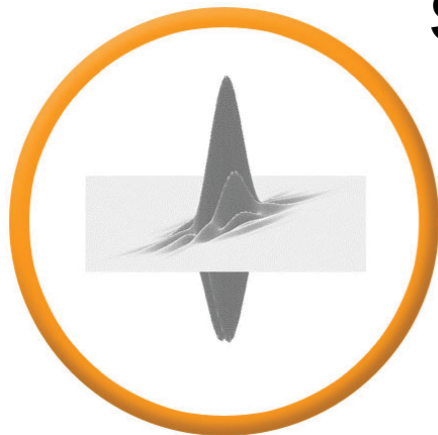




Compressive seismic imaging with simultaneous acquisition

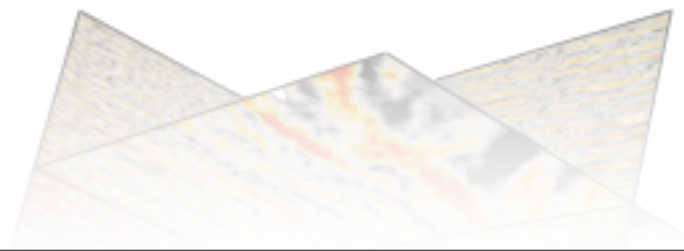
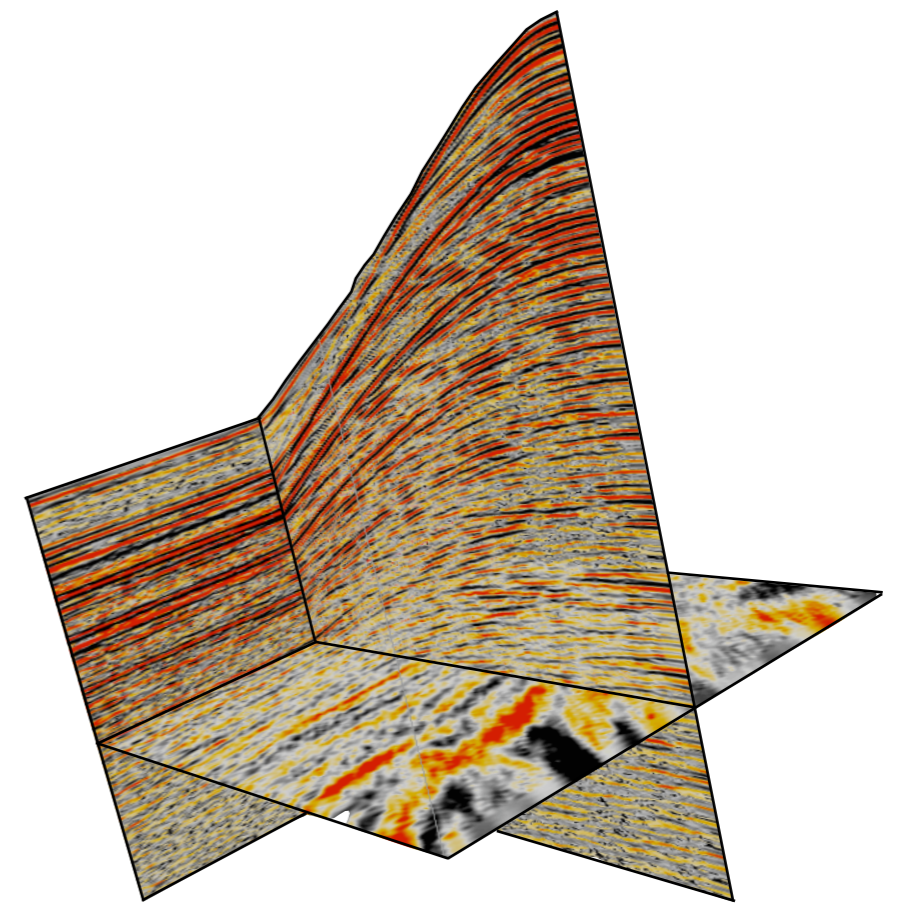


Felix J. Herrmann*

fherrmann@eos.ubc.ca

Joint work with Yogi Erlangga, and Tim Lin

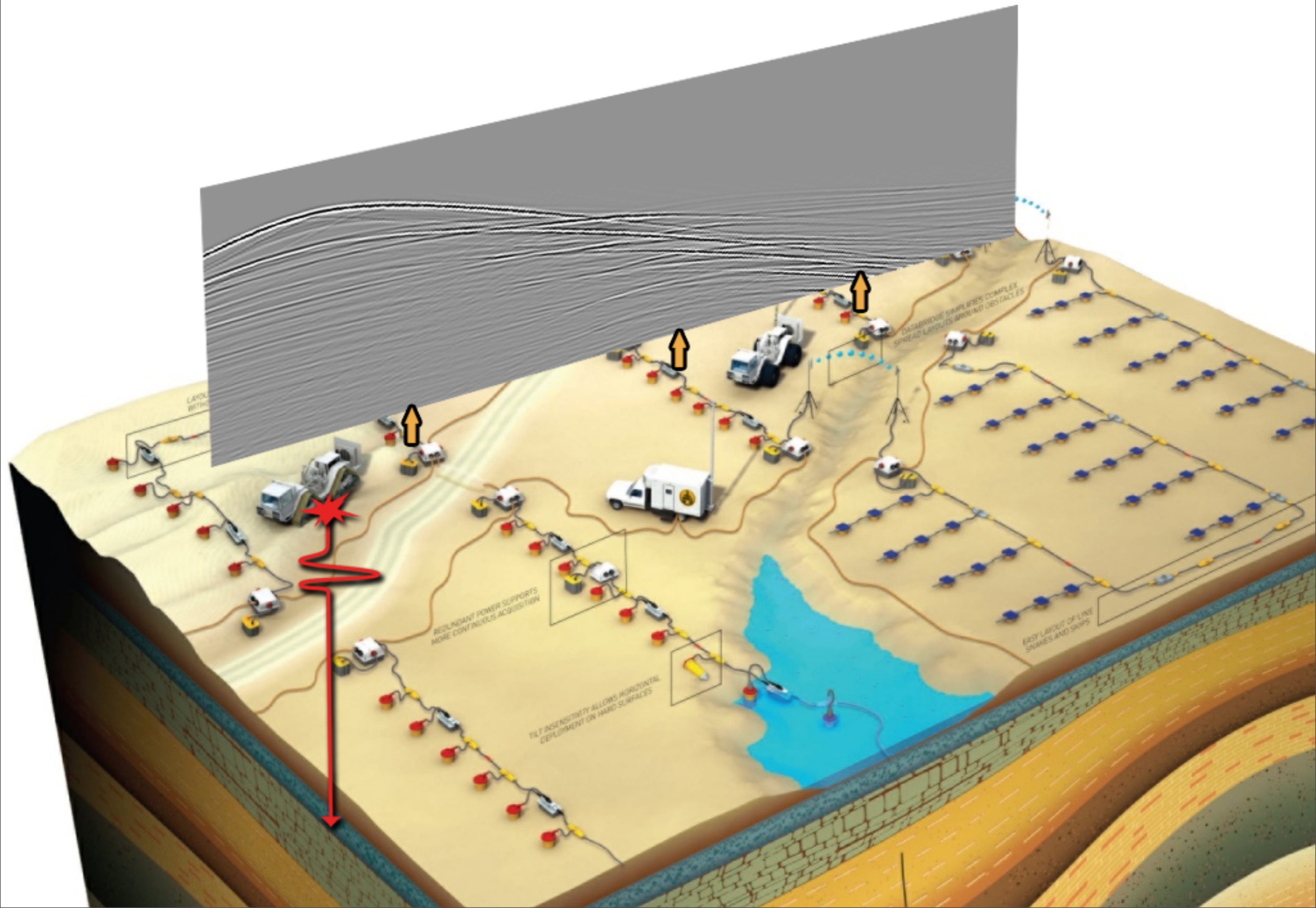
***Seismic Laboratory for Imaging & Modeling**
Department of Earth & Ocean Sciences
The University of British Columbia



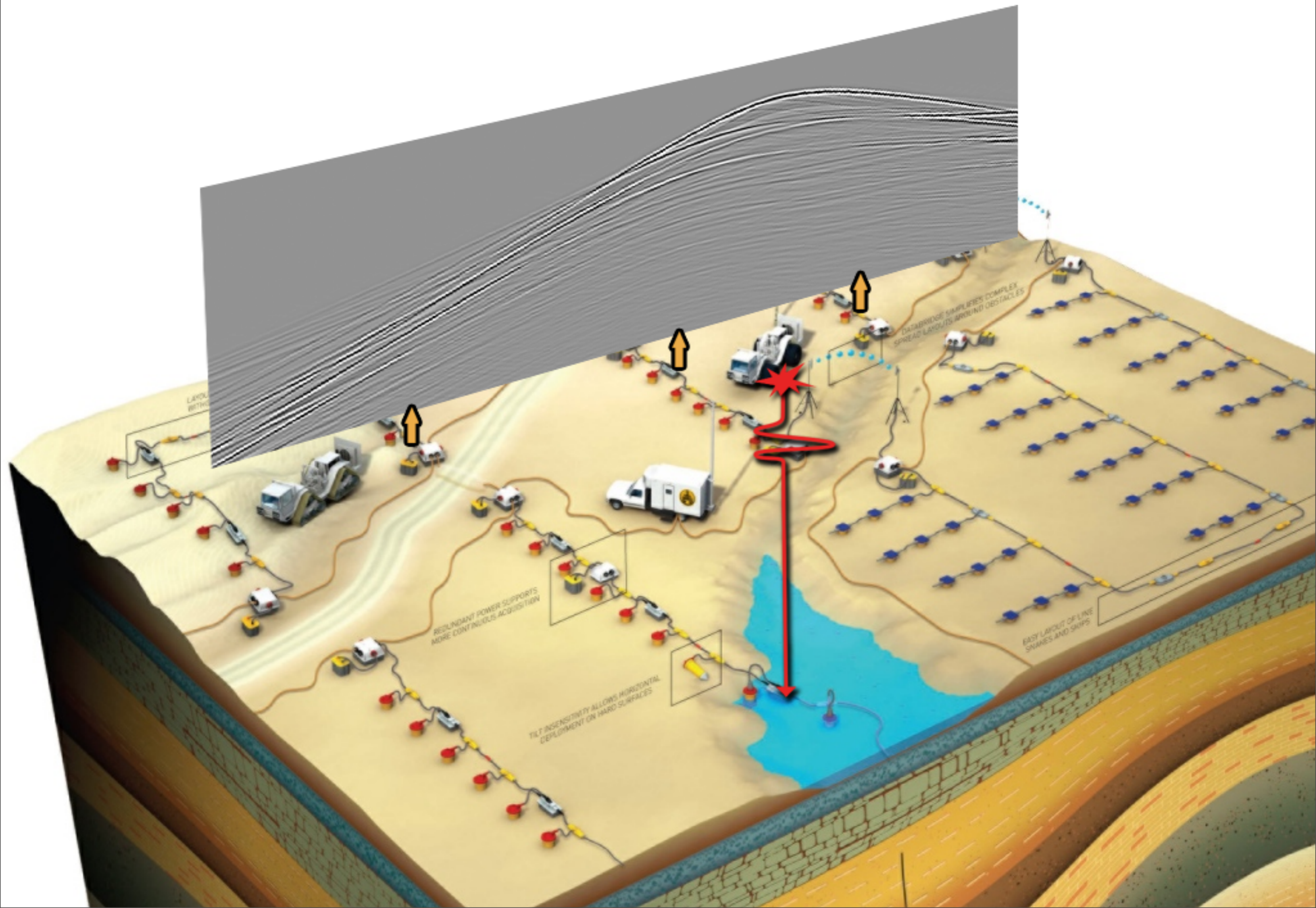
Seismic acquisition



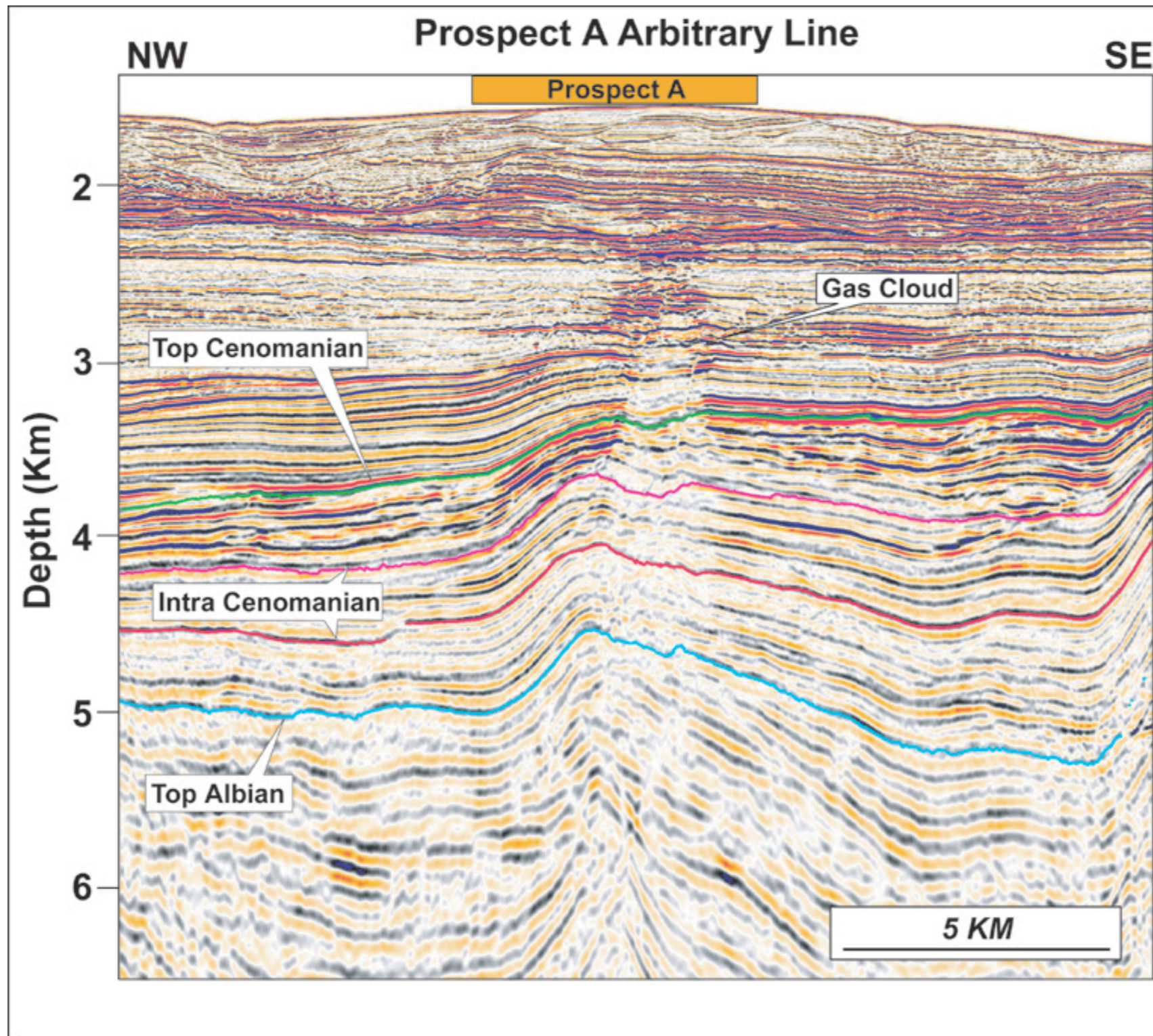
Individual shots



Individual shots



After imaging



Observations

- Seismic imaging methods are mostly based on *linearizations*
- Seismic imaging methods are despite the *spectral gap* able to
 - locate major singularities
 - assign some sense of reflection strength
- Seismic images
 - are derived from *multiexperiment* data (petabytes) \Leftrightarrow *redundancy*
 - permit sparse representation by multiscale & multidirection transforms that capture the “wavefront set” of the subsurface reflectors (e.g. curvelets)
- Seismic images do *not capture the whole picture!*
- *There is a push for full waveform inversion ...*

Seismic imaging & inversion

***Multiexperiment* PDE-constrained optimization problem:**

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

P = Total multi-source and multi-frequency data volume

D = Detection operator

U = Solution of the Helmholtz equation

H = Discretized multi-frequency Helmholtz system

Q = Unknown seismic sources

m = Unknown model, e.g. $c^{-2}(x)$

Wavefield simulations

Based on discretization of the Helmholtz equation:

$$\mathcal{H}u = -\Delta u - \omega^2 mu = q$$

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

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

Δf frequency sample interval

Adjoint state methods [Plessix '06 & many others]

For each *separate* source \mathbf{q} solve the **unconstrained problem**:

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

where ***model updates*** \Leftrightarrow ***migrated image***

$$\delta\mathbf{m} = \Re \left(\sum_{\omega} \omega^2 \sum_s \bar{\mathbf{u}} \odot \mathbf{v} \right) = \mathbf{K}^*[\mathbf{m}, \mathbf{Q}]\delta\mathbf{d}$$

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

involve single ***implicit*** solves of Helmholtz system

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

with

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

Challenges: there are many ...

Helmholtz system is *indefinite & ill conditioned* => lack of convergence
indirect Krylov solvers

Multiexperiment setup with *multiple right-hand-sides* is *computationally prohibitive* as part of *iterative* Newton methods

Inversion problem can be both *over-* and *underdetermined* [Symes, '09]

- data cannot be explained fully
- there are local minima
- many velocity models may explain data within some error

Proposed ideas to tackle *multimodality* by *extensions & focusing* make the situation worse by additional *degrees of freedom*

Indirect solver

Preconditioner [Erlangga, Oosterlee, Vuik, 2006]

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

Deflation operator [Erlangga, Nabben, '08, FJH, Erlangga, '08]

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

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

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

Similar computational complexity as TDFD ...

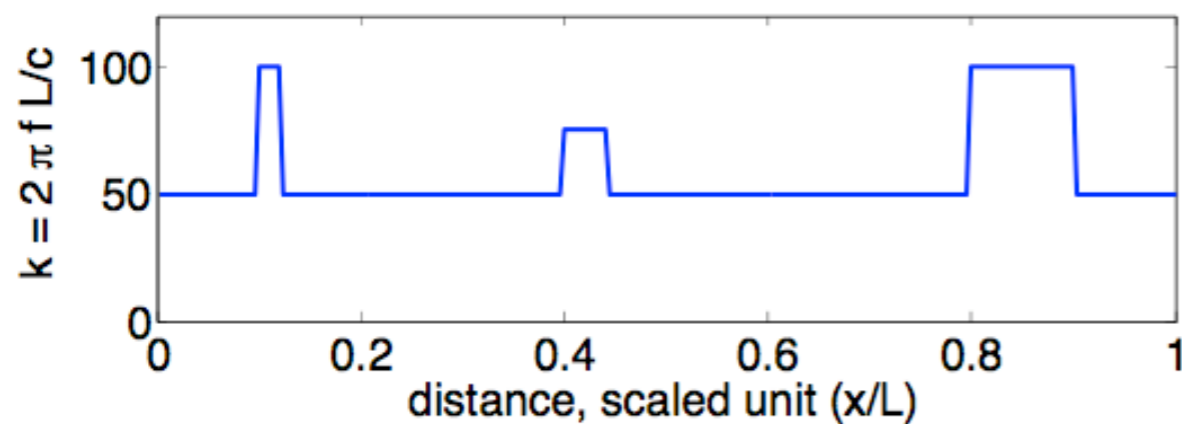
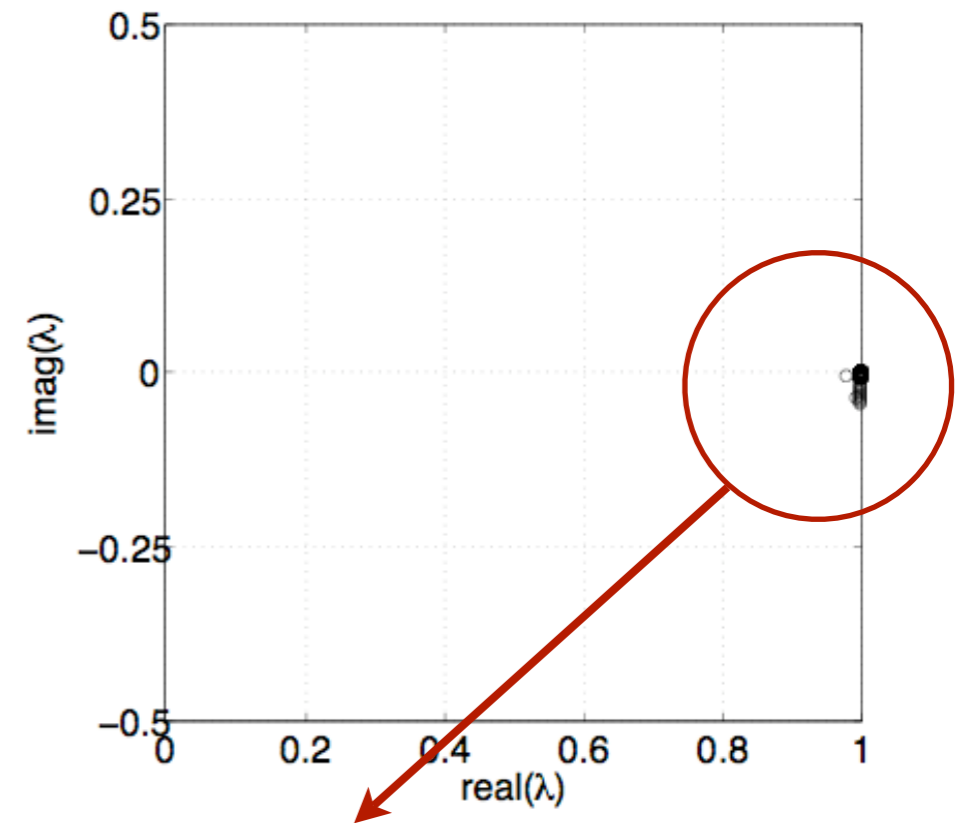
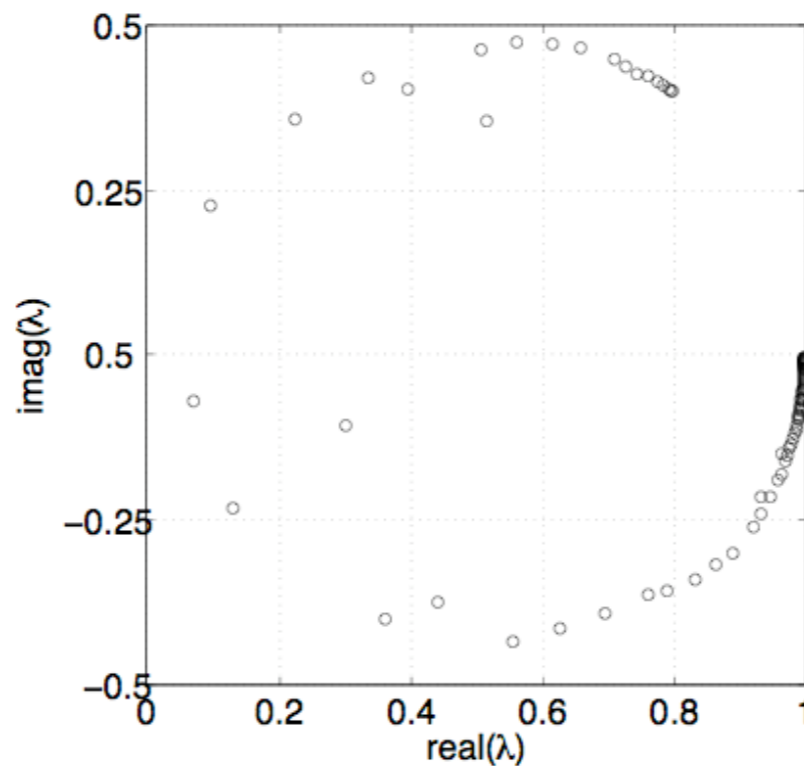
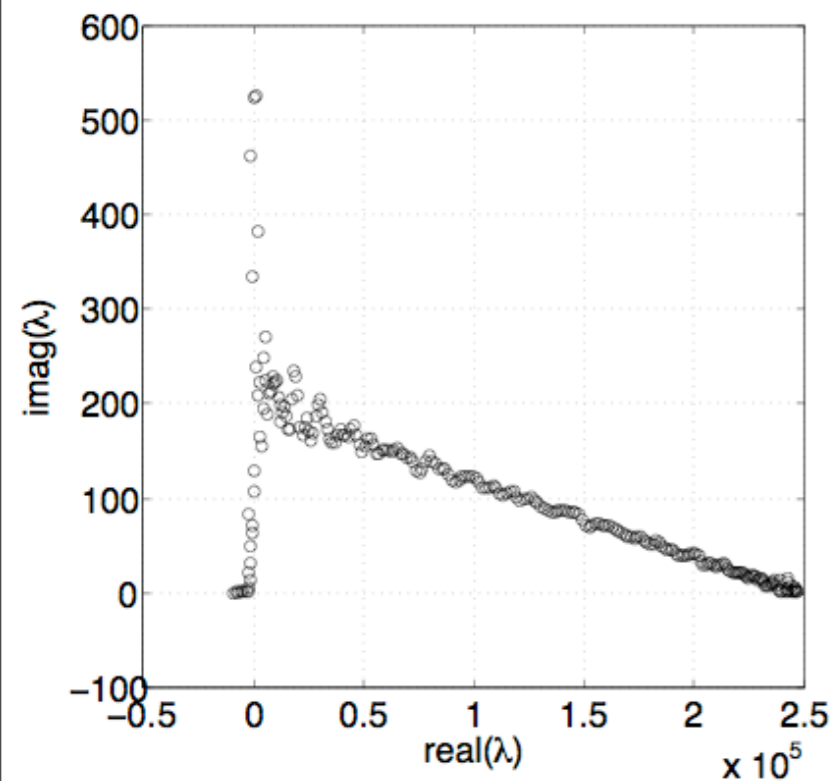
Behavior eigenvalues

1D non-constant wavenumber k , hard model $k = (50, 100)$

H

HM^{-1}

$HM^{-1}Q$



Clustering around one
For constant, smooth, or hard model, one can expect the same convergence rate

Challenges: there are many ...

- ✓ Helmholtz system is *indefinite* & *ill conditioned* => lack of convergence *indirect* Krylov solvers

Multiexperiment setup with *multiple right-hand-sides* is *computationally prohibitive*

Inversion problem can be both *over-* and *underdetermined*

- data cannot be explained fully
- there are local minima
- many velocity models may explain data within some error

Proposed plans to tackle *multimodality* by *extensions & focusing* make the situation worse by additional *degrees of freedom*

System-size reduction

- Apply CS to reduce cost of *wavefield* simulation with Helmholtz
 - use *simultaneous sources* instead of *separated sources*
 - leverage transform-domain sparsity & randomized subsampling by **one-norm sparsity promotion**
 - reduce size Helmholtz system
 - sources (number of right-hand sides)
 - angular frequencies (number of blocks)

- Apply CS to reduce cost of computing *image volumes* by multi-dimensional correlations via *explicit* matrix-matrix multiplies
 - randomize and subsample wavefields in **model space**
 - leverage transform-domain sparsity and focusing in the *model space* by **joint sparsity promotion with mixed (1,2) norms**
 - reduce costs of storage and explicit matrix-matrix multiplies
 - sources (right-hand sides), receivers, depth
 - angular frequencies (blocks)

Relation to existing work

- **Simultaneous & continuous acquisition:**

- *Efficient Seismic Forward Modeling using Simultaneous Random Sources and Sparsity* by N. Neelamani and C. Krohn and J. Krebs and M. Deffenbaugh and J. Romberg, '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.

- **Full-waveform inversion:**

- *3D prestack plane-wave, full-waveform inversion* by Vigh and Starr, '08

- **Wavefield extrapolation:**

- *Compressed wavefield extrapolation* by T. Lin and F.J.H, '07
- *Compressive wave computations* by L. Demanet (SIA '08 MS79 & Preprint)

Tools

Compressive sensing based on Johnson-Lindenstrauss embeddings

- *Compressive sensing* [Donoho, '06, Candes, Romberg, Tao, '06]

$$\mathbf{b} = \mathbf{RM}\mathbf{x} \quad \text{[randomized subsampling]}$$

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{RM}\mathbf{x} - \mathbf{b}\|_2 \leq \sigma$$

$$\tilde{\mathbf{x}} \approx \mathbf{x}$$

Fast matrix computations based on Johnson-Lindenstrauss embeddings

- Improved Approximation Algorithms for Large Matrices via Random Projections by Tamás Sarlós, '08

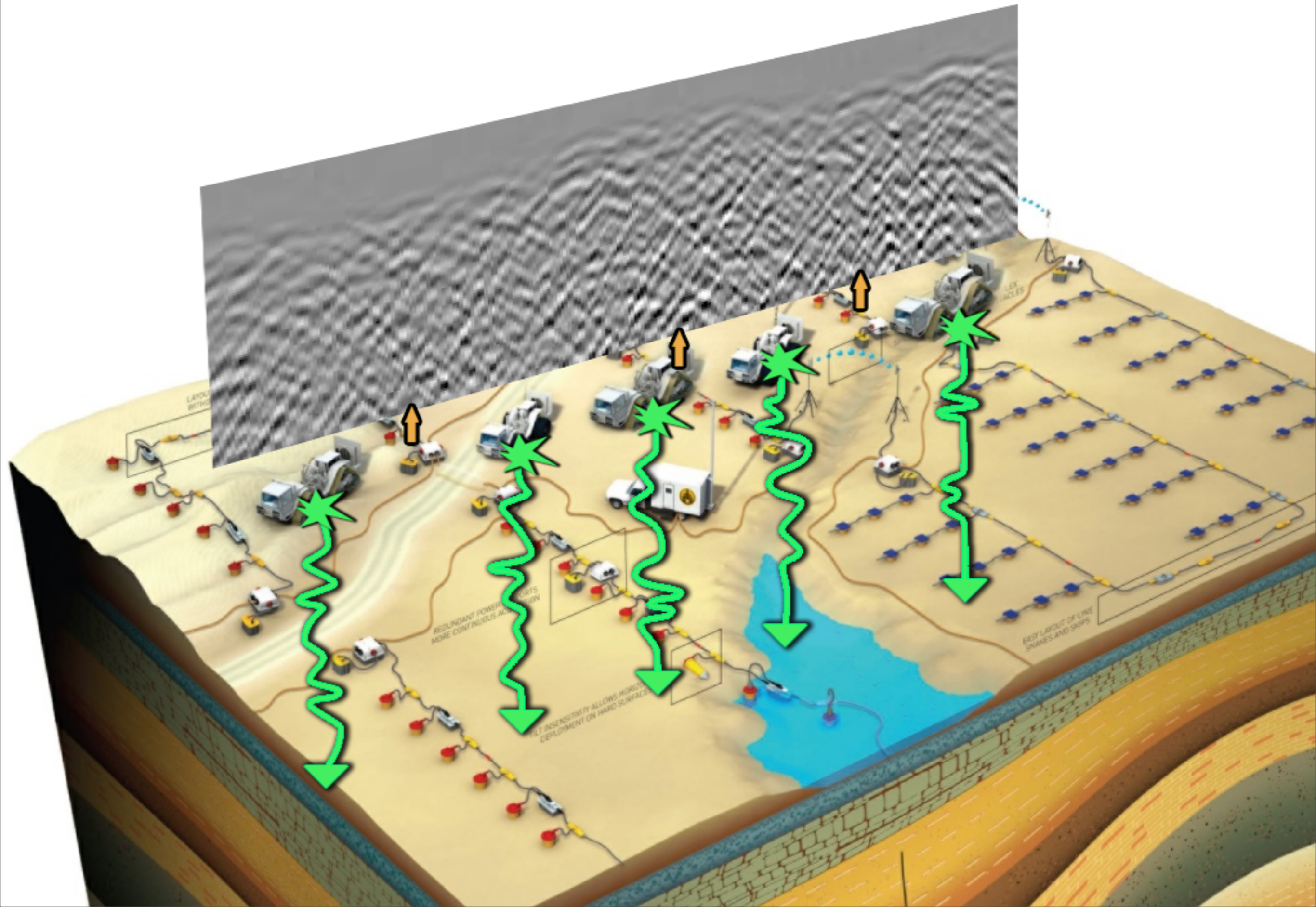
$$\mathbf{AB} \approx \mathbf{A} (\mathbf{RM})^* (\mathbf{RM}) \mathbf{B}$$

Joint sparsity-promotion with mixed (1,2) norm minimization

- *Joint-sparse recovery from multiple measurements* by E. van den Berg and M. Friedlander, '09

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

Simultaneous & continuous sources



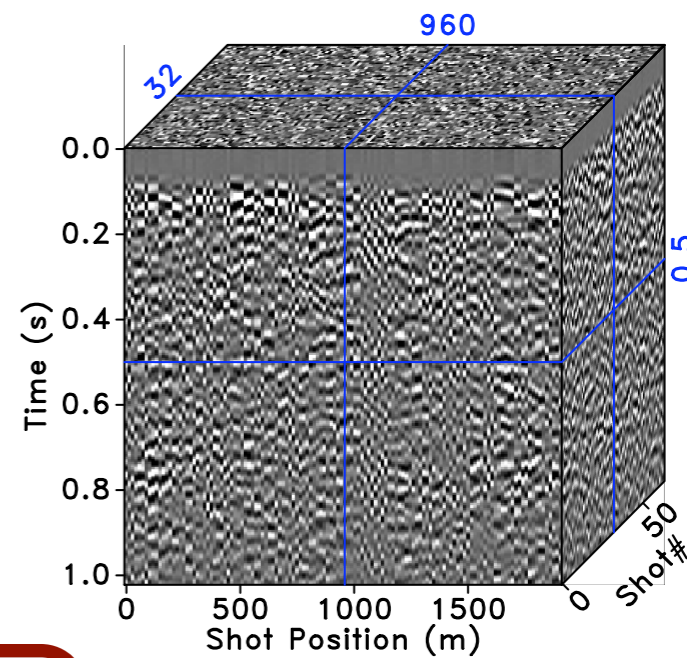
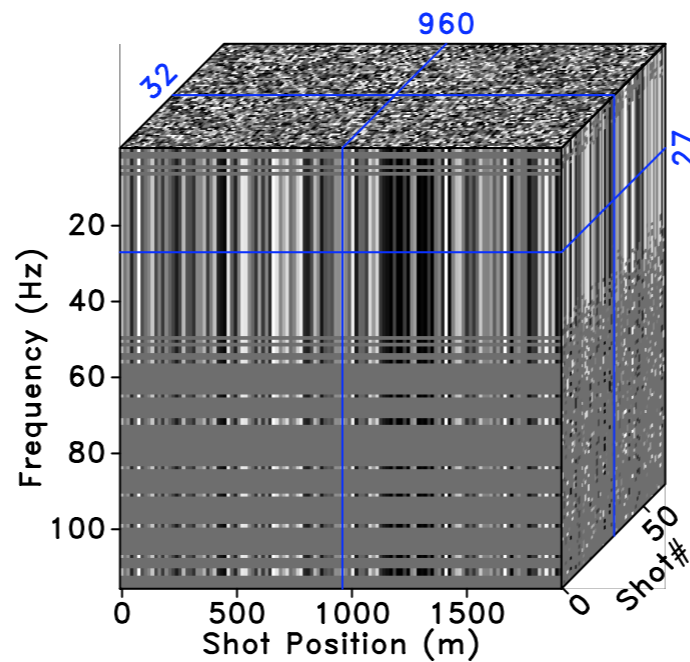
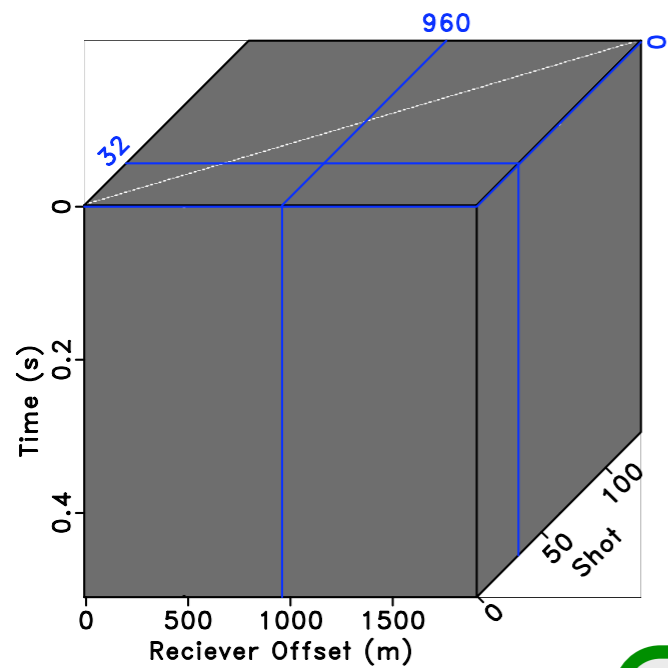
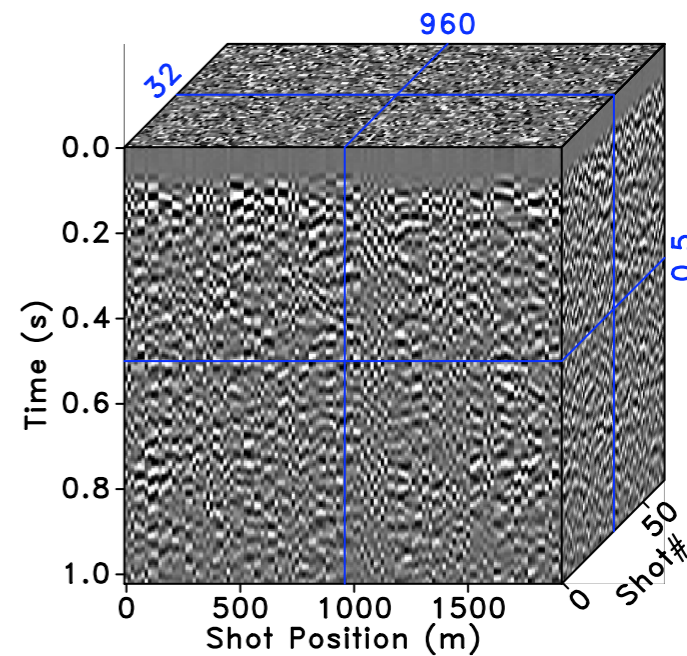
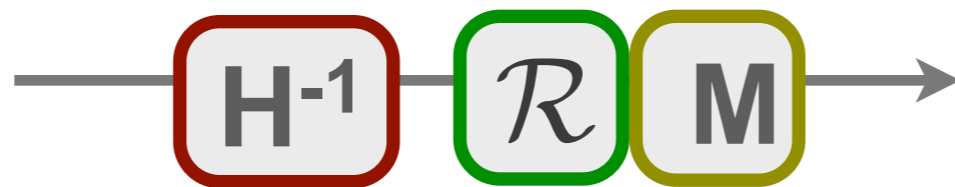
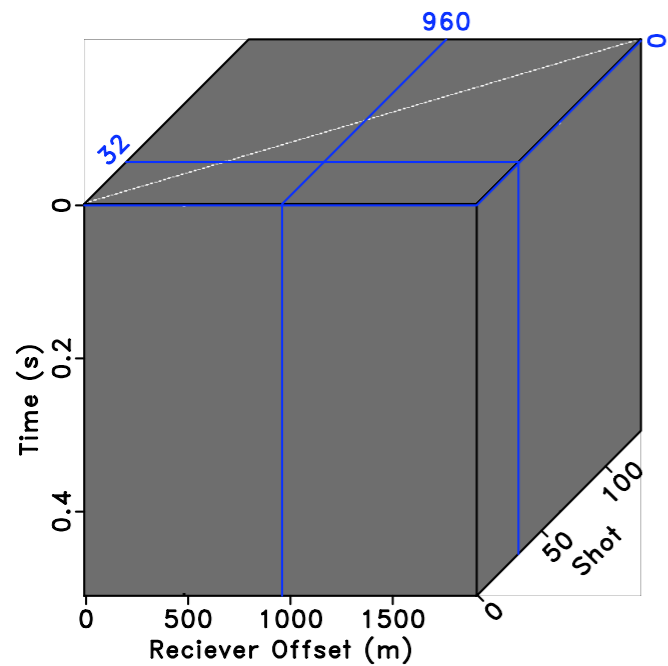
System-size reduction [FJH, Lin, and Erlangga, '09]

Subsample along **source** and **frequency** coordinates

Use **fast** transform-based sampling algorithms such as **scrambled Fourier** [Romberg, '08] or **Hadamard** ensembles [Gan et. al., '08]
sub sampler

$$\mathbf{RM} = \underbrace{\begin{bmatrix} \mathbf{R}_1^\Sigma \otimes \mathbf{I} \otimes \mathbf{R}_1^\Omega \\ \vdots \\ \mathbf{R}_{n_{s'}}^\Sigma \otimes \mathbf{I} \otimes \mathbf{R}_{n_{s'}}^\Omega \end{bmatrix}}_{\text{random phase encoder}} \underbrace{\left(\mathbf{F}_2^* \text{diag} \left(e^{i\hat{\theta}} \right) \otimes \mathbf{I} \right) \mathbf{F}_3}_{\theta_w = \text{Uniform}([0, 2\pi])}$$

- Different random restriction for each $n'_s \ll n_s$ simultaneous experiments
- Restriction reduces system size



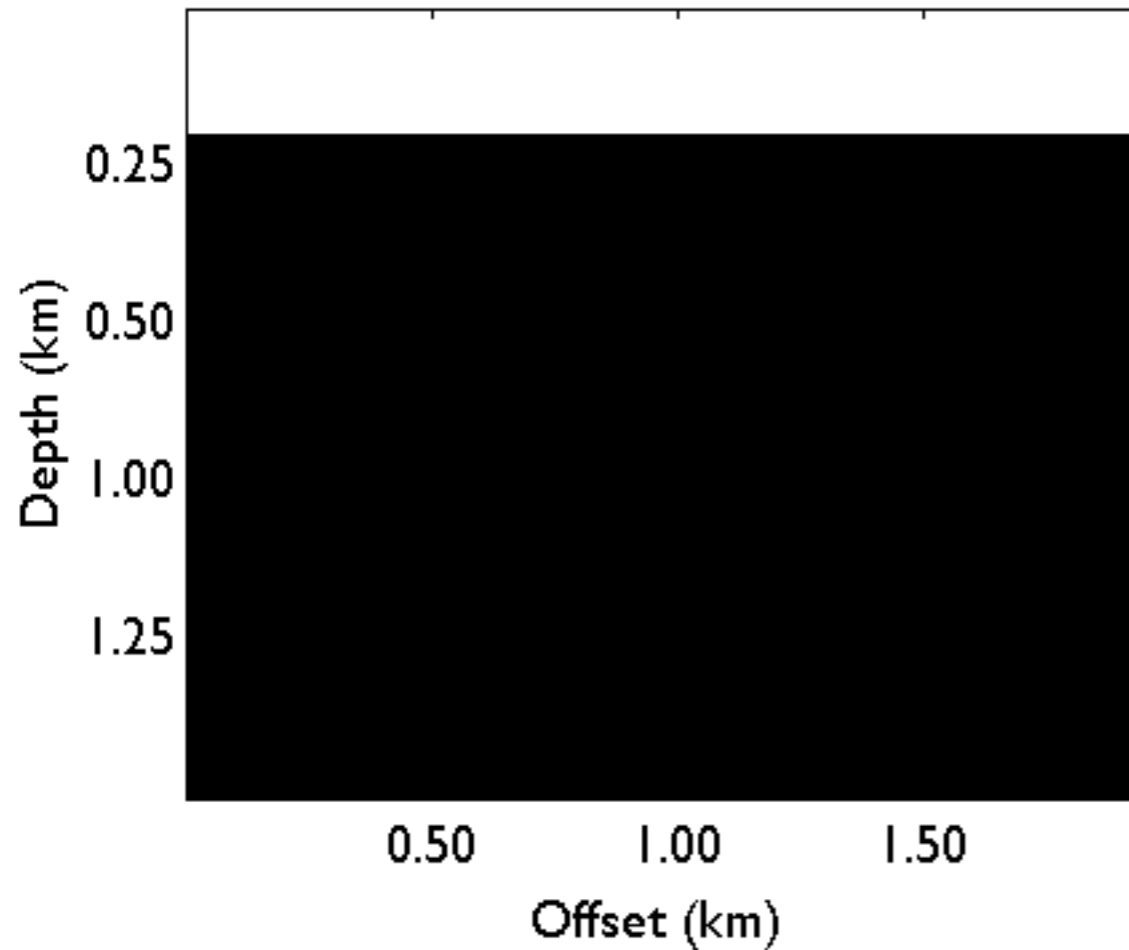
Sparsifying transform [Demanet '06]

- Use fast discrete 2-D Curvelet transform based on wrapping along shot and receiver coordinates
 - compresses highly geometrical features of monochromatic wavefields
 - incoherent with compressive-sampling matrix that acts along the source coordinate
- Use fast discrete wavelet transform along the time coordinate
 - compresses front-like features arriving along the time direction
 - reasonable incoherent with sampling of angular frequencies
- Combine both transforms through a **Kronecker** product

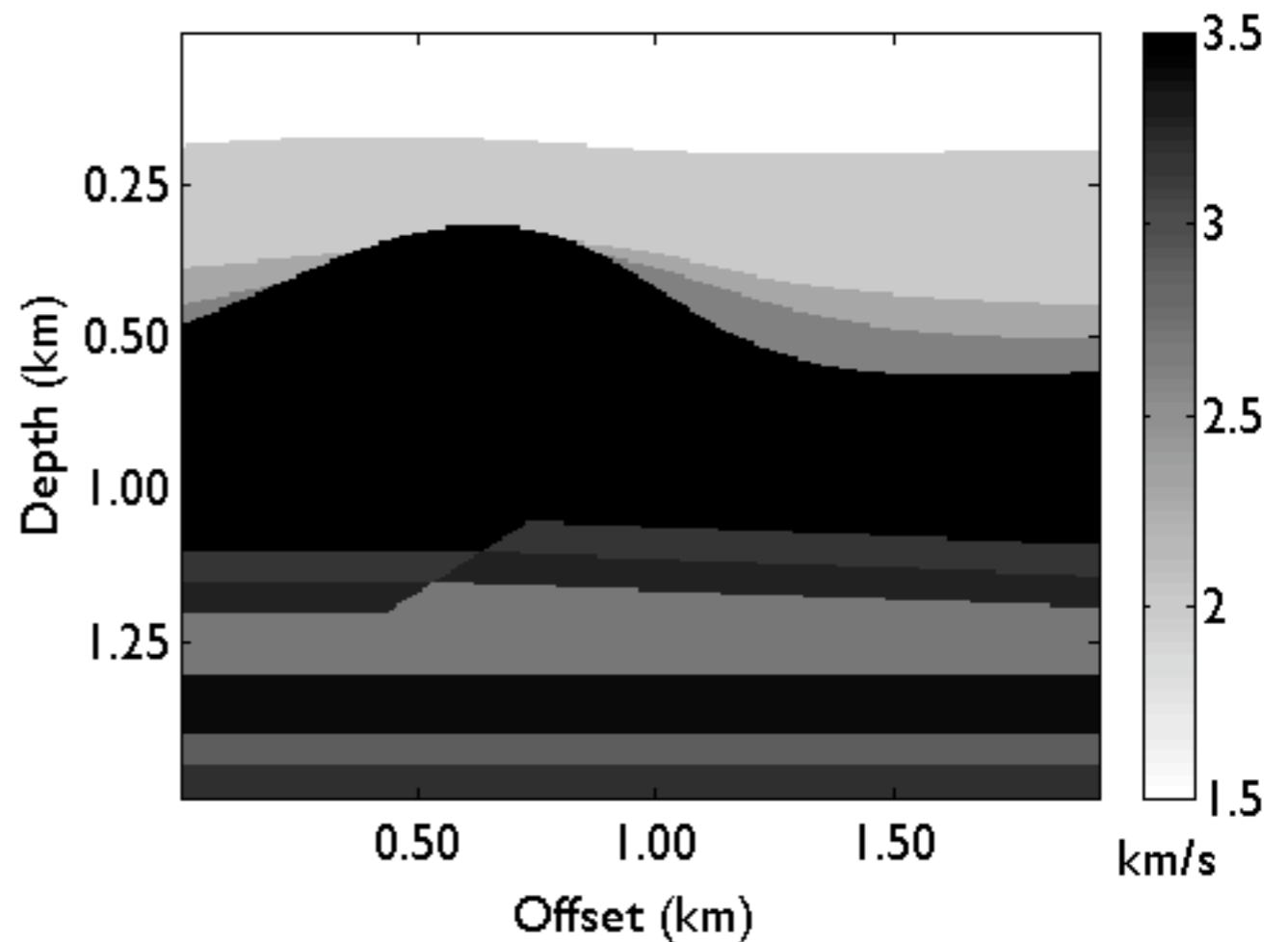
$$\mathbf{S} = \mathbf{C}_{2d} \otimes \mathbf{W}$$

Velocity models

simple model

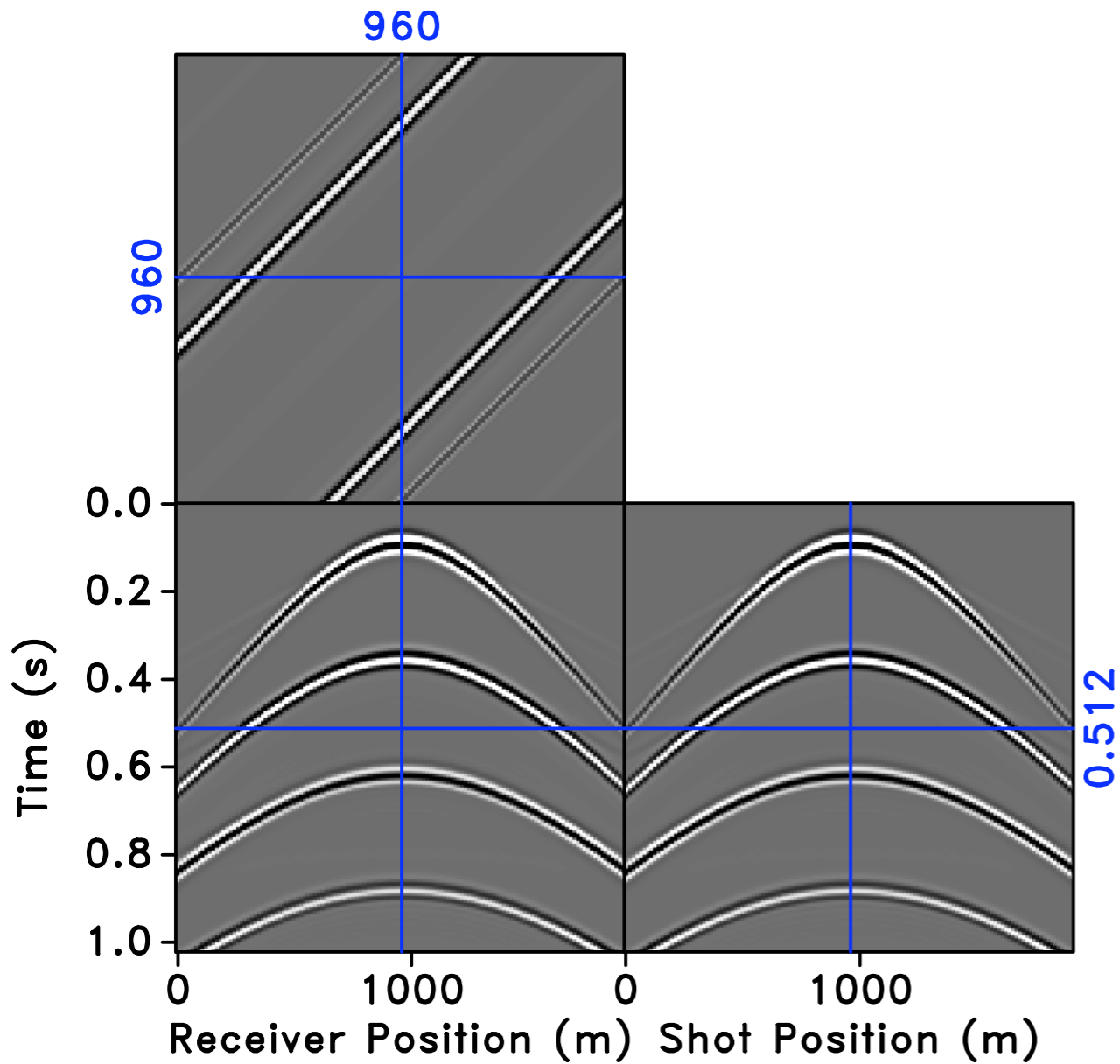


complex model

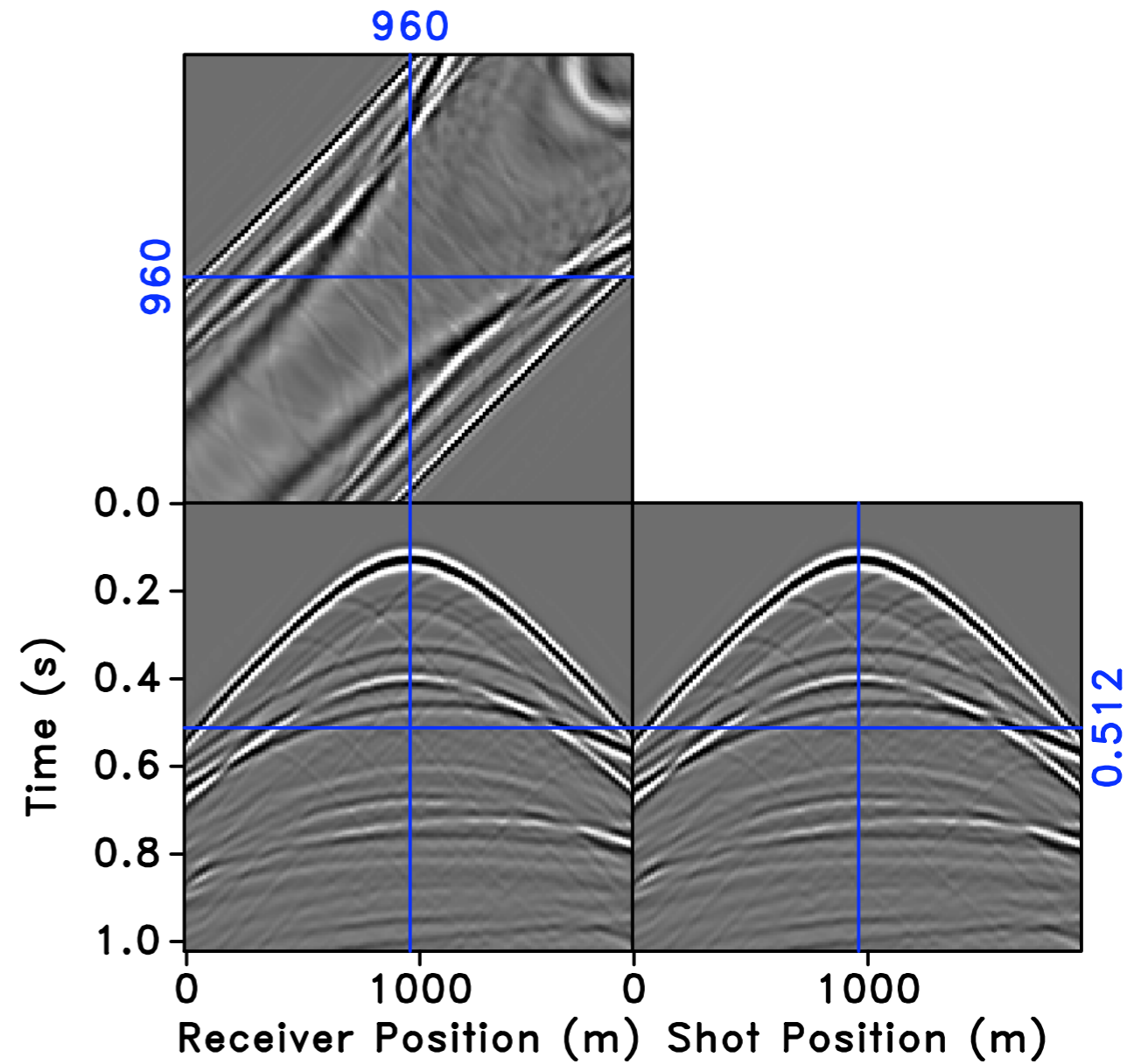


Green's functions

simple model

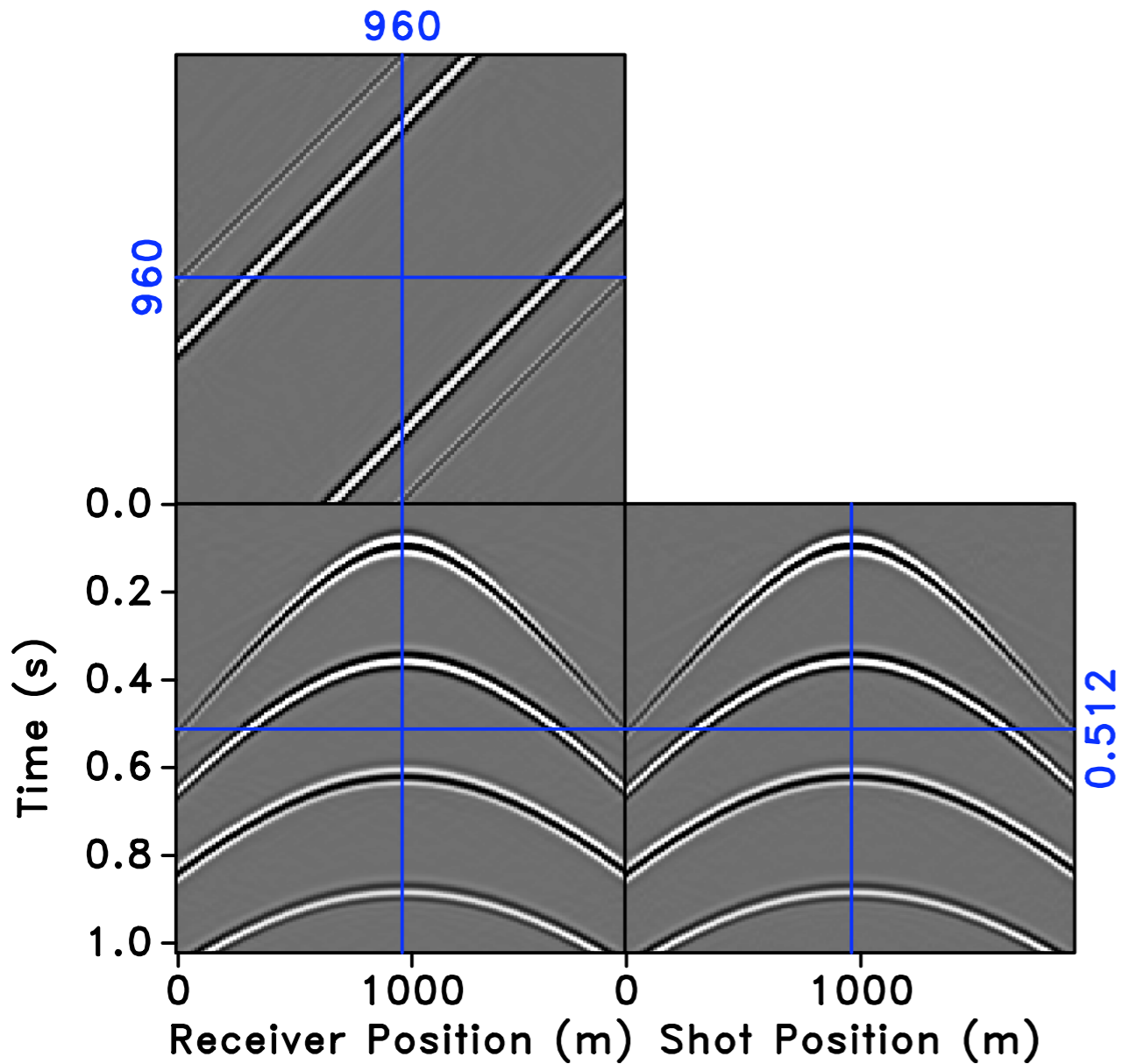


complex model



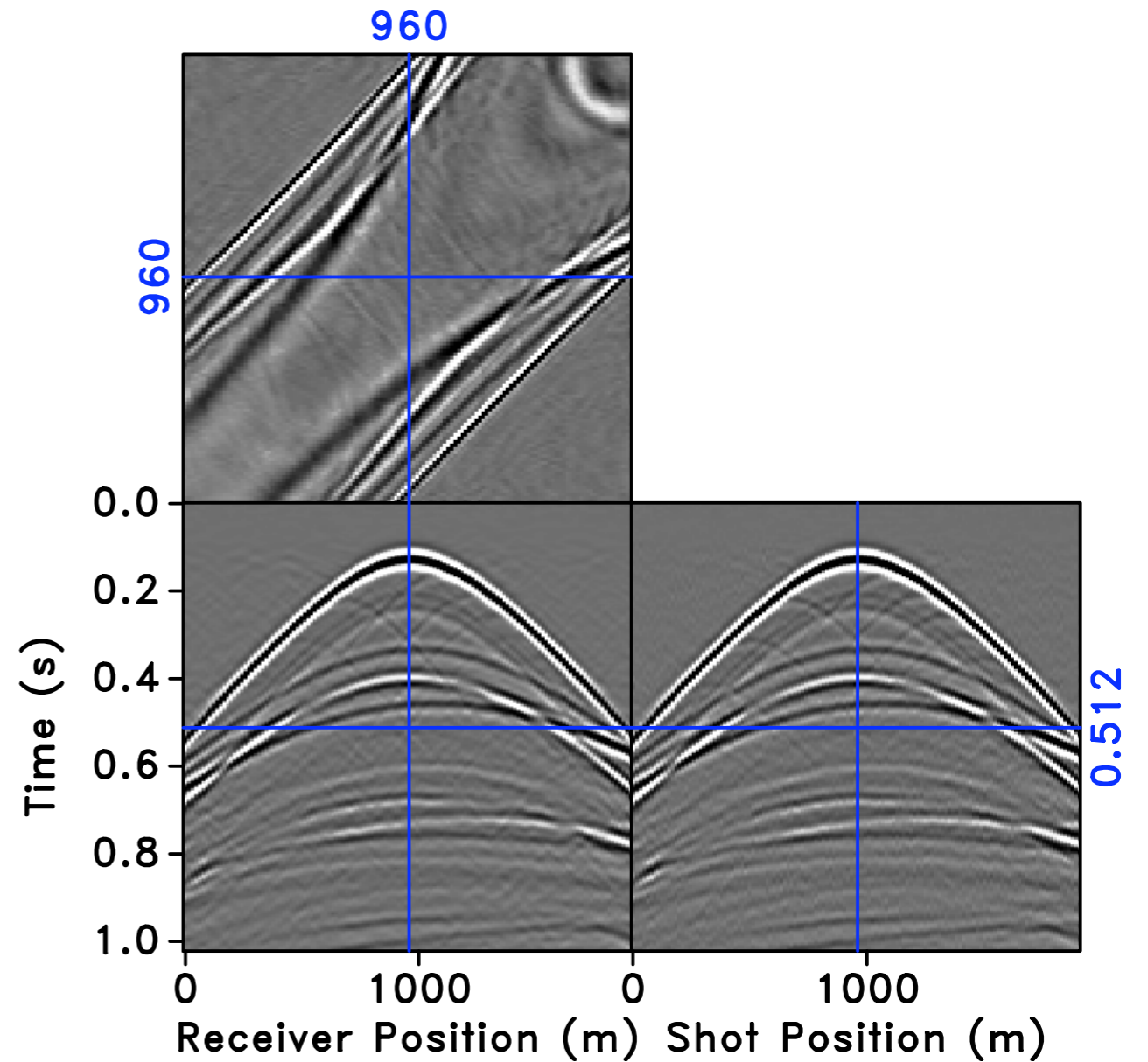
Recovered data

simple model



28.1dB

complex model



18.2dB

300 SPGL1 iteration

Observations & outlook

- **CS** provides a *linear sampling paradigm*
 - breaks *subsampling-related coherent interferences* by turning them into *harmless noise*
 - **degree** of *subsampling commensurate* with transform-domain **sparsity**
 - *subsampling* of solutions to PDEs
- Works as long as *recovery costs* are **smaller** than *simulation-cost reductions*
- **Robust** (via sparsity promotion) instance of exploiting **invariance** = “*sparsity conservation*” of *multiscale transform* under certain **solution operators**
- Bottom line: **numerical modeling costs** are **no** longer determined by the **size** of the **discretization** but by **transform-domain compressibility** of the **solution ...**

Challenges: there are many ...

✓ Helmholtz system is *indefinite* & *ill conditioned* => lack of convergence *indirect* Krylov solvers

±✓ Multiexperiment setup with *multiple right-hand-sides* is *computational prohibitive*

Inversion problem can be both *over-* and *underdetermined*

- data cannot be explained fully
- there are local minima
- many velocity models may explain data within some error

Proposed plans to tackle *multimodality* by *extensions & focusing* make the situation worse by additional *degrees of freedom*

System-size reduction

- Apply CS to reduce cost of *wavefield* simulation with Helmholtz
 - use *simultaneous sources* instead of *separated sources*
 - leverage transform-domain sparsity & randomized subsampling by **one-norm sparsity promotion**
 - reduce size Helmholtz system
 - sources (number of right-hand sides)
 - angular frequencies (number of blocks)
- Apply CS to reduce cost of computing *image volumes* by multi-dimensional correlations via *explicit* matrix-matrix multiplies
 - randomize and subsample wavefields in **model space**
 - leverage transform-domain sparsity and focusing in the *model space* by **joint sparsity promotion with mixed (1,2) norms**
 - reduce costs of storage and explicit matrix-matrix multiplies
 - sources (right-hand sides), receivers, depth
 - angular frequencies (blocks)

Tools

Compressive sensing based on Johnson-Lindenstrauss embeddings

- *Compressive sensing* [Donoho, '06, Candes, Romberg, Tao, '06]

$$\mathbf{b} = \mathbf{RM}\mathbf{x} \quad \text{[randomized subsampling]}$$

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{RM}\mathbf{x} - \mathbf{b}\|_2 \leq \sigma$$

$$\tilde{\mathbf{x}} \approx \mathbf{x}$$

Fast matrix computations based on Johnson-Lindenstrauss embeddings

- Improved Approximation Algorithms for Large Matrices via Random Projections by Tamás Sarlós, '08

$$\mathbf{AB} \approx \mathbf{A} (\mathbf{RM})^* (\mathbf{RM}) \mathbf{B}$$

Joint sparsity-promotion with mixed (1,2) norm minimization

- *Joint-sparse recovery from multiple measurements* by E. van den Berg and M. Friedlander, '09

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

Differential semblance

- Invoke *physical* principle of **focusing** [Claerbout & many others] \Leftrightarrow *mathematical* principle of **extensions** [Symes]
- Motivated by Symes' differential semblance principle [Symes '09]:
 “Amongst all possible quadratic forms in the data, parameterized by velocity, of the form

$$\min_{\mathbf{m}} \left\| \left(\overbrace{P_h \delta \mathbf{I}(\cdot, h; \mathbf{m}, \delta \mathbf{d})}^{\text{image volume}} \right) \right\|_2 \quad \text{with } \overbrace{P_h \cdot}^{\text{annihilator}} = \mathbf{h} \cdot,$$

\uparrow
 redundant coordinate

only differential semblance is smooth jointly as function of smooth perturbations in velocity and finite energy perturbations in data [Stolk & Symes, '03]”

- Forms the basis of **nonlinear** migration velocity analysis on **linearized** data [Symes, '09].

Image volume

Compute multi-D **cross-correlations** on **multiexperiment** solutions of the forward- and reverse-time Helmholtz systems--i.e,

$$\delta\mathbf{I}(m, h, t) = \left(\bar{\mathbf{U}} * \mathbf{V}^T \right)$$

with

$$\mathbf{U}_f = [\mathbf{u}_1 \cdots \mathbf{u}_{n_f}] \text{ and } \mathbf{V}_f = [\mathbf{v}_1 \cdots \mathbf{v}_{n_f}]$$

and

$$\left(\bar{\mathbf{U}} * \mathbf{V}^T \right) := \mathbf{T}_{(x_s, x_r, \omega) \mapsto (m, h, t)} \begin{bmatrix} \bar{\mathbf{U}}_1 & & \\ & \ddots & \\ & & \bar{\mathbf{U}}_{n_f} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \vdots \\ \mathbf{V}_{n_f}^T \end{bmatrix}$$

where

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

High dimensional and highly **redundant** ...

Imaging condition

Claerbout's imaging principle:

$$\begin{aligned}\delta \mathbf{m} &= \delta \mathbf{I}(\cdot, h = 0, t = 0) \\ &= \mathbf{K}^* \delta \mathbf{d}\end{aligned}$$

- implicit in adjoint state method
- Image volume
 - very large because of additional degree of freedom
 - expensive to store

System-size reduction by CS

For each angular frequency, subsample with CS matrix

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

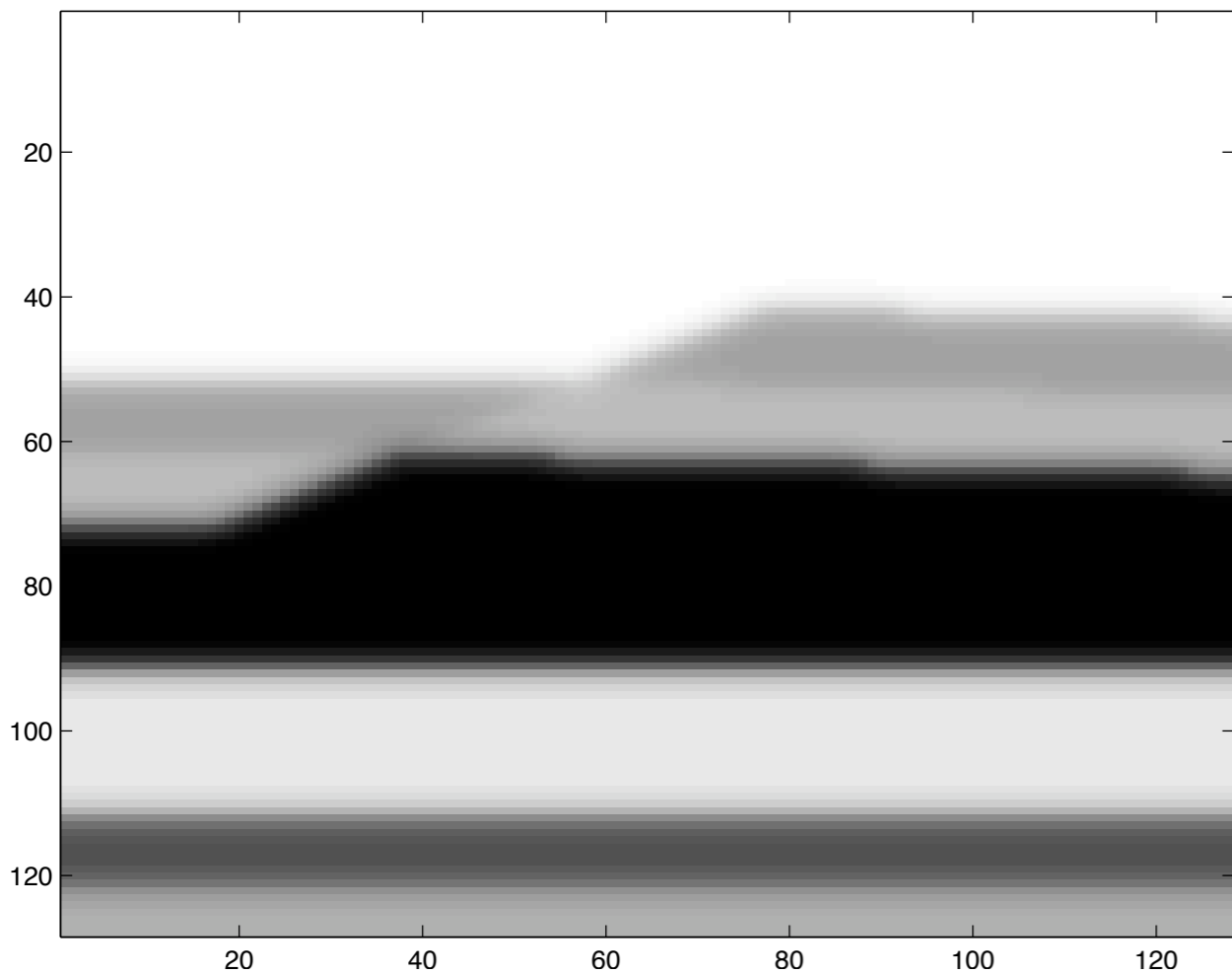
with

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

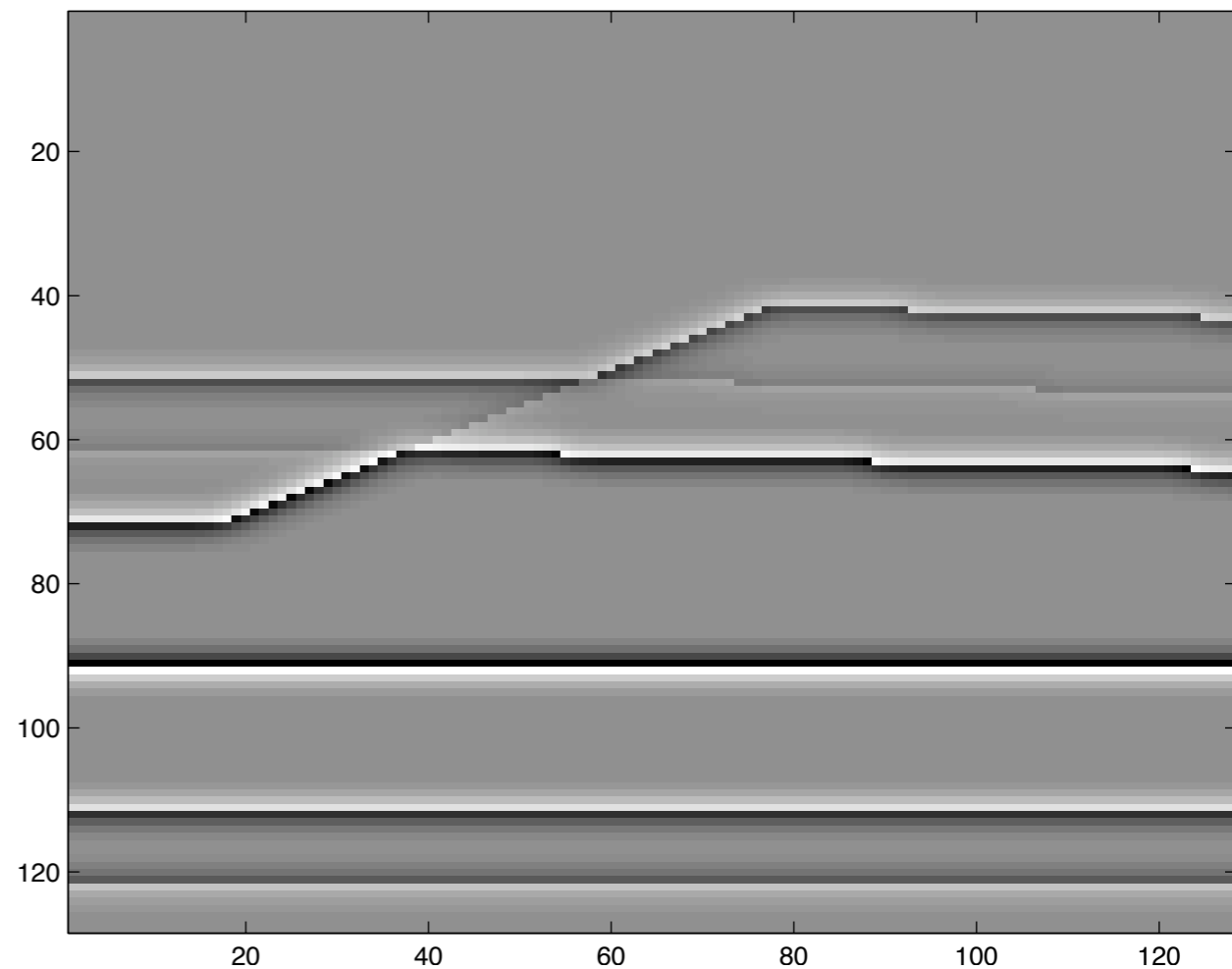
Model-space CS subsampling along subsurface source, receiver, and depth coordinates yielding an *approximate extended* image

$$\delta\mathbf{I}(m, h, t) \approx \left(\bar{\mathbf{U}}(\mathbf{RM})^* * \mathbf{RMV}^T \right)$$

Example



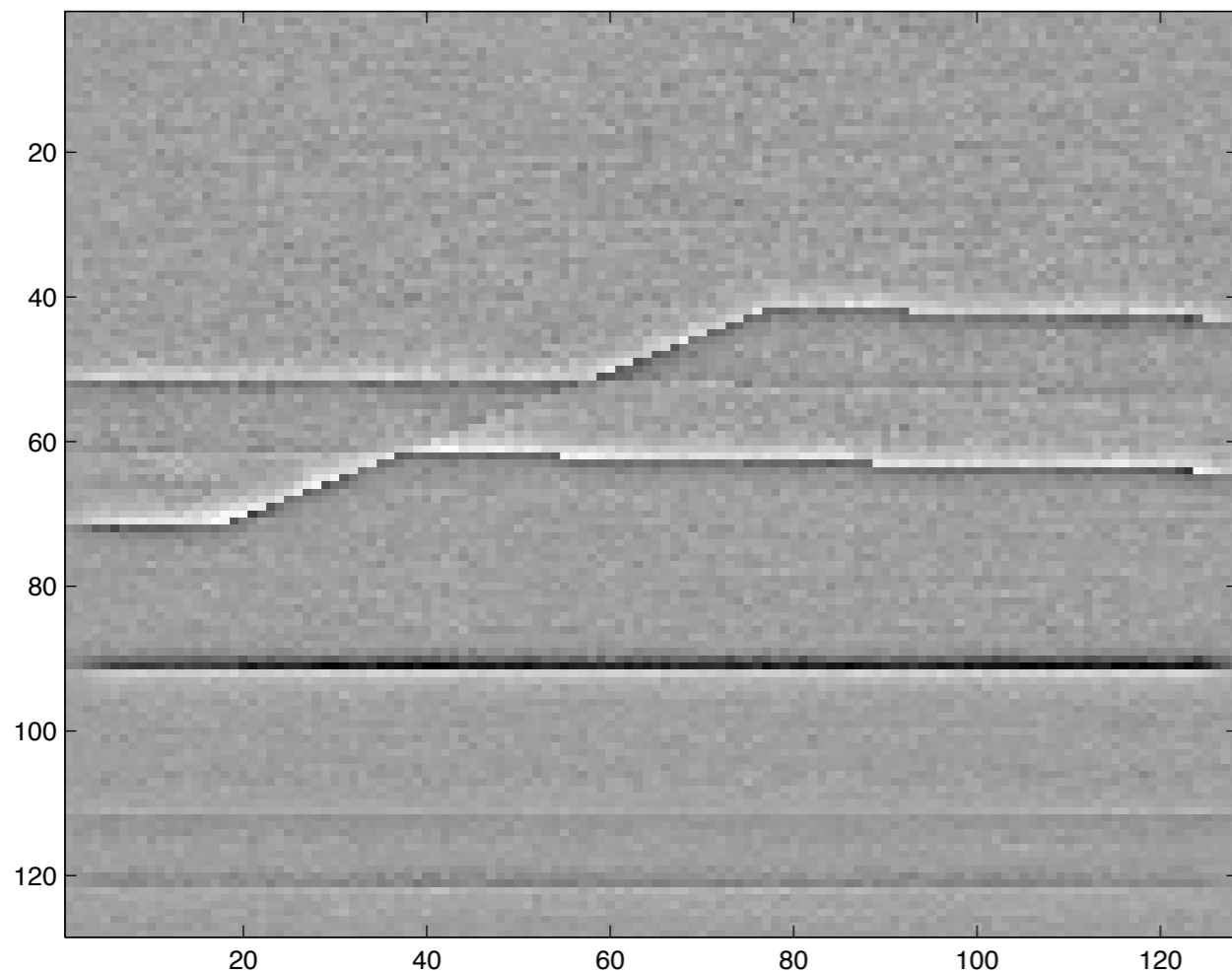
background velocity model



perturbation

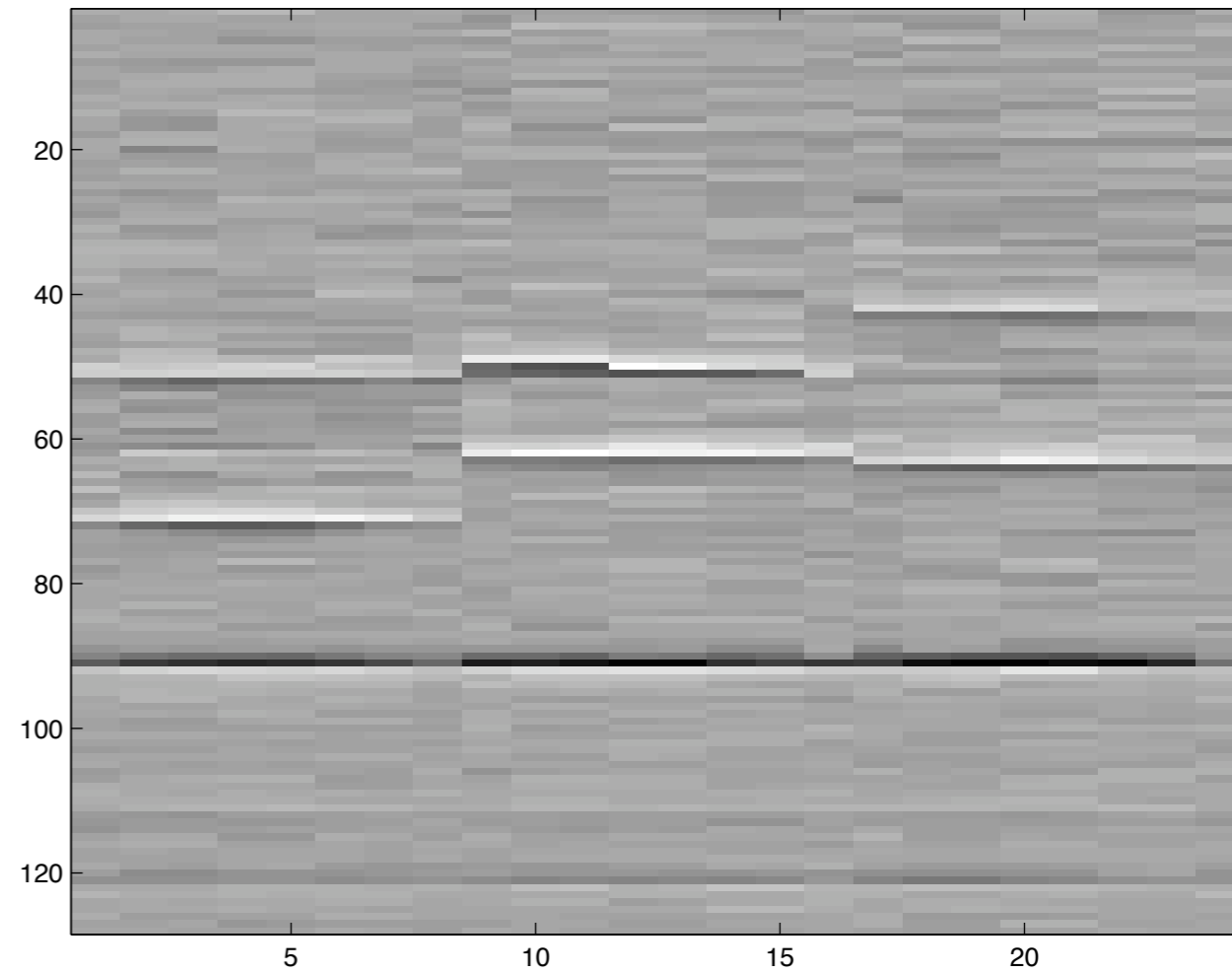
Example: matched filter

migrated CS image



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

migrated CS cigs



$$\delta\mathbf{I}([m_1, m_2, m_3], h, t = 0)$$

Recovery from 64-fold subsampling ...

- Noisy
- Not focused

Tools

Compressive sensing based on Johnson-Lindenstrauss embeddings

- *Compressive sensing* [Donoho, '06, Candes, Romberg, Tao, '06]

$$\mathbf{b} = \mathbf{RM}\mathbf{x} \quad \text{[randomized subsampling]}$$

$$\tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{RM}\mathbf{x} - \mathbf{b}\|_2 \leq \sigma$$

$$\tilde{\mathbf{x}} \approx \mathbf{x}$$

Fast matrix computations based on Johnson-Lindenstrauss embeddings

- Improved Approximation Algorithms for Large Matrices via Random Projections by Tamás Sarlós, '08

$$\mathbf{AB} \approx \mathbf{A} (\mathbf{RM})^* (\mathbf{RM}) \mathbf{B}$$

Joint sparsity-promotion with mixed (1,2) norm minimization

- *Joint-sparse recovery from multiple measurements* by E. van den Berg and M. Friedlander, '09

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

Extended Born & focusing

Define *extended linearized* forward model [Symes, '09]:

$$\bar{\mathbf{K}}[\mathbf{m}, \mathbf{Q}]\delta\mathbf{I} \approx \delta\mathbf{D}$$

- multiexperiment form *amenable* for *joint sparsity promotion*
- introduce penalty term that penalizes *defocusing*

Form augmented system with **focusing**:

$$\bar{\mathbf{K}}\delta\mathbf{I} \approx \delta\mathbf{D} \quad \text{data fit}$$

$$\lambda^2 \mathbf{P}_h \delta\mathbf{I} \approx \mathbf{0} \quad \text{focusing}$$

with $\mathbf{P}_h \cdot = \mathbf{h} \cdot$ *annihilator* that increasingly *penalizes* non-zero offsets.

Solution involves multi-D “deconvolution” (adjoint of cross correlation):

$$(\mathbf{U}^* \star \delta\mathbf{I}) \approx \mathbf{V}^T$$

Compressed linearized inversion

Compressively sample augmented system that includes sparsity synthesis operator--i.e,

$$\begin{aligned} \mathbf{RM} (\mathbf{U}^* \star \mathbf{S}^* \mathbf{X}) &\approx \mathbf{RMV}^T & \mathbf{AX} &\approx \mathbf{B} \\ P_h \mathbf{X} &\approx \mathbf{0} \end{aligned}$$

with the sparsifying transform \mathbf{S} for each offset h given by the curvelet or wavelet transform

Recover focused solution by mixed (1,2)-norm minimization.

Promote sparsity amongst images though one-norm on columns

Penalize energy amongst rows => focusing

Joint-sparsity promotion [van den berg & Friedlander, '09]

Recover focused solution by mixed (1,2)-norm minimization:

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

with

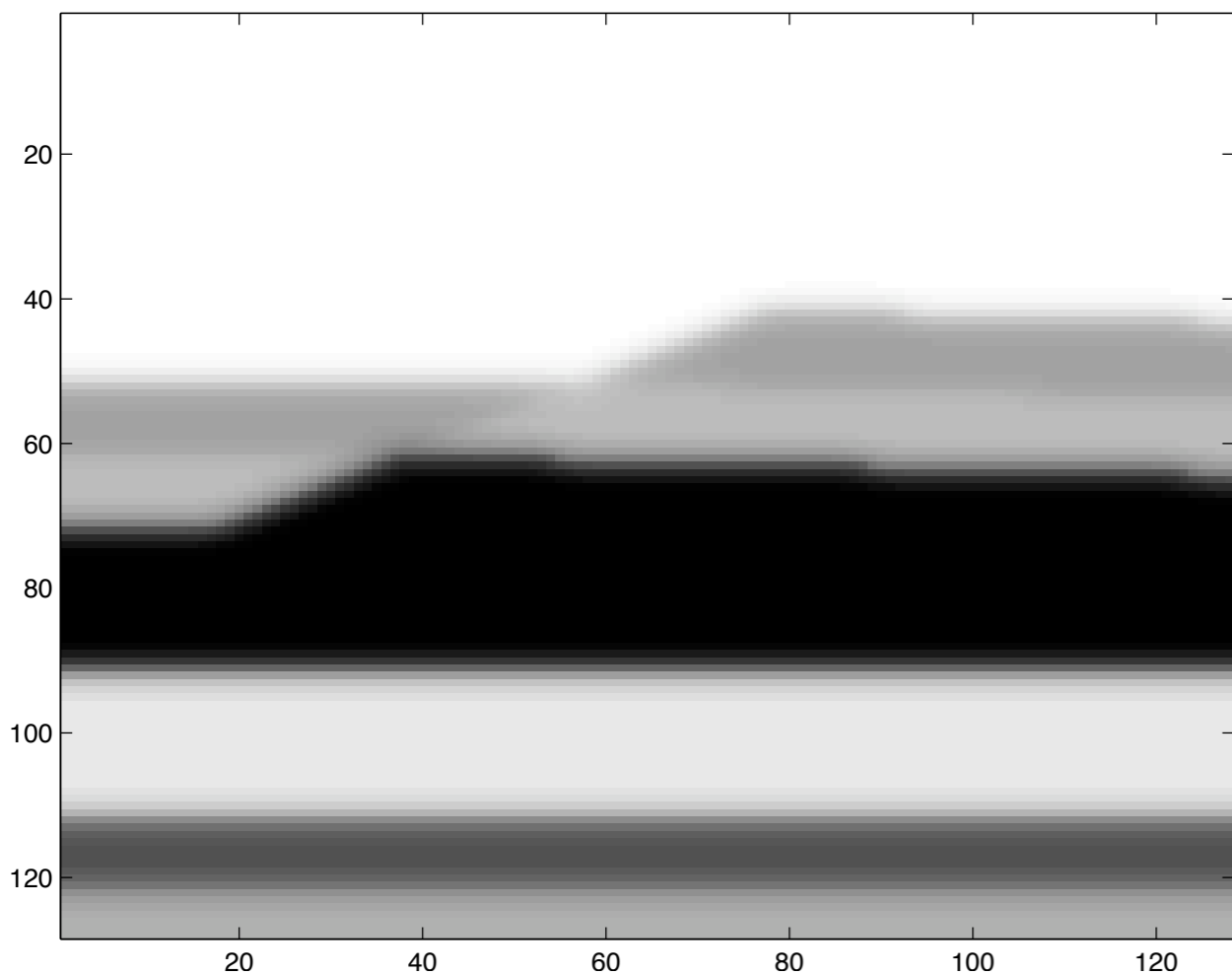
$$\|\mathbf{X}\|_{1,2} := \sum_{i \in \text{rows}(\mathbf{X})} \|\text{row}_i(\mathbf{X})^*\|_2$$

and

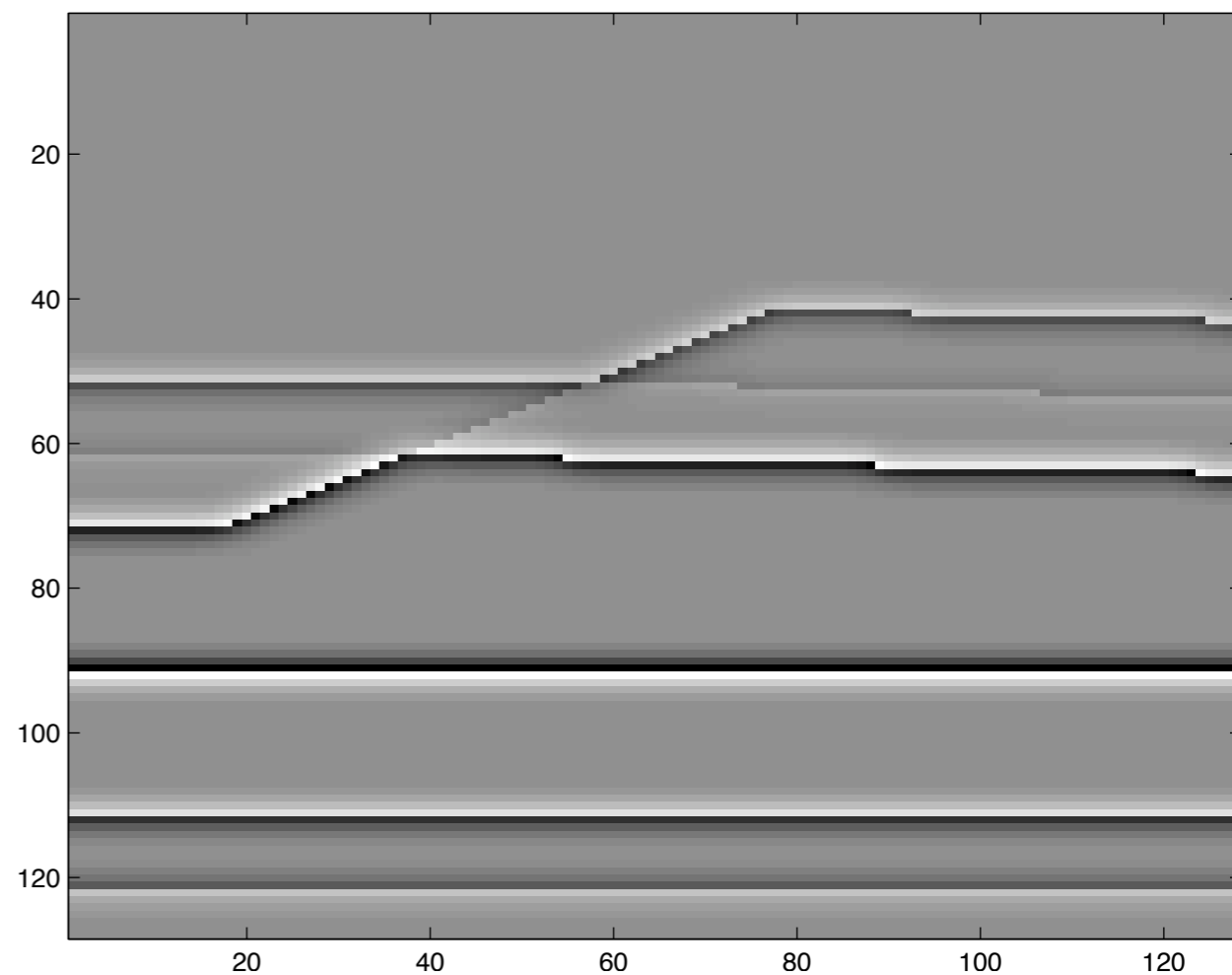
$$\|\mathbf{X}\|_{2,2} := \left(\sum_{i \in \text{rows}(\mathbf{X})} \|\text{row}_i(\mathbf{X})^*\|_2^2 \right)^{\frac{1}{2}}.$$

Solved with SPGL1.

Example



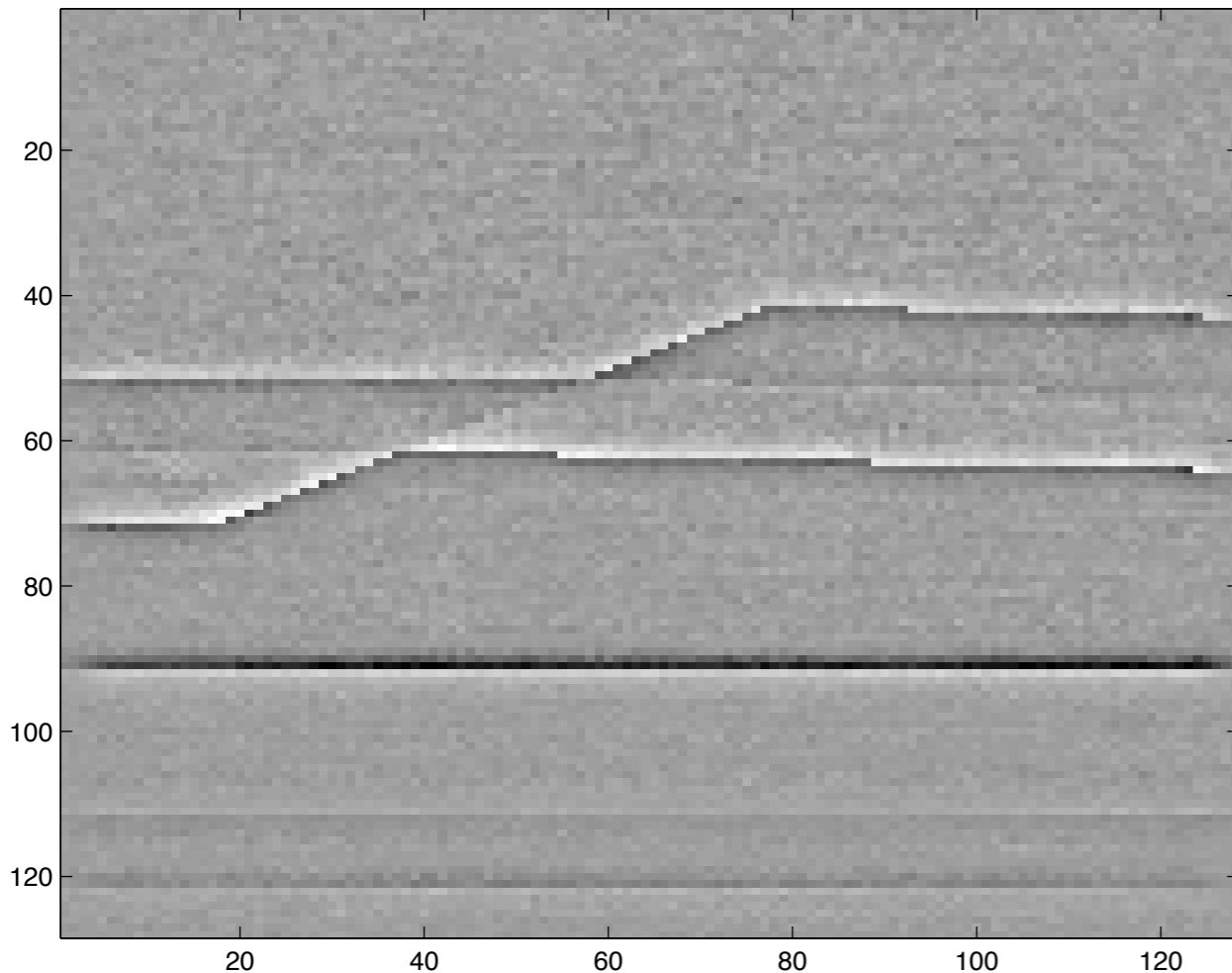
background velocity model



perturbation

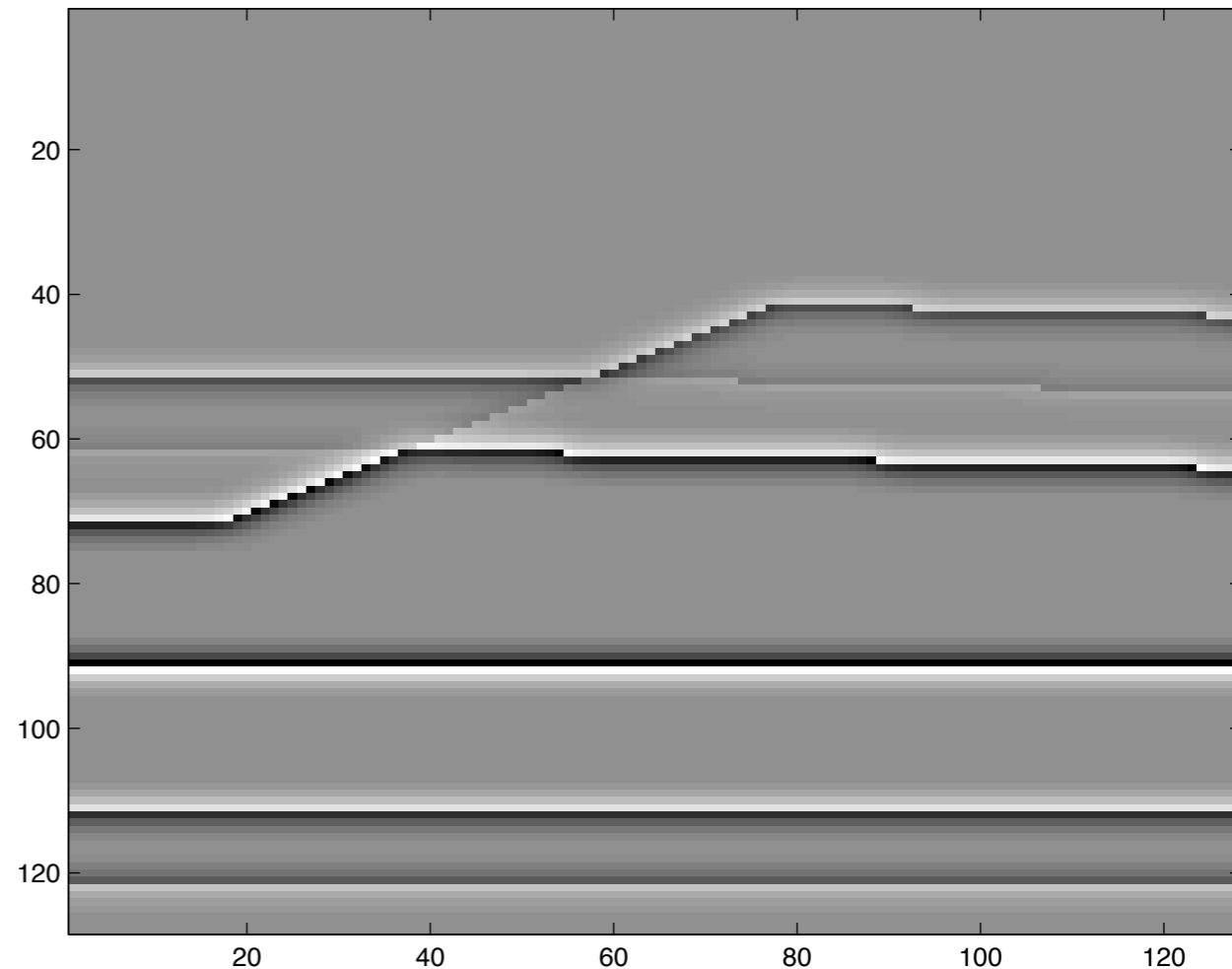
Example

migrated CS image



matched filter

inverted CS image

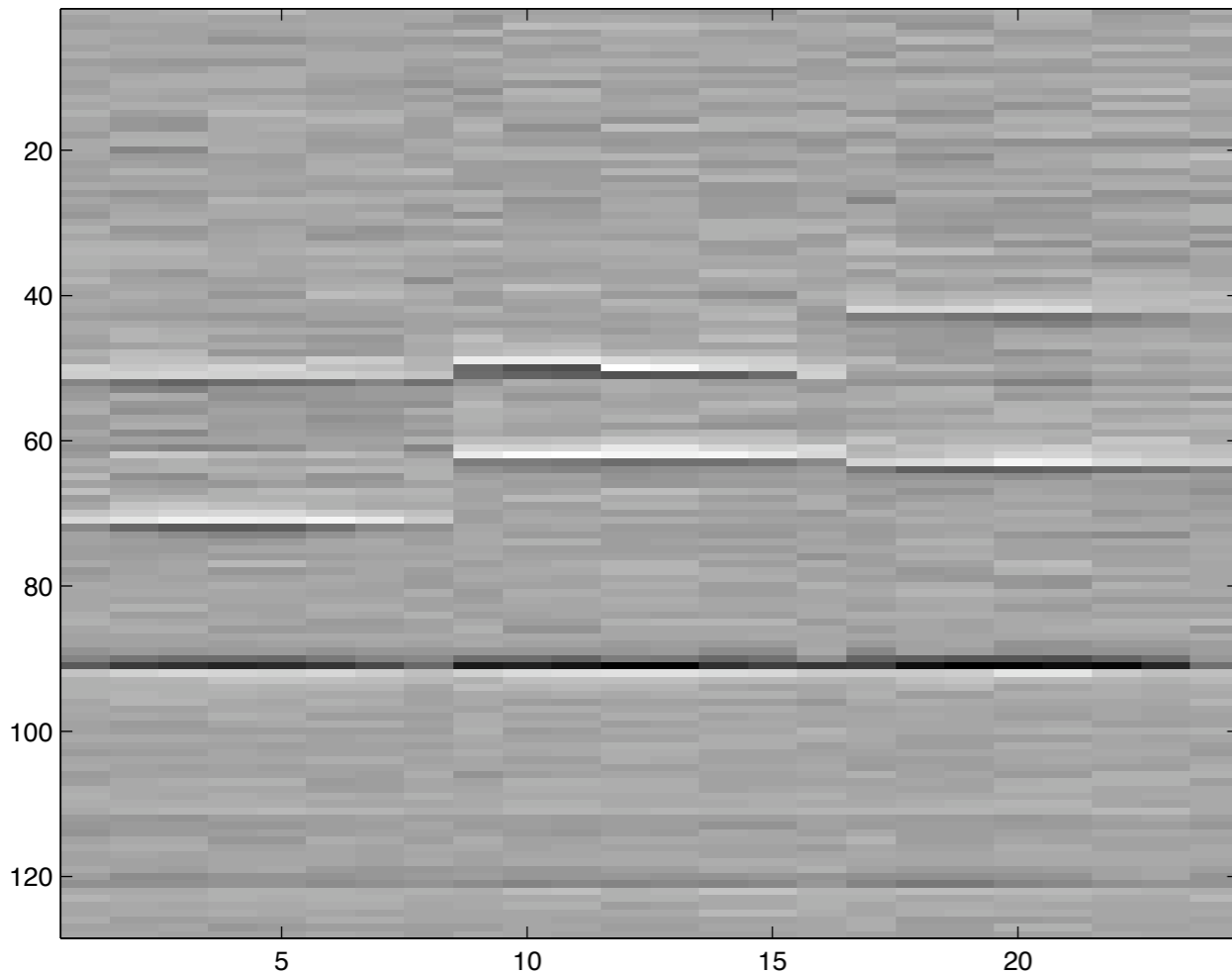


sparsity promotion

Recovery from 64-fold subsampling ...

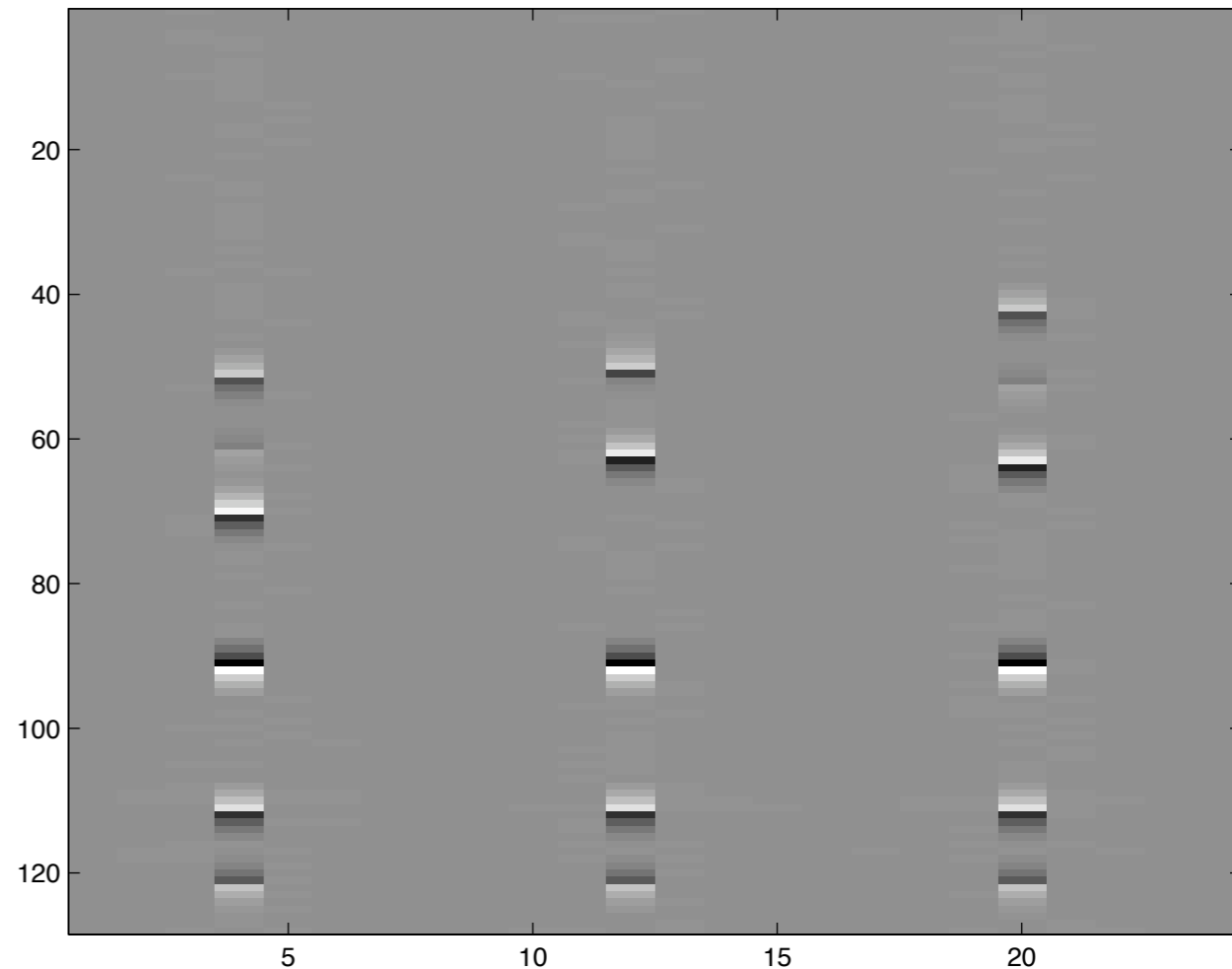
Example

migrated CS cigs



matched filter

inverted CS cigs



sparsity promotion

Common-image gathers are focussed.

Observations & outlook

- CS allows for a ***compression*** of data volumes without ***significant loss of information*** yielding a ***reduction*** in ***computational costs***
- CS has ***direct*** implications for seismic acquisition--from ***sequential*** to ***simultaneous*** acquisition
- ***Joint*** sparsity promotion allows for ***focusing***
- **Speculation:** Proposed approach may be suitable to handle Symes's proposal to add a degree of freedom yielding a ***nonlocal forward*** model in tandem with an inverse problem that penalizes ***nonlocality*** through ***focusing ...***

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!