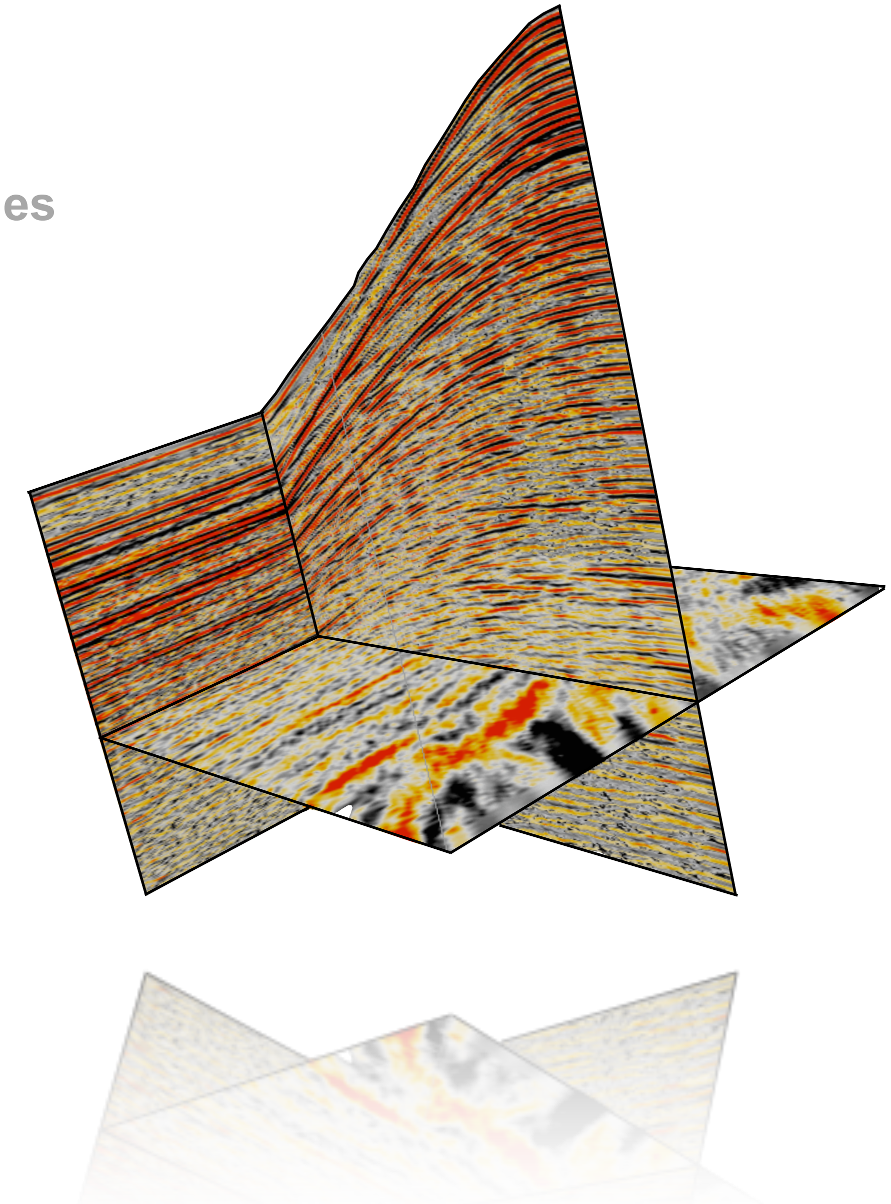
Beating Nyquist by randomized sampling

Felix J. Herrmann*
fherrmann@eos.ubc.ca

Joint work with Gang Tang, Reza Shahidi, Gilles Hennenfent, and Tim Lin

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

slim.eos.ubc.ca

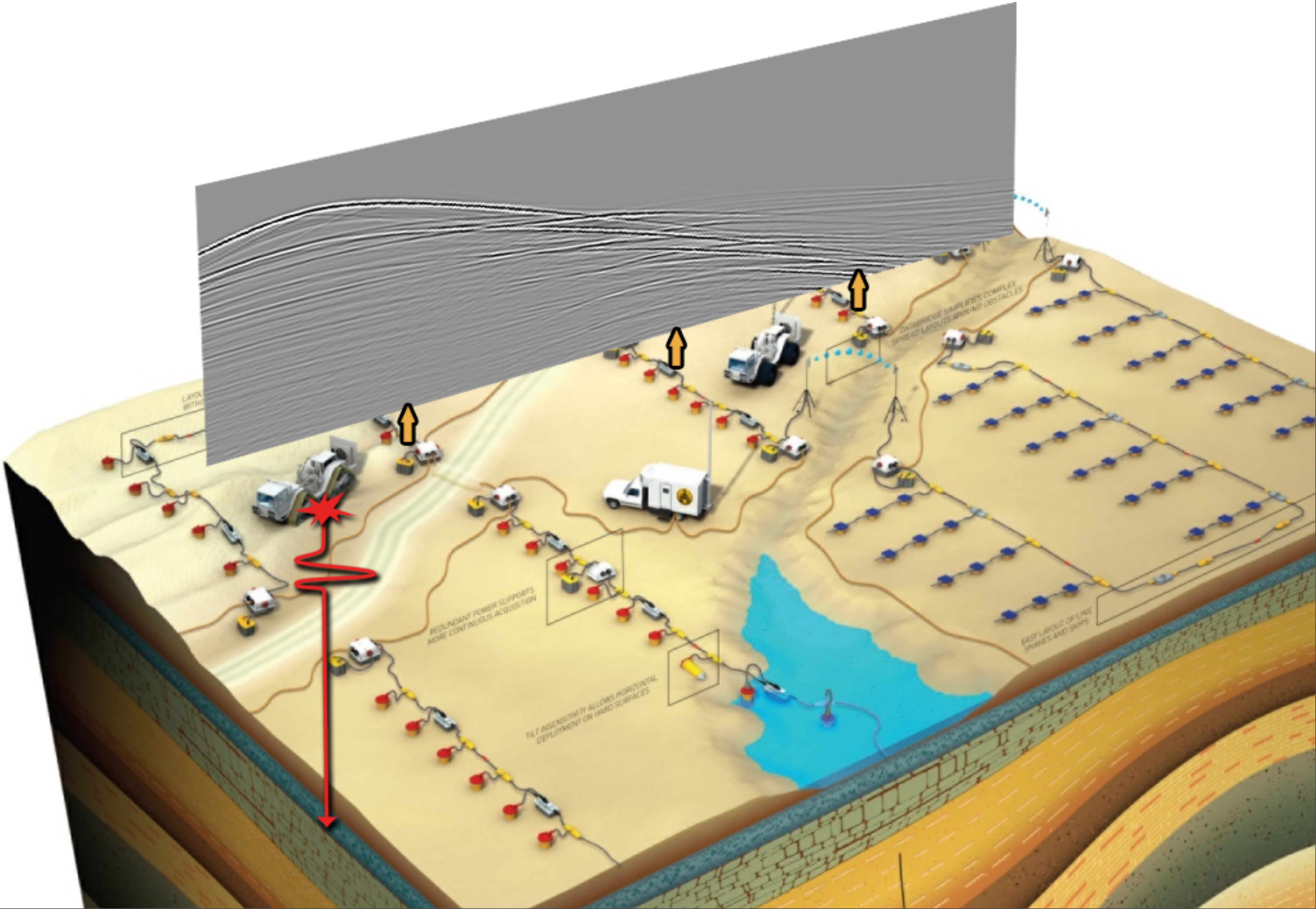


Vienna July 24, 2009

image courtesy ION (www.iongeo.com)

Seismic acquisition

Individual shots



Individual shots

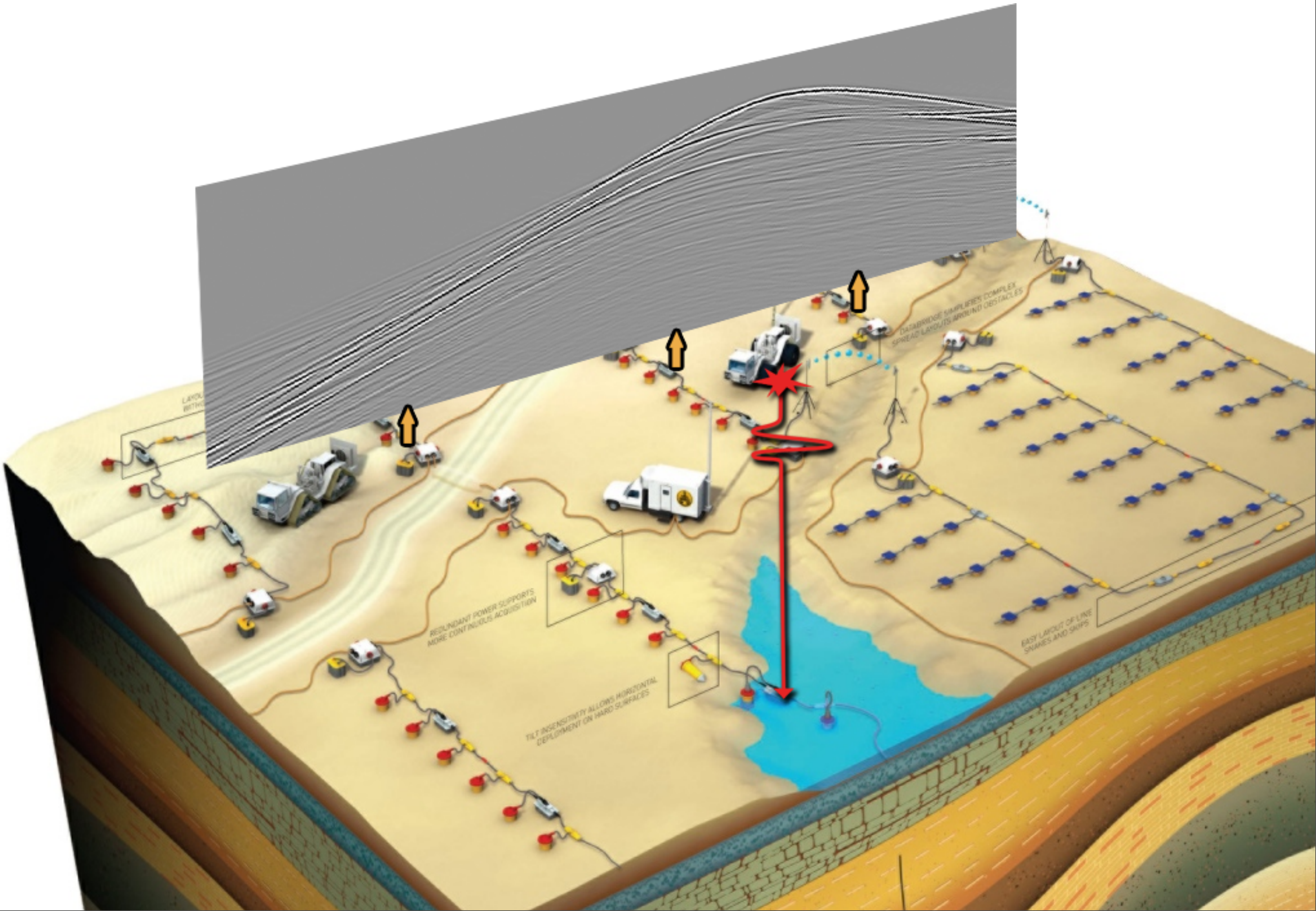
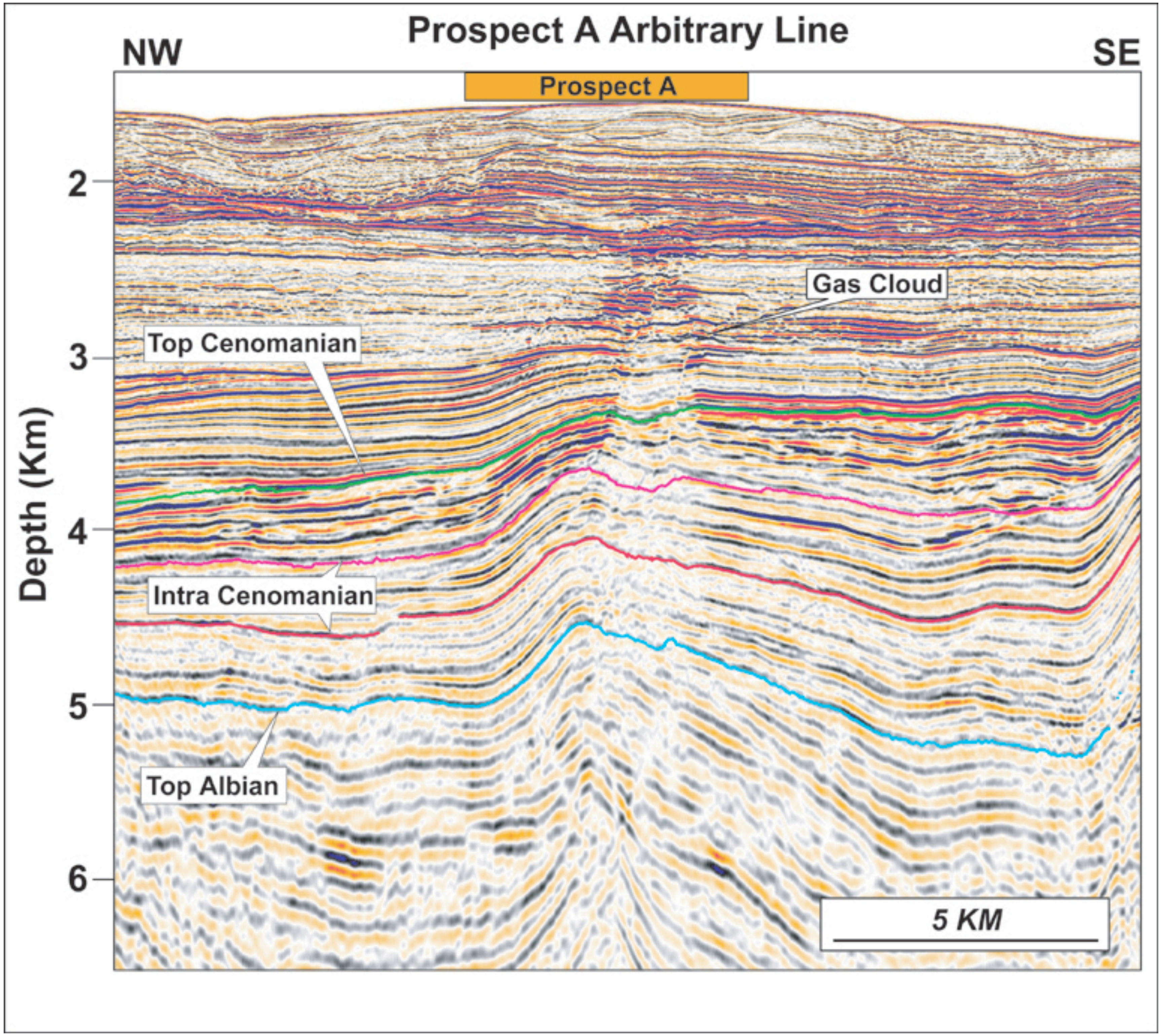


image courtesy vancoenergy (www.vancoenergy.com)

After imaging



Motivation

- **Seismic data processing, modeling & inversion:**
  - firmly rooted in Nyquist's sampling paradigm for high-dimensional wavefields
  - too *pessimistic* for signals with *structure,* i.e, there exists some sparsifying transform (e.g. Fourier, curvelets)

- **Recent theoretical & hardware developments**
  - Alternative multiscale, localized & directional transform domains that compress seismic data
  - New nonlinear sampling theory that supersedes the overly pessimistic Nyquist sampling criterion
  - New autonomous data acquisition devices that allow for more flexibility during acquisition
  - New simultaneous & continuous recording

- **Today's agenda:**
  - Extensions of ***jittered*** sampling to higher-D through ***randomized*** *blue-noise* sampling
  - Connections between ***randomized*** *simultaneous* acquisition and *compressive sampling*
  - Incorporation of additional ***physics***, e.g. include surface operator

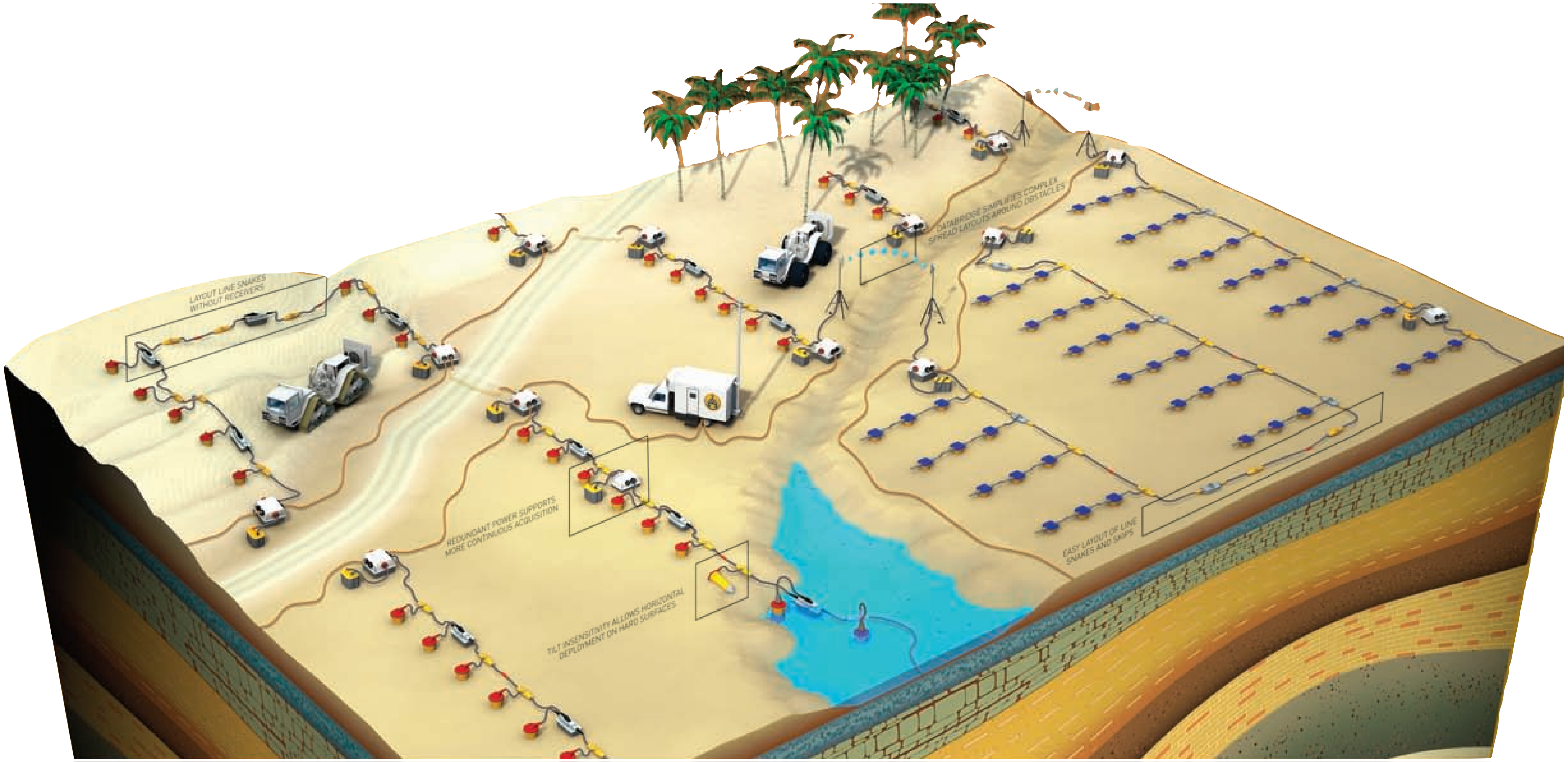
Motivation cont'd

- **Solution strategy:**
  - *leverage new* paradigm of *compressive sensing* (CS)
    - identify wavefield reconstruction from missing sources & receivers or from simultaneous acquisition as instances of CS
    - reduce acquisition, simulation, and inversion costs by ***randomization*** and deliberate ***subsampling***
  - recovery from sample **rates ≈ acquisition & computational costs** *proportional* to **transform-domain sparsity** of *data* or *model*

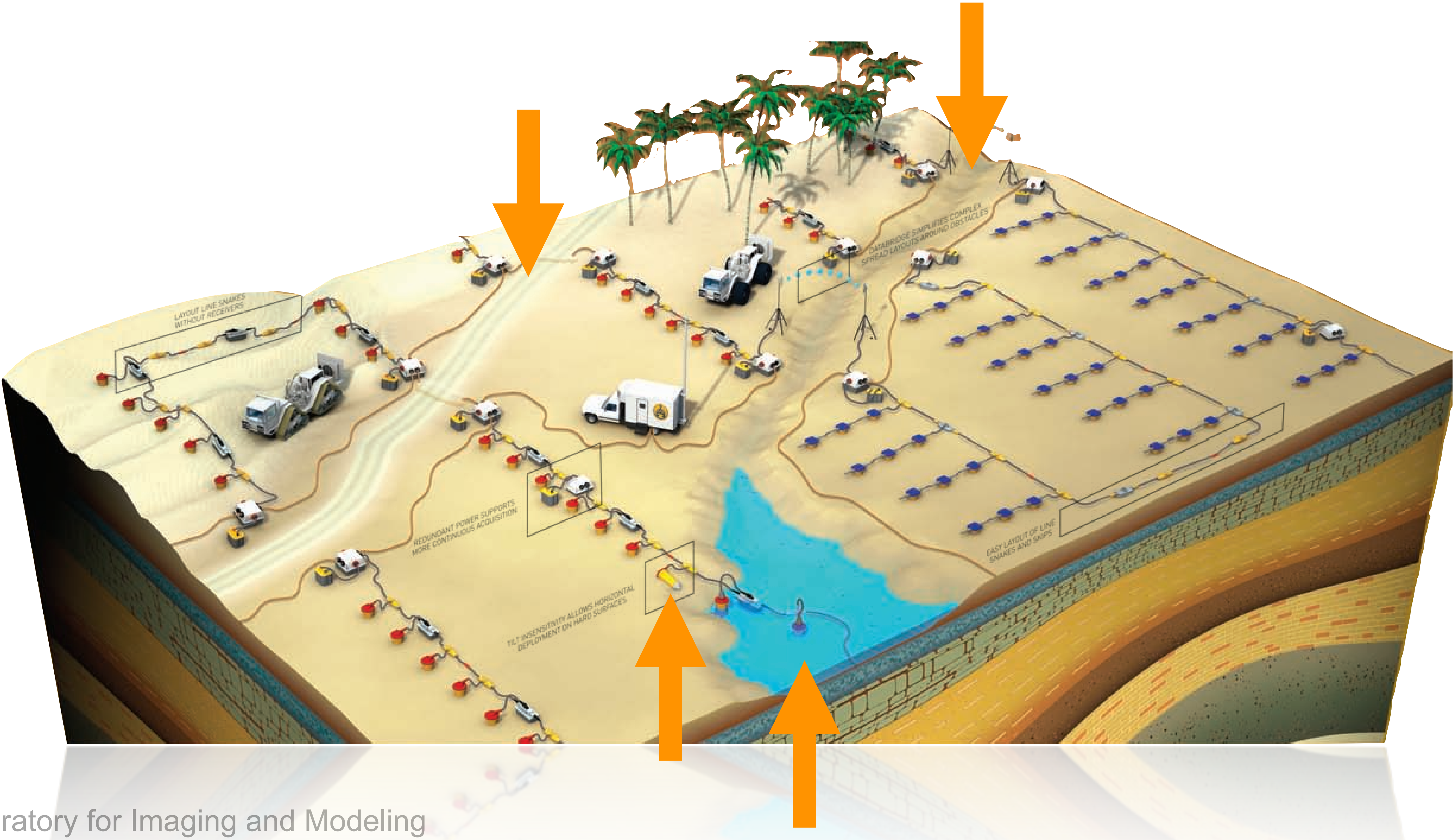
- **Remove the "curse of dimensionality" by removing constructive aliases/interferences**
  - breaking the ***periodicity of regular sampling***
  - using ***incoherent sources***

- **Turn problems into a "simple" denoising problems ....**
  - use spatial blue-noise sampling techniques from computer graphics community
  - use randomized phase encoding for simultaneous source-function design

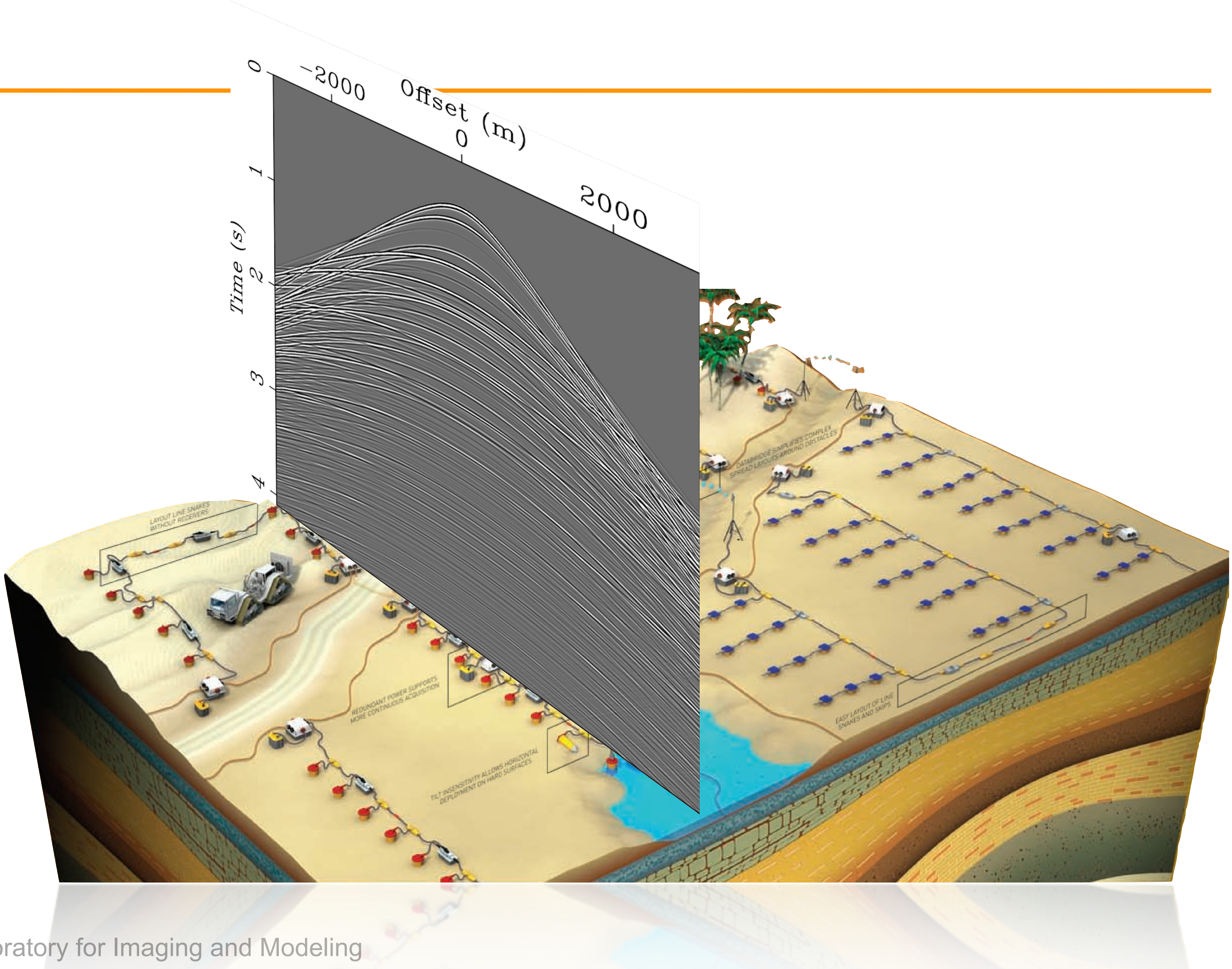
Acquisition image courtesy ION (www.iongeo.com)



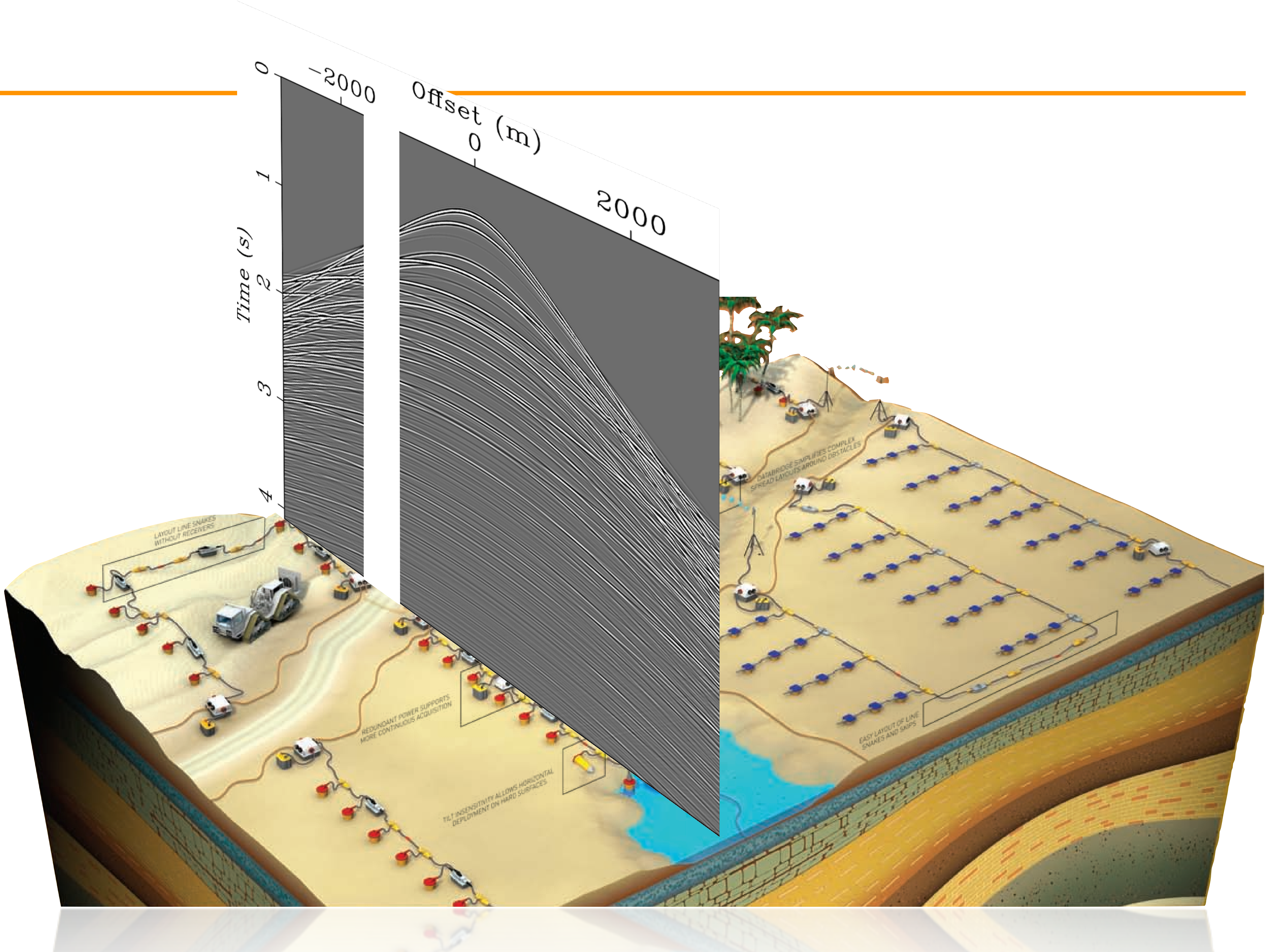
Acquisition image courtesy ION (www.iongeo.com)

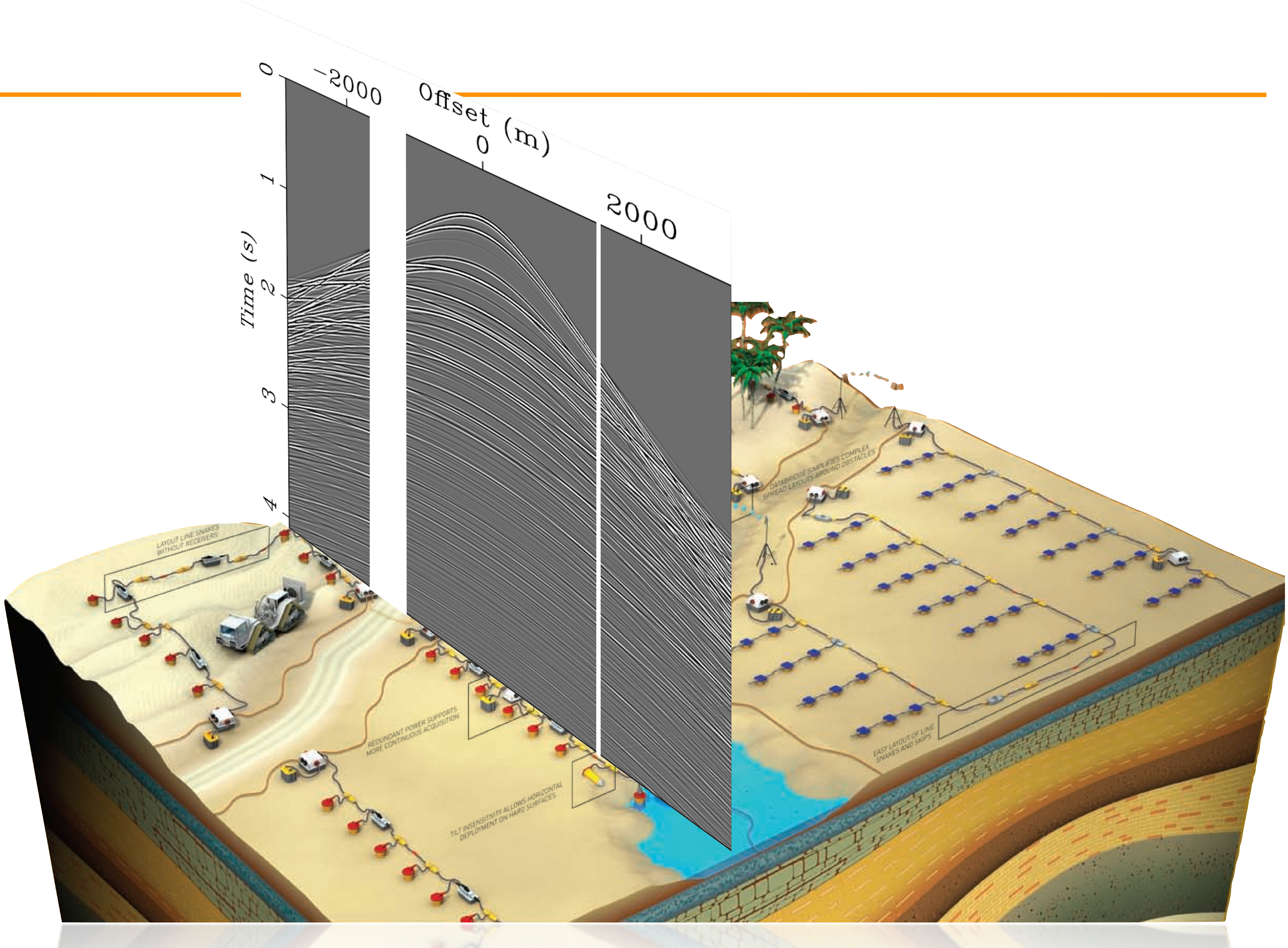


Acquisition image courtesy ION (www.iongeo.com)



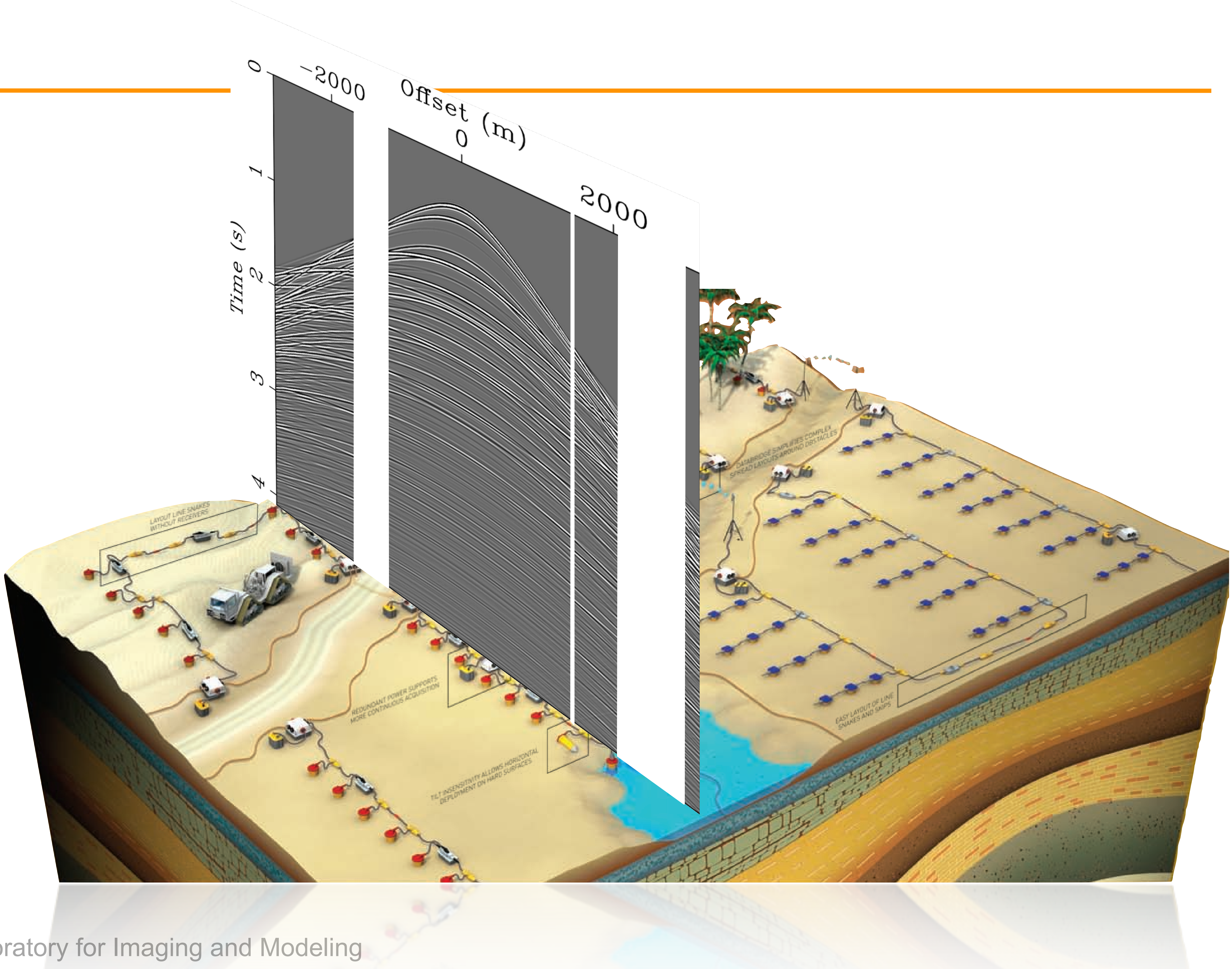
Acquisition image courtesy ION (www.iongeo.com)

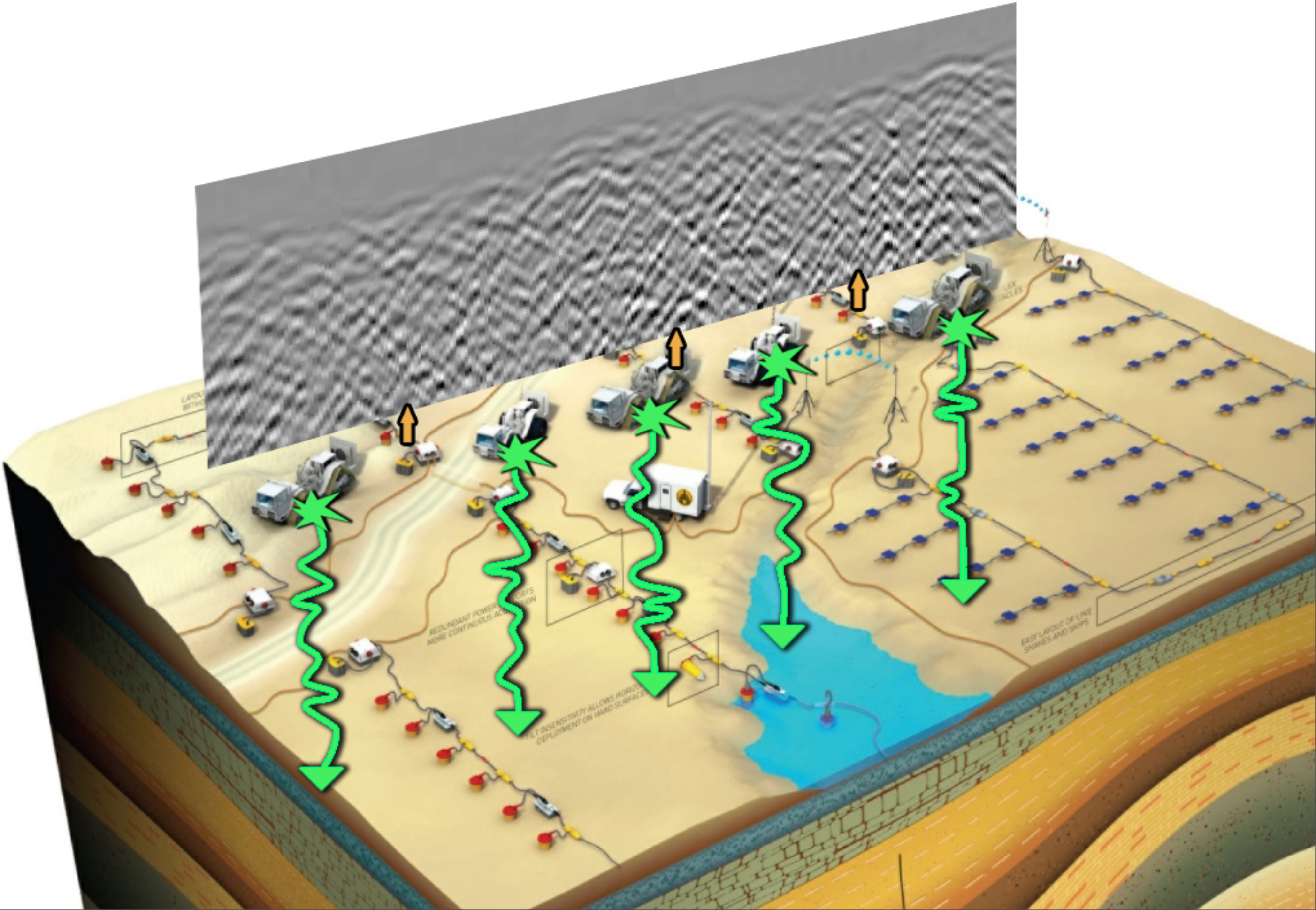




Acquisition image courtesy ION (www.iongeo.com)

Acquisition image courtesy ION (www.iongeo.com)




REDUNDANT POWER SUPPORTS
MORE CONTINUOUS ACQUISITION
TILT INSENSITIVITY ALLOWS HORIZ
DEPLOYMENT ON HARD SURFACE
EASY LAYOUT OF LINE
SNAKES AND SKIPS

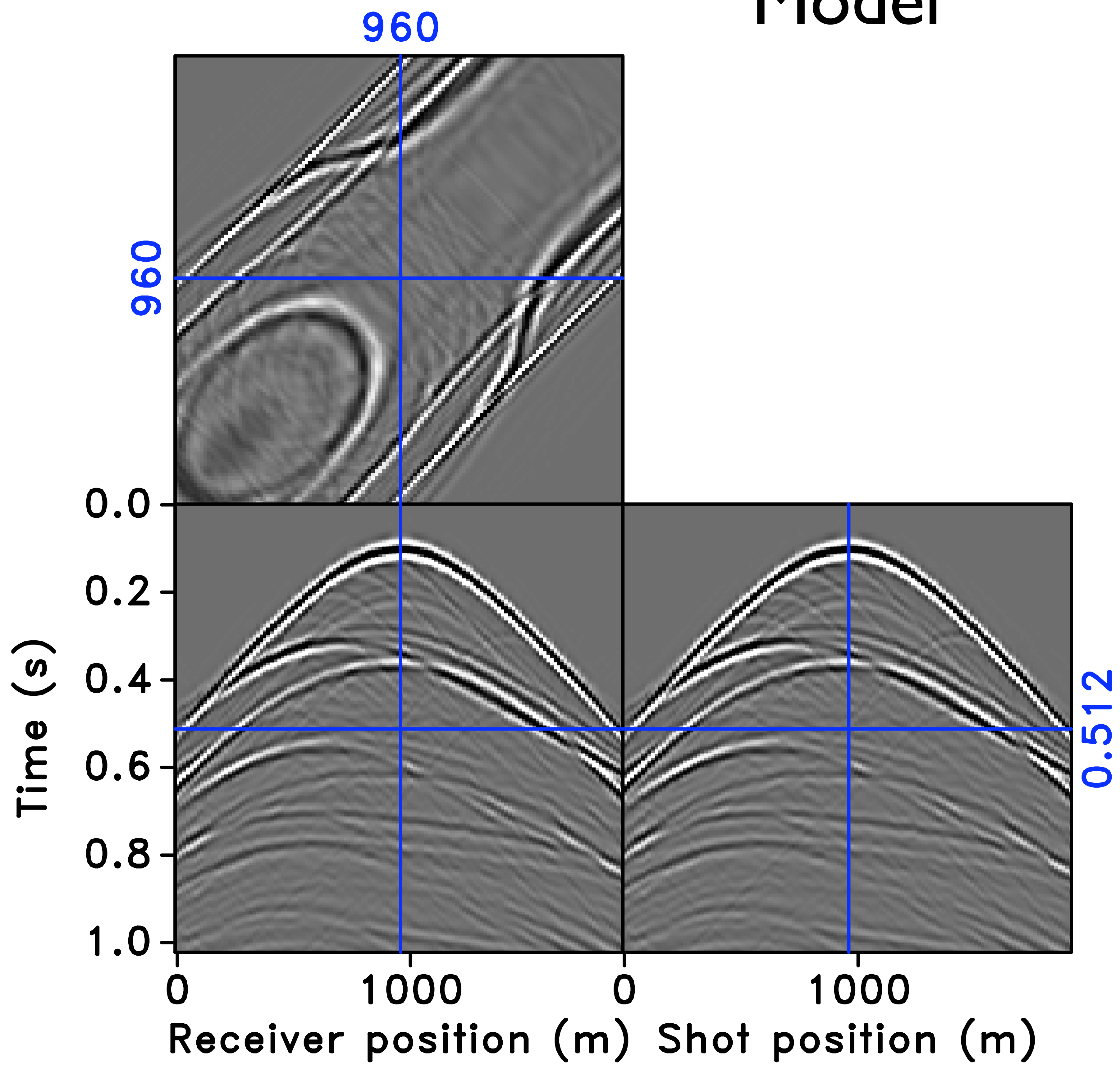
Questions

- **What is better to *periodically* sample *sequential* impulsive sources or to sample at *randomized* positions?**

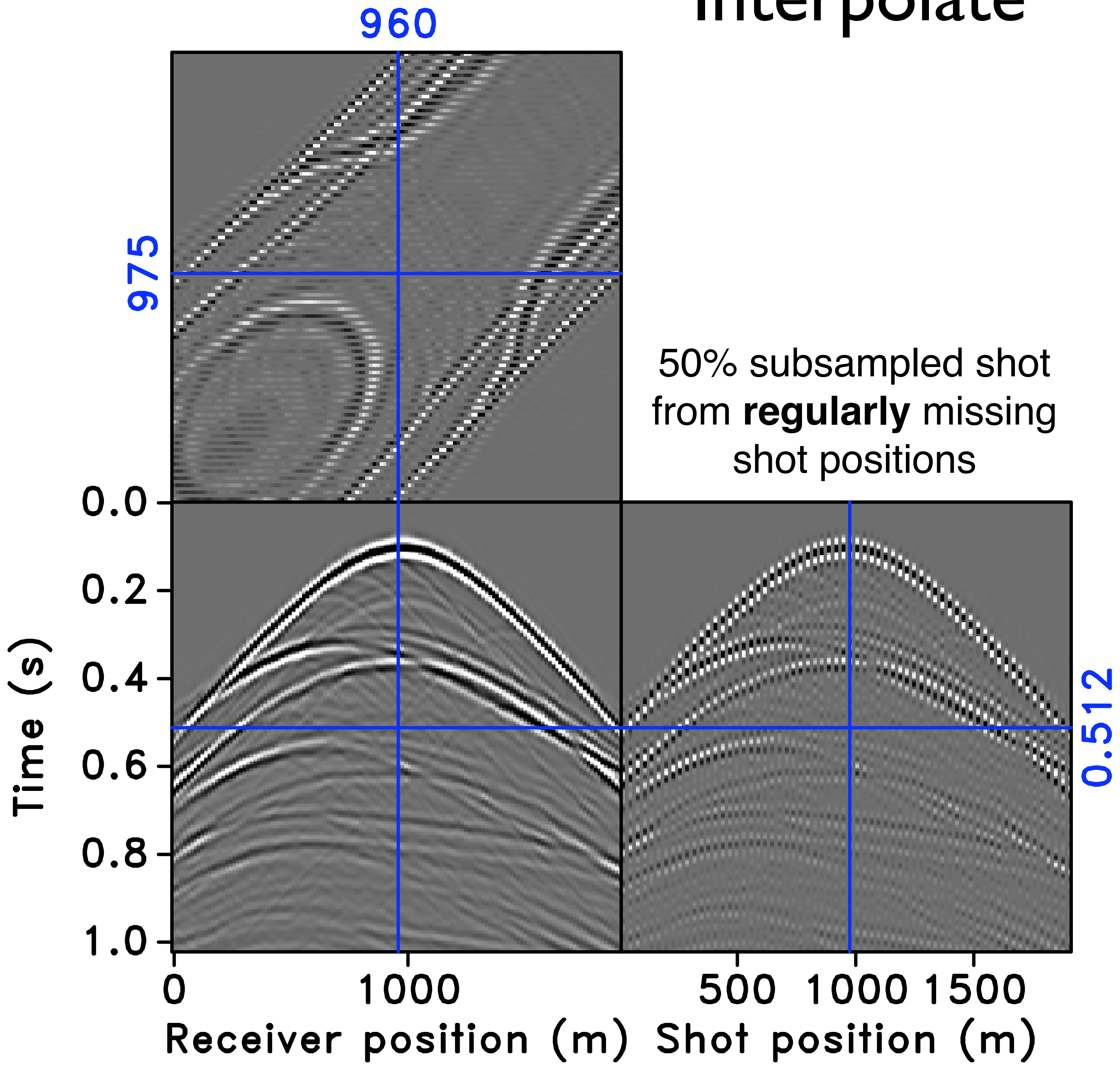
- **What is better? Having missing *single-source* or *missing randomized incoherent simultaneous* experiments?**

- Comparison between different undersampling strategies for source experiments:
  - ***Randomized*** jittered shot positions
  - ***Randomized*** simultaneous shots

Model

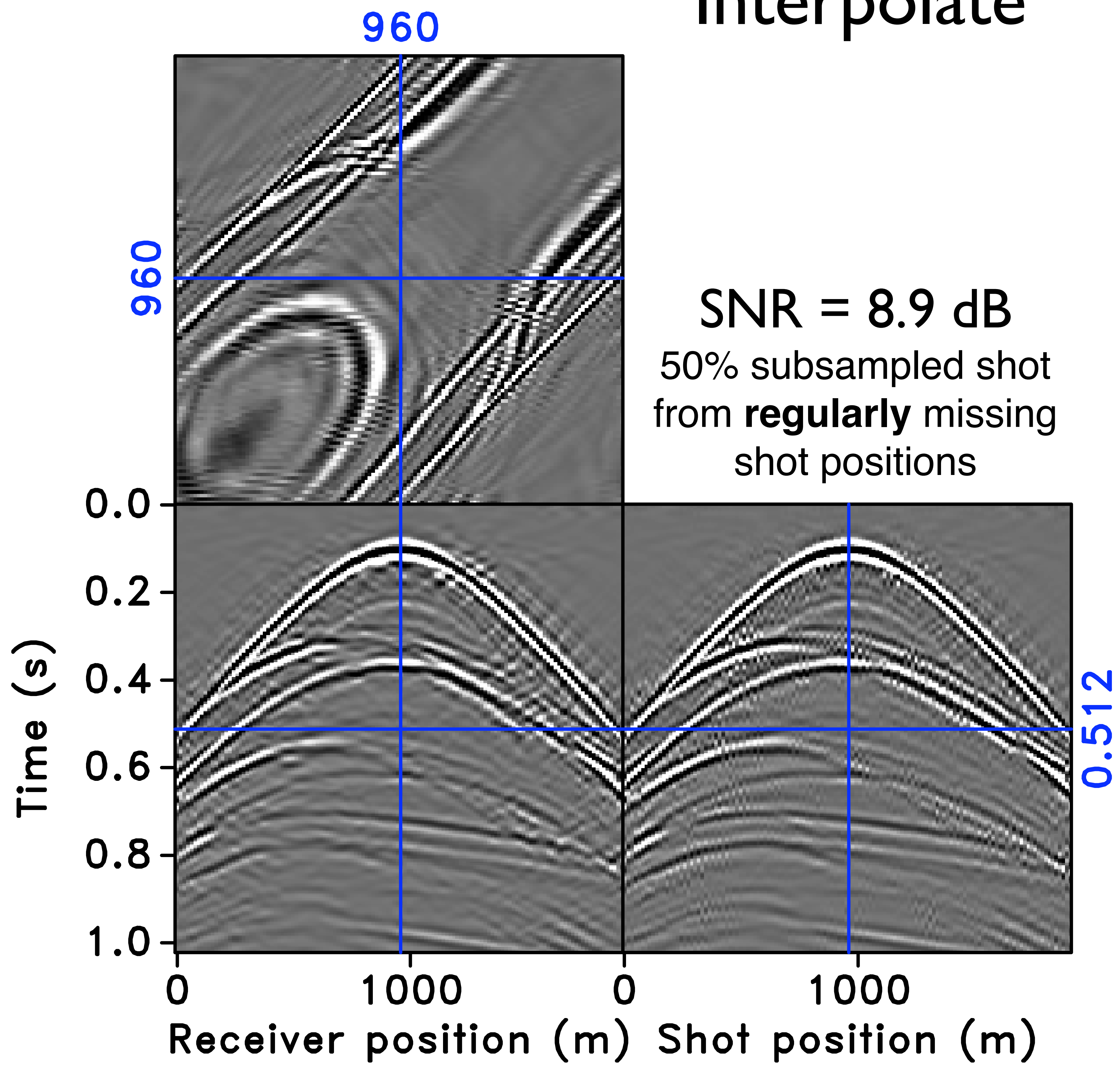

960
960
0.512
0.0
0.2
0.4
0.6
0.8
1.0
Time (s)
0
1000
0
1000
Receiver position (m)
Shot position (m)

Interpolate



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

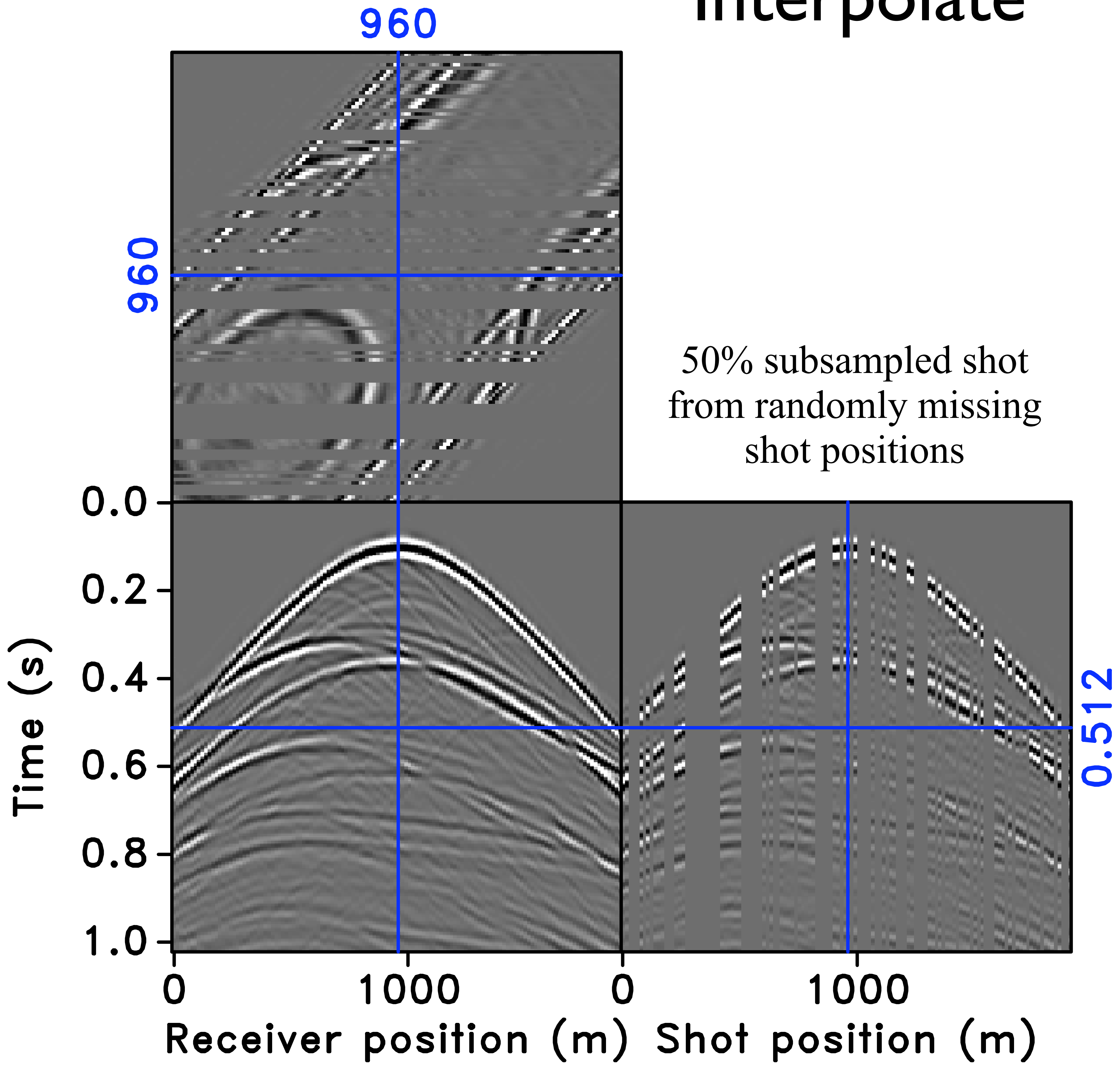
Interpolate



SNR = 8.9 dB

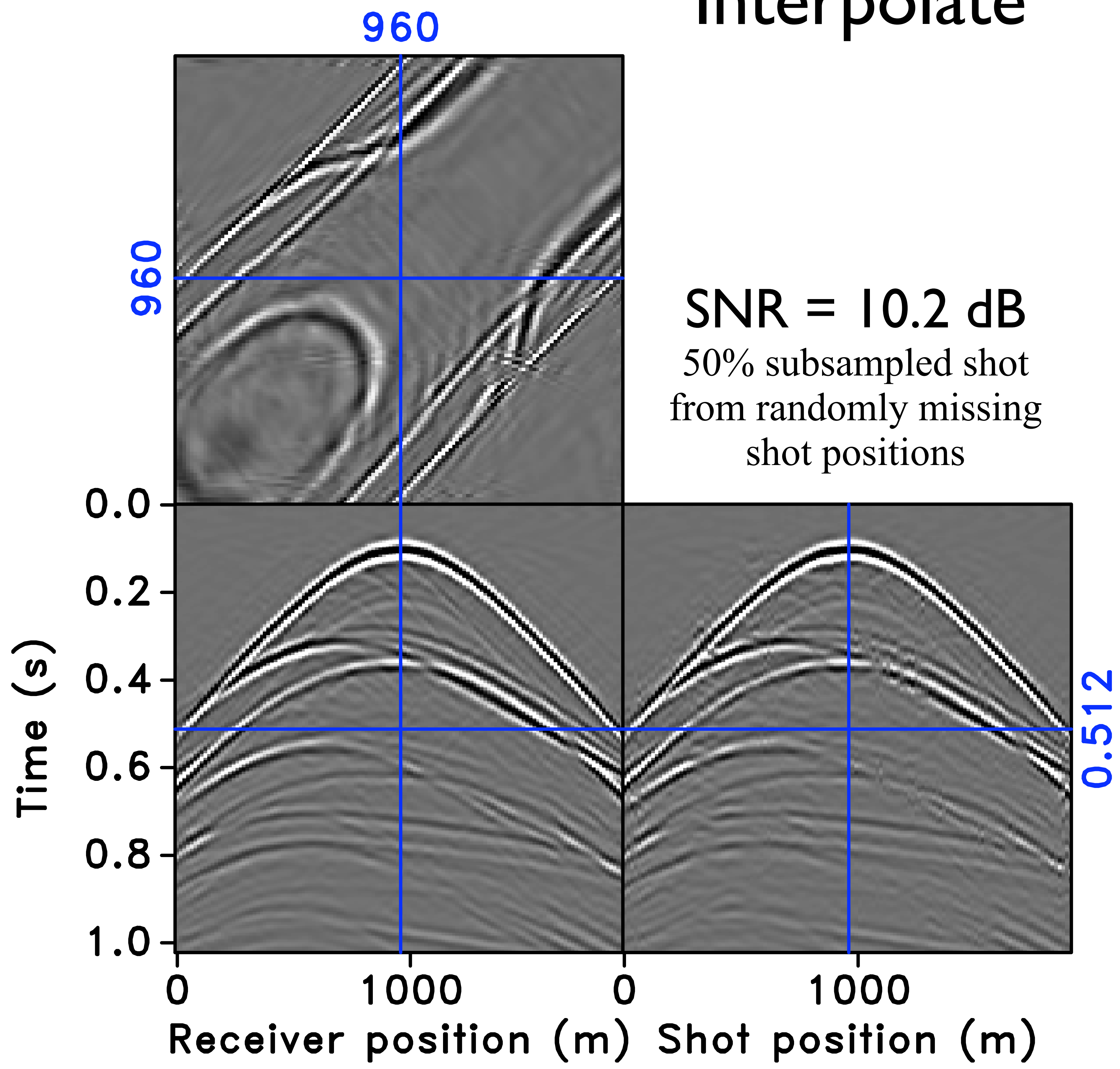
50% subsampled shot from **regularly** missing shot positions

Interpolate



50% subsampled shot
from randomly missing
shot positions

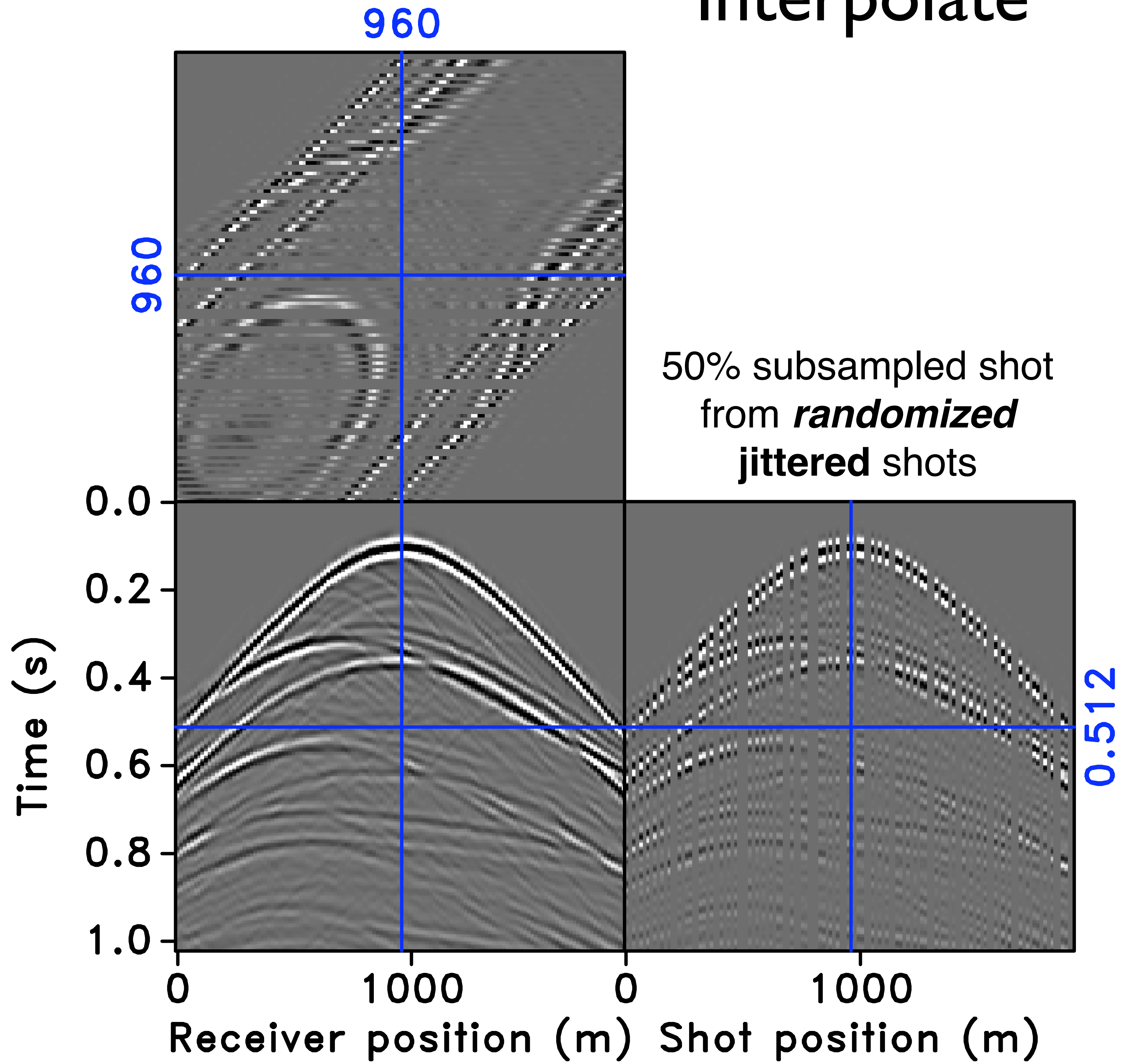
Interpolate



SNR = 10.2 dB

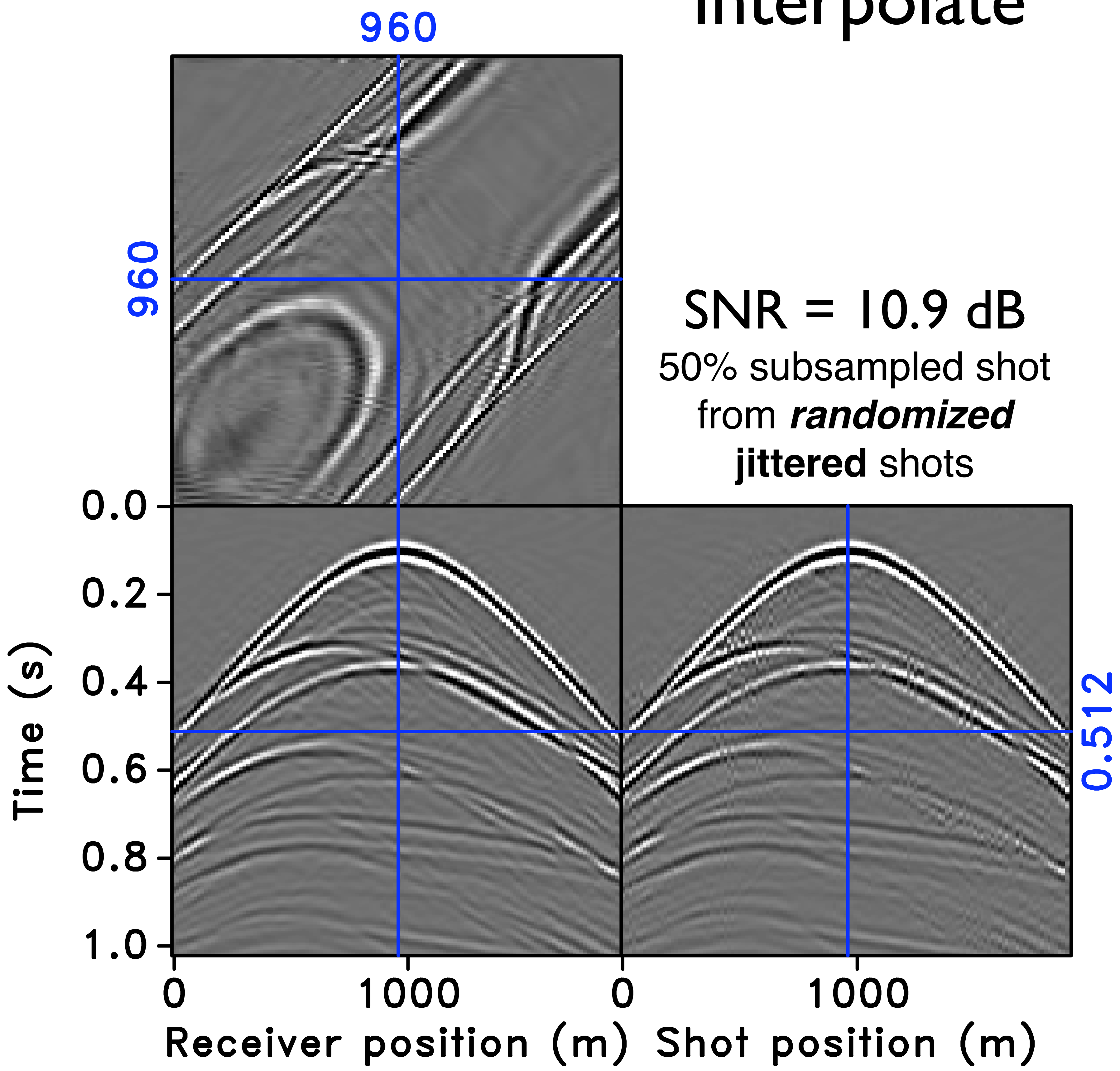
50% subsampled shot
from randomly missing
shot positions

Interpolate



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

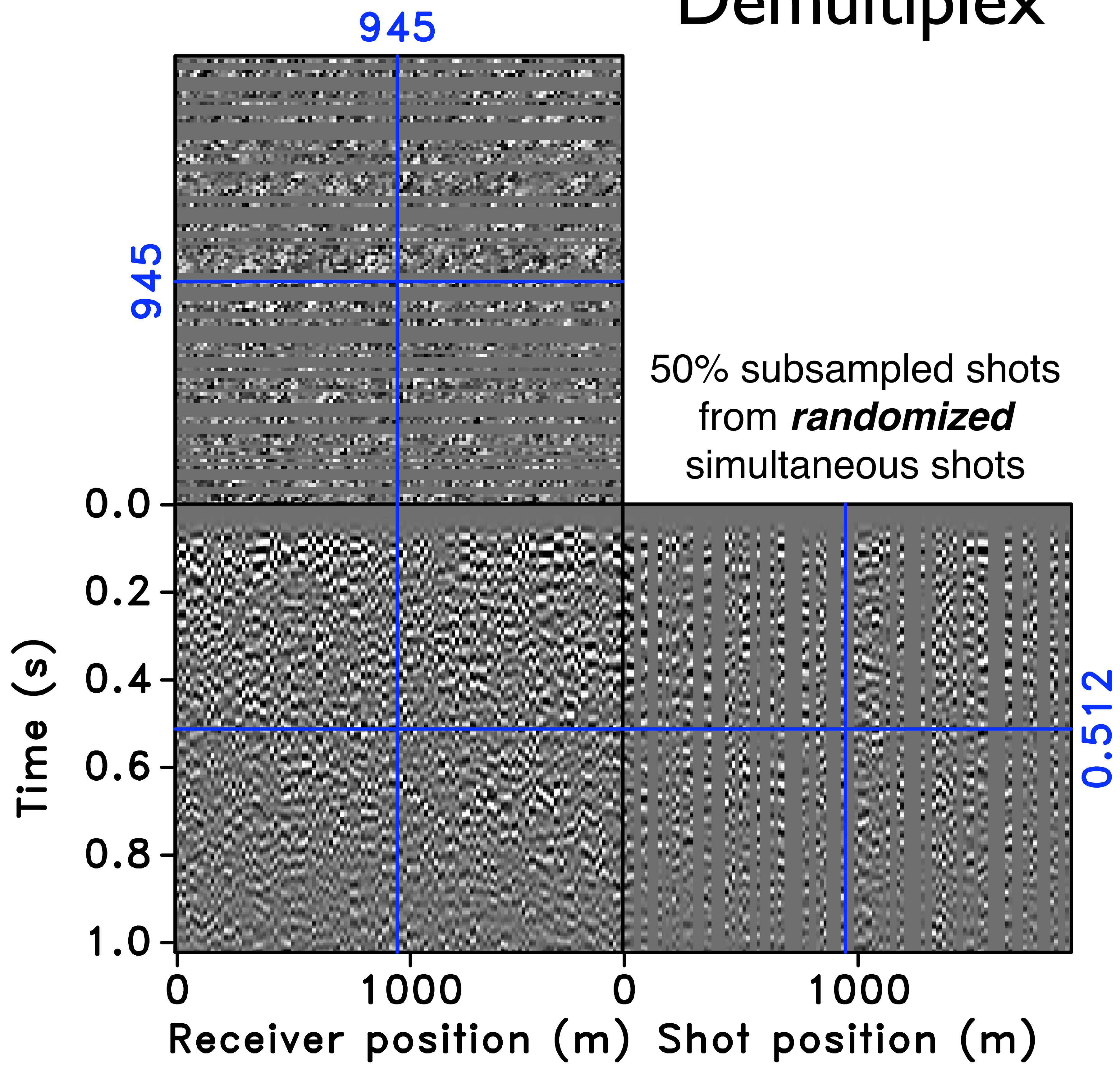
Interpolate



SNR = 10.9 dB

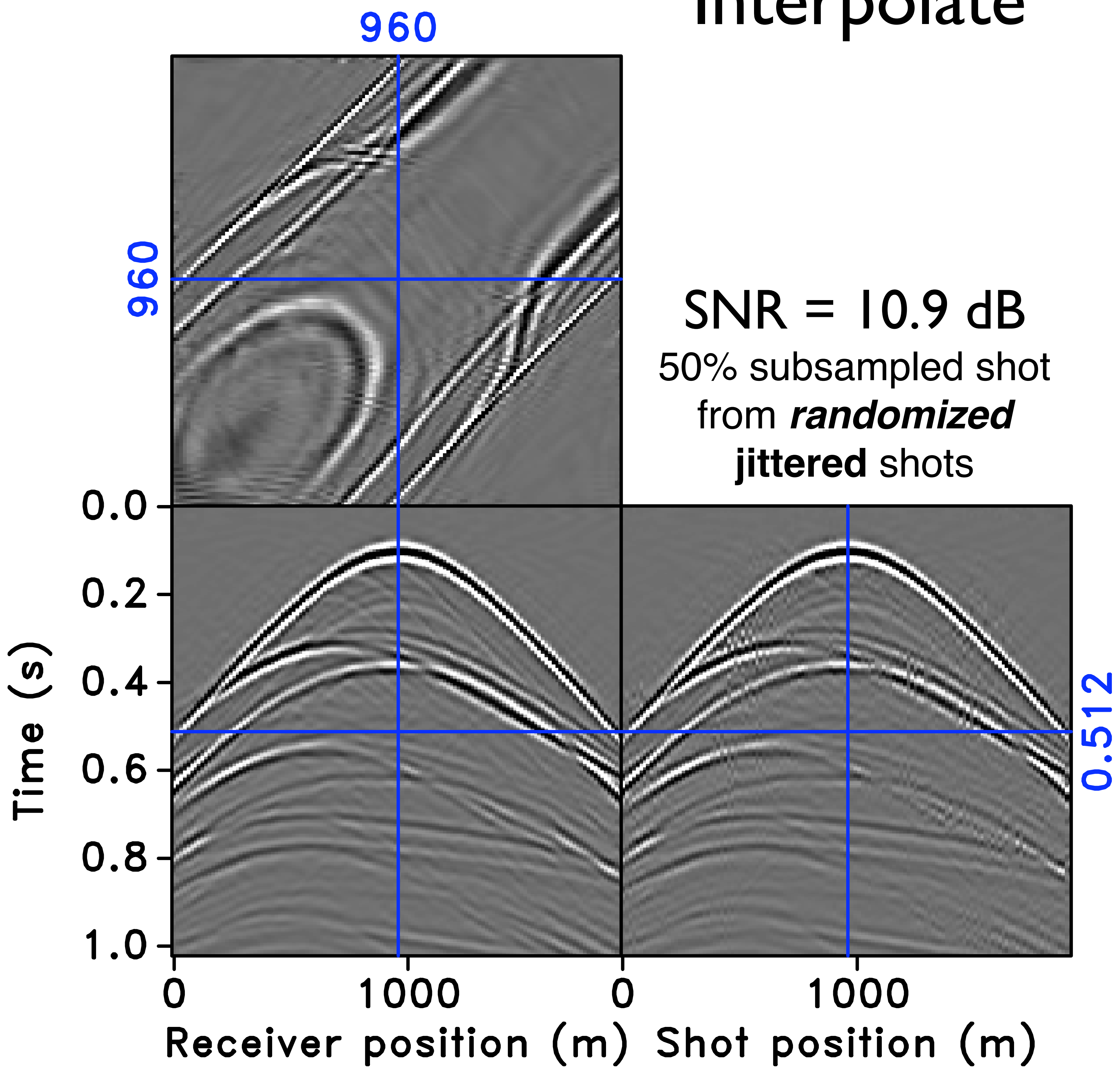
50% subsampled shot
from ***randomized***
jittered shots

Demultiplex



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

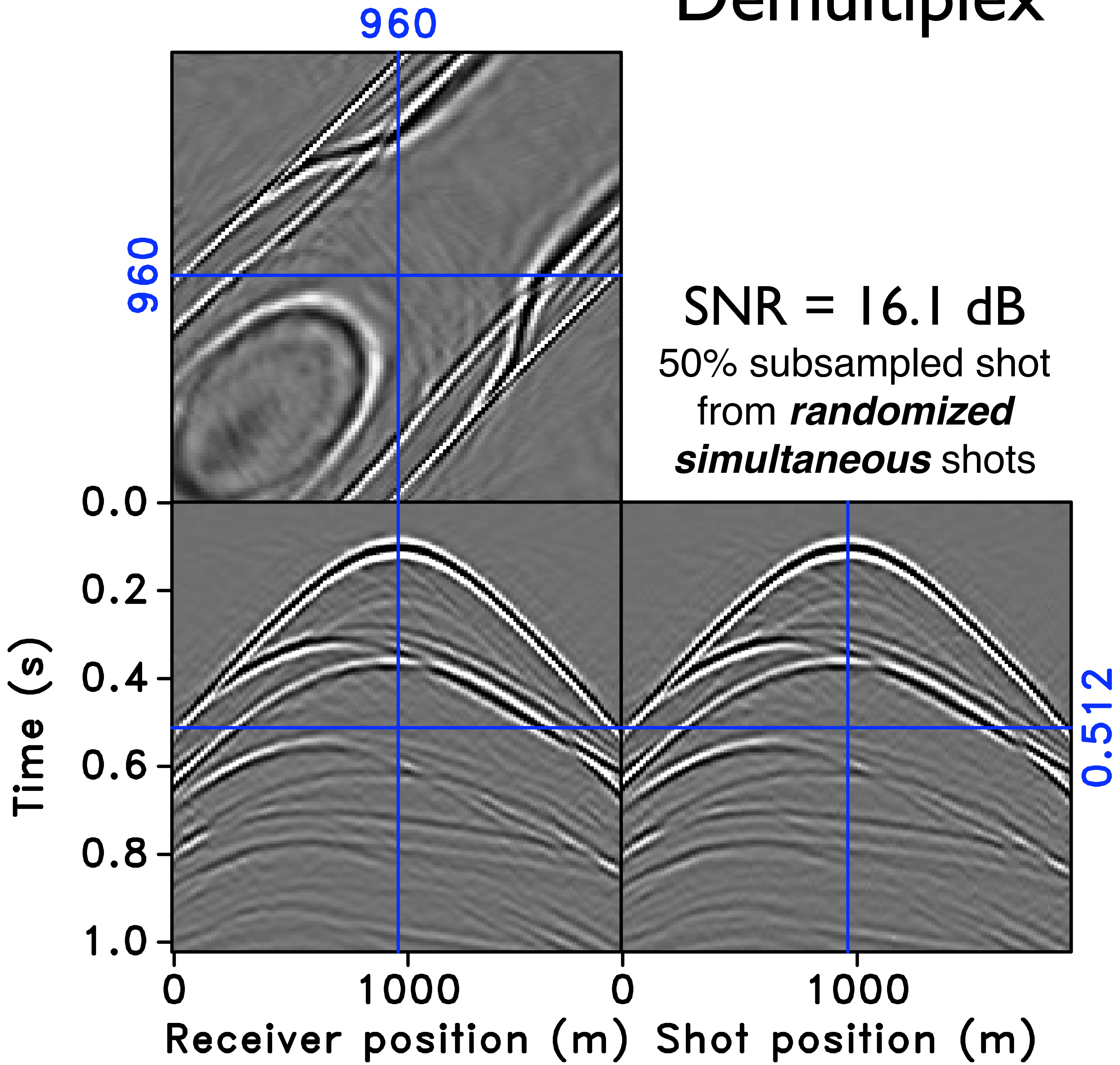
Interpolate



SNR = 10.9 dB

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

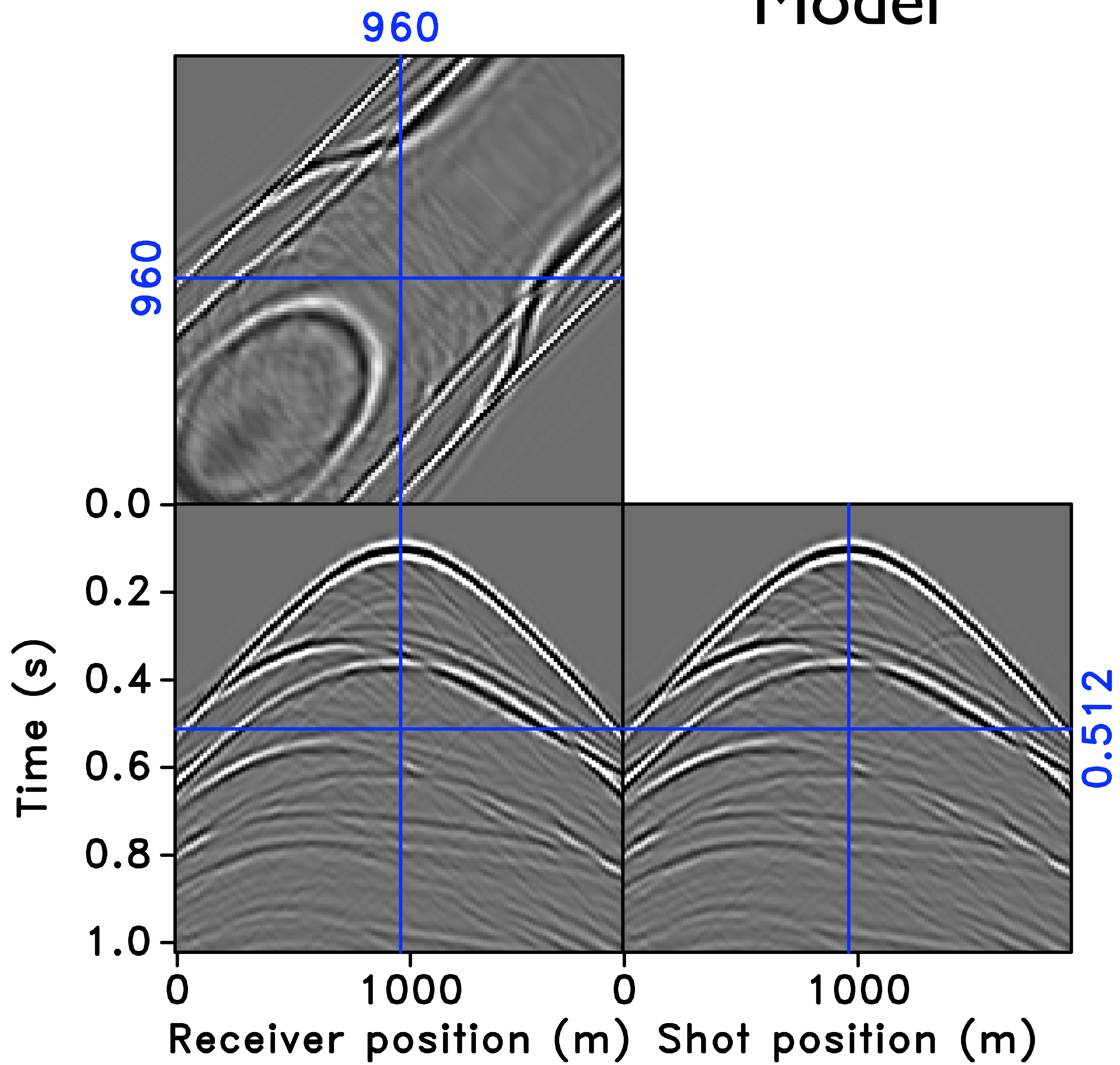
Demultiplex



SNR = 16.1 dB

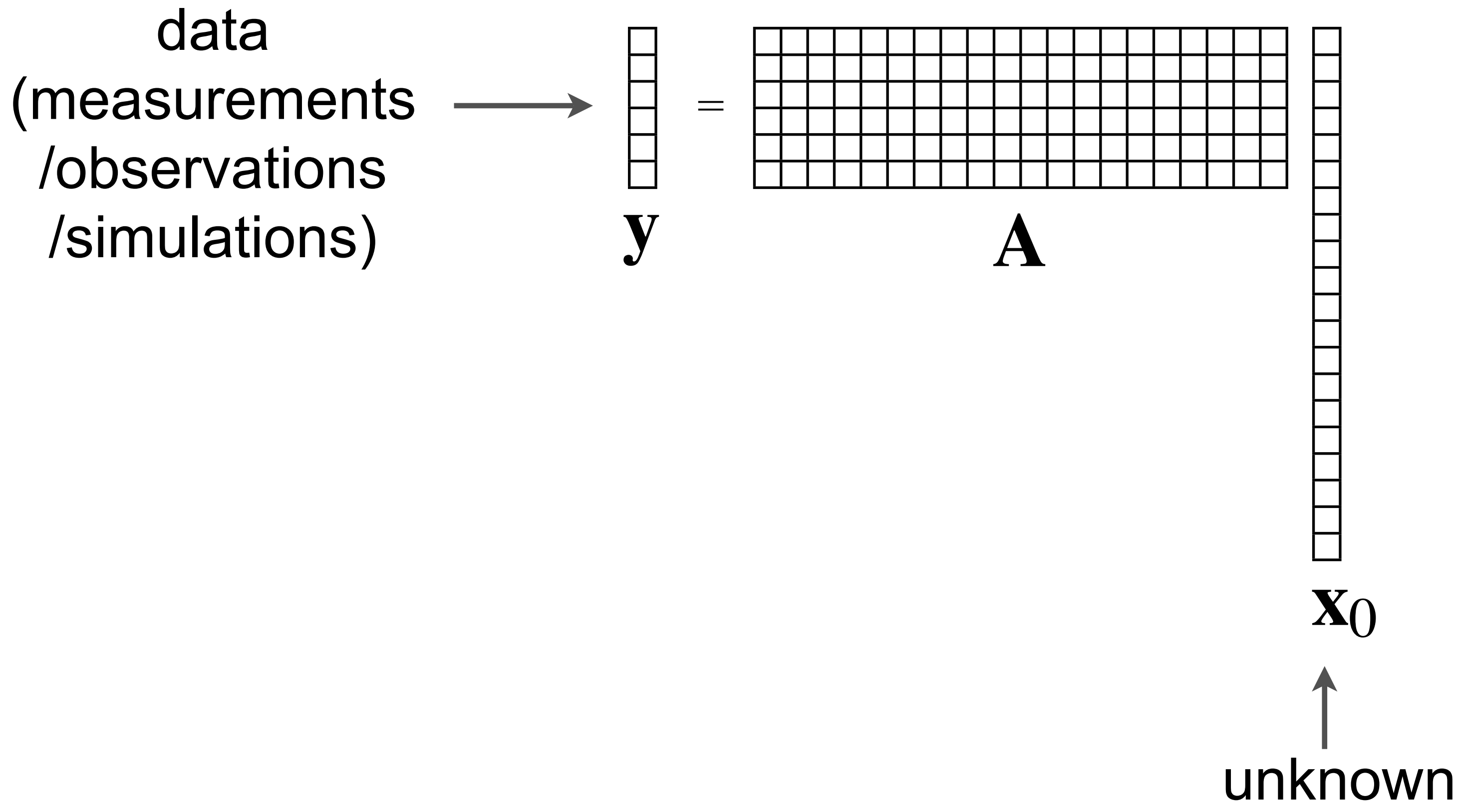
50% subsampled shot
from ***randomized***
simultaneous shots

Model


960
960
0.512
0.0
0.2
0.4
0.6
0.8
1.0
Time (s)
0
1000
0
1000
Receiver position (m)
Shot position (m)

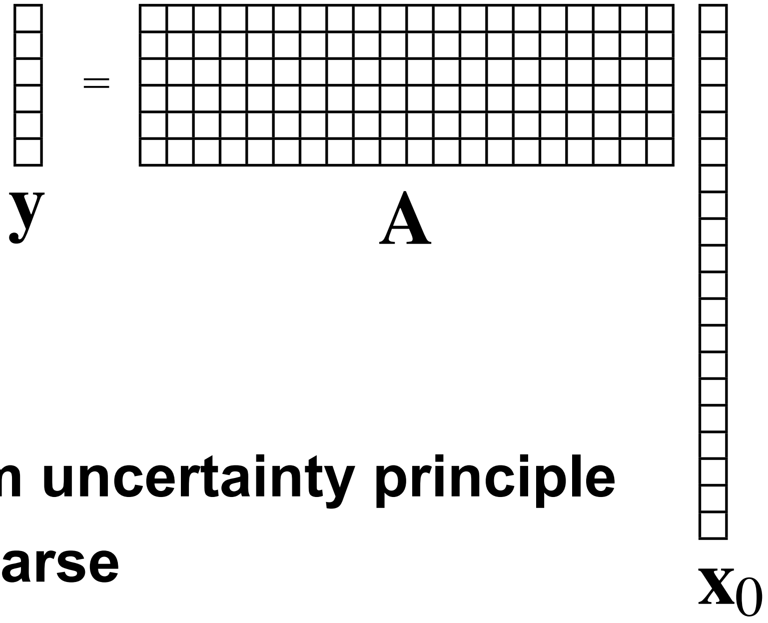
Problem statement

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 the **uniform uncertainty principle**
  - **x**$_0$ is **sufficiently sparse**

- recovery procedure

$$\underbrace{\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} = \sum_n |x_i|}_{\text{"sparsity"}} \quad \text{subject to} \quad \underbrace{\mathbf{Ax} = \mathbf{y}}_{\text{perfect reconstruction}}$$

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

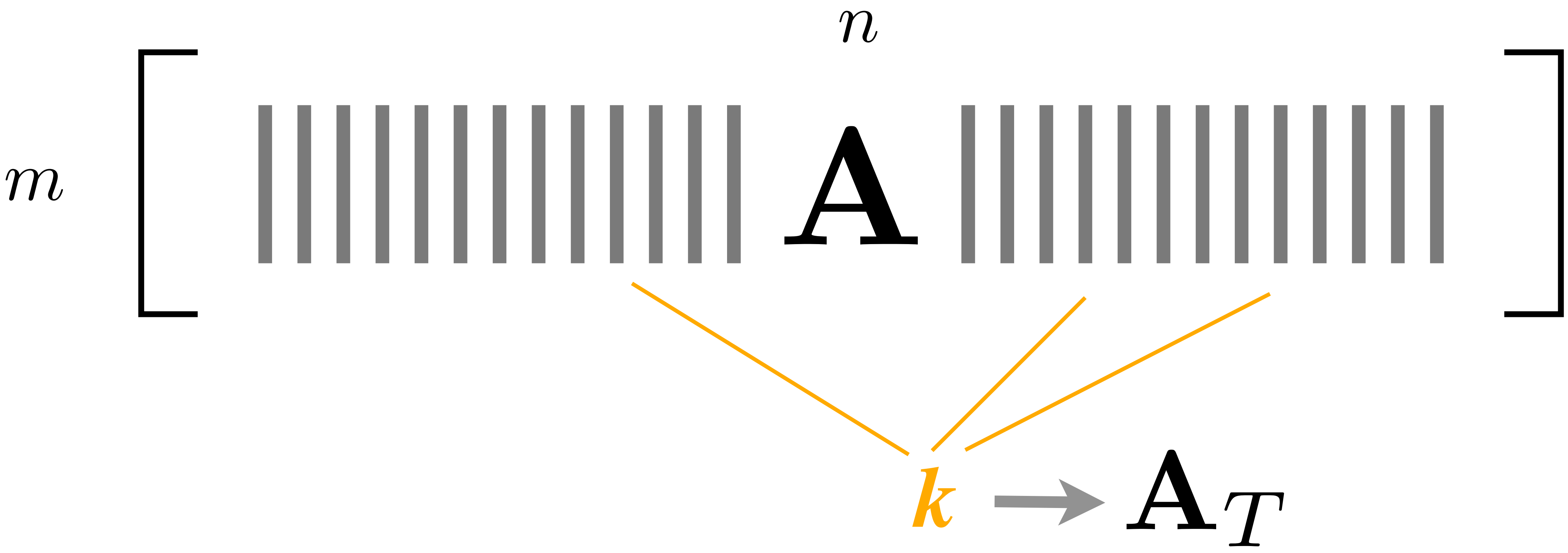
[Candès et al.'06]
[Donoho'06]

Designing CS acquisition matrix

- Restricted Isometry Property holds $m \geq C \cdot k \log(n/k)$

$$(1-\delta_k)\|\mathbf{x}_T\|_{\ell 2} \leq \|\mathbf{A}_T\mathbf{x}\|_{\ell 2} \leq (1+\delta_k)\|\mathbf{x}_T\|_{\ell 2}$$

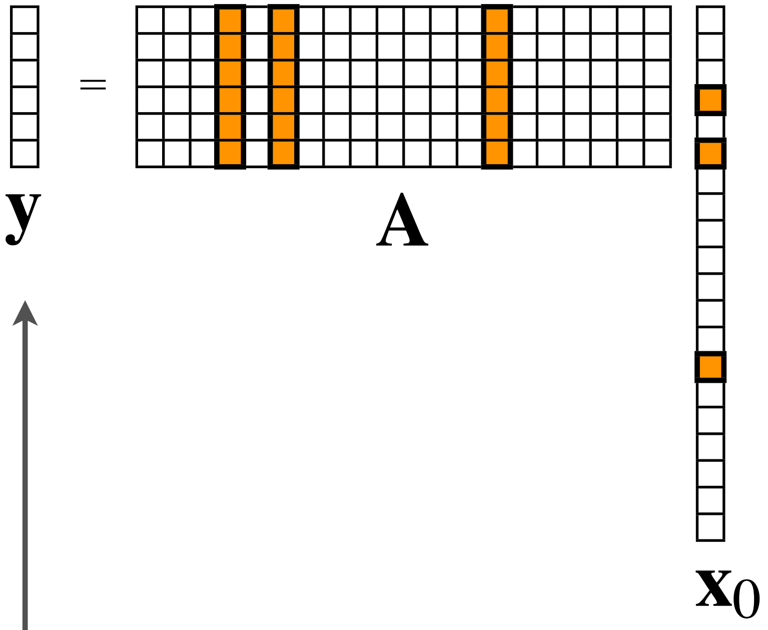
- Bounds singular values of **A**



- Subsets T of columns of **A** behave approximately as an orthonormal basis => stable
- Construction of **A** depends on ***randomization*** => spread of energy
- **A** like a matrix with iid zero-centered random Gaussian entries

Simple example

restriction operator

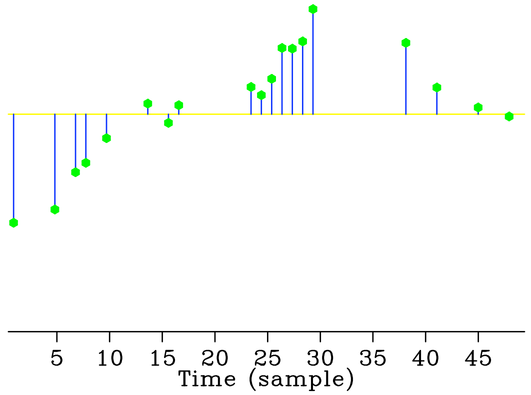


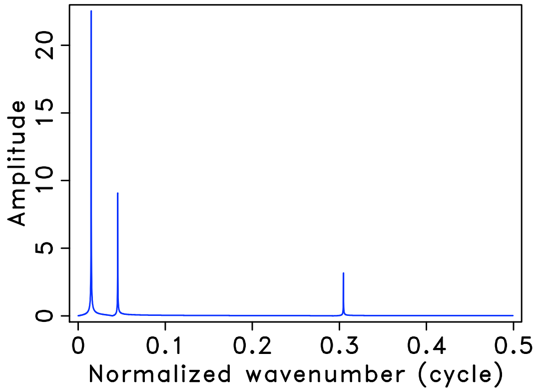
with $\mathbf{A} := \mathbf{R}\mathbf{F}^H$

Fourier transform

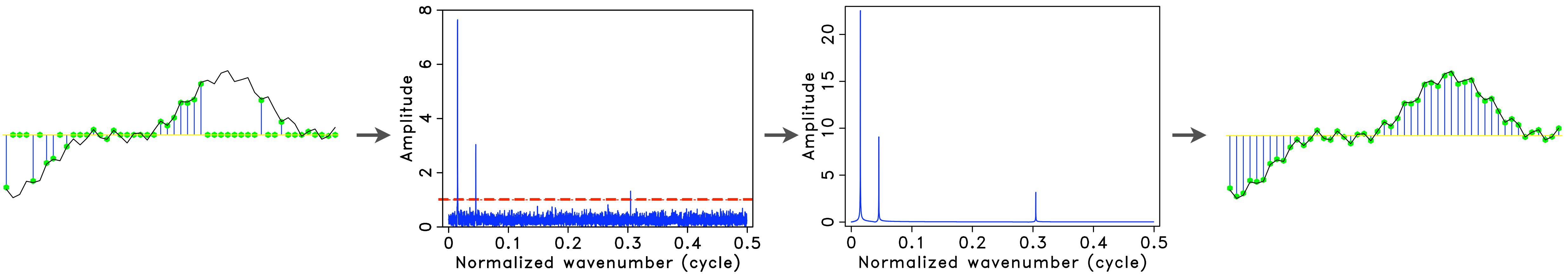
signal

Fourier coefficients (sparse)





NAIVE sparsity-promoting recovery



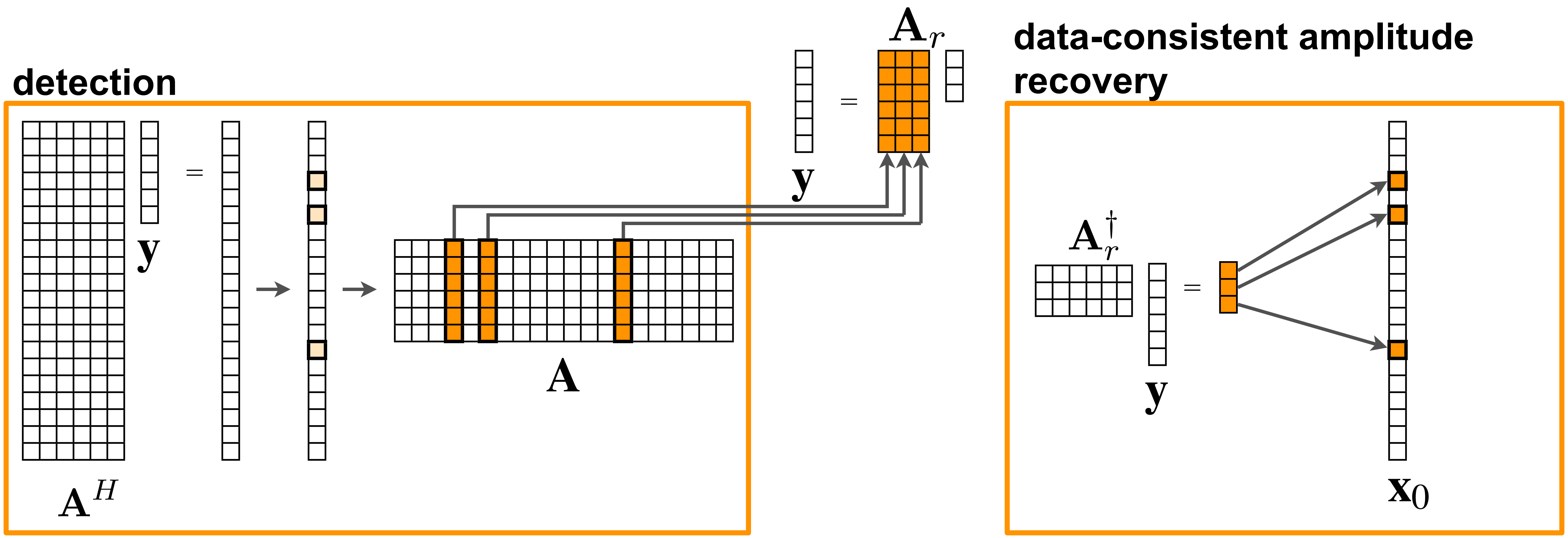
inverse Fourier transform

detection + data-consistent amplitude recovery

Fourier transform

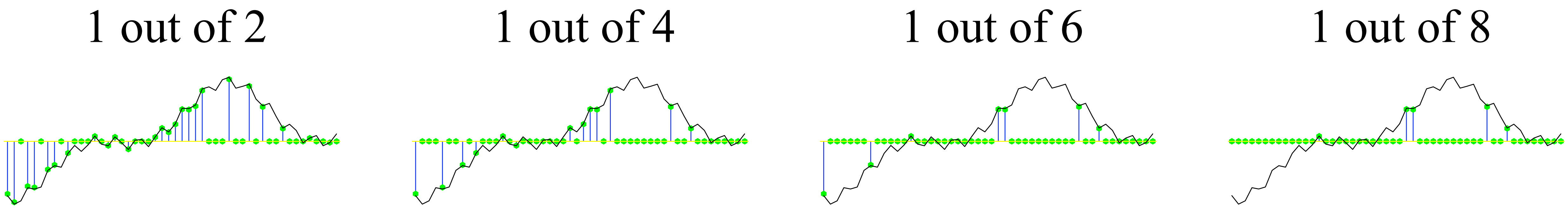
detection

data-consistent amplitude recovery



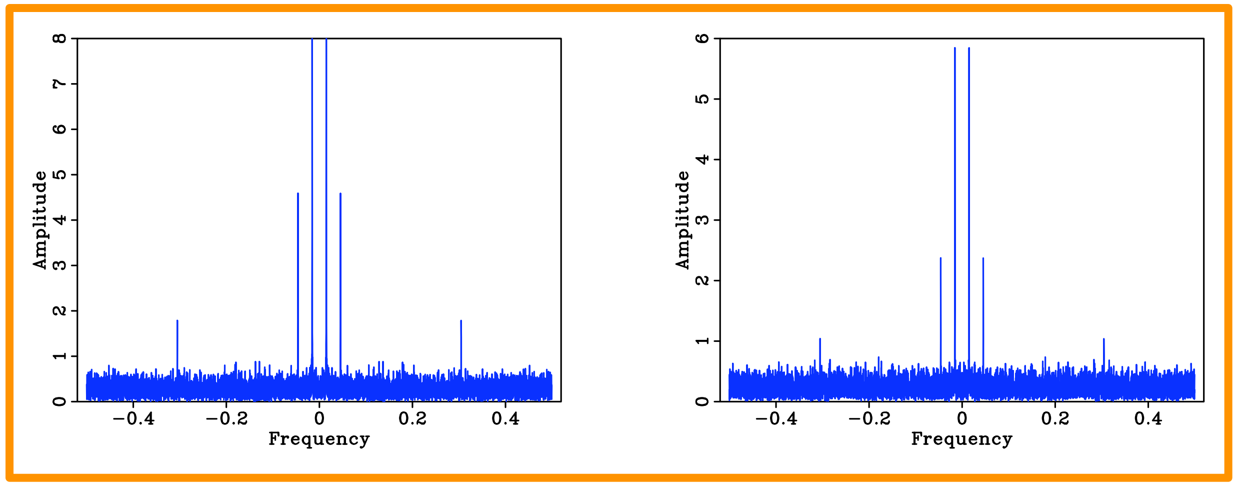
Undersampling “noise”

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

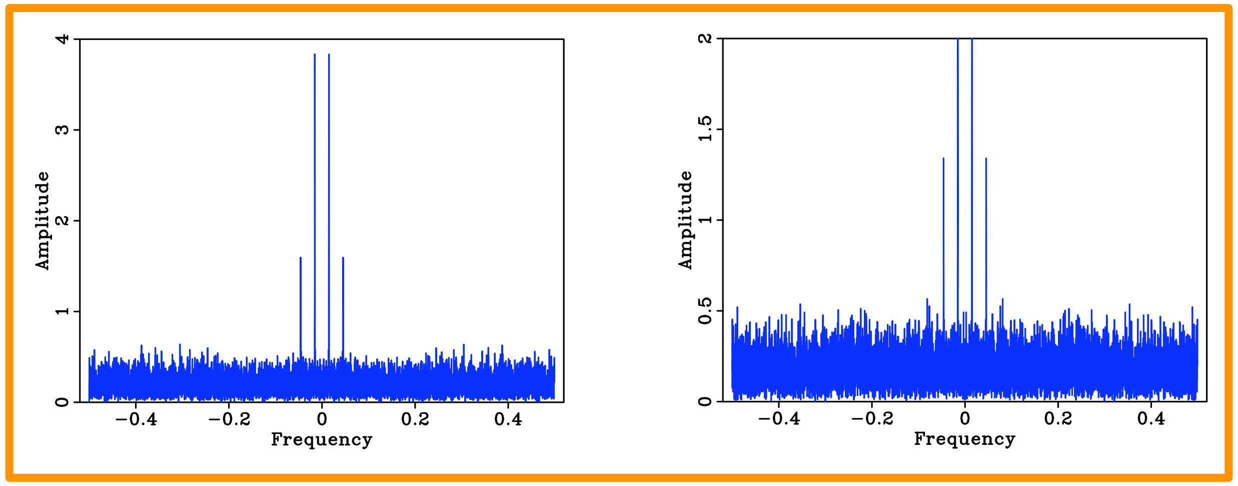


less acquired data

3 detectable Fourier modes



2 detectable Fourier modes



Extensions

- Use CS principles to select *physically* appropriate
  - ***randomized*** *restriction* matrix **R = downsampler**
  - *measurement* matrix either **I**, or ***random*** **phase encoder**, or ***randomized physics***
  - sparsifying transform **S** (e.g. curvelets)
  - driven by signal type, physics, and type of acquisition (e.g. fMRI vs seismic)
- Sparse signal representation:

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0$$

with

$$\mathbf{A} = \underset{\textbf{restriction matrix}}{\mathbf{R}}\,\underset{\textbf{measurement matrix}}{\mathbf{M}}\,\underset{\textbf{sparsity matrix}}{\mathbf{S}}^{H}$$

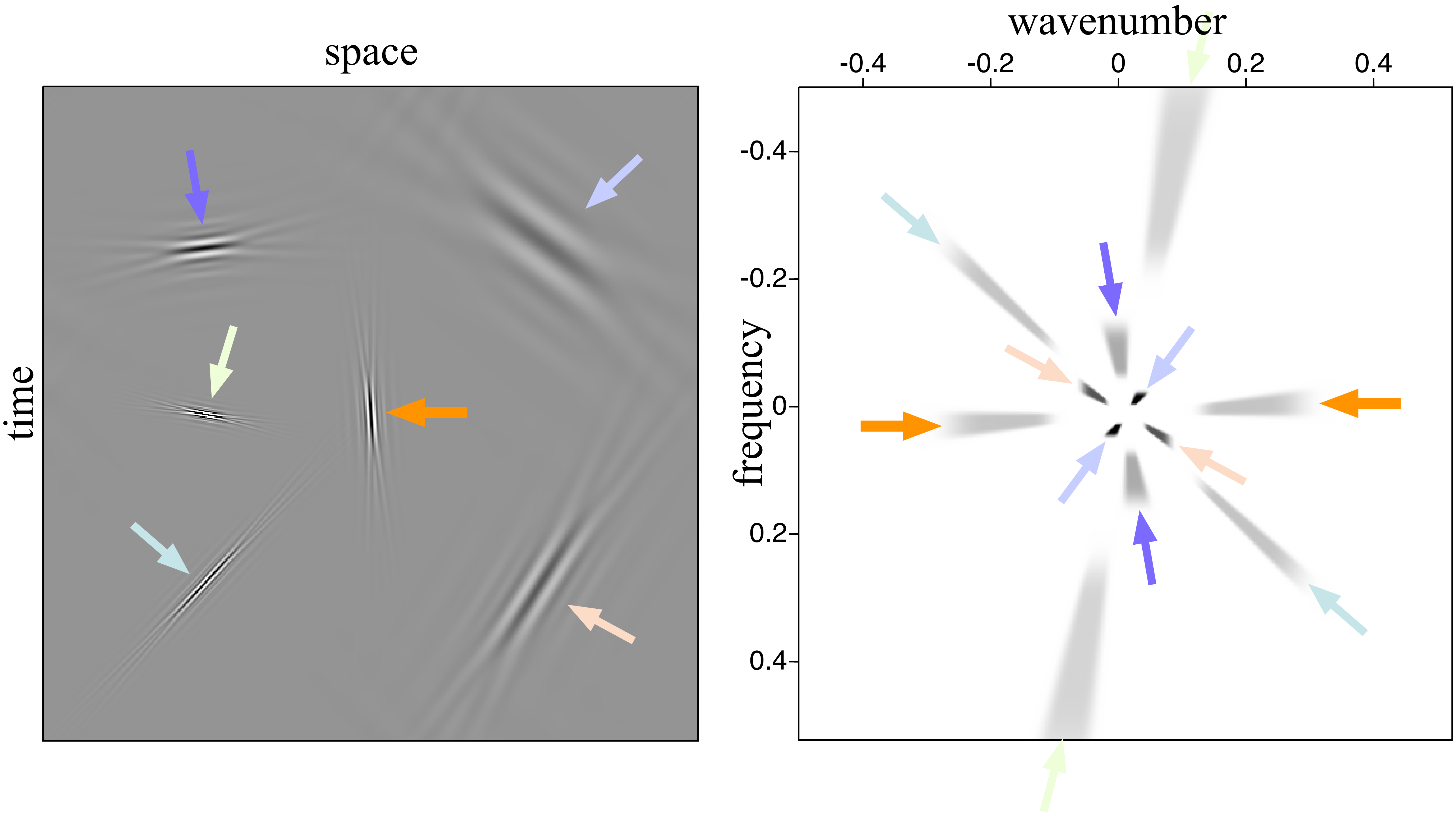
Selection turns *aliases/coherent subsampling artifacts* into harmless ***noise*** ...

Problem: CS does not yet provide practical design principles ...

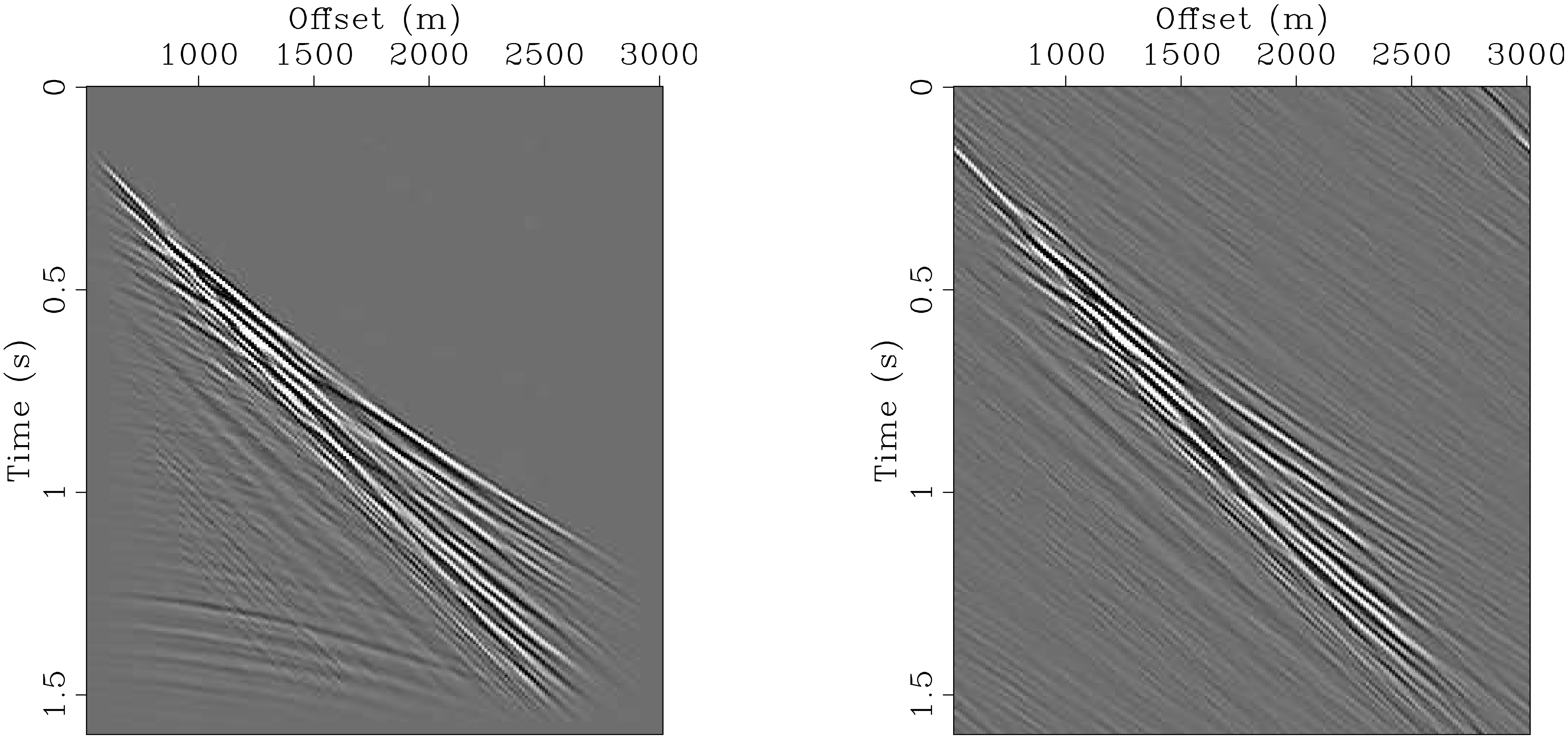
Key elements

- ☐ *sparsifying transform*
  - typically *strictly* **localized** in the Fourier space
  - rapid decay physical space to handle the complexity of seismic data
  - mutual incoherence
- ☐ *advantageous **randomized** coarse sampling*
  - generates incoherent random undersampling “noise” in sparsifying domain
  - ***spatial sampling*** that does not create large gaps
    - because of the limited spatiotemporal extent of transform elements used for the reconstruction
  - ***randomized*** subsampling of ***simultaneous-source*** experiments
    - does not create large interferences
    - leads to compression of linear systems
      - reduction # right-hand-sides
      - spectral representation operators
- ☐ *sparsity-promoting solver*
  - requires few matrix-vector multiplications

2D discrete curvelets

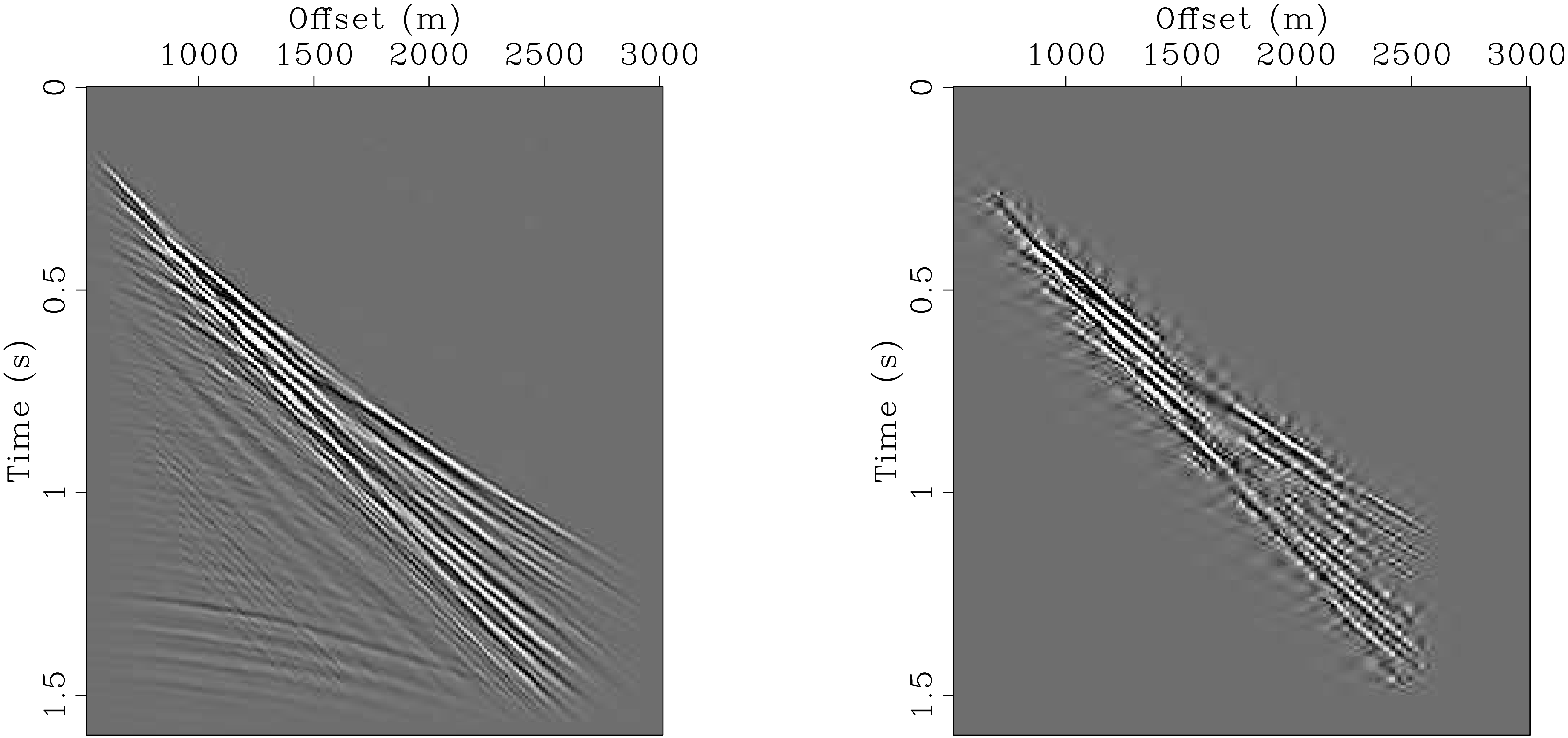


Fourier reconstruction



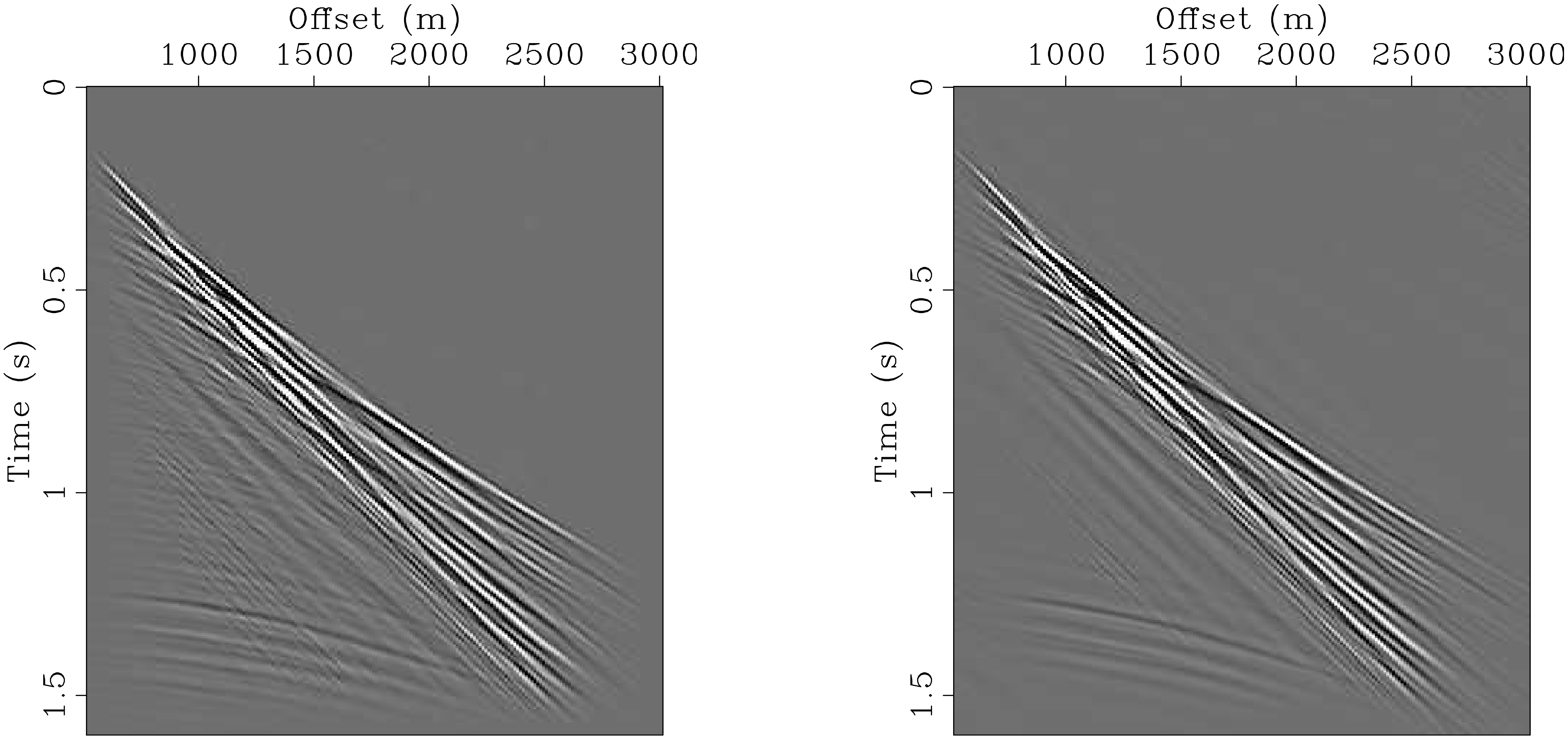
1 % of coefficients

Wavelet reconstruction



1 % of coefficients

Curvelet reconstruction



1 % of coefficients

Key elements

☑ *sparsifying transform*

- typically *strictly* **localized** in the Fourier space
- rapid decay physical space to handle the complexity of seismic data
- mutual incoherence

☐ *advantageous* ***randomized*** *coarse sampling*

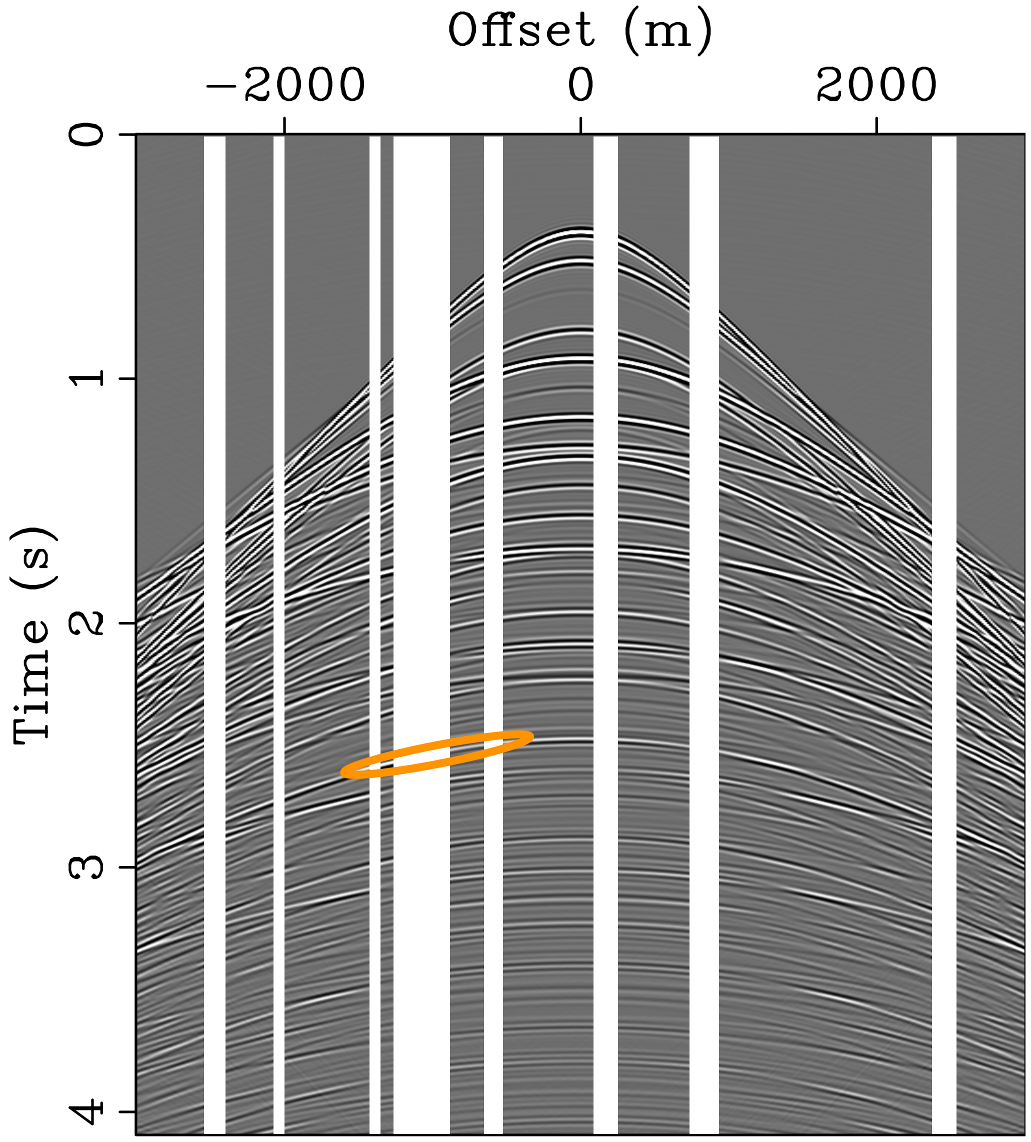
- generates incoherent random undersampling “noise” in sparsifying domain
- ***spatial sampling*** that does not create large gaps
  - because of the limited spatiotemporal extent of transform elements used for the reconstruction
- ***randomized*** subsampling of ***simultaneous-source*** experiments
  - does not create large interferences
  - leads to compression of linear systems
    - reduction # right-hand-sides
    - spectral representation operators

☐ *sparsity-promoting solver*

- requires few matrix-vector multiplications

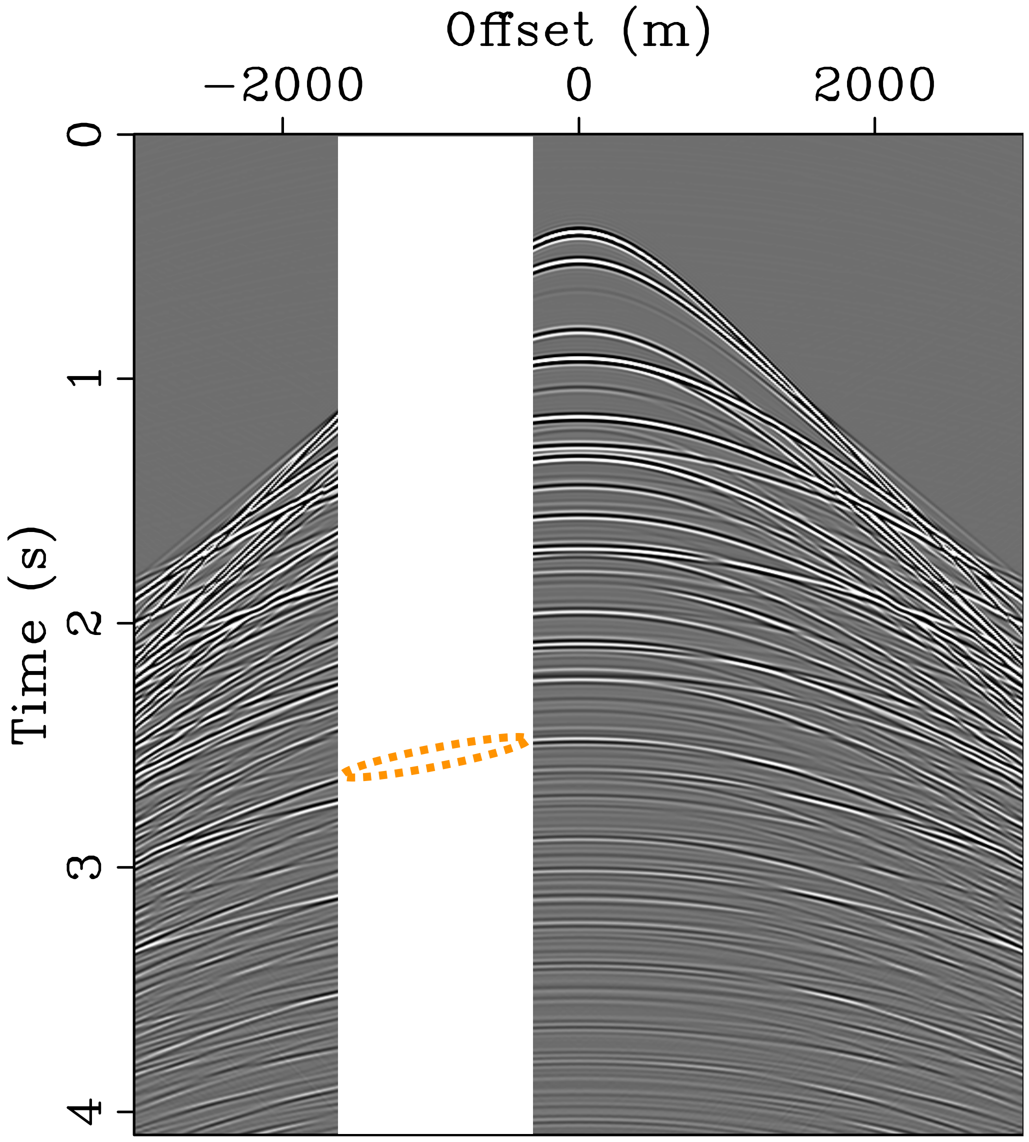
Localized transform elements & gap size

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



Data

✓



Data

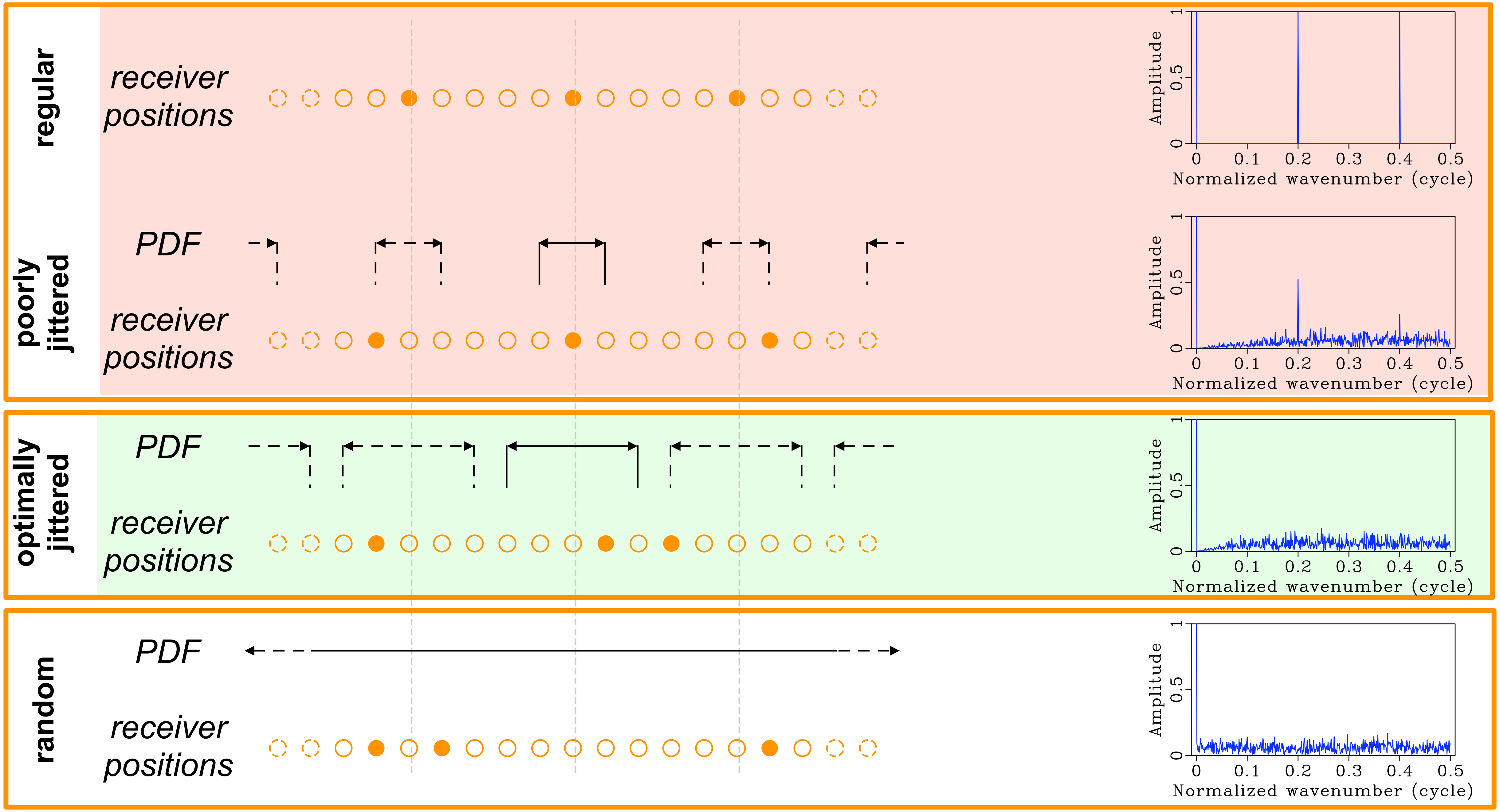
✗

Discrete *randomized* jittered undersampling

Type

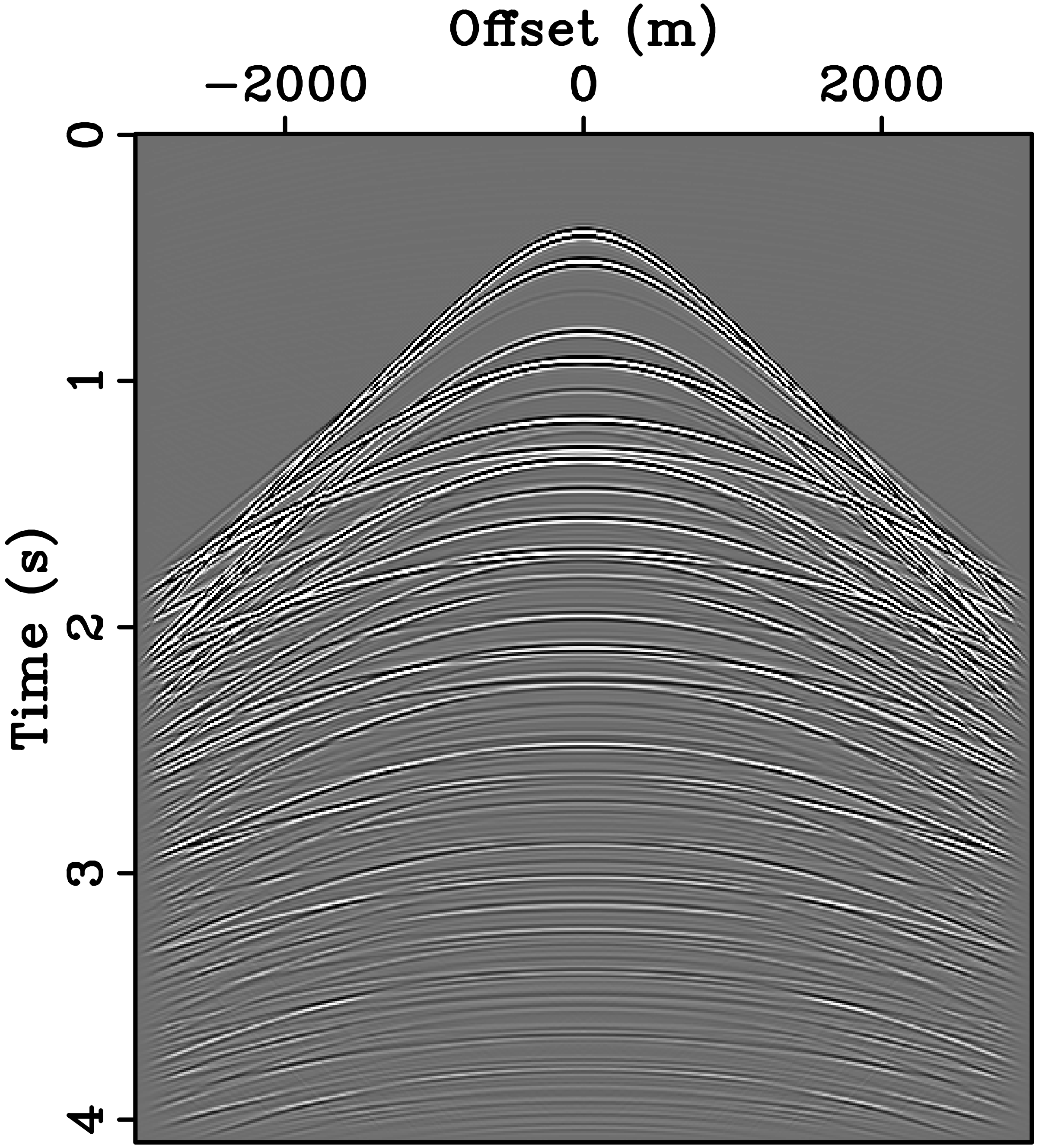
Sampling scheme

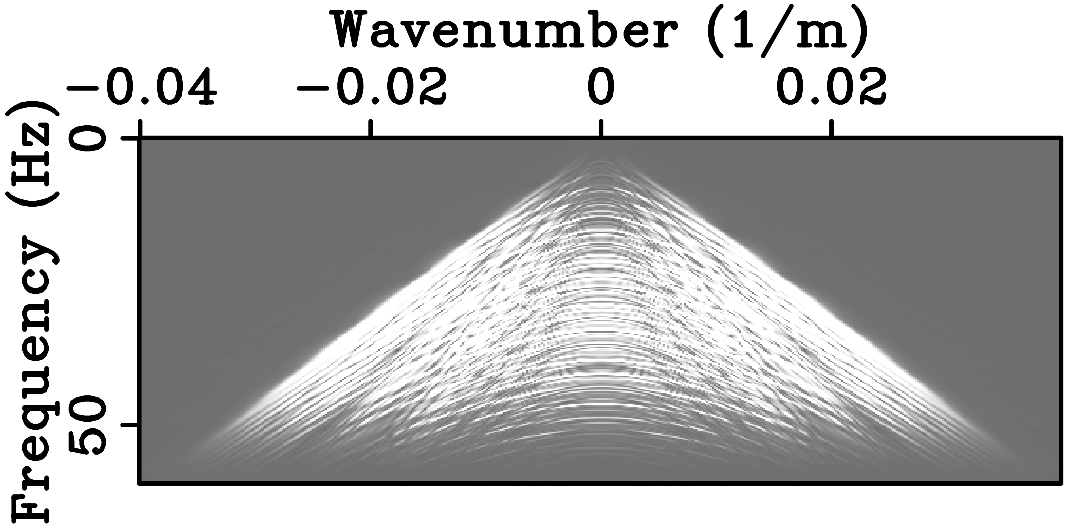
Typical spatial convolution kernel



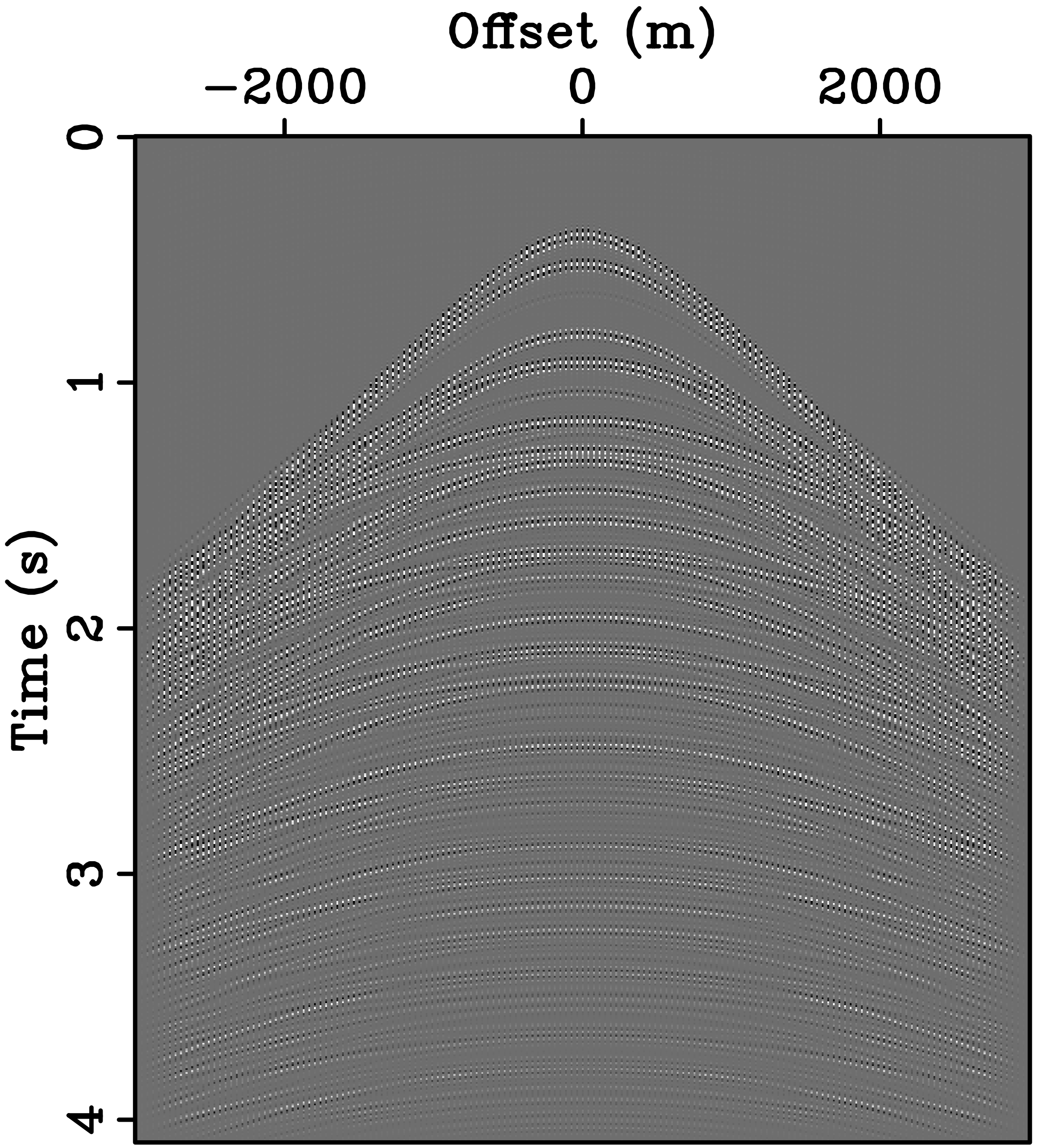
[Hennenfent and FJH ‘08]

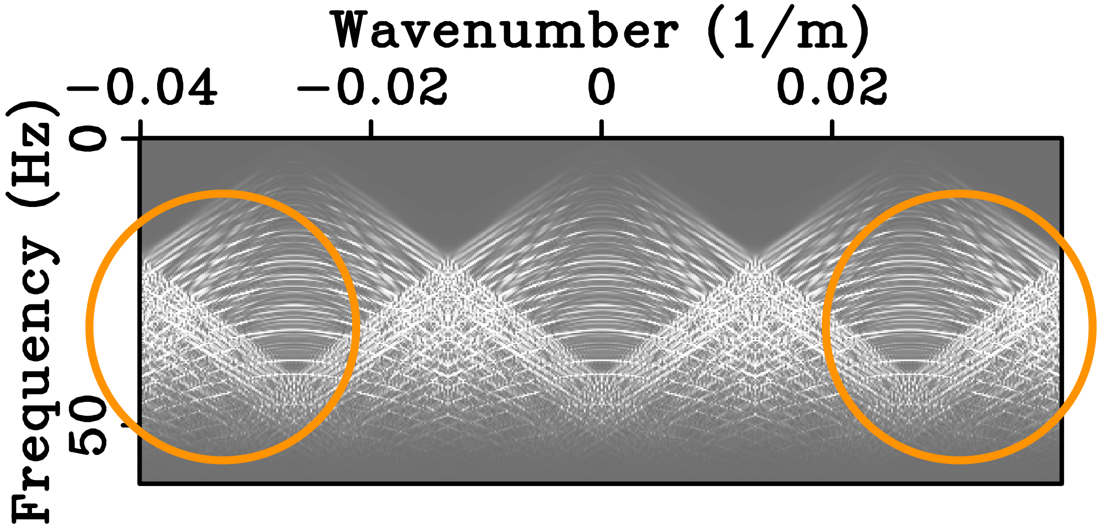
Model


Offset (m)
−2000
0
2000
Time (s)
0
1
2
3
4

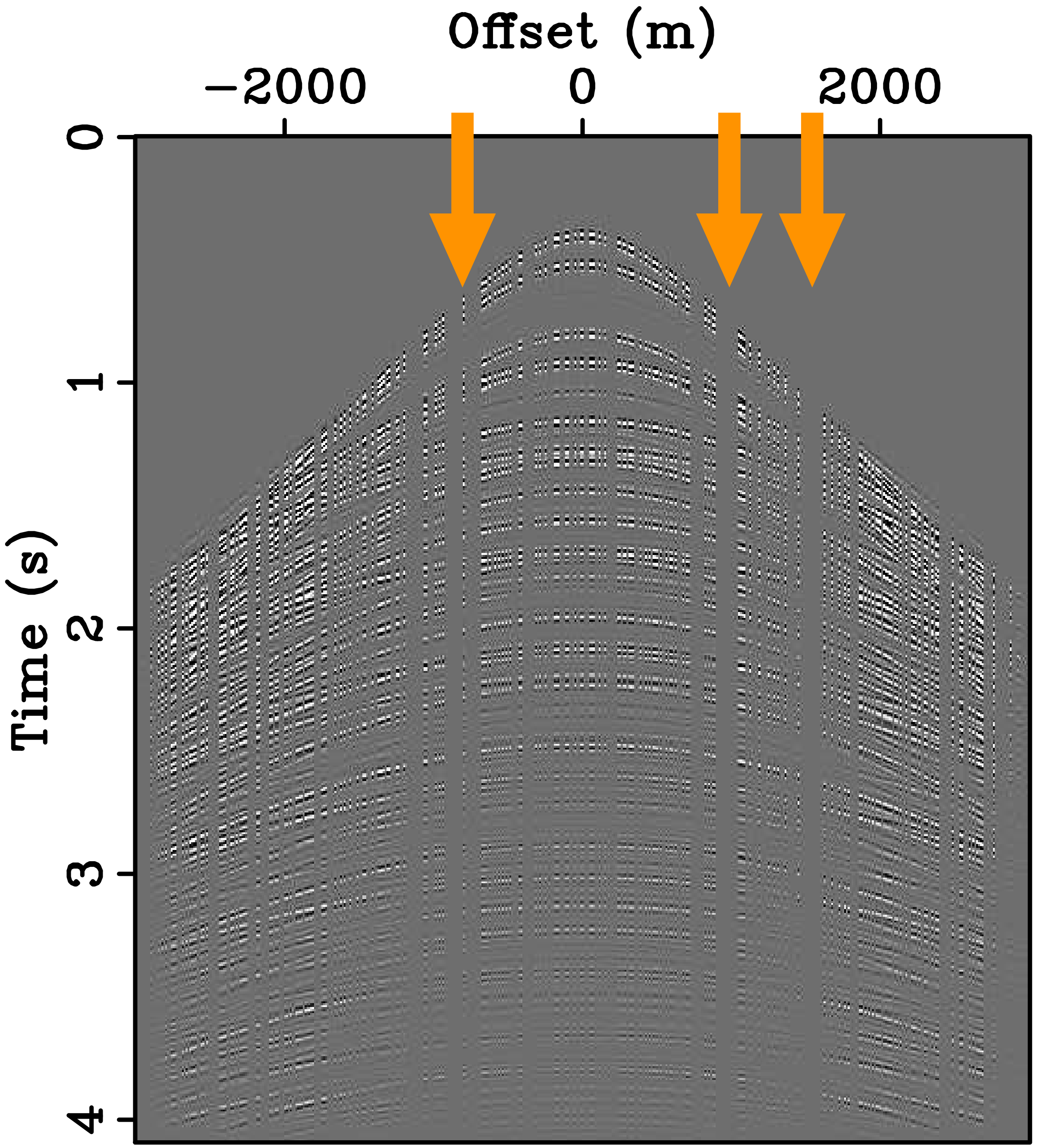

Wavenumber (1/m)
−0.04
−0.02
0
0.02
Frequency (Hz)
0
50

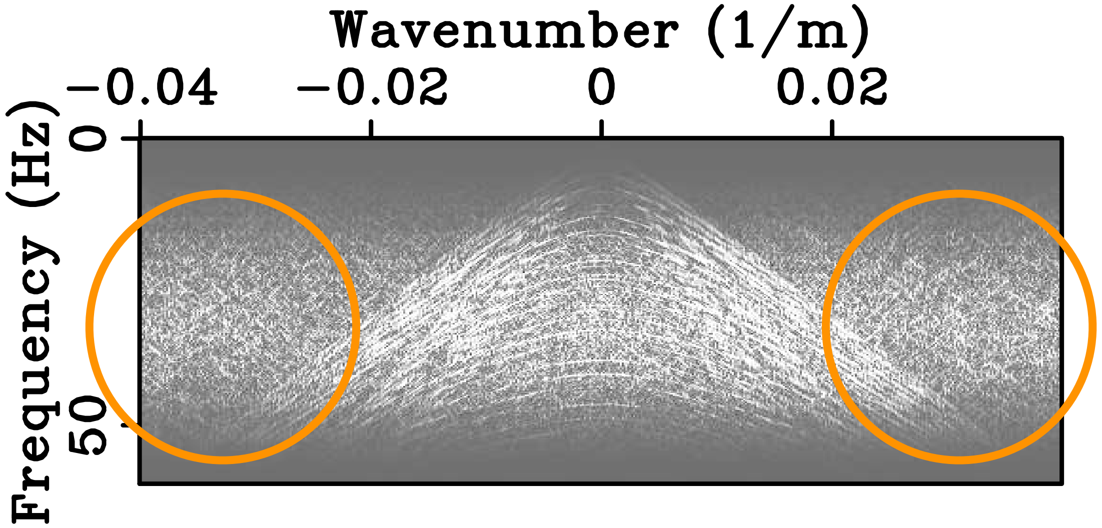
Regular 3-fold undersampling



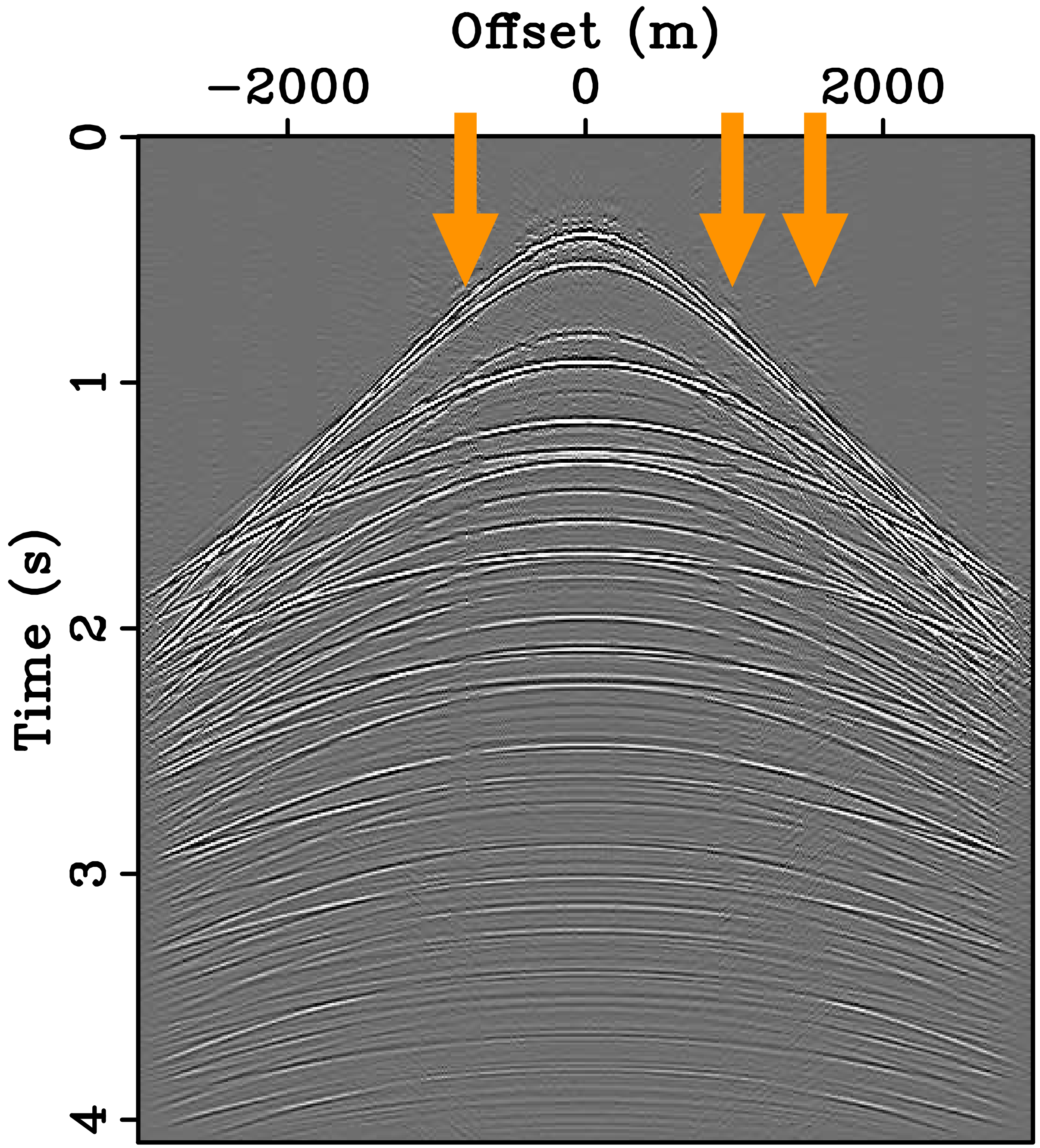


Random 3-fold undersampling

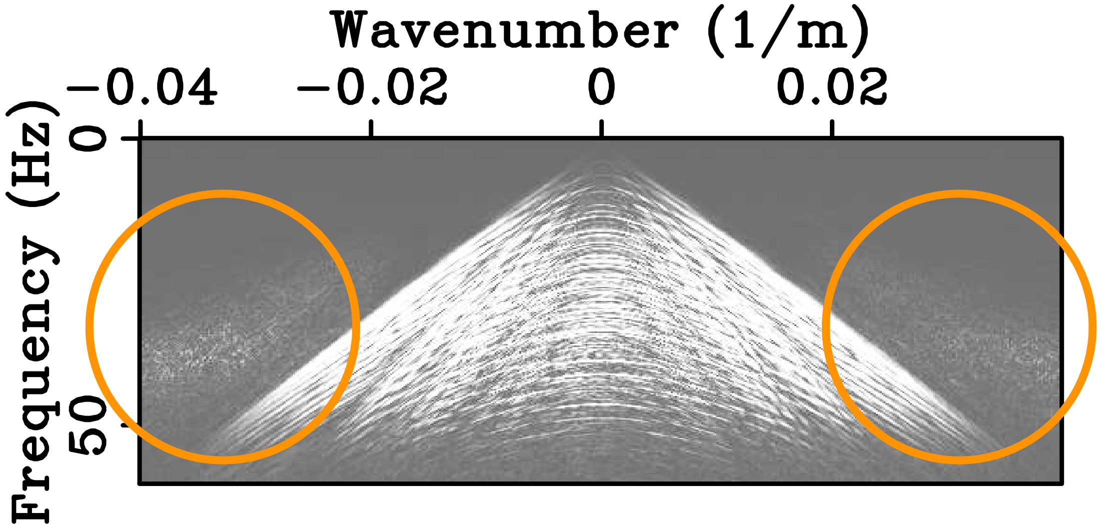




CRSI from random 3-fold undersampling

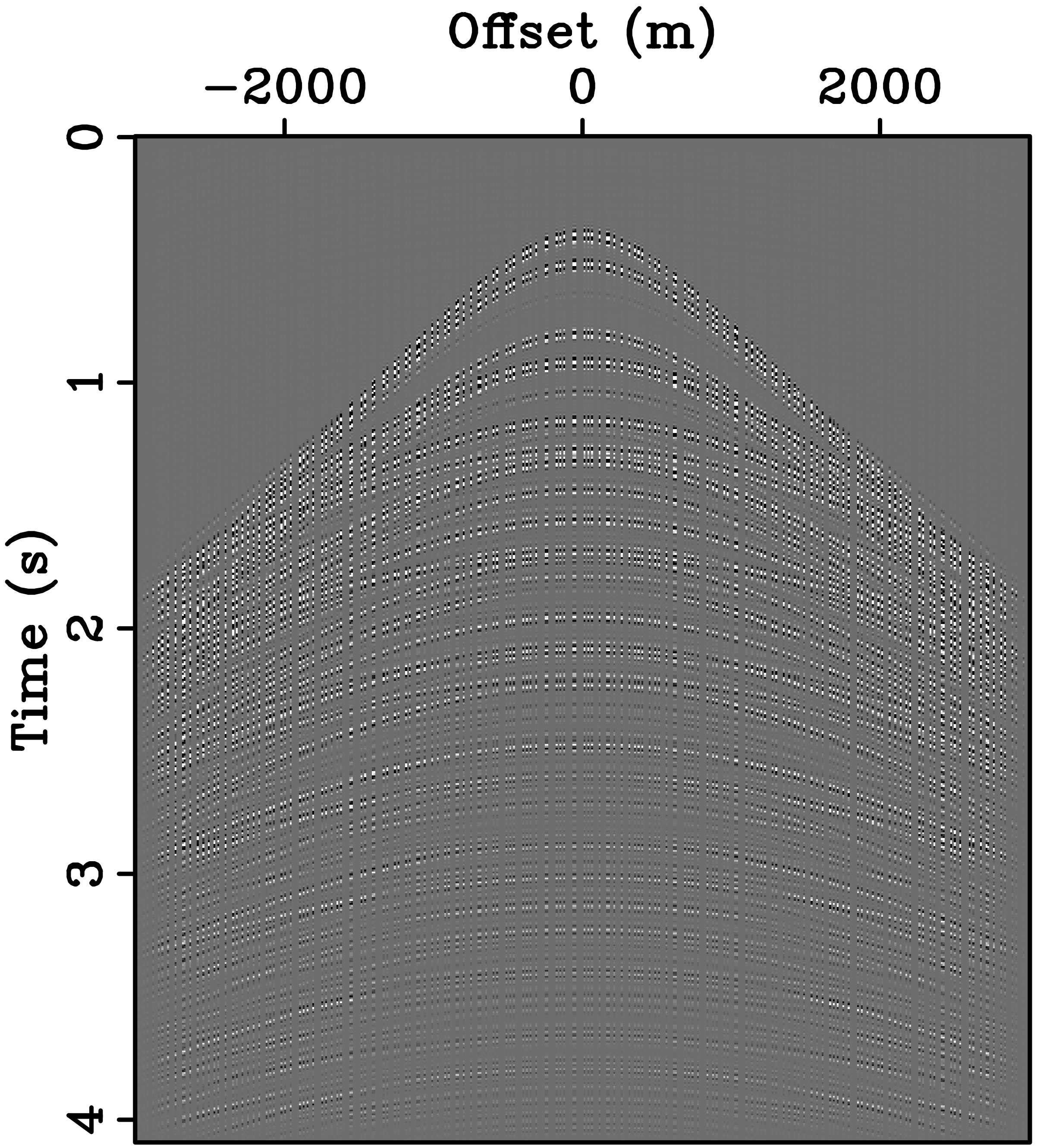


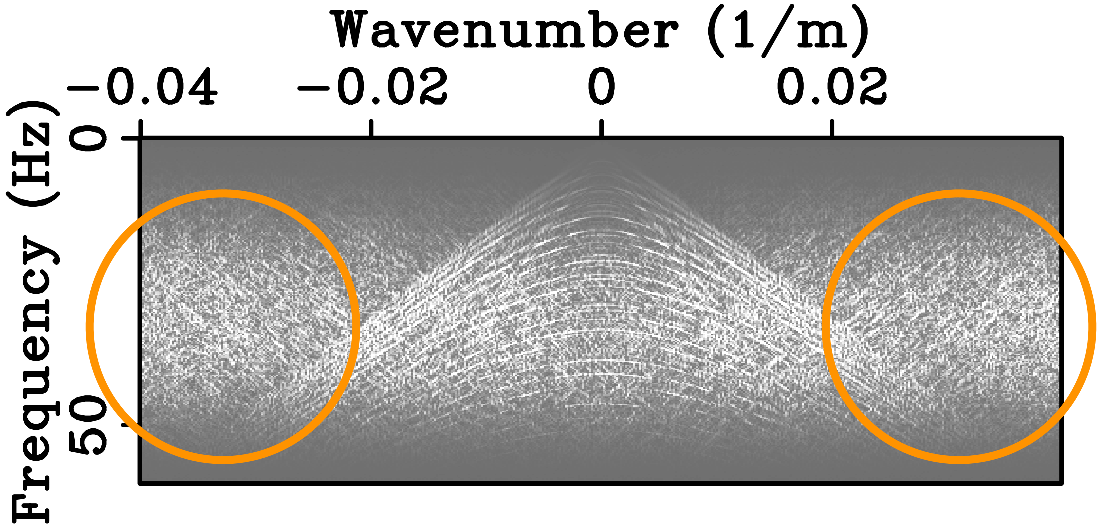
SNR = 9.72 dB



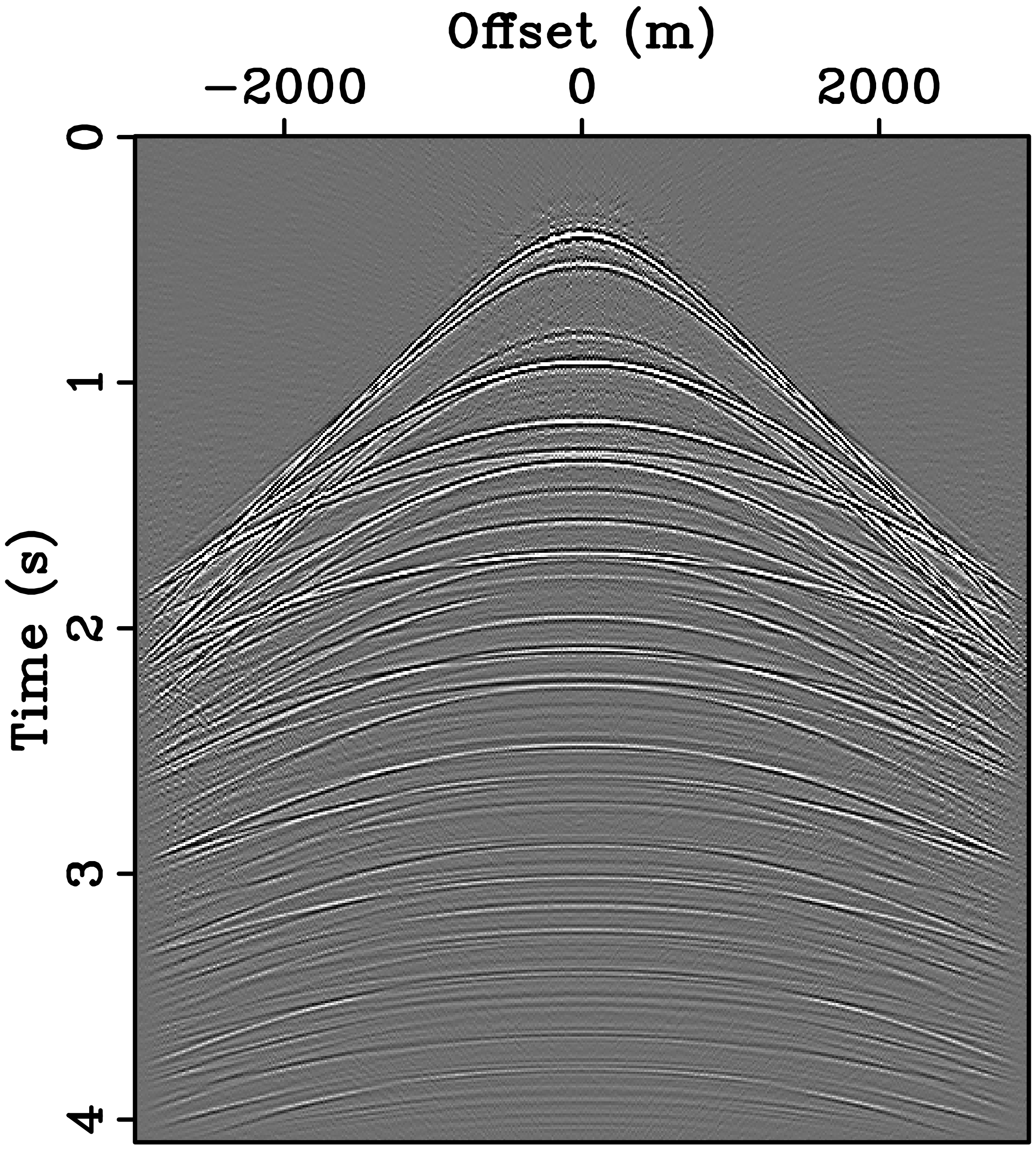
$$\mathrm{SNR} = 20 \times \log_{10}\left(\frac{\|\mathrm{model}\|_2}{\|\mathrm{reconstruction\ error}\|_2}\right)$$

Optimally-jittered 3-fold undersampling

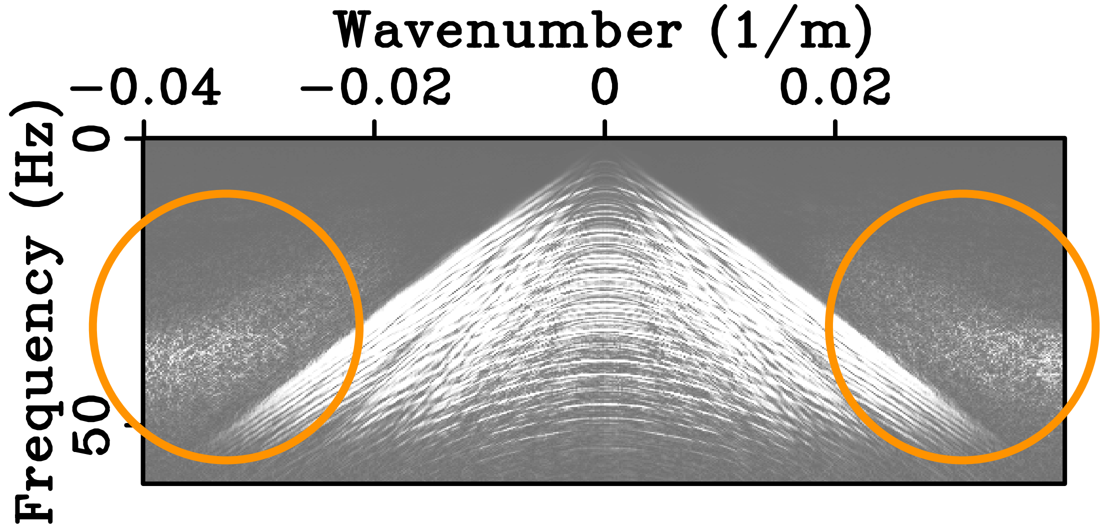




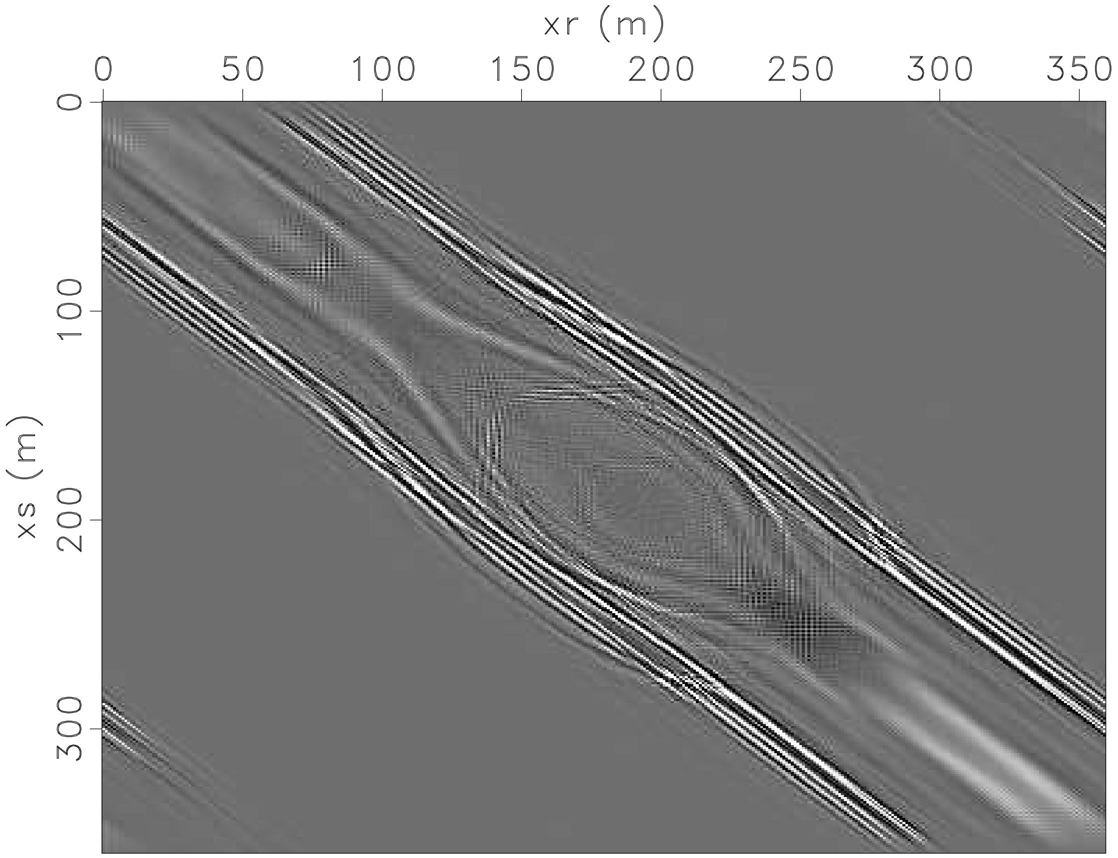
CRSI from opt.-jittered 3-fold undersampling



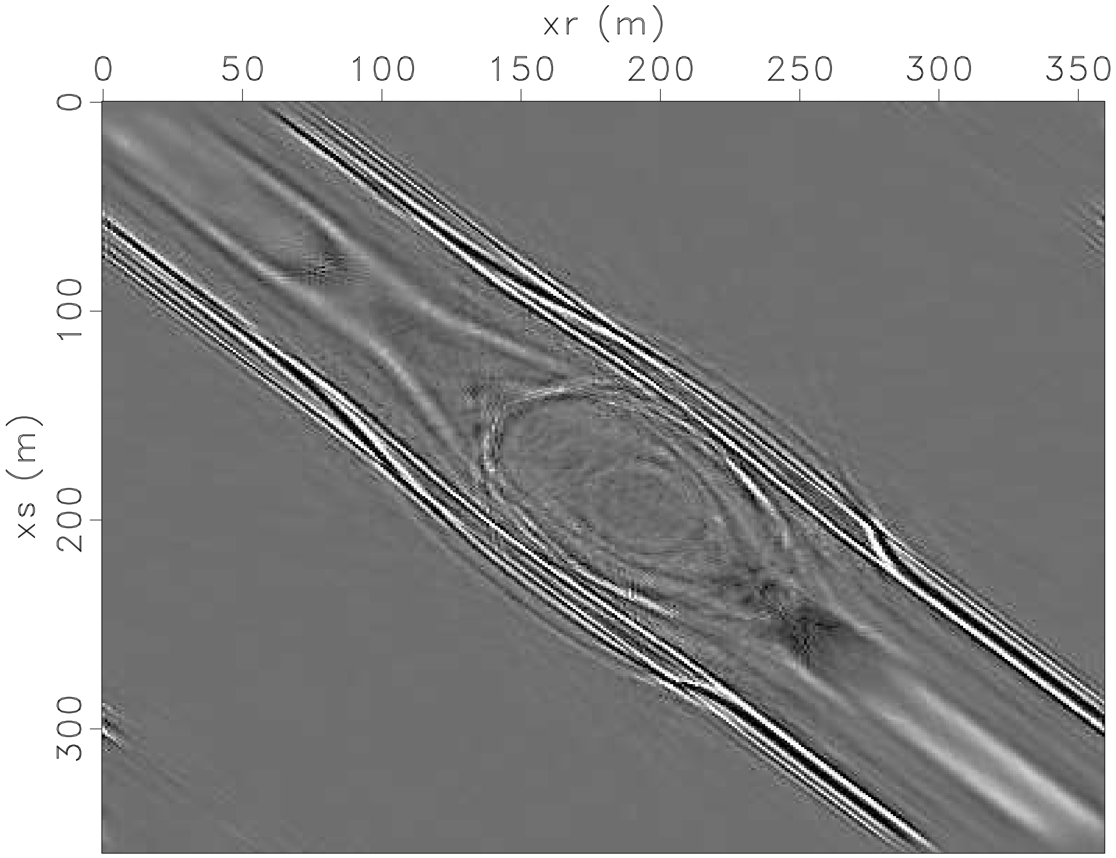
SNR = 10.42 dB



Regular vs uniform randomized 2D sampling



CRSI reconstruction from regular 2-D sampling (25% of data taken)
SNR: 4.161 dB

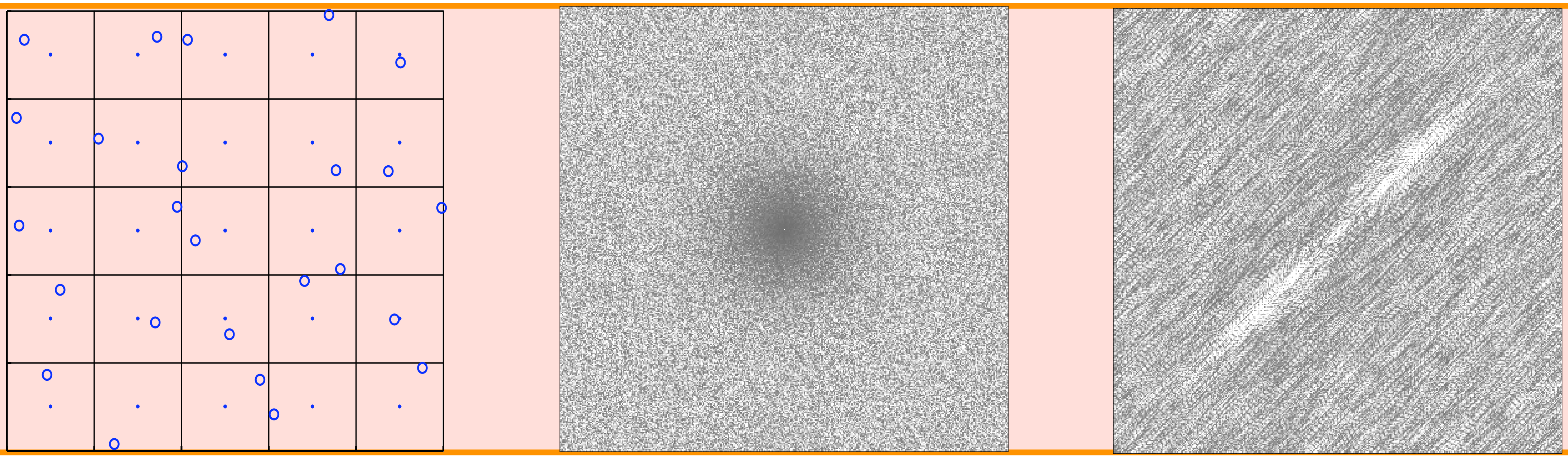


CRSI reconstruction from randomized 2-D sampling (25% of data taken)
SNR: 9.979 dB

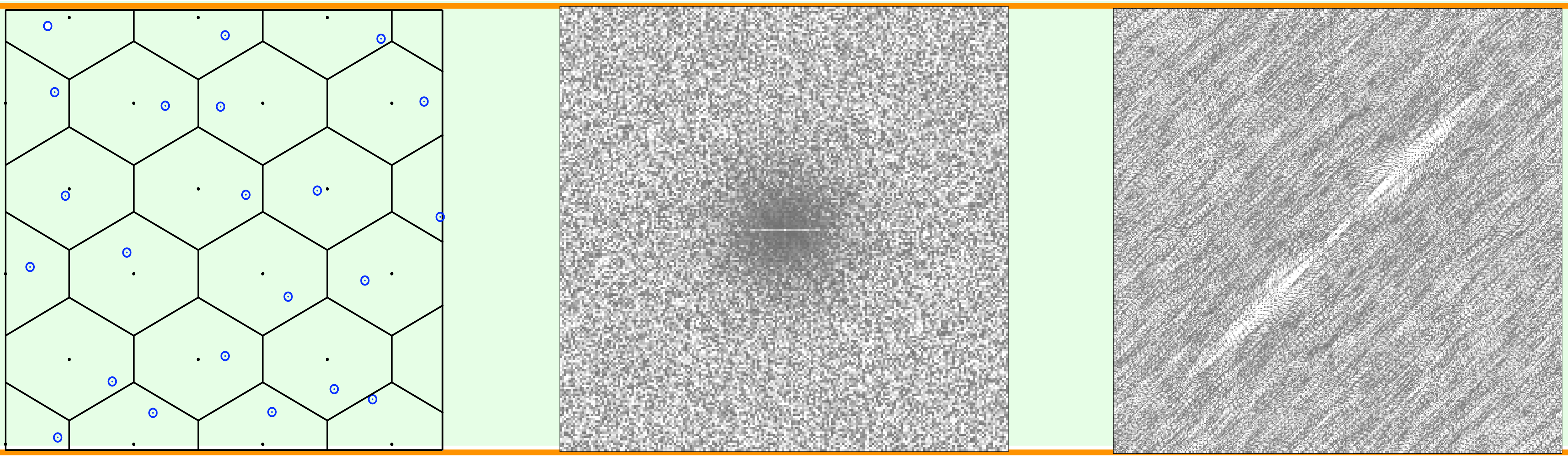
2-D discrete random *jittered* sampling

Type **Sampling scheme** **Mask's spectra** **Samples' spectra**

Square

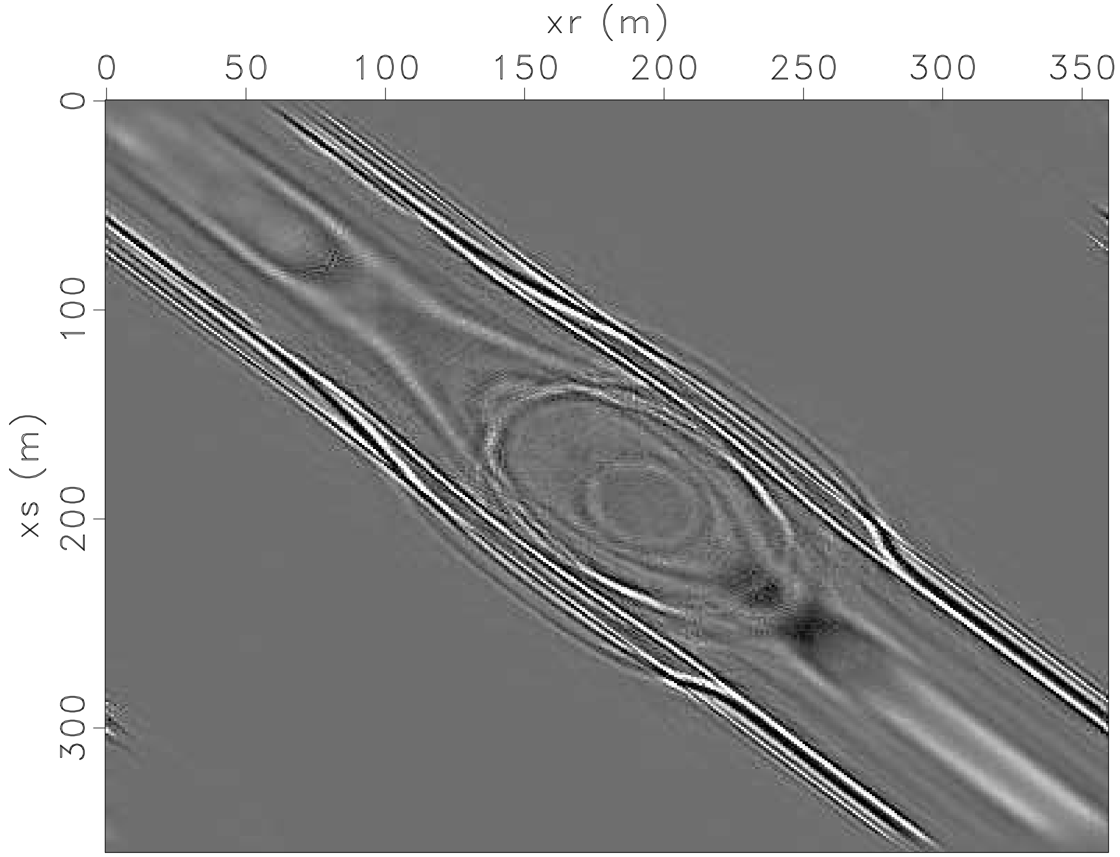


Hexagon

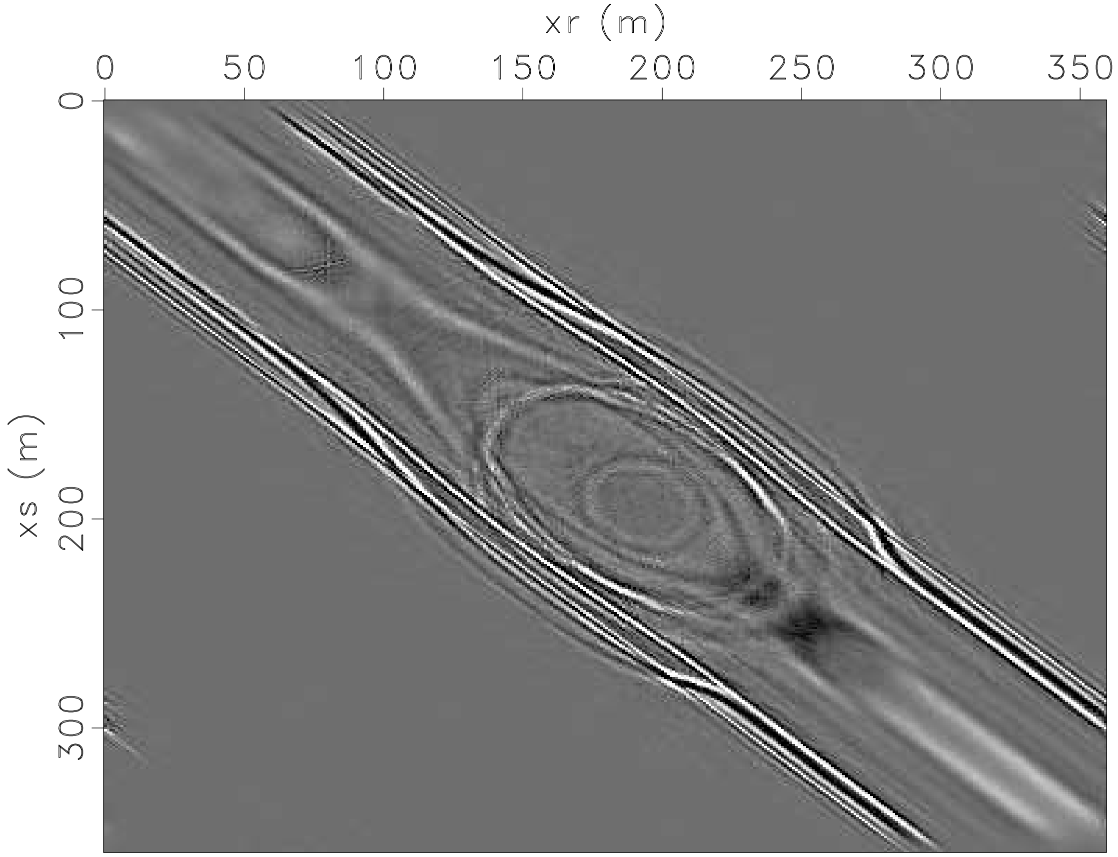


2-D discrete random jittered subsampling

- Cartesian & hexagonal jittered reconstructions almost the same.

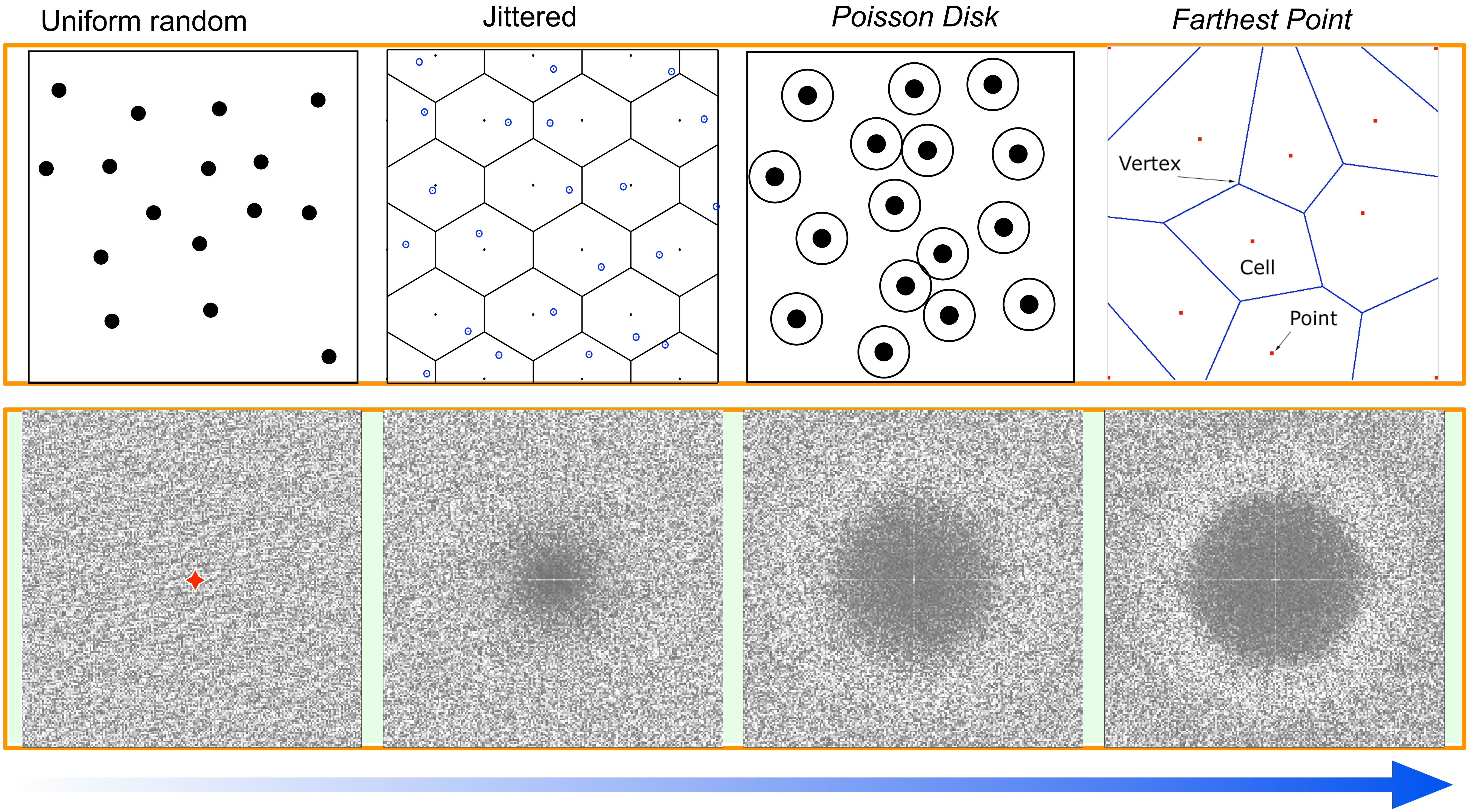


CRSI Recovery (Cartesian)
SNR = 10.820



CRSI Recovery (Hexagonal)
SNR = 10.904

Blue-noise spectra from 2D sampling methods



Uniform random

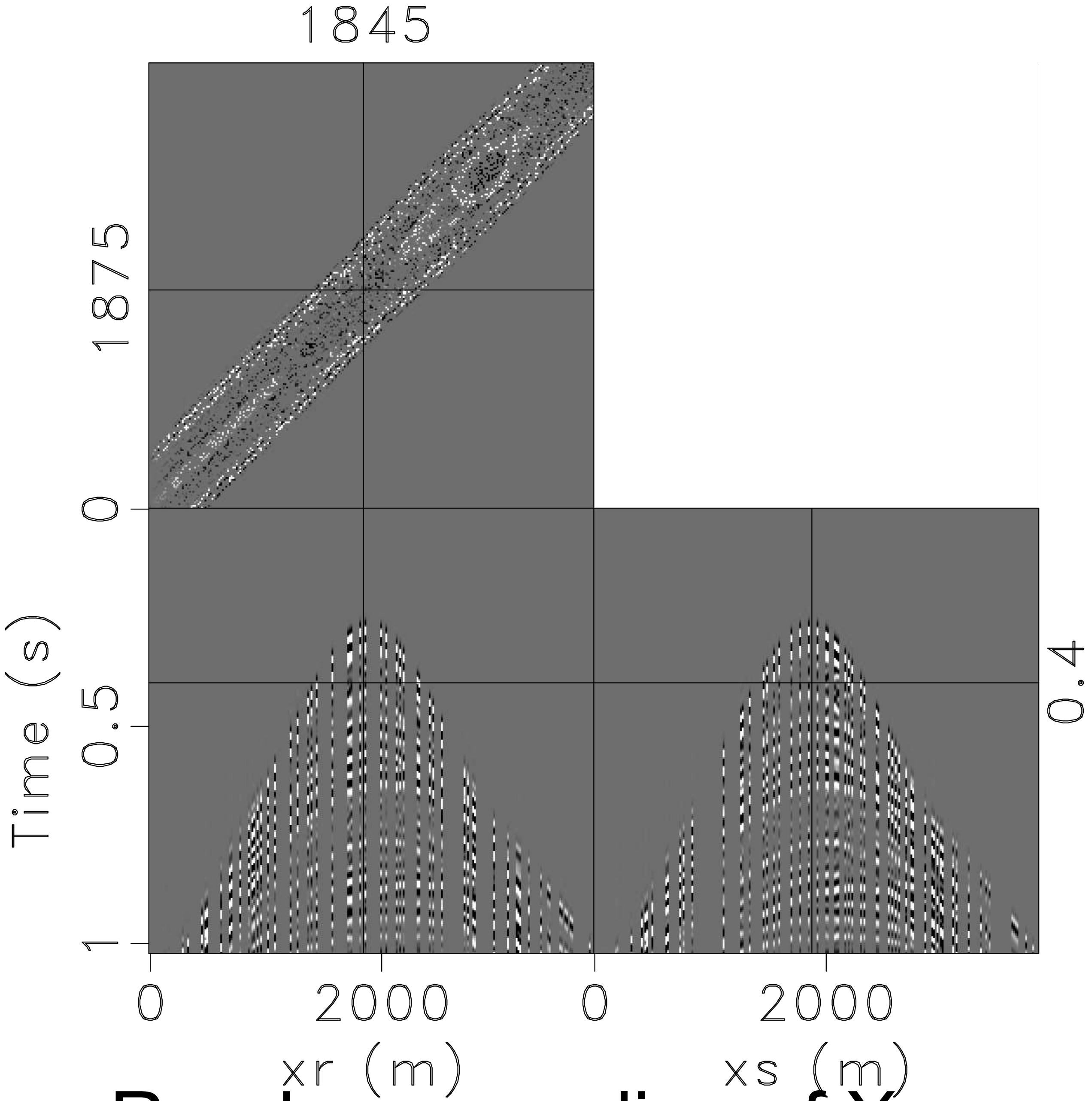
Jittered

Poisson Disk

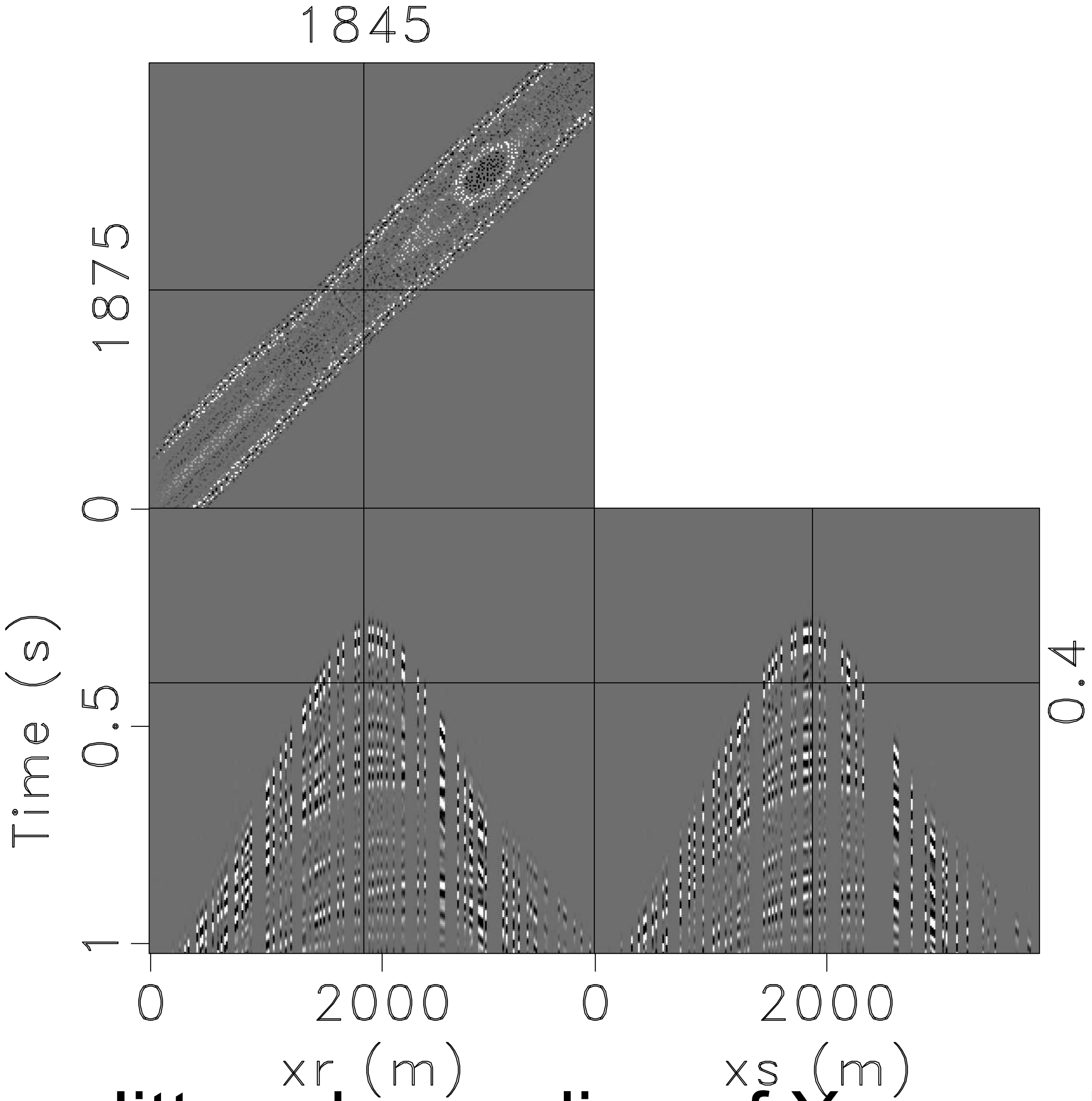
Farthest Point

Spectra become increasingly *“blue”*

Randomized 2D uniform vs jittered

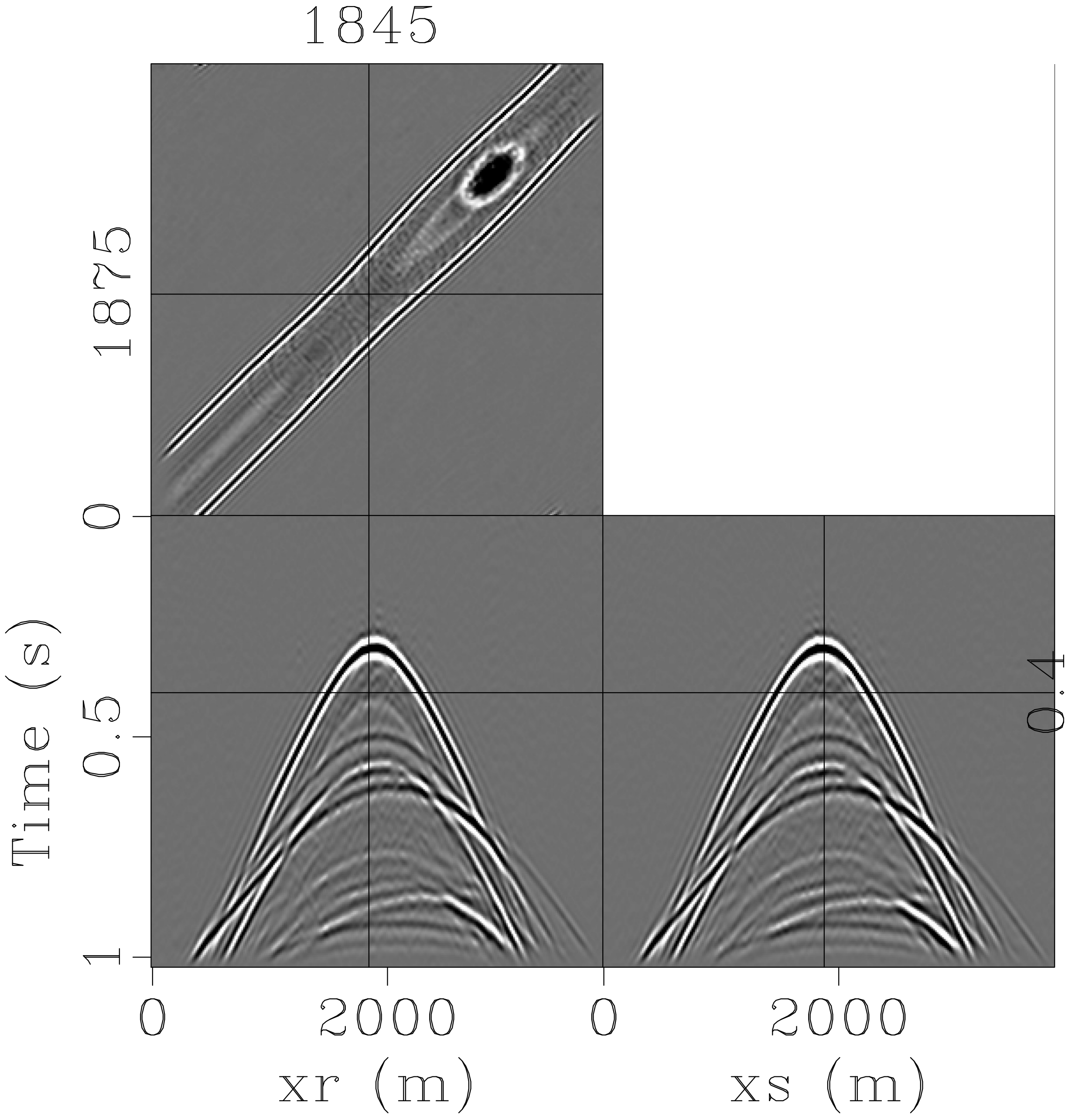


Random sampling of X-spread data (25% total data sampled)

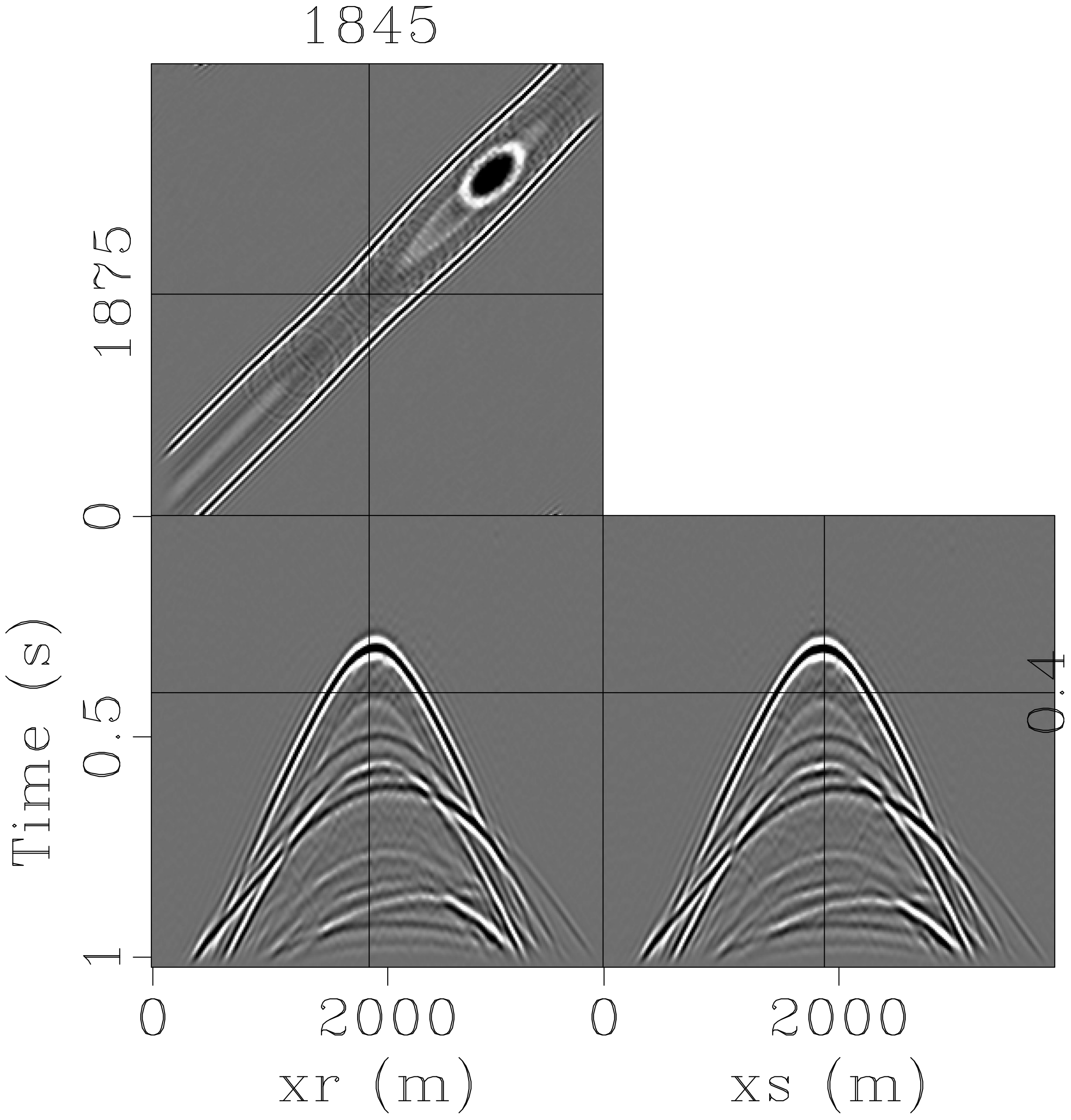


Jittered sampling of X-spread data (25% sampled)

Randomized 2D uniform vs jittered - reconstruction

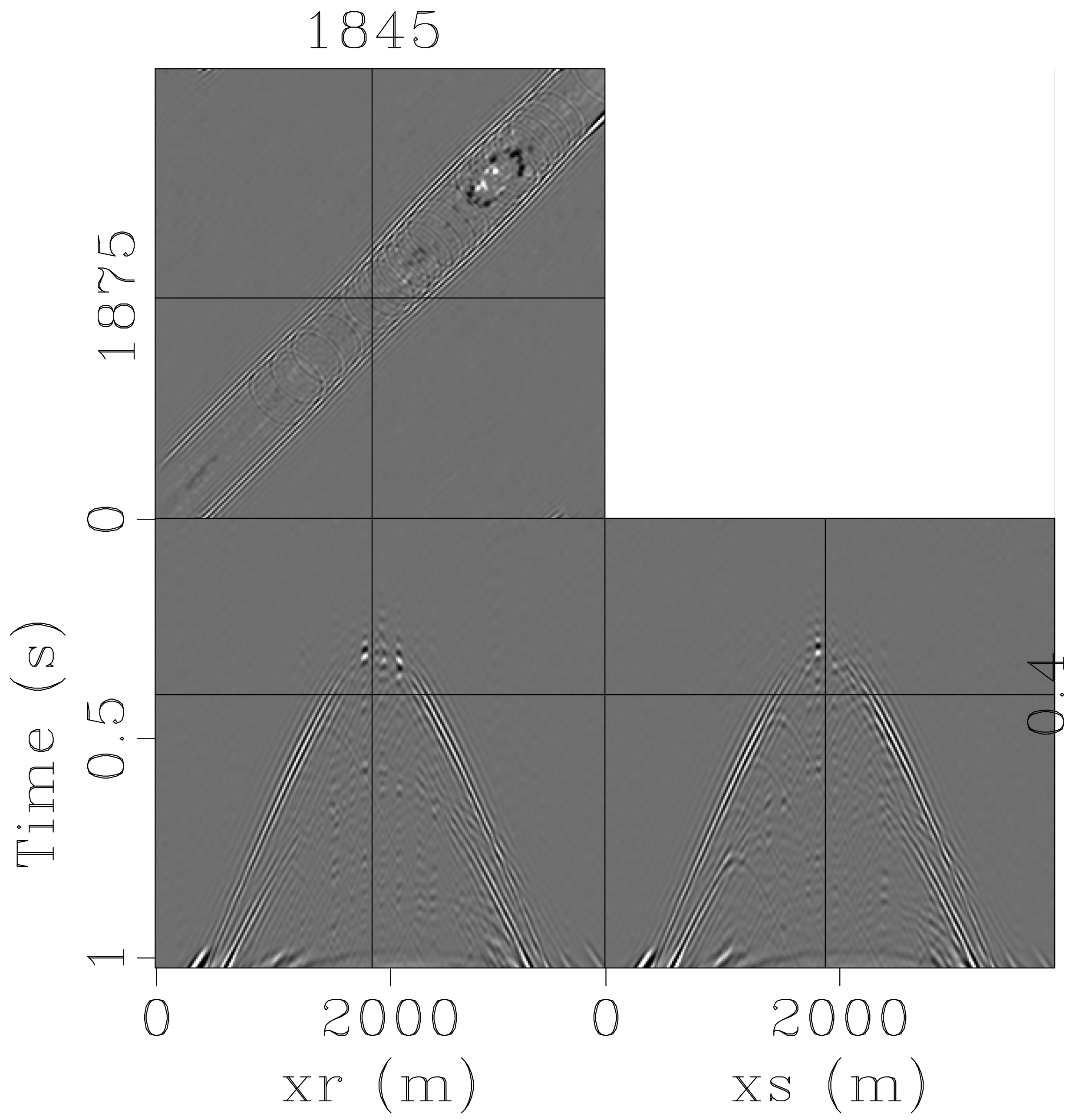


CRSI reconstruction from uniform samples
SNR=8.134

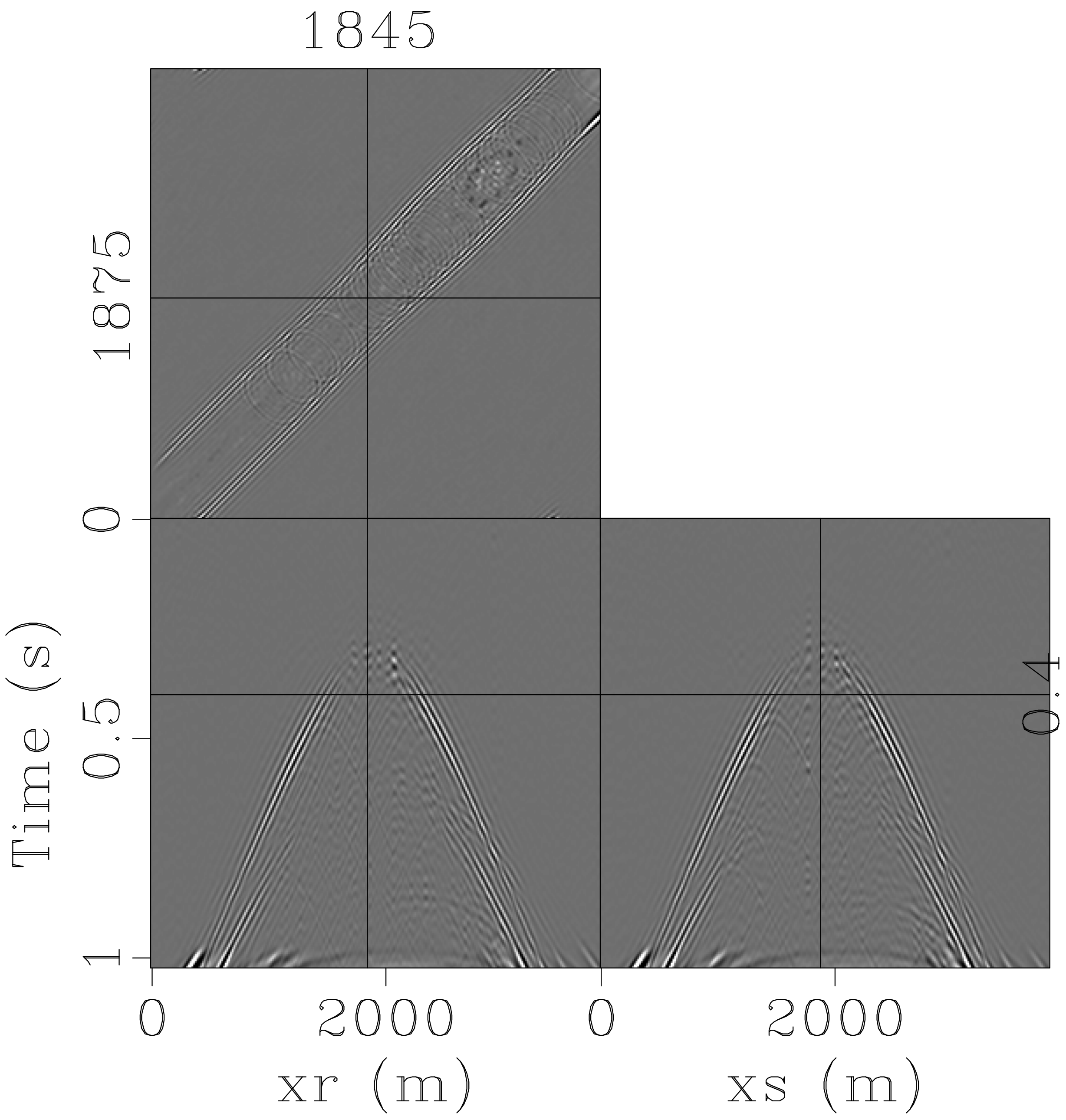


CRSI recon., 2D jittered (hexagonal) samples
SNR=8.434

Randomized 2D uniform vs jittered - residues

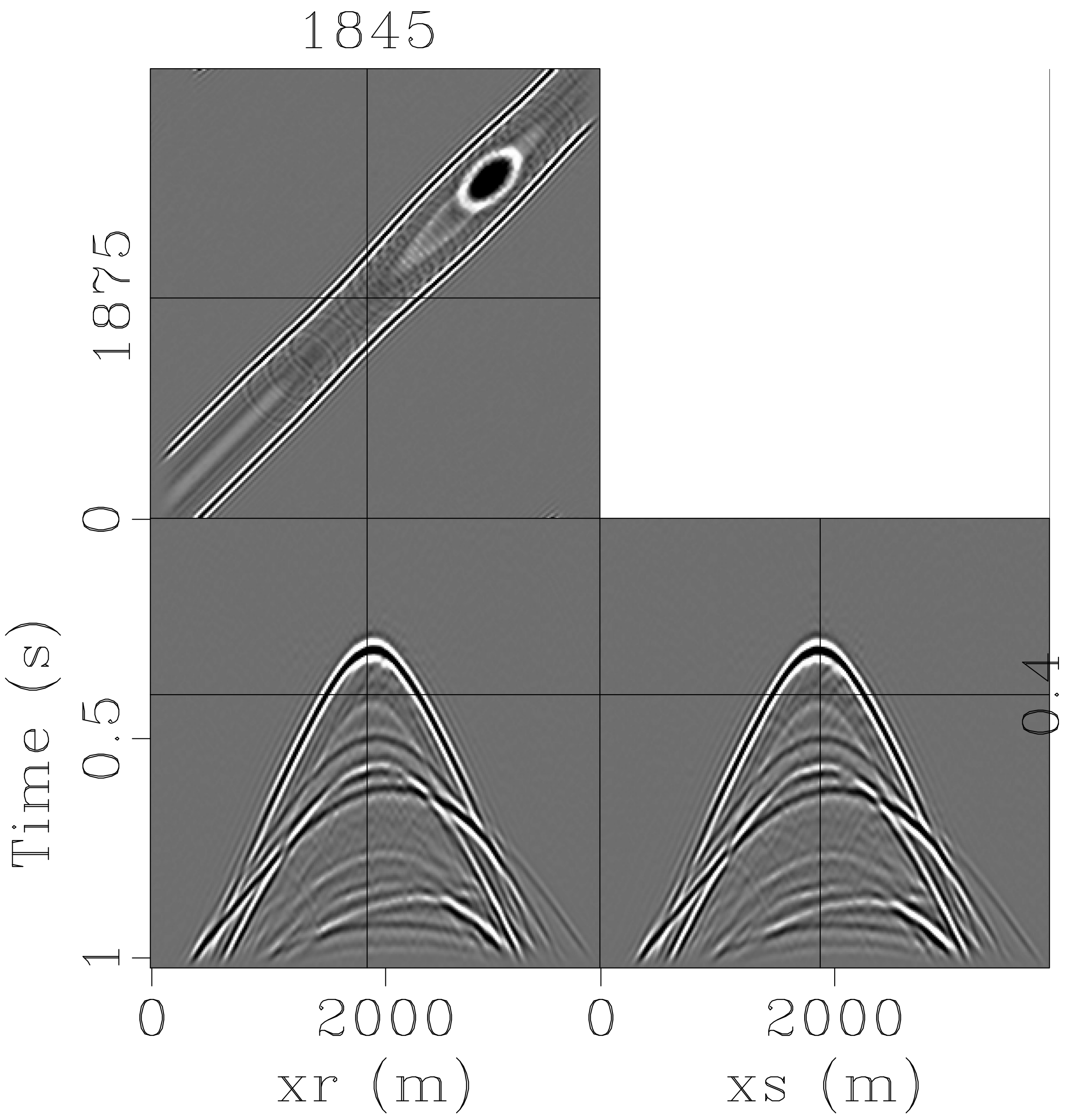


CRSI recon. residual from random samples

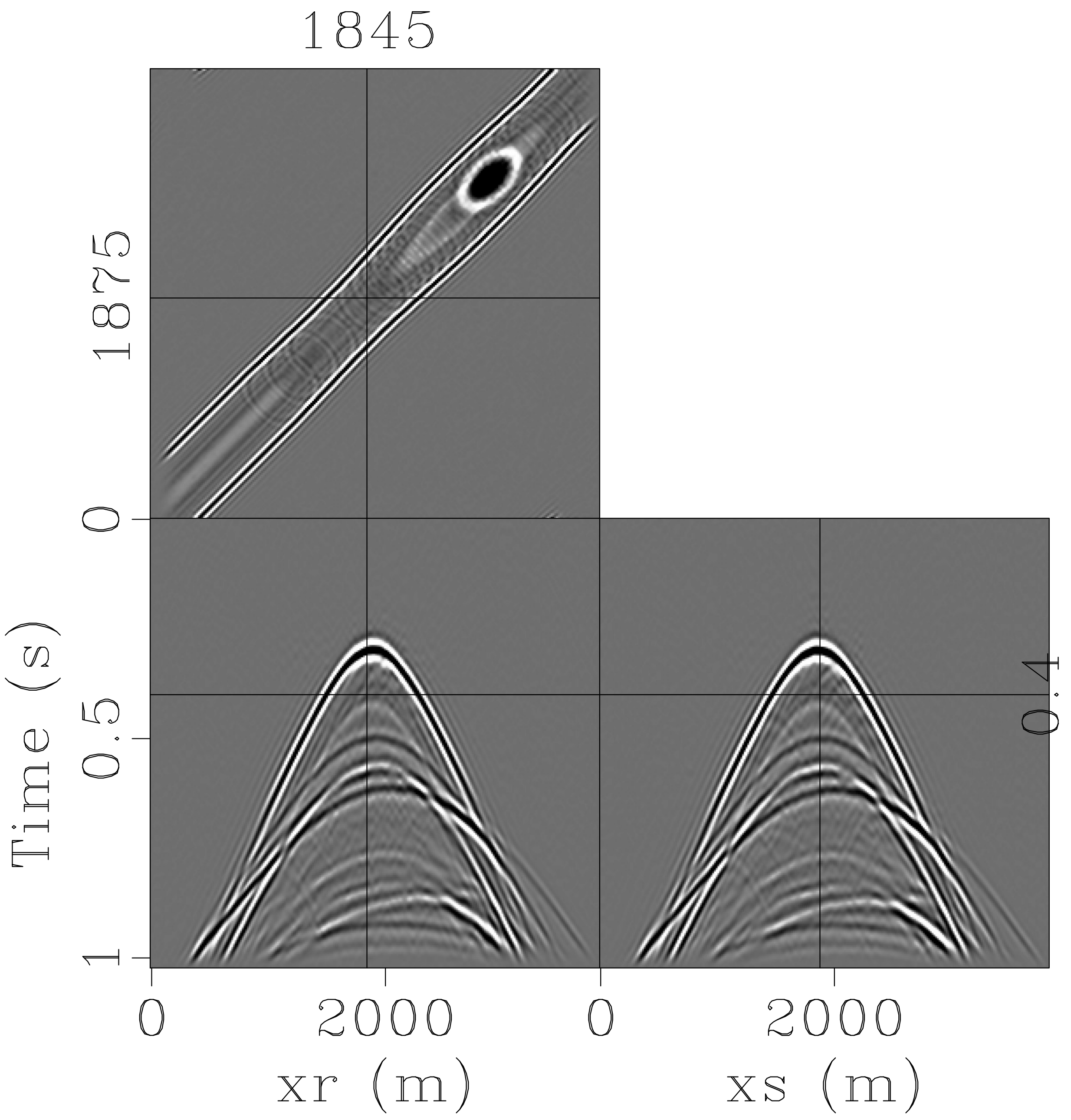


CRSI recon. residual from jittered (hexagonal) samples

Farthest point vs Poisson disk - reconstruction

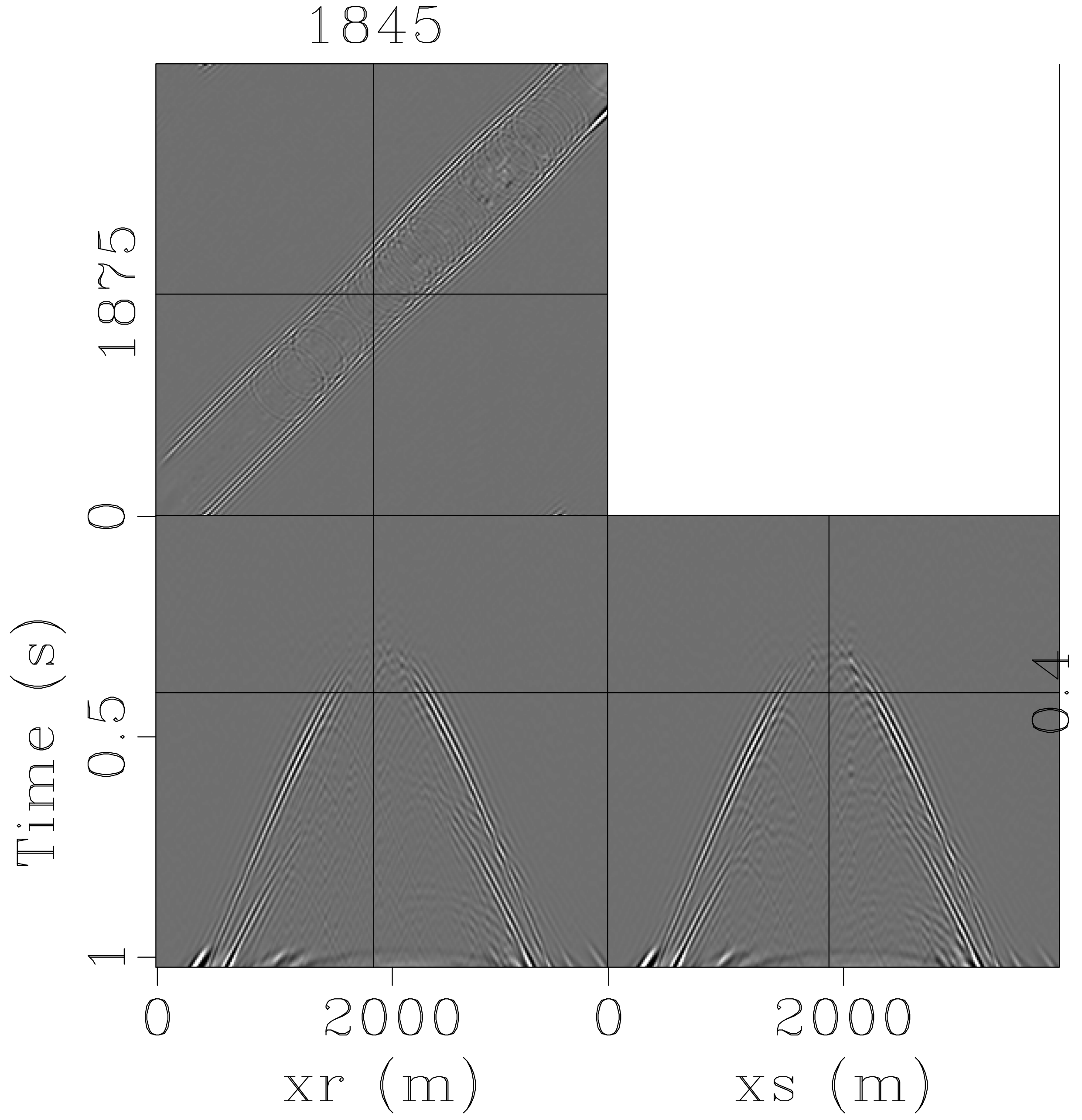


CRSI reconstruction from Farthest Point samples, SNR=8.496



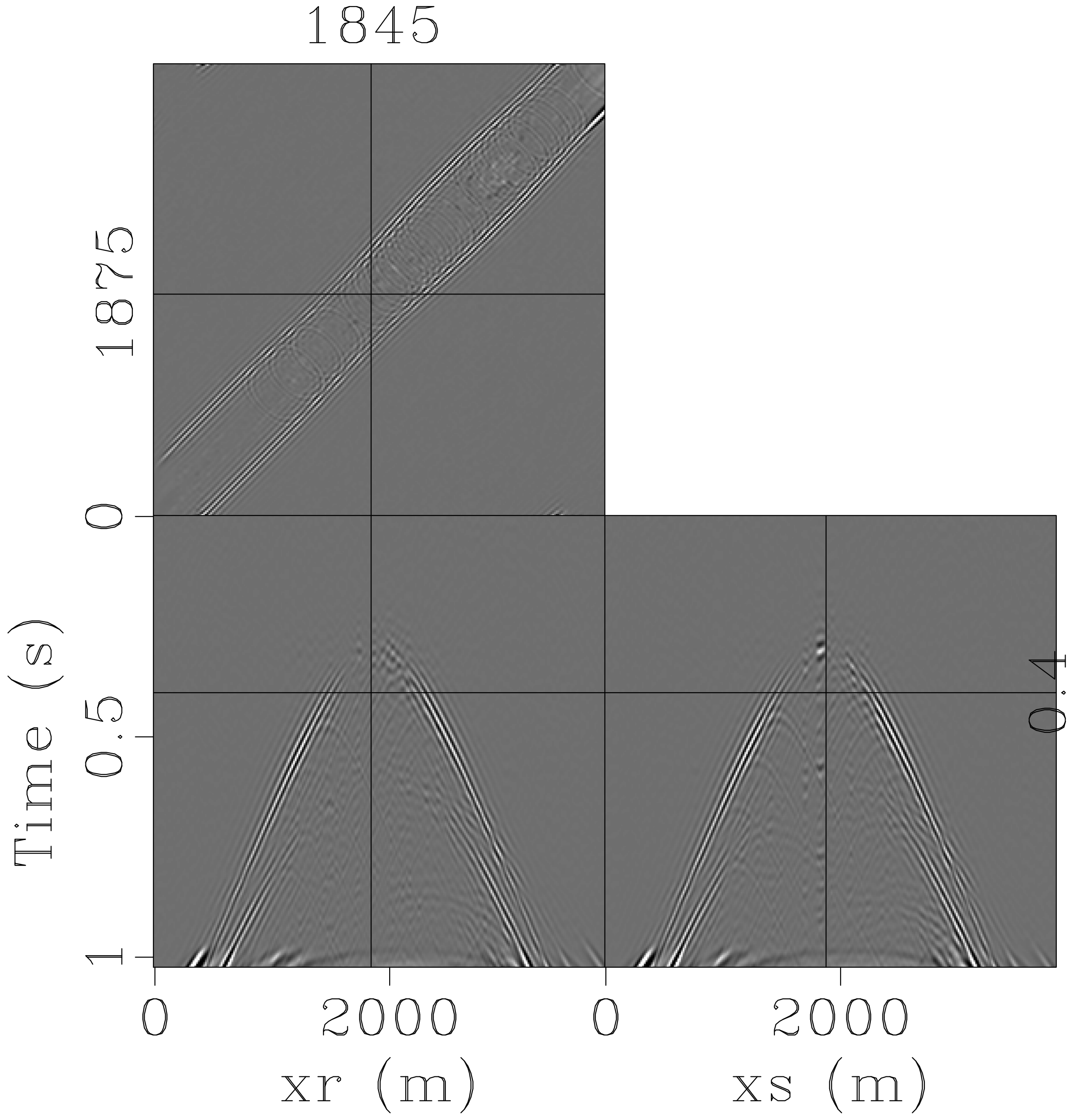
CRSI recon. from Poisson Disk samples SNR=8.483

Farthest point vs Poisson disk - residual



CRSI recon. residual from

Farthest Point samples



CRSI recon. residual from

Poisson Disk samples

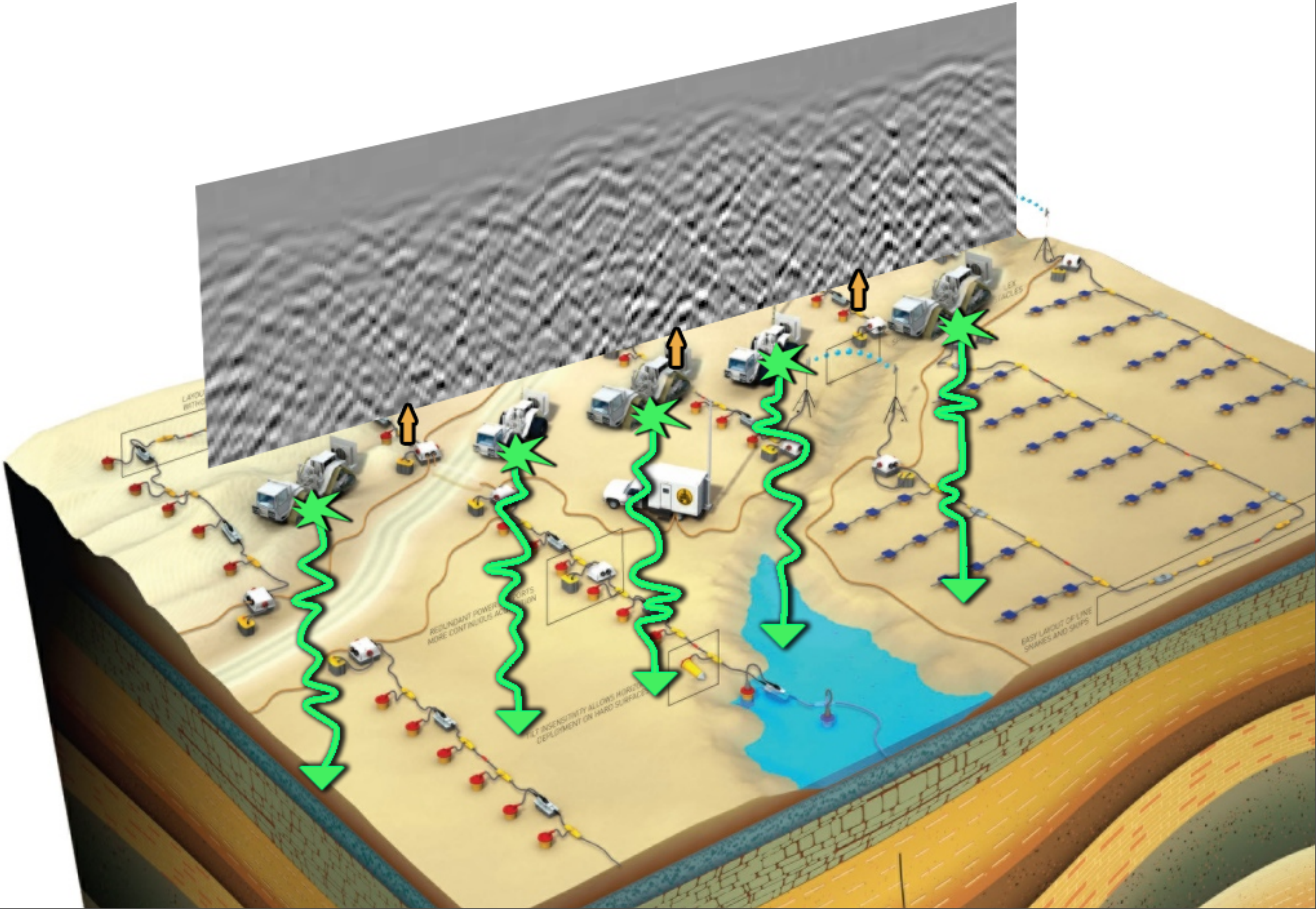
Observation & extensions

- Findings from 1D jittered sampling extend to higher dimensions
  - randomized is better than regular subsampling
  - Cartesian versus hexagonal sampling are equivalent for optimal jittered sampling
  - Furthest point and Poisson sampling lead to similar results

- Gap-size control
  - jittered sampling gives explicit control max distance between adjacent samples
  - farthest point and Poisson disk also have bounds but not explicit

- Future extensions
  - variable density sampling
  - ungridded
  - exploring symmetry (e.g. reciprocity)

- Open math. questions
  - extension of CS results to frames
  - practical design principles

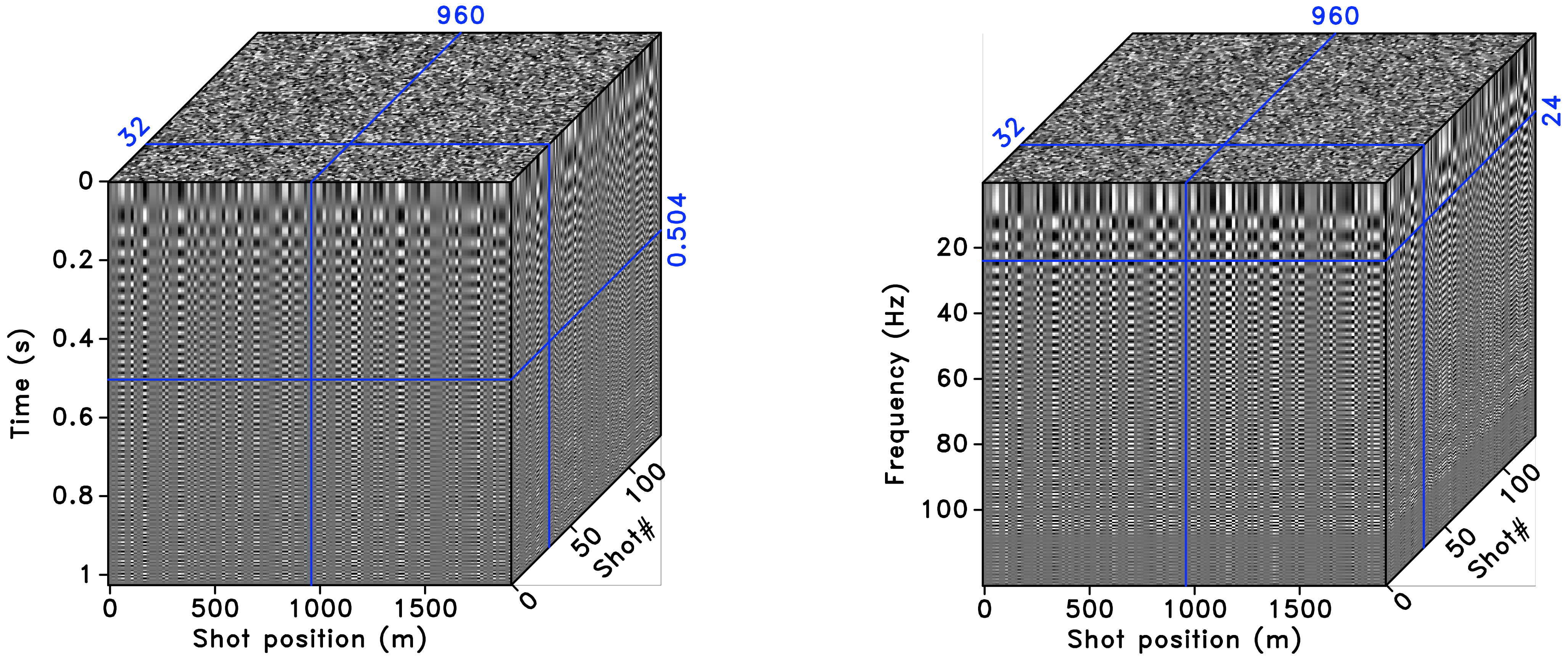

REDUNDANT POWER
MORE CONTINUOUS AC
TILT INSENSITIVITY ALLOWS HORIZ
DEPLOYMENT ON HARD SURFACE
EASY LAYOUT OF LINE
SNAKES AND SKIPS

Relation to existing work

- **Simultaneous & continuous acquisition:**
  - *A new look at marine simultaneous sources* by C. Beasley, ‘08
  - *Simultaneous Sourcing without Compromise* by R. Neelamani & C.E. Krohn, ’08.
  - *Changing the mindset in seismic data acquisition* by A. Berkout, ’08
  - *Independent simultaneous sweeping - A method to increase the productivity of land seismic crews* by D. Howe, M. Foster, T. Allen, B. Taylor, and I. Jack, ‘08

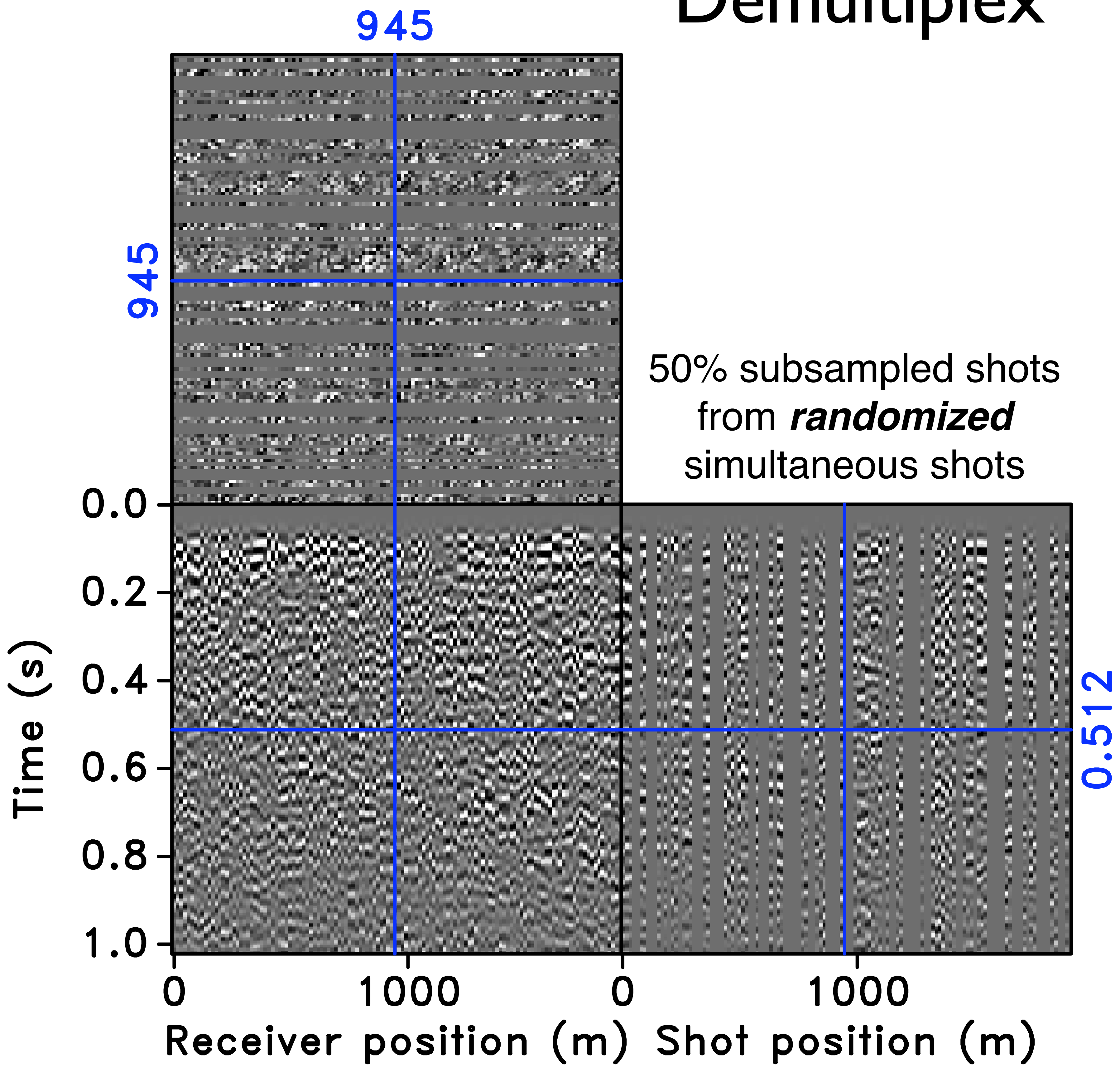
Recovery from simultaneous data

$$\mathbf{RM} = \overbrace{\left[\mathbf{R}^{\Sigma}\otimes\mathbf{I}\otimes\mathbf{I}\right]}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_{\mathrm{s}}^{*}\mathrm{diag}\left(e^{\hat{i}\boldsymbol{\theta}}\right)\mathbf{F}_{\mathrm{s}}\otimes\mathbf{I}\otimes\mathbf{I}\right)}^{\text{random phase encoder}}$$



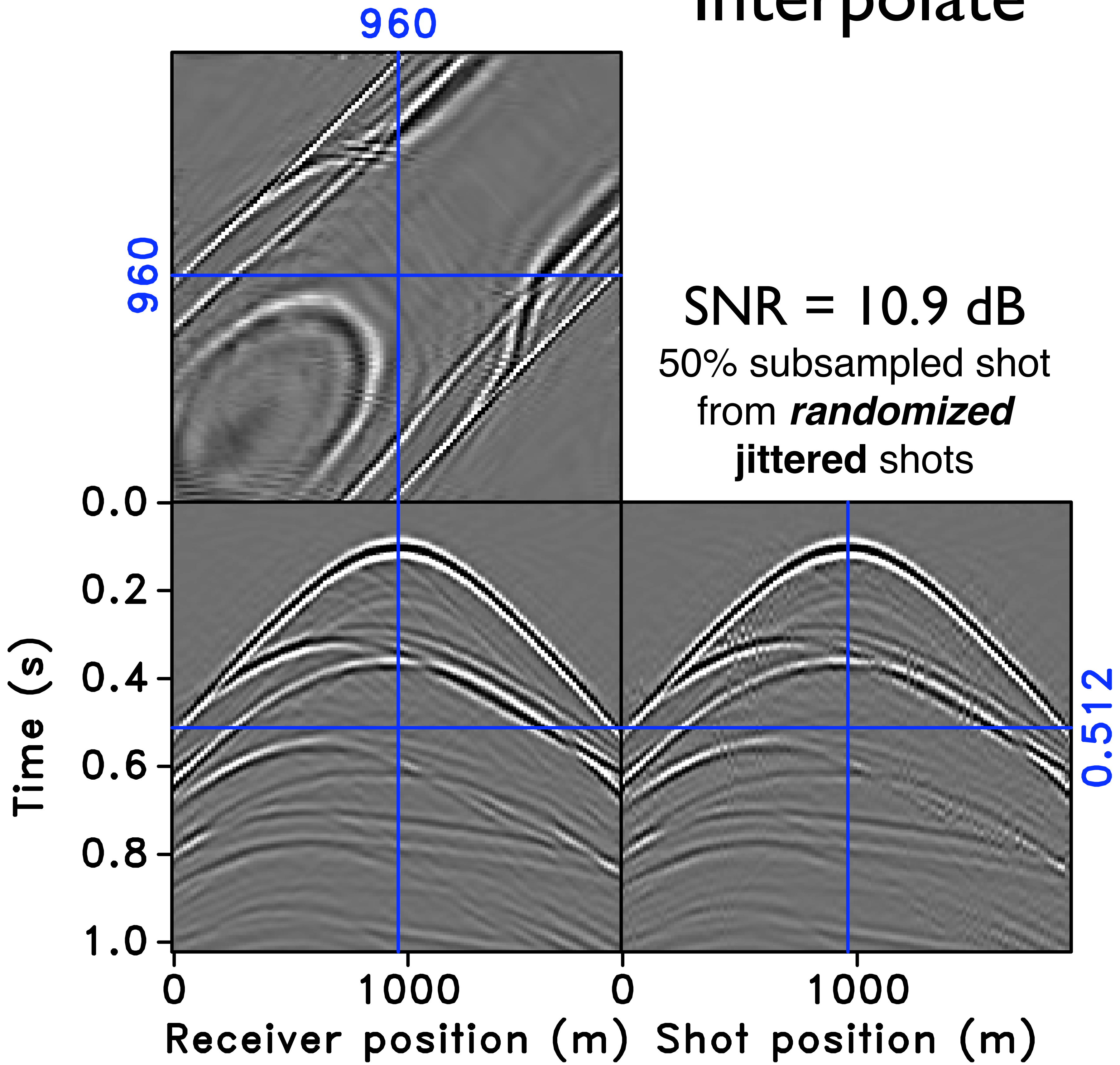
- Linearly ramping seismic sweep, 5 to 110 Hz
- Simultaneous source at all positions, each randomly phase encoded

Demultiplex



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

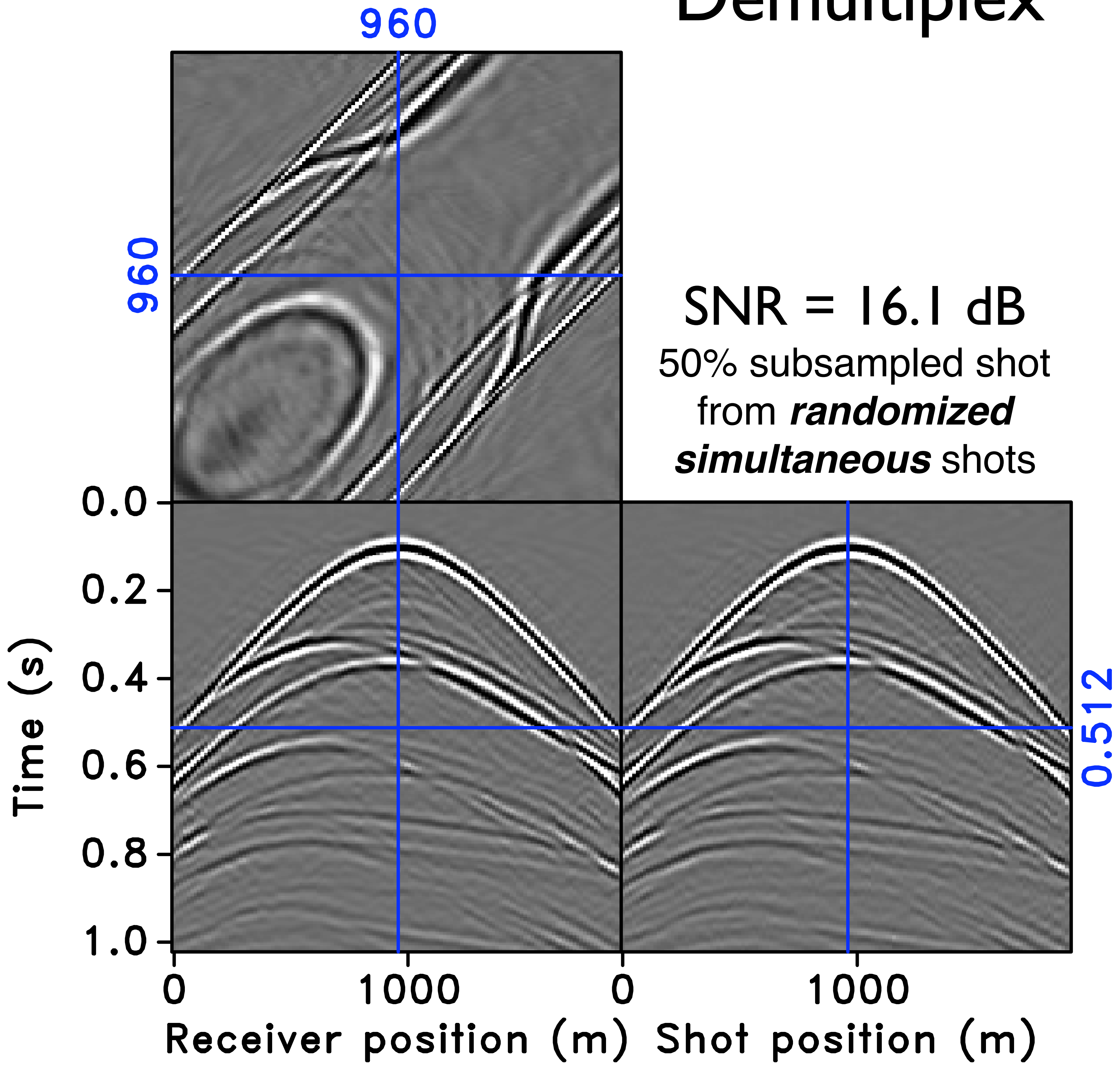
Interpolate



SNR = 10.9 dB

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

Demultiplex



SNR = 16.1 dB

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

Key elements

☑ *sparsifying transform*

- typically *strictly* **localized** in the Fourier space
- rapid decay physical space to handle the complexity of seismic data
- mutual incoherence

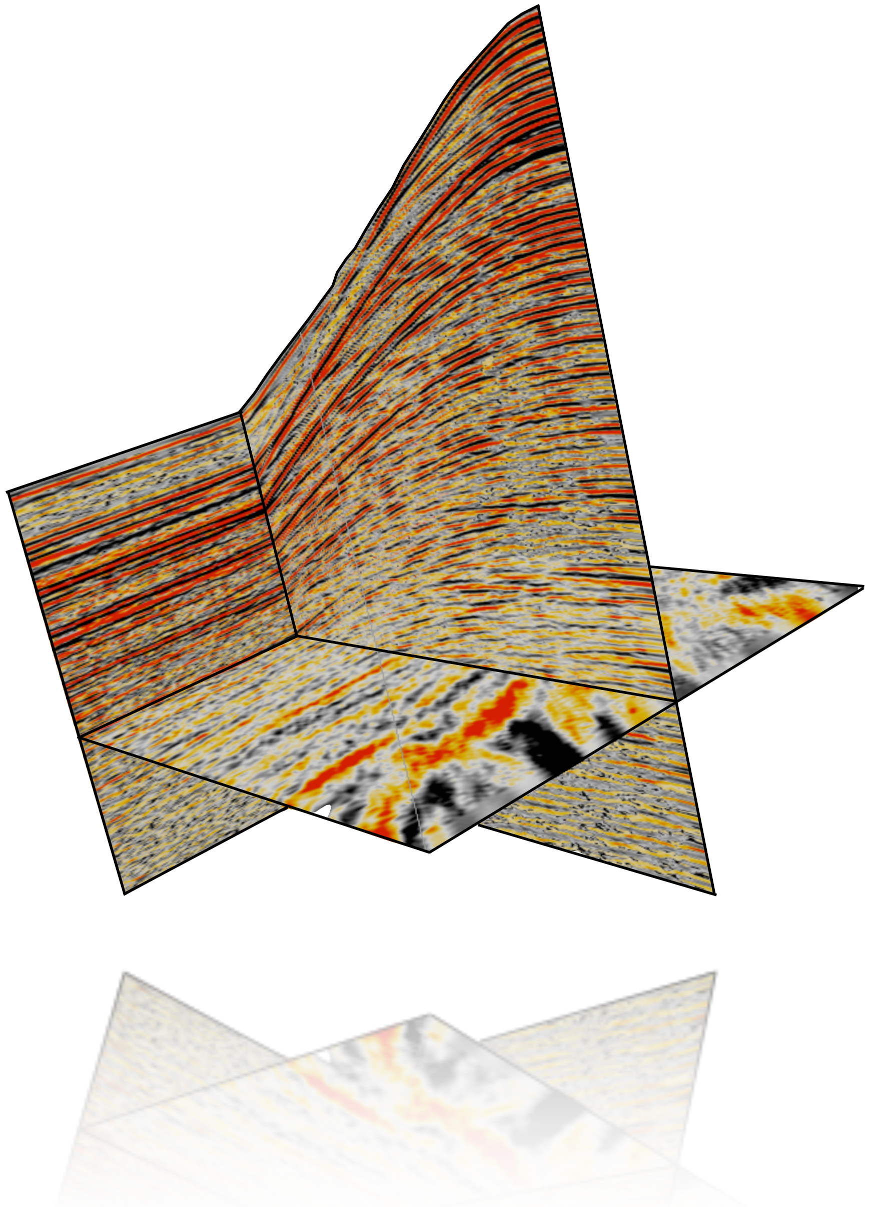
☑ *advantageous* ***randomized*** *coarse sampling*

- generates incoherent random undersampling “noise” in sparsifying domain
- ***spatial sampling*** that does not create large gaps
  - because of the limited spatiotemporal extent of transform elements used for the reconstruction
- ***randomized*** subsampling of ***simultaneous-source*** experiments
  - does not create large interferences
  - leads to compression of linear systems
    - reduction # right-hand-sides
    - spectral representation operators

☐ *sparsity-promoting solver*

- requires few matrix-vector multiplications

Recent developments: recovery of surface-free data from simultaneous acquisition



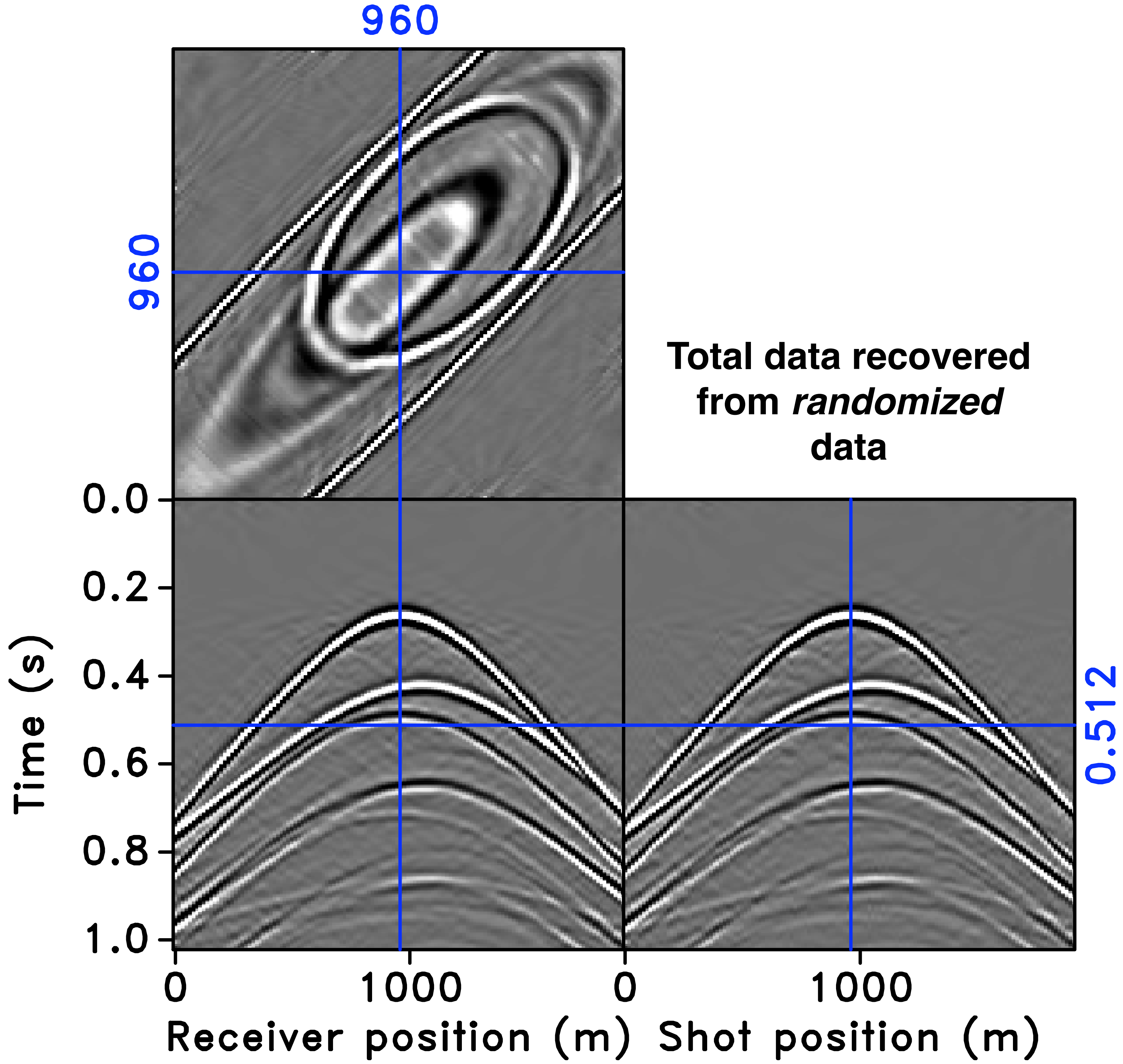
Vienna July 24, 2009

Estimation of primaries from simultaneous data

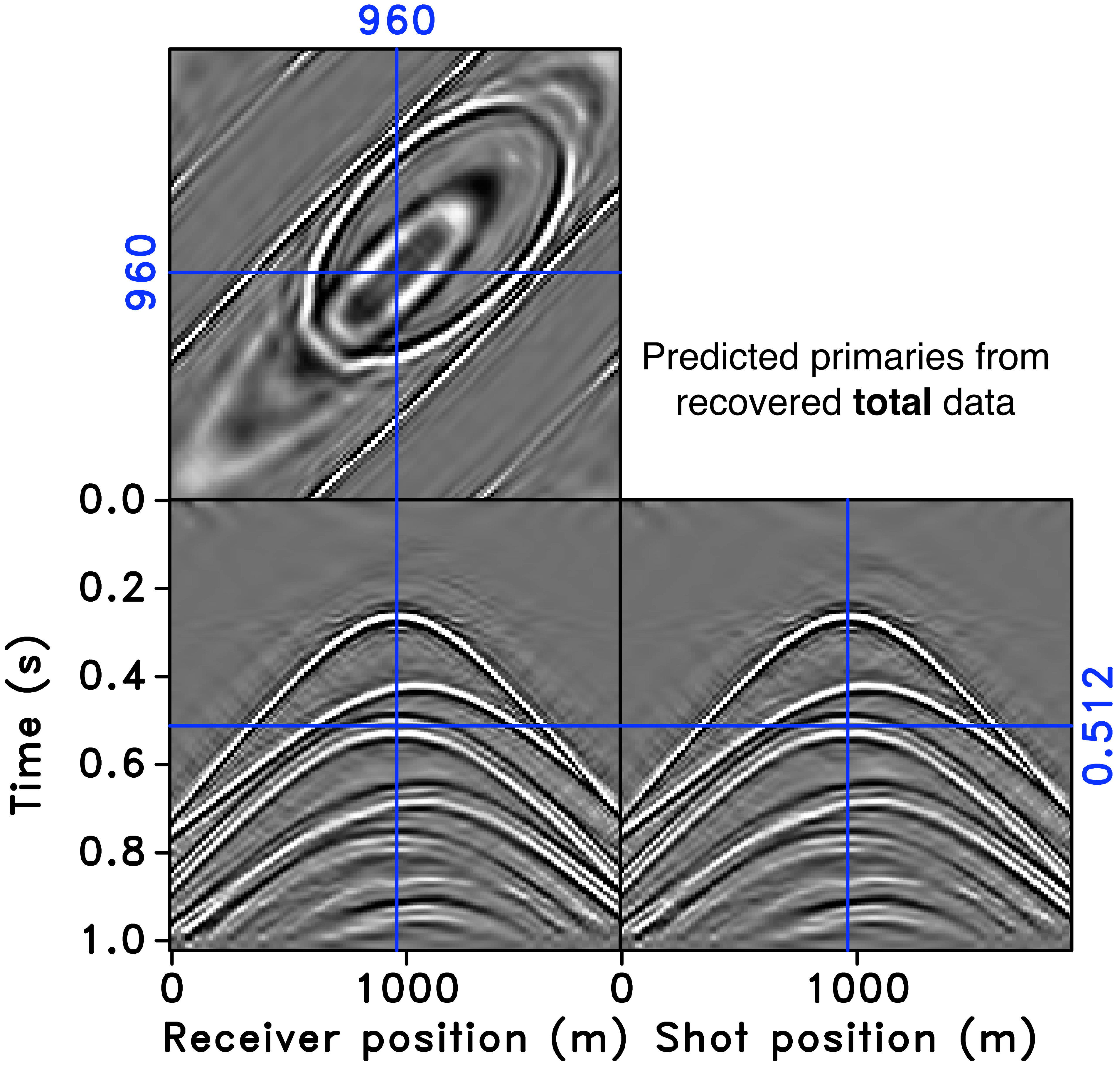
- Include multiple prediction operator
  - operator that generates all surface-related multiples
  - invert the operator as part of the sparsity-promoting recovery

$$\mathbf{RM} = \overbrace{\left[\mathbf{R}^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{I}\right]}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_{\mathrm{s}}^{*} \mathrm{diag}\left(e^{\hat{i}\boldsymbol{\theta}}\right) \mathbf{F}_{\mathrm{s}} \otimes \mathbf{I} \otimes \mathbf{I}\right)}^{\text{random phase encoder}} \overbrace{\mathbf{G}}^{\text{multiple prediction}}$$

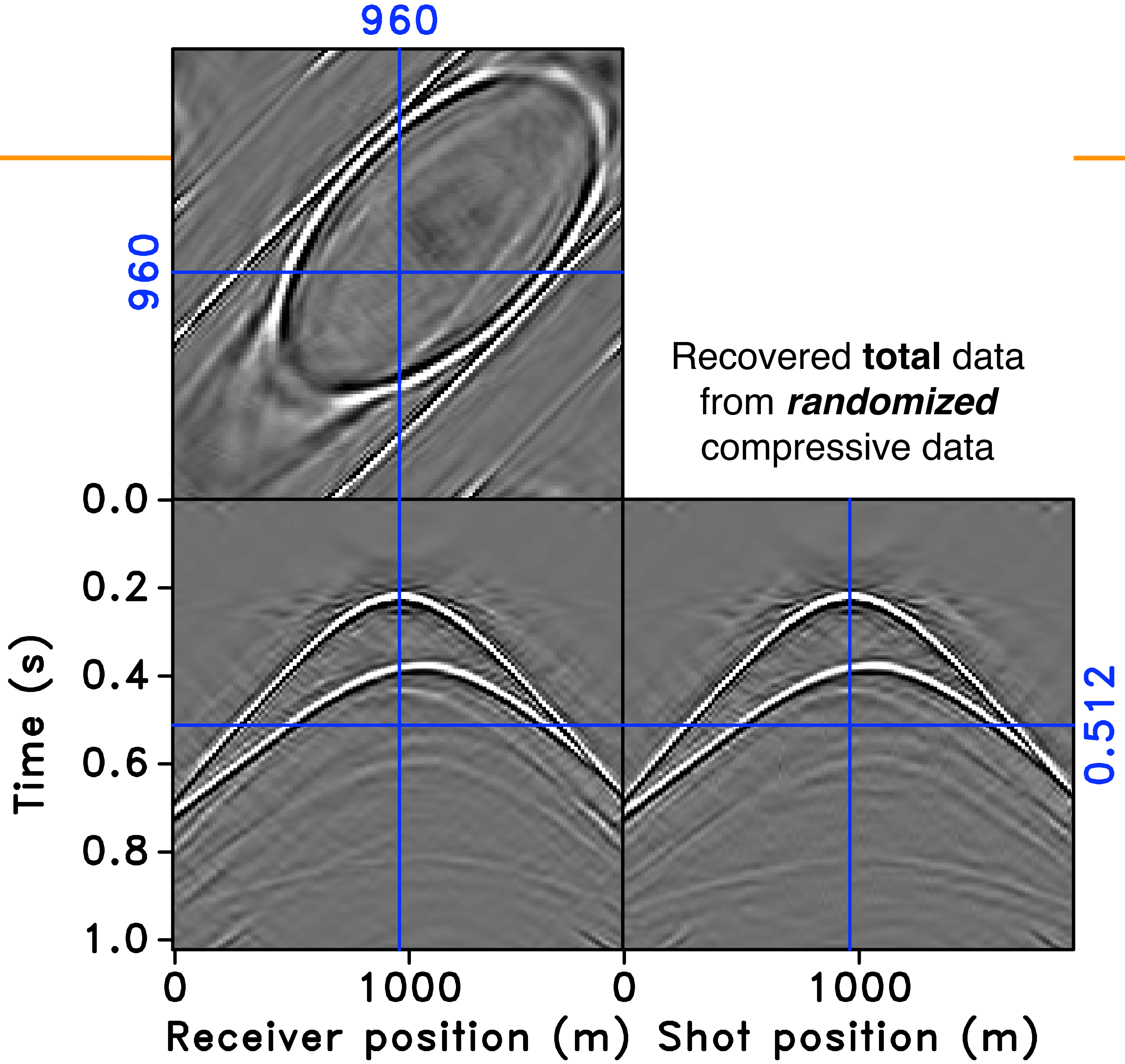
- **Primary prediction through wavefield inversion:**
  - *Elimination of free-surface related multiples without need of the source wavelet* by L. Amundsen, ‘01
  - *Primary estimation by sparse inversion and its application to near offset reconstruction* by G. van Groenenstijn and D. Verschuur, ’09



Total data recovered from *randomized* data

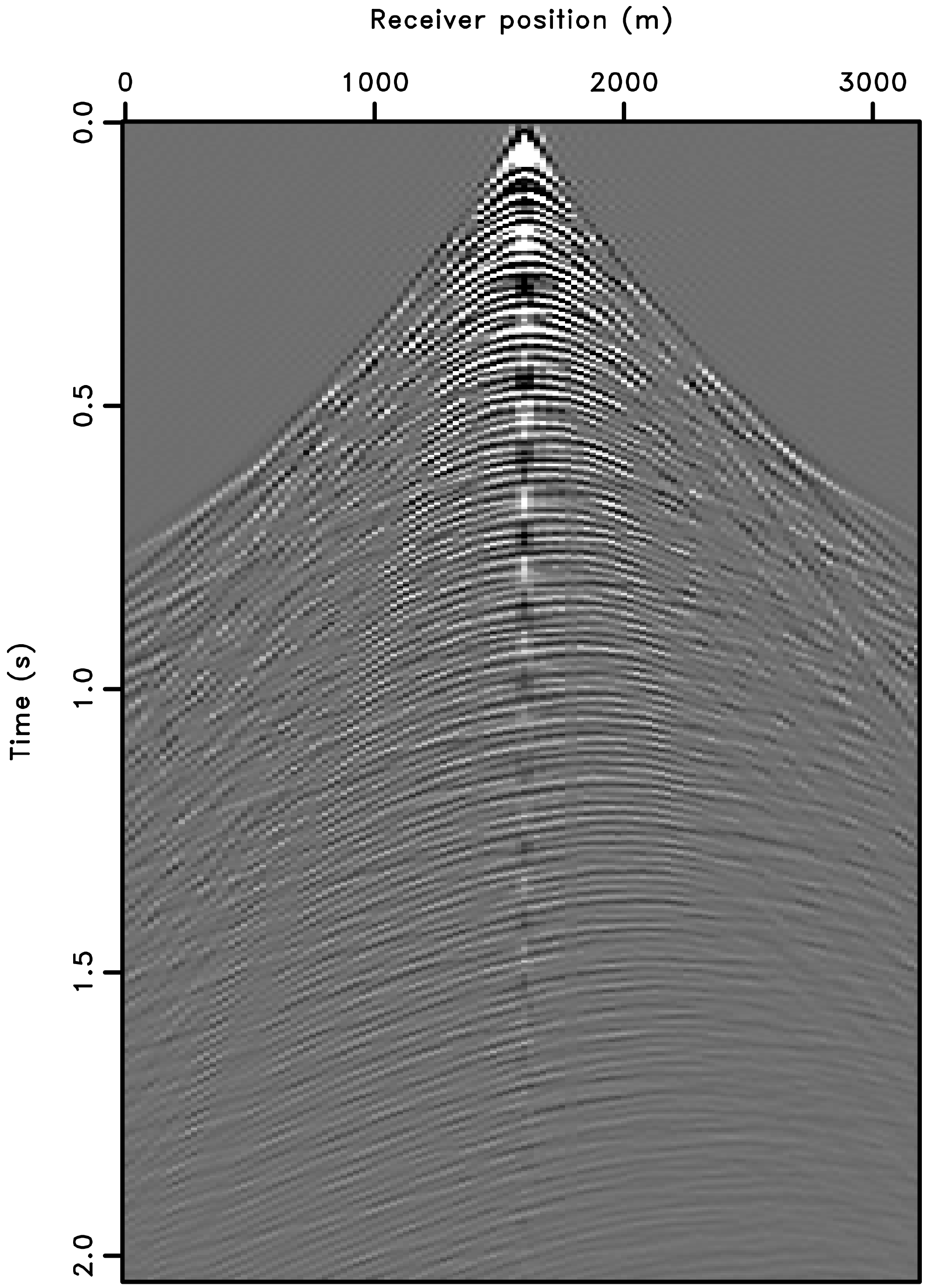


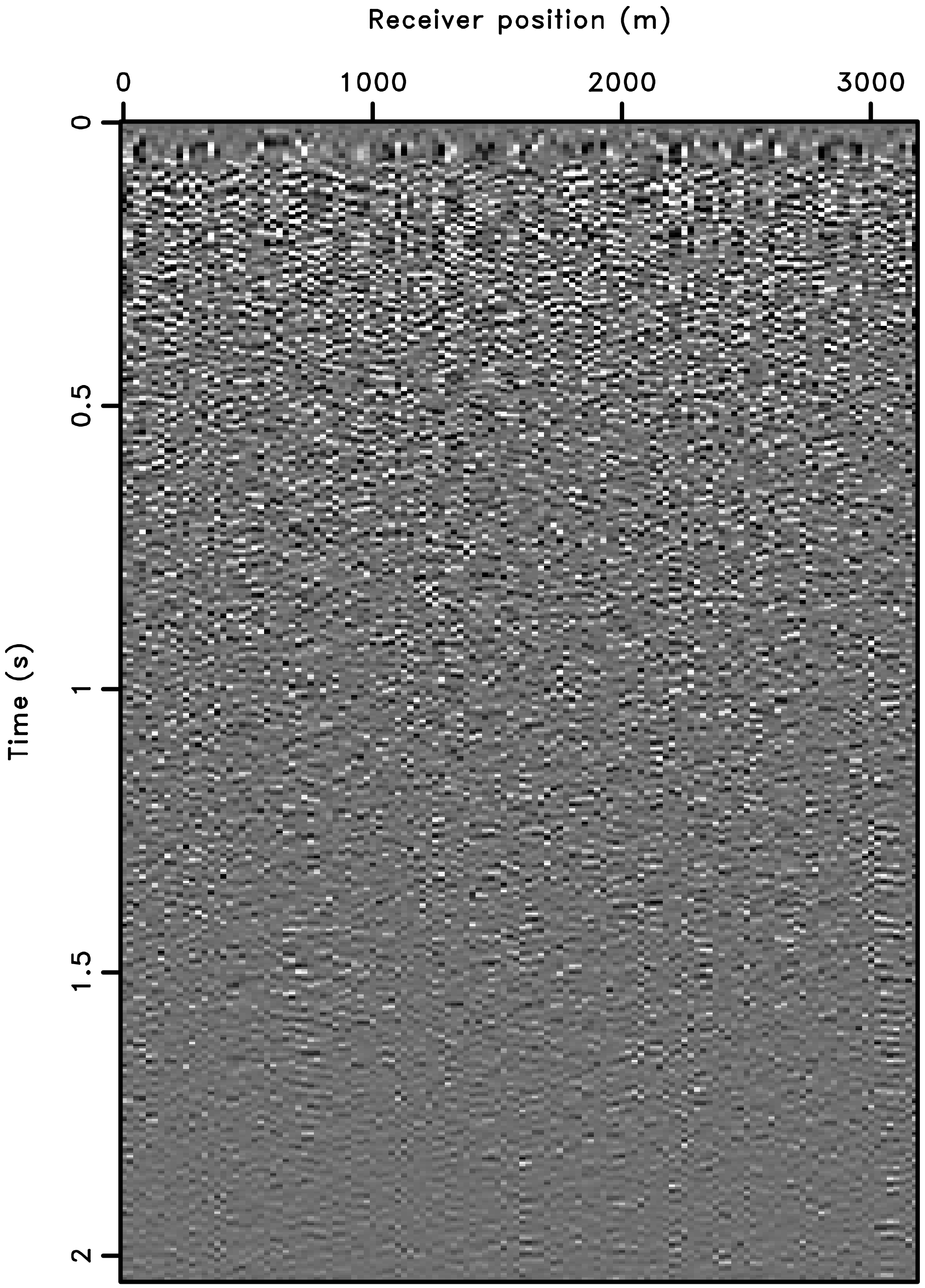
Predicted primaries from
recovered **total** data



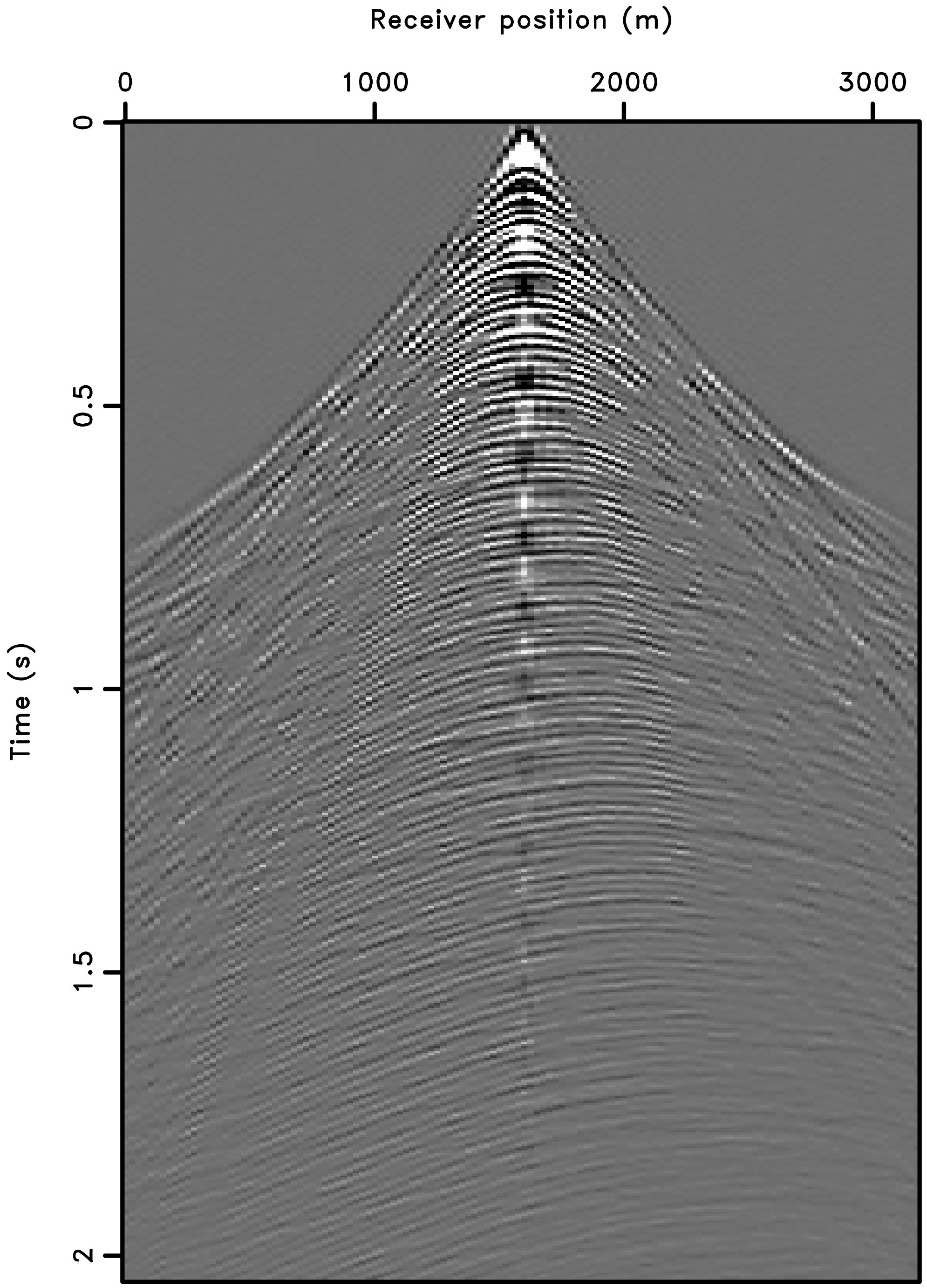
Recovered **total** data
from ***randomized***
compressive data

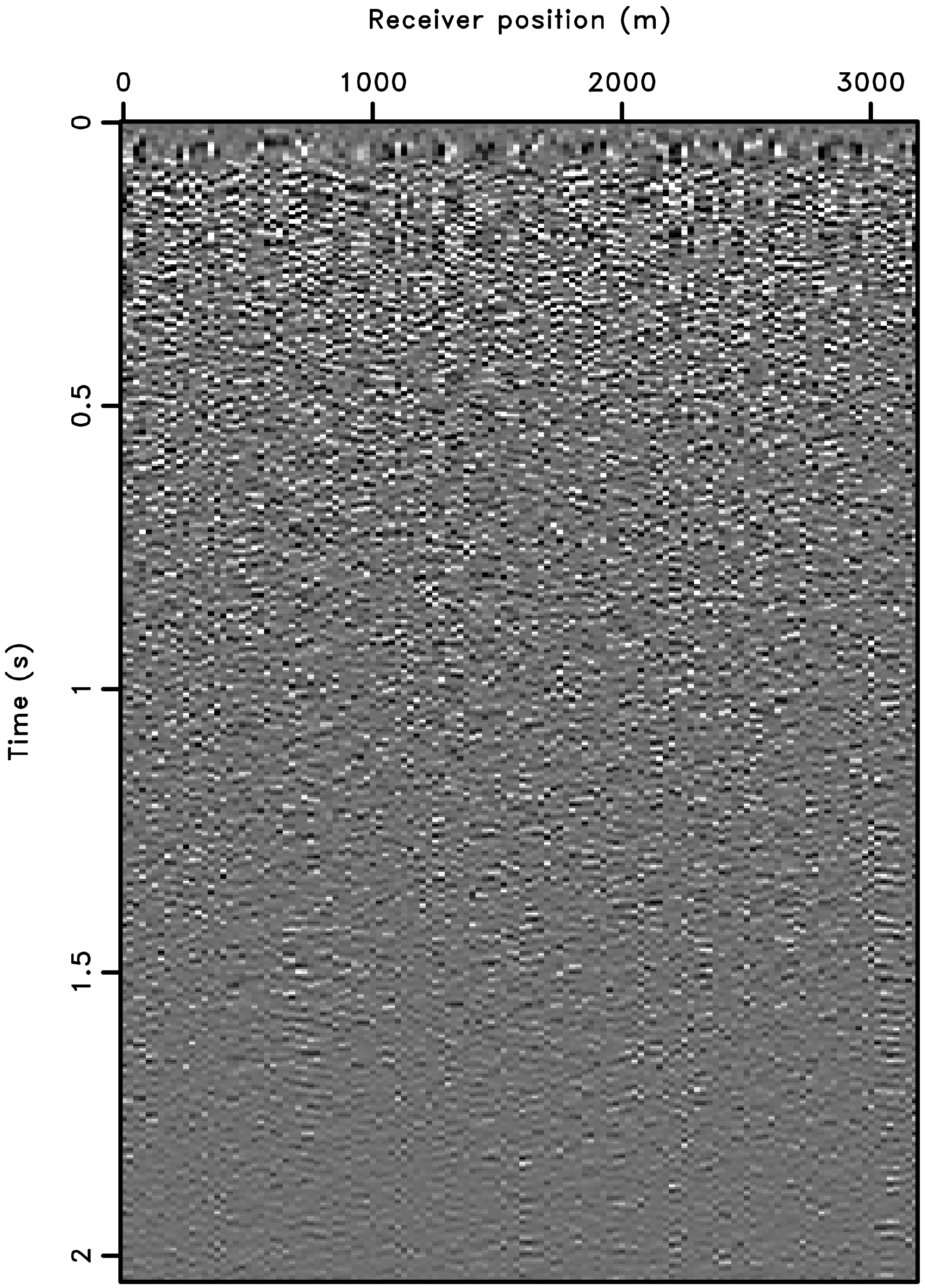
Real Marine data





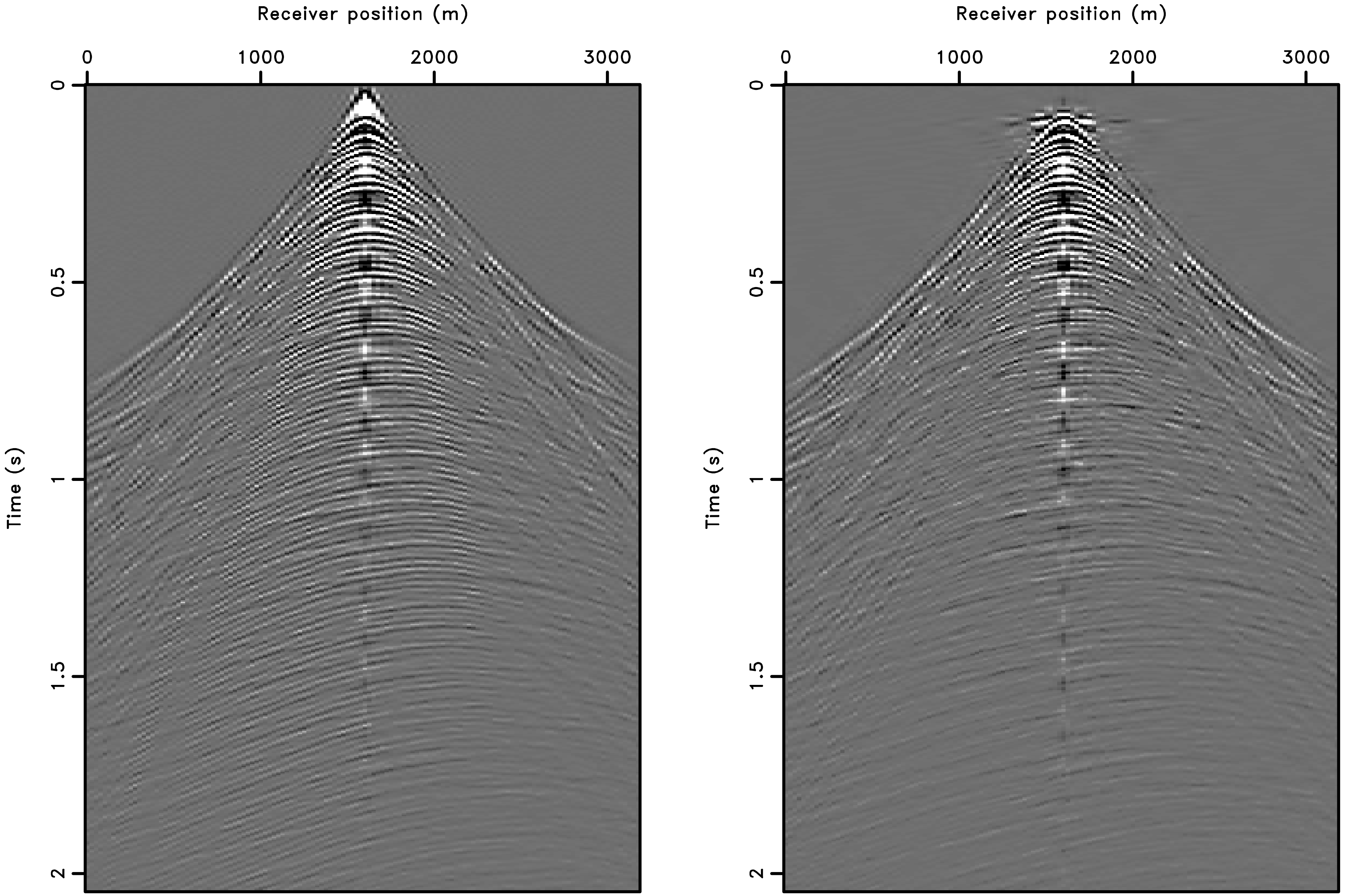
Recovery Real Marine data





85x

Primary prediction Real Marine data



Conclusions

- **Randomization** is essential for recovery from incomplete data

- Good **randomized** sampling
  - with blue-noise characteristics give good curvelet recovery
  - with simultaneous sources gives excellent curvelet recovery

- **Randomization** leads to
  - “acquisition” of *smaller* data volumes that carry the ***same*** *information* or
  - to ***improved*** *inferences* from data using the *same* resources

- **Bottom line: acquisition costs** are **no** longer determined by the ***size*** of the ***discretization*** but by ***transform-domain sparsity*** of the sampled wavefield **...**

Acknowledgments

- E. van den Berg and M. P. Friedlander for *SPGL1* (www.cs.ubc.ca/labs/scl/spgl1) & *Sparco* (www.cs.ubc.ca/labs/scl/sparco)
- Sergey Fomel and Yang Liu for Madagascar (rsf.sf.net)
- E. Candes and the Curvelab team

This work was carried out as part of the Collaborative Research & Development (CRD) grant DNOISE (334810-05) funded by the Natural Science and Engineering Research Council (NSERC) and matching contributions from BG, BP, Chevron, ExxonMobil and Shell. FJH would also like to thank the Technische University for their hospitality.

slim.eos.ubc.ca

and... Thank you!