Randomized sampling strategies

Felix J. Herrmann

SLIM

Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Drivers & impediments

"Acquisition costs"

- Full-waveform inversion requires ***hifi*** data

"Data deluge"

- *"Curse of dimensionality"* compounded by over restrictive "*Nyquist sampling criterion*"

"*Limits on computational resources*"

- End of "*Moore's law*"

SLIM

Wish list

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

- sampling *criteria* that are *dominated 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

Strategy

Adapt recent ***compressive sensing*** (CS)

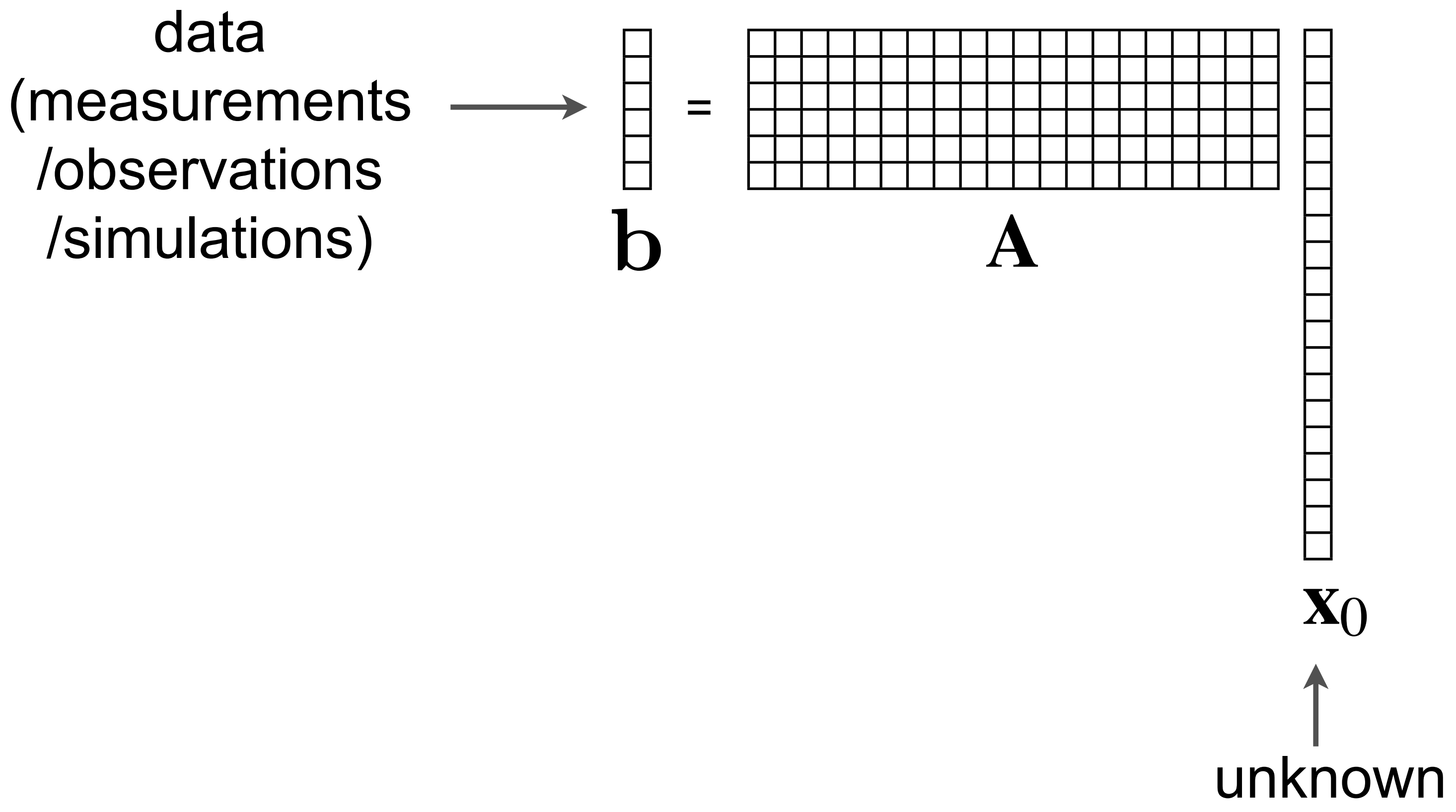
- *randomized* subsampling - turn *aliases/interferences* into *noise*
- *sparsity* promotion - removes subsampling *noise* by exploiting signal *structure*

This is really an *"acquisition" design problem*

Let's have a look at a stylized *recovery* problem first...

Problem statement

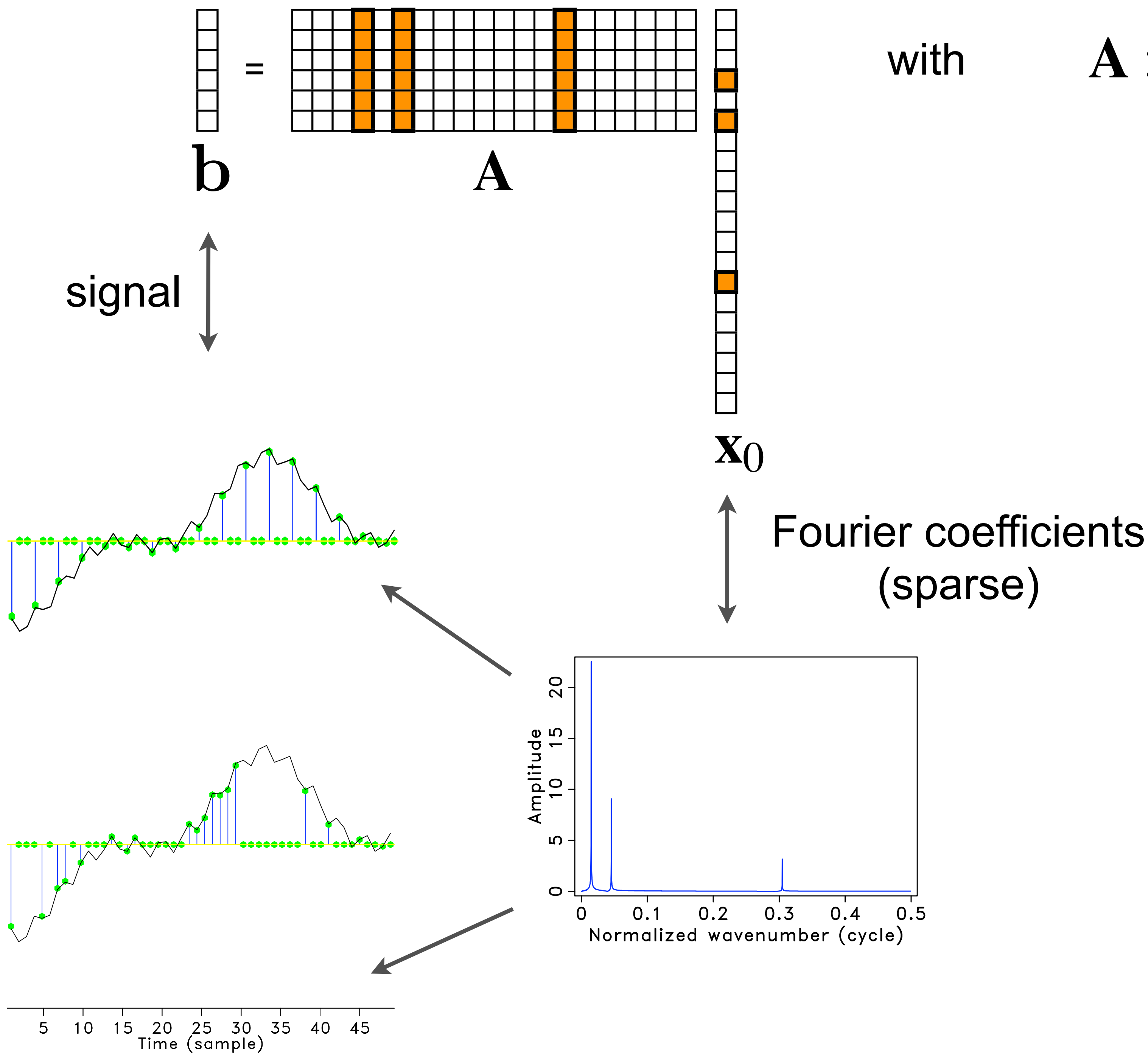
Consider the following (severely) *underdetermined* system of *linear* equations:



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

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

Sparse recovery

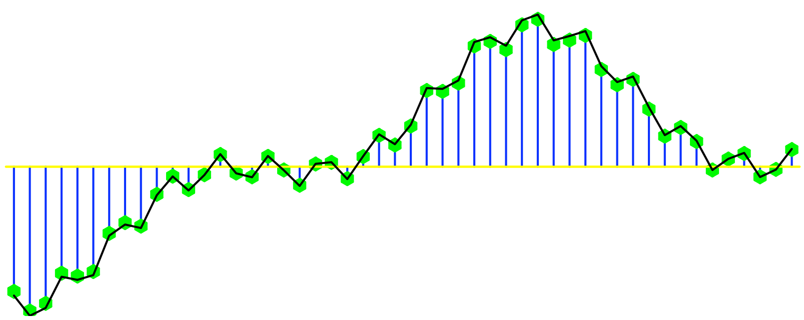


with $\mathbf{A} := \mathbf{R}\mathbf{F}^H$

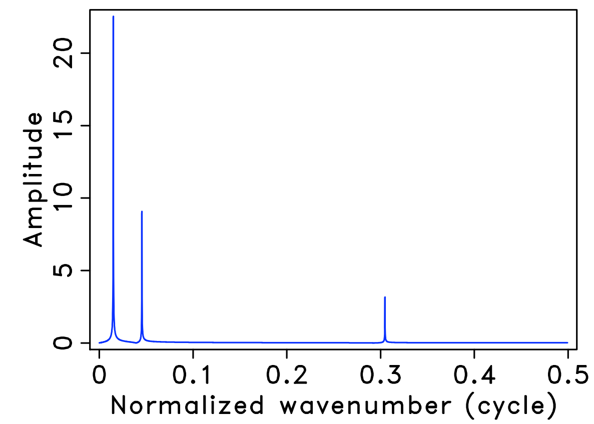
restriction operator

Fourier transform

Coarse sampling schemes

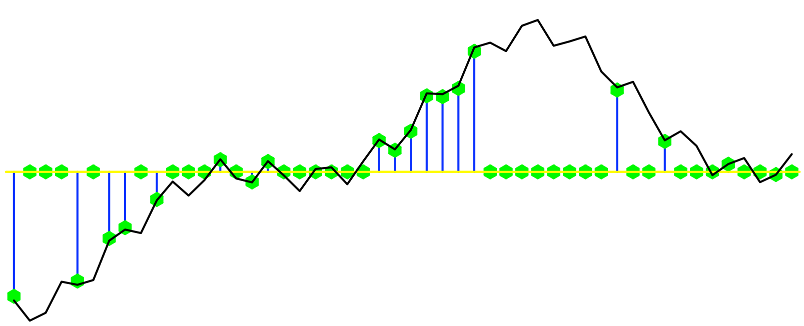


Fourier → transform

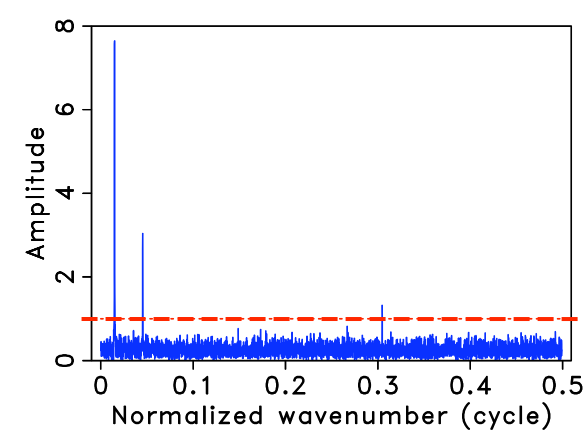


few significant coefficients

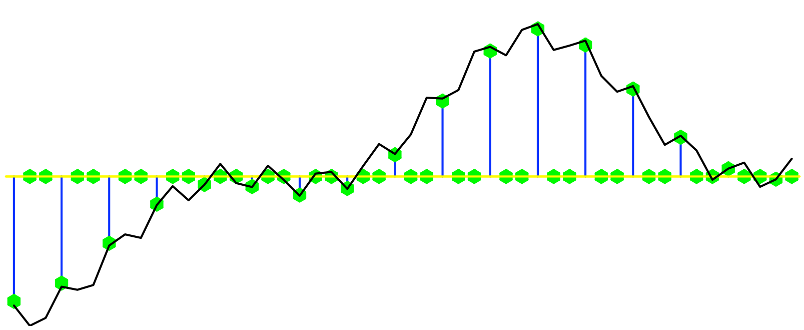
3-fold under-sampling



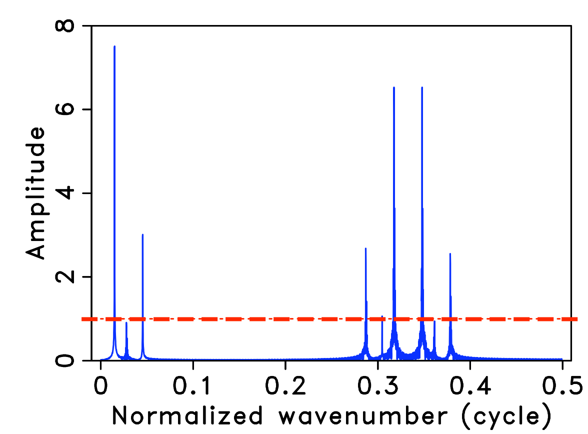
Fourier → transform



✓ significant coefficients detected



Fourier → transform



✗ ambiguity

[Hennenfent & Herrmann, '08]

Sparse one-norm recovery

SLIM

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 $\mathrm{n} \ll N$

Study recovery as a function of

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

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

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**

Let's adapt this theory to seismic acquisition and processing

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

Extension

Extend CS framework:

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

restriction matrix   measurement matrix   sparsity matrix

Expected to perform well when

$$\mu = \max_{1 \le i \ne j \le N} |\left(\mathbf{R}\mathbf{M}\mathbf{s}^i\right)^H \mathbf{R}\mathbf{M}\mathbf{s}^j|$$

Generalizes to *redundant* transforms for cases where

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

Open research topic...

Empirical performance analysis

Selection of the appropriate sparsifying transform

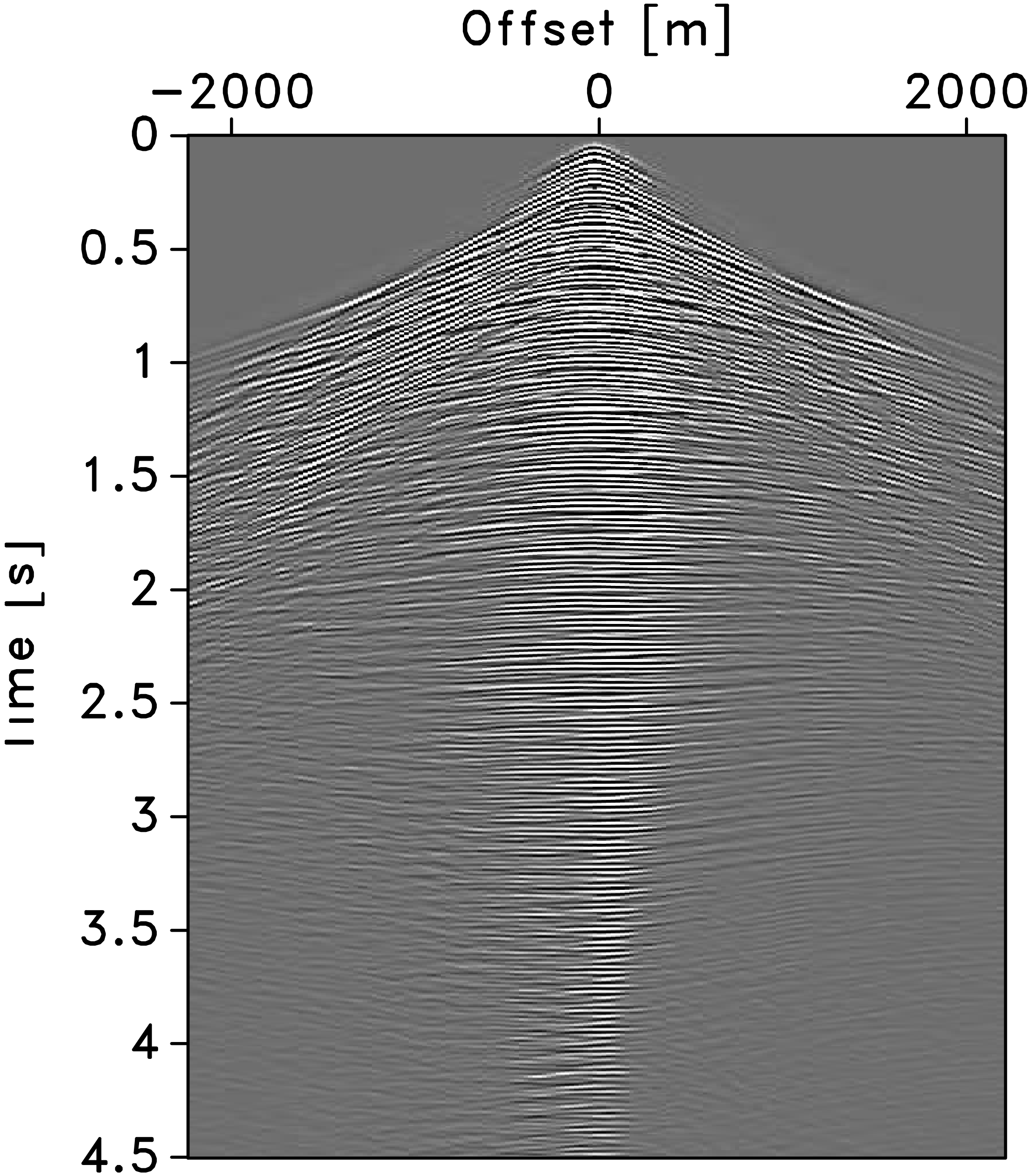
- nonlinear approximation error

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

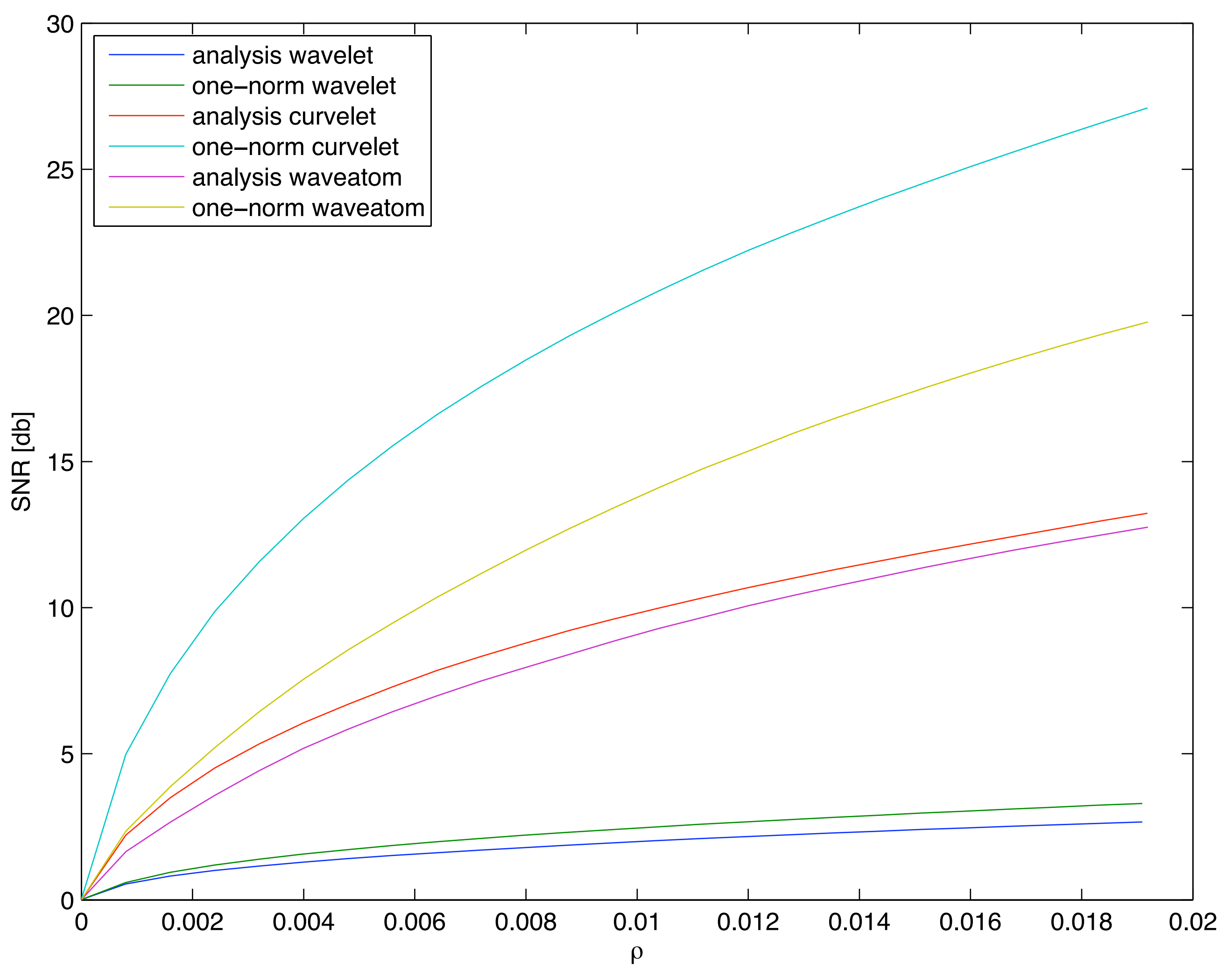
[FJH, '10]

Nonlinear approximation error

common receiver gather

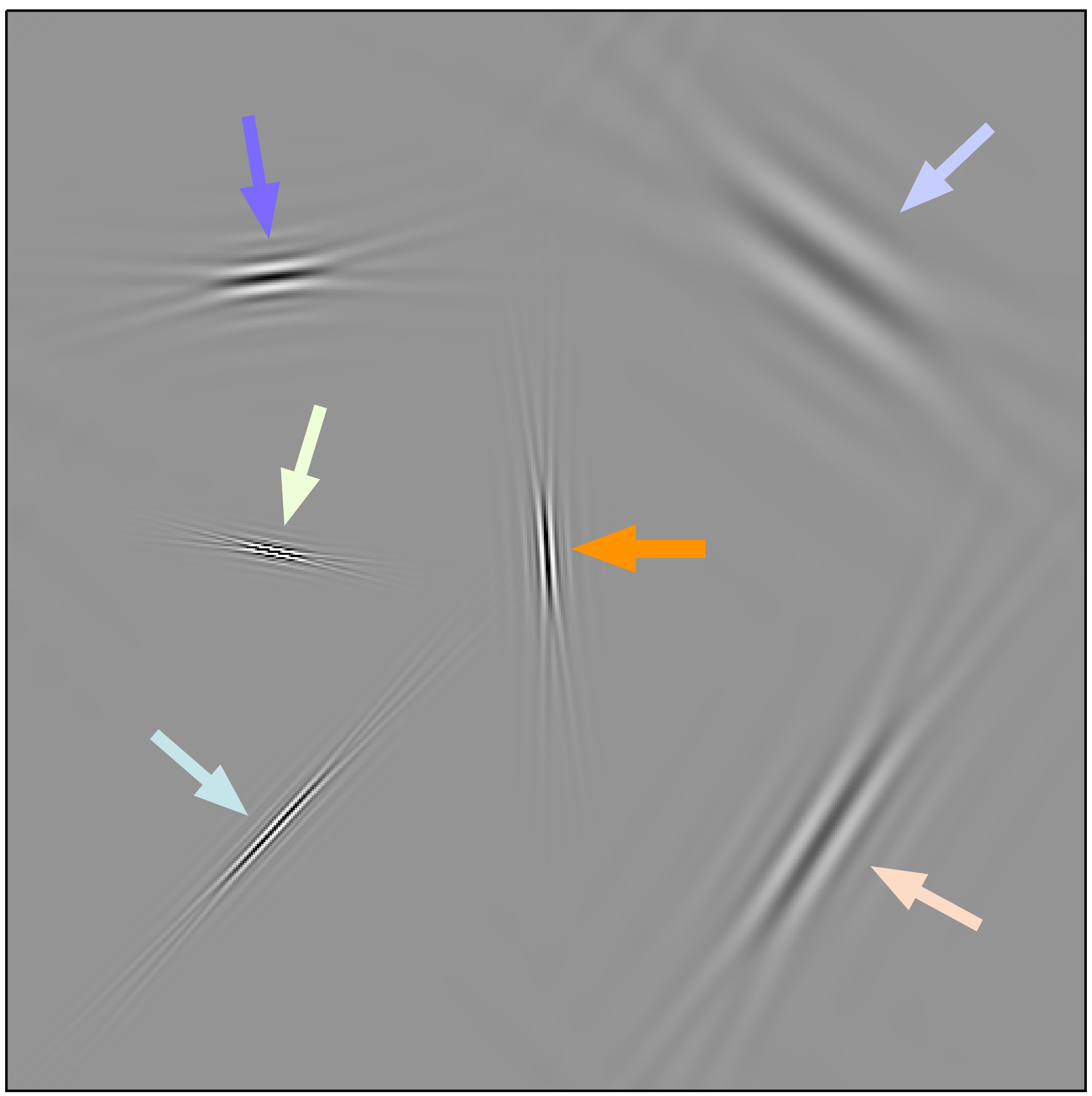


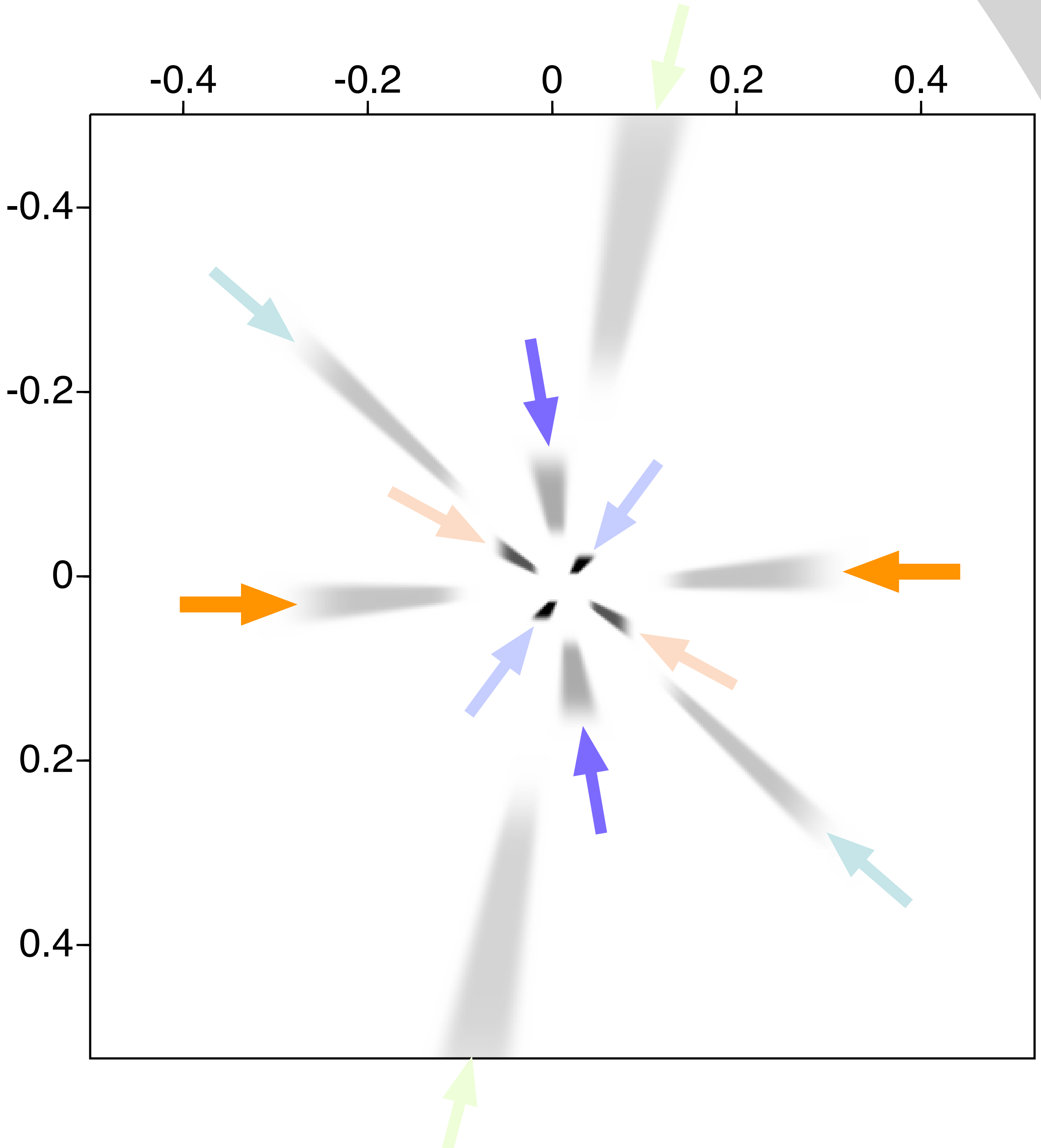
recovery error



[Demanet et. al., '06]

Curvelets





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*

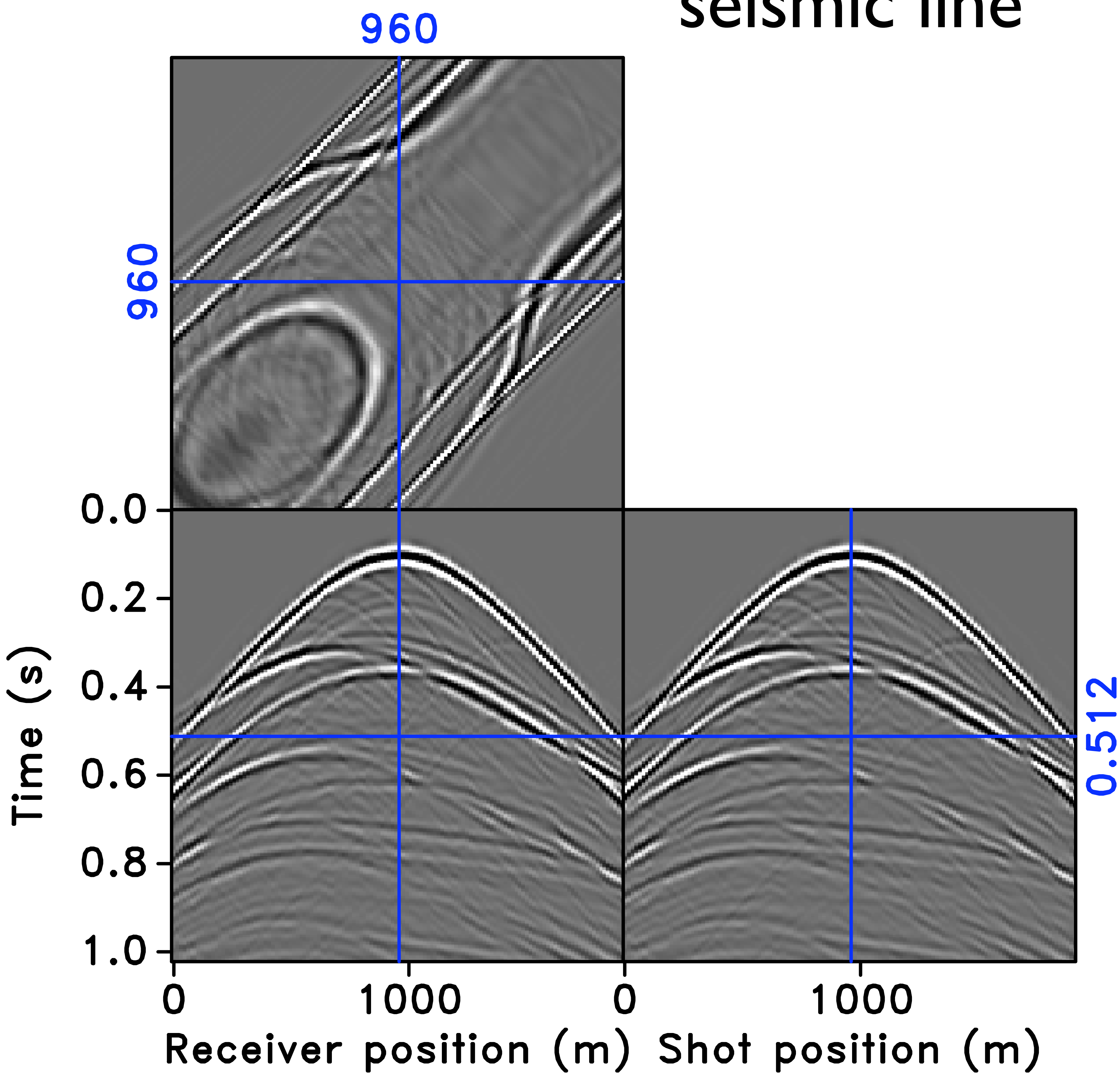
- requires few matrix-vector multiplications

Case study I

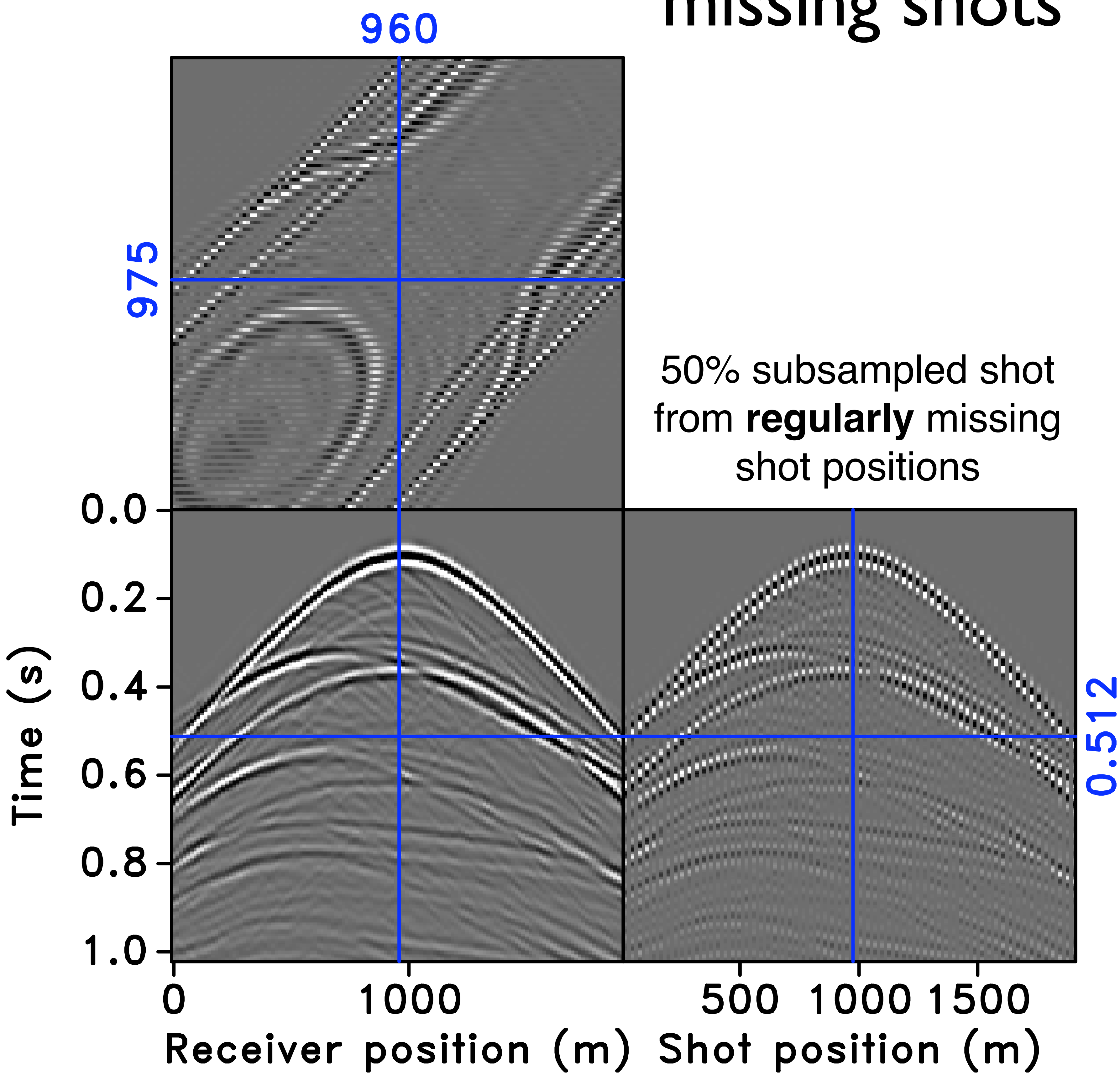
Acquisition design according to Compressive Sensing

- ***Periodic*** *subsampling* vs ***randomized jittered*** sampling of ***sequential*** sources in 2- and 3-D

seismic line

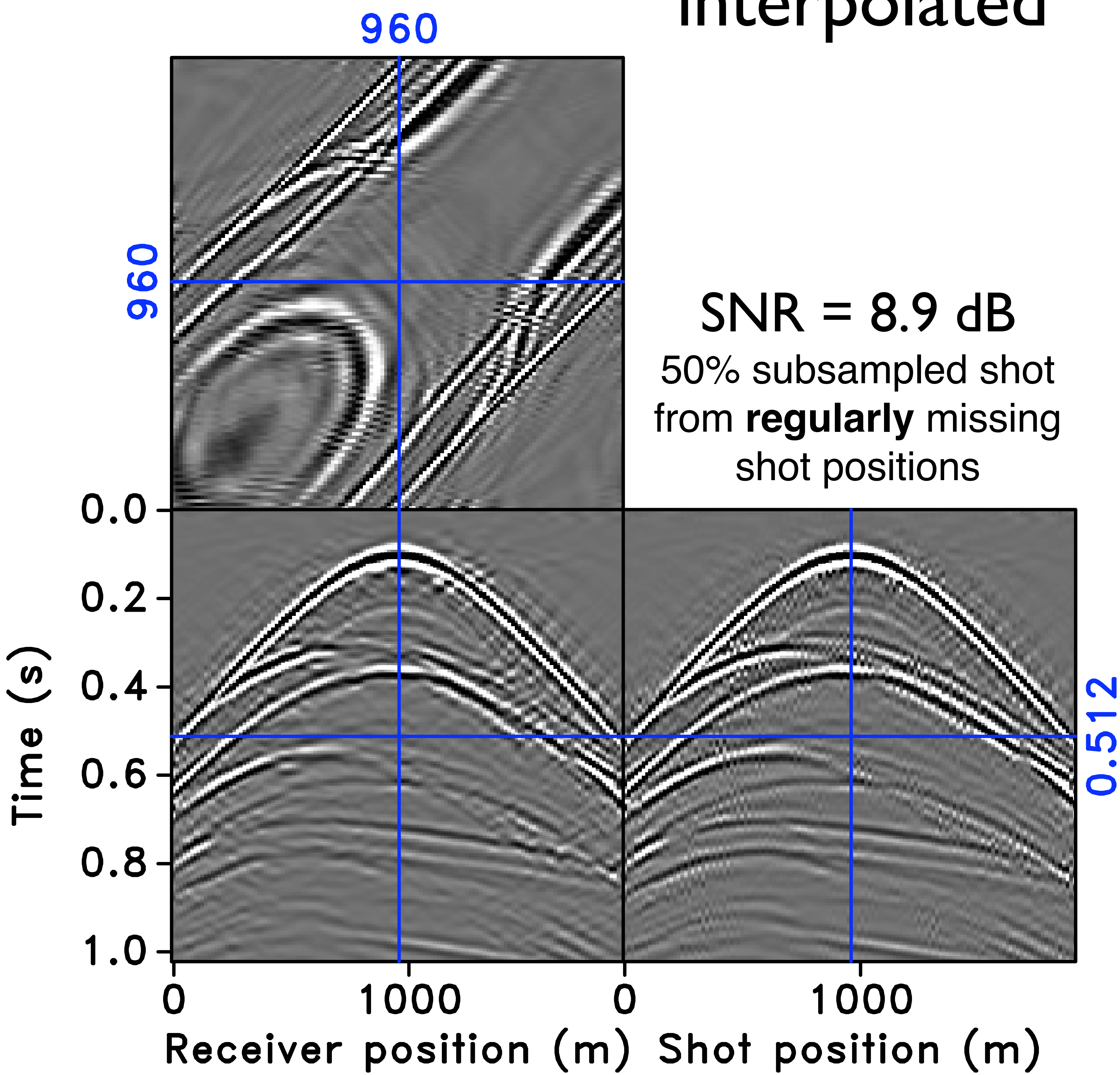


missing shots



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

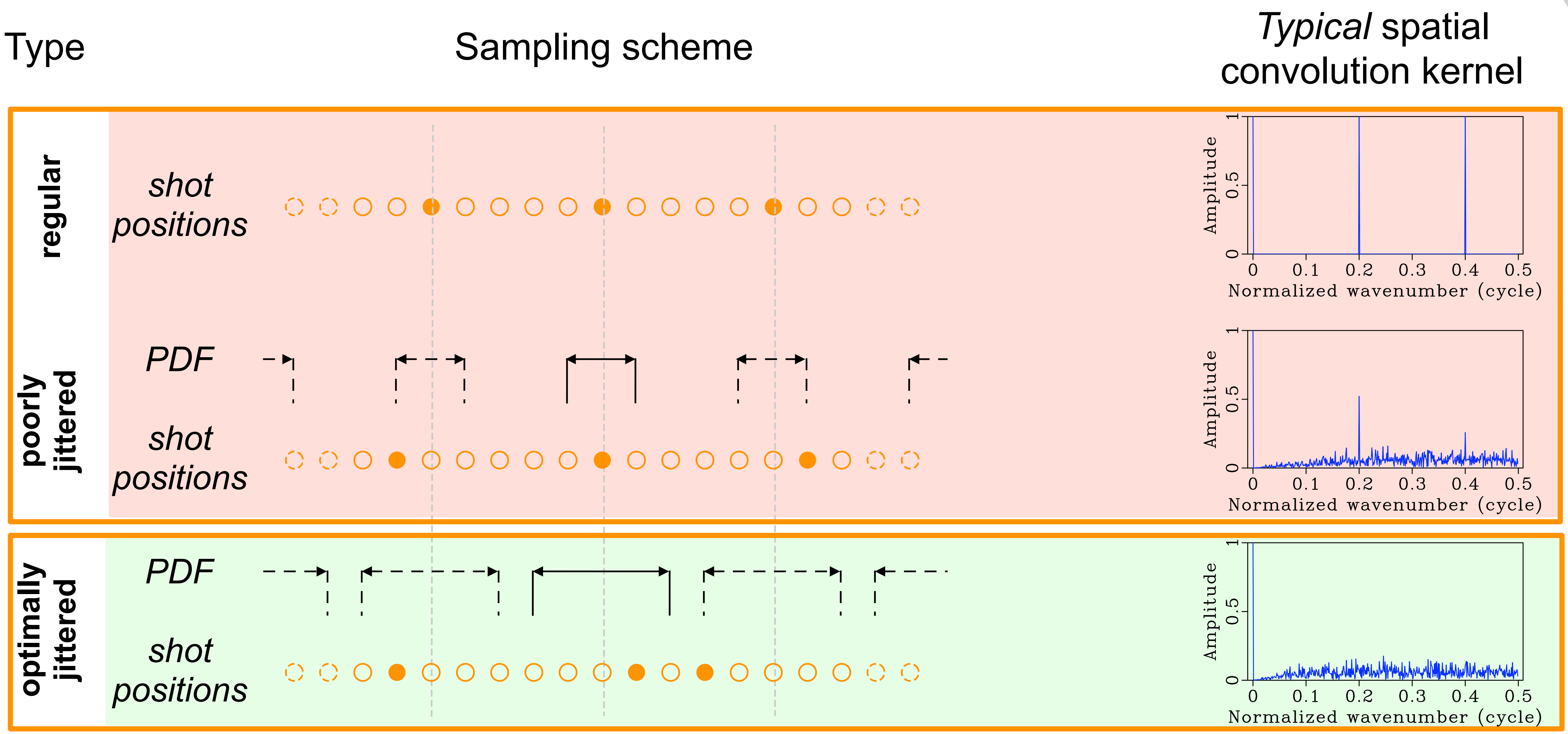
interpolated



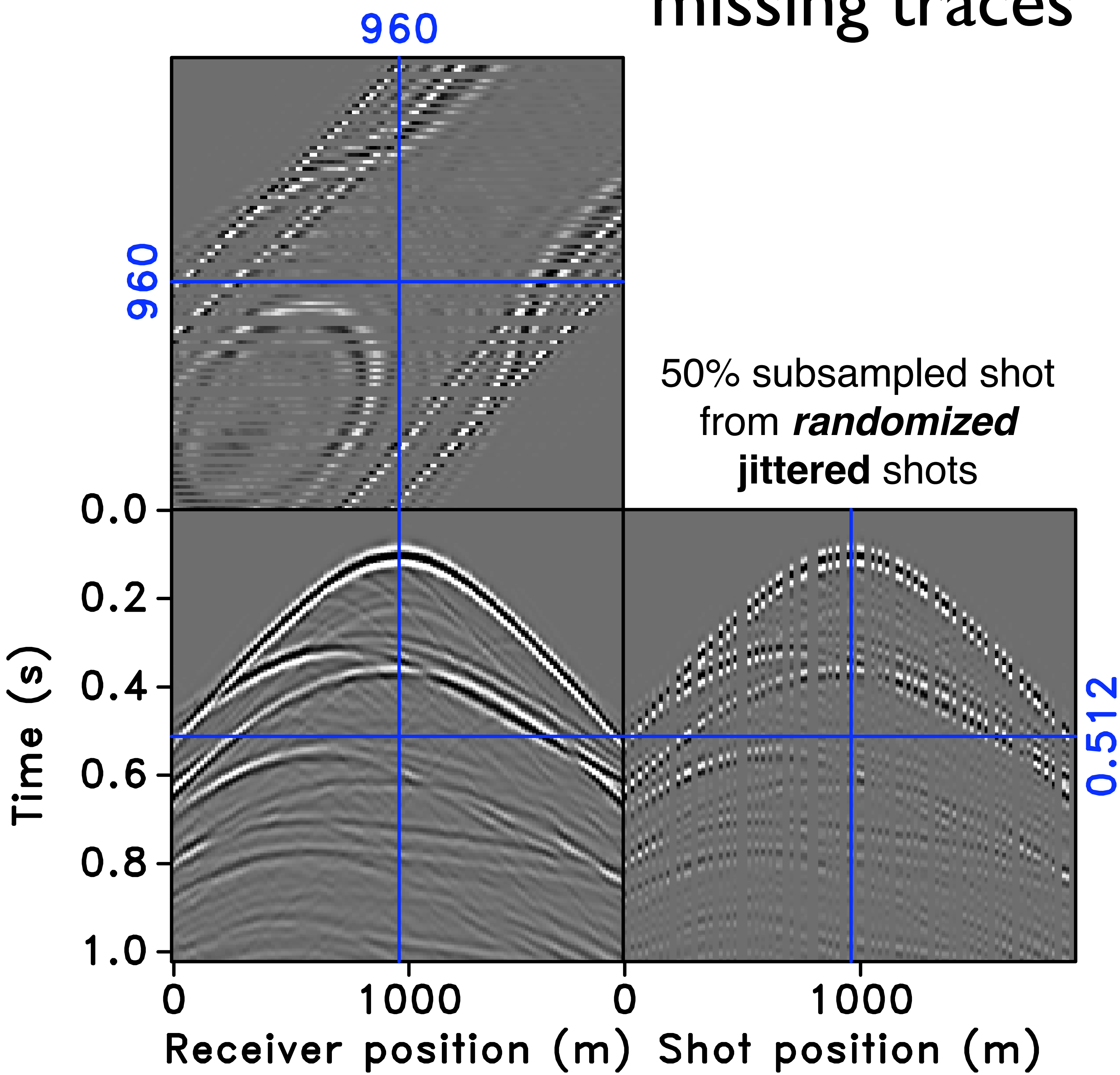
SNR = 8.9 dB

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

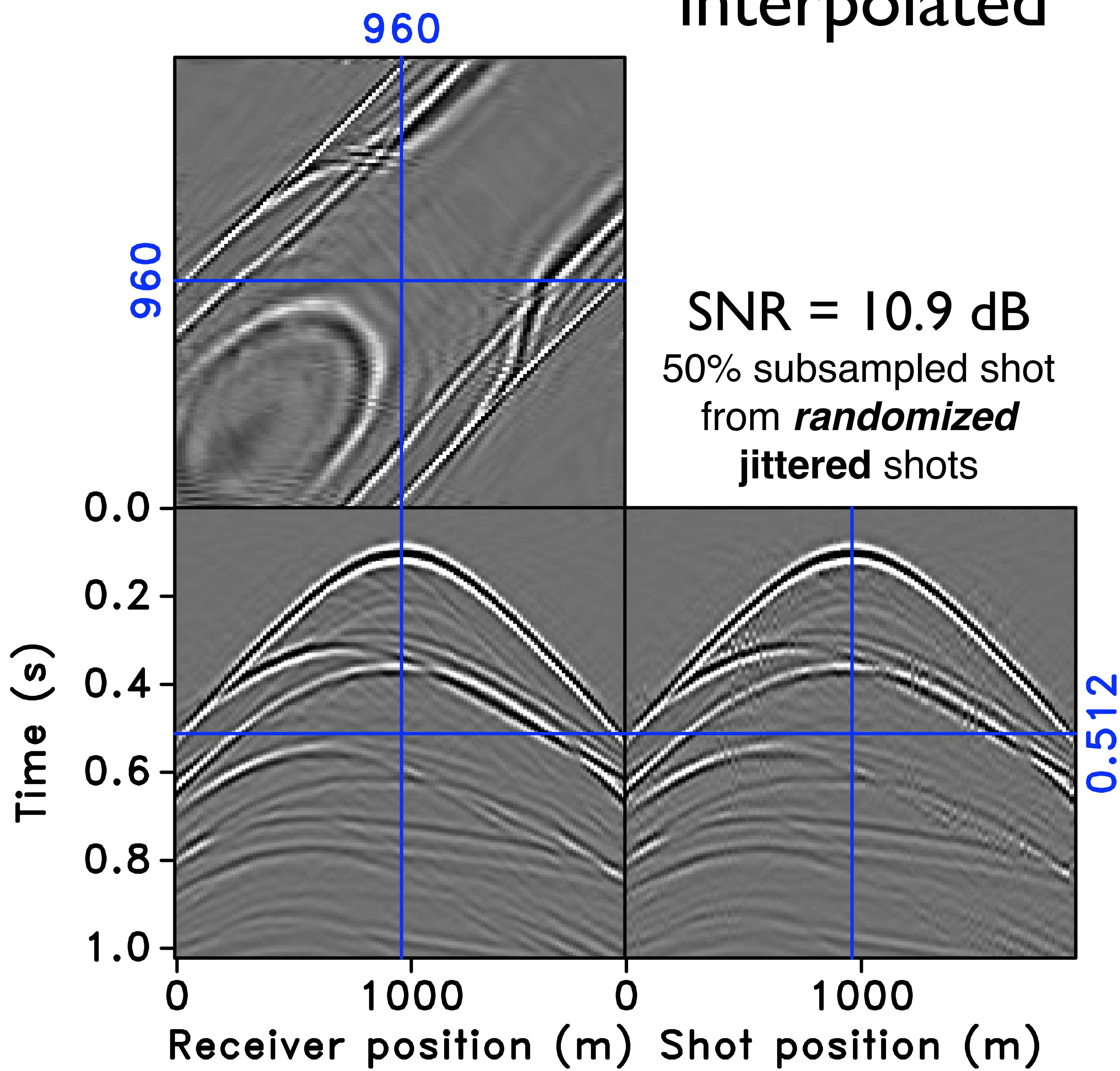
Jittered sampling



missing traces



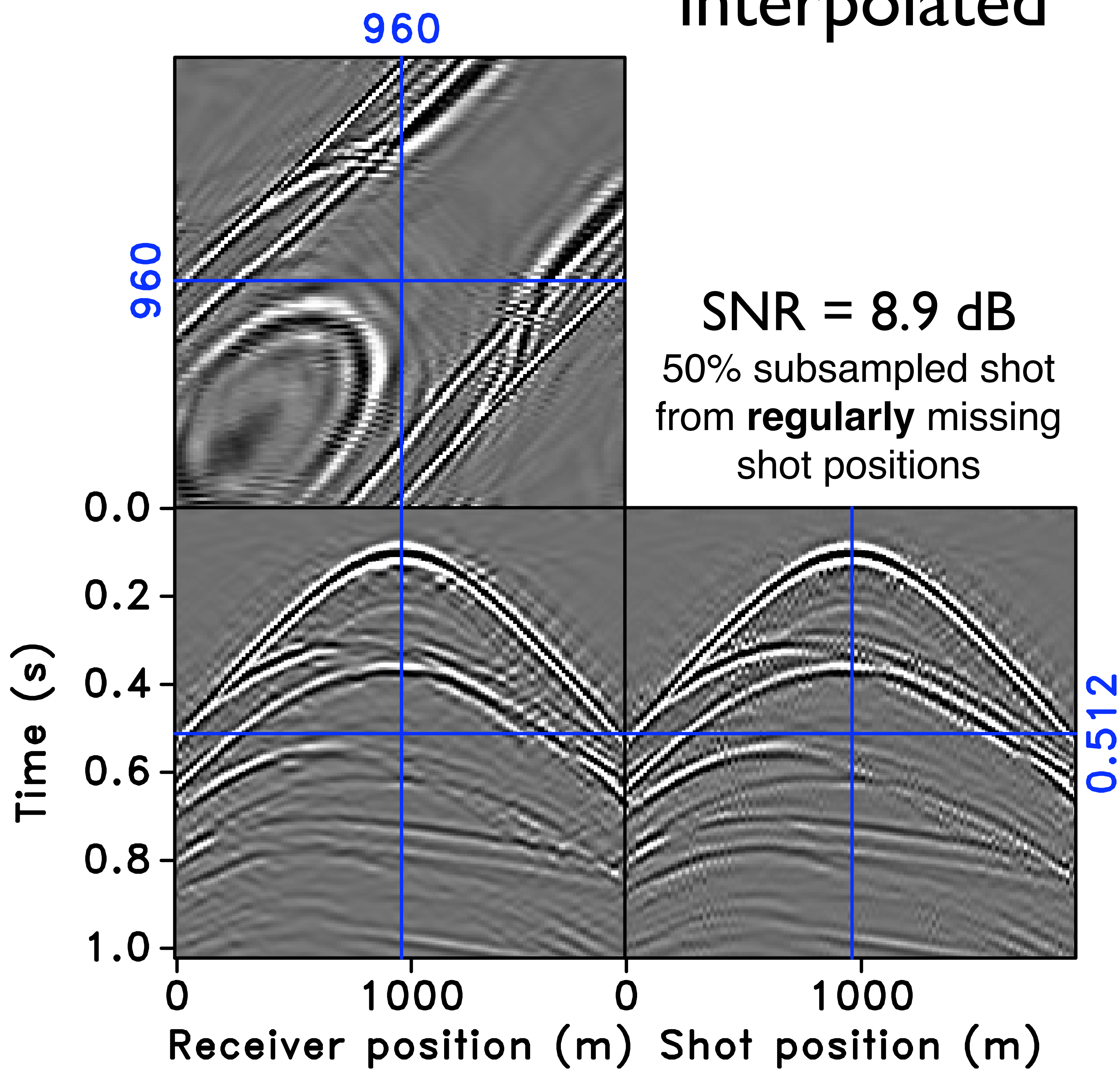
interpolated



SNR = 10.9 dB

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

interpolated



SNR = 8.9 dB

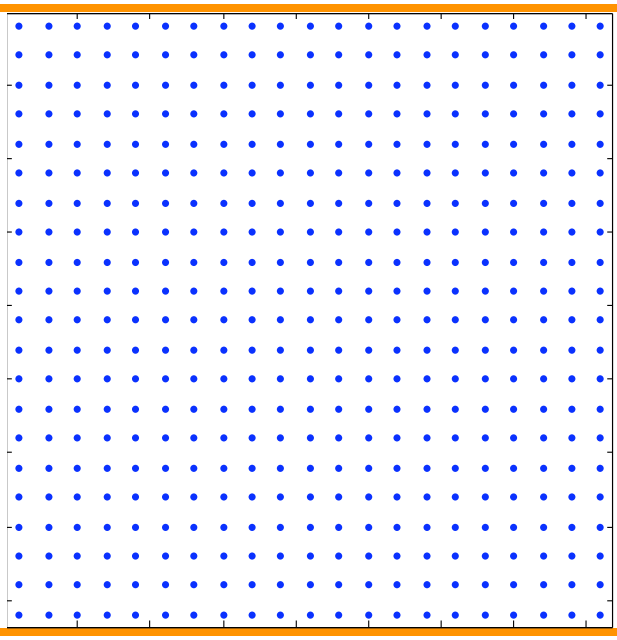
50% subsampled shot from **regularly** missing shot positions

1 & 2-D jittered samplings

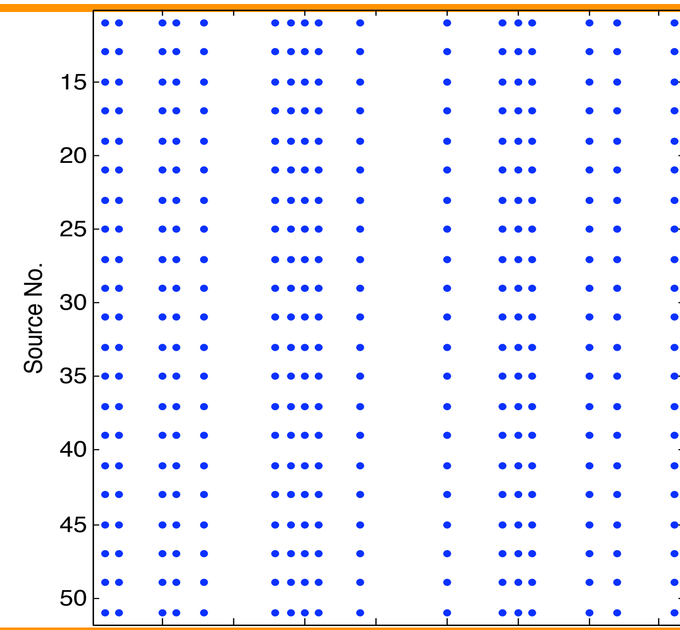
[Tang et. al., '09-'10]

SLIM

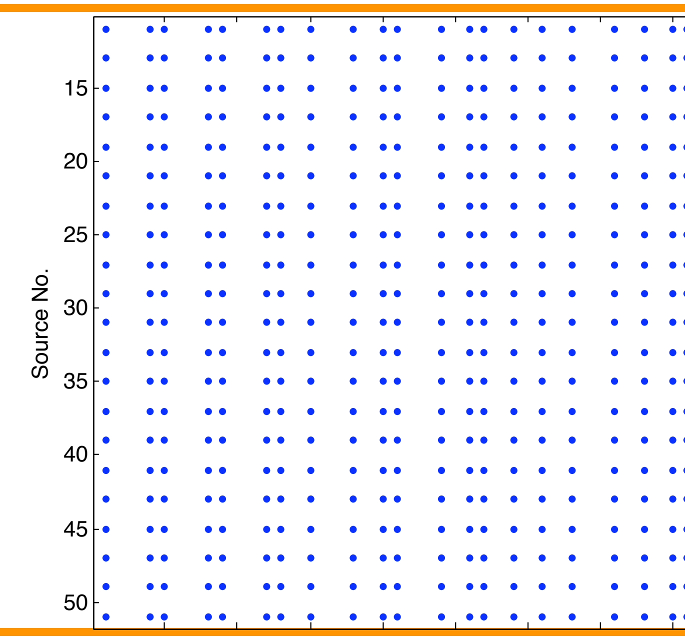
regular



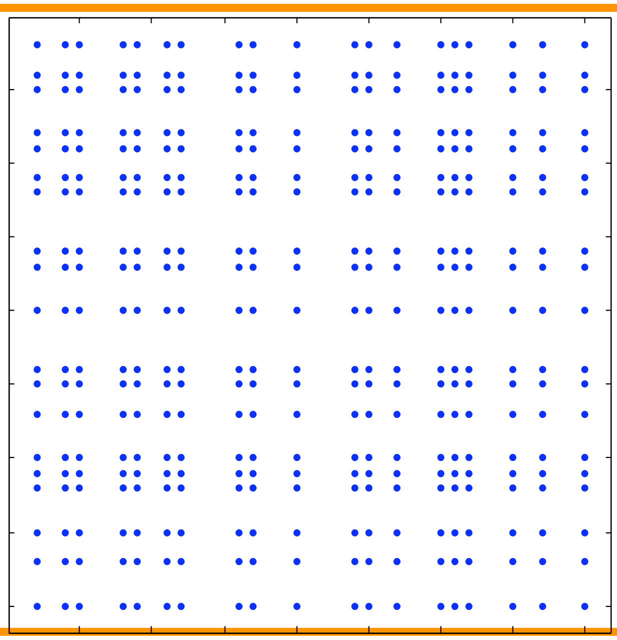
uniform



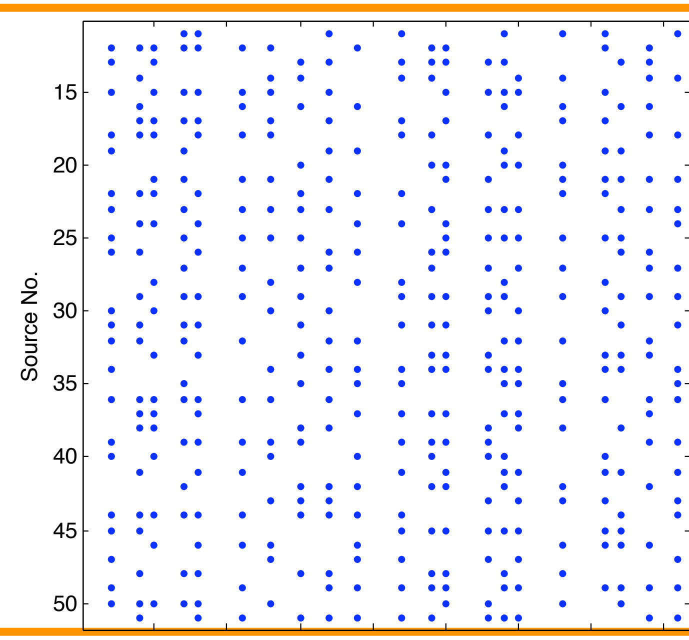
jittered



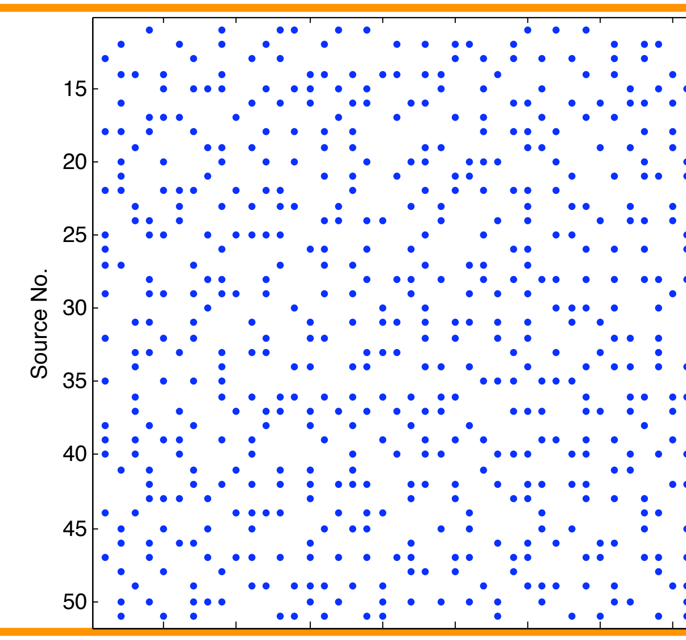
separable 2d jittered



non-seperable 2d jittered



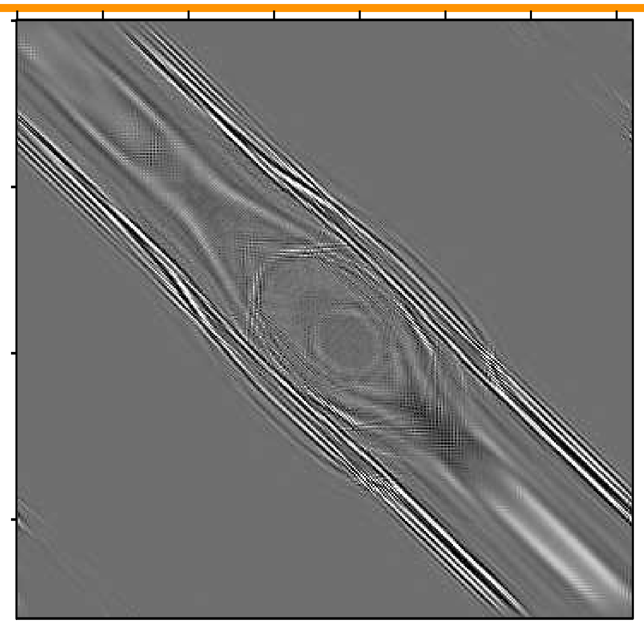
fully 2d jittered



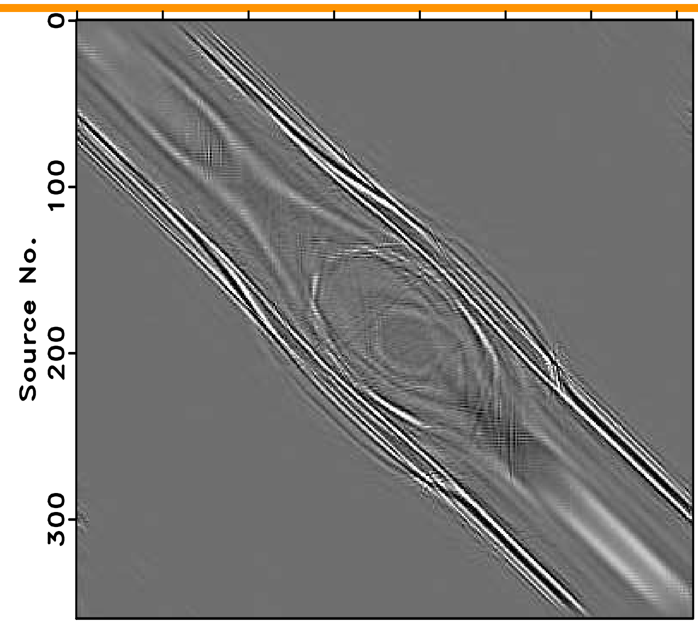
Spectra become increasingly *"blue"*

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

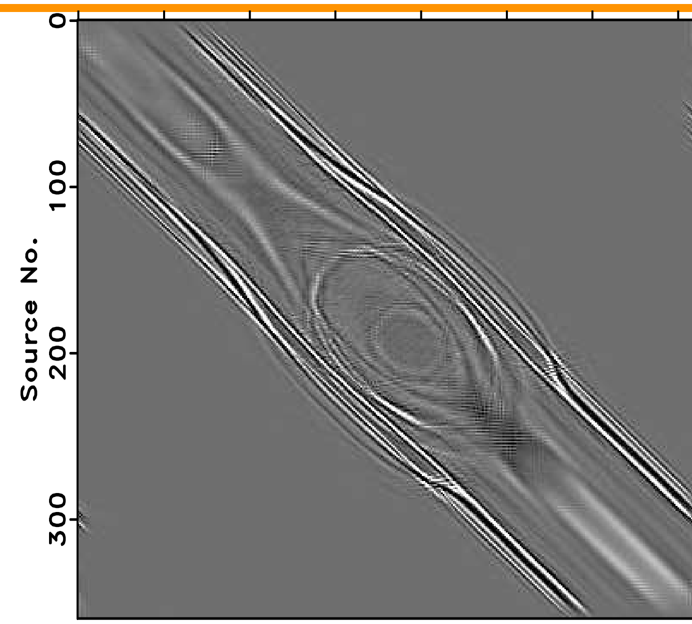
regular
(3.91 dB)



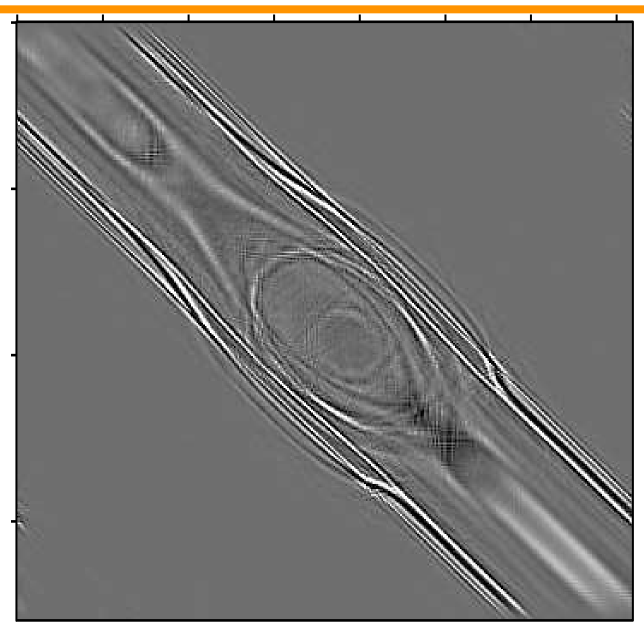
uniform
(7.20 dB)



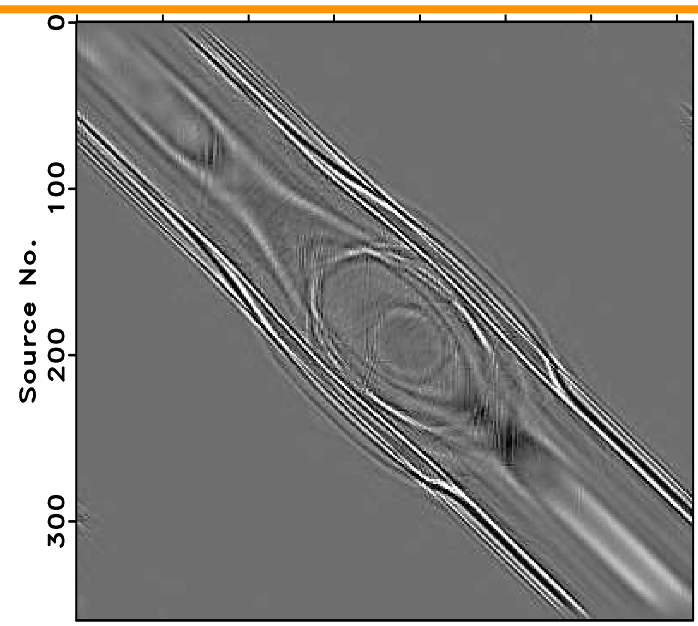
jittered
(8.94 dB)



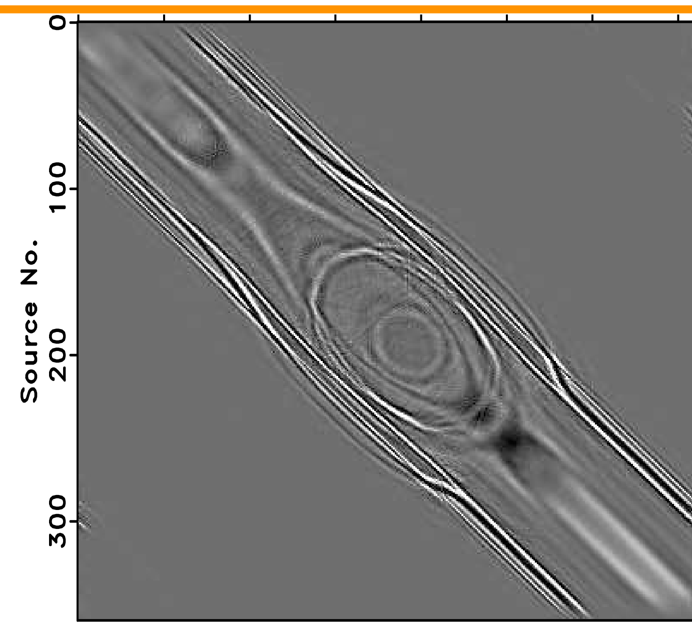
separable
2d jittered (9.45 dB)



non-seperable
2d jittered (10.03 dB)



fully 2d
jittered (10.86 dB)



Spectra become increasingly *"blue"*

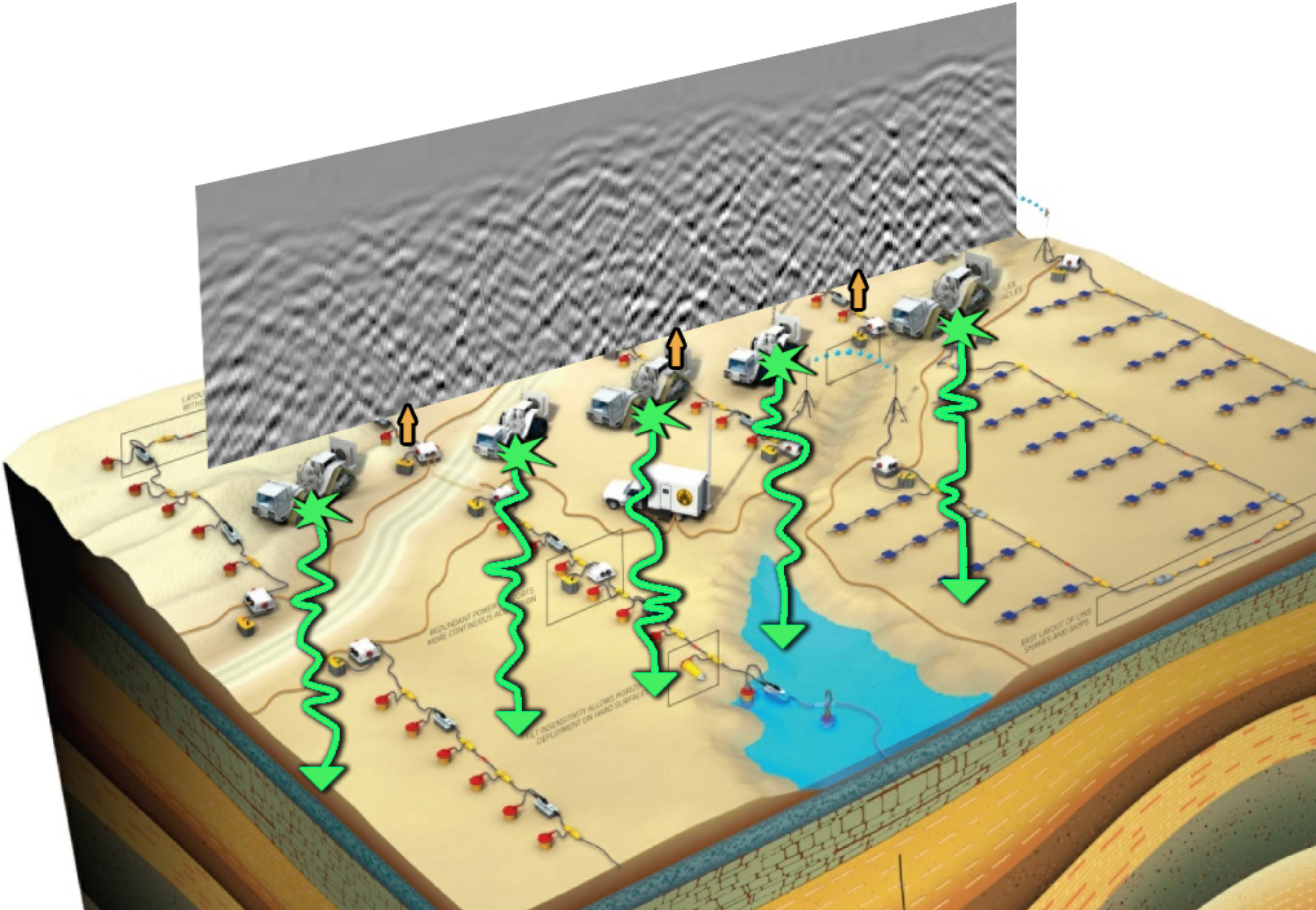
Case study II

[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

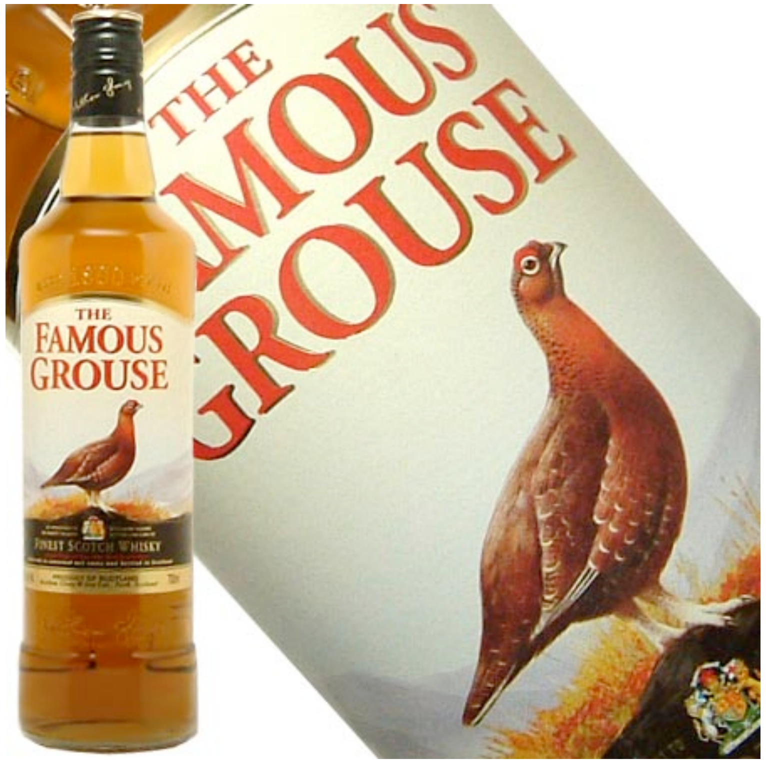
Simultaneous & incoherent sources



Unblending/ Demultiplexing

✓

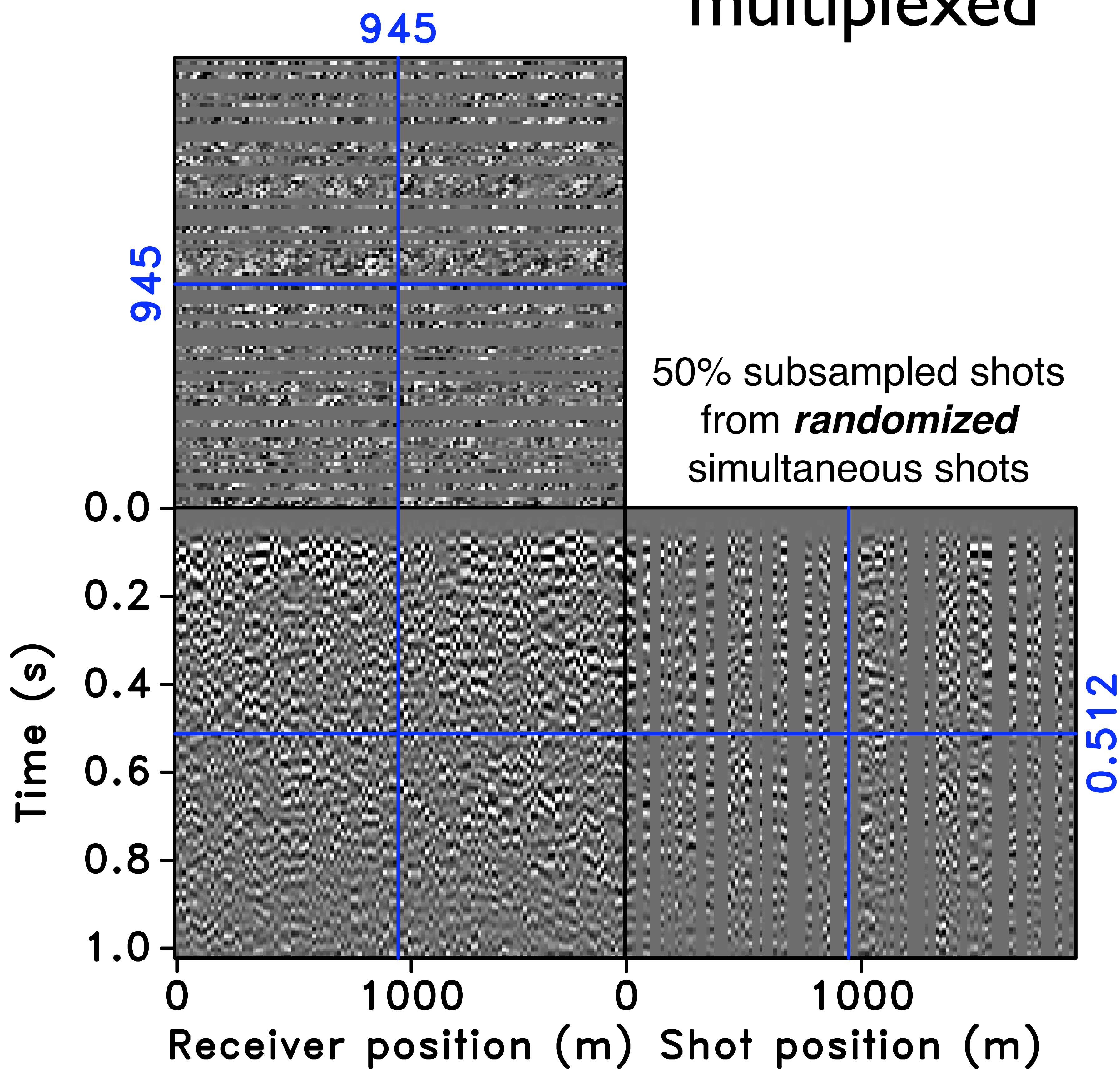
?



$$$$$$$$$$ $

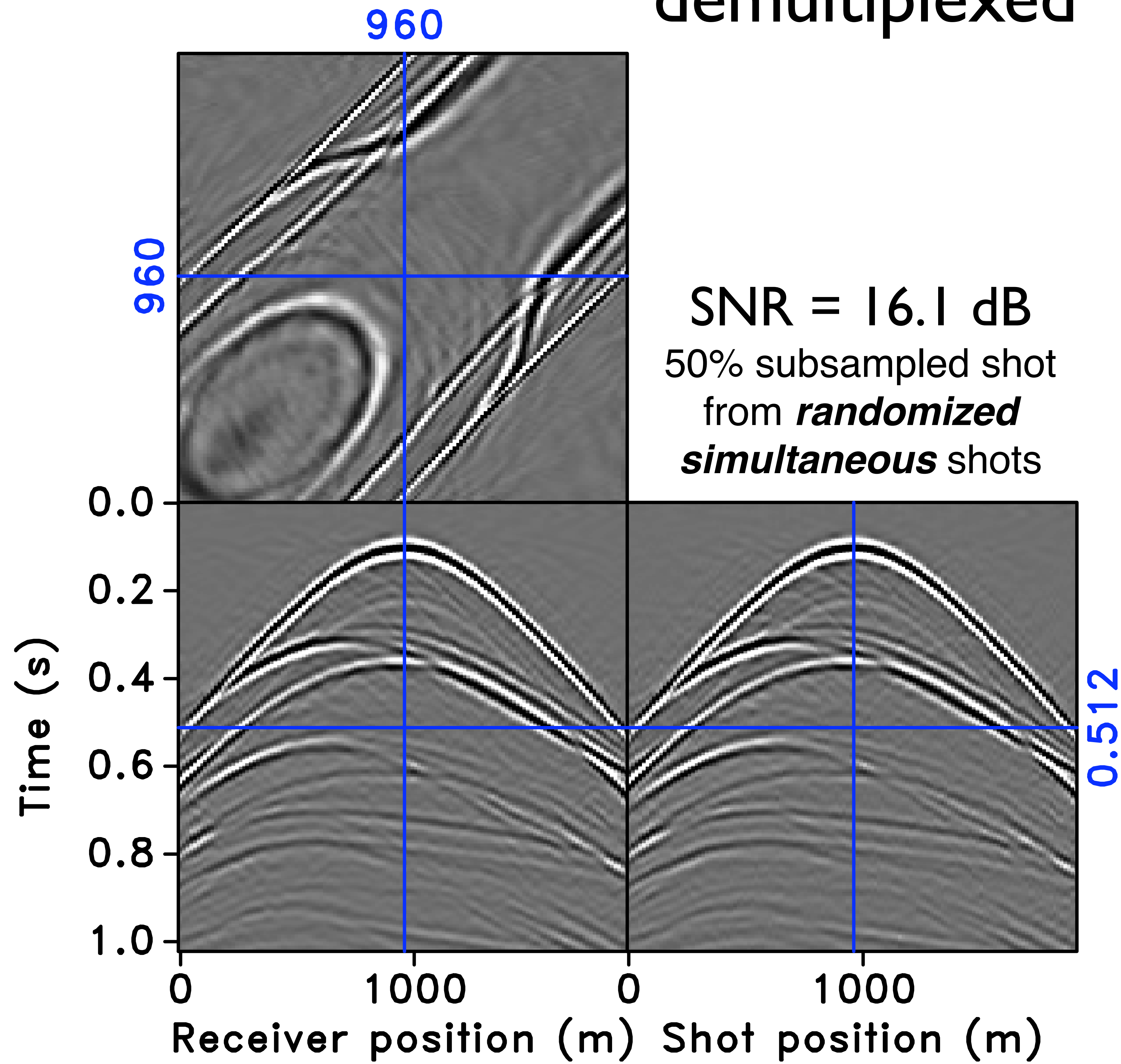
Blending versus *unblending* ...

multiplexed



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

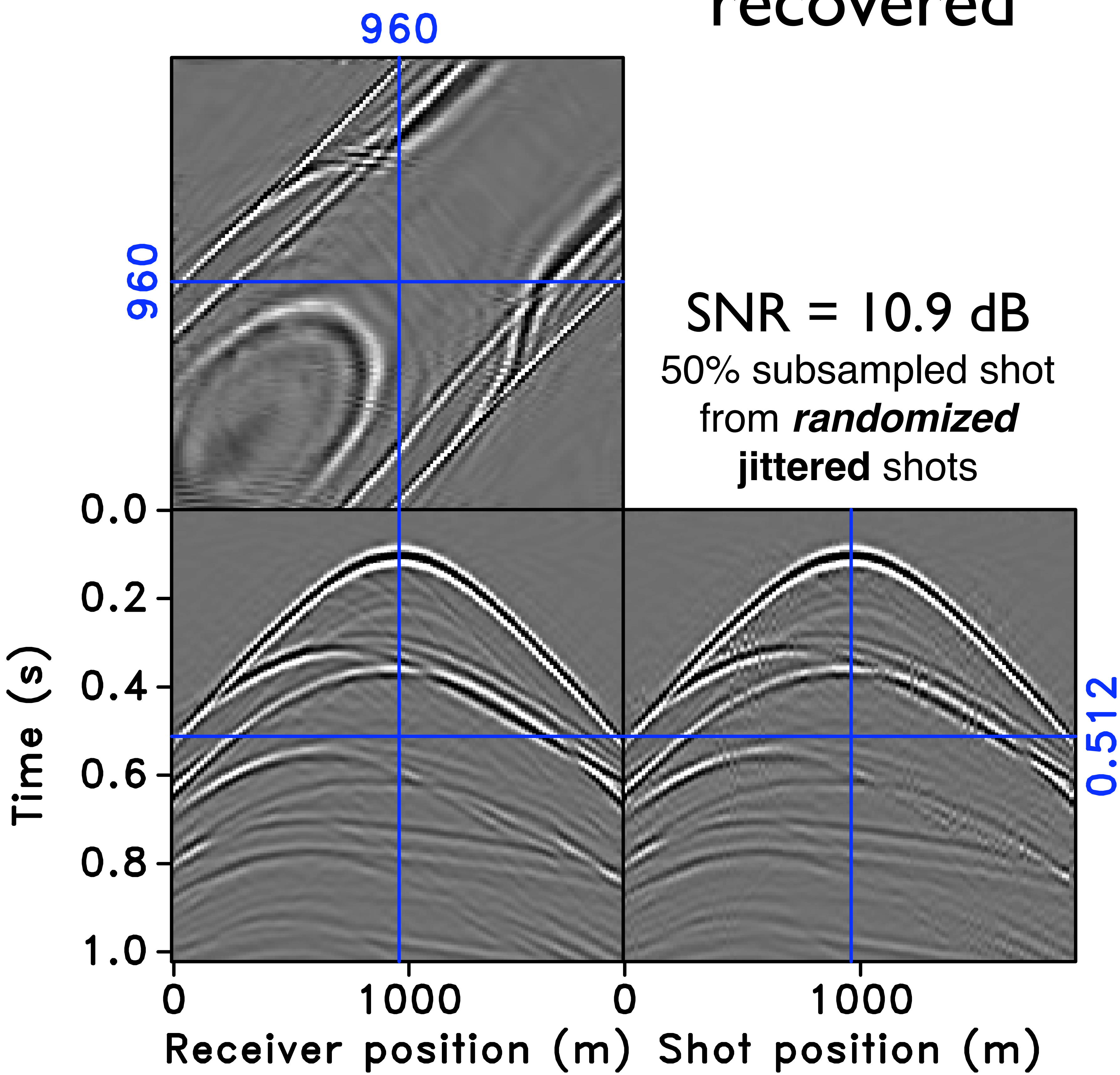
demultiplexed



SNR = 16.1 dB

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

recovered



SNR = 10.9 dB

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

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

$$\mathrm{SNR}(\delta) = -20\log\frac{\|\mathbf{f}-\tilde{\mathbf{f}}_\delta\|}{\|\mathbf{f}\|}$$

wi[redacted] $\delta =$

[FJH, '10]

Multiple experiments

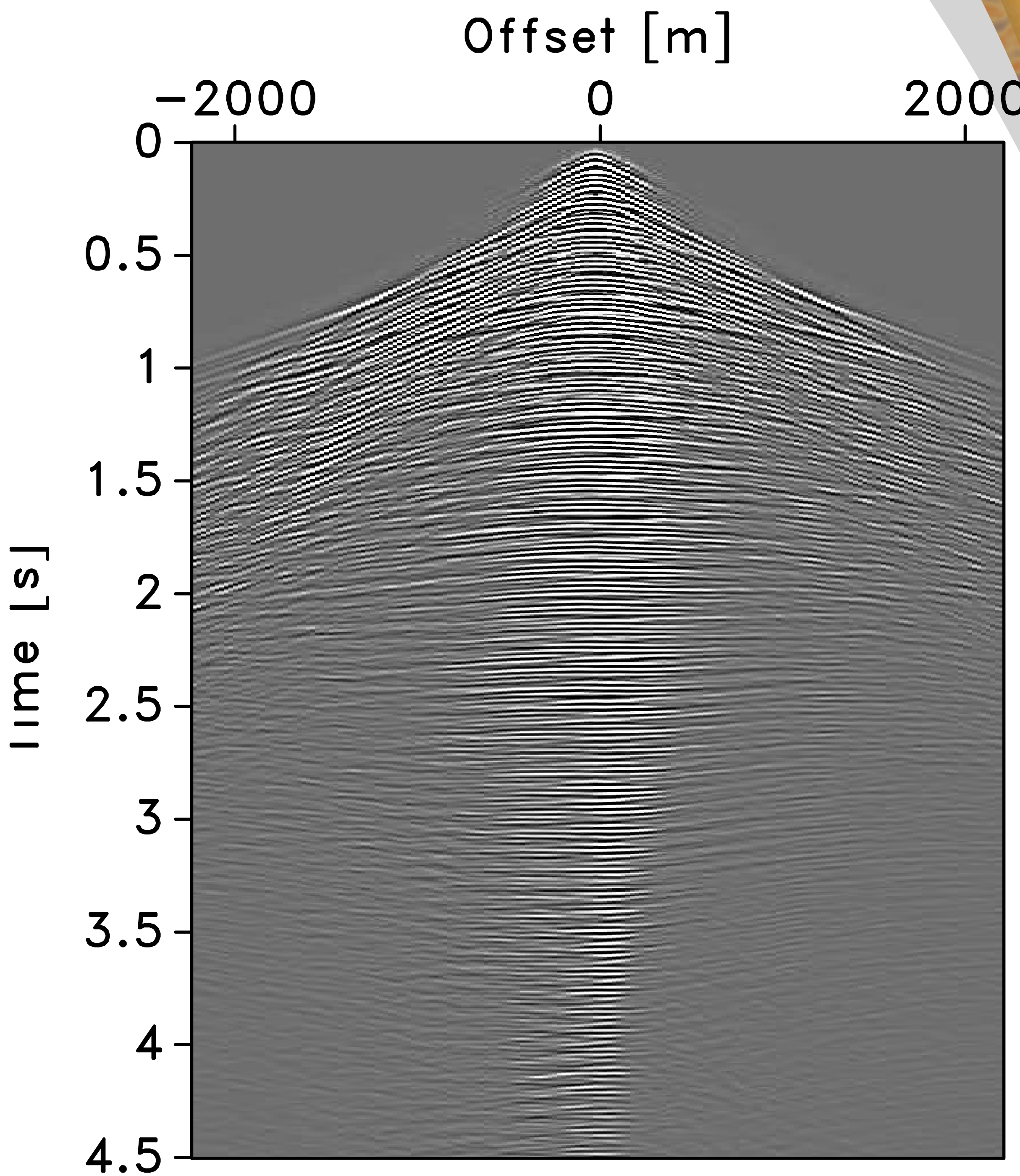
Generate 25 random experiments for varying subsampling ratios:

- sequential sources
- simultaneous sources

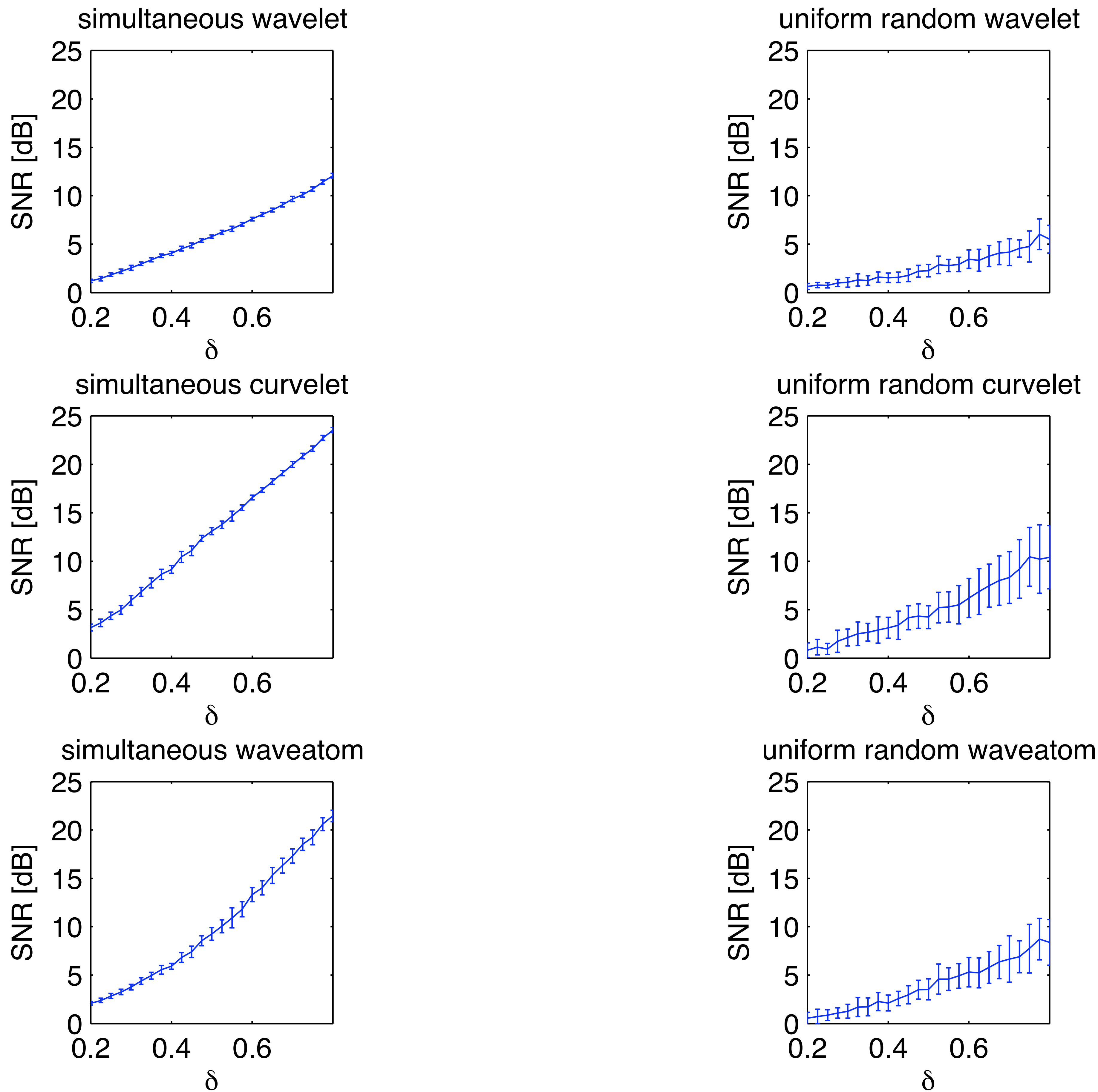
Study recovery errors & oversampling ratios for

- Wavelets
- Curvelets
- Waveatoms

common receiver gather



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

$$\mathrm{SNR}(\delta) = -20\log\frac{\|\mathbf{f}-\tilde{\mathbf{f}}_\delta\|}{\|\mathbf{f}\|}$$

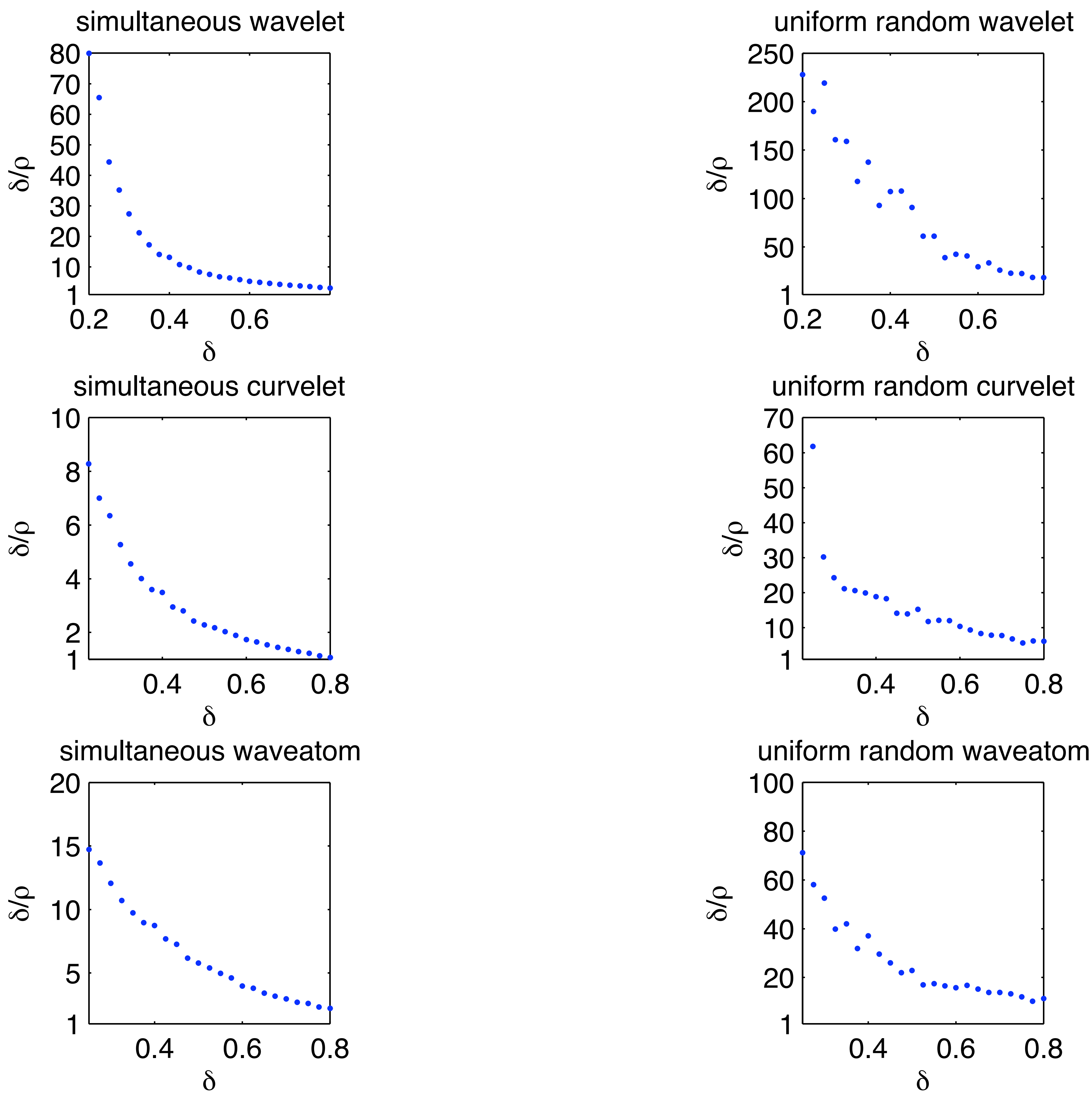
wi[redacted] $\delta =$

- 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

- [x] *sparsifying transform*

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

- [x] *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

Reality check

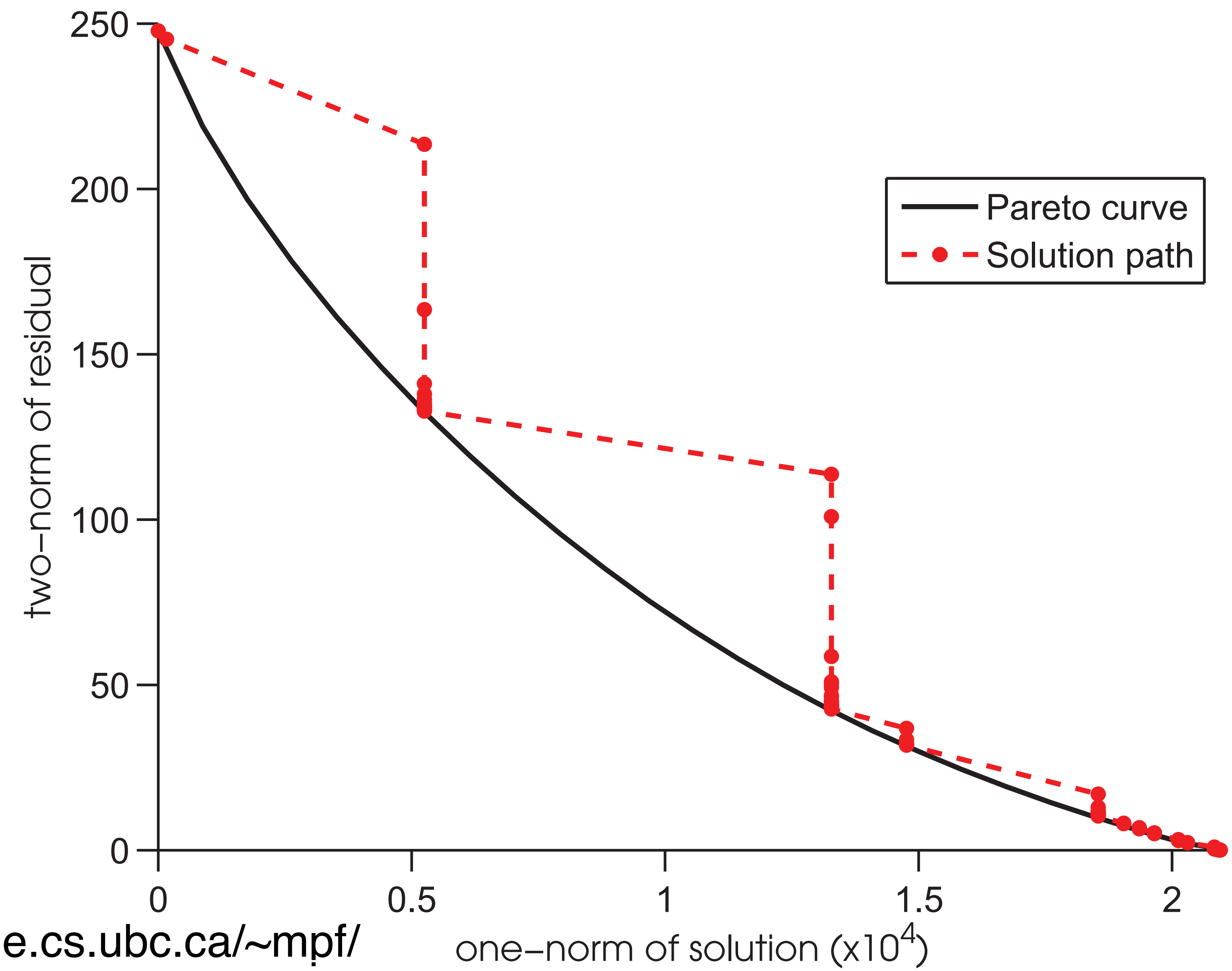
"When a traveler reaches a fork in the road, the l_1-norm tells him to take either one way or the other, but the l_2 -norm instructs him to head off into the bushes."

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

☑ *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

Observations

Controllable error for reconstruction from *randomized* subsamplings

Curvelets and *simultaneous* acquisition perform the best

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*

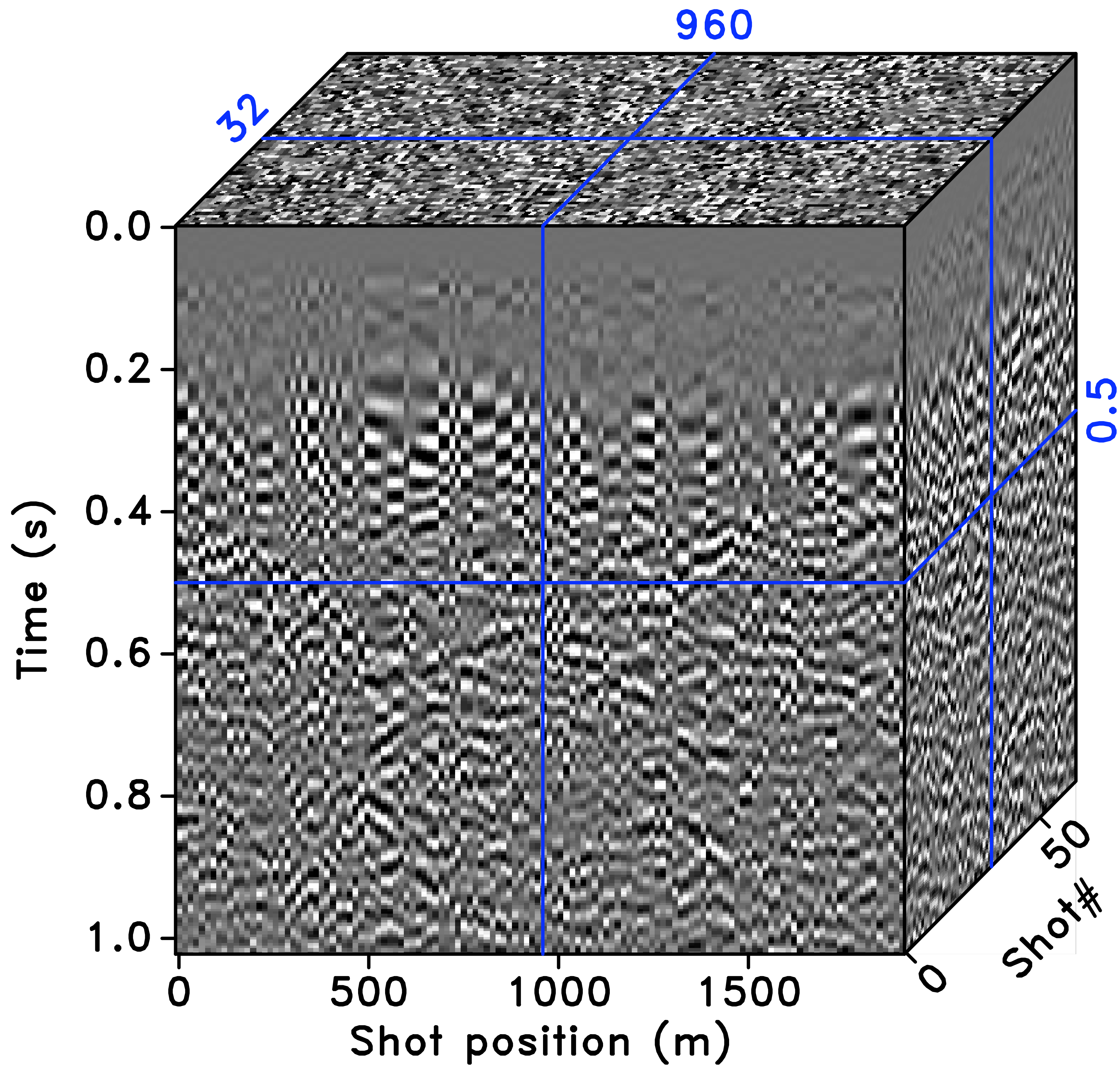
Case study II

Processing according to CS

- CS recovery from *simultaneous* data, followed by *primary* estimation

vs.

- *Primary* estimation *directly* from *simultaneous* data

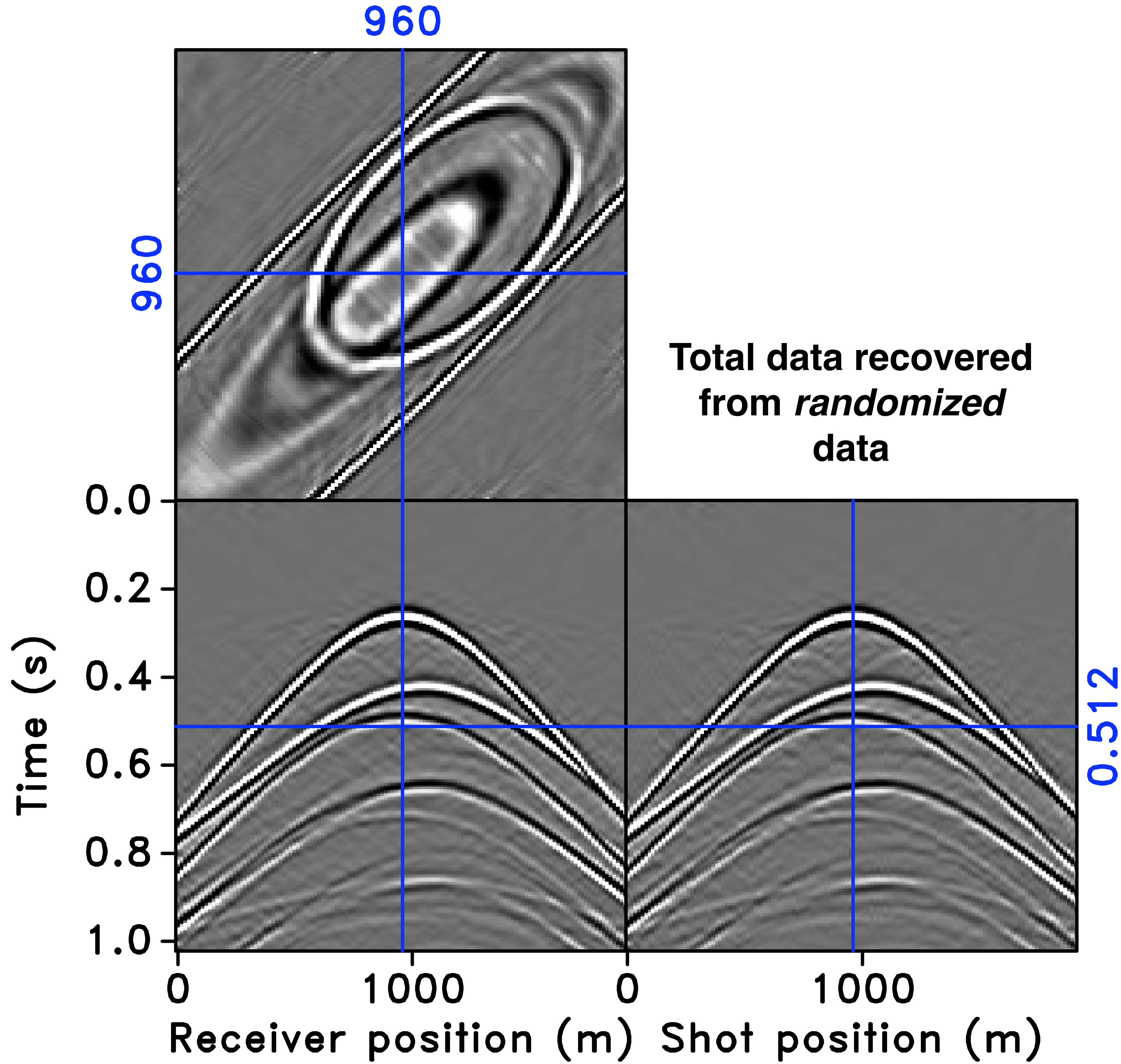

960
32
0.5
0.0
0.2
0.4
0.6
0.8
1.0
Time (s)
0
500
1000
1500
Shot position (m)
0
50
Shot#

CS

Use to demultiplex

$$\mathbf{A} = \left(\mathbf{R}\begin{bmatrix}\textbf{Gaussian}\\ \textbf{matrix}\end{bmatrix} \otimes \mathbf{I}\right)\mathbf{S}^*$$

(Randomized simultaneous sources)



Total data recovered from *randomized* data

[Groenestijn et. al. '09]
[Lin and Herrmann, '09]

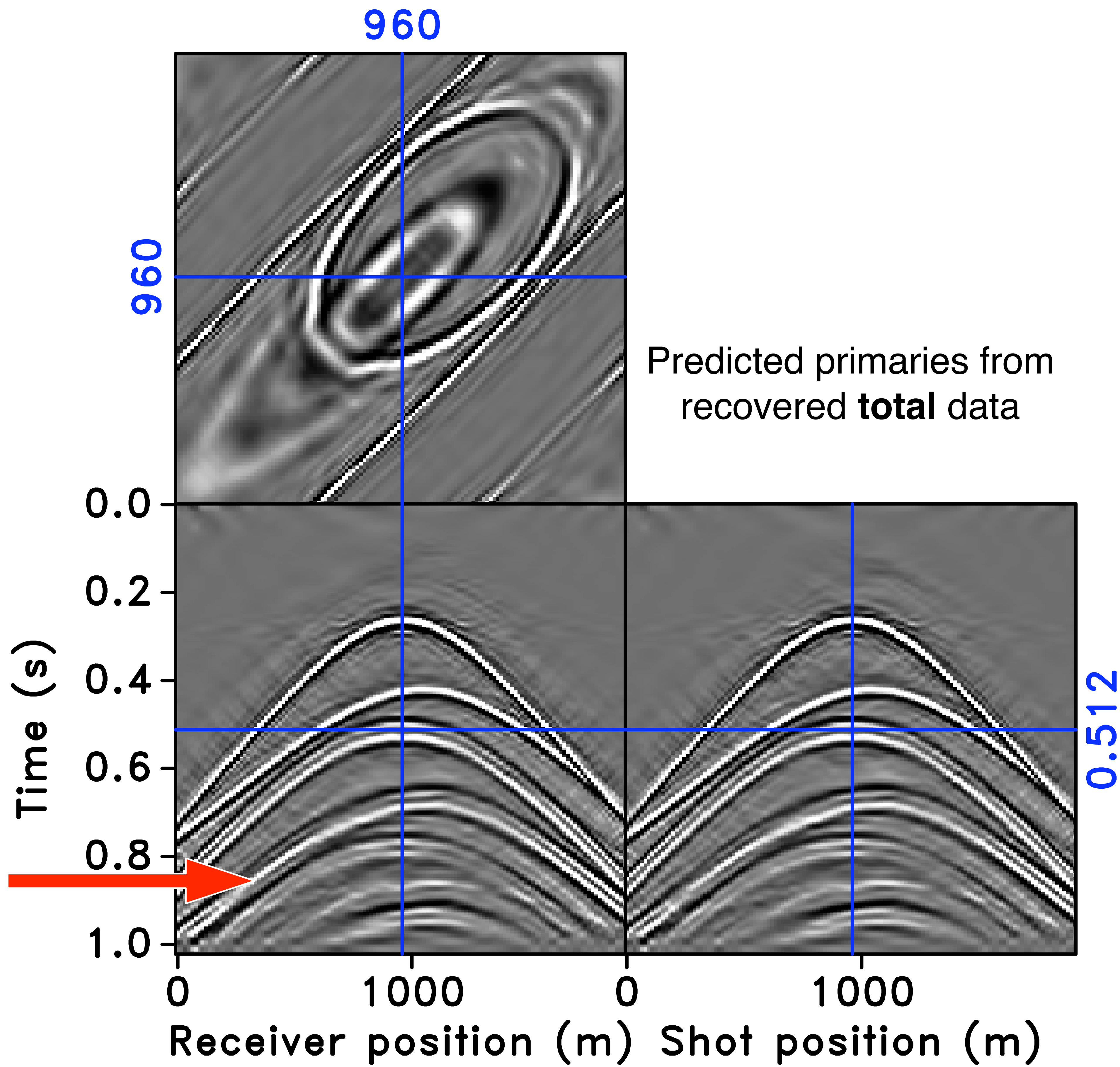
SLIM

Physical principle

Modeling the surface:

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

Inversion "focusses" multiples onto primaries ...



Predicted primaries from
recovered **total** data

[Lin & FJH, ’09]

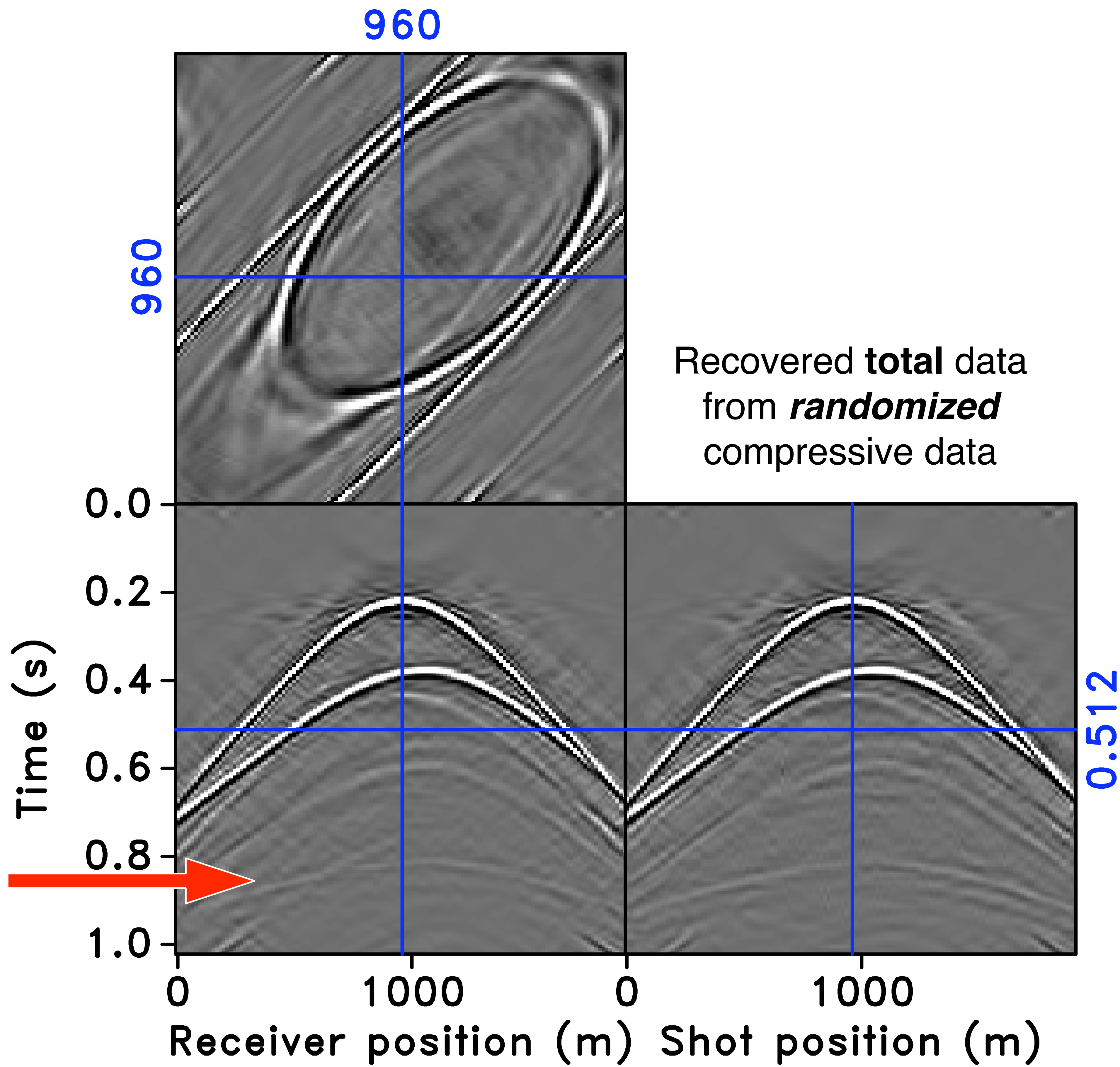
SLIM

Extension CS

Use to demultiplex & predict

$$\mathbf{A} = \Big[\mathbf{R}\Big]\Big[\text{Gaussian noise}\Big]\overbrace{\Big[\mathbf{M}\Big]}^{\text{randomized physics}}\Big[\mathbf{S}^{\dagger}\Big]$$

(**M** models free surface & source function)


960
960
Recovered **total** data
from ***randomized***
compressive data
0.0
0.2
0.4
0.6
0.8
1.0
Time (s)
0.512
0
1000
0
1000
Receiver position (m)
Shot position (m)

Conclusions

Sparse wavefield recovery benefits from

- *randomization*
- *sparsification*
- *inclusion of physics*

Recovery has a *controlable* error

Leads to acquisition & processing where costs are ***no*** *longer dominated by resolution & size* but by transform-domain sparsity & recovery error

Acknowledgments

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).

This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BP, Petrobras, and WesternGeco.

Thank you

slim.eos.ubc.ca

Further reading

Compressive sensing

- *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information by Candes, 06*.
- *Compressed Sensing* by D. Donoho, '06

Simultaneous acquisition

- *A new look at simultaneous sources* by Beasley et. al., '98.
- *Changing the mindset in seismic data acquisition* by Berkhout '08.

Transform-based seismic data regularization

- *Interpolation and extrapolation using a high-resolution discrete Fourier transform* by Sacchi et. al, '98
- *Non-parametric seismic data recovery with curvelet frames* by FJH and Hennenfent.,'07
- *Simply denoise: wavefield reconstruction via jittered undersampling* by Hennenfent and FJH, '08

Estimation of surface-free Green's functions:

- *Estimating primaries by sparse inversion and application to near-offset data reconstruction* by Groenestijn, '09
- *Unified compressive sensing framework for simultaneous acquisition with primary estimation* by T. Lin & FJH, '09

Review

- *Randomized sampling and sparsity: getting more information from fewer samples* by FJH, '09-'10