Seismology meets compressive sampling

Felix J. Herrmann

Seismic Laboratory for Imaging and Modeling
Department of Earth and Ocean Sciences
University of British Columbia (Canada)
slim.eos.ubc.ca

IPAM,

UCLA, October 29

SLIM
Seismic Laboratory for Imaging and Modeling

General trend

(Seismic) data sets are becoming larger and larger

Demand for more information to be inferred from data

Data collection is expensive

Distilling information is time consuming

Industry ripe for recent developments in applied harmonic analysis and information theory

Today’s topics

Problems in seismic imaging

- acquisition, processing & imaging costs

Compressive sampling in exploration seismology

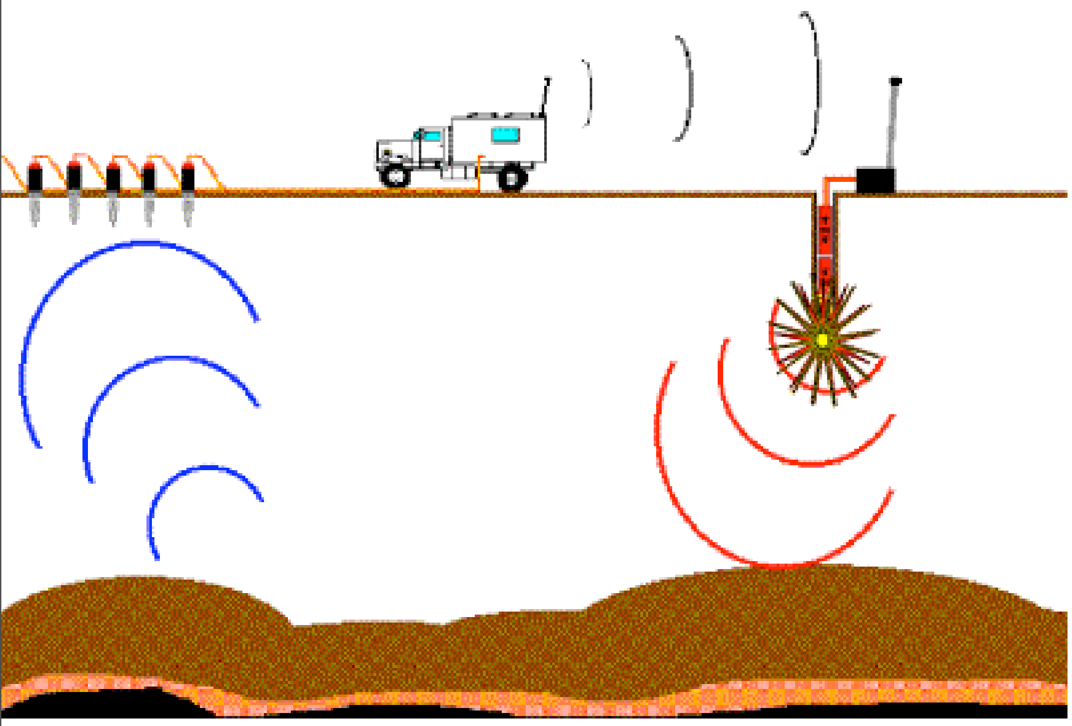
- wavefield recovery from jittered sampling
- compressive wavefield extrapolation
- road ahead: compressive computations

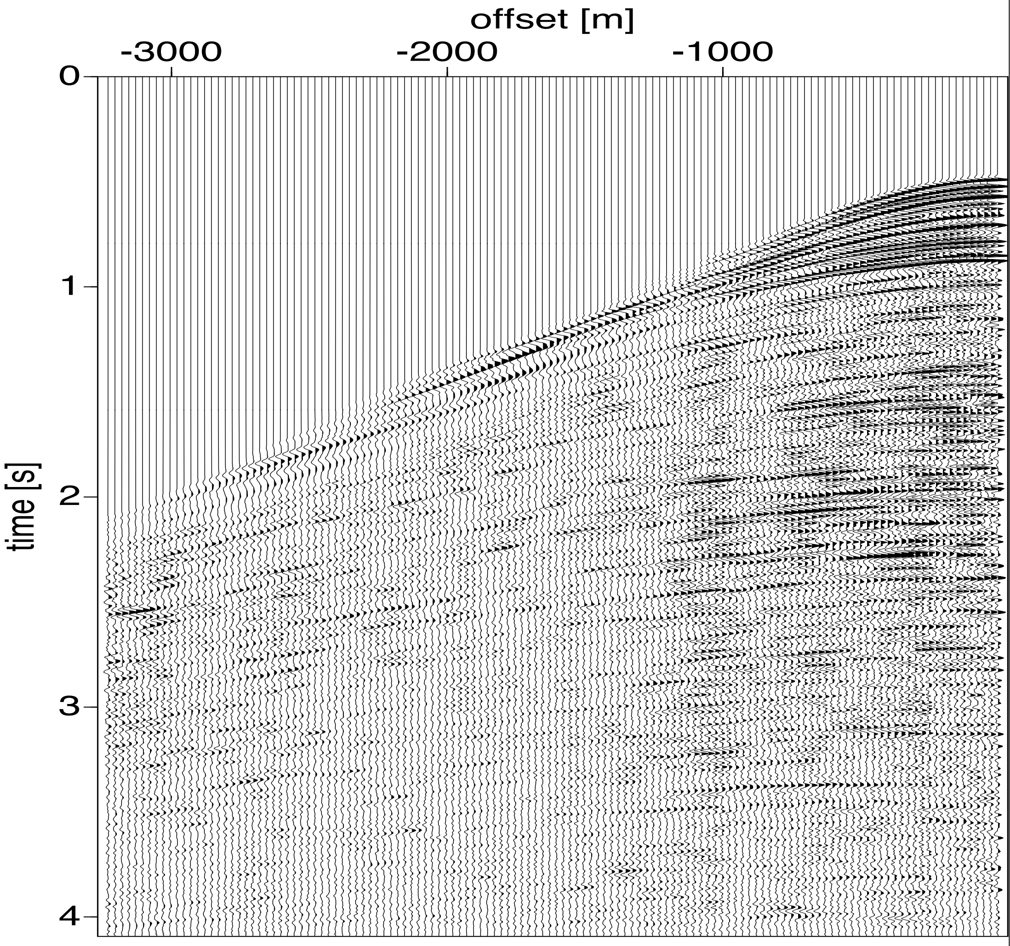
DNOISE: an academic-industry-NSERC partnership

- truly interdisciplinary academic collaboration
- knowledge dissemination

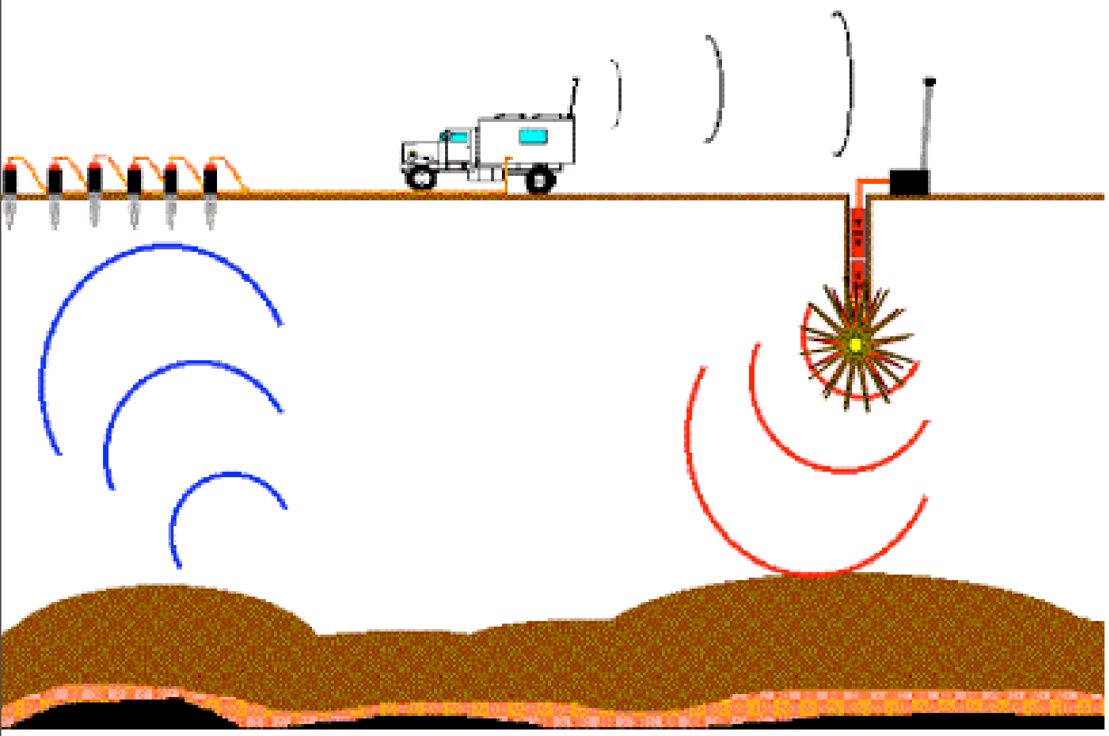
SLIM Seismic Laboratory for Imaging and Modeling

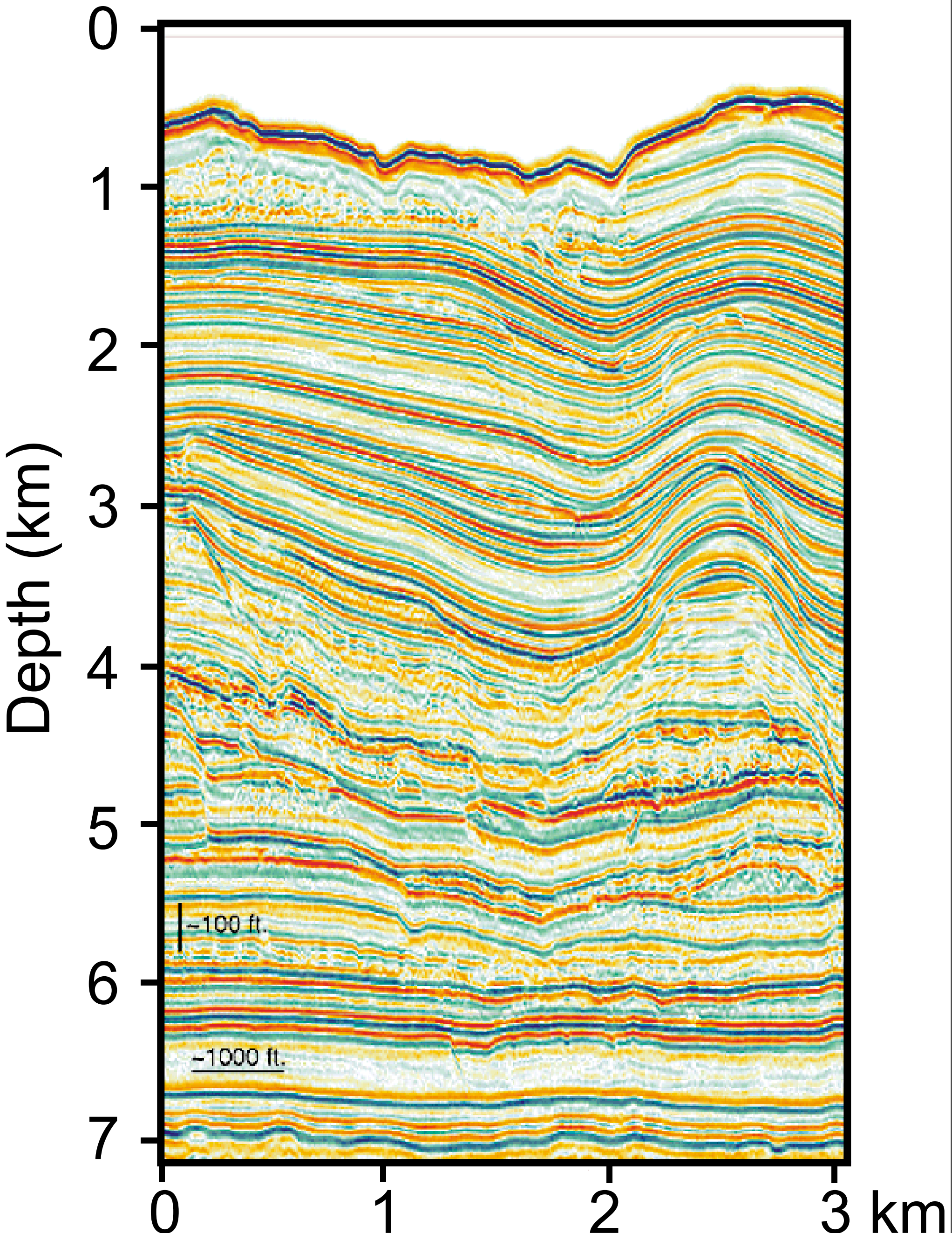
Seismic data acquisition





Exploration seismology





- **create images of the subsurface**
- **need for higher resolution/deeper**
- **clutter and data incompleteness**
- **image repeatability <=> monitoring**

Today's challenges

Seismic data volumes are

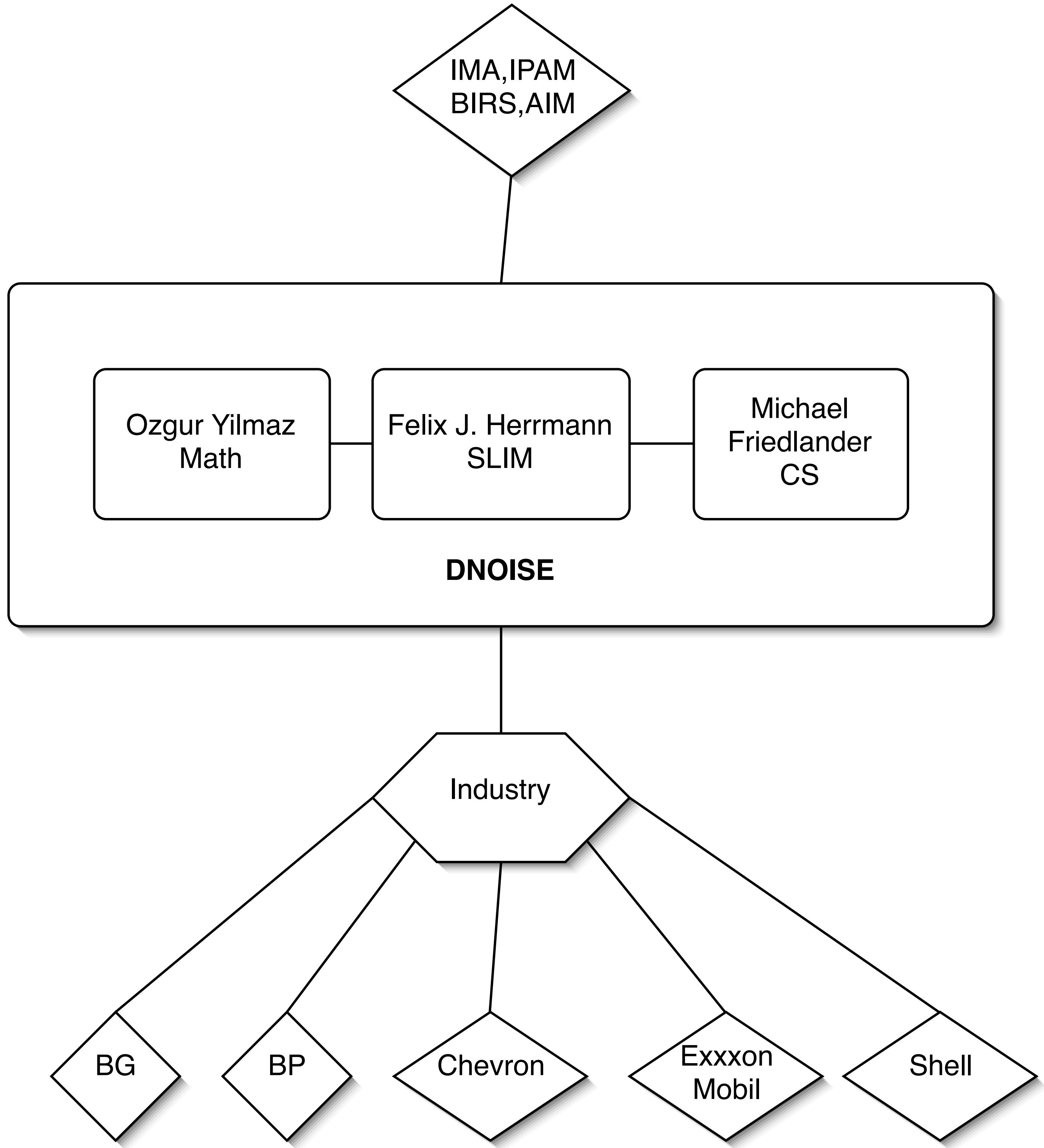
- extremely large (5-D, tera-peta bytes)
- incomplete and noisy
- operators expensive to apply

Physics & mathematics not fully understood

- linearization
- PDE constrained optimization is remote

Infusion of math has been a bumpy road

- inward looking
- after "the fact" proofs
- really understand problems that can not be tailored
- industry wants results not proofs ...



SLIM Seismic Laboratory for Imaging and Modeling

My research program

Successfully leverage recent developments in applied computational harmonic analysis and information theory

- multi-directional transforms such as curvelets
  - new construction that did not exist in seismology
- theory of compressive sampling
  - existed before BUT without proof & (fundamental) understanding
- theory of pseudodifferential operators
  - “invented” independently without proofs

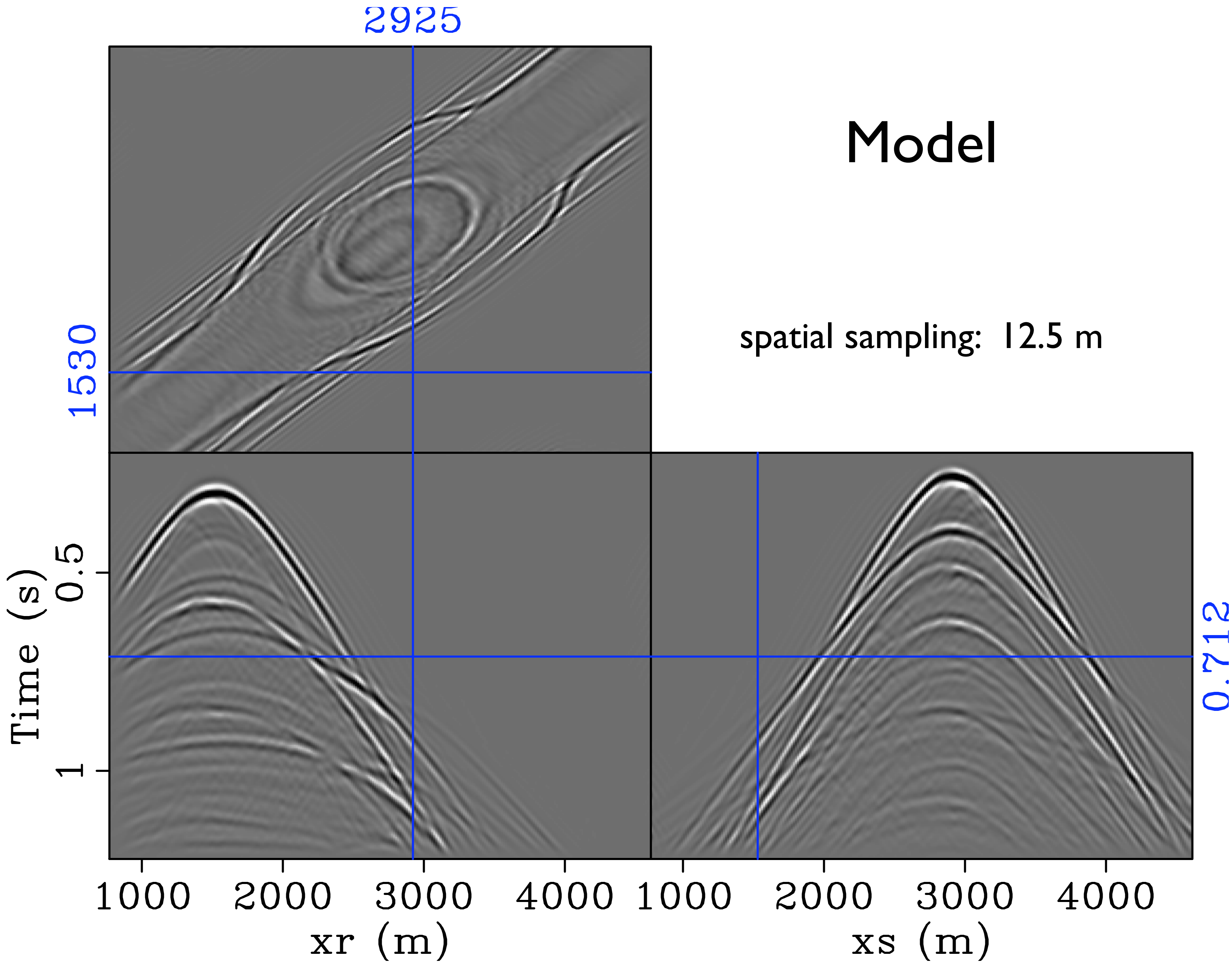
Combining these developments underlies the success of my research program

Seismic wavefield reconstruction

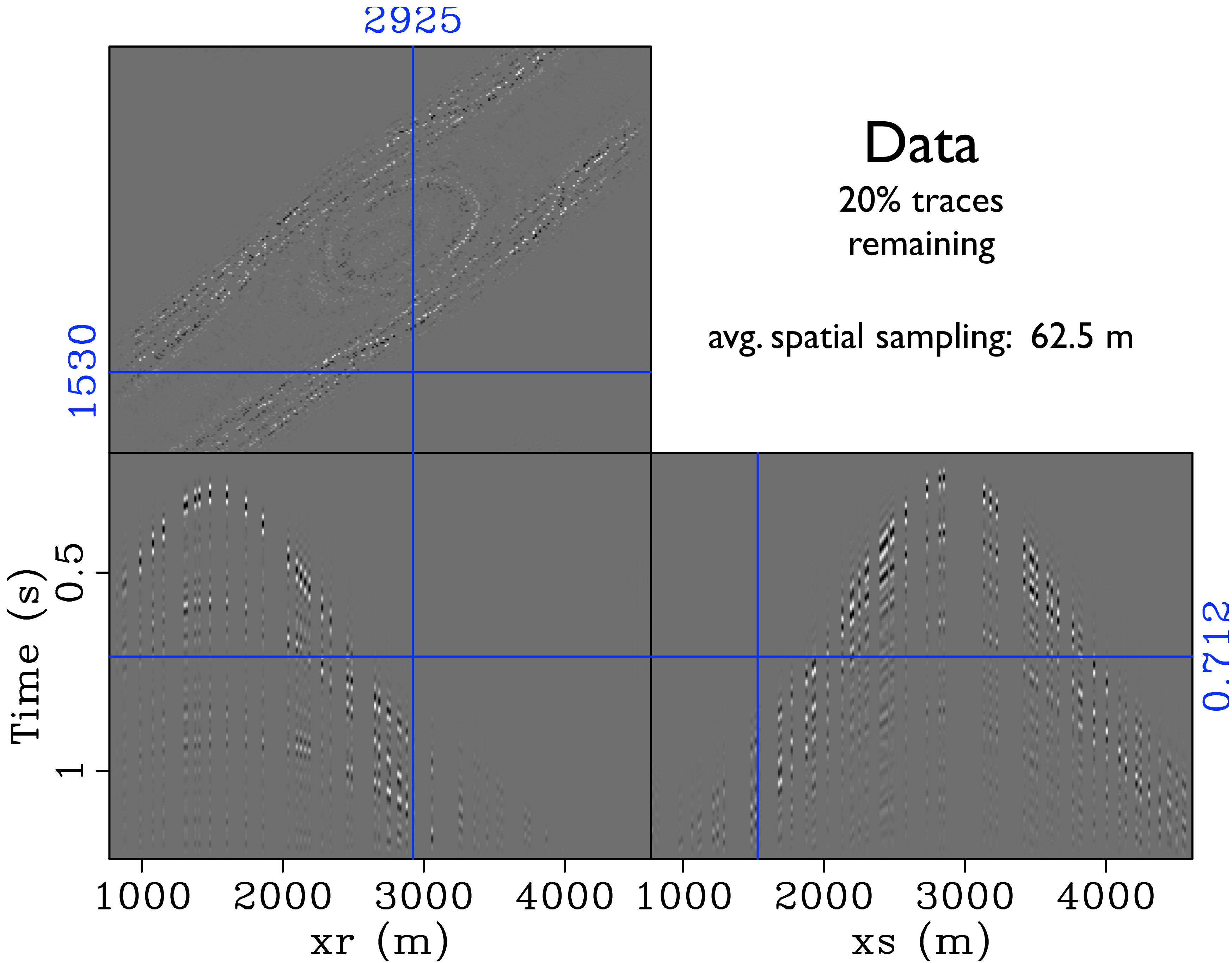
joint work with Gilles Hennenfent

"Curvelet-based seismic data processing: a multiscale and nonlinear approach" to appear in Geophysics, "Non-parametric seismic data recovery with curvelet frames" & "Simply denoise: wavefield reconstruction via jittered undersampling"

Model



spatial sampling: 12.5 m



Data

20% traces remaining

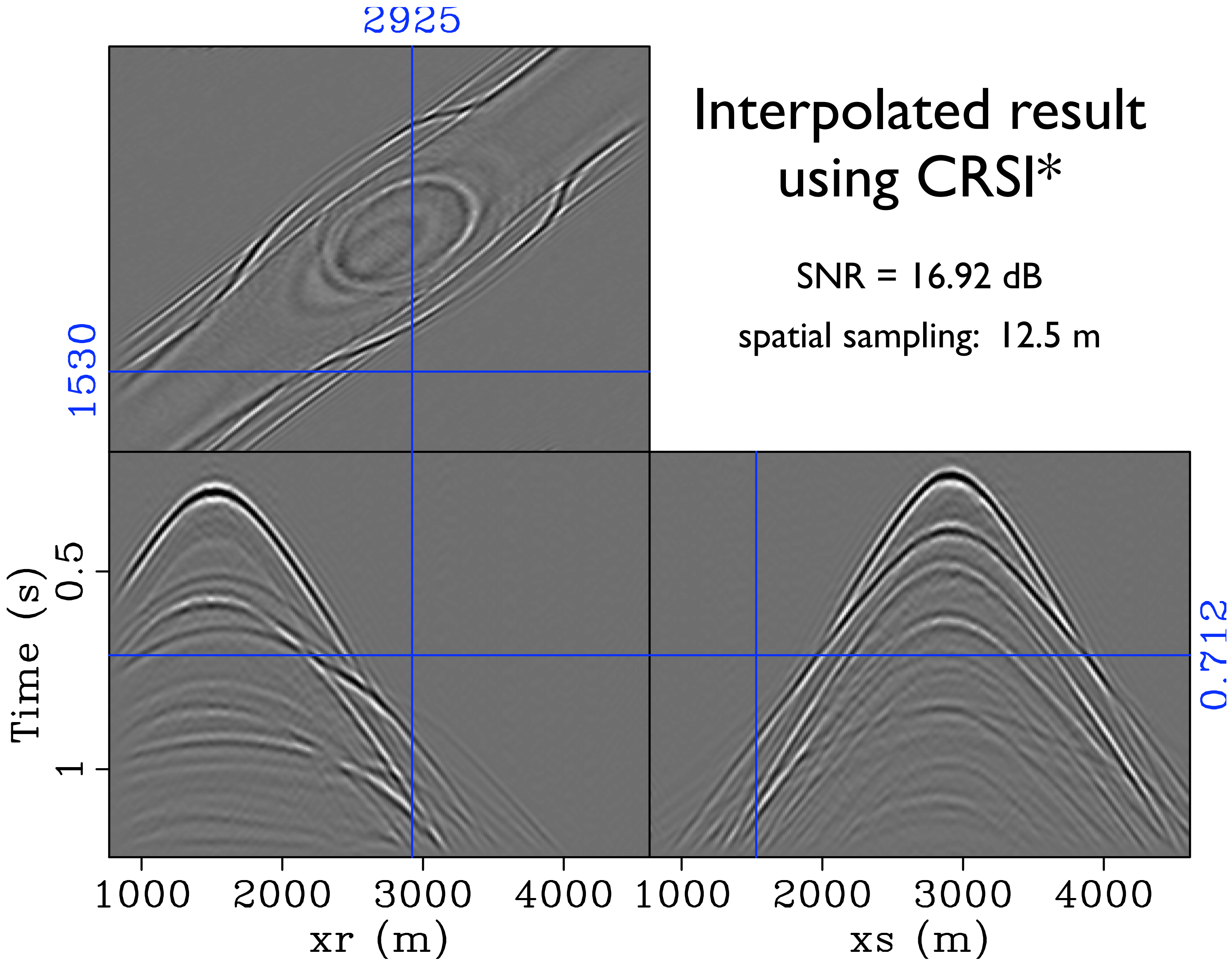
avg. spatial sampling: 62.5 m

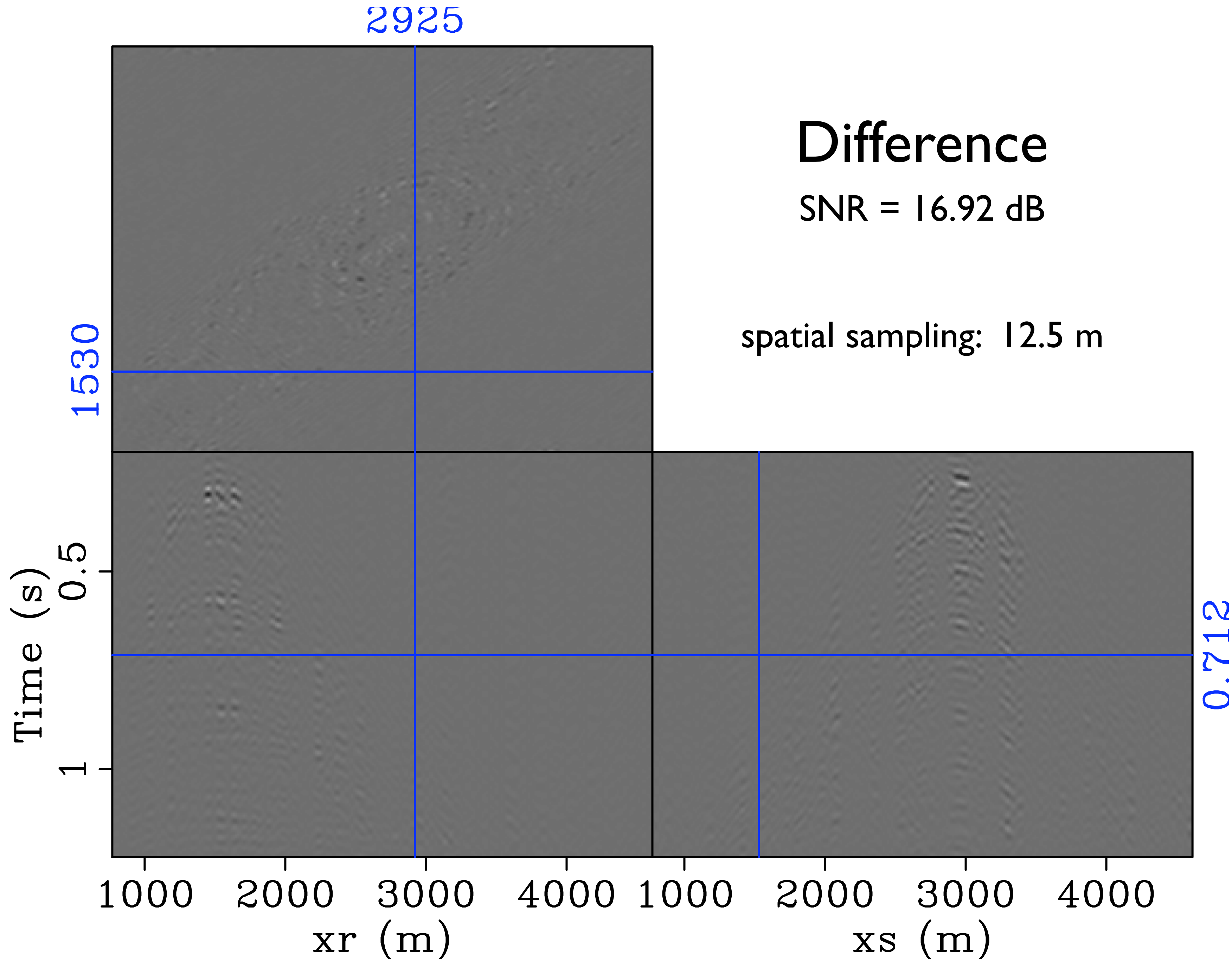
* **CRSI**: Curvelet Reconstruction with Sparsity-promoting Inversion

Interpolated result using CRSI*

SNR = 16.92 dB

spatial sampling: 12.5 m





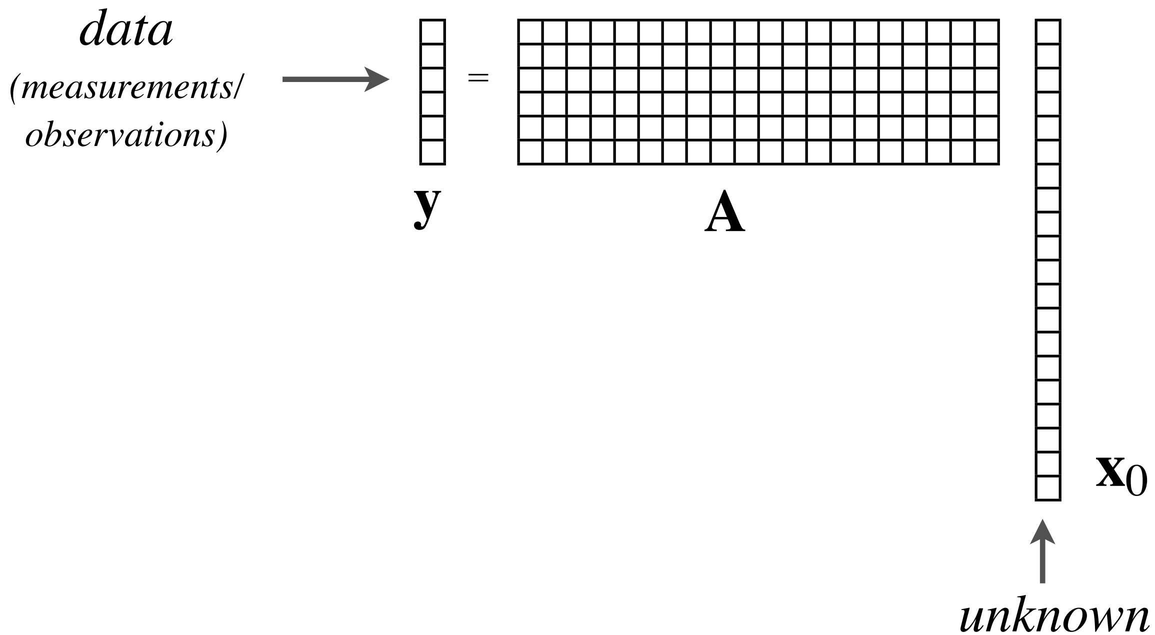
Difference

SNR = 16.92 dB

spatial sampling: 12.5 m

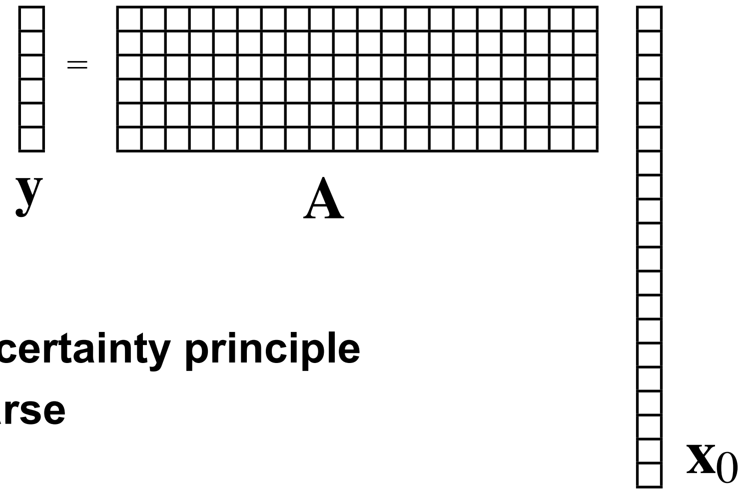
Problem statement

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 a type of **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, E., J. Romberg, and T. Tao, 2006b, Stable signal recovery from incomplete and inaccurate measurements: Communications On Pure and Applied Mathematics, **59**, 1207–1223.

Nonlinear wavefield sampling

- ***sparsifying transform***
  - typically **localized** in the time-space domain to handle the complexity of seismic data
  - preserves edges/wavefronts

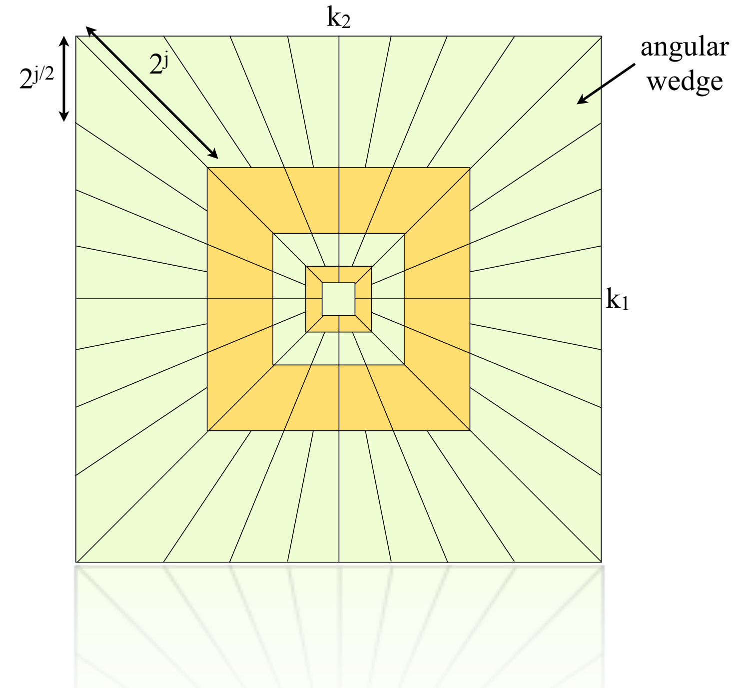
- *advantageous coarse sampling*
  - generates incoherent random undersampling “noise” in the sparsifying domain
  - does not create large gaps
    - because of the limited spatiotemporal extend of transform elements used for the reconstruction

- *sparsity-promoting solver*
  - requires few matrix-vector multiplications
  - scales to number of unknowns exceeding 2^{30} (“small”)

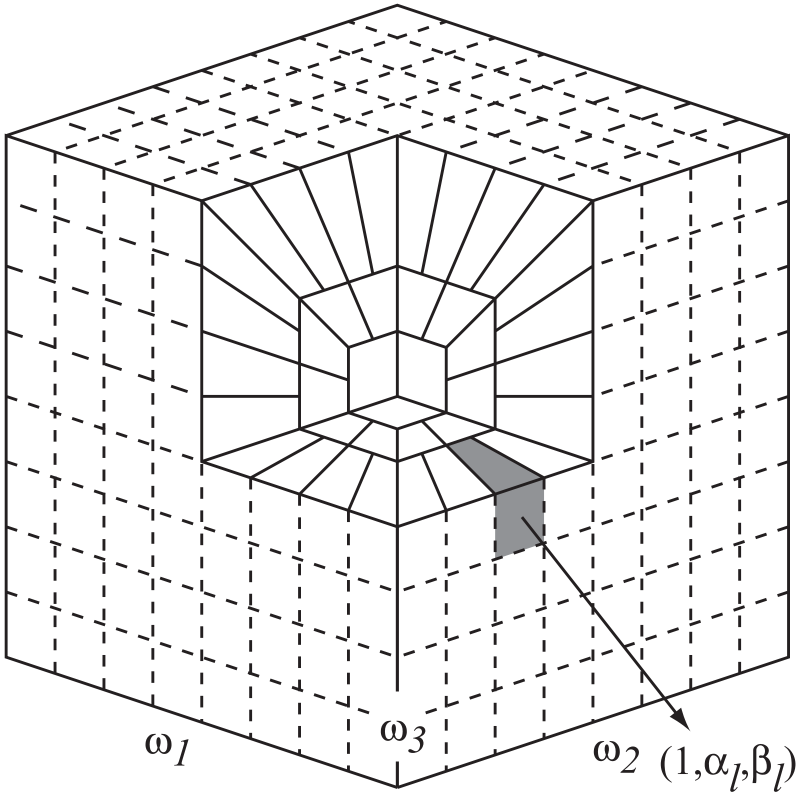
Representations for seismic data

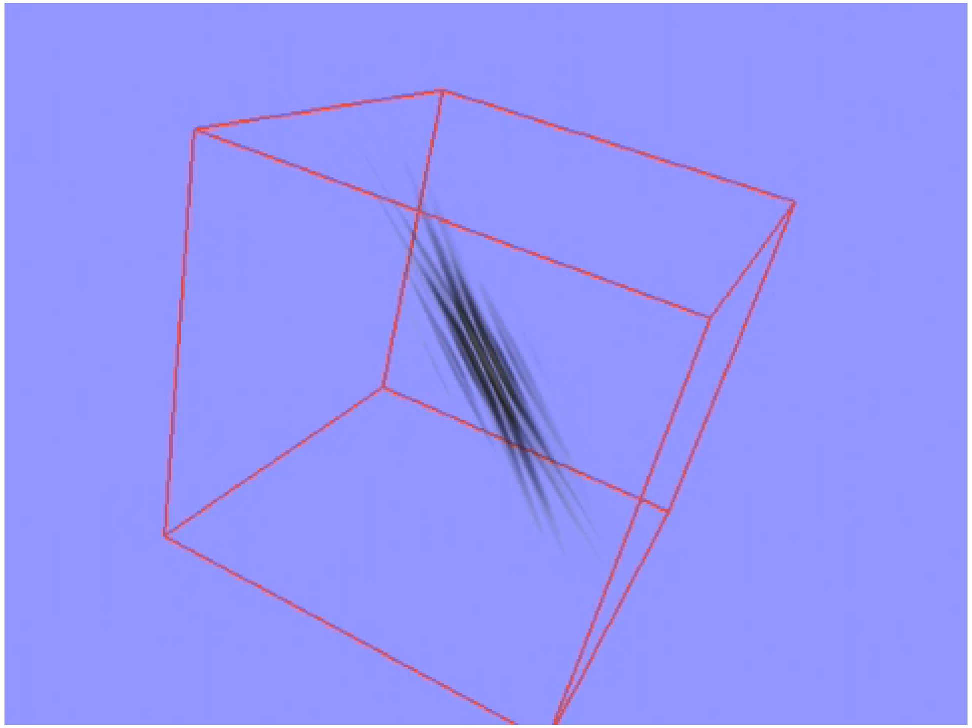
Transform	Underlying assumption
FK	plane waves
linear/parabolic Radon transform	linear/parabolic events
wavelet transform	point-like events (1D singularities)
curvelet transform	**curve-like events (2D singularities)**

- curvelet transform
  - **multi-scale**: tiling of the FK domain into dyadic coronae
  - **multi-directional**: coronae sub-partitioned into angular wedges, # of angle doubles every other scale
  - **anisotropic**: parabolic scaling principle
  - **local**



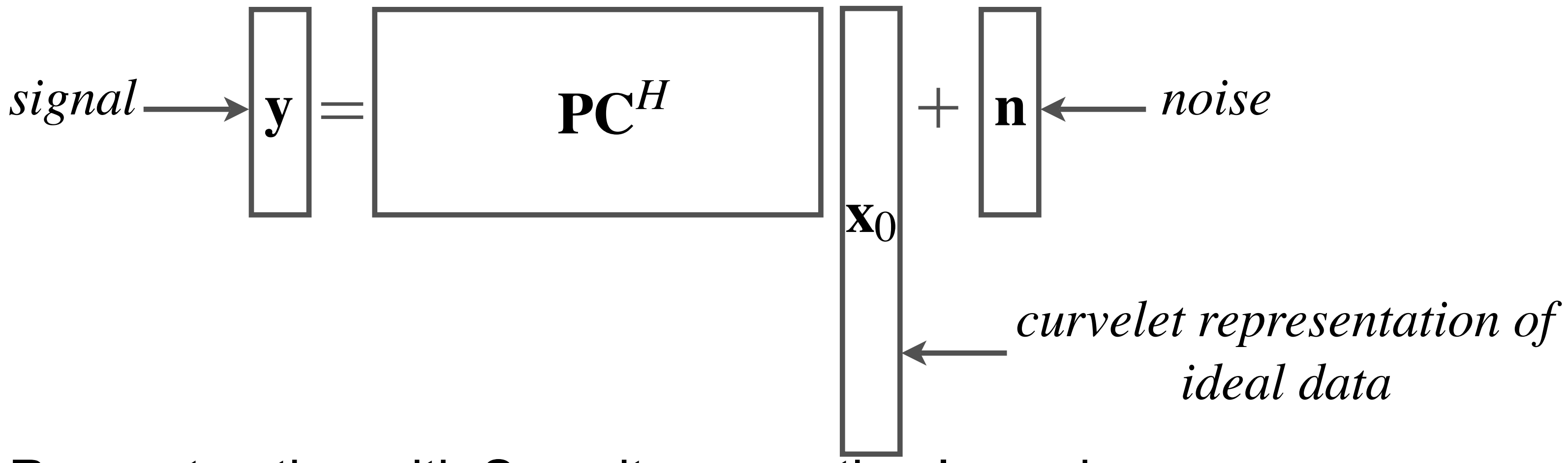
3D curvelets





CRSIn

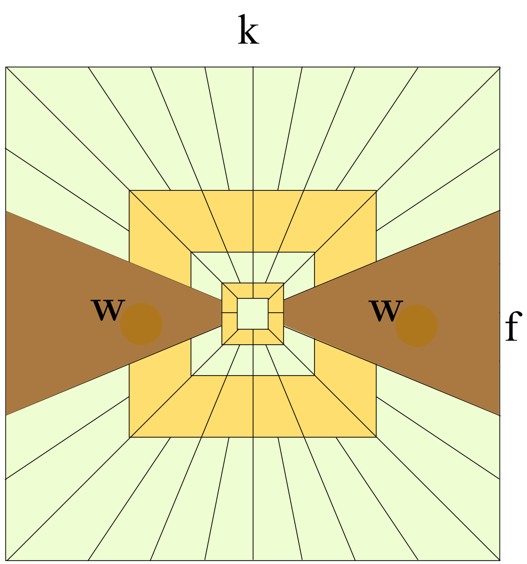
- reformulation of the problem



- Curvelet Reconstruction with Sparsity-promoting Inversion
  - look for the **sparsest/most compressible**, **physical** solution

KEY POINT OF THE RECOVERY

$$
(P_1)\begin{cases} \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \overbrace{\|\mathbf{W}\mathbf{x}\|_1}^{\text{sparsity constraint}} \quad \text{s.t.} \quad \overbrace{\|\mathbf{y} - \mathbf{P}\mathbf{C}^H\mathbf{x}\|_2}^{\text{data misfit}} \leq \varepsilon \\ \tilde{\mathbf{f}} = \mathbf{C}^H \tilde{\mathbf{x}} \end{cases}
$$



Nonlinear wavefield sampling

- *sparsifying transform*
  - typically **localized** in the time-space domain to handle the complexity of seismic data
  - preserves edges/wavefronts

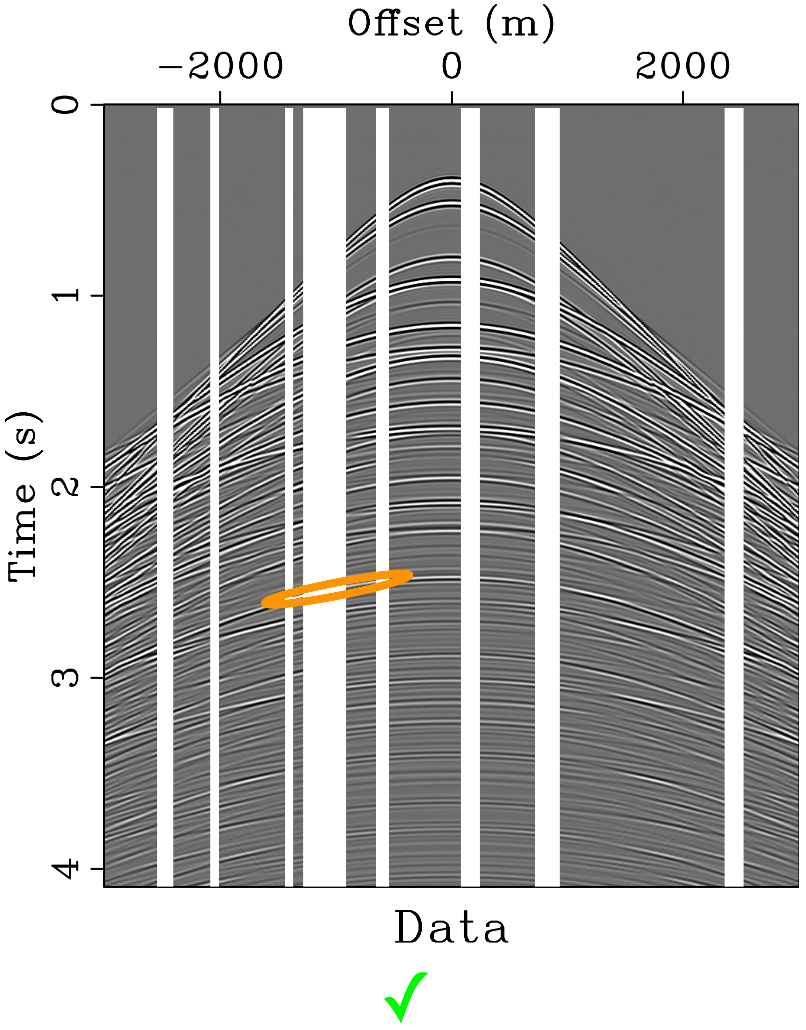
- ***advantageous coarse sampling***
  - generates incoherent random undersampling “noise” in the sparsifying domain
  - does not create large gaps
    - because of the limited spatiotemporal extend of transform elements used for the reconstruction

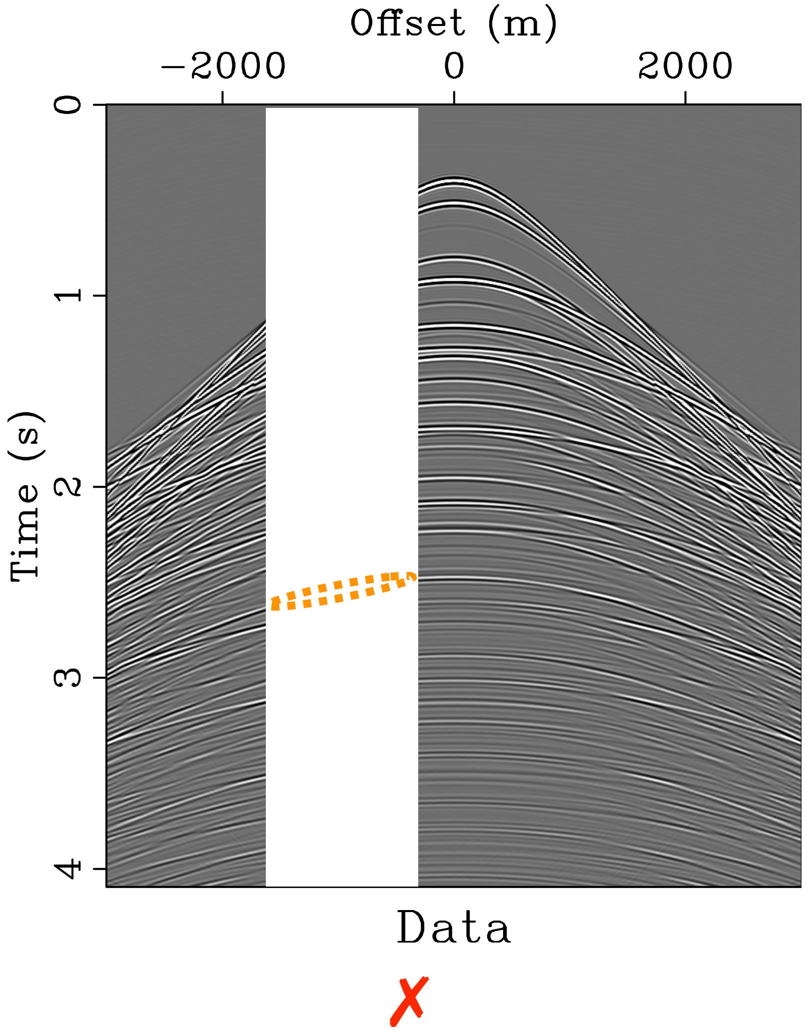
- *sparsity-promoting solver*
  - requires few matrix-vector multiplications
  - scales to number of unknowns exceeding 2^{30} (“small”)

Lustig et. al 2007

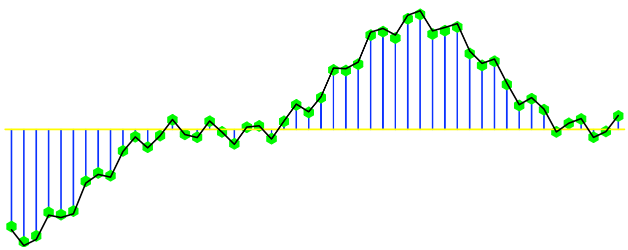
Localized transform elements & gap size

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

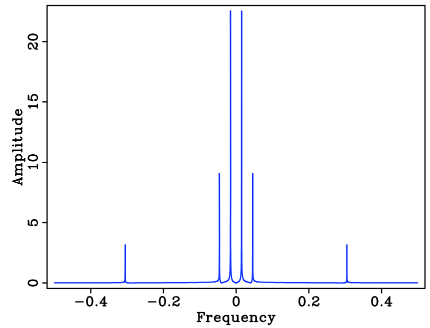




Sampling

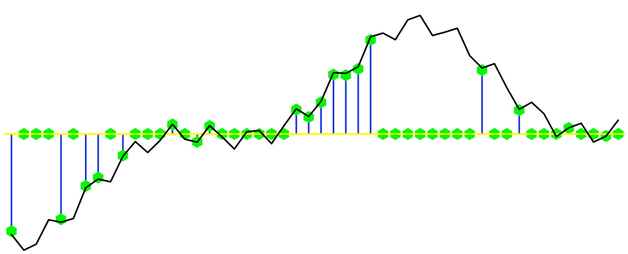


Fourier transform

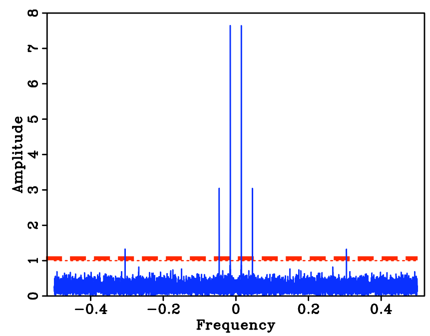


few significant coefficients

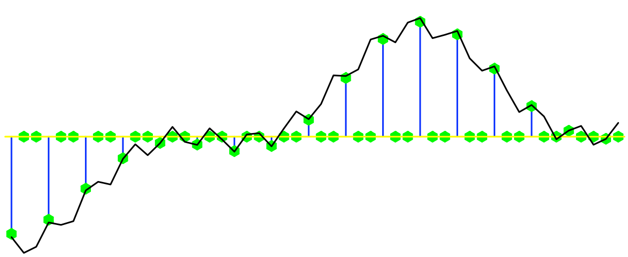
3-fold under-sampling



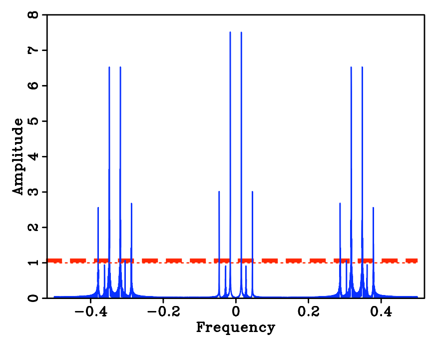
Fourier transform



✓ significant coefficients detected



Fourier transform



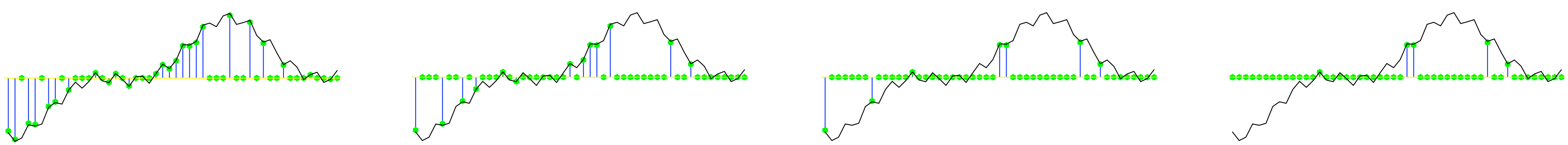
✗ ambiguity

Undersampling "noise"

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

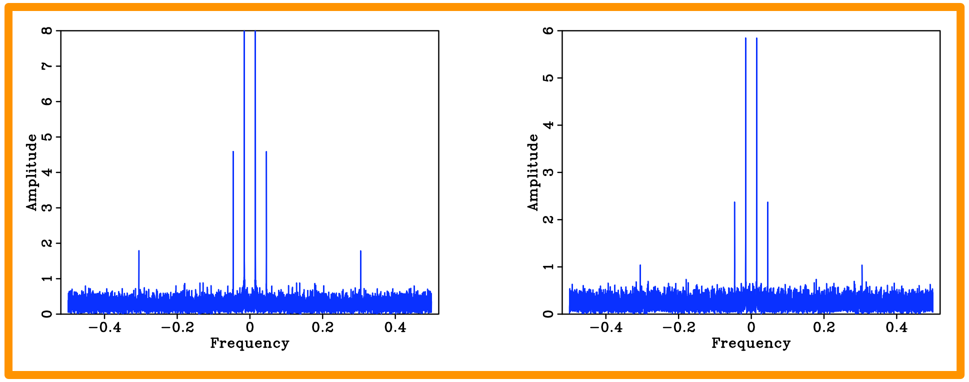
D.L. Donoho et.al. '06

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

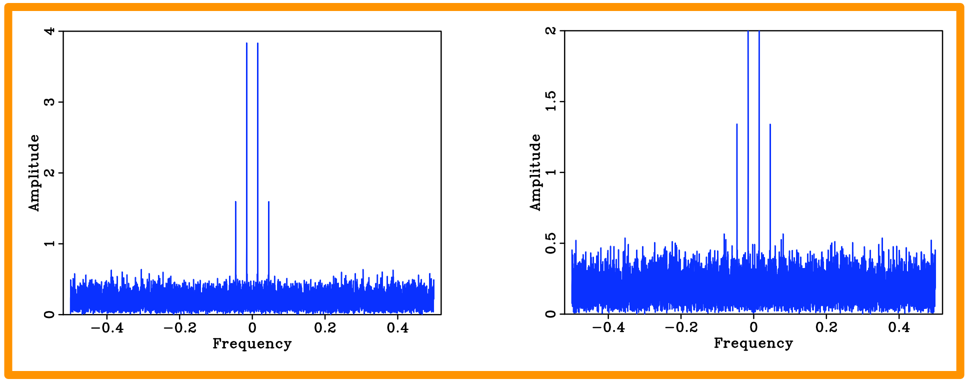


less acquired data

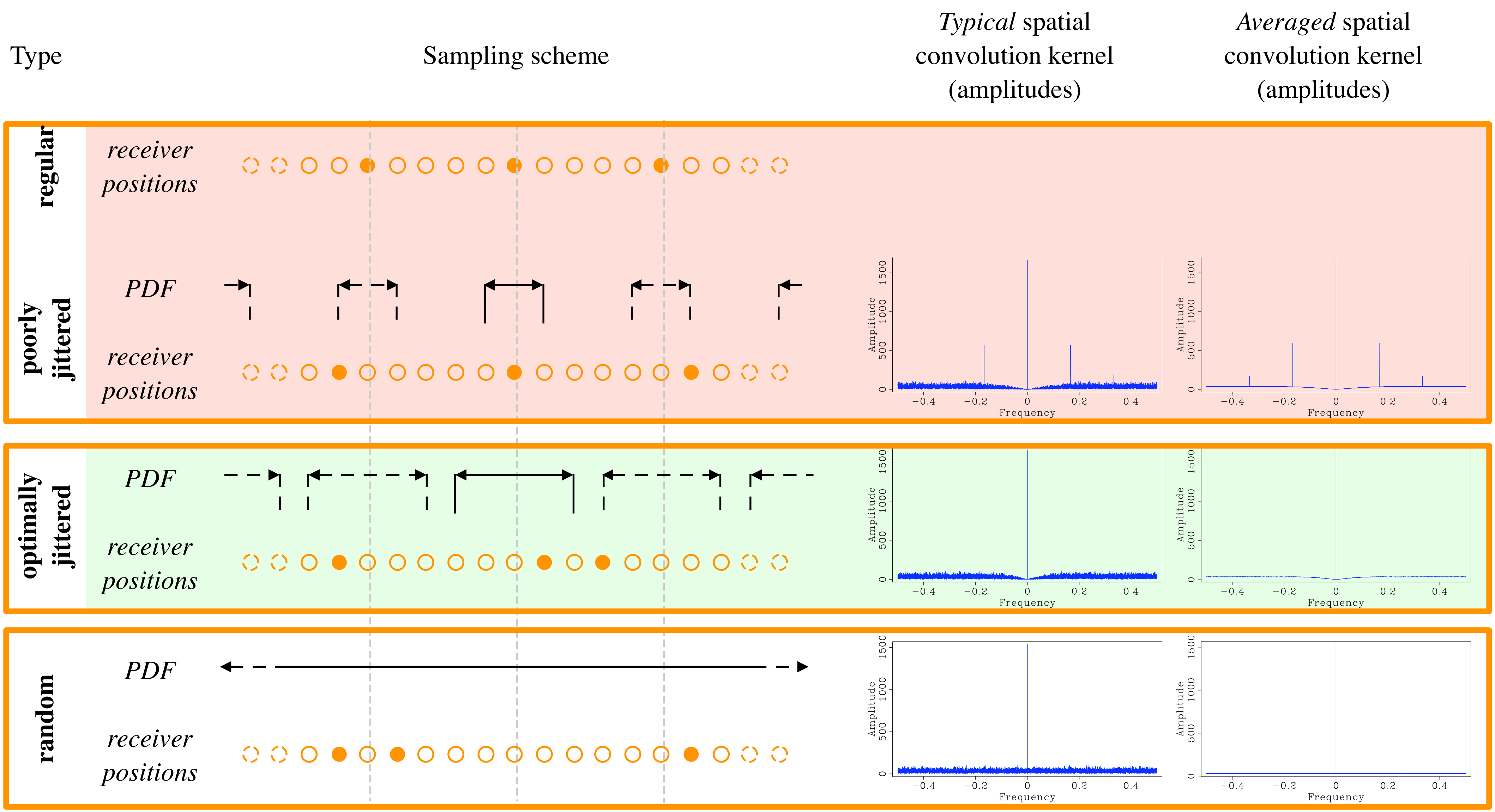
3 detectable Fourier modes



2 detectable Fourier modes

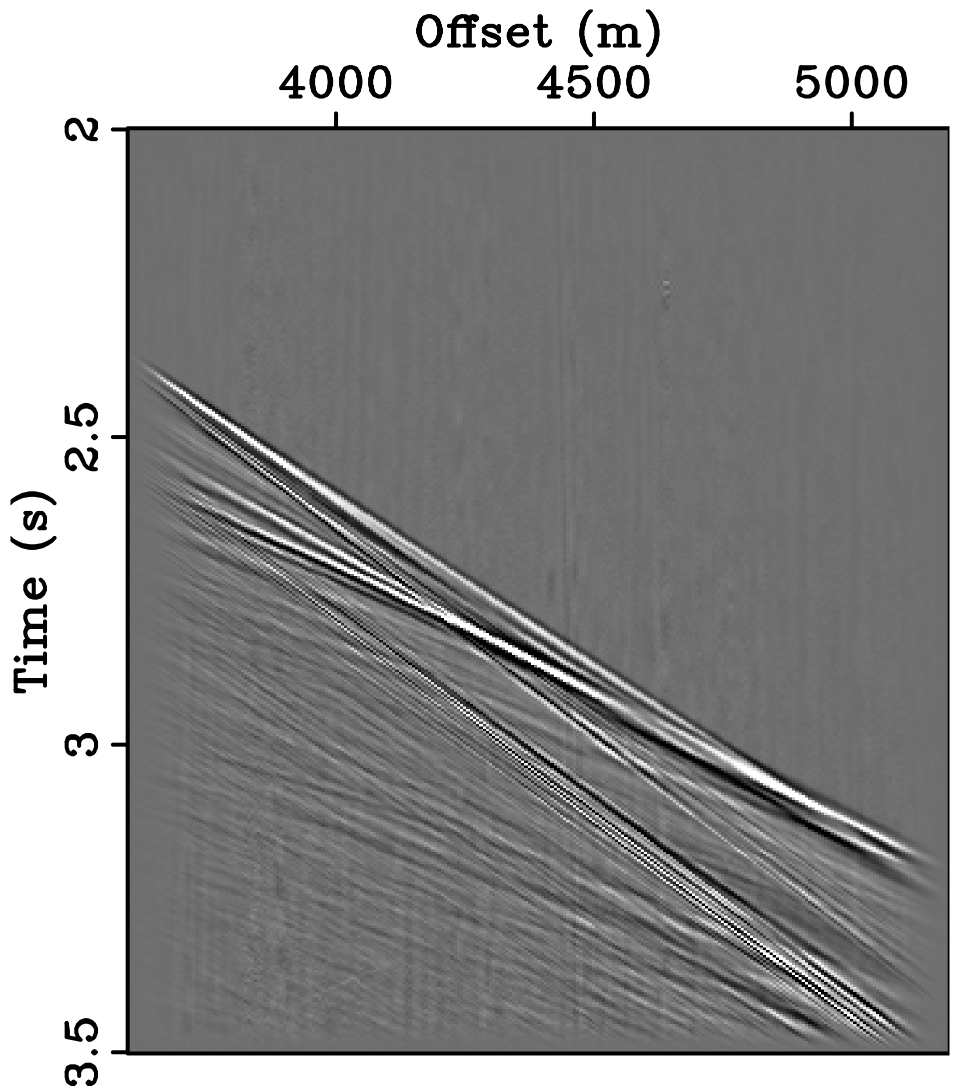


Discrete random jittered undersampling

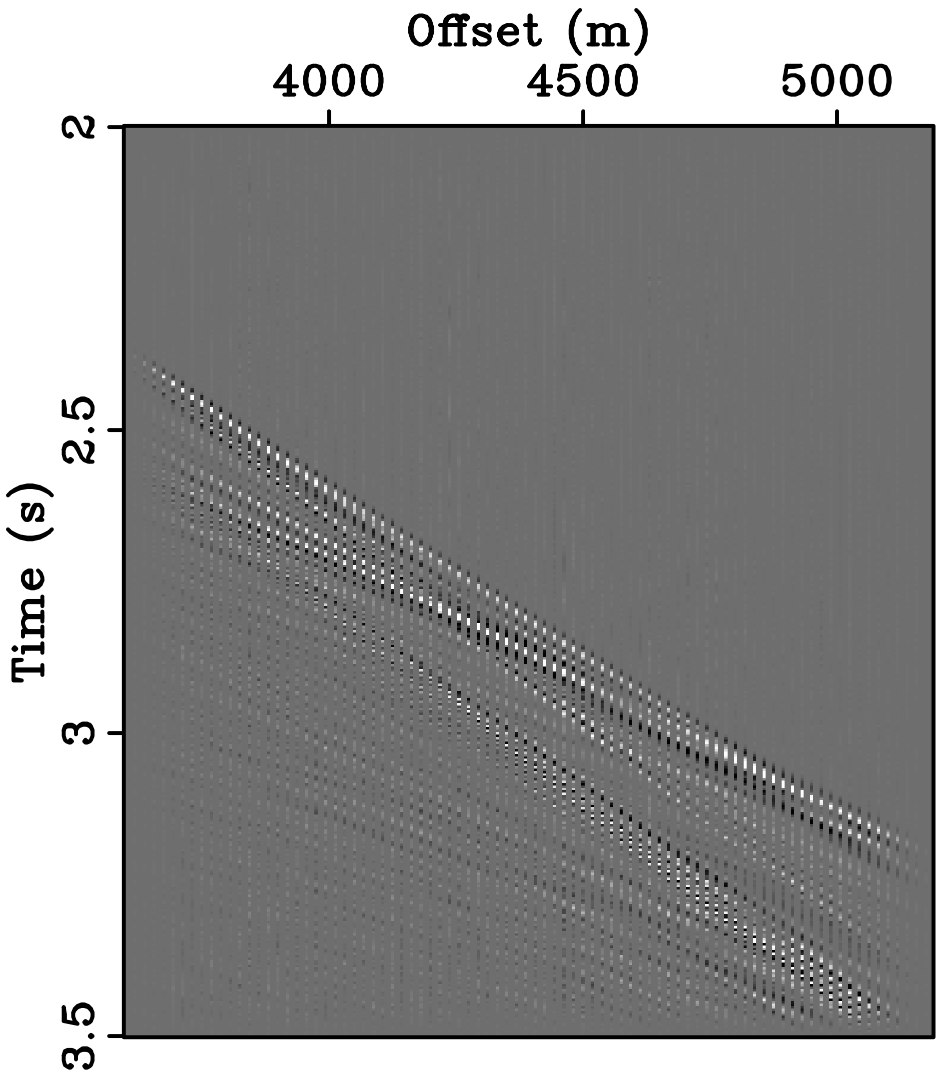


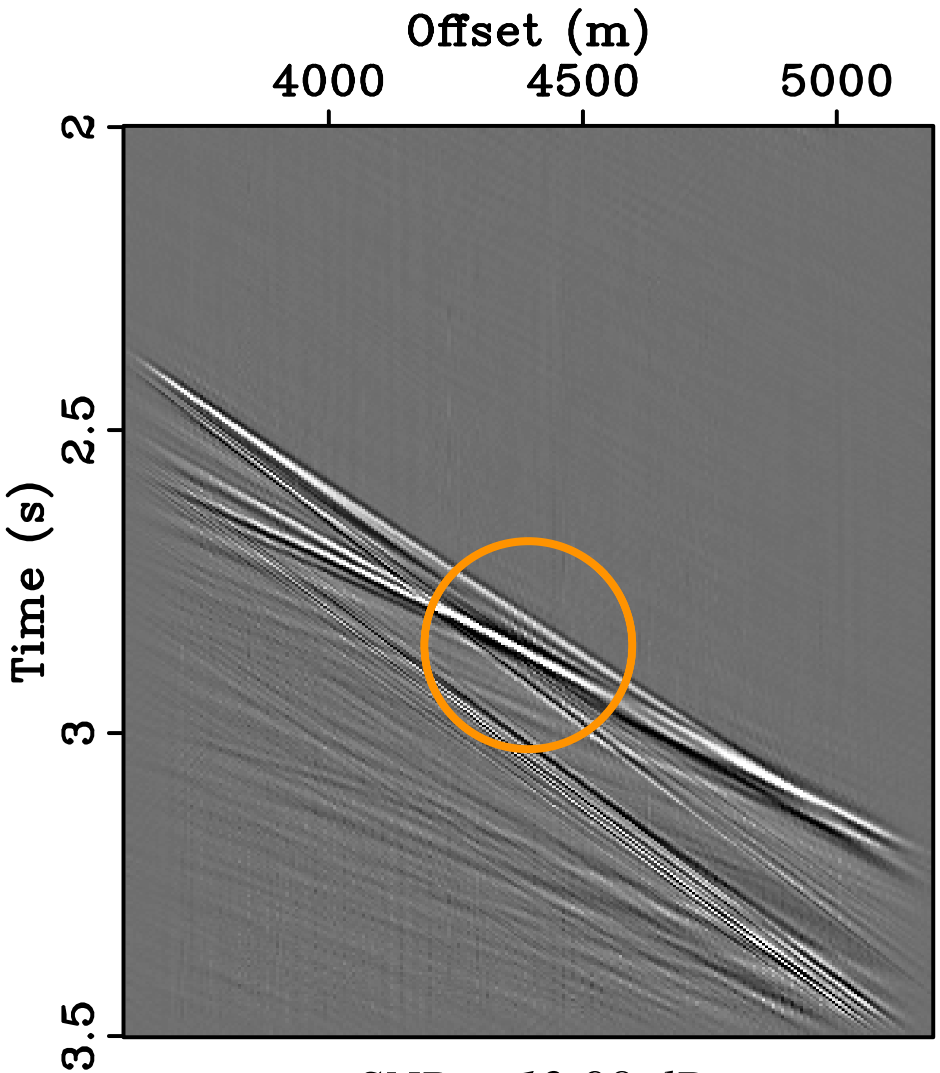
[Hennenfent and Herrmann '07]

Model



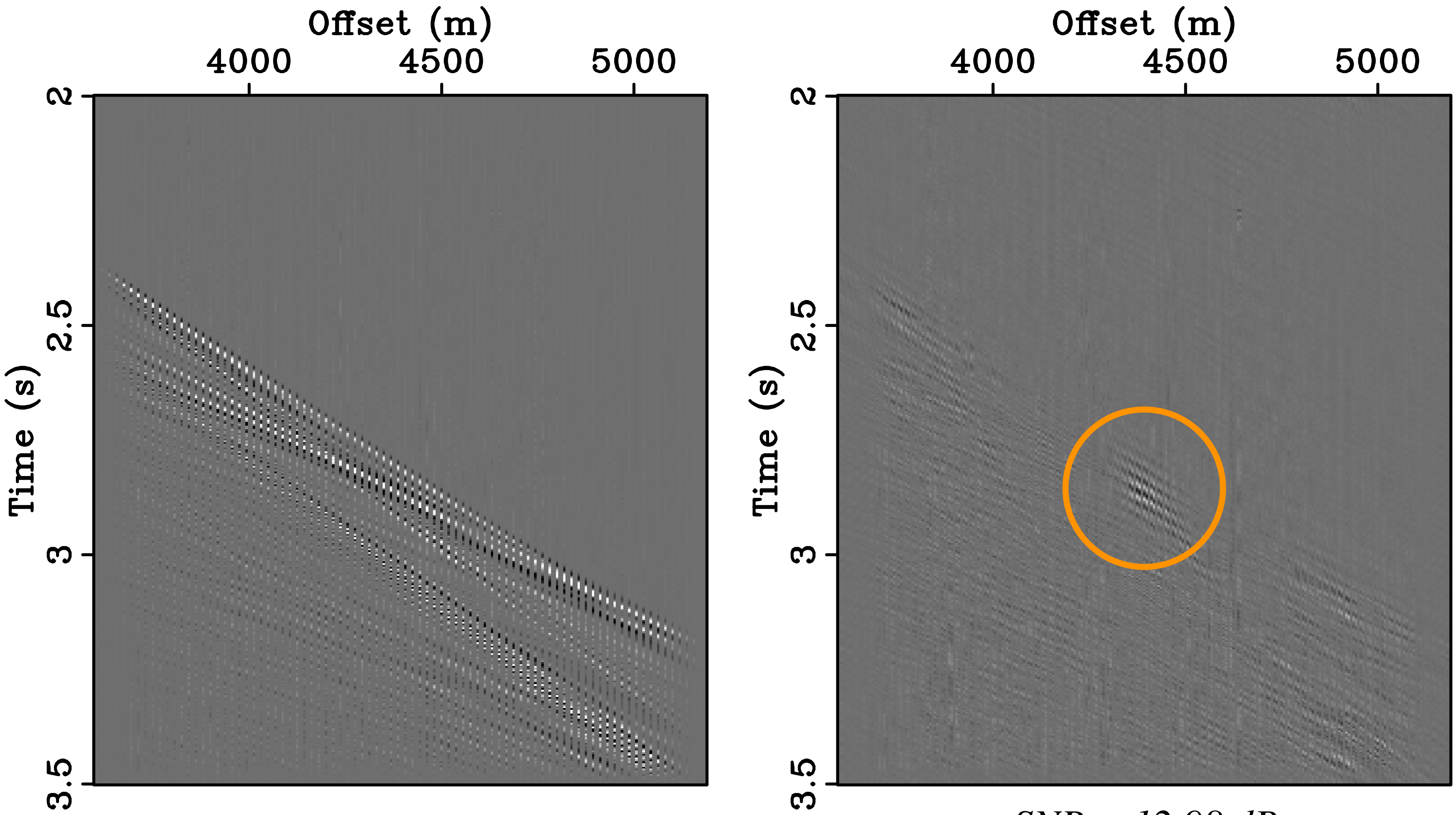
Regular 3-fold undersampling





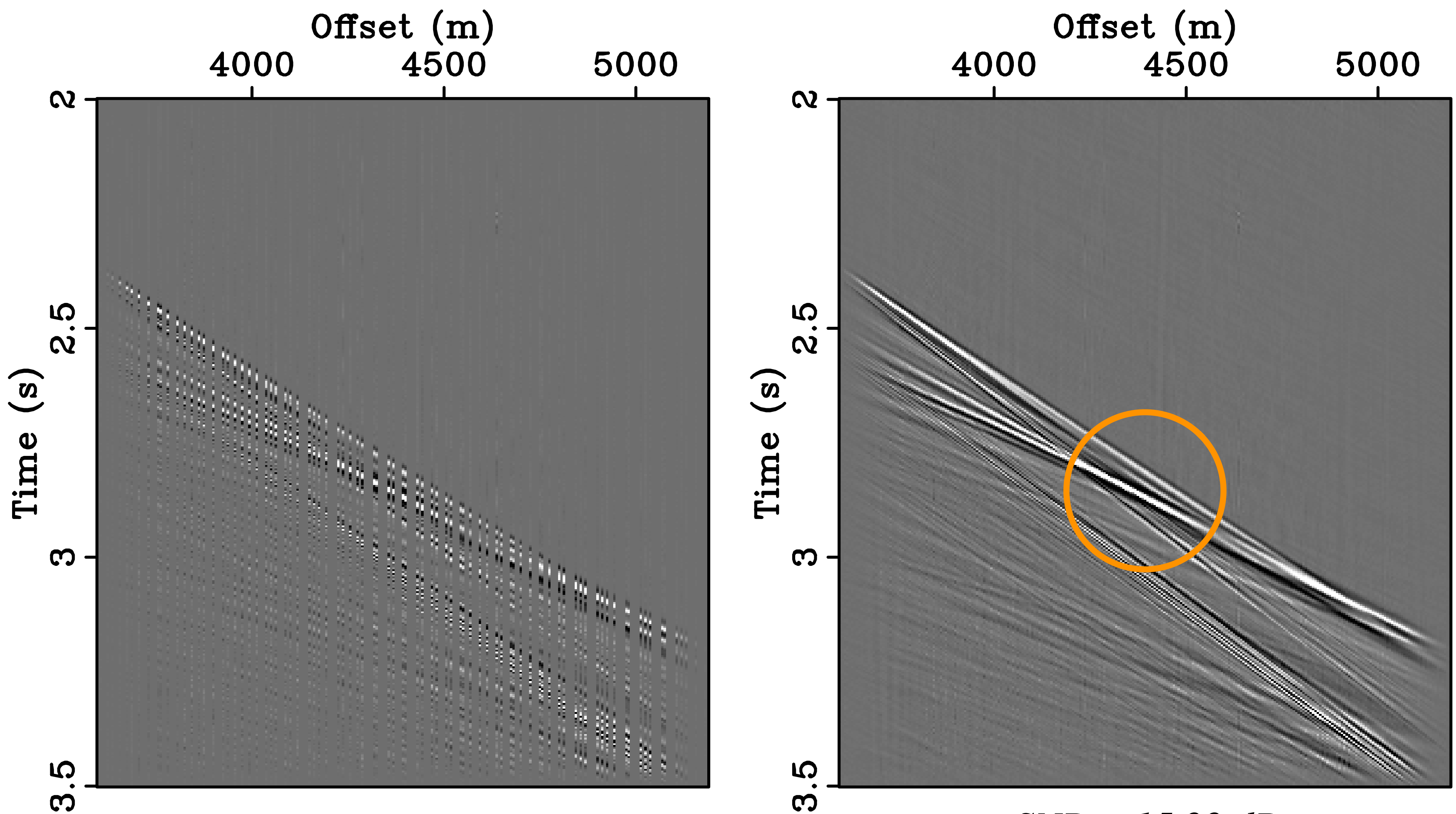
SNR = 12.98 dB

Regular 3-fold undersampling



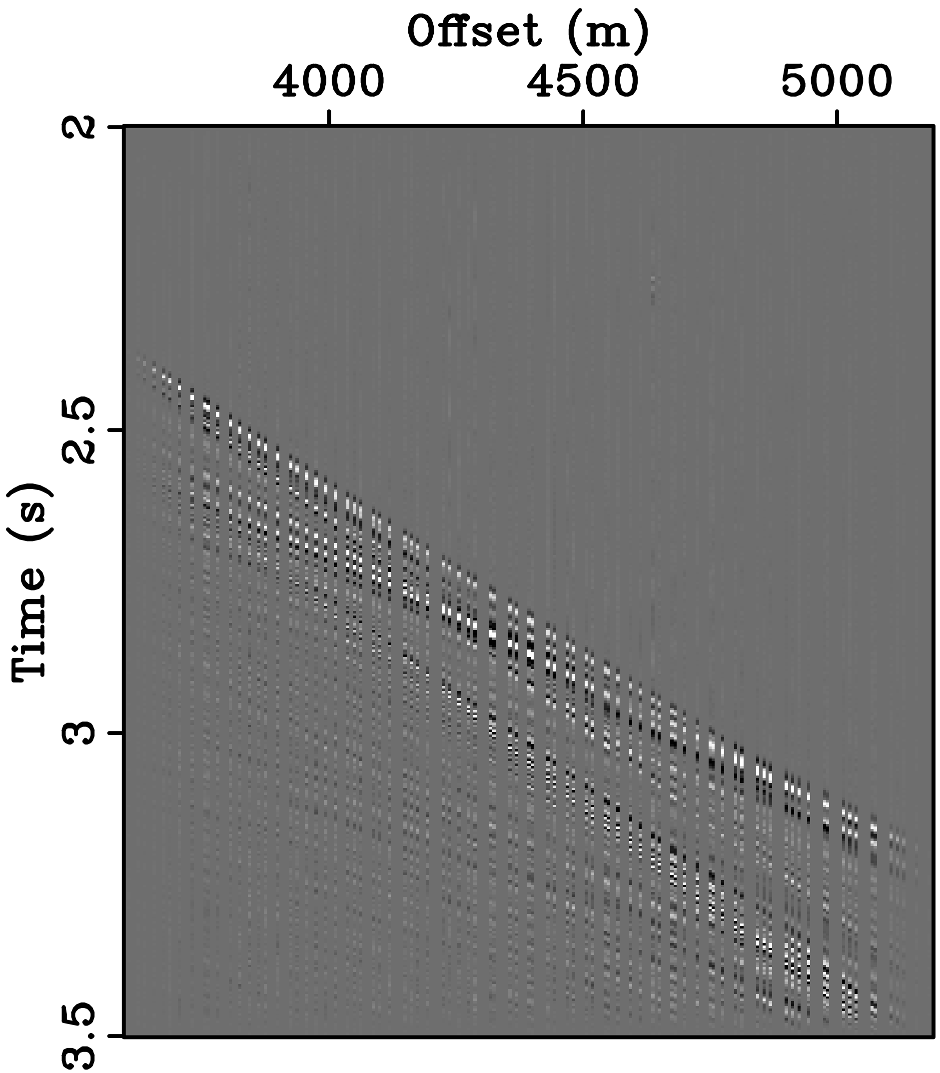
SNR = 12.98 dB

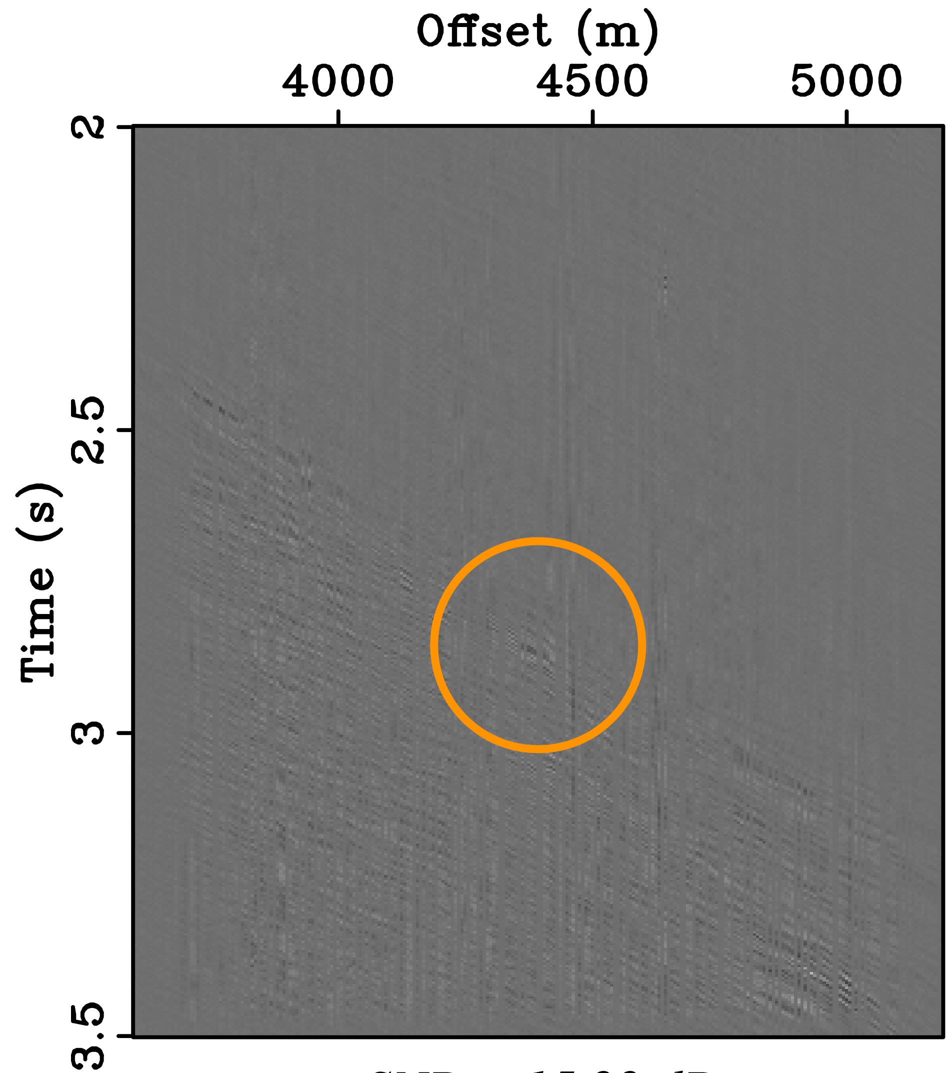
Optimally-jittered 3-fold undersampling



SNR = 15.22 dB

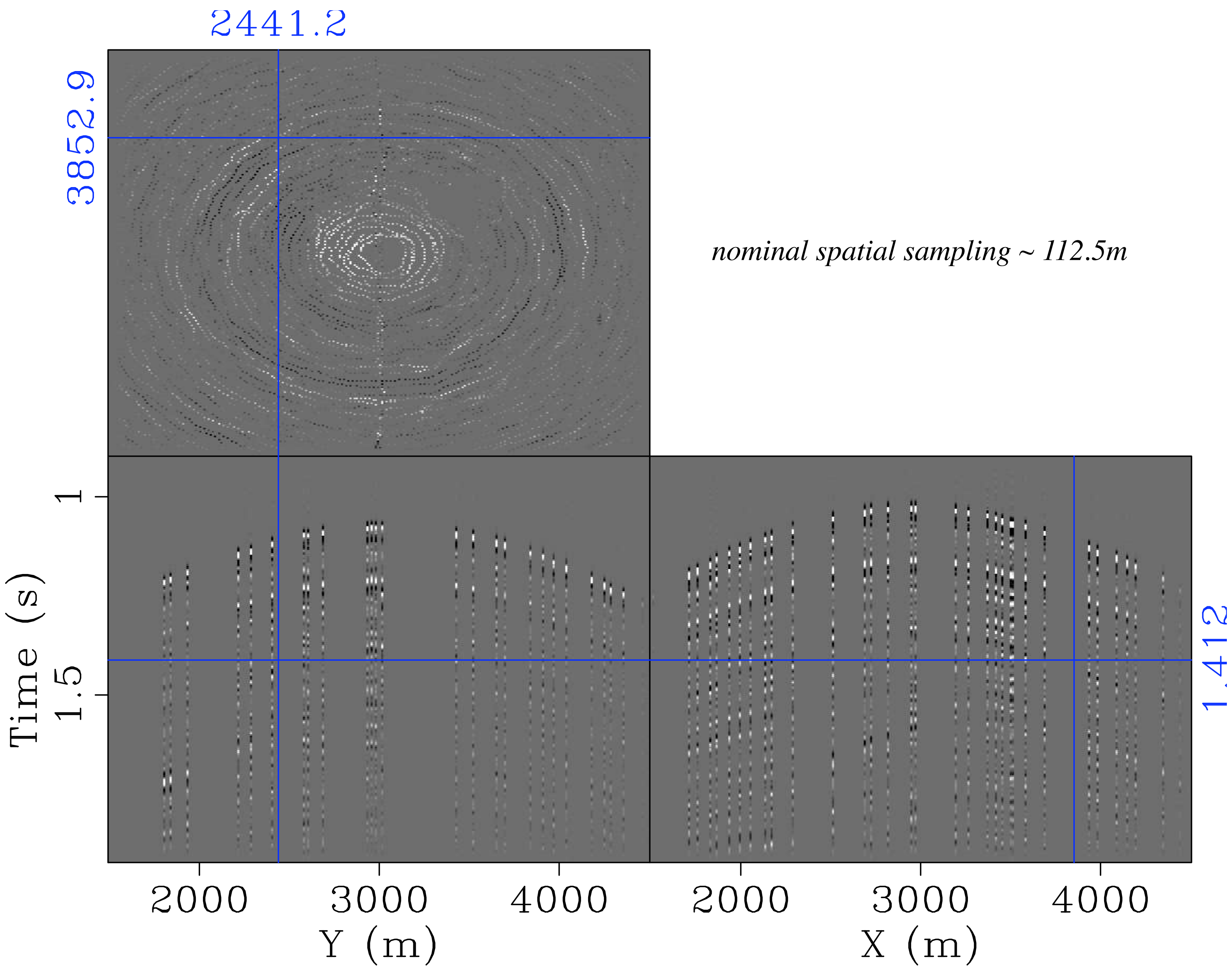
Optimally-jittered 3-fold undersampling





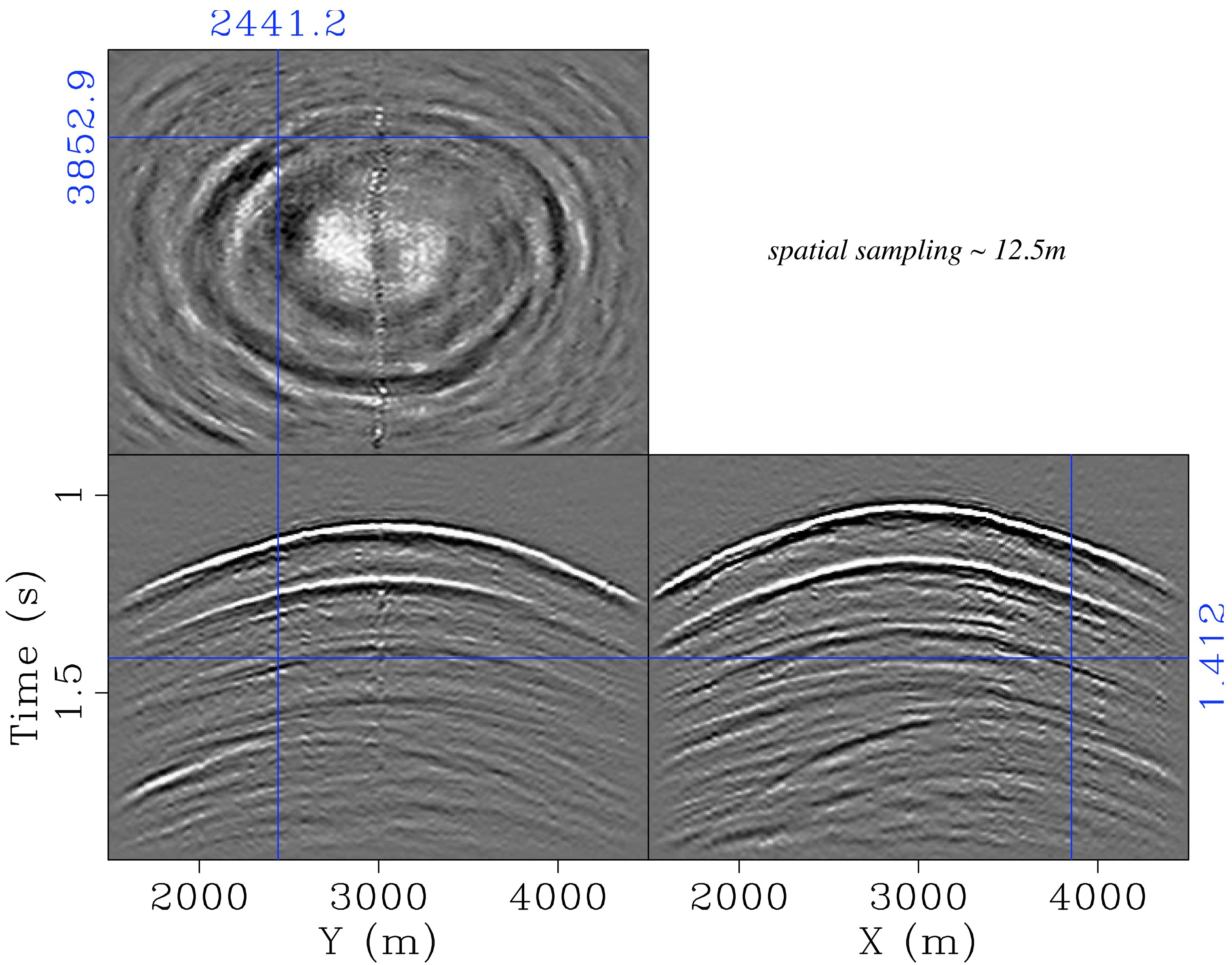
SNR = 15.22 dB

Data



nominal spatial sampling ~ 112.5m

CRSI



spatial sampling ~ 12.5m

Observations

- *sparsity* is a powerful property that offers striking benefits for signal reconstruction BUT it is not enough

- in the sparsifying domain, *interpolation is a denoising problem*
  - regular undersampling:
    harmful coherent undersampling "noise", i.e., aliases
  - **random & jittered undersamplings:**
    harmless incoherent random undersampling "noise"

- nonlinear wavefield sampling
  - sparsifying transform: **curvelet transform**
  - coarse sampling scheme: **jittered undersampling**
  - sparsity-promoting solver: **iterative soft thresholding with cooling**

- open problem: optimal (non-random) sampling schemes, large-scale solvers & hard CS results for frames

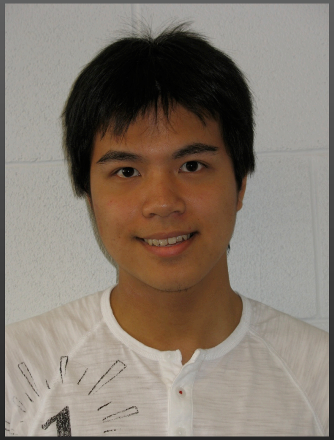
observations continued

- CS ideas already existed in exploration seismology (Sacchi '98)

- New insights give solid proofs that
  - (hopefully) help convince management
  - engineers will do their implementations => innovation
- Results for seismic wavefield reconstruction
  - very encouraging
  - **industry calls for commercialization/industrialization**
  - looking into a startup
- Real-life implementation requires substantial investment
  - understanding the real problem & QC
  - infrastrcuture
  - solution that scales
- Real-life implementations require
  - parallelization of algorithms
  - massive IO
  - run on 10.000 CPU plus clusters ...

Compressed wavefield extrapolation

joint work with Tim Lin



“Compressed wavefield extrapolation” in Geophysics

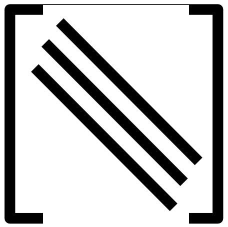
Problem statement

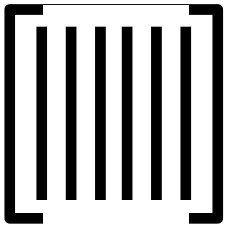
- Goal: employ the 1-Way wavefield extrapolation based on factorization of the Helmholtz operator

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

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

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

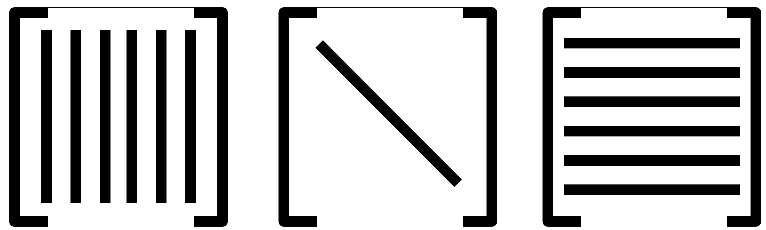
$\mathbf{H}_2 =$ 

$\mathbf{H}_1 =$ 

Modal domain

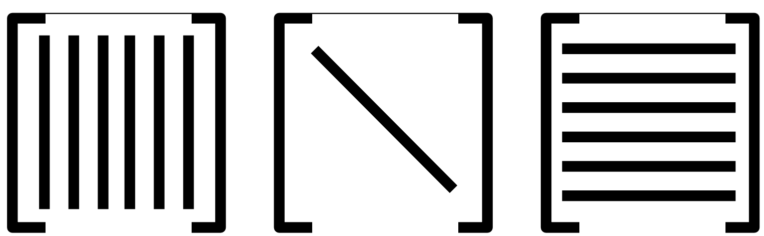
- In this case W is computed by eigenvalue decomposition

$$\mathbf{H}_2 = \mathbf{L}\boldsymbol{\Lambda}\mathbf{L}^{\mathrm{T}} =$$



$$\mathbf{L} \quad \boldsymbol{\Lambda} \quad \mathbf{L}^{\mathrm{T}}$$

$$\mathbf{W}^{\pm} =$$



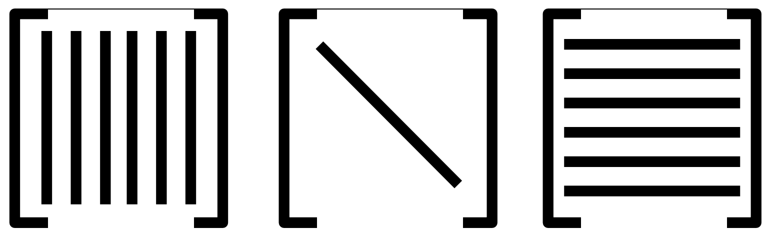
$$\mathbf{L} \quad e^{-j\sqrt{\boldsymbol{\Lambda}}\Delta x_3} \quad \mathbf{L}^{\mathrm{T}}$$

- requires, per frequency:
  - 1 eigenvalue problem ($O(n^4)$)
  - 2 full matrix-vector for eigenspace transform ($O(n^2)$)

Our approach

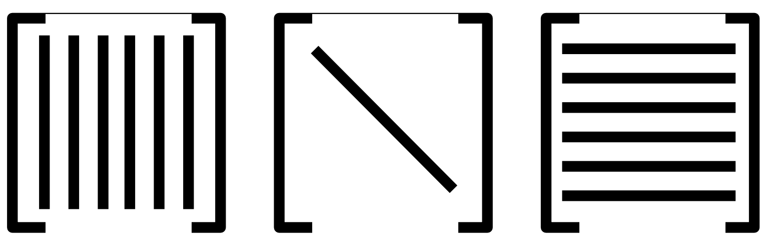
- Computation requires similar approach to $\mathbf{W}^{\pm}$

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



$\mathbf{L}$ $\quad \mathbf{\Lambda} \quad$ $\mathbf{L}^{\mathrm{T}}$

$$\mathbf{S} =$$

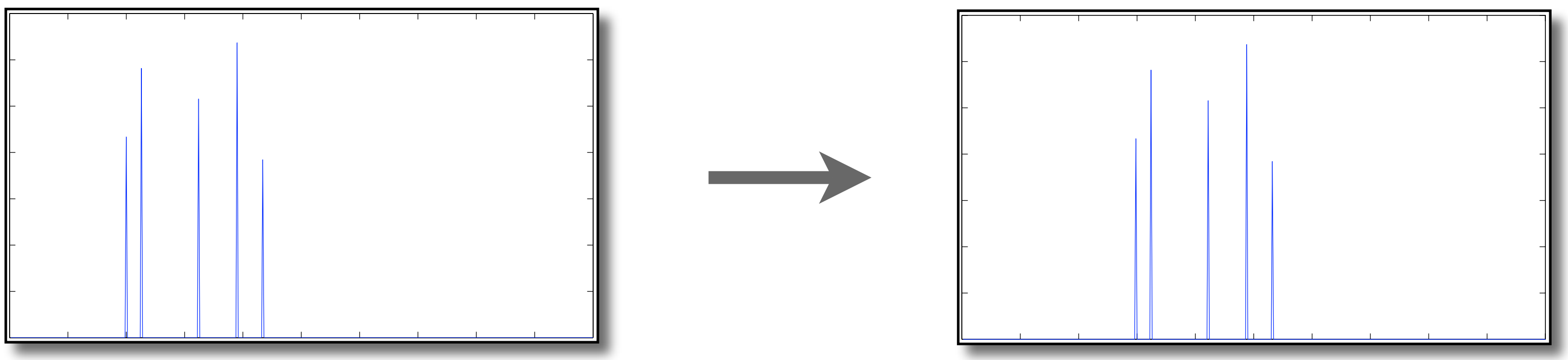


$\mathbf{L}$ $\quad e^{-j\frac{\Delta x}{2\pi}\mathbf{\Lambda}} \quad$ $\mathbf{L}^{\mathrm{T}}$

- However, for $\mathbf{D}$, $\mathbf{L} = \mathrm{DFT}$, so computation trivial with FFT

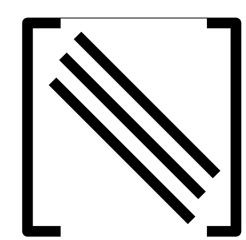
Our approach

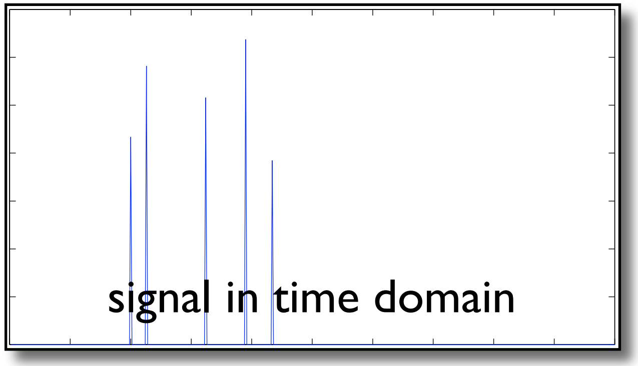
- Consider a related, but simpler problem: shifting (or translating) signal

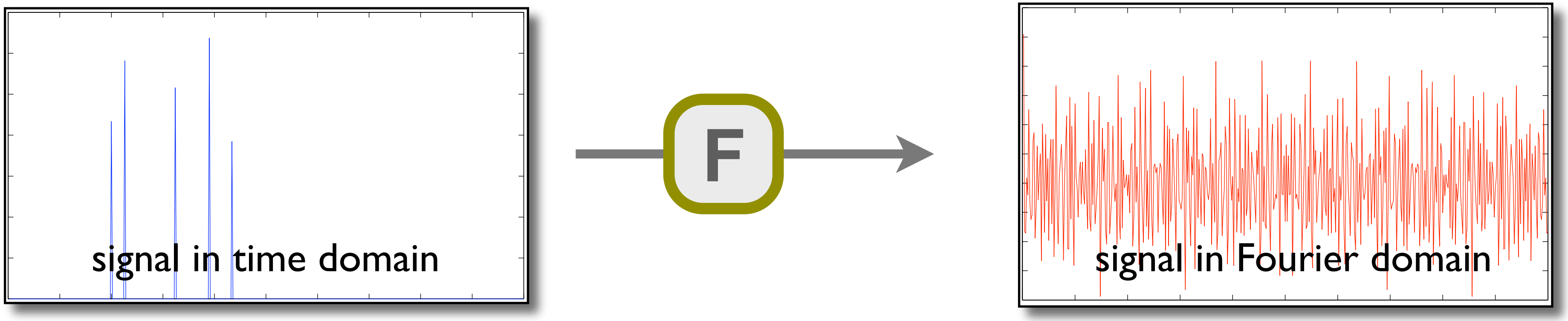


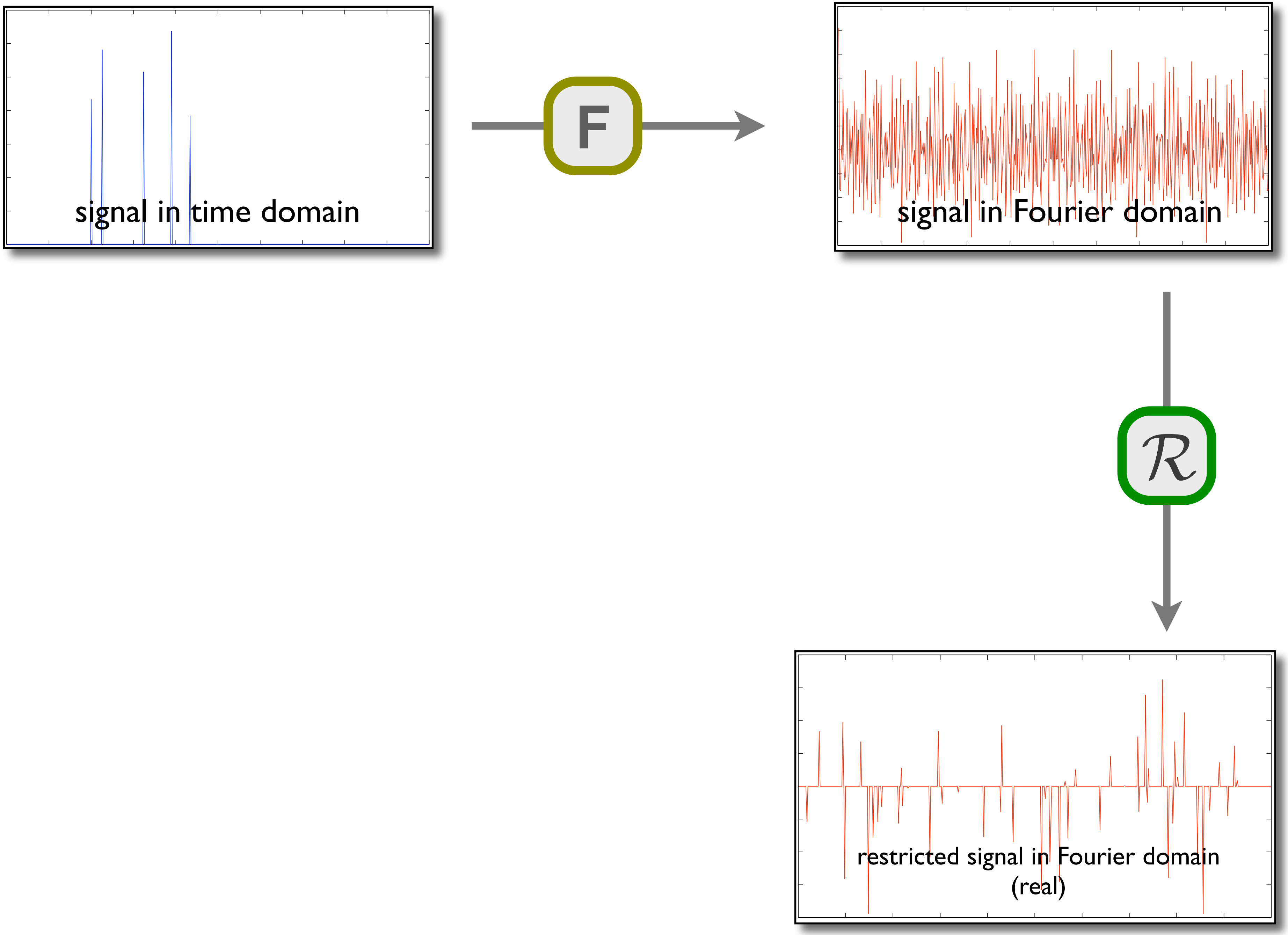
- operator is $\mathbf{S} = e^{-j\frac{\Delta x}{2\pi}\mathbf{D}}$
- $\mathbf{D}$ is differential operator

$\mathbf{D} =$

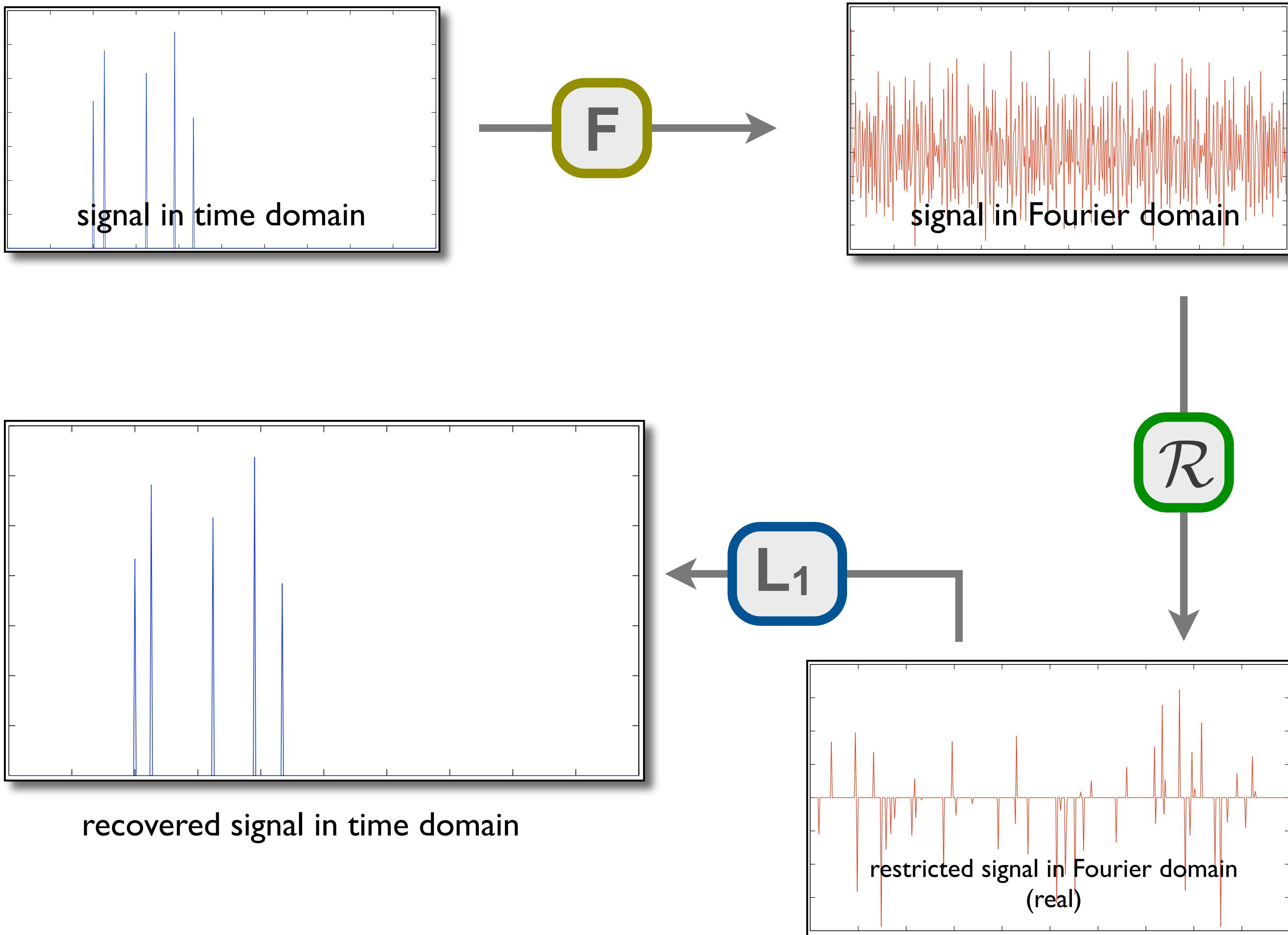



signal in time domain

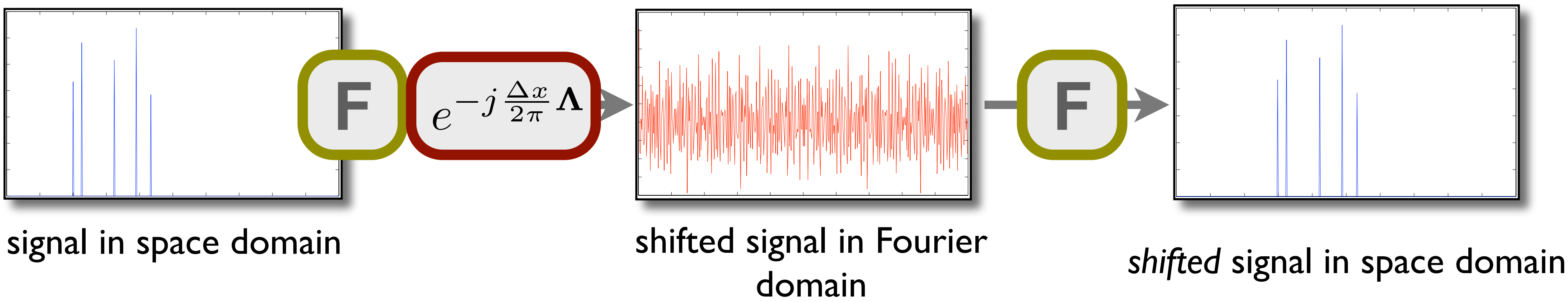

signal in time domain
F
signal in Fourier domain



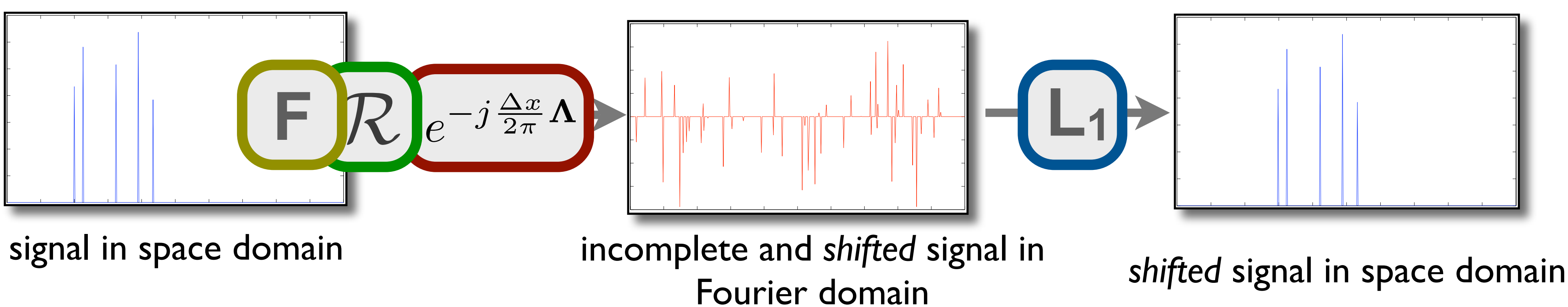
SLIM
Seismic Laboratory for Imaging and Modeling


signal in time domain
F
signal in Fourier domain
$\mathcal{R}$
restricted signal in Fourier domain
(real)
L_1
recovered signal in time domain

Straightforward Computation



Compressed Processing



Compressed ~~Sensing~~ "Computation"

- In a nutshell:
  - Trades the cost of L1 solvers for a compressed operator that is cheaper to compute, store, and synthesize

- L1 solver research is currently a hot topic in applied mathematics

Tibshirani, R., 1996, Least absolute shrinkage and selection operator, Software: http://www-stat.stanford.edu/~tibs/lasso.html.

Candès, E. J., and J. Romberg, 2005, ℓ_1-magic. Software: http://www.acm.caltech.edu/limagic/.

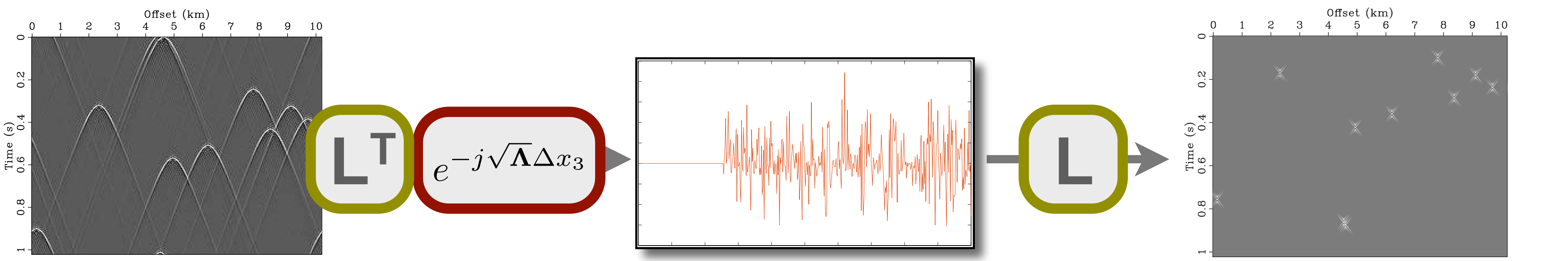
Donoho, D. L., I. Drori, V. Stodden, and Y. Tsaig, 2005, SparseLab, Software: http://sparselab.stanford.edu/.

Figueiredo, M., R. D. Nowak, and S. J. Wright, 2007, Gradient projection for sparse reconstruction, Software: http://www.lx.it.pt/~mtf/GPSR/.

Koh, K., S. J. Kim, and S. Boyd, 2007, Simple matlab solver for l1-regularized least squares problems, Software: http://www-stat.stanford.edu/~tibs/lasso.html.

SLIM
Seismic Laboratory for Imaging and Modeling

Straightforward 1-Way inverse Wavefield Extrapolation

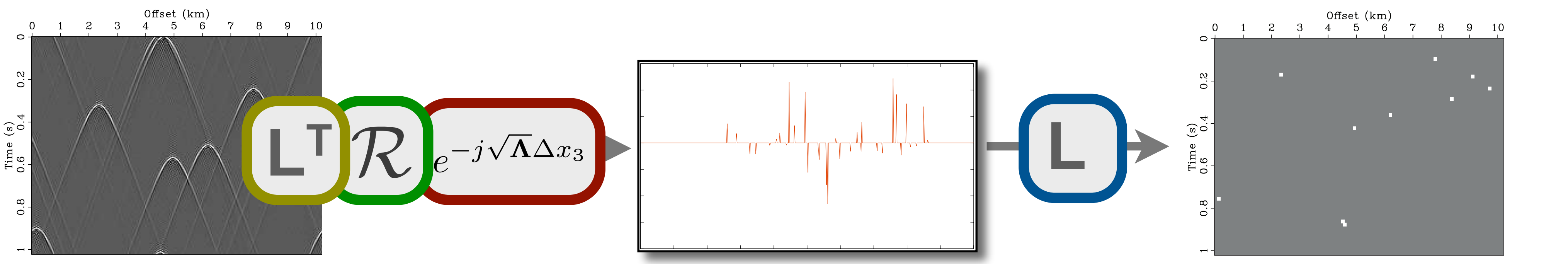


wavefield in space-time domain

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

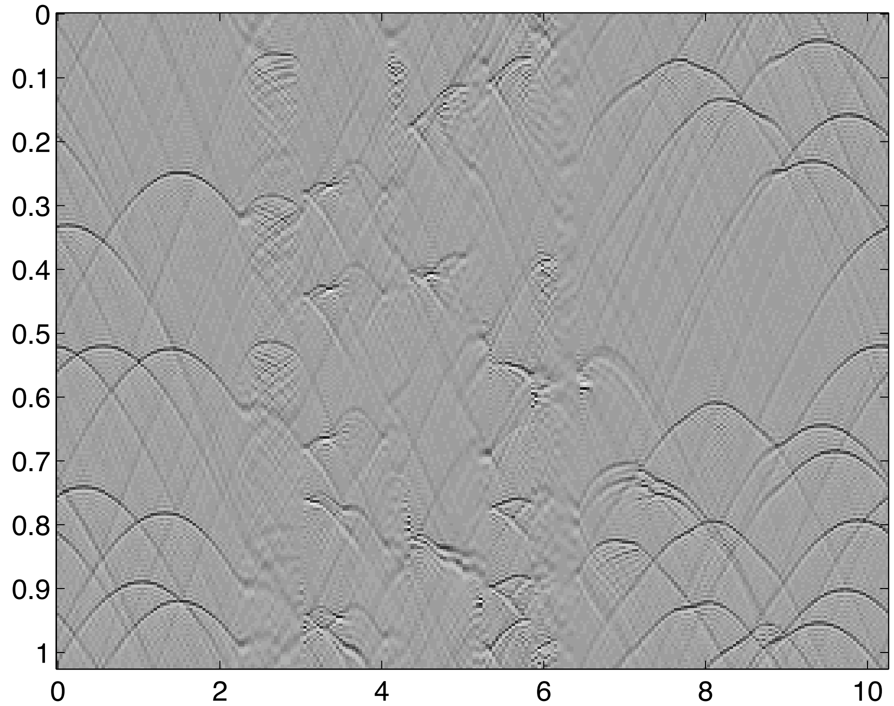
Compressed wavefield extrapolation

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

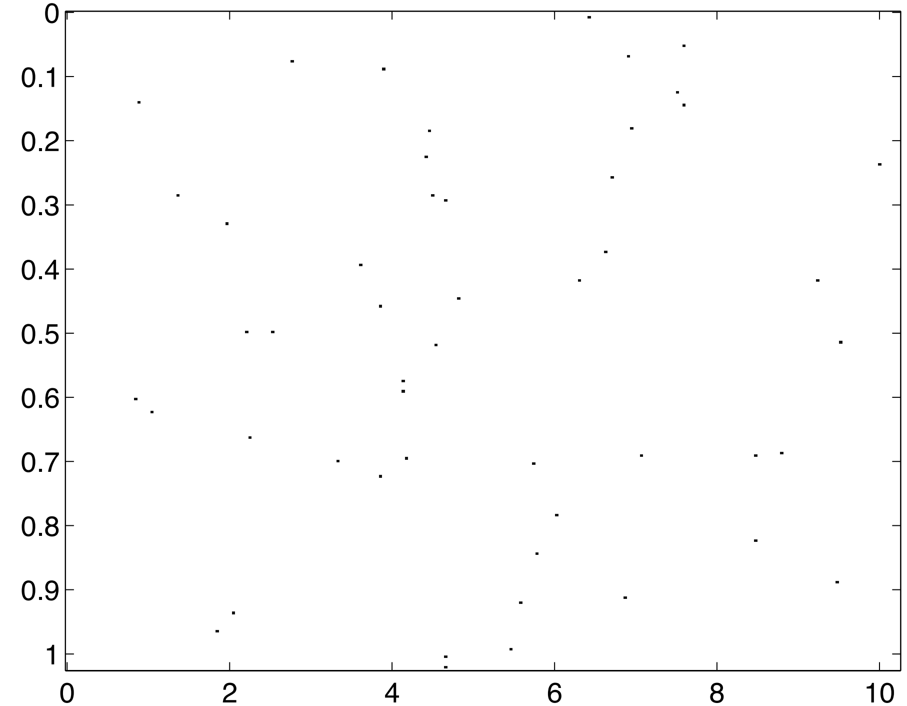
- Randomly subsample & phase rotation in Modal domain
- Recover by norm-one minimization
- Capitalize on
  - the incoherence modal functions and point scatterers
  - reduced explicit matrix size
  - constant velocity <=> Fourier recovery

Compressed wavefield extrapolation

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



propagated 1.5km down



recovered though L1 inverson

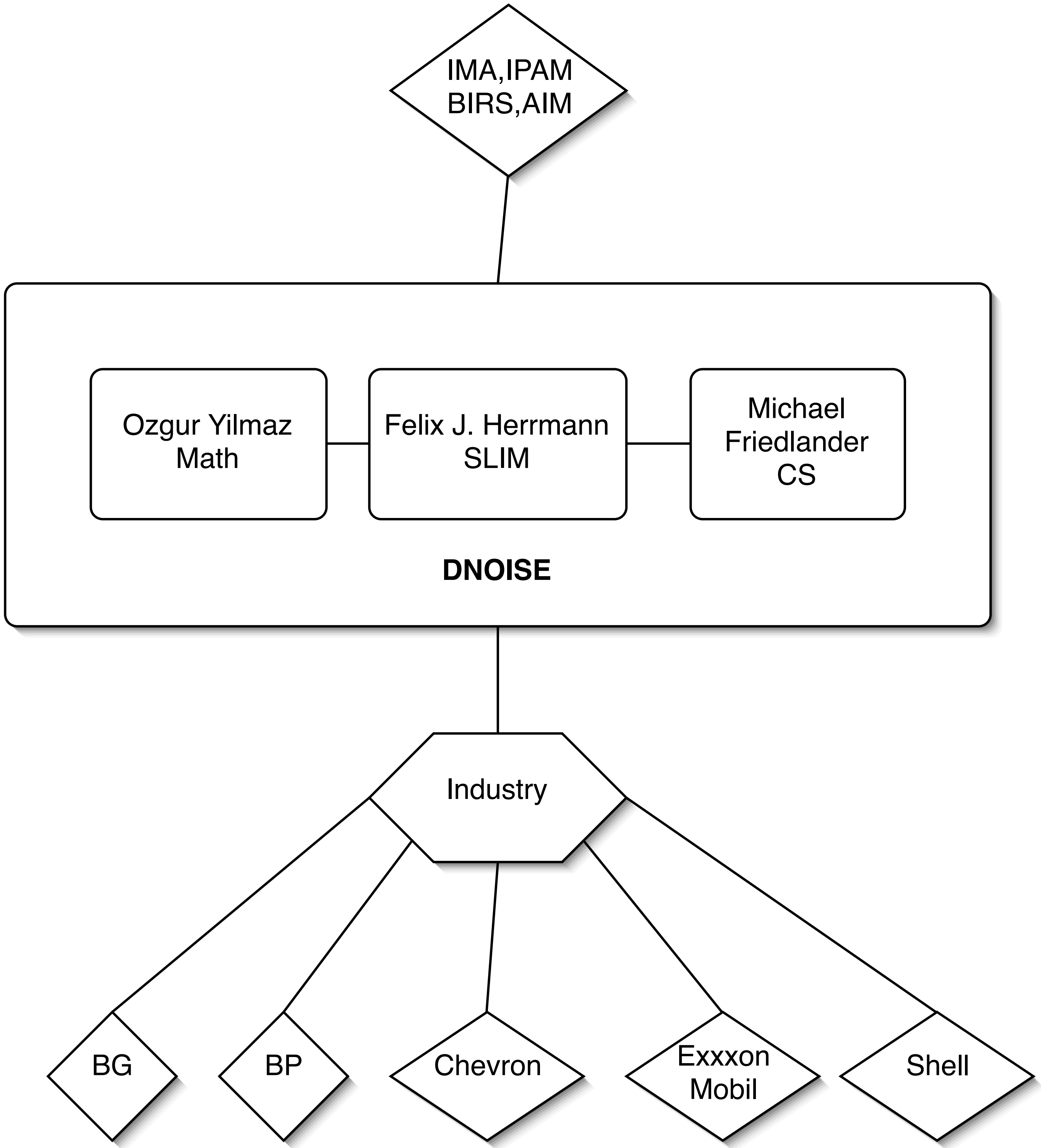
Restricted L transform to ~0.01 of original coefficients

SLIM Seismic Laboratory for Imaging and Modeling

Observations

- Compressed wavefield extrapolation
  - reduction in synthesis cost
  - mutual coherence curvelets and eigenmodes
  - performance of norm-one solver
  - keep the constants under control ...
- Open problems
  - fast “random” eigensolver
  - incoherence eigenfunctions and sparsity transform
- Double-role CS matrix is cool ... upscaling to “real-life” is a challenge ....

DNOISE: an academic-industry-NSERC partnership


IMA,IPAM
BIRS,AIM
Ozgur Yilmaz
Math
Felix J. Herrmann
SLIM
Michael
Friedlander
CS
DNOISE
Industry
BG
BP
Chevron
Exxxon
Mobil
Shell

Industry consortia

Since early 80's in exploration seismology

Consortia work on common set of problems

No secret research

Hurdles

- data access
- QC
- IT infrastructure
- University Liaison offices
- being interdisciplinary sounds easier than it is ...

SLIM
Seismic Laboratory for Imaging and Modeling

DNOISE

DNOISE: Dynamic nonlinear optimization for imaging in seismic exploration

- NSERC Collaborative Research & Development Grant
- Matches SINBAD Consortium supported by industry
  - organized by ITF (non-profit technology broker in the UK)
  - supported by BG, BP, Chevron, ExxonMobil and Shell
  - $70 k annually per company
  - total budget $500-600 k annually
- Involves
  - Dr. Michael Friedlander (CS) and Ozgur Yilmaz (Math) as co-PI's
  - 2-3 postdocs
  - 8 graduate students
  - 2 undergraduate students
  - 2 programmers
  - 1 part-time admin person

Challenges

Development of common language amongst
- Geophysics
- Computer Science
- Math

Difference in mentality/approach
- Geophysicist throws everything at a problem and if it works ... it works
- Mathematicians/computer scientists
  - narrow problem to proof theorems
  - may not be relevant
  - do not necessary understand what "deliverables" are
  - do not speak the same language

Knowledge dissemination

Dissemination

SPARCO: a test suite for norm-one problems

- framework for setting up small-size CS problems
- first step towards performance benchmarks
- www.cs.ubc.ca/labs/scl/sparco

SLIMPy: “compiler” for abstract numerical algorithms

- operator overloading in Python
- integration with scalable seismic processing packages

Madagascar: public-domain seismic processing package

- reproducible research
- slim.eos.ubc.ca/
- rsf.sourceforge.nethyperlink

SLIM
Seismic Laboratory for Imaging and Modeling

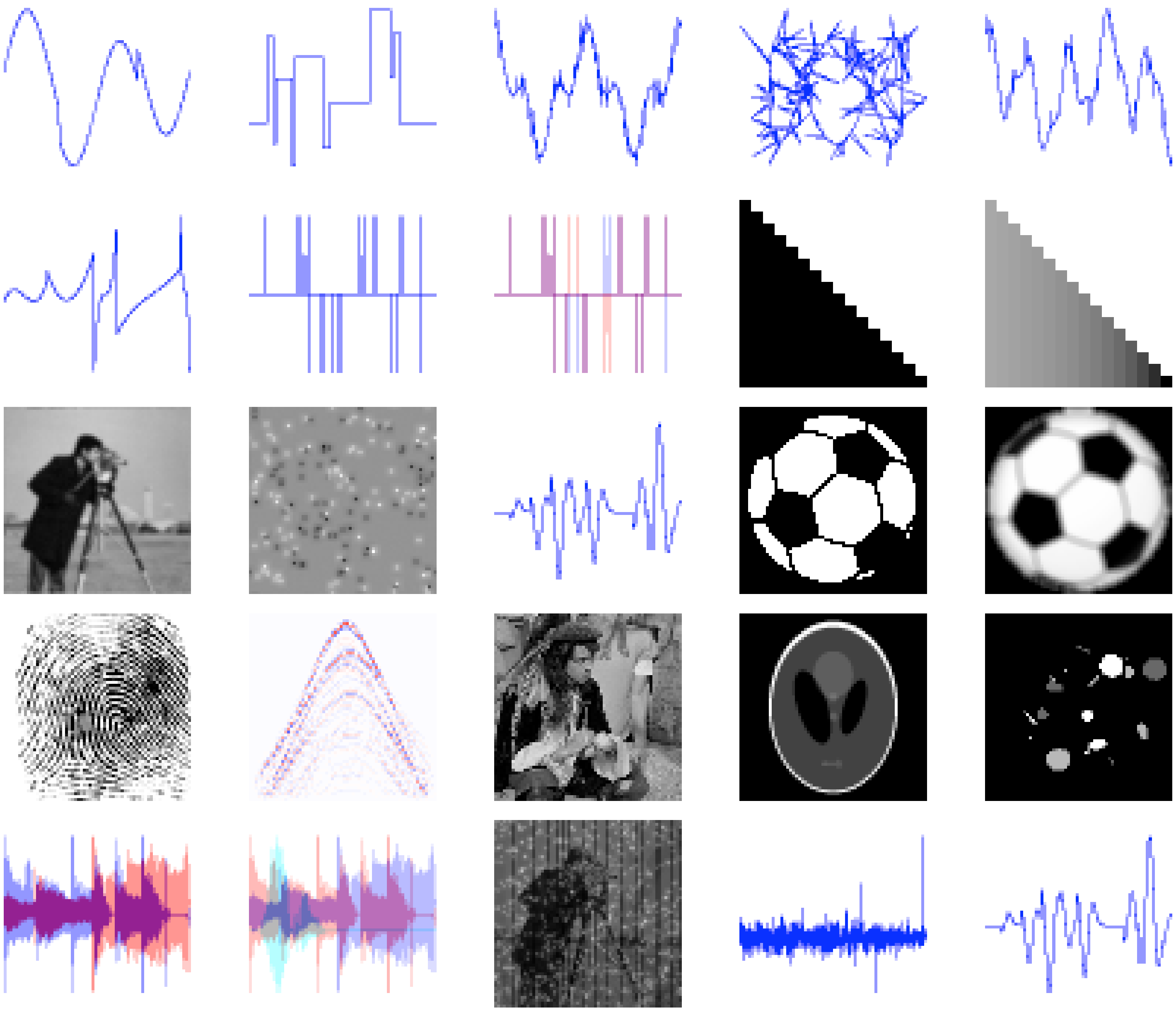
Nonlinear wavefield sampling

- *sparsifying transform*
  - typically **localized** in the time-space domain to handle the complexity of seismic data
  - preserves edges/wavefronts

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

- ***sparsity-promoting solver***
  - requires few matrix-vector multiplications
  - scales to number of unknowns exceeding 2^{30} ("small")

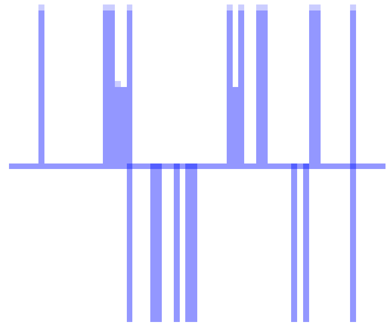
SPARCO: Sparse Reconstruction Test Suite



http://www.cs.ubc.ca/labs/scl/sparco

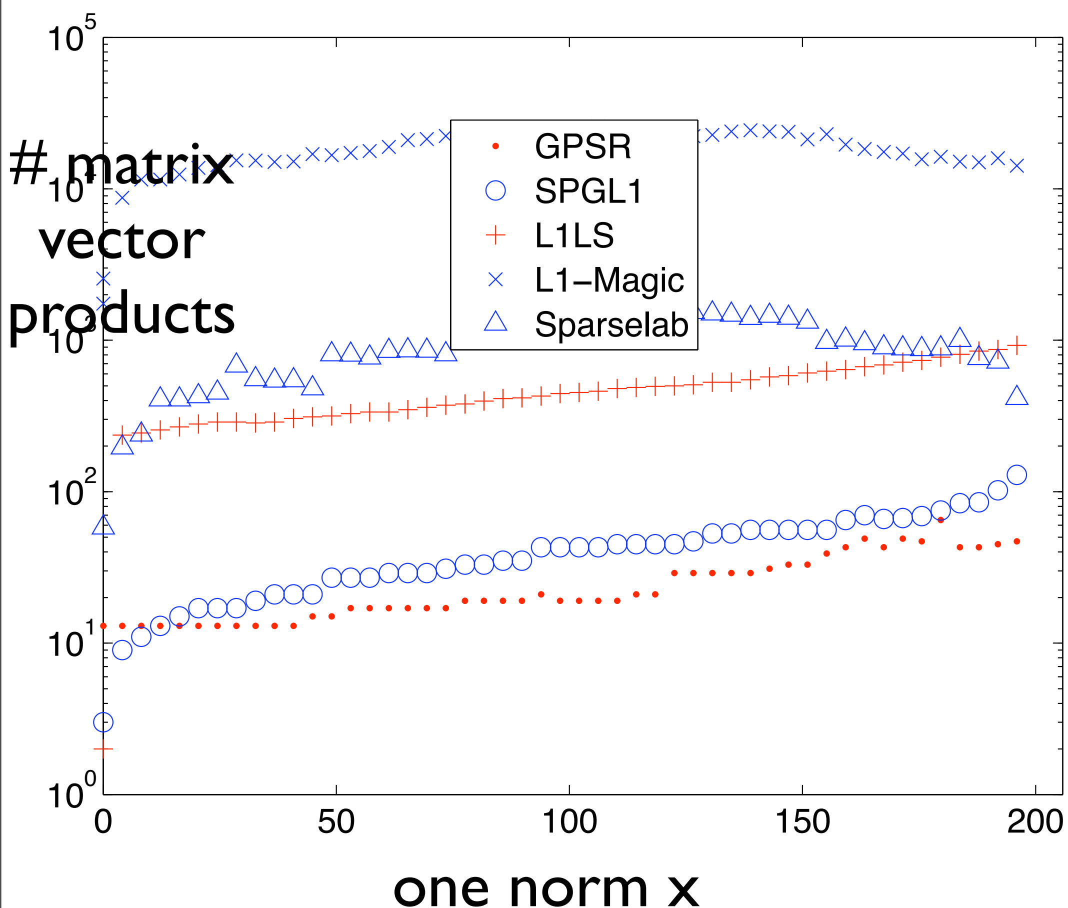
Gaussian ensemble, spikes signal

$A =$ Gaussian

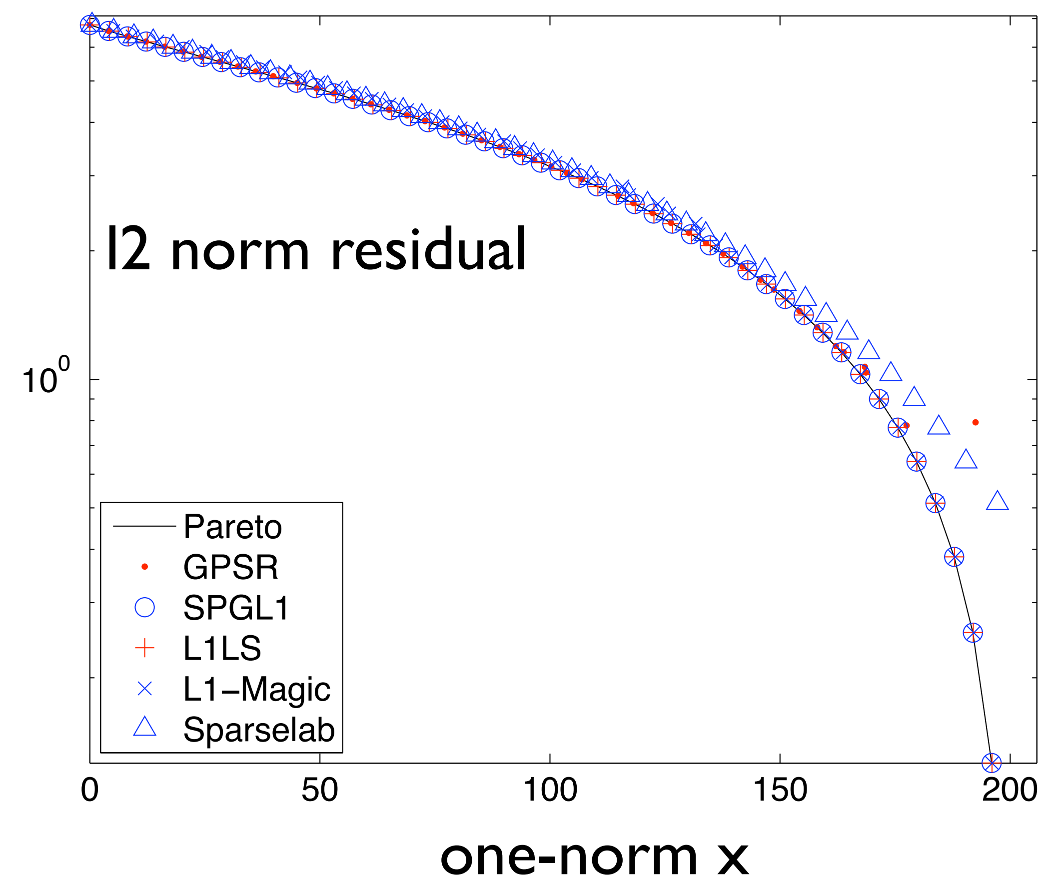
$b =$ 

1200×5120
Candés, Romberg, & Tao '05

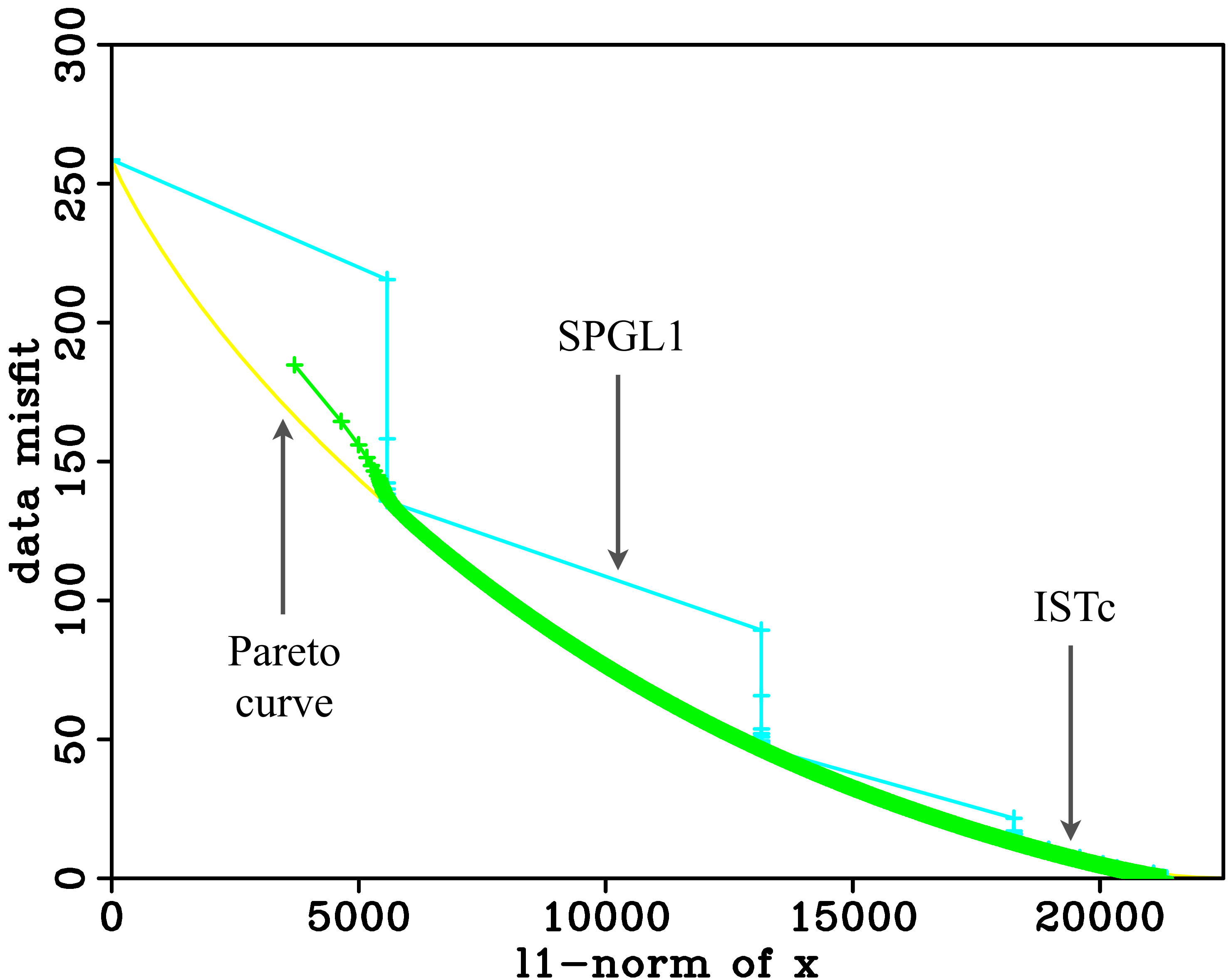
Matrix-vector products



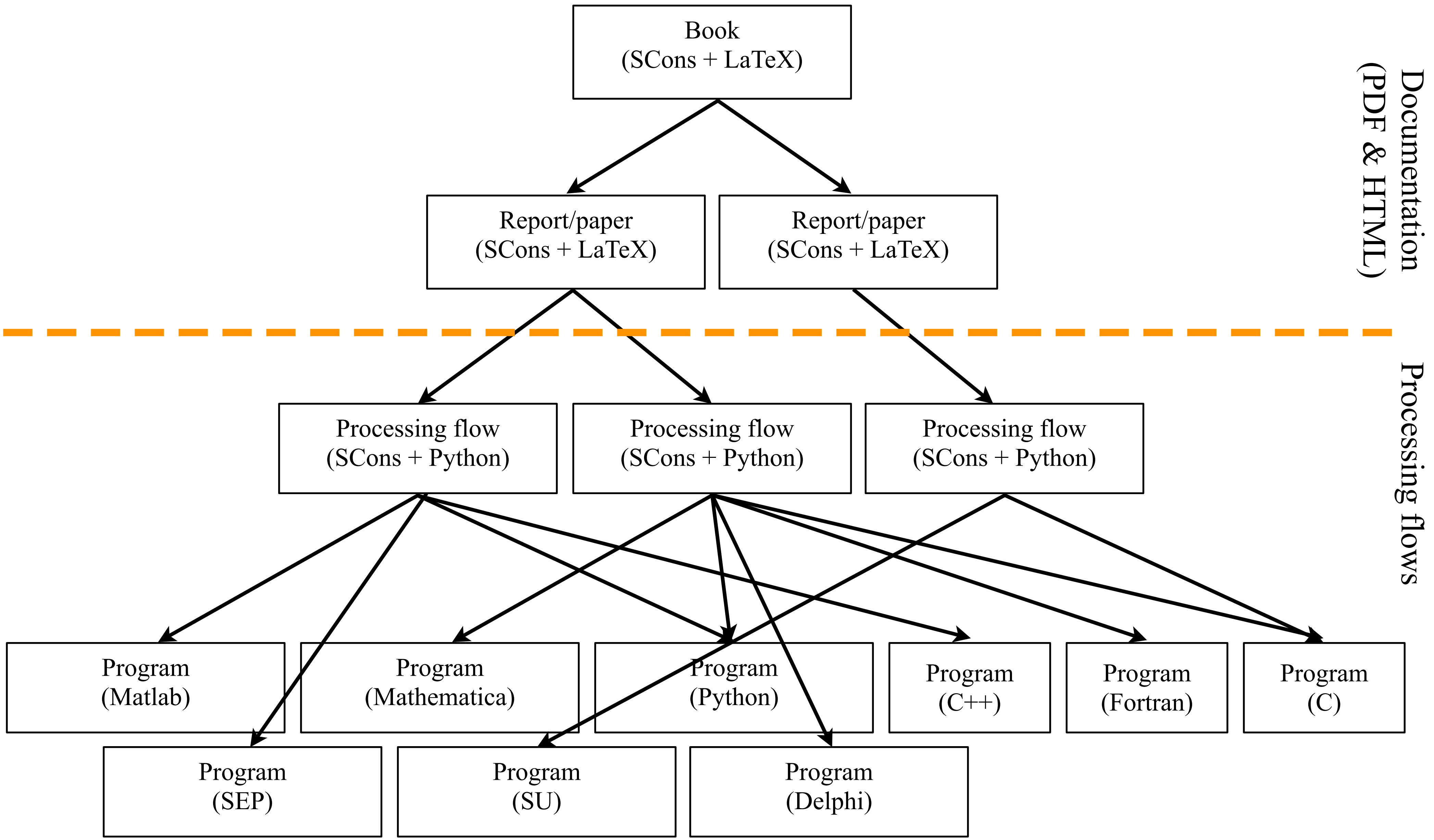
Pareto curve



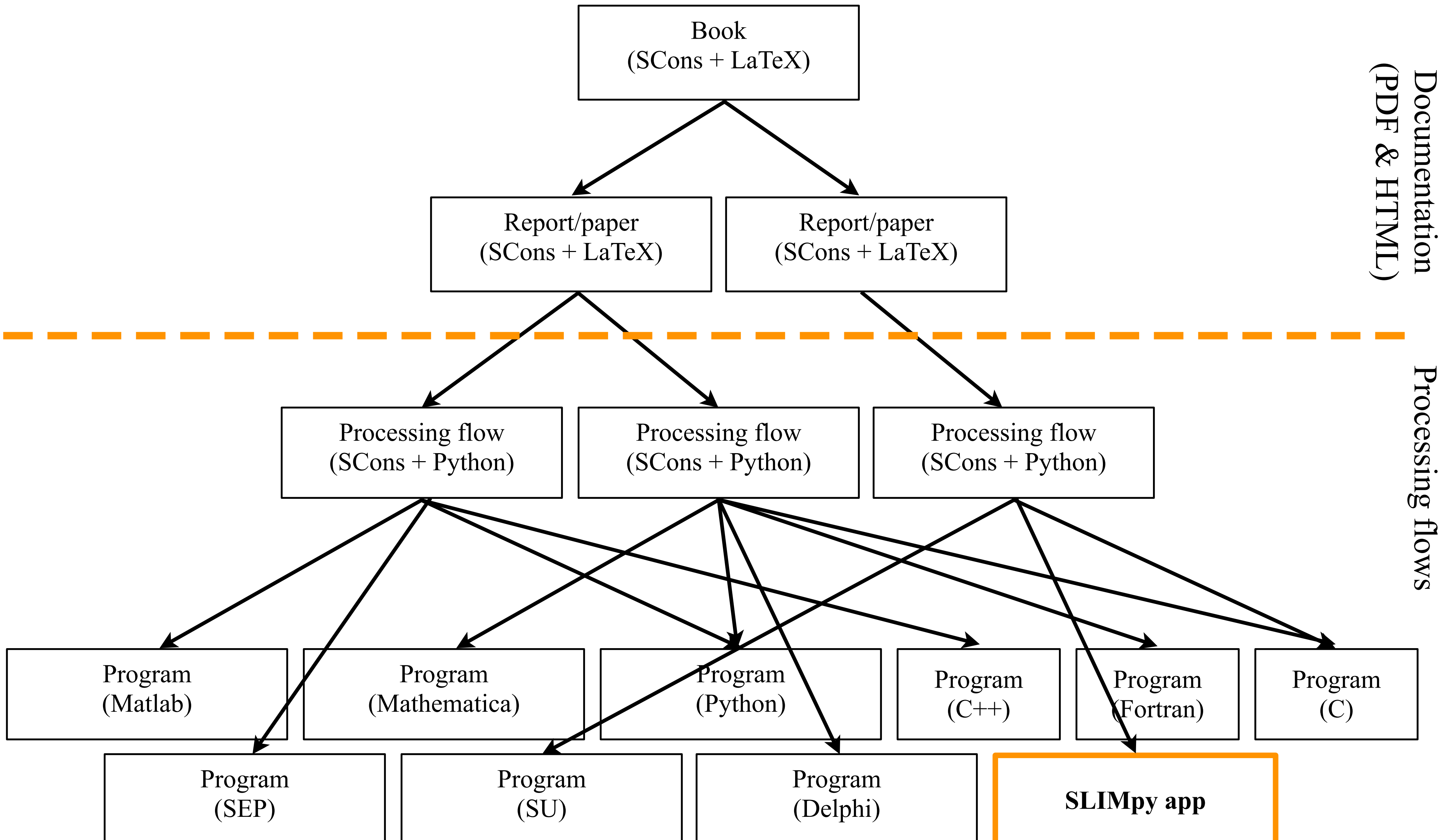
Optimization paths



Madagascar


Book
(SCons + LaTeX)
Report/paper
(SCons + LaTeX)
Report/paper
(SCons + LaTeX)
Documentation
(PDF & HTML)
Processing flow
(SCons + Python)
Processing flow
(SCons + Python)
Processing flow
(SCons + Python)
Processing flows
Program
(Matlab)
Program
(Mathematica)
Program
(Python)
Program
(C++)
Program
(Fortran)
Program
(C)
Program
(SEP)
Program
(SU)
Program
(Delphi)

Madagascar


Book
(SCons + LaTeX)
Report/paper
(SCons + LaTeX)
Report/paper
(SCons + LaTeX)
Documentation
(PDF & HTML)
Processing flow
(SCons + Python)
Processing flow
(SCons + Python)
Processing flow
(SCons + Python)
Processing flows
Program
(Matlab)
Program
(Mathematica)
Program
(Python)
Program
(C++)
Program
(Fortran)
Program
(C)
Program
(SEP)
Program
(SU)
Program
(Delphi)
SLIMpy app

Abstraction

Let data be a vector $y \in \mathbb{R}^n$.
Let $\mathbf{A}_1 := \mathbf{C}^T \in \mathbb{C}^{n \times M}$ be the inverse curvelet transform and $\mathbf{A}_2 := \mathbf{F}^H \in \mathbb{C}^{n \times n}$ the inverse Fourier transform.

Define $\mathbf{A} := \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T \end{bmatrix}^T$

Solve

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon$$

```
y = vector('data.rsf')

A1 = fdct2(domain=y.space).adj()

A2 = fft2(domain=y.space).adj()

A = aug_oper([A1, A2])

solver = GenThreshLandweber(10,5,thresh=None)

x=solver.solve(A,y)
```

Conclusions

Math institutes have been instrumental

- exposure to the latest of the latest
- establish a research network

Success research program depends on

- understanding the problems
- engineering & software development
- disseminate results (reproducible research)

Science:Extension CS towards

- more general (nonlinear) problems
- compressive computations ....

For the future: Redirection of emphasis away from “Let’s gather as much data as we can and let’s analyze it all” to “What are we looking for and how can we best sample....”

SLIM Seismic Laboratory for Imaging and Modeling

Acknowledgments

The audience for listening and the organizers for putting this great workshop together ....

The authors of CurveLab (Demanet, Ying, Candes, Donoho)

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE (334810-05) of F.J.H. This research was carried out as part of the SINBAD project with support, secured through ITF (the Industry Technology Facilitator), from the following organizations: BG Group, BP, Chevron, ExxonMobil and Shell.