Compressive sensing and sparse recovery in exploration seismology

Felix J. Herrmann

SLIM

Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Drivers

Our incessant

- demand for *hydrocarbons* while we are *no* longer finding oil...
- desire to understand the Earth's inner workings

Push for improved *seismic inversion* to

- create *more high-resolution* information
- from *noisier* and *incomplete* data

Controversial statements

Size of our *discretizations* is *dictated* by

- a *far* too *pessimistic Nyquist-sampling criterion* compounded by the *curse of dimensionality*
- our *insistence* to sample *periodically* and/or *sequentially*

Our desire to work with *all* data

- leads to "over emphasis" on data *collection* & *full-data processing*
- prohibits *inversion* that requires *multiple* passes through *data*

Wish list

Acquisition & *inversion* costs determined by *structure of data* & *complexity* of the *subsurface*

- sampling *criteria* that are *dictated by* transform-domain *sparsity* and *not* by the *size* of the *discretization*

Controllable error that depends on

- degree of *subsampling* / *dimensionality* reduction
- available computational resources

Main themes

1. Randomized sampling and sparsifying transforms
2. Convex optimization
3. Randomized dimensionality reduction
4. Tutorials & Software

Theme I: Randomized sampling and sparsifying transforms

SLIM
Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Additional faculty
Ozgur Yilmaz
Associate professor Math

- M.A., Bogazici University, Turkey
- Ph.D., Princeton
- Applied harmonic analysis
- Signal processing
- Information theory

SLIM
Seismic Laboratory for Imaging and Modeling

MSc. Student Haneet Wason

- BSc Geophysics, University of Calgary
- Shearlets, curvelets, and simultaneous-source data

SLIM
Seismic Laboratory for Imaging and Modeling

Postdoc
Rayan Saab

- M.A.Sc. in Electrical Eng., UBC
- B.E. in Computer and Communications Eng., American Univ. of Beirut
- Blind Source Separation
- Statistical Signal Processing
- Discrete Optimization
- Seismic and Biomedical Signal Processing

SLIM
Seismic Laboratory for Imaging and Modeling

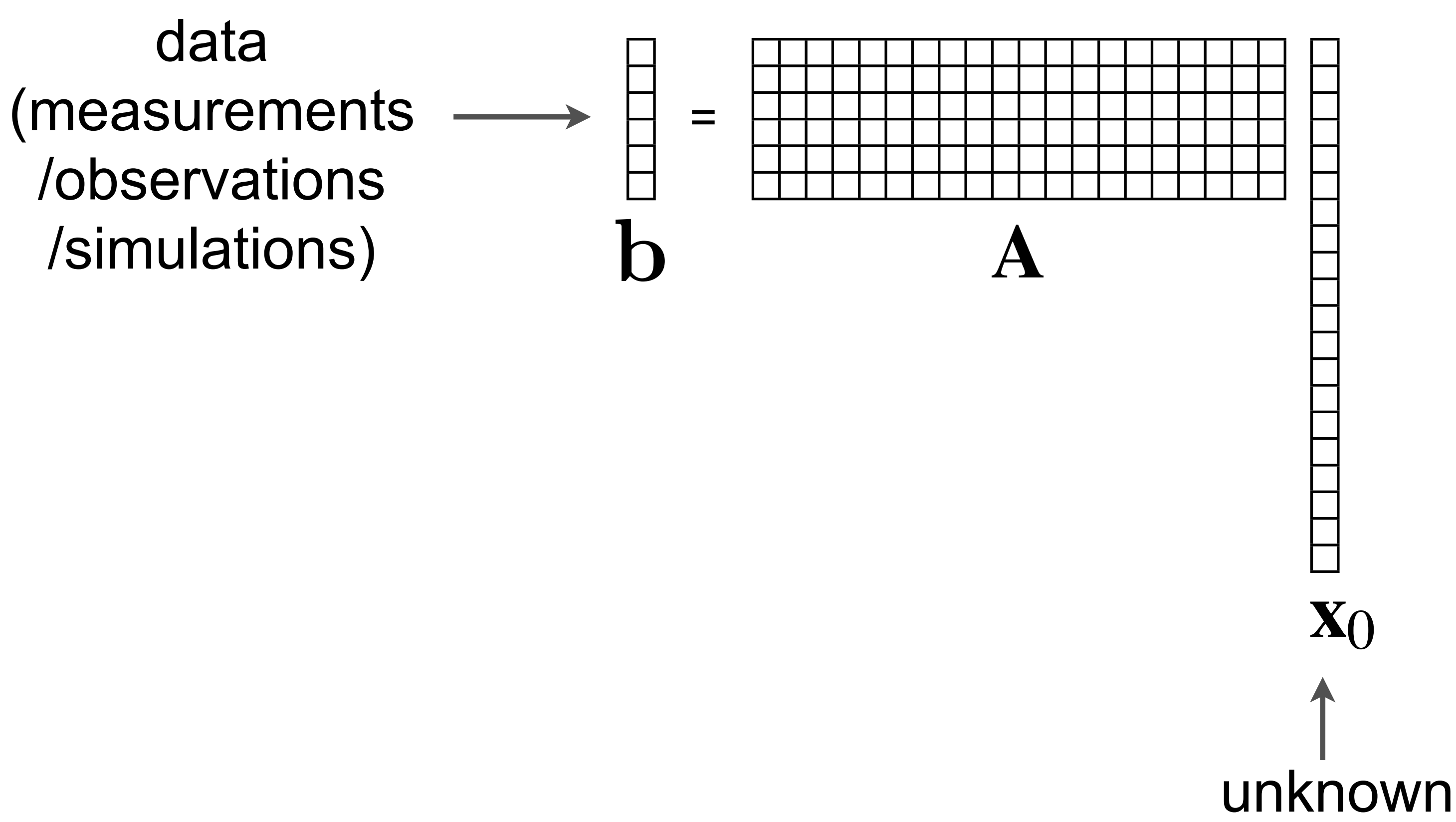
Post Doctoral Fellow
Hassan Mansour

- Joined September 2009
- Ph.D. from Electrical Engineering, UBC, 2009 in video coding.
- Current research interests: L1-minimization and compressed sensing.

SLIM
Seismic Laboratory for Imaging and Modeling

Problem statement

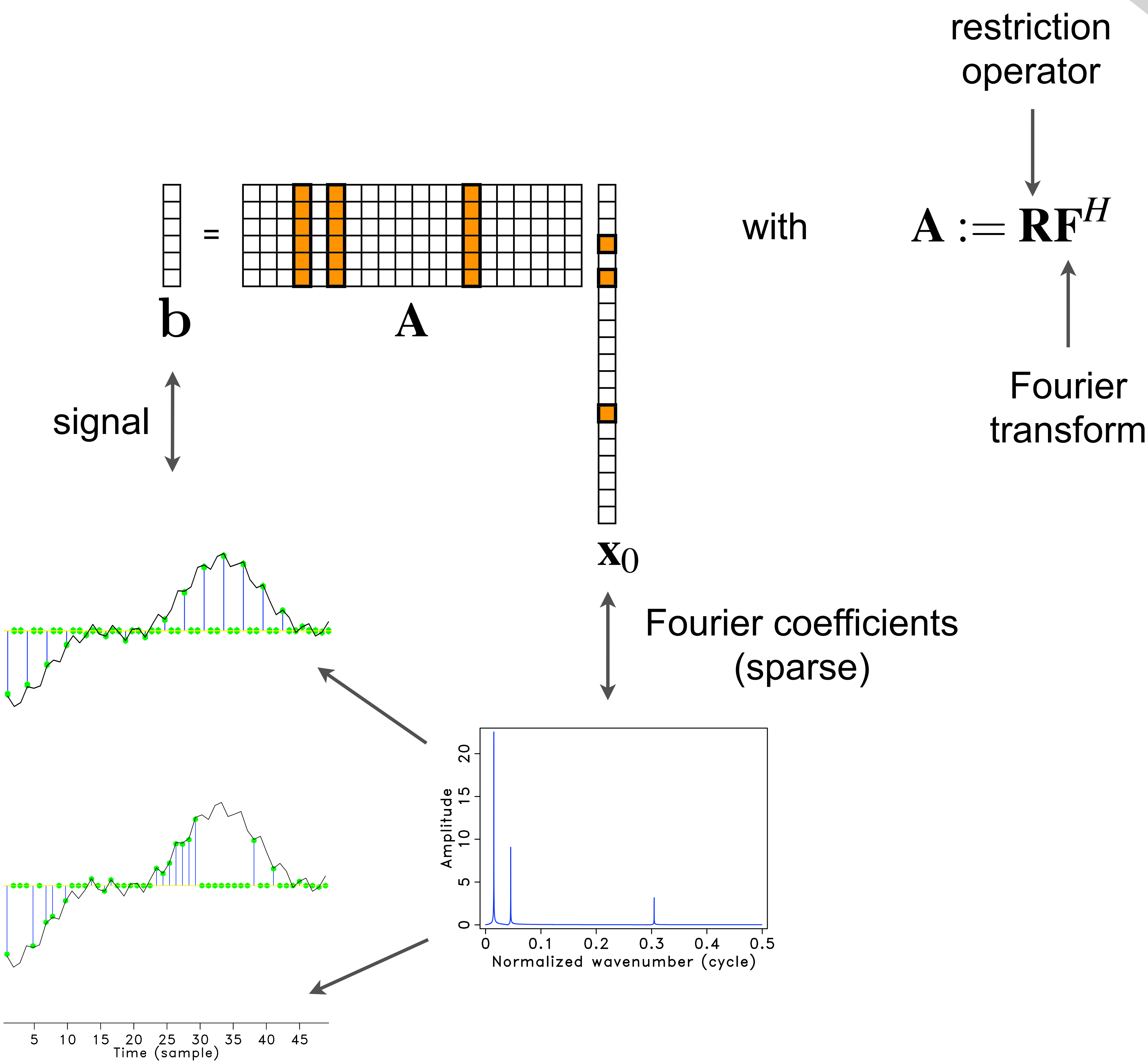
Consider the following (severely) *underdetermined* system of *linear* equations:



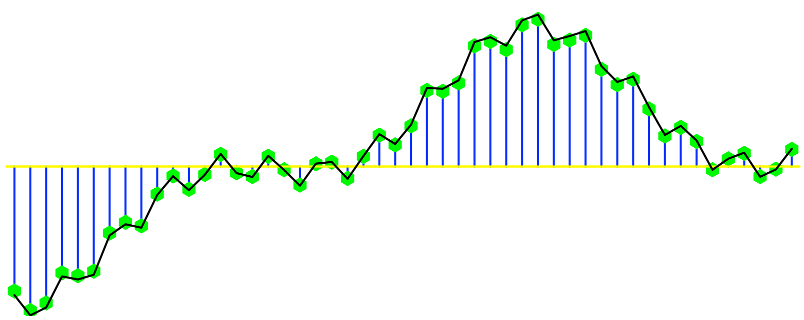
Is it possible to recover $\mathbf{x}_0$ accurately from $\mathbf{b}$?

The new field of *Compressive Sensing* attempts to answer this.

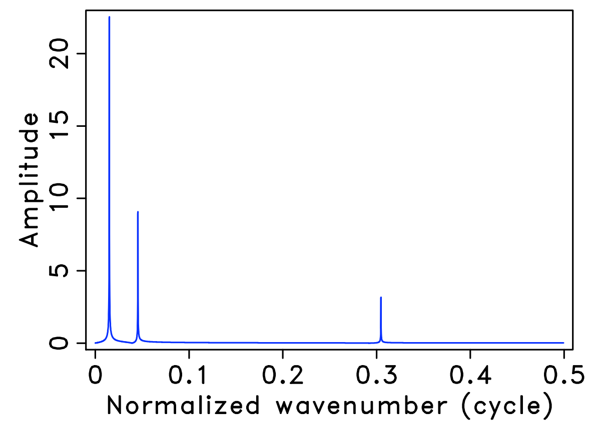
Sparse recovery



Coarse sampling schemes

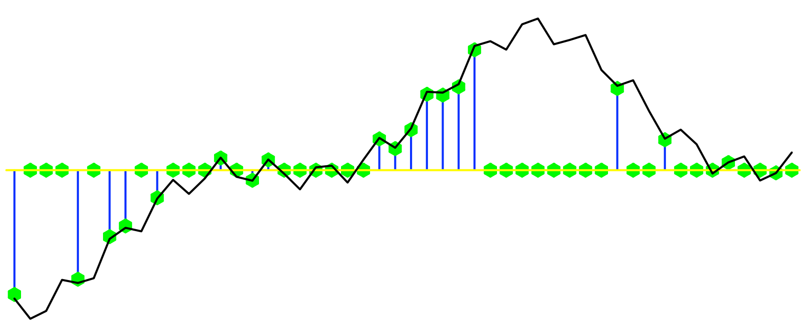


Fourier transform

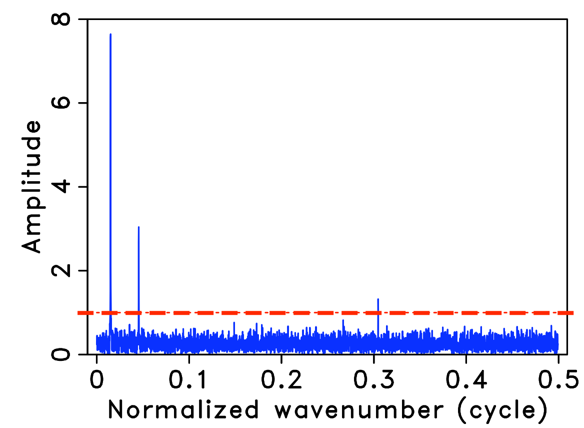


few significant coefficients

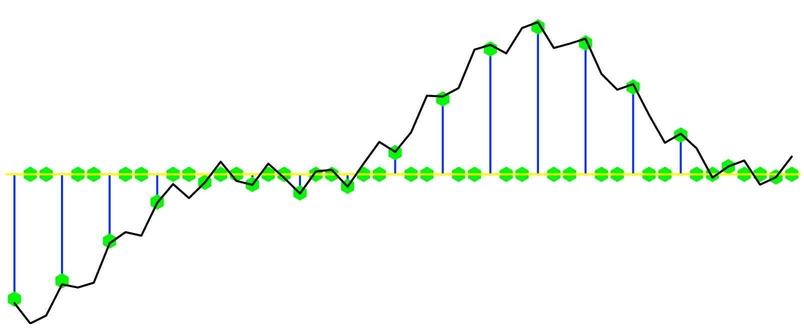
3-fold under-sampling



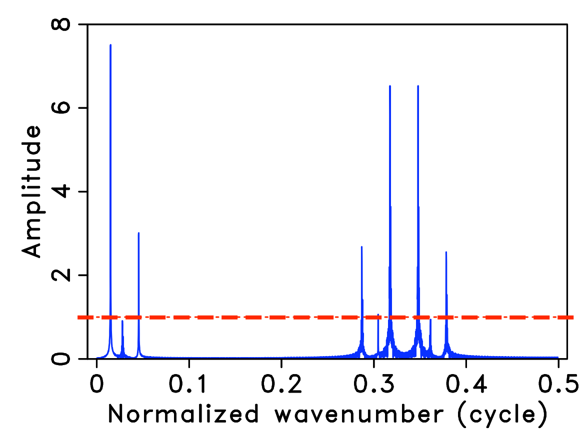
Fourier transform



✓ significant coefficients detected



Fourier transform

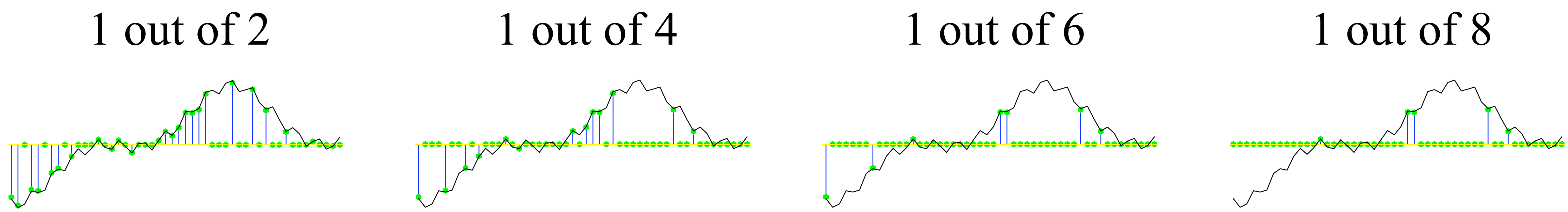


✗ ambiguity

[Hennenfent & Herrmann, ’08]

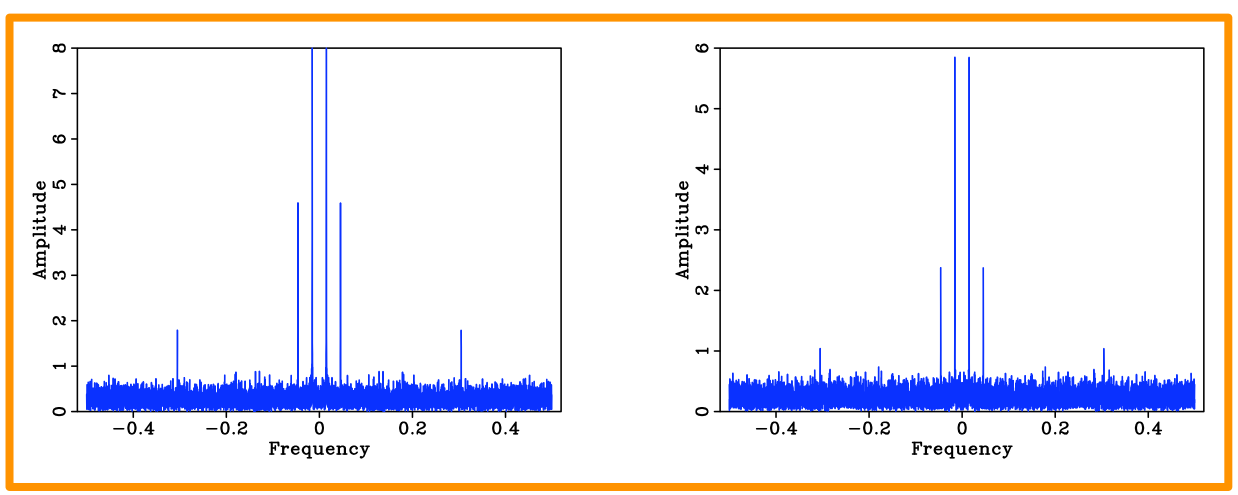
Undersampling “noise”

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

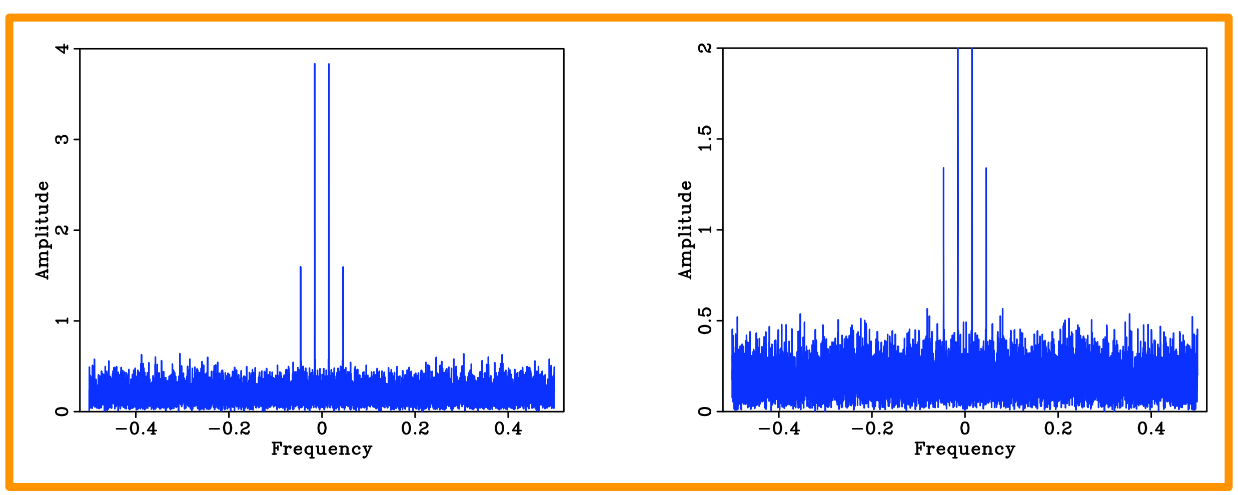


less acquired data

3 detectable Fourier modes



2 detectable Fourier modes



Sparse one-norm recovery

Signal model

$$\mathbf{b} = \mathbf{A}\mathbf{x}_0 \quad \text{where} \quad \mathbf{b} \in \mathbb{R}^n$$

and $\mathbf{x}_0$ *k* sparse

Sparse one-norm recovery

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \stackrel{\text{def}}{=} \sum_{i=1}^{N} |x[i]| \quad \text{subject to} \quad \mathbf{b} = \mathbf{A}\mathbf{x}$$

with $\mathrm{n} \ll N$ where N is the *ambient dimension*

Study recovery as a function of

- the subsampling ratio n/N
- “over sampling” ratio k/n

[Sacchi ’98]
[Candès et.al, Donoho, ’06]

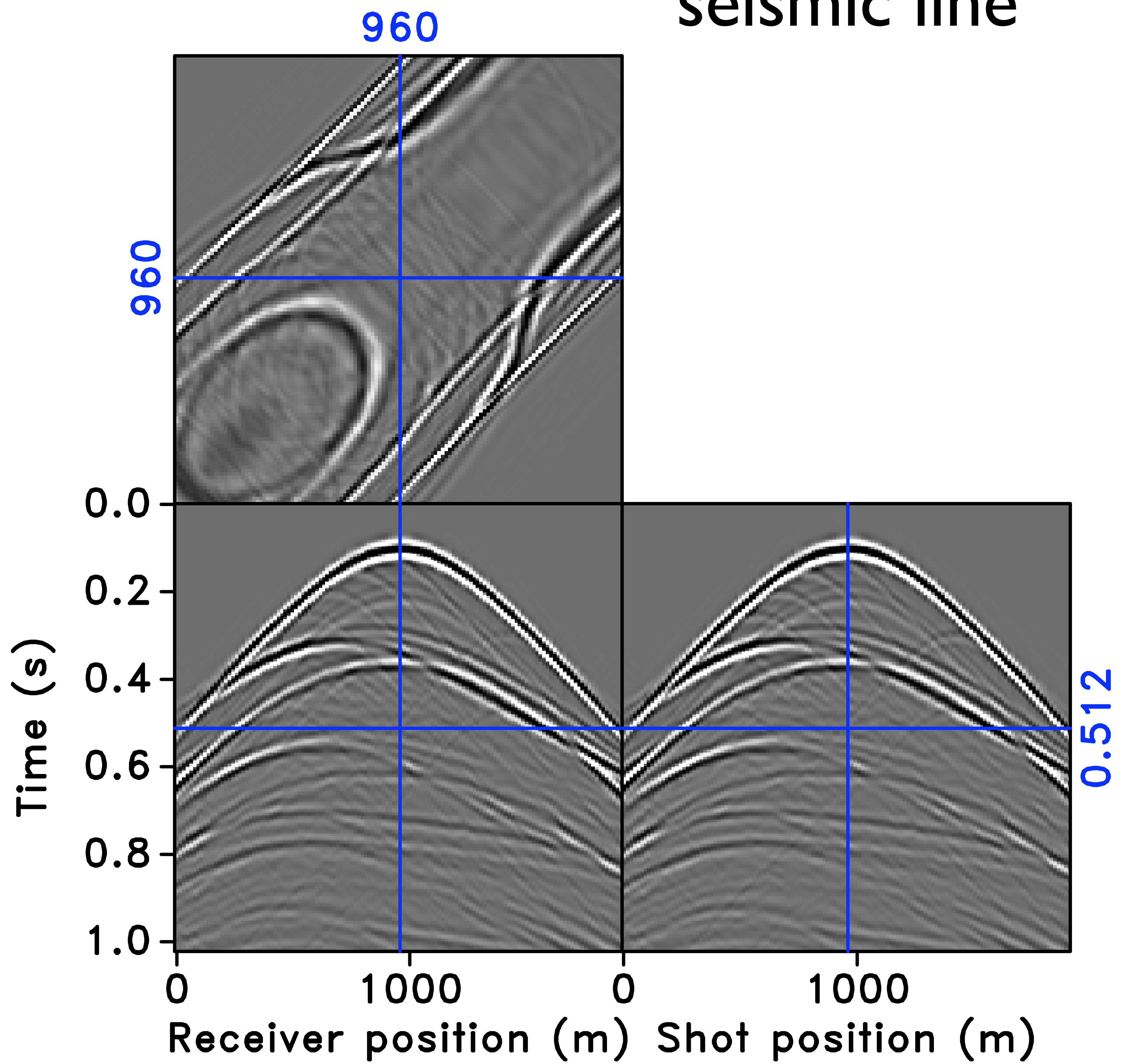
Case study I

Acquisition design according to Compressive Sensing

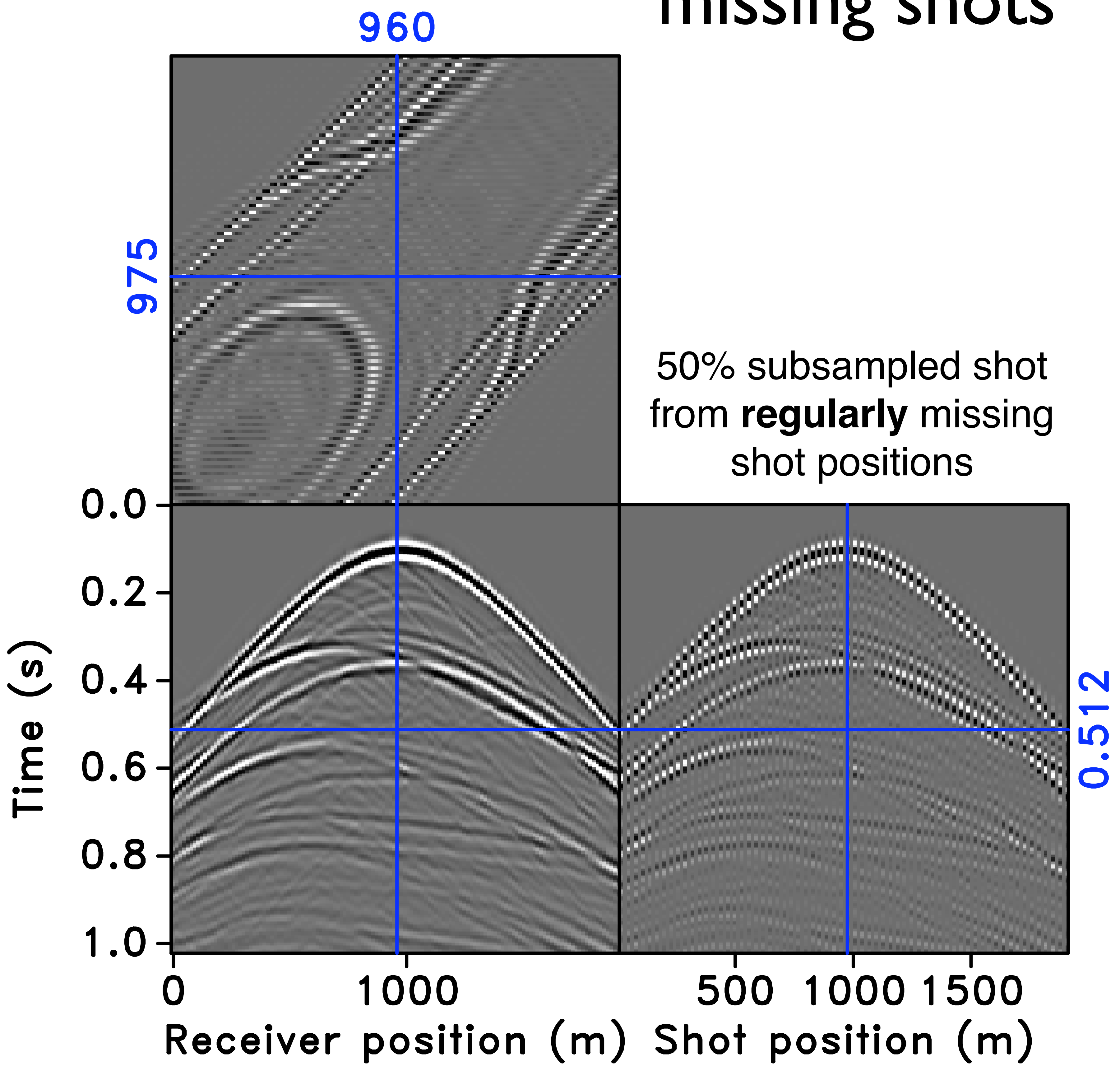
- ***Periodic*** *subsampling* vs ***randomized jittered*** sampling of ***sequential*** sources
- *Subsampling with randomized jittered* ***sequential*** sources vs *randomized phase-encoded* ***simultaneous*** sources

[Hennenfent & Herrmann, '08]

seismic line

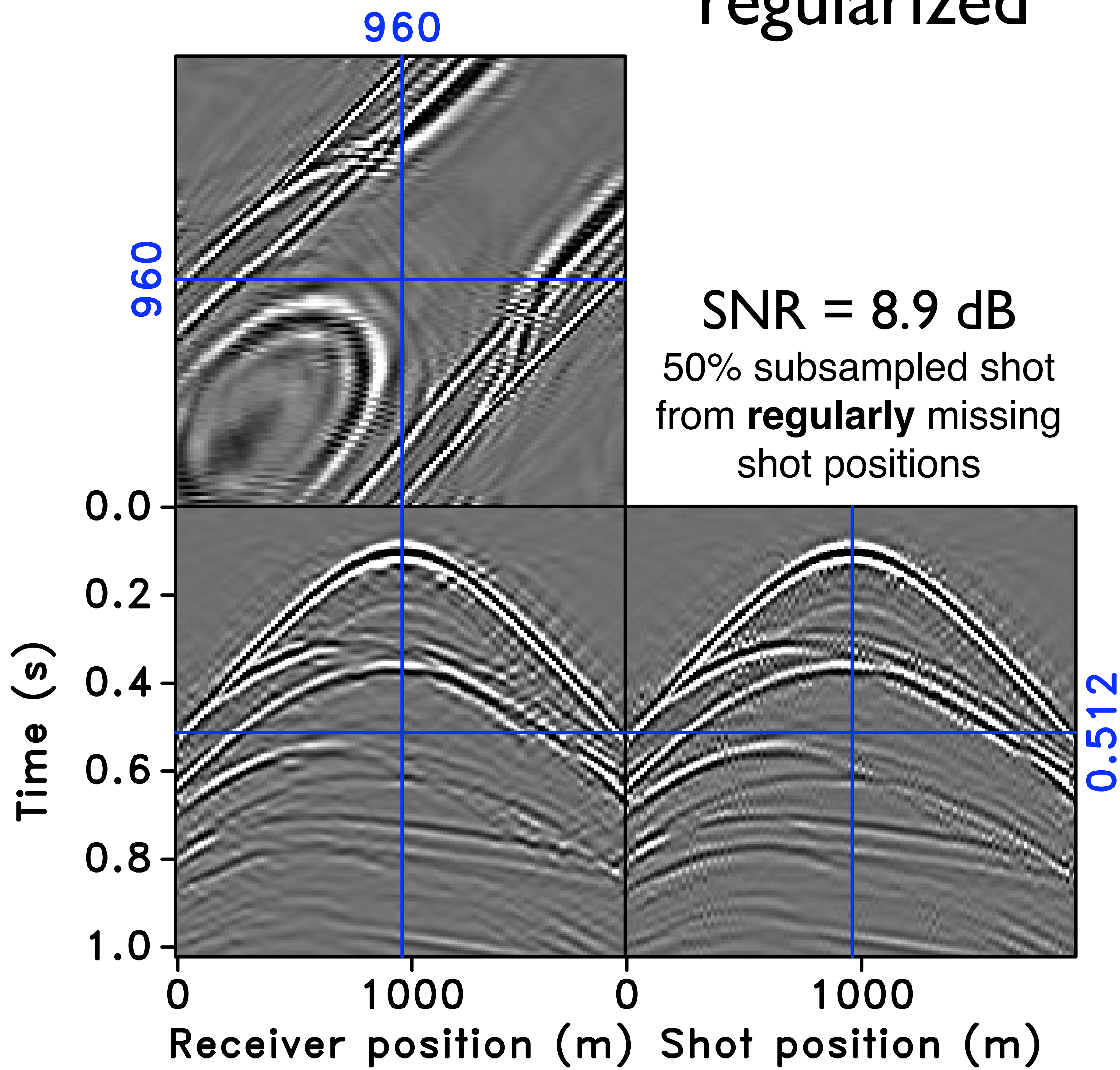


missing shots



50% subsampled shot from **regularly** missing shot positions

regularized



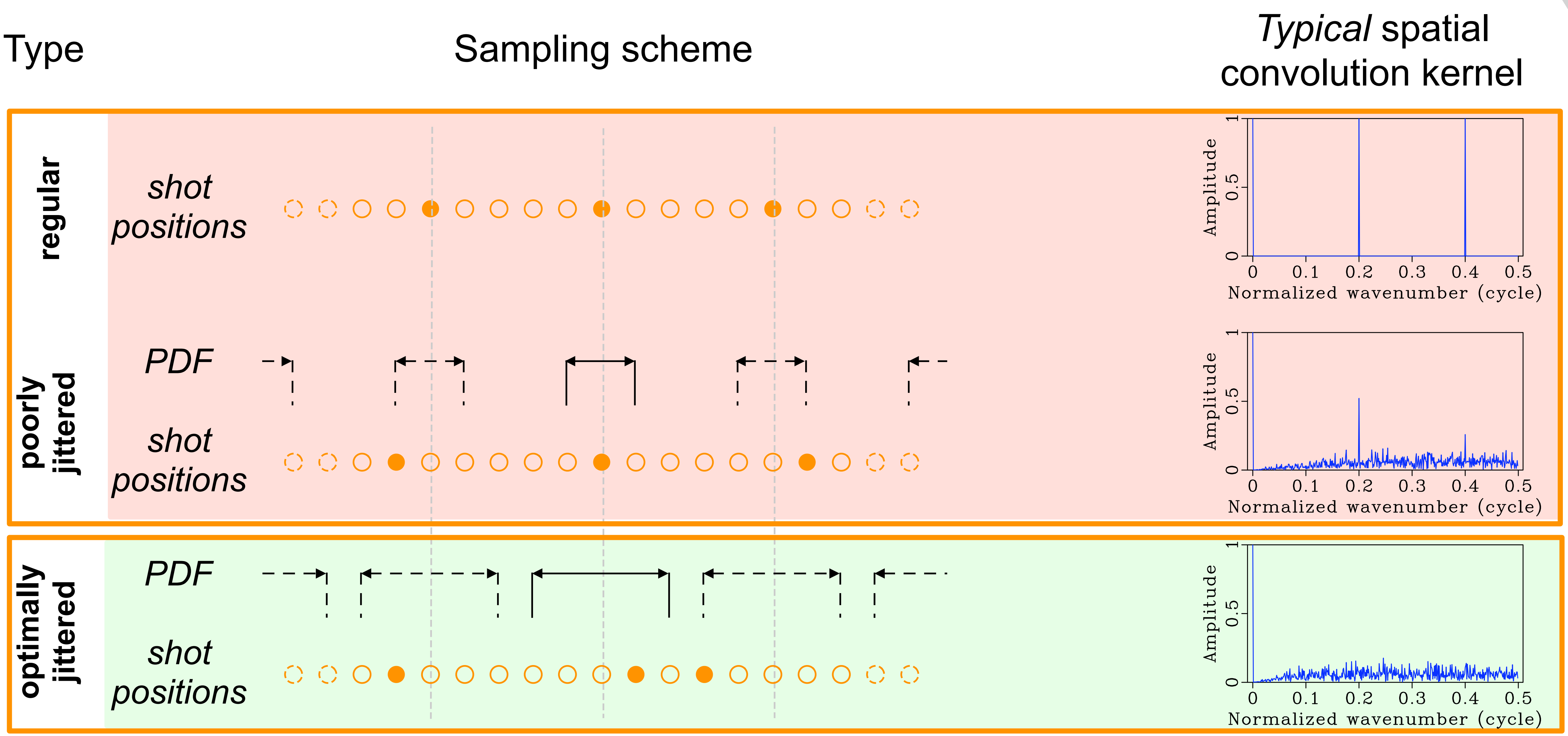
SNR = 8.9 dB

50% subsampled shot from **regularly** missing shot positions

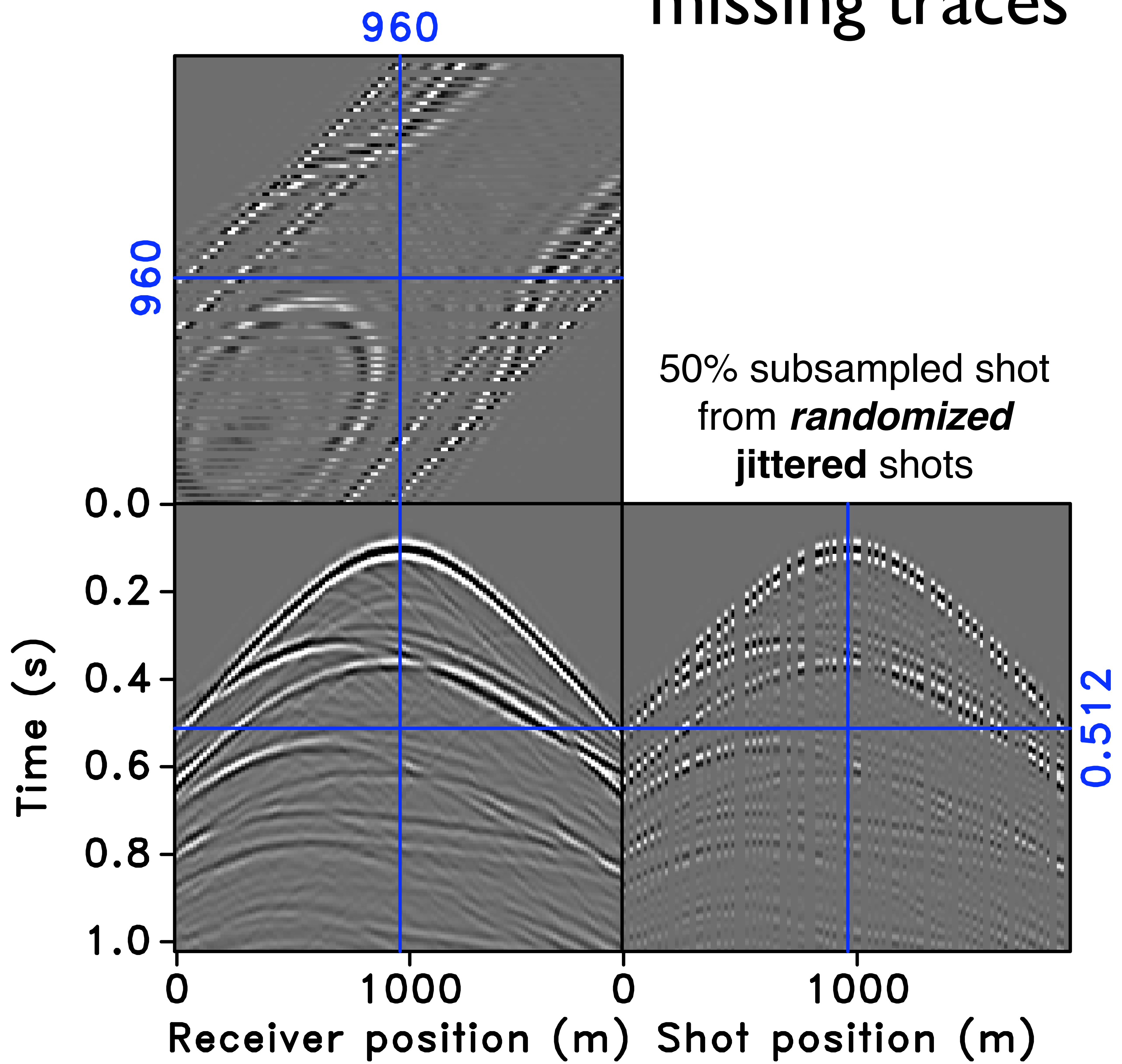
[Hennenfent & FJH, ’08]
[Gang et.al., ‘09]

SLIM

Jittered sampling

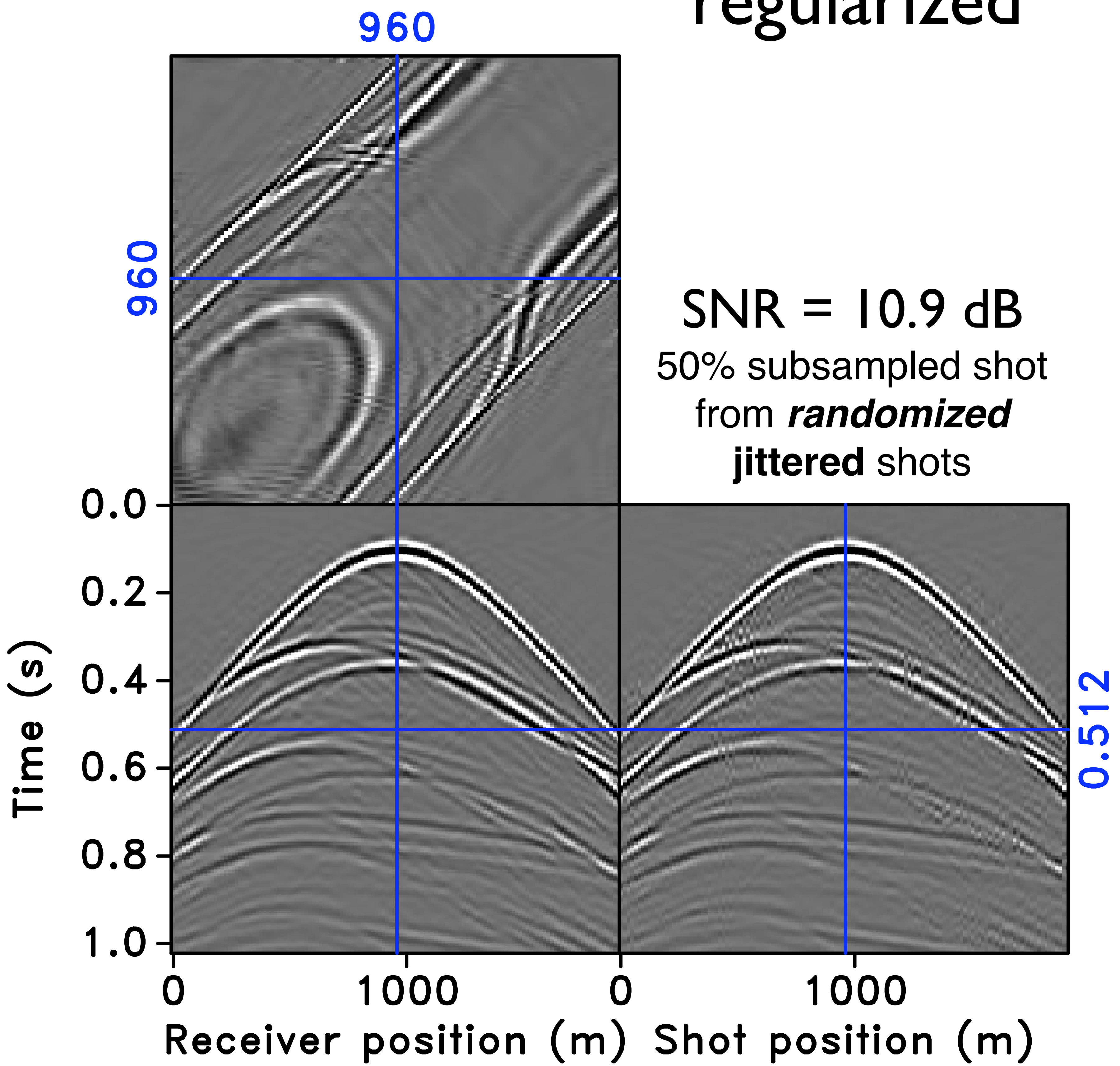


missing traces



50% subsampled shot from ***randomized*** **jittered** shots

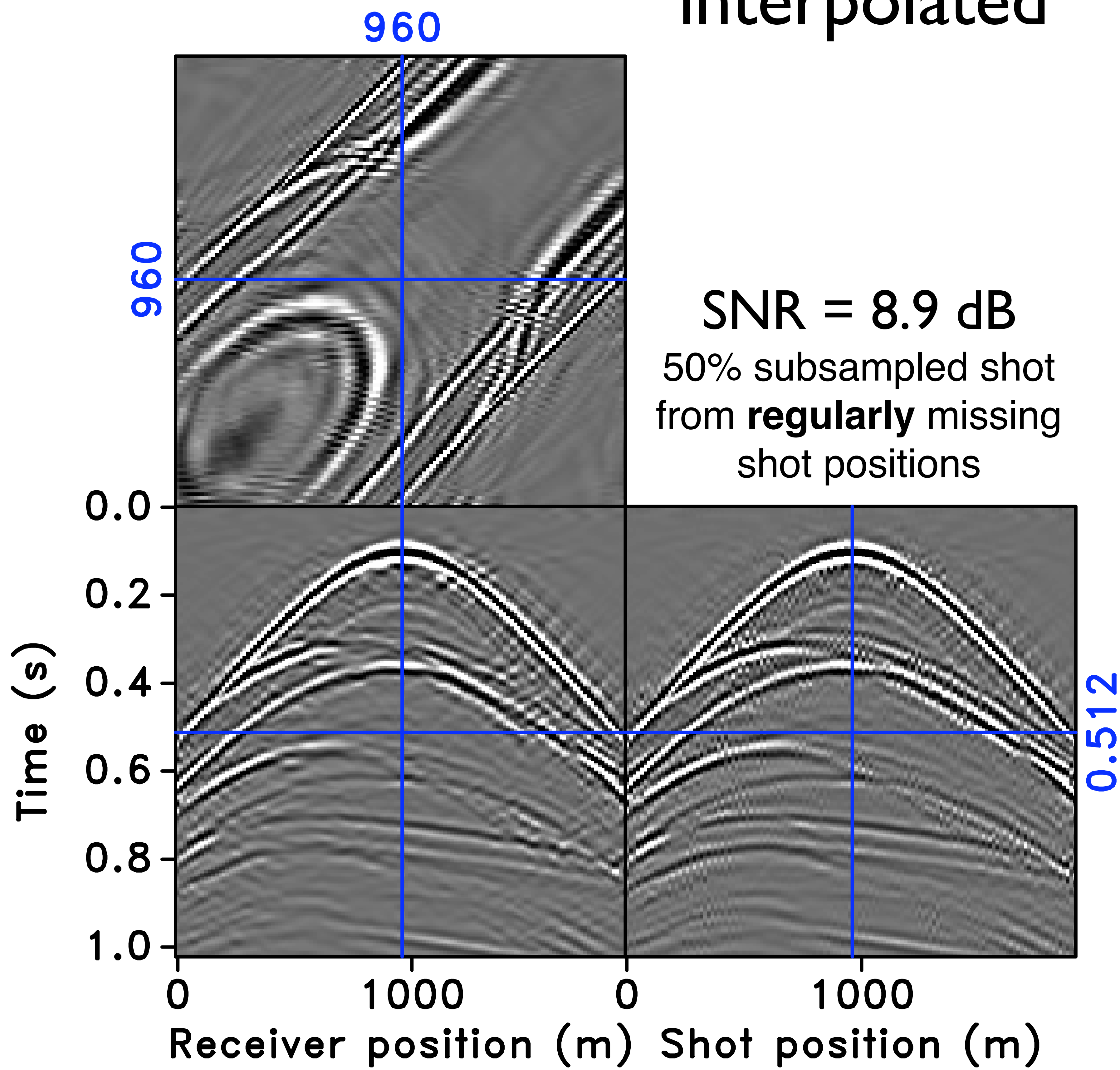
regularized



SNR = 10.9 dB

50% subsampled shot from ***randomized*** **jittered** shots

interpolated



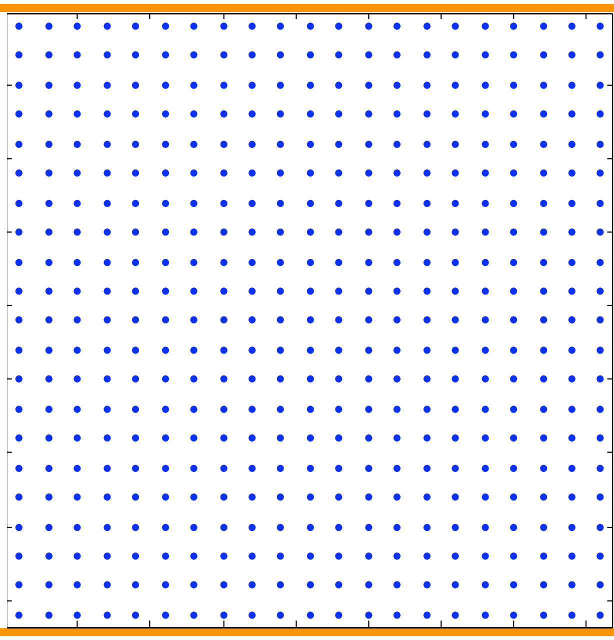
SNR = 8.9 dB

50% subsampled shot from **regularly** missing shot positions

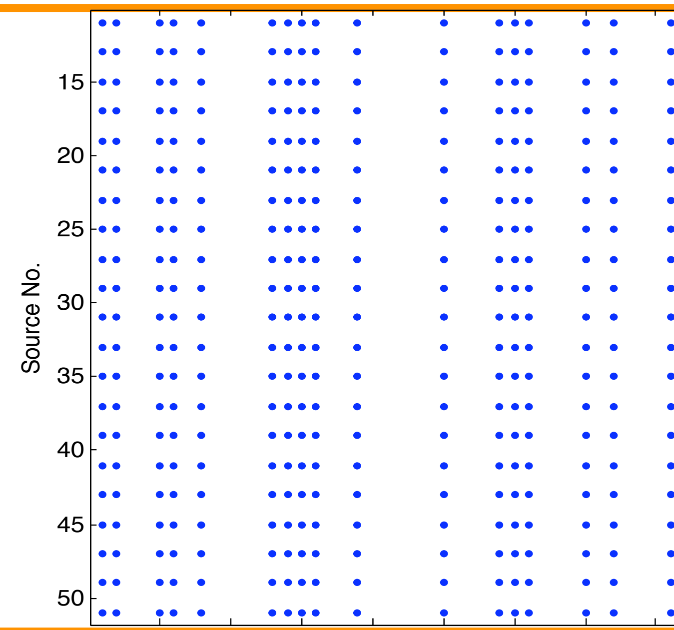
1 & 2-D jittered samplings

[Tang et. al., '09-'10]

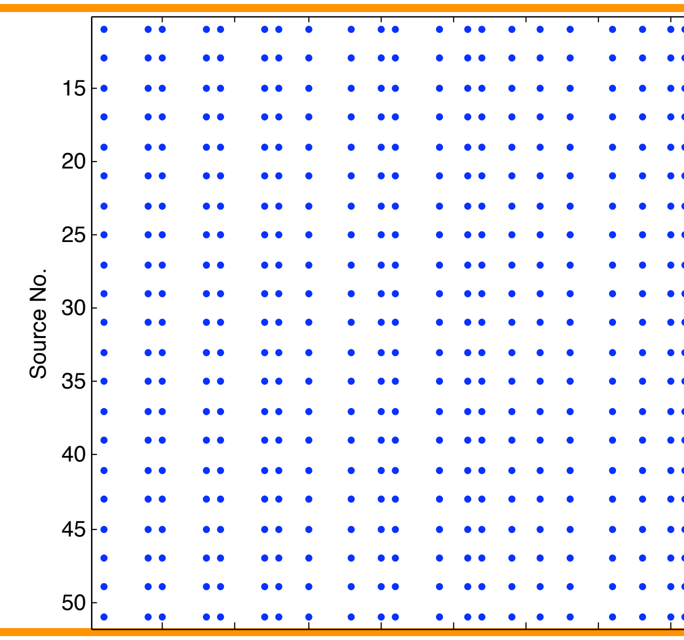
regular



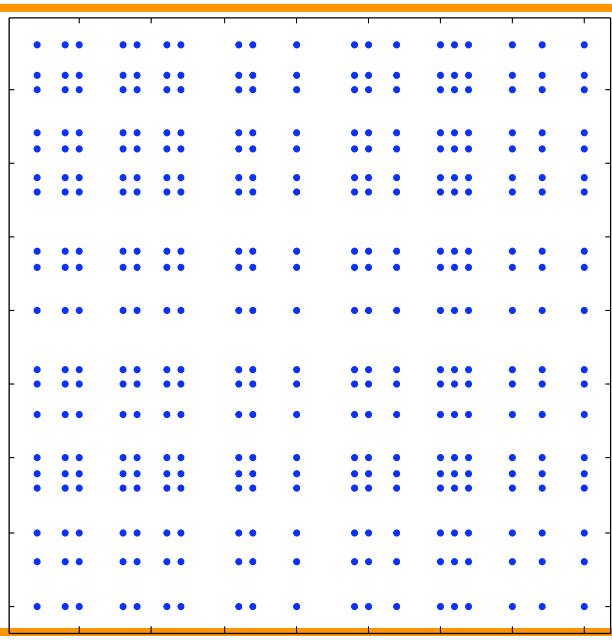
uniform



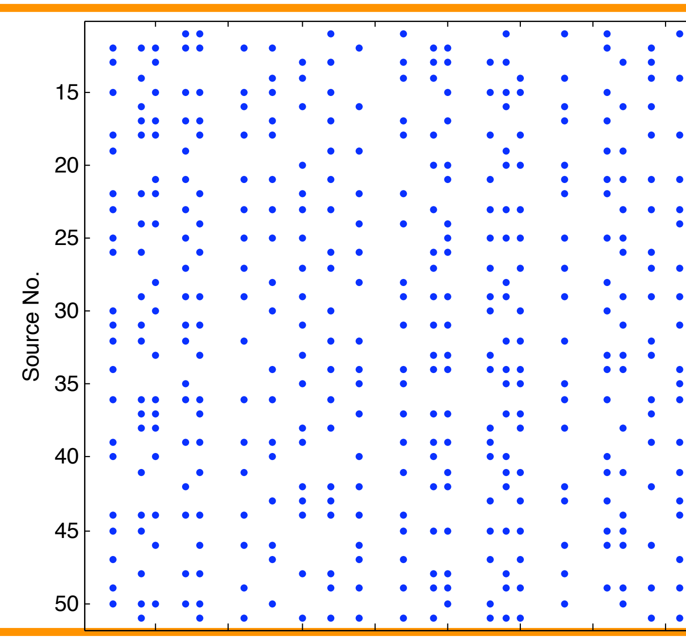
jittered



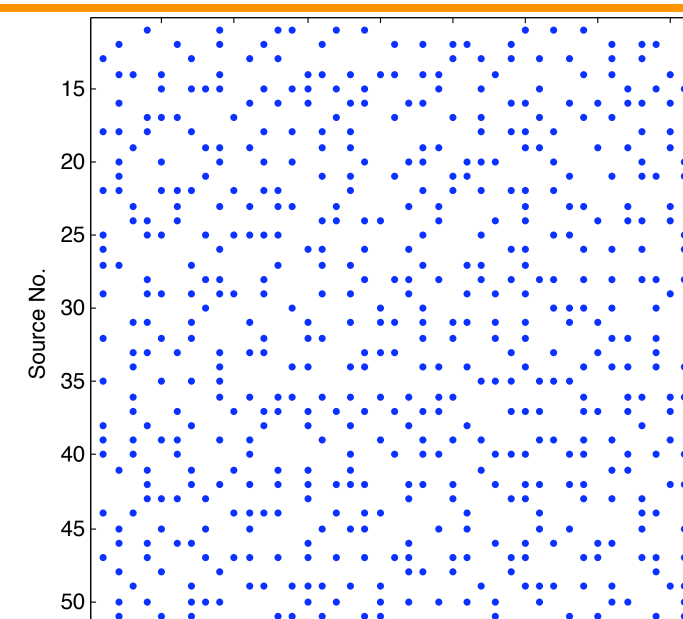
separable 2d jittered



non-seperable 2d jittered



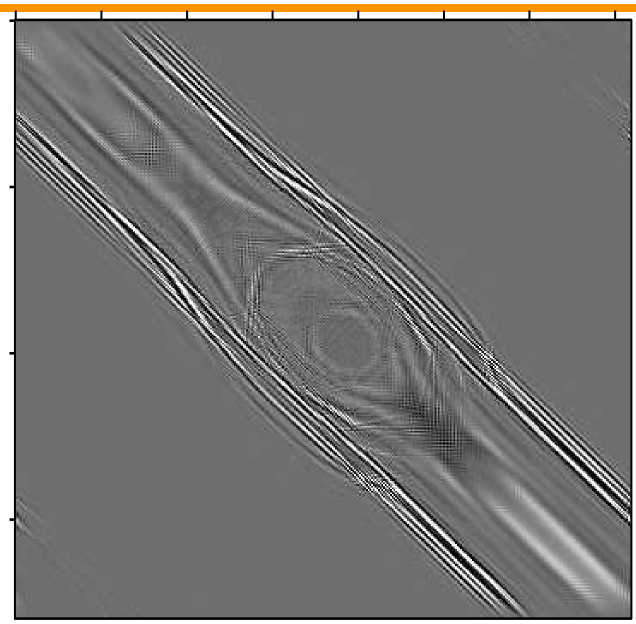
fully 2d jittered



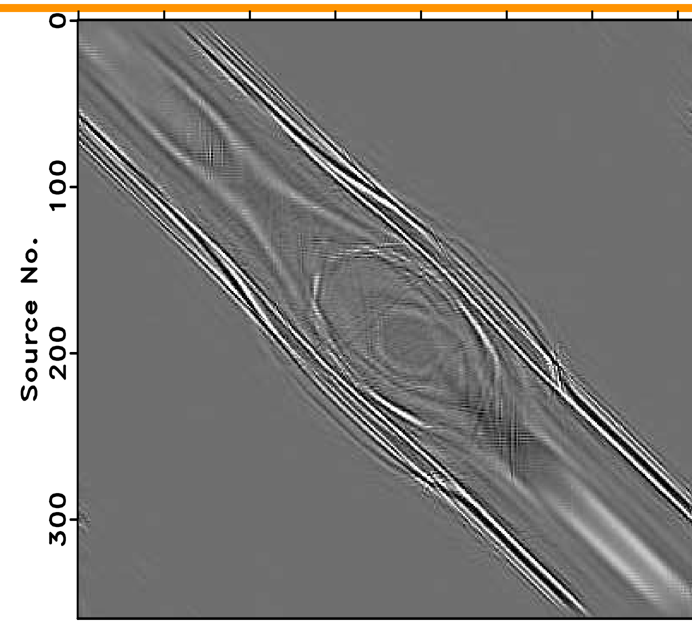
Spectra become increasingly "*blue*"

Recovery from 1-2D jittered samplings (25%)

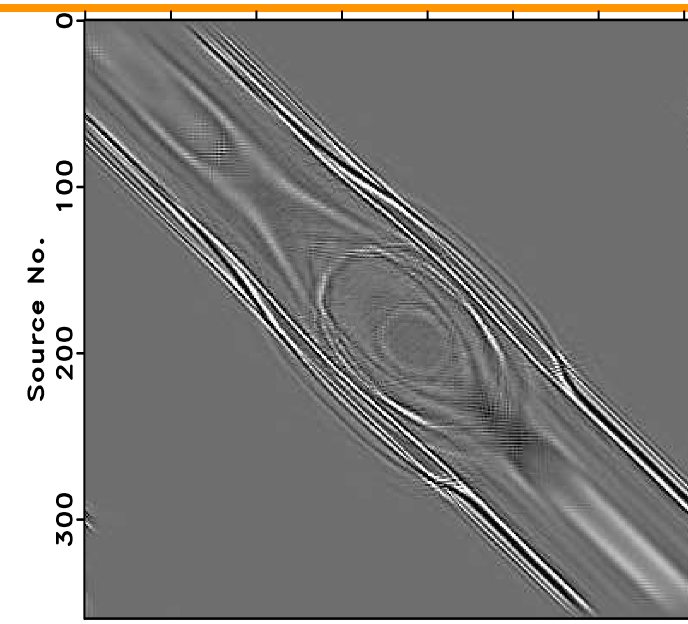
regular
(3.91 dB)



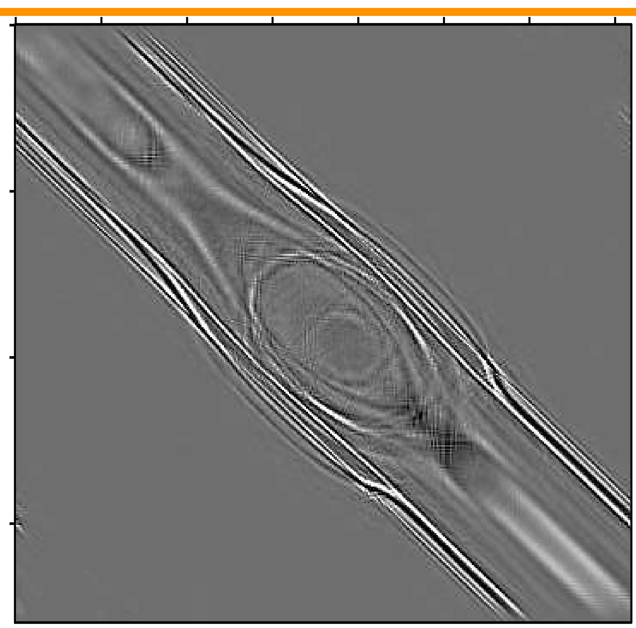
uniform
(7.20 dB)



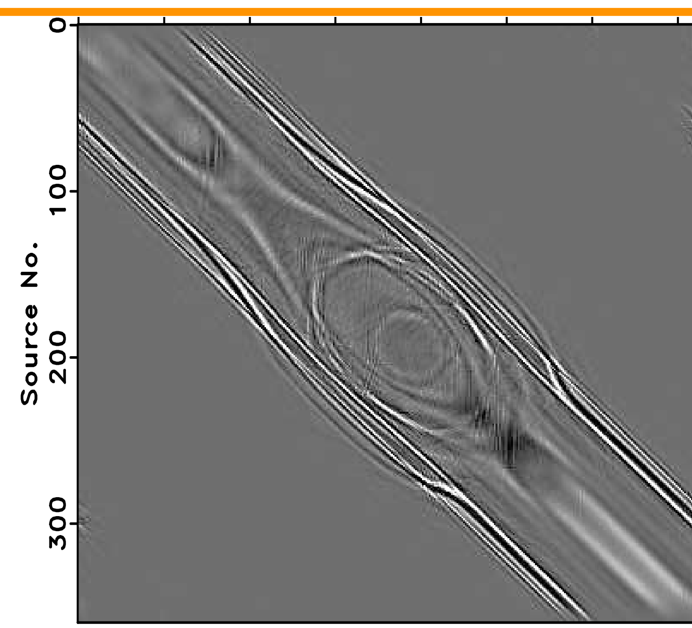
jittered
(8.94 dB)



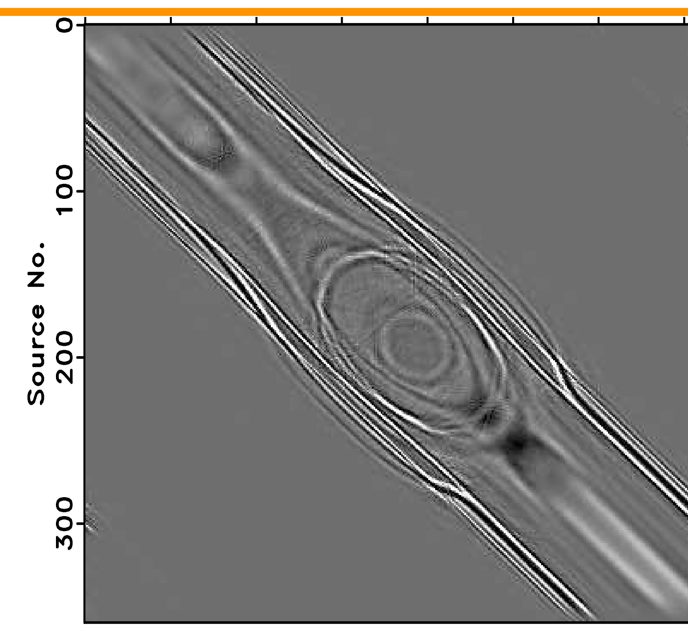
separable
2d jittered (9.45 dB)



non-seperable
2d jittered (10.03 dB)



fully 2d
jittered (10.86 dB)



Spectra become increasingly *"blue"*

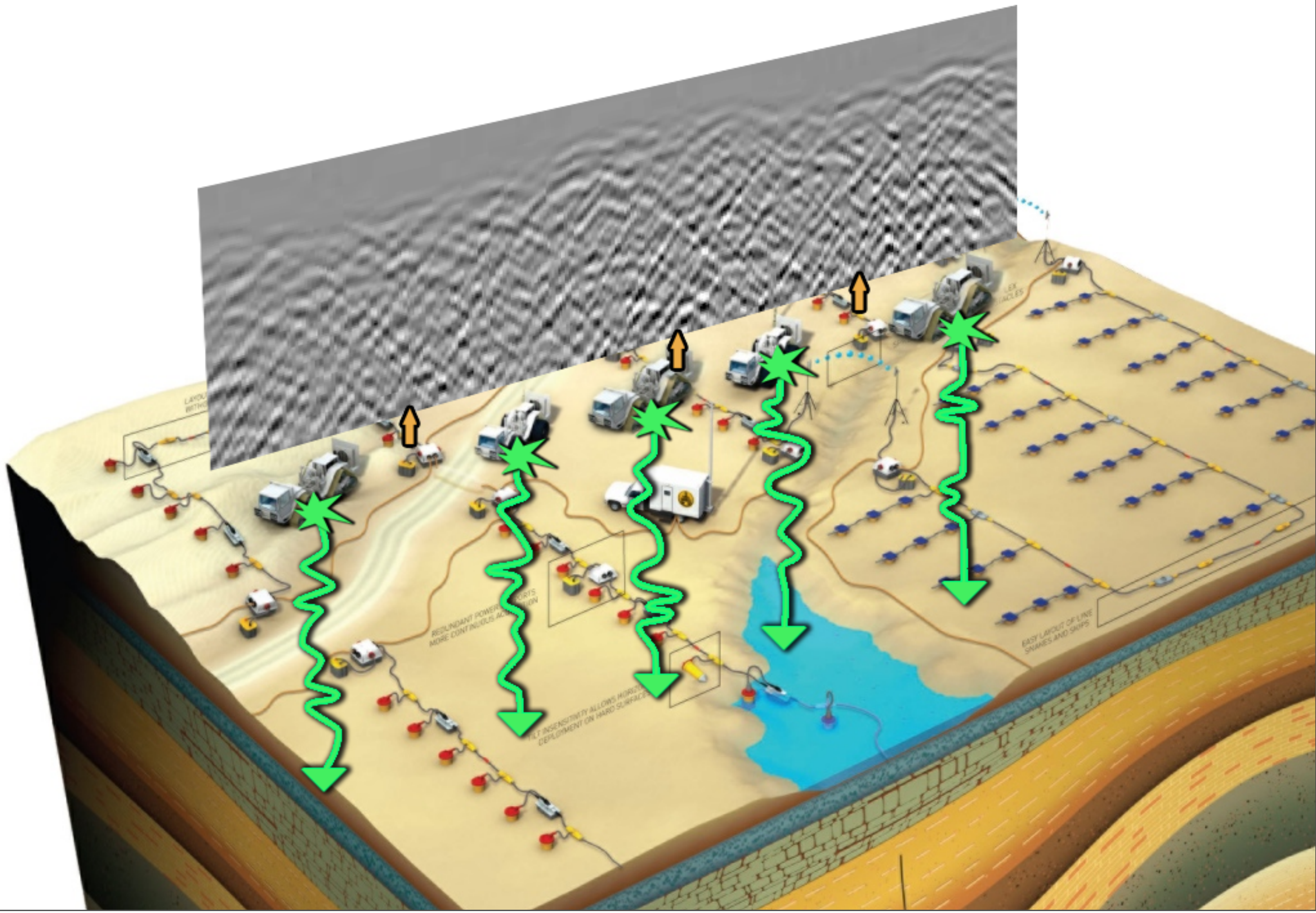
Case study II

[Beasley et. al., ’98]
[Berkhout ’08]
[Herrmann ’09-’10]

Acquisition design according to Compressive Sensing

- *Subsampling with randomized jittered* ***sequential*** sources vs *randomized phase-encoded* ***simultaneous*** sources

Simultaneous & incoherent sources



Unblending/ Demultiplexing

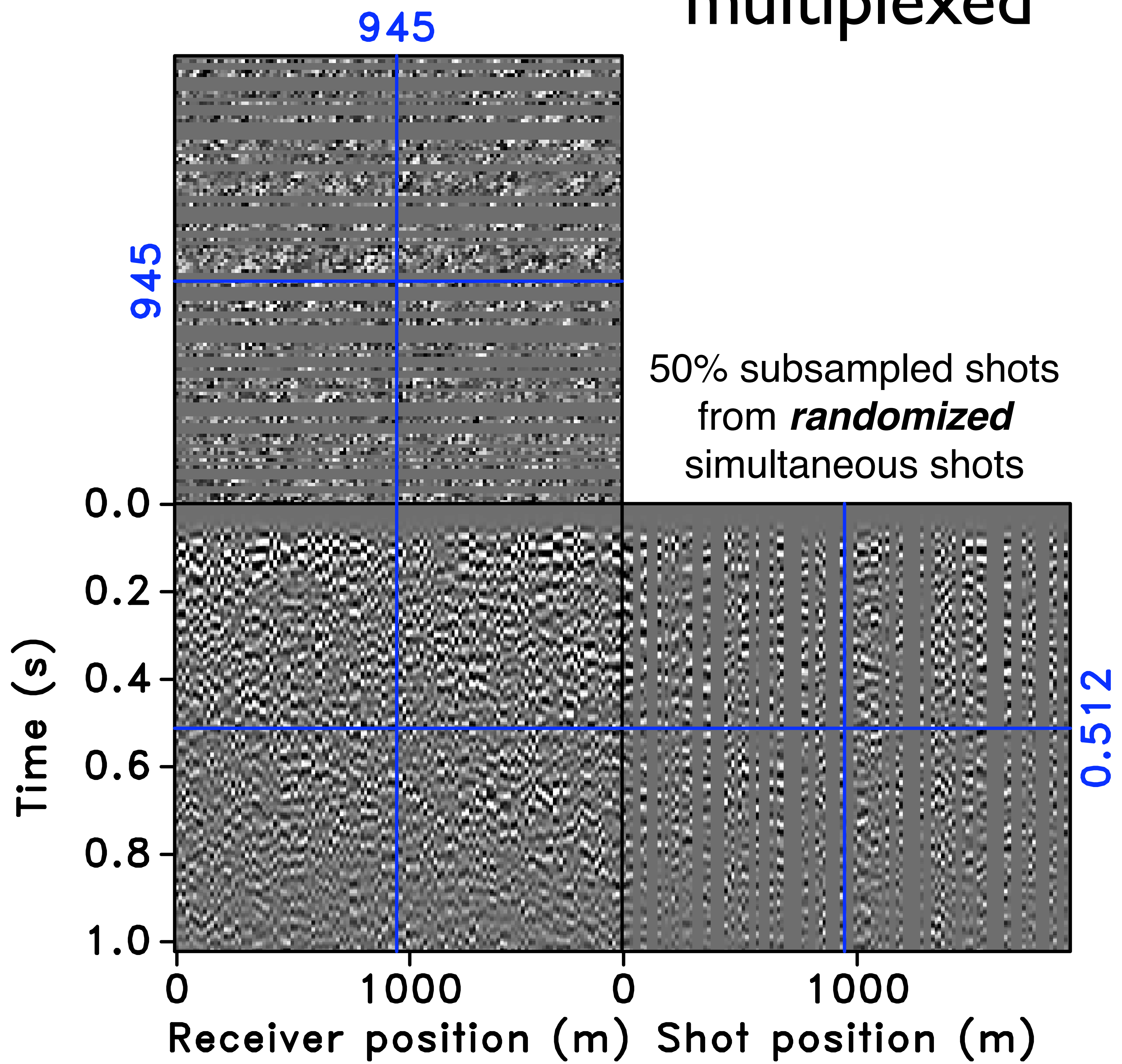
✓ →

? ←

$$$$$$$$$ $

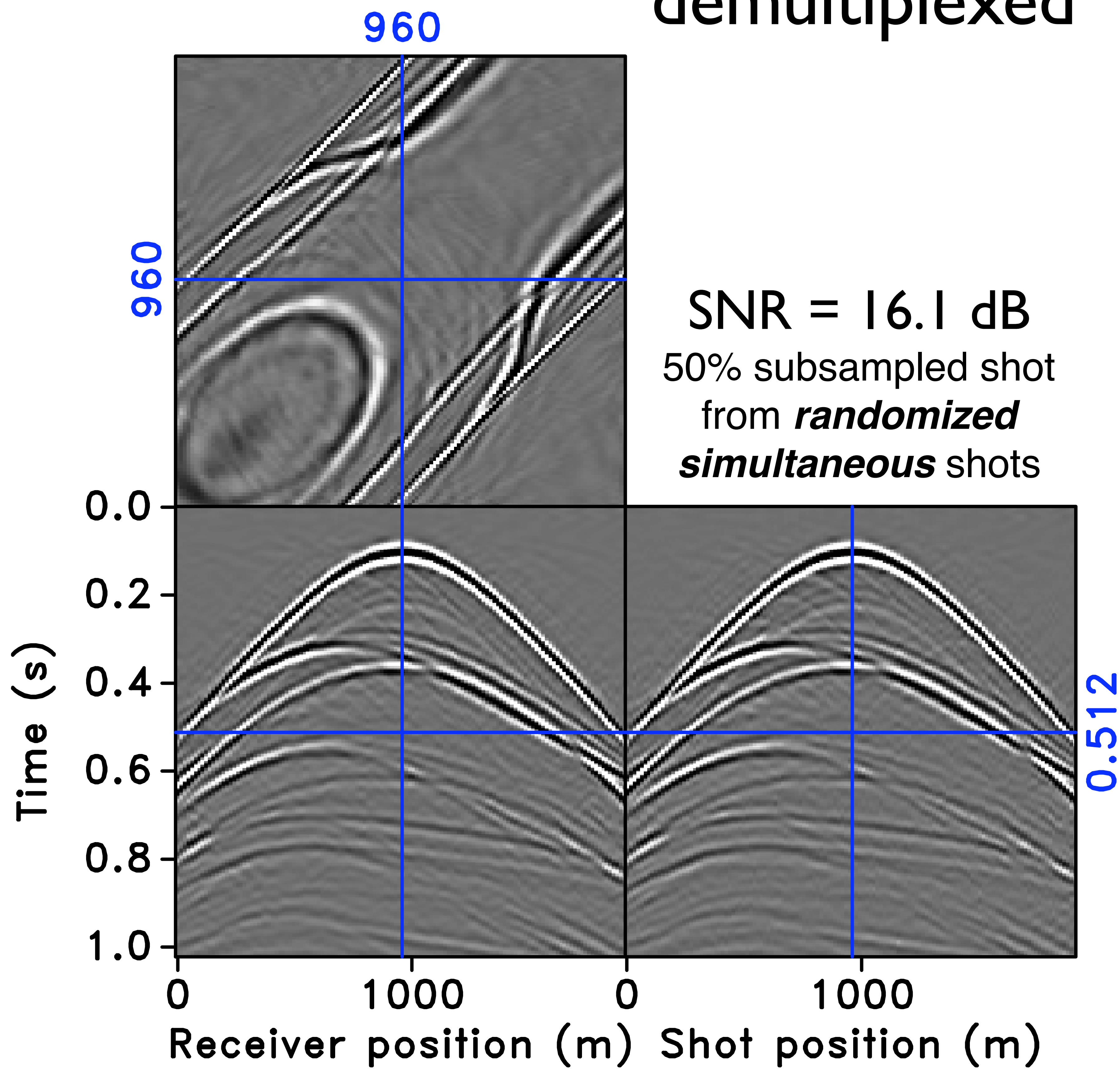
Blending versus *unblending* ...

multiplexed



50% subsampled shots from ***randomized*** simultaneous shots

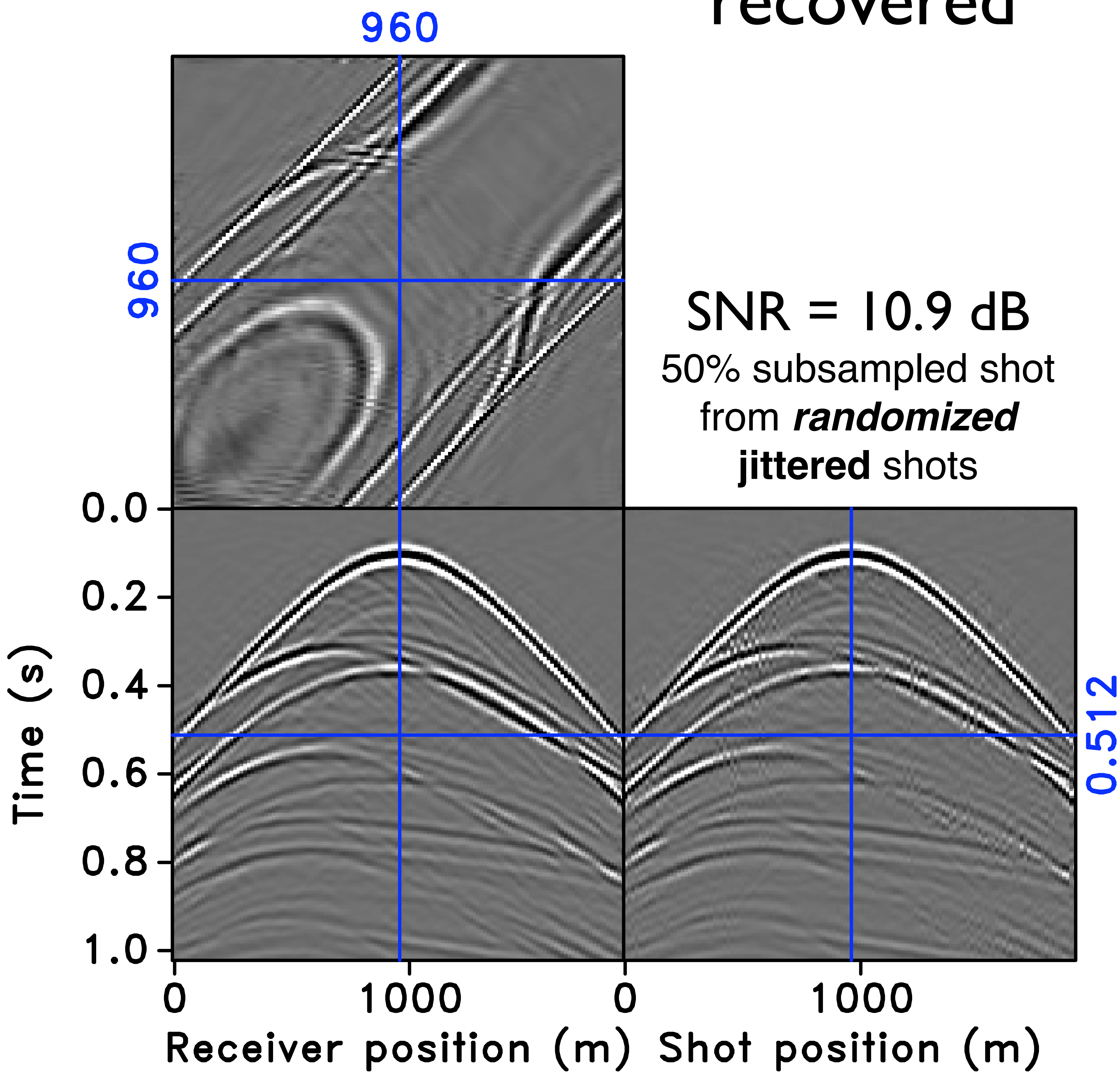
demultiplexed



SNR = 16.1 dB

50% subsampled shot from ***randomized simultaneous*** shots

recovered



SNR = 10.9 dB

50% subsampled shot from ***randomized*** **jittered** shots

Recent results

Recovery of seismic lines based

- on "separable" *sparsifying* transform

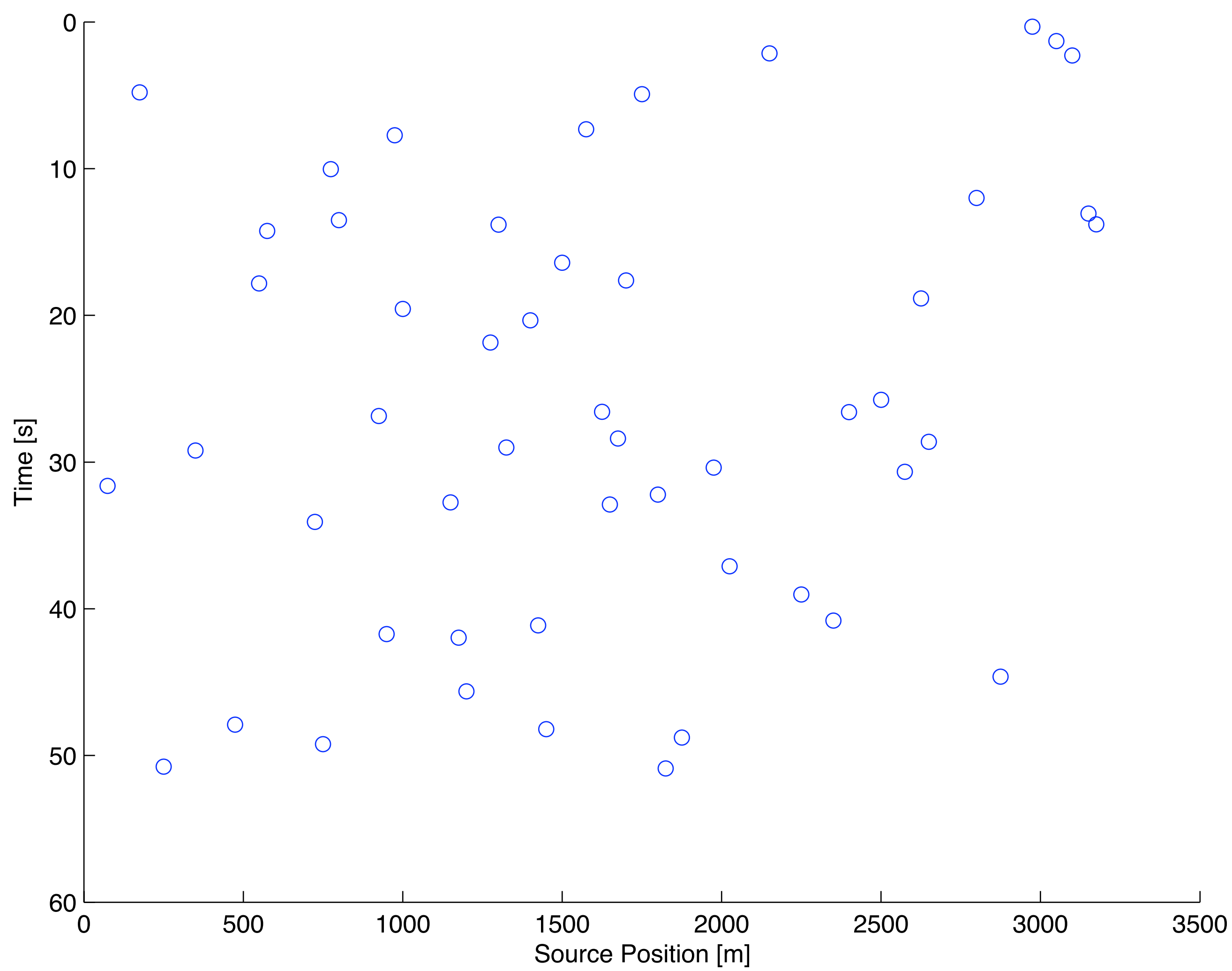
$$\mathbf{S} = \mathbf{C} \otimes \mathbf{W}$$

- *favorable* simultaneous acquisition

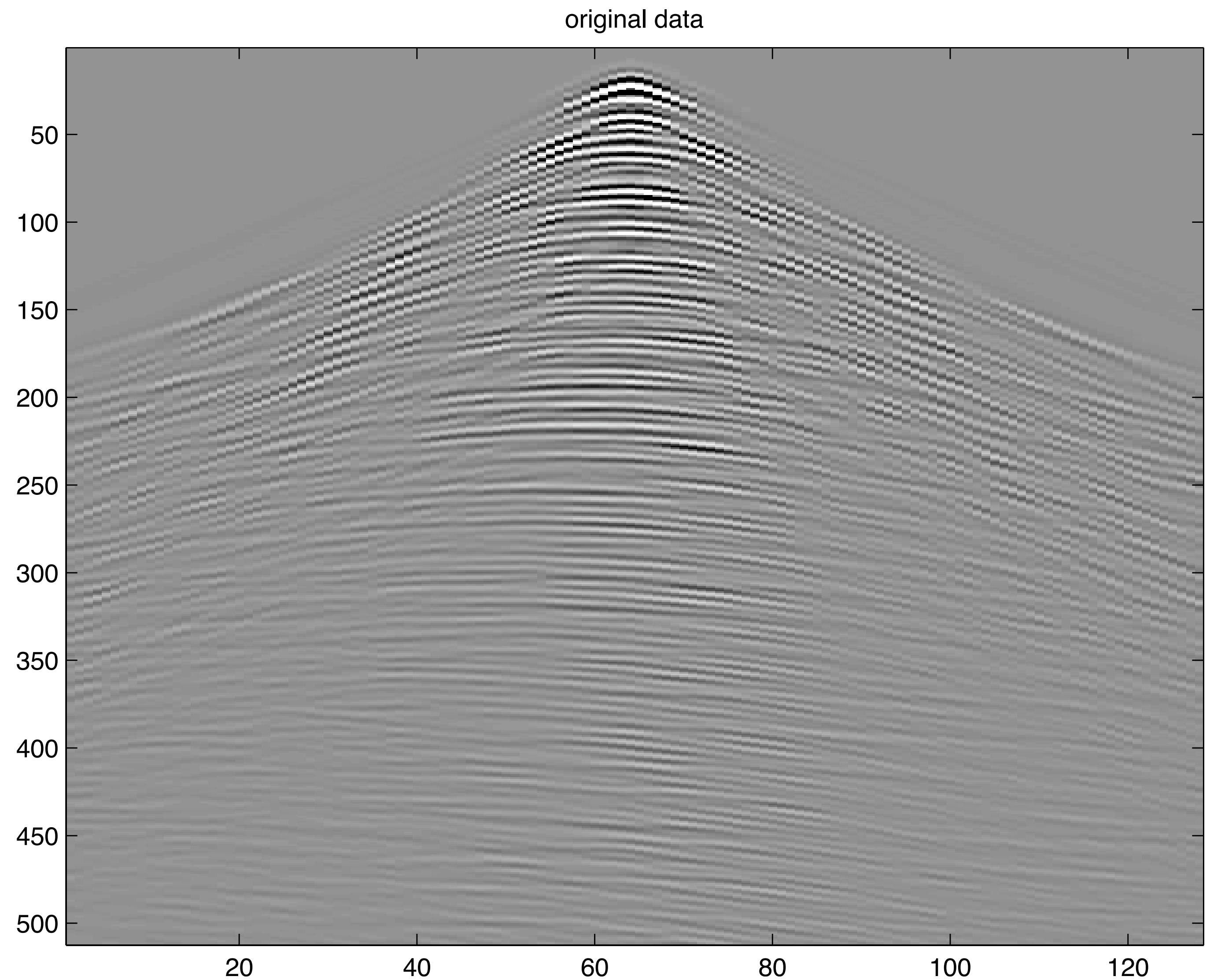
Consider "Marine" case

Simultaneous sources

Marine case

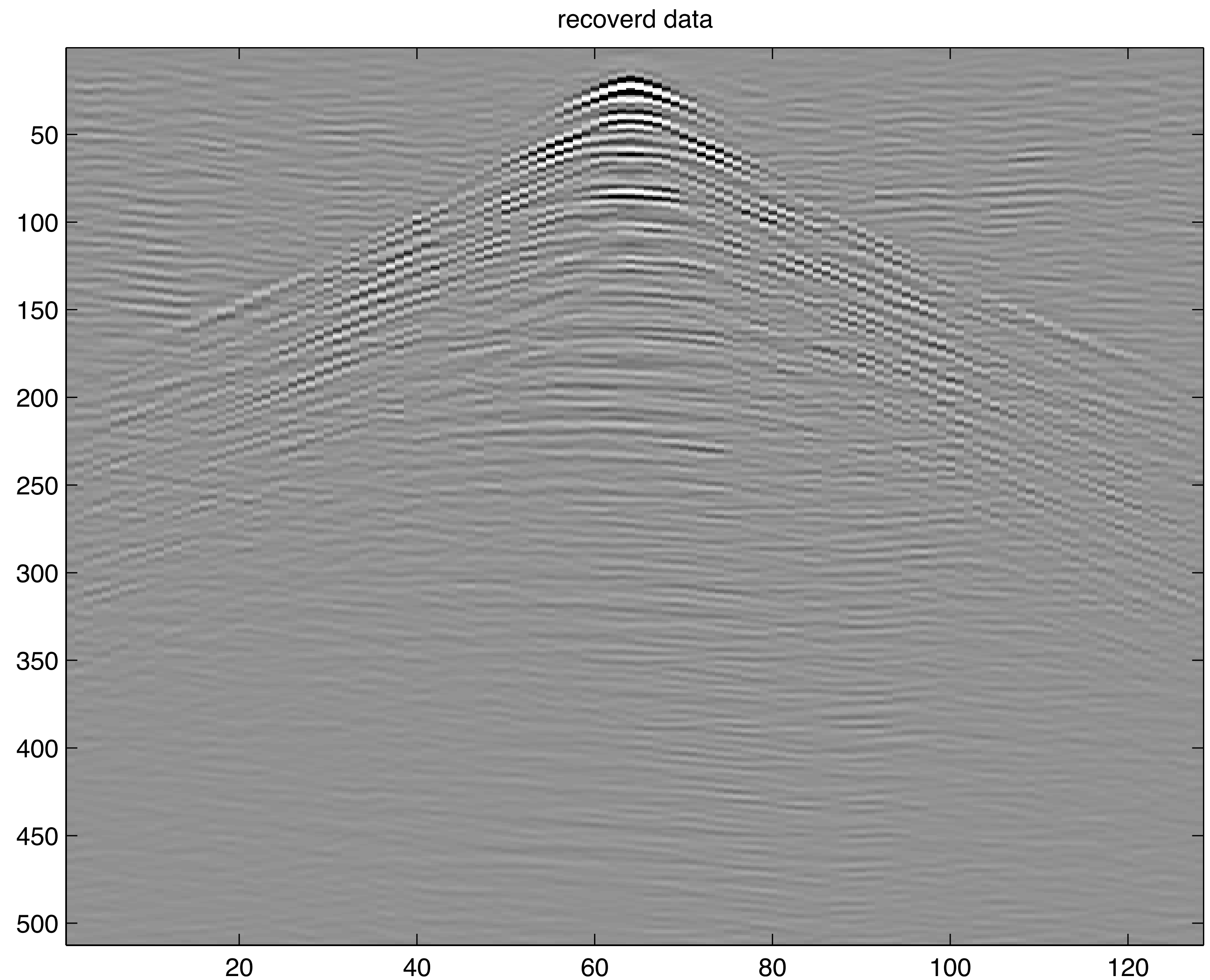


Original data



Recovered data

40 % of shots in 20 % of recording time



The math of Compressive Sensing [Candès et.al, '06]

Recovery is *possible* & *stable* as long as each *subset S* of *k* columns of $\mathbf{A} \in \mathbb{R}^{n \times N}$ with $k \leq N$ the # of nonzeros *approximately* behaves as an *orthogonal* basis.

In that case, we have

$$(1 - \hat{\delta}_k) \|\mathbf{x}_S\|_{\ell_2}^2 \leq \|\mathbf{A}_S \mathbf{x}_S\|_{\ell_2}^2 \leq (1 + \hat{\delta}_k) \|\mathbf{x}_S\|_{\ell_2}^2,$$

where *S* runs over all sets with cardinality $\leq k$

- the smaller the *restricted isometry constant* (RIP) $\hat{\delta}_k$ the more *energy* is captured and the more *stable* the *inversion* of **A**
- determined by the *mutual coherence* of the cols in **A**

Math cont'ed

[Candès et.al, '06]

RIP constant is bounded by

$$\hat{\delta}_k \leq (k-1)\mu$$

where

$$\mu = \max_{1 \leq i \neq j \leq N} |\mathbf{a}_i^H \mathbf{a}_j|$$

Matrices with small $\hat{\delta}_k$ contain subsets of *k* *incoherent* columns.

Gaussian random matrices with *i.i.d.* entries have this property.

One-norm solvers recover $\mathbf{x}_0$ as long it is *k* sparse and

$$k \leq C \cdot \frac{n}{\log_2(N/n)},$$

yields an *oversampling ratio* of

$$\boxed{n/k \approx C \cdot \log_2 N}$$

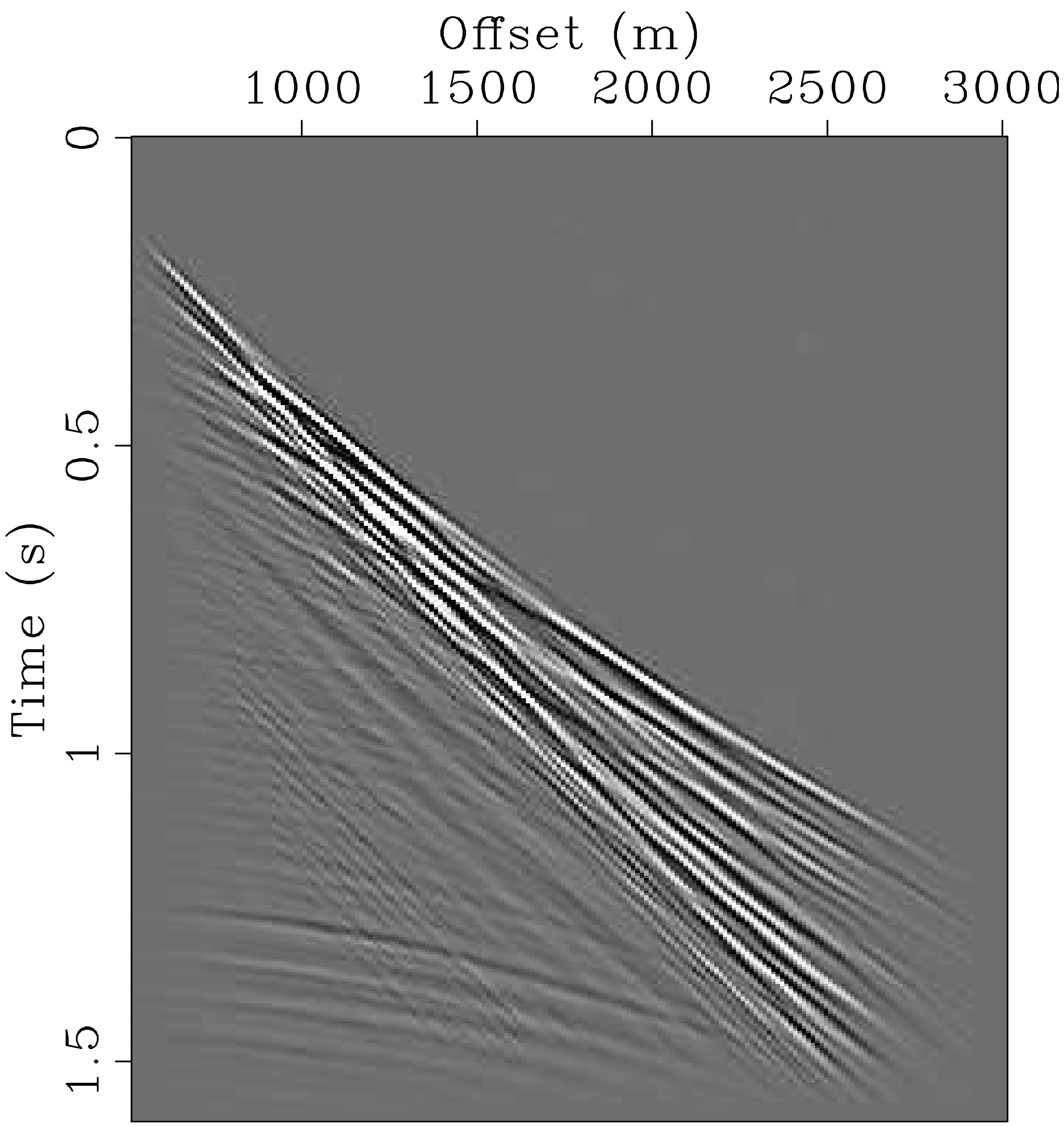
Key elements

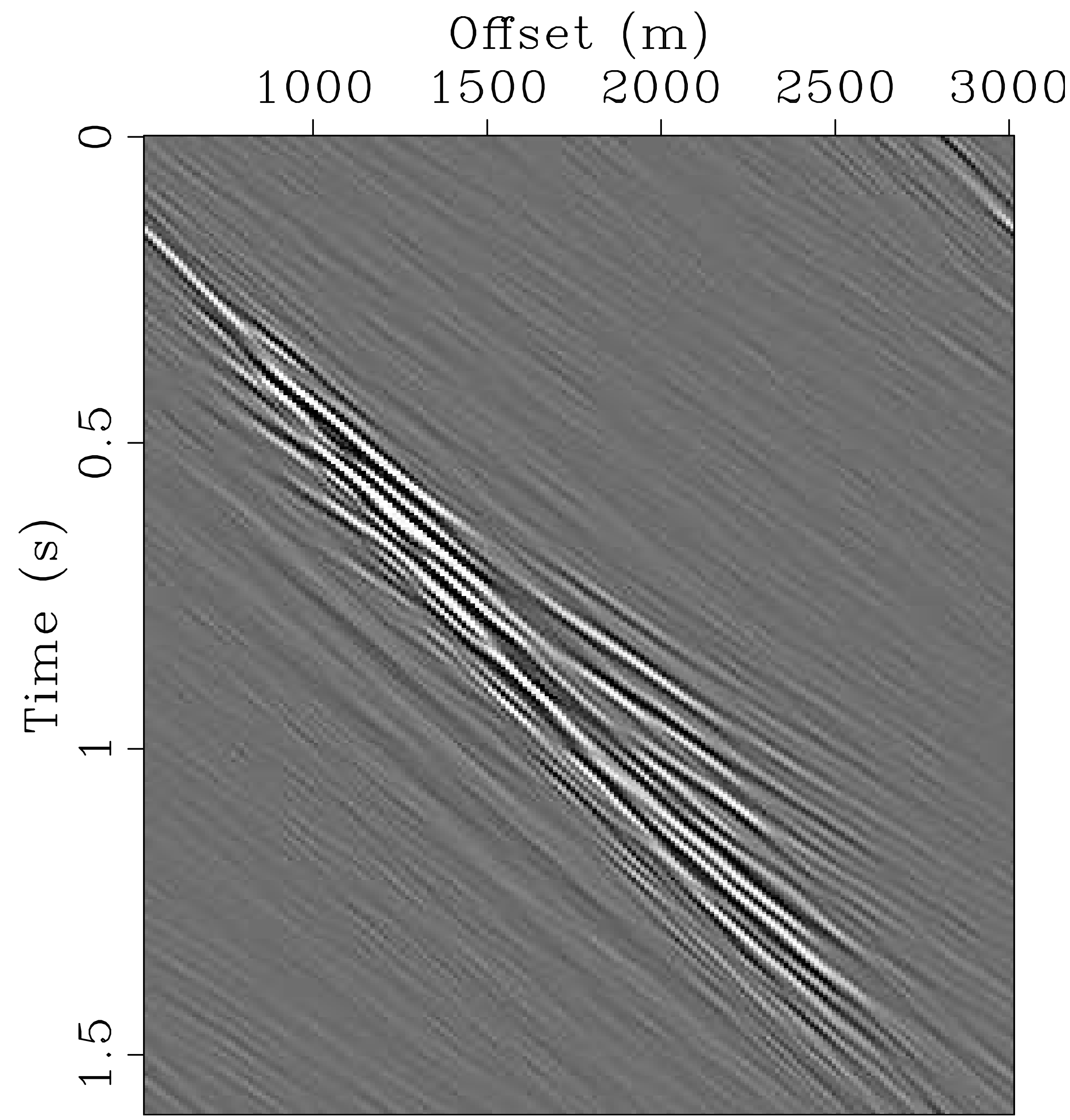
- *sparsifying transform*
  - typically **localized** in the time-space domain to handle the complexity of seismic data

- *advantageous coarse randomized sampling*
  - generates incoherent random undersampling "noise" in the sparsifying domain

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

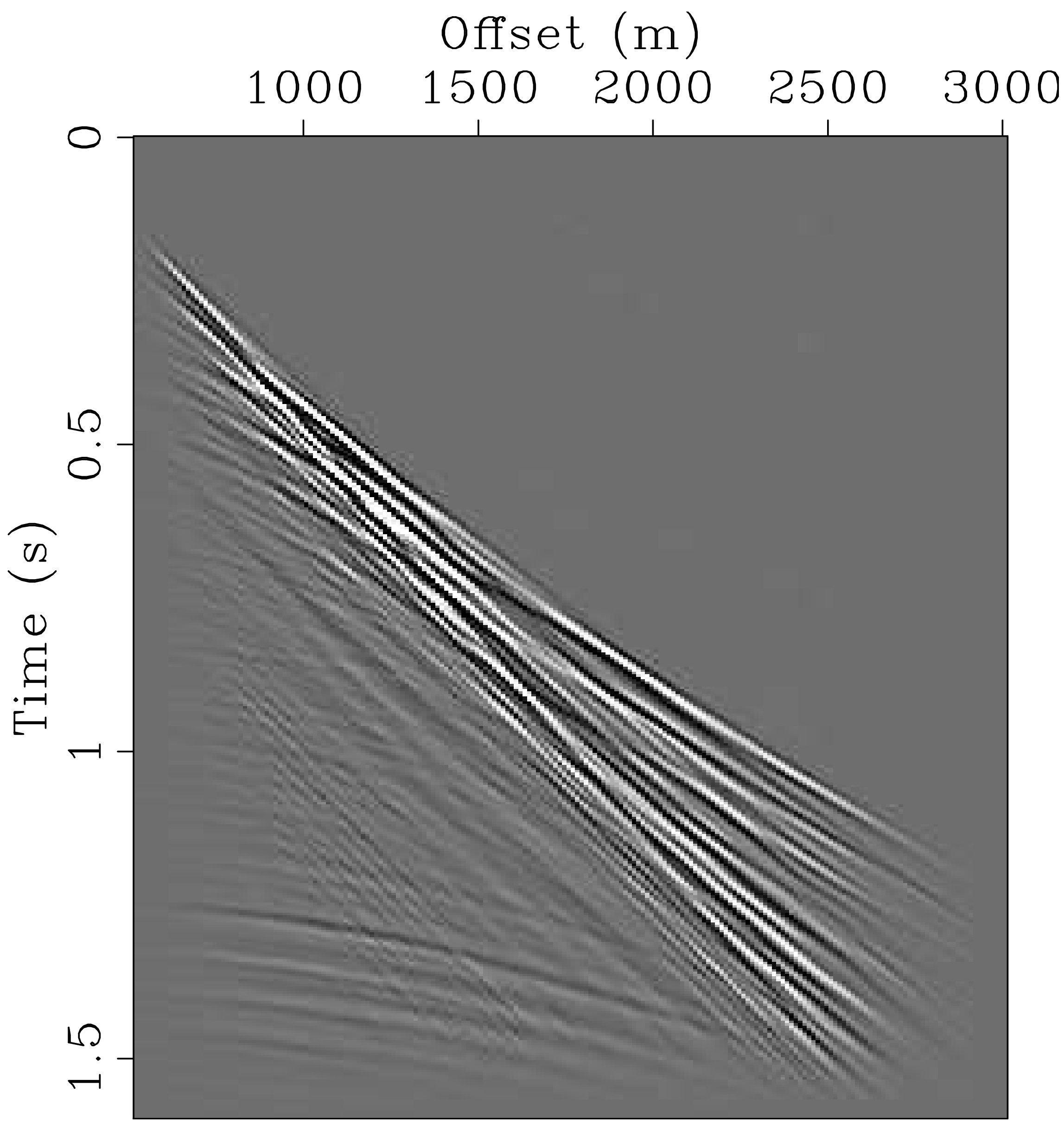
Fourier reconstruction

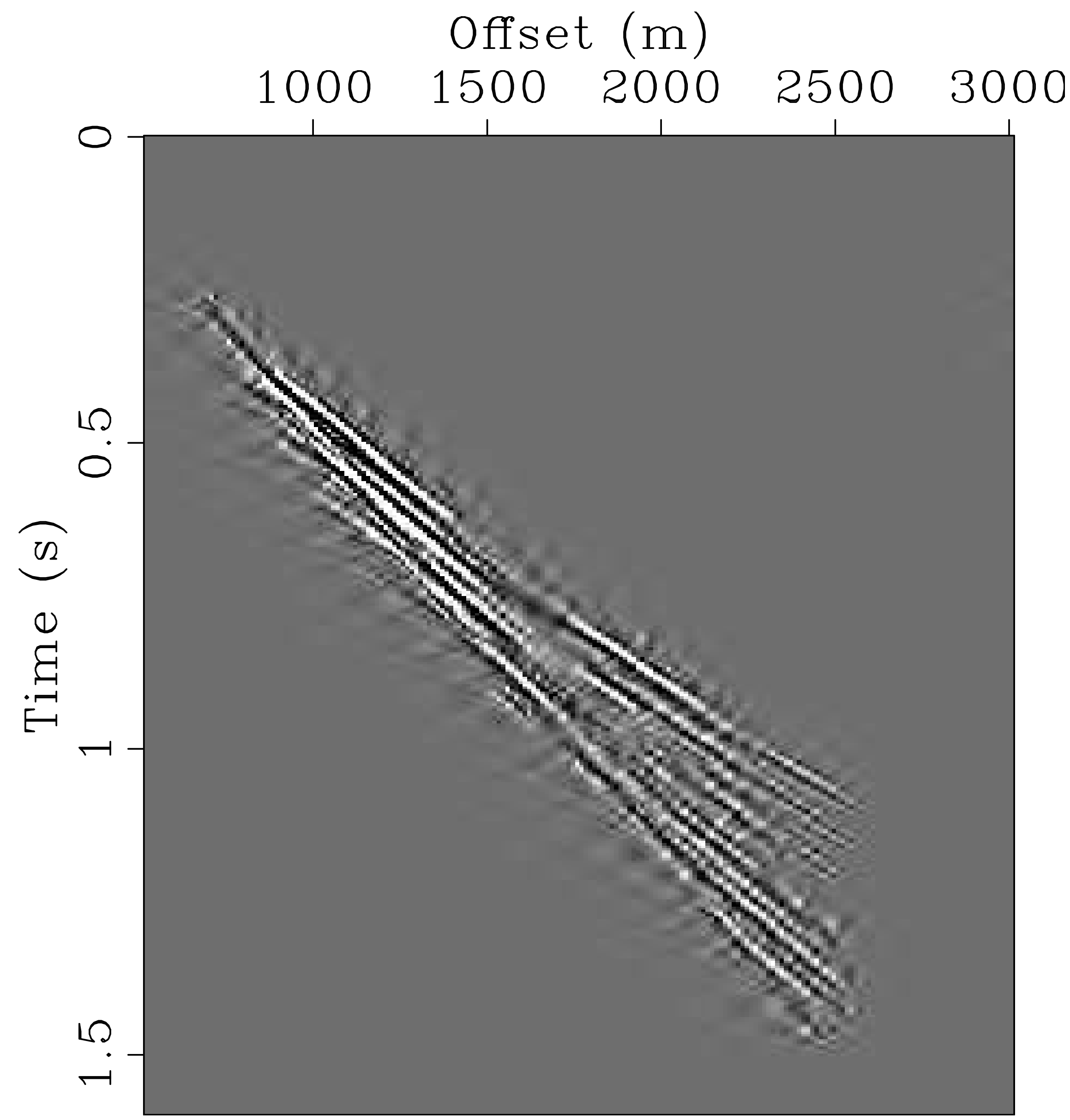




1 % of coefficients

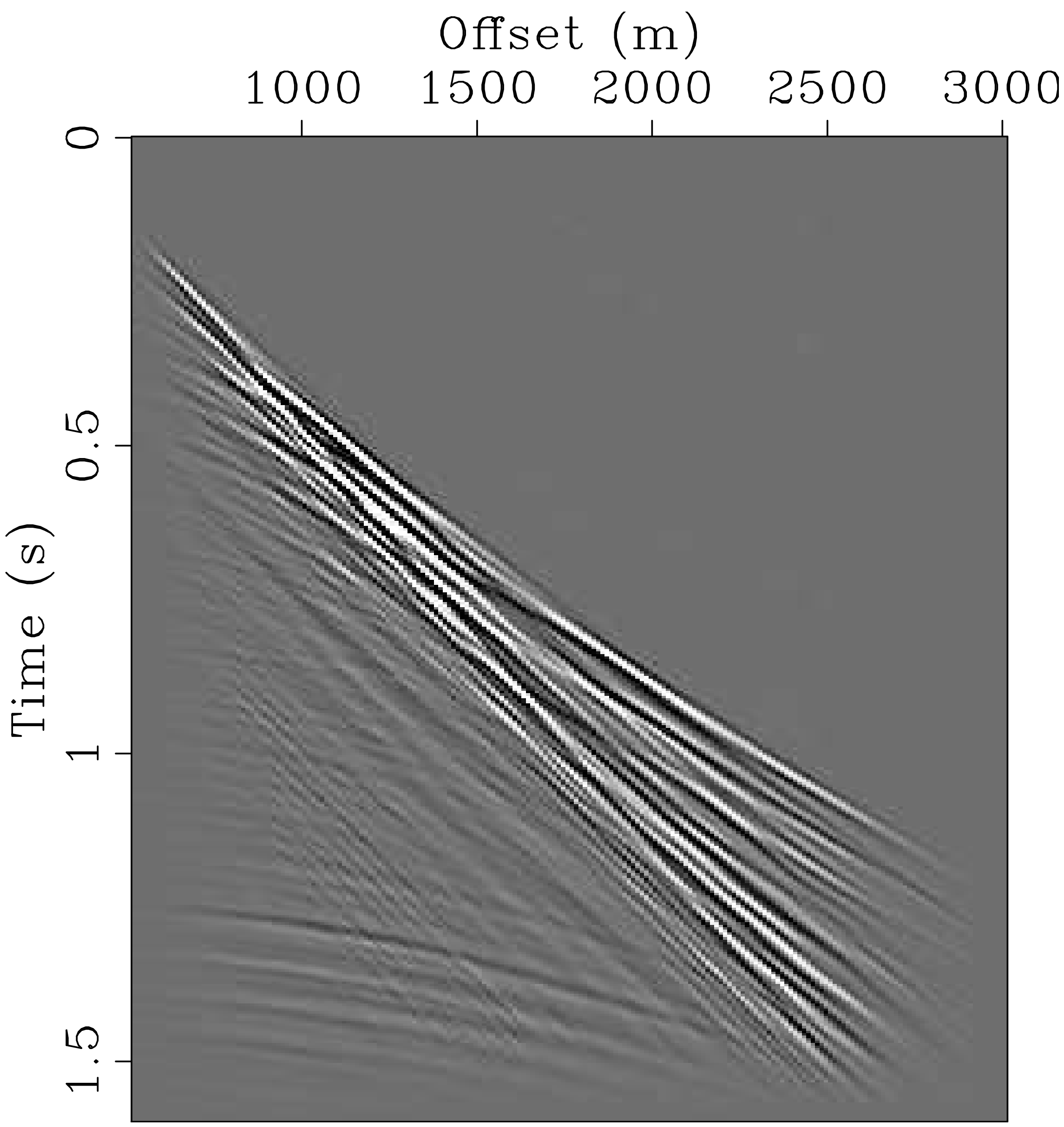
Wavelet reconstruction

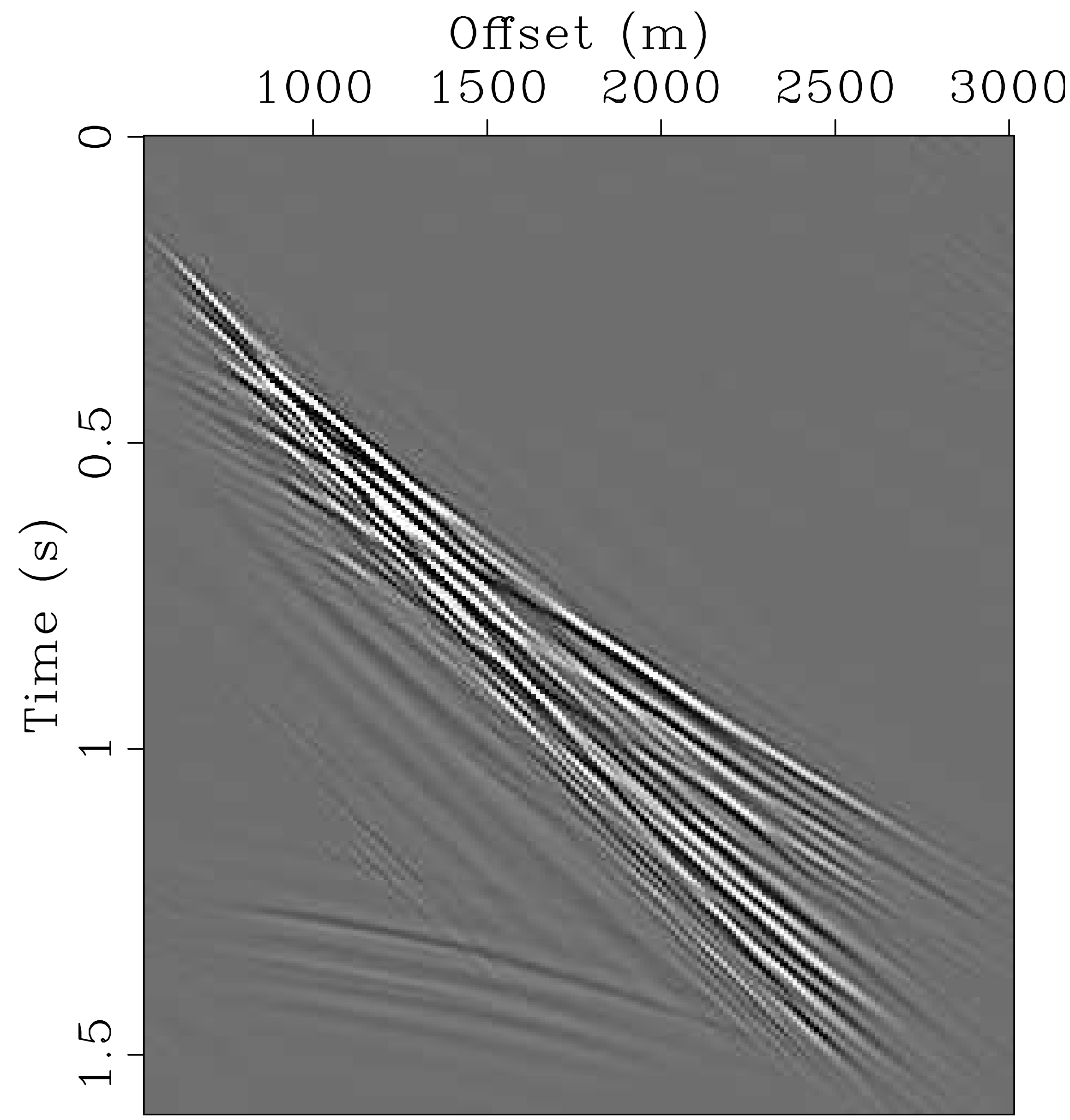




1 % of coefficients

Curvelet reconstruction



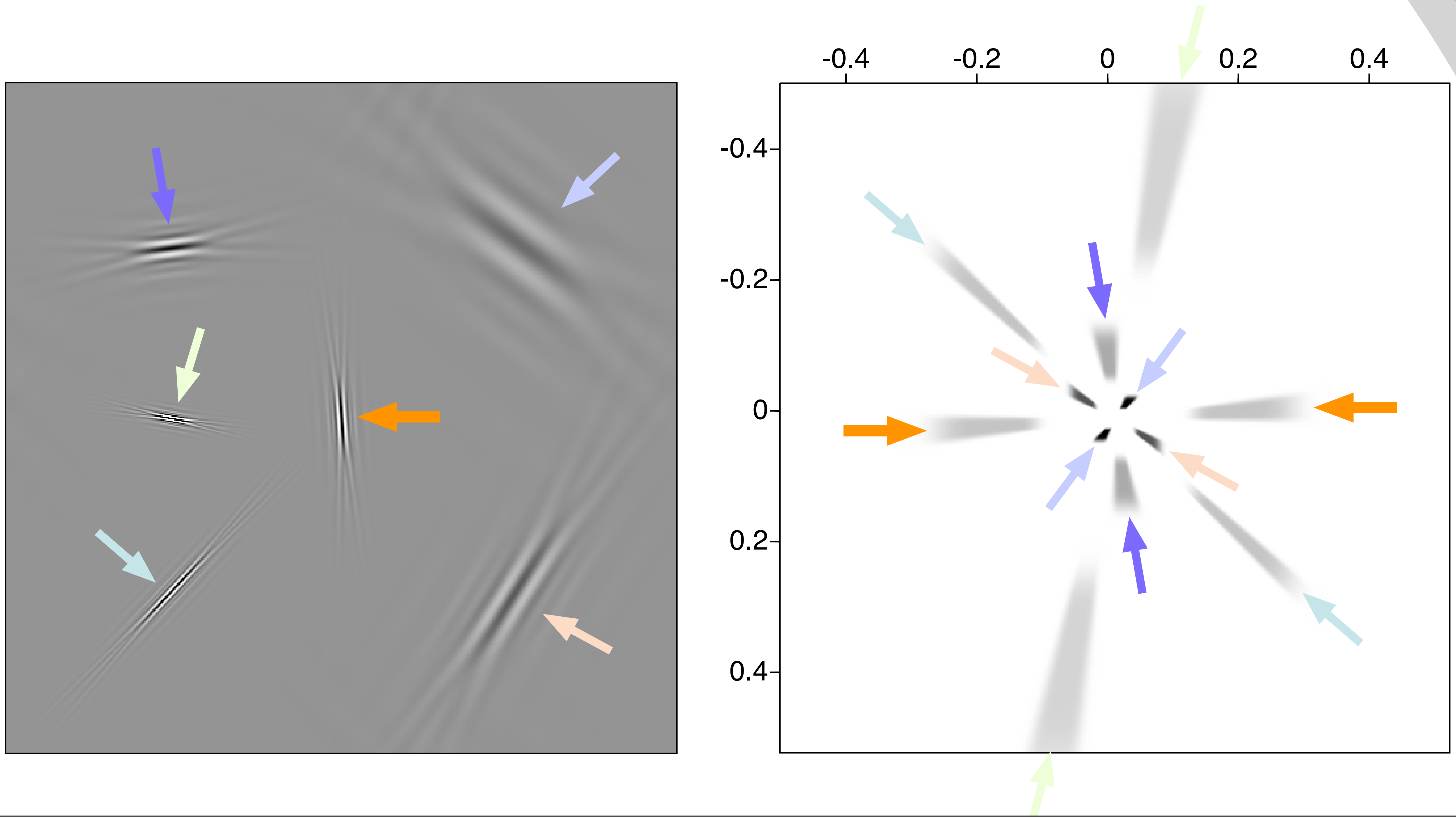


1 % of coefficients

SLIM

[Demanet et. al., '06]

Curvelets



Extension

Extend CS framework:

$$\mathbf{A} := \mathbf{RMS}^H$$

restriction matrix — measurement matrix — sparsity matrix

Expected to perform well when

$$\mu = \max_{1 \leq i \neq j \leq N} \left| \left(\mathbf{RMs}^i\right)^H \mathbf{RMs}^j \right|$$

Generalizes to *redundant* transforms for cases where

- max of RIP constants for **M**, **S** are small [Rauhut et.al, '06]
- or $\mathbf{SS}^H\mathbf{x}$ remains sparse for **x** sparse [Candès et.al, '10]

Open research topic...

Empirical performance analysis

Selection of the appropriate sparsifying transform

- nonlinear approximation error

$$\mathrm{SNR}(\rho) = -20 \log \frac{\|\mathbf{f} - \mathbf{f}_\rho\|}{\|\mathbf{f}\|} \quad \text{with} \quad \rho = k/P$$

- recovery error

$$\mathrm{SNR}(\delta) = -20 \log \frac{\|\mathbf{f} - \tilde{\mathbf{f}}_\delta\|}{\|\mathbf{f}\|} \quad \text{with} \quad \delta = n/N$$

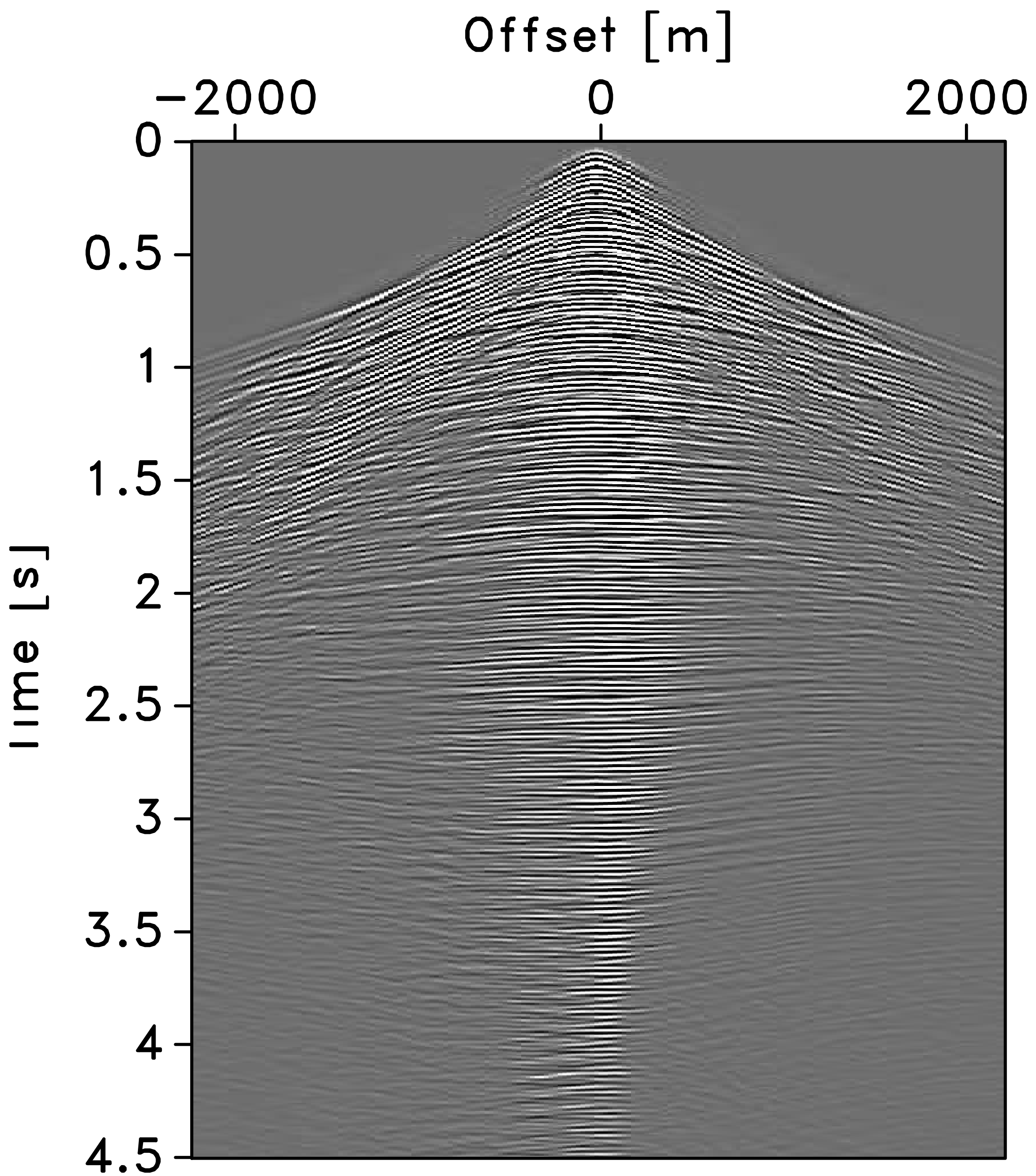
- oversampling ratio

$$\delta/\rho \quad \text{with} \quad \rho = \inf\{\tilde{\rho} : \quad \overline{\mathrm{SNR}}(\delta) \leq \mathrm{SNR}(\tilde{\rho})\}$$

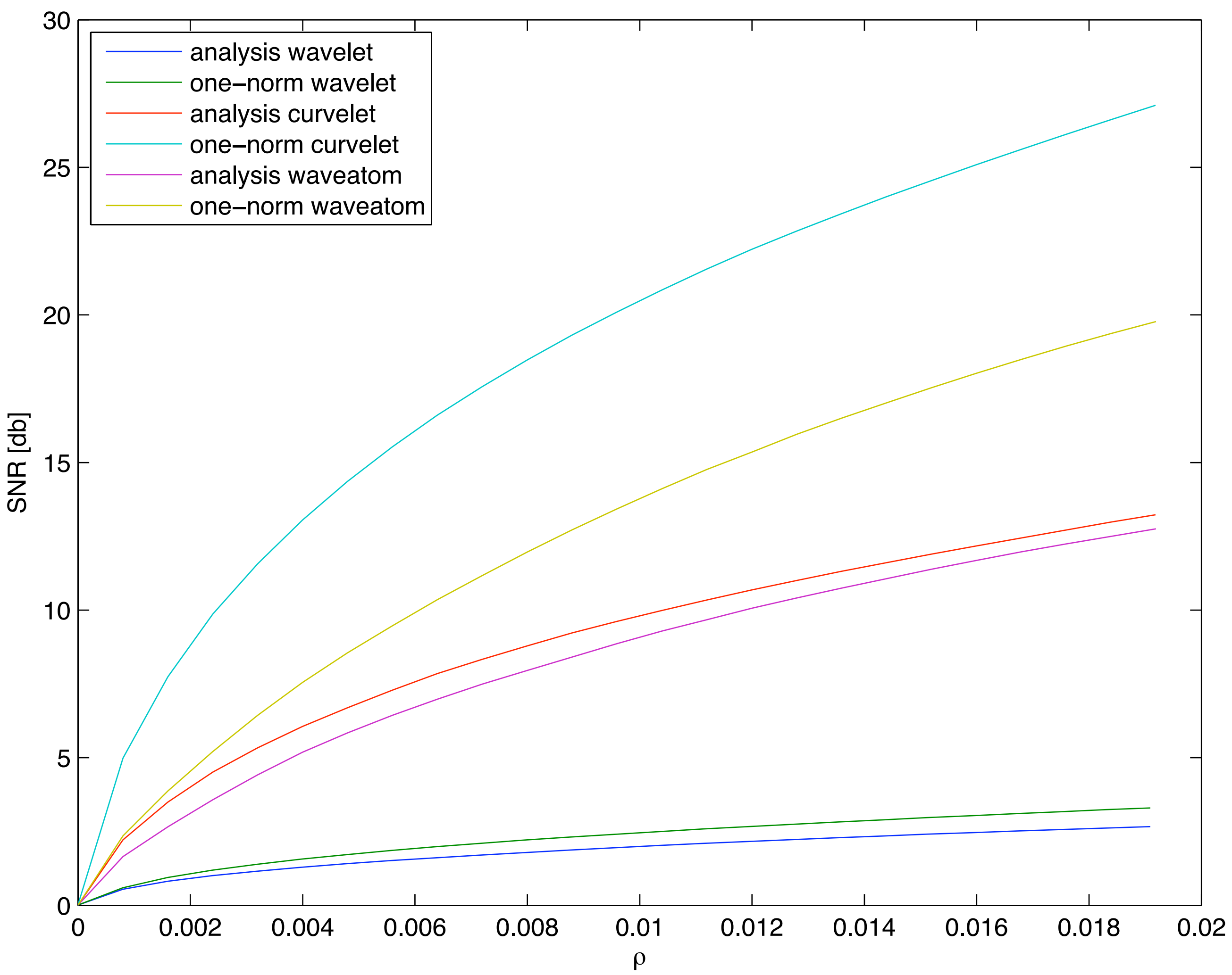
[FJH, '10]

Nonlinear approximation error

common receiver gather



recovery error



[FJH, '10]

Key elements

☑ *sparsifying transform*

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

☐ *advantageous coarse sampling*

- generates incoherent random undersampling "noise" in the sparsifying domain

☐ *sparsity-promoting solver*

- requires few matrix-vector multiplications

Key elements

SLIM

☑ *sparsifying transform*

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

☑ *advantageous coarse sampling*

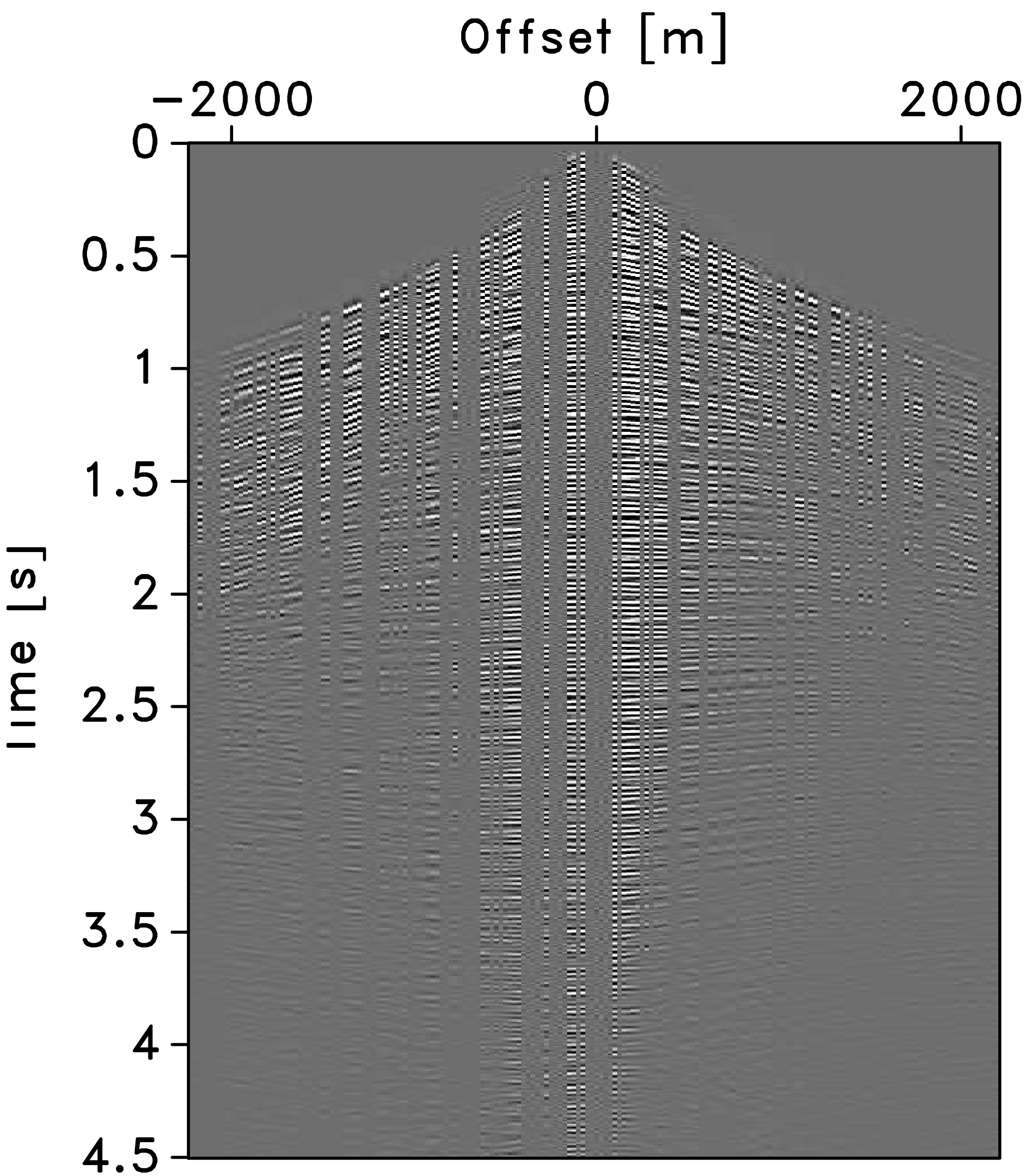
- generates incoherent random undersampling "noise" in the sparsifying domain
- does not create large gaps for measurement in the physical domain
- does not create coherent interferences in simultaneous acquisition

☐ *sparsity-promoting solver*

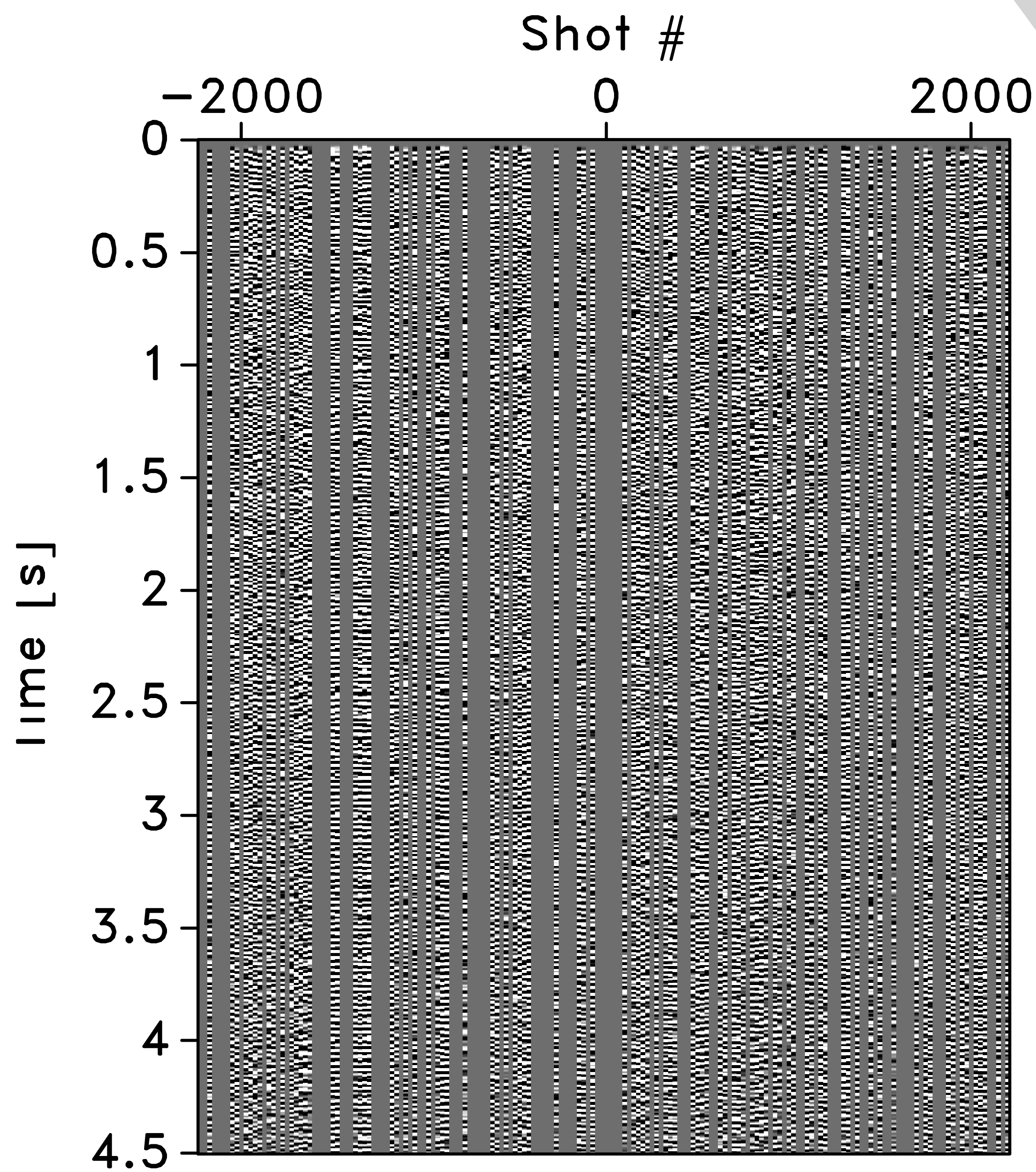
- requires few matrix-vector multiplications

Data

missing shots

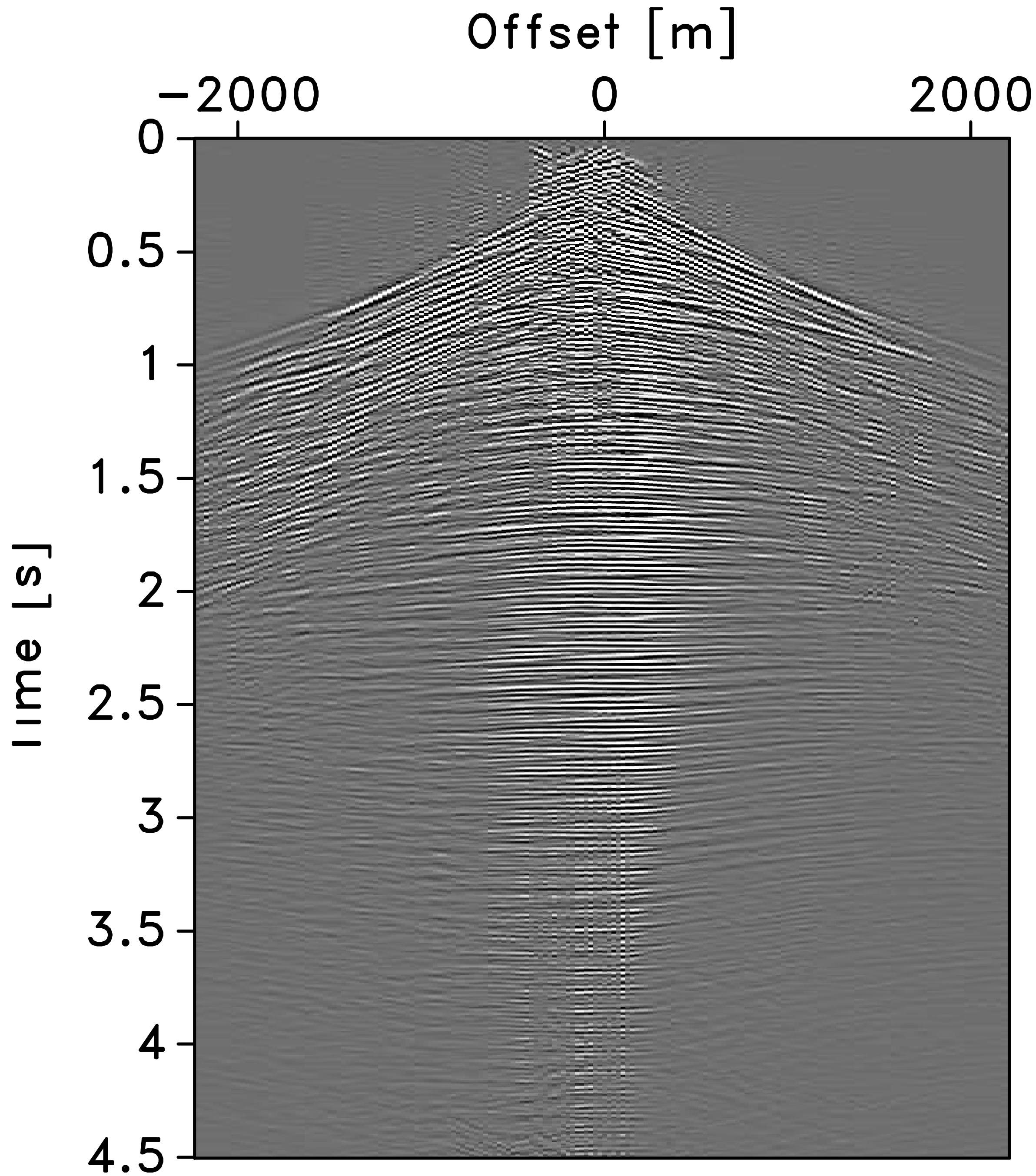


sim. shots

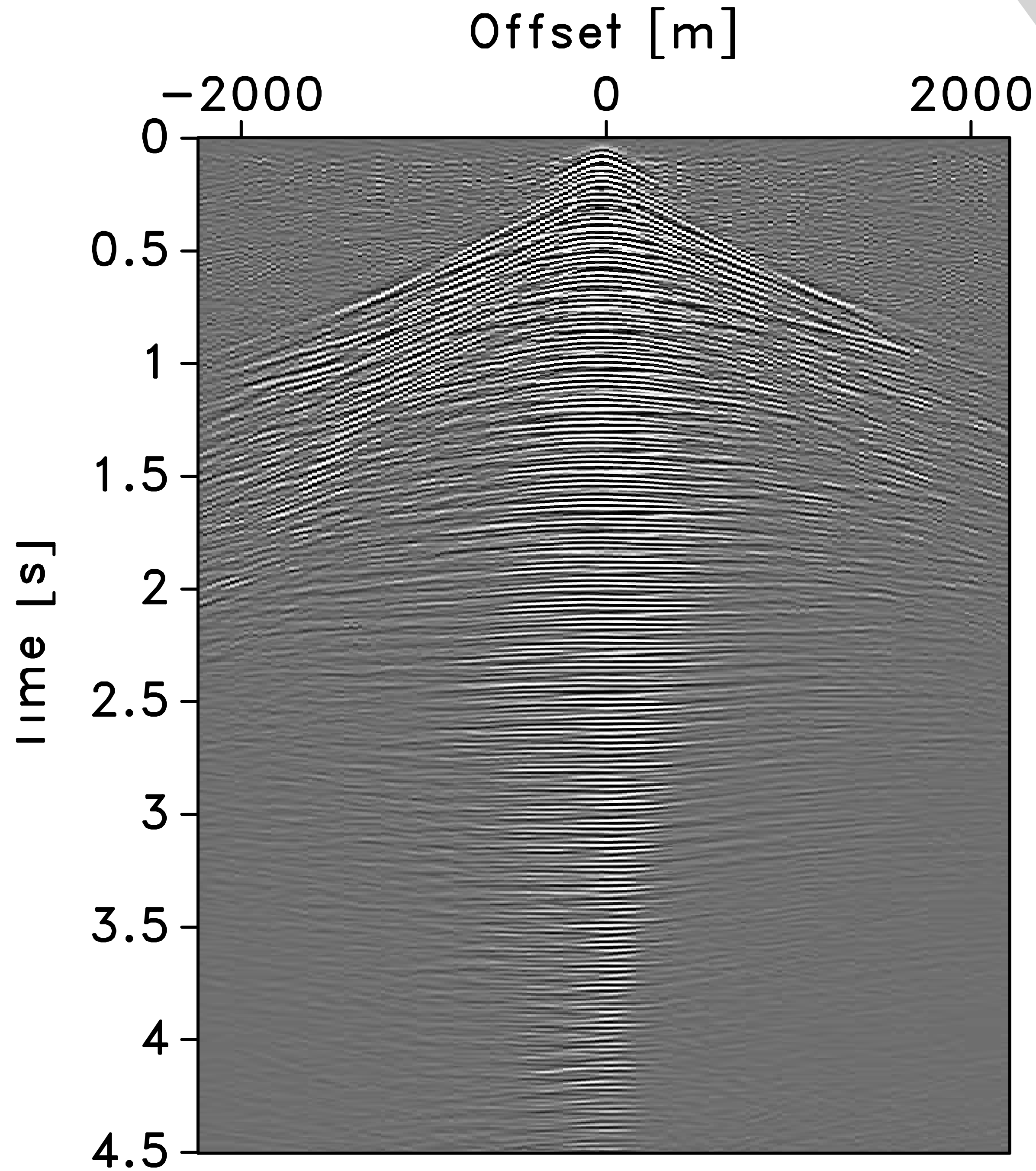


Sparse recovery

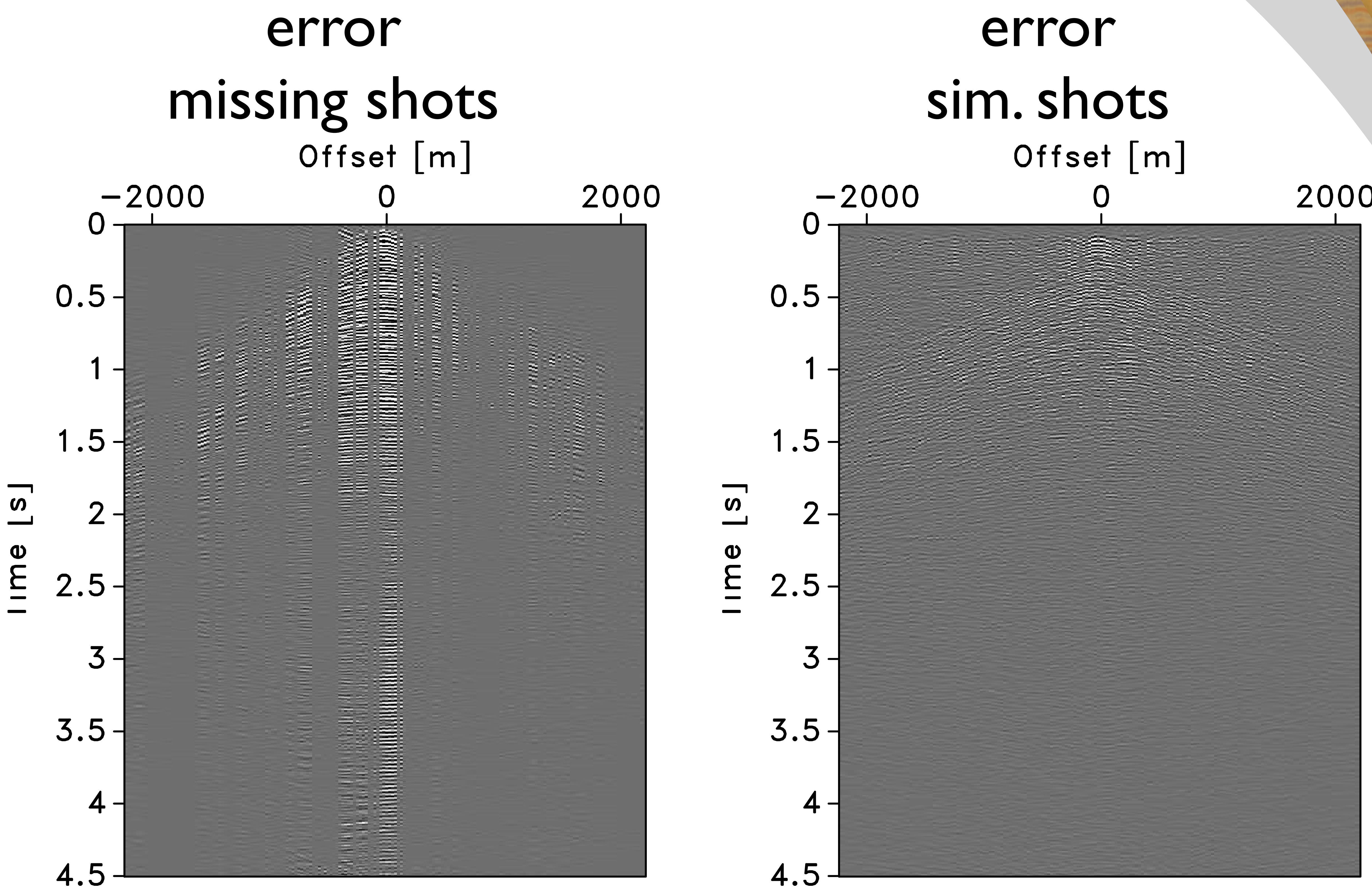
recovery
missing shots



recovery
sim. shots



Sparse recovery error



Empirical performance analysis

Selection of the appropriate sparsifying transform

- nonlinear approximation error

$$\mathrm{SNR}(\rho) = -20 \log \frac{\|\mathbf{f} - \mathbf{f}_\rho\|}{\|\mathbf{f}\|} \quad \text{with} \quad \rho = k/P$$

➡ recovery error

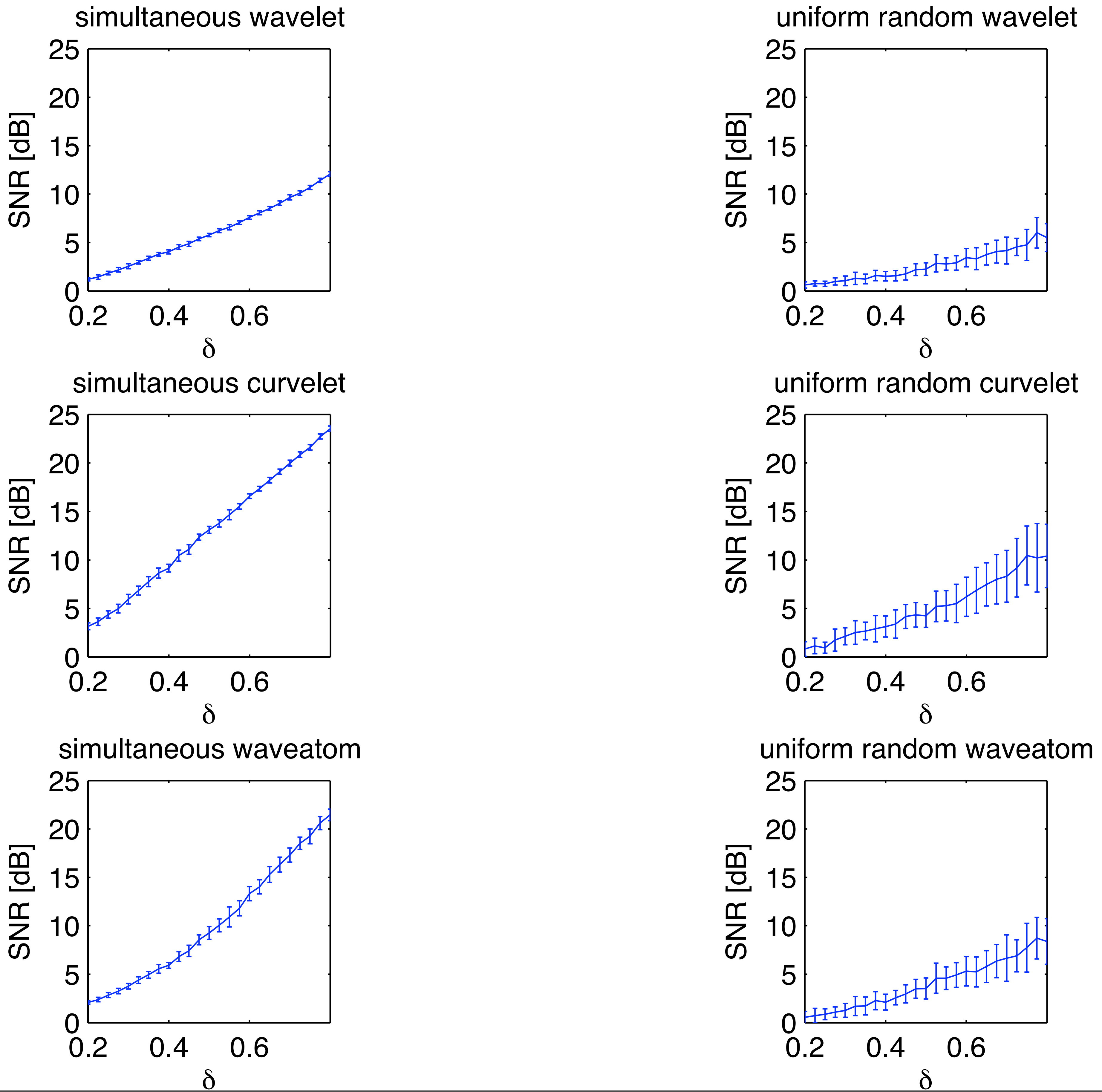
$$\mathrm{SNR}(\delta) = -20 \log \frac{\|\mathbf{f} - \tilde{\mathbf{f}}_\delta\|}{\|\mathbf{f}\|} \quad \text{with} \quad \delta = n/N$$

- oversampling ratio

$$\delta/\rho \quad \text{with} \quad \rho = \inf\{\tilde{\rho} : \quad \overline{\mathrm{SNR}}(\delta) \le \mathrm{SNR}(\tilde{\rho})\}$$

[FJH, '10]

Multiple experiments



Empirical performance analysis

Selection of the appropriate sparsifying transform

- nonlinear approximation error

$$\mathrm{SNR}(\rho) = -20\log\frac{\|\mathbf{f}-\mathbf{f}_\rho\|}{\|\mathbf{f}\|} \quad \text{with} \quad \rho = k/P$$

- recovery error

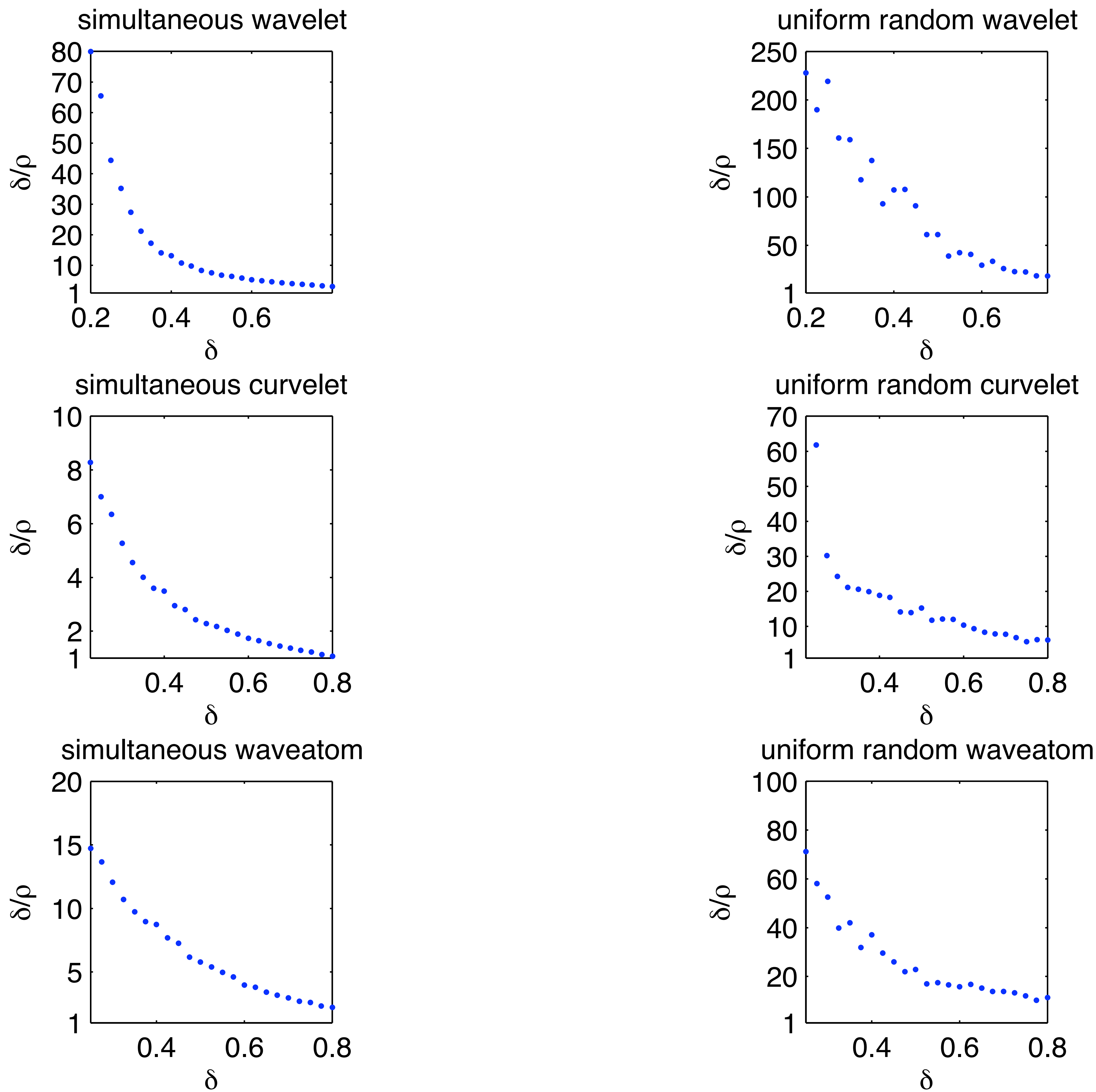
$$\mathrm{SNR}(\delta) = -20\log\frac{\|\mathbf{f}-\tilde{\mathbf{f}}_\delta\|}{\|\mathbf{f}\|} \quad \text{with} \quad \delta = n/N$$

➡ oversampling ratio

$$\delta/\rho \quad \text{with} \quad \rho = \inf\{\tilde{\rho} : \quad \overline{\mathrm{SNR}}(\delta) \leq \mathrm{SNR}(\tilde{\rho})\}$$

[FJH, '10]

Oversampling ratios



Key elements

SLIM

☑ *sparsifying transform*

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

☑ *advantageous coarse sampling (mixing)*

- generates incoherent random undersampling "noise" in the sparsifying domain
- does not create large gaps for measurement in the physical domain
- does not create coherent interferences in simultaneous acquisition

☐ *sparsity-promoting solver*

- requires few matrix-vector multiplications

Reality check

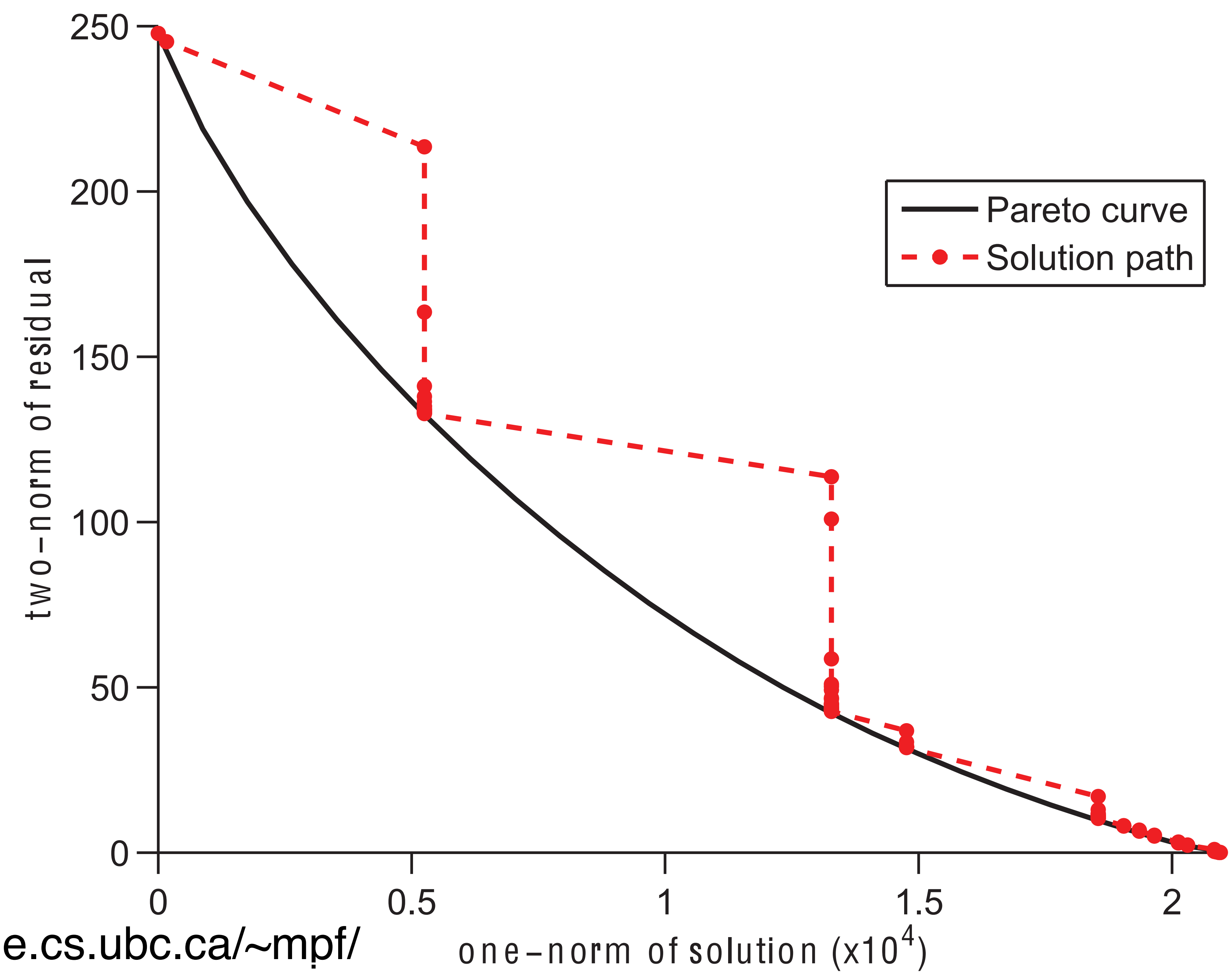
"When a traveler reaches a fork in the road, the l_1-norm tells him to take either one way or the other, but the l_2 -norm instructs him to head off into the bushes."

John F. Claerbout and Francis Muir, 1973

[van den Berg & Friedlander, '08]
[Hennenfent, FJH, et. al, '08]

SLIM

One-norm solver



from http://people.cs.ubc.ca/~mpf/

Key elements

SLIM

☑ *sparsifying transform*

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

☑ *advantageous coarse sampling (mixing)*

- generates incoherent random undersampling "noise" in the sparsifying domain
- does not create large gaps for measurement in the physical domain
- does not create coherent interferences in simultaneous acquisition

☑ *sparsity-promoting solver*

- requires few matrix-vector multiplications

DNOISE contributions

Application CS to

- seismic data regularization
- acquisition design that favors curvelet-based recovery
  - recovery from *jittered* sampling
  - demultiplexing *simultaneous* sources

Observations

Controllable error for reconstruction from *randomized* subsamplings

Oversampling compared to *conventional compression* is small

Combination of *sampling* & *encoding* into a *single* ***linear*** step has profound implications

- *acquisition* costs ***no*** longer determined by *resolution & size*
- *but* by *transform-domain sparsity & recovery error*

3-D Curvelets and simultaneous acquisition perform the best

Talks

Wed 10:00-10:40 AM Ozgur Yilmaz. Sparse approximations and compressive sensing: an overview

Wed 10:40-11:00 AM Haneet Wason. Sequential source data recovery from simultaneous acquisition through transform-domain sparsity promotion

Wed 11:00-11:30 AM Chuck Mosher (ConocoPhillips). Operator localization with Generalized Windowed Transforms

Wed 11:30-12:00 PM Rayan Saab. Compressed sensing using Kronecker products

Wed 12:00-12:30 PM Hassan Mansour. Recovering compressively sampled signals using partial support information

Theme II: Convex optimization

SLIM
Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Additional faculty
Michael Friedlaender
Associate professor CS

- Fellow Argonne
- B.A., Cornell,   MSc. & Ph.D., Stanford
- Numerical optimization
- Numerical linear algebra
- Design & implementation of constrained optimization
- Scientific computing

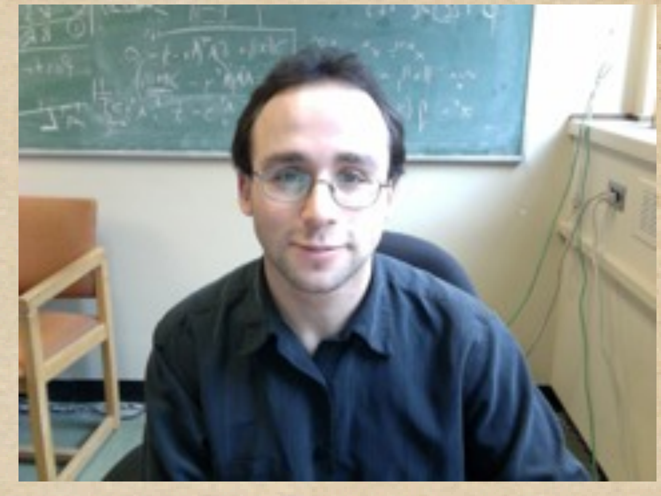
SLIM
Seismic Laboratory for Imaging and Modeling

M. Sc. Student
Tim Lin

- Graduate in Hon. Physics, UBC
- Joined SLIM in 2006 as summer co-op student and now an M. Sc. student
- Compressive Wavefield Modeling and Migration
- Imaging with extensions by Symes

SLIM
Seismic Laboratory for Imaging and Modeling

PDF
Aleksandr Aravkin



- PhD from University of Washington, Math
- Joined SLIM in 2010
- Convex optimization
- Nonlinear inversion with sparsity promotion
- Huber and other norms

SLIM
Seismic Laboratory for Imaging and Modeling

Ph.D. Student
Xiang Li

- M.Sc. in Geophysics from Jilin University (awarded 2009).
- Sparsity-promoting migration with phase encoding
- Dimensionality-reduced FWI with compressive updates

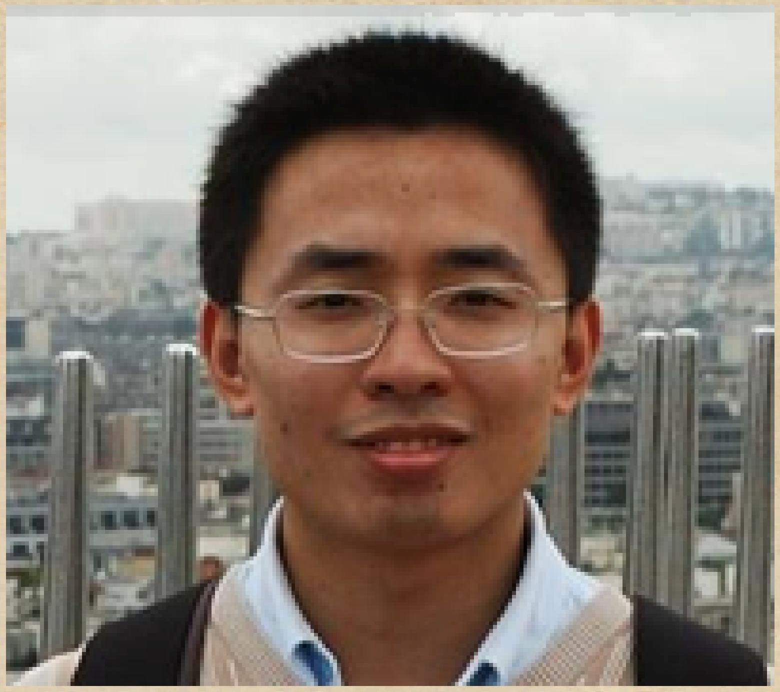
SLIM
Seismic Laboratory for Imaging and Modeling

M. Sc. Student
Mufeed AlMatar

- Previously worked for Aramco in Saudi Arabia
- Curvelet-matched EPSI

SLIM
Seismic Laboratory for Imaging and Modeling

Ph.D. Student
Tu Ning



- M.Sc. from Tsinghua University in 2009 with research related to seismic attenuation characterization.
- Current research interests: imaging with surface-related multiples, compressive sensing, sparse representations, and related applications in seismic exploration.

SLIM
Seismic Laboratory for Imaging and Modeling

Convex optimization

Efficient solvers for problems of the type

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

and

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

are *instrumental* to the success of approach to seismic data processing, imaging, and inversion

SLIM

Prerequisites

The solver needs to

- scale to extremely large problems
- be frugal with # of matrix-vector products
- be like ‘black box’

Solution strategy

Use continuation method to solve a series of one-norm problems

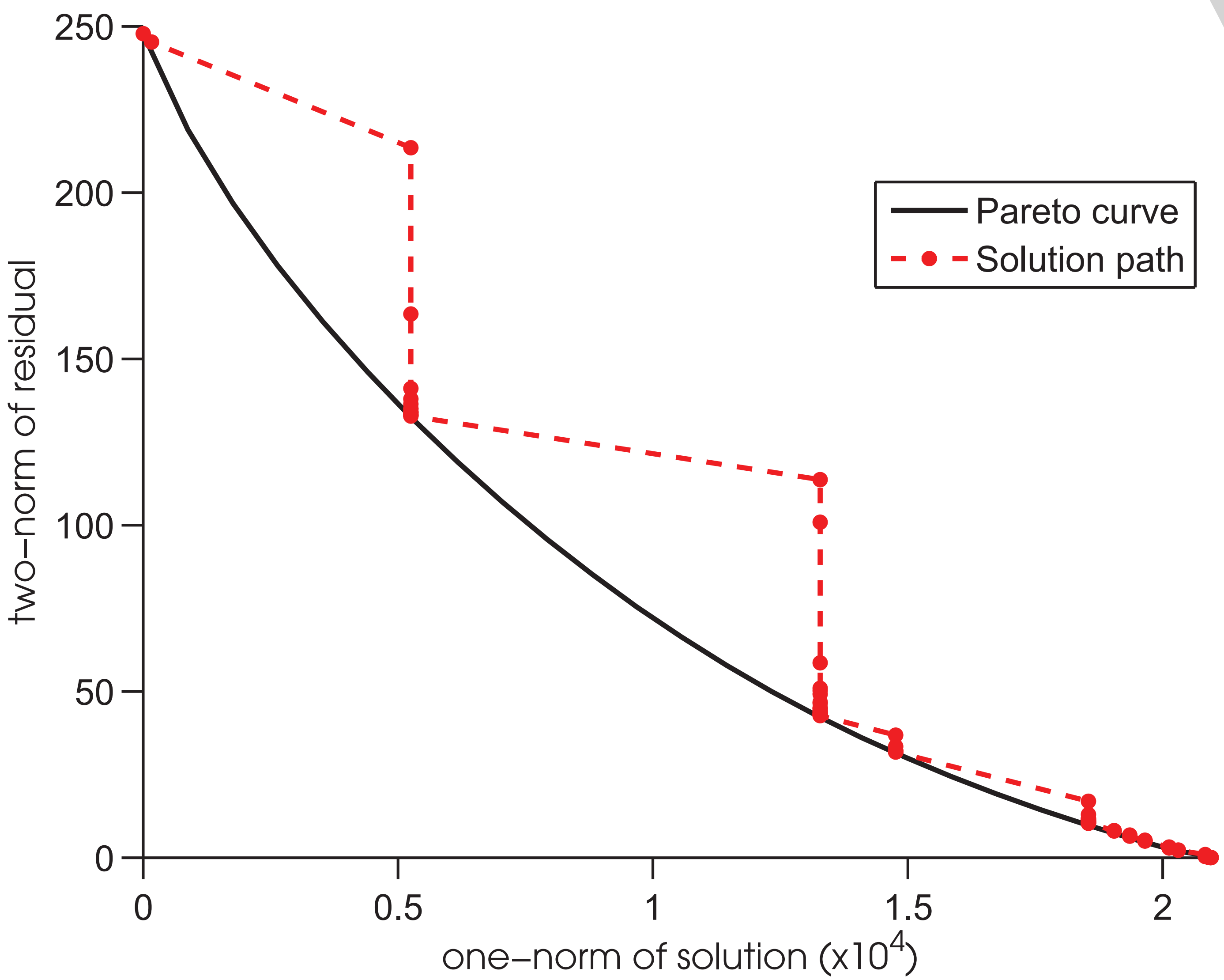
- use properties of the Pareto curve
- divide problem into several subproblems that offer control on the components that enter into the solution
- solution to the subproblem offer flexibility to
  - solve ‘overdetermined’ problems through *subsampling*
  - solve alternating optimization problems

Solution strategy

- Allow components to enter the solution slowly by solving a series of LASSO's

$$\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \tau_k$$

- Solution of these subproblems is a *natural intermediate stopping criterion*

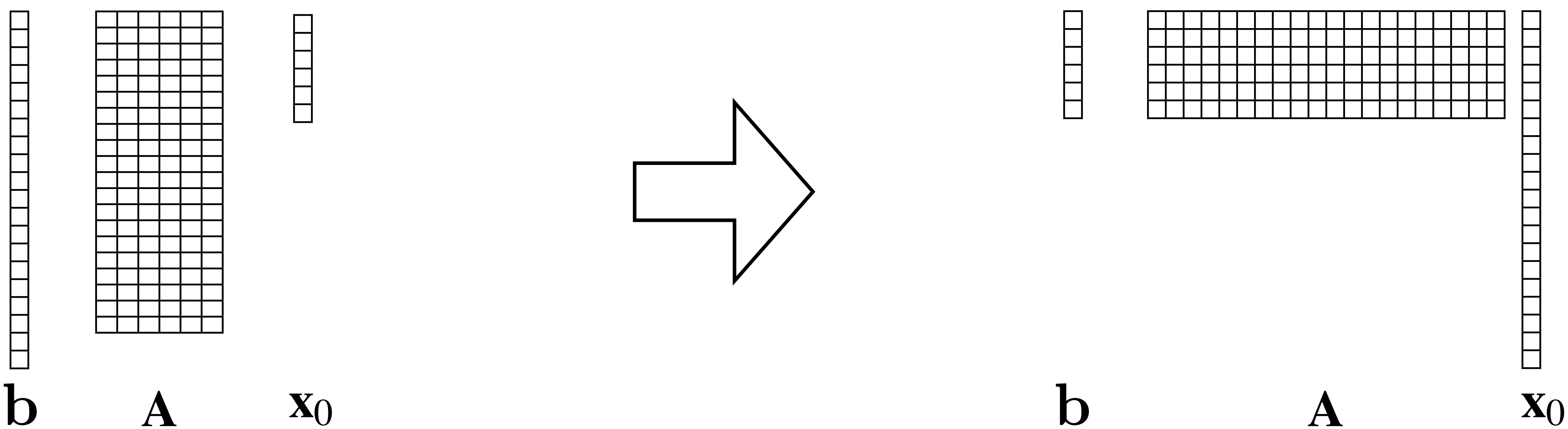


Sparse inversion

$$\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}_{\text{total}}\|_2 \qquad \min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}_k \mathbf{x}\|_2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \tau_k$$

overdetermined
expensive

underdetermined
cheap



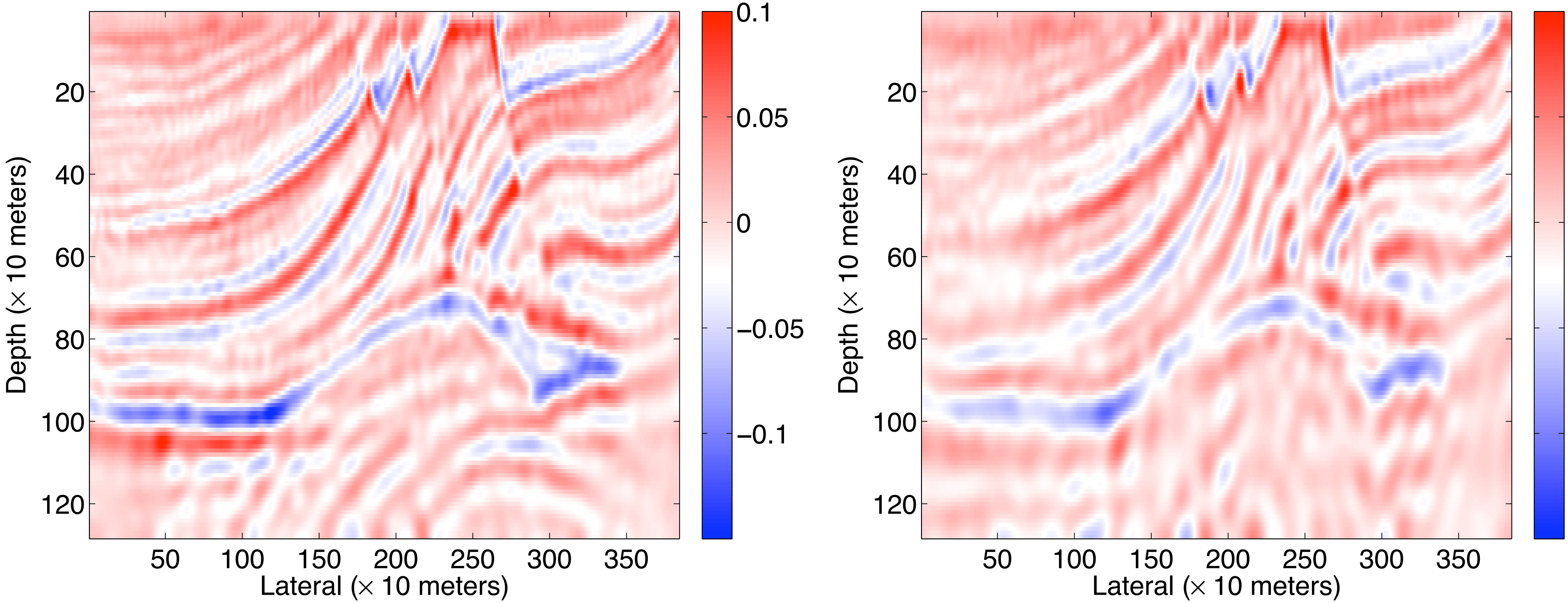
Choose a new set of *simultaneous* sources after each subproblem is solved

SLIM

Linearized sparse inversion

8 simultaneous shots 3 random frequencies

sparse recovery with renewal ***sparse* recovery without renewal**



Block-coordinate descents

$$\min_{\mathbf{x},\mathbf{q}\in\mathcal{C}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{b}-\mathbf{A}[\mathbf{q}]\mathbf{x}\|_2 \leq \sigma$$

$\mathcal{C}$ the set of short wavelets

with **A** and **q** articulating

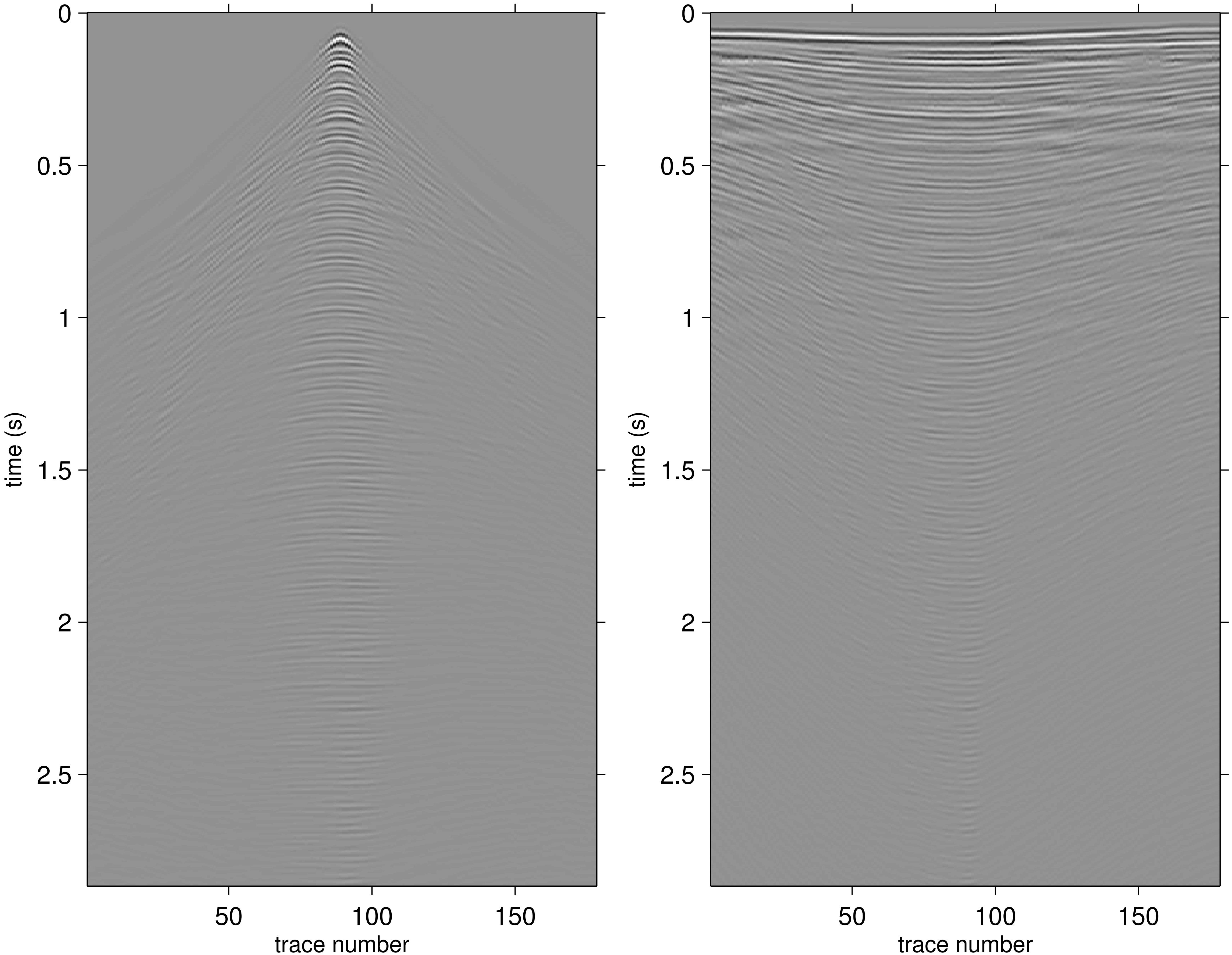
$$\overbrace{\mathbf{P}}^{\text{upgoing wavefield}} \approx \underbrace{\mathbf{G}}_{\text{surface-free impulse response}} \overbrace{[\mathbf{Q}-\mathbf{P}]}^{\text{downgoing wavefield}}$$

where

$$\mathbf{Q} = \mathbf{I}\mathbf{q}$$

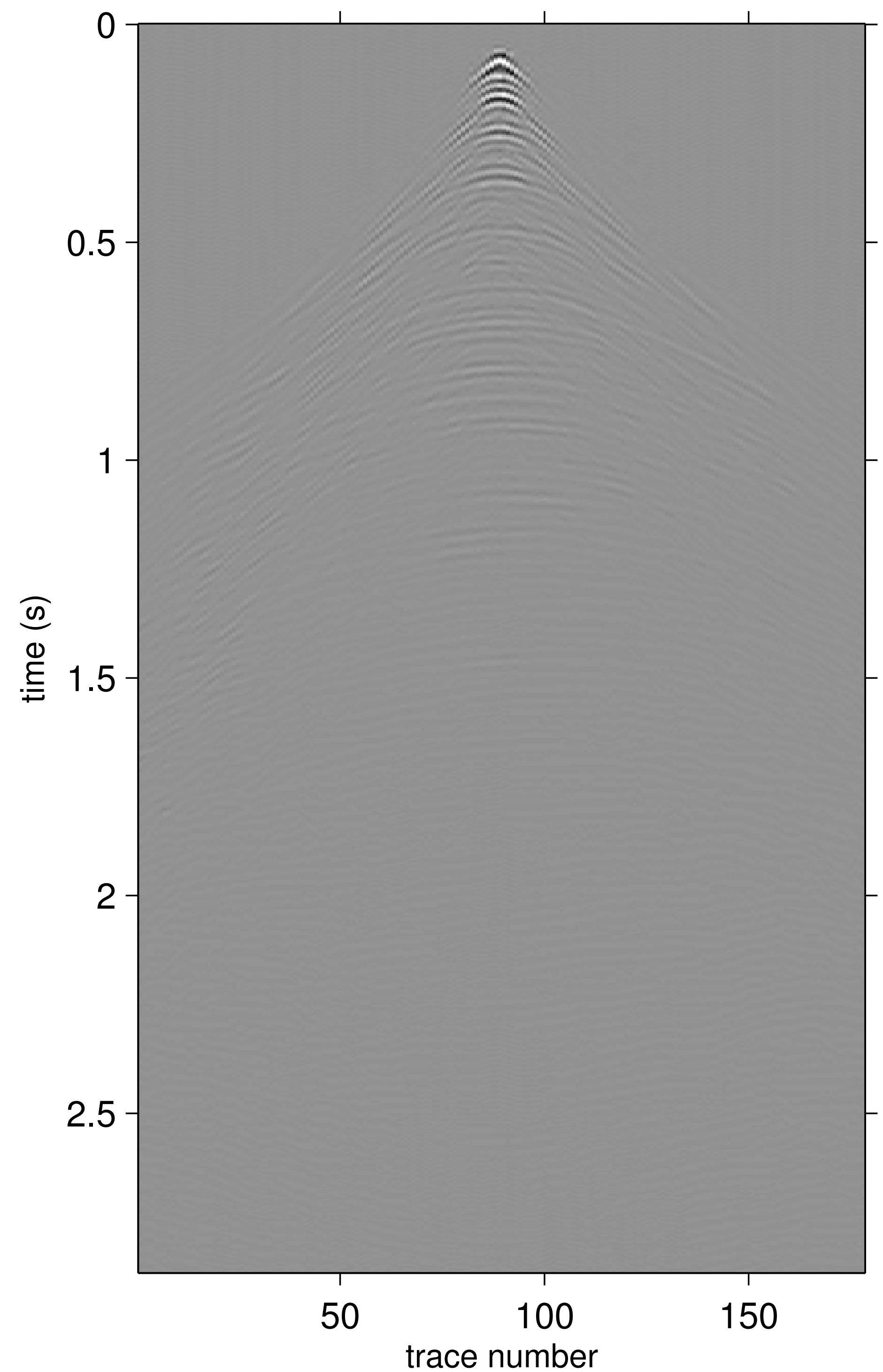
Sparsity vs L1

Gulf of Suez

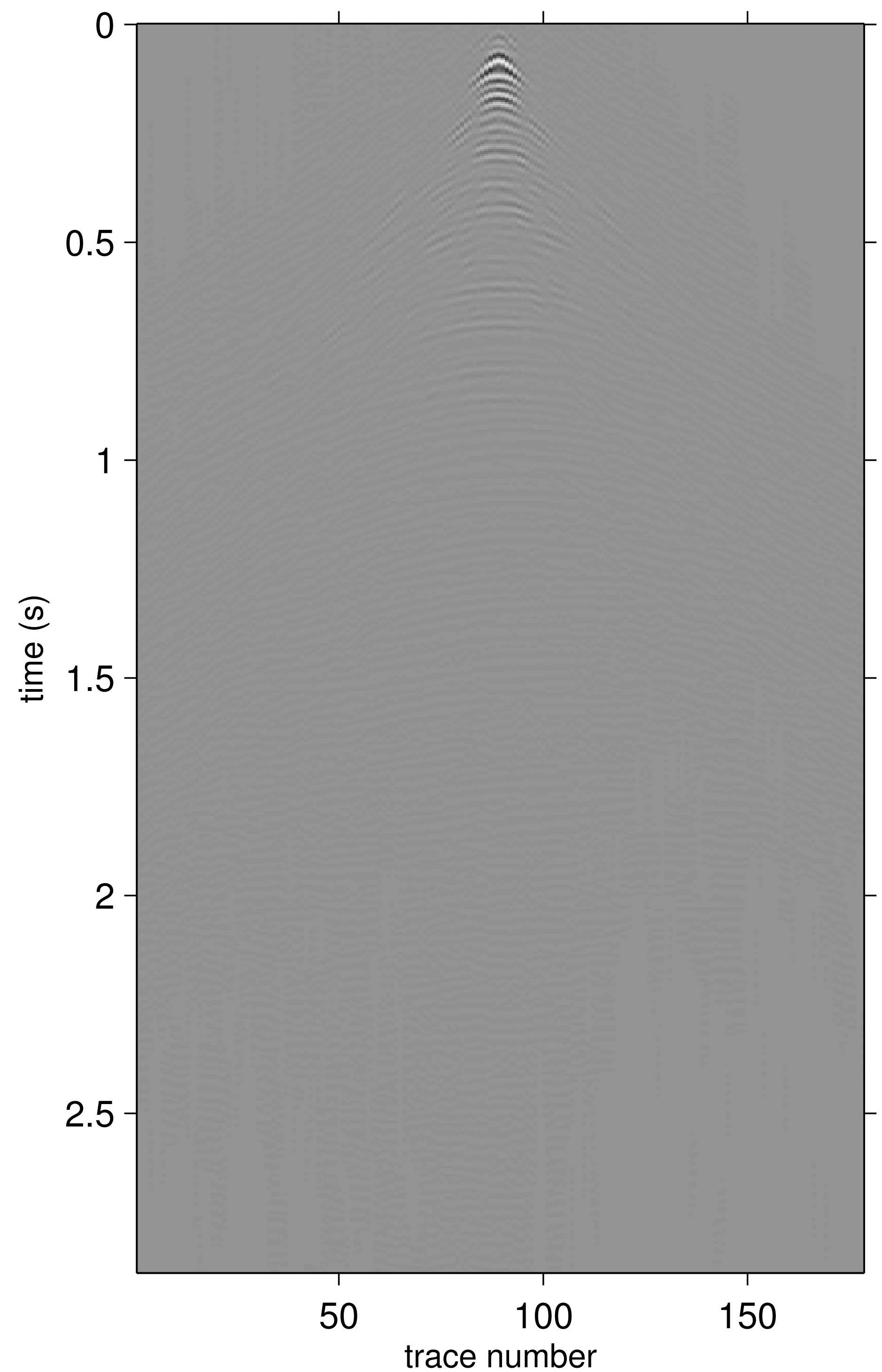


Sparsity vs L1

Gulf of Suez
(one shot)



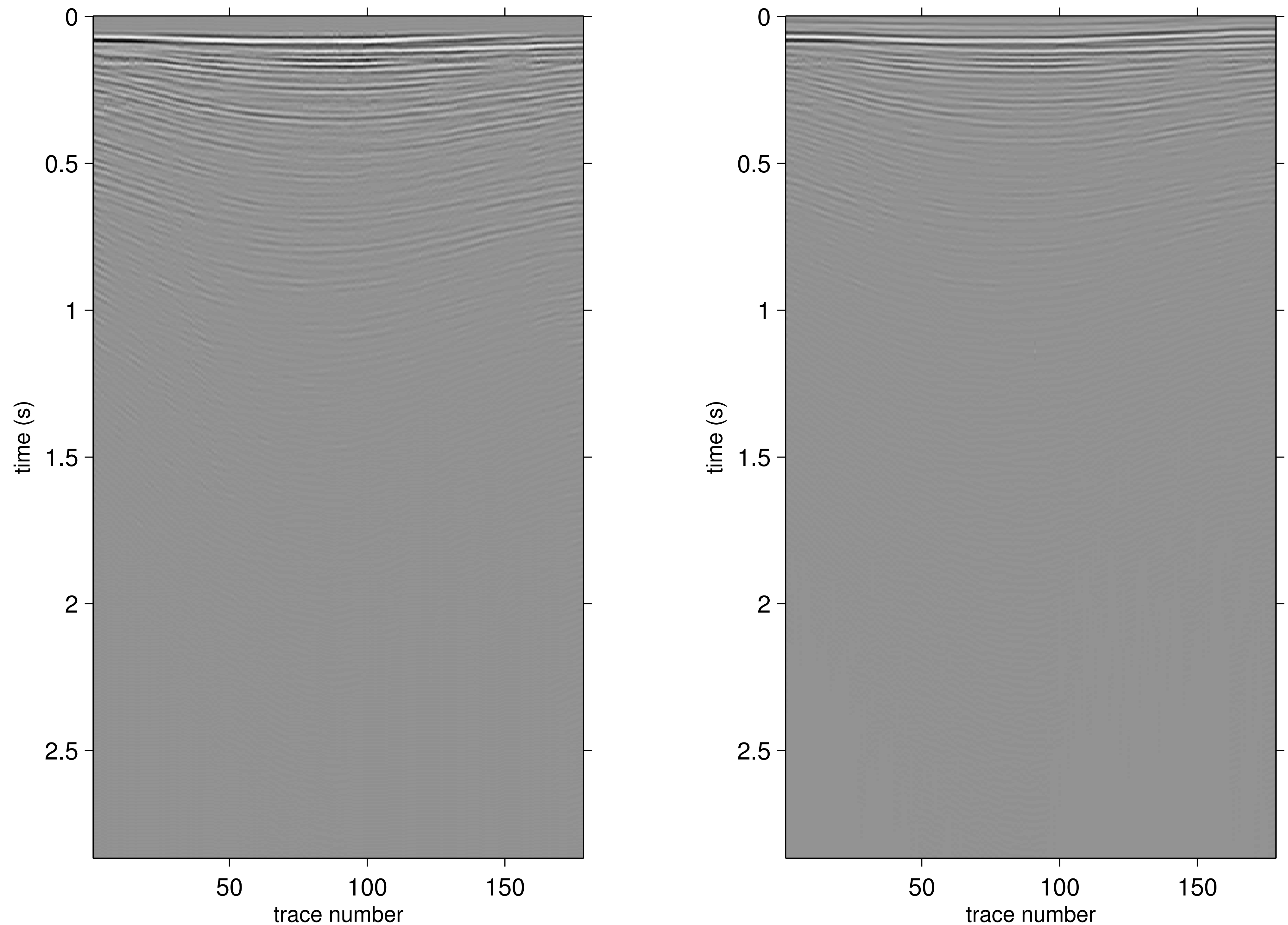
Sparse EPSI



ℓ_1 EPSI

Sparsity vs L1

Gulf of Suez
(zero-offset)

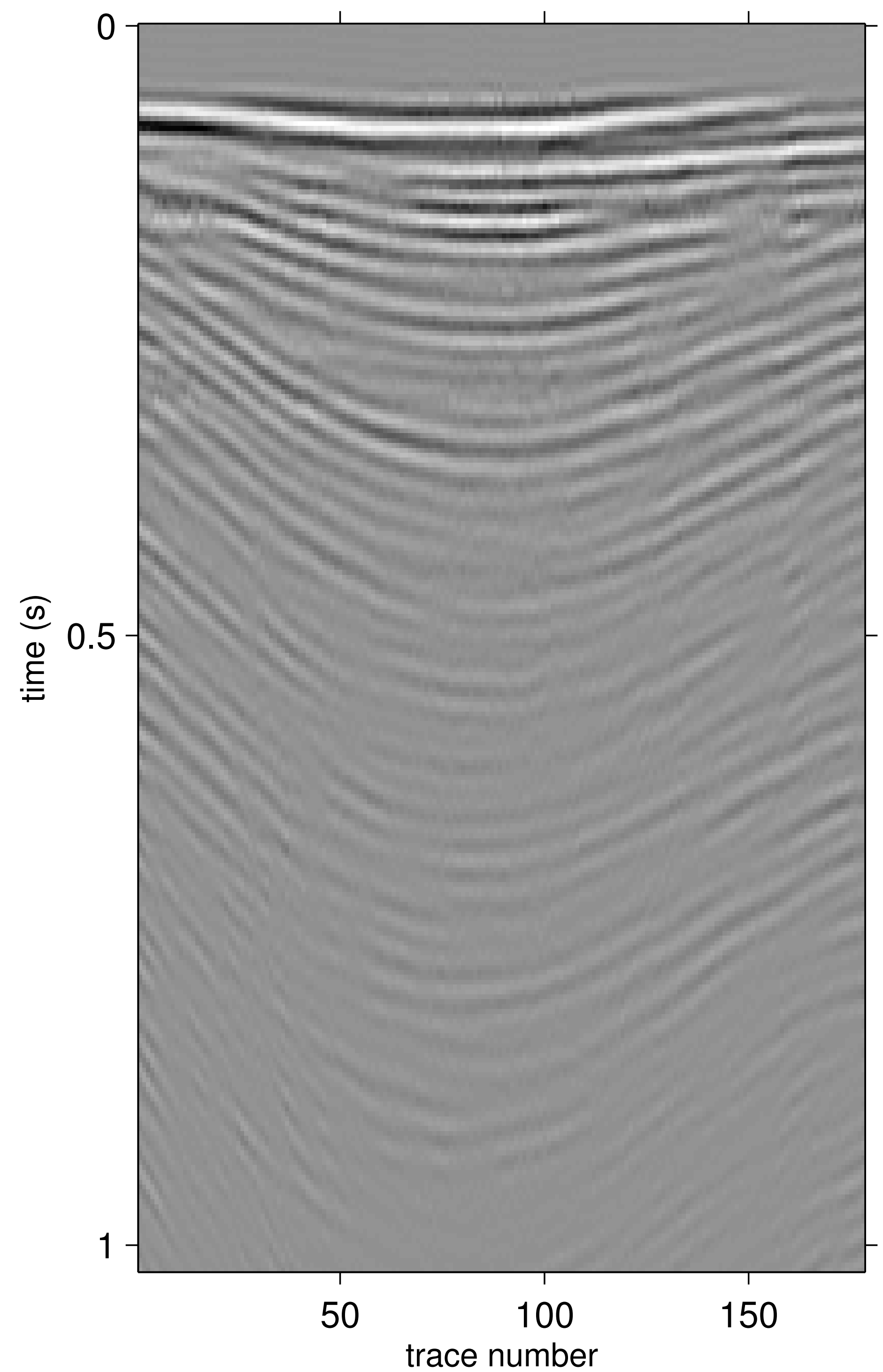


Sparse EPSI

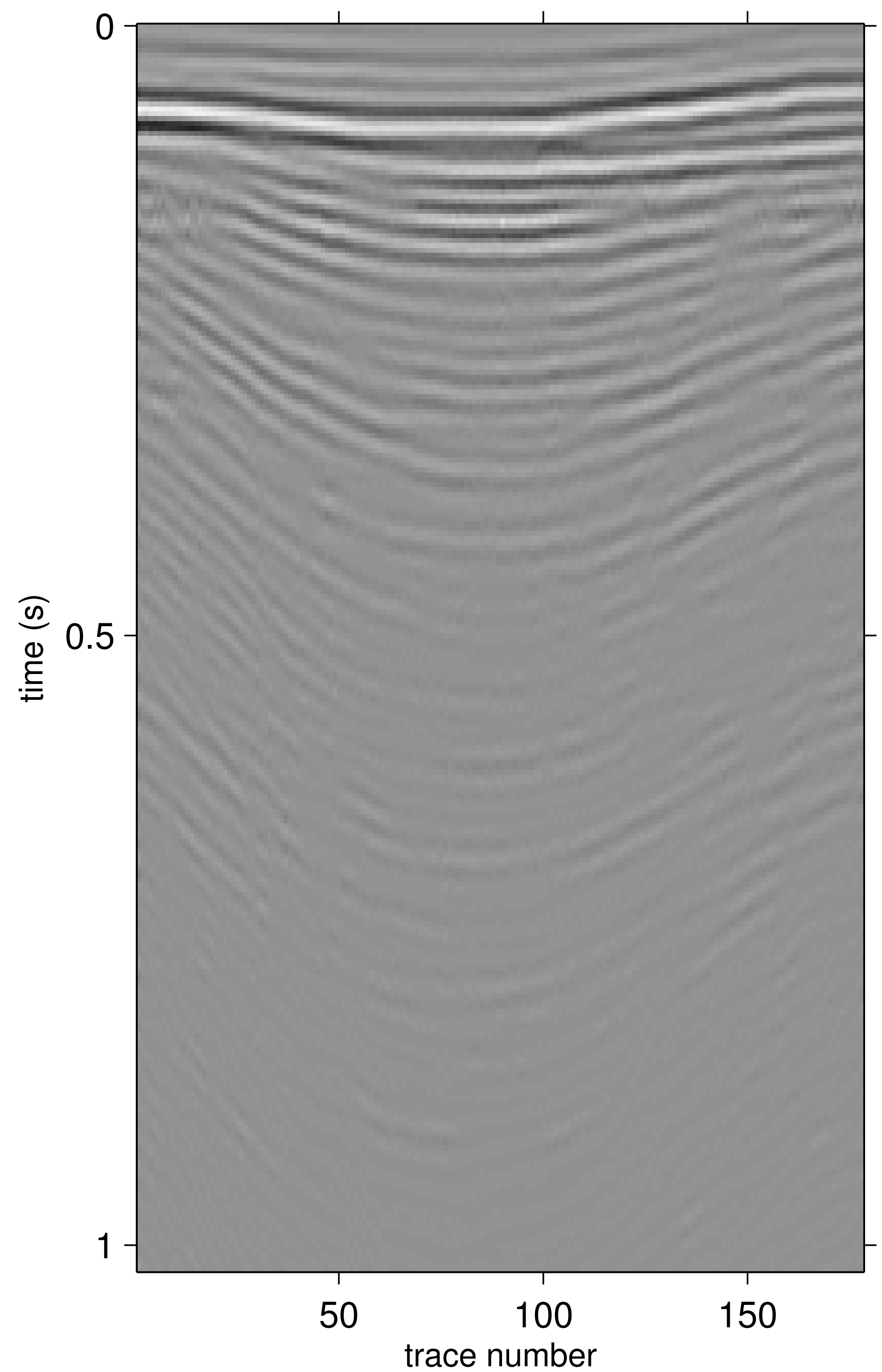
ℓ_1 EPSI

Sparsity vs L1

Gulf of Suez
(zero-offset zoomed)



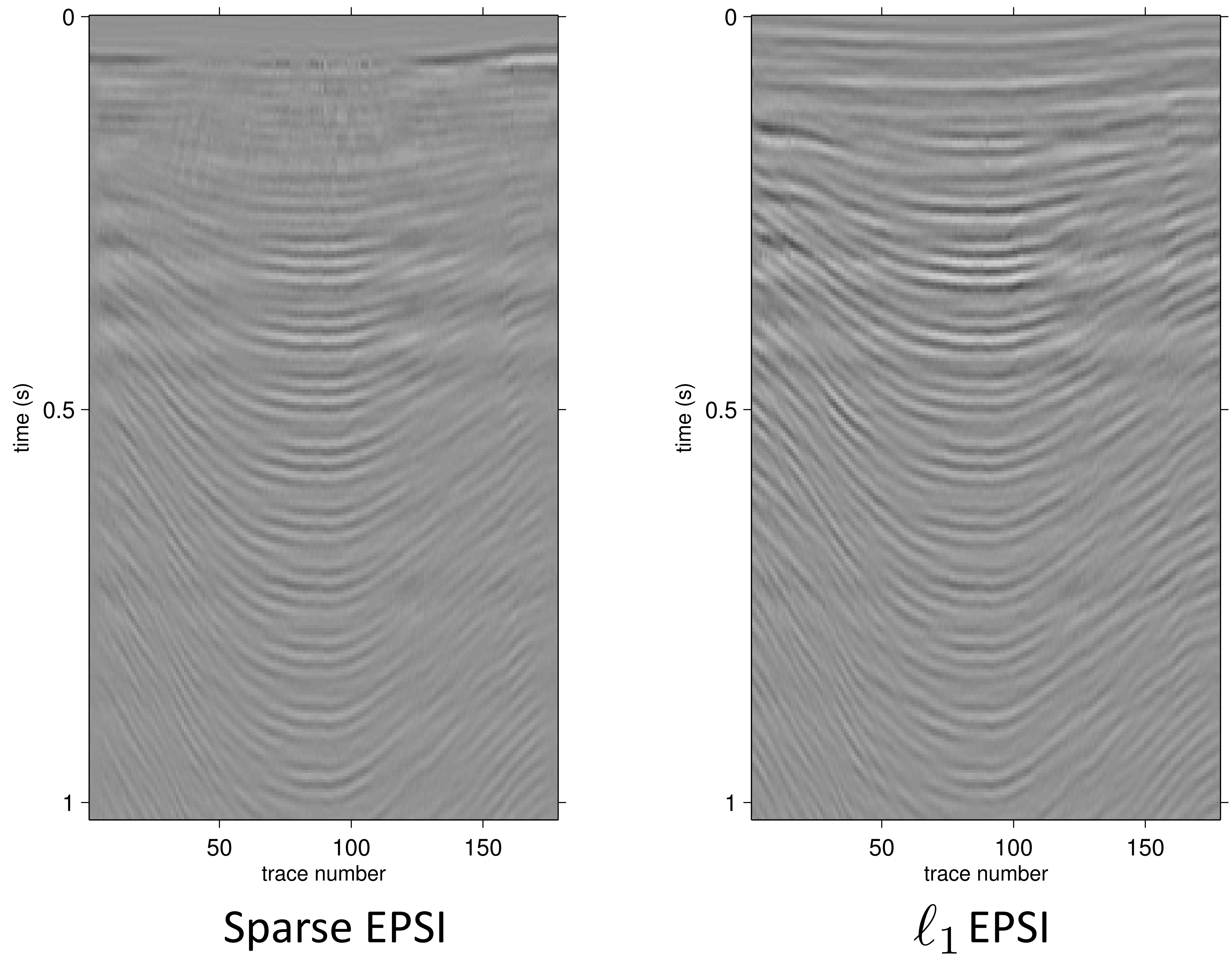
Sparse EPSI



ℓ_1 EPSI

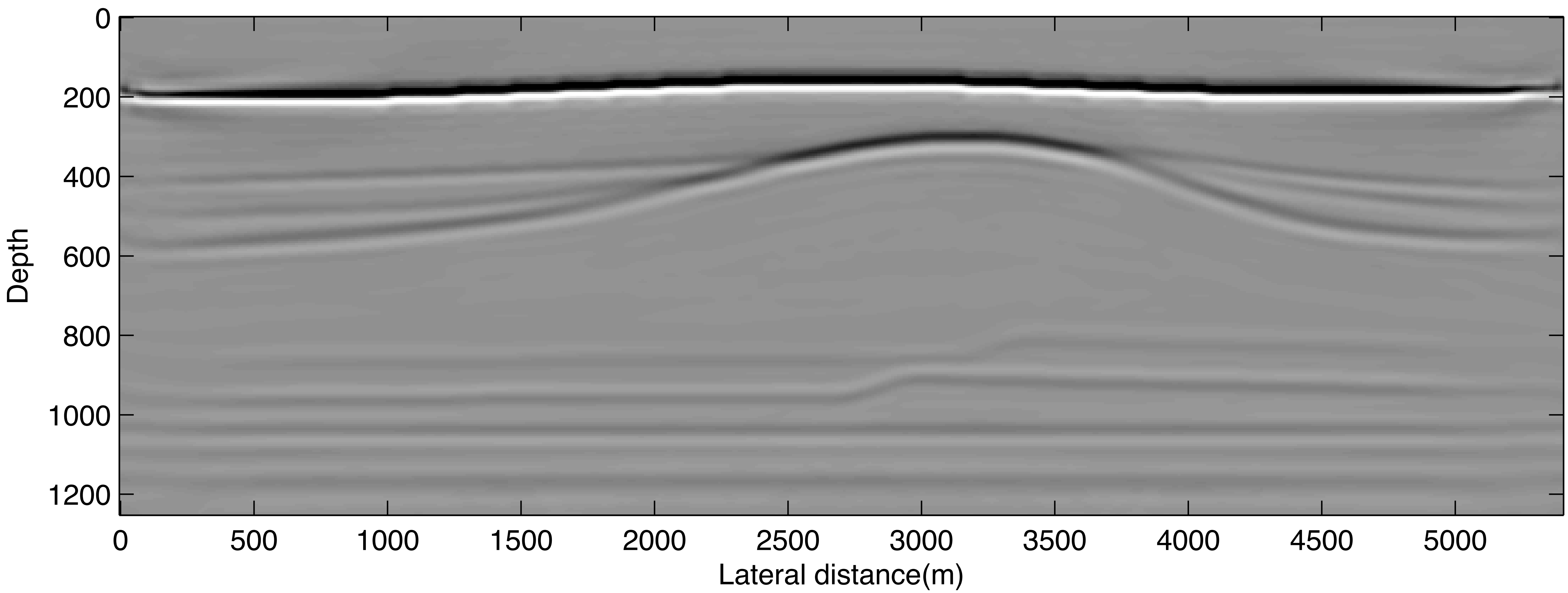
Sparsity vs L1

Data minus estimated primary
(zero-offset)

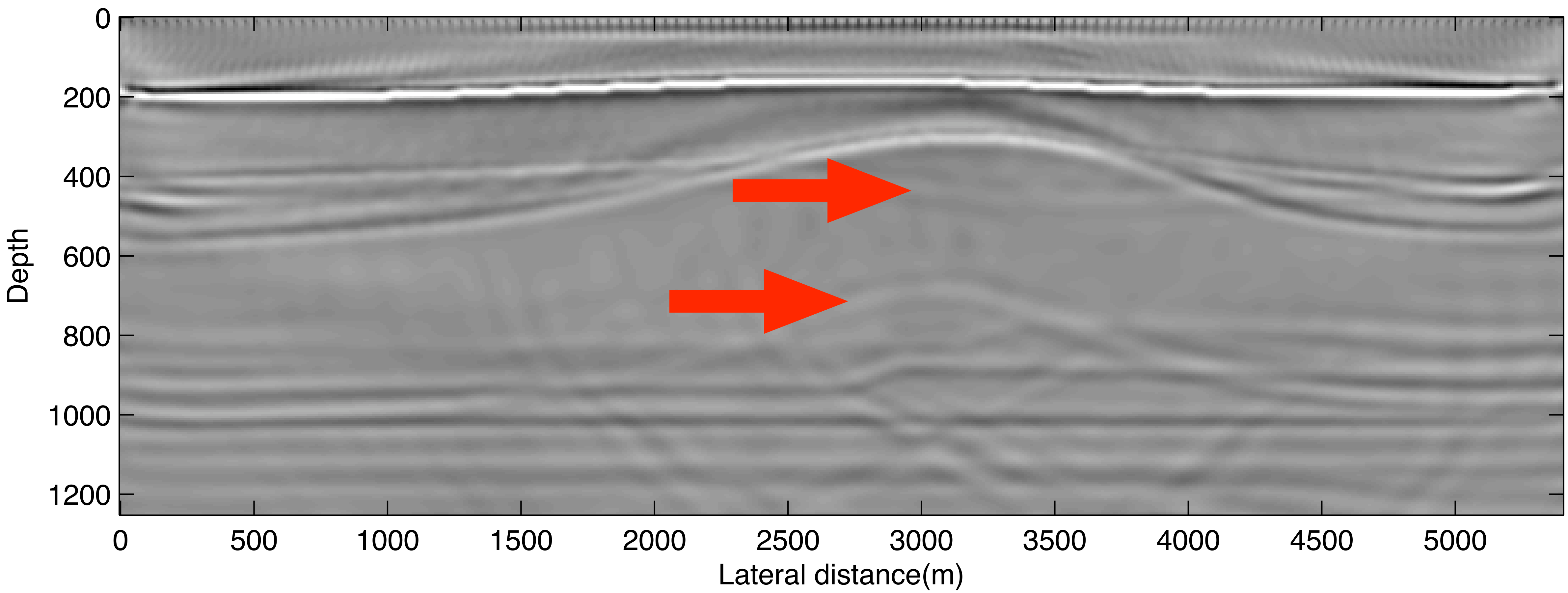


SLIM

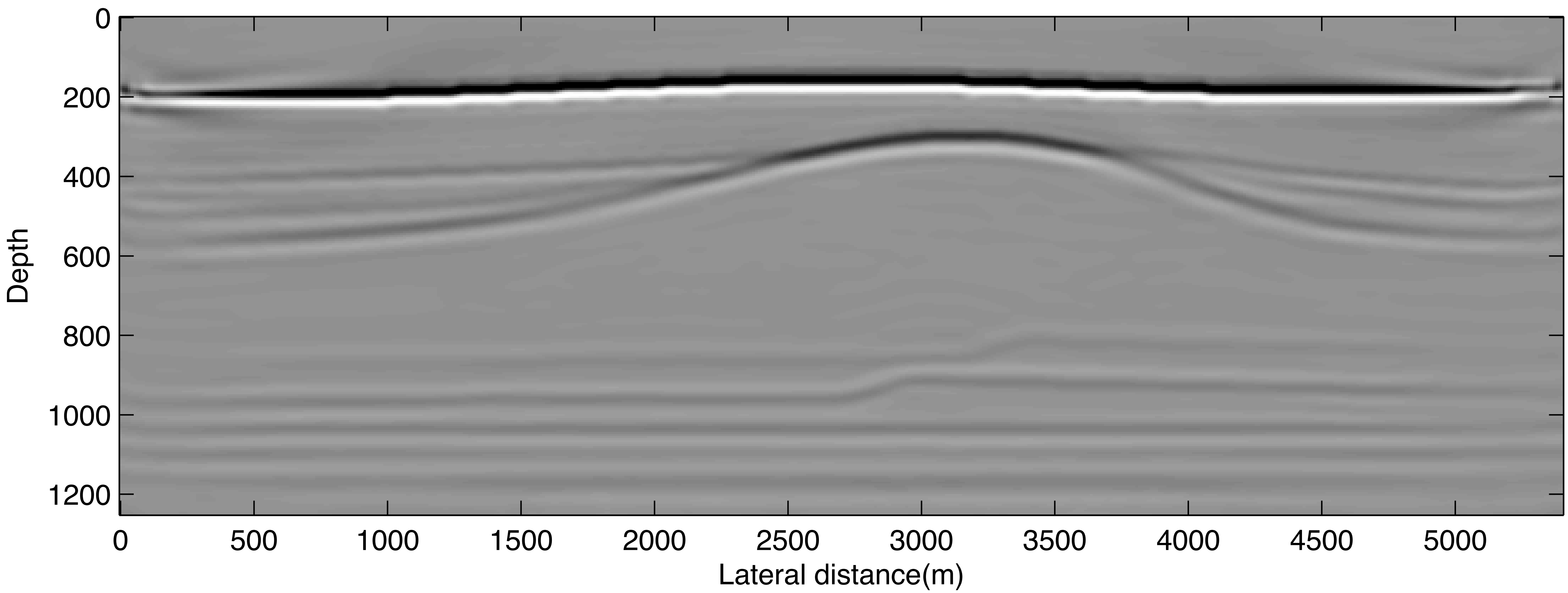
Migration of multiple-free data



Sparse inversion of data with multiples



Sparse inversion with EPSI



DNOISE contributions

Fast solvers for large-scale convex optimization problems

(Bi-)Convex formulations for

- sparse linearized inversion
- source- & Green's function estimation by sparse inversion
- linearized inversion with surface-related multiples

Observations

Curvelet-domain sparsity promotion adds "robustness".

Incorporating more physics improves results.

First steps towards a new generation of processing and imaging technologies.

Dimensionality reduction, which allows us to do sparse inversions.

Talks

Wed 01:30-02:00 PM Michael Friendlander. Algorithms for Sparse Optimization

Wed 02:00-02:40 PM Tim Lin. Sparse optimization and the L1norm

Wed 02:40-03:00 PM Aleksandr Aravkin. Introduction to convex composite optimization

Wed 03:30-03:55 PM Xiang Li. Compressive imaging

Wed 03:55-04:40 PM Tim Lin. Leveraging informed blind deconvolution techniques for the estimation ...

Wed 04:40-05:00 PM Mufeed Ak-Matar. Estimation of surface-free data by curvelet-domain ...

Wed 05:00-05:30 PM Ning Tu. Sparsity promoting migration with surface-related multiples

Theme III: Randomized dimensionality reduction

SLIM

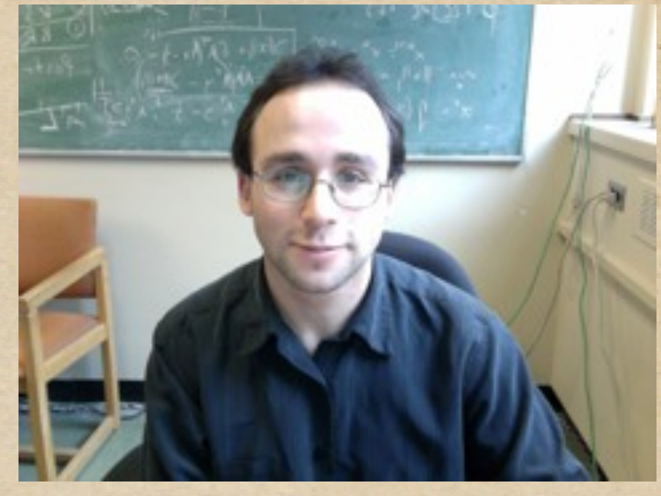
Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Post-doctoral Fellow
Mark Schmidt

- PhD UBC, CS
- Joined SLIM in 2010
- Machine learning
- Convex optimization

SLIM
Seismic Laboratory for Imaging and Modeling

PhD Student
Aleksandr Aravkin



- PhD from University of Washington, Statistics
- Joined SLIM in 2010
- Convex optimization
- Nonlinear inversion with sparsity promotion
- Kalman filtering & smoothing
- Huber and other norms

SLIM
Seismic Laboratory for Imaging and Modeling

Post-doctoral fellow
Peyman Moghaddam

- B.Sc. & M.Sc. in Electrical Eng., Tehran Polytechnic, Iran
- Reverse-time migration
- Migration preconditioning
- Stochastic optimization & FWI

SLIM
Seismic Laboratory for Imaging and Modeling

Post-doctoral fellow Tristan van Leeuwen

- B.Sc./MSc University of Utrecht
- PhD Delft University of Technology
- Joined SLIM in 2010
- Correlation-based migration velocity analysis
- Stochastic optimization & FWI

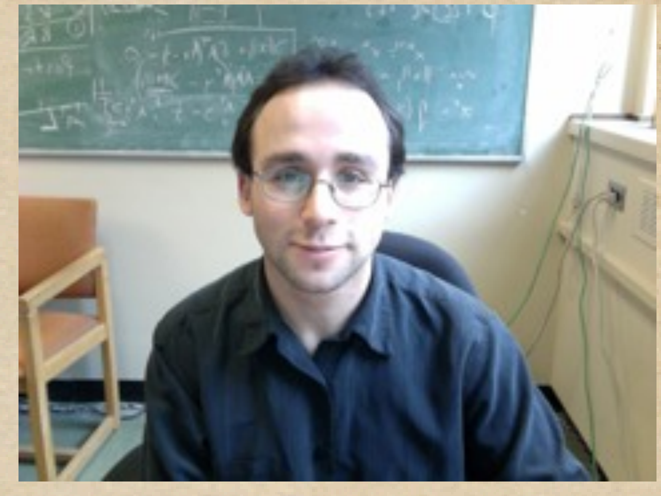
SLIM Seismic Laboratory for Imaging and Modeling

Ph.D. Student
Xiang Li

- M.Sc. in Geophysics from Jilin University (awarded 2009).
- Sparsity-promoting migration with phase encoding
- Dimensionality-reduced FWI with compressive updates

SLIM
Seismic Laboratory for Imaging and Modeling

PDF
Aleksandr Aravkin



- PhD from University of Washington, Math
- Joined SLIM in 2010
- Convex optimization
- Nonlinear inversion with sparsity promotion
- Kalman filtering & smoothing
- Huber and other norms

SLIM
Seismic Laboratory for Imaging and Modeling

Confronting the ‘data deluge‘

Remove major impediment of ‘data overload’ through *randomized dimensionality reduction*

- *artificial randomized* simultaneous sources
- *stochastic approximation*
- *compressive sensing*

Bottom line: *reduction of # PDE solves in FWI*

- makes *FWI computationally feasible* for *realistic* data volumes

Stochastic optimization

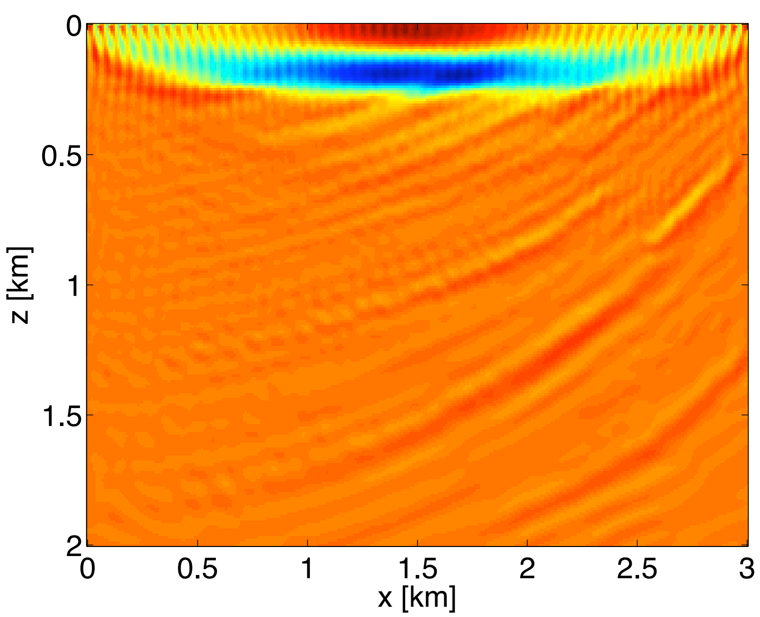
Stochastic average approximation:

- replace sequential sources by fewer # of simultaneous sources
- converges to full solution if # of random sources increases
- corresponds to Monte-Carlo sampling so error decays slowly

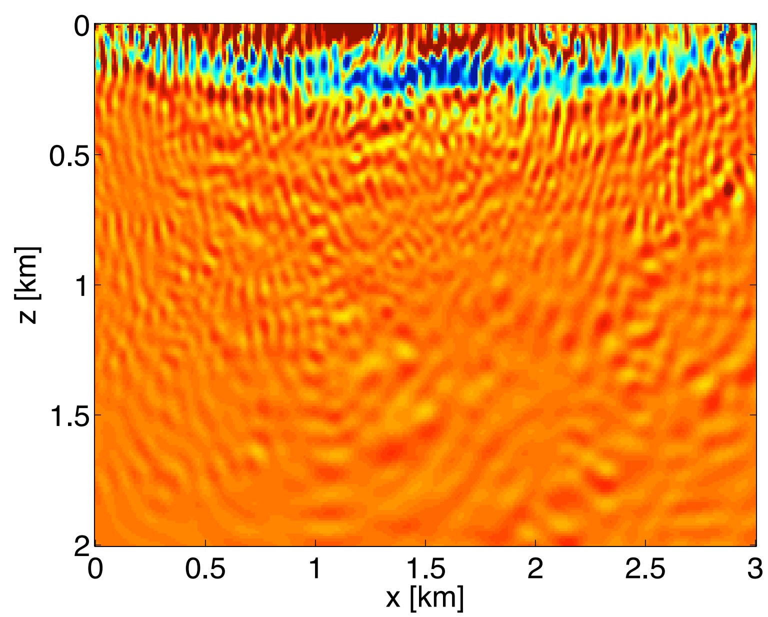
Stylized example

Search direction for batch size K:

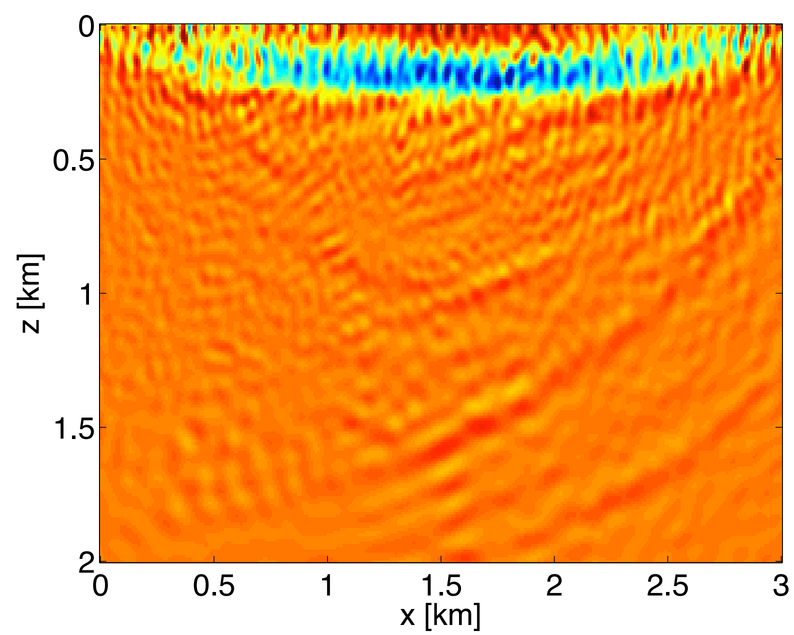
$$\mathbf{g}_K \approx \frac{1}{K} \sum_{j=1}^{K} \nabla \mathcal{F}^*[\mathbf{m}; \underline{\mathbf{q}}_j] \delta \underline{\mathbf{d}}_j$$



full



K=1



K=5

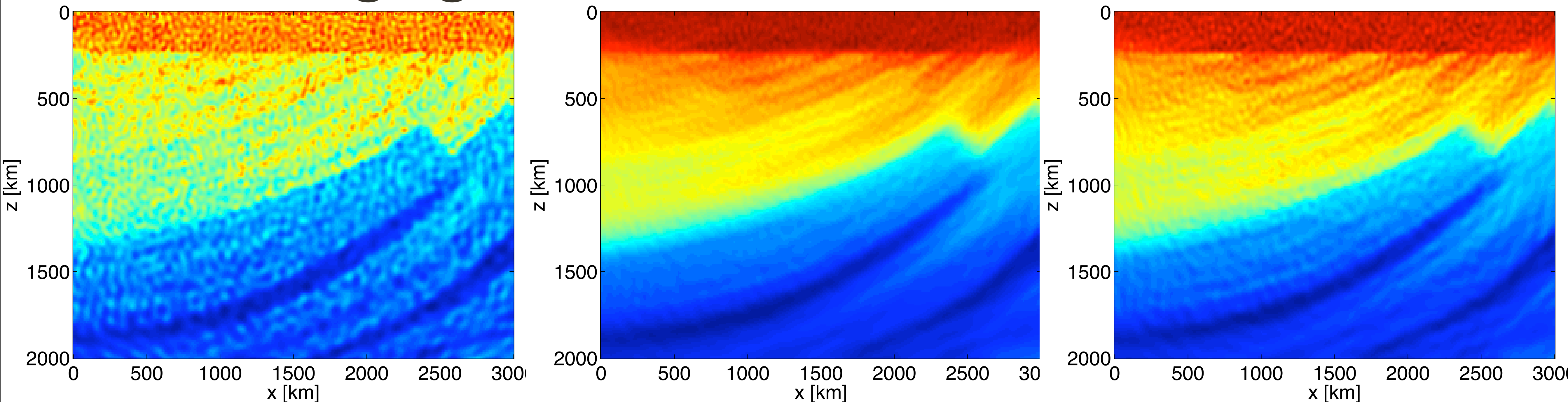
Stochastic optimization

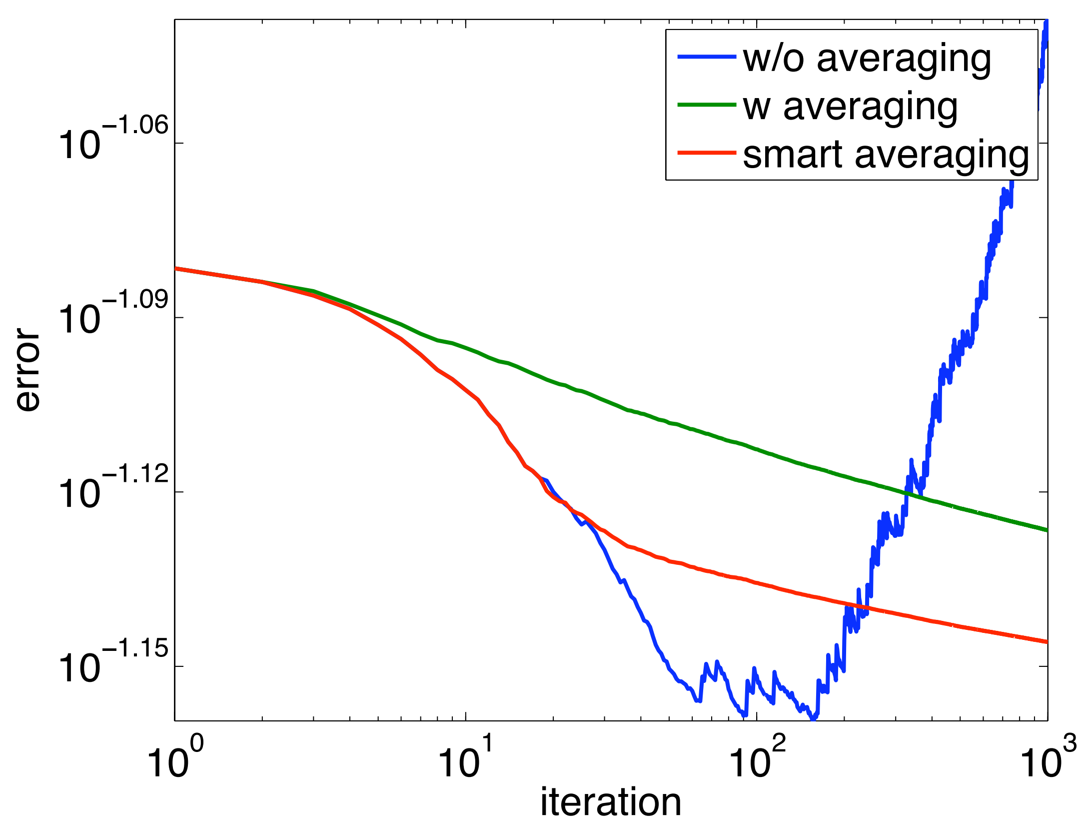
Stochastic approximation:

- replace *deterministic gradients* by *stochastic* gradients
- draw a new set of *simultaneous* sources for each gradient
- converges to solution after sufficient # of iterations
- becomes unstable when noise added

K=1 [noisy case]

w/o averaging — w averaging — smart averaging





Compressive sensing

FWI with *compressively* recovered updates

- consider *phase-encoded* Gauss-Newton updates as separate *compressive-sensing* experiments
- remove *interferences* by *curvelet-domain sparsity* promotion
- exploit properties of Pareto curves in combination with stochastic optimization
- turn ‘overdetermined’ problems into ‘undetermined’ problems via *random source encoding*

Xiang Li's examples

DNOISE contributions

Stochastic optimization formalizes *ad hoc* phase encodings proposed in the geophysical literature.

Combination of the *stochastic approximation* with *compressive sensing* leads to a *dimensionality reduced* formulation of Gauss Newton.

Formulation allows us to tap into *convex* solvers via *convex-composite* arguments.

Concrete prototypes for *dimensionality reduced* FWI.

Observations

Curvelet-domain sparsity promotion allows us to mitigate noise induced by

- source-cross talk
- renewals

Dimensionality reduction, which allows us to do sparse FWI.

Talks

SLIM

Thu 09:00-09:30 AM Felix J. Herrmann. Dimensionality reduction for full-waveform inversion

Thu 09:30-10:00 AM Mark Schmidt. Hybrid stochastic-deterministic methods

Thu 10:30-11:00 AM Peyman Moghaddam. Randomized full-waveform inversion

Thu 11:00-11:30 AM Tristan van Leeuwen. Waveform inversion by stochastic optimization

Thu 11:30-12:00 PM Xiang Li. Full-waveform inversion with randomized L1 recovery for the model updates

Thu 12:00-12:30 PM Aleksandr Aravkin & Tristan van Leeuwen. Exploiting sparsity in full-waveform inversion: nonlinear basis pursuit denoise algorithm

Tutorials & Software

SLIM

Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Research Faculty
Henryk Modzelewski (Ph.D.)

- Ph.D. in Atmospheric Sciences, UBC
- Scientific programming
  - High-Performance Computing
  - Development: MPI and Python
- System administration

SLIM
Seismic Laboratory for Imaging and Modeling

Nameet Kumar (undergraduate COOP)

- Physics UBC
- Scientific programmer for pSPOT

SLIM
Seismic Laboratory for Imaging and Modeling

Sebastien Pacteau (undergraduate interm)

- Geophysics
- Scientific programmer for Kroneckers in pSPOT

Recent developments

Purchase of Parallel matlab turned underutilized cluster into fully utilized problem solver.

OO programming allows us to

- seamlessly build in parallelization
- fast prototyping
- compartmentalize our applications into logical units, linear operators, solvers, etc.
- streamlines *dissemination* of our *technology*

Contributions

Development of a parallel OO framework for linear operators

Large-scale for convex optimization problems

Software release for primary-estimation by sparse inversion

Planned releases for linearized inversion (with multiples)

Talks

Thu 01:30-01:50 PM Henryk Modzeleweski. Software releases and architecture

Thu 01:50-02:10 PM Michael Friedlander. Introduction to Spot – A Linear-Operator Toolbox

Thu 02:10-02:30 PM Nameet Kumar. Parallelizing operations with ease using Parallel SPOT

Thu 02:30-02:50 PM Sebastien Pactau. Parallel SPOT and Kronecker products

Thu 02:50-03:10 PM Tim Lin. Software release: Estimation of primaries by l1 inversion