Curvelet-based non-linear adaptive subtraction with sparseness constraints

Felix J Herrmann, Peyman P Moghaddam
(EOS-UBC)

Earth and Ocean Sciences

Context

Spiky/minimal structure deconvolution (Claerbout, Ulrych, Oldenburg, Sacchi, Trad etc.)

Sparse Radon (Ulrych, Sacchi, Trad, etc.)

FFT/Focus transform-based Interpolation (Duyndam, Zwartjes, Verschuur, Berkhout)

Redundant dictonaries/Morphological component separation/Pursuits (Mallat, Chen, Donoho, Starck, Elad)

L2-migration (Nemeth, Chavent, de Hoop, Hu, Kuehl)

2-D/3-D Curvelets Non-linear synthesis (Durand, Starck, Candes, Demanet, Ying)

Anisotropic Diffusion (Osher)

Goals

Processing & imaging scheme

- **increases resolution & SNR**
- **preserves edges = freq. content**
- **works with and extends existing**
  - **noise removal/signal separation**
  - **imaging schemes**

Develop the right *language* to deal with SNR ≤ 0 ....

General Framework

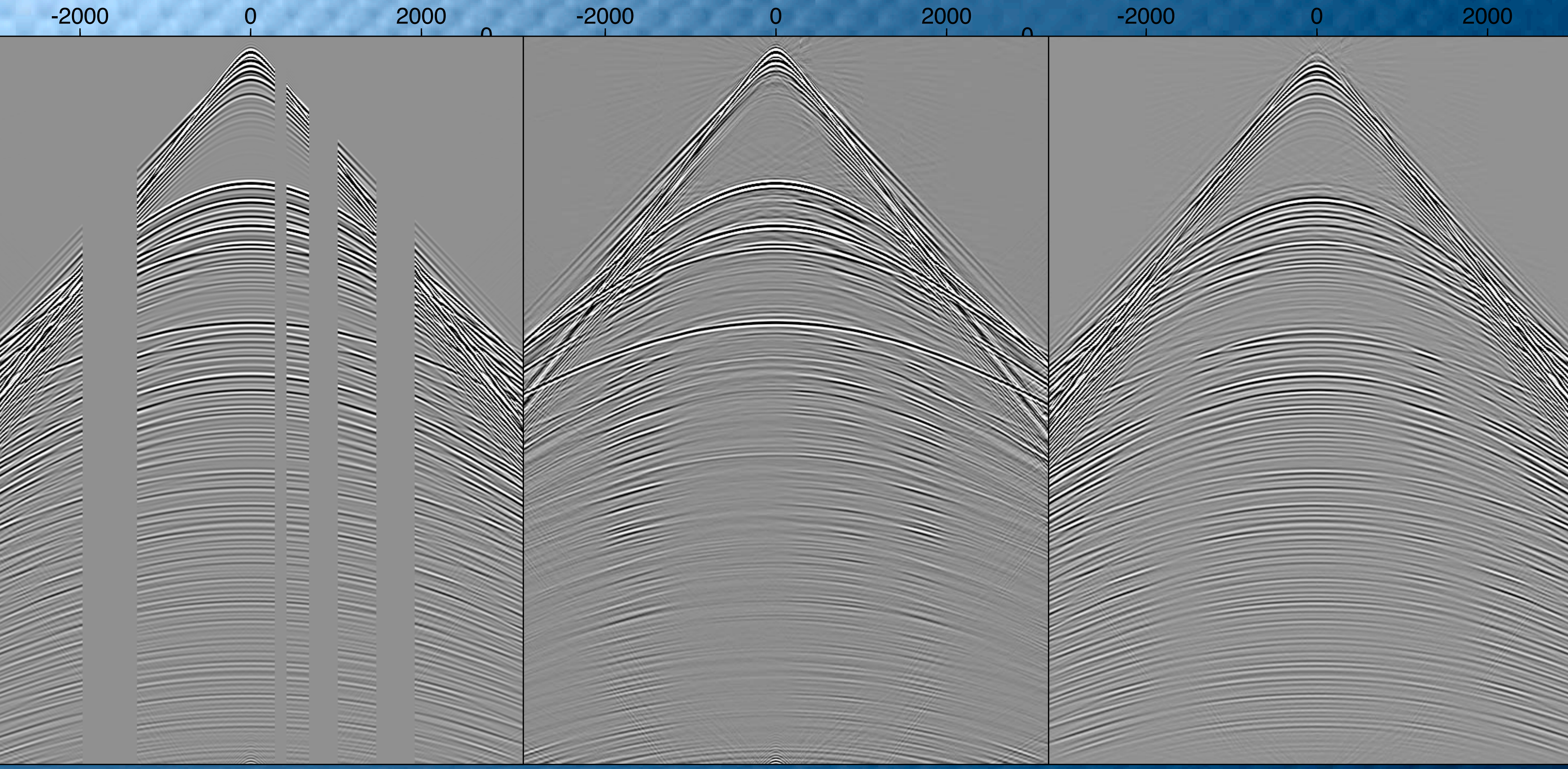
Divide-and-conquer approach:

1. **Sparseness with thresholding**
2. **Continuity with constrained optimization**

Main focus:

- **use of Curvelets as optimal Frames**
- **(iterative) thresholding**

Appetizer



missing data with multiples **estimated primaries** **estimated multiples**

Wish list

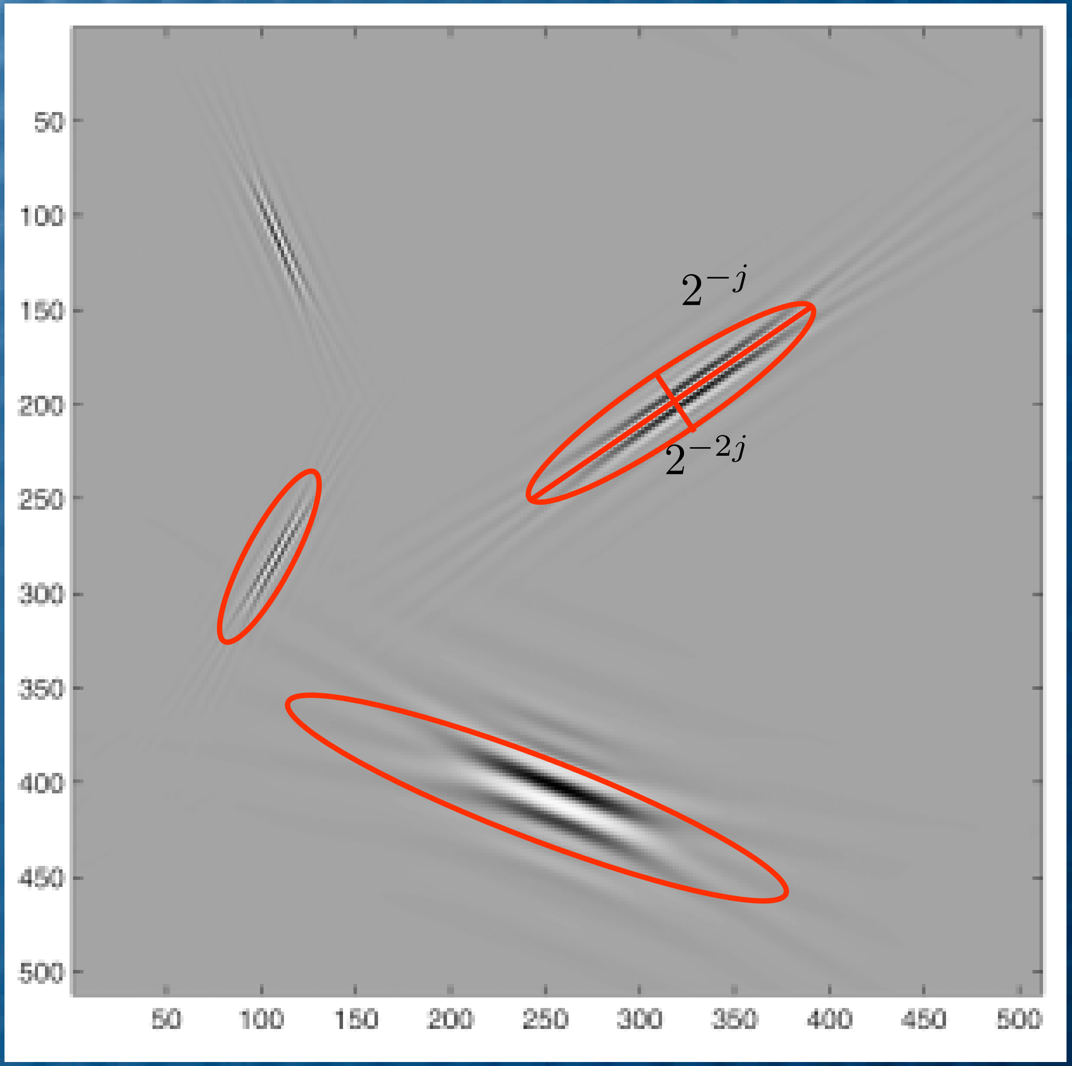
Seek a *transform domain* that is

- ***relative insensitive* to *local phase***
- ***sparse & local* (position/dip)**
- ***optimal* for *curved* events**
- ***well-behaved* under *operators***
- ***near diagonalizes Covariance***

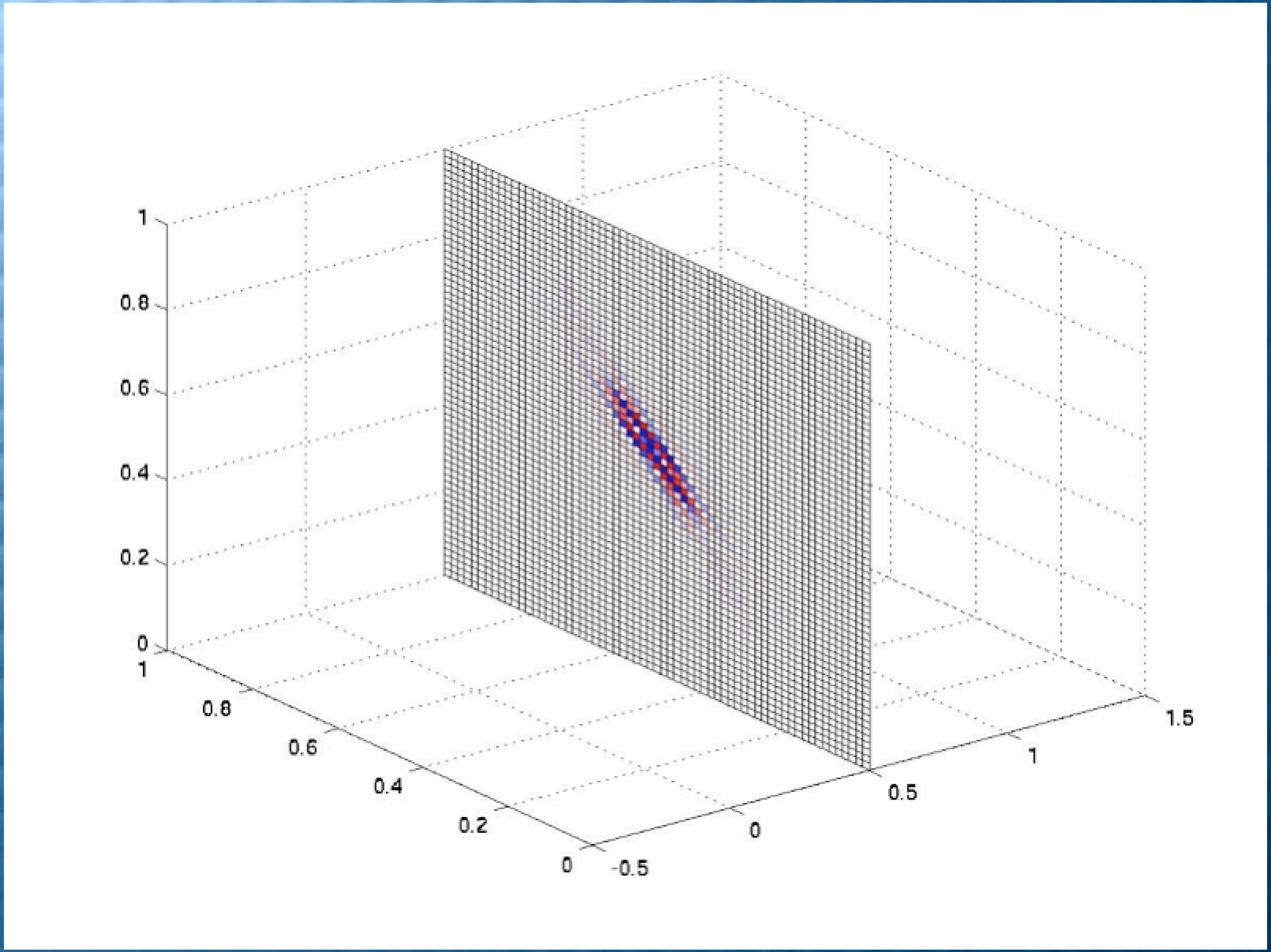
Aim to bring out those high frequencies with *ultra-low* SNR<0!

Why curvelets

- Nonseparable
- Local in 2-D space
- Local in 2-D Fourier
- Anisotropic
- Multiscale
- Almost orthogonal
- Tight frame $\mathrm{B}^T\mathrm{B} = \mathrm{I}$
- Optimal

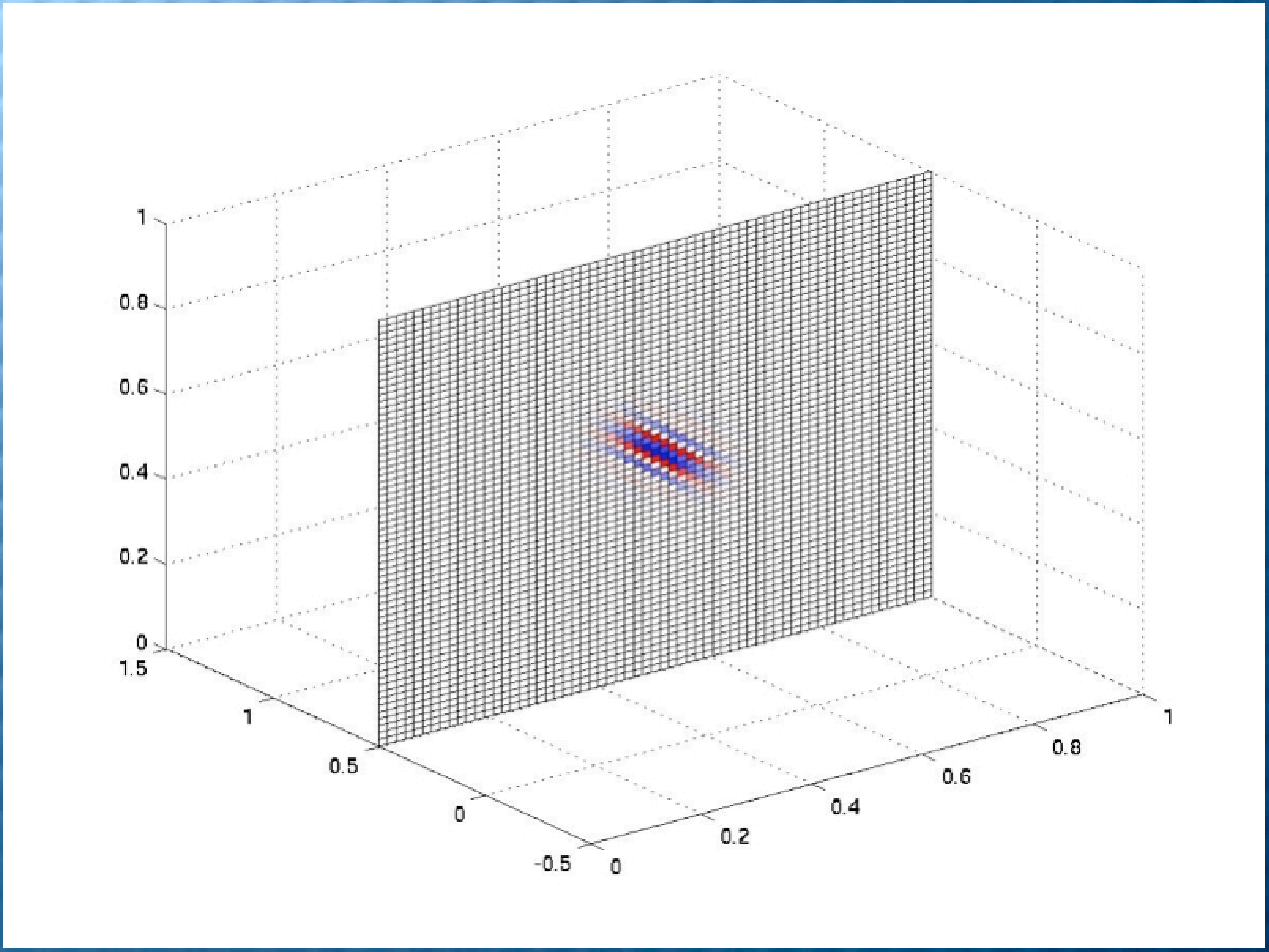


3-D Curvelets



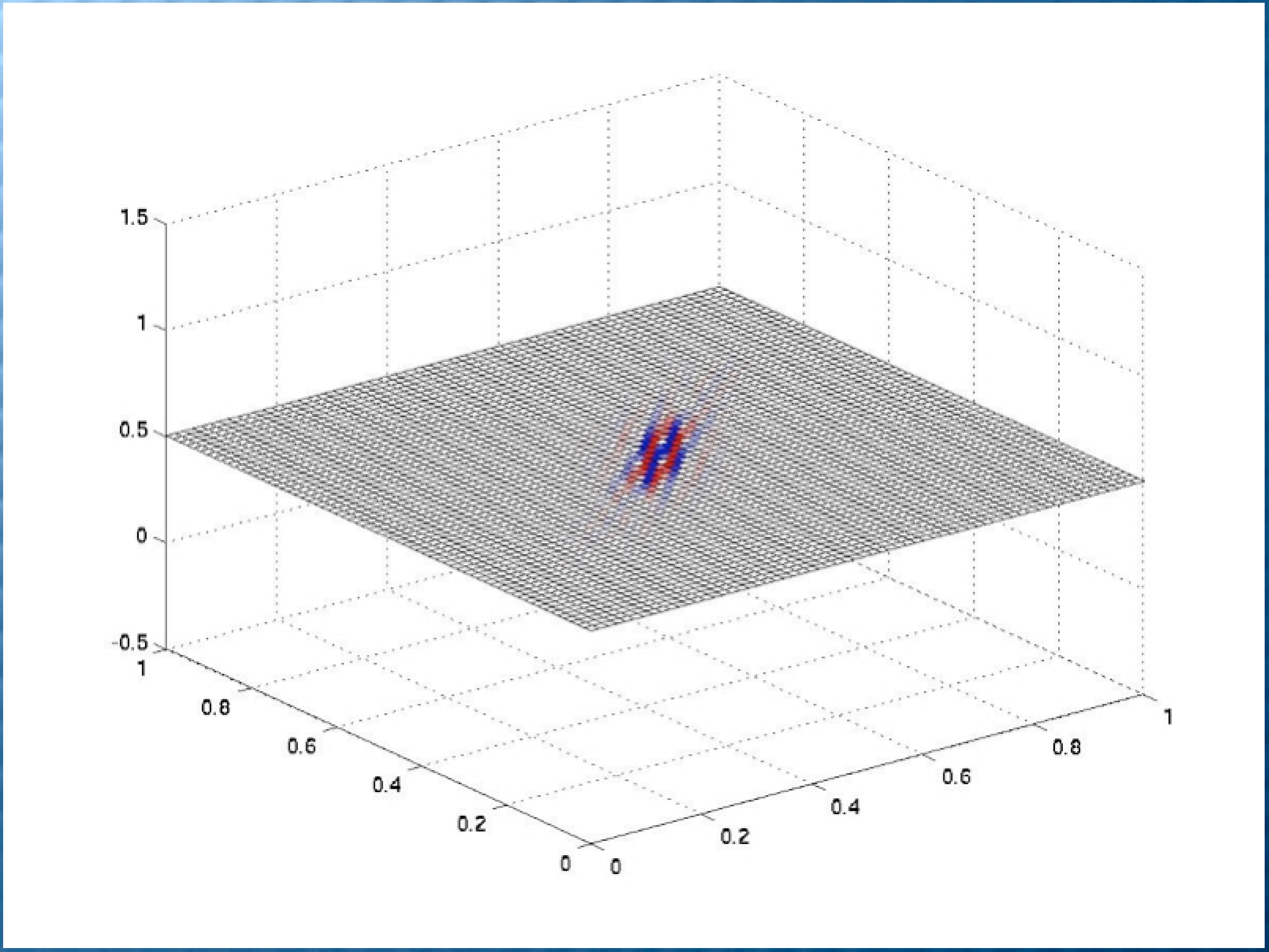
Thanks to Demanet & Ying

3-D Curvelets



Thanks to Demanet & Ying

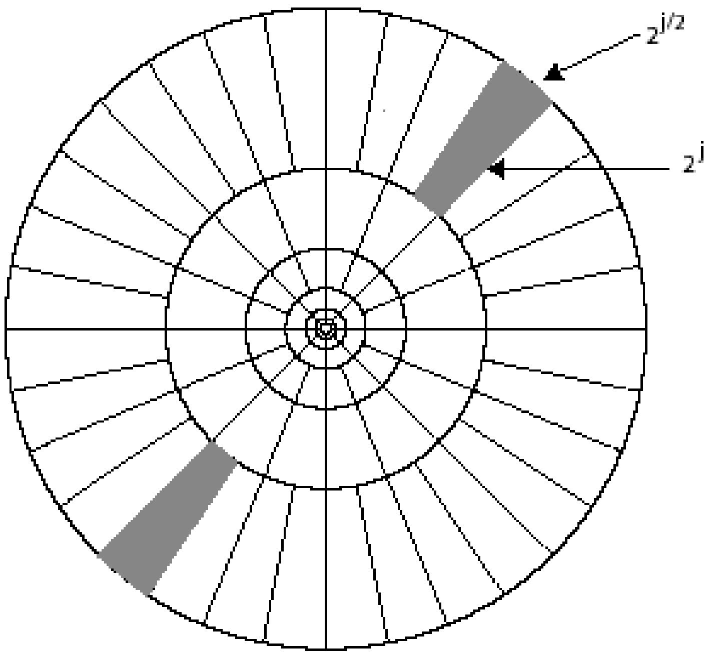
3-D Curvelets



Thanks to Demanet & Ying

Why curvelets

$$W_j = \{\zeta, \quad 2^j \le |\zeta| \le 2^{j+1}, |\theta - \theta_J| \le \pi \cdot 2^{\lfloor j/2 \rfloor}\}$$



second dyadic partitioning

Fourier/SVD/KL

$$||f - \tilde{f}_m^F|| \propto m^{-1/2}, \; m \to \infty$$

Wavelet

$$||f - \tilde{f}_m^W|| \propto m^{-1}, \; m \to \infty$$

Optimal data adaptive

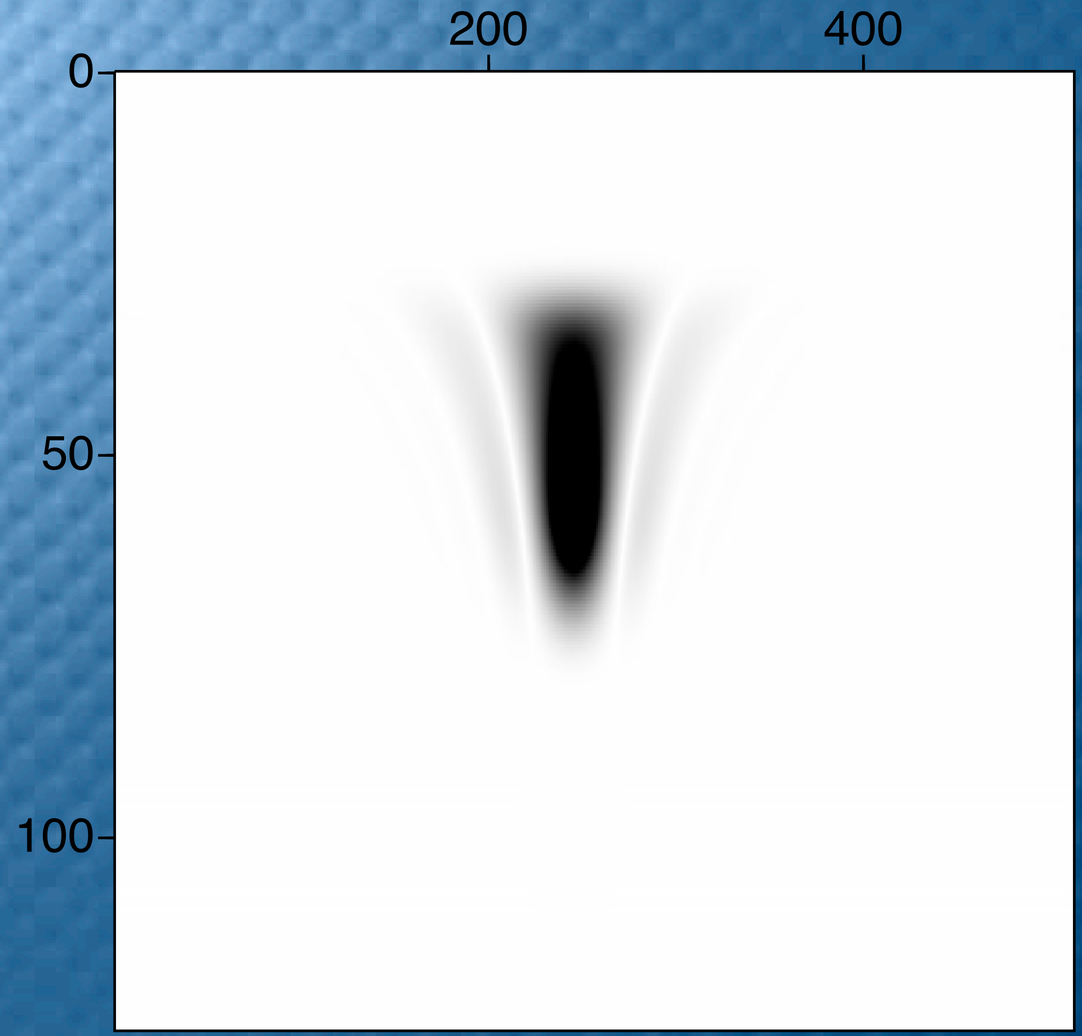
$$||f - \tilde{f}_m^A|| \propto m^{-2}, \; m \to \infty$$

Close to optimal Curvelet

$$||f - \tilde{f}_m^C|| \le C \cdot m^{-2} (\log m)^3, \; m \to \infty$$

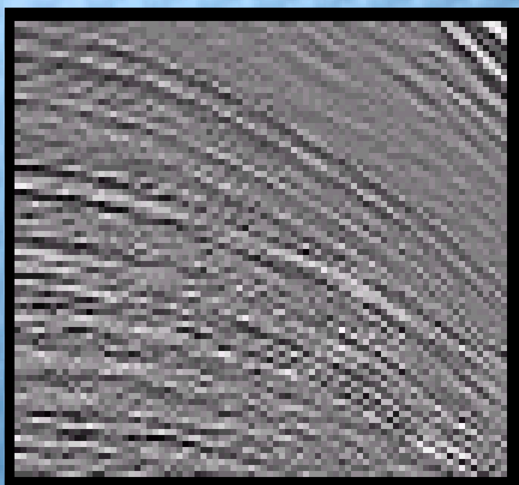
source: Candes'01, Stein '90

Why curvelets



Curvelet in FK-domain

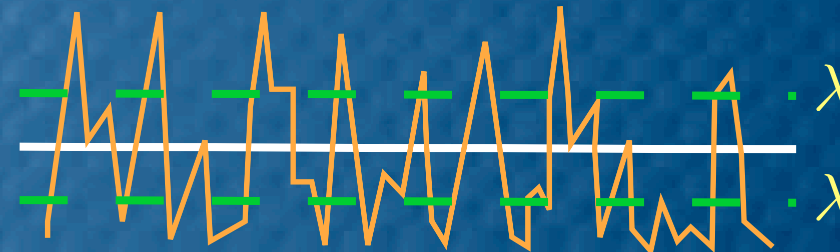
Denoising



Input data with noise

Curvelet transform

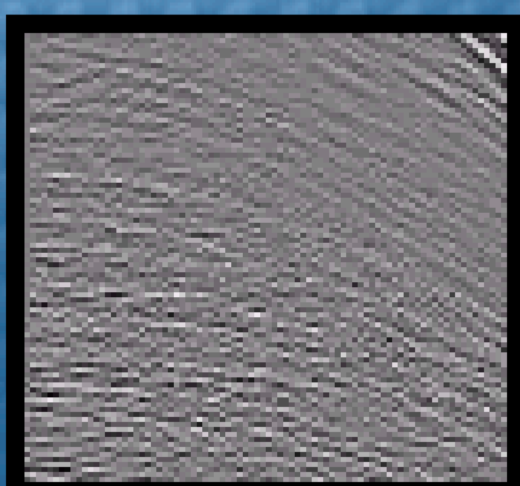
$\mathbf{Bd}$



Curvelet coeff. data

Threshold

$S_\lambda(\mathbf{Bd})$

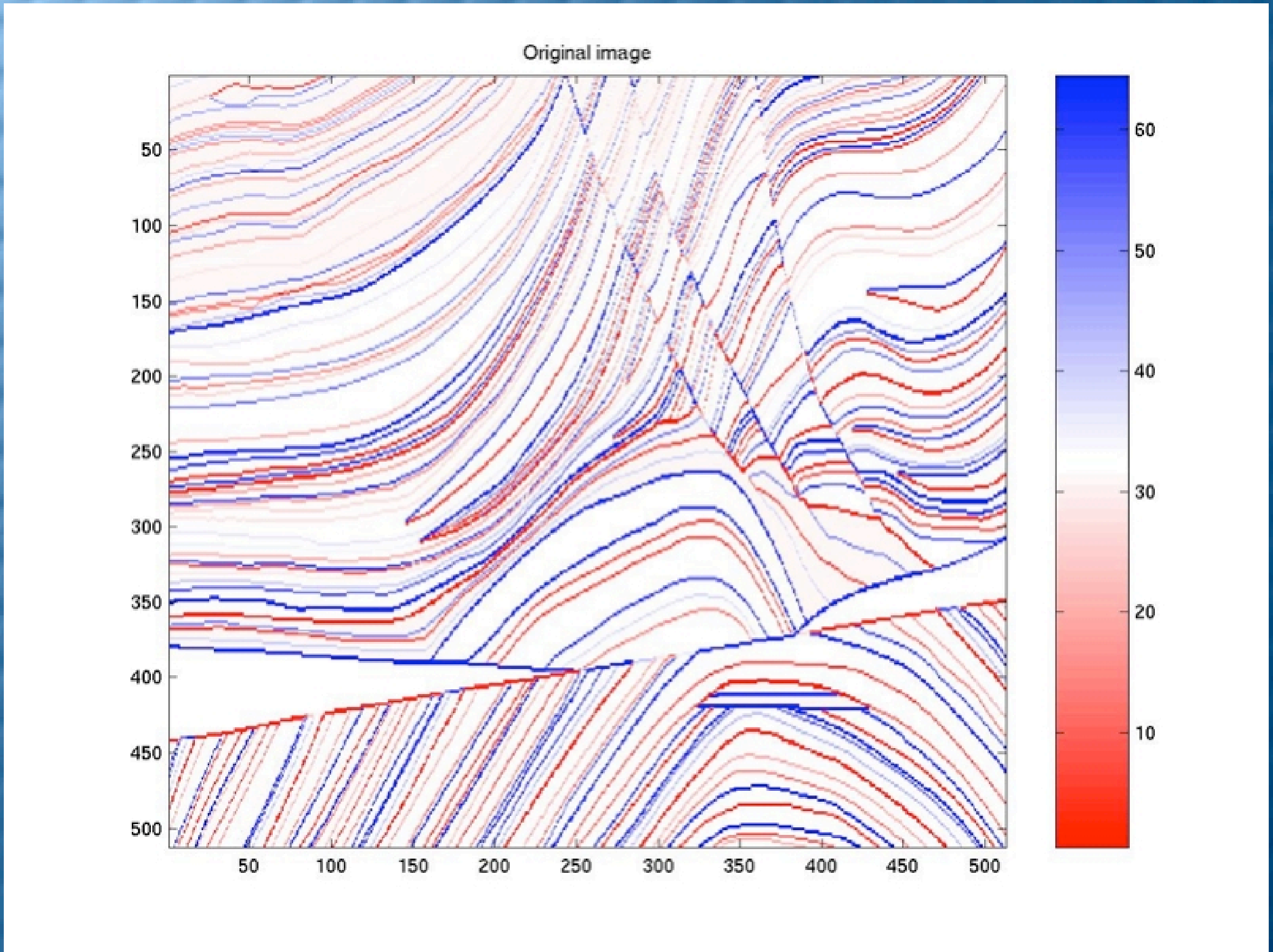


Filtered input data

Inv. curvelet transform

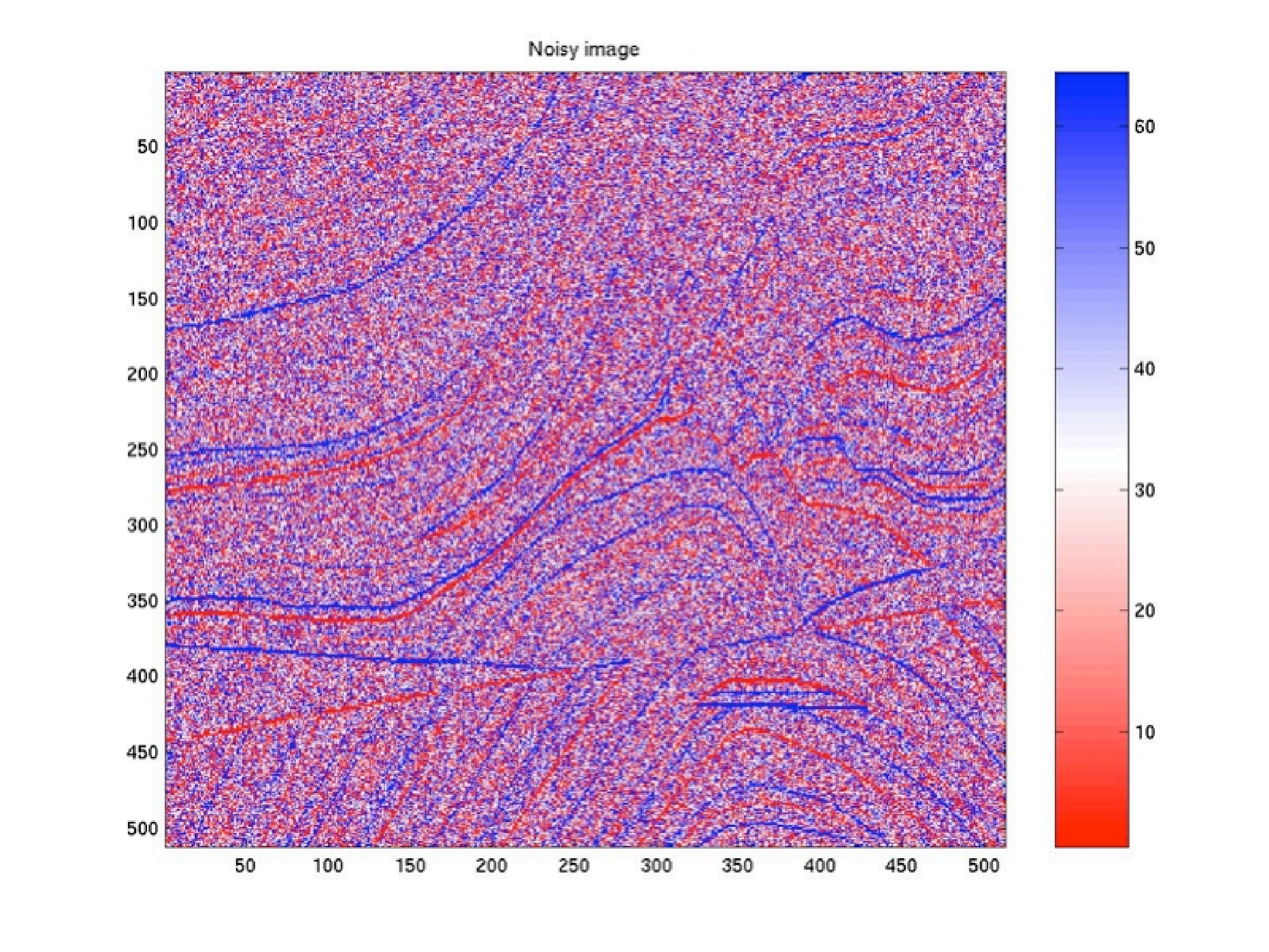
denoised Curvelet coeff.

Denoising

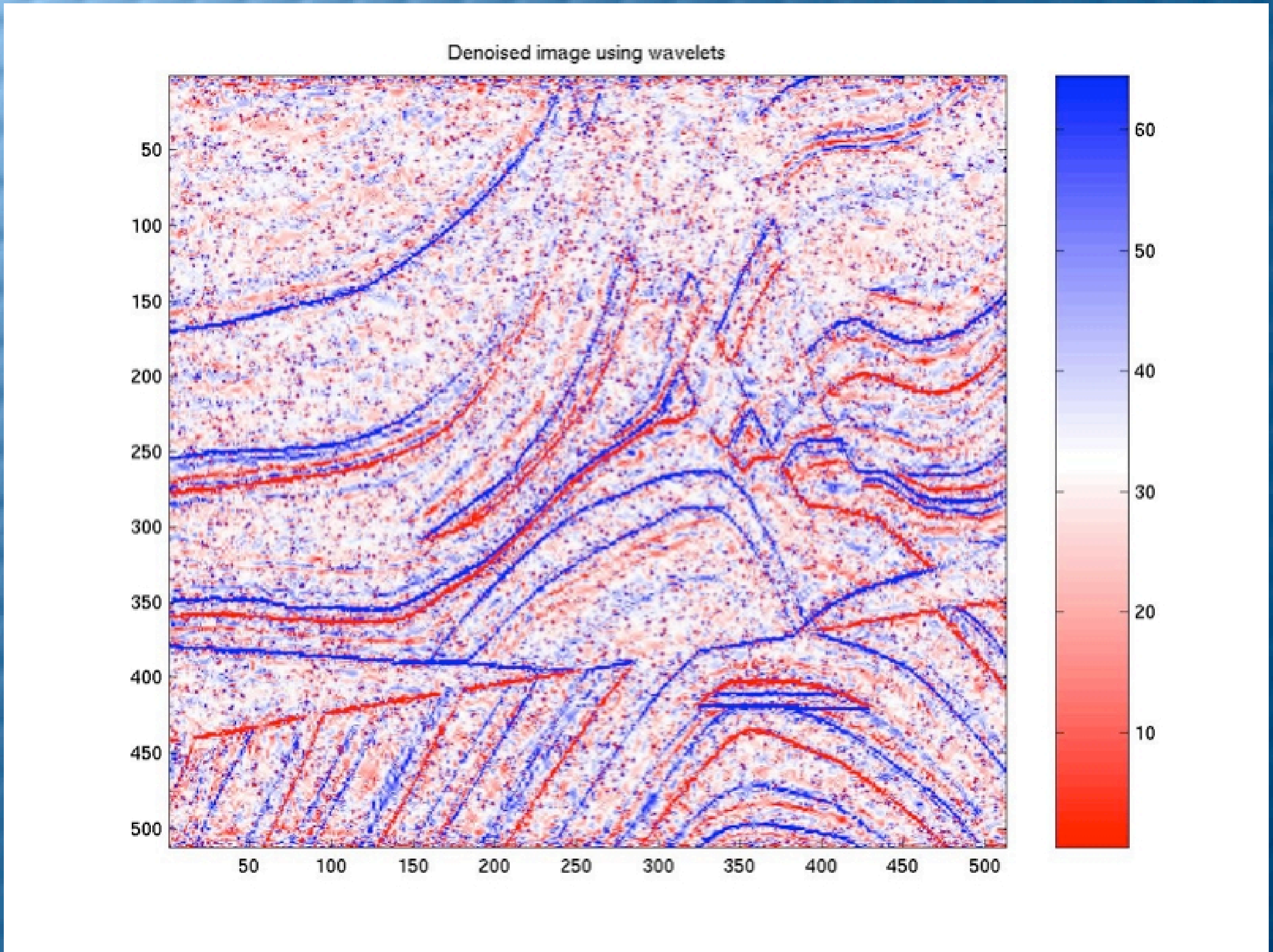


Original image
50
100
150
200
250
300
350
400
450
500
60
40
30
20
10

Denoising

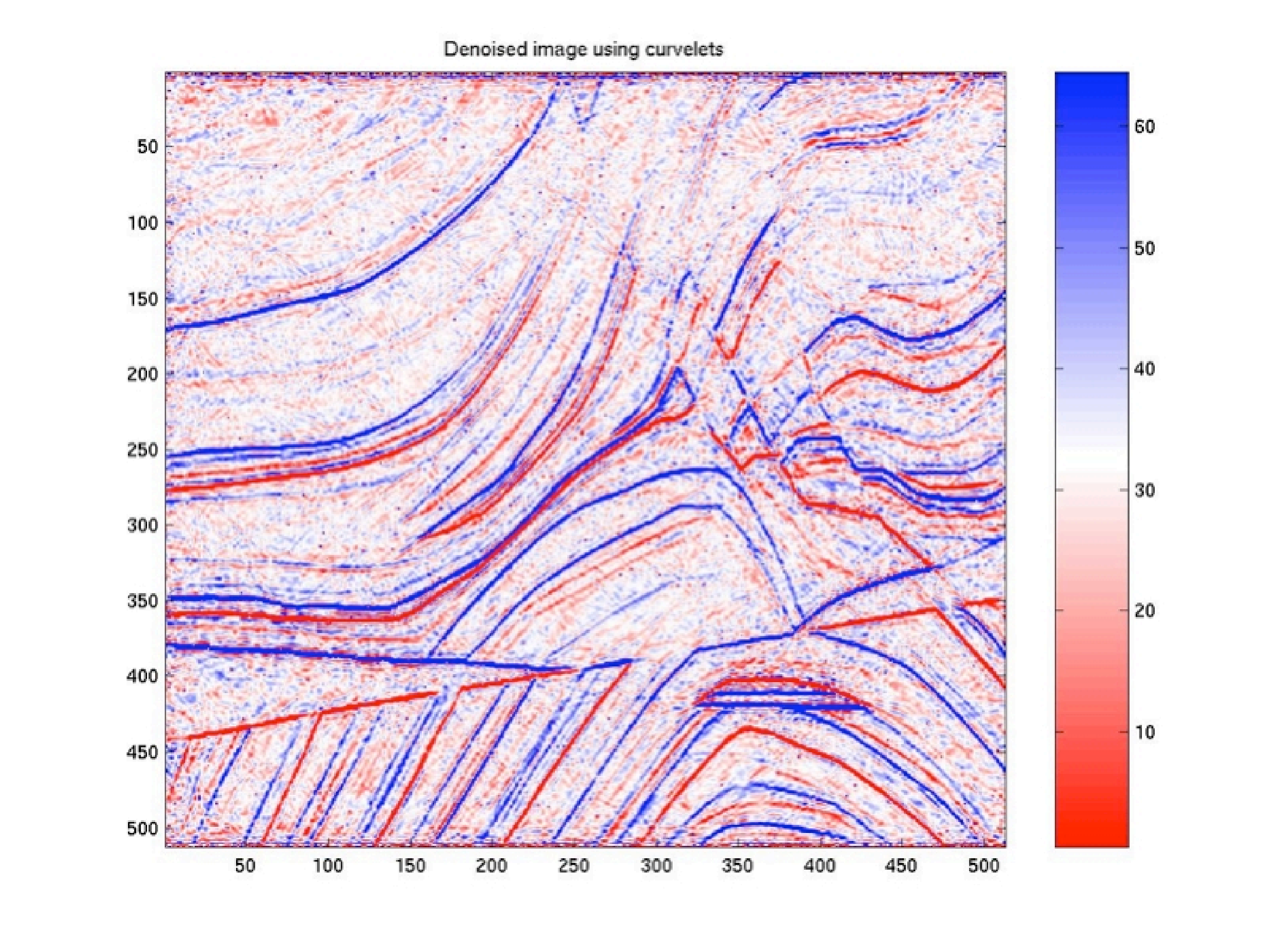

Noisy image
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

Wavelets

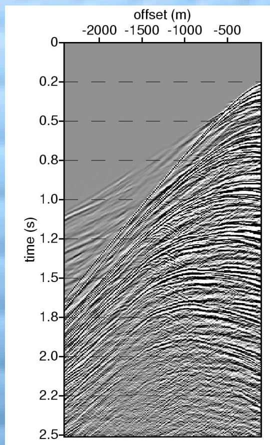


Denoised image using wavelets
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

Curvelets

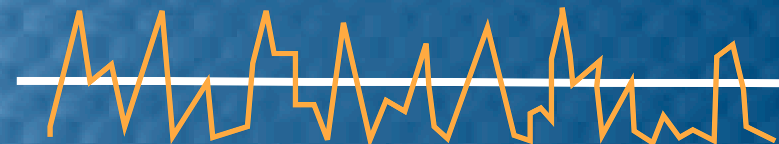

Denoised image using curvelets
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

Coherent noise suppression with curvelets



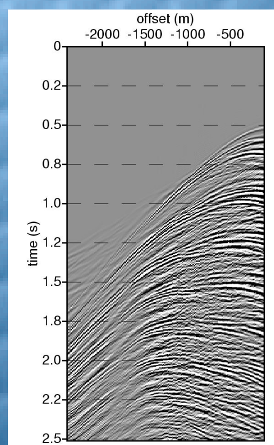
Input data with noise

Curvelet transform



Curvelet coeff. data

$\mathbf{Bd}$



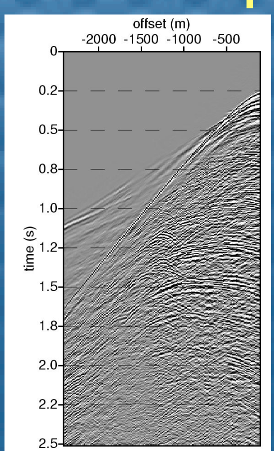
predicted noise

Threshold

Γ

$|\mathbf{B}\hat{\mathbf{n}}|$

Curvelet coeff. pred. noise



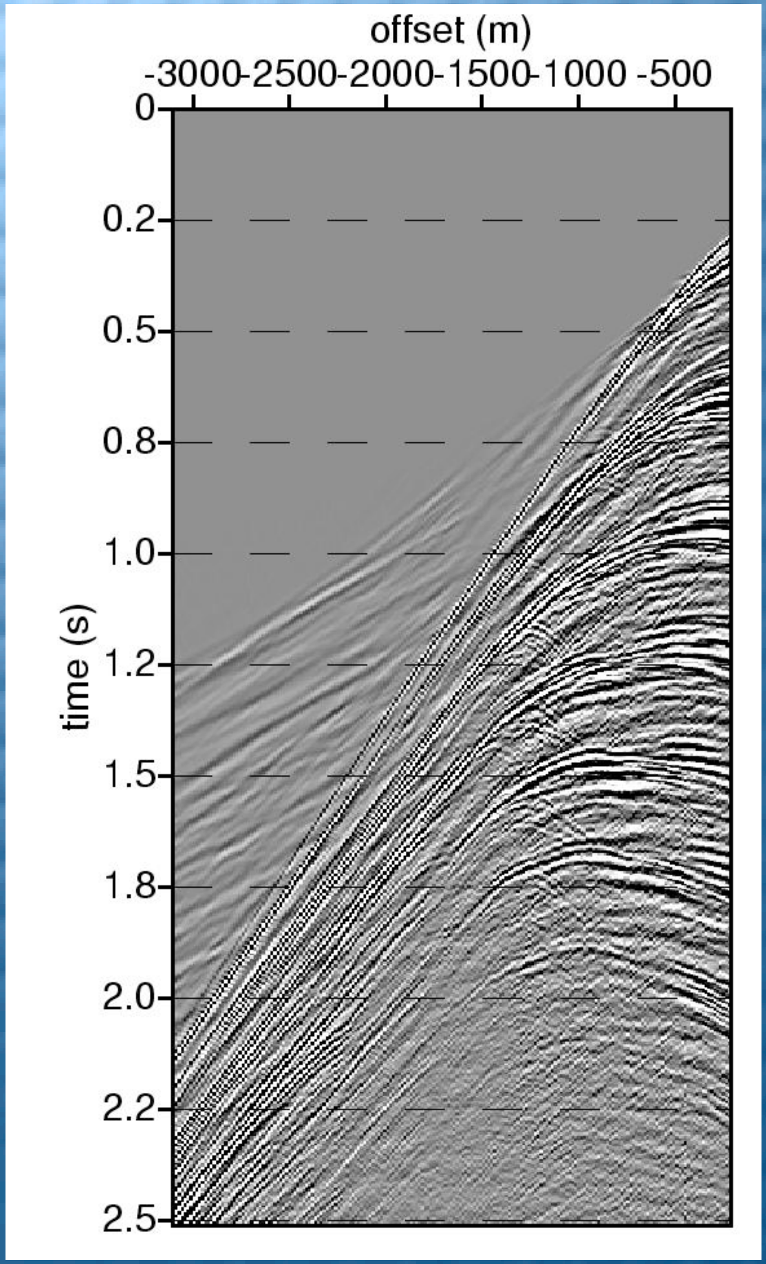
Filtered input data

Inv. curvelet transform

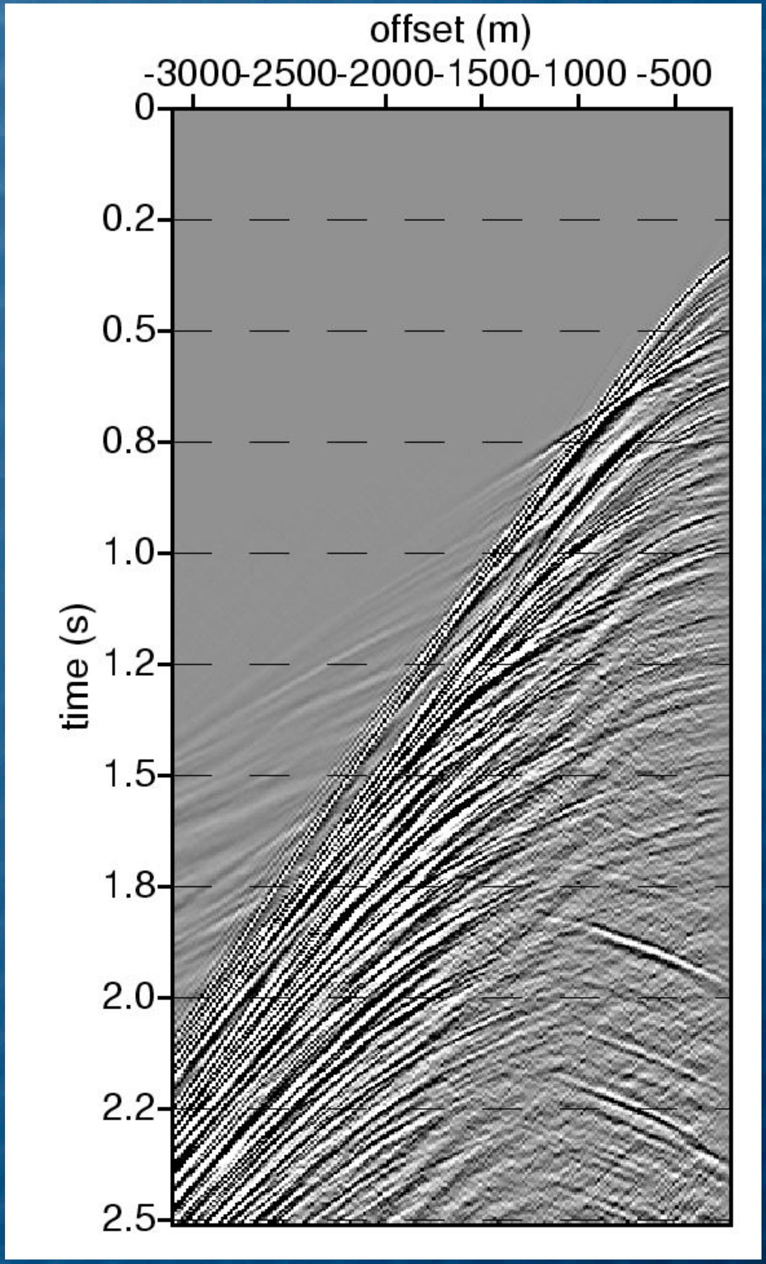
$S_{\lambda\Gamma}(\mathbf{Bd})$

Curvelet coeff. model

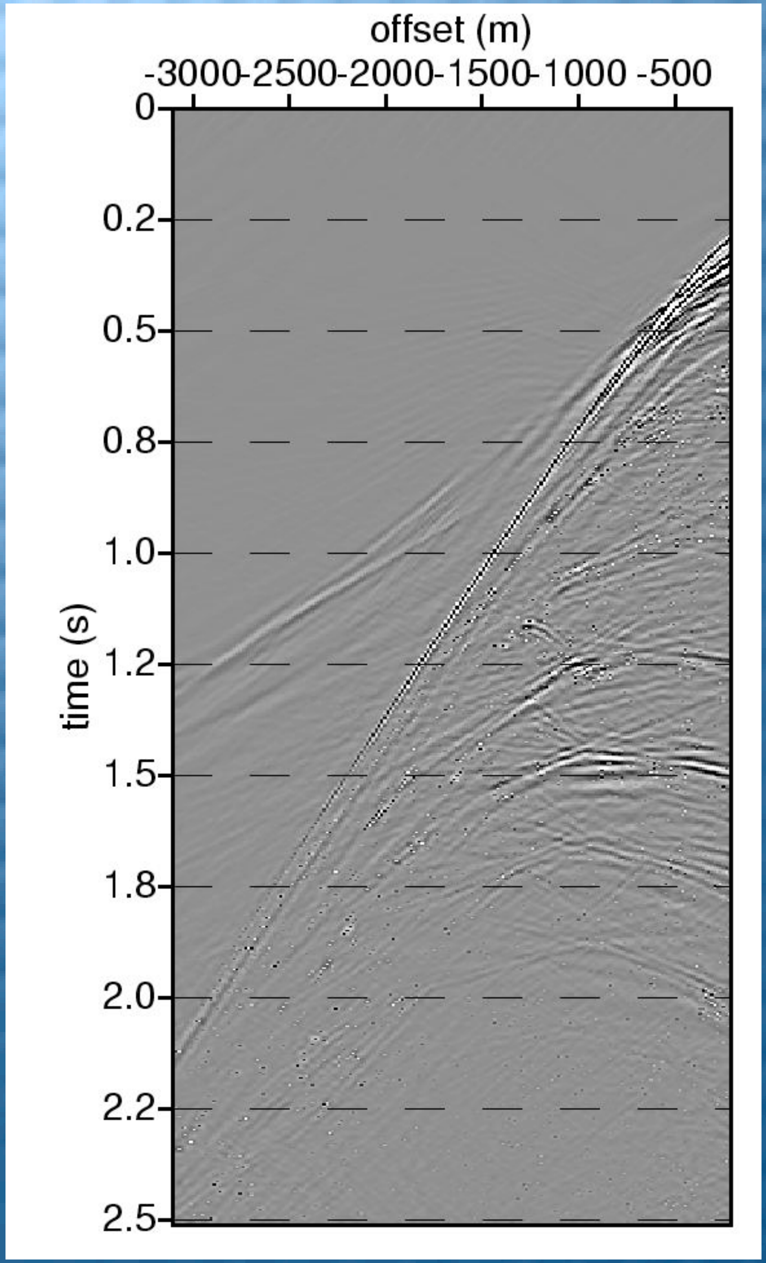
Multiple suppression with curvelets



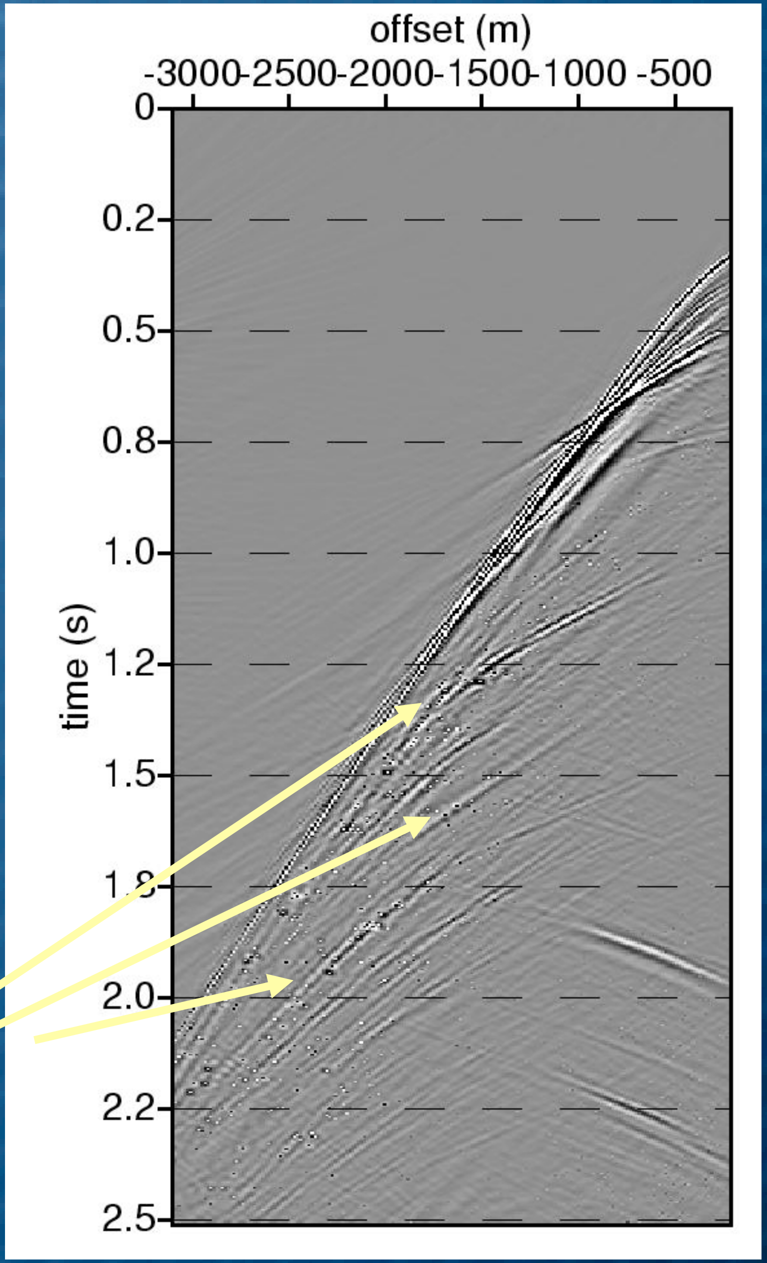
Input with multiples



Multiple suppression with curvelets



Output curvelet filtering with stronger threshold



Preserved primaries

Denoising

Denoising ⇔ *signal separation*

- **Multiple & Ground-roll removal**
- **Migration denoising (H & M '04)**
- **4-D difference cubes**

Model "noise"

Strategies:

- **Weighted thresholding**
- **Iterated weighted thresholding**

L2/L1-matched filter

Matched filter:

$$\overbrace{\hat{\mathbf{n}}}^{\text{denoised}} : \min_{\mathbf{\Phi}} = \| \underbrace{\mathbf{d}}_{\text{noisy data}} - \overbrace{\mathbf{\Phi}}^{\text{matched filter}} \overset{t}{*} \underbrace{\mathbf{m}}_{\text{pred. noise}} \|_p$$

- **p=1 enhances sparseness**
- **residue is the denoised data**
- **risk of over fitting**

Loose primary reflection events ...

Guitton&
Verschuur'04

Colored denoising

$$\overbrace{\mathbf{d}}^{\text{noisy data}} = \underbrace{\mathbf{m}}_{\text{noise-free}} + \overbrace{\mathbf{n}}^{\text{col. noise}}$$

Denoising:

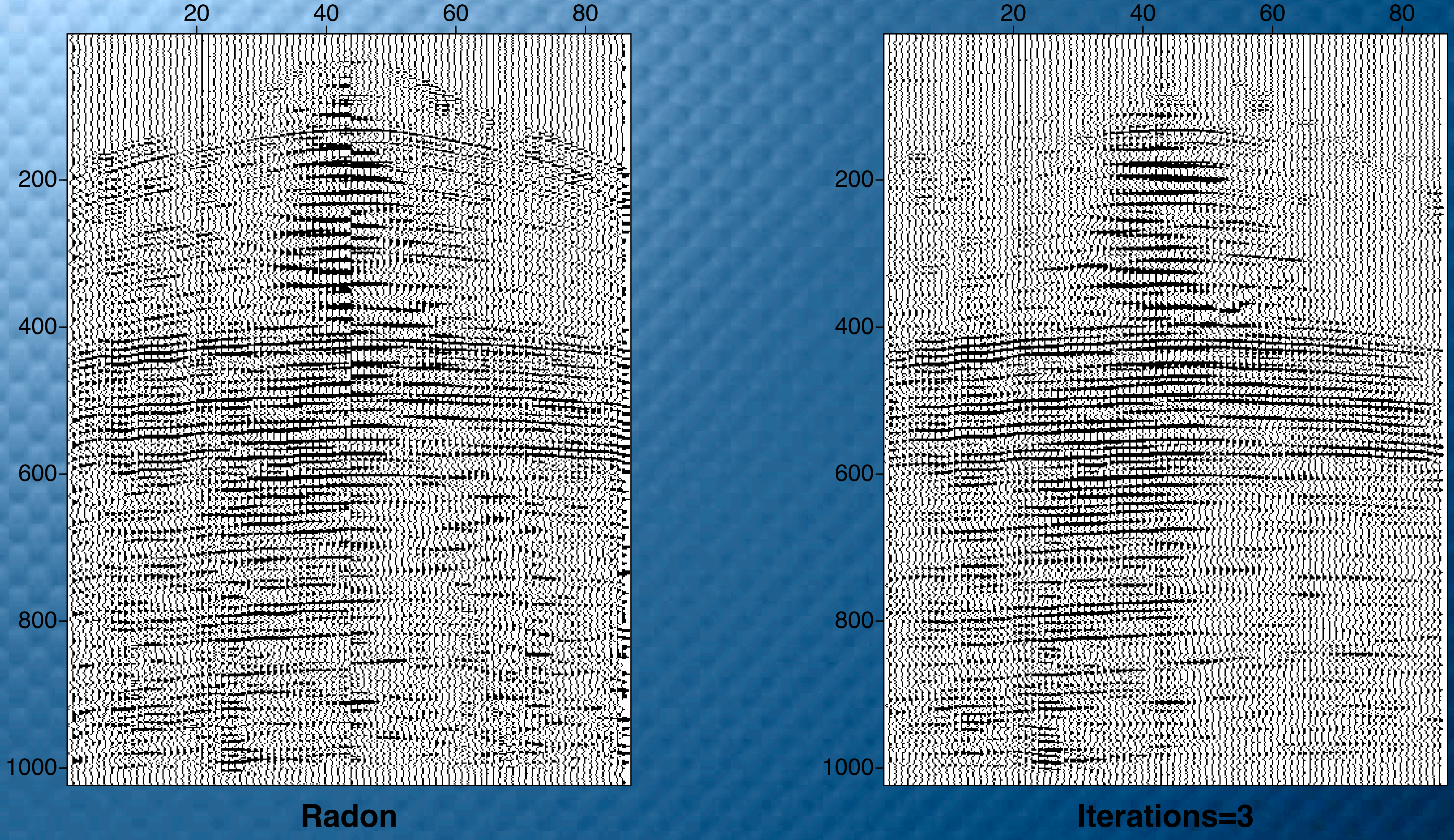
$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} \frac{1}{2}\|\mathbf{C}_n^{-1/2}(\mathbf{d}-\mathbf{m})\|_2^2 + J(\mathbf{m})$$

with covariance

$$\mathbf{C}_n \equiv \mathbf{E}\{\mathbf{n}\mathbf{n}^T\}$$

and both m, n **related to PDE**

Ground-roll removal with curvelets



Weighted thresholding

Covariance model & noise near diagonal:

$$\mathbf{B}\mathbf{C}_{n\,\mathrm{or}\,m}\mathbf{B}^T \approx \mathbf{\Gamma}^2 \quad \text{near diagonal}$$

For ortho basis and app. noise prediction:

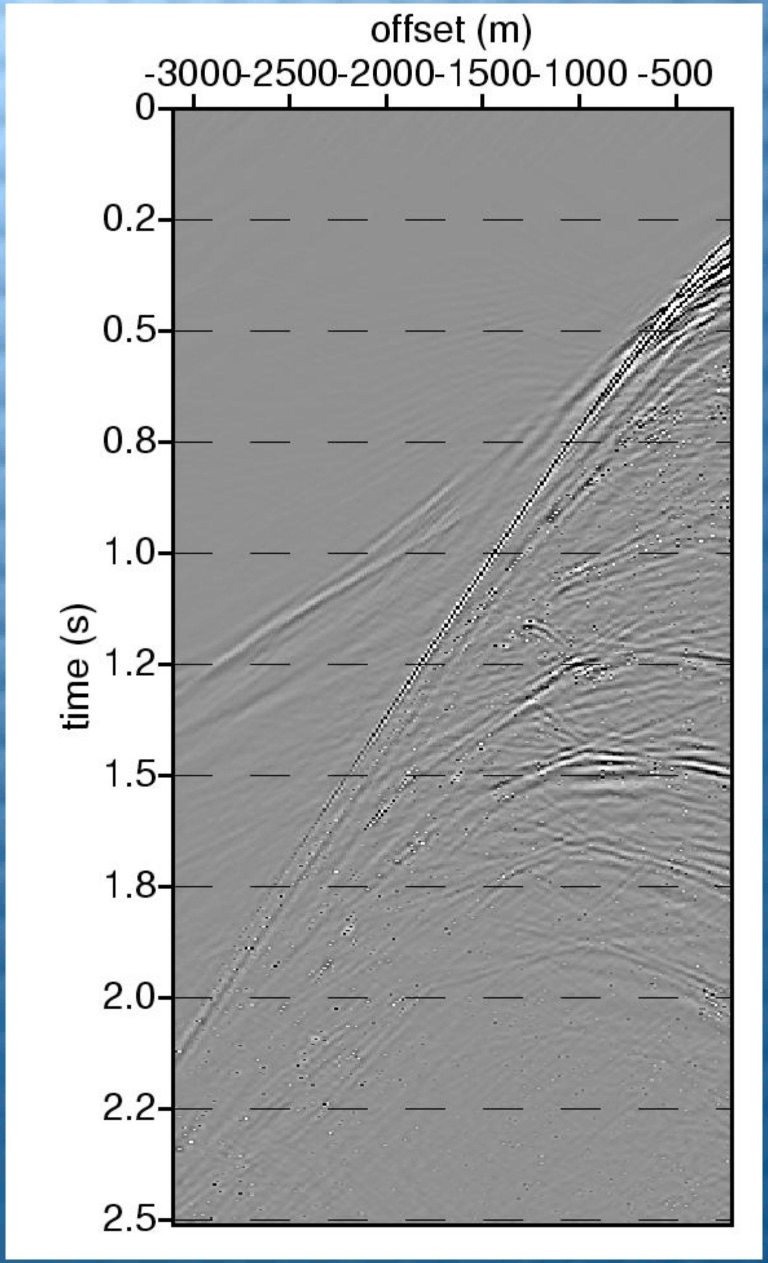
$$\boxed{\hat{m} = \mathbf{B}^T S_{\lambda\Gamma}\left(\mathbf{Bd}\right)}$$

equivalent to

$$\boxed{\hat{\mathbf{m}} = \mathbf{B}^T \arg\min_{\tilde{m}} \frac{1}{2}\|\mathbf{\Gamma}^{-1}\left(\tilde{\mathbf{d}} - \tilde{\mathbf{m}}\right)\|_2^2 + \|\tilde{\mathbf{m}}\|_{1,\lambda}}$$

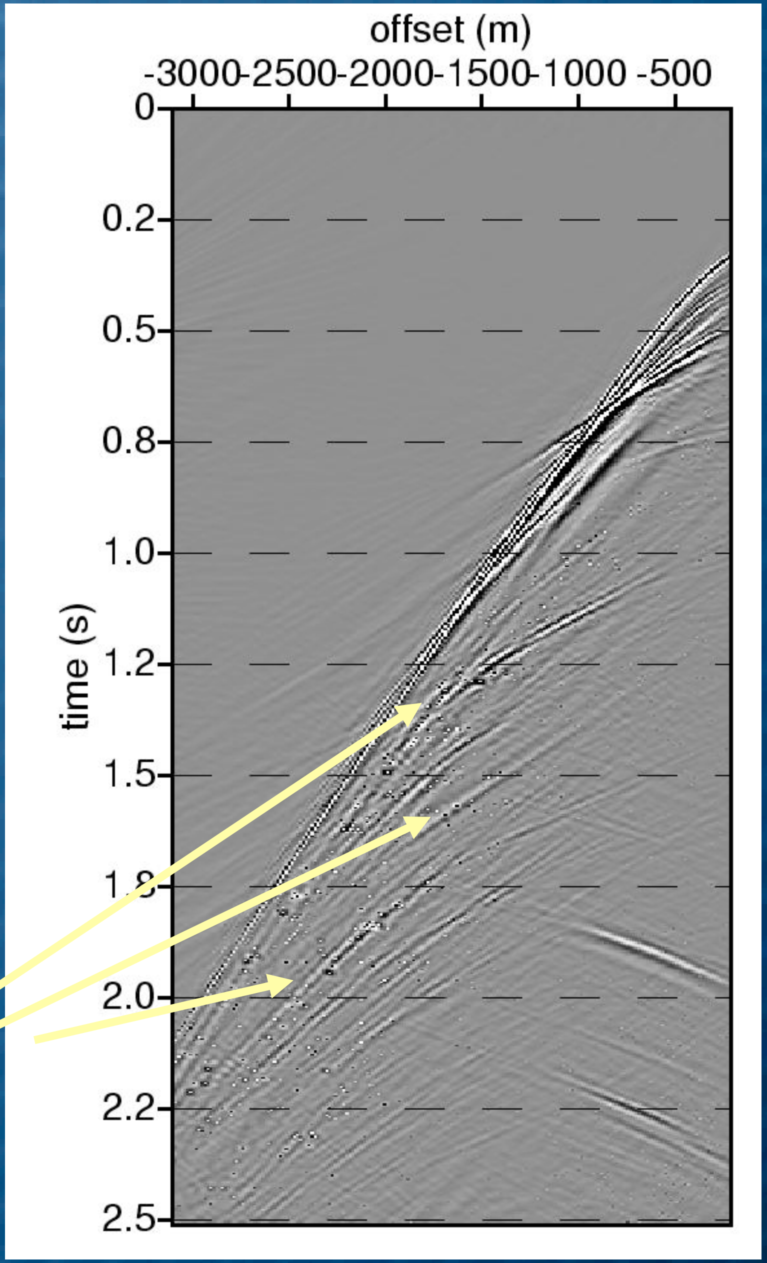
$$\tilde{\mathbf{d}} = \mathbf{Bd}, \quad \tilde{\mathbf{m}} = \mathbf{Bm} \quad \text{and} \quad \mathbf{\Gamma} = [\mathrm{diag}\{\mathrm{diag}\{\mathbf{B}\hat{\mathbf{n}}\}\}]^{1/2}$$

Multiple suppression with curvelets

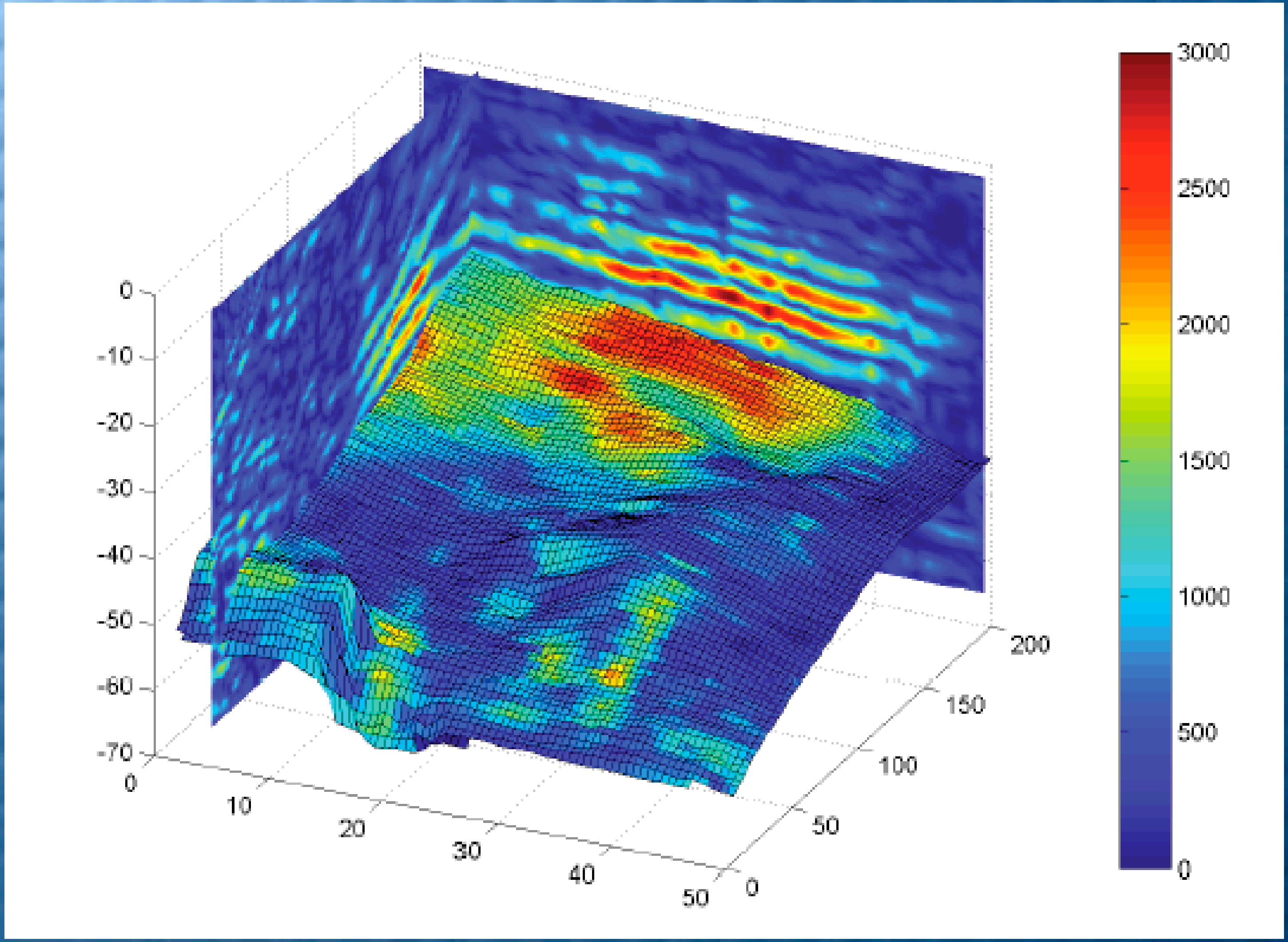


Output curvelet filtering with stronger threshold

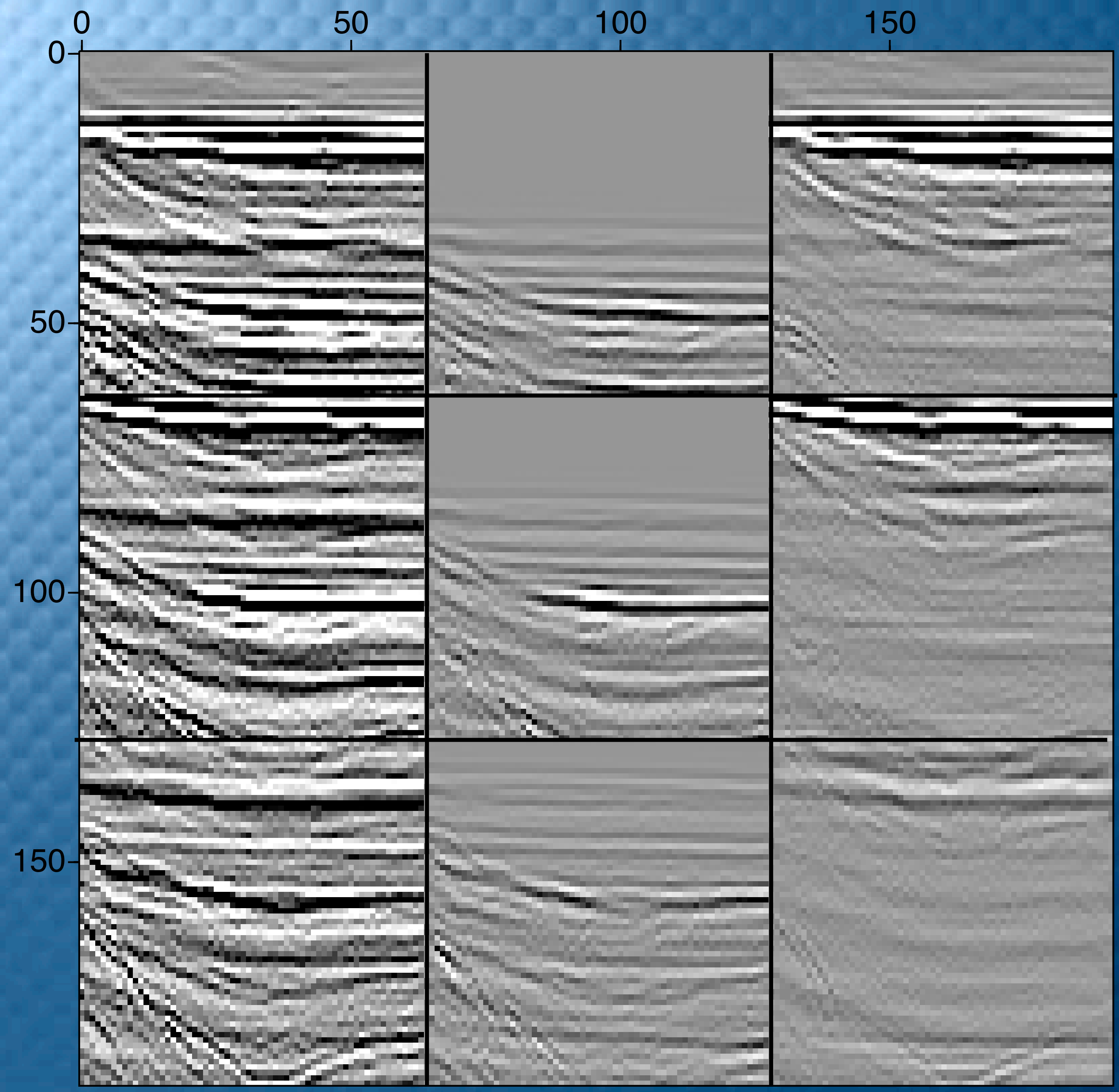
Preserved primaries



4-D difference



3-D Curvelets



Iterative thresholding

Curvelets are Frames:

- **redundant (factor 7.5-4)**
- **thresholding does *not* solve:**

$$\hat{\mathbf{m}} = \mathrm{B}^T \arg\min_{\tilde{m}} \frac{1}{2} \|\boldsymbol{\Gamma}^{-1} \left(\tilde{\mathbf{d}} - \tilde{\mathbf{m}}\right) \|_2^2 + \|\tilde{\mathbf{m}}\|_{1,\lambda}$$

Alternative formulation by iterative thresholding!

Iterative thresholding

Solves signal-separation problem:

$$\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n} \quad \text{and} \quad \mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$$

by minimizing LP-program

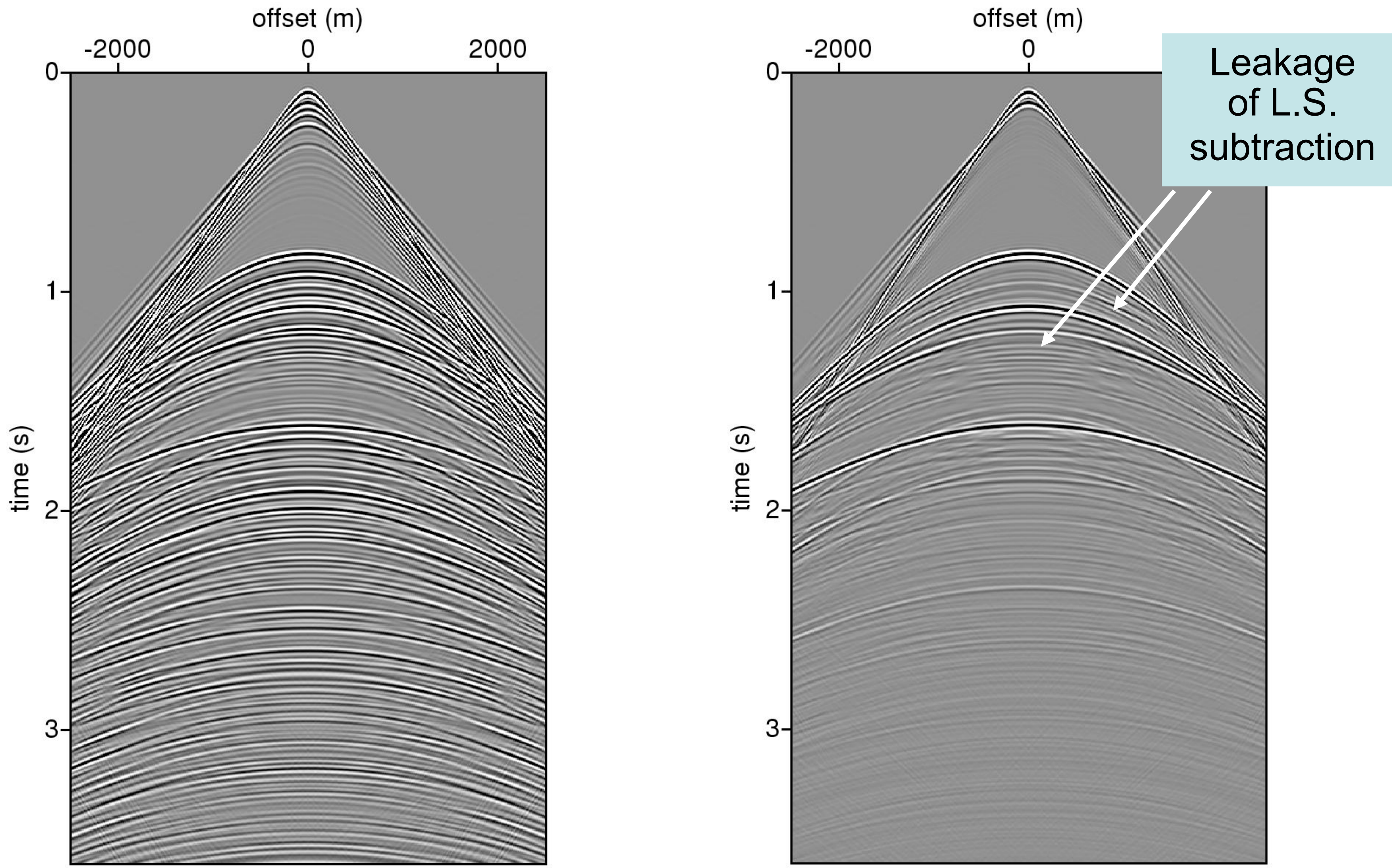
$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{s} - \mathbf{\Phi}\mathbf{x}\|_2^2 + \|\mathbf{x}_1\|_{w_1,1} + \|\mathbf{x}_2\|_{w_2,1}$$

with

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{bmatrix}^T \quad \text{and} \quad \mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 \end{bmatrix}$$

Starck *et al.;* Daubechies *et al., Donoho 2004*
(also Zwartjes, Sacchi, Trad & Ulrich)

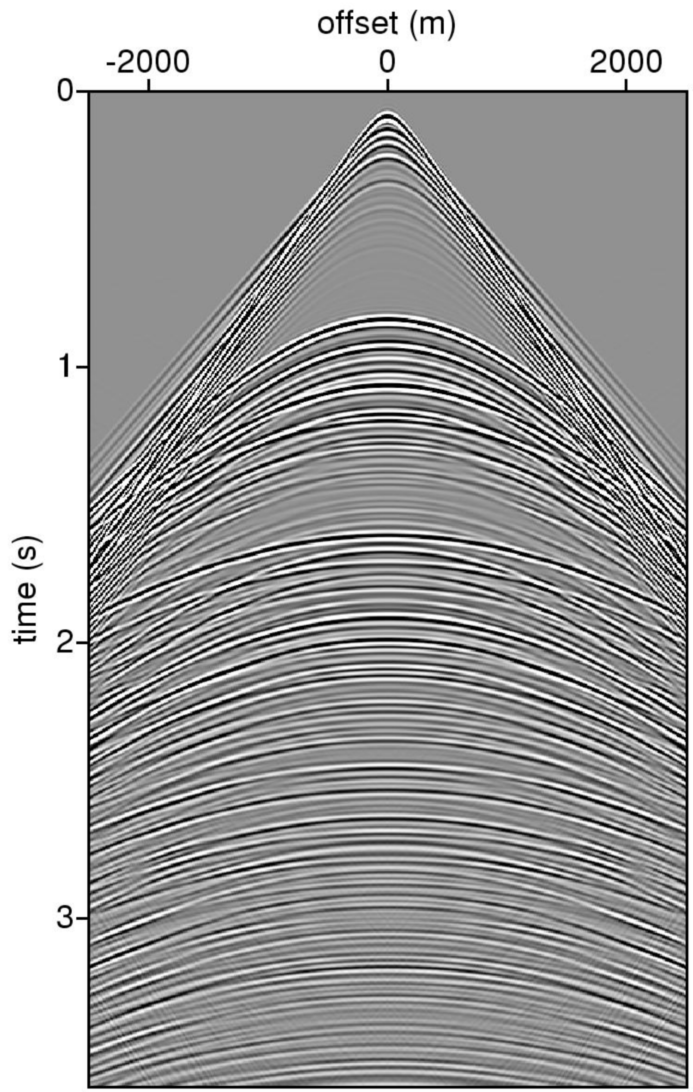
Example shallow water environment



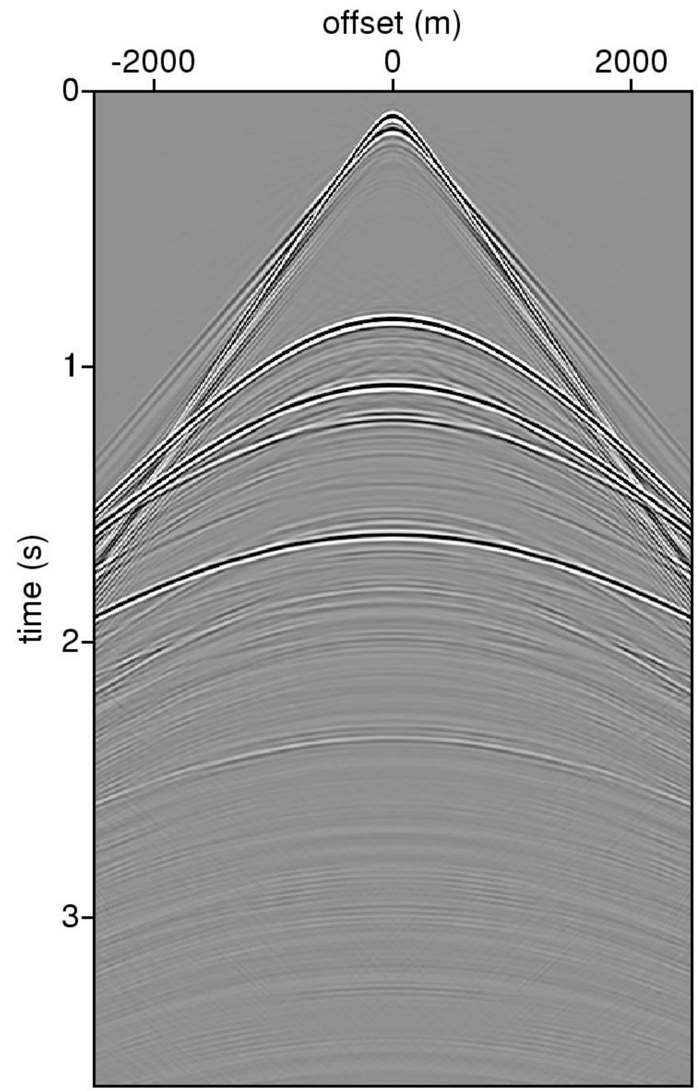
Input data with multiples

Result L.S. subtraction

Example shallow water environment

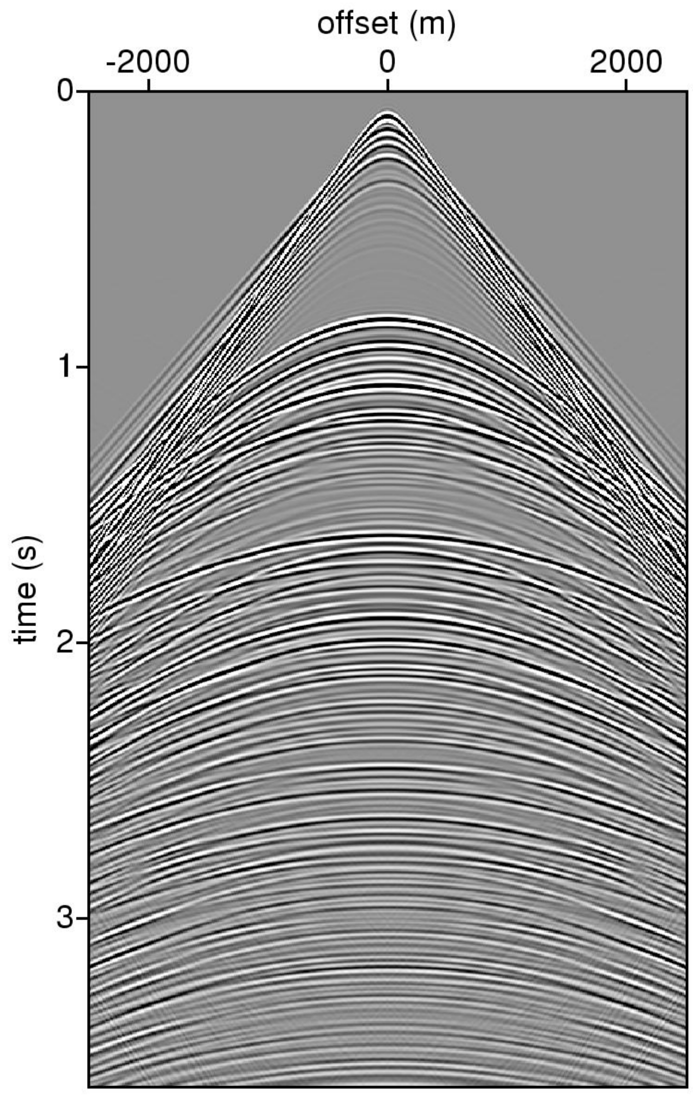


Input data with multiples

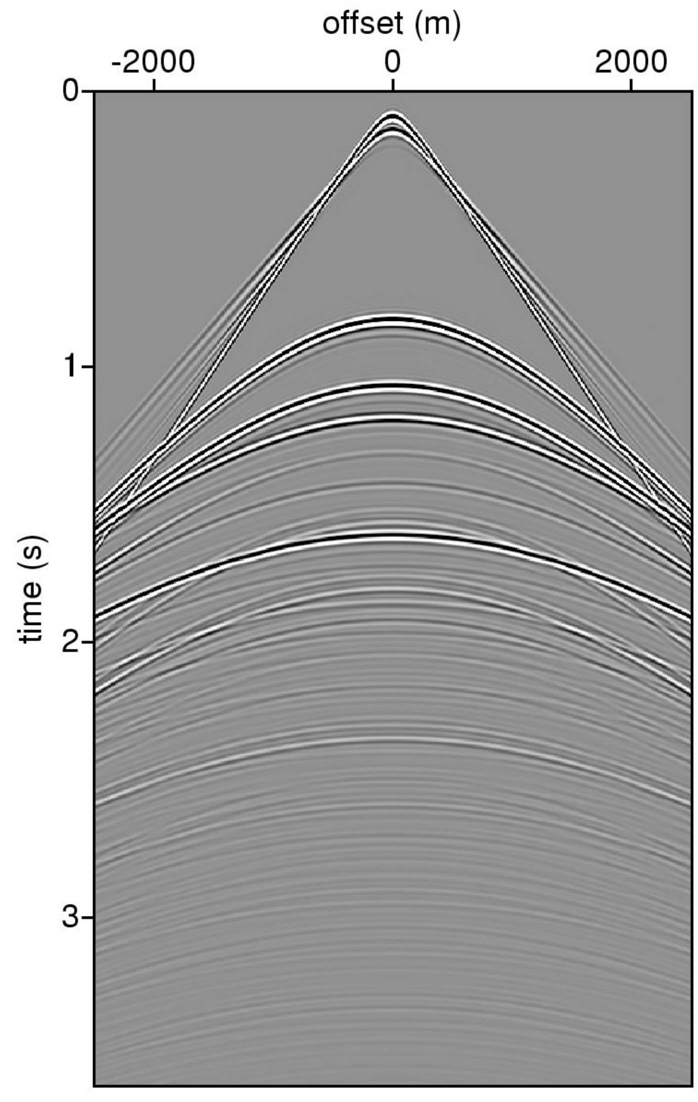


Result curvelet-based subtraction

Example shallow water environment

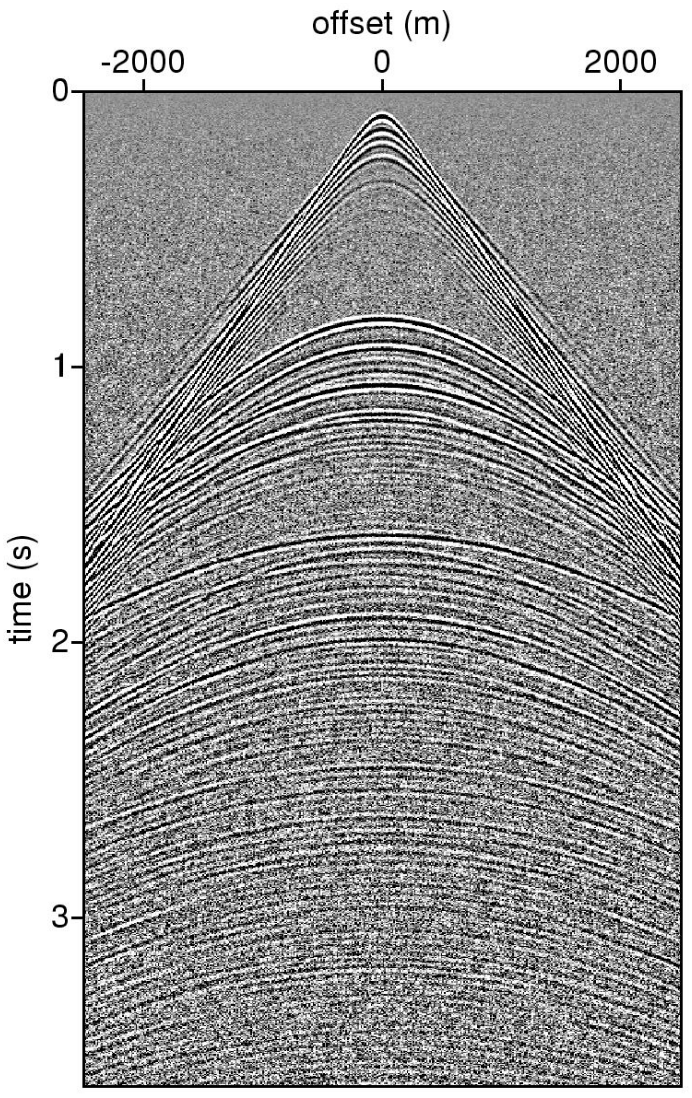


Input data with multiples

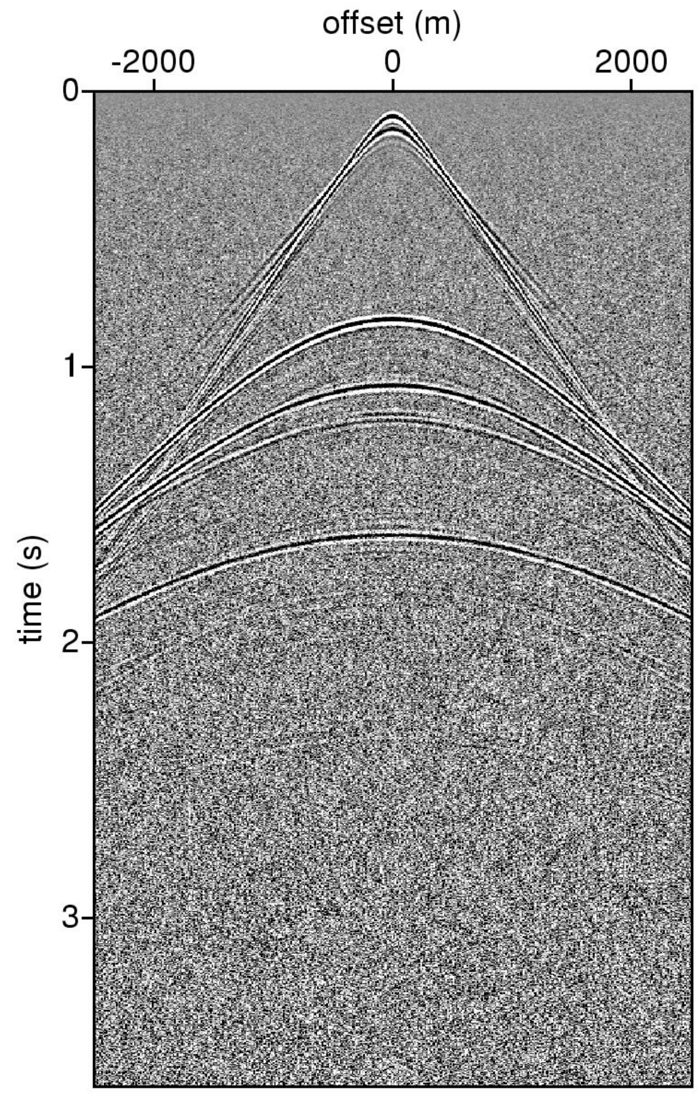


True primaries

Example shallow water environment



Input data with multiples and noise



Result curvelet-based subtraction

Global optimization

After thresholding:

- **remove artifacts & ‘miss fires’**
- **normal operator (inversion)**

Impose additional *penalty functional*

- ***prior* information**
- ***sparseness & continuity***

Defining the right norm is crucial ...

Global optimization

Formulate constrained optimization:

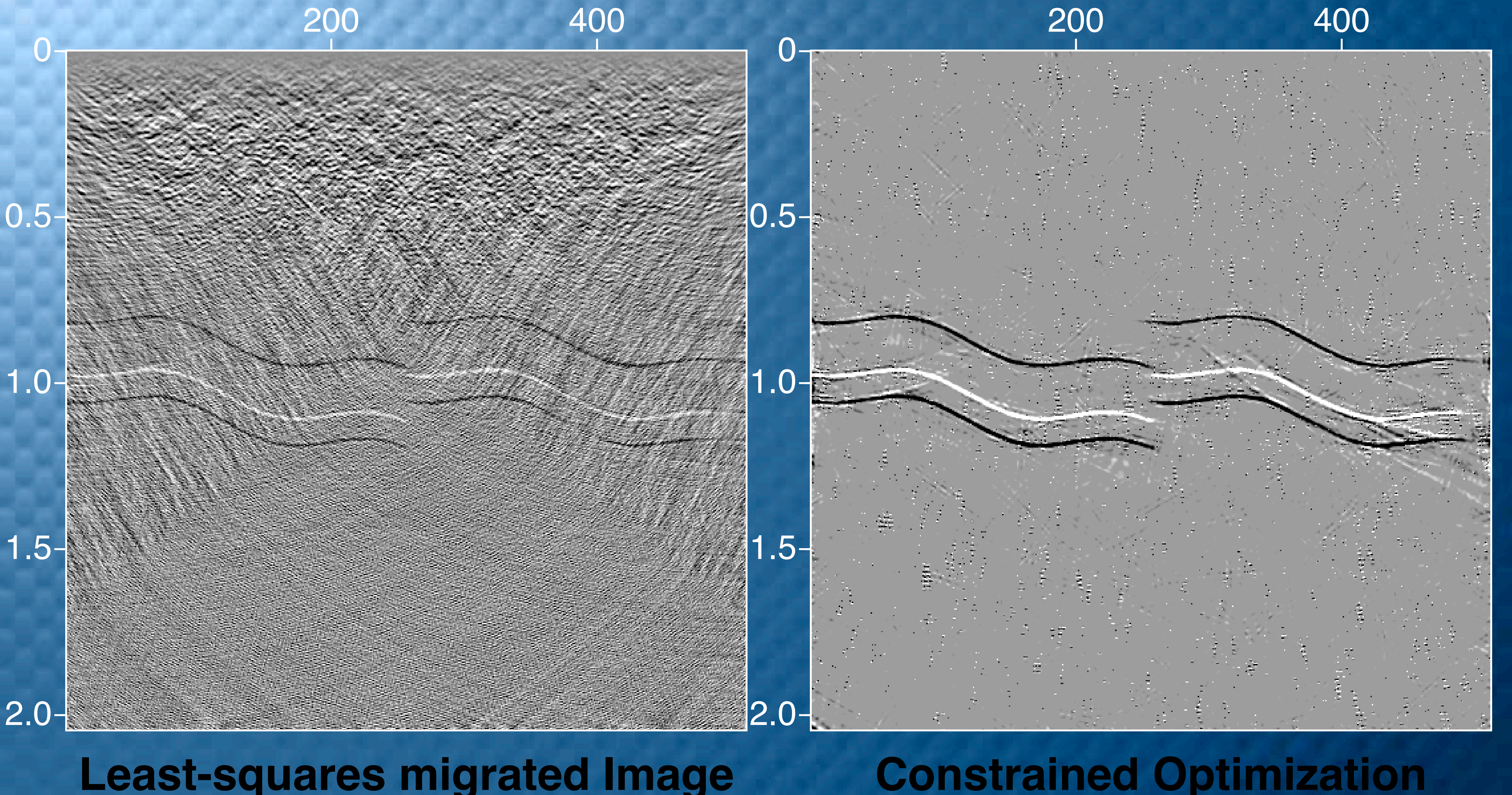
$$\hat{\mathbf{m}} : \quad \min_{m} J(\mathbf{m}) \quad \text{s.t.} \quad |\tilde{\mathbf{m}} - \hat{\tilde{\mathbf{m}}}_0|_\mu \leq \mathbf{e}_\mu, \quad \forall \mu$$

with

$$\hat{\mathbf{m}}_0 = \mathbf{B}^\dagger \Theta_{\boldsymbol{\lambda}\boldsymbol{\Gamma}} \left(\tilde{\mathbf{d}} \right)$$

and with e_μ threshold and noise-dependent *tolerance* on curvelet coeff.

Imaging with Curvelets



Penalty functionals

Anisotropic diffusion for imaging:

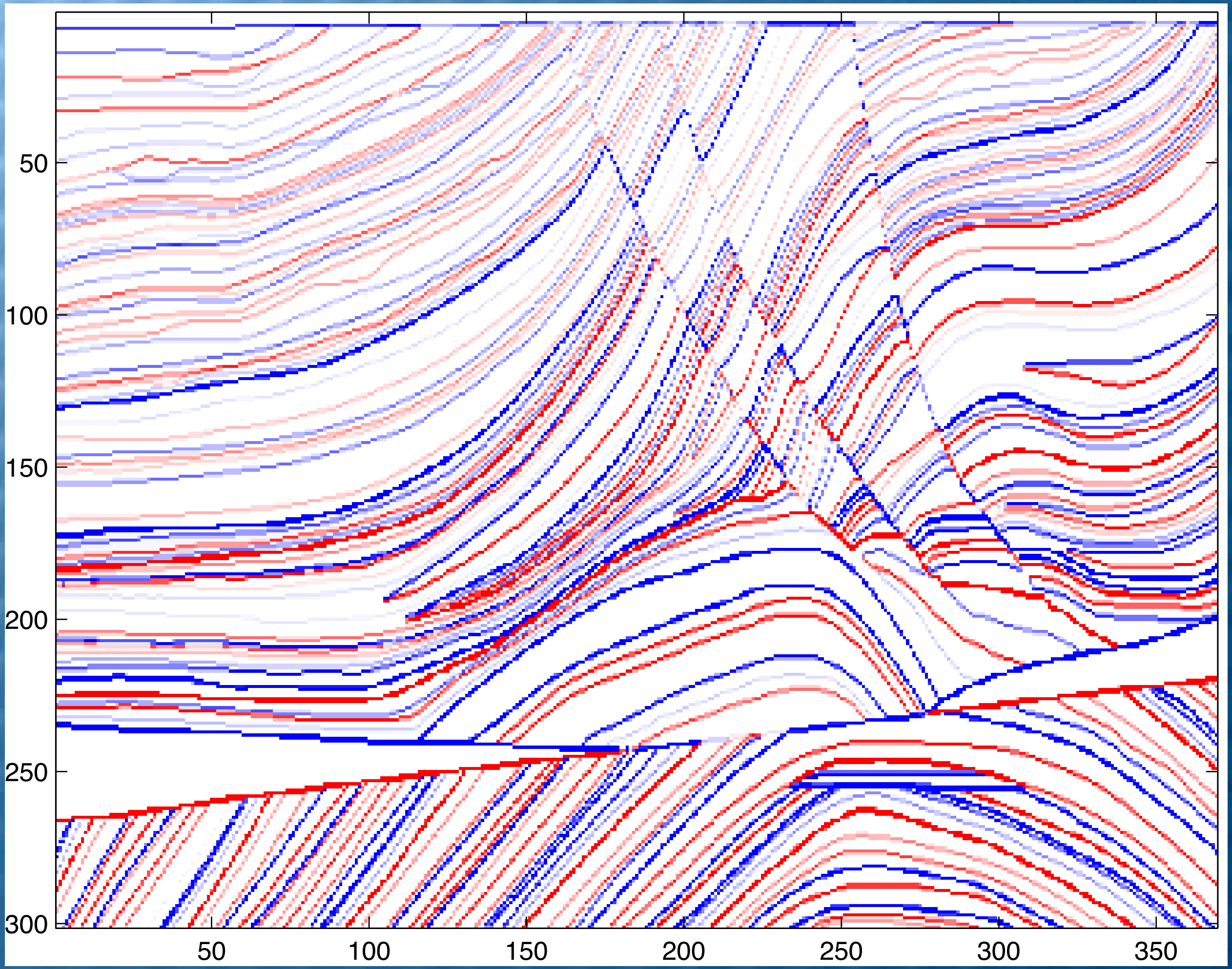
$$J(\mathbf{m}) = \|\mathbf{\Lambda}^{1/2}\mathbf{\nabla}\mathbf{m}\|_p$$

with

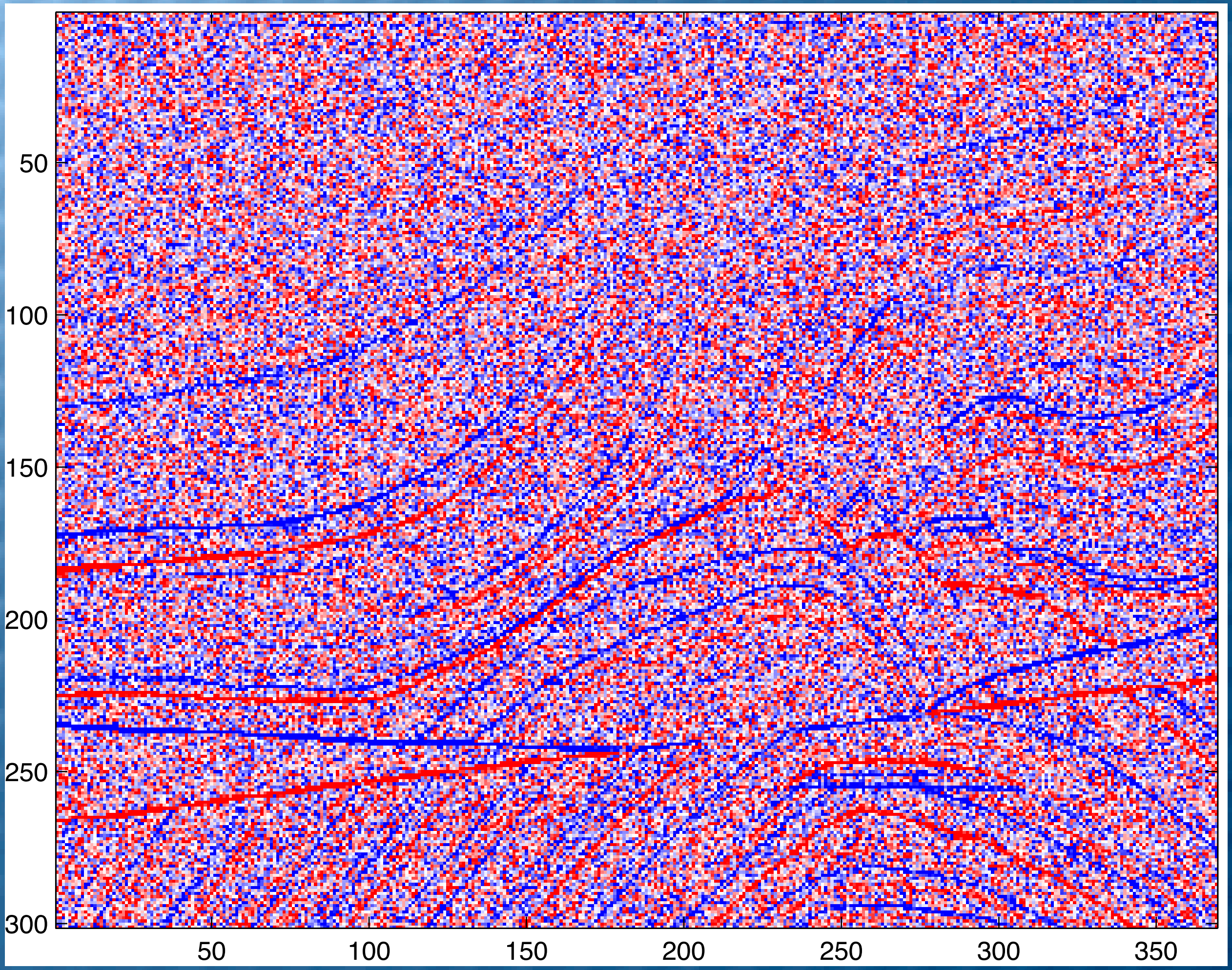
$$\mathbf{\Lambda}[\bar{c}] = \frac{1}{|\nabla\bar{c}|^2 + 2\beta_2}\left\{\begin{pmatrix}\partial_{x_2}\bar{c}\\ -\partial_{x_1}\bar{c}\end{pmatrix}\begin{pmatrix}\partial_{x_2}\bar{c} & -\partial_{x_1}\bar{c}\end{pmatrix} + \beta_2\mathbf{I}\right\}.$$

brings out wavefront set.

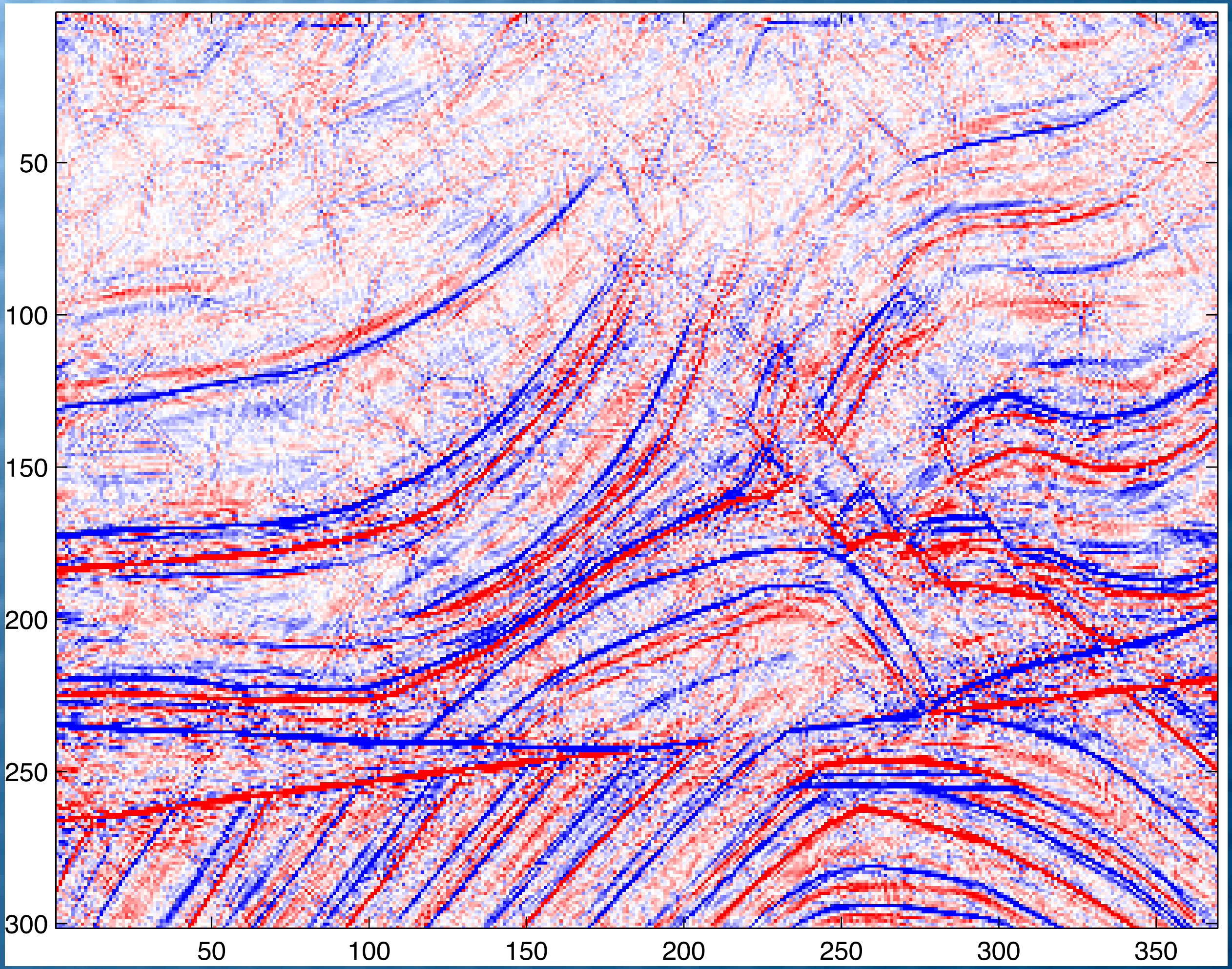
Denoising



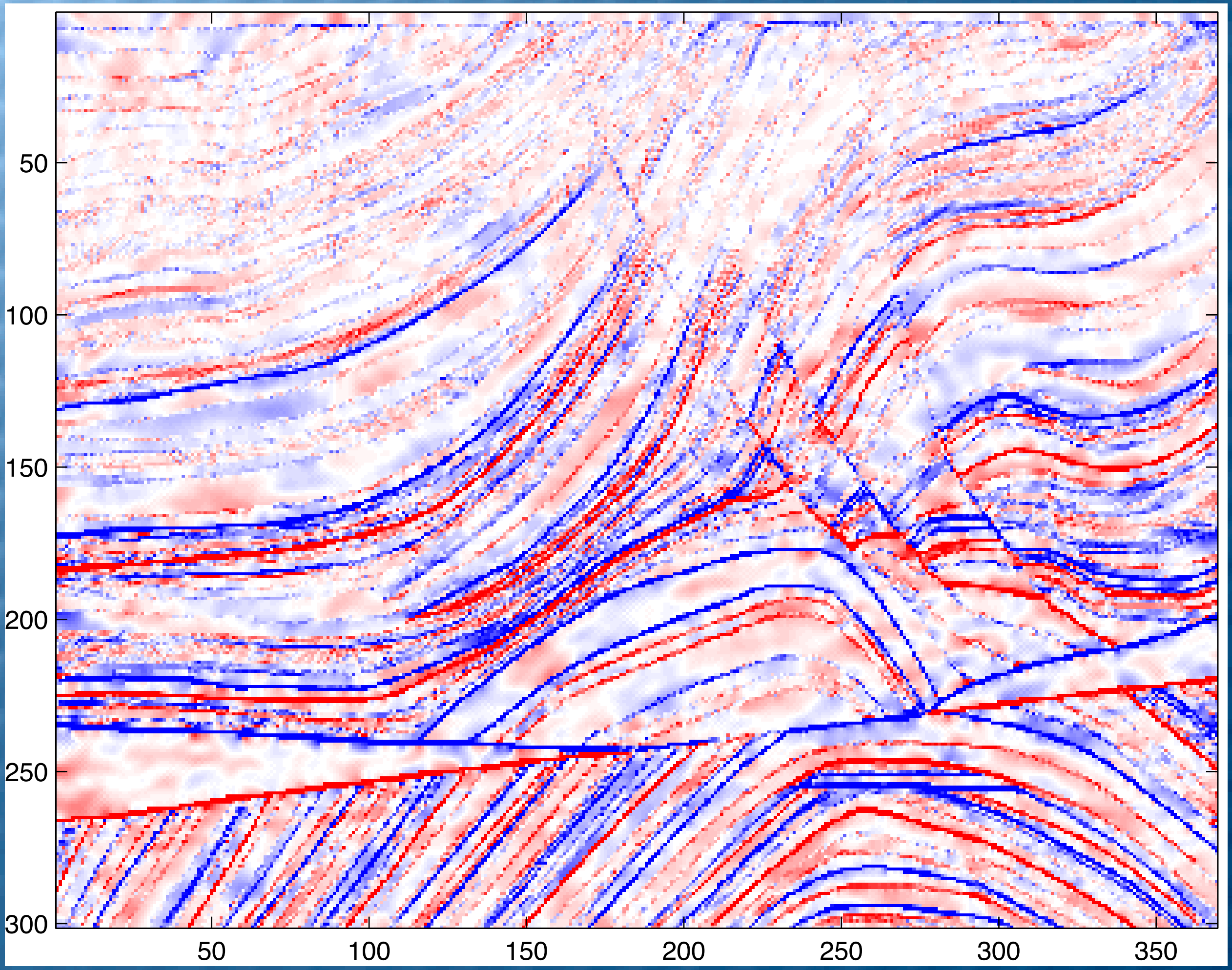
Denoising



Denoising



Denoising



Conclusions

Presented a framework that

- **Stable signal separation *via* thresholding:**
  - Ground-roll & Multiple removal (Wednesday)
  - Compute 4-D difference Cubes
- **improves imaging & inversion:**
  - deals with incoherent noise & missing data
  - sparseness constrained imaging (H & P ‘04)
  - extended to 3-D

Acknowledgements

Candes, Donoho, Demanet, Ying for making their Curvelet code available.

Partially supported by a NSERC Grant.

NSERC Investing in people, discovery and innovation
CRSNG Investir dans les gens, la découverte et l'innovation