Curvelet imaging and processing: adaptive multiple elimination

Felix J Herrmann, Peyman P Moghaddam (EOS-UBC) and Eric J. Verschuur (Delphi, TUD)



Thanks to Emmanuel Candes

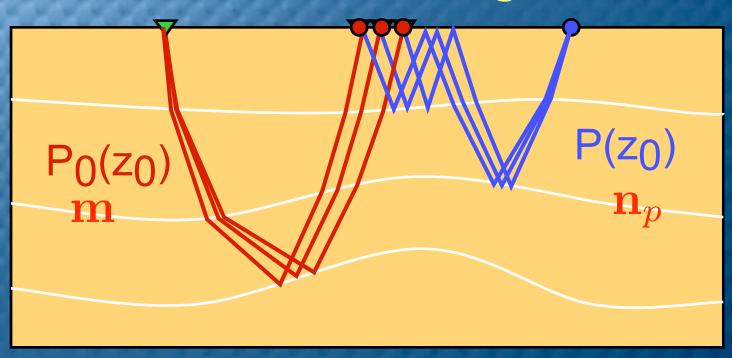
Motivation

Multiple removal:

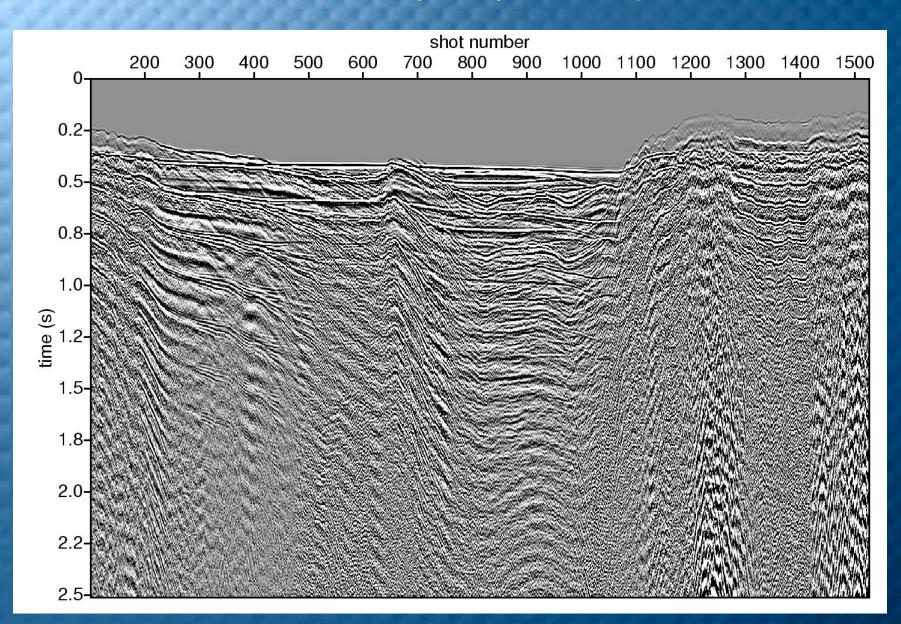
- **★ Geometry effects in 3D surface**multiple prediction
- ★ New approach to adaptive subtraction robust under
 - phase rotations
 - misalignments

Surface multiple elimination

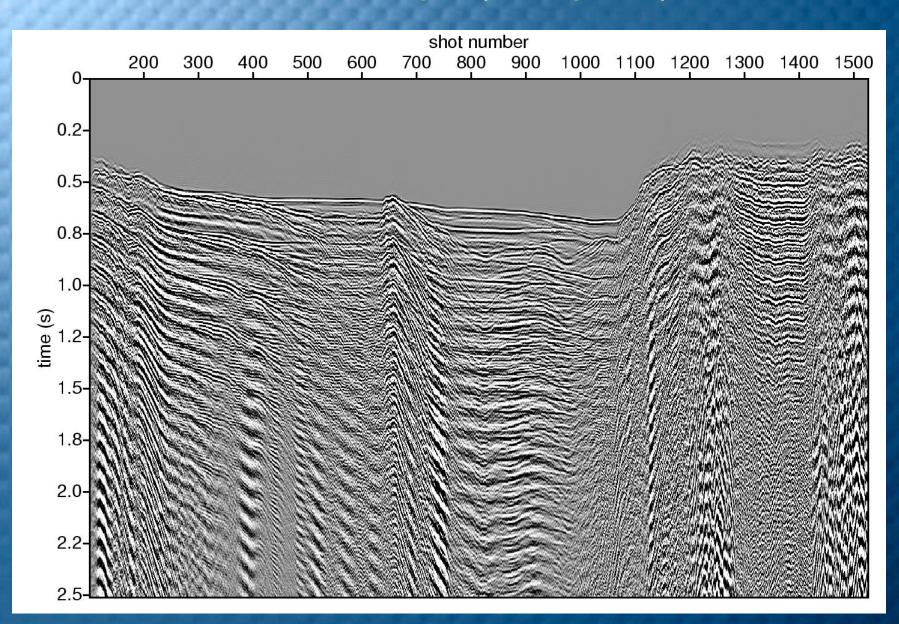
Multiple prediction: data convolution along surface



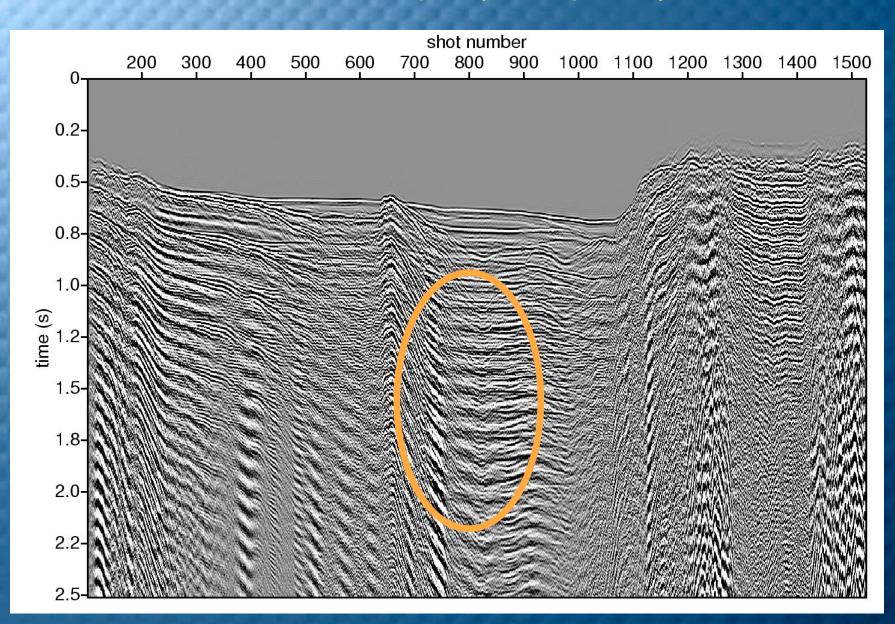
Common offset (500 m) with multiples



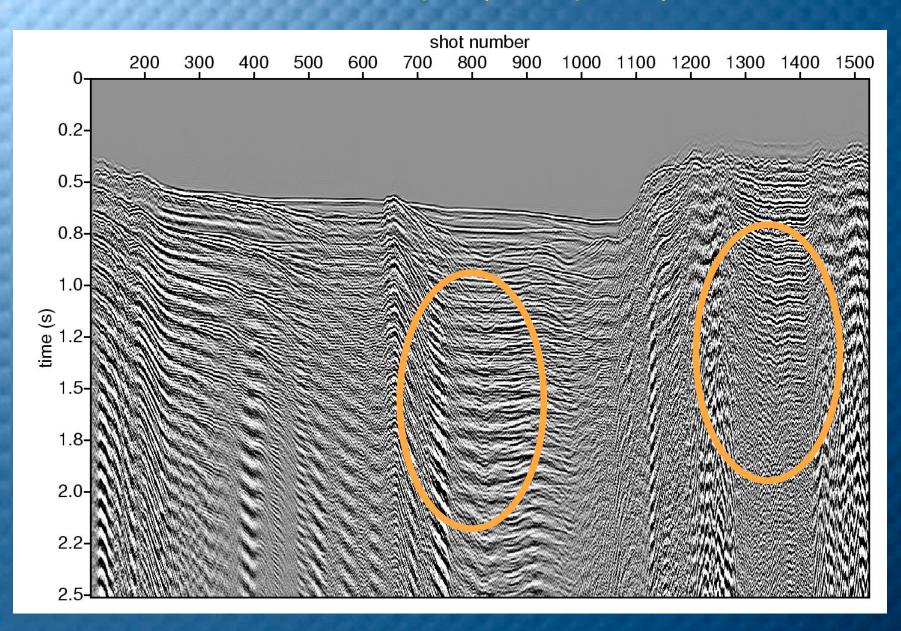
Predicted multiples (no adaptation)



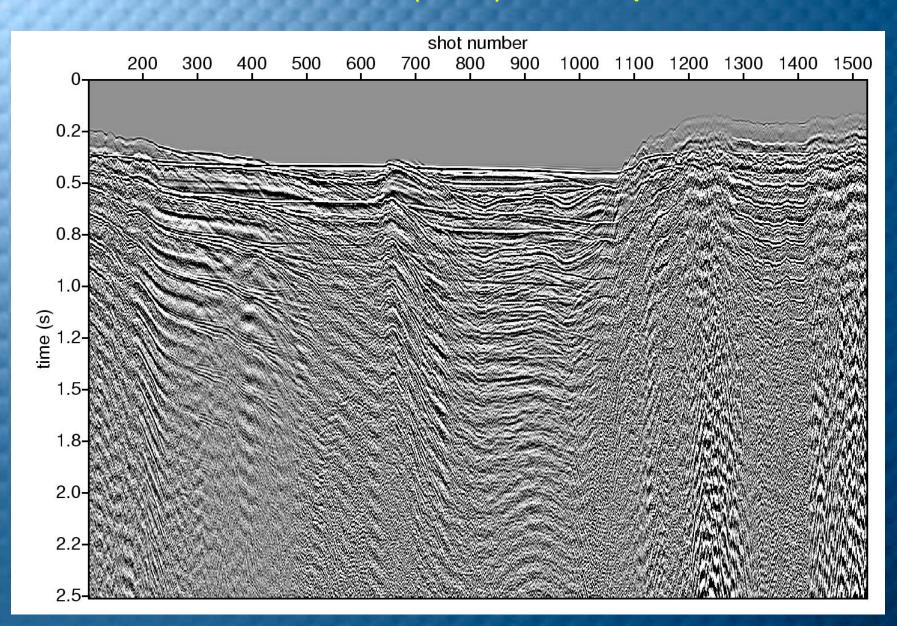
Predicted multiples (no adaptation)



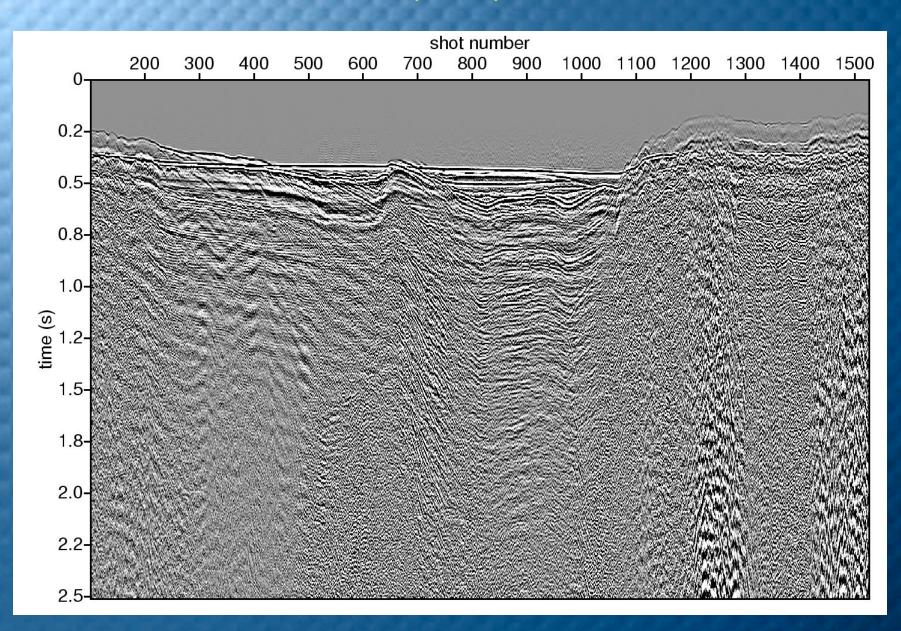
Predicted multiples (no adaptation)



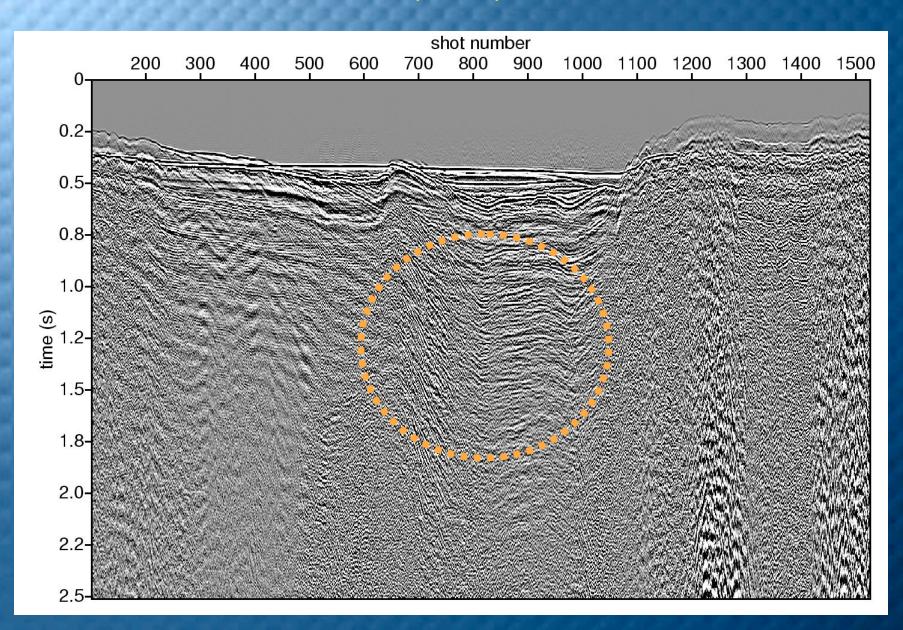
Common offset (500 m) with multiples



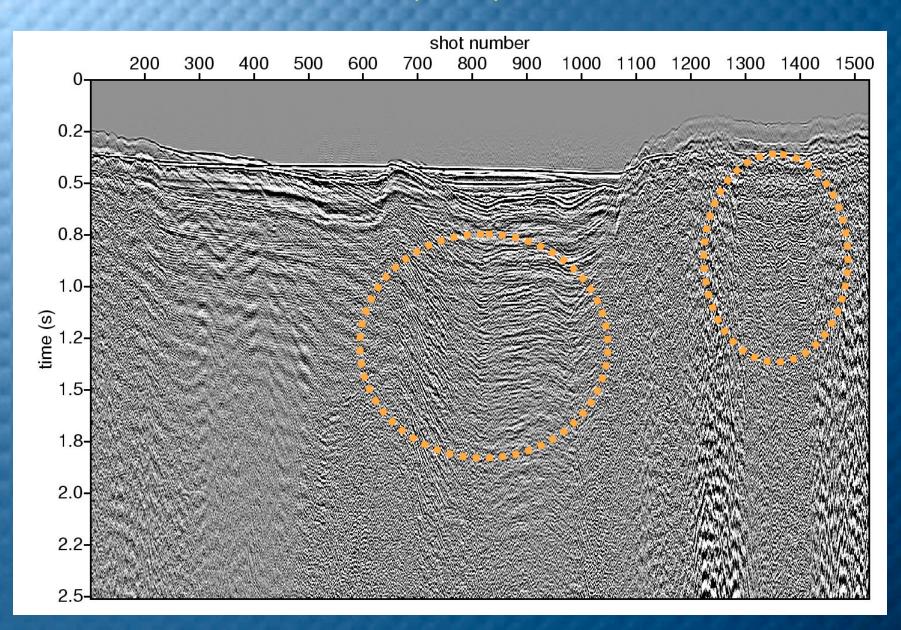
Common offset (500 m) after 2D SRME



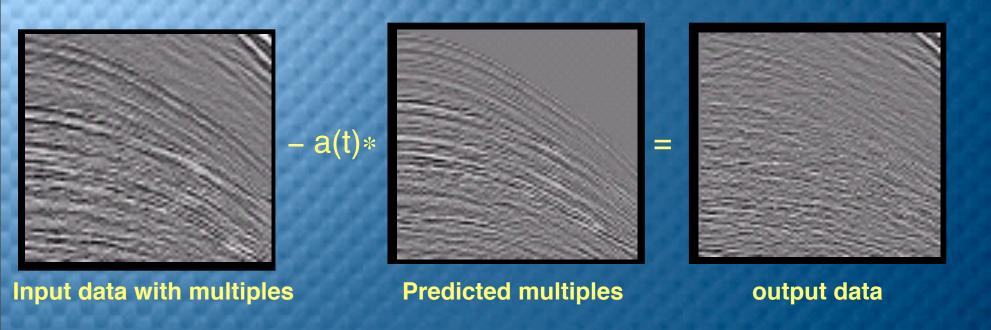
Common offset (500 m) after 2D SRME



Common offset (500 m) after 2D SRME



Least-squares subtraction



Adaptive subtraction based on minimum energy in the output

Adaptive subtraction

Matched filter:

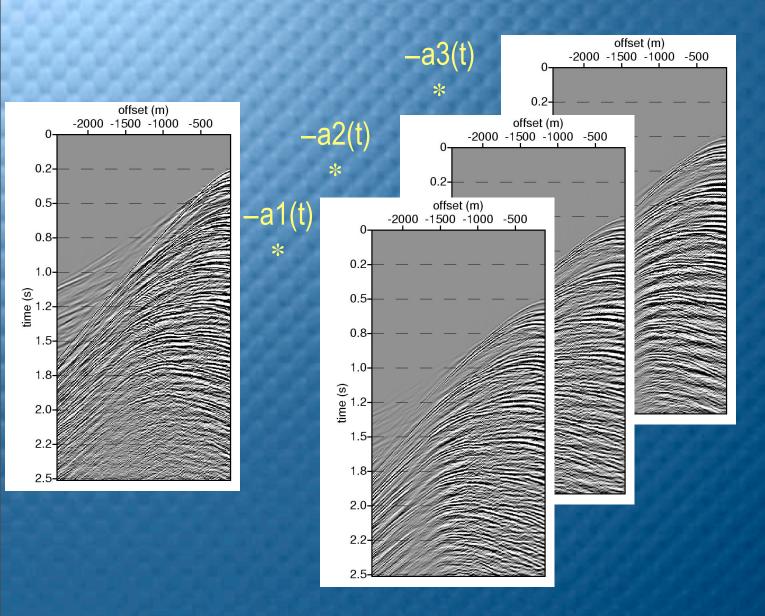
n

matched filter

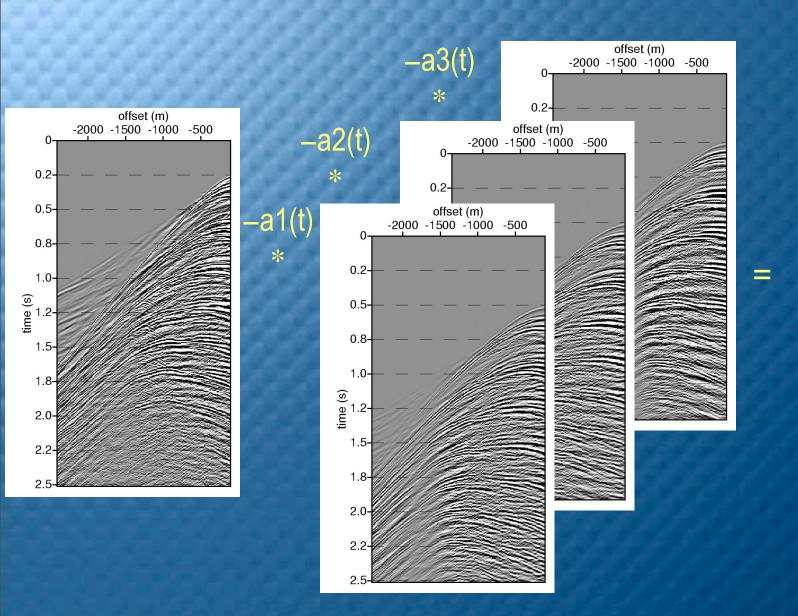
- residue is the denoised data
- risk of over fitting

May loose primary reflection events ...

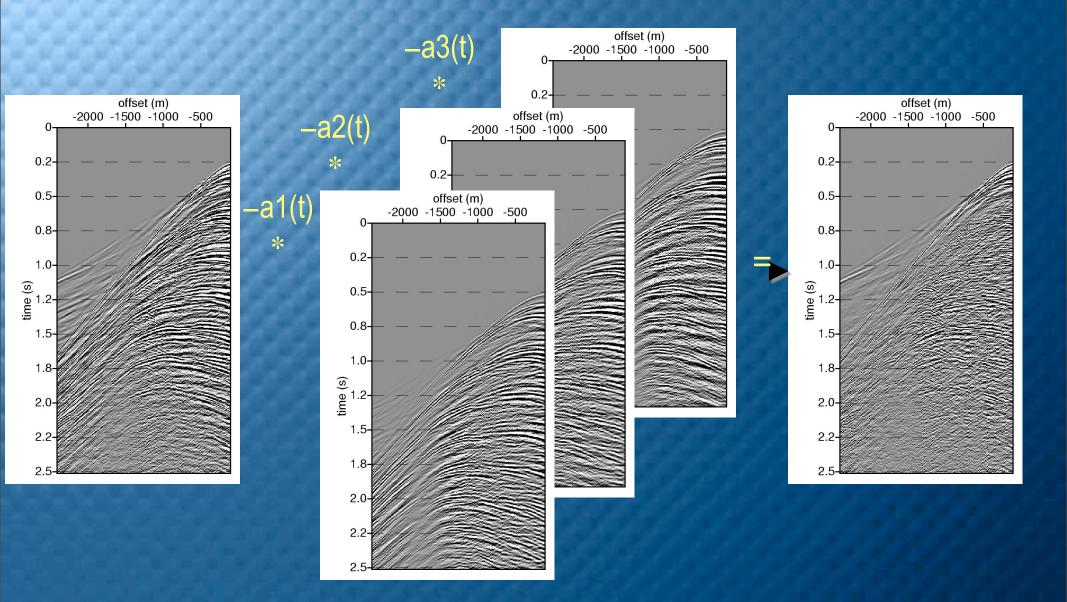
Multi-gather subtraction



Multi-gather subtraction



Multi-gather subtraction



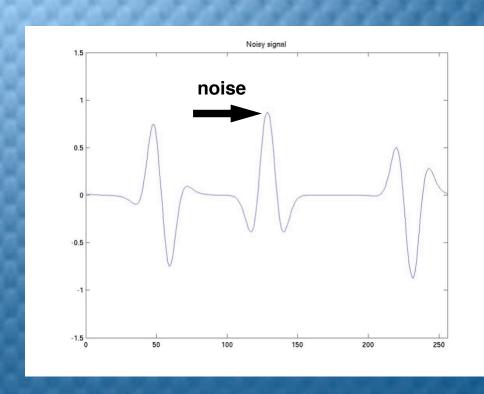
Least-squares subtraction

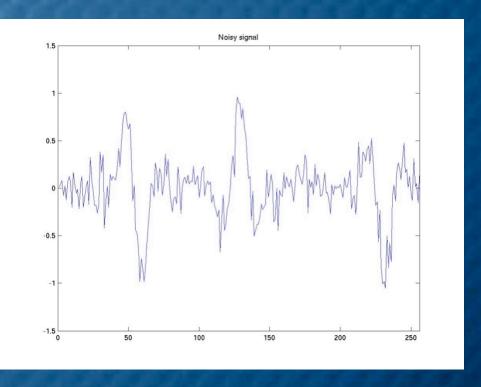
- Can be effective in many situations
- Relatively fast and easy to apply
- But minimum energy assumption is not always valid
- Danger of distorting the primaries due to overfitting
- Explore other subtraction domains that

Other domains

- sparse & local
- relatively insensitive to
 - phase rotations
 - misalignments
- almost diagonalize Covariance operator of multiples & primaries

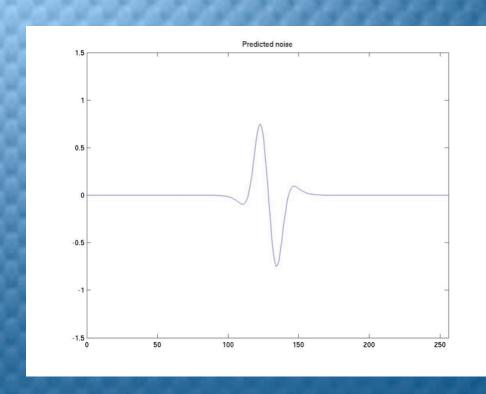
We know that wavelets are unconditional bases

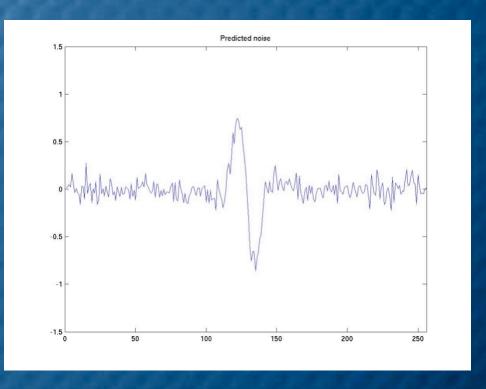




signal + coherent noise

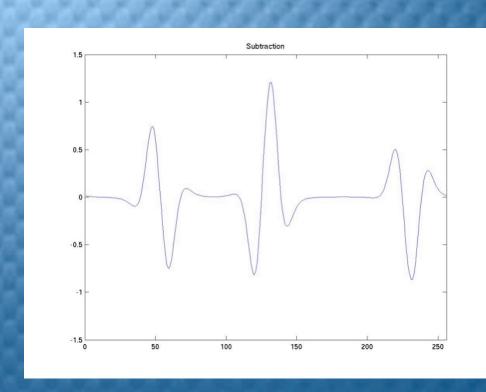
signal + coherent & incoherent noise

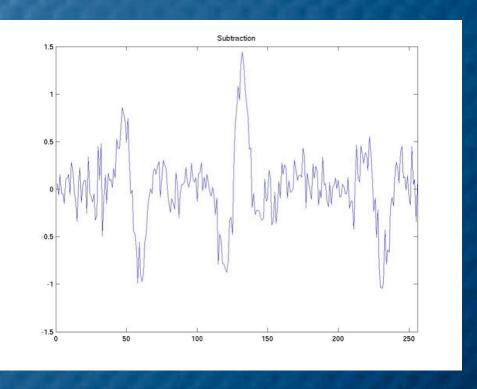




'wrongly' predicted noise

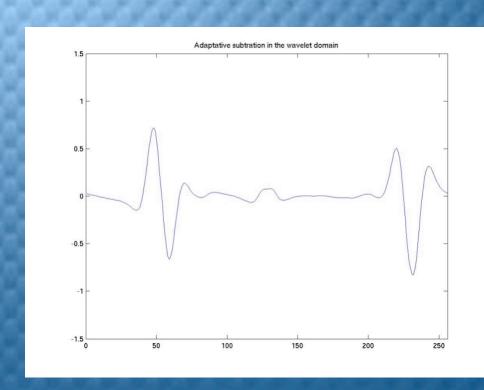
'wrongly' noisy predicted noise

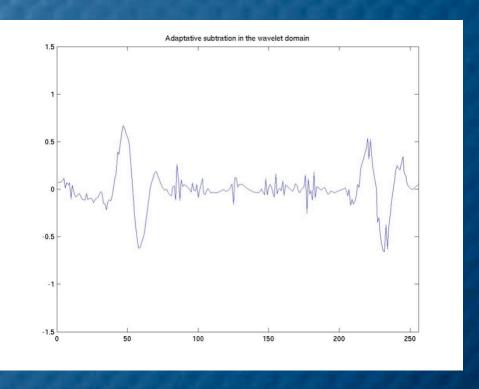




noise-free subtraction

noisy subtraction





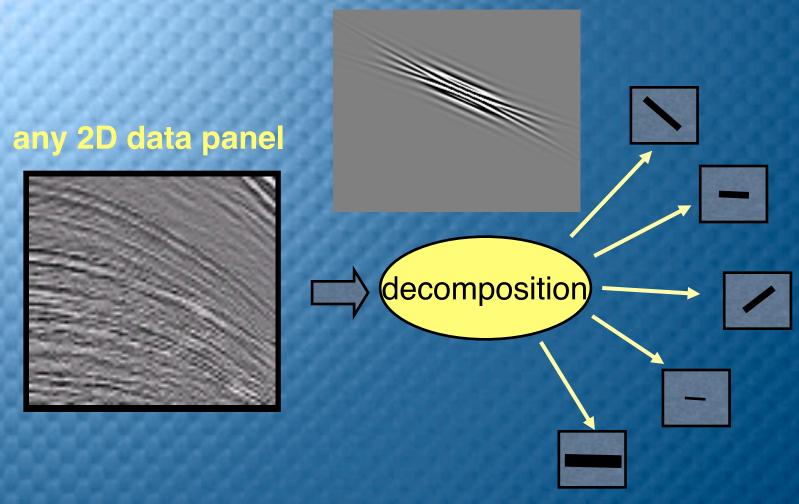
noise-free adaptive subtraction

noisy adaptive subtraction

Wavelets

- Represent piece-wise smooth functions at "no" additional cost.
- Do not have to know where the singularities are.
- Are unconditional basis.
- Only good for point-scatterers or horizon/vertically-aligned reflectors.
- Lack directional selectivity.
- Do NOT work well with waves.

Curvelet domain



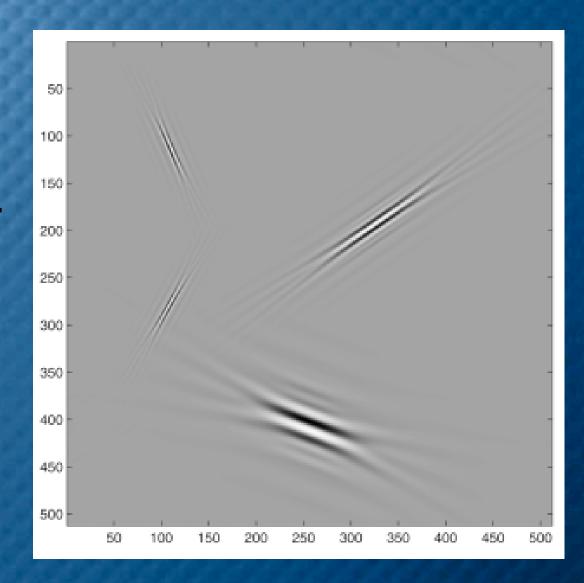
- Almost orthogonal decomposition into multiscale basis functions with local frequency and local dip properties
- Natural basis for wave equations
- Consist of plane wavelets invariant under convolution

Curvelet domain

- Curvelet transform developed by Candes and Donoho (2002)
- Geophysical applications:
 - Multiple & ground-roll suppression
 - New imaging algorithms
 - 4D data matching
- Curvelets sense local dip and local frequency content and can discriminate on these properties

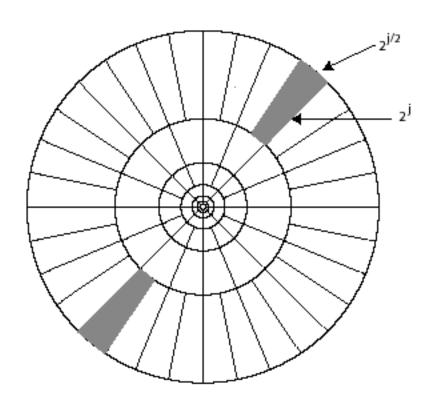
Why curvelets

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



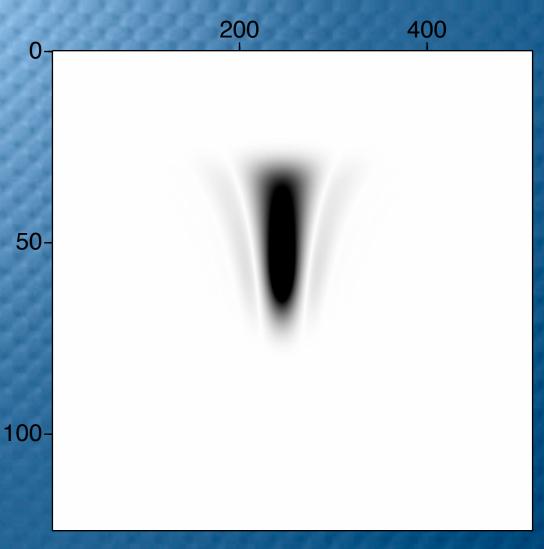
Why curvelets

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



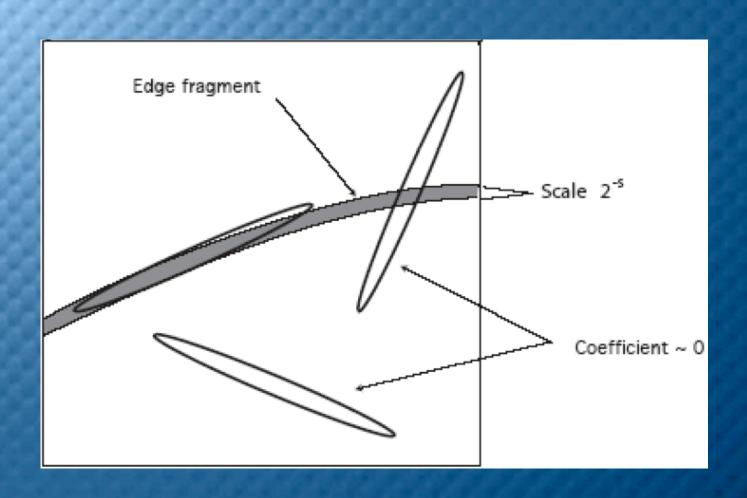
second dyadic partitioning

Why curvelets

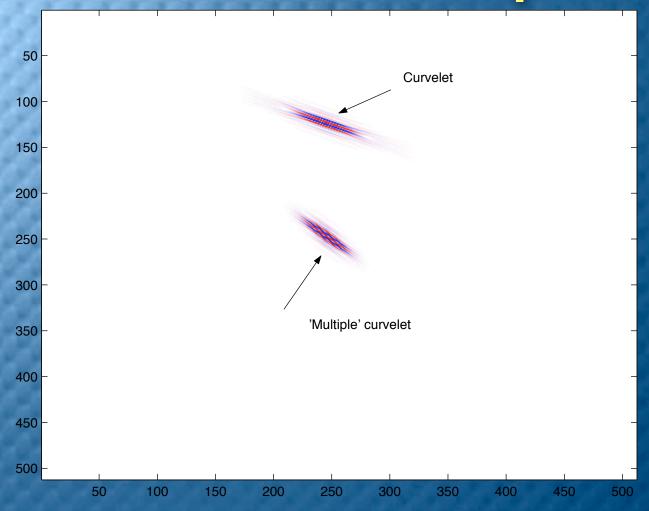


Curvelet in FK-domain

Dot products



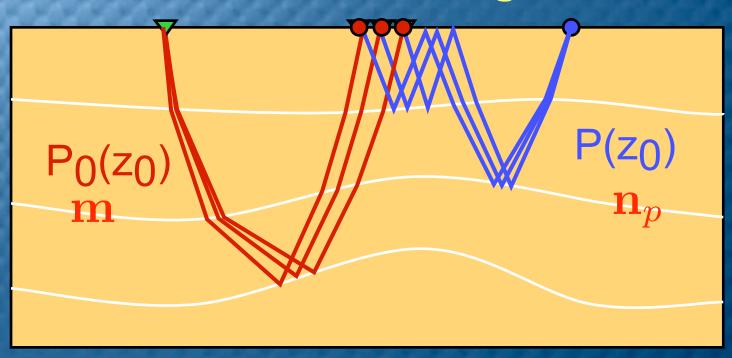
Curvelet 'Multiples'



- Almost diagonalize Green's functions (Candes & Demanet '04)
- Natural basis for wave equations
- Invariant under convolution, i.e. 'multiple multiple' = curvelet-like

Surface multiple elimination

Multiple prediction: data convolution along surface



What do we do?

Use as alternative to matched adapt. sub.

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

to denoise

$$\mathbf{d}$$
 = \mathbf{m} + \mathbf{n} noise noise-free

with a simple *mute* with λ control parameter

$$\lambda = 3 \leftrightarrow 90\%$$
 confidence interval

Non-linear adaptive subtraction

Extend to colored noise:

$$d$$
 = m + n noise-free

Solve

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

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:

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

'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

$$\hat{\mathbf{m}} = \mathbf{B}^\dagger \mathbf{\Gamma} \Theta_\lambda \left(\mathbf{\Gamma}^{-1} \mathbf{B} \mathbf{d}
ight) = \mathbf{B}^\dagger \Theta_{oldsymbol{\lambda} \mathbf{\Gamma}} \left(\mathbf{ ilde{d}}
ight).$$

No matched filter required!

Non-linear estimation

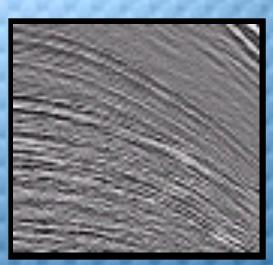
Hard thresholding for p=0:

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

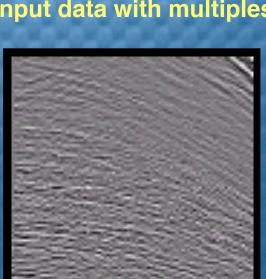
Soft thresholding for p=1:

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

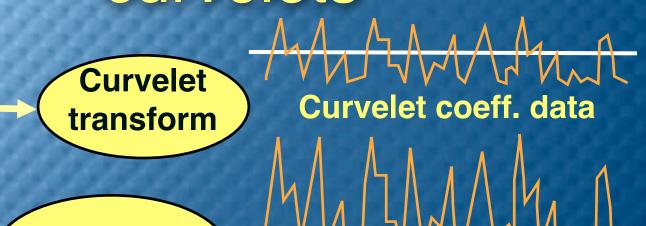
Threshold



Input data with multiples



Filtered input data



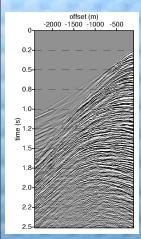


Apply mute filter

Curvelet coeff. pred. multiples

Bd

 $\lambda |\mathbf{Bn}_p|$



Input data vith multiples

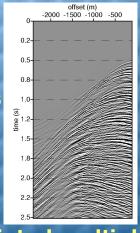


Threshold

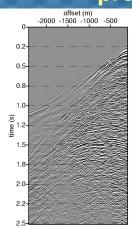
Curvelet coeff. data



Curvelet coeff. pred. multiples



predicted multiples





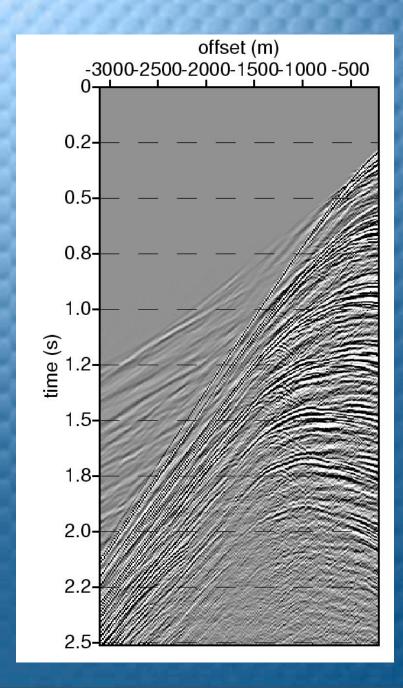


Curvelet coeff. primaries

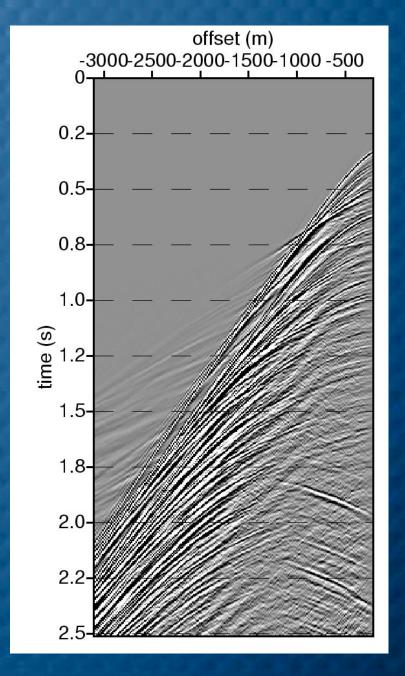
 $oldsymbol{\Theta}_{\lambda\Gamma}\left(\mathbf{Bd}
ight)$

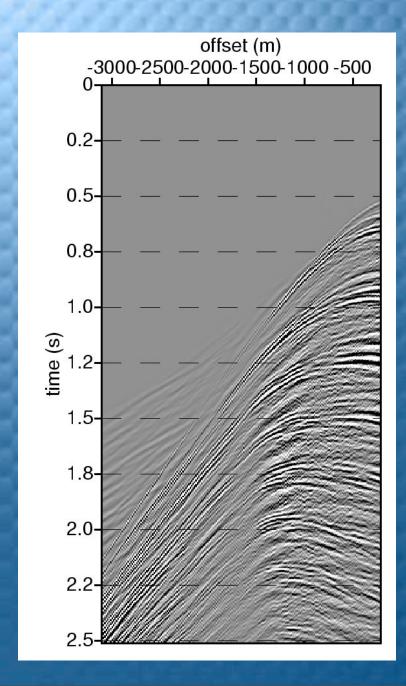
Bd

Filtered input data

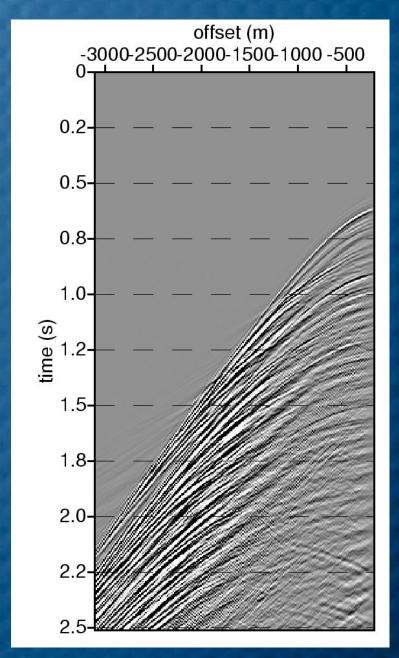


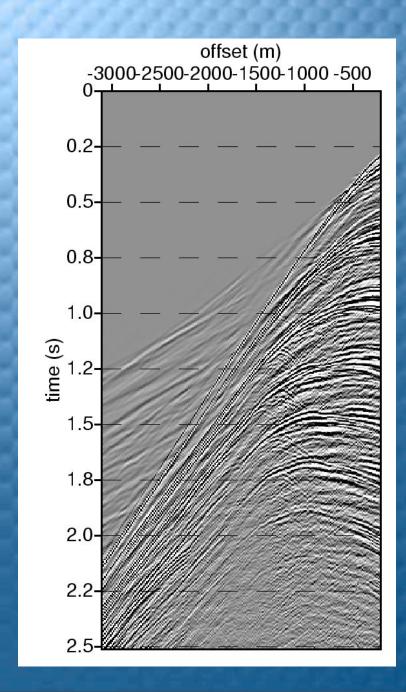
Input with multiples



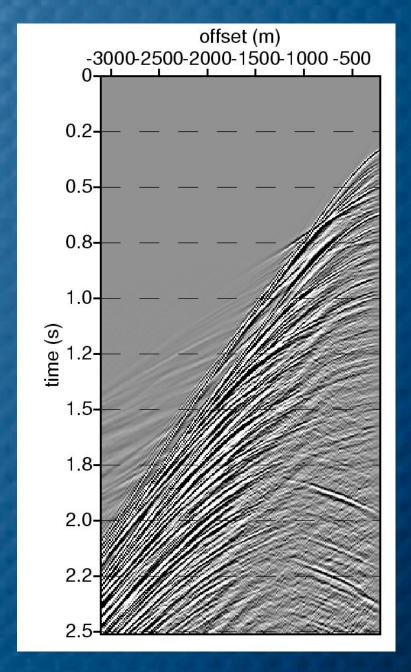


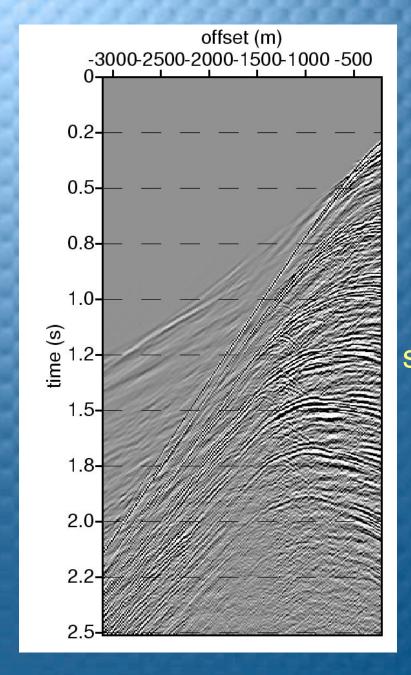
predicted multiples



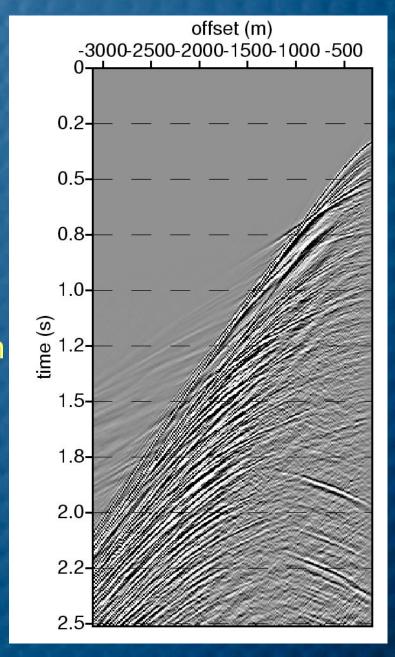


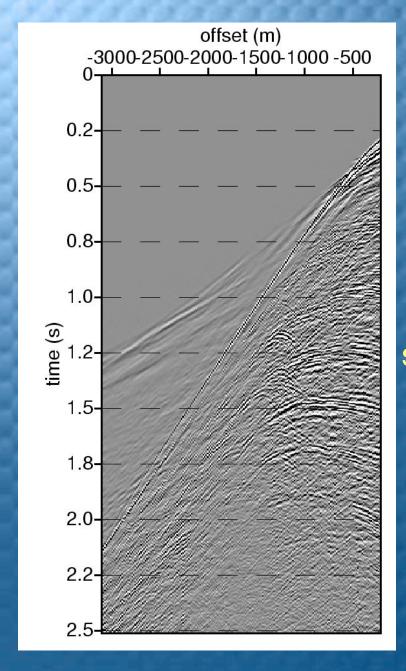
Input with multiples



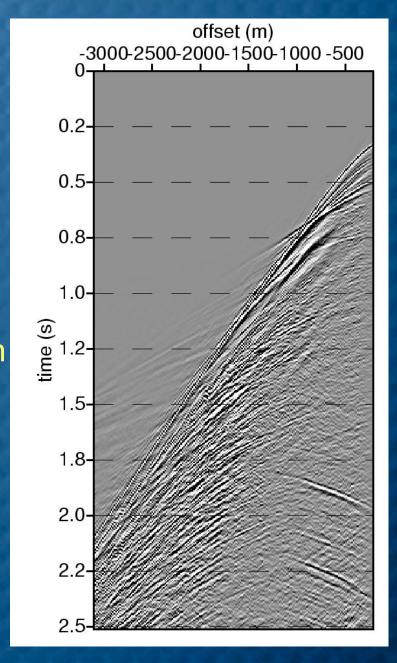


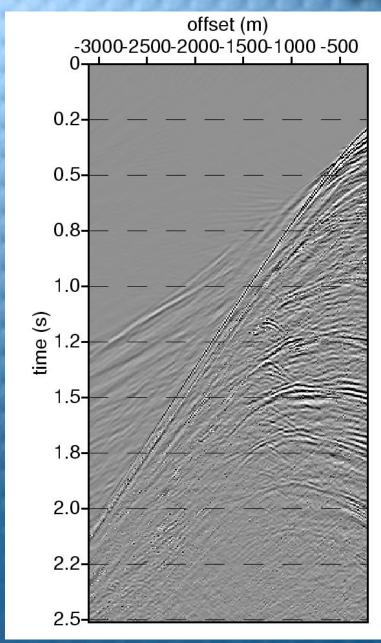
Output
SRME
L2
subtraction



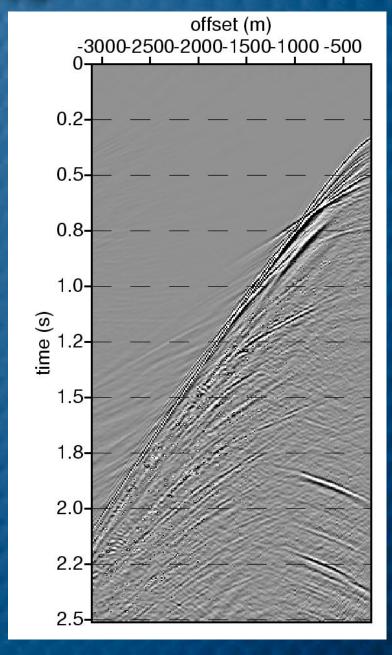


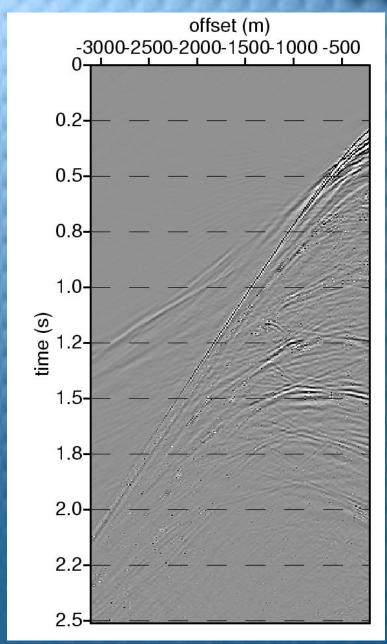
Output SRME multi-L2 subtraction



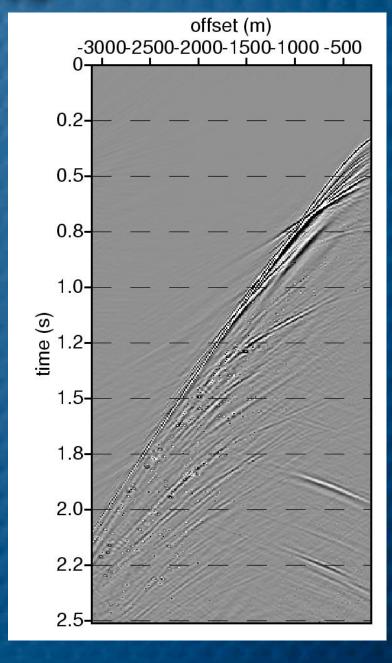


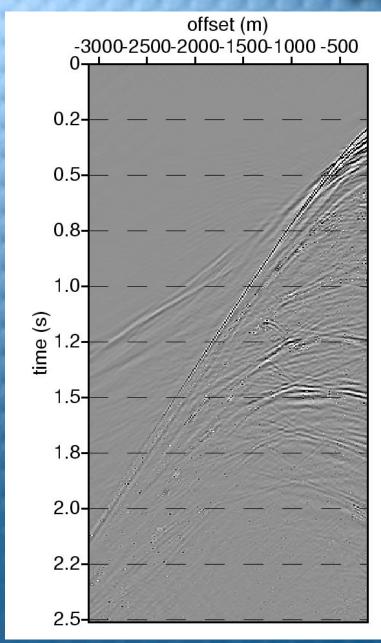
Output curvelet filtering





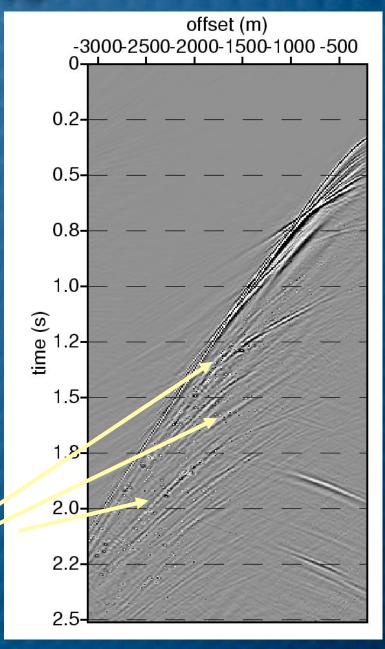
Output curvelet filtering with stronger threshold

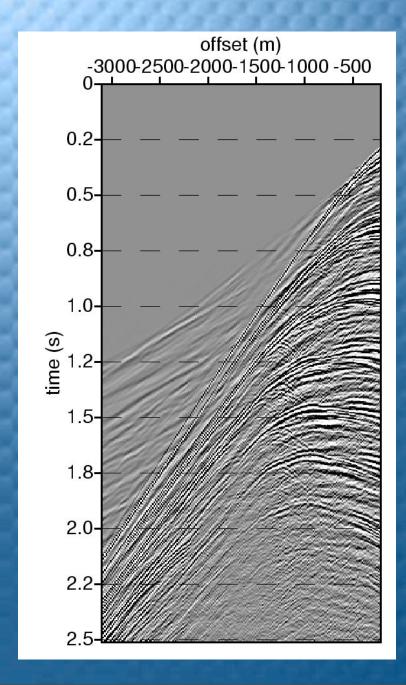




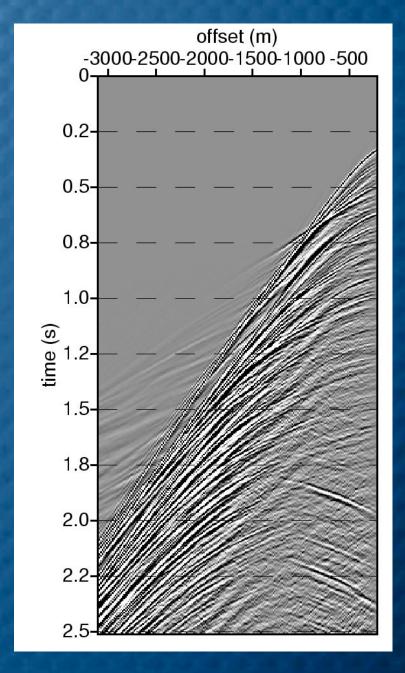
Output curvelet filtering with stronger threshold

Preserved primaries





Input with multiples



Noise prediction

Used 3D SRME to predict noise

Can use other noise predictions (e.g. Radon)

We do NOT subtract rather mute

- put to zero or preserve coherent features
- * less sensitive to errors

Impose additional (sparseness) norms

Preserve the edges & primaries!

Global optimization

Formulate constrained optimization:

$$\hat{\mathbf{m}}: \min_{m} J(\mathbf{m}) \quad \text{s.t.} \quad |\tilde{\mathbf{m}} - \hat{\tilde{\mathbf{m}}}_{0}|_{\mu} \leq \mathbf{e}_{\mu}, \quad \forall \mu$$

with

$$\hat{\mathbf{m}}_0 = \mathbf{B}^\dagger \Theta_{oldsymbol{\lambda}oldsymbol{\Gamma}} \left(ilde{\mathbf{d}}
ight)$$

and with e_{μ} threshold and noisedependent *tolerance* on curvelet coeff.

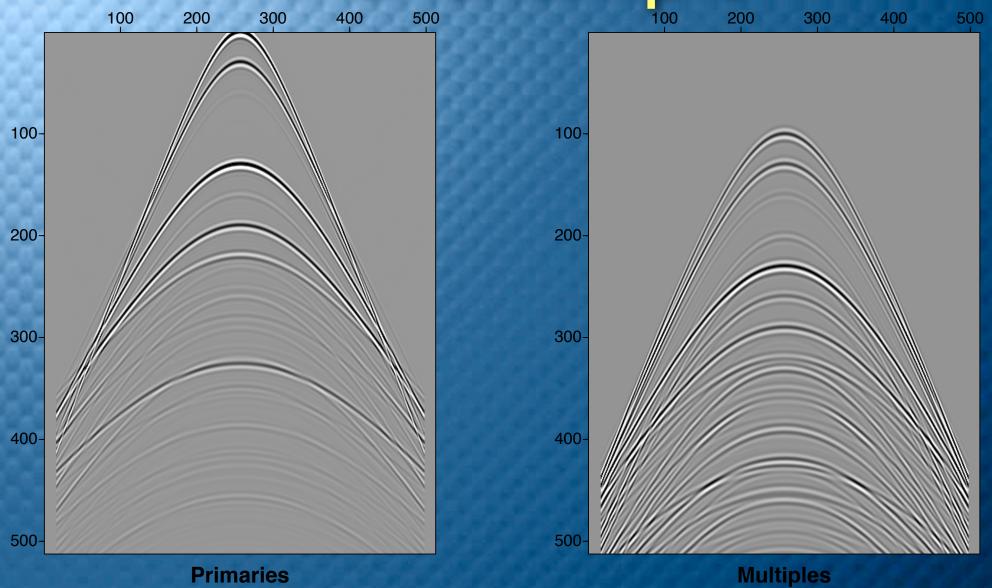
Global optimization

Set tolerances

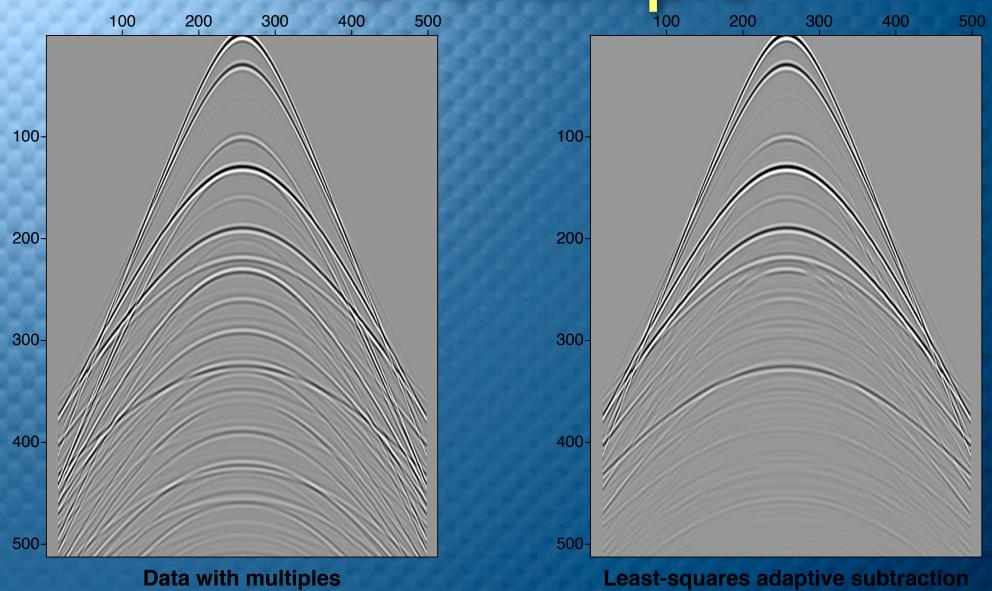
$$\mathbf{e}_{\mu} = egin{cases} \Gamma_{\mu} & ext{if} & |\mathbf{\hat{\tilde{m}}}_{0}|_{\mu} & \geq & |\lambda\Gamma|_{\mu} \ oldsymbol{\lambda}\Gamma_{\mu} & ext{if} & |\mathbf{\hat{\tilde{m}}}_{0}|_{\mu} & < & |\lambda\Gamma|_{\mu} \end{cases}$$

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

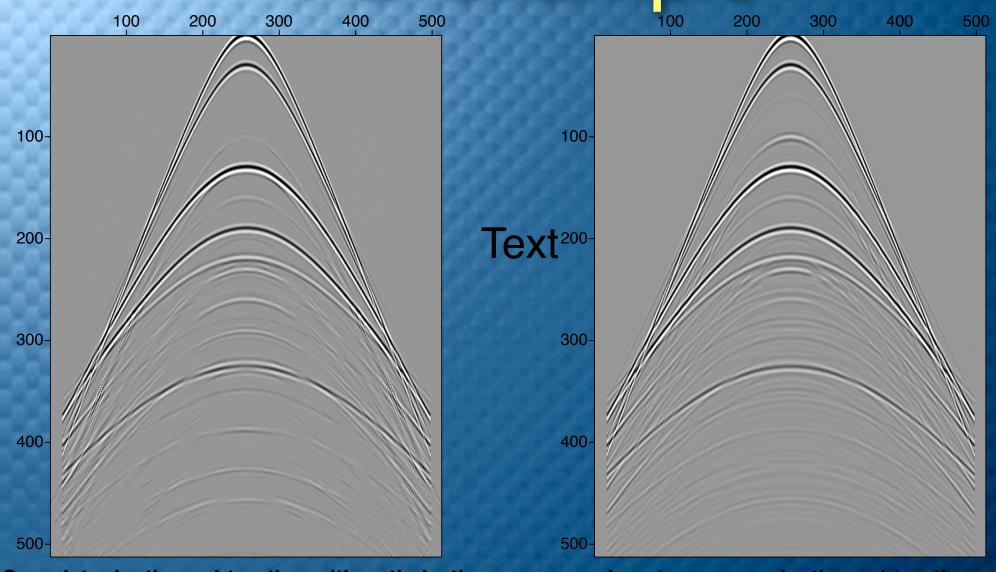
1-D example



1-D example



1-D example



Curvelet adaptive subtraction with optimization

Least-squares adaptive subtraction

Conclusions

- For 3D SRME the acquisition geometry determines the prediction quality and possibilities
- When it is a second to a second the adaptive subtraction toolbox and hence the quality of the end result
- Non-linear Curvelet thresholding adds robustness
- Finding a norm that enhances the wave-front set is an open problem
- Estimating the seismic wavelet should be feasible

Acknowledgements

Frank Kempe (Cray) for conducting the 3D SRME tests on the Cray-X1 and George Stephenson (Cray) for his support

Candes & Donoho for making their Curvelet code available.

Partially supported by a NSERC Grant.

