Curvelet imaging & processing: an overview

Felix J. Herrmann (EOS-UBC)
felix@eos.ubc.ca
www.eos.ubc.ca/~felix

thanks to: Gilles, Peyman and Candes, Sacchi

Earth and Ocean Sciences UBC

Research program

How **to improve seismic images?**

What **is in the image?**

Why **is it in the image?**

Curvelet imaging & processing series is devoted to the *'How'*.

Goals

Processing & imaging scheme

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

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

Wish list

Seek a *transform domain* that is

- ***relative insensitive* to *local phase***
- ***sparse & local* (position/dip)**
- ***optimal* for *curved* reflectors**
- ***well-behaved* under *operators* (e.g imaging)**

Aim to bring out those high frequencies with *low* SNR!

Basic idea

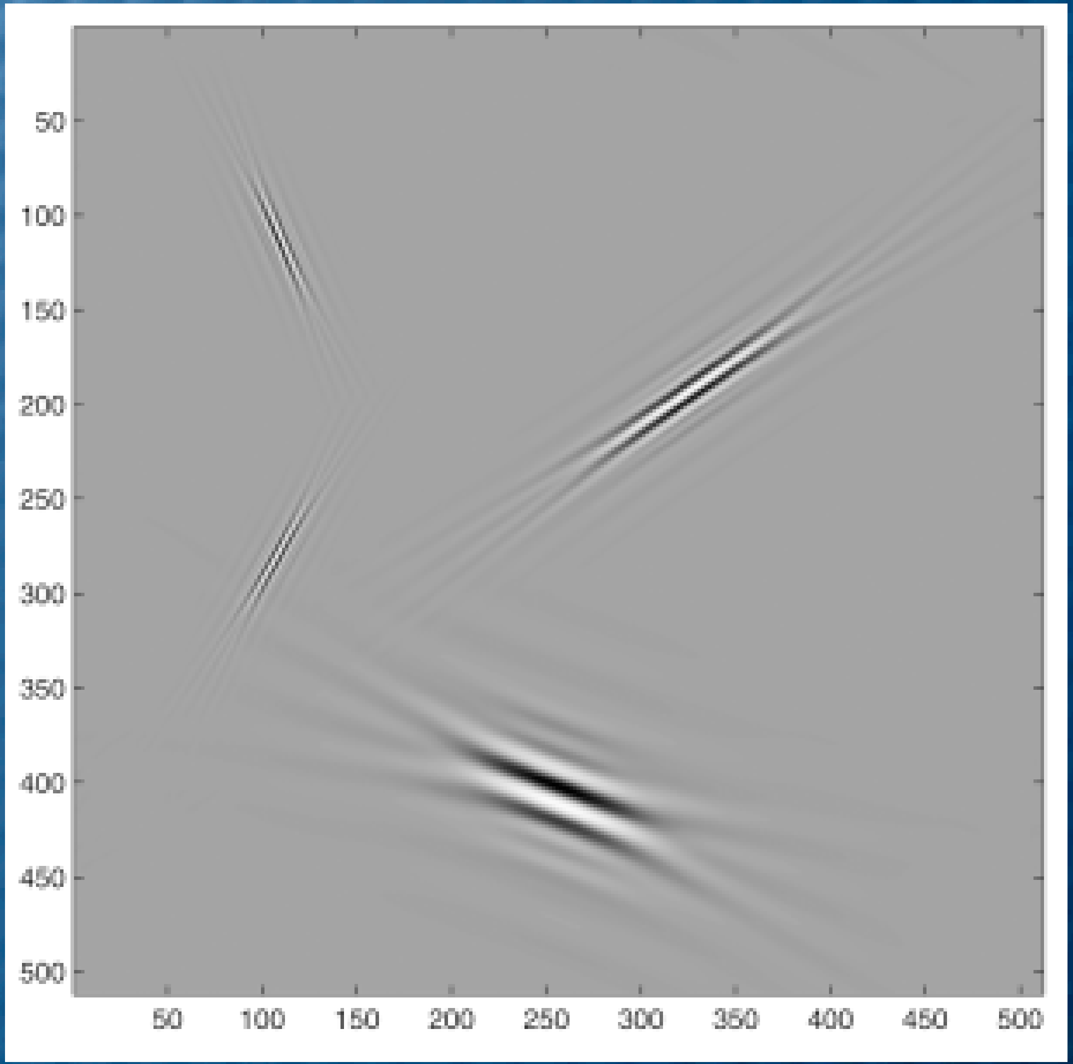
Build on the premise that you stand a much better chance of solving a problem when the model is represented optimally ...

- **local**
- **sparse**
- **multi-scale and multi-directional**

Well behaved under operators (e.g migration)!

Why curvelets

- Nonseparable
- Local in 2-D space
- Local in 2-D Fourier
- Anisotropic
- Multiscale
- Almost orthogonal
- Tight frame
- Optimal



Curvelets

Curvelets/Contourlets:

- **Anisotropic scaling law:**

$$\textbf{width} \equiv 2^{-2j} \qquad \textbf{length}^2 \equiv 2^{-j}$$

- **Directional selective:**

$$\textbf{\#orientations} = 1/\sqrt{\text{scale}}$$

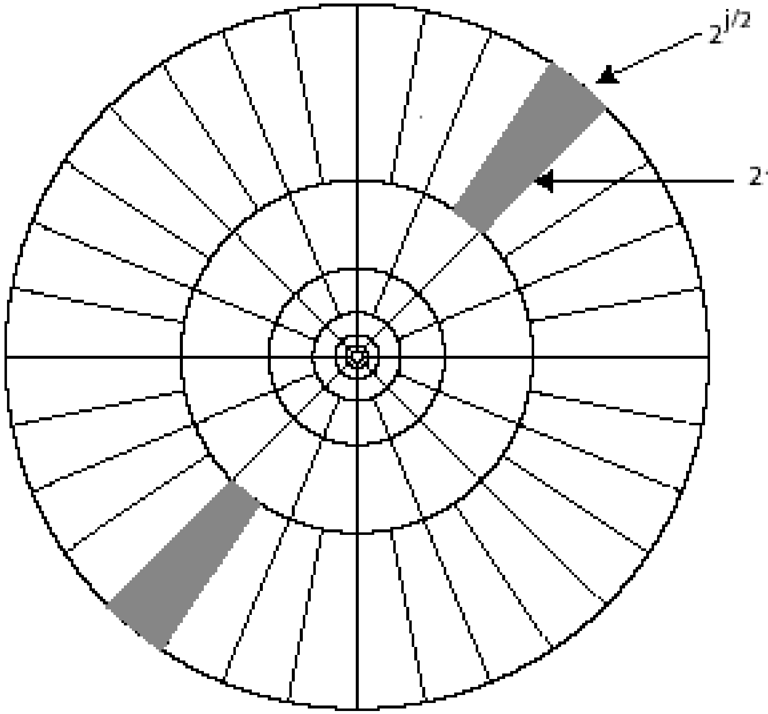
- **Close to optimal for functions with singularities on C^2-curves:**

$$\|m - \tilde{m}_m^{\text{improved wavelet}}\|_2 \qquad C \cdot m^{-2} \left(\log m\right)^3$$

Candes '02, Do '02

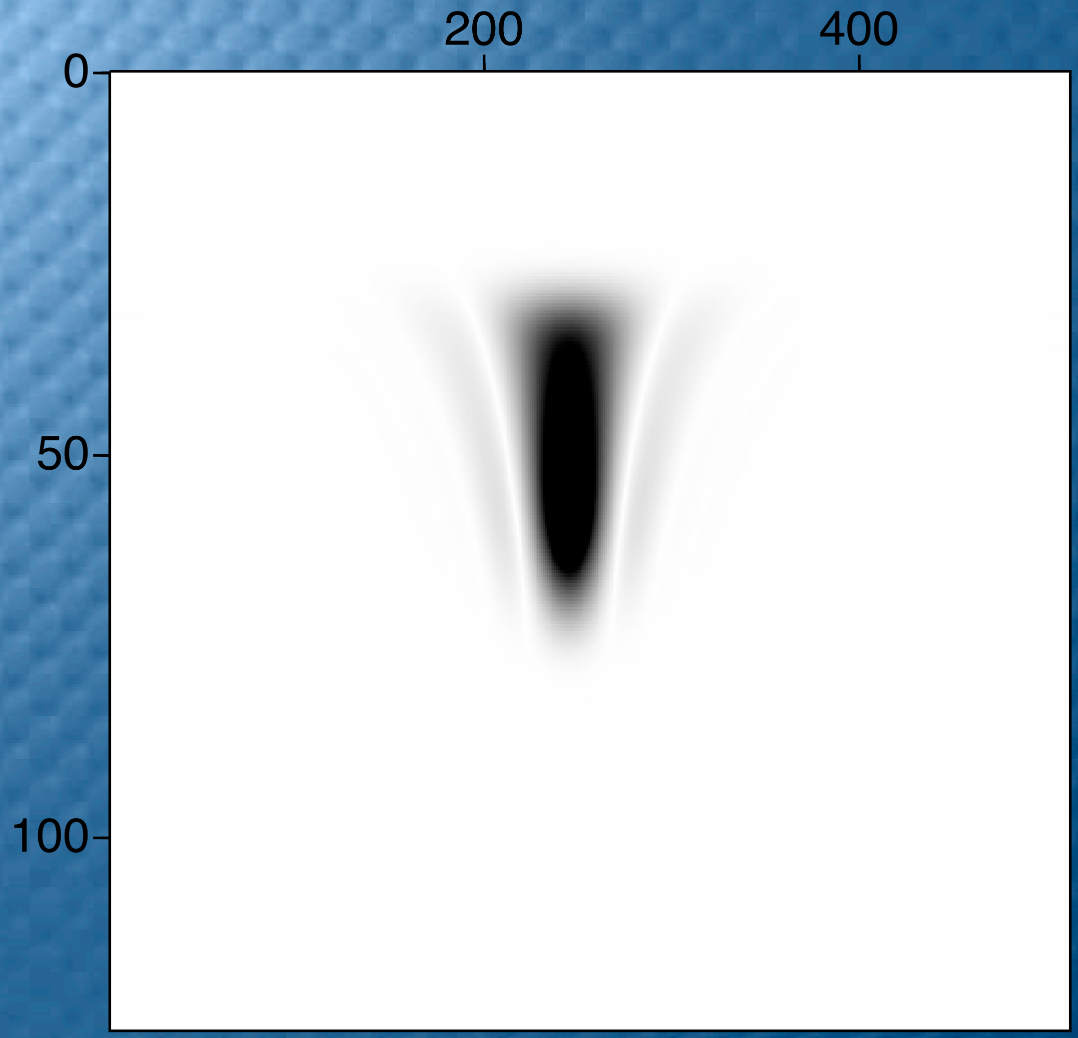
Why curvelets

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



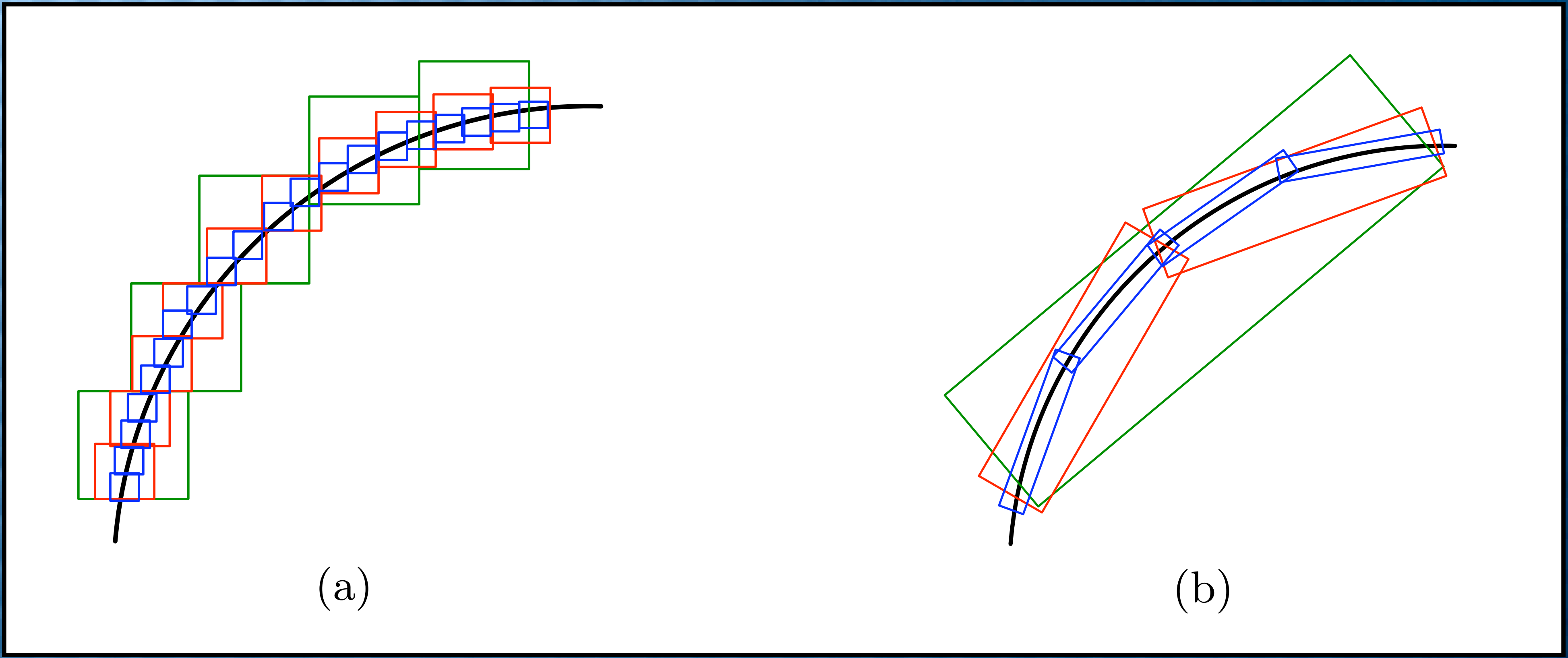
second dyadic partitioning

Why curvelets



Curvelet in FK-dom□in

Approximation rates



Do ‘01

Other domains

Most techniques are ***global:***

Non-adaptive:

- FFT
- Radon

Adaptive:

- Principle & Independent components
- SVD & KL

Approximation rates

- **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**

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

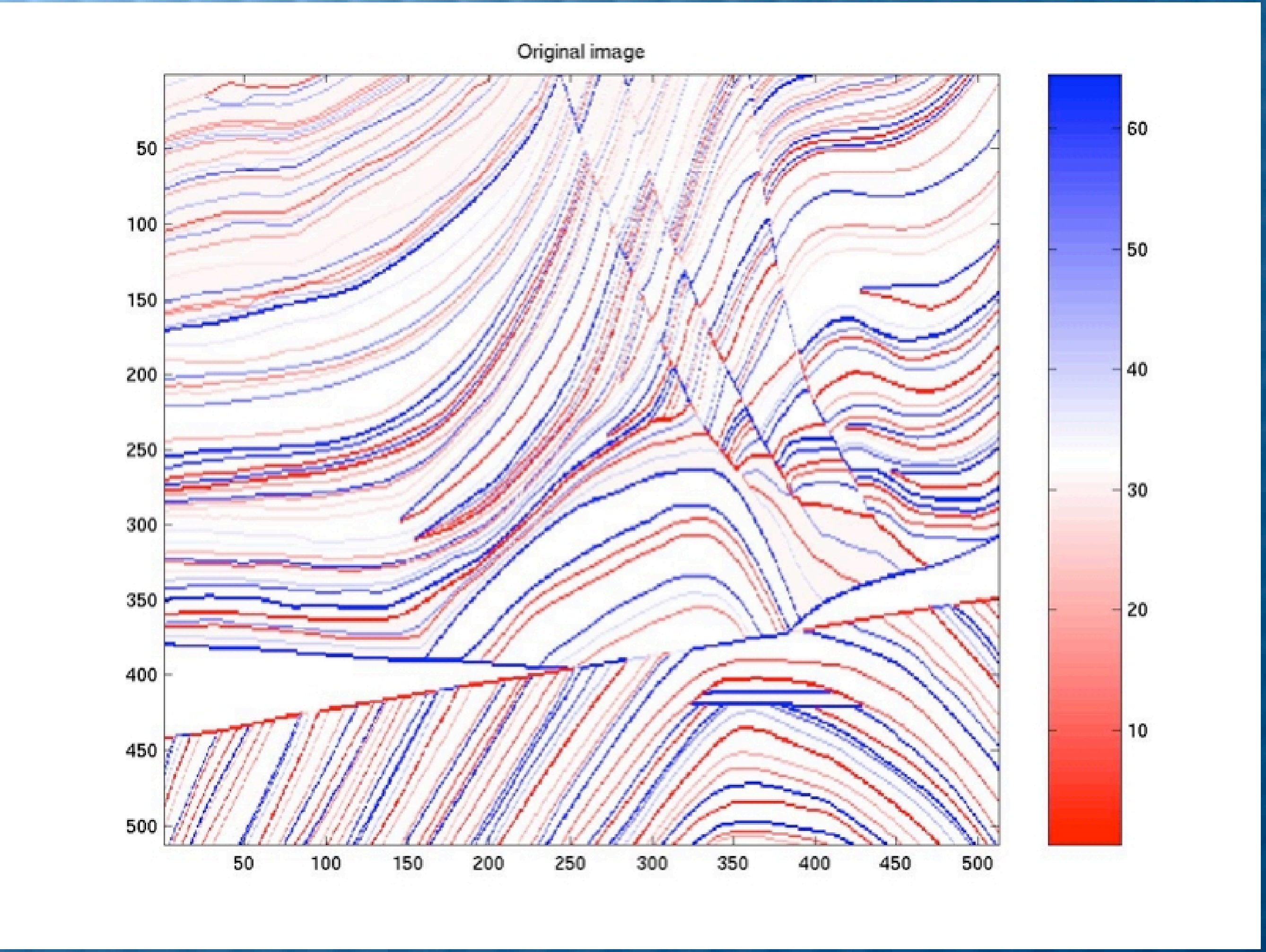
- **Close to optimal Curvelet**

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

Gain orders of magnitude ... that's ruler of the game

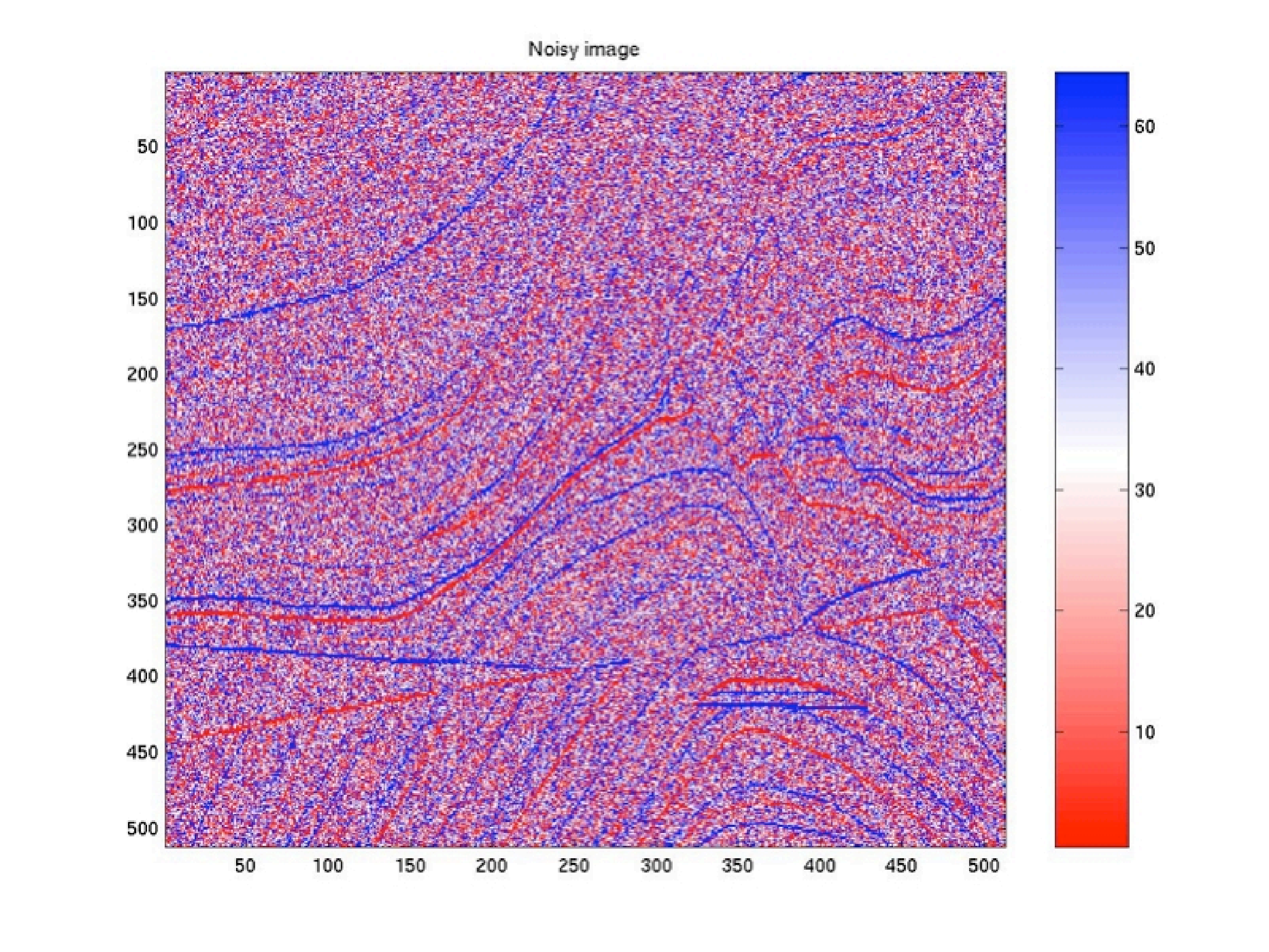
Candes '02

Denoising



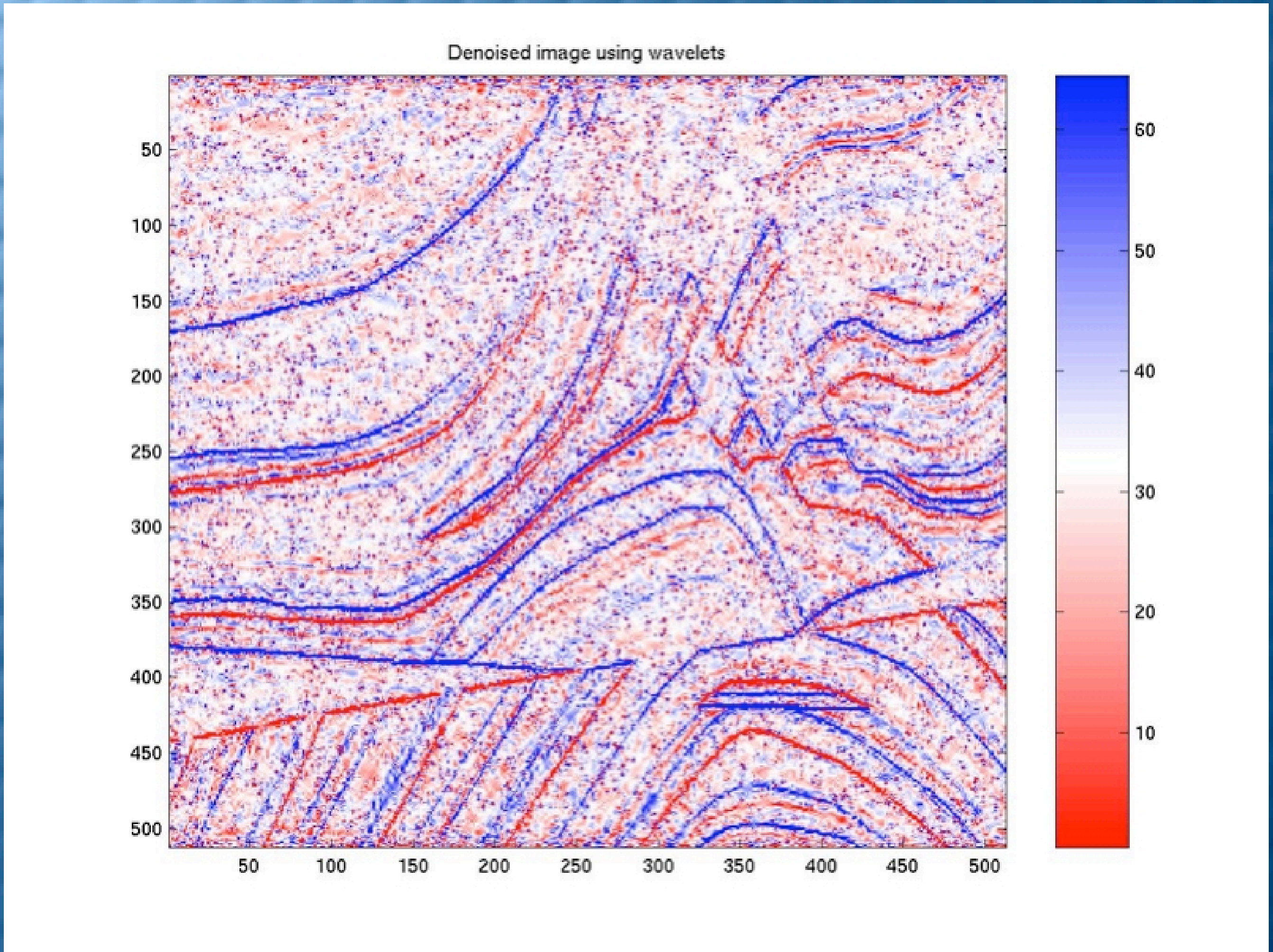
Original image
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

Denoising



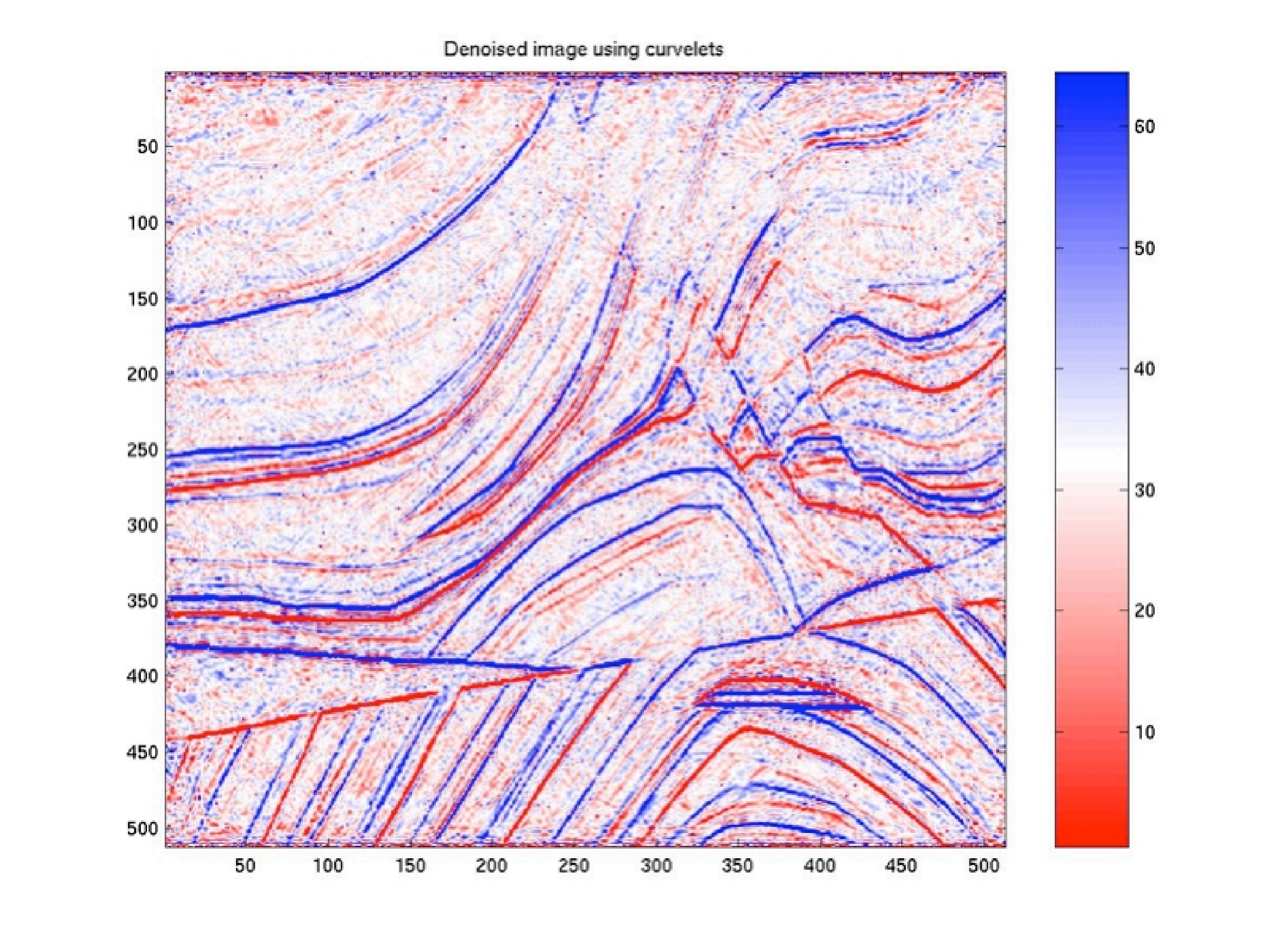
Noisy image
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

Wavelets



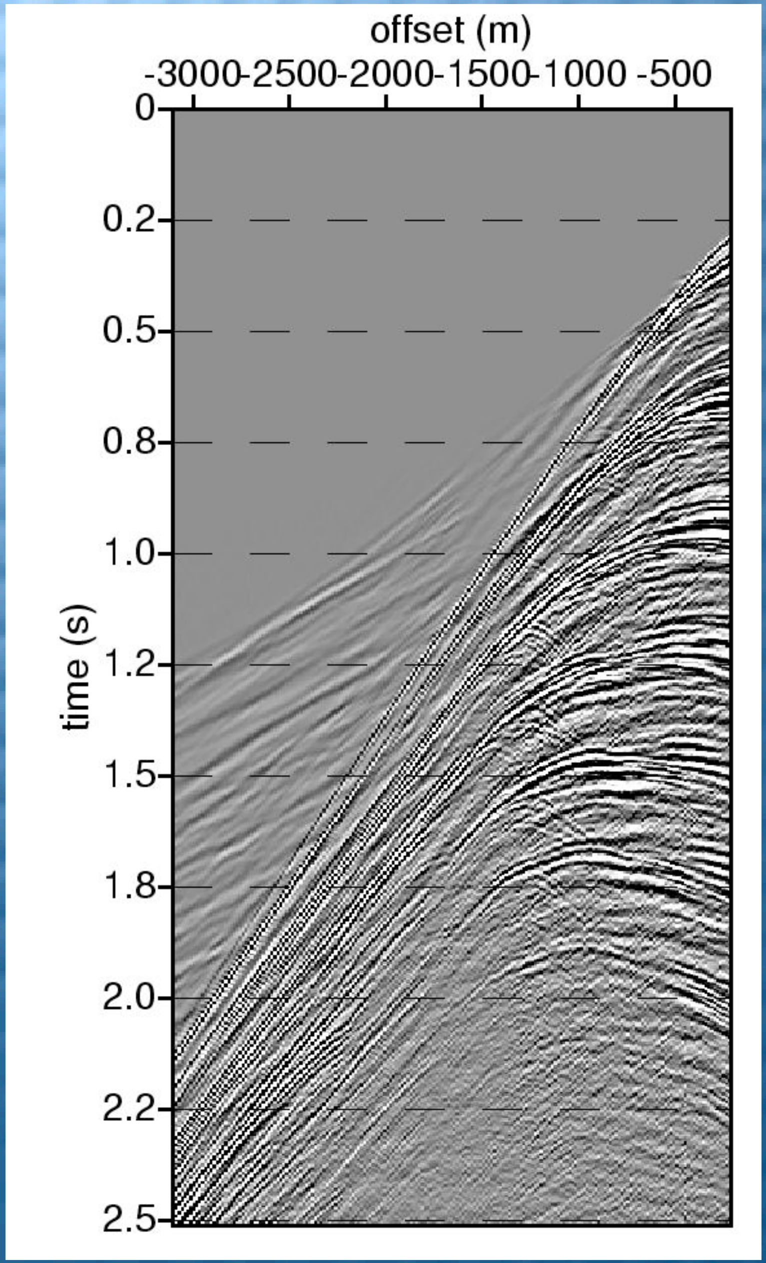
Denoised image using wavelets
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

Curvelets

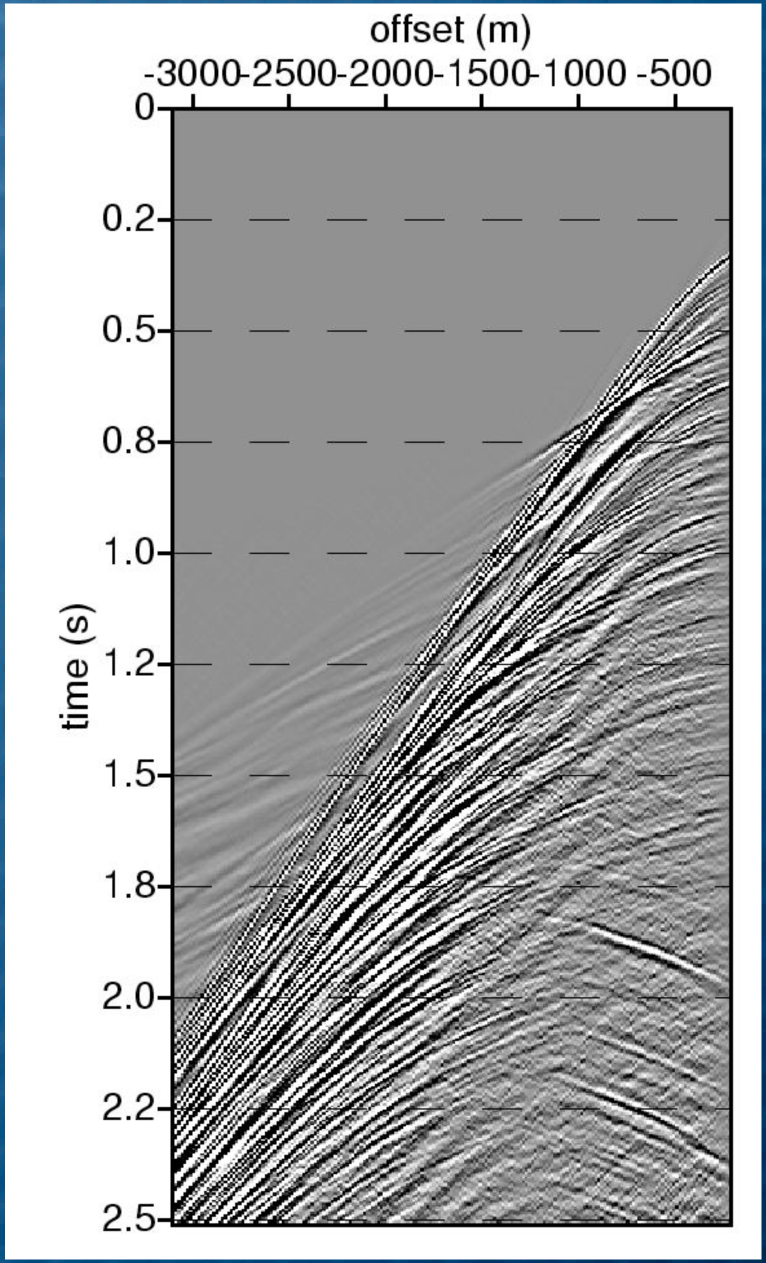


Denoised image using curvelets
50
100
150
200
250
300
350
400
450
500
50
100
150
200
250
300
350
400
450
500
60
50
40
30
20
10

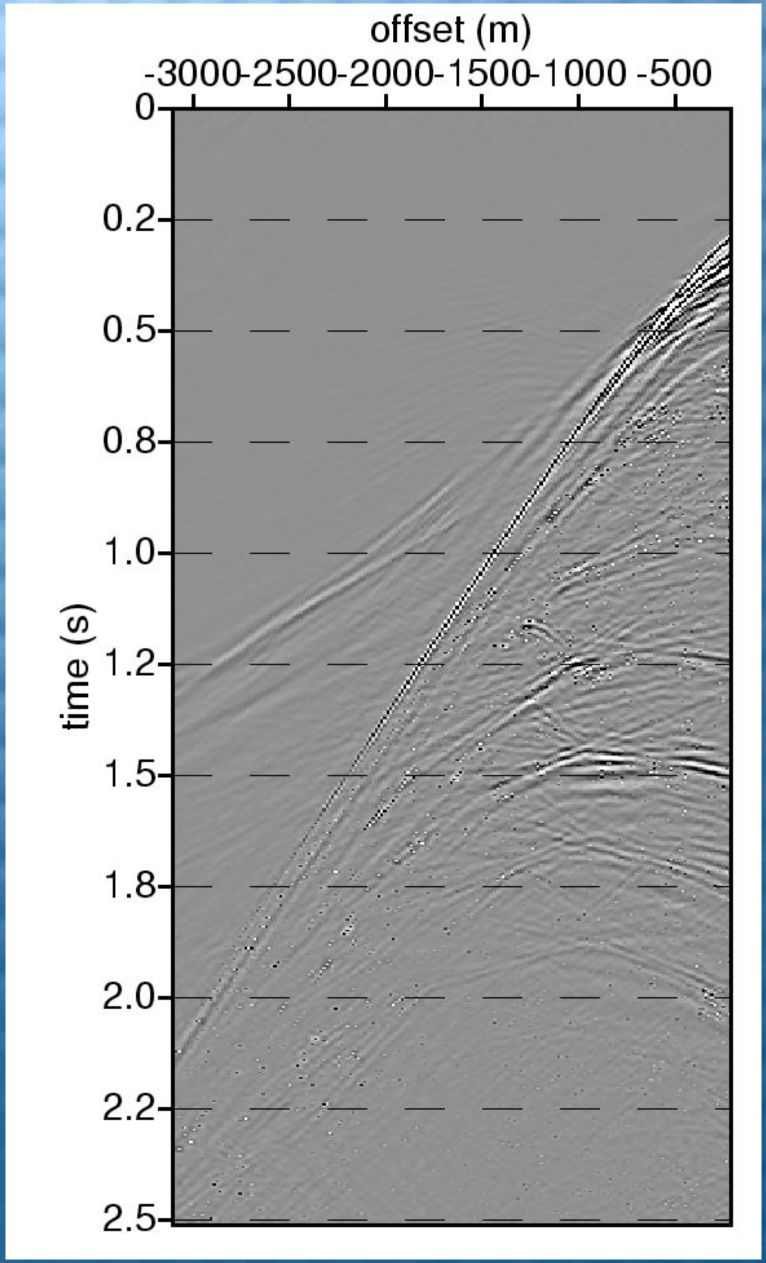
Multiple suppression with curvelets



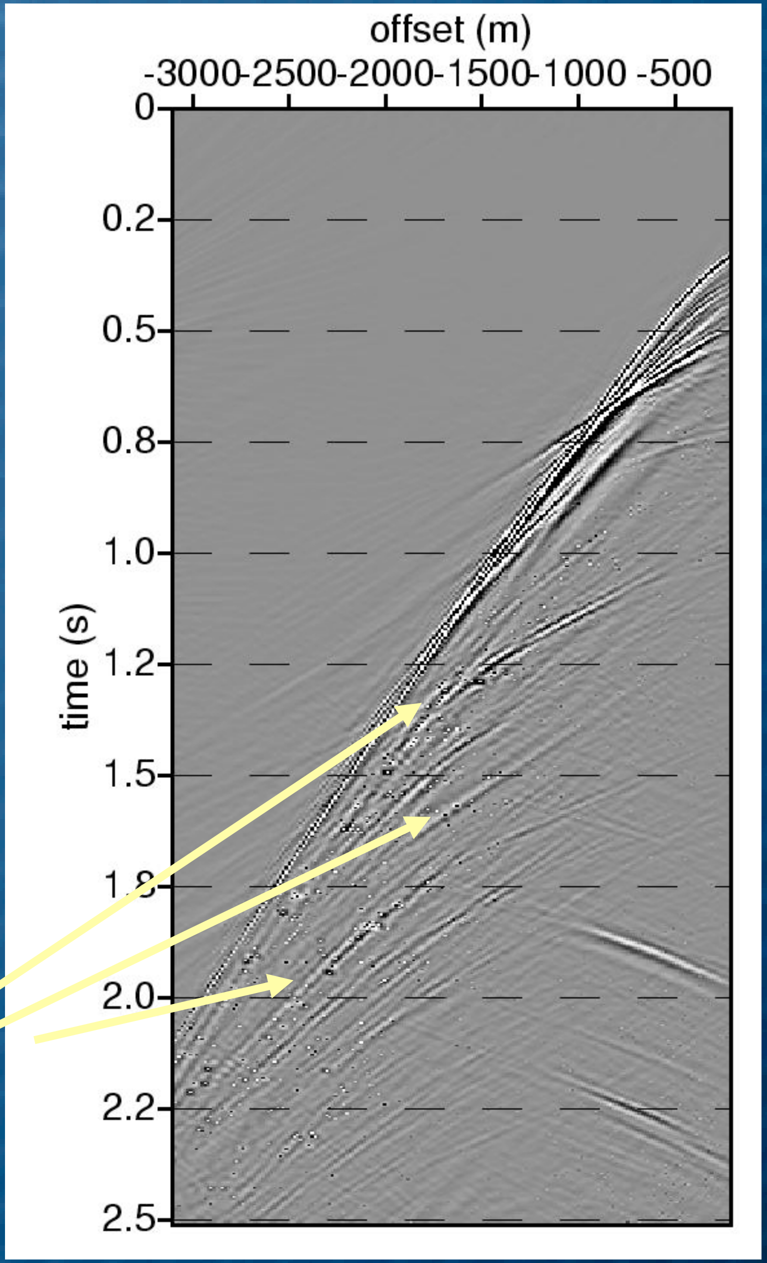
Input with multiples



Multiple suppression with curvelets



Output curvelet filtering with stronger threshold



Preserved primaries

What did we do?

Used

$$\hat{\mathbf{m}} = \mathbf{B}^{\dagger}\boldsymbol{\Theta}_{\lambda\Gamma}\left(\mathbf{B}\mathbf{d}\right) \text{ with } \Gamma = \left|\mathbf{B}\mathbf{n}_{\mathbf{p}}\right|$$

to denoise

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

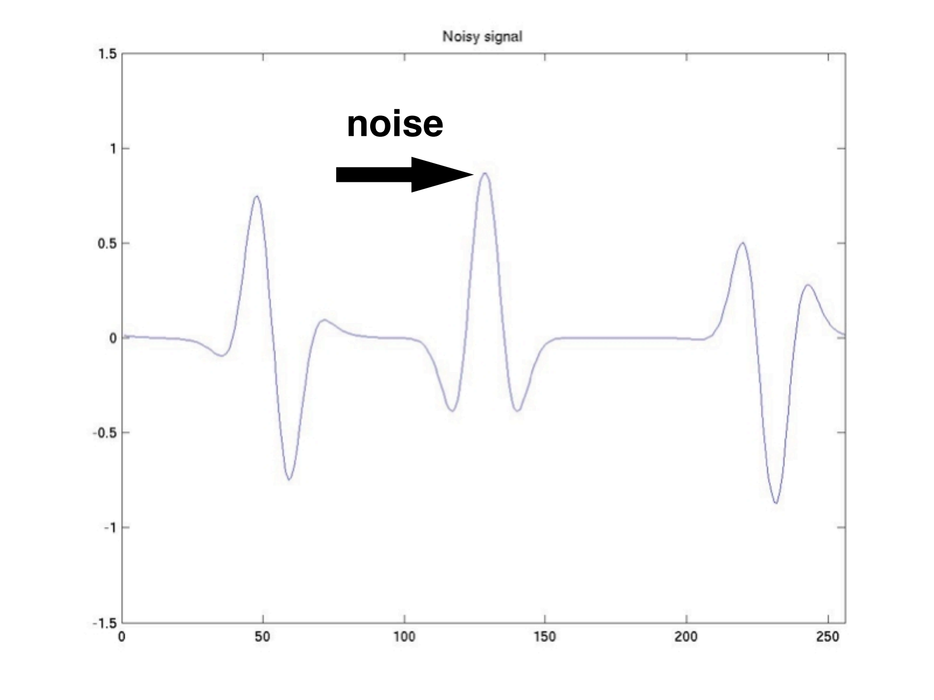
with a simple ***mute*** with λ control parameter.

Some theory

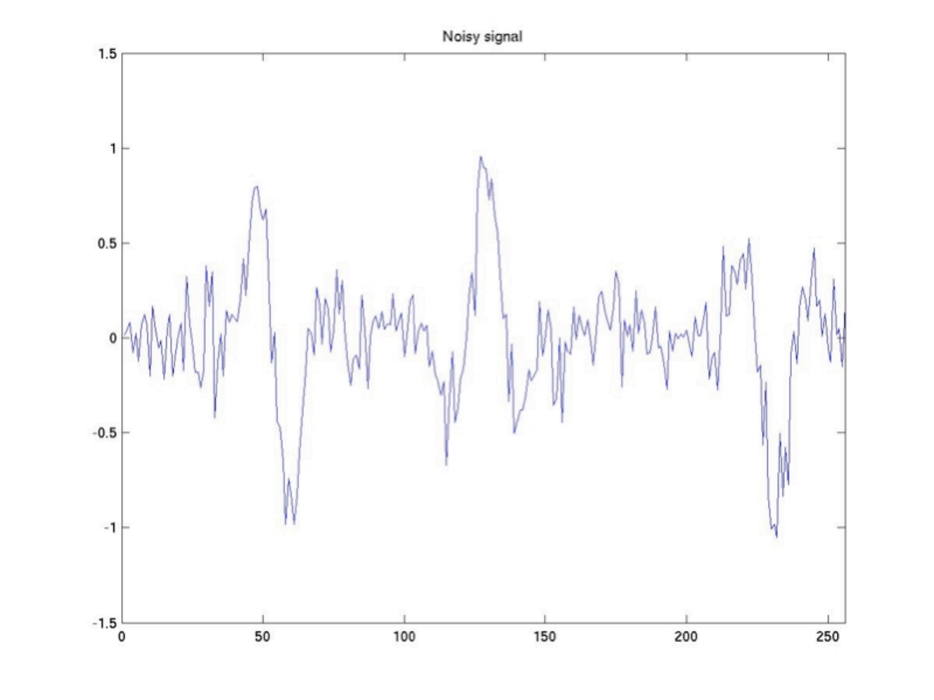
- ***Unconditional* bases:**
  - **norm *always* shrinks when shrinking coef.**
- ***Denoising*** by ***minimax* estimation:**
  - ***diagonal*** thresholding (non-linear)
- ***Extension*** to ***colored noise:***
  - ***almost diagonalizes Covariance***

Ideal *edge-preserving* tool for seismic processing.

Unconditional bases

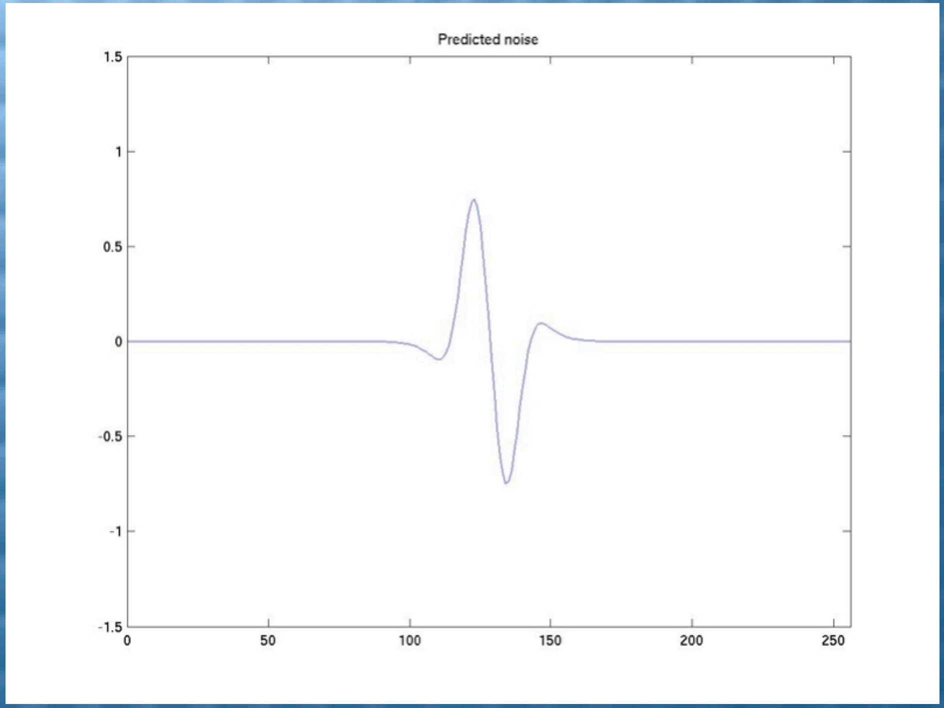


signal + coherent noise

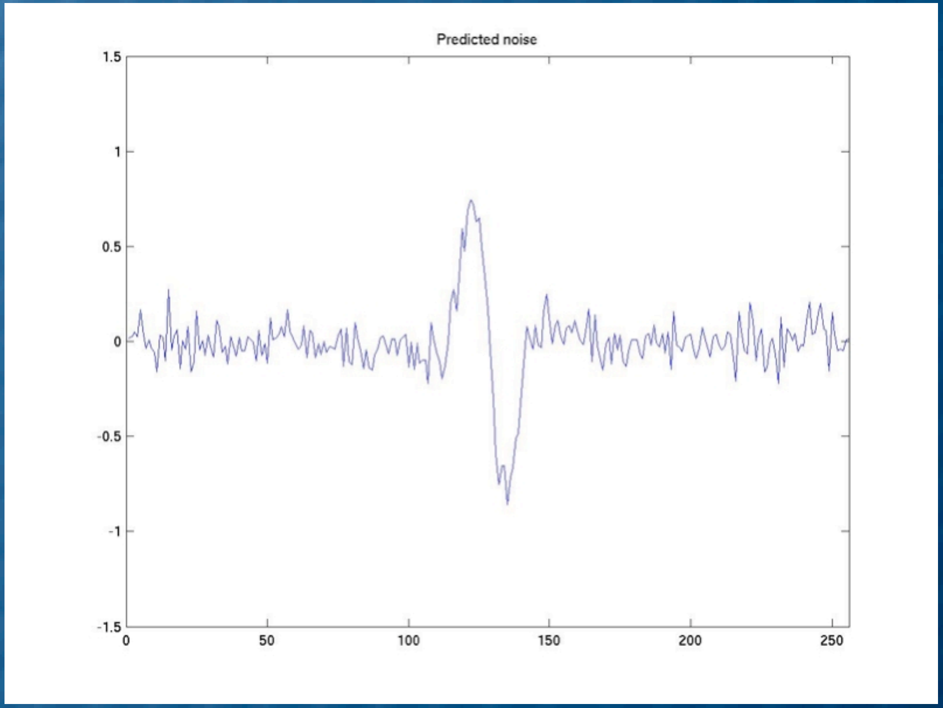


signal + coherent & incoherent noise

Unconditional bases

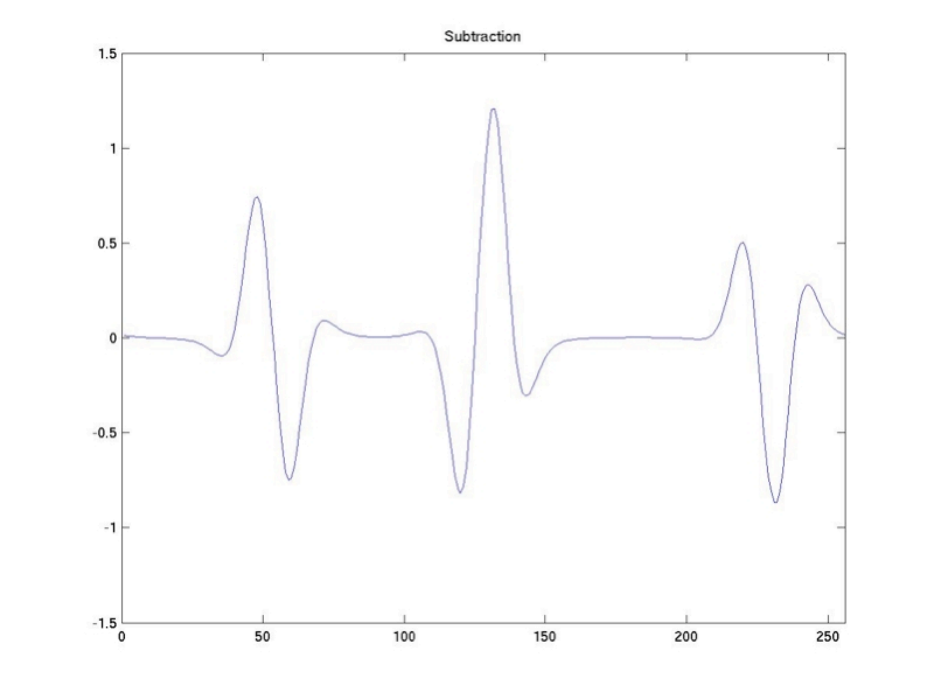


‘wrongly’ predicted noise

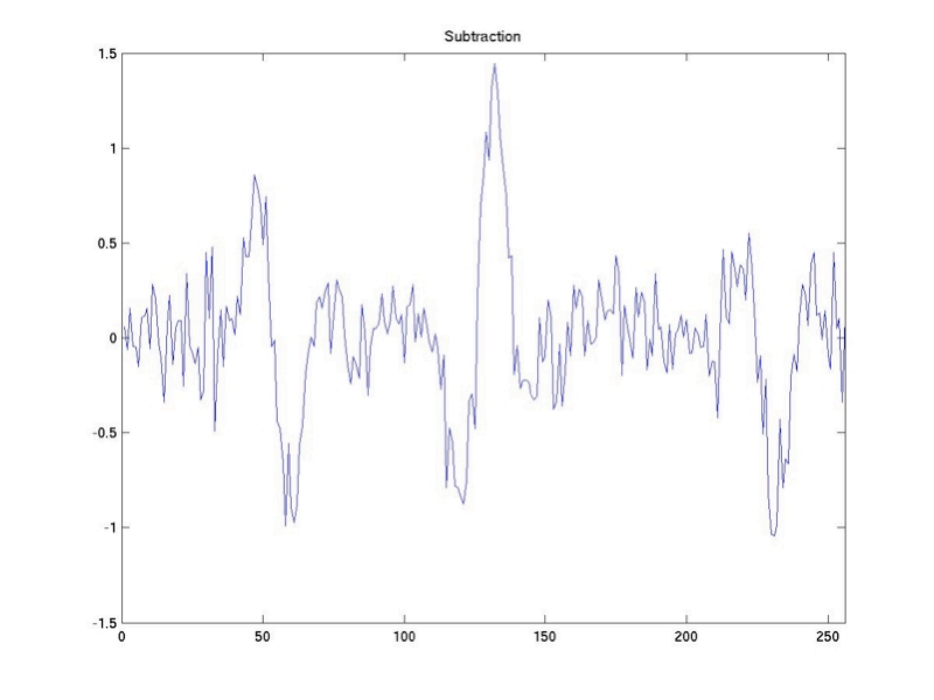


‘wrongly’ noisy predicted noise

Unconditional bases

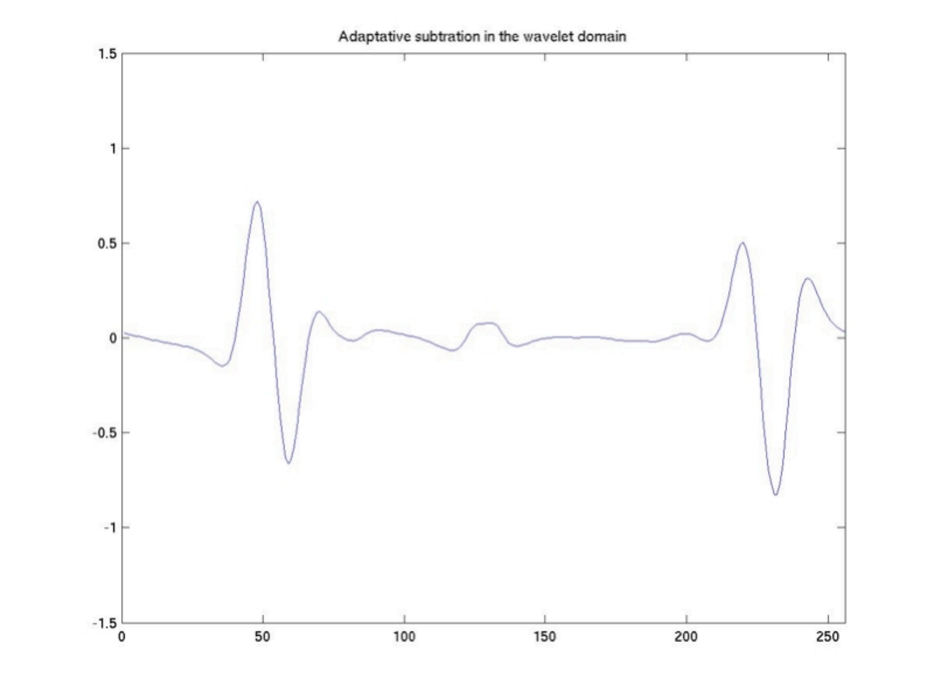


noise-free subtraction

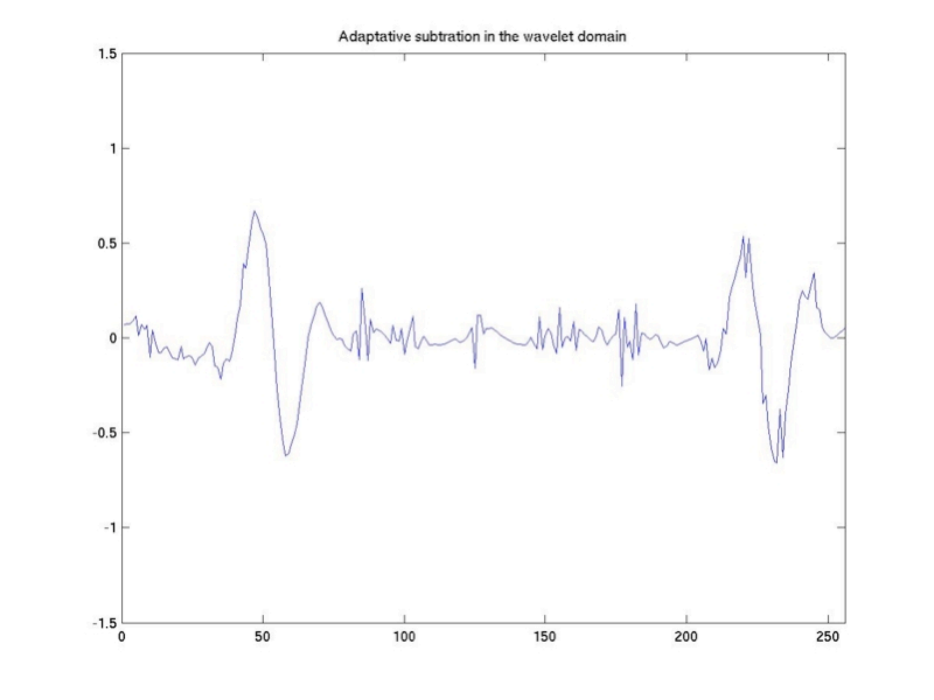


noisy subtraction

Unconditional bases



noise-free adaptive subtraction



noisy adaptive subtraction

Unconditional bases

Works because wavelets are

- ***local* in space**
  - ***accounts for bulk of the phase***
- ***local* in frequency**
  - ***adapts* to amplitude spectrum**

Adds robustness to the equation ...

Minimax estimation

$$\hat{\mathrm{m}} = \mathbf{B}^{\dagger}\Theta_t\left(\mathbf{B}\mathbf{d}\right)$$

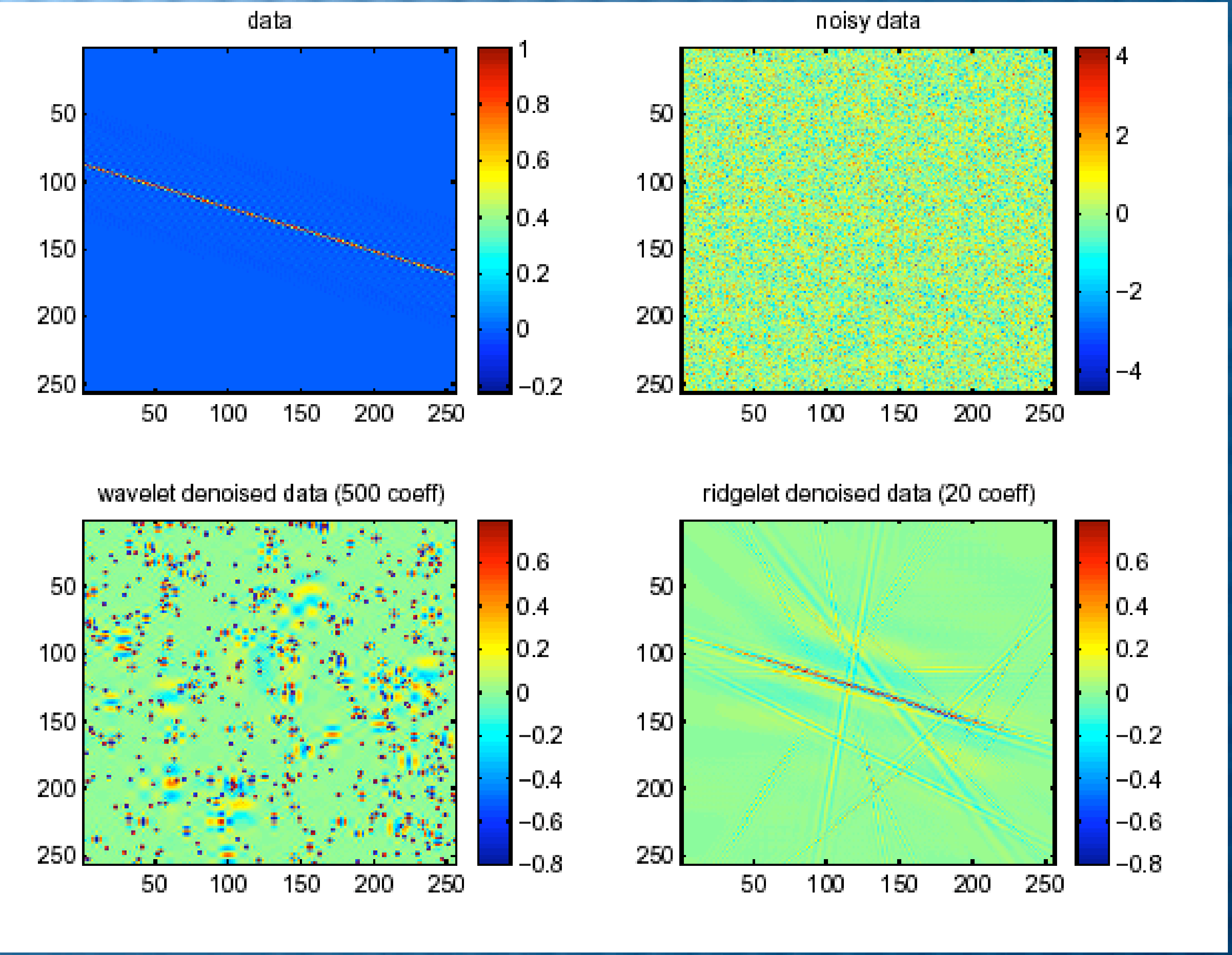
- approximates **minimax**, minimizes max. risk without *prior* **info**
- **Bayes for 'least favorable' *prior***
- **preserves edges**
- **optimal/unconditional** basis functions

Mallat '97
Donoho '92

Wavelets

- Represent **piece-wise smooth** functions at **“no”** additional cost.
- Do **not** have to know where the **singularities** are.
- **Only** good for **point-scatterers** or **horizon/vertically-aligned reflectors.**

Directional wavelets



courtesy WaveleLab + Beamlab '02

Non-linear estimation

Reformulate

$$\hat{\mathbf{m}} \; : \; \min_{\mathbf{m}} \frac{1}{2}\|\mathbf{d} - \mathbf{m}\|_2^2$$

into

$$\hat{\tilde{\mathbf{m}}} \; : \; \min_{\tilde{\mathbf{m}}} \frac{1}{2}\|\tilde{\mathbf{d}} - \tilde{\mathbf{m}}\|_2^2 + \lambda^2\|\tilde{\mathbf{m}}\|_p$$

with

$$\tilde{\mathbf{m}} \triangleq \mathbf{B}\mathbf{m} \text{ and } \tilde{\mathbf{d}} \triangleq \mathbf{B}\mathbf{d}$$

Non-linear estimation

Hard thresholding for p=0:

$$\Theta_\lambda^h\left(\tilde{\mathbf{d}}\right) \triangleq \begin{cases} \tilde{\mathbf{d}} & \text{if} \quad |\tilde{\mathbf{d}}| > \lambda \\ 0 & \text{if} \quad |\tilde{\mathbf{d}}| \leq \lambda \end{cases}$$

Soft thresholding for p=1:

$$\Theta_\lambda^s\left(\tilde{\mathbf{d}}\right) \triangleq \begin{cases} \text{sign}(\tilde{\mathbf{d}})(|\tilde{\mathbf{d}}| - \lambda)_+ & \text{if} \quad |\tilde{\mathbf{d}}| > \lambda \\ 0 & \text{if} \quad |\tilde{\mathbf{d}}| \leq \lambda \end{cases}$$

Adaptive subtraction

Adaptive subtraction by matched filter:

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

- **residue is the denoised data**
- **risk of over fitting**

May loose primary reflection events ...

Guitton& Verschuur'04

Non-linear adaptive subtraction

Extend to *colored noise:*

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

Solve

$$\hat{\mathbf{m}} : \min_{\mathbf{m}} \frac{1}{2} \|\mathbf{C}_n^{-1/2} \left(\mathbf{d} - \mathbf{m}\right)\|_2^2$$

with

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

Non-linear adaptive subtraction

Recast in *Curvelet* domain:

$$\hat{\tilde{\mathbf{m}}} \ : \ \min_{\tilde{\mathbf{m}}} \frac{1}{2}\|\mathbf{C}_{\tilde{n}}^{-1/2}\left(\tilde{\mathbf{d}} - \tilde{\mathbf{m}}\right)\|_2^2 + \lambda^2\|\tilde{\mathbf{m}}\|_p$$

Use ***unconditional-basis*** **property:**

$$\mathbf{C}_{\tilde{n}} \triangleq \mathbf{E}\{\tilde{\mathbf{n}}\tilde{\mathbf{n}}^T\} \approx \mathrm{diag}\left(\mathrm{diag}\left(\mathbf{C}_{\tilde{n}}\right)\right) \triangleq \Gamma^2$$

‘Challenge’ to find the Γ’s

Non-linear adaptive subtraction

Solve

$$\hat{\tilde{\mathbf{m}}} \; : \; \min_{\tilde{\mathbf{m}}} \frac{1}{2} \|\mathbf{\Gamma}^{-1}\left(\tilde{\mathbf{d}} - \tilde{\mathbf{m}}\right)\|_2^2 + \lambda^2 \|\tilde{\mathbf{m}}\|_p$$

can be written as

$$\hat{\mathbf{m}} = \mathbf{B}^{\dagger}\mathbf{\Gamma}\Theta_{\lambda}\left(\mathbf{\Gamma}^{-1}\mathbf{B}\mathbf{d}\right) = \mathbf{B}^{\dagger}\Theta_{\boldsymbol{\lambda}\mathbf{\Gamma}}\left(\tilde{\mathbf{d}}\right).$$

No matched filter required!

Noise prediction

Model or predict noise and set

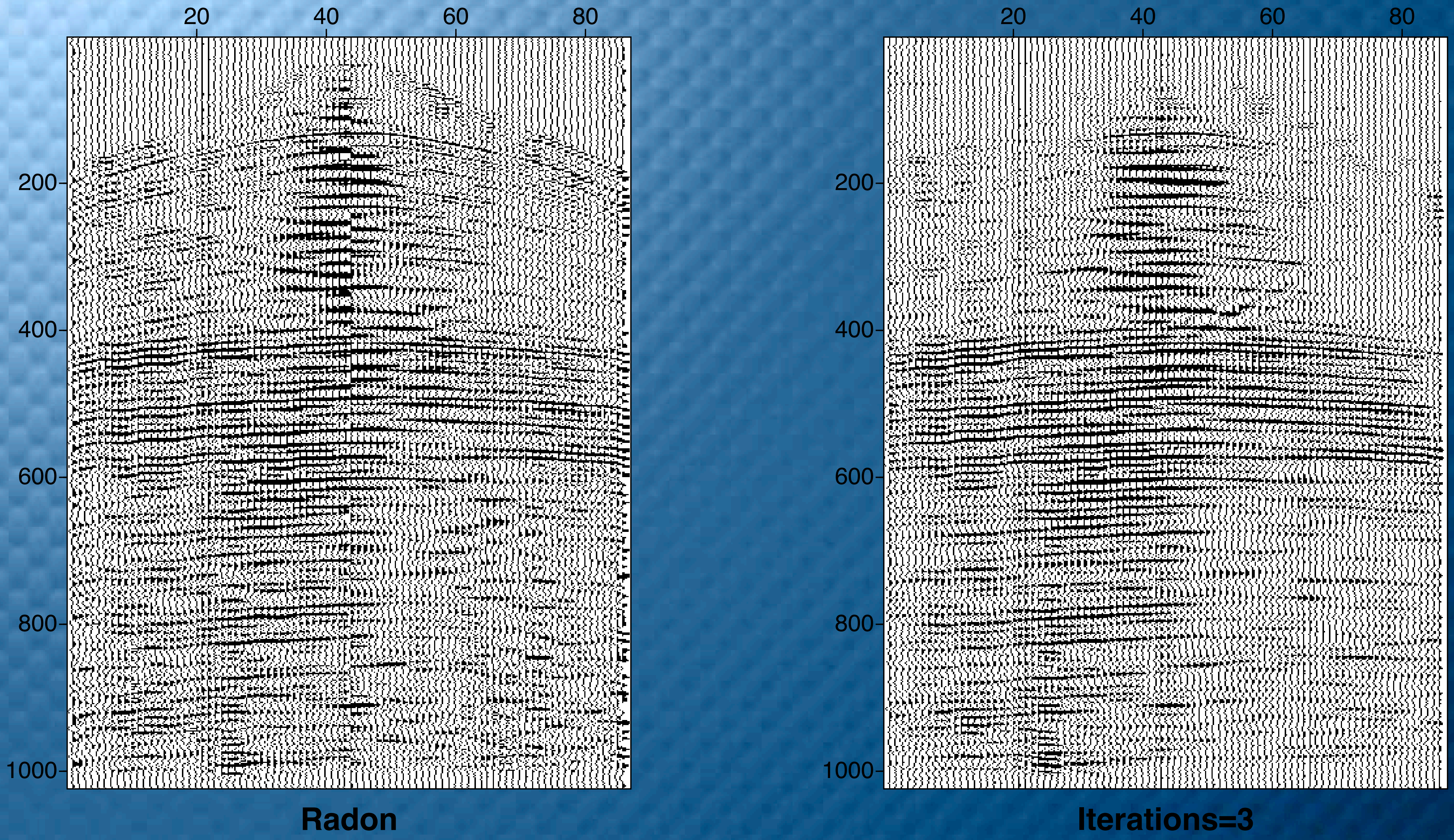
$$\Gamma = |\mathbf{B}\mathbf{n}_p|$$ **with n_p predicted noise**

- **ground roll, multiples, 4-D vintages**
- **conventional techniques (e.g. Radon)**

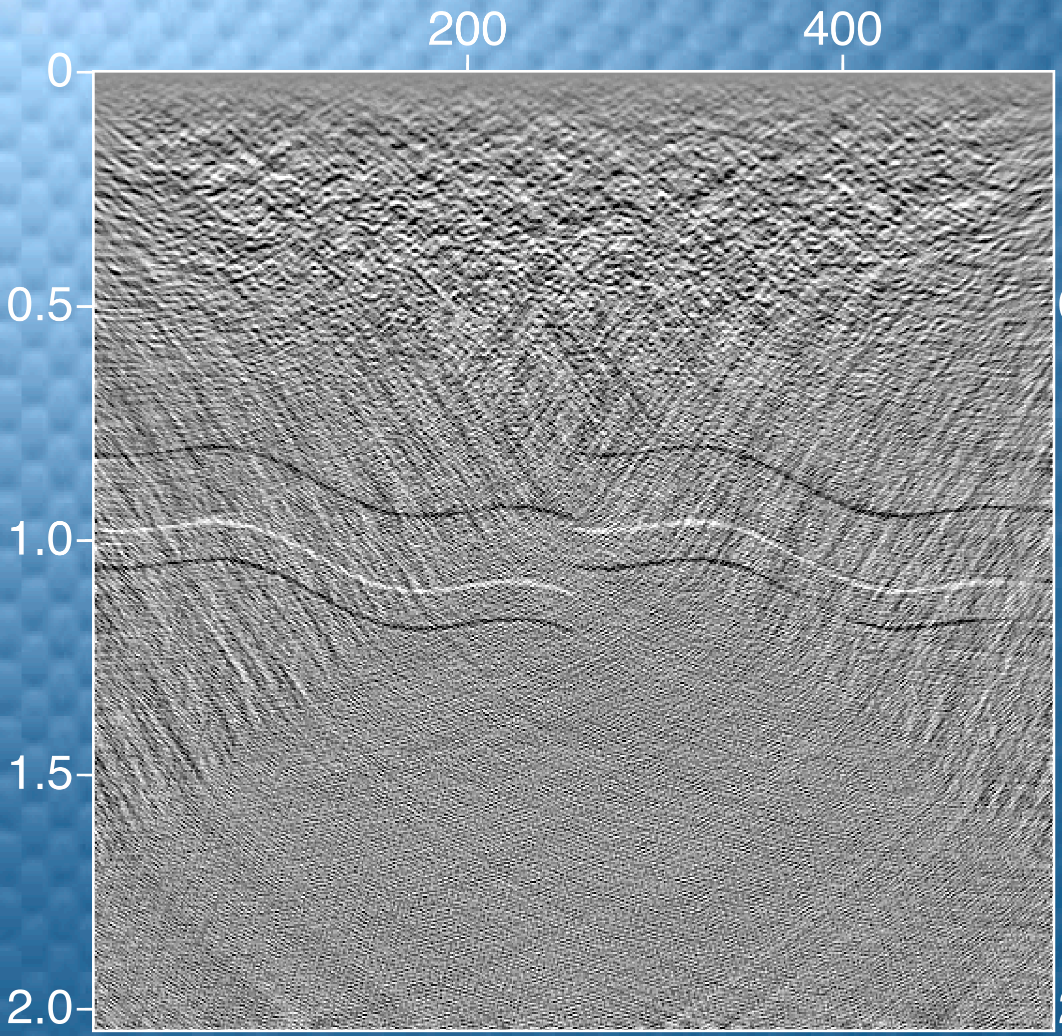
Monte-Carlo sample diag. covariance

- **migration**

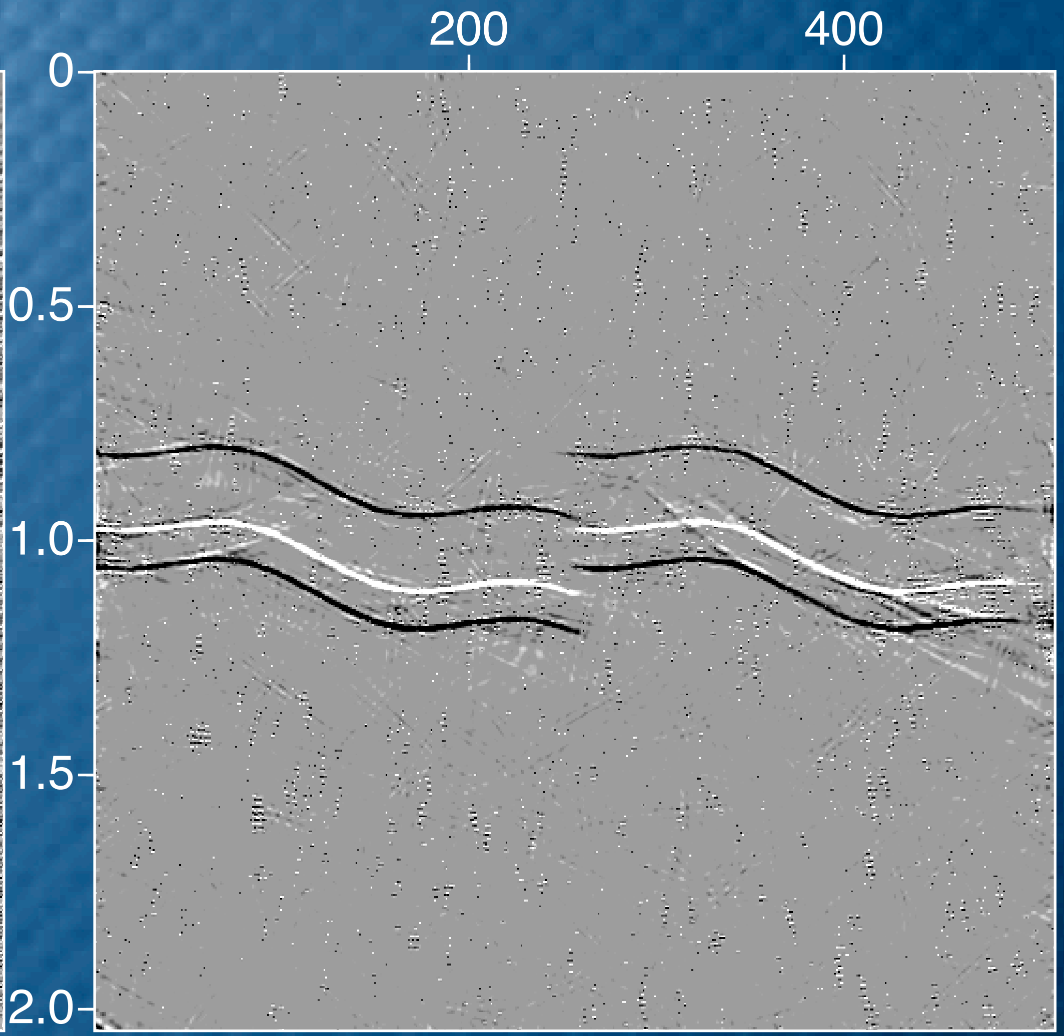
Ground-roll removal with curvelets



Imaging with Curvelets

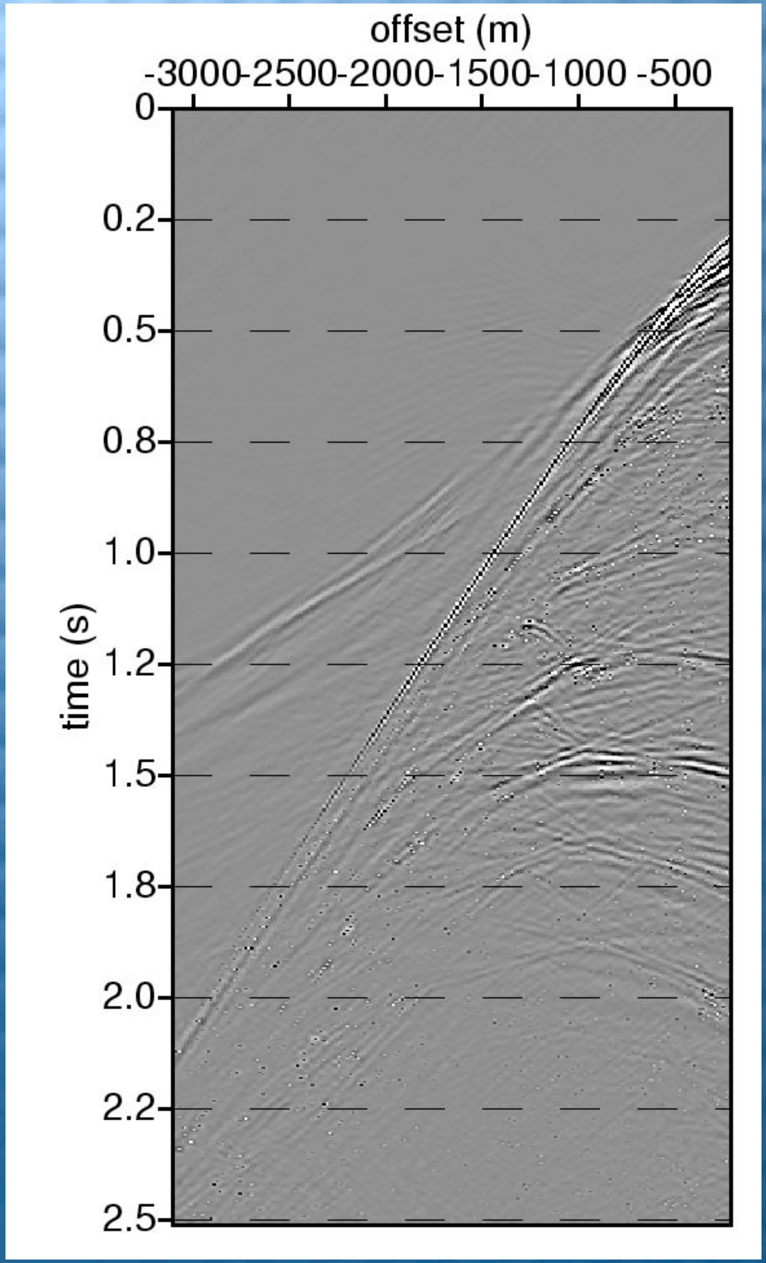


Least-squares migrated Image

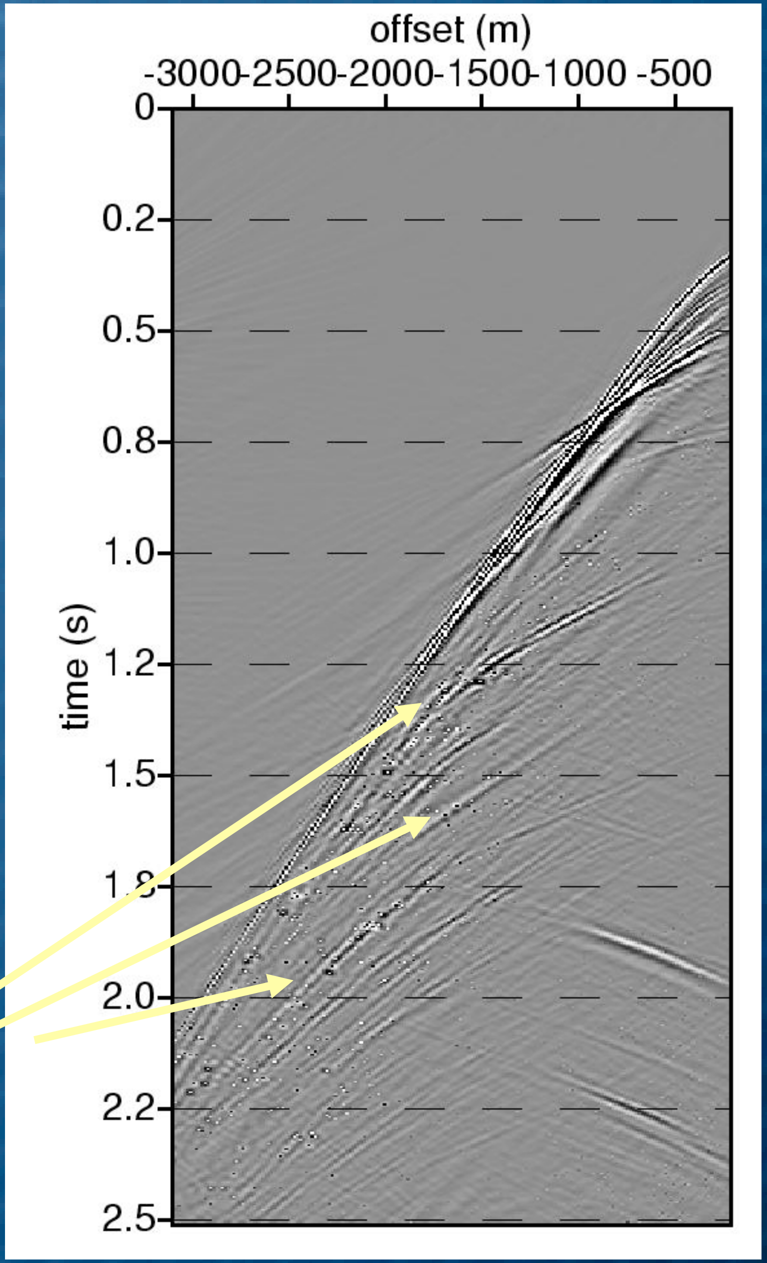


Constrained Optimization

Multiple suppression with curvelets



Output curvelet filtering with stronger threshold



Preserved primaries

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

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

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

Global optimization

Set tolerances

$$\mathbf{e}_\mu = \begin{cases} \boldsymbol{\Gamma}_\mu & \text{if} \quad |\hat{\tilde{\mathbf{m}}}_0|_\mu \quad \geq \quad |\lambda\boldsymbol{\Gamma}|_\mu \\ \boldsymbol{\lambda\Gamma}_\mu & \text{if} \quad |\hat{\tilde{\mathbf{m}}}_0|_\mu \quad < \quad |\lambda\boldsymbol{\Gamma}|_\mu \end{cases}$$

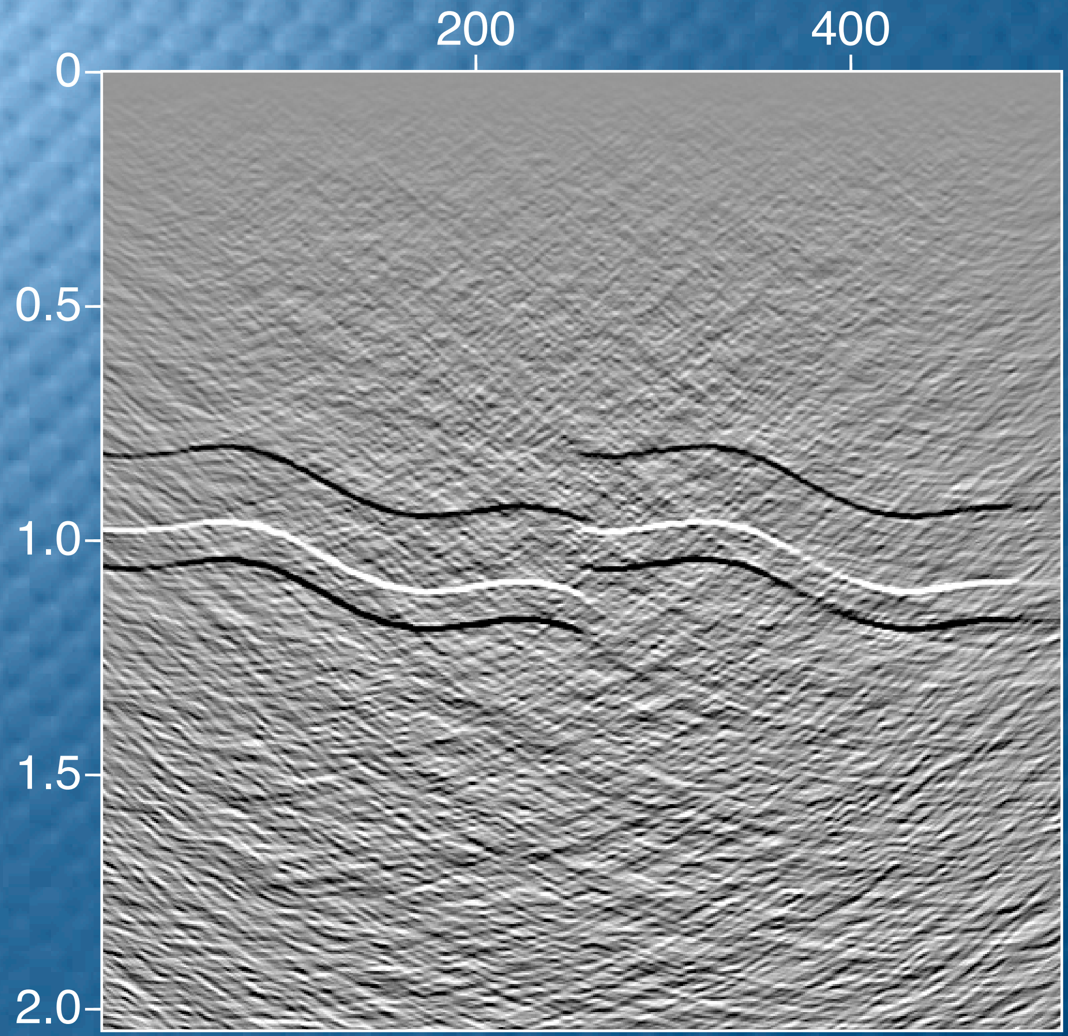
with $\sqcap$ **defining the confidence interval,**
e.g $\sqcap = 3$ **corresponds to 95 %.**

Global optimization

Constrained optimization problem:

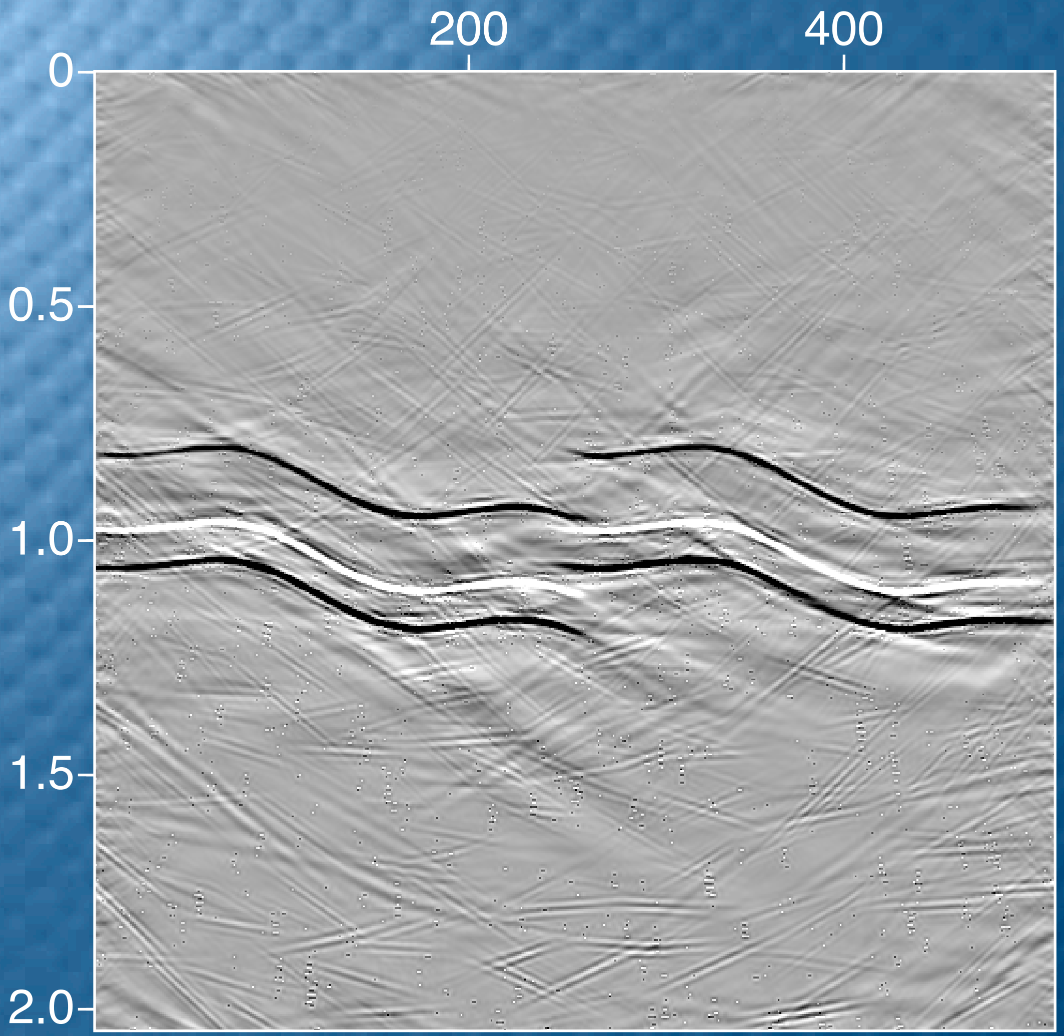
- **Uses augmented Lagrangian (Nocedal and Wright 2001)**
- **L1-penalty function**
- **Initial Lagrangian multipliers given by gradient of** $\hat{m}_0$
- **Uses Steepest Decent and line search**

Examples



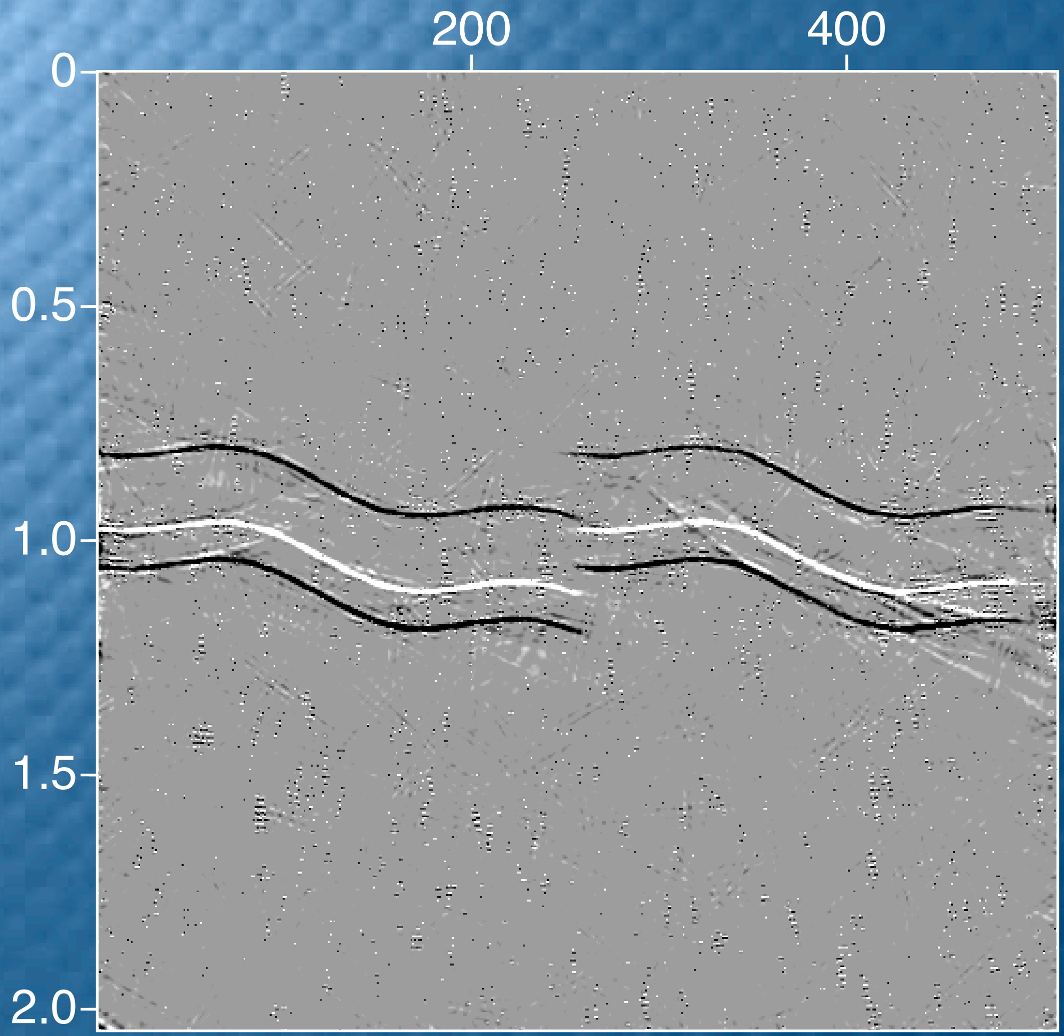
Noisy Image

Examples



Denoised after Thresholding

Examples



Constrained Optimization

Applications

Presented a framework that

- **removes (coherent) noise *via* adaptive subtraction:**
  - Ground-roll & Multiple removal (Thursday)
  - Compute 4-D difference Cubes (Wednesday)
- **improves imaging & inversion:**
  - sparseness constrained imaging (next & Thursday)

- **ample opportunity to expand!**

Conclusions

- Succeeded in partly full filling our “dream”.
- Devised a robust estimation method relatively insensitive to local phase.
- Right norm is an important & open problem.
    - exploit redundancy seismic data
    - function spaces & learning functionals
- So far focussed on ‘simple denoising’ let’s include an operator ....

Acknowledgements

Candes & Donoho 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