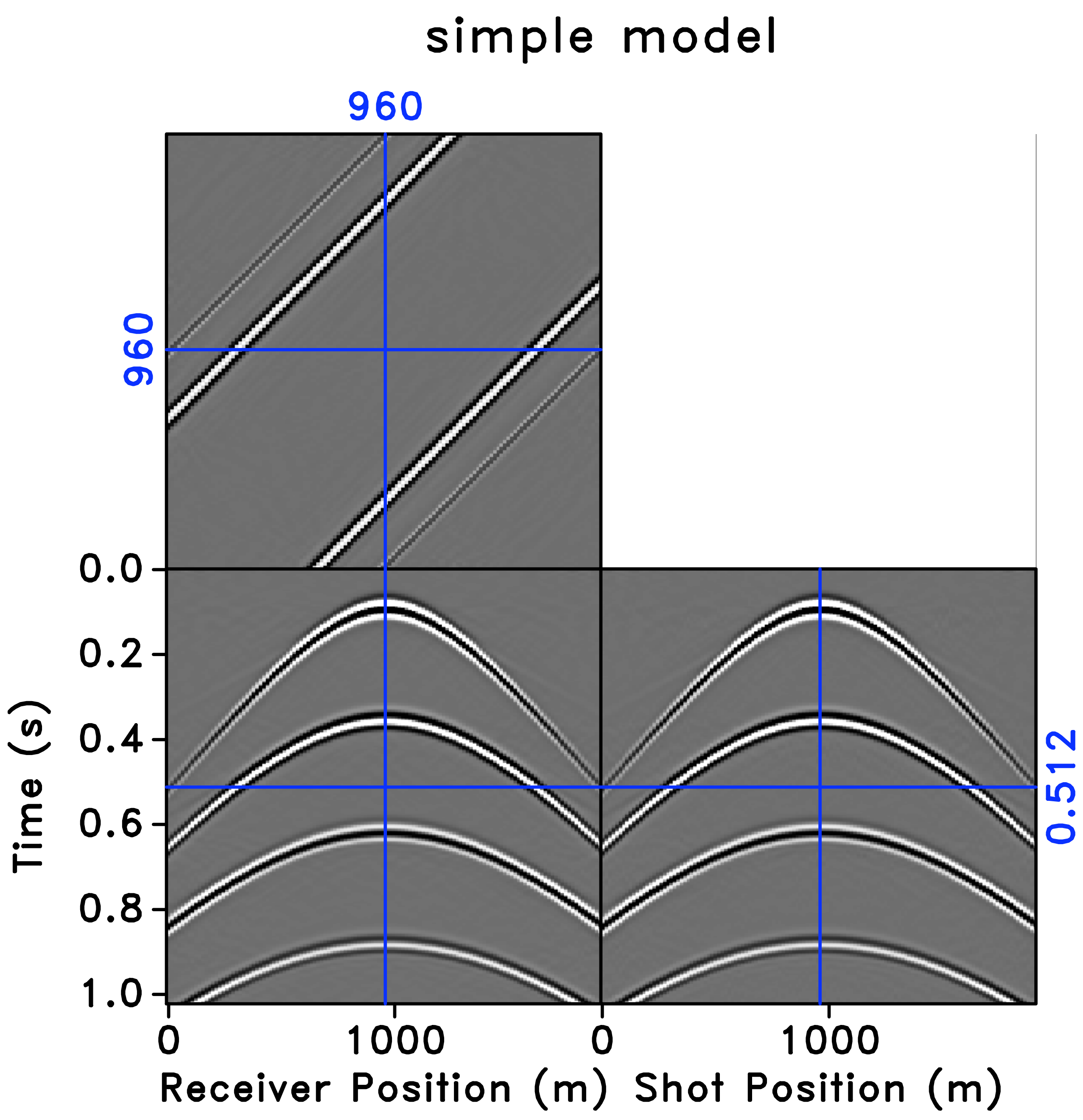
Matched filter



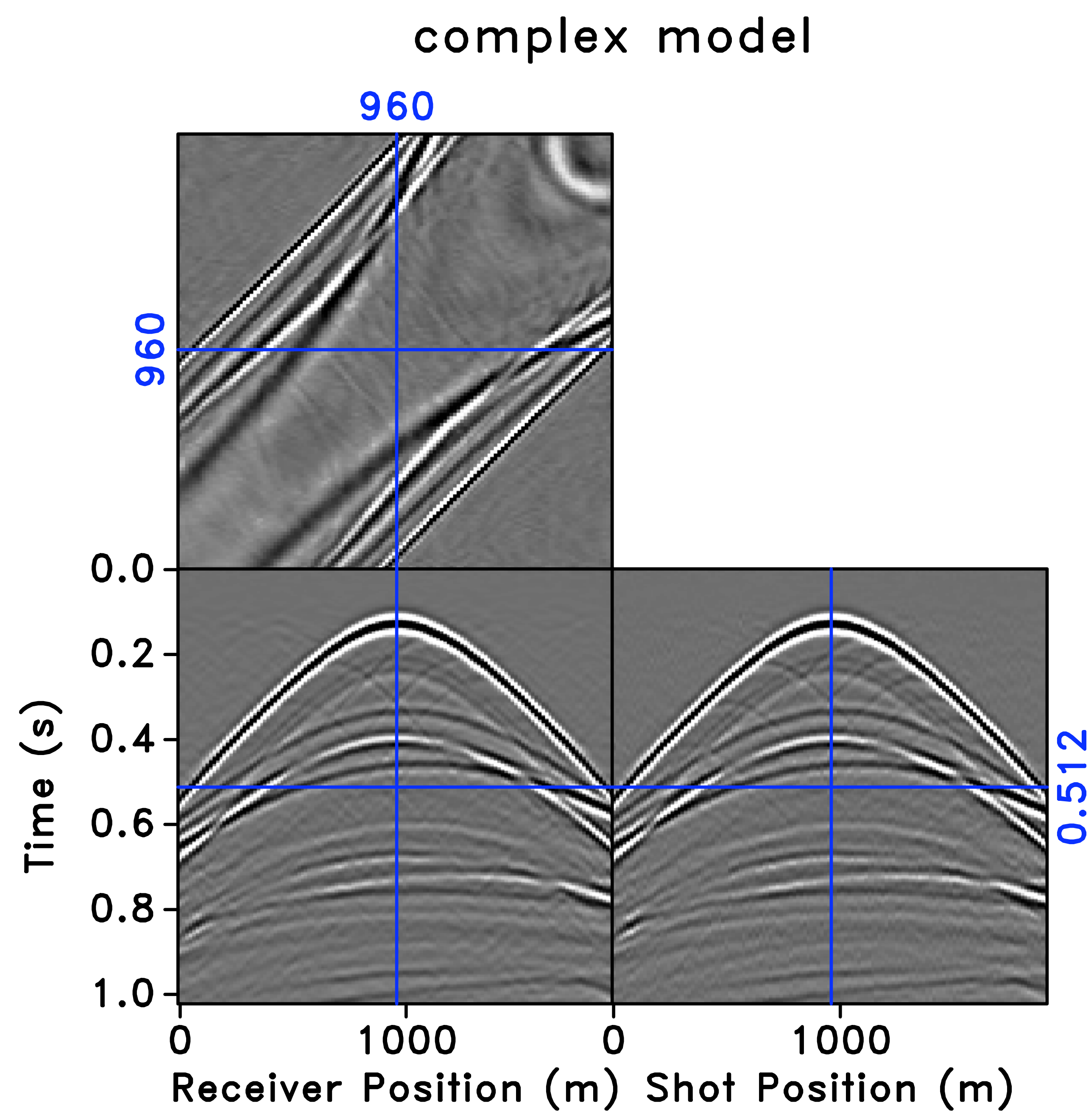


SLIM
Seismic Laboratory for Imaging and Modeling

Recovered data



28.1dB

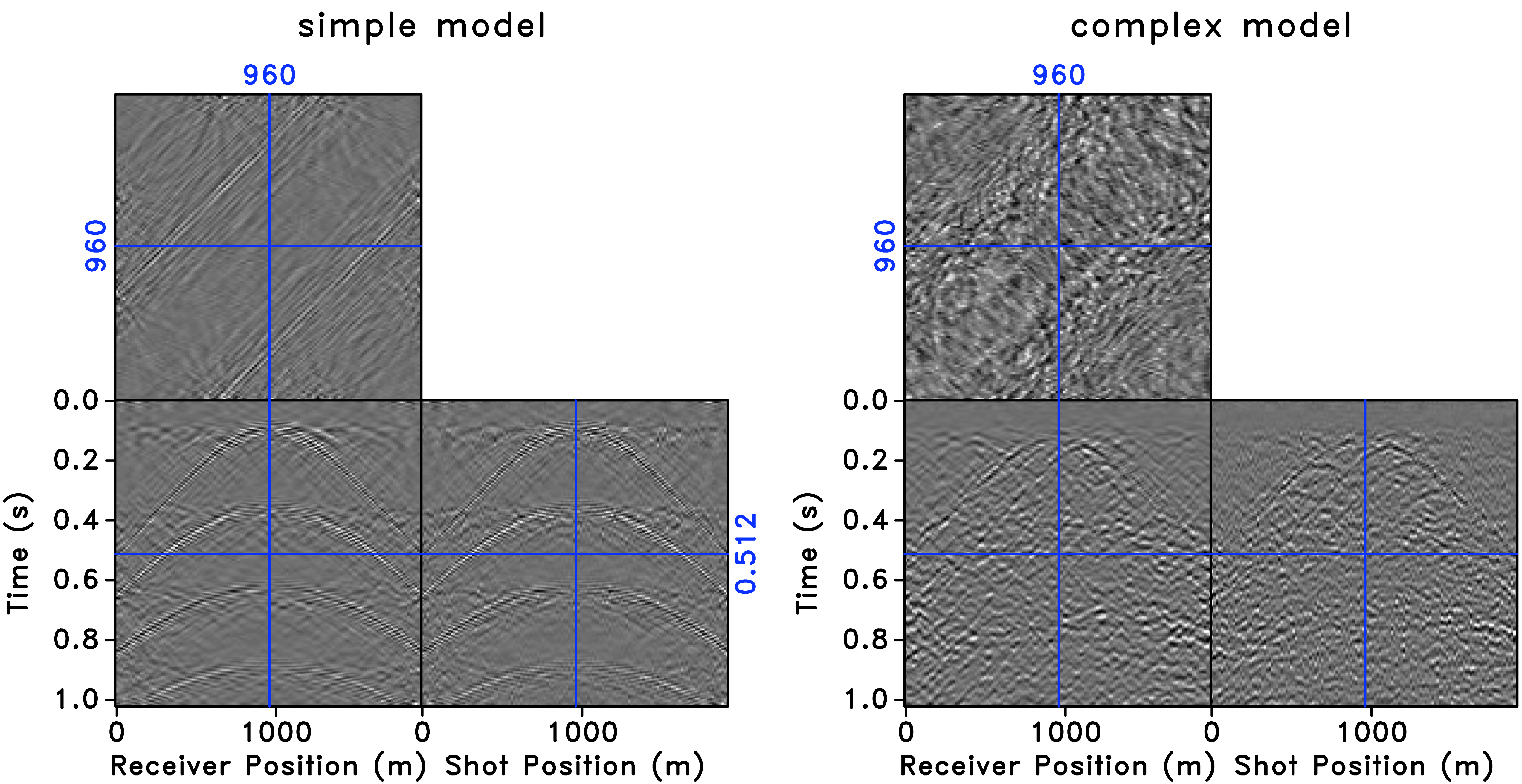


18.2dB

300 SPGL1 iteration

SLIM
Seismic Laboratory for Imaging and Modeling

Difference



28.1dB

18.2dB

300 SPGL1 iteration

SLIM Seismic Laboratory for Imaging and Modeling

Sample ratio SNR (dB)

problem size 2^{22}

Total computed data fraction

Frequencies / # Shots

	0.25	0.15	0.07
2	14.3	12.1	8.6
1	18.2	14.5	10.2
0.5	22.2	16.5	10.7

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

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

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

Complexity analysis cont'd

Cost sparsity promoting optimization problem dominated by matrix-vector products

- Sparsity transform is $\mathcal{O}(n^3 \log n)$
- 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. Random Convolutions by Romberg '08)

- 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

Observations

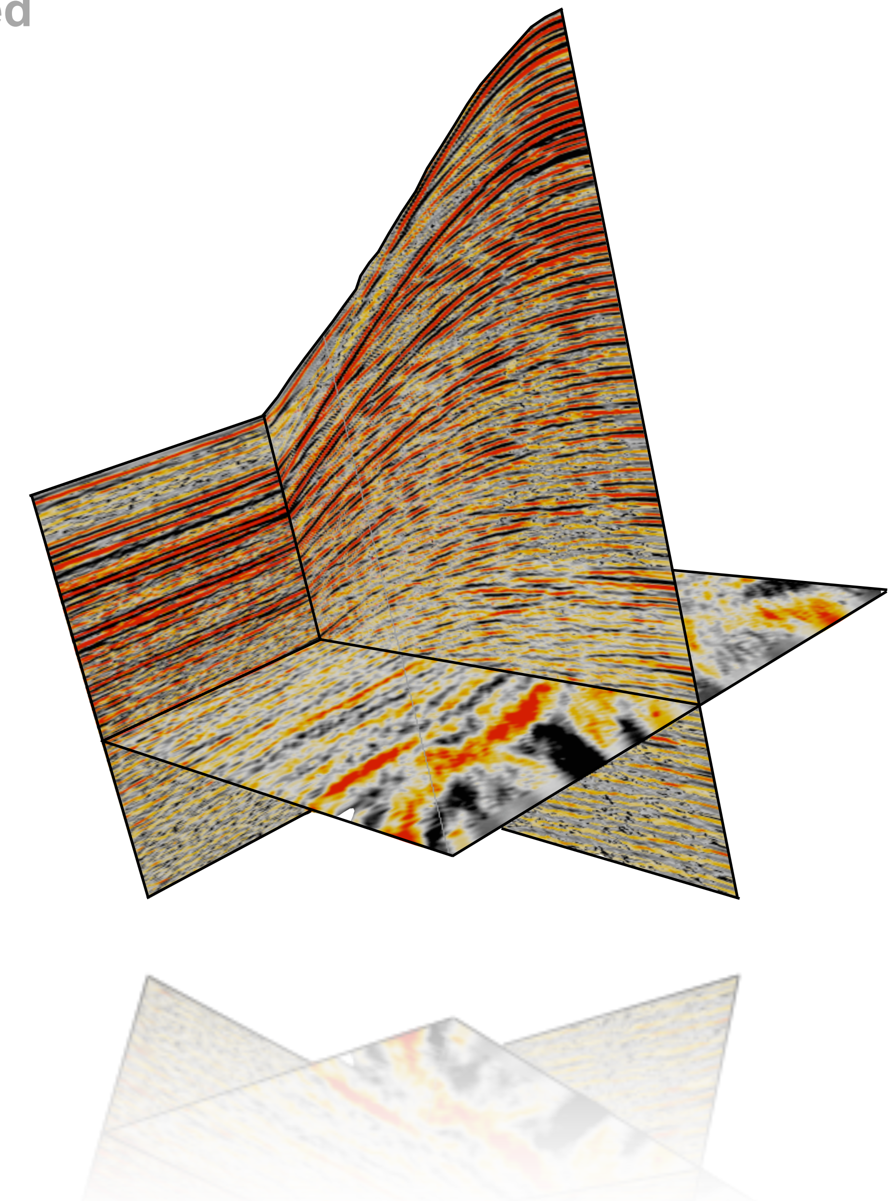
- Simultaneous sources lead to a significant reduction of simulation costs
  - reduction of the number of right-hand sides
  - extension of preconditioning to multiple right-hand sides

- According to CS computational costs are proportional to tranform-domain sparsity of the solution
  - CS projects the information to a smaller subspace
  - at the expense of solving a cheaper one-norm recovery
  - incorporation of sparsity promotion in PDE-constrained optimization
  - betting on development of large-scale convex optimization problems

- Question: Can these ideas be extended to model space CS?

Compressive imaging by wavefield inversion with group sparsity

Felix J. Herrmann, 2009. SEG abstract submitted today ...



Delft University of Technology
Delft, April 8, 2009

Motivation

- New insights ... Extended (linearize) forward modeling with image volumes => prestack migration
- Requires multi-dimensional correlations of wavefields
  - full matrix-matrix multiplies
- Seek a solution based on wavefield inversion (multi-dimensional deconvolution)
- *Joint sparsity promotion*
- *Success migration-velocity analysis*
- *Address the non-uniqueness problem ...*

Multi-source Adjoint state method

Solve PDE-constraint optimization problem

$$\min_{\mathbf{U}\in\mathcal{U},\,\mathbf{m}\in\mathcal{M}} \frac{1}{2}\|\mathbf{P}-\mathbf{D}\mathbf{U}\|_2^2 \quad \text{subject to} \quad \mathbf{H}[\mathbf{m}]\mathbf{U}=\mathbf{F}$$

Involves the solution of

$$\mathbf{H}[\mathbf{m}]\mathbf{U}=\mathbf{F}, \quad \text{and} \quad \mathbf{H}[\mathbf{m}]^*\mathbf{V}=\Delta\mathbf{R},$$

with

$$\Delta\mathbf{R} := \mathbf{D}^*\left(\mathbf{P}-\mathbf{D}\mathbf{U}\right)$$

Prestack imaging

Replace model updates with imaging an ***extended*** imaging volume

$$\mathbf{I}(m,h) = \mathbf{T}^{\Delta h}_{(x_s,x_r)\mapsto(m,h)}\left(\mathbf{U}\mathbf{V}^*\right),$$

with $m = \frac{1}{2}(x_s + x_r) \quad \text{and} \quad h = \frac{1}{2}(x_s - x_r)$

Penalize ***defocusing*** via minimizing

$$\|\mathbf{H}\mathbf{I}(\cdot,\, h)\|_2$$

with H an operator that increasingly penalizes the non-zero offsets.

Remark: conventional imaging principle (same for time)

$$\delta\mathbf{m} = \mathbf{I}(\cdot, h = 0)$$

Imaging by wavefield inversion

Replace multi-dimensional ***correlation***-based imaging principle by imaging through multi-dimensional ***deconvolution***

Add a focusing principle

Solve for the extended image from

$$\begin{aligned} \mathbf{U}^*\mathbf{S}^*\mathbf{X} &\approx \mathbf{V}^* \\ \mathrm{H}\mathbf{X} &\approx \mathbf{0} \end{aligned} \qquad \text{focuses}$$

with the sparsifying transform (curvelets/wavelets along depth-midpoint)

$$\mathbf{S} := \mathrm{vec}^{-1}\left(\left(\mathbf{Id} \otimes \mathbf{C}\right)\mathbf{T_0}\right)\mathrm{vec}\left(\cdot\right)\cdot$$

and $\mathbf{T_0}$ source/receiver-midpoint offset mapping supplemented with the imaging condition for t=0 (adjoint of the summing over frequency)

Combine with sparsity promotion.

Joint sparsity promotion

Recent generalization of sparsity promotion to ***joint***-sparsity promotion [van den Berg and Friedlander, '08]

Solution of mixed (1,2)-norm matrix-valued problem:

$$\tilde{\mathbf{X}} = \arg\min_{\mathbf{X}} \|\mathbf{X}\|_{1,2} \quad \text{subject to} \quad \|\mathbf{A}\mathbf{X} - \mathbf{B}\|_{2,2} \leq \sigma,$$

with $$\|\mathbf{X}\|_{1,2} := \sum_{i \in \text{rows}(\mathbf{X})} \|\text{row}_i(\mathbf{X})^*\|_2$$

and $$\|\mathbf{X}\|_{2,2} := \left(\sum_{i \in \text{rows}(\mathbf{X})} \|\text{row}_i(\mathbf{X})^*\|_2^2 \right)^{\frac{1}{2}}.$$

Compressive imaging

Subsample using CS

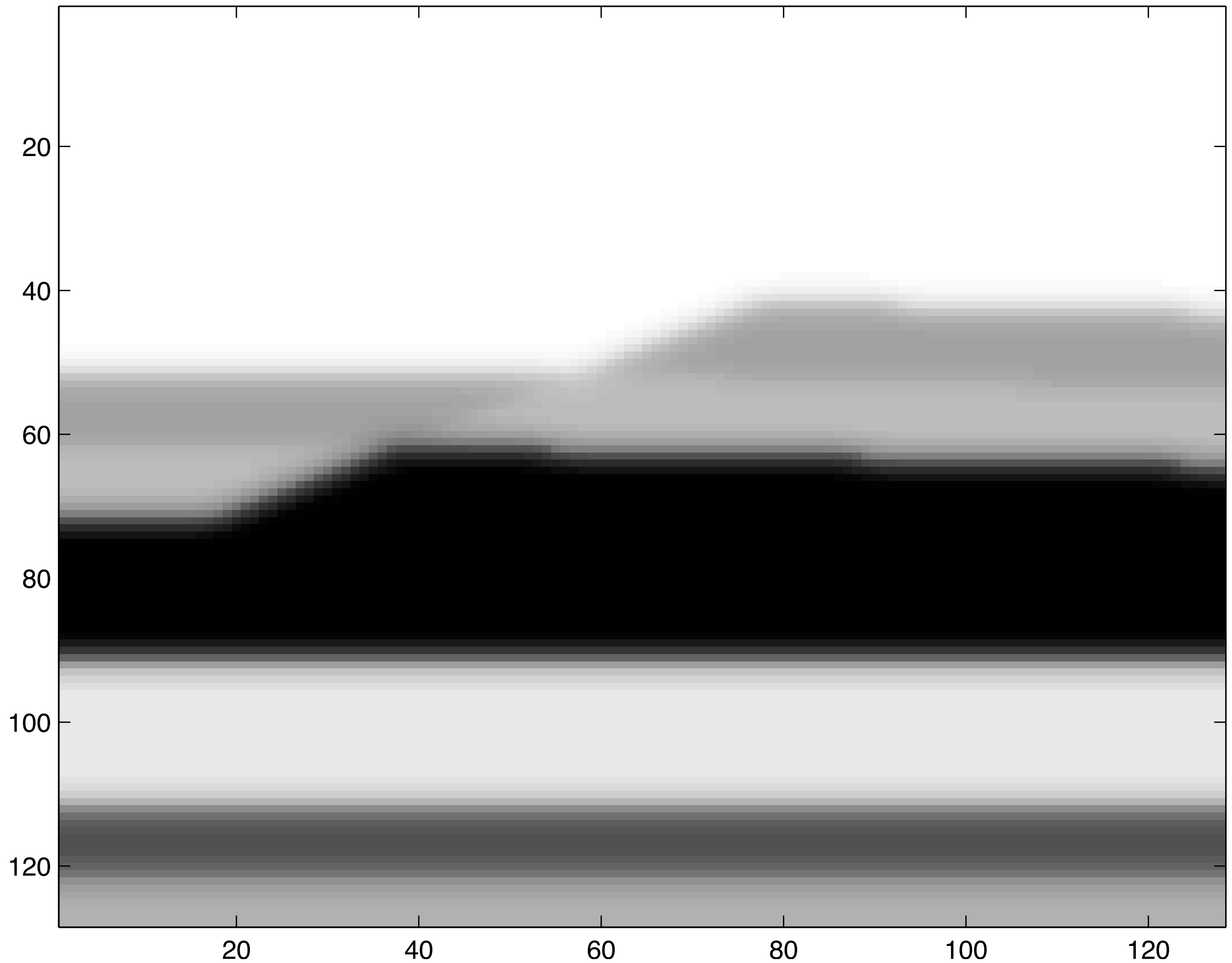
$$\mathbf{RM} := \overbrace{\begin{bmatrix} \mathbf{R}_1^{\sigma} \otimes \mathbf{R}_1^{\rho} \otimes \mathbf{R}_1^{\zeta} \\ \vdots \\ \mathbf{R}_{n'_f}^{\sigma} \otimes \mathbf{R}_{n'_f}^{\rho} \otimes \mathbf{R}_{n'_f}^{\zeta} \end{bmatrix}}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_3^* \left(e^{\hat{i}\boldsymbol{\theta}}\right)\right) \mathbf{F}_3}^{\text{random phase encoder}} ,$$

with

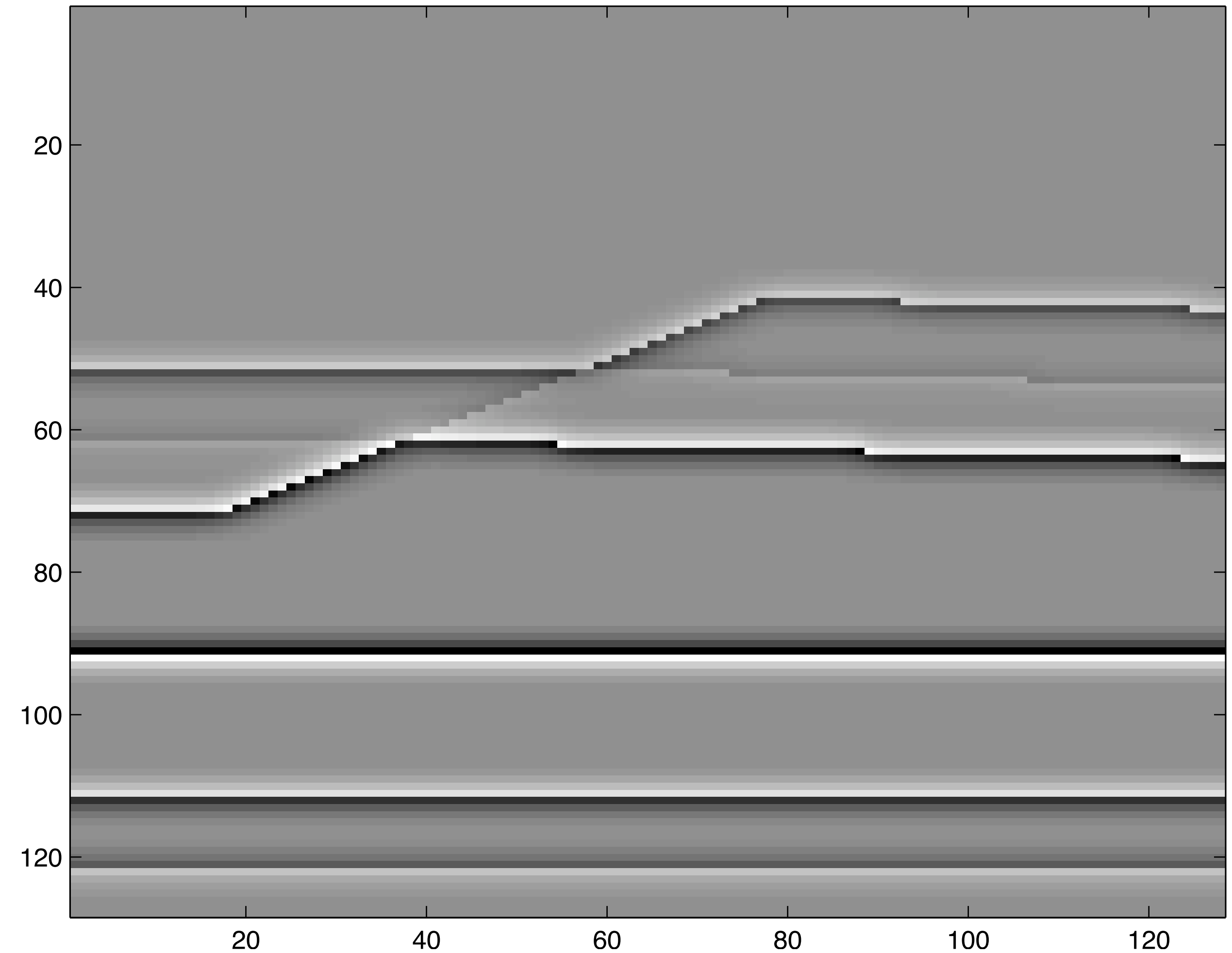
$$n'_f \times n_\sigma \times n_\rho \times n_\zeta \ll n_f \times n_s \times n_r \times n_z$$

Model-space CS subsampling along source, receiver, and depth coordinates.

Example

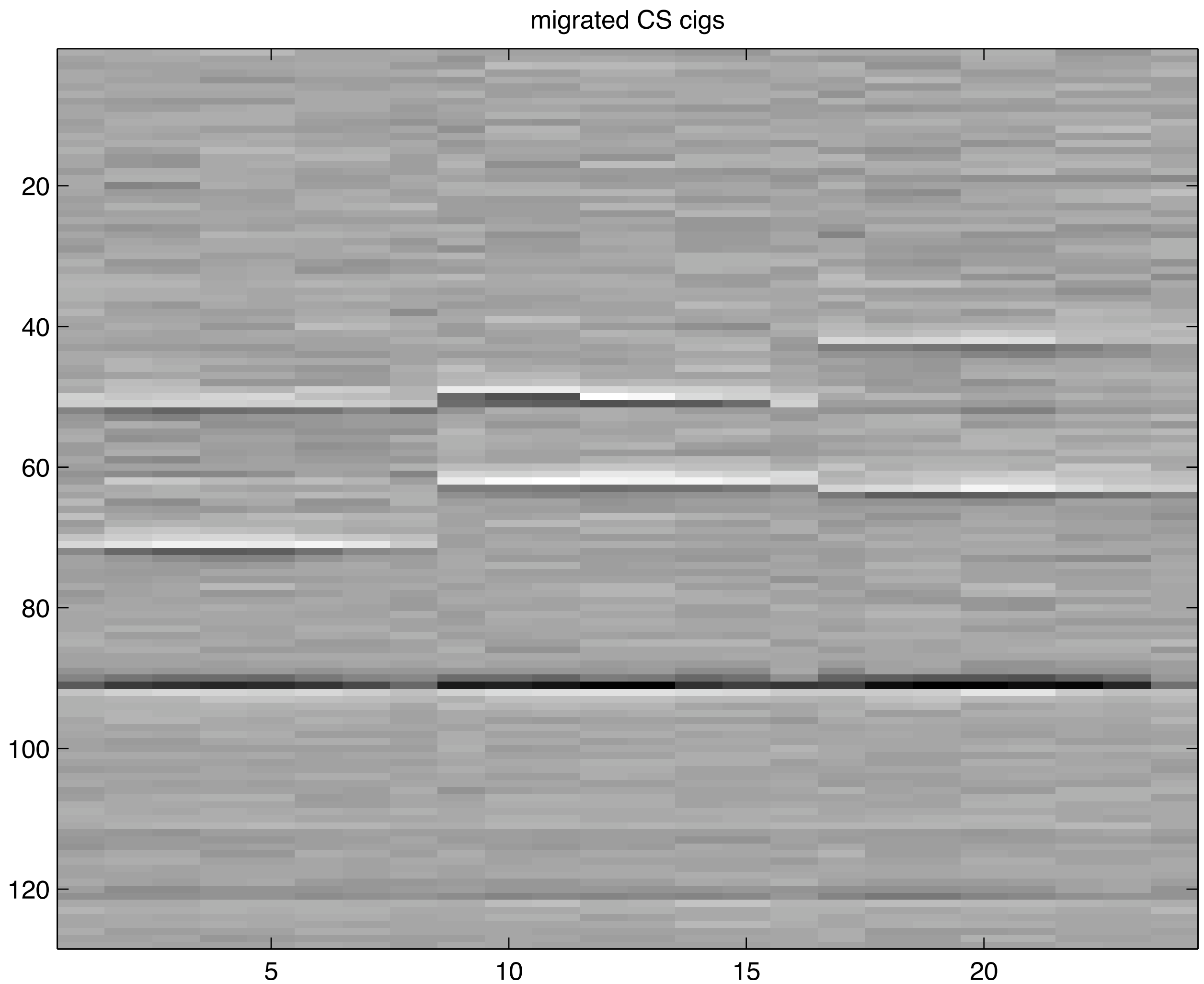


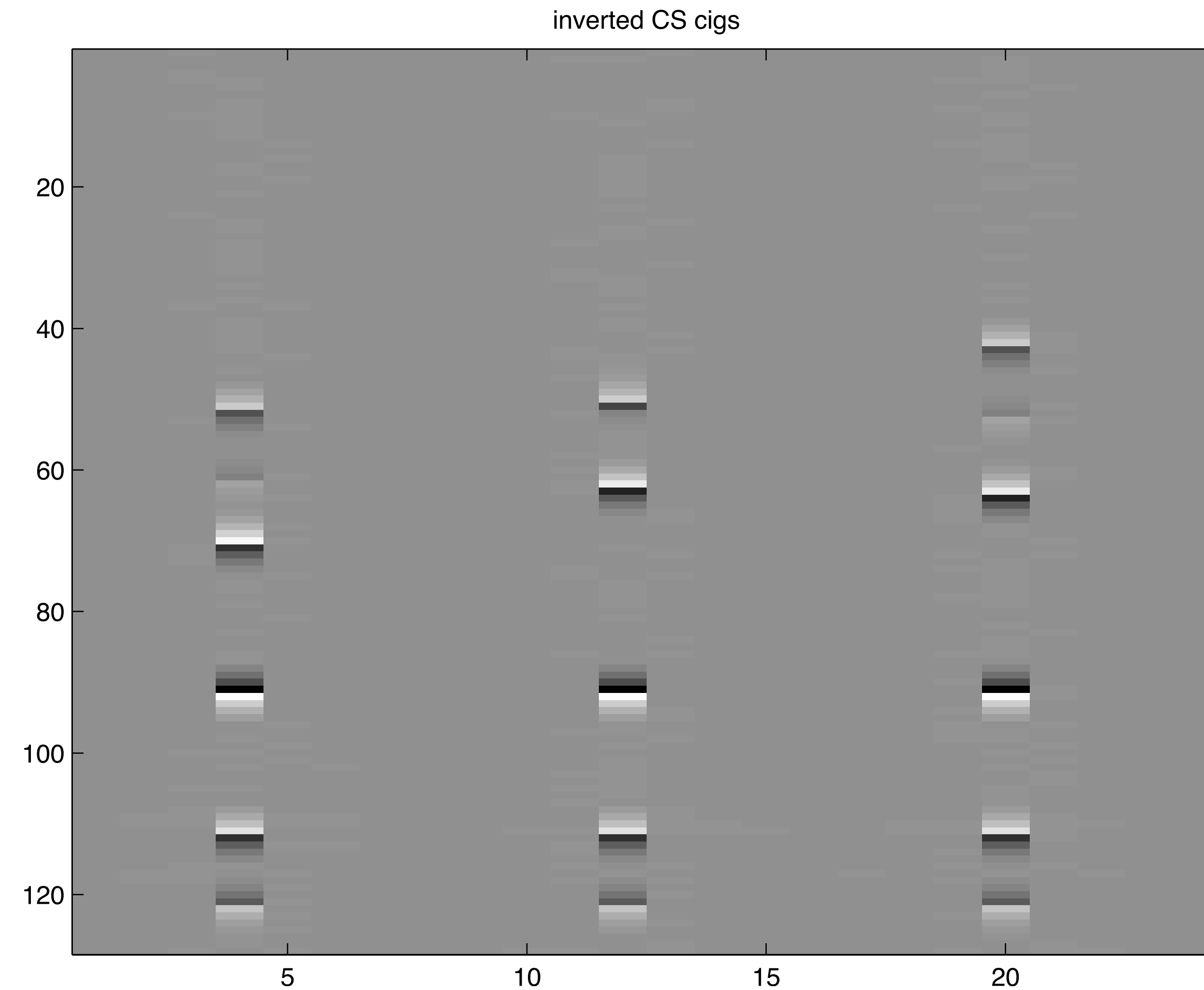
background velocity model



perturbation

Example

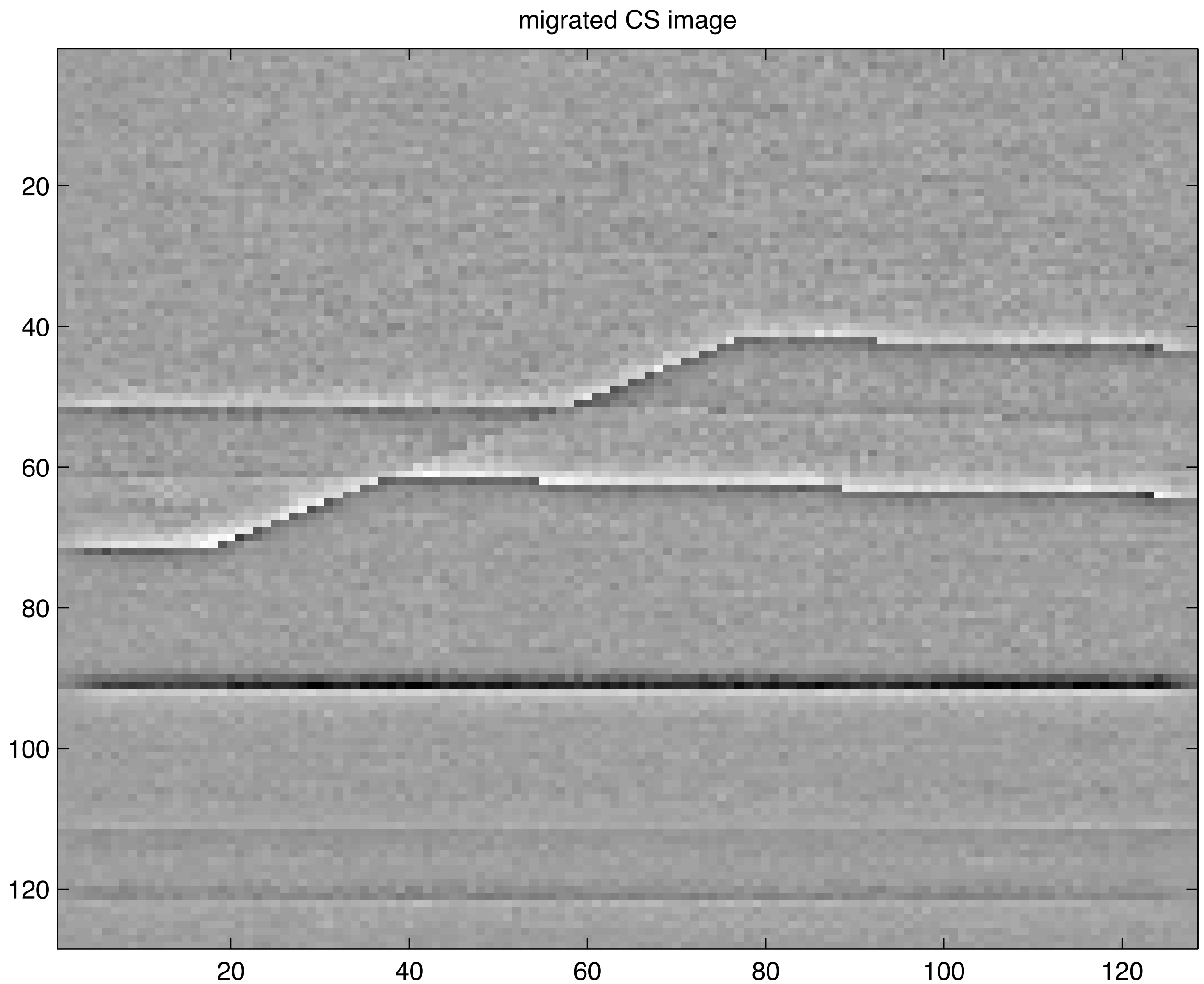




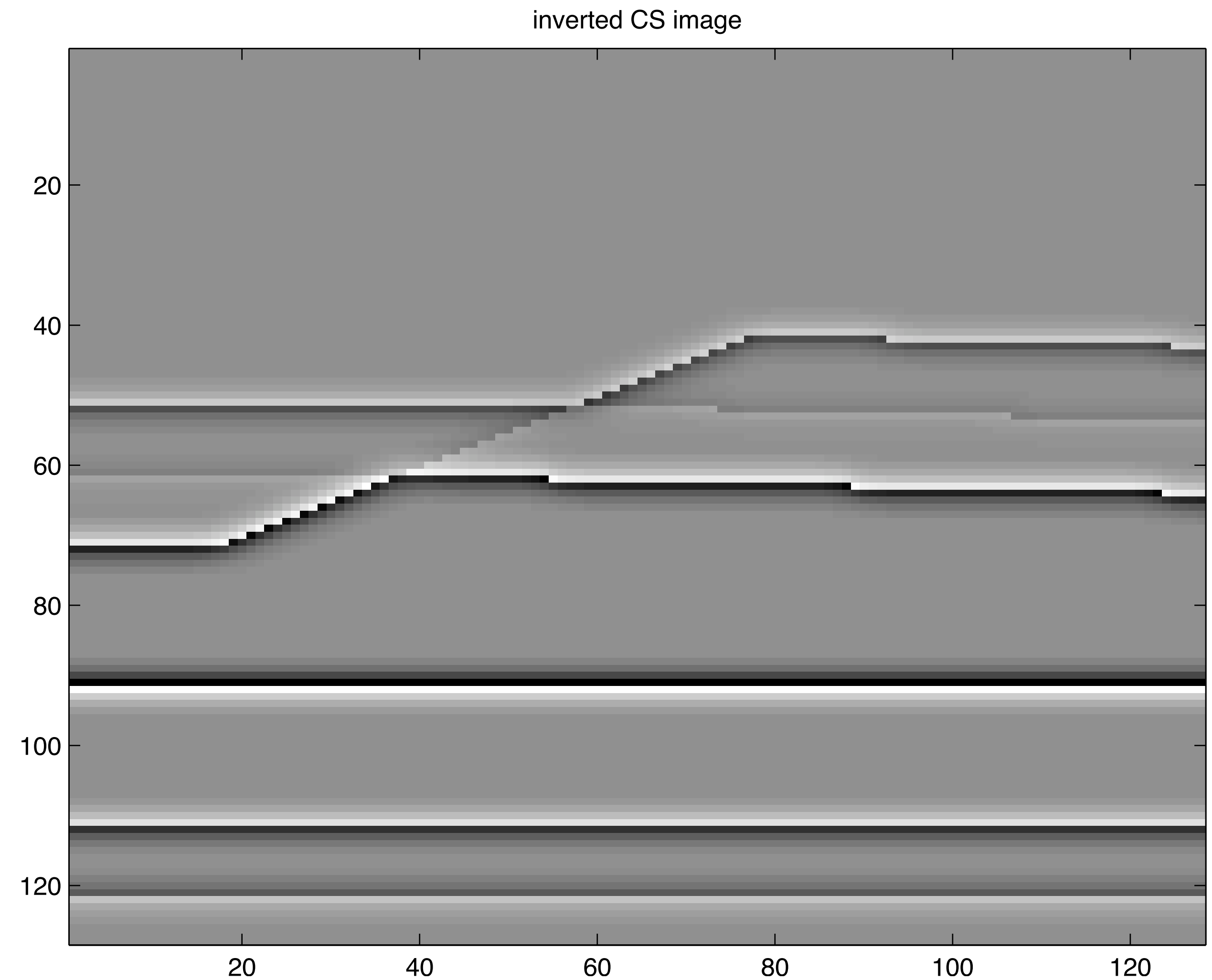
correlation based **wavefield inversion**

Common-image gathers

Example



correlation based



wavefield inversion

Conclusions & outlook

- **CS** provides a **new** *linear sampling paradigm*
  - ***degree*** of *subsampling commensurate* with transform-domain ***sparsity***
  - subsampling of seismic data volumes
    - missing source-receiver locations
    - simultaneous acquisition
  - *subsampling* of solutions to PDEs

- **CS** leads to
  - "acquisition" of *smaller* data volumes that carry the ***same*** *information* or
  - to ***improved*** *inferences* from data using the *same* resources

- Bottom line: **acquisition & numerical modeling costs** are **no** longer determined by the ***size*** of the ***discretization*** but by the ***transform-domain compressibility*** of the **solution ...**

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 and Yang Liu for Madagascar (rsf.sf.net)
- E. Candes and the Curvelab team

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.

slim.eos.ubc.ca

and... Thank you!