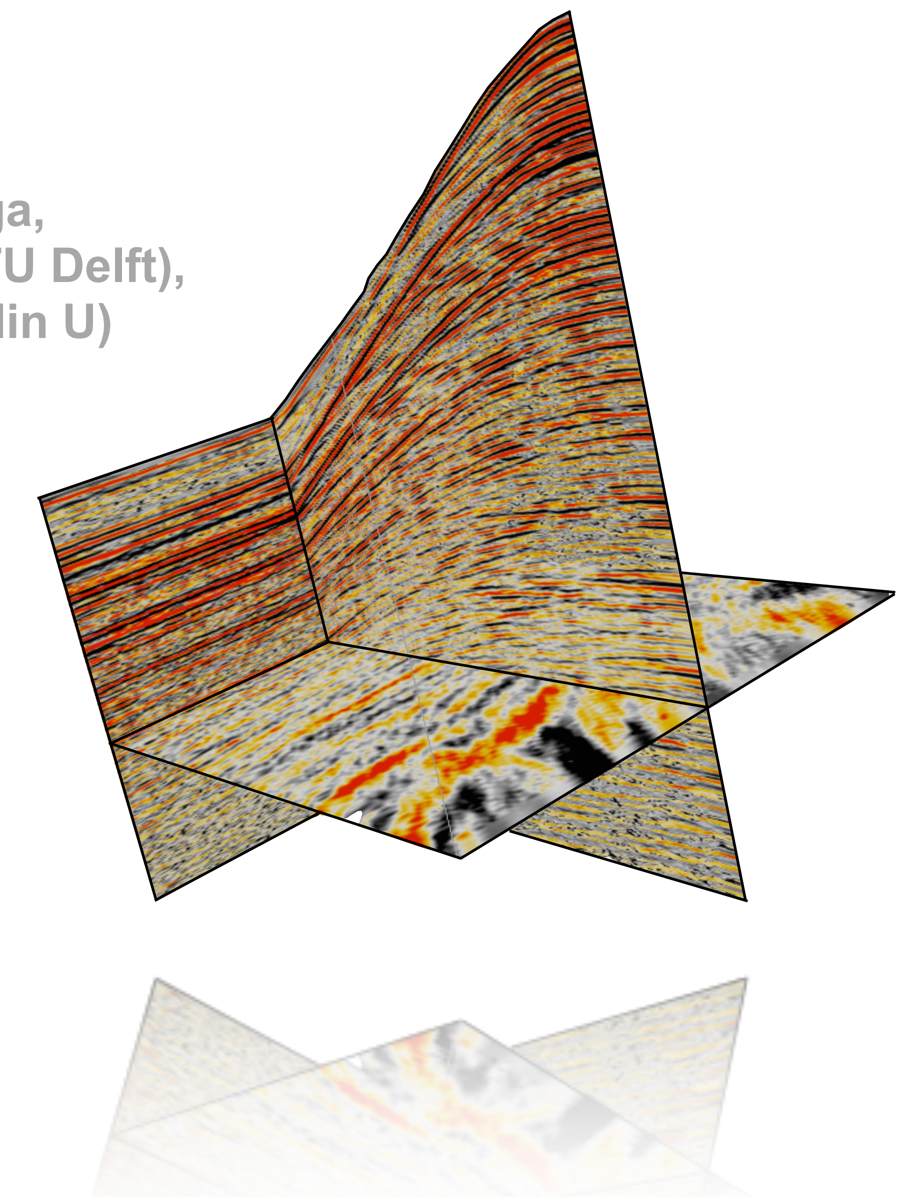
# Seismic noise: the good, the bad, & the ugly
# A Curvelet Approach

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

**Joint work with Yarham Carson, Yogi Erlangga, Gilles Hennenfent, Tim Lin, Eric Verschuur (TU Delft), Dave Wilkinson (Chevron), and Deli Wang (Jilin U)**

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



SEG
Las Vegas, November 9-14

# Strategies

- Noise *shaping* to facilitate separation by *denoising*
  - turn *coherent* sub-Nyquist aliases into *incoherent* noise by jittered subsampling
  - turn *coherent* interferences of simultaneous acquisition into *incoherent* noise
  - (sub)sample according to the principles of *Compressive Sensing*
- *Adaptive* curvelet-domain *matching* compounded with curvelet-domain Bayes separation
  - primary-multiple separation & surface-wave removal
- *Adaptive* noise shaping by diagonalization of the model-space covariance operator
  - exploit *invariance* of curvelets under certain operators
- *Denoising* via curvelet-domain sparsity promotion

# Today's agenda

- The good: sparsity-promoting recovery from compressive sampling
  - incomplete data
  - blended data

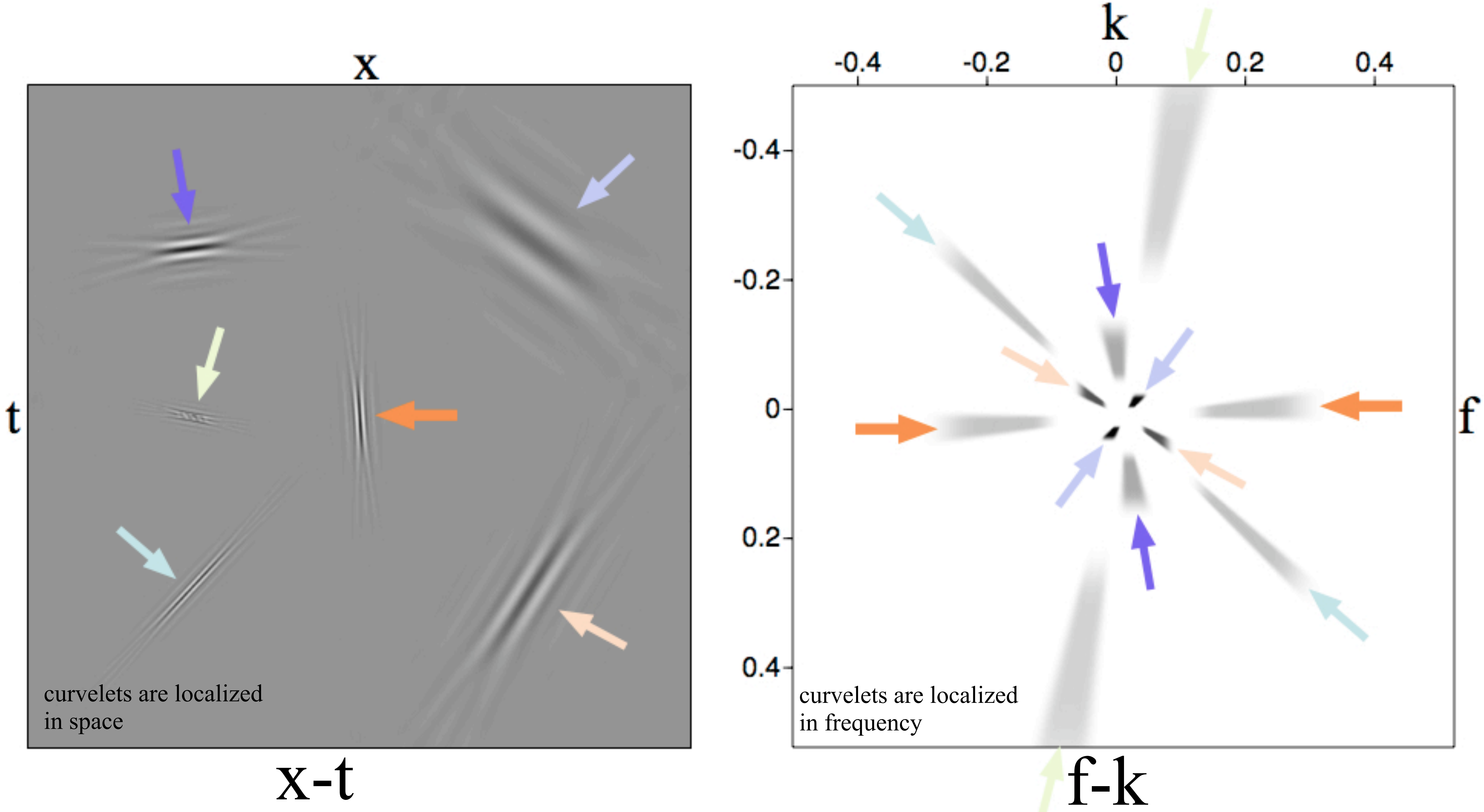
- The bad: separation of coherent signal components by sparsity promotion
  - multiple removal
  - surface-wave removal

- The ugly: migration amplitude recovery from noisy data

# Key properties: curvelets

- **Detection of wavefronts** without requiring prior information on the dips or on the velocity model
  - curvelets are localized functions that oscillate in one direction and are smooth in the other
  - anistropic scaling => needle like for small scales
- **Approximate invariance of curvelets** under the action of wave propagation
  - curvelets are transformed to ‘curvelets’ under the action of the migration-demigration (normal) operator
  - curvelets are transformed to ‘curvelets’ under the action of wave propagation
- **Compression** of curvelets
  - seismic data
  - seismic images

# Some 2-D curvelet examples



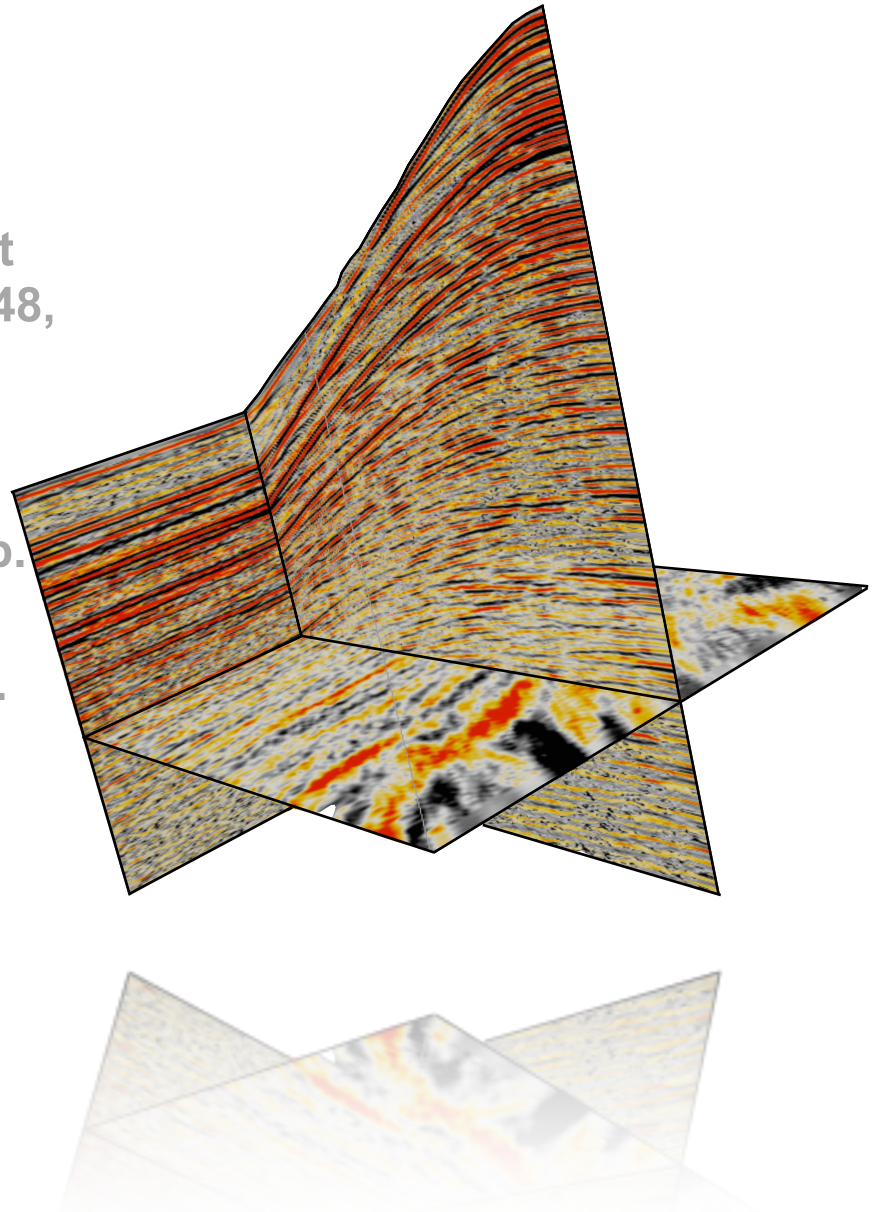
# The good: recovery from incomplete (blended) data

## Signal is sparse in the curvelet domain while noise is not ....

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

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

Felix J. Herrmann, Yogi Erlangga, and Tim T. Y. Lin. Seismic Laboratory for Imaging and Modeling. The university of British Columbia Technical Report. TR-2008-3. Compressive simultaneous full-waveform simulation.

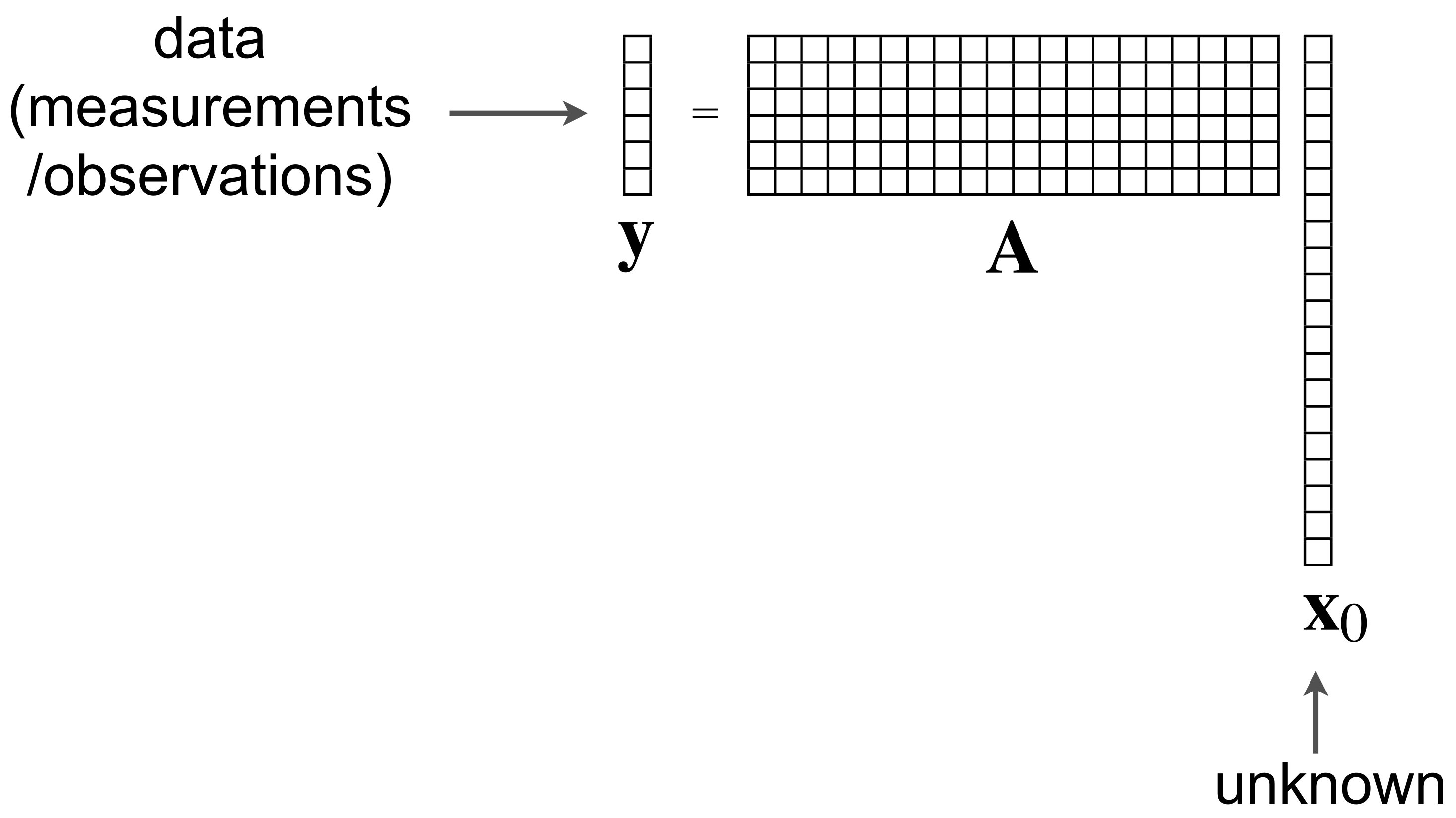


# Key ingredients

- Recognize recovery from *missing* traces and *blended* data as instances of *compressive sampling*
- Design subsampling schemes that result in *incoherent* interferences
  - break the *periodicity* in *sourcer-receiver* placement
  - *encode* the sources part of *simultaneous* acquisition
- Select the appropriate sparsifying transform
- Solve a sparsity-promoting program => *denoise*

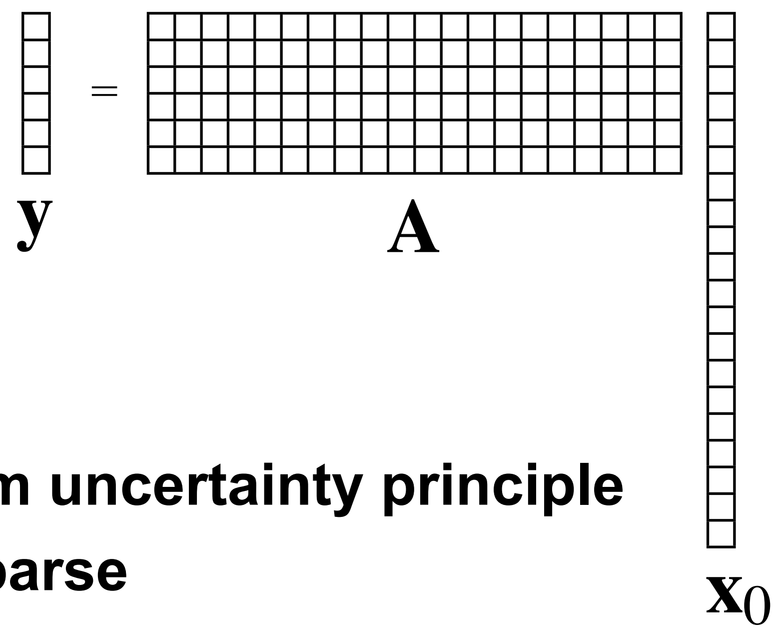
# Problem statement

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**
  - $\mathbf{x}_0$ is **sufficiently sparse**

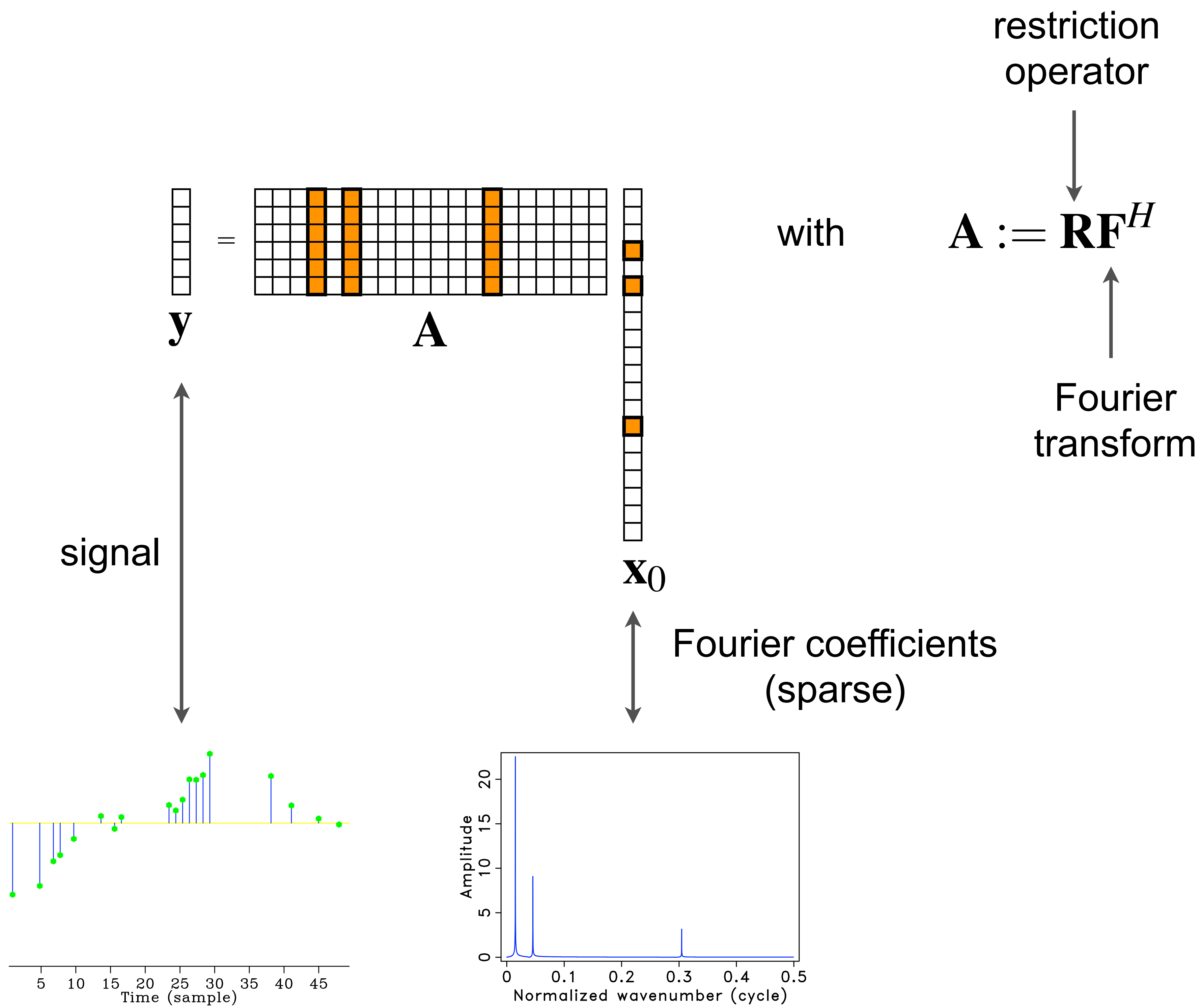
- procedure

$$\underbrace{\min_{\mathbf{x}} \|\mathbf{x}\|_1}_{\text{sparsity}} \quad \text{s.t.} \quad \underbrace{\mathbf{A}\mathbf{x}=\mathbf{y}}_{\text{perfect reconstruction}}$$

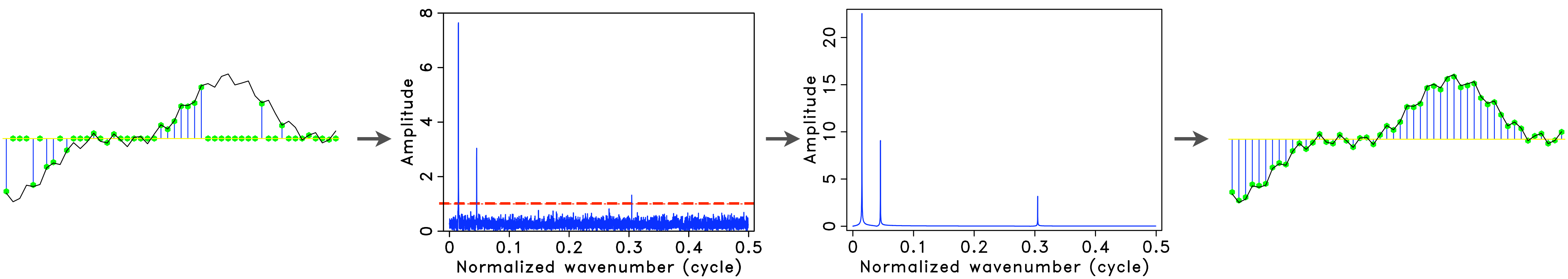
- performance
  - *S*-sparse vectors recovered from roughly on the order of *S* measurements (to within constant and *log* factors)

[Candès et al.'06]
[Donoho'06]

# Simple example



# NAIVE sparsity-promoting recovery



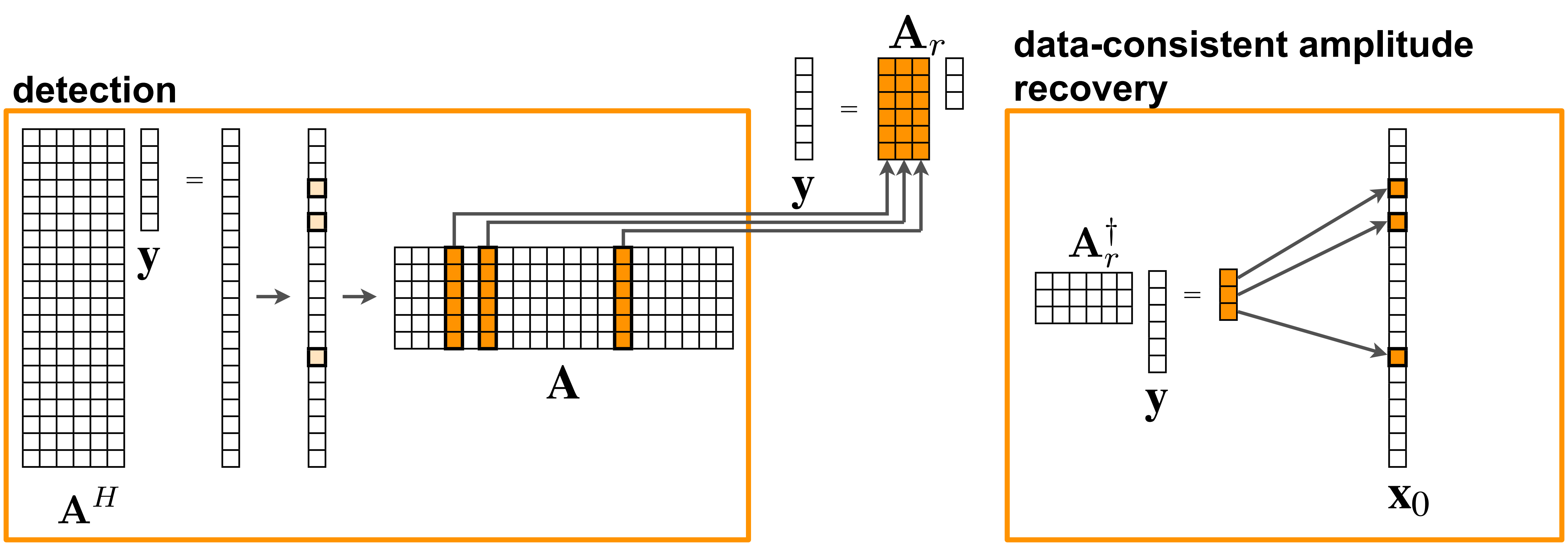

inverse Fourier transform

detection + data-consistent amplitude recovery

Fourier transform

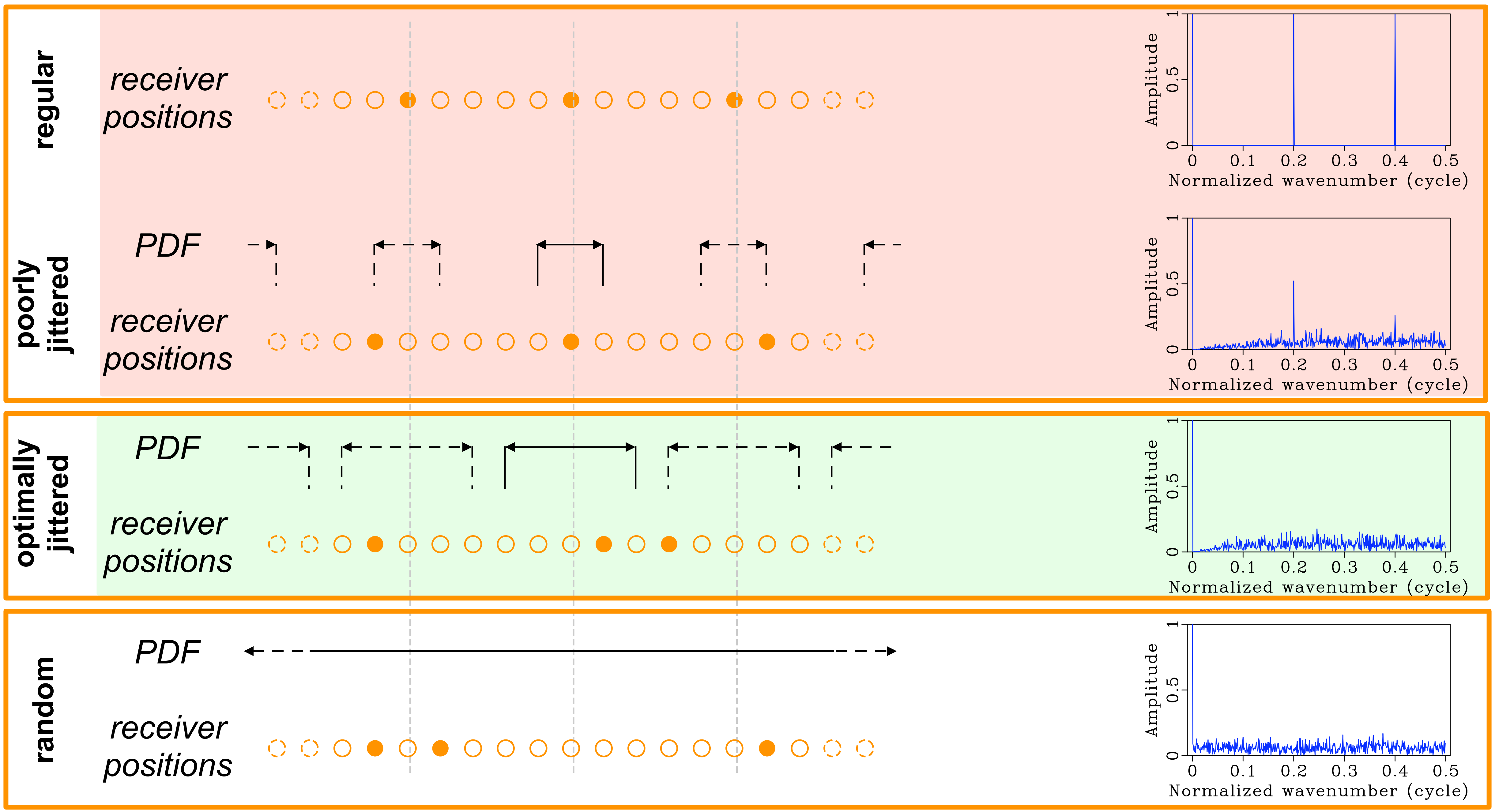
**detection**

**data-consistent amplitude recovery**



# Discrete random jittered undersampling

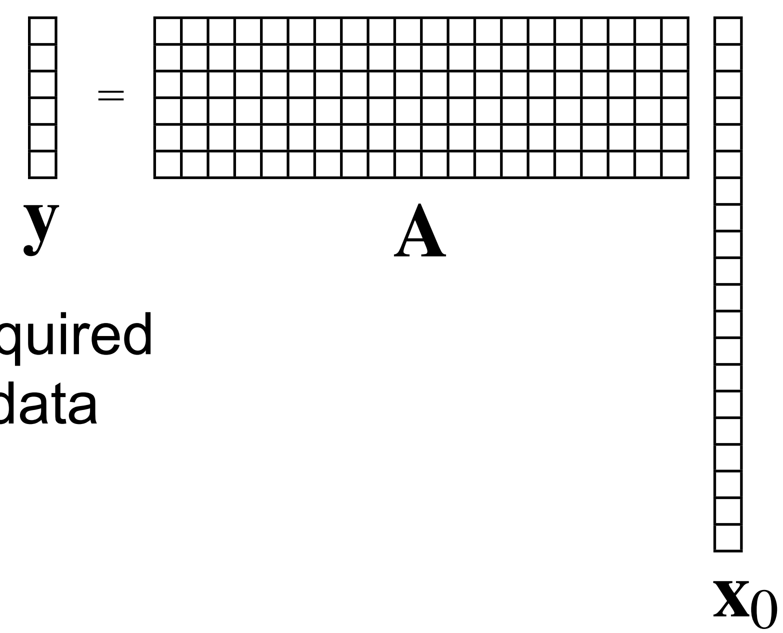
Type | Sampling scheme | *Typical* spatial convolution kernel




[Hennenfent and Herrmann‘08]

# Sparsity-promoting wavefield reconstruction

restriction operator




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

acquired data

complete wavefield (transform domain)

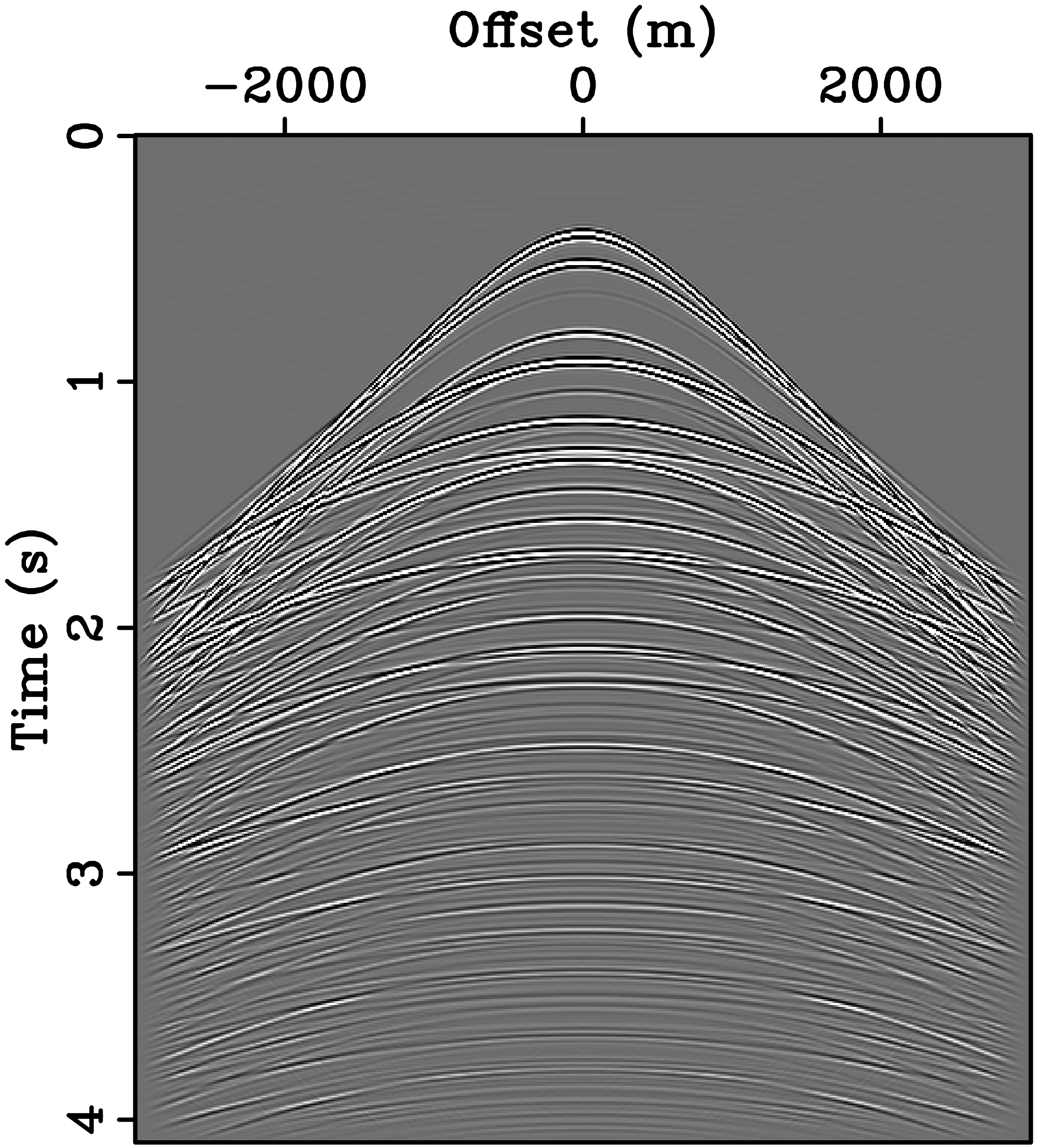
sparsifying transform for seismic data

Interpolated data given by $\tilde{\mathbf{f}} = \mathbf{S}^H \tilde{\mathbf{x}}$ with

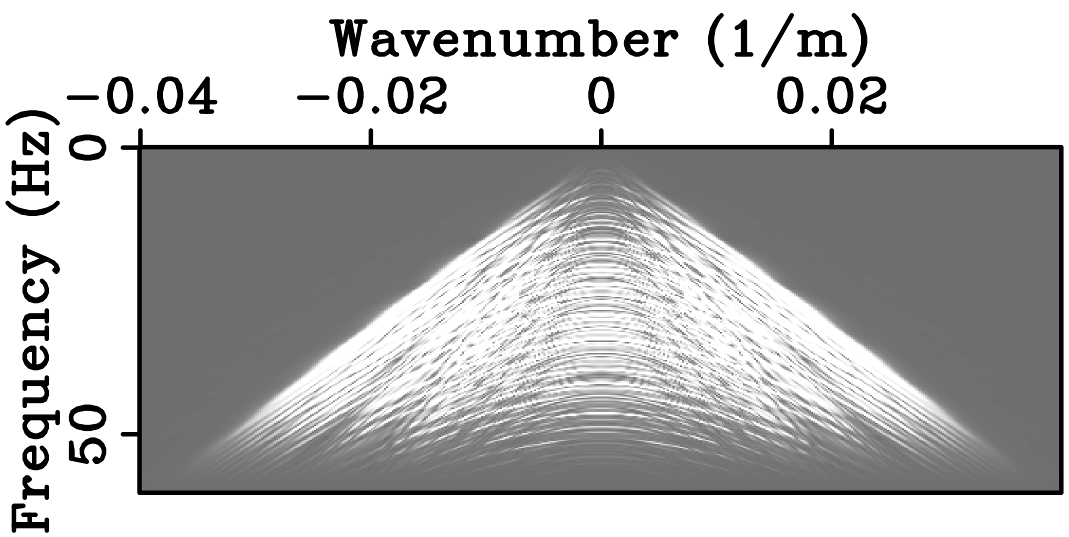
$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} ||\mathbf{x}||_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

[Sacchi et al.‘98]
[Xu et al.‘05]
[Zwartjes and Sacchi‘07]
[Herrmann and Hennenfent‘08]

# Model

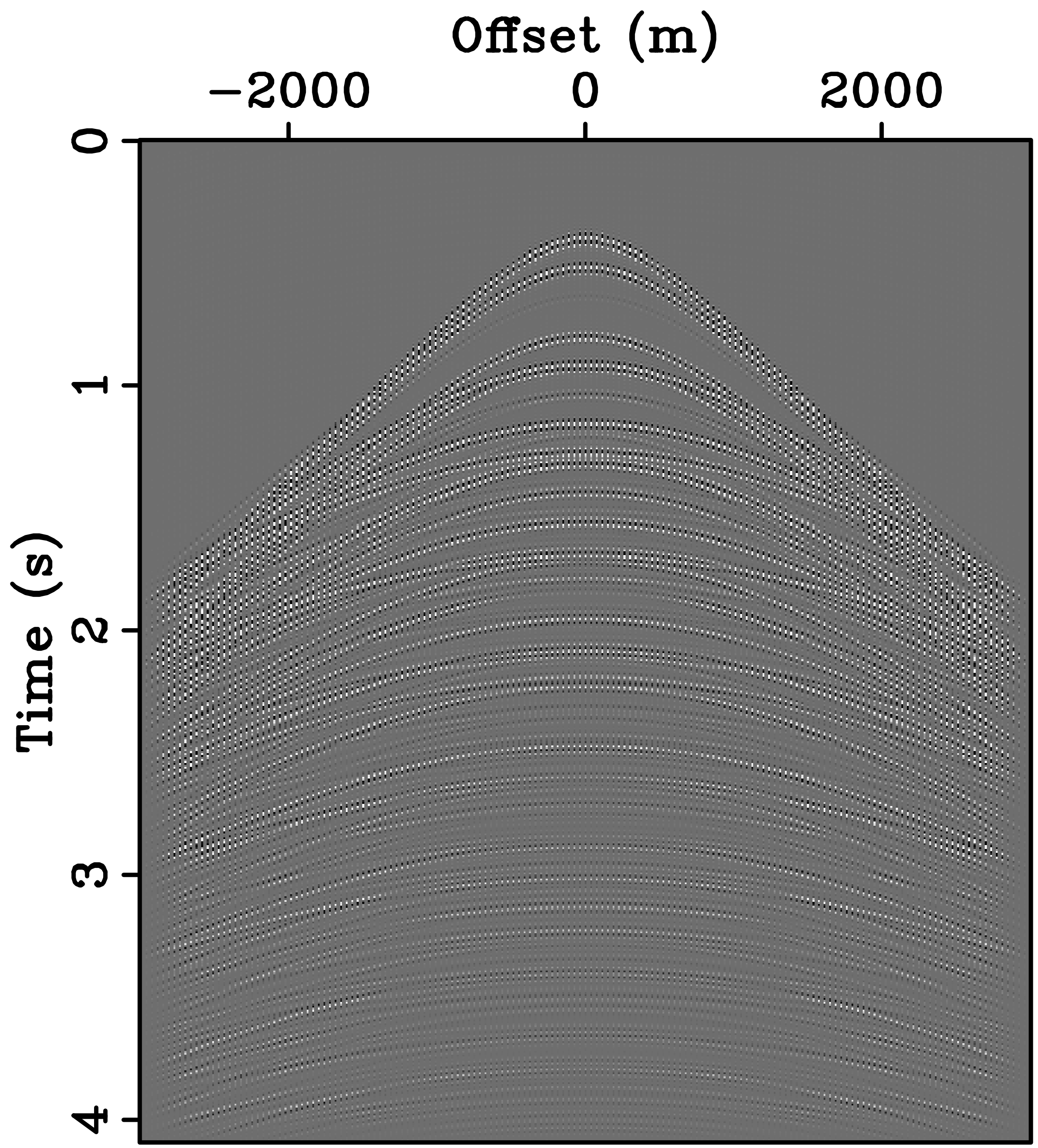


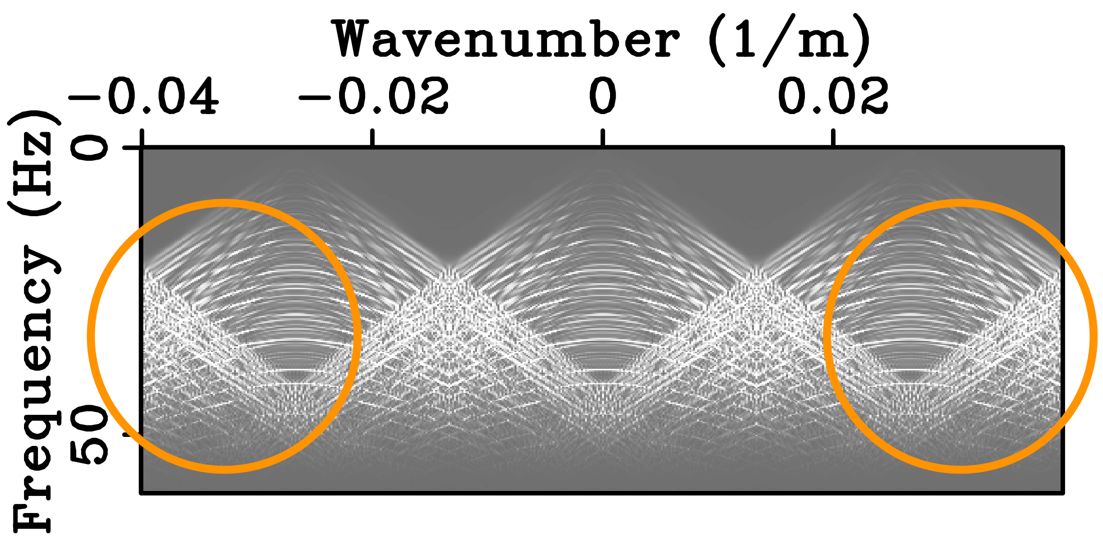
Offset (m)
−2000
0
2000
Time (s)
0
1
2
3
4

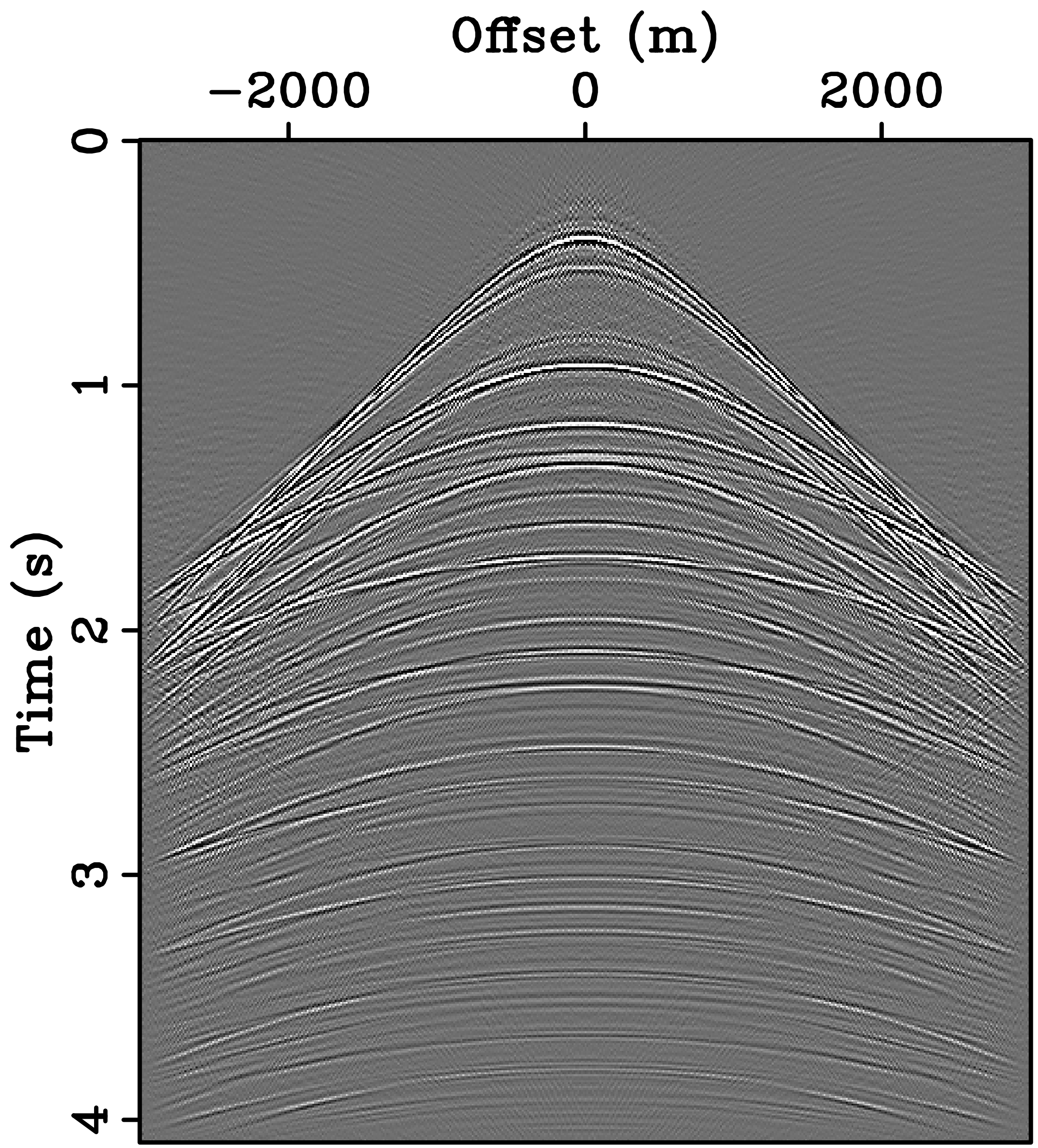
Wavenumber (1/m)
−0.04
−0.02
0
0.02
Frequency (Hz)
0
50

# Regular 3-fold undersampling

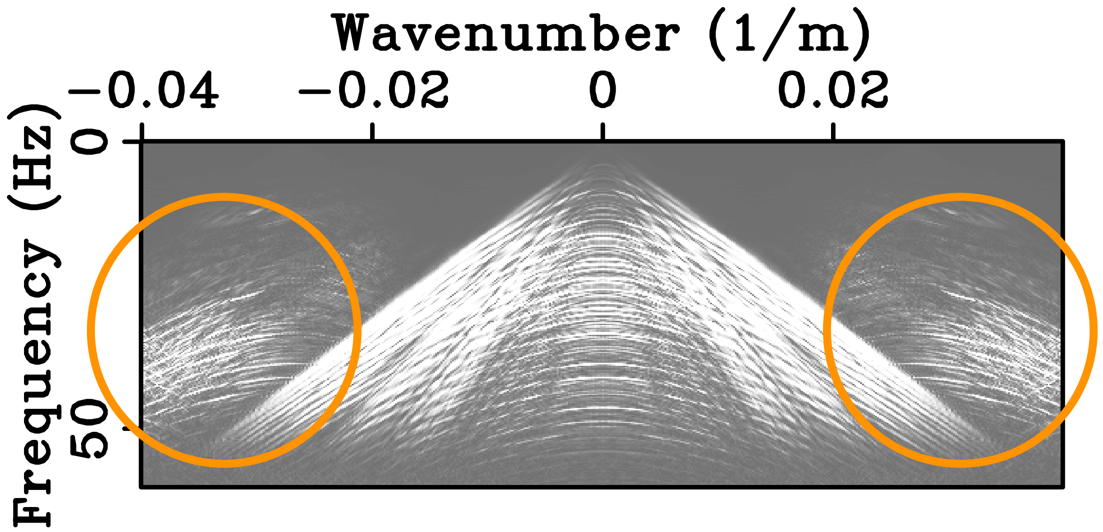


# CRSI from regular 3-fold undersampling

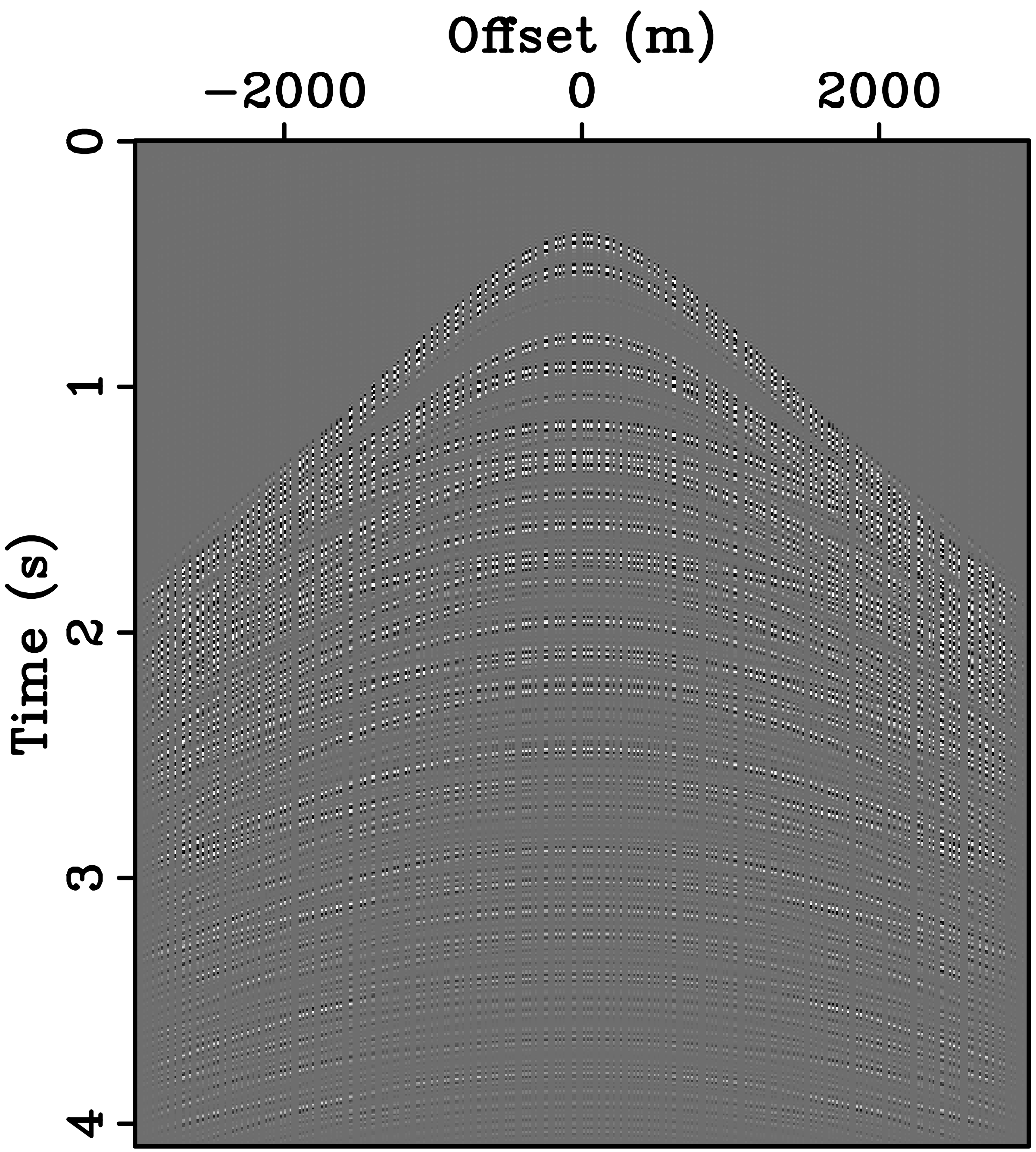


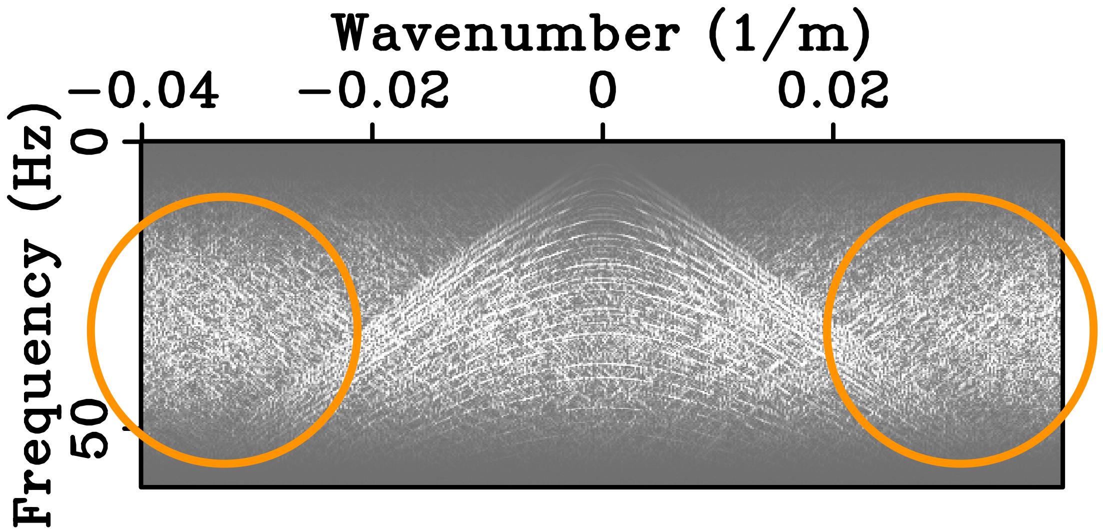

SNR = 6.92 dB



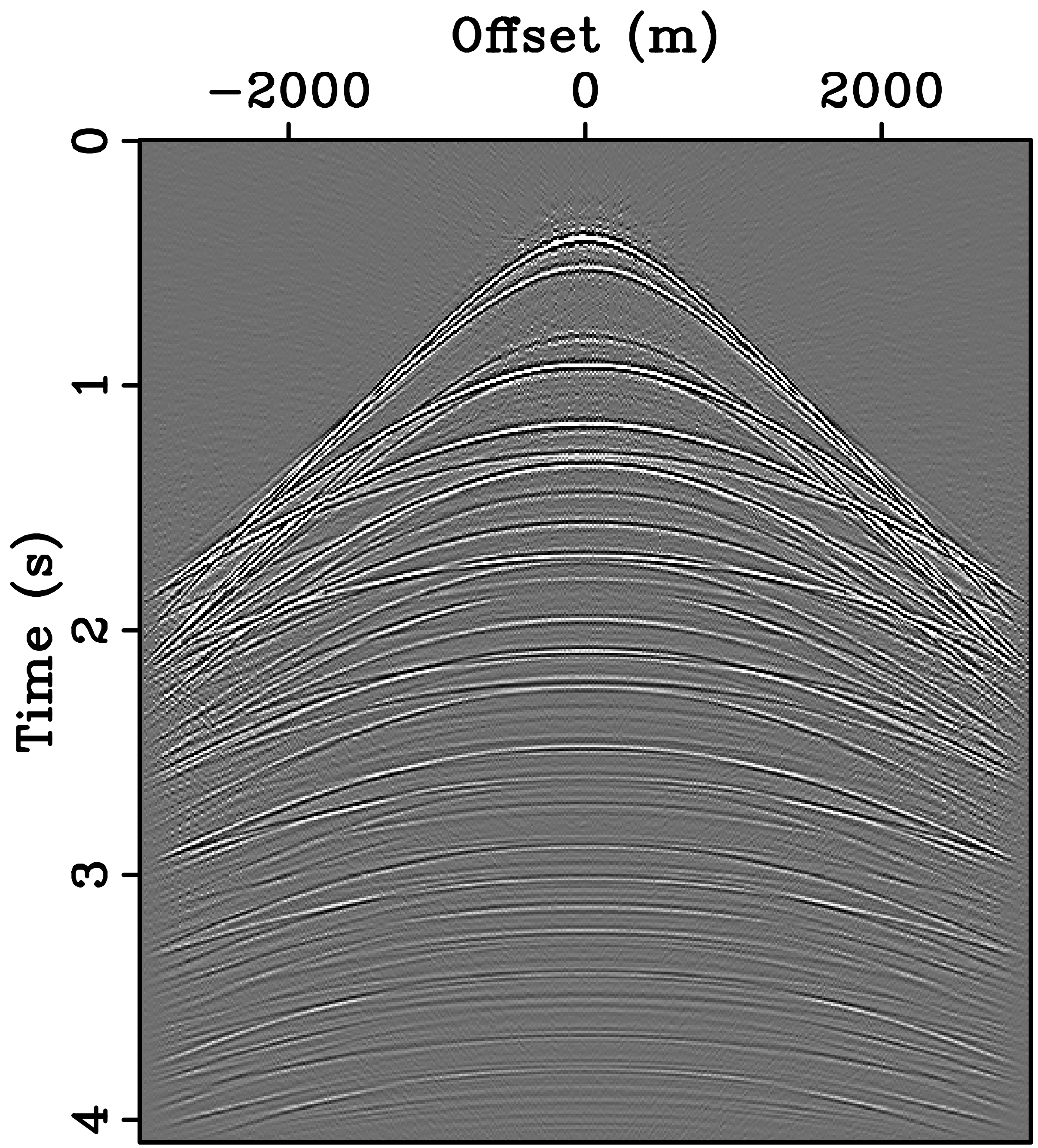

$$\mathrm{SNR} = 20 \times \log_{10}\left(\frac{\|\mathrm{model}\|_2}{\|\mathrm{reconstruction\ error}\|_2}\right)$$

# Optimally-jittered 3-fold undersampling

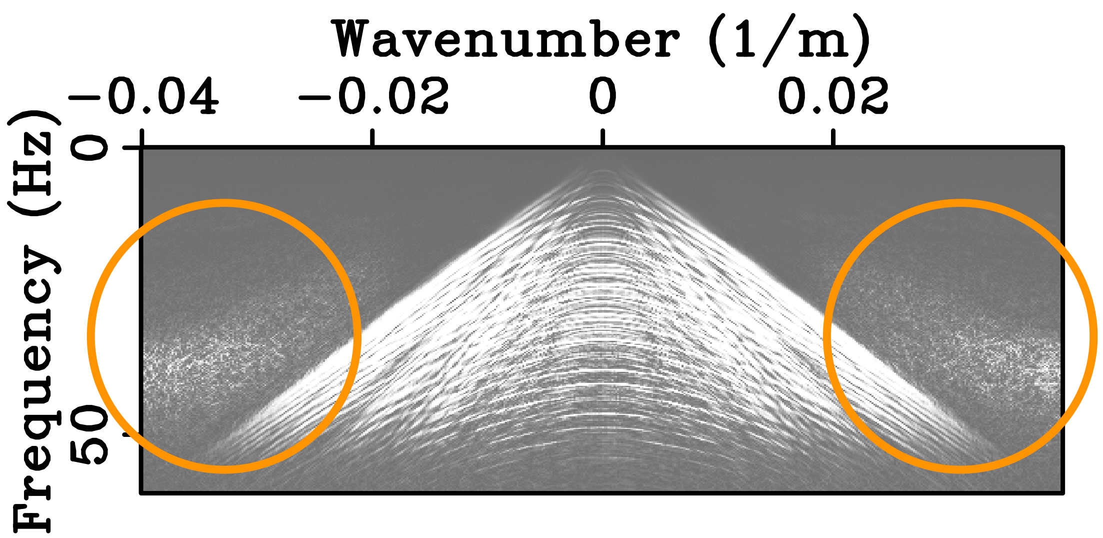


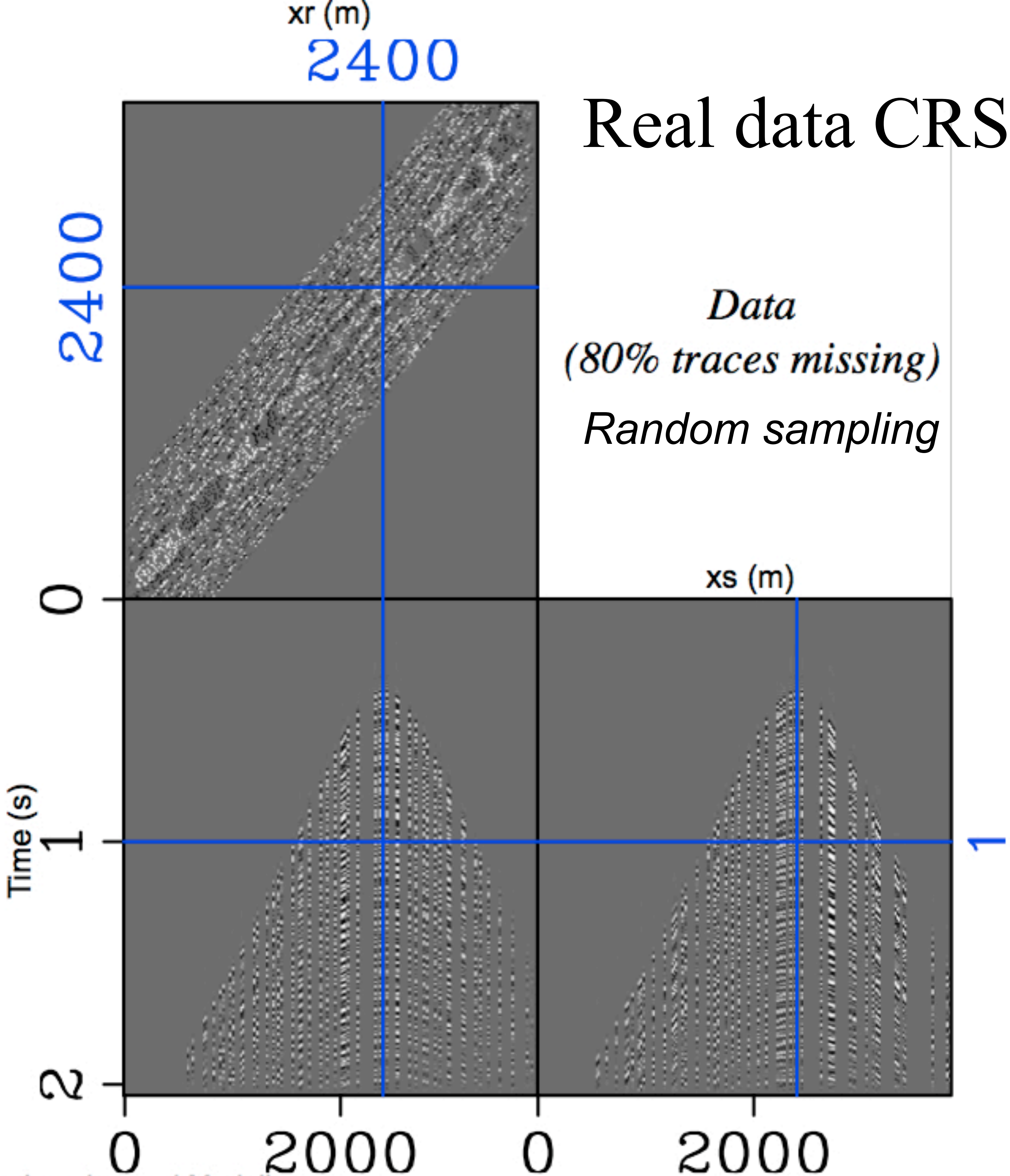
# CRSI from opt.-jittered 3-fold undersampling




SNR = 10.42 dB

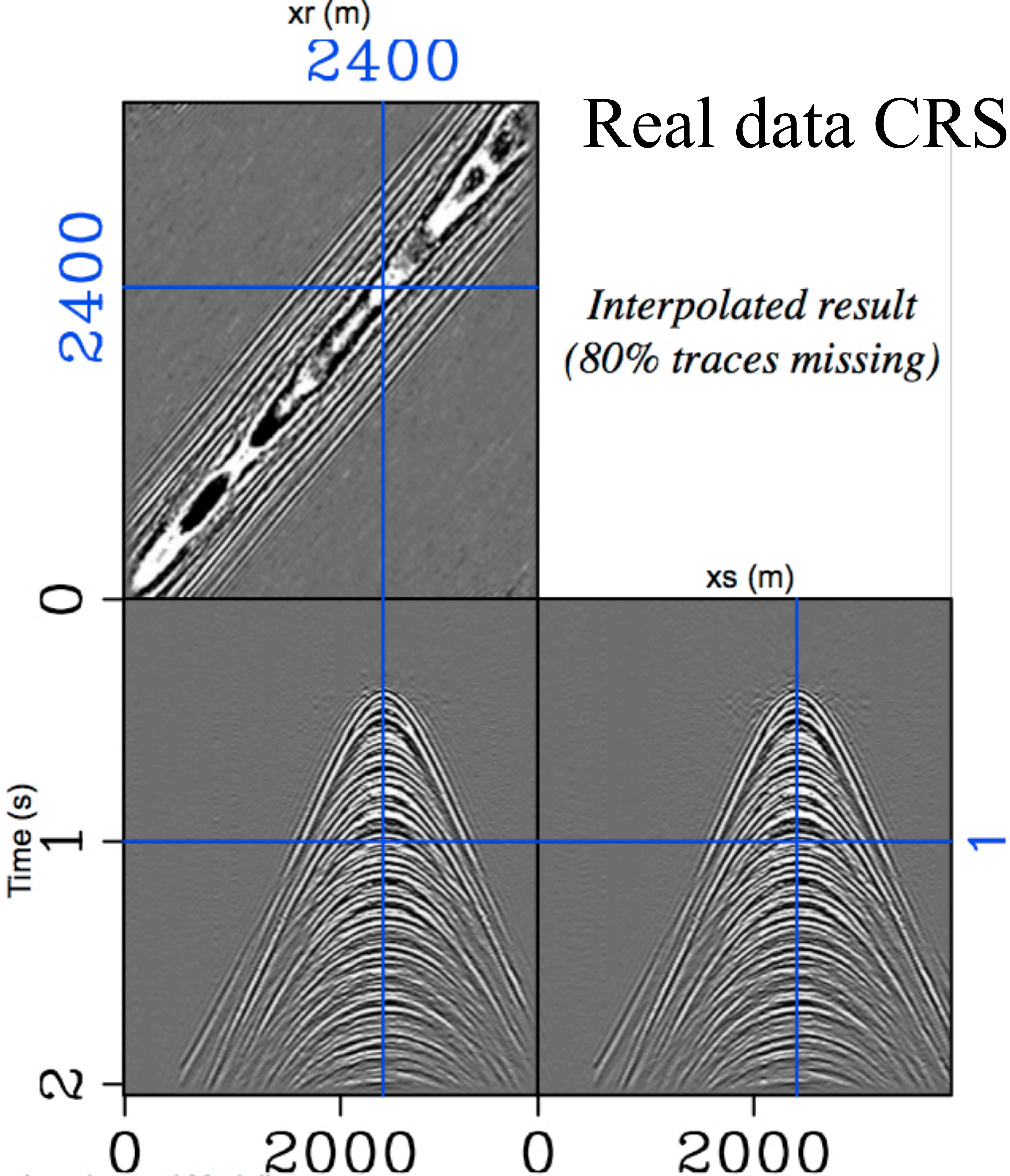


# Real data CRSI example




*Data*
*(80% traces missing)*
*Random sampling*

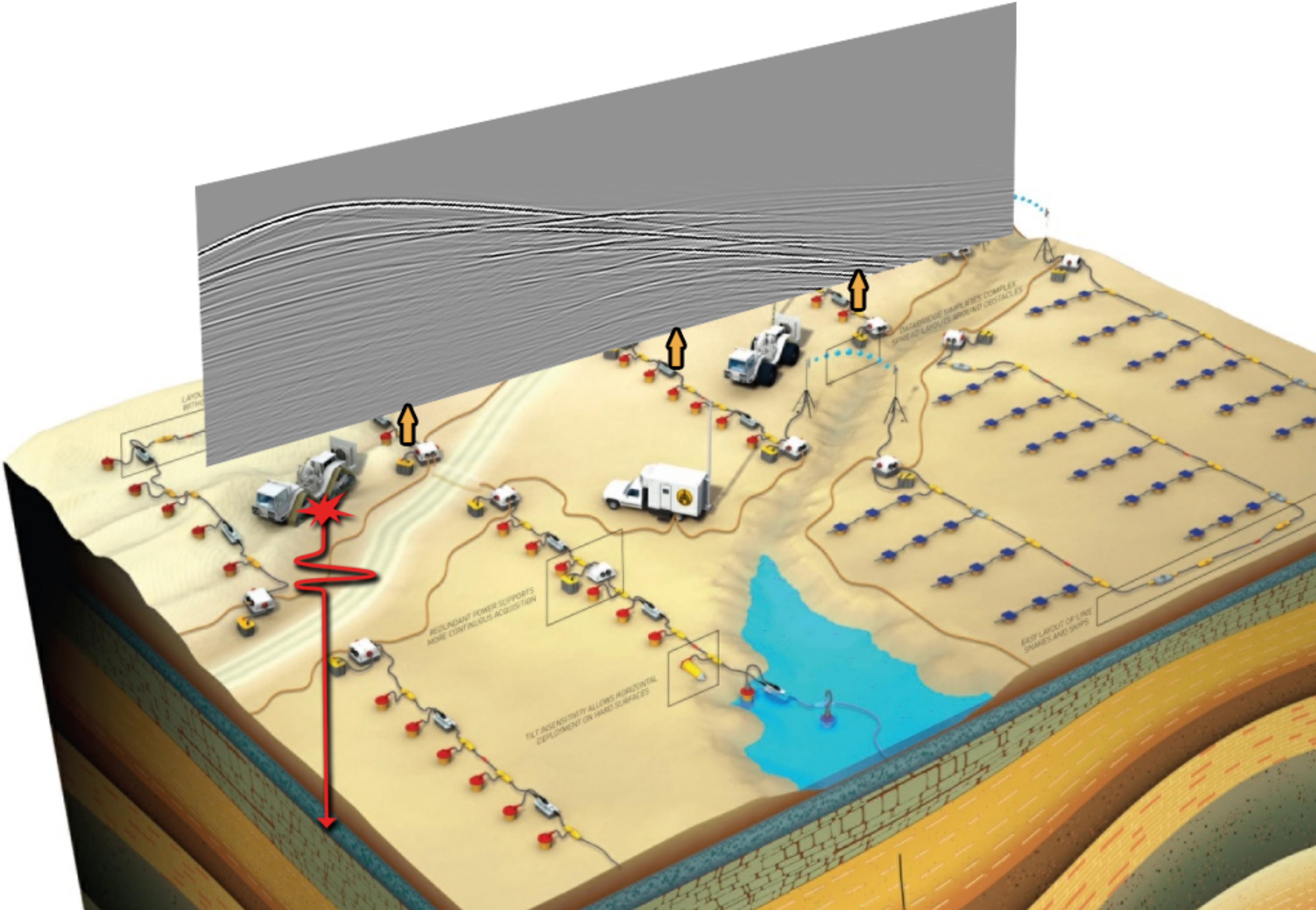
# Real data CRSI example



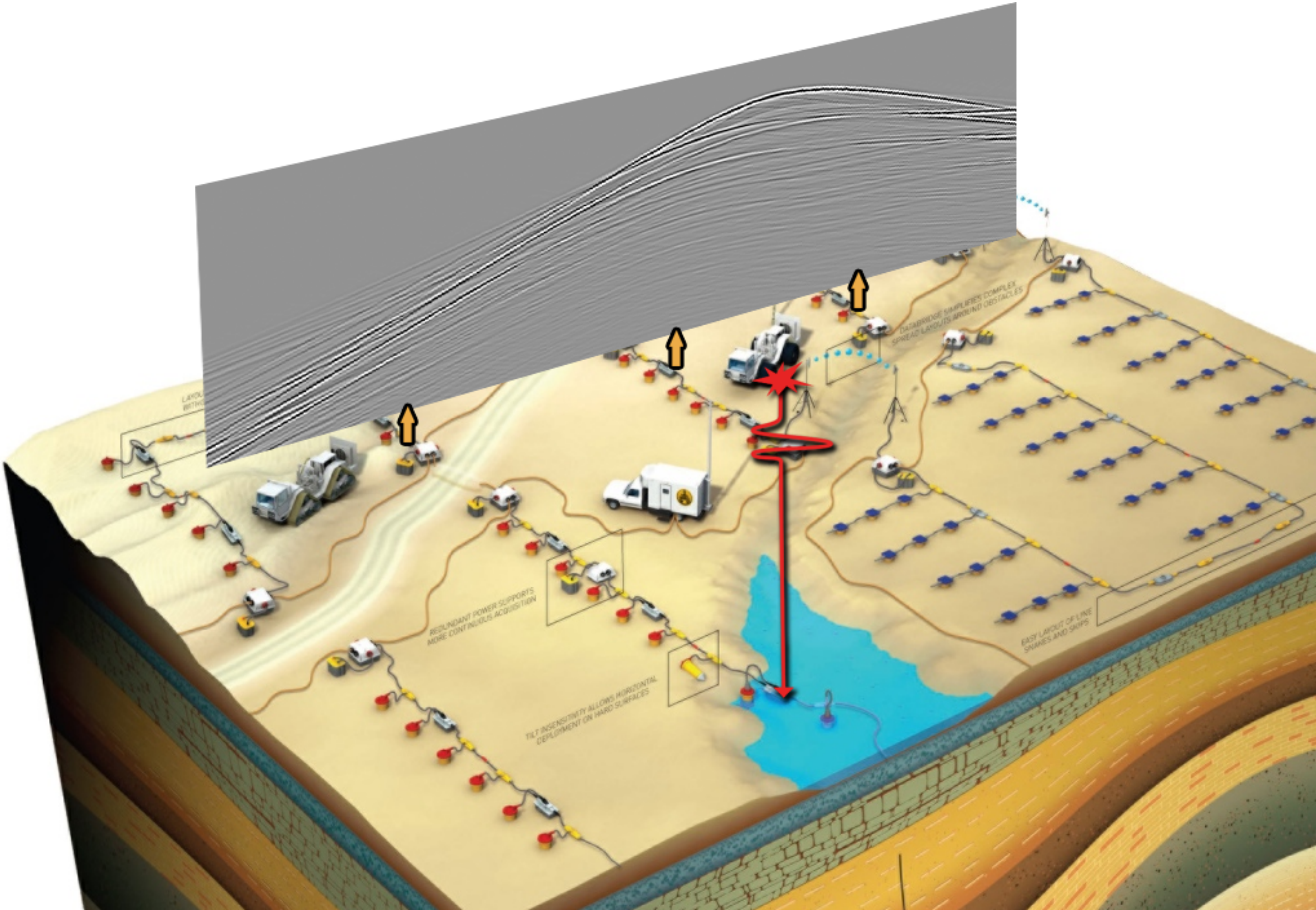

*Interpolated result*
*(80% traces missing)*

Acquisition image courtesy ION (www.iongeo.com)

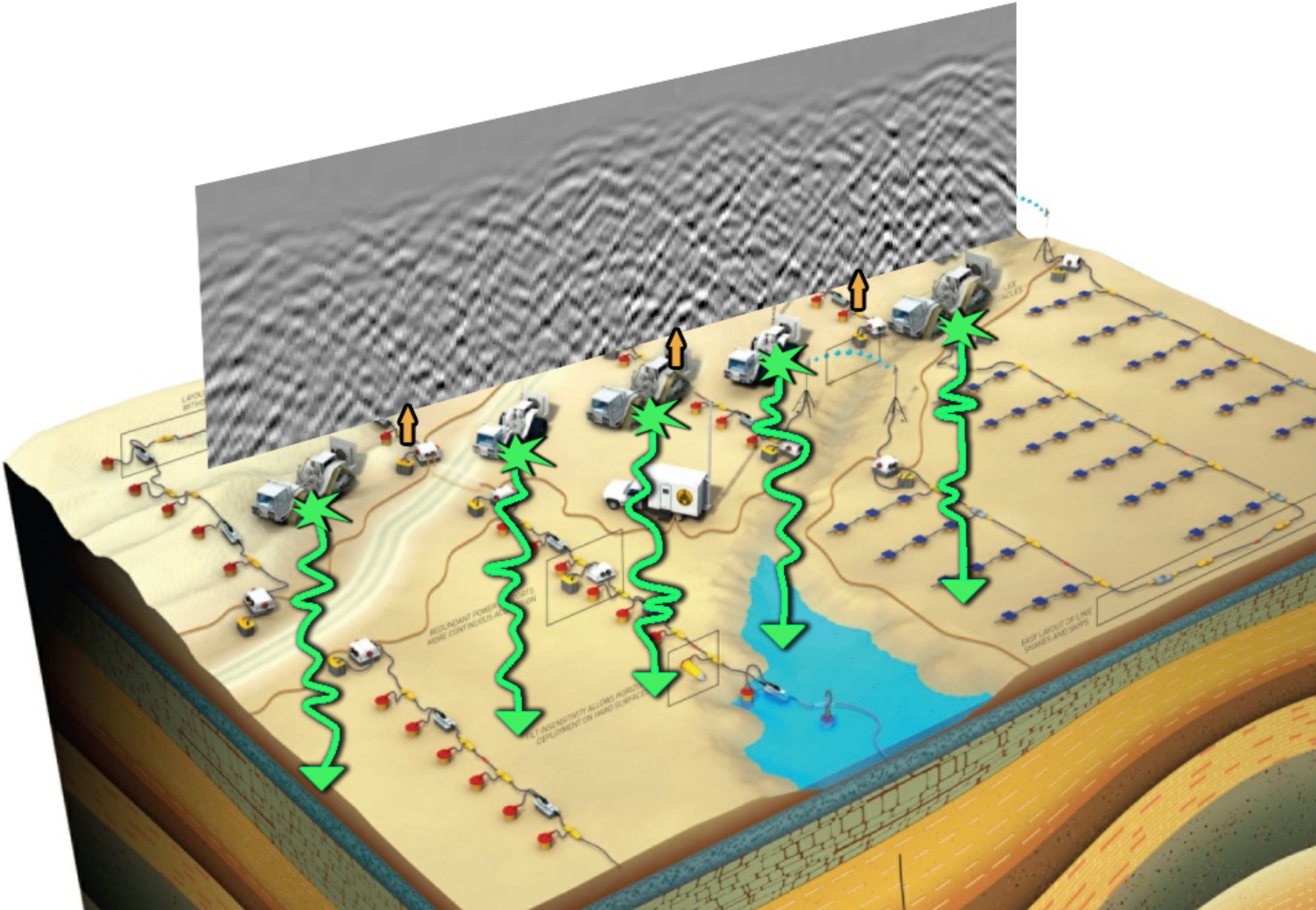
# Individual shots



# Individual shots



# Simultaneous & continuous shots



# Simultaneous modeling & acquisition

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

- **New paradigm:**
  - ***simultaneous*** *&* ***continuous*** source experiments in the field
  - ***simultaneous*** (continuous) simulations in the computer
  - **continuous** simultaneous simulations are equivalent to **multiple** simultaneous experiments
  - **Con:** *postprocessing necessary to separate into individual shots*

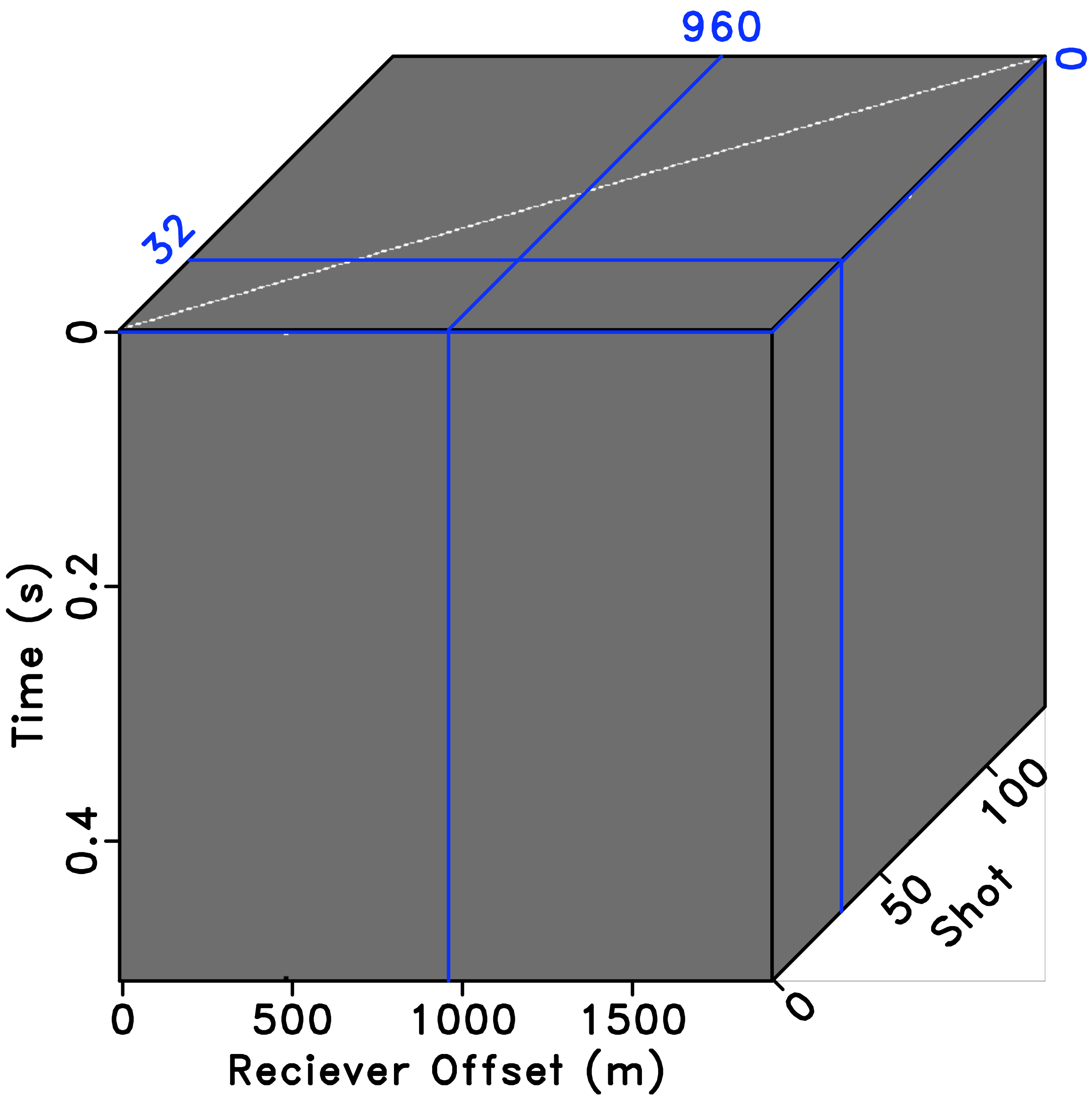
- **Key observation: this is *really* another instance of CS ...**

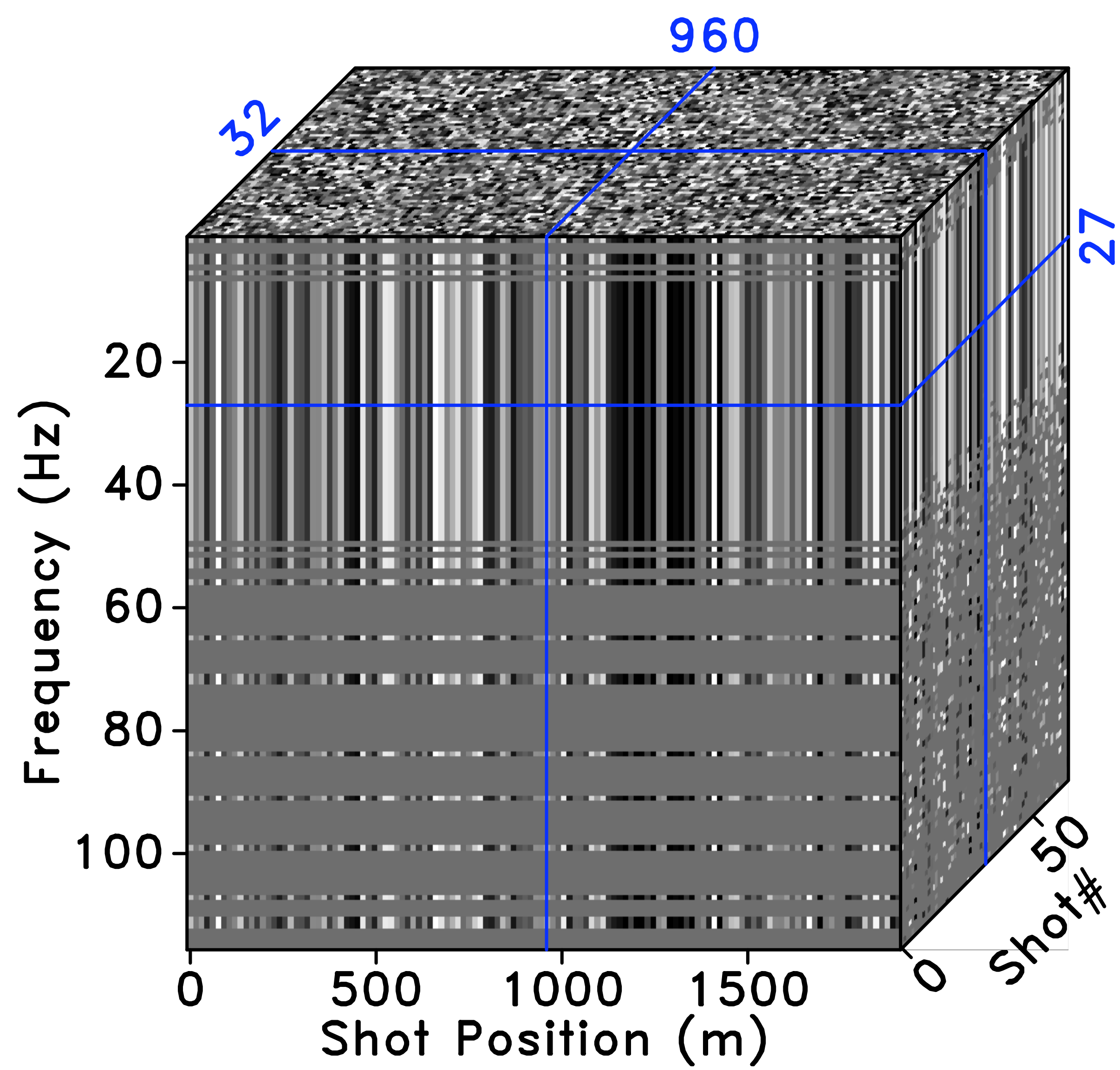
# Recovery from simultaneous acquisition

$$\mathbf{RM} = \overbrace{\begin{bmatrix} \mathbf{R}_1^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{R}_1^{\Omega} \\ \vdots \\ \mathbf{R}_{n_{s'}}^{\Sigma} \otimes \mathbf{I} \otimes \mathbf{R}_{n_{s'}}^{\Omega} \end{bmatrix}}^{\text{sub sampler}} \overbrace{\left(\mathbf{F}_2^* \operatorname{diag}\left(e^{\hat{i}\boldsymbol{\theta}}\right) \otimes \mathbf{I}\right) \mathbf{F}_3}^{\text{random phase encoder}},$$

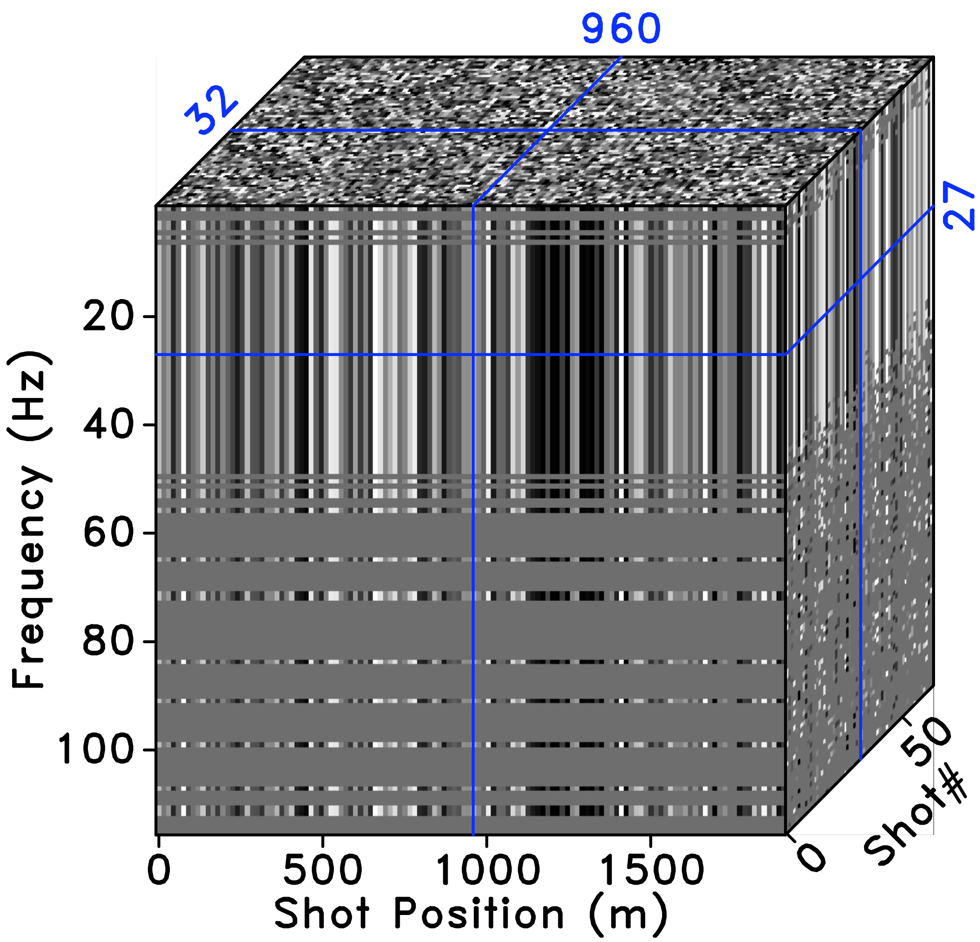
- Compressive sampling consists of
  - random phase encoding in the Fourier domain along the shot & receiver coordinates
  - subsample (by restriction) along the source and angular frequency coordinates
  - reduced data volume with reduced # of shots (now simultaneous) & frequencies
  - specifically adapted for time-harmonic processing

# Compressive sampling source wavefield

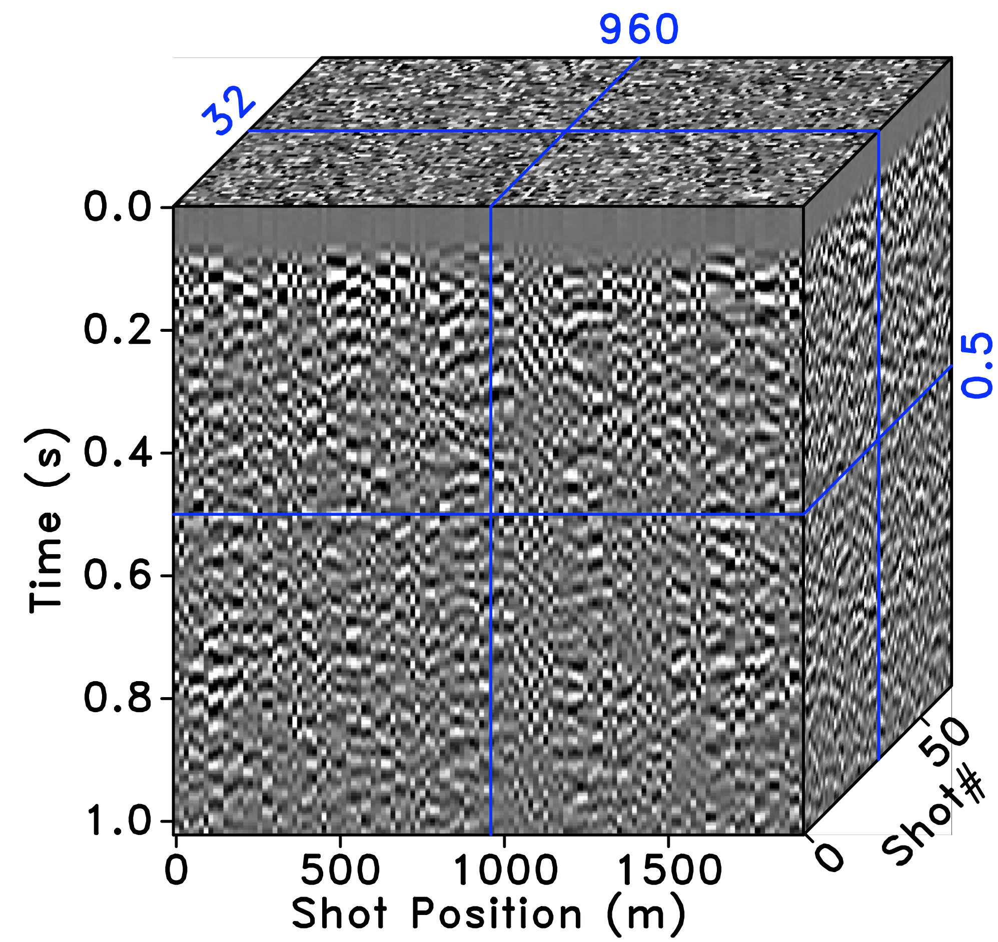


# Compressively sampled solution Helmholtz

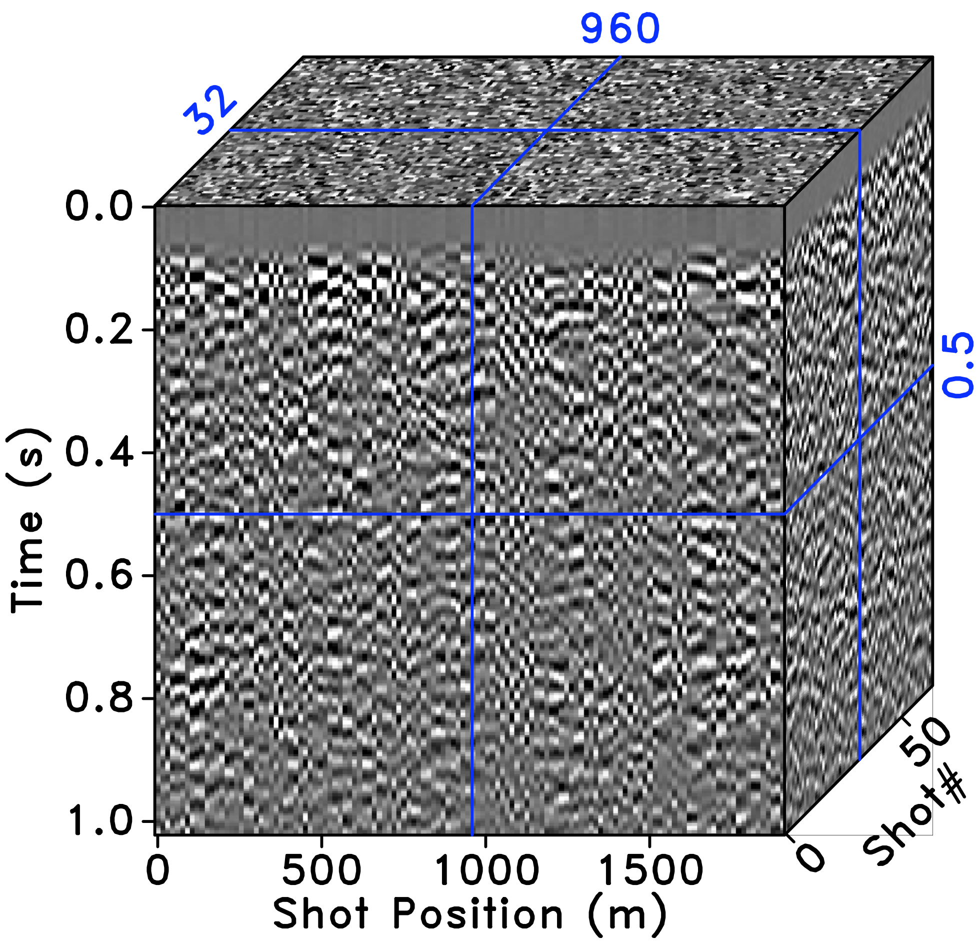


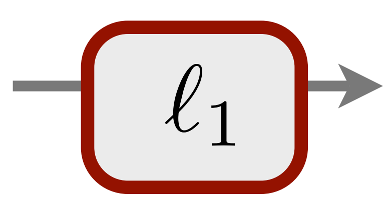

CS-sampled
sources

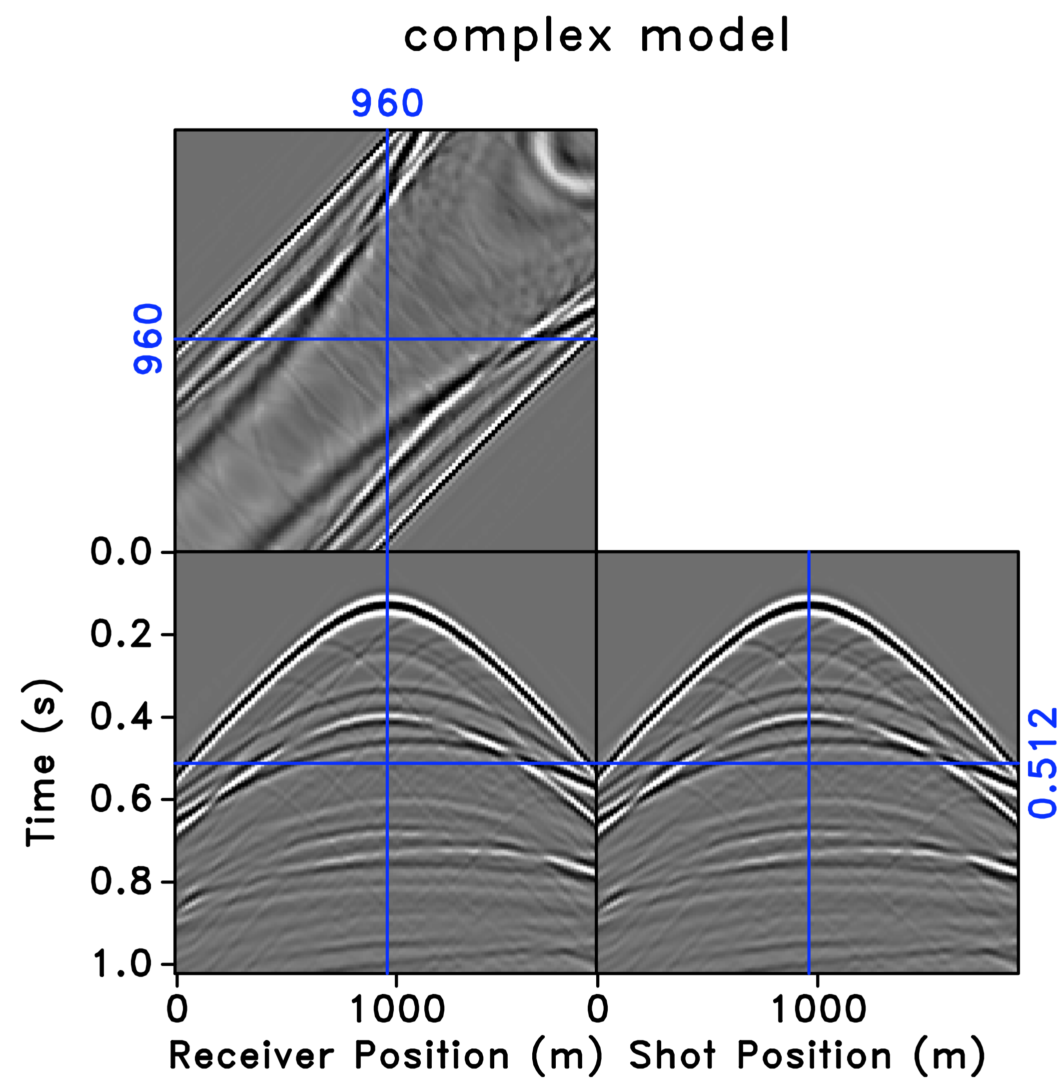




CS-sampled
solution

# Recovered solution by sparsity promotion




**CS-sampled solution**

# The bad: coherent noise separation

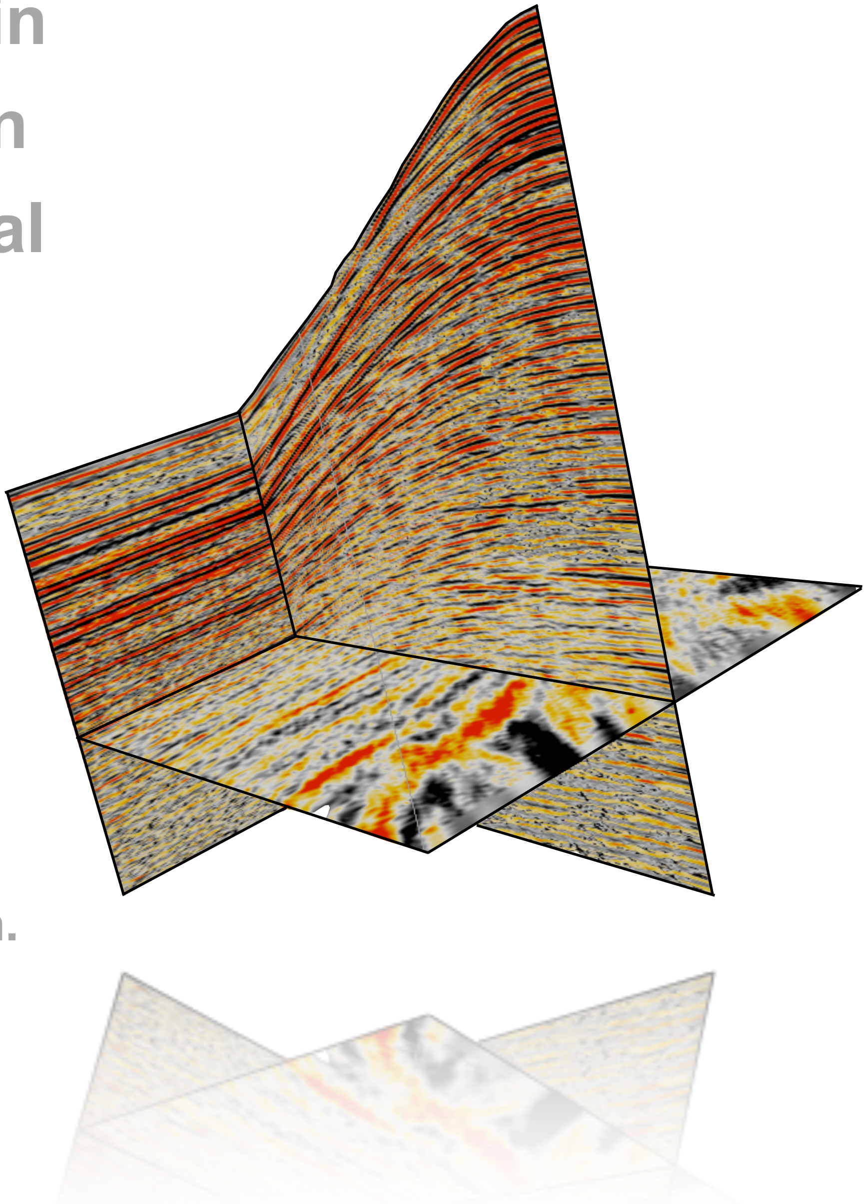
Signal is sparse in curvelet domain

Noise is sparse in curvelet domain

But have estimate of noise + signal and noise have different characteristics in the curvelet domain

D. Wang, R. Saab, O. Yilmaz and F J. Herrmann. Bayesian wavefield separation by transform-domain sparsity promotion. Geophysics, Vol 73, No. 5, A33-A38, 2008.

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



# Curvelet-based Bayesian separation

*Forward* model: [Saab et. al '07, Wang et.al '07, '08]

$$\mathbf{b} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n}$$ (total data)

$$\mathbf{b}_1 = \mathbf{A}\mathbf{x}_1 + \mathbf{n}_1$$ (predicted primaries)

$$\mathbf{b}_2 = \mathbf{A}\mathbf{x}_2 + \mathbf{n}_2$$ (predicted multiples)

where

$\mathbf{x}_1$ curvelet coefficients of *primaries*

$\mathbf{x}_2$ curvelet coefficients of *multiples*

$\mathbf{A}$ inverse curvelet transform

# Curvelet-based Bayesian separation

Involves the solution of the following nonlinear problem:

$$\mathbf{P_w}: \quad \begin{cases} \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \lambda_1 \|\mathbf{x}_1\|_{1,\mathbf{w}_1} + \lambda_2 \|\mathbf{x}_2\|_{1,\mathbf{w}_2} + \\ \|\underline{\mathbf{A}\mathbf{x}_2 - \mathbf{b}_2}\|_2^2 + \eta \|\mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) - \mathbf{b}\|_2^2 \\ \tilde{\mathbf{s}}_1 = \mathbf{A}\tilde{\mathbf{x}}_1 \quad \text{and} \quad \tilde{\mathbf{s}}_2 = \mathbf{A}\tilde{\mathbf{x}}_2. \end{cases}$$

where

- $\mathbf{b}_2$ predicted multiples
- $\mathbf{A}$ inverse discrete curvelet transforms
- $\tilde{\mathbf{s}}_{1,2}$ estimated primaries(1)and multiples(2)
- $\lambda_{1,2}$ and $\eta$ are control parameters

Can be solved by iterative soft thresholding.

SLIM Seismic Laboratory for Imaging and Modeling

# Curvelet-based Bayesian separation

Given initial estimates of $\mathbf{x}_1^0$ and $\mathbf{x}_2^0$, the $n^{\mathrm{th}}$ iteration of the algorithm proceeds as follows

$$\begin{aligned}
\mathbf{x}_1^{n+1} &= \mathbf{T}_{\frac{\lambda_1 \mathbf{w}_1}{2\eta}} \left[\mathbf{A}^T\mathbf{b}_2 - \mathbf{A}^T\mathbf{A}\mathbf{x}_2^n + \mathbf{A}^T\mathbf{b}_1 - \mathbf{A}^T\mathbf{A}\mathbf{x}_1^n + \mathbf{x}_1^n\right] \\
\mathbf{x}_2^{n+1} &= \mathbf{T}_{\frac{\lambda_2 \mathbf{w}_2}{2(1+\eta)}} \left[\mathbf{A}^T\mathbf{b}_2 - \mathbf{A}^T\mathbf{A}\mathbf{x}_2^n + \mathbf{x}_2^n + \frac{\eta}{\eta+1}(\mathbf{A}^T\mathbf{b}_1 - \mathbf{A}^T\mathbf{A}\mathbf{x}_1^n)\right]
\end{aligned}$$

where $\mathbf{T_u} : \mathbb{R}^{|\mathcal{M}|} \mapsto \mathbb{R}^{|\mathcal{M}|}$ is the elementwise soft-thresholding operator, i.e.,

$$T_{u_\mu}(v_\mu) := \frac{v_\mu}{|v_\mu|} \cdot \max(0, |v_\mu| -| u_\mu|)$$

SLIM
Seismic Laboratory for Imaging and Modeling

# Curvelet-based Bayesian separation

Parametrization:

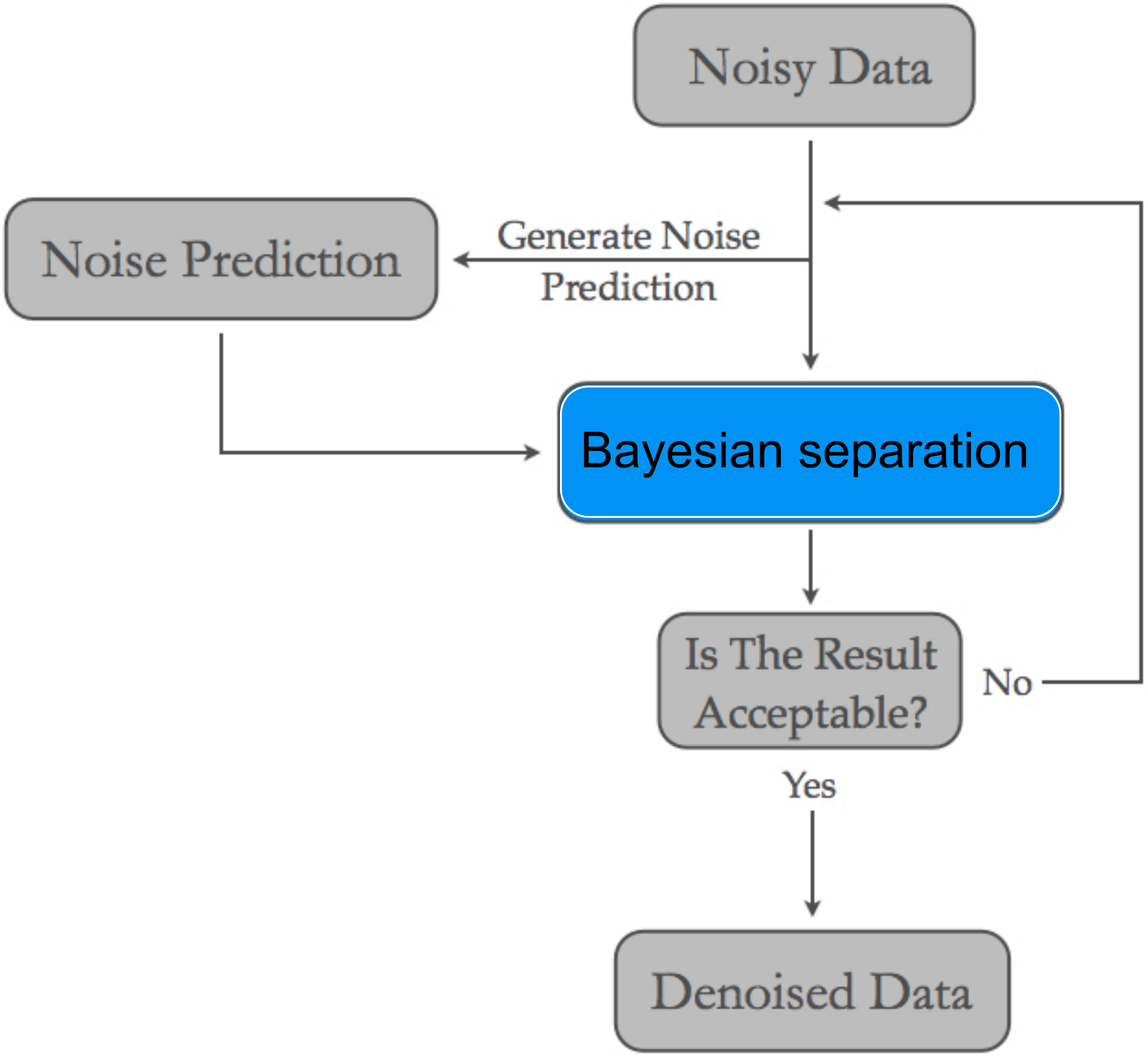
| | |
|---|---|
| $\eta$ | Prediction confidence parameter |
| $\lambda_1$ | Expected reflector sparsity |
| $\lambda_2$ | Expected surface wave sparsity |

Limiting case:

$\eta \rightarrow \infty$ Total lack of confidence => block-relaxation

SLIM
Seismic Laboratory for Imaging and Modeling

# Coherent noise-removal workflow



# Noise prediction & adaptation via matching

input data

$$\mathbf{m}_{\text{predicted}} = \mathbf{P}\mathbf{p} \text{ (multi-D convolution)}$$

conservative Fourier matching

$$\mathbf{m}_0 = \mathbf{F}\mathbf{m}_{\text{predicted}} \text{ with } \mathbf{F} = \mathcal{F}^H \text{diag}\left(\hat{\mathbf{f}}\right)\mathcal{F}$$

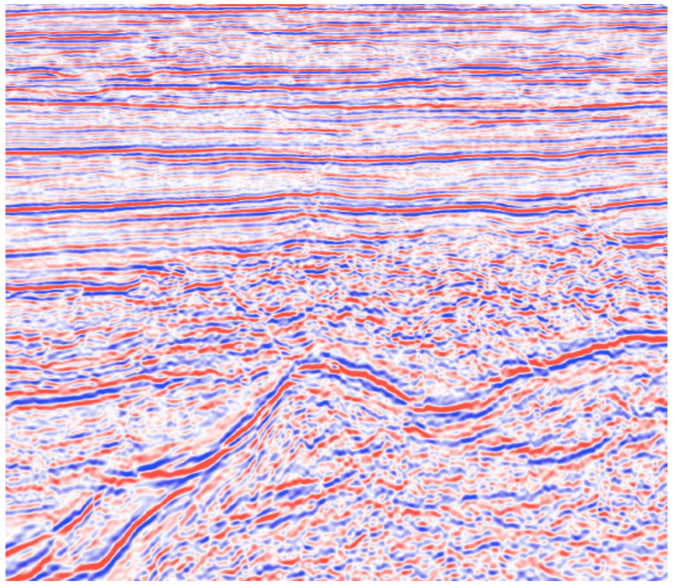
curvelet-domain matching

$$\begin{aligned} \mathbf{b}_2 &= \mathbf{B}\mathbf{m}_0 \\ \text{with } \mathbf{B} &= \mathbf{C}^T \text{diag}\left(e^{\mathbf{z}}\right)\mathbf{C}\mathbf{m}_0 \\ &\approx \mathcal{F}^H b(x,k)\mathcal{F}\mathbf{m}_0 \end{aligned}$$

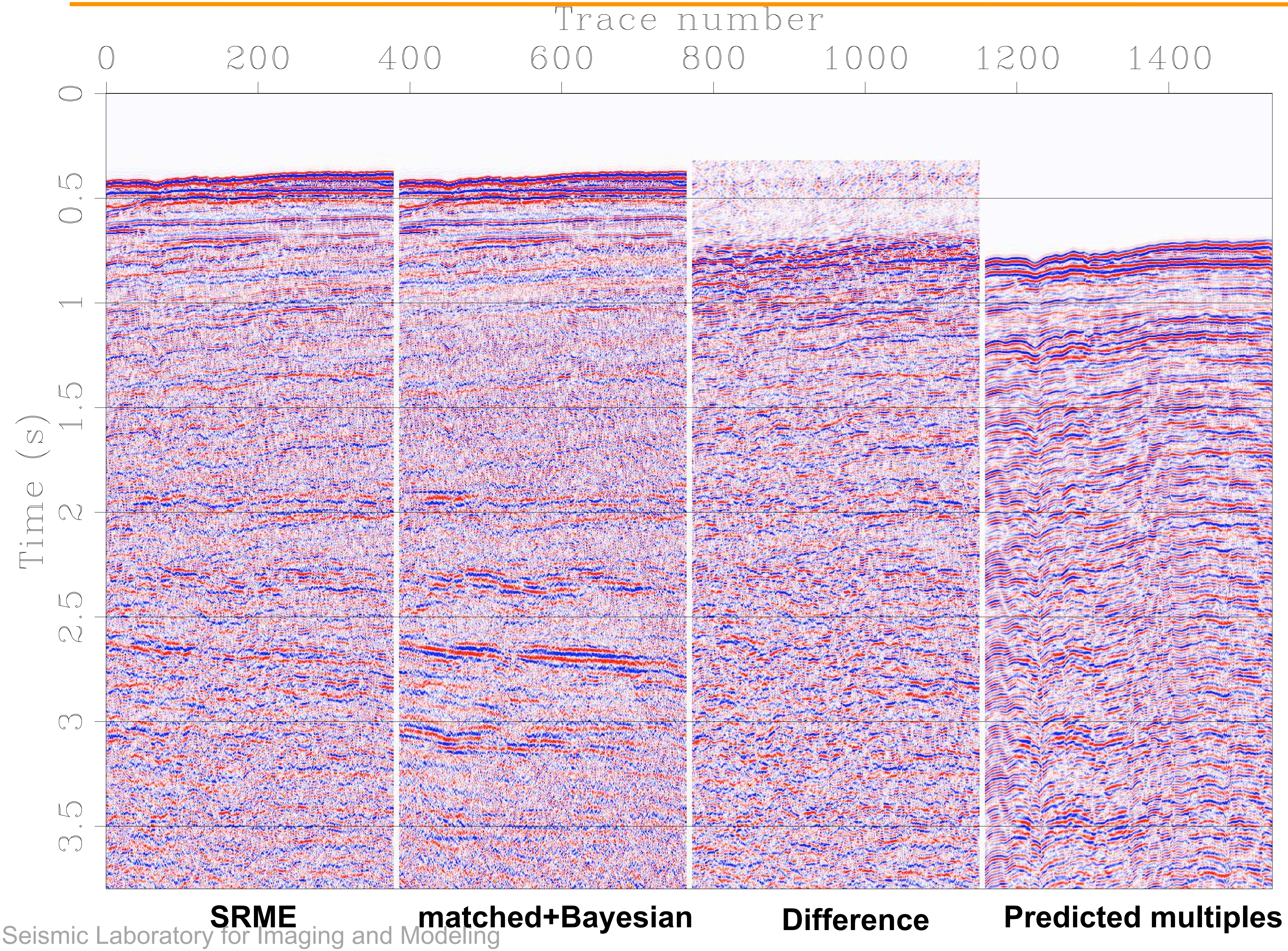
Bayesian separation

$$\mathbf{P}_{\mathbf{w}}: \begin{cases} \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \lambda_1\|\mathbf{x}_1\|_{1,\mathbf{w}_1} + \lambda_2\|\mathbf{x}_2\|_{1,\mathbf{w}_2} + \\ \|\mathbf{A}\mathbf{x}_2 - \mathbf{b}_2\|_2^2 + \eta\|\mathbf{A}(\mathbf{x}_1+\mathbf{x}_2) - \mathbf{b}\|_2^2 \\ \tilde{\mathbf{s}}_1 = \mathbf{A}\tilde{\mathbf{x}}_1 \quad \text{and} \quad \tilde{\mathbf{s}}_2 = \mathbf{A}\tilde{\mathbf{x}}_2. \end{cases}$$

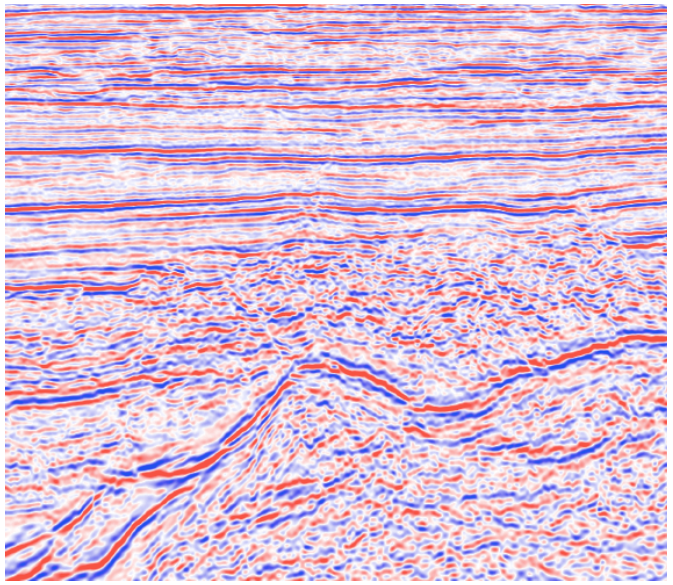
# Primary-multiple seperation work

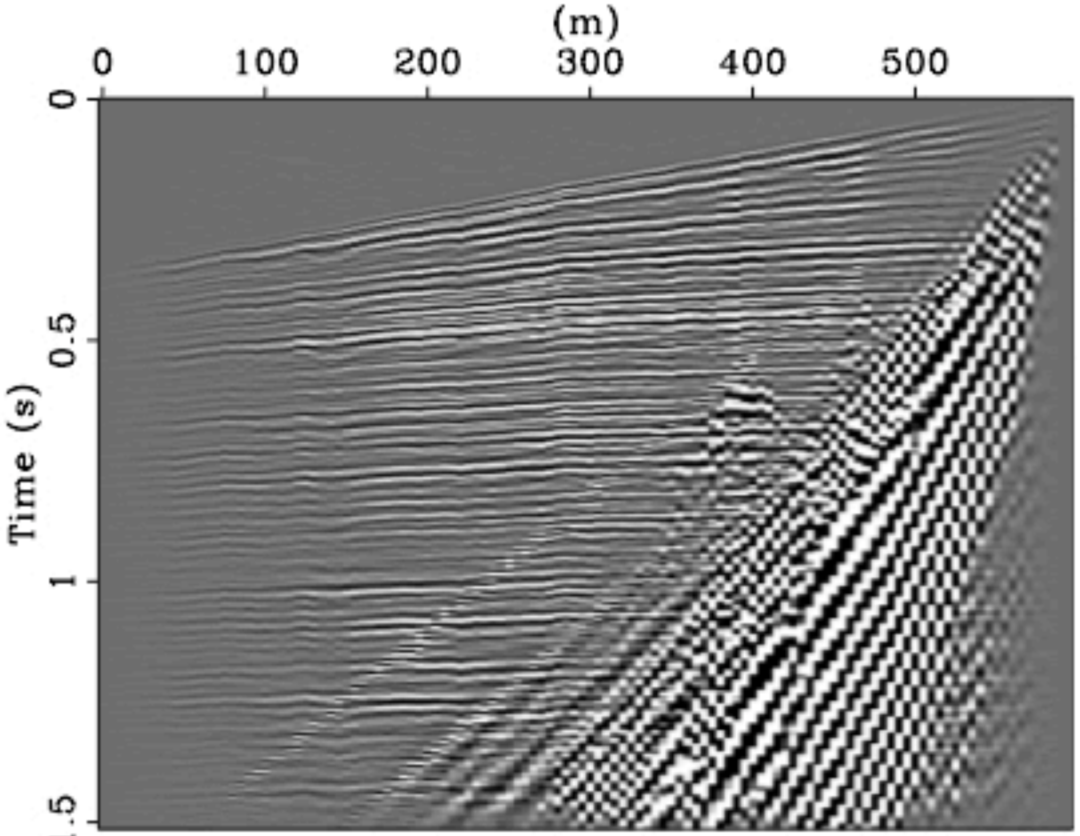


# Real-data example

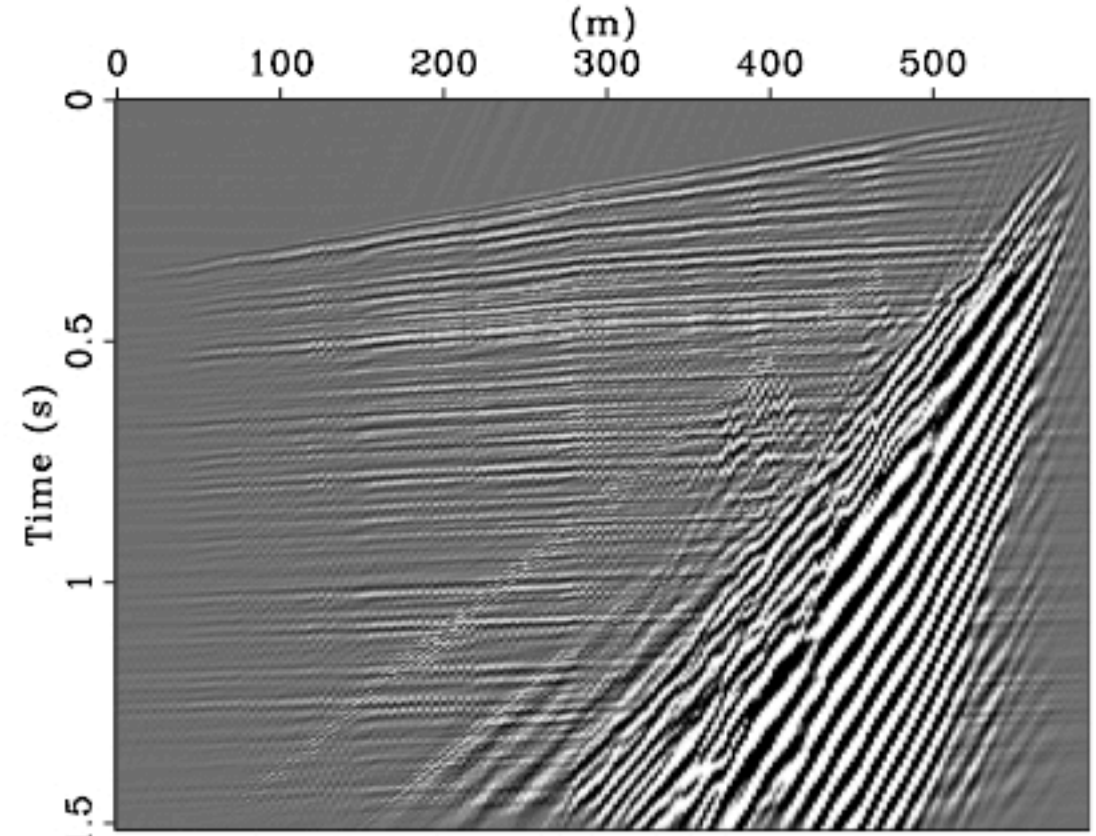


# Ground-roll removal work

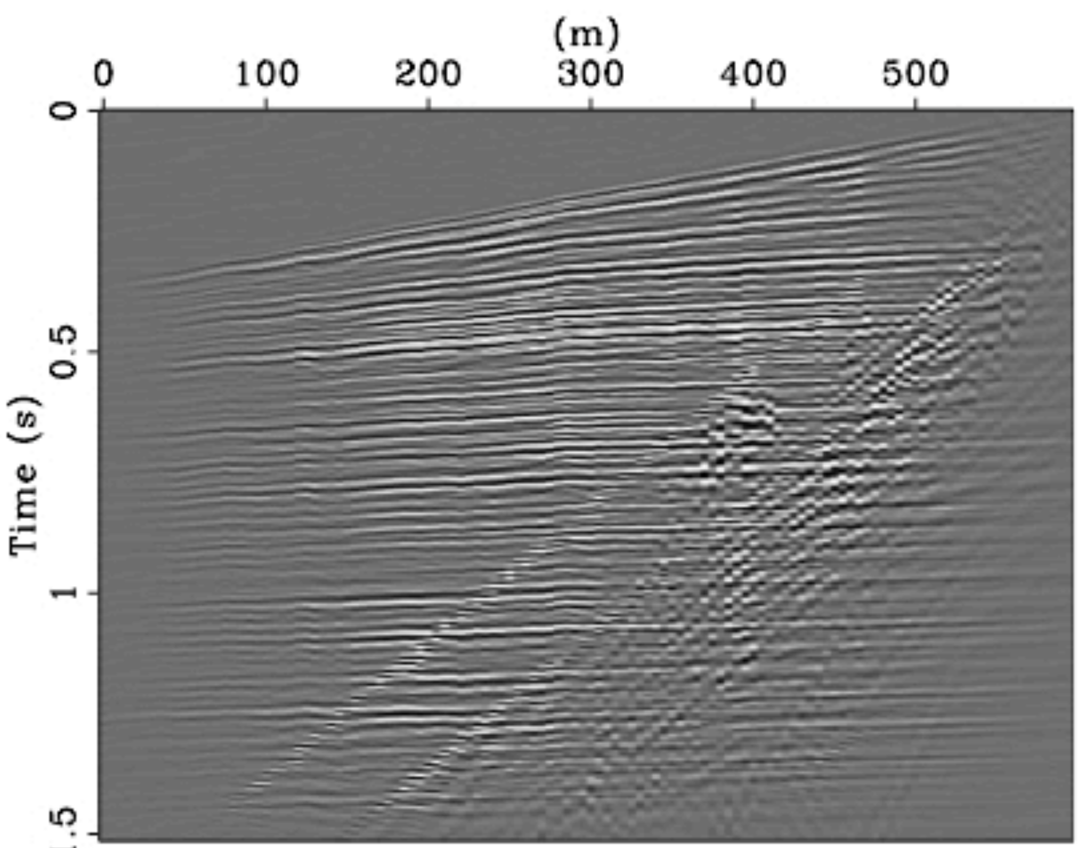




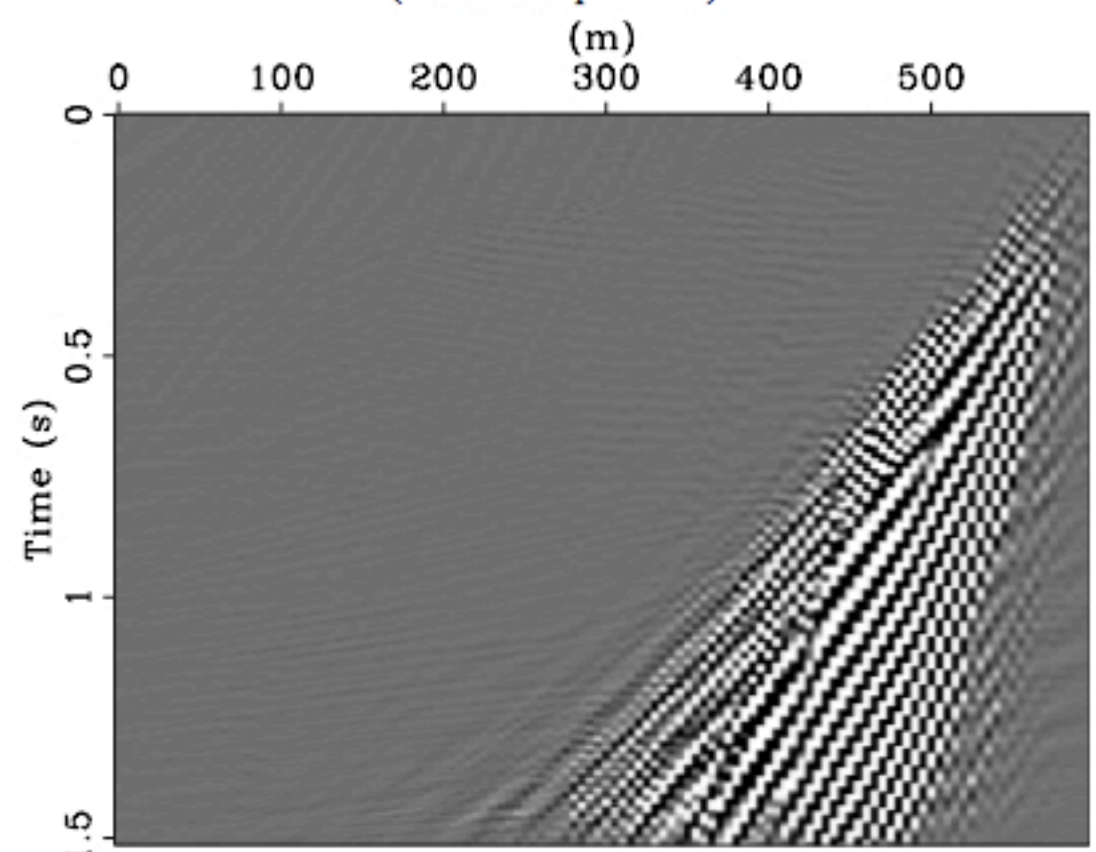

Data (5m) - aliased groundroll




Data (2.5m) - dealiased groundroll
(CRSI interpolation)



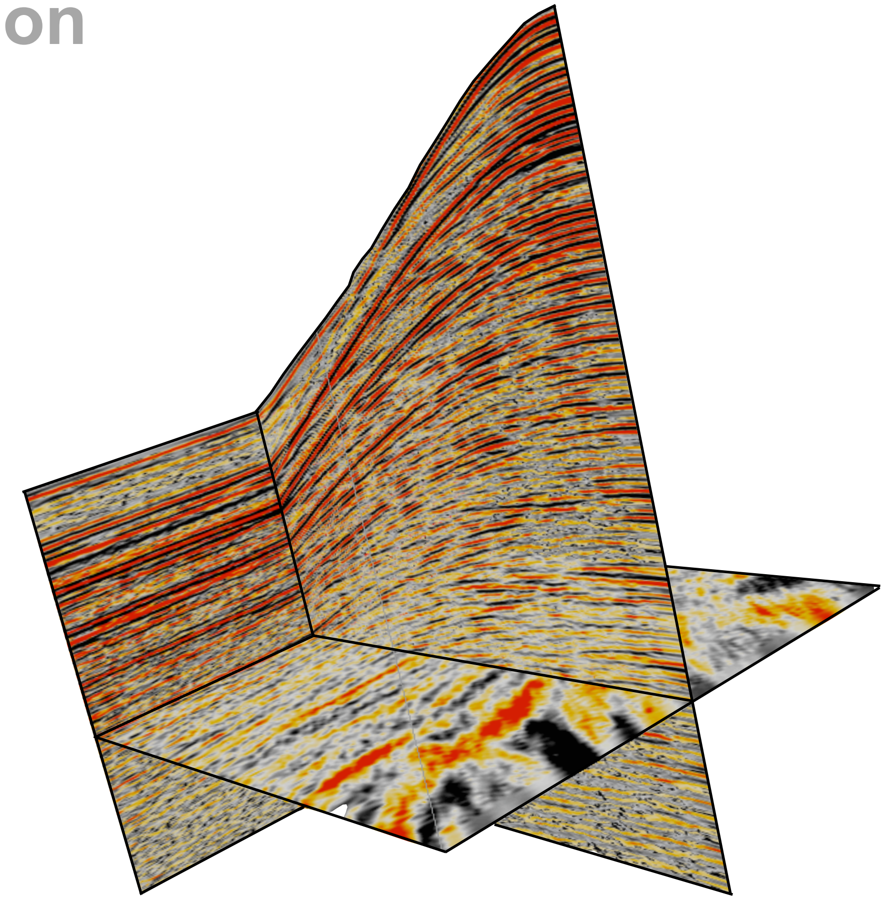

Curvelet denoised
data (5m)




Estimated
groundroll (5m)

# The ugly: noisy migration-amplitude recovery

- Coloring of noise by the migration operator
- Utilize near diagonalization of curvelets to approximate the covariance of the noise



# Migration-amplitude recovery

- Migration of noisy data yields a colored noisy image, i.e.,

$$\mathbf{y} = \mathbf{K}^T\mathbf{d} \text{ with } \mathbf{d} = \mathbf{K}\mathbf{m} + \mathbf{n}$$

- This noisy image contains a non-stationary colored noise term, i.e.,

$$\mathbf{e} = \mathbf{K}^T\mathbf{n}$$

$$\begin{aligned}
\mathbf{y} &= \text{noisy migrated image} \\
\mathbf{K} &= \text{linearized Born scattering operator} \\
\mathbf{m} &= \text{unknown reflectivity} \\
\mathbf{d} &= \text{noisy data} \\
\mathbf{n} &= \text{white Gaussian noise, i.e., } n_i \in N(0, \sigma)
\end{aligned}$$

# Migration-amplitude recovery

Use approximation

$$\mathbf{\Psi r} \approx \mathbf{A}\mathbf{A}^T\mathbf{r} \text{ with } \mathbf{A} = \mathbf{C}^T\mathbf{D}_\Psi^{1/2}$$

based on the diagonalization

$$\mathbf{\Psi} \approx \mathbf{C}^T\mathbf{D}_\Psi\mathbf{C} \text{ with } \operatorname{diag}(\mathbf{d}_\Psi)$$

Approximate forward model

$$\mathbf{y} \approx \mathbf{A}\mathbf{x}_0 + \mathbf{e}$$

Solve

$$\mathbf{P}_\epsilon : \quad \begin{cases} \tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon \\ \tilde{\mathbf{m}} = \left(\mathbf{A}^T\right)^\dagger \tilde{\mathbf{x}} \end{cases}$$

# Migration-amplitude recovery

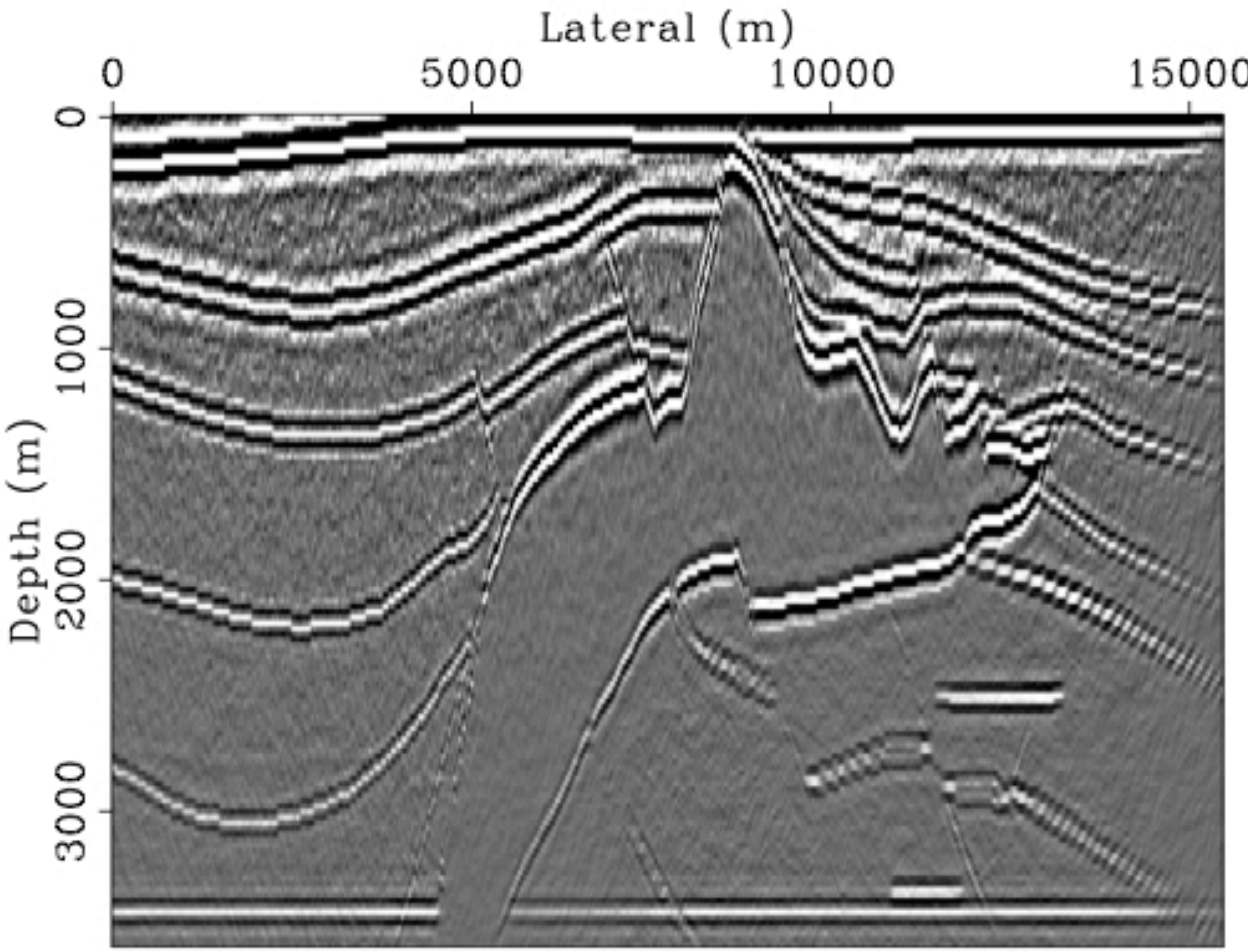
- Diagonal estimated by *adaptive* curvelet-domain matched filtering of a *migrated* and *remigrated* reference vector, i.e.,

$$\tilde{\mathbf{z}} = \arg\min_{\mathbf{z}} \frac{1}{2}\|\mathbf{\Psi}\mathbf{r} - \mathbf{C}^T \mathrm{diag}\left(\mathbf{C}\mathbf{r}\right) e^{\mathbf{z}}\|_2^2 + \gamma\|\mathbf{L}_{\mathcal{C}} e^{\mathbf{z}}\|_2^2$$

$$\mathbf{d}_{\Psi} = e^{\mathbf{z}}$$

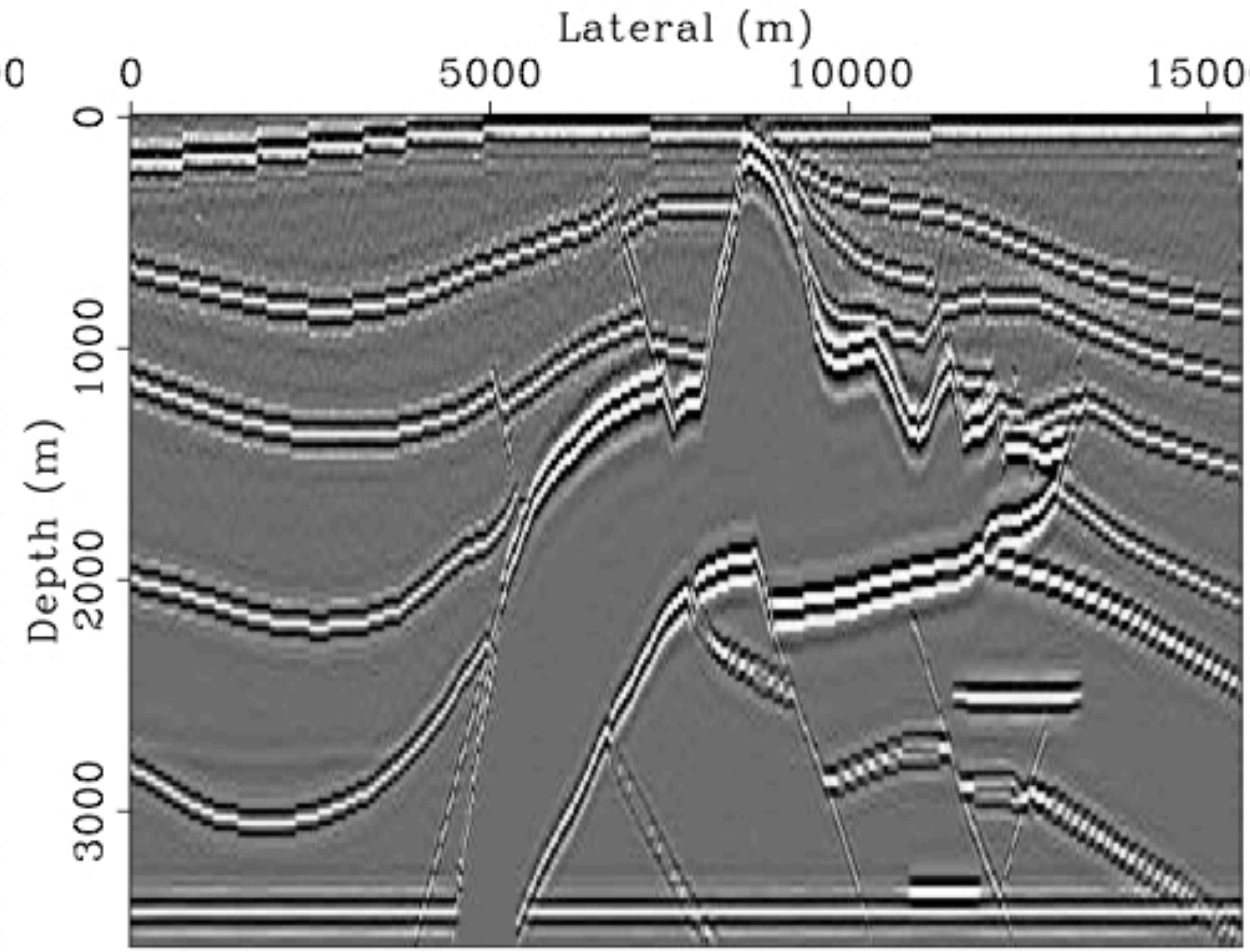
  - $\mathbf{L}_{\mathcal{C}}$ *curvelet-domain* sharpening operator that promotes smoothness
  - guarantees the solution to be *positive*
  - avoids *overfitting* => scaling contains action Hessian only

# Noisy example



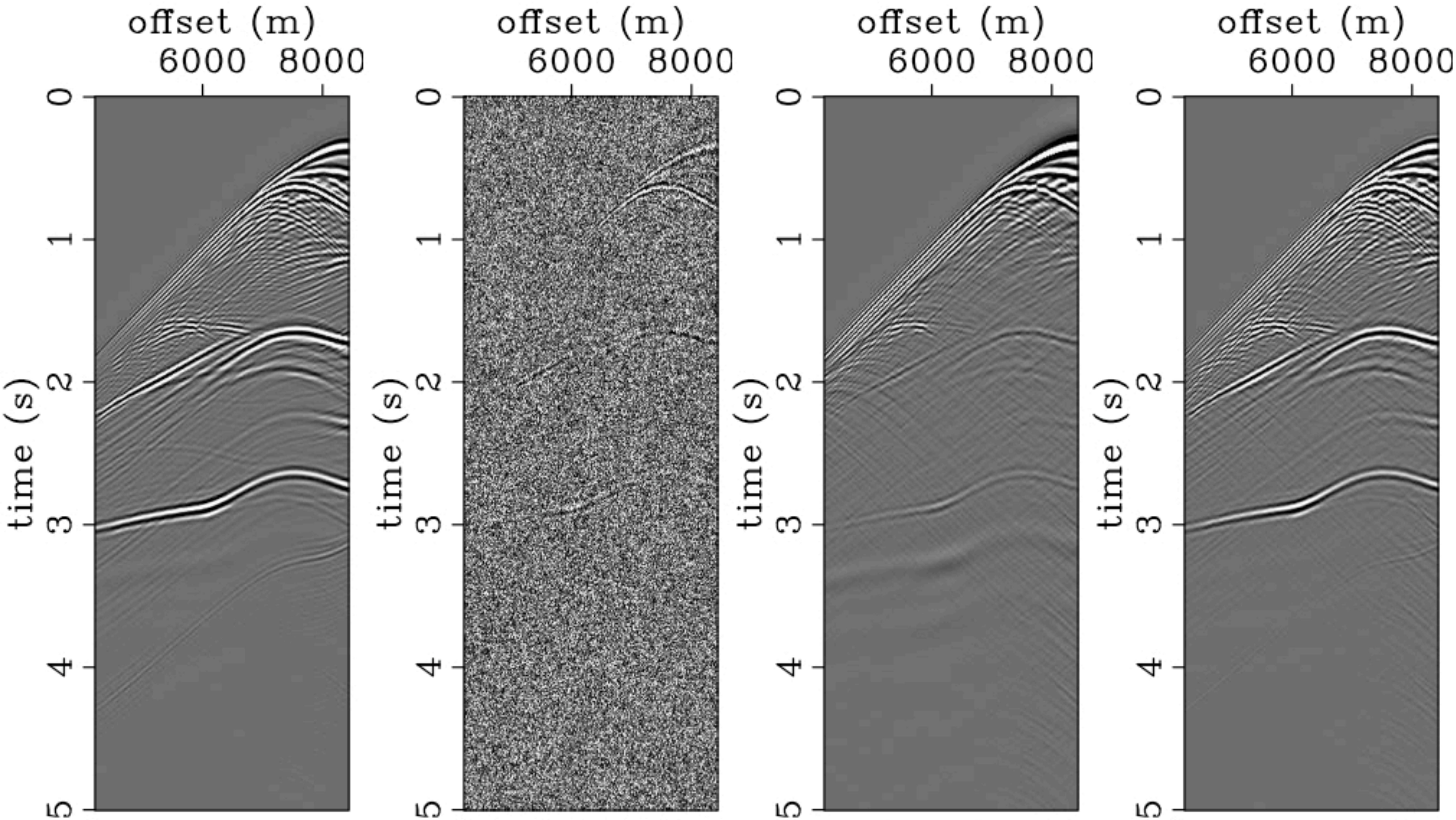

*Migrated data*

Noisy data




*Amplitude-corrected & denoised migrated data*

CRSI denoise




*Noise-free data*

*Noisy data (3 dB)*

*Data from migrated image*

*Data from amplitude-corrected & denoised migrated image*

# Conclusions

- Curvelets ability to detect wavefronts and their approximate invariance (=> sparsity) under wave propagation allow for successful removal for different types of noise from seismic data

- By compounding the curvelet transform with certain Matrices, different denoising problems can be cast into the same optimization problem. This optimization problem promotes sparsity in the curvelet domain which allows for of various noise and signal components

# Acknowledgements

- SLIM team members
- ExxonMobil for the real test dataset
- D. J. Verschuur for the Delphi dataset
- E. Candes, L. Demanet, and L. Ying for CurveLab
- S. Fomel and the developers of Madagascar
- W. Symes for the reverse-time migration code
- M. O'Brien, S. Gray, and J. Dellinger for the SEG AA' dataset

This presentation was carried out as part of the SINBAD project with financial support, secured through ITF, from the following organizations: BG, BP, Chevron, ExxonMobil, and Shell. SINBAD is part of the collaborative research & development (CRD) grant number 334810-05 funded by the Natural Science and Engineering Research Council (NSERC).