THE UNIVERSITY OF BRITISH COLUMBIA | VANCOUVER



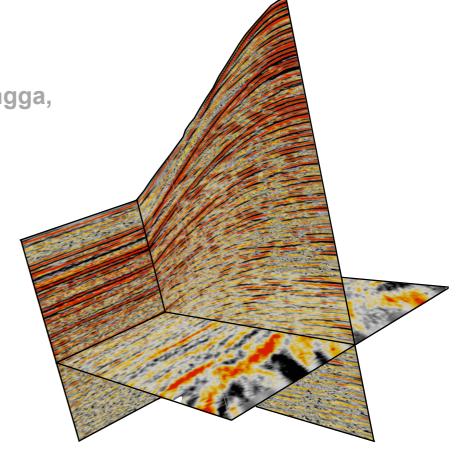
Compressive sensing: a paradigm shift for the imaging sciences?

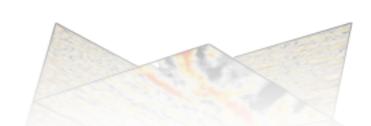
Felix J. Herrmann*

fherrmann@eos.ubc.ca

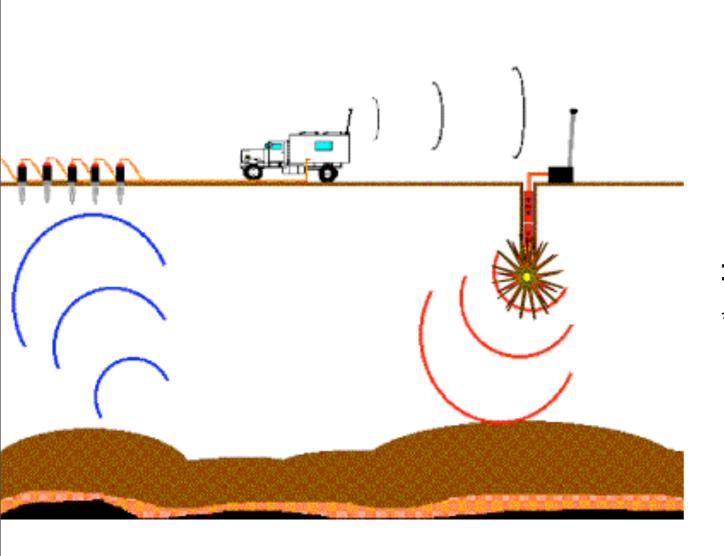
Joint work with Gilles Hennenfent, Yogi Erlangga, and Tim Lin

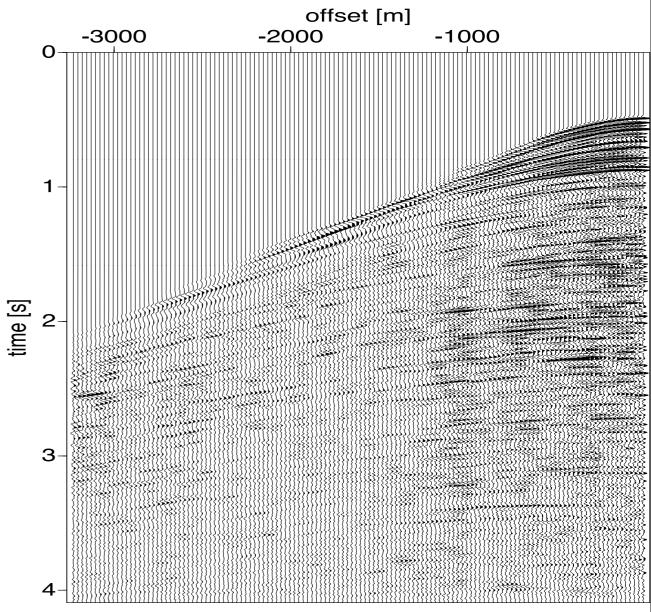
*Seismic Laboratory for Imaging & Modeling
Department of Earth & Ocean Sciences
The University of British Columbia



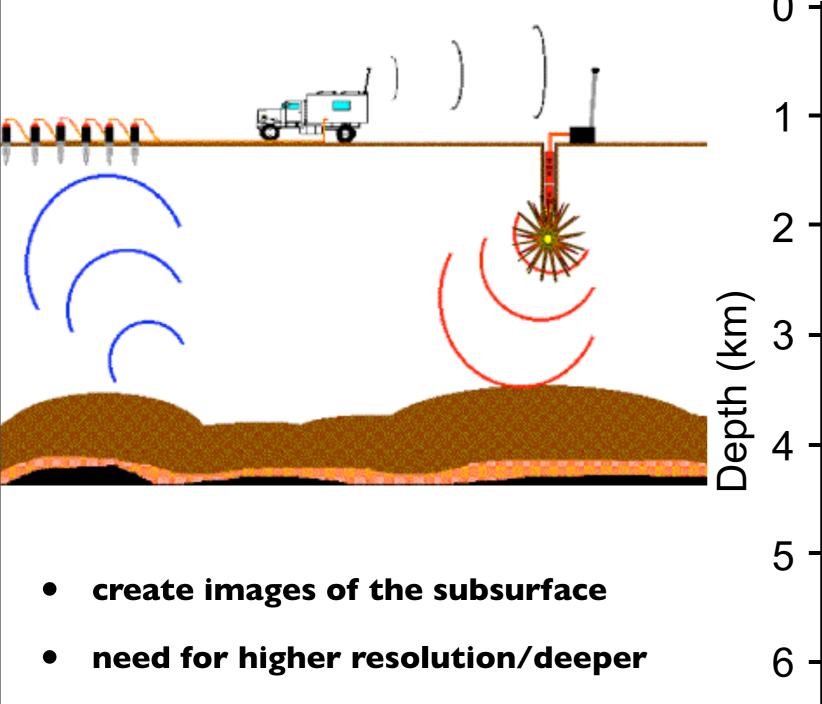


Seismic data acquisition

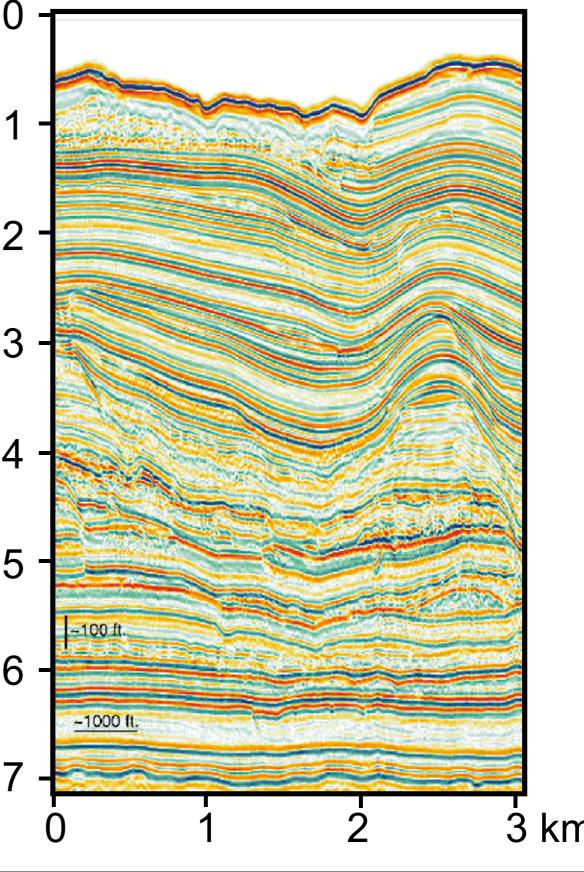




Exploration seismology



• clutter, data incompleteness, and large data size are problems



Motivation

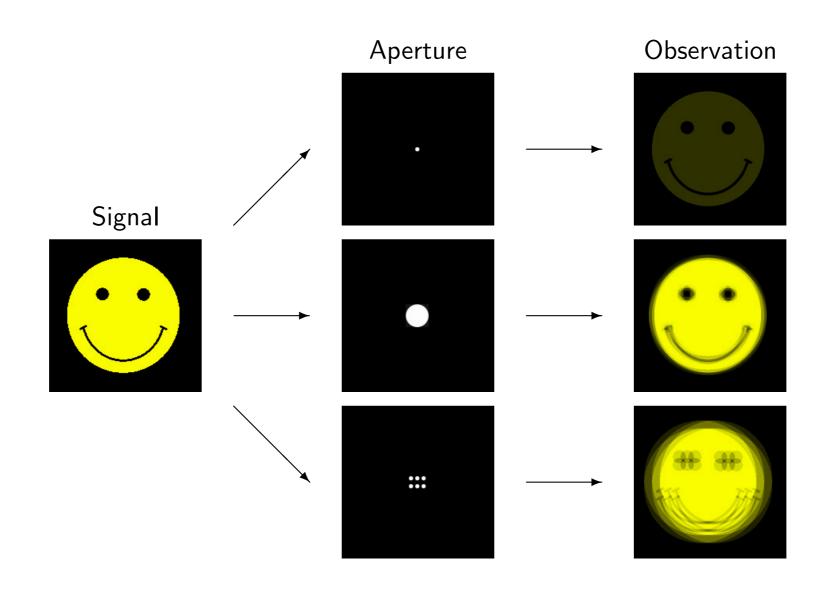
• Current state of affairs:

- (Seismic) data acquisition, processing, modeling, and imaging are firmly rooted in the paradigm of regular Nyquist sampling
- Acquisition is based on regularly-sampled & source-separated data volumes
- Major impediment related to the size of seismic data volumes
 - costs of acquisition & modeling dependent on size of the discretization
 - difficult to form explicit (imaging) operators
 - difficult to solve implicit (imaging) operators

Recent theoretical developments

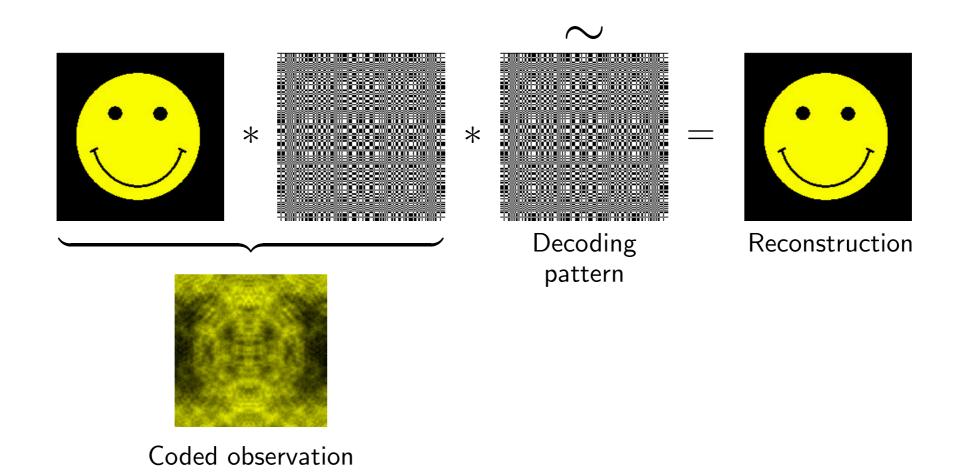
- New nonlinear sampling theory--known as Compressive Sensing (CS)--that supersedes the overly pessimistic Nyquist sampling criterion
- New CS-based acquisition techniques are being developed (MRI, AD, Radar, etc)
- New continuous and simultaneous acquisition strategies <=> instances of CS
- Success of compressive sensing hinges on
 - transform-domain sparsity (e.g. sparsity attained by multiscale/directional curvelets)
 - linear subsampling schemes that reduce interferrences
 - nonlinear recovery by sparsity promotion

An example [adapted from Roummel F. Marcia and Rebecca M. Willett, 2008]



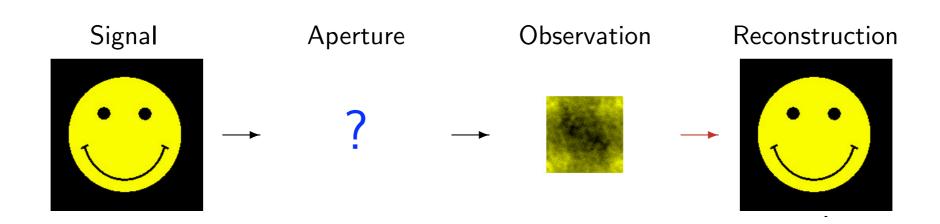
- Large pin holes (samples) blur
- Regularly-sampled multiple pinholes lead to coherent interferences

An example [adapted from Roummel F. Marcia and Rebecca M. Willett, 2008]



- Coded aperture gives brighter image
- Requires linear decoding
- No reduction in acquired data volumes

An example [adapted from Roummel F. Marcia and Rebecca M. Willett, 2008]



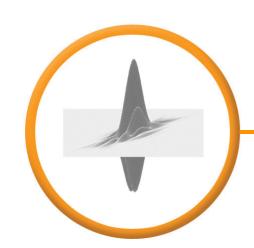
- Reduce the dimensionality by subsampling
 - create observations that contain the same information (encoded)
 - facilitate recovery by nonlinear transform-domain sparsity promotion
- Design physically realizable encoding schemes that break interferences
- Design nonlinear recovery techniques that exploit transformdomain sparsity/compressibility of certain structured signals
 - multiscale & directional wavelets
 - multiscale, multidirectional, and anisotropic curvelets

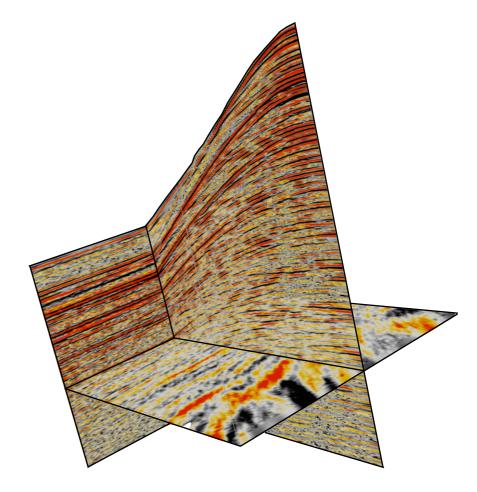
Today's agenda

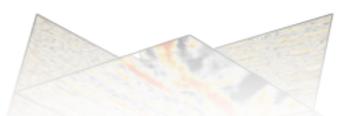
- Brief introduction to compressive sensing
 - sparsifying transform
 - favorable (random) acquisition (acquisition-grid and source-function design)
 - nonlinear recovery by sparsity promotion
- CS applied to wavefield reconstruction from incomplete acquisitions
 - use the "Dirac" basis as the sampling domain
 - reduction of the number of sources & receivers
- 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



Compressive sensing

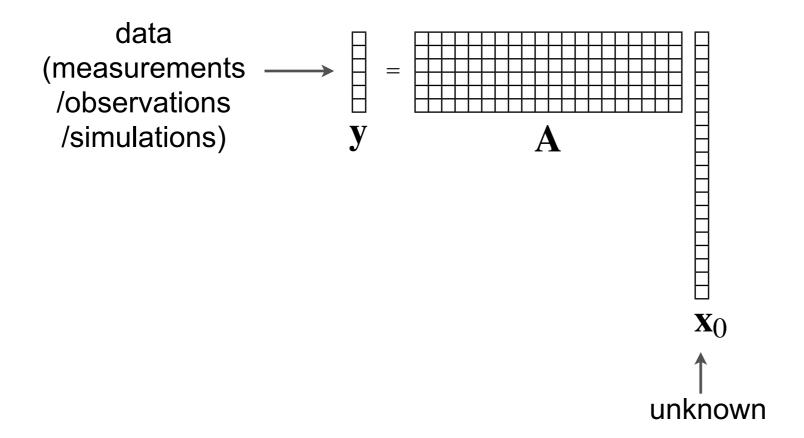






Problem statement

Consider the following (severely) underdetermined system of linear equations



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

Perfect recovery

$$\mathbf{y}$$

- conditions
 - A obeys the uniform uncertainty principle
 - x₀ is sufficiently sparse

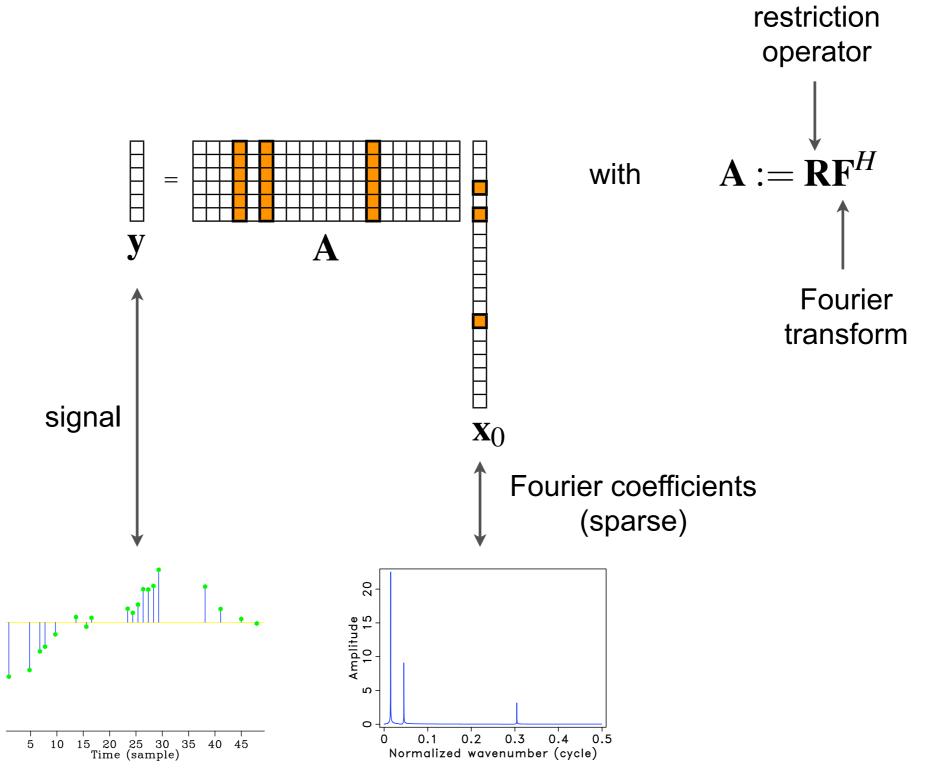
 \mathbf{x}_0

procedure

$$\underbrace{\min_{\mathbf{x}} \|\mathbf{x}\|_{1}}_{\text{sparsity}} \quad \text{s.t.} \quad \underbrace{\mathbf{A}\mathbf{x} = \mathbf{y}}_{\text{perfect reconstruction}}$$

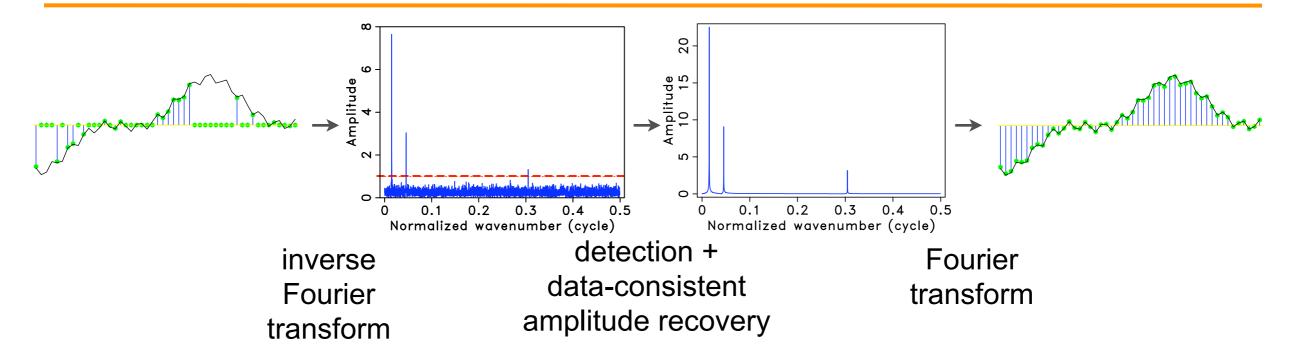
- performance
 - S-sparse vectors recovered from roughly on the order of S measurements (to within constant and log factors)

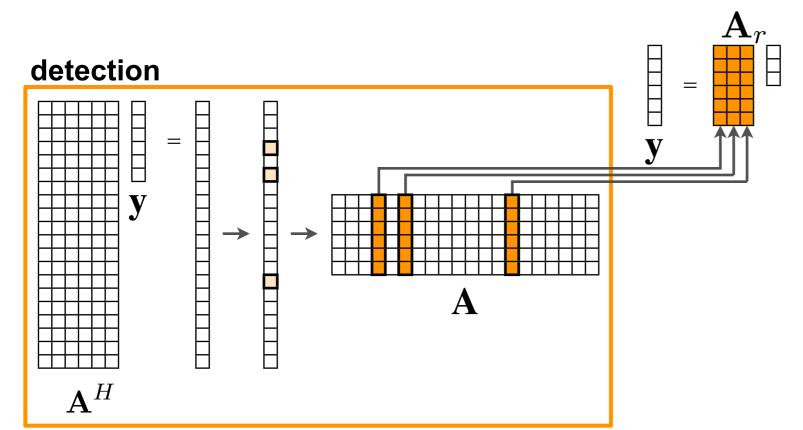
Simple example



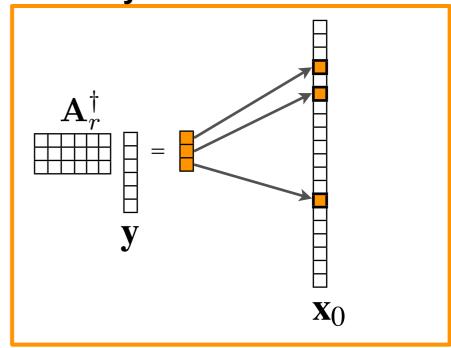
Seismic Laboratory for Imaging and Modeling

NAIVE sparsity-promoting recovery



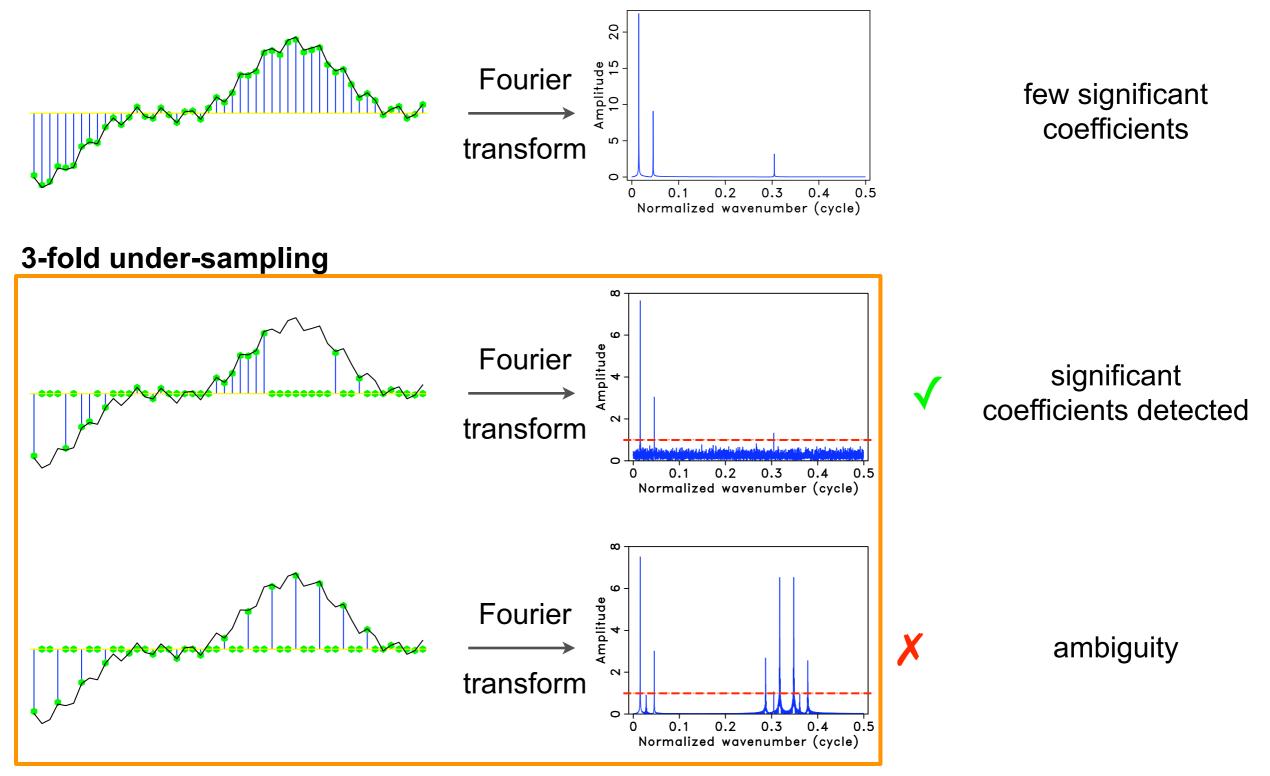


data-consistent amplitude recovery



Seismic Laboratory for Imaging and Modeling

Coarse sampling schemes



Extensions

Incomplete and noisy measurements:

$$\mathbf{y} = \mathbf{\widehat{R}}$$
 \mathbf{M} $\mathbf{m} + \mathbf{n}$

Measurement

y incomplete (compressively sampled) and noisy datam the unknown model

M "arbitrary" measurement matrix

Fourier, eigenfunctions simultaneous sources

R restriction matrix

n Gaussian noise

Extensions

with

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

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0$$
 $\mathbf{A} = \mathbf{R}\mathbf{M}\mathbf{S}^H$ restriction measurement sparsity

matrix matrix matrix

Selection is aimed at turning *aliases/coherent subsampling artifacts* into harmless *noise* ...

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),

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

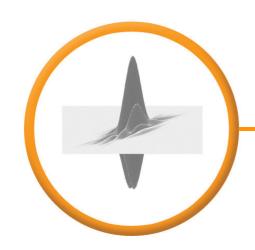
and appropriate subsampling

- controls leakage
- maps interferences into noise
- importance sample in the band
- Compressibility, i.e.,

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

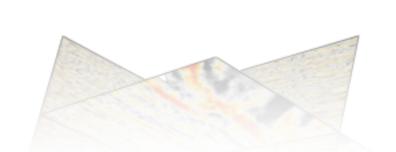


Wavefield reconstruction



Herrmann, F. J. and Hennenfent, G. Non-parametric seismic data recovery with curvelet frames, Geop. J. Int., Vol. 173, No. 1, pp. 233-248, 2008

Hennefent, G. and Herrmann, F. J. Simply denoise: wavefield reconstruction via jittered under-sampling. Geophysics, Vol. 73, No. 3, pp. V19–V28, 2008.



Key elements

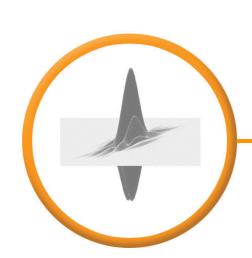
- sparsifying transform
 - typically localized in the time-space domain to handle the complexity of seismic data

- advantageous coarse sampling
 - generates incoherent random undersampling "noise" in the sparsifying domain
 - does not create large gaps
 - because of the limited spatiotemporal extent of transform elements used for the reconstruction

- ☐ sparsity-promoting solver
 - requires few matrix-vector multiplications

THE UNIVERSITY OF BRITISH COLUMBIA | VANCOUVER

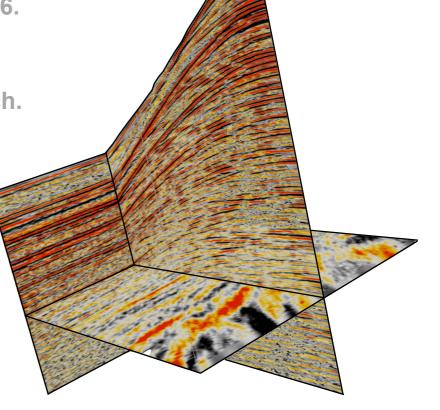


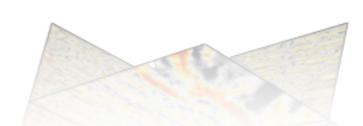


Curvelet transform

Hennenfent, G and Herrmann, F. J., Seismic Denoising with Nonuniformly Sampled Curvelets. Comp. in Sc. & Eng., vol. 8, no. 3, pp. 16-25, 2006.

Herrmann, F. J., Wang, D., Hennenfent, G. and Moghaddam, P. Curvelet-based seismic data processing: a multiscale and nonlinear approach. Geophysics, Vol. 73, No. 1, pp. A1–A5, 2008.



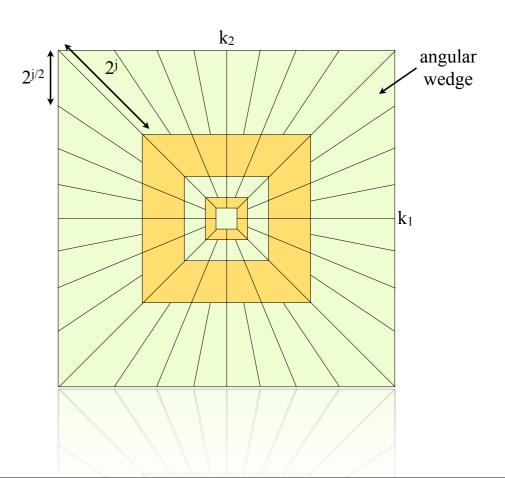


Representations for seismic data

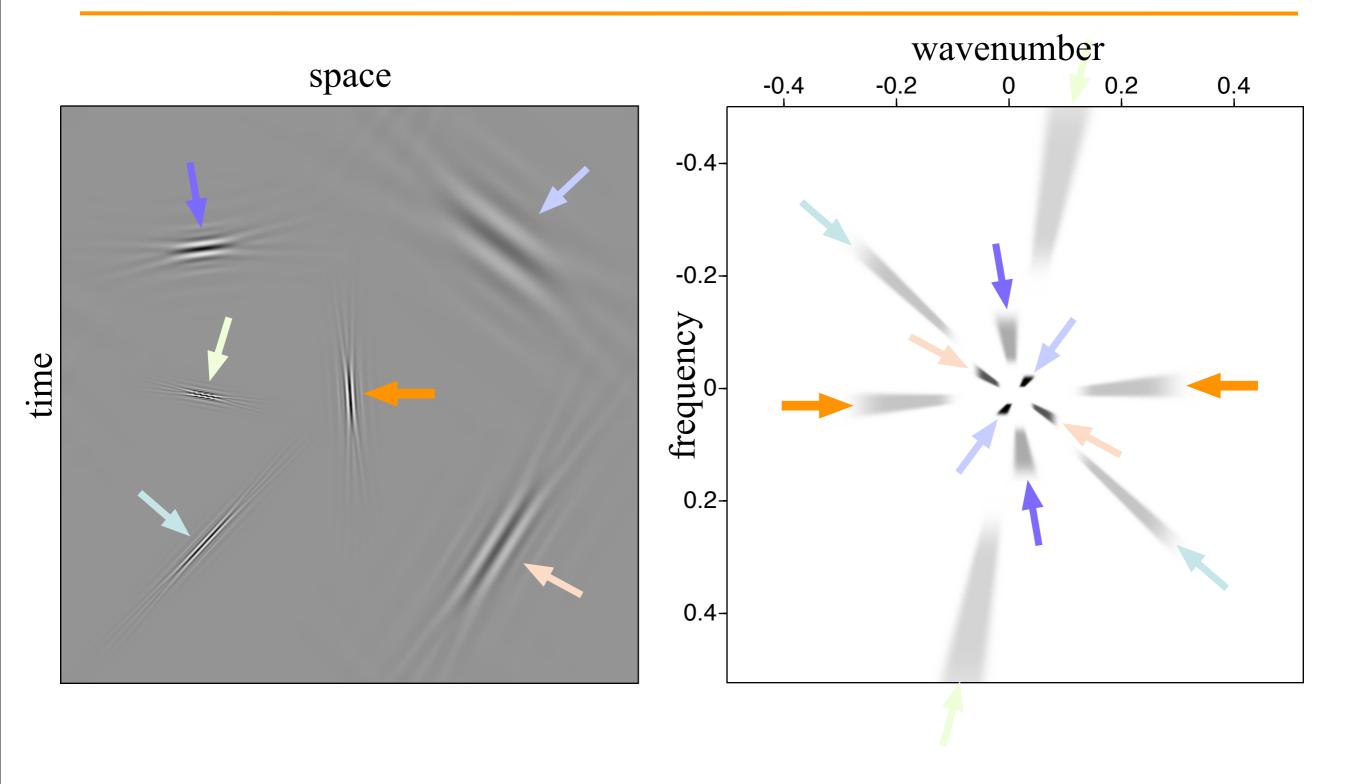
Transform	Underlying assumption
FK	plane waves
linear/parabolic Radon transform	linear/parabolic events
wavelet transform	point-like events (1D singularities)
curvelet transform	curve-like events (2D singularities)

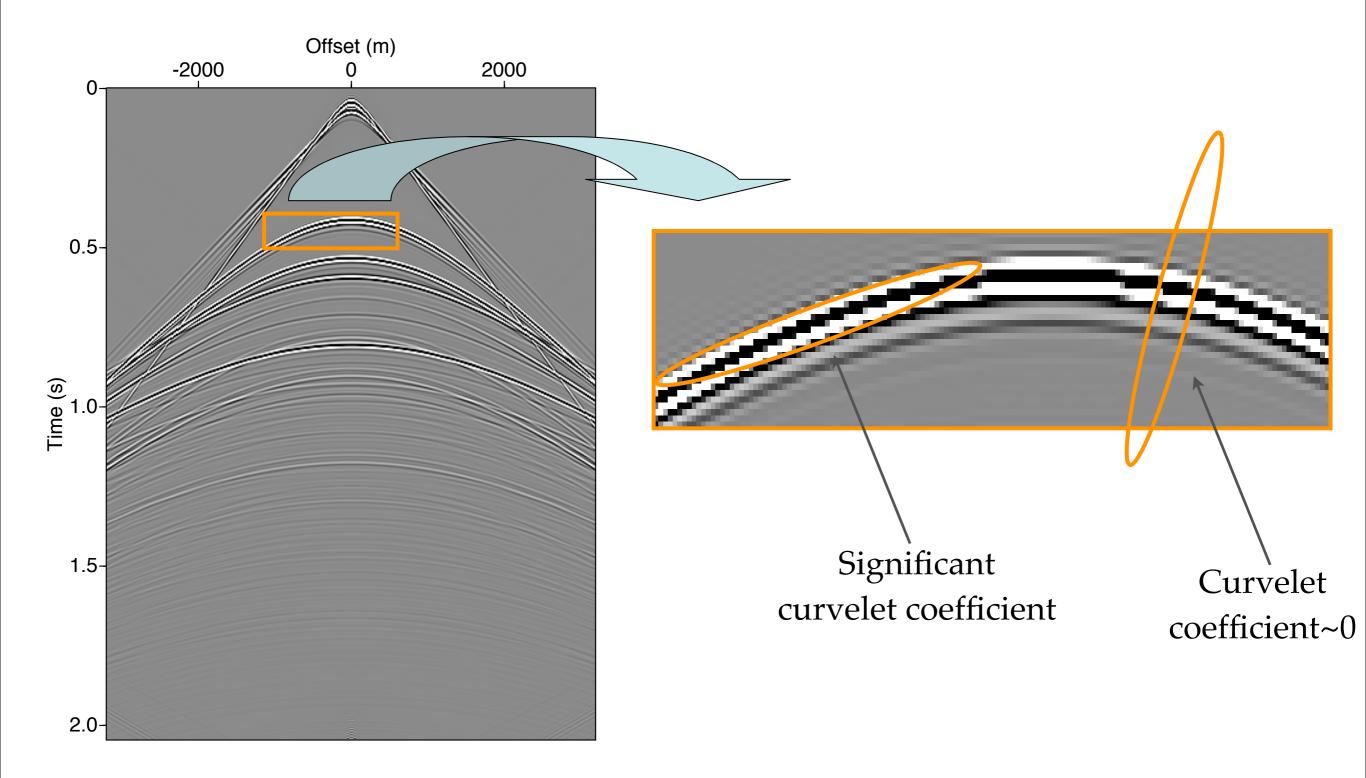
curvelet transform

- multiscale: tiling of the FK domain into dyadic coronae
- multidirectional: coronae subpartitioned into angular wedges, # of angles doubles every other scale
- anisotropic: parabolic scaling principle
- local

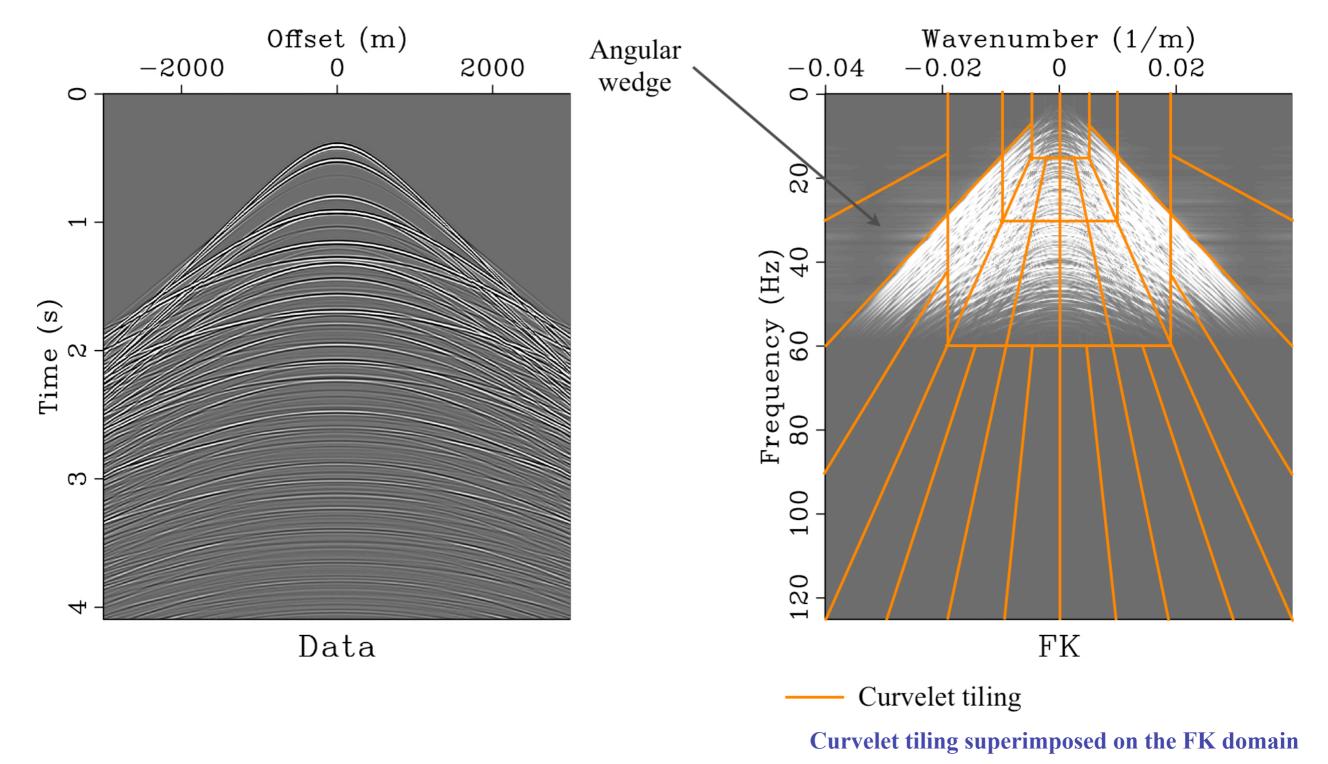


2D discrete curvelets



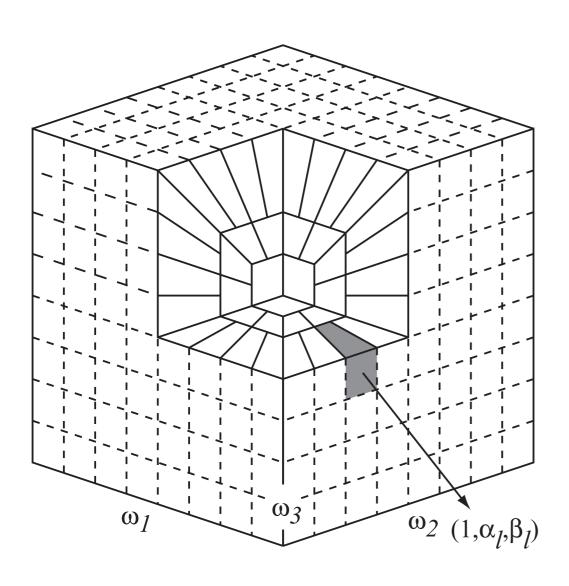


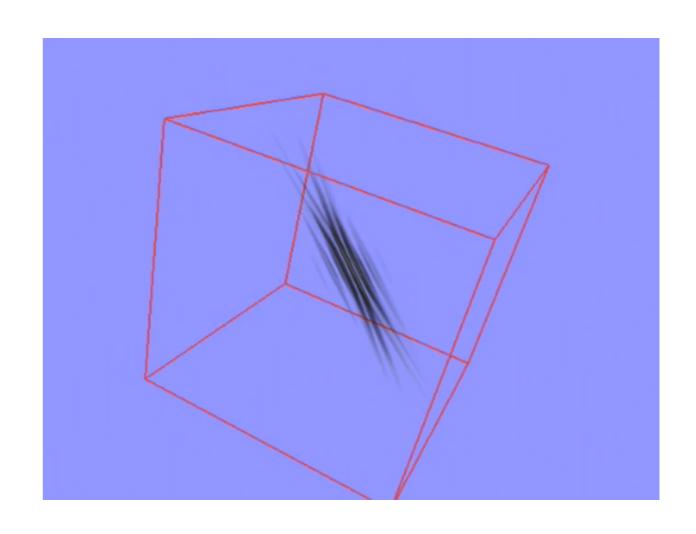
Curvelet tiling & seismic data



Seismic Laboratory for Imaging and Modeling

3D discrete curvelets





Key elements



 typically localized in the time-space domain to handle the complexity of seismic data

advantageous coarse sampling

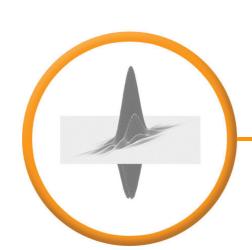
- generates incoherent random undersampling "noise" in the sparsifying domain
- does not create large gaps
 - because of the limited spatiotemporal extent of transform elements used for the reconstruction

☐ sparsity-promoting solver

- requires few matrix-vector multiplications

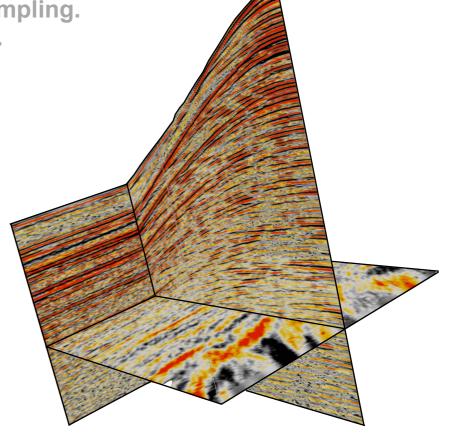
THE UNIVERSITY OF BRITISH COLUMBIA | VANCOUVER

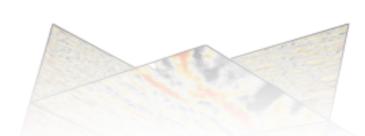




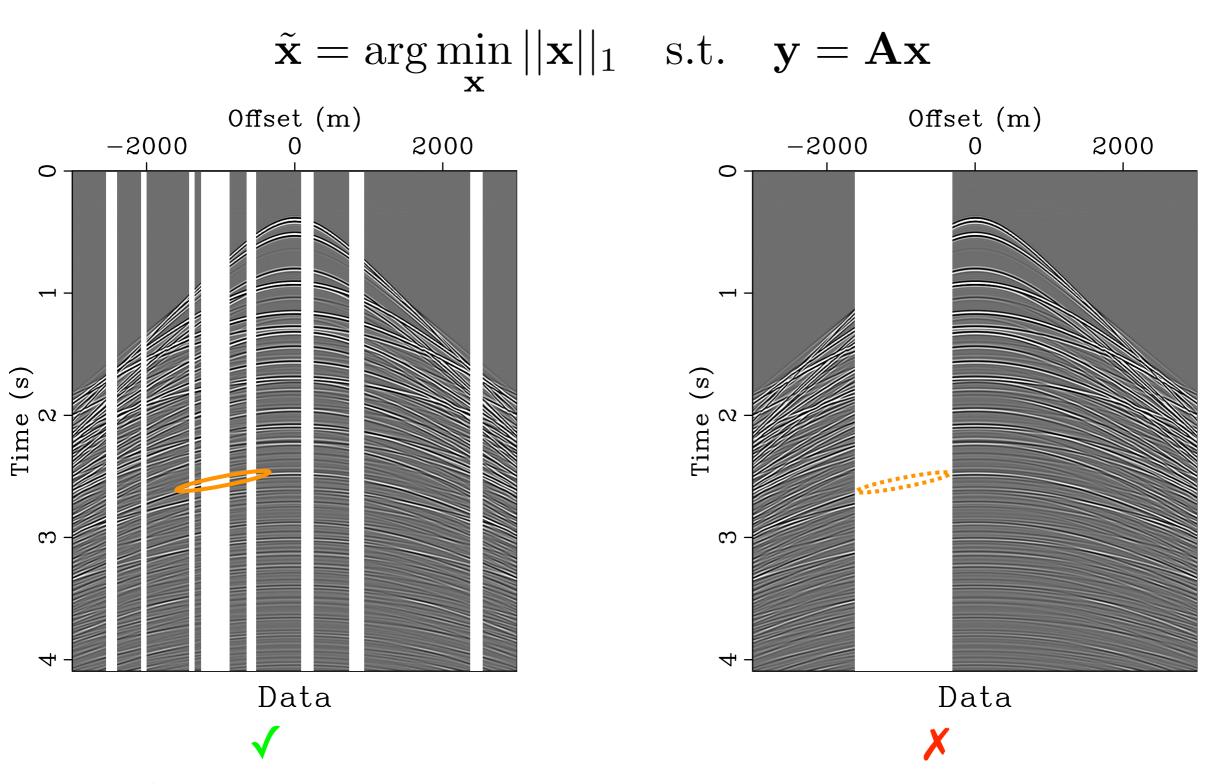
Jittered undersampling

Hennefent, G. and Herrmann, F. J. Simply denoise: wavefield reconstruction via jittered under-sampling. Geophysics, Vol. 73, No. 3, pp. V19–V28, 2008.

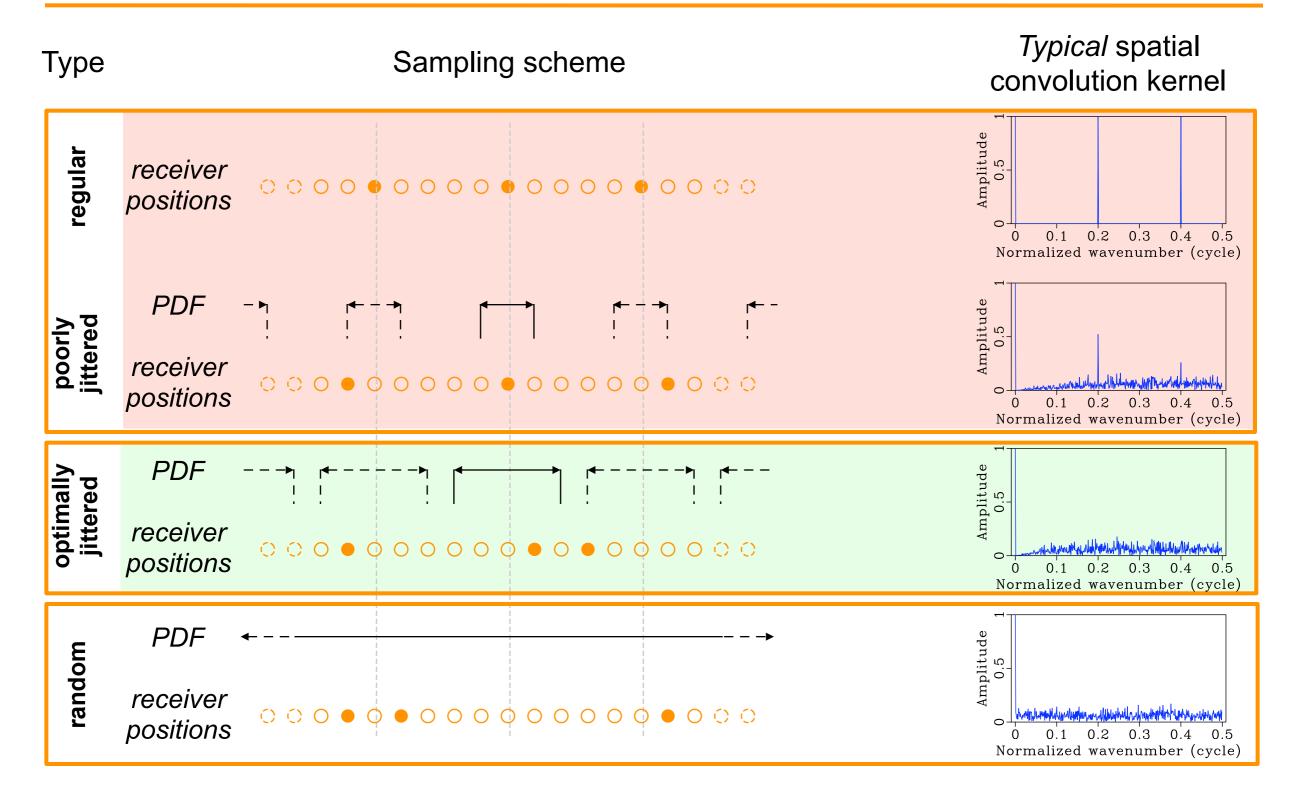




Localized transform elements & gap size



Discrete random jittered undersampling



Key elements

sparsifying transform

 typically localized in the time-space domain to handle the complexity of seismic data

Madvantageous coarse sampling

- generates incoherent random undersampling "noise" in the sparsifying domain
- does not create large gaps
 - because of the limited spatiotemporal extent of transform elements used for the reconstruction

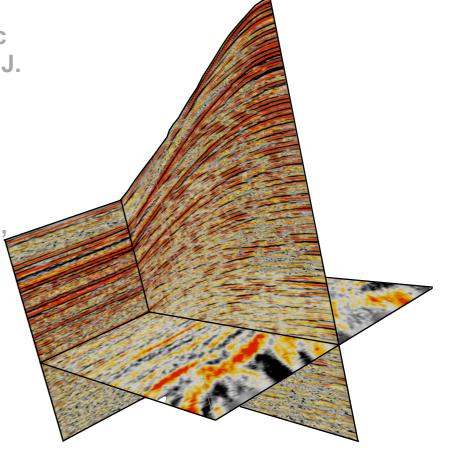
- ☐ sparsity-promoting solver
 - requires few matrix-vector multiplications

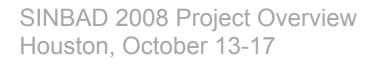


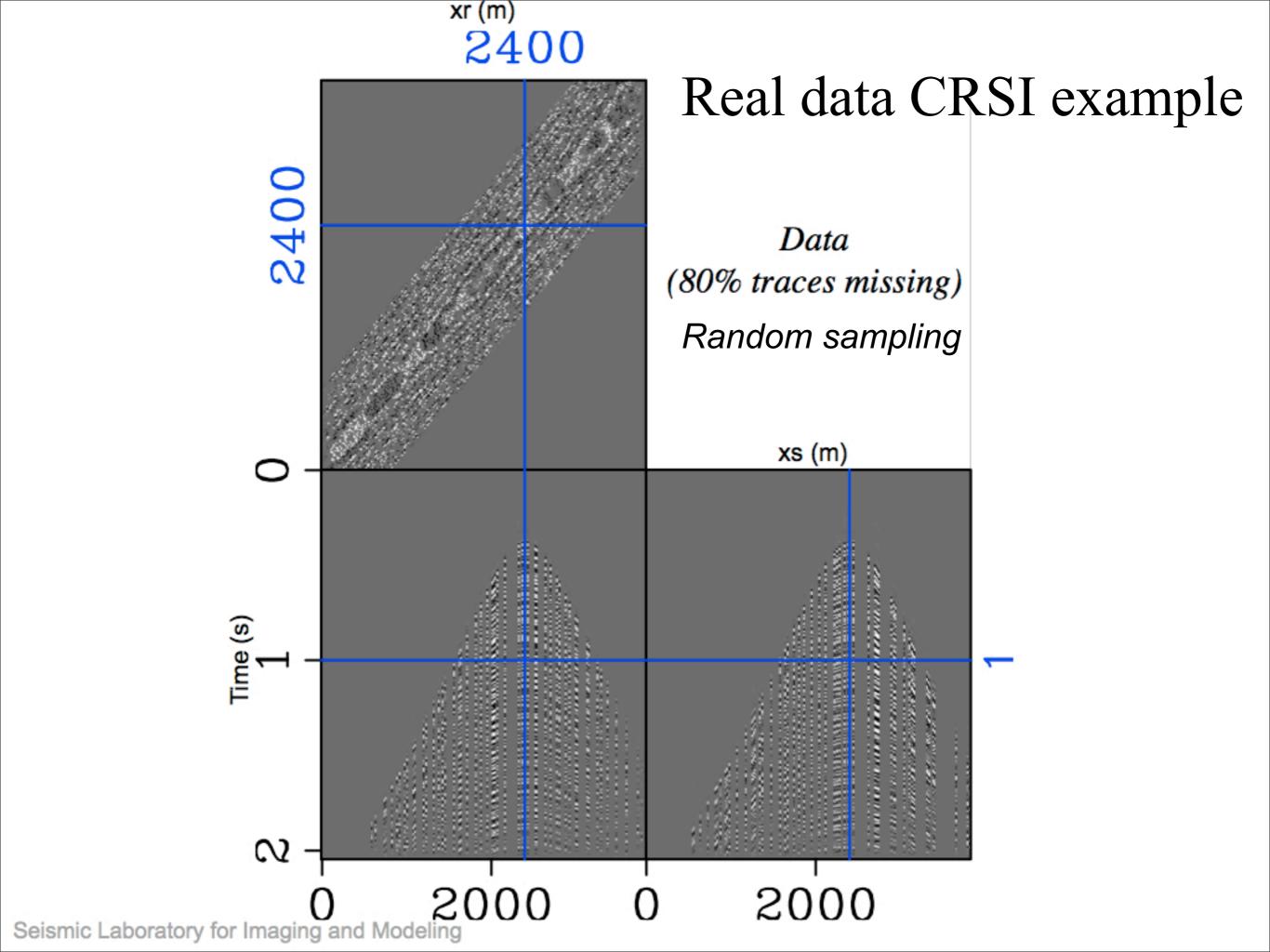
Sparsity-promoting solvers

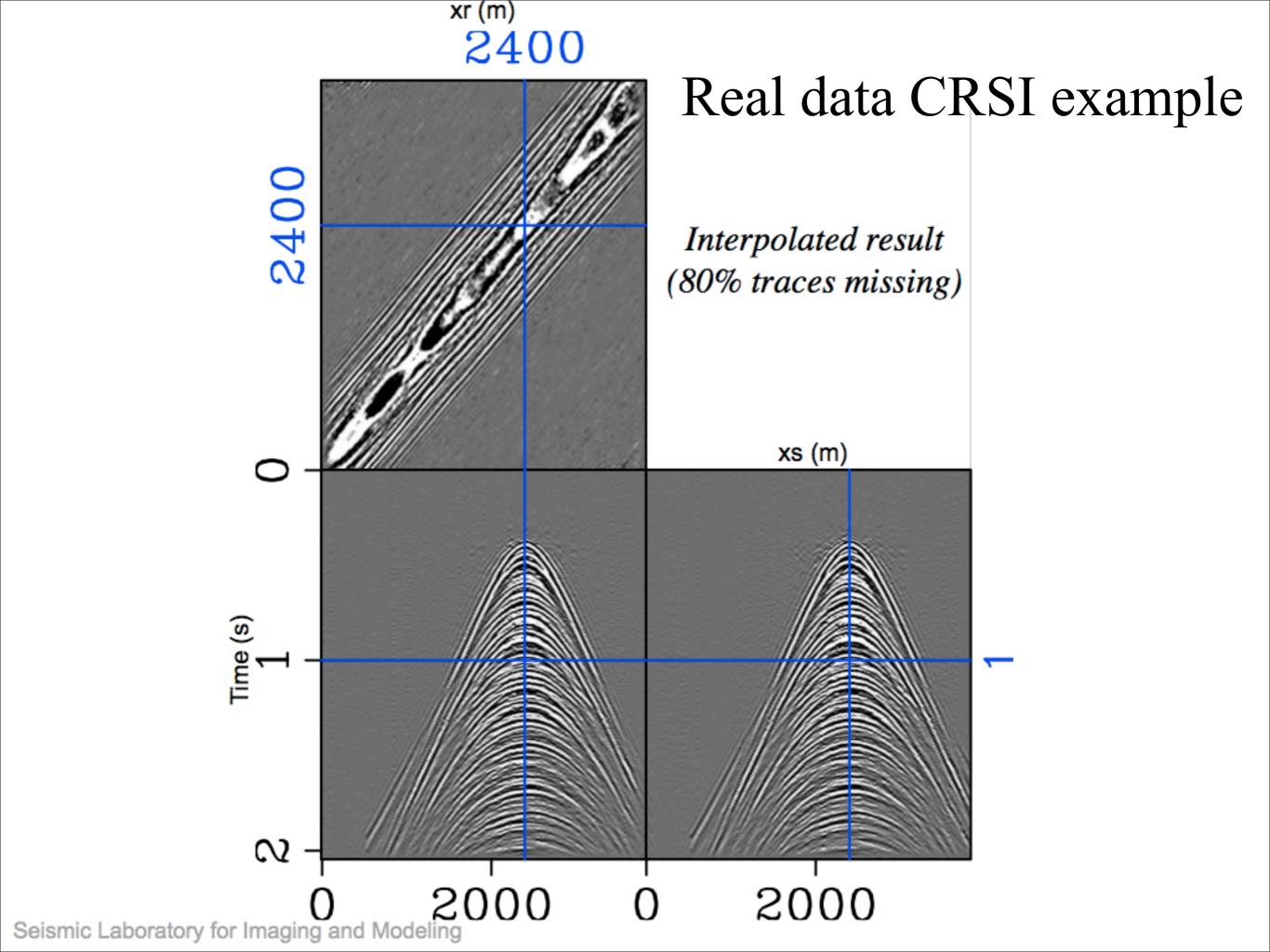
Herrmann, F. J. and Hennenfent, G. Non-parametric seismic data recovery with curvelet frames, Geop. J. Int., Vol. 173, No. 1, pp. 233-248, 2008.

G. Hennenfent, E. van den Berg, M. P. Friedlander, and F. J. Herrmann. New insights into one-norm solvers from the Pareto curve. Geophysics, Vol. 73, No. 4, 2008.









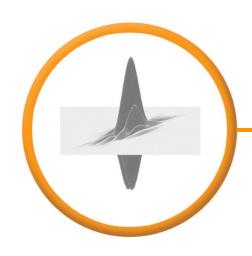
Observations

- Random subsampling breaks constructive interferences, i.e. aliases
- Turns aliases into incoherent noise
- Works by virtue of
 - incoherence (correlations) between the rows of the Dirac measurement basis and the columns of the Fourier synthesis basis
 - maximum spreading of Diracs in Fourier domain
 - maximum leakage
 - independence amongst columns of A, i.e., there exists a subset of columns of A that forms an orthonormal basis
- According to theory of compressive sensing
 - recovery stable w.r.t. noise
 - measurement & sparsity bases can be more general
- Apply these principles to compressive wavefield computations
 - reduce memory imprint
 - reduce the number of sources = number of right-hand sides of linear systems

THE UNIVERSITY OF BRITISH COLUMBIA | VANCOUVER



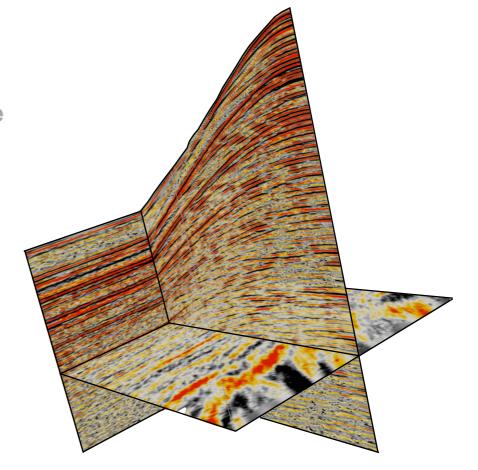
Compressive one-way wavefield extrapolation

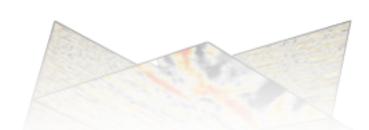


T. T. Y. Lin and F. J. Herrmann. Compressed wavefield extrapolation. Geophysics,

72(5):SM77-SM93, 2007.

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





Delft University of Technology Delft, February 2, 2009

Introduction

Goal: employ the complete one-Way Helmholtz operator for w

Grimbergen, J., F. Dessing, and C. Wapenaar, 1998, Modal expansion of one-way operator on laterally varying media: Geophysics, **63**, 995–1005.

$$\mathbf{w}^{\pm} = e^{\mp j\Delta x \mathbf{H}_1}$$
 $\mathbf{H}_2 = \mathbf{H}_1 \mathbf{H}_1$

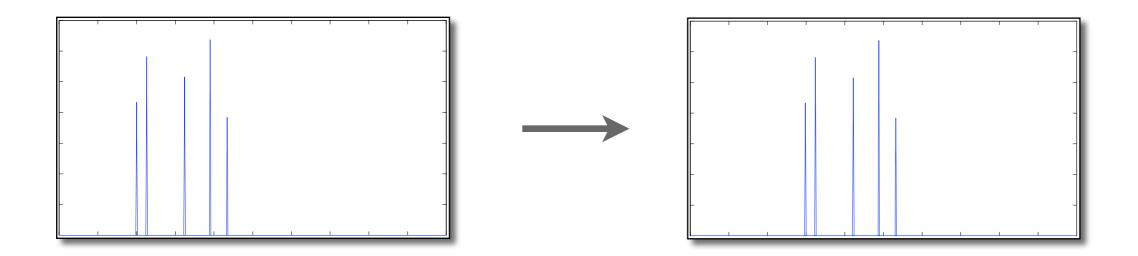
- Problem: computation & storage complexity
 - lacktriangle creating and storing \mathbf{H}_2 is trivial
 - however \mathbf{H}_1 is not trivial to compute and store

$$\mathbf{H}_2 = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1 \end{bmatrix}$$



Our approach

 Consider a related, but simpler problem: shifting (or translating) a signal



- lacksquare operator is $\mathbf{T}_{\Delta x} = e^{-j \frac{\Delta x}{2\pi} \mathbf{D}}$
- D is differential operator

$$\mathbf{D} = igg[igg|_{}^{}$$



Our approach

Computation requires similar approach to

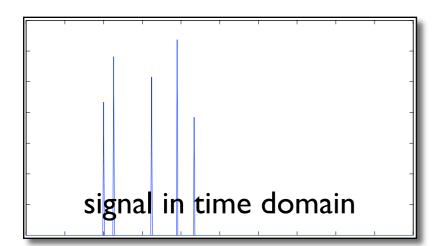


$$\mathbf{D} = \mathbf{L} \mathbf{\Lambda} \mathbf{L}^{\mathrm{T}} = \begin{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{L} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{L} \end{bmatrix}$$

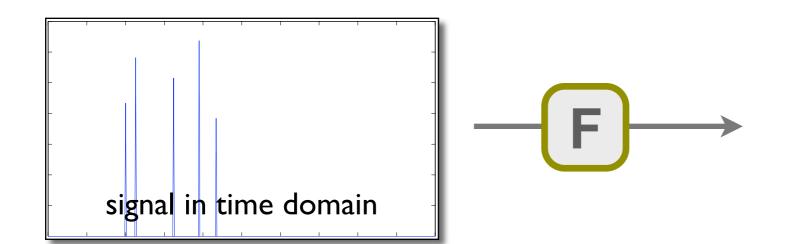
$$\mathbf{T}_{\Delta x} = \begin{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{L} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{L} \end{bmatrix} \begin{bmatrix} \mathbf{L} \\ \mathbf{L} \end{bmatrix}$$

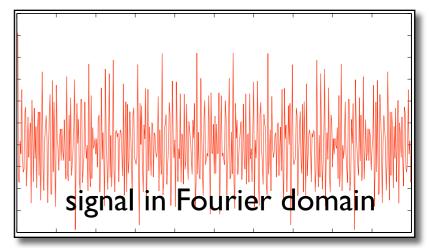
$$\mathbf{L} = e^{-j\frac{\Delta x}{2\pi}\mathbf{\Lambda}} \mathbf{L}^{\mathrm{T}}$$

 $lue{f \Box}$ However, for ${f D}$, ${f L}={
m DFT}$, so computation trivial with FFT

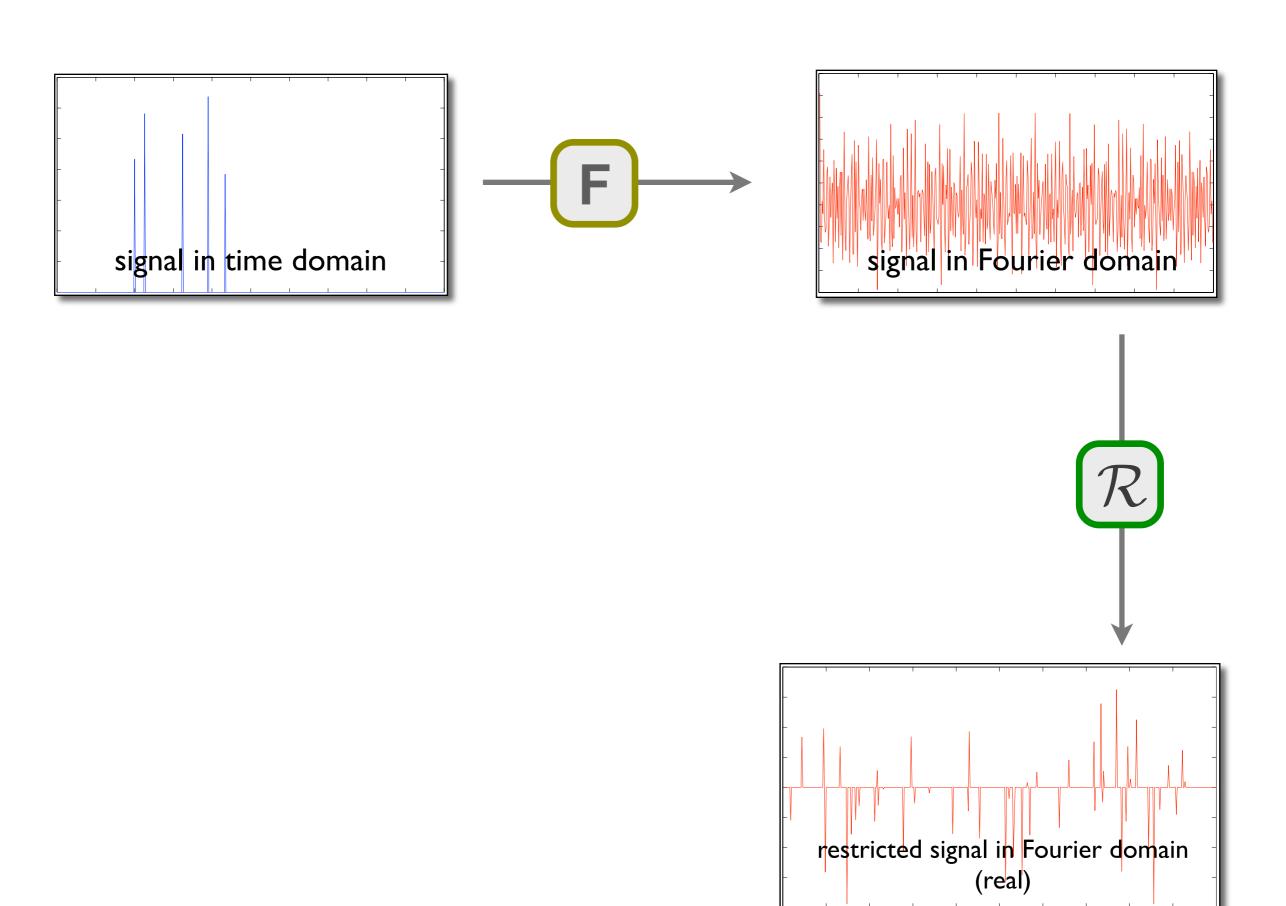




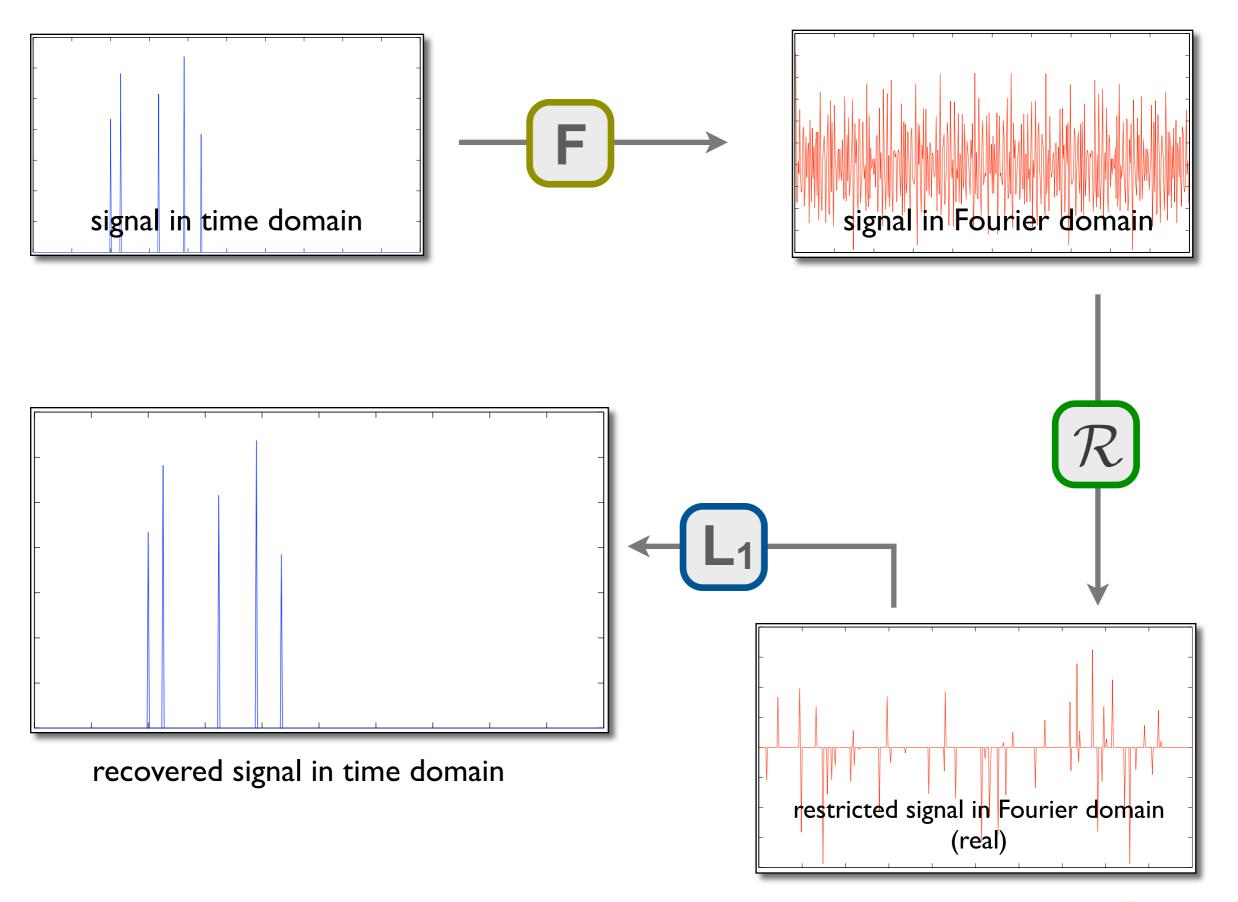






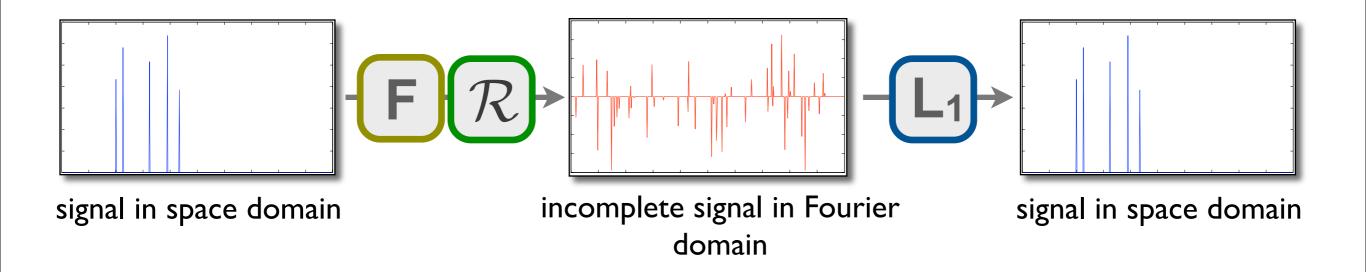




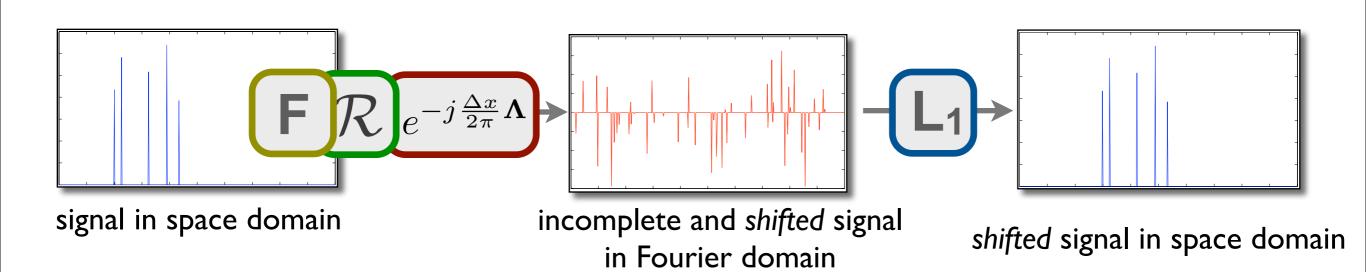




Compressed Sensing

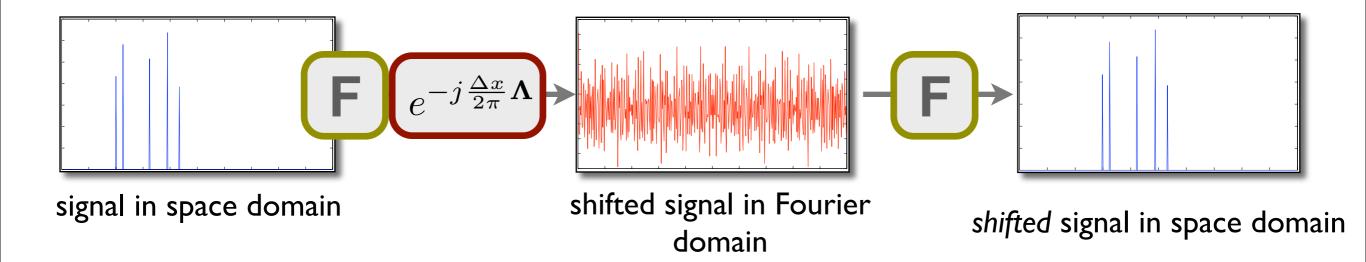


Compressed Computations

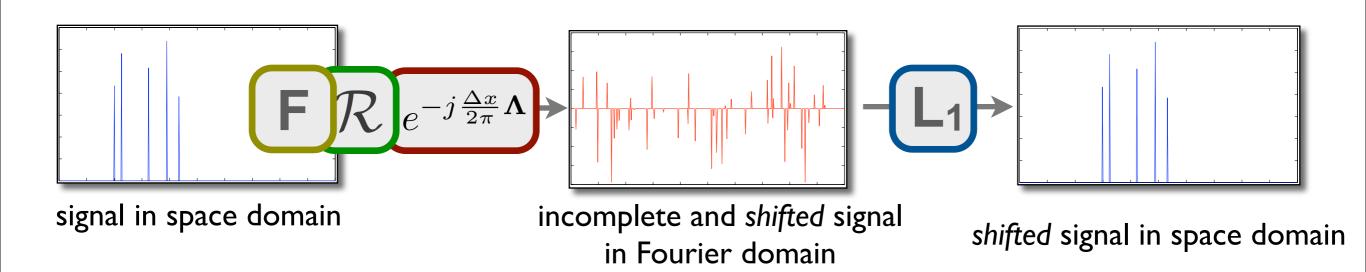




Straightforward Computation

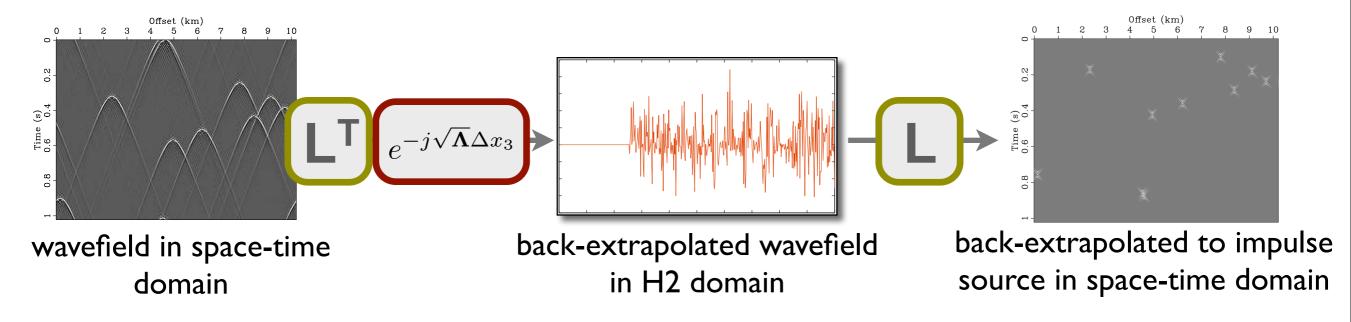


Compressed Computation

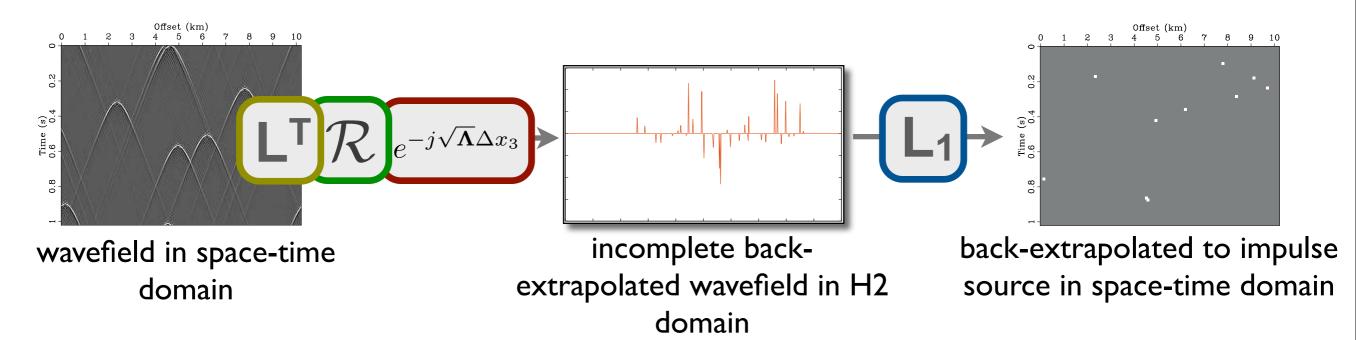




Straightforward 1-Way inverse Wavefield Extrapolation



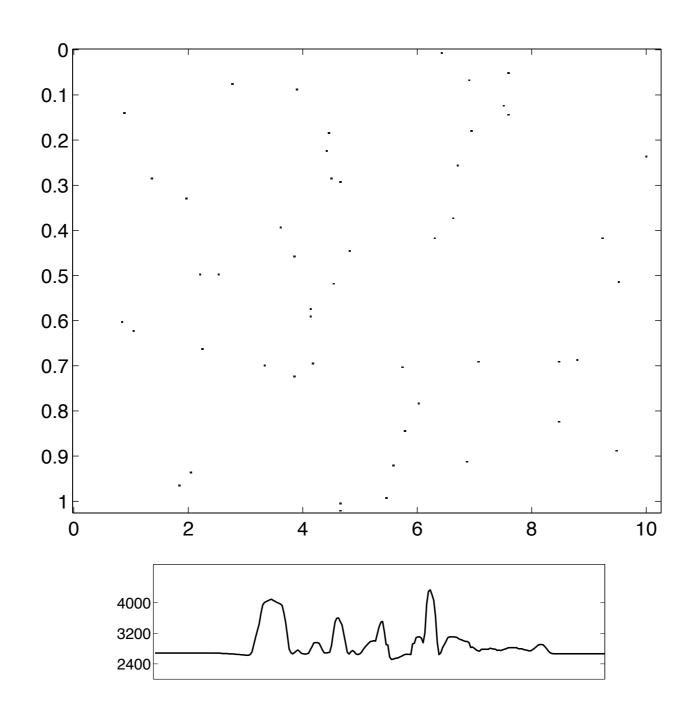
Compressed 1-Way Wavefield Extrapolation





Compressed wavefield extrapolation

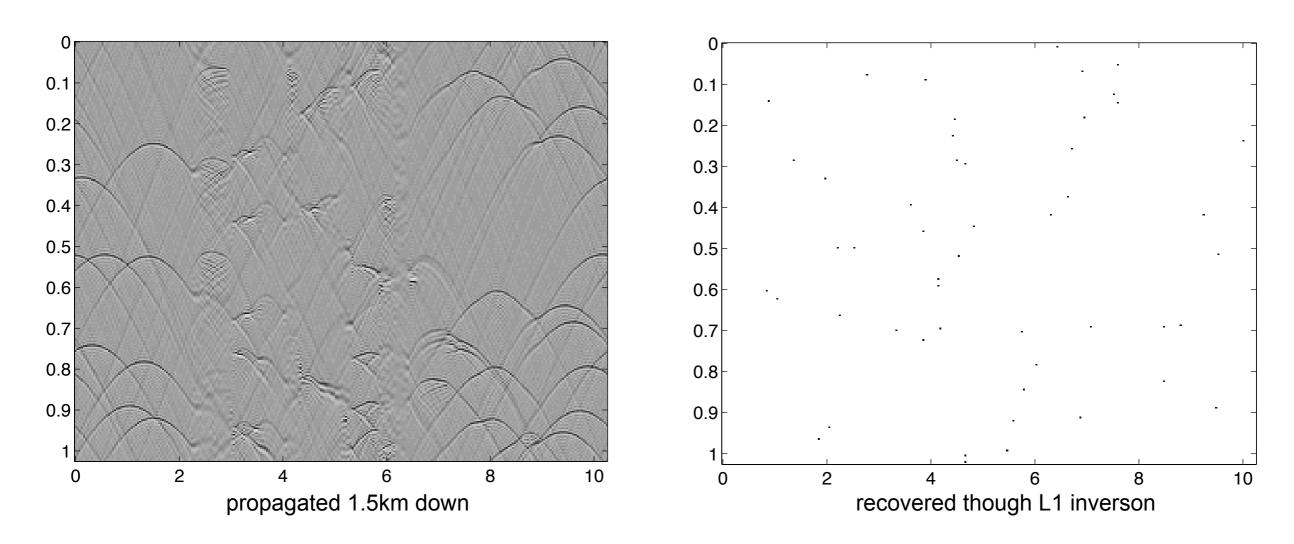
simple 1-D space/time propagation example with point scatters





Compressed wavefield extrapolation

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

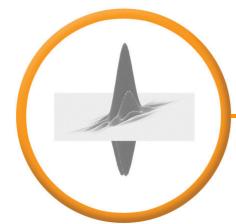


Restricted L transform to ~0.01 of original coefficients

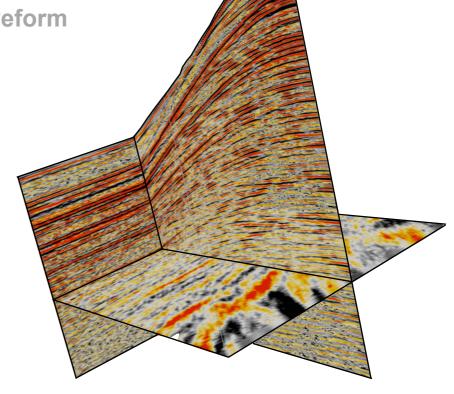


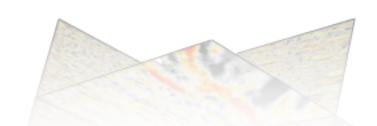


Simultaneous simulation & recovery



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.





Wavefield computations

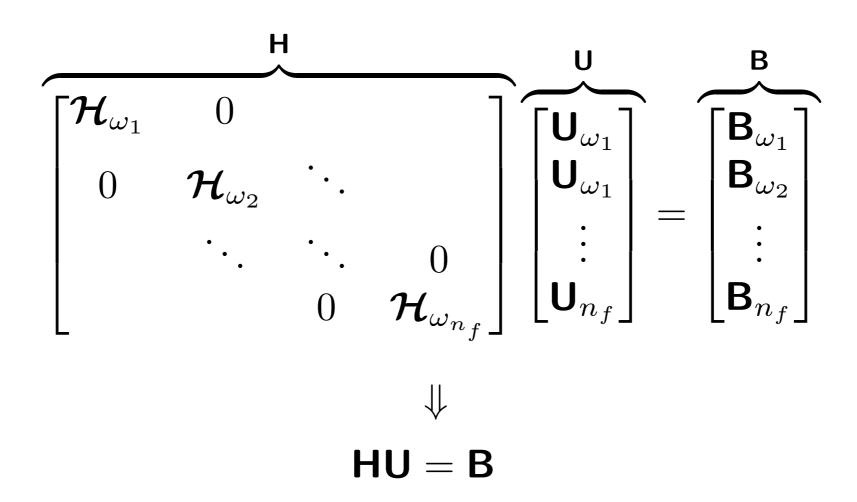
- ullet For n_s shots and n_f frequencies, the linear (Helmholtz) systems are independent
- The multi-shot and multi-frequency problem is embarrassingly parallel

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

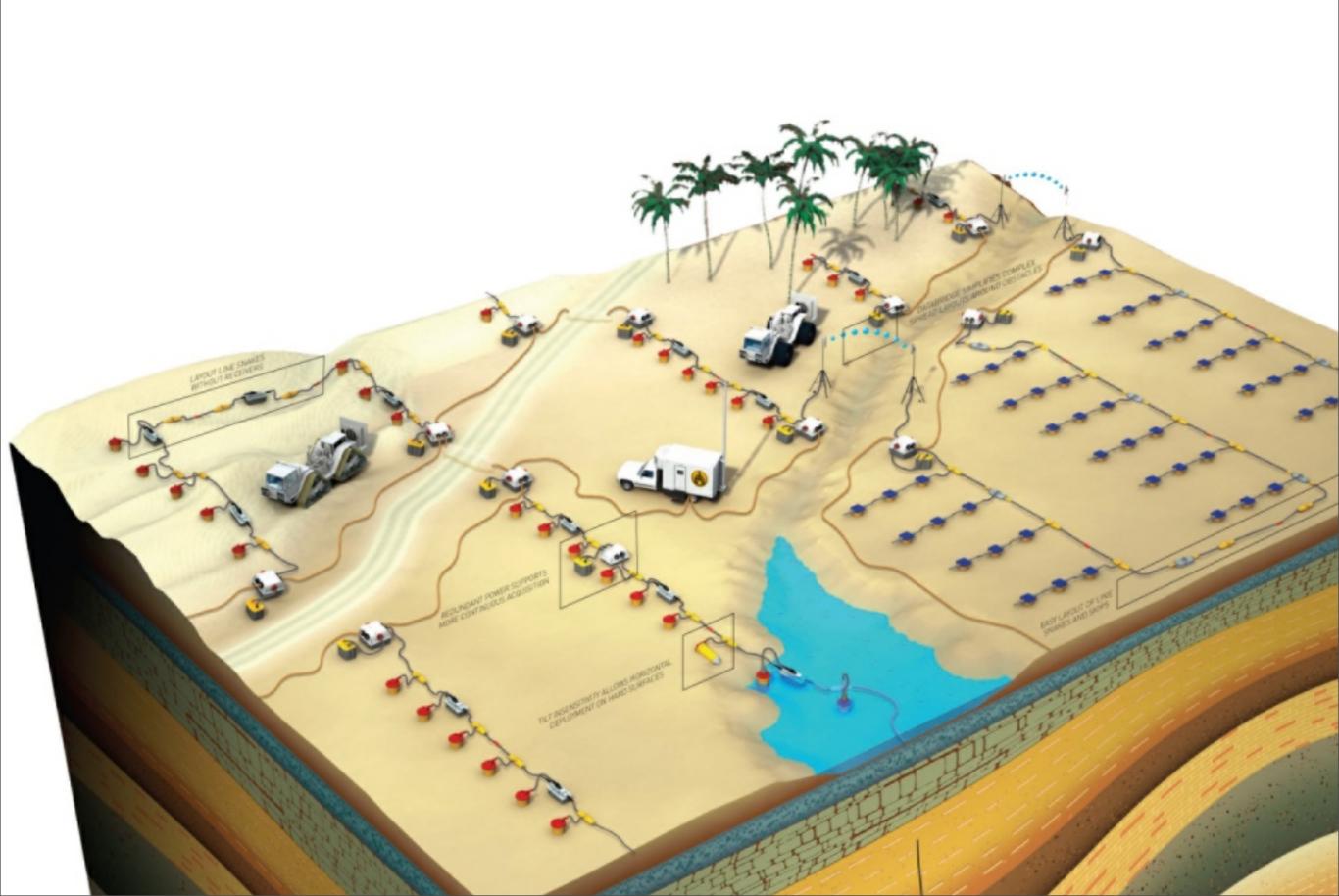
$$\mathcal{H}_{\omega_j} := \mathcal{H}(\omega_j), \quad \omega_j = 2\pi j \Delta f, \quad j = 1, \dots, n_f$$

 Δf frequency sample interval

Wavefield computations



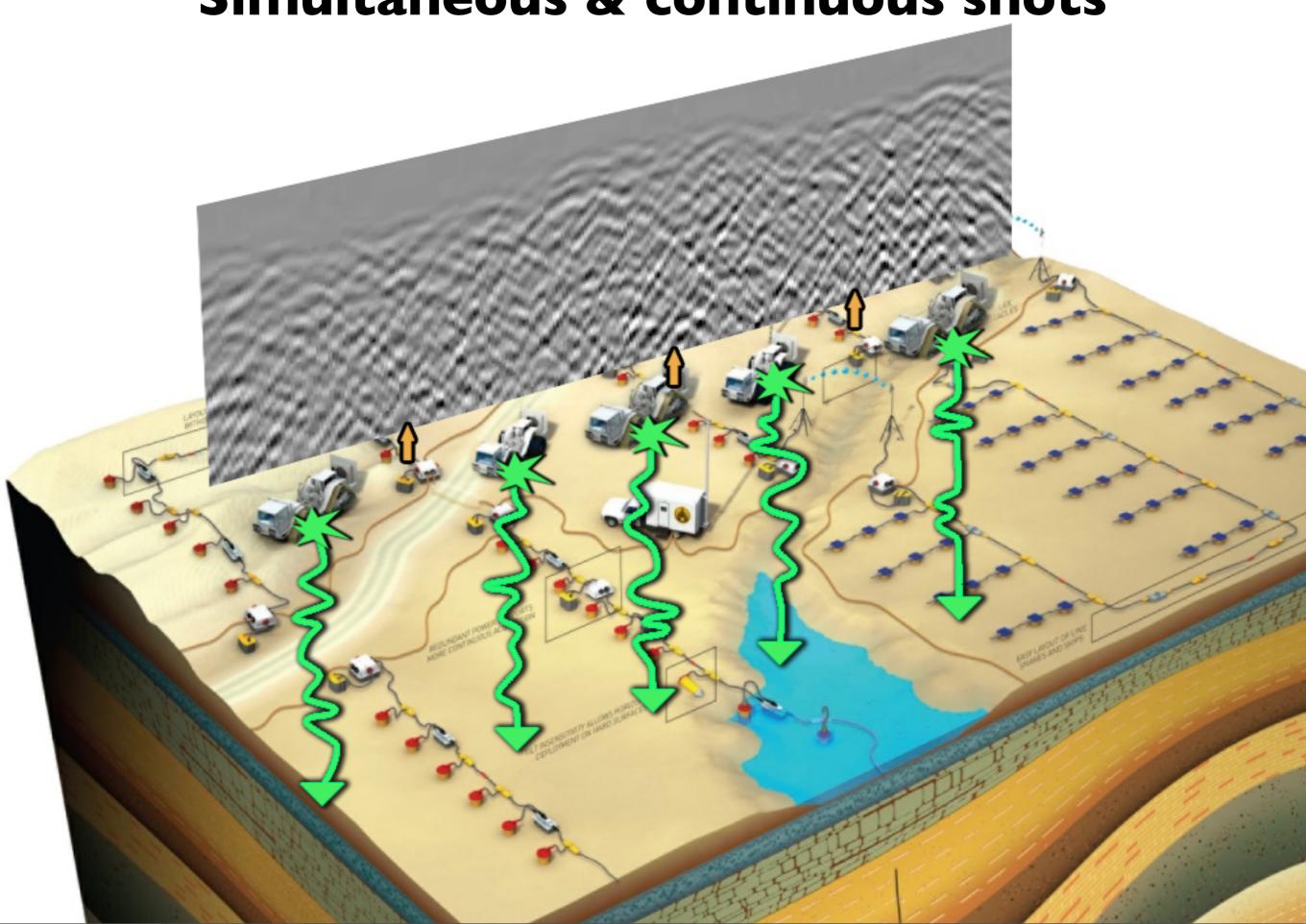
Natural for CS setting!



Individual shots

Individual shots

Simultaneous & continuous shots



Simultaneous modeling & acquisition

Current paradigm:

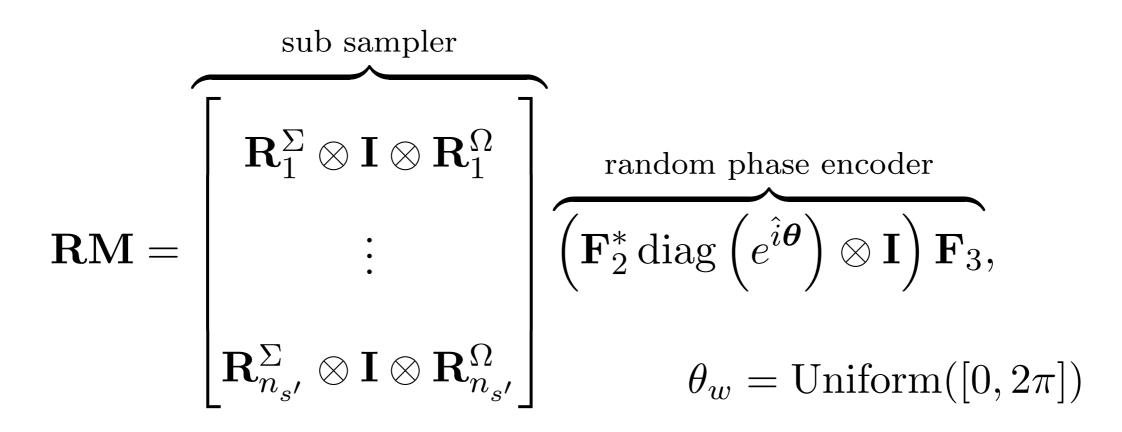
- separate single-source experiments in the field
- separate single-shot simulations in the computer
- Con: expensive

New paradigm:

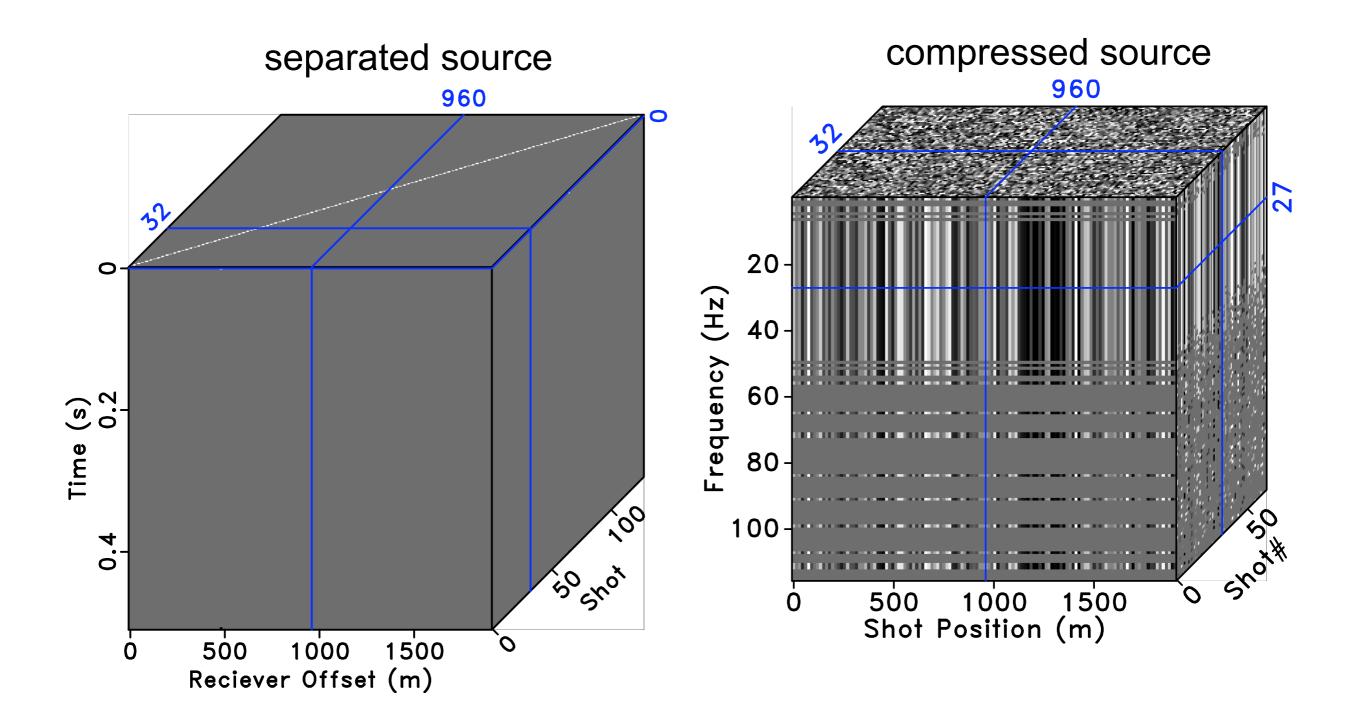
- simultaneous & continuous source experiments in the field
- simultaneous (continuous) simulations in the computer
- continuous simultaneous simulations are equivalent to multiple simultaneous experiments
- Con: postprocessing necessary to separate into individual shots
- Key observation: this is really another instance of CS ...
 - design new simultaneous acquisitions & recovery schemes!

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 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



Applying to Shot Sources





Equivalence

Show equivalence between

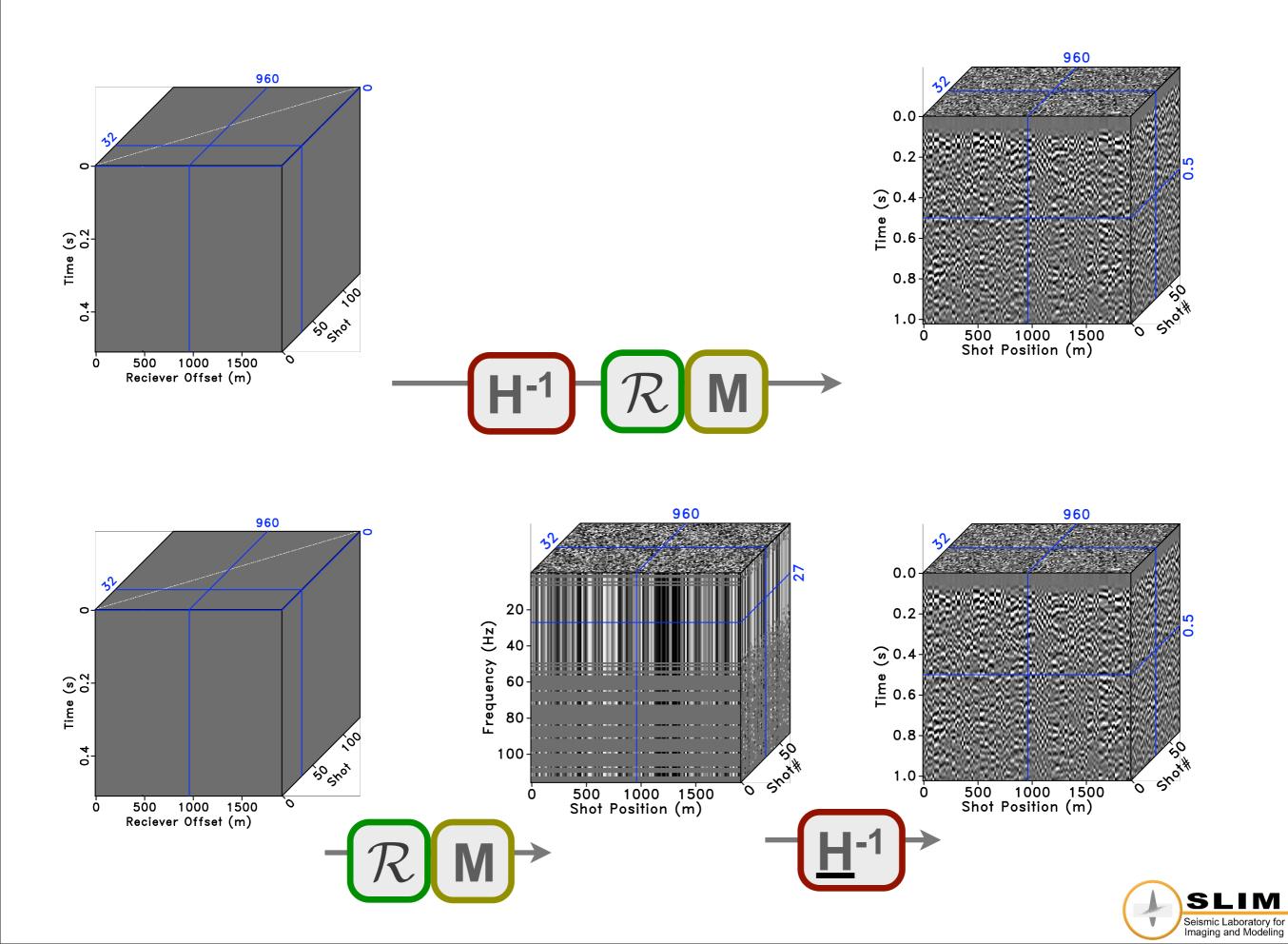
- CS sampling of *full* solution for separate single-source (sweep) experiments
- Solution of reduced system after CS sampling the collective single-shot source wavefield => simultaneous source experiments

$$\begin{cases} \mathbf{B} = \mathbf{D}^* \quad \mathbf{\underline{s}} \\ \mathbf{H}\mathbf{U} = \mathbf{B} \\ \mathbf{y} = \mathbf{R}\mathbf{M}\mathbf{D}\mathbf{U} \end{cases} \iff \begin{cases} \underline{\mathbf{B}} = \mathbf{B} \\ \mathbf{\underline{H}}\mathbf{U} \\ \underline{\mathbf{y}} = \mathbf{B} \end{cases}$$

$$\begin{cases} \underline{B} = \underline{D}^* & \underline{RMs} \\ & \text{simul. shots} \\ \underline{HU} = \underline{B} \\ \underline{y} = \underline{DU} \end{cases}$$

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





CS

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

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 $(SPG\ell_1, Friedlander \& van den Berg, `07-'08)$
- CS matrix has to lead to physically realizable source wavefield for modeling & acquisition

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

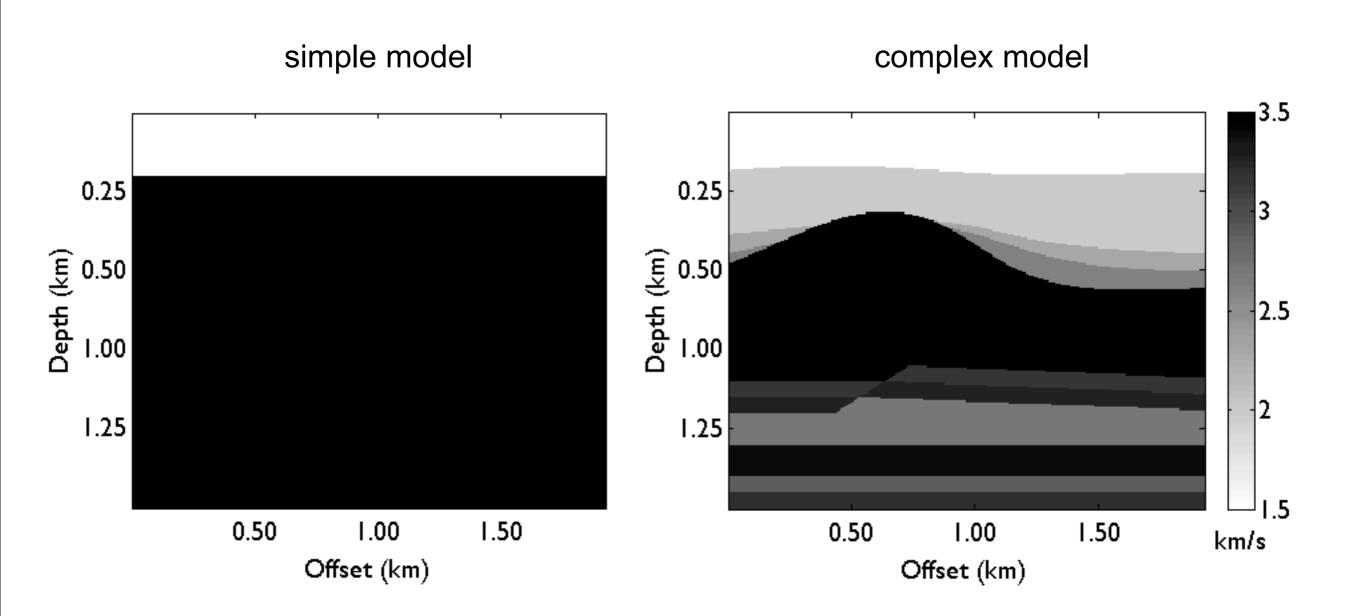
- Wavelet sparsity on temporal-frequency coordinate
- 2D Curvelet sparsity on shot and receiver plane

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

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

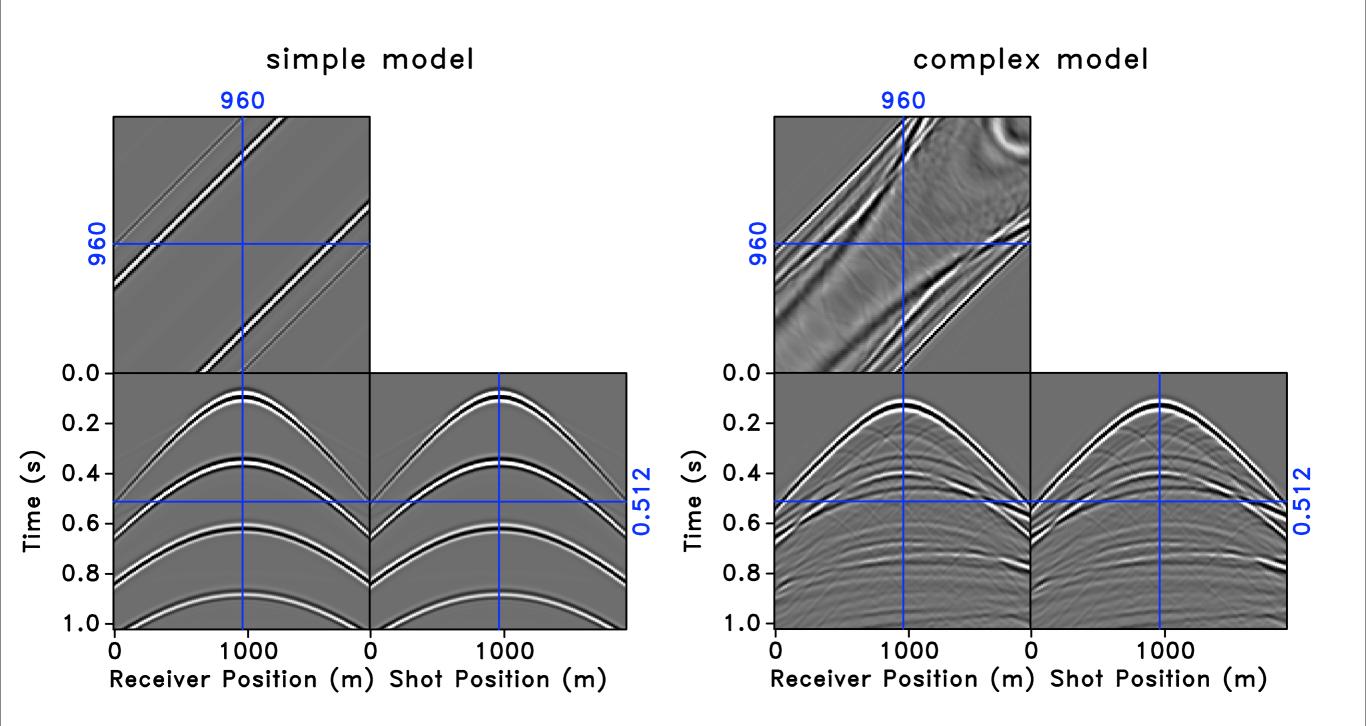


Velocity models



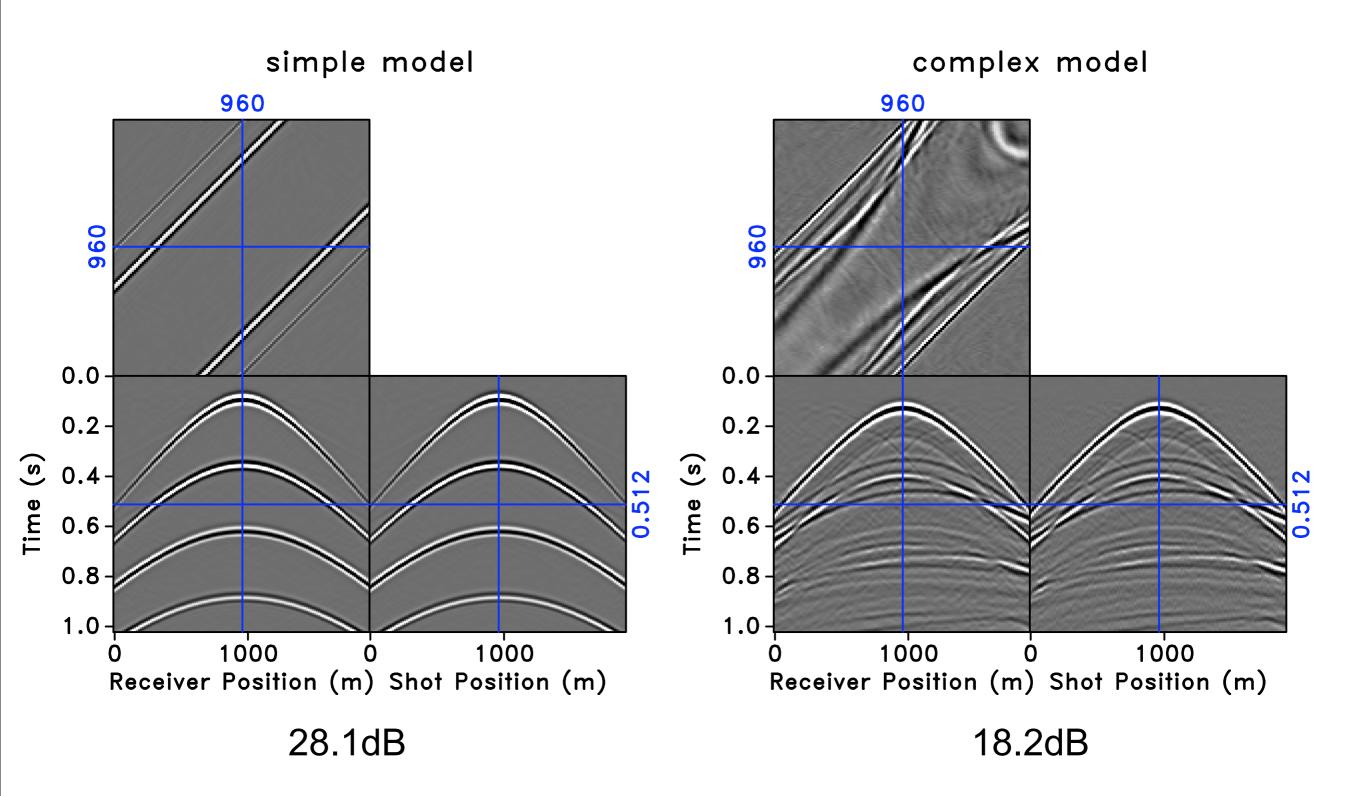


Green's functions



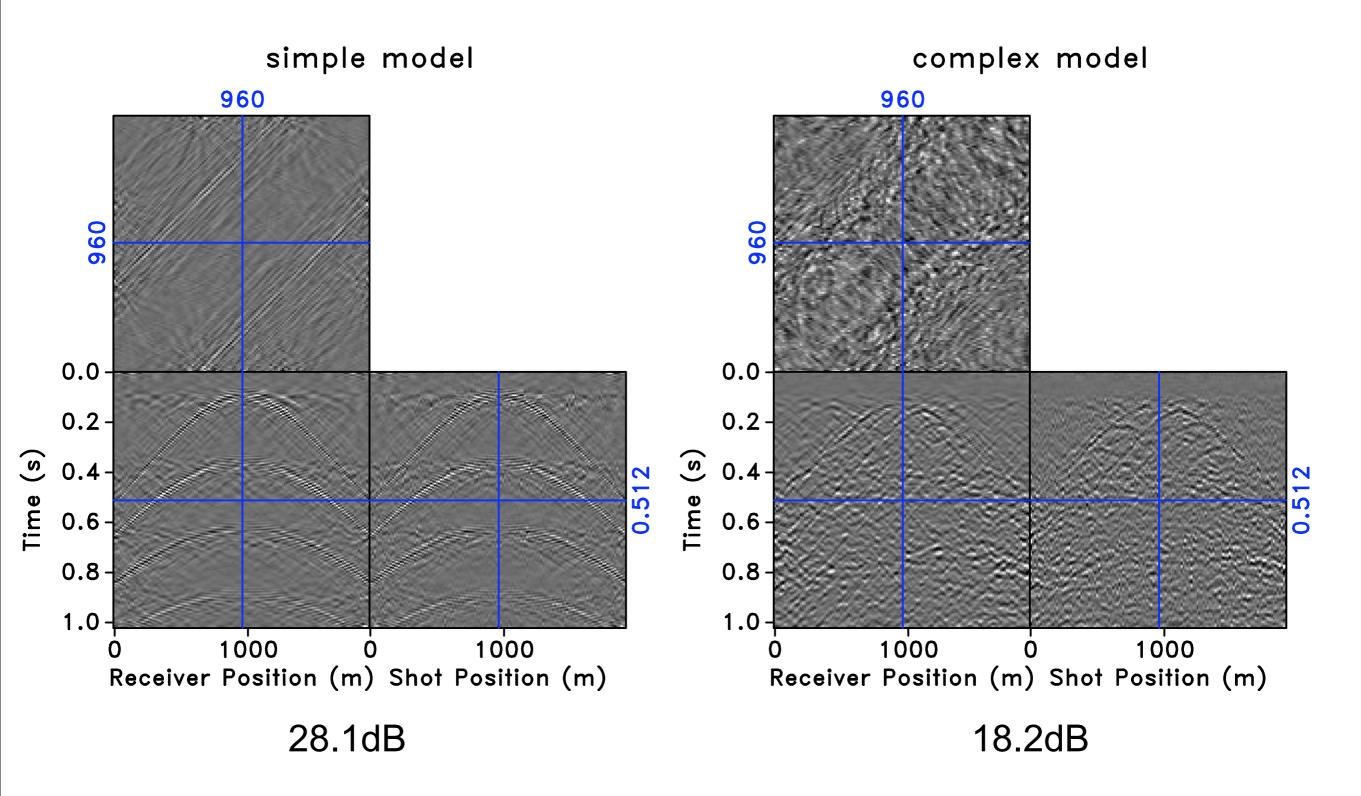


Recovered data





Difference







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!