

# Recent developments in primary-multiple separation

Felix J. Herrmann



## Motivation

## Motivation

Primary-multiple separation step is crucial

- moderate prediction errors
- 3-D complexity & noise

Inadequate separation leads to

- remnant multiple energy
- deterioration primary energy

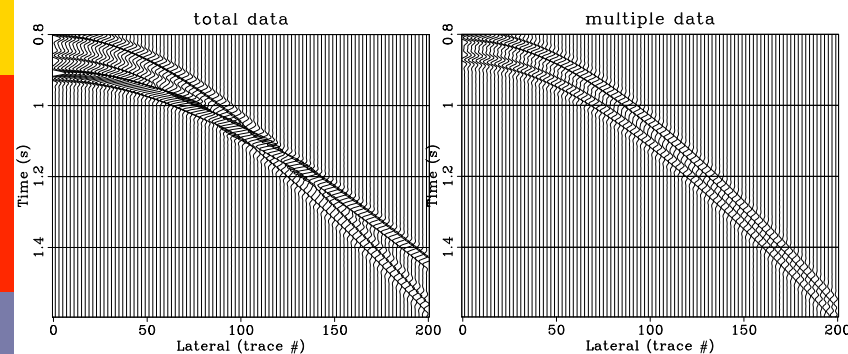
Introduce a transform-based technique

- stable
- insensitive to moderate shift, phase rotations

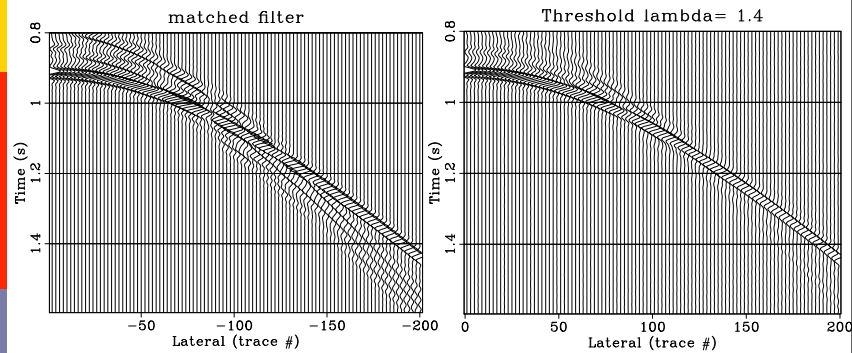
Exploit sparsity and parameterization transformed domain



## Move-out error



## Move-out error



## The problem

Sparse signal model:

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \mathbf{n},$$

with

$$\mathbf{A} = [\mathbf{A}_1 \quad \mathbf{A}_2] \quad \text{and} \quad \mathbf{x}_0 = [\mathbf{x}_{01} \quad \mathbf{x}_{02}]^T$$

- augmented synthesis and sparsity vectors
- index 1 <-> primary
- index 2 <-> multiple

## The solution

The weighted norm-one optimization problem:

$$\mathbf{P}_w : \begin{cases} \min_{\mathbf{x}} \|\mathbf{x}\|_{w,1} & \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon \\ \hat{\mathbf{s}}_1 = \mathbf{A}_1 \hat{\mathbf{x}}_1 & \text{and } \hat{\mathbf{s}}_2 = \mathbf{A}_2 \hat{\mathbf{x}}_2 \\ \text{given: } \check{\mathbf{s}}_2 & \text{and } \mathbf{w}(\mathbf{y}, \check{\mathbf{s}}_2) \end{cases}$$

with

$$\mathbf{w} := [\mathbf{w}_1, \mathbf{w}_2]^T$$

$$\mathbf{A} := [\mathbf{C}^T, \mathbf{C}^T]$$

$$\check{\mathbf{s}}_2 := \text{predicted multiples}$$

$$\check{\mathbf{s}}_1 := \mathbf{S} - \check{\mathbf{S}}_2$$

## Solution cont'd

The weights

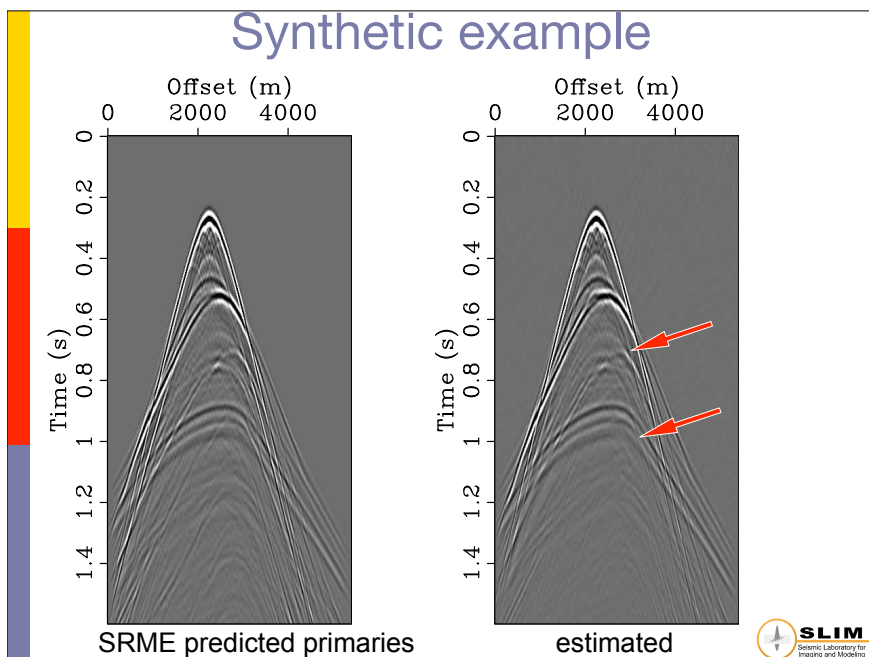
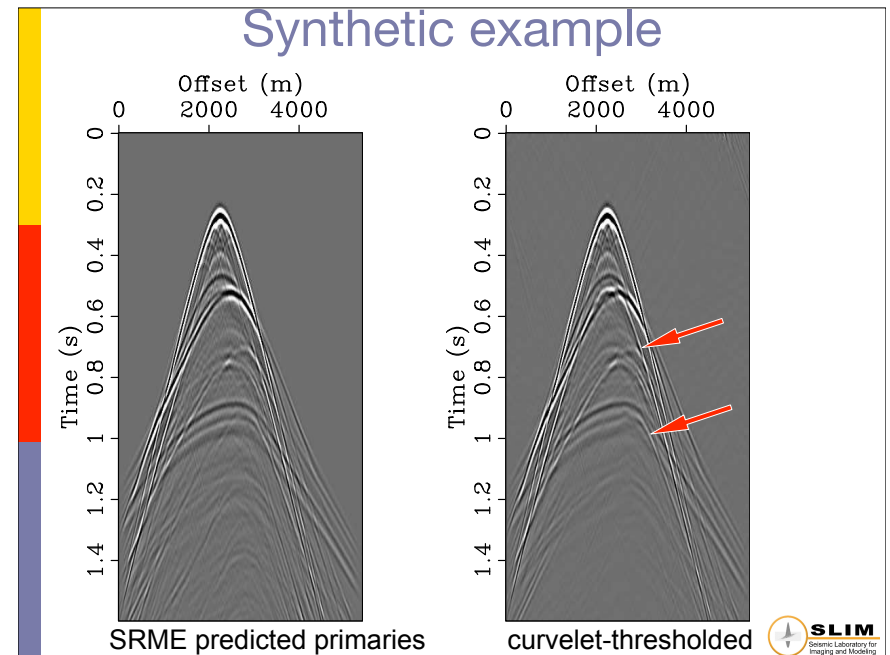
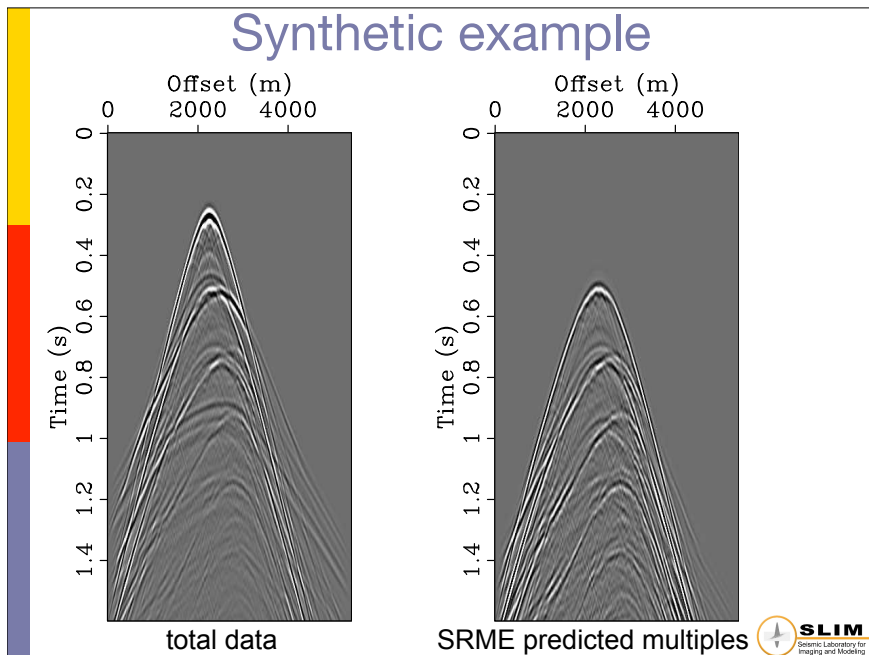
$$\begin{cases} \mathbf{w}_1 := \max(\sigma \cdot \sqrt{2 \log N}, C_1 |\check{\mathbf{u}}_1|) \\ \mathbf{w}_2 := \max(\sigma \cdot \sqrt{2 \log N}, C_2 |\check{\mathbf{u}}_2|) \end{cases}$$

with

$$\check{\mathbf{u}}_1 \approx \mathbf{C}\check{\mathbf{s}}_1$$

$$\check{\mathbf{u}}_2 \approx \mathbf{C}\check{\mathbf{s}}_2$$

- during minimization signal components are driven apart
- curvelet compression helps
- separates on the basis of position, scale and direction



## Problem

Single thresholding is too aggressive  
 Iterative method tends to drive all energy to the primaries

- nontrivial weighting
- nontrivial stop criterion

**Solutions:**

- Non-adaptive Bayesian framework motivated by results from "Blind source deconvolution"
  - sparsity & decorrelation in the curvelet domain
- Adaptive curvelet-domain matched filtering
  - smoothness in the phase space  $\Leftrightarrow$  curvelet domain

## Outline

**Non-adaptive** curvelet domain primary-multiple separation

- formulation of the primary-multiple separation problem
- the curvelet transform
- Bayesian formulation, taking inaccurate predictions into account
- Solution with iterative thresholding algorithm

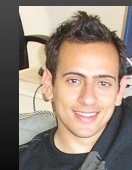
**Adaptive** curvelet-domain matched filtering

- formulation of the SRME-base primary-multiple separation problem
- Phase space formulation taking nonstationary amplitude variations into account
- Curvelet-base matched filtering by imposing symbol smoothness in phase space



## Non-adaptive curvelet domain primary-multiple separation

Joint work with Eric Verschuur, Deli Wang, Rayan Saab and Ozgur Yilmaz.



## Problem formulation

Consider measurements as the sum of primaries, multiples and noise

$$\mathbf{b}_1 = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n}$$

Given a possibly erroneous prediction for the multiples (e.g. via SRME)

$$\mathbf{b}_2 = \mathbf{s}_2 + \mathbf{n}_2$$

Required: To estimate and hence separate primaries and multiples.



## Curvelet transform

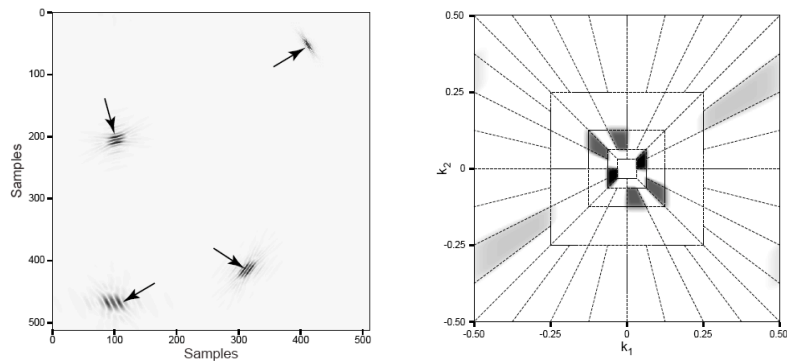
Formulate the separation problem in the curvelet domain where both the primaries and multiples can be modeled as **sparse** vectors.

Curvelets:

- little plain waves, multiscale and multidirectional, optimal for detecting wavefronts
- sparse on primaries and multiples
- parameterized by position, angle and scale (frequency band)



## Curvelet transform



## Problem formulation revisited

Let  $\mathbf{A} = \mathbf{C}^T$  be the inverse Curvelet transform.

Let  $\mathbf{x}_1$  and  $\mathbf{x}_2$  be the curvelet coefficients of the primaries and multiples, respectively.

Write for the forward model

$$\begin{aligned} \mathbf{b}_2 &= \mathbf{s}_2 + \mathbf{n}_2 \implies \mathbf{b}_2 = \mathbf{A}\mathbf{x}_2 + \mathbf{n}_2 \\ \mathbf{b}_1 &= \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{n} \implies \mathbf{b}_1 = \mathbf{A}\mathbf{x}_1 + \mathbf{n} - \mathbf{n}_2. \end{aligned}$$

## Problem formulation revisited

### Objective:

To estimate the primaries and multiples by estimating their curvelet coefficients.

### Method:

- derive variational problem for  $\mathbf{x}_1$  and  $\mathbf{x}_2$  with a Bayesian formulation.
- impose sparsity as a **prior** on the coefficients
- solve for the coefficients to fit the data and the prior

## Maximum *A posteriori* estimation

We want to maximize  $P(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{b}_1, \mathbf{b}_2)$

Apply Bayes' rule

$$\begin{aligned} P(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{b}_1, \mathbf{b}_2) &= \frac{P(\mathbf{x}_1, \mathbf{x}_2)P(\mathbf{b}_1 | \mathbf{x}_1, \mathbf{x}_2)P(\mathbf{b}_2 | \mathbf{x}_1, \mathbf{x}_2)}{P(\mathbf{b}_1, \mathbf{b}_2)} \\ &= \frac{P(\mathbf{x}_1, \mathbf{x}_2)P(\mathbf{n})P(\mathbf{n}_2)}{P(\mathbf{b}_1, \mathbf{b}_2)}. \end{aligned}$$

.... bottom line worry about the numerator:

$$\arg \max_{\mathbf{x}_1, \mathbf{x}_2} P(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{b}_1, \mathbf{b}_2) = \arg \max_{\mathbf{x}_1, \mathbf{x}_2} P(\mathbf{x}_1, \mathbf{x}_2)P(\mathbf{n})P(\mathbf{n}_2).$$

## Key assumptions

Curvelet coefficients of seismic data are sparse, i.e., mostly close to zero with few important non-zero coefficients,

- reasonable distribution to impose as a **prior** on the curvelet coefficients is a Laplacian/Cauchy distribution  $\iff \|\mathbf{x}\|_1$
- noise and prediction errors are modeled as Gaussian noise  $\iff \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$



## Optimization problem

Rewrite,

$$\arg \max_{\mathbf{x}_1, \mathbf{x}_2} P(\mathbf{x}_1, \mathbf{x}_2 | \mathbf{b}_1, \mathbf{b}_2) = \arg \min_{\mathbf{x}_1, \mathbf{x}_2} f(\mathbf{x}_1, \mathbf{x}_2)$$

We want to minimize:

$$f(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1\|_{1, \mathbf{w}_1} + \|\mathbf{x}_2\|_{1, \mathbf{w}_2} + \|\mathbf{A}\mathbf{x}_2 - \mathbf{b}_2\|_2^2 + \eta \|\mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) - (\mathbf{b}_1 + \mathbf{b}_2)\|_2^2,$$

with the weights:

$$\begin{aligned} \mathbf{w}_1 &= \lambda_1 |\mathbf{A}^T \mathbf{b}_2| \\ \mathbf{w}_2 &= \lambda_2 |\mathbf{A}^T \mathbf{b}_1| \end{aligned}$$



## Optimization problem properties

Control parameters:

- $\eta$  controls the **tradeoff** between fitting the total data and fitting the predicted multiples
  - $\eta \rightarrow 0 \iff$  denoise multiples
  - $\eta \rightarrow \infty \iff$  old formulation
- $\eta$  controls trust in the prediction versus total data
- $\lambda_1$  and  $\lambda_2$  control the **sparsity** of the coefficient vectors of the primaries and multiples
- ratio  $\lambda_{1,2}$  versus  $\eta$  controls **sparsity versus data mismatch**



## Separation algorithm

Minimize the objective function with the iterative algorithm:

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

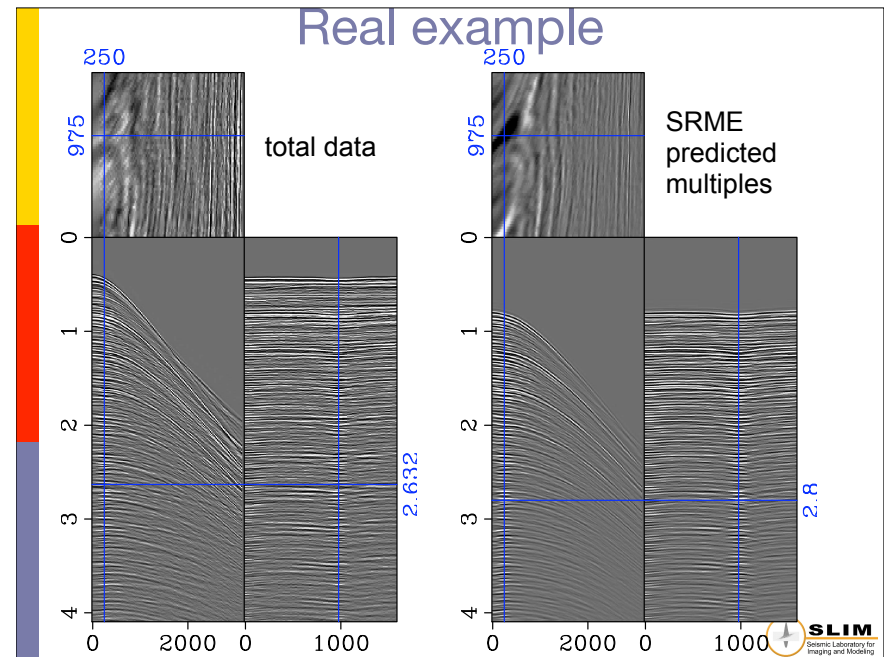
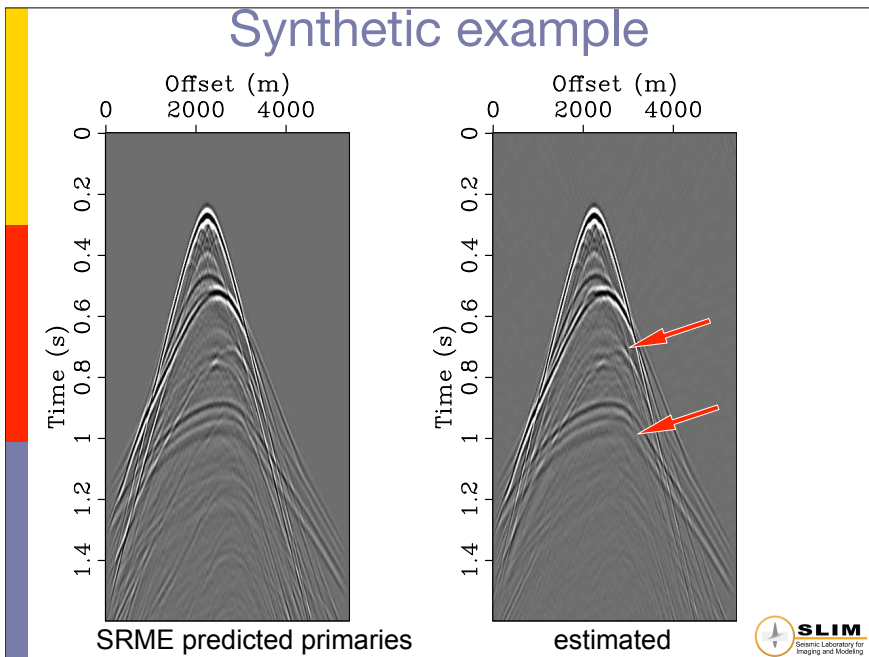
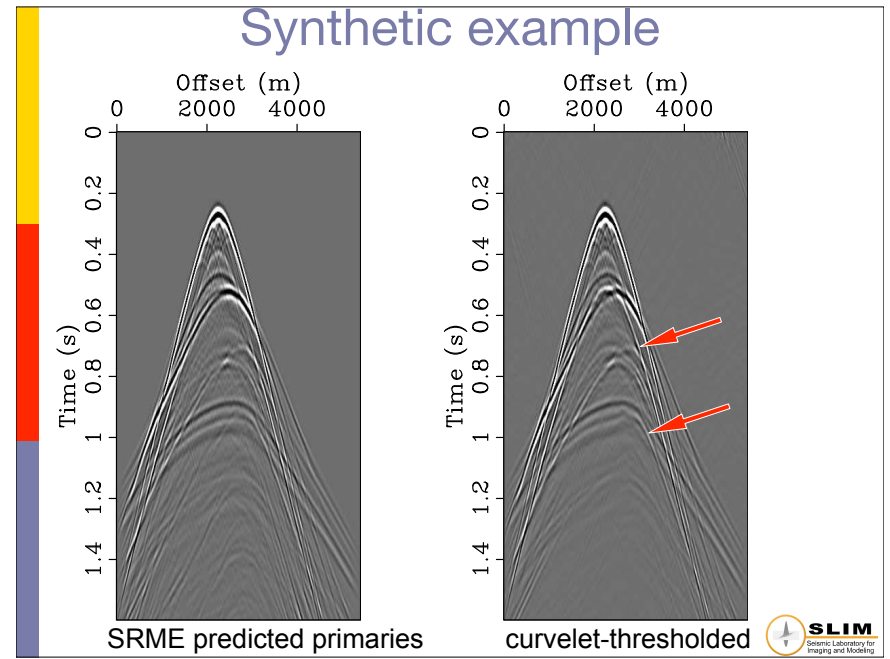
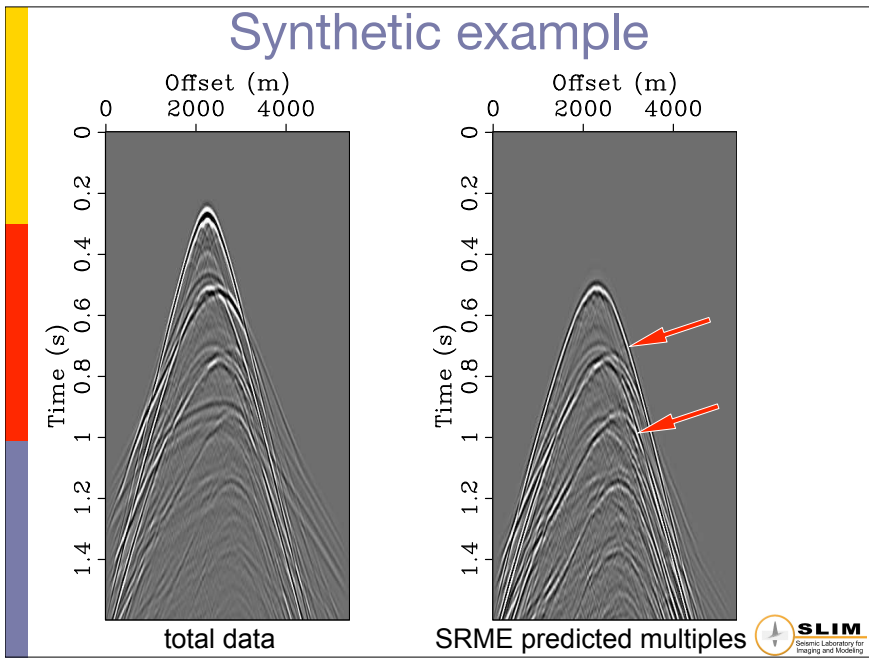
where

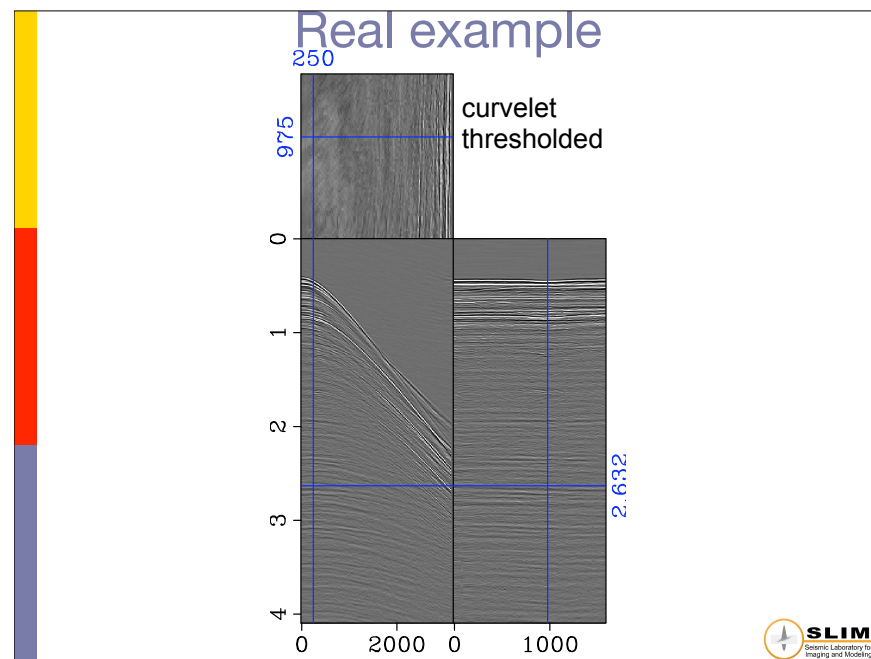
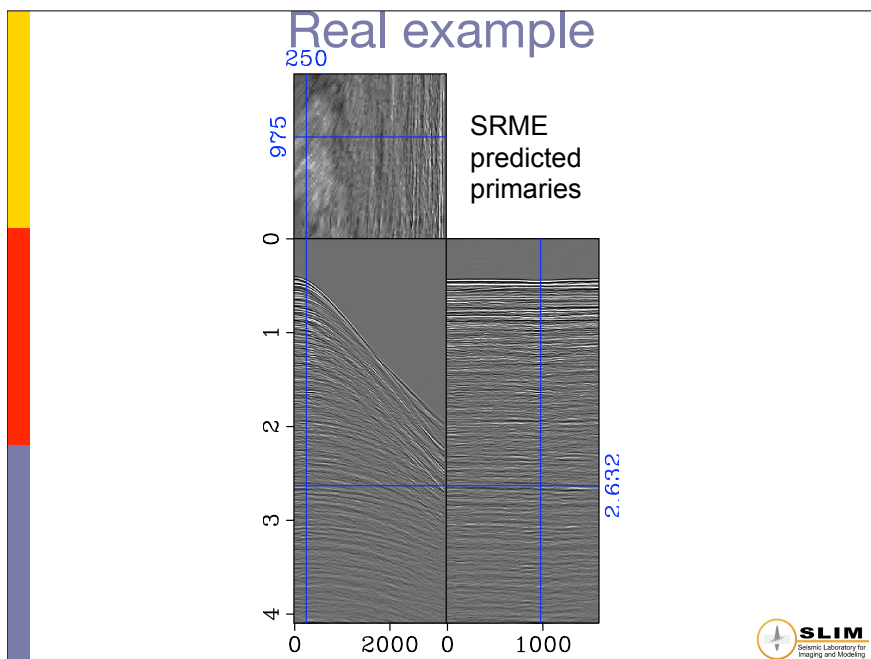
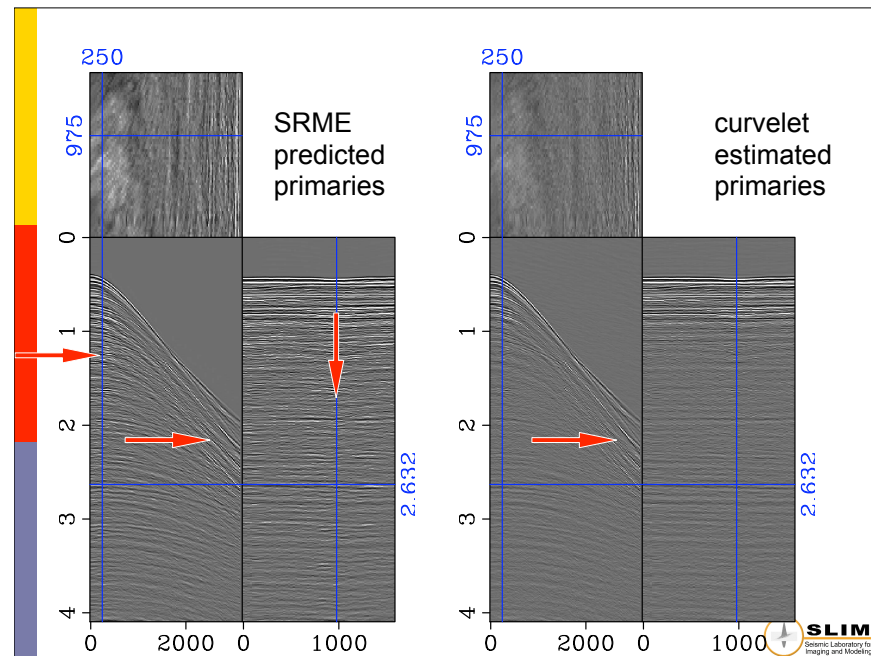
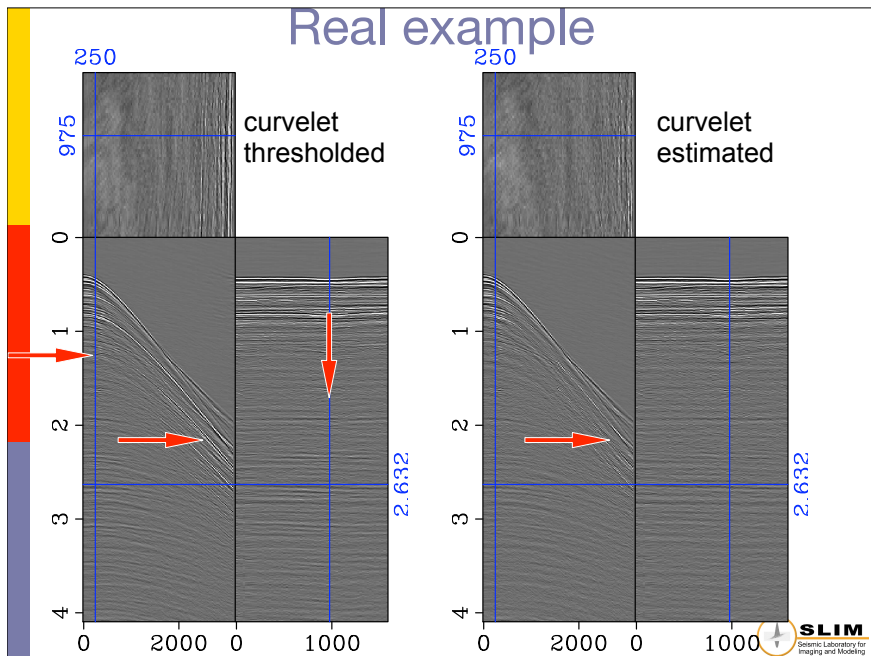
$$S_{\alpha_\mu}(v_\mu) = \text{sgn}(v_\mu) \cdot \max(0, |v_\mu| - |\alpha_\mu|)$$

is the **elementwise** soft thresholding operator.

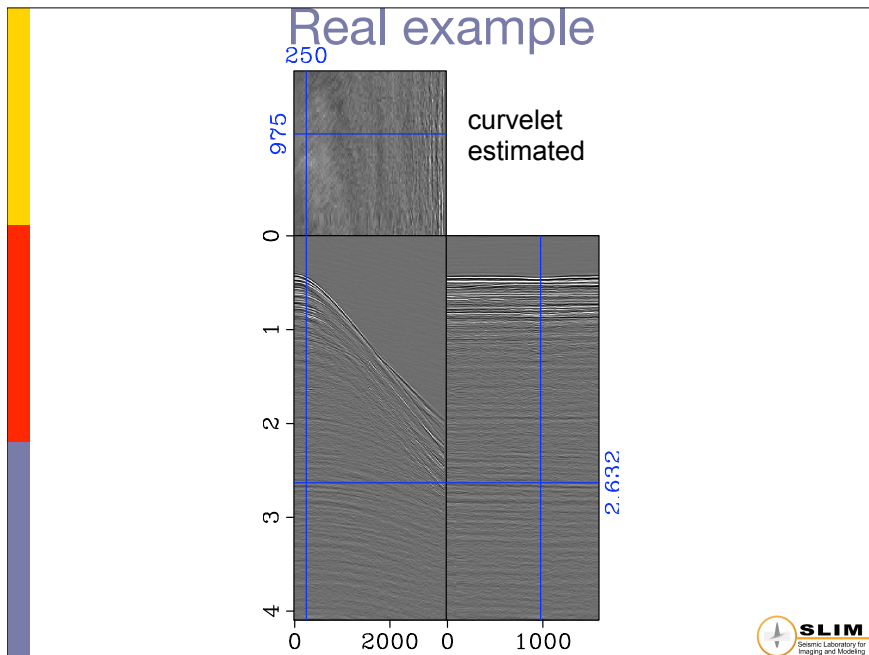
Provably this converges for positive weights.











## Observations

Inclusion of the additional equation for the predicted multiples prevents zero solutions for either signal component.

Formulation based on a solid Bayesian argument.  
The algorithm provably converges.

Method is not adaptive but gives control over

- denoising the multiples versus solving the old problem
- sparsity versus data fit
- tradeoff trust in prediction versus trust in total data

SLIM  
Seismic Laboratory for  
Imaging and Modeling

## Adaptive curvelet-domain matched filtering

Joint work with Deli Wang, Cody Brown and Peyman Moghaddam

SLIM  
Seismic Laboratory for  
Imaging and Modeling

## Motivation

**Kinematics** are generally well predicted.  
Non-adaptive curvelet-domain separation adds **robustness**.

Large errors in the location, dip and amplitude of the predicted multiples remain a problem.

Present an **adaptive** curvelet-domain separation based on **matched** filtering which assumes that

- the "seismic wavelet" has been removed
- the variations in the multiple predictions versus the true multiples vary **slowly** in **phase space**
- kinematics** are roughly correct

Design a technique that exploits the invariance of curvelets under a certain class of operators.

SLIM  
Seismic Laboratory for  
Imaging and Modeling

## Multiple prediction SRME

SRME-multiple prediction

$$\Delta \mathbf{p} \mapsto \check{\mathbf{m}}^{(1)}(s, r, t) = (\Delta \mathbf{P} \mathcal{A} *_t \Delta \mathbf{p})(s, r, t)$$

with

$\Delta \mathbf{p}$  = vector with the primaries

$\check{\mathbf{m}}^{(1)}$  = vector with predicted first-order multiples

$\Delta \mathbf{P}$  =  $\mathbf{F}^H$  block diag $\{\Delta p\} \mathbf{F}$

$\mathbf{F}$  = temporal Fourier transform

$\mathcal{A}$  = inverse wavelet.

In practice,  $\mathbf{p} \mapsto \Delta \mathbf{p}$ ,  $\mathbf{P} \mapsto \Delta \mathbf{P}$ , with  $\mathbf{p}$  the total data, so

$$\check{\mathbf{m}}^{(1)} \approx \mathbf{P} \mathcal{A} \mathbf{p}$$

## Multiple prediction SRME

Matched filter

$$\tilde{\mathbf{a}} = \arg \min_{\mathbf{a}} \|\mathbf{p} - \mathbf{a} *_t \check{\mathbf{m}}^{(1)}\|_2$$

yielding

$$\tilde{\Delta \mathbf{p}} = \mathbf{p} - \tilde{\mathbf{a}} *_t \check{\mathbf{m}}^{(1)}$$

$$\tilde{\mathcal{A}} = \text{block diag}\{\tilde{\mathbf{a}}\}$$

for each offset.

## Problem

Assumes the filter to be **stationary** (diagonal in Fourier space)

Source characteristics may change with offset.

Wavelet changes as a function of  $(s, r, t)$ .

Windowed matched-filtering techniques have been proposed

- window sizes arbitrary
- under fit (remnant primary energy)
- over fit (removal of primary energy)
- no control over the variations of the estimated filters amongst different windows

Propose a curvelet-domain matched filtering approach.

## Curvelet-domain matched filtering

Naive solution,

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} \|\mathbf{p} - \mathbf{P}[\mathbf{v}]\mathbf{u}\|_2$$

with

$\tilde{\mathbf{u}}$  = curvelet coefficients of the matched multiples

$\mathbf{P}[\mathbf{v}]$  = operator dependent on  $\mathbf{v}$

$\mathbf{v}$  = curvelet coefficients of  $\check{\mathbf{m}}^{(1)}$

- nonlinear is a problem
- underdetermined
- no control over estimated coefficients

## Curvelet-domain matched filtering

Assume,

$$\mathbf{m}^{(1)} = \mathbf{B}\check{\mathbf{m}}^{(1)}$$

Operator decomposition

$$\mathbf{B} = \mathcal{A}\Psi$$

with

$$(\Psi f)(x) = \int_{\mathbb{R}^d} e^{-ix \cdot \xi} a(x, \xi) \hat{f}(\xi) d\xi$$

a zero-order Pseudodifferential operator.



## Curvelet-domain matched filtering

Motivated by related work

- "Prediction of internal multiples" by F. ten Kroode who introduces an obliquity factor.
- "A method for inverse scattering based on the generalized Bremmer coupling series" by A. Malcolm and M. de Hoop who introduce certain weighting factors
- "Amplitude and kinematic corrections of migrated images for nonunitary imaging operators" by A. Guitton & "Optimal Scaling for Reverse Time Migration" by Symes who introduce a diagonal smooth scaling.



## Curvelet-domain matched filtering

Observe that "classical" matched filter

- absorbs the seismic wavelet
- deals with stationary phase rotations (Hilbert) and differential operators (derivatives)
- yields an inverse wavelet that is
  - of compact support in the time domain
  - smooth in the Fourier domain

Reasonable to assume

$$\begin{aligned} \mathbf{m}^{(1)} &= \mathbf{B}\check{\mathbf{m}}^{(1)} \\ &= \mathcal{A}\Psi\check{\mathbf{m}}^{(1)} \\ &= \Psi\mathcal{A}\check{\mathbf{m}}^{(1)} \\ &= \Psi\check{\mathbf{m}}_0^{(1)} \end{aligned}$$

with  $\Psi$  a zero-order PsDO.



## Curvelet-domain matched filtering

From work on migration amplitude recovery we have

**Theorem 1.** *The following estimate for the error holds*

$$\|(\Psi(x, D) - C^T \mathbf{D}_\Psi C)\varphi_\mu\|_{L^2(\mathbb{R}^n)} \leq C'' 2^{-|\mu|/2},$$

where  $C''$  is a constant depending on  $\Psi$ .

Allows for a curvelet domain diagonalization of  $\Psi$ ,

$$\mathbf{m}^{(1)} = \mathbf{C}^T \text{diag}\{\tilde{\mathbf{u}}\} \mathbf{C}\check{\mathbf{m}}_0^{(1)}$$



## Curvelet-domain matched filtering

Solve

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} \|\mathbf{p} - \mathbf{P}[\mathbf{v}]\mathbf{u}\|_2$$

with

$\tilde{\mathbf{u}}$  = curvelet coefficients of the matched multiples

$$\mathbf{P}[\mathbf{v}] = \mathbf{C}^T \text{diag}\{\mathbf{v}\}$$

$\mathbf{v}$  = curvelet coefficients of  $\check{\mathbf{m}}_0^{(1)}$

remains

- underdetermined system
- no control on the smoothness of the coefficients

**Assume symbol  $a(x, \zeta)$  to be smooth!**



## Curvelet-domain matched filtering

Solve

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} \frac{1}{2} \|\mathbf{b} - \mathbf{P}\mathbf{u}\|_2^2 + \eta^2 \|\mathbf{L}\mathbf{u}\|_2^2$$

with

$$\mathbf{L} = [\mathbf{D}_1 \quad \mathbf{D}_2 \quad \mathbf{D}_\theta]$$

equivalent to

$$\begin{bmatrix} \mathbf{P} \\ \eta\mathbf{L} \end{bmatrix} \mathbf{u} = \begin{bmatrix} \mathbf{p} \\ \mathbf{0} \end{bmatrix}$$

or

$$\mathbf{A}\mathbf{u} = \mathbf{b}$$



## Curvelet-domain matched filtering

Procedure

Calculate:  $\mathbf{v} = \mathbf{C}\check{\mathbf{m}}_0^{(1)}$ .

Set:  $\eta = \eta_{min}$ ;

**while**  $\exists (\tilde{u}_\mu)_{\mu \in \mathcal{M}} < 0$  **do**

Solve

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} \frac{1}{2} \|\mathbf{p} - \mathbf{P}\mathbf{u}\|_2^2 + \eta^2 \|\mathbf{L}\mathbf{u}\|_2^2$$

Increase the Lagrange multiplier

$\lambda = \eta + \Delta\eta$

**end while**

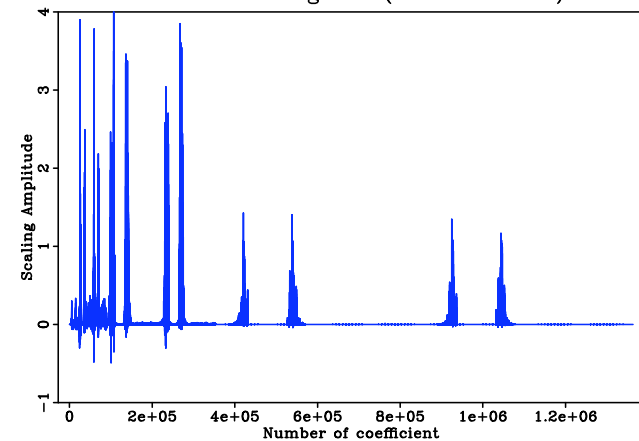
New estimate for the primaries

$$\begin{aligned} \tilde{\Delta}\mathbf{p} &= \mathbf{p} - \Psi\mathcal{A}\check{\mathbf{m}}_0^{(1)} \\ &\approx \mathbf{p} - \mathbf{P}\tilde{\mathbf{u}} \end{aligned}$$



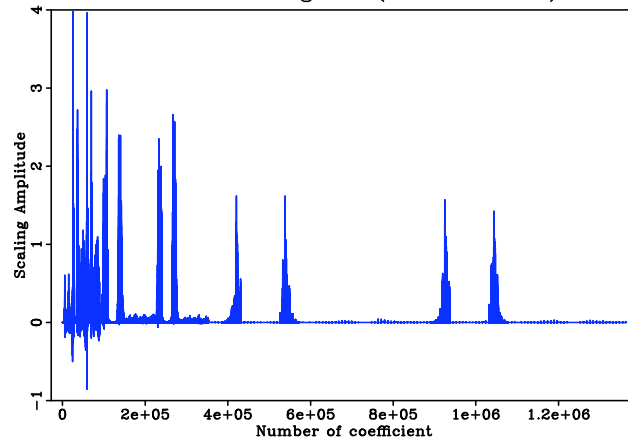
## Curvelet-domain matched filtering

Estimated diagonal (Lambda=0.05)



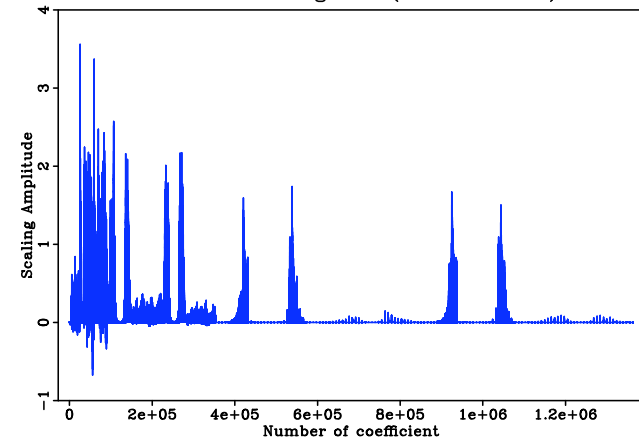
## Curvelet-domain matched filtering

Estimated diagonal (Lambda=0.25)



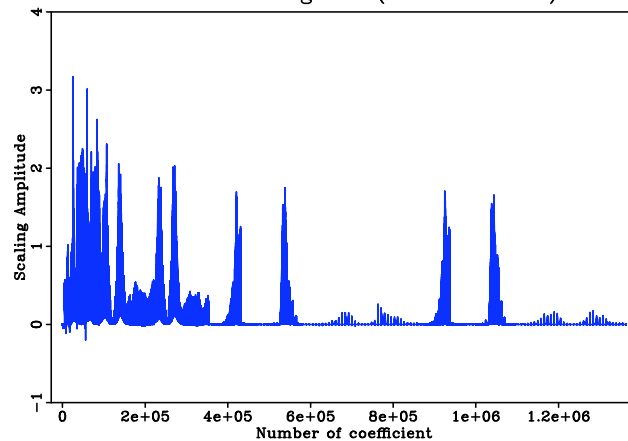
## Curvelet-domain matched filtering

Estimated diagonal (Lambda=0.5)



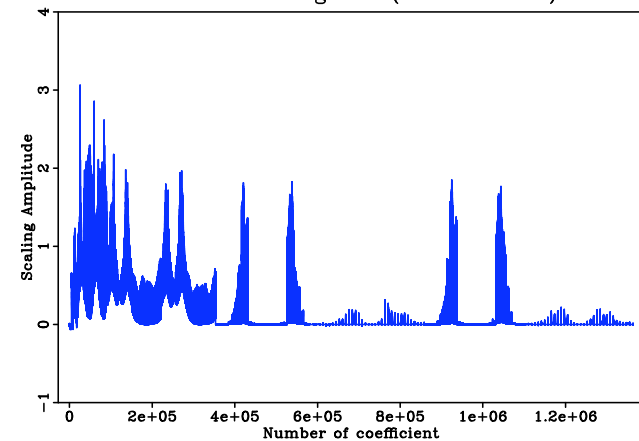
## Curvelet-domain matched filtering

Estimated diagonal (Lambda=0.75)

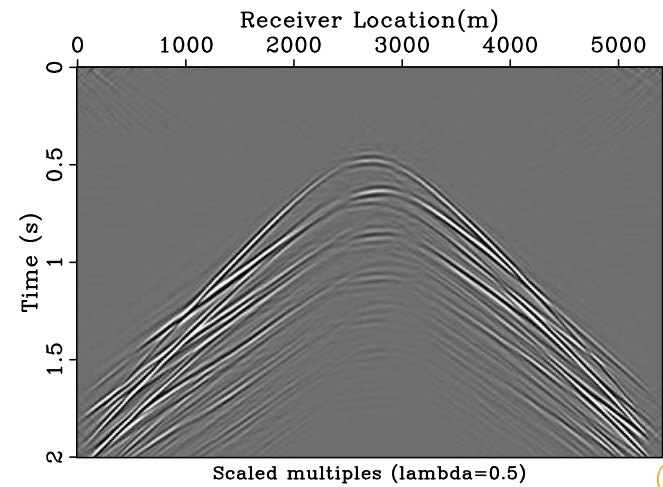
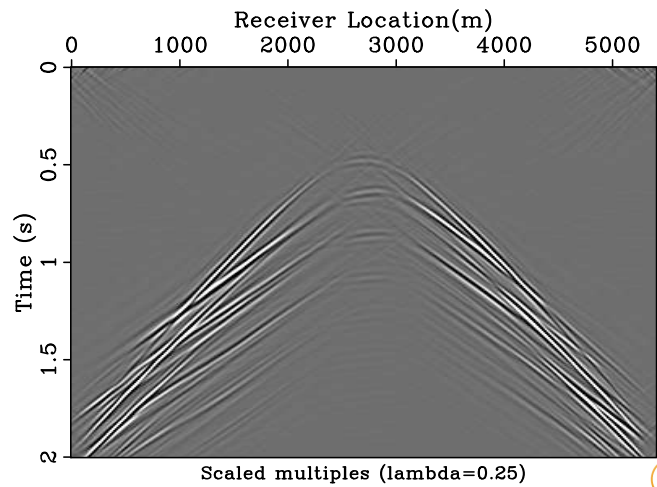
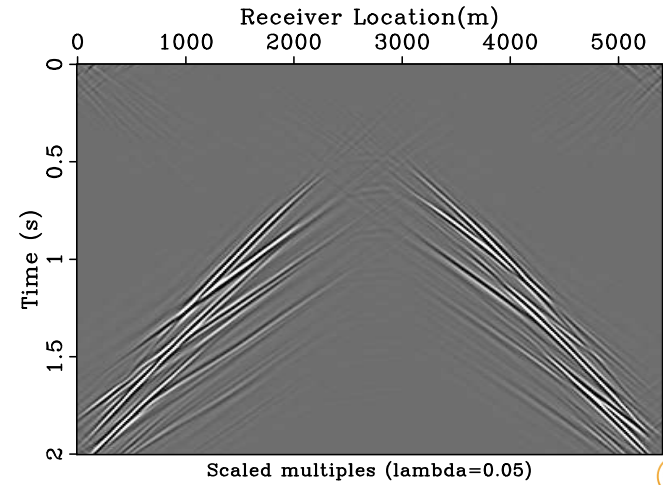
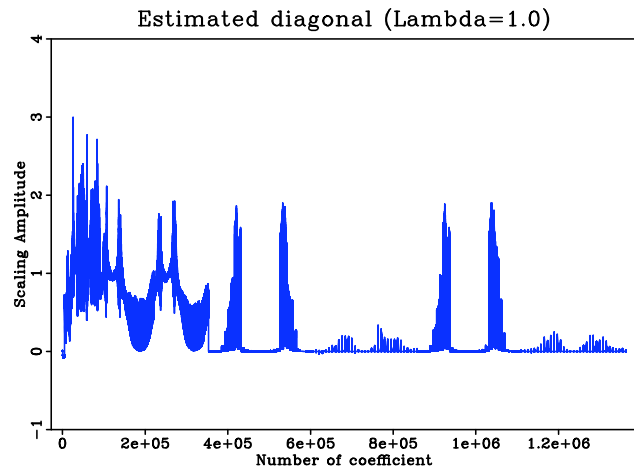


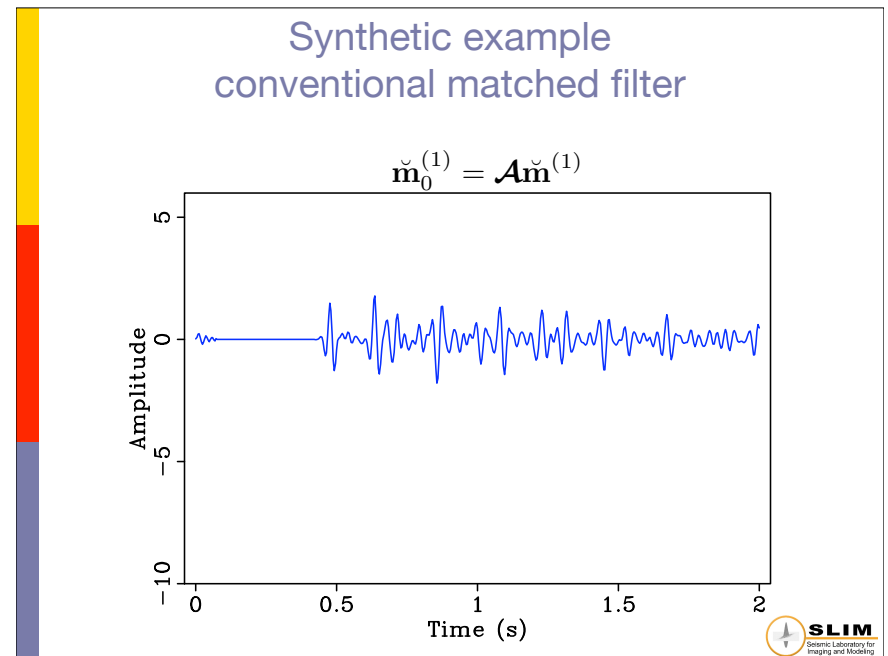
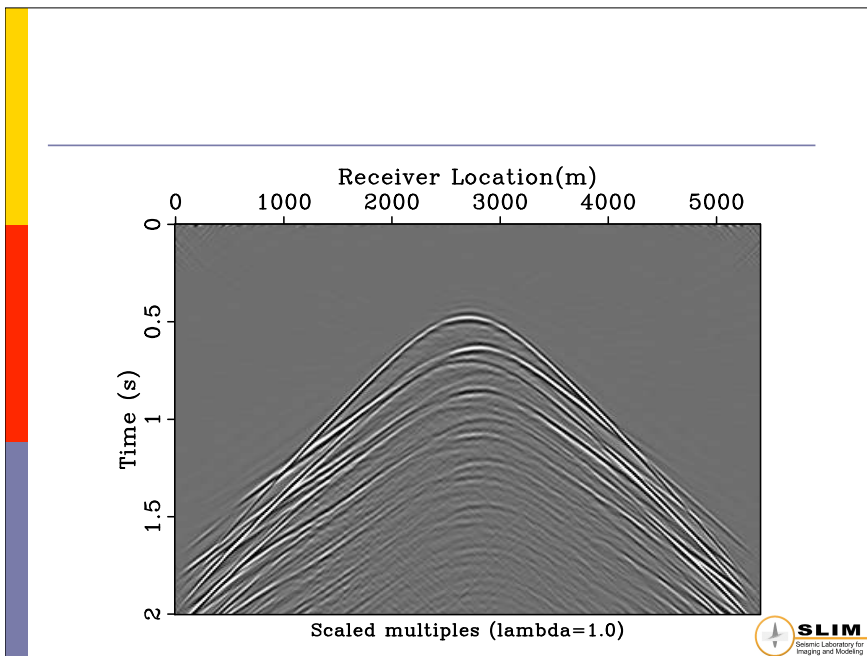
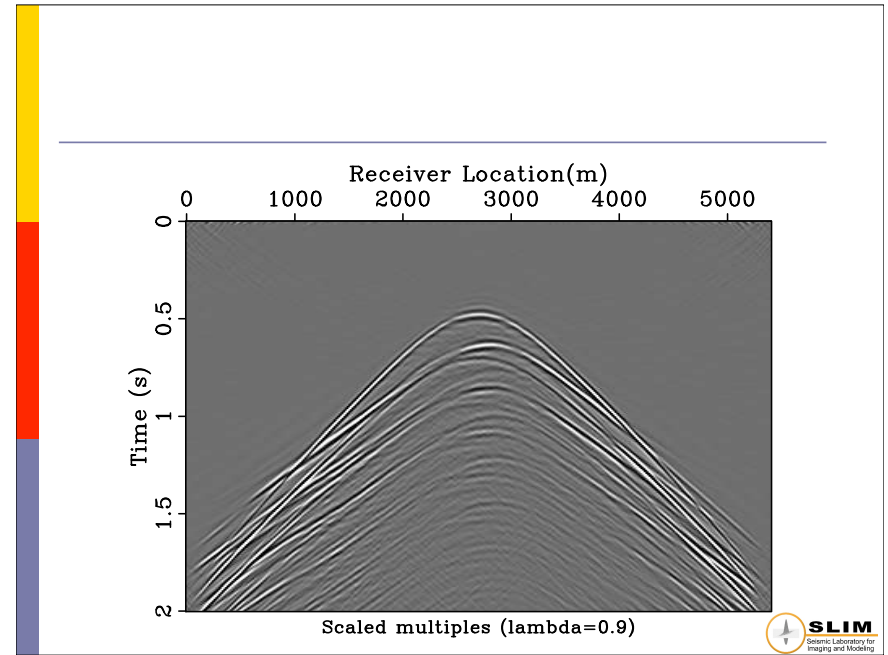
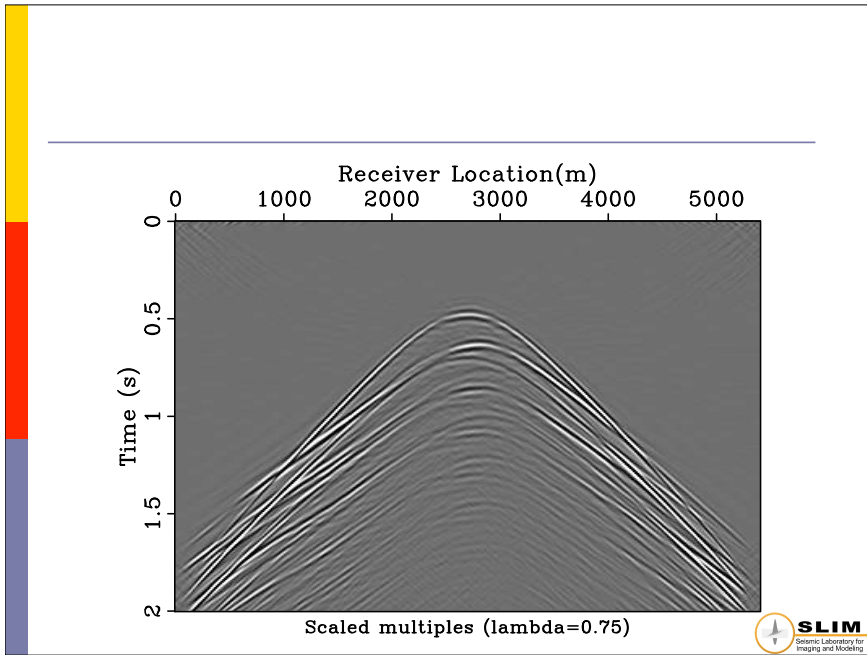
## Curvelet-domain matched filtering

Estimated diagonal (Lambda=0.9)

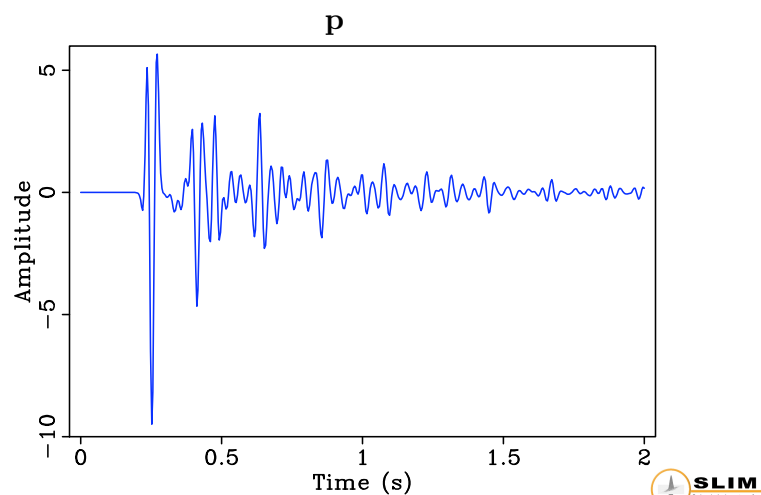


# Curvelet-domain matched filtering

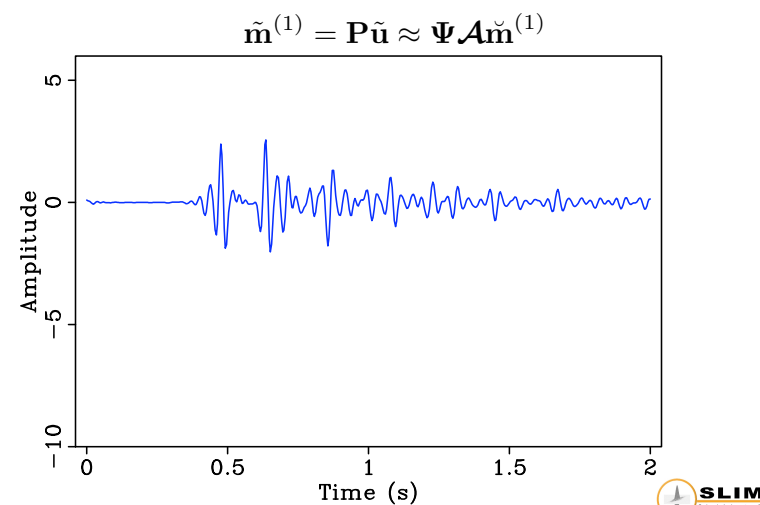




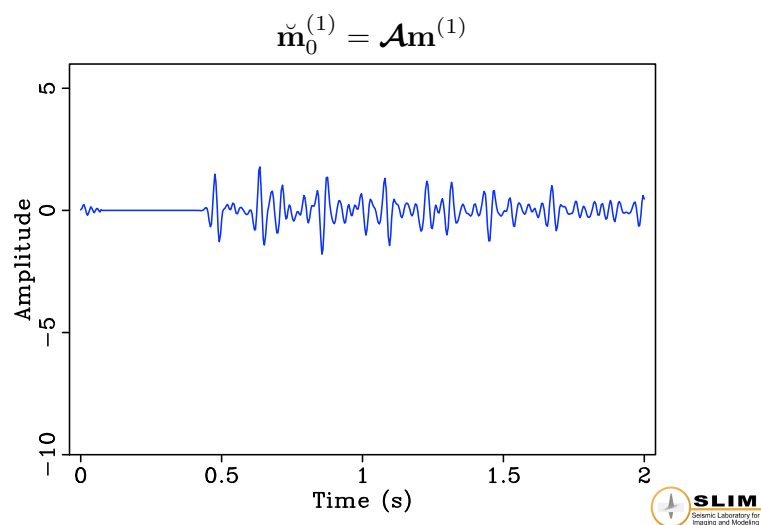
### Synthetic example total data



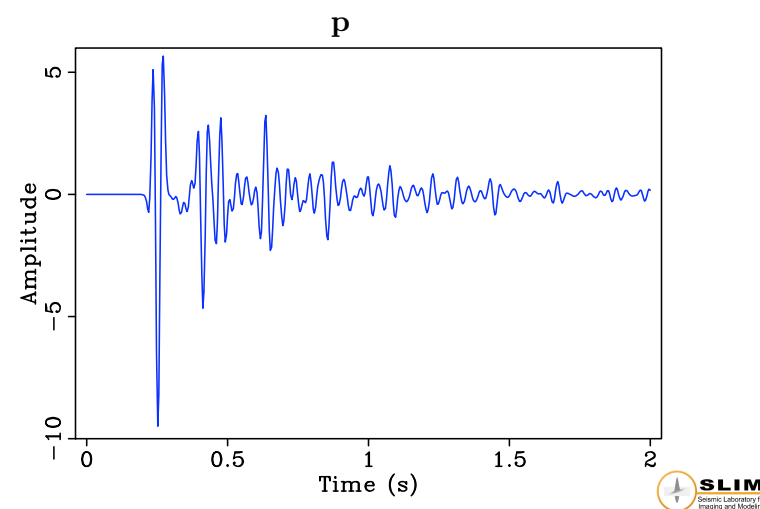
### Synthetic example curvelet-domain matched multiples



### Synthetic example conventional matched filter

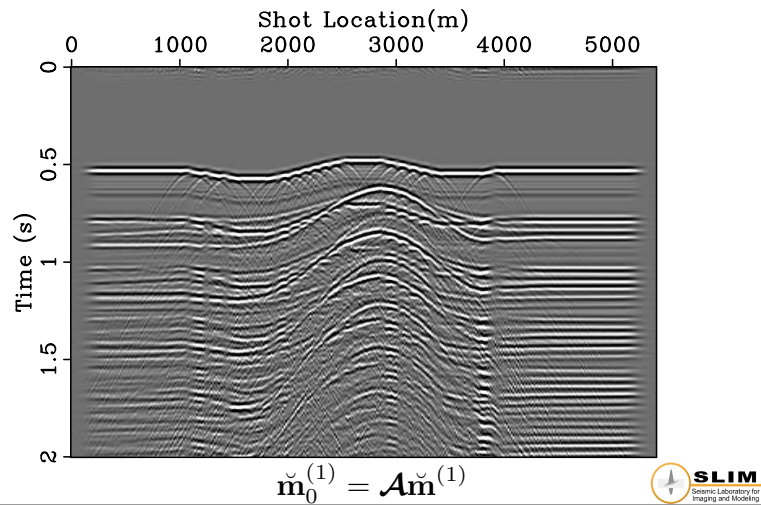


### Synthetic example total data

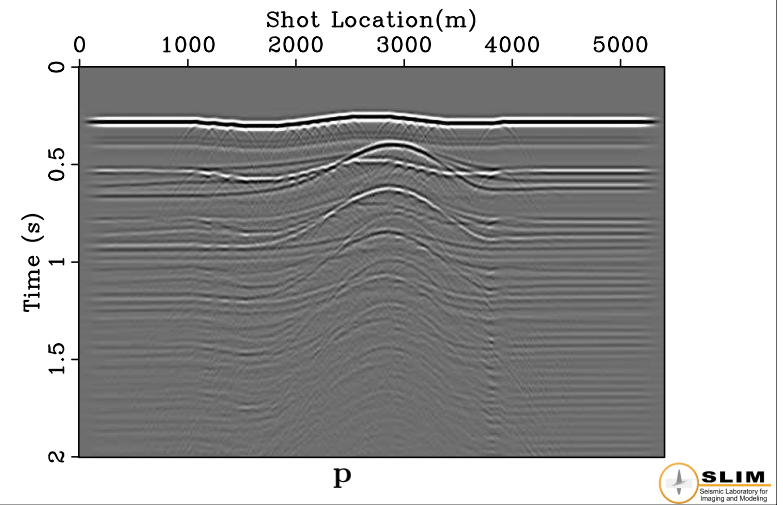




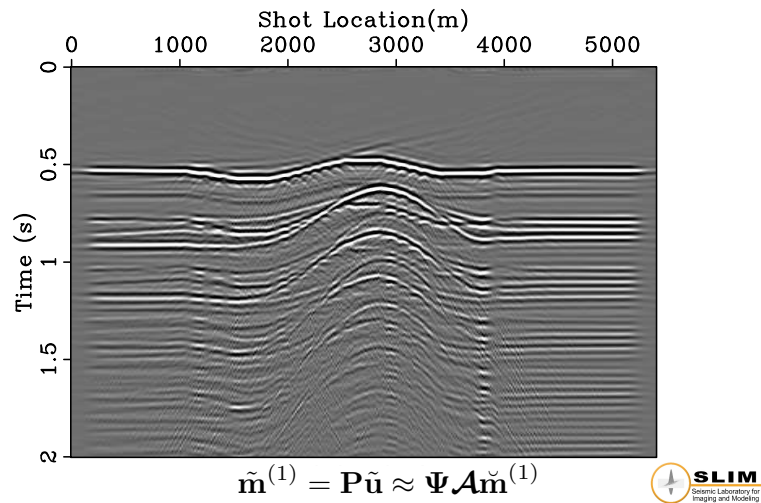
### Synthetic example conventional matched filter



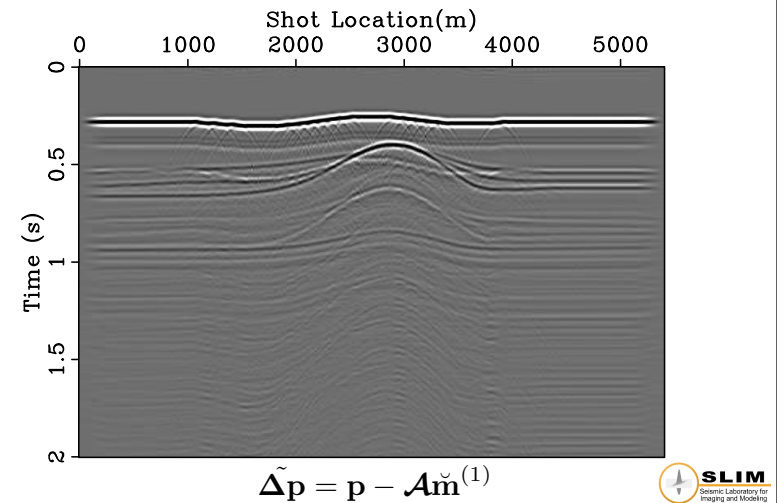
### Synthetic example total data



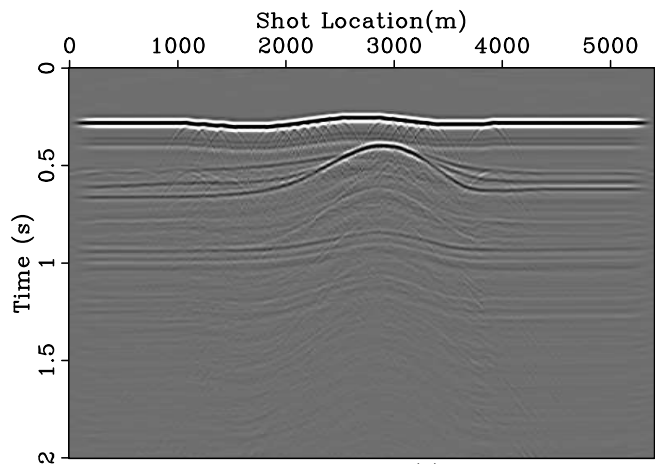
### Synthetic example curvelet-domain matched multiples



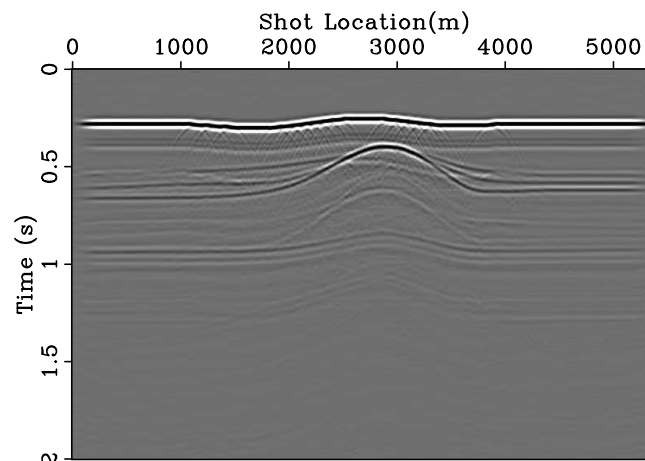
### Synthetic example conventional estimate for primaries



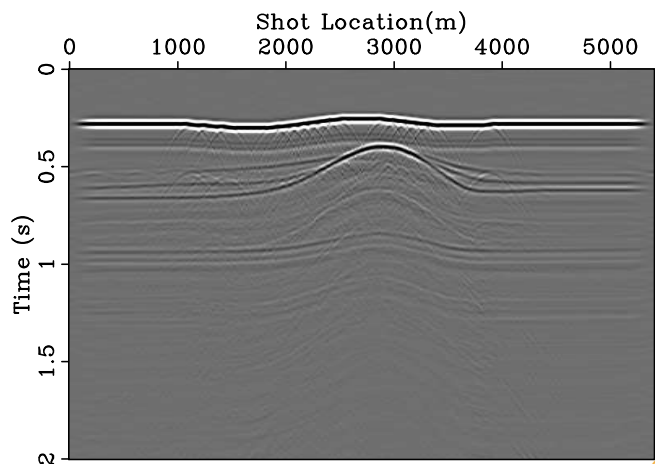
### Synthetic example curvelet-based estimate for primaries



$$\tilde{\Delta p} = p - \Psi \mathcal{A} \tilde{m}^{(1)} \approx p - P \tilde{u}$$



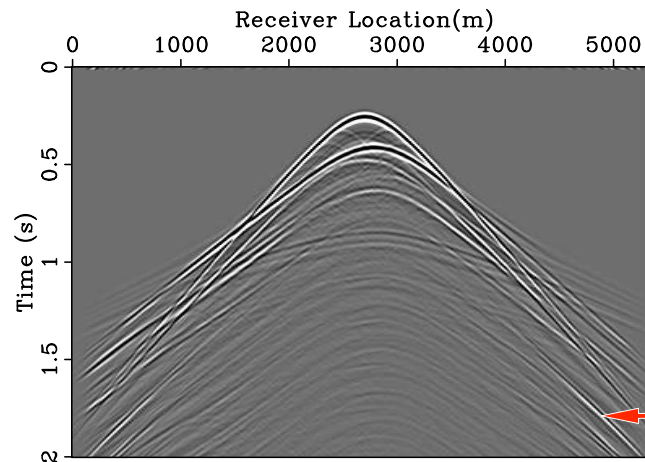
Bayesian Primaries (LSQR multiples)



Bayesian Primaries (Scaled multiples)



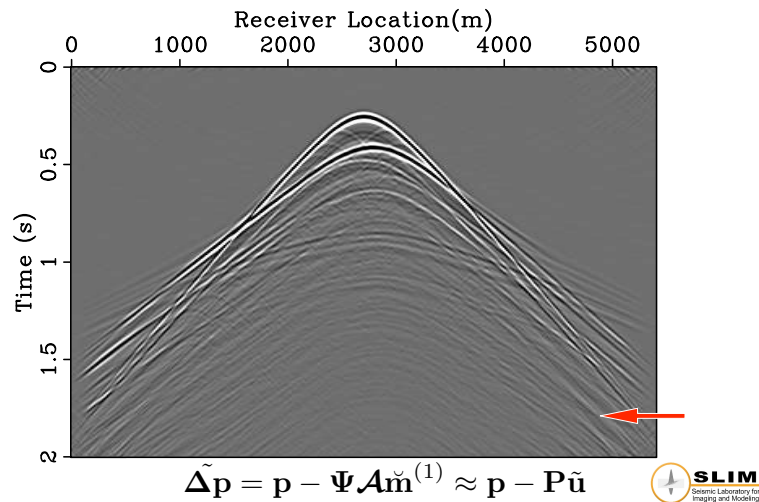
### Synthetic example conventional estimate for primaries



$$\tilde{\Delta p} = p - \mathcal{A} \tilde{m}^{(1)}$$



## Synthetic example curvelet-based estimate for primaries



## Observations

Smoothness penalty on the curvelet coefficients

- behaves as expected (becomes positive)

Alternative to Bayesian-based separation.

## Diffraction



## Curvelet-based Multi-term Multiple subtraction

Most of the application of surface-related multiple attenuation are 2D

Most of 3D effects come from diffractive structures

Split the data into a specular and a diffraction part can handle 3D effects

Multiple prediction scheme can be rewritten as:

$$\mathbf{M} = [\Delta \mathbf{P}_r + \Delta \mathbf{P}_d][\mathbf{P}_r + \mathbf{P}_d]$$

and we get four terms multiples:

$$\mathbf{M}_{rr} = \Delta \mathbf{P}_r \mathbf{P}_r$$

$$\mathbf{M}_{rd} = \Delta \mathbf{P}_r \mathbf{P}_d$$

$$\mathbf{M}_{dr} = \Delta \mathbf{P}_d \mathbf{P}_r$$

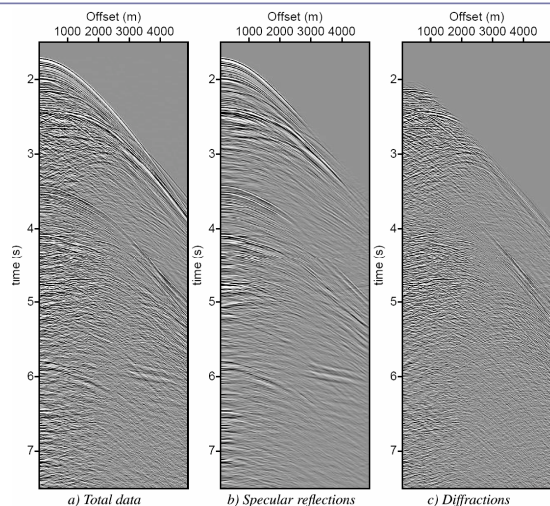
$$\mathbf{M}_{dd} = \Delta \mathbf{P}_d \mathbf{P}_d$$

Where:

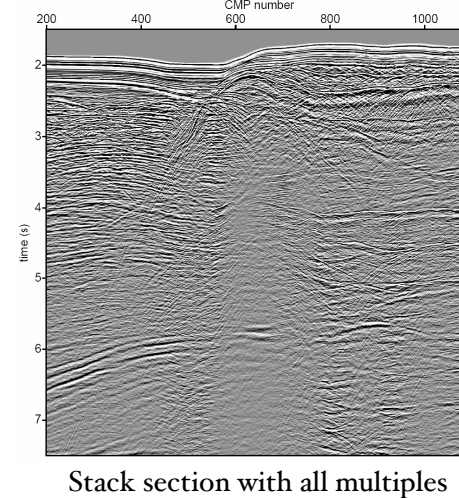
$r$  : represent reflection part

$d$  : represent diffraction part

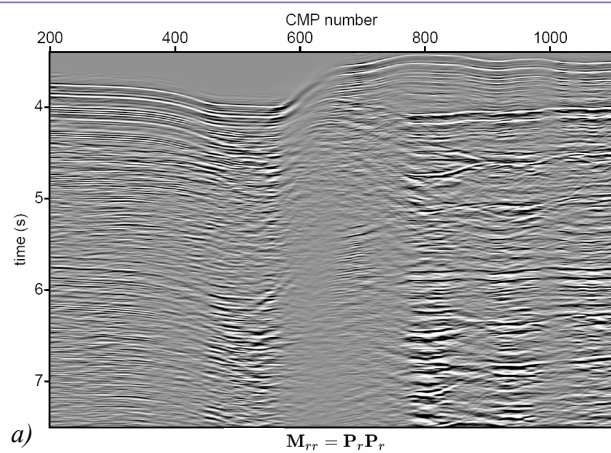
## Curvelet-based multi-term multiple separation



## Curvelet-based multi-term multiple separation

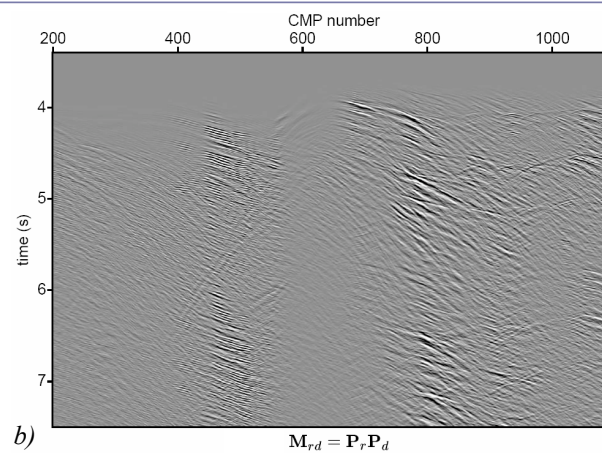


## Curvelet-based multi-term multiple separation



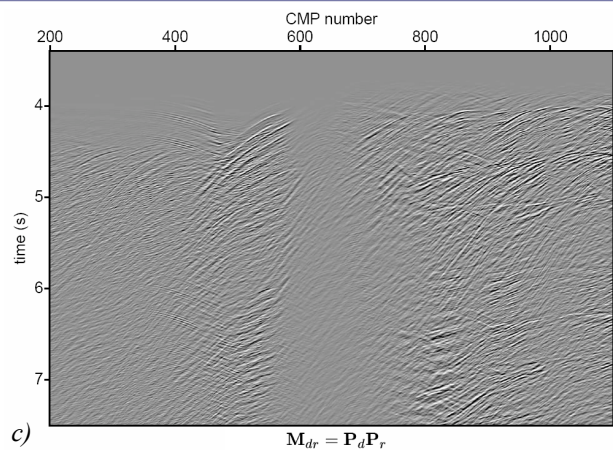
Stack section of the first-term multiples

## Curvelet-based multi-term multiple separation

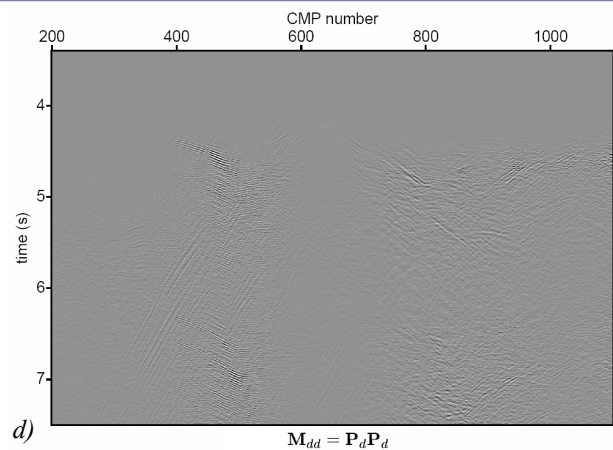


Stack section of the second-term multiples

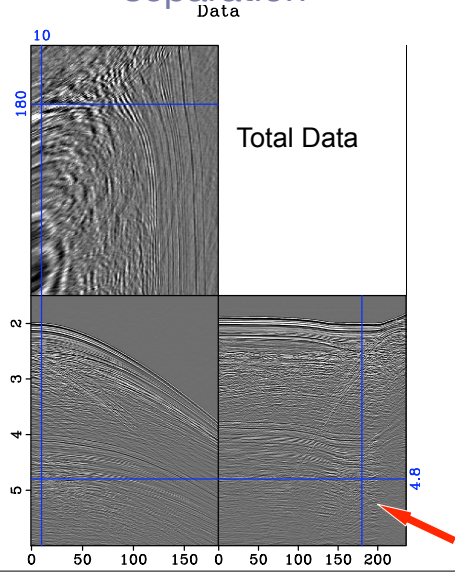
## Curvelet-based multi-term multiple separation



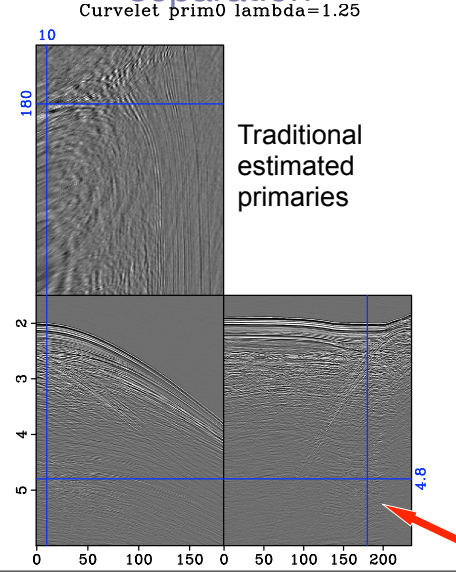
## Curvelet-based multi-term multiple separation



## Curvelet-based multi-term multiple separation

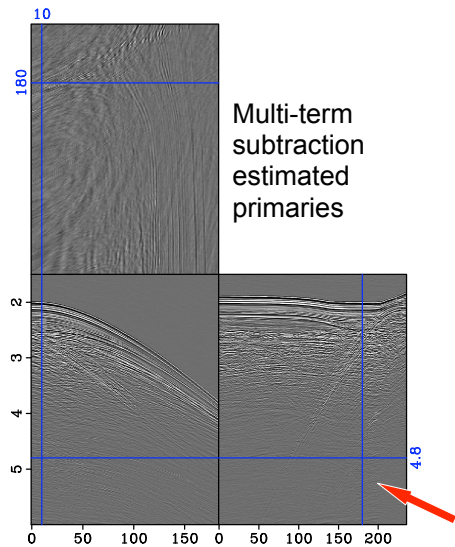


## Curvelet-based multi-term multiple separation



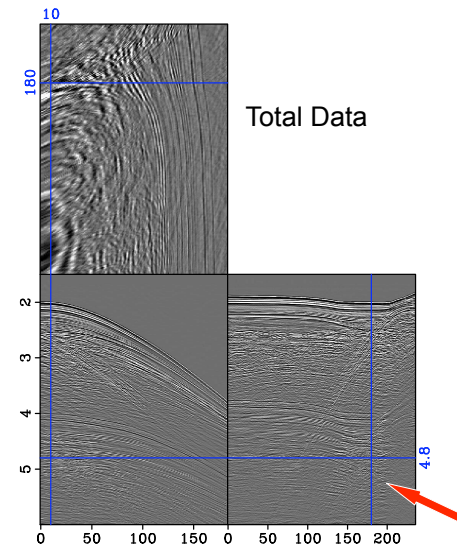
## Curvelet-based multi-term multiple separation

Curvelet1-4  $\lambda=1.25$



## Curvelet-based multi-term multiple separation

Data



## Observations

Single-threshold based separation with multiterm prediction

Each term is handled separately

Improved multiple separation compared to the "all in one separation"

Increased flexibility to handle 3-D diffractions



## Extensions

part of SINBAD II