Machine learning applications to geophysical data analysis

Ben Bougher, August 15th 2016

SLIM

University of British Columbia

Contributions

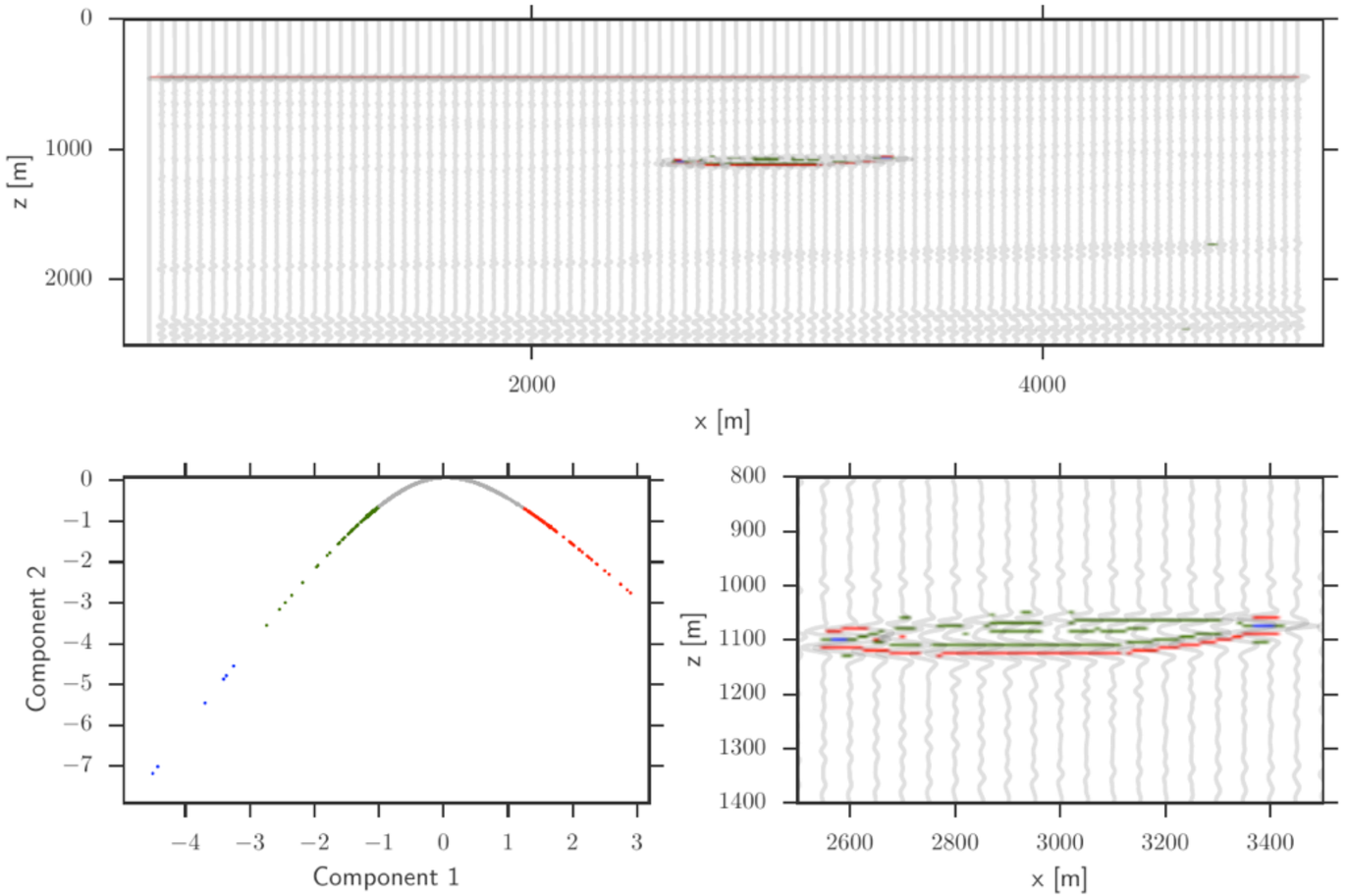
Predicting stratigraphic units from well logs using supervised learning

- Novel use of the scattering transform as a feature representation of well logs.
- CSEG 2015 expanded abstract, honorable mention for best student talk.
- Published article in the CSEG Recorder (Jan 2015).

Reflection seismology as an unsupervised learning problem

- Generalized and automated a hydrocarbon exploration analysis workflow.
- Reservoir discovery as convex optimization.
- Accepted abstract, to be presented @SEG (Houston, 2016)

Reflection seismology as an unsupervised learning problem



Motivation:

Inability to discover hydrocarbons directly from seismic data.

Problem:

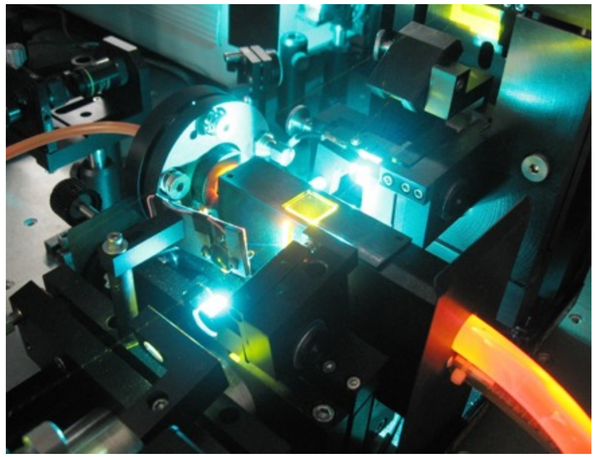
Automatically segment potential hydrocarbon reserves from seismic images.

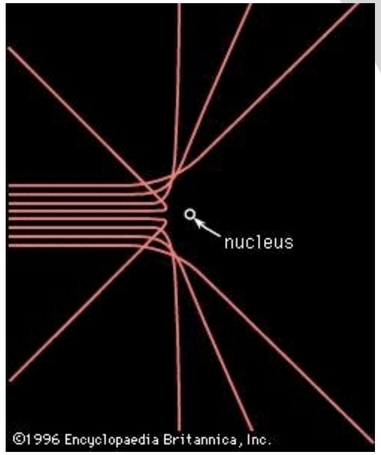
Approaches:

Physics driven (conventional)

Data driven (thesis contributions)

Scattering physics



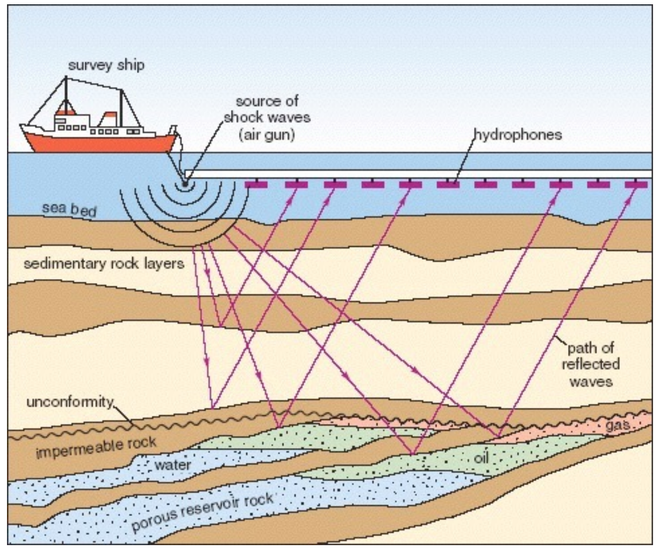


Ubiquitous in experimental physics.

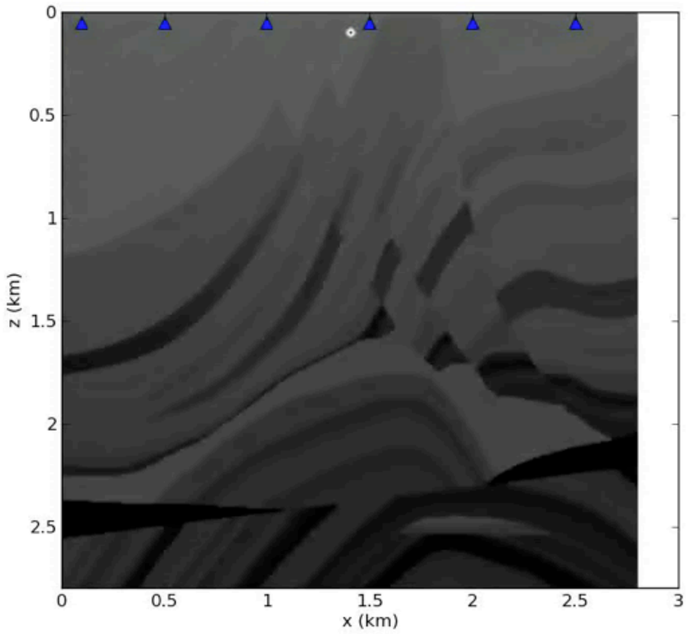
Measure the scattering pattern from a known source incident on a material.

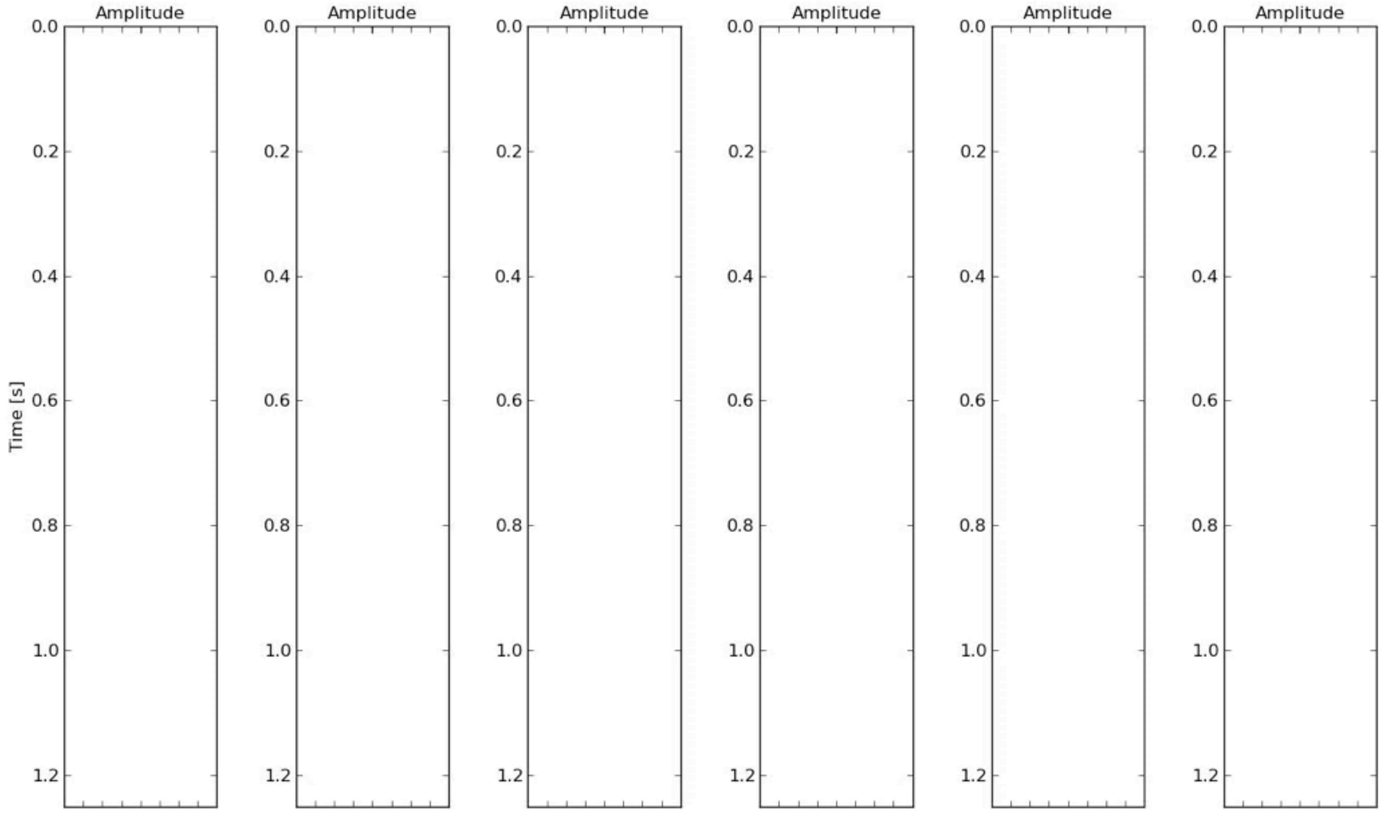
Performed in highly controlled and calibrated laboratories (laser sources, temperature controlled, vacuums, etc...).

Reflection seismology is a scattering experiment in an uncontrolled environment.

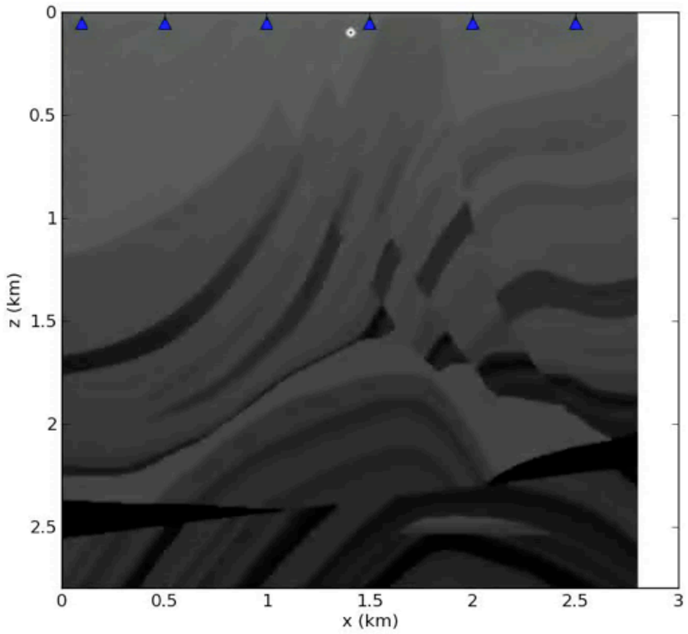


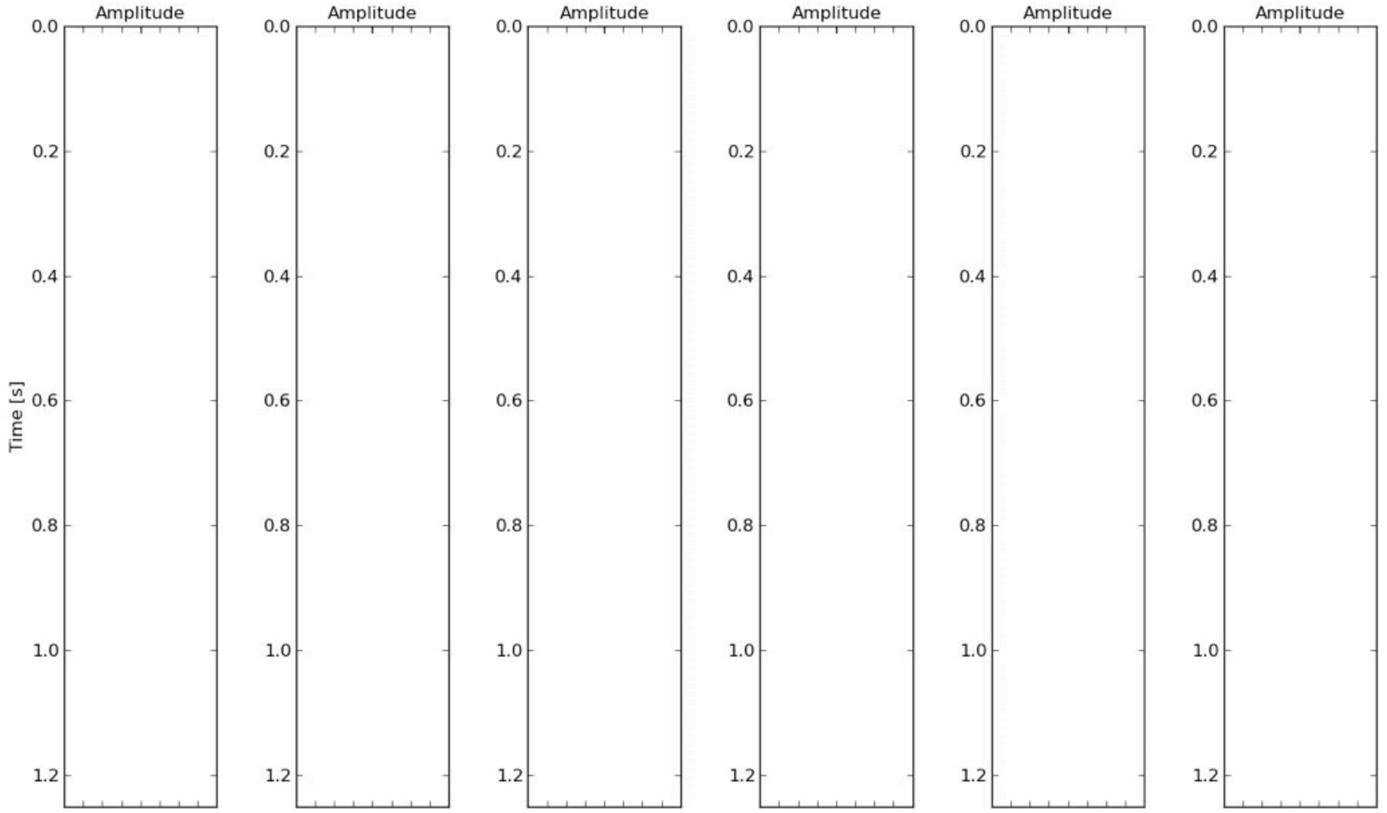
Seismic experiment



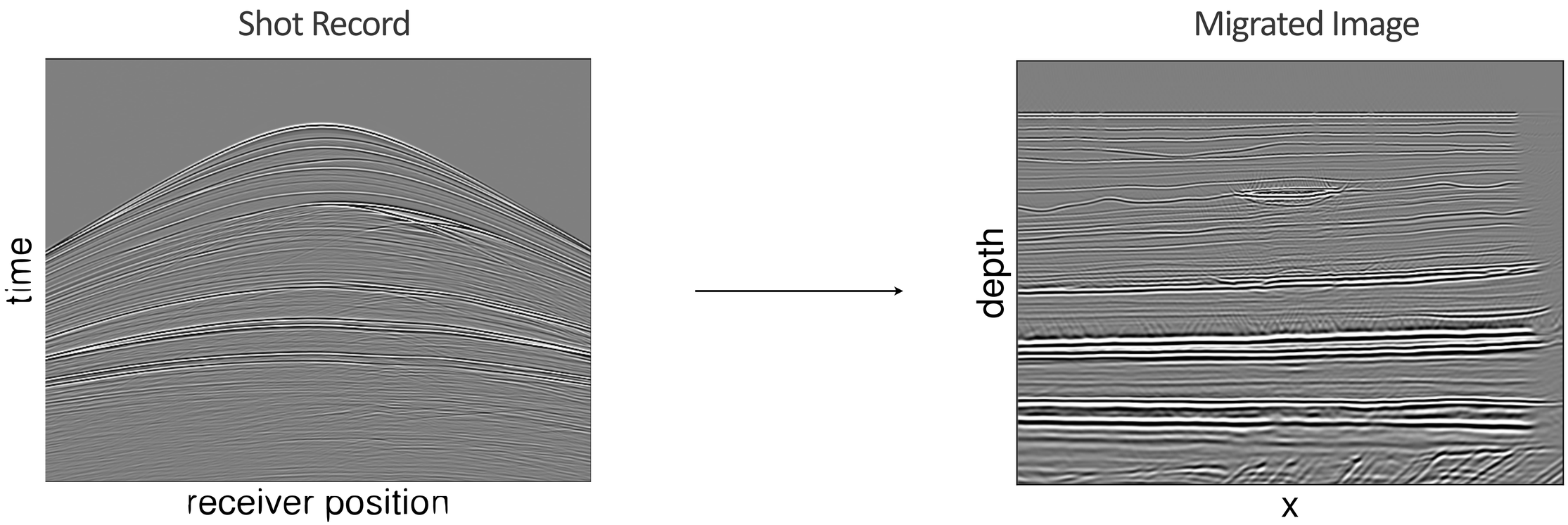


Seismic experiment



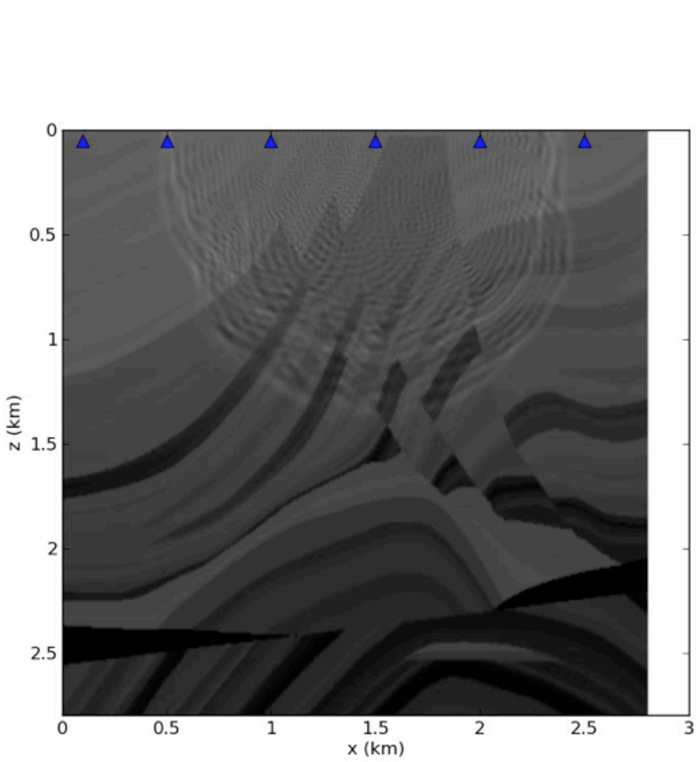


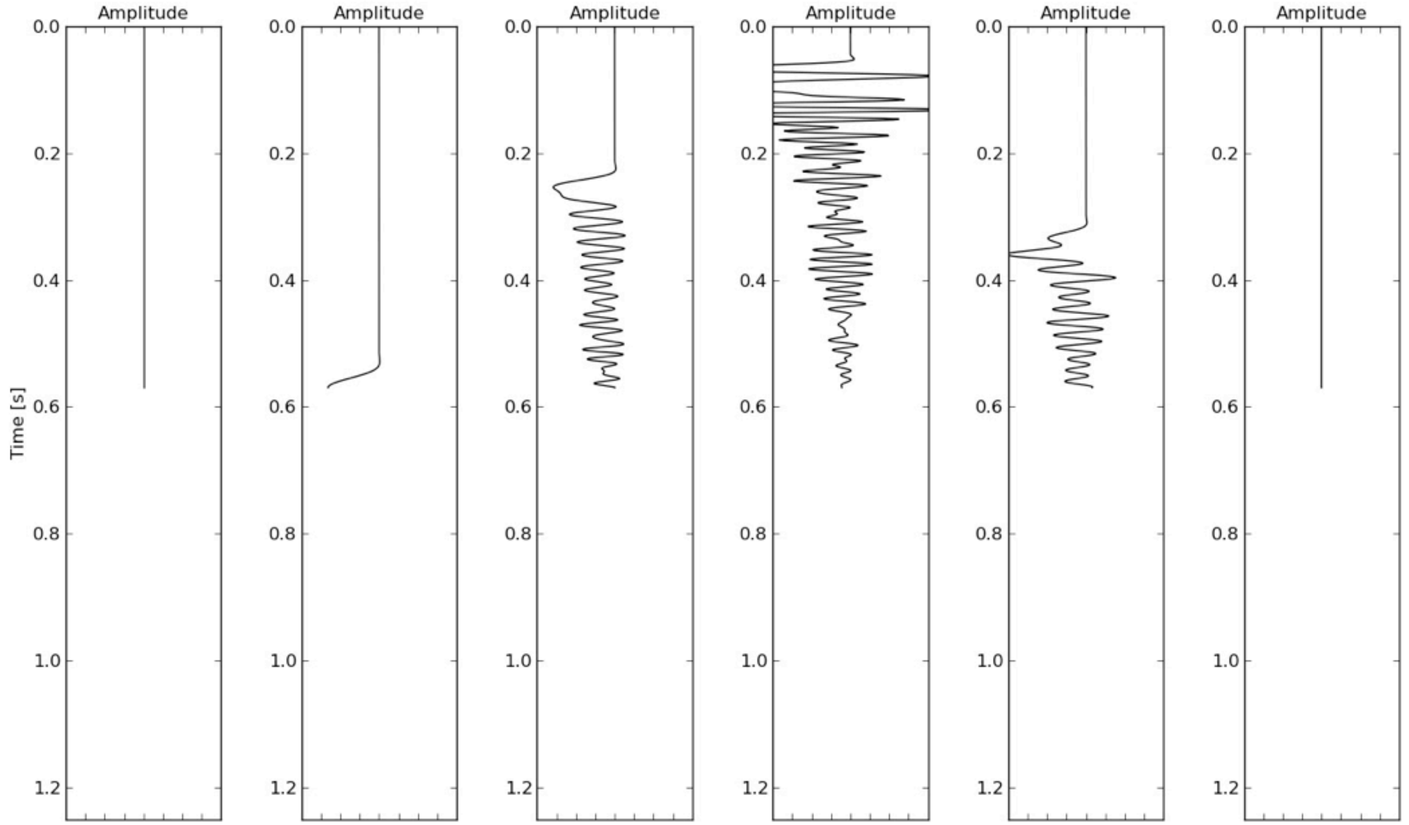
Migration



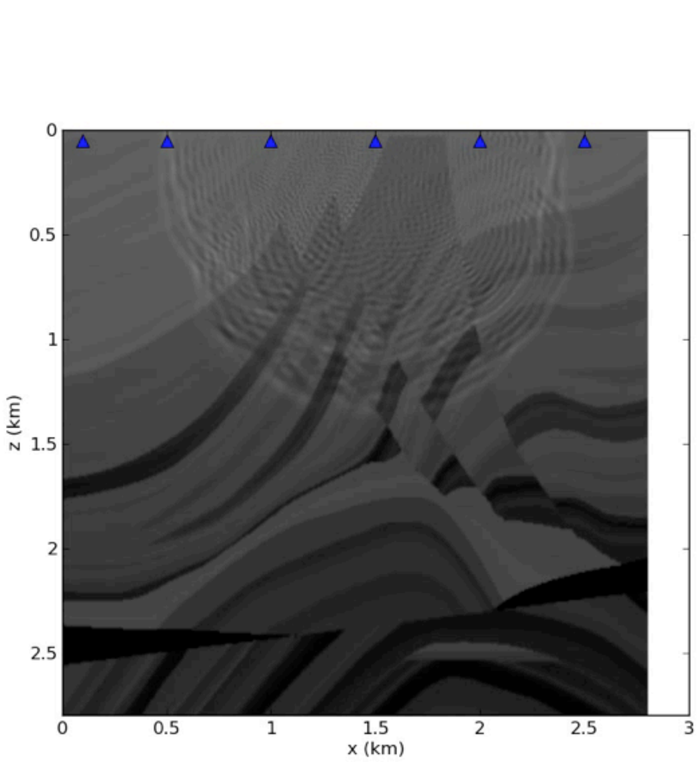
Migration maps **shot records** of **reflections** recorded at the surface into **images** of the subsurface.

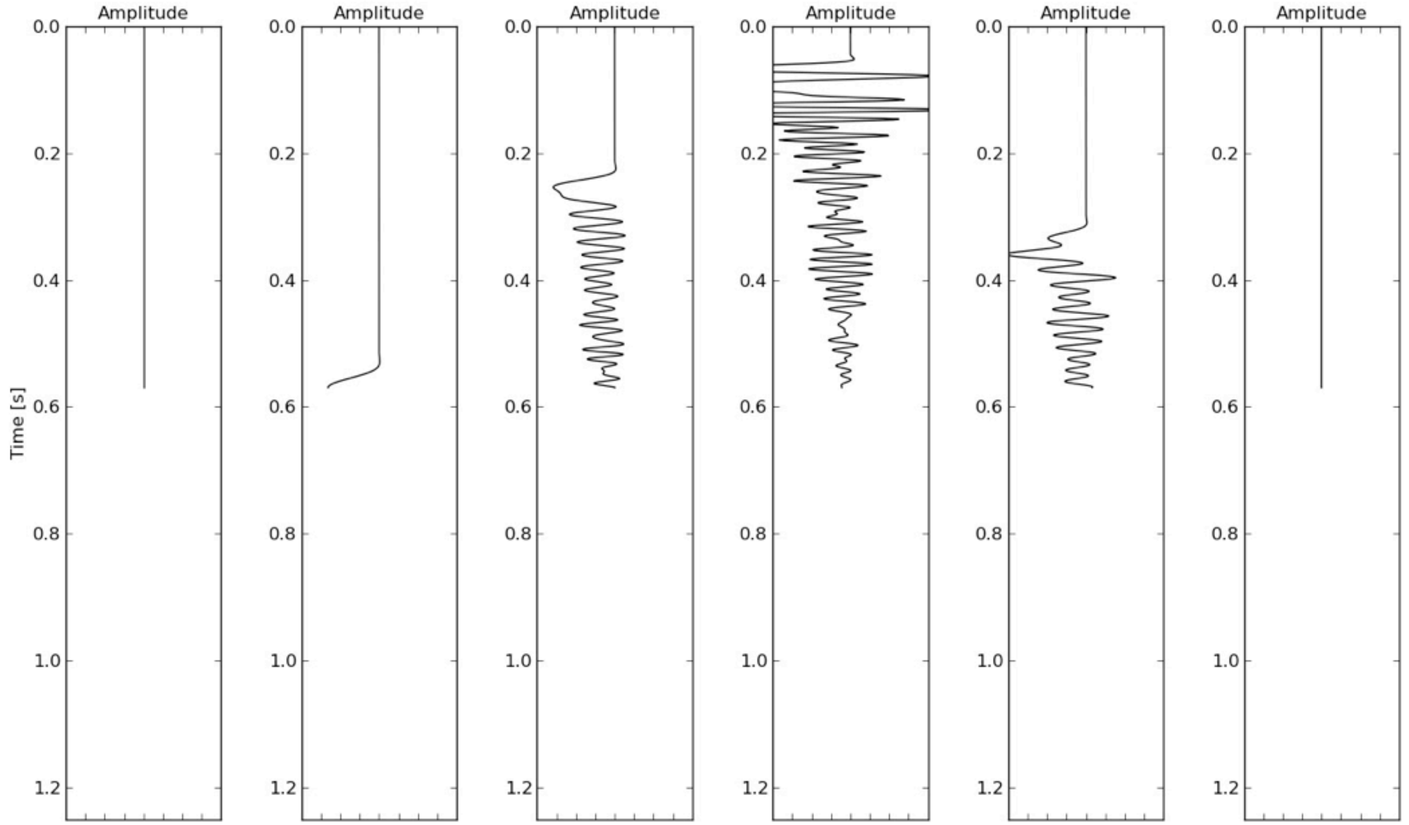
Migration





Migration

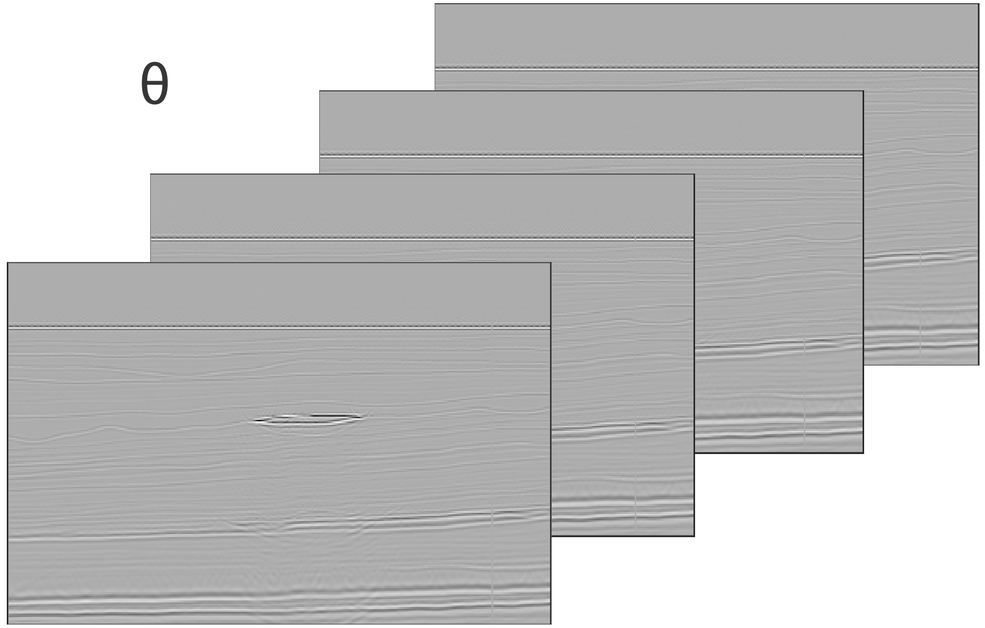


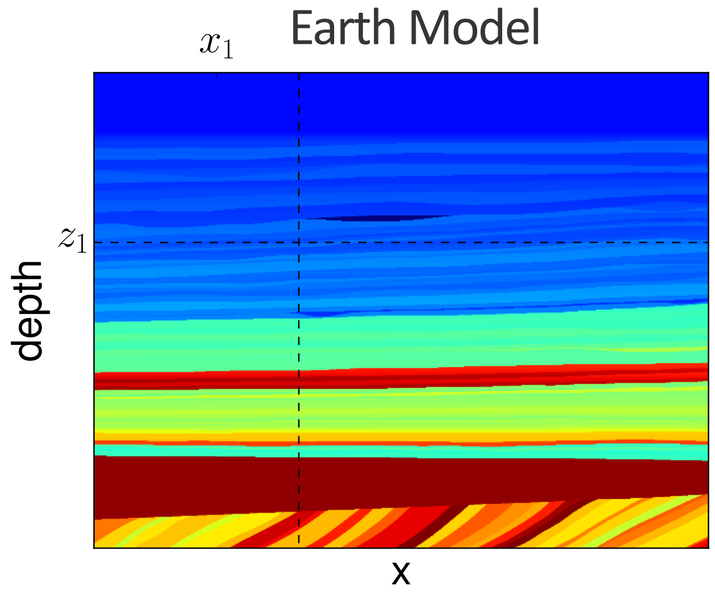


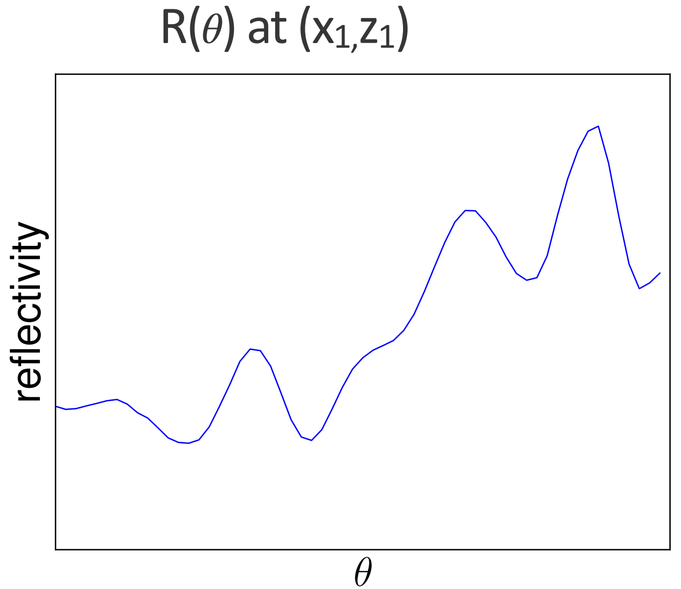
Angle domain common image gathers

Problem:
Need angle dependent reflectivity responses

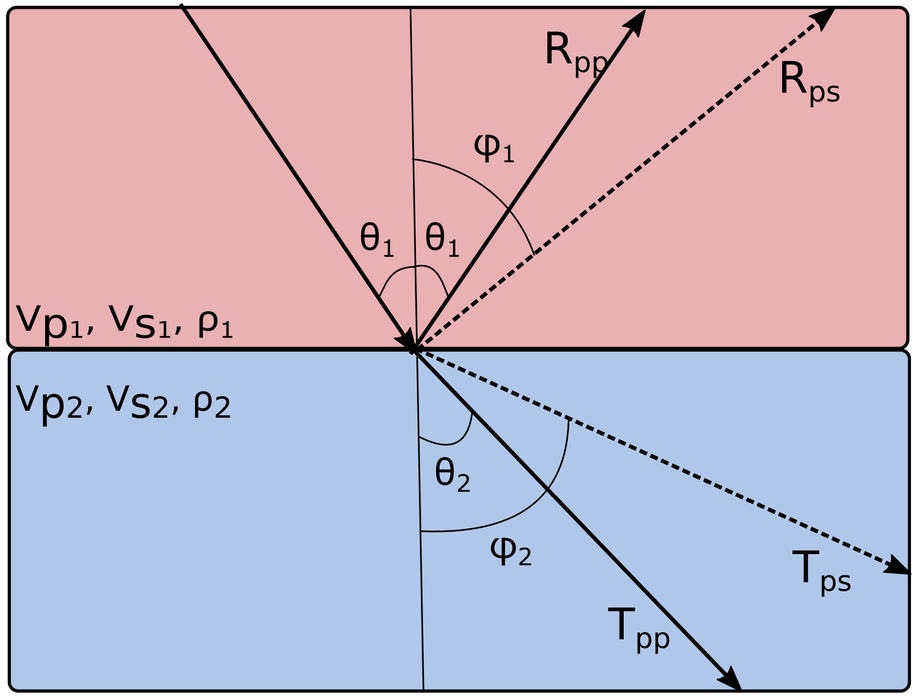
Solution:
Angle domain common image gather migration







Scattering theory (Zoeppritz)



Problem:
Relate angle dependent reflectivity to rock physics.

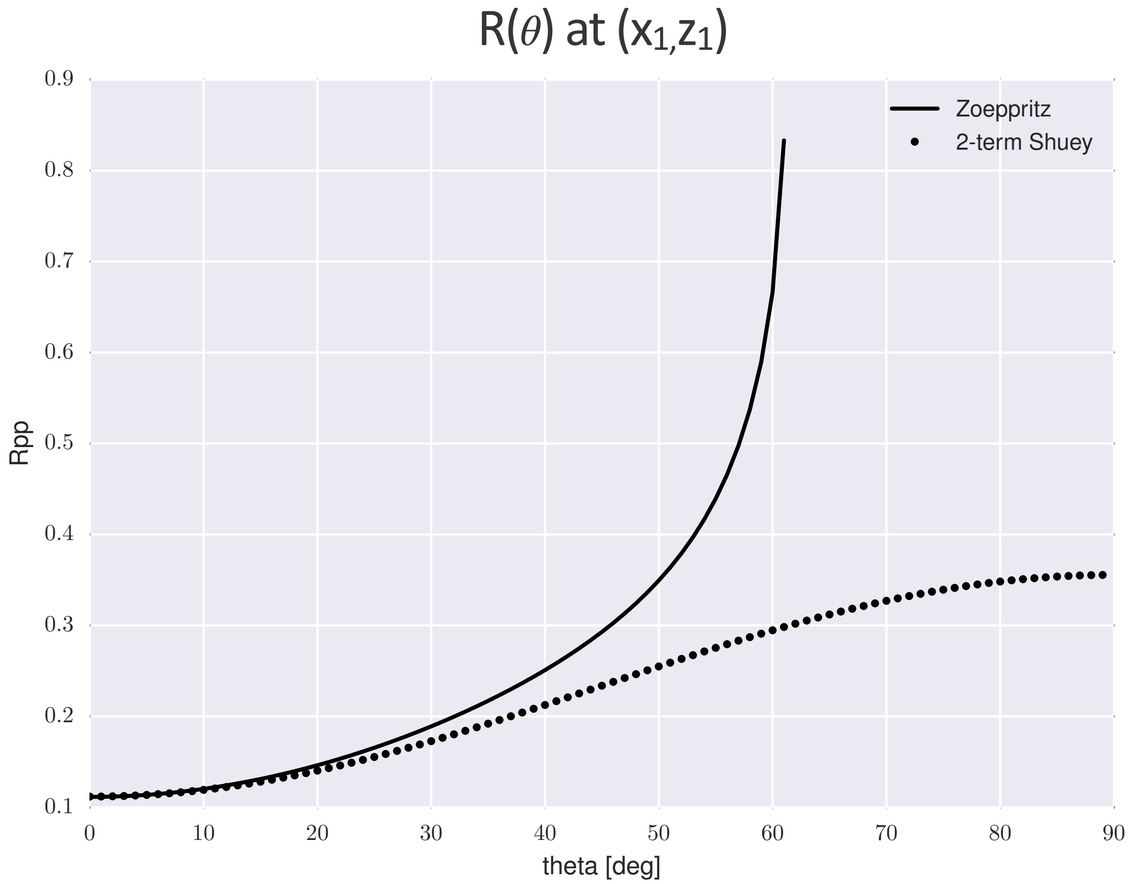
Assumption: Ray theory approximation.

Solution: $R(\theta) \propto V_P, V_S, \rho$

Problem: Non-linear, not useful for inversion.

$$\begin{bmatrix} R_{PP} \\ R_{PS} \\ T_{PP} \\ T_{PS} \end{bmatrix} = \begin{bmatrix} -\sin\theta_1 & -\cos\phi_1 & \sin\theta_2 & \cos\phi_2 \\ \cos\theta_1 & -\sin\phi_1 & \cos\theta_2 & -\sin\phi_2 \\ \sin 2\theta & \frac{V_{P1}}{V_{S1}} cos2\phi_1 & \frac{\rho_2 V_{S2}{}^2 V_{P1}}{\rho_1 V_{S1}^2 V_P 2}\cos 2\theta_1 & \frac{\rho_2 V_{S2} V_{P1}}{\rho_1 V_{S1}^2}\cos 2\phi_2 \\ -cos\phi_2 & \frac{V_{S1}}{V_{P1}}\sin 2\phi_1 & \frac{\rho_2 V_{P2}}{\rho_1 V_{P1}} & \frac{\rho_2 V_{S2}}{\rho_1 V_{P1}}\sin 2\phi_2 \end{bmatrix}^{-1} \begin{bmatrix} \sin\theta_1 \\ \cos\theta_1 \\ \sin 2\theta_1 \\ \cos 2\phi_1 \end{bmatrix}$$

Scattering theory (Shuey)



Shuey approximation

$$R_{pp}(\theta) = i(\Delta V_P, \Delta\rho) + g(\Delta V_P, \Delta V_S, \Delta\rho)\sin^2\theta$$

Limitations:

Small perturbations over a background trend

valid < 30 degrees

Benefits:

linear for i and g

invert using simple least squares

Shuey term inversion as a projection

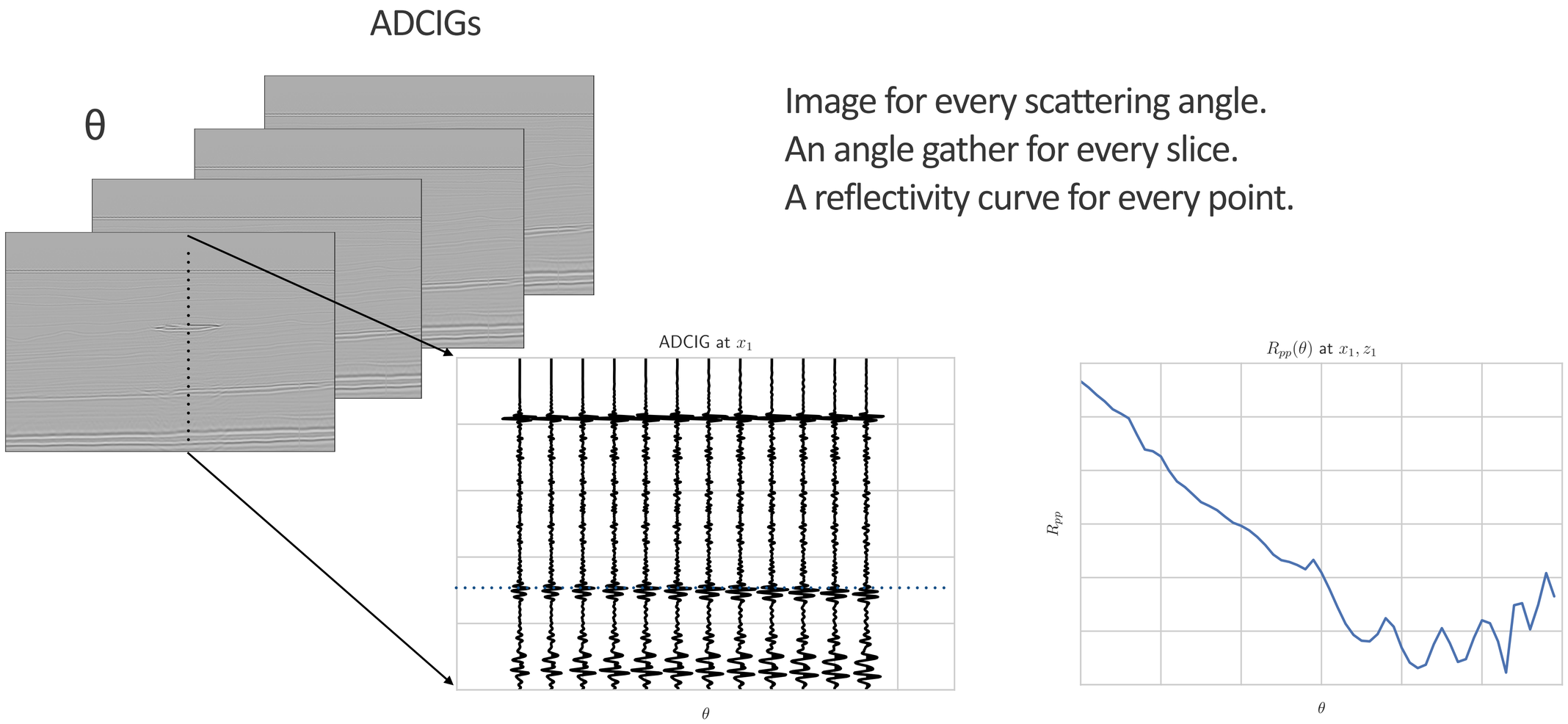
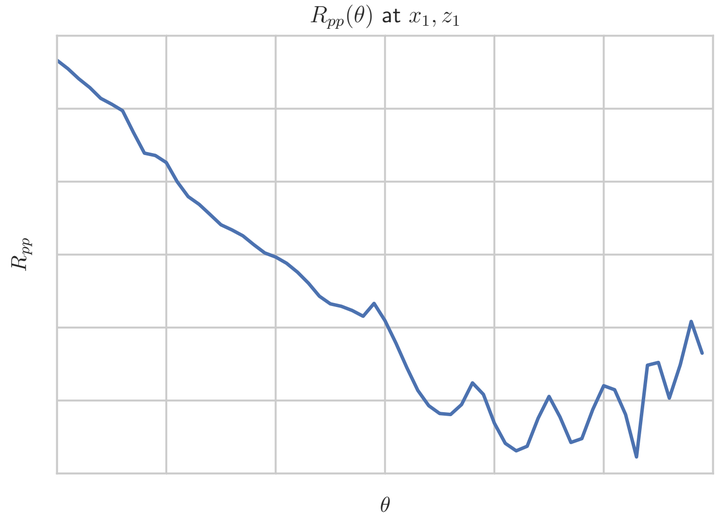
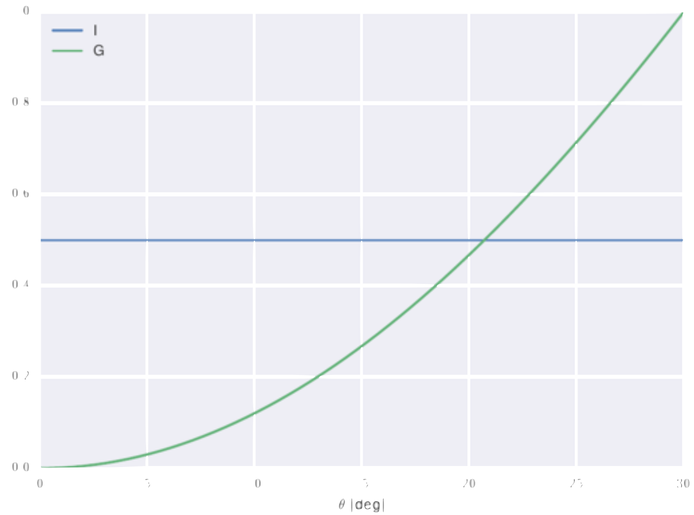


Image for every scattering angle.
An angle gather for every slice.
A reflectivity curve for every point.

Shuey term inversion as a projection

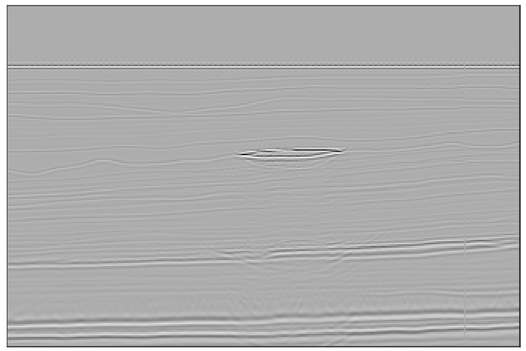




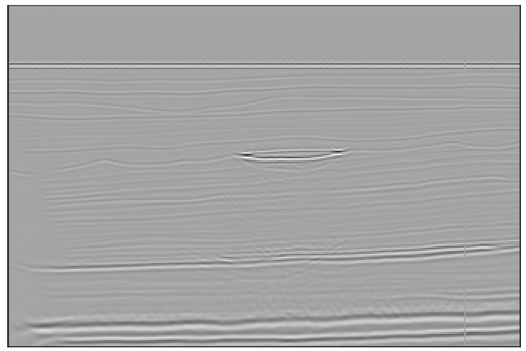
Fit the reflectivity curve using a linear combination of the Shuey vectors.
Each reflectivity curve is projected down to two coefficients.

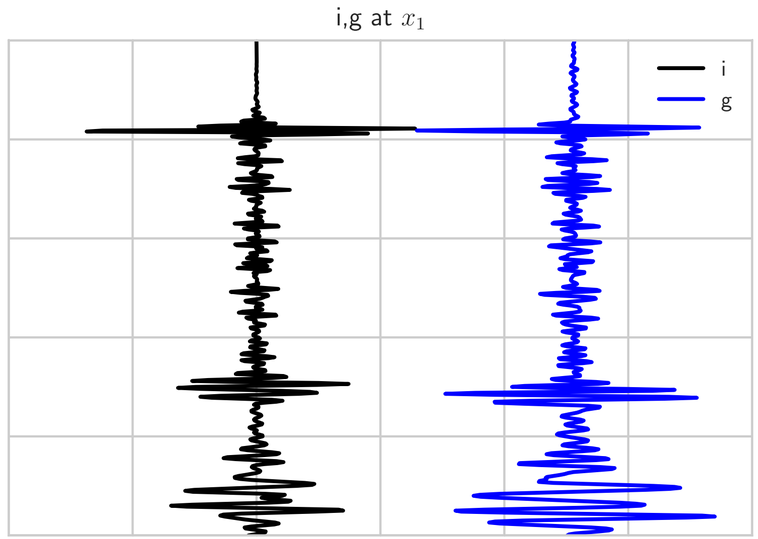
Shuey term inversion as a projection

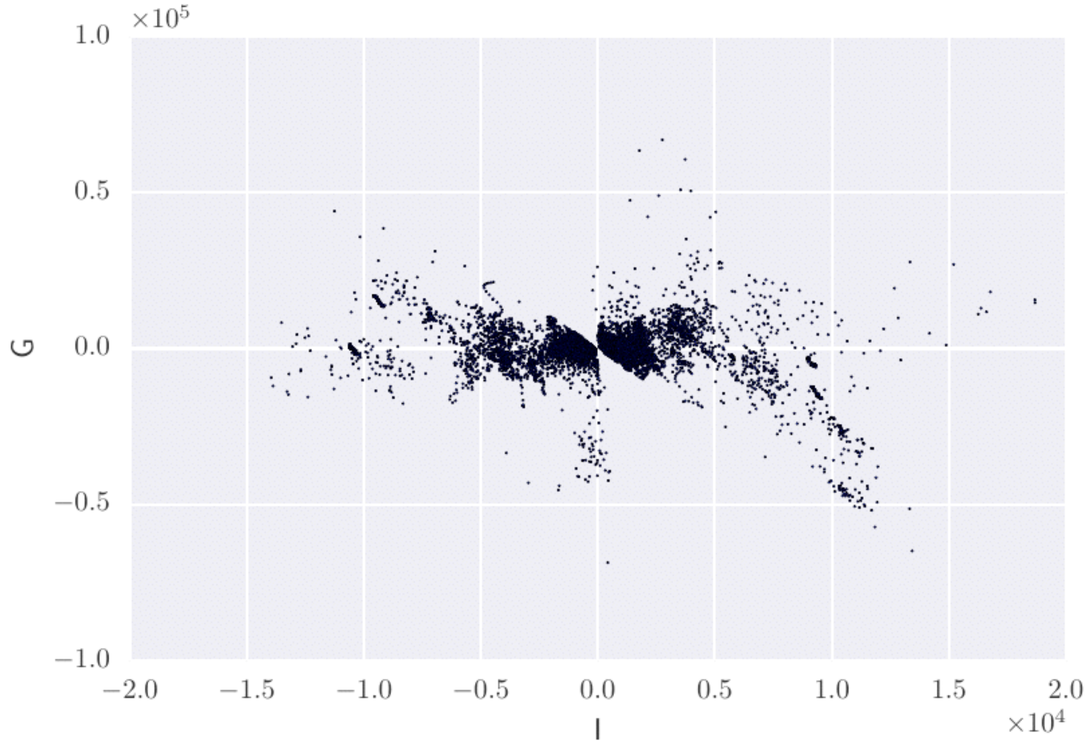
i



g



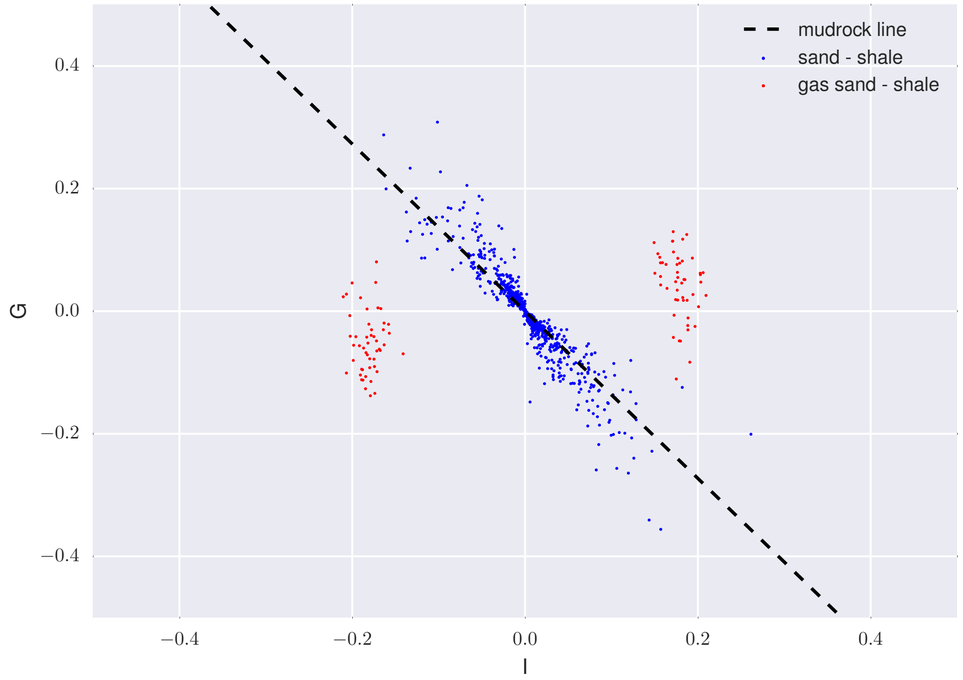




Reduced the dimensionality of the angle gathers to two coefficients.

Projection coefficients can be plotted to analyze the multivariate relationships.

Relation to hydrocarbons



Brine-saturated sands and shales follow a mudrock line:

$$g = \frac{i}{1+k}\left[1 - 4\frac{\langle V_S \rangle}{\langle V_P \rangle}\left(\frac{2}{m} + k\frac{\langle V_S \rangle}{\langle V_P \rangle}\right)\right]$$

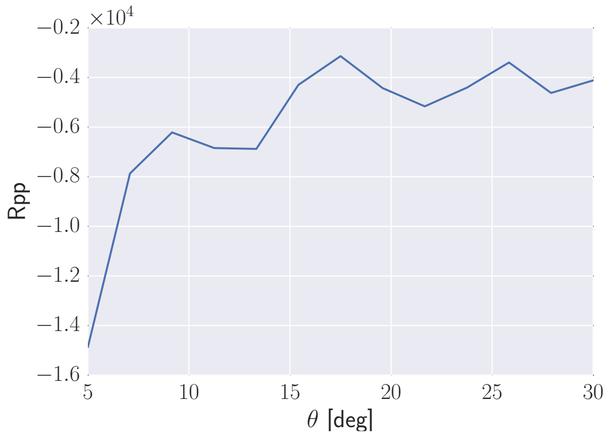
Hydrocarbon saturated sands deviate from this trend.

Hydrocarbon reserves are found from outliers of a crossplot!

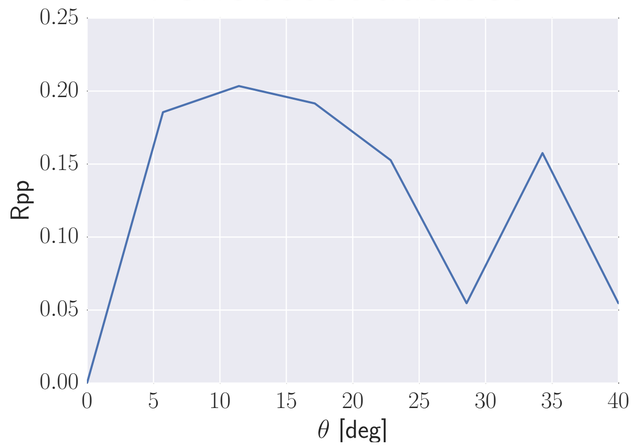
*m, c, k are geological parameters determined empirically from well logs/laboratory measurements

Reality bites

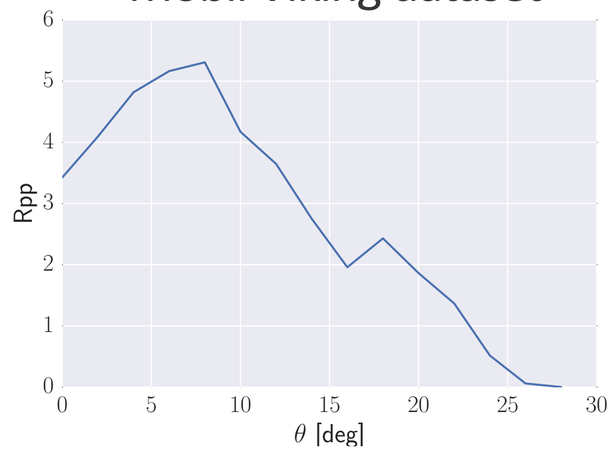
BG dataset

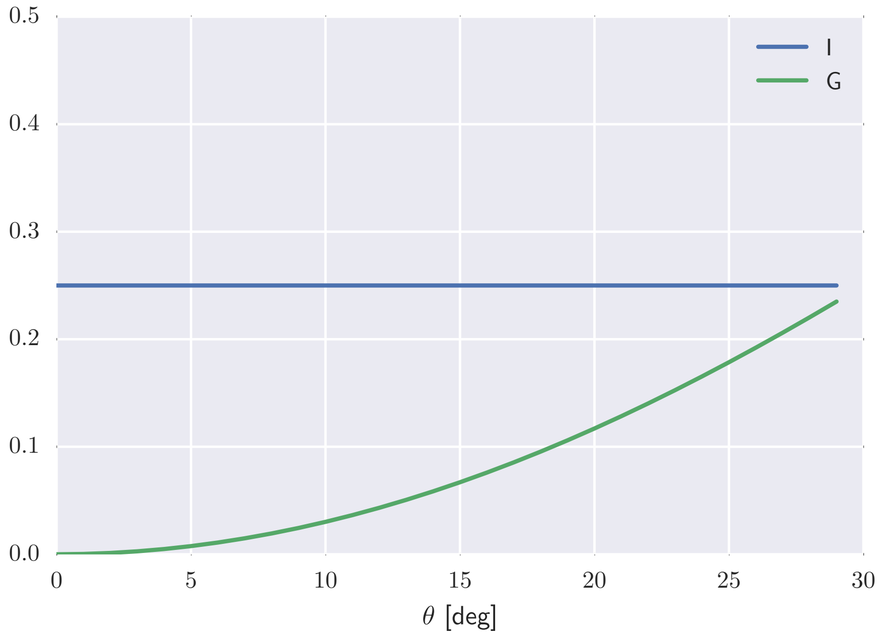


Penobscot dataset



Mobil Viking dataset





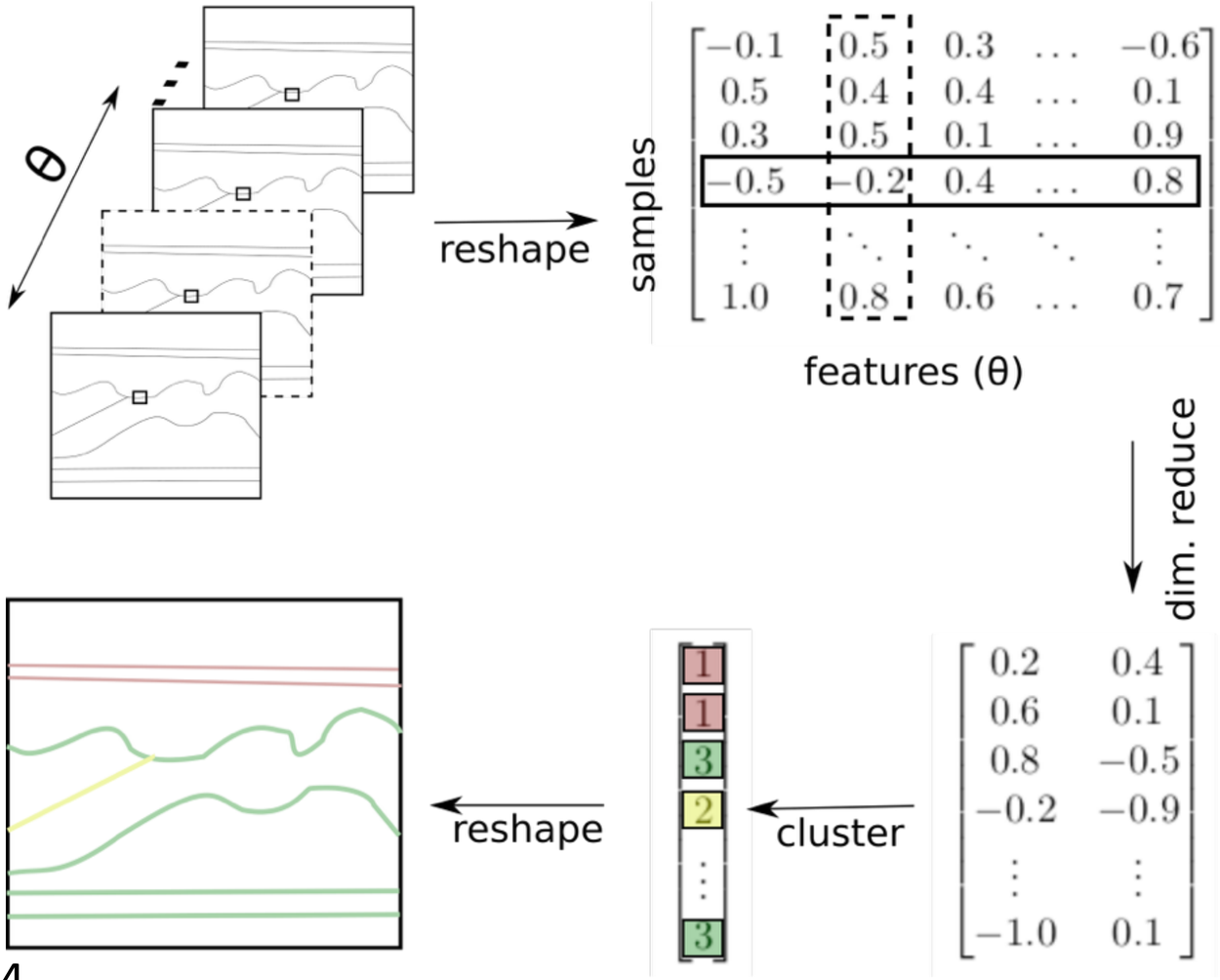
Problem:

Shuey components can't explain the features in real data.

Solution:

Use unsupervised machine learning to find better projections.

Unsupervised learning problem



$$X \in \mathbb{R}^{n \times d}$$

n is number of samples in the image,
d is the number of angles

Principle component analysis (PCA)

Eigendecomposition of the covariance matrix:

$$C = X^T X = \frac{1}{n} \sum_i^n \mathbf{x}_i \mathbf{x}_i^T$$

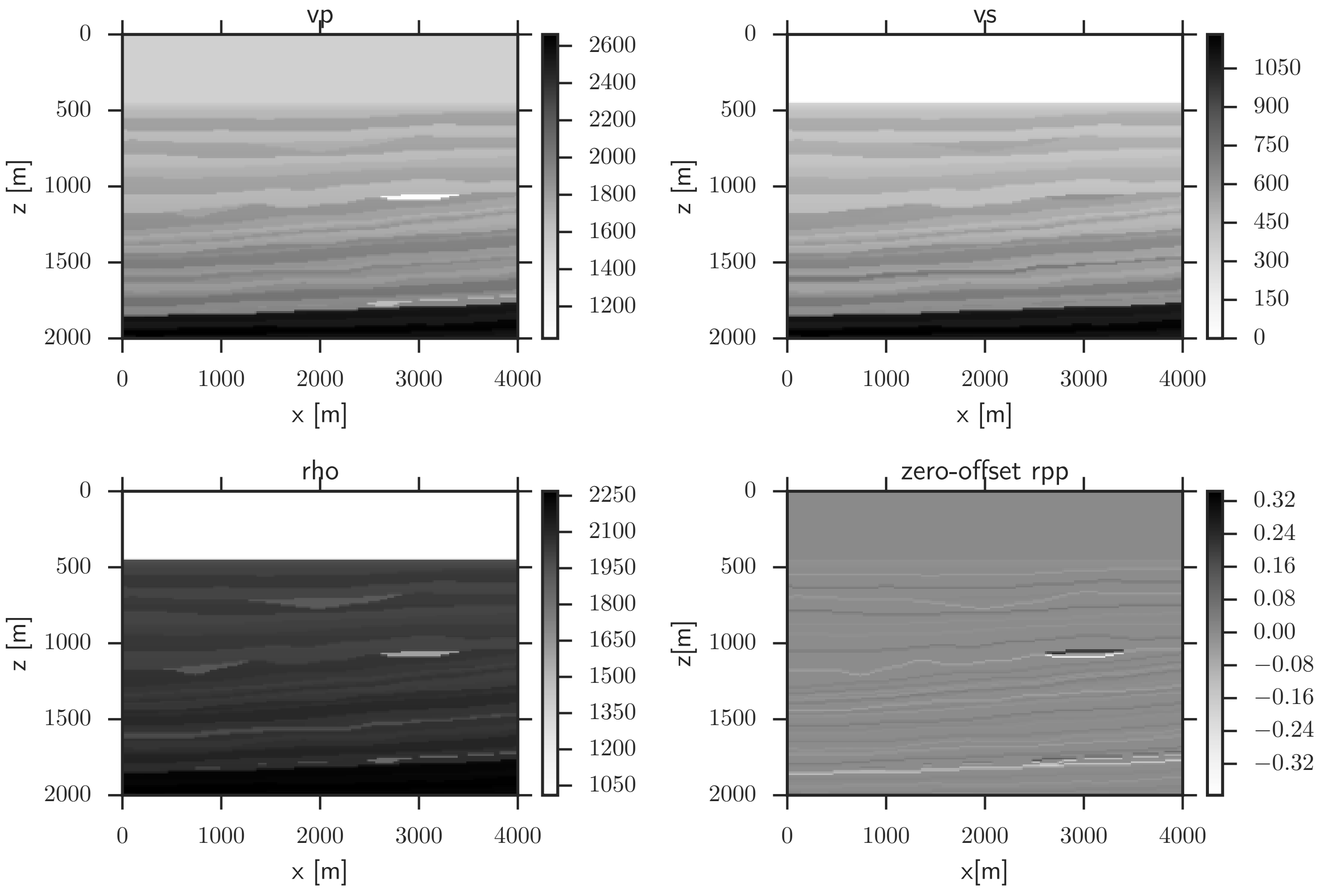
Project onto the eigenvectors with the two largest eigenvalues.

Maximizes the variance (a measure of information).

DEMO

http://ec2-54-224-182-64.compute-1.amazonaws.com/#/pca

Marmousi II Earth model

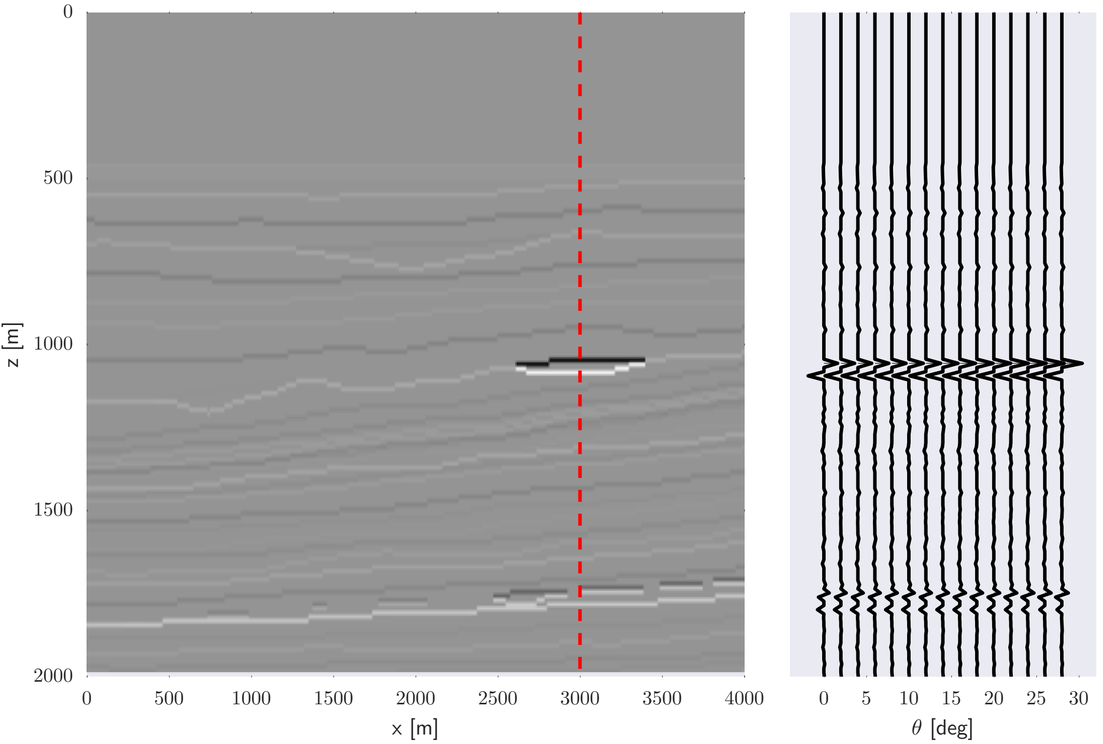


Specifically made for testing amplitude vs. offset analysis

Contains gas-saturated sand embedded in shales and brine-sands.

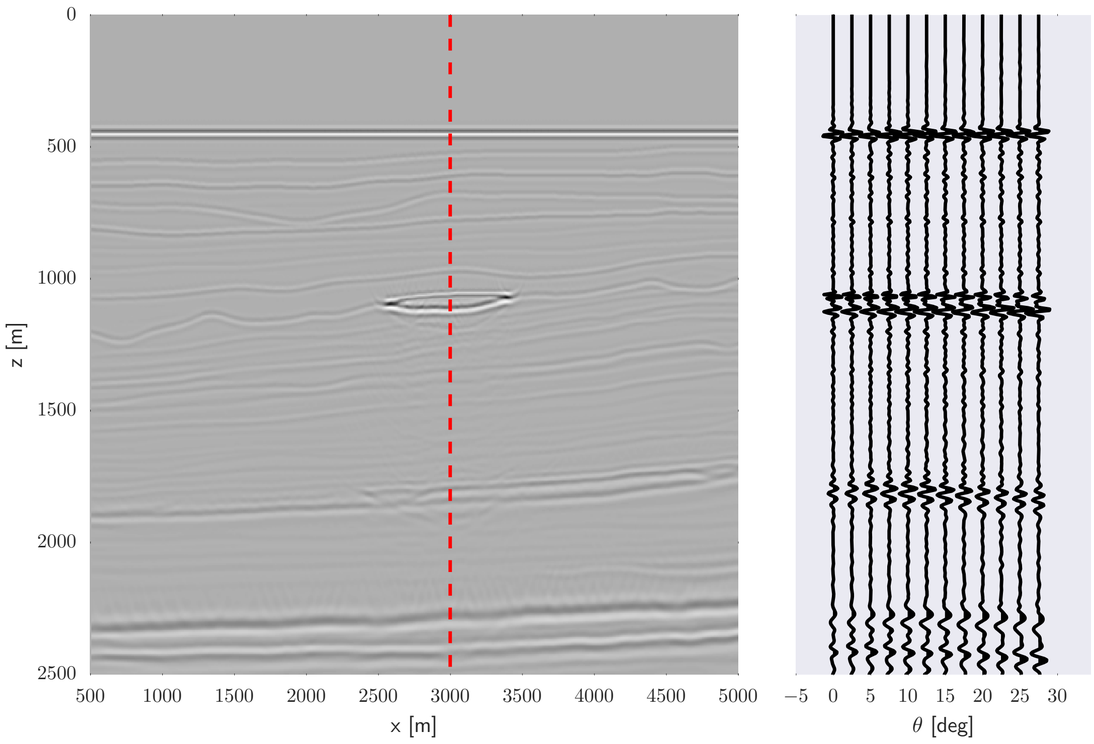
Seismic modeling

True reflectivity



Physically consistent with the Zoeppritz equations

Migrated seismic



Migrated visco-acoustic survey

DEMO

http://ec2-54-224-182-64.compute-1.amazonaws.com/#/avo

Kernel PCA

Problem:

Find a non-linear projection that provides better discrimination of trends and outliers.

Solution:

Use the “kernel-trick” to compute PCA in a high-dimensional non-linear feature space

Kernel trick

PCA can be calculated from the Gramian inner product matrix:

$$XX^T = \begin{pmatrix} \langle \mathbf{x}_1, \mathbf{x}_1 \rangle & \langle \mathbf{x}_1, \mathbf{x}_2 \rangle & \dots & \langle \mathbf{x}_1, \mathbf{x}_n \rangle \\ \langle \mathbf{x}_2, \mathbf{x}_1 \rangle & \langle \mathbf{x}_2, \mathbf{x}_2 \rangle & \dots & \langle \mathbf{x}_2, \mathbf{x}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{x}_n, \mathbf{x}_1 \rangle & \langle \mathbf{x}_n, \mathbf{x}_2 \rangle & \dots & \langle \mathbf{x}_n, \mathbf{x}_n \rangle \end{pmatrix}$$

Replace $\langle \mathbf{x}_i, \mathbf{x}_j \rangle$ with a kernel $\kappa(\mathbf{x}_i, \mathbf{x}_j)$

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^T \mathbf{x}_j + b)^c = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$$

Example c=2, b=1

$$\phi(\mathbf{x}) = [1, \sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1 x_2]$$

DEMO

http://ec2-54-224-182-64.compute-1.amazonaws.com/#/avo

Recap

Situation: Exploring seismic data for anomalous responses.

Problem: Physical model can not explain real world data.

Solution: Learn useful projections directly from the data.

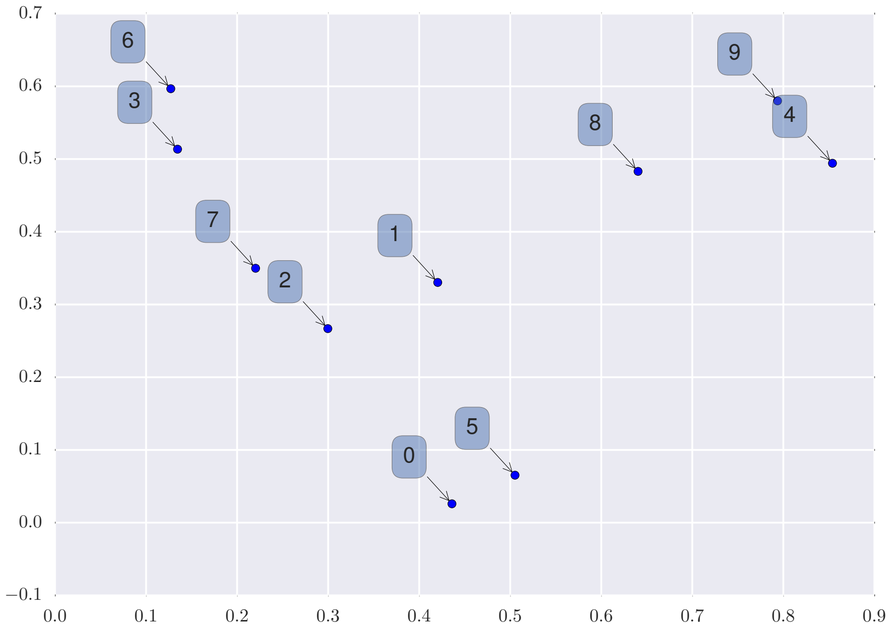
Assessment:

- PCA is equivalent for physically consistent data, but more robust to processing/acquisition artifacts.
- Kernel PCA makes outliers linearly separable from the background.

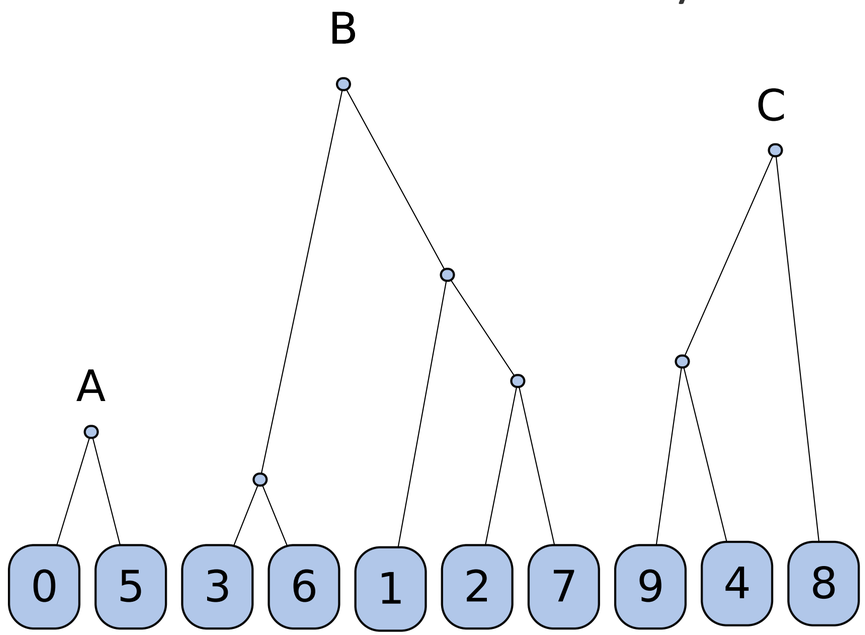
Next: Automatic segmentation (clustering)

Hierarchical clustering

Unclassified data



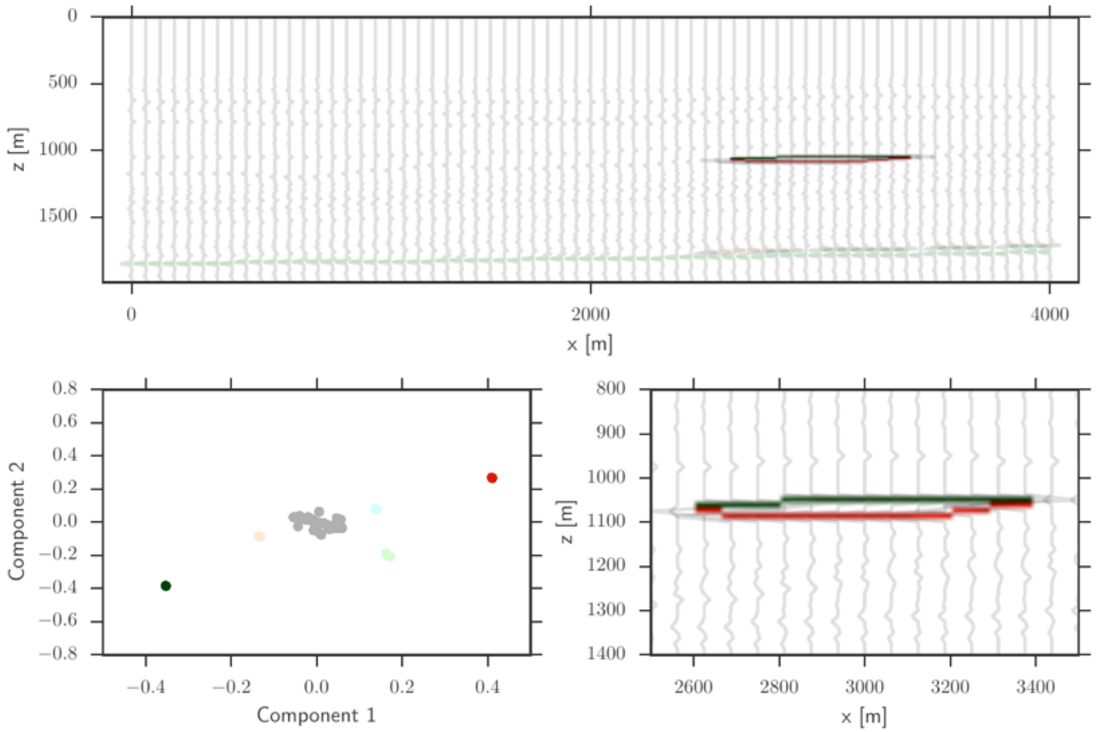
Classification hierarchy



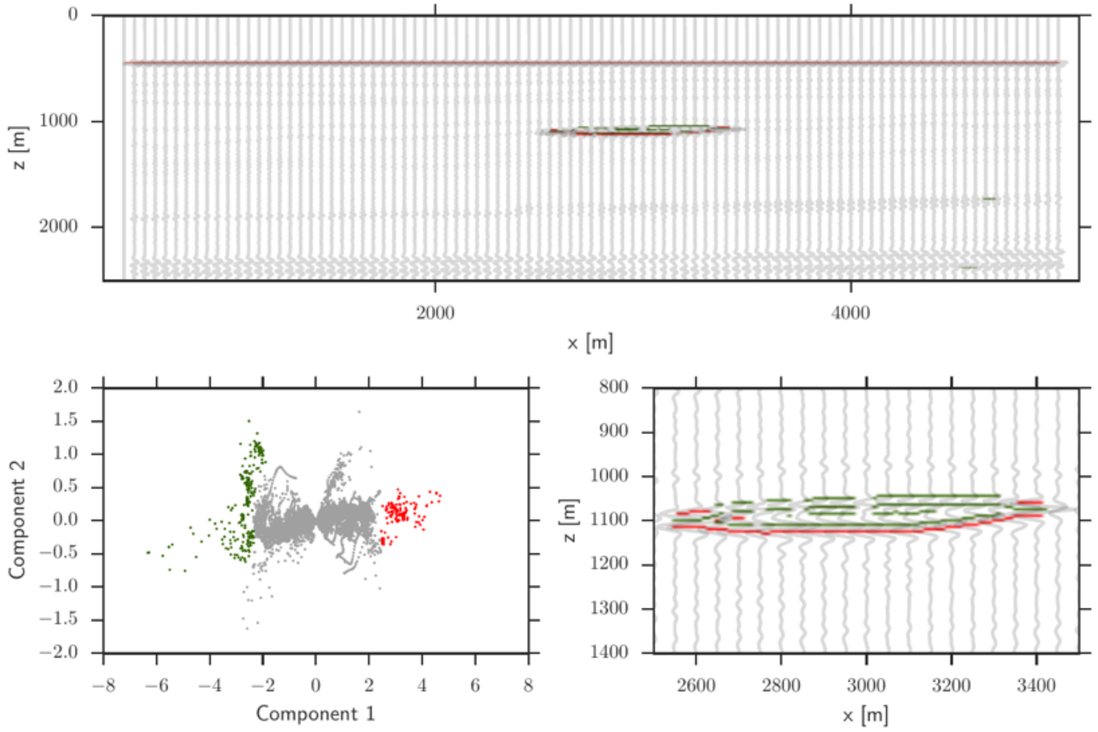
Each point is initially considered a cluster.
Each iteration merges the closest clusters into a larger cluster.
Builds a hierarchy until a *defined* number of classes is reached.

Results - PCA

Clustered physically consistent data

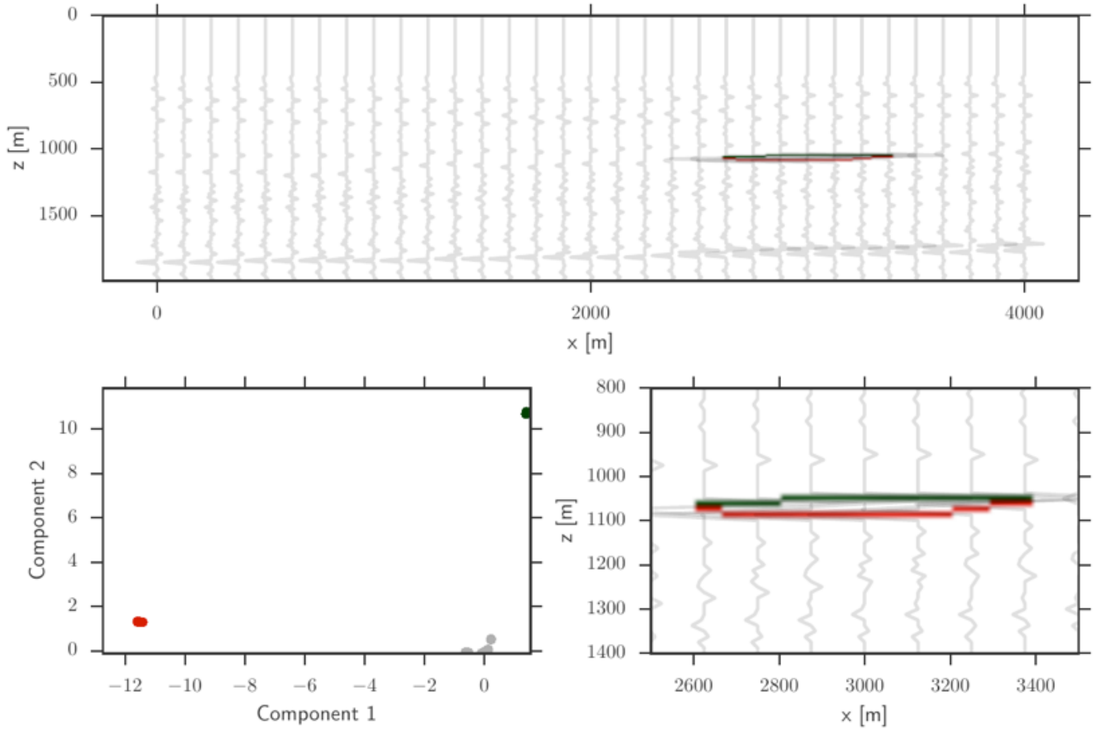


Clustered migrated data

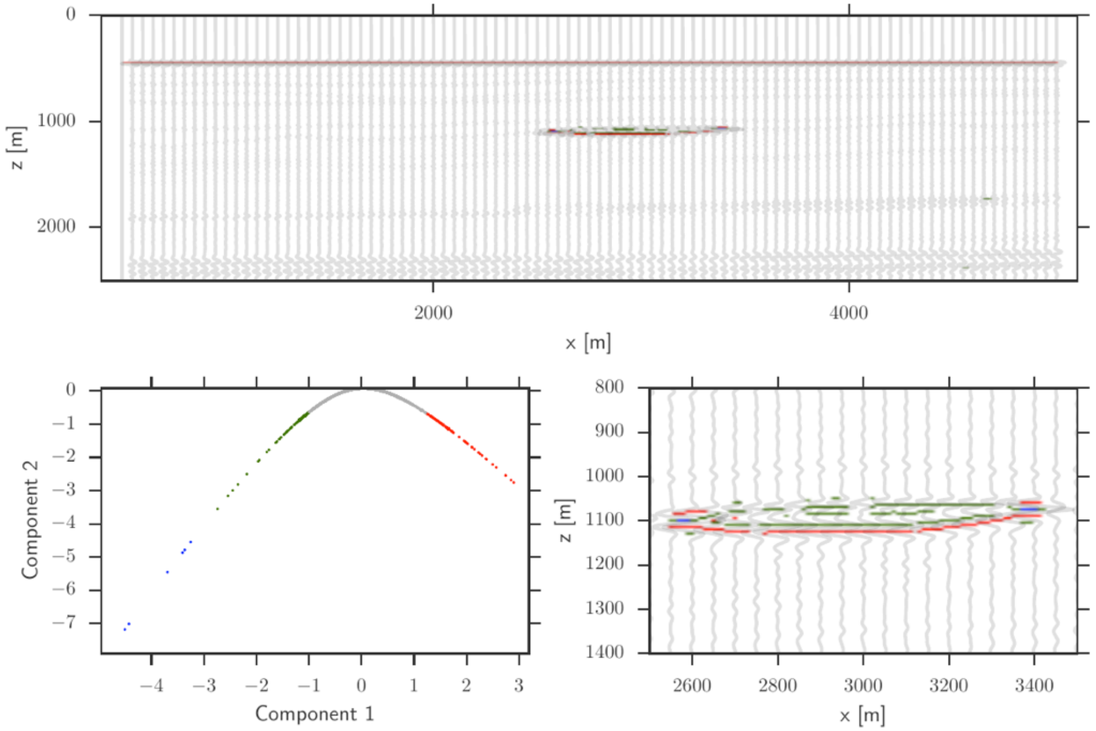


Results - Kernel PCA

Clustered physically consistent data



Clustered migrated data



Summary

Successes:

Each projection could segment the reservoir.

Kernel PCA provided advantageous multivariate geometries (linearly separable).

Challenges:

Clustering is highly sensitive to user chosen parameters.

Kernel PCA is computationally expensive and lacks interpretation.

Robust PCA

Problem:

Find a sparse set of outlying reflectivity responses against a background trend.

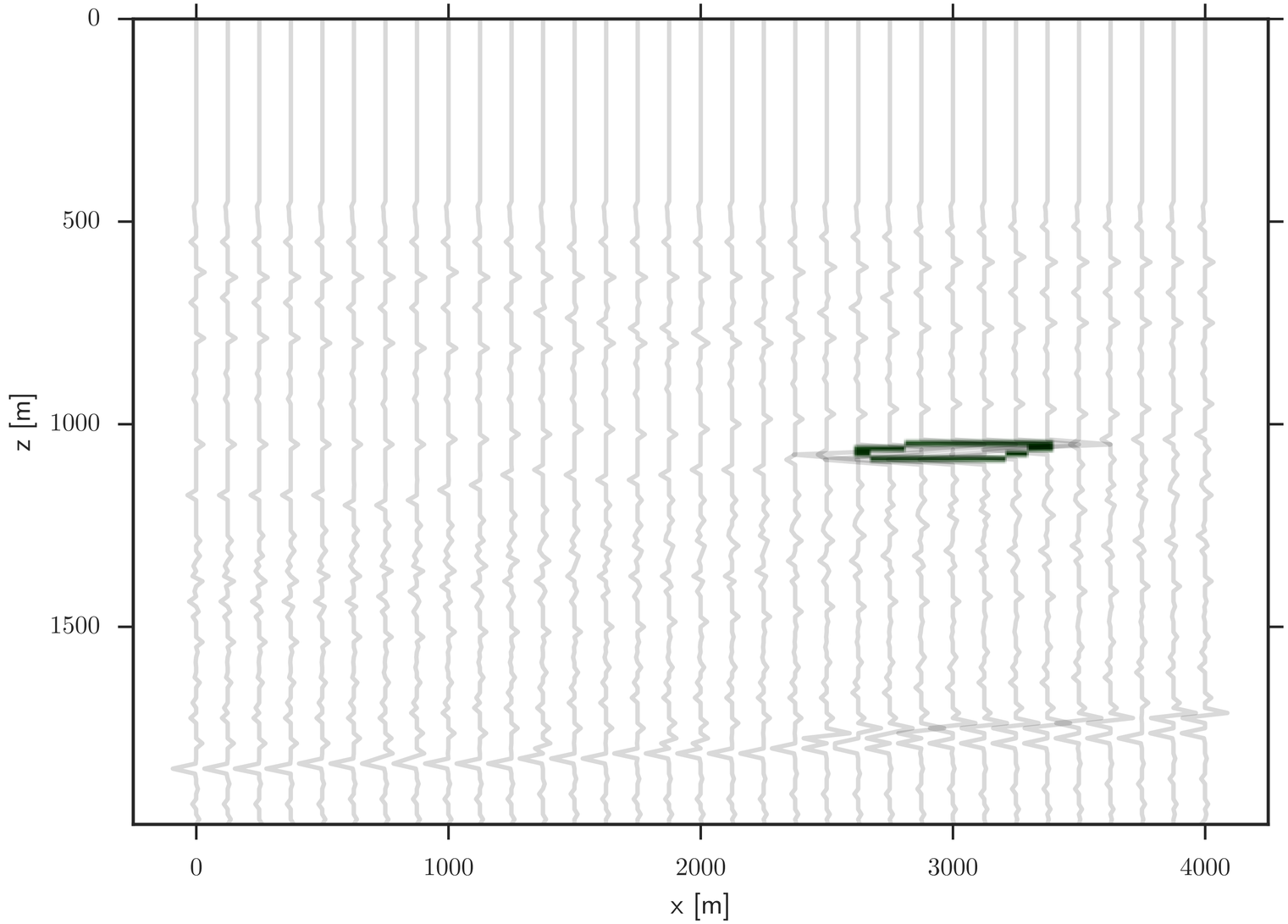
Assumption:

The background trend of similar curves is highly redundant, which forms a low rank matrix.

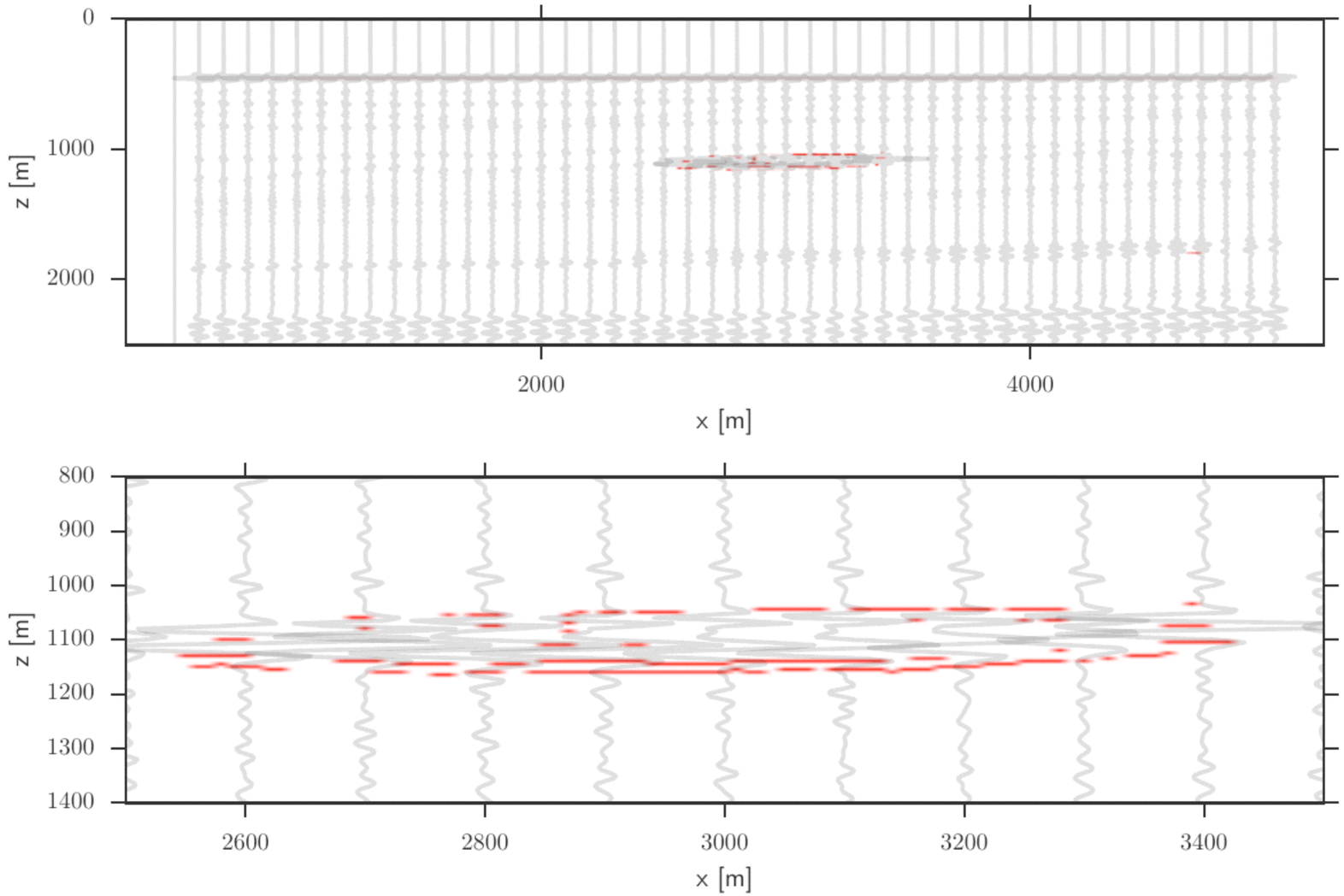
Solution:

$$\min_{L,S} \quad \|L\|_* + \lambda \|S\|_{1,\infty} \text{ s.t.} \quad L + S = X$$

Results - physically consistent data



Results - migrated seismic



Summary

Successes:

Segmentation of reservoir in both images.

Physically interpretable segmentation without clustering.

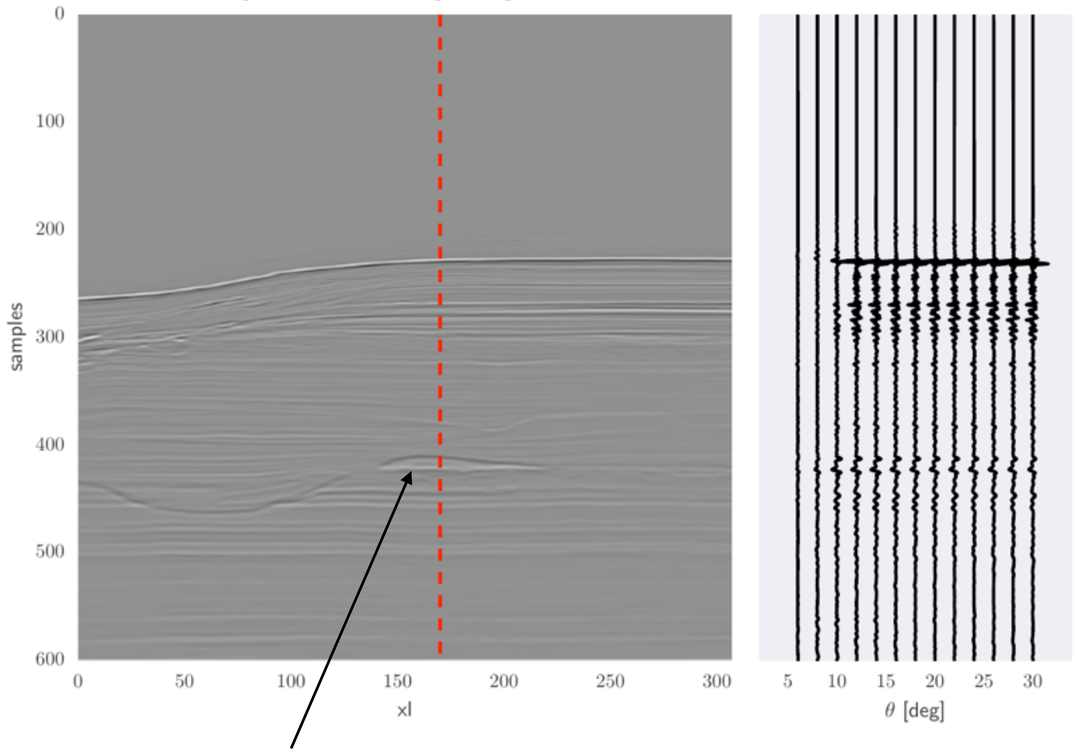
Challenges:

Requires tuning of one optimization trade off parameter.

Convergence sensitive to the rank of outliers, not well understood.

Comparison on field data

Migrated angle gathers



Data provided by BG group.
Interpreted to contain a potential gas reserve.

Compare unsupervised methods to BG group's approach: Dynamic Intercept Gradient Inversion (DIGI).

Note: Clustering approaches were not directly useful.

DIGI-inverse problem

$$\begin{bmatrix} \mathbf{d} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} W & W\sin^2\theta \\ \lambda\nabla & \lambda\nabla \\ W(\theta_{me})\cos(\chi_{me}) & W(\theta_{me})\sin(\chi_{me}) \end{bmatrix} \begin{bmatrix} \mathbf{i} \\ \mathbf{g} \end{bmatrix}$$

Convolution forward model: $d(x,t,\theta) = w(x,t,\theta) * r(x,t,\theta)$

- *ill-posed,* $\lambda\nabla$ term forces a smooth answer

Further augmented by extended elastic reflectivity (EER) term:

$$EER(\chi_{me}) = i\cos(\chi_{me}) + g\sin(\chi_{me})$$

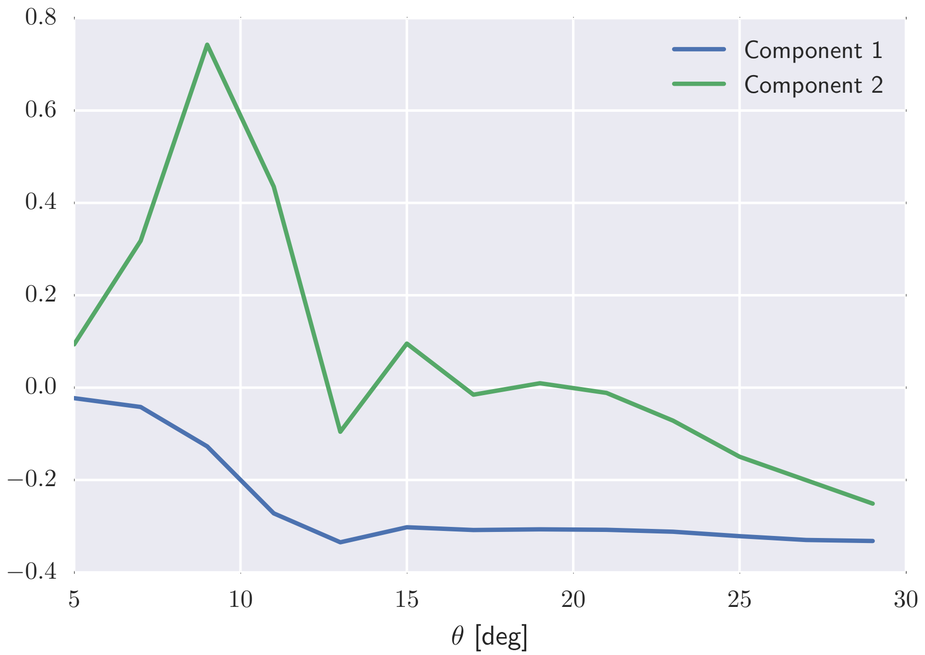
- promotes correlation between *i* and *g*
- χ_{me} is related a priori geological information

System is solved using the conjugate gradient based algorithm LSQR.

PCA extended DIGI

$$\begin{bmatrix} \mathbf{d} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} W & W\sin^2\theta \\ \lambda\nabla & \lambda\nabla \\ W(\theta_{me})\cos(\chi_{me}) & W(\theta_{me})\sin(\chi_{me}) \end{bmatrix} \begin{bmatrix} \mathbf{i} \\ \mathbf{g} \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{d} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} W\mathbf{c_1} & W\mathbf{c_2} \\ \lambda\nabla & \lambda\nabla \\ W(\theta_{me})\cos(\chi_{me}) & W(\theta_{me})\sin(\chi_{me}) \end{bmatrix} \begin{bmatrix} \mathbf{i} \\ \mathbf{g} \end{bmatrix}$$

Same inverse problem, but use the two largest principle components instead of the Shuey terms.



Minimum energy projection

$$EER(\chi_{me}) = i\cos(\chi_{me}) + g\sin(\chi_{me})$$

forms an image where large values correspond to uncorrelated *i* and *g* terms.

Thresholding this image will therefore segment outliers.

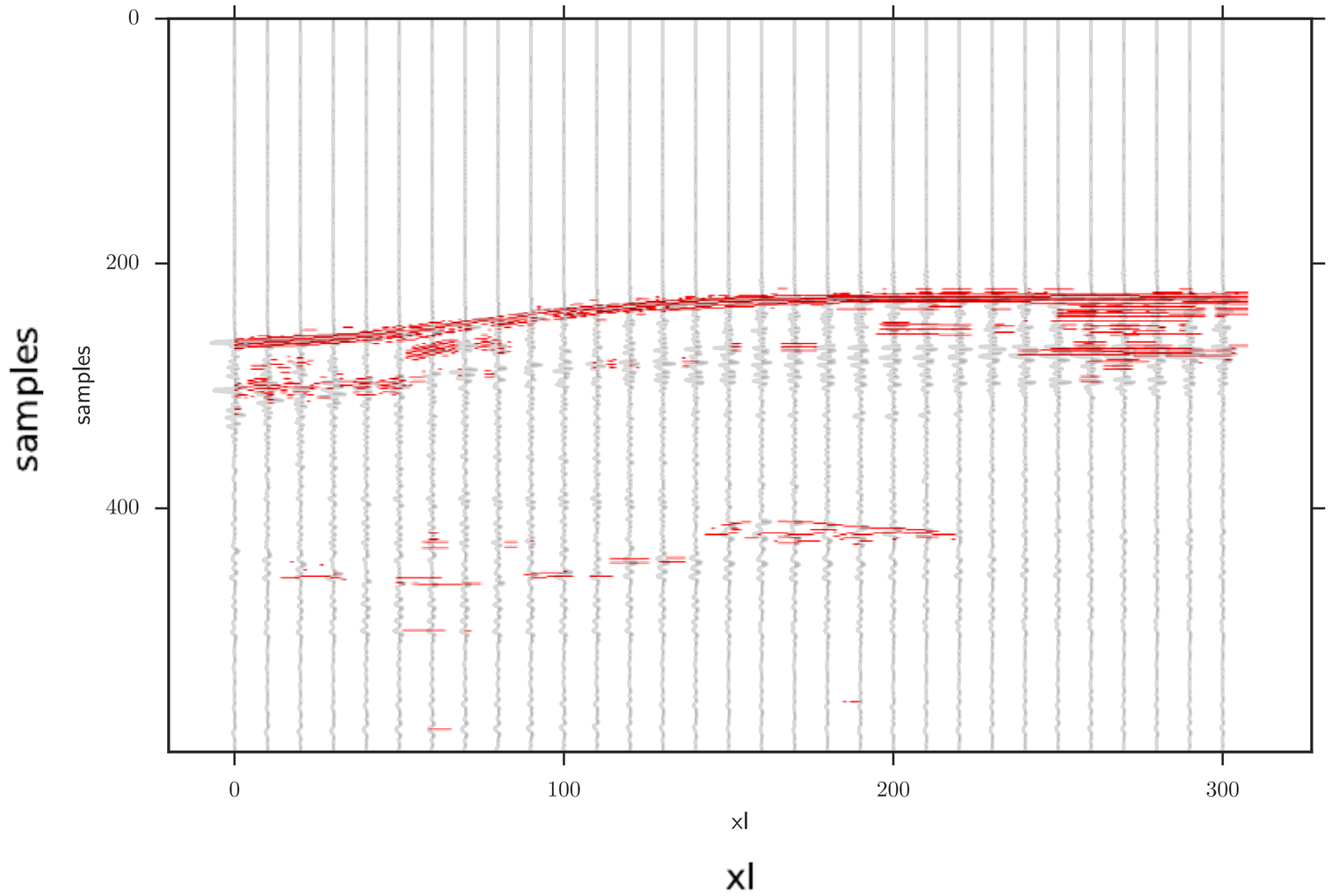
Comparison method

Manually threshold the image to segment the potential hydrocarbon reserve while maintaining the least amount of spurious segmentation.

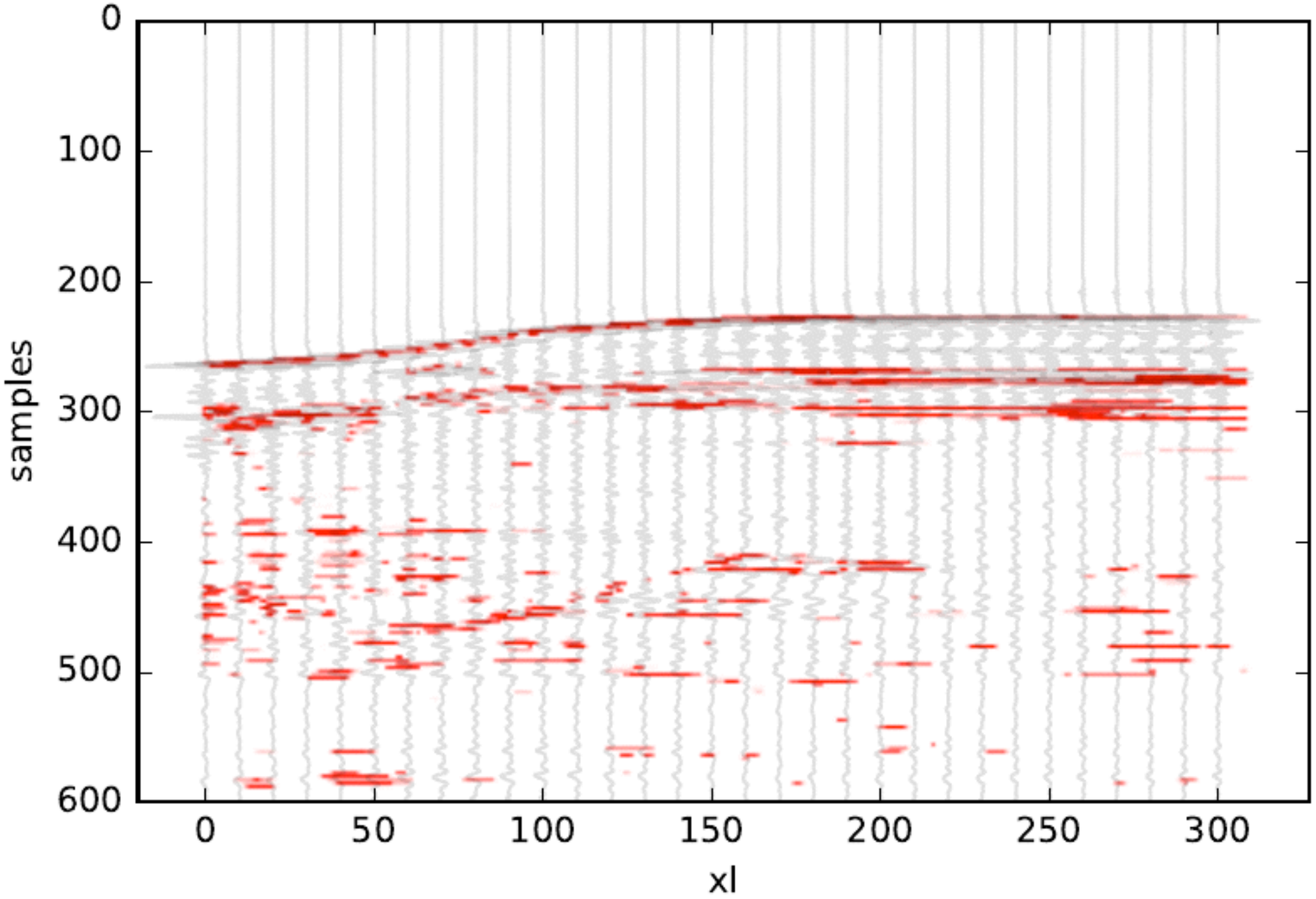
Compare:

- Robust PCA
- DIGI
- PCA extended

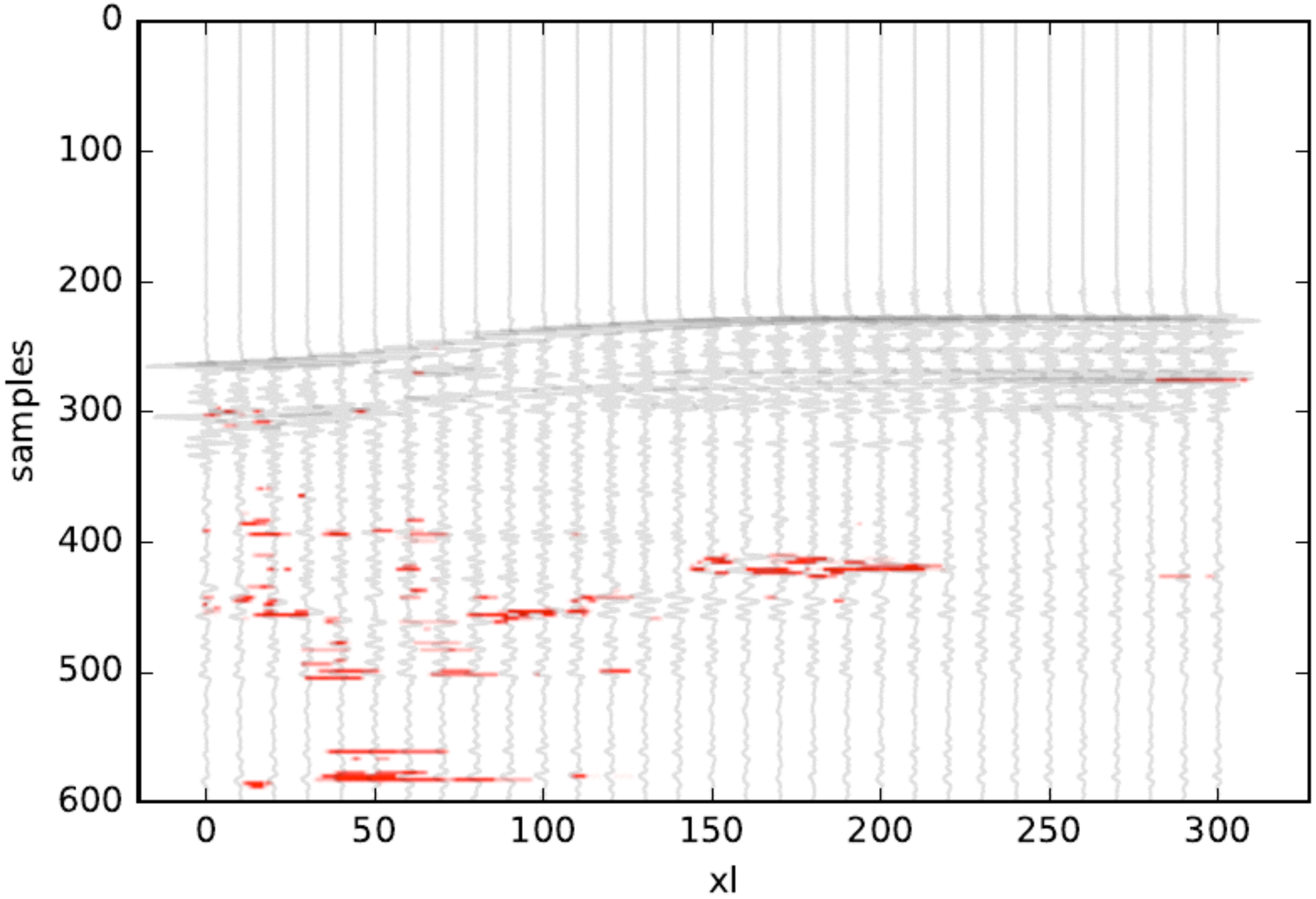
Results - robust PCA



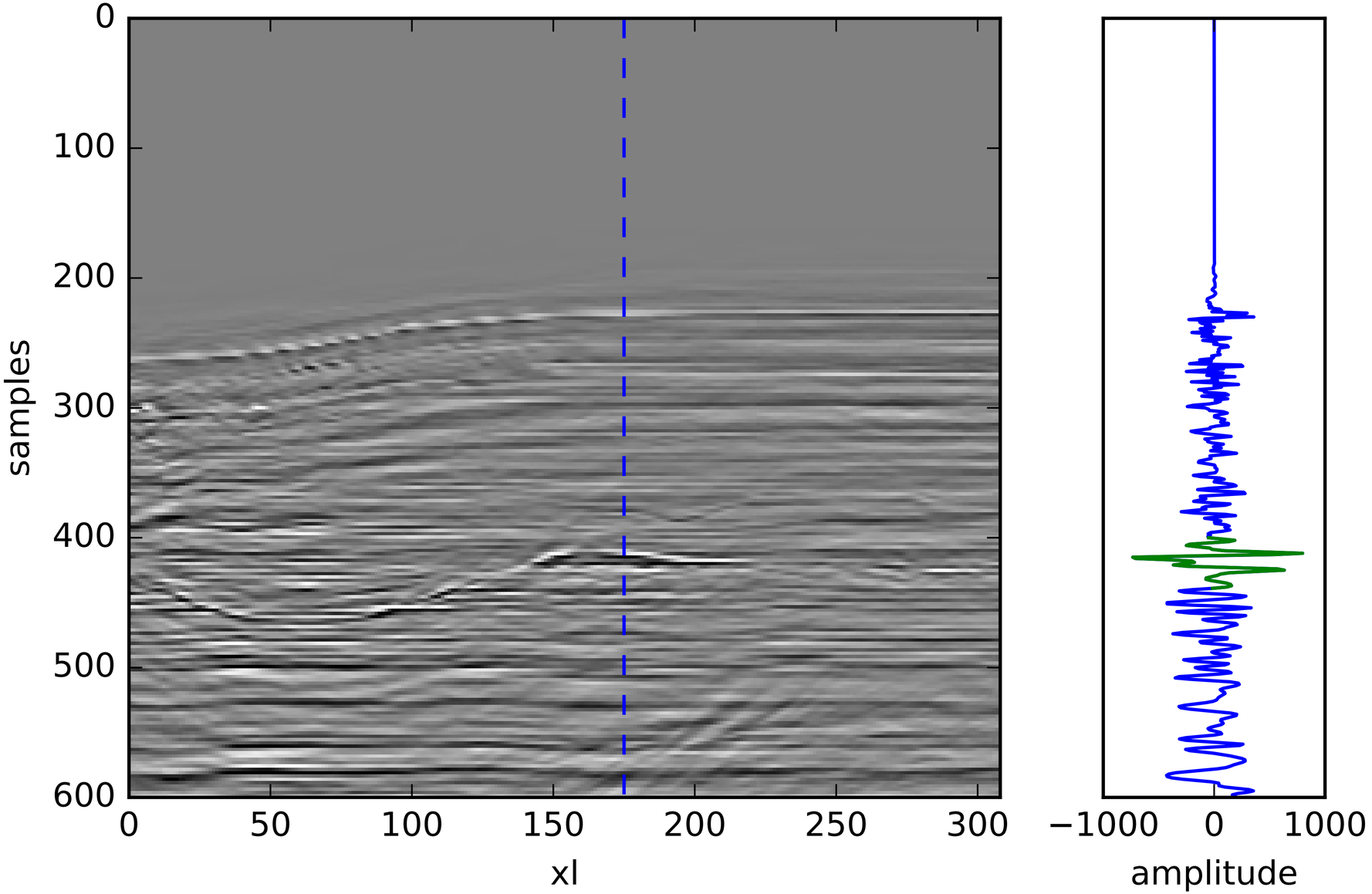
Results - DIGI



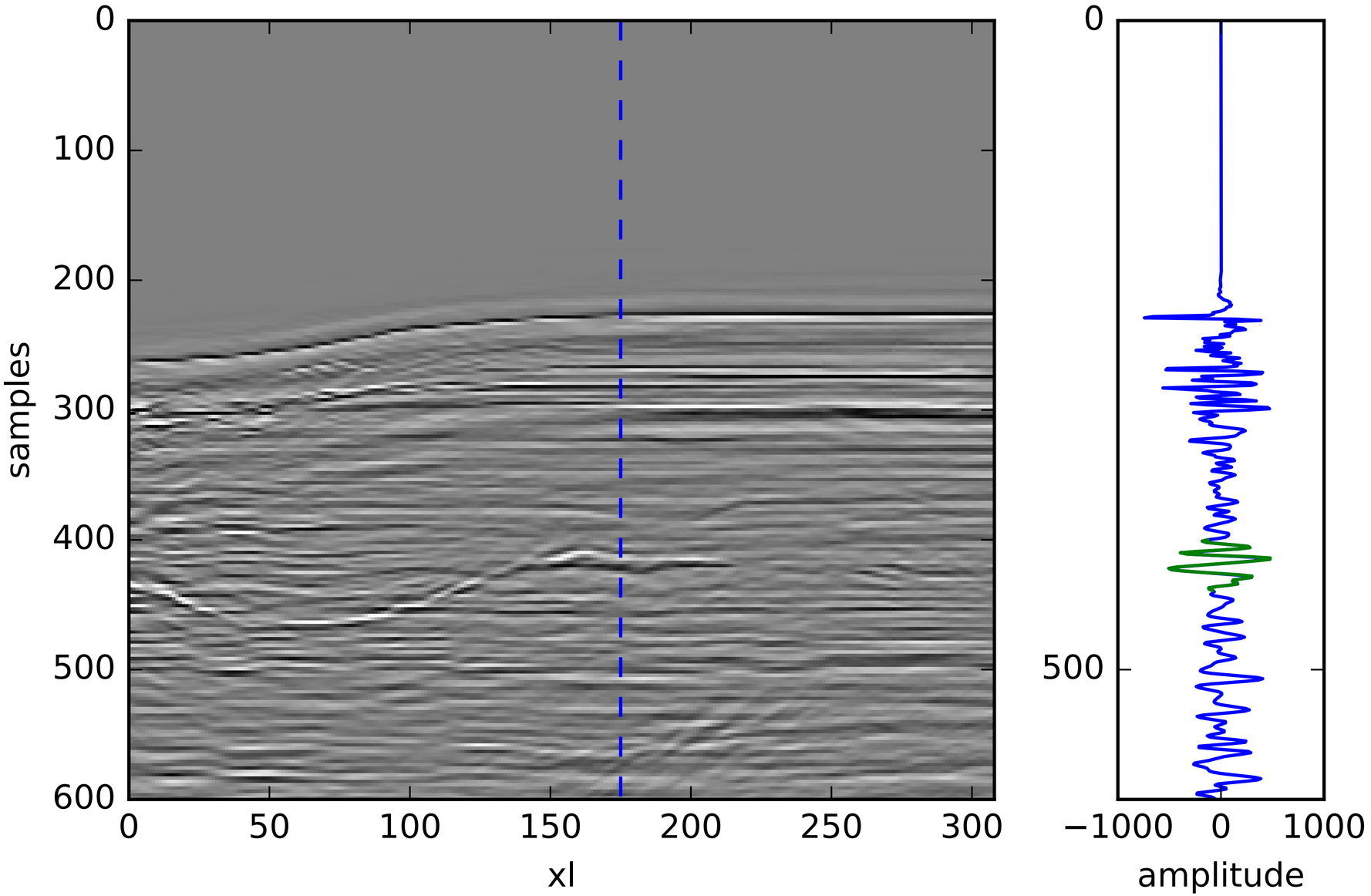
Results - PCA extend DIGI



Results - EER using PCA extended DIGI

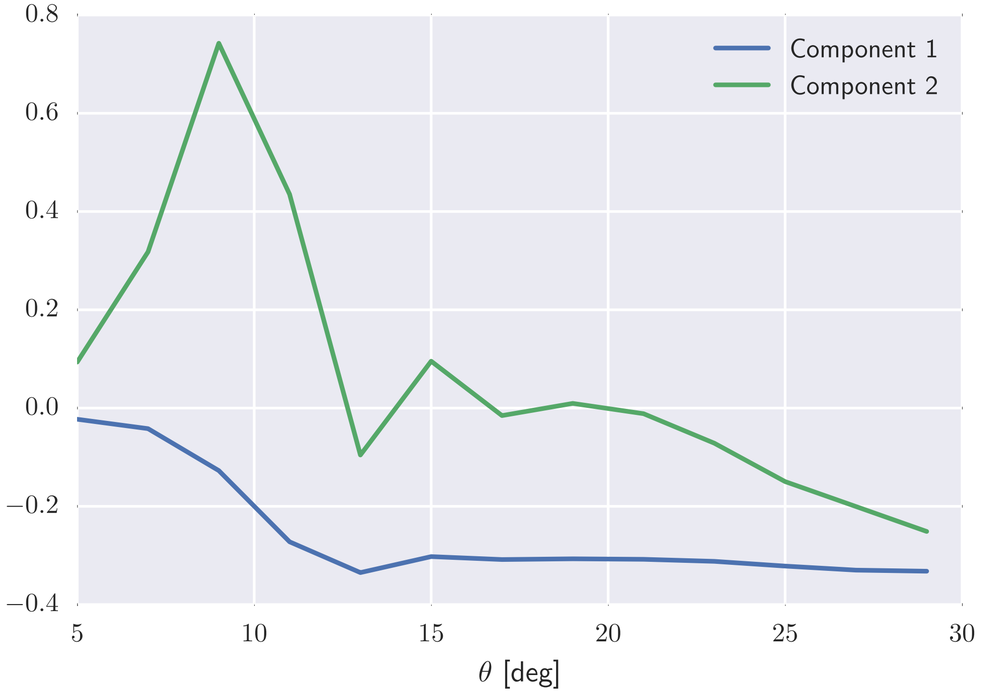


Results - EER using DIGI

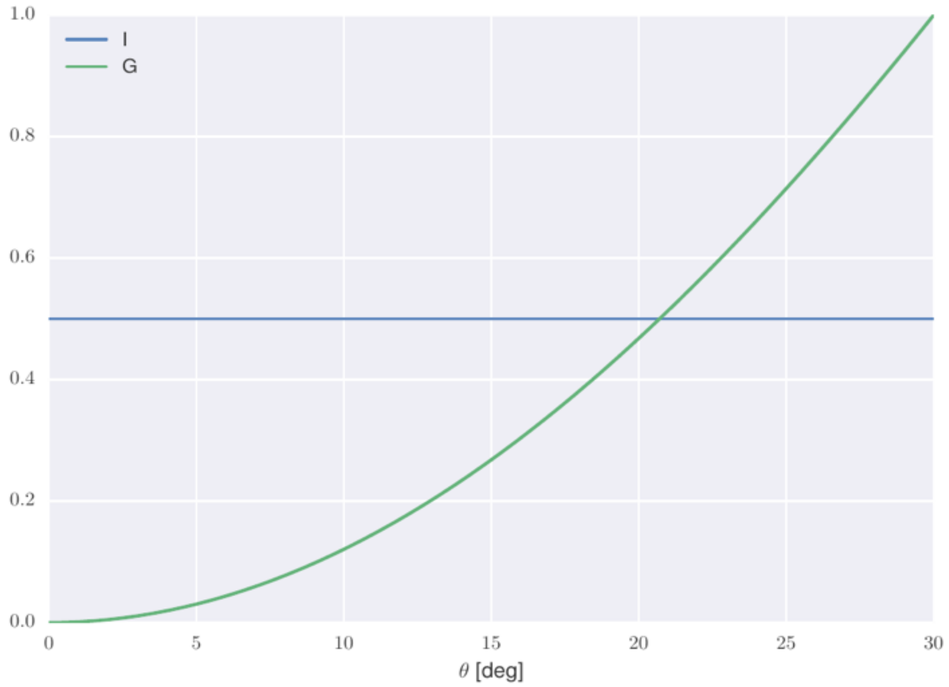


Why the difference?

Principle components



Shuey terms



Principle components can explain more features in the data.

Summary

- Robust PCA provided the best image segmentation.
- PCA extended DIGI better separated the potential reservoir from the background trend.
- The extracted principle components showed significantly different shapes than the Shuey vectors.

Epilogue

Outcome:

Generalized a conventional analysis approach using unsupervised learning models.

Successful in segmented potential hydrocarbon reserves from seismic data

Future:

More data, standardized datasets

Quantitative benchmarks

Thanks

References

[1]Zoeppritz, K., 1919, VII b. Über Reflexion und Durchgang seismischer Wellen durch Unstetigkeitsflächen: Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, 1919, 66–84

[2]Shuey, R. T., 1985, A simplification of the Zoeppritz equations: Geophysics, 50, 609–614.

[3]Gardner, G. H. F., L. W. Gardner, and A. R. Gregory, 1974, Formation velocity and density; the diagnostic basics for stratigraphic traps: Geophysics, 39, 770–780.

[4]Castagna, J. P., M. L. Batzle, and R. L. Eastwood, 1985, Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks: Geophysics, 50, 571–581.

[5]Castagna, J. P., H. W. Swan, and D. J. Foster, 1998, Framework for AVO gradient and intercept interpretation: Geophysics, 63, 948–956.

[6]Edgar, J., and J. Selvage, 2013, Dynamic Intercept-gradient Inversion