

Compressive seismic imaging

Felix J. Herrmann

joint work with

Tim Lin*, Peyman Moghaddam*, Gilles Hennenfent*,

Deli Wang* & Chris Stolk (Universiteit Twente)

*Seismic Laboratory for Imaging and Modeling

slim.eos.ubc.ca

von Neuman Meeting 2007,

Snowbird, July 11

Motivation

Seismic imaging involves extremely large high-dimensional data (petabytes= 2^{50} bytes)

Application and synthesis of disc. operators expensive

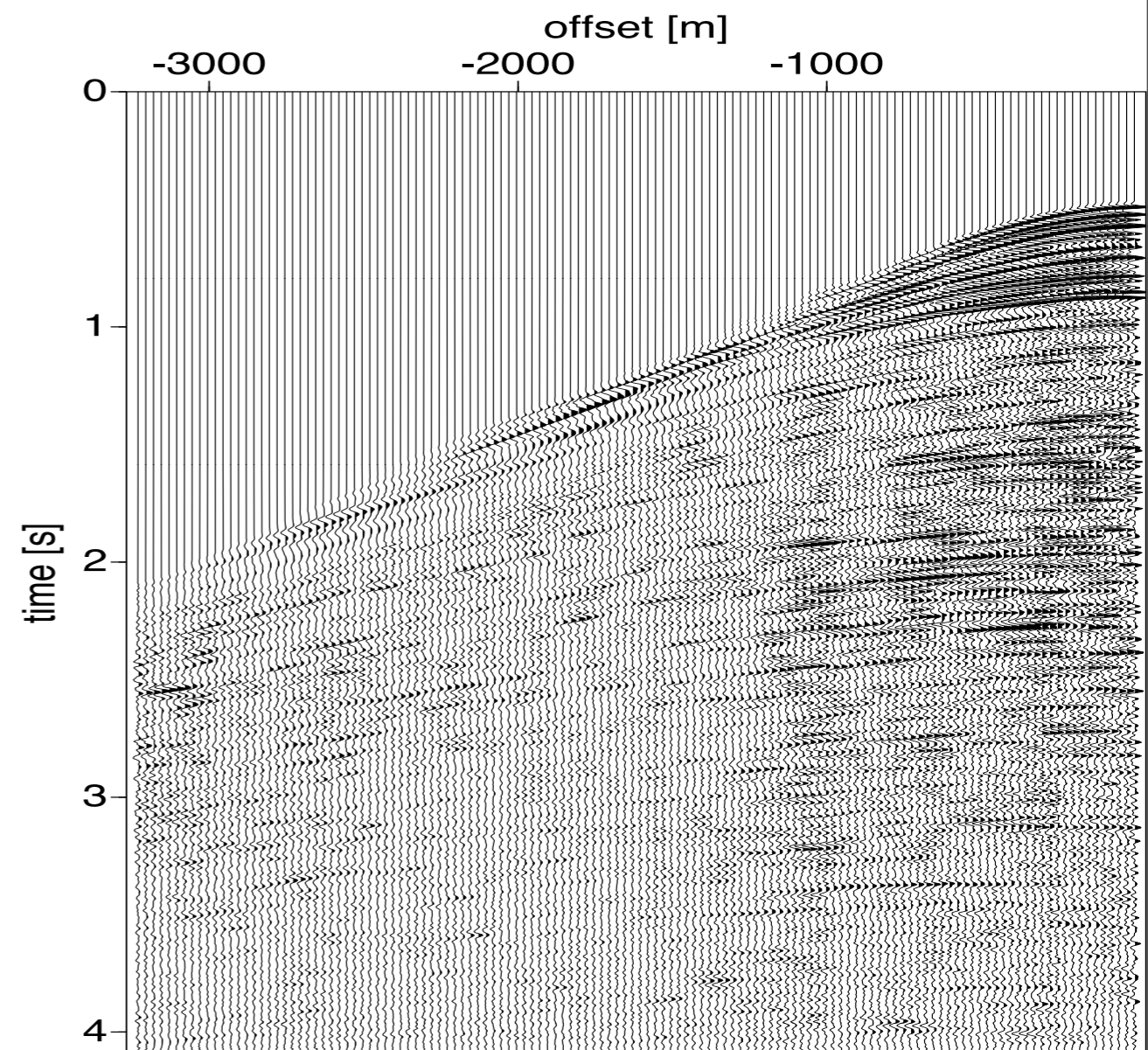
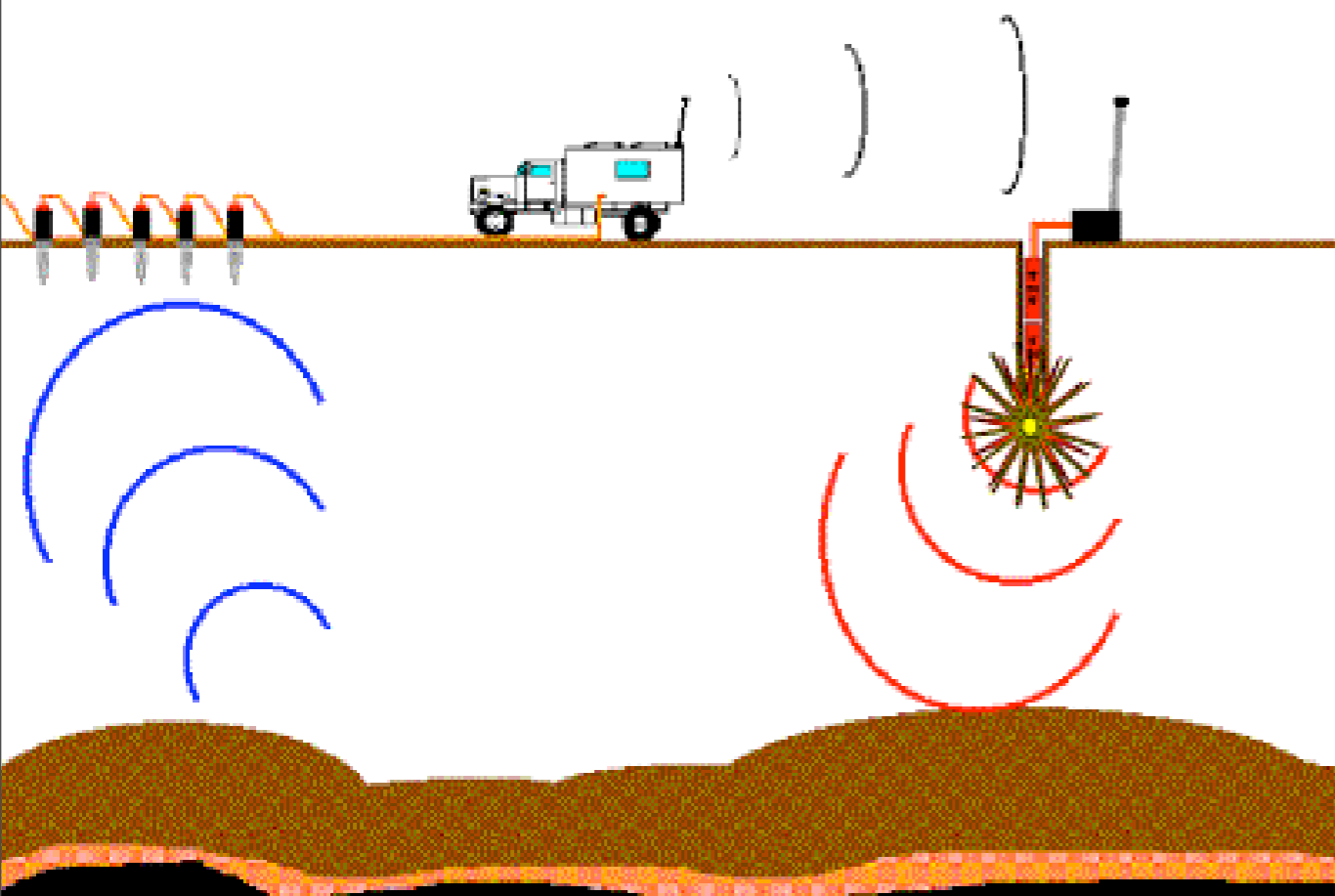
Imaging operators are near unitary (pseudolocal)

Compress the action of the operators ...

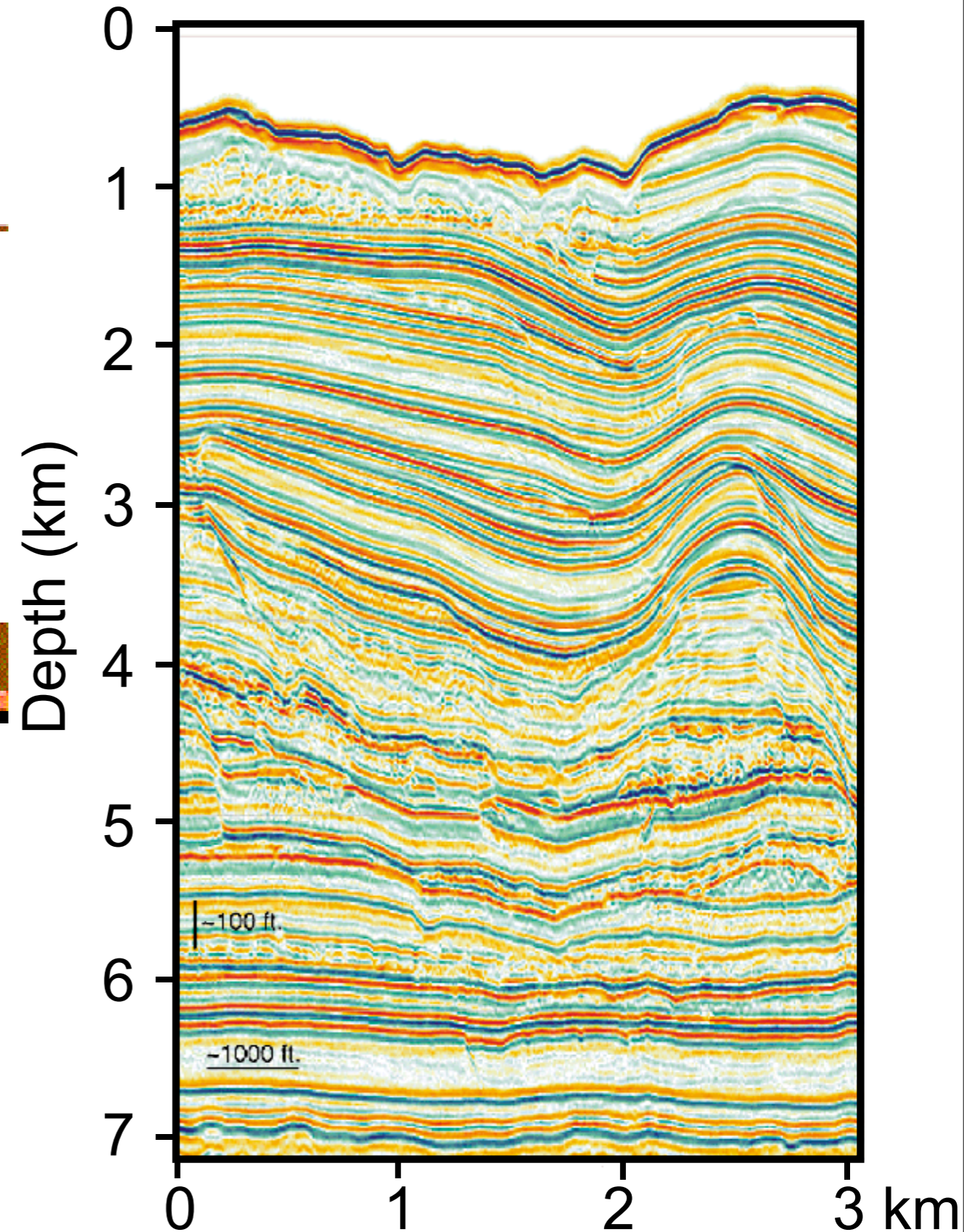
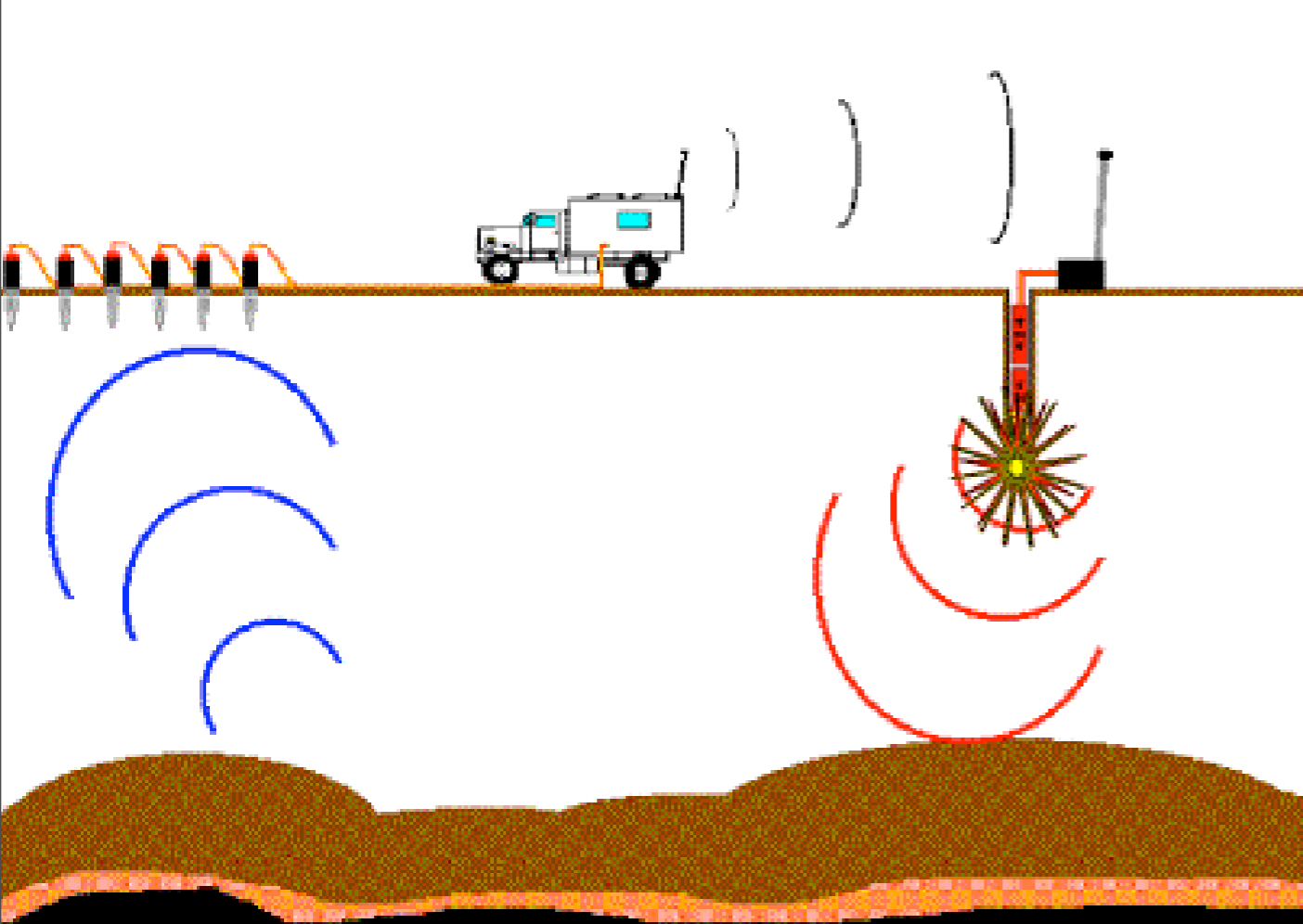
Inspiration:

- Quasi-SVD/Wavelet-Vaguelette
 - sparsity on the model
 - app invariance under the operator \Leftrightarrow diagonalization
- Compressive sampling
 - sparsity on the model
 - incoherence measurement "basis" and sparsity frame
 - diagonalization of the operator by the measurement basis

Seismic data acquisition



Exploration seismology



- **create images of the subsurface**
- **need for higher resolution/deeper**
- **clutter and data incompleteness are problems**

Forward problem

$$F[c]u := \left(\frac{1}{c^2(x)} \cdot \frac{\partial^2}{\partial t^2} - \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2} \right) \mathbf{u}(x, t) = f(x, t)$$

- second order hyperbolic PDE
- interested in the singularities of

$$m = c - \bar{c}$$

Inverse problem

Minimization:

$$\tilde{m} = \arg \min_m \|d - F[m]\|_2^2$$

After linearization (Born app.) forward model with noise:

$$d(x_s, x_r, t) = (K[\bar{c}]m)(x_s, x_r, t) + n(x_s, x_r, t)$$

Conventional imaging:

$$(K^T d)(x) = (K^T K m)(x) + (K^T n)(x)$$

$$y(x) = (\Psi m)(x) + e(x)$$

Ψ is prohibitively expensive to invert

evaluation of $K[\bar{c}]$ involves expensive wavefield extrapolators

Approximate inversion Gramm matrix by scaling/Quasi-SVD

Joint work with Chris Stolk* and
Peyman Moghaddam



Mathematics Department,
Twente University, the Netherlands

“Sparsity- and continuity-promoting seismic imaging
with curvelet frames” to appear in ACHA

Related work

Wavelet-Vaguelette/Quasi-SVD methods based on

- homogeneous operators
- absorb “square-root” of the Gram matrix in WVD’s
- Wavelets/curvelets near diagonalize the operator and are sparse on the model
 - Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition (Donoho ‘95)
 - Recovering Edges in Ill-posed Problems: Optimality of curvelet Frames (Candes & Donoho ‘00)

Scaling methods based on a diagonal approximation of Ψ , assuming

- smoothness on the symbol and conormality reflectors
 - Illumination-based normalization (Rickett ‘02)
 - Amplitude preserved migration (Plessix & Mulder ‘04)
 - Amplitude corrections (Guitton ‘04)
 - Amplitude scaling (Symes ‘07)

Hessian/Normal operator

[Stolk 2002, ten Kroode 1997, de Hoop 2000, 2003]

Alternative to expensive least-squares migration.

In high-frequency limit Ψ is a pseudo-differential operator

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

- composition of two Fourier integral operators
- pseudolocal (near unitary)
- singularities are preserved
- symbol is smooth for smooth velocity models \bar{c}

Corresponds to a spatially-varying dip filter after appropriate preconditioning (\Rightarrow zero order PsDO).

Approximation

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 Ψ .

Allows for the decomposition

$$\begin{aligned} (\Psi \varphi_\mu)(x) &\simeq (C^T \mathbf{D}_\Psi C \varphi_\mu)(x) \\ &= (A A^T \varphi_\mu)(x) \end{aligned}$$

with $A := \sqrt{\mathbf{D}_\Psi} C$ and $A^T := C^T \sqrt{\mathbf{D}_\Psi}$.

Solution

Solve

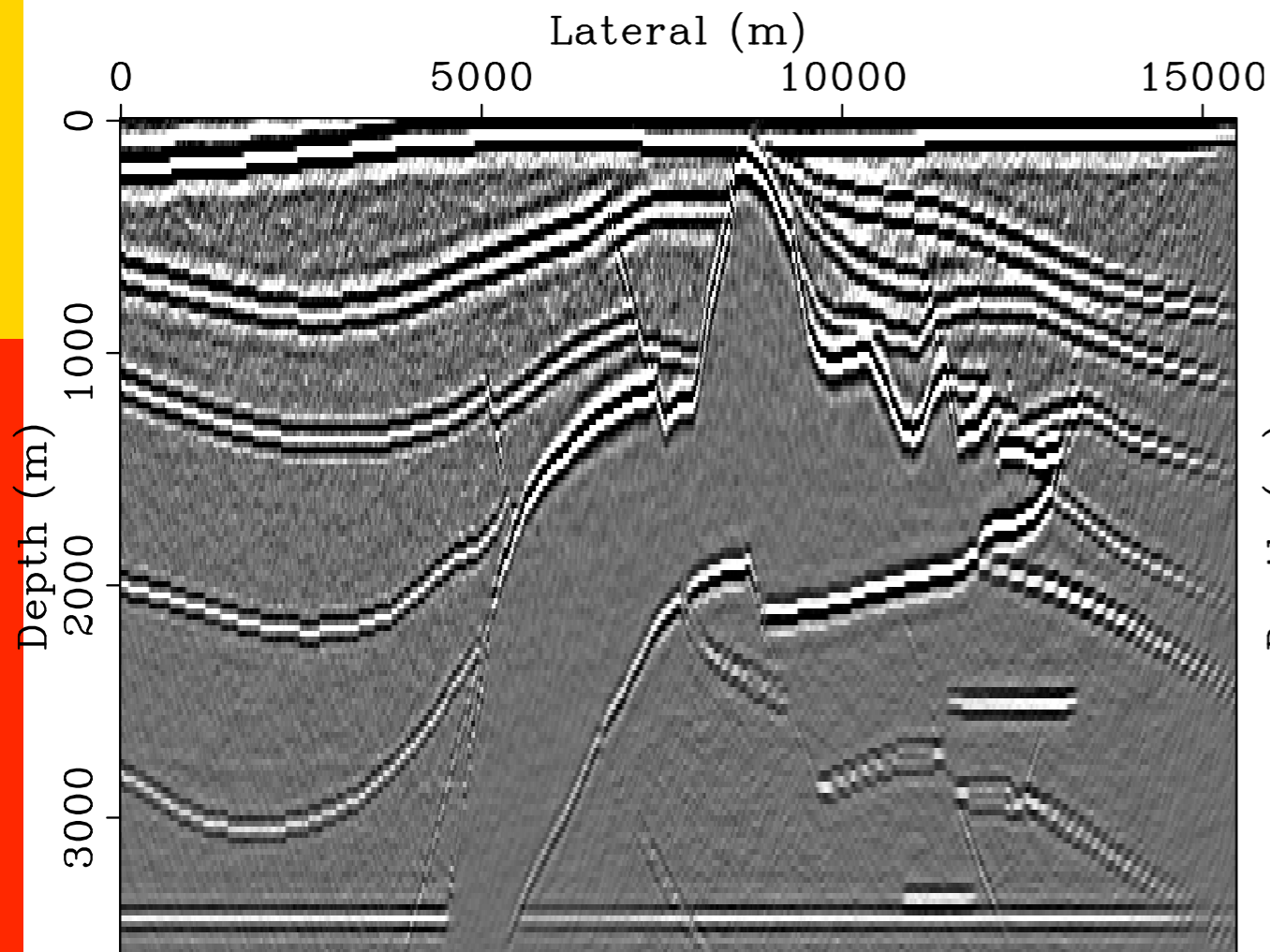
$$\mathbf{P} : \begin{cases} \min_{\mathbf{x}} J(\mathbf{x}) & \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon \\ \tilde{\mathbf{m}} = (\mathbf{A}^H)^\dagger \tilde{\mathbf{x}} \end{cases}$$

with

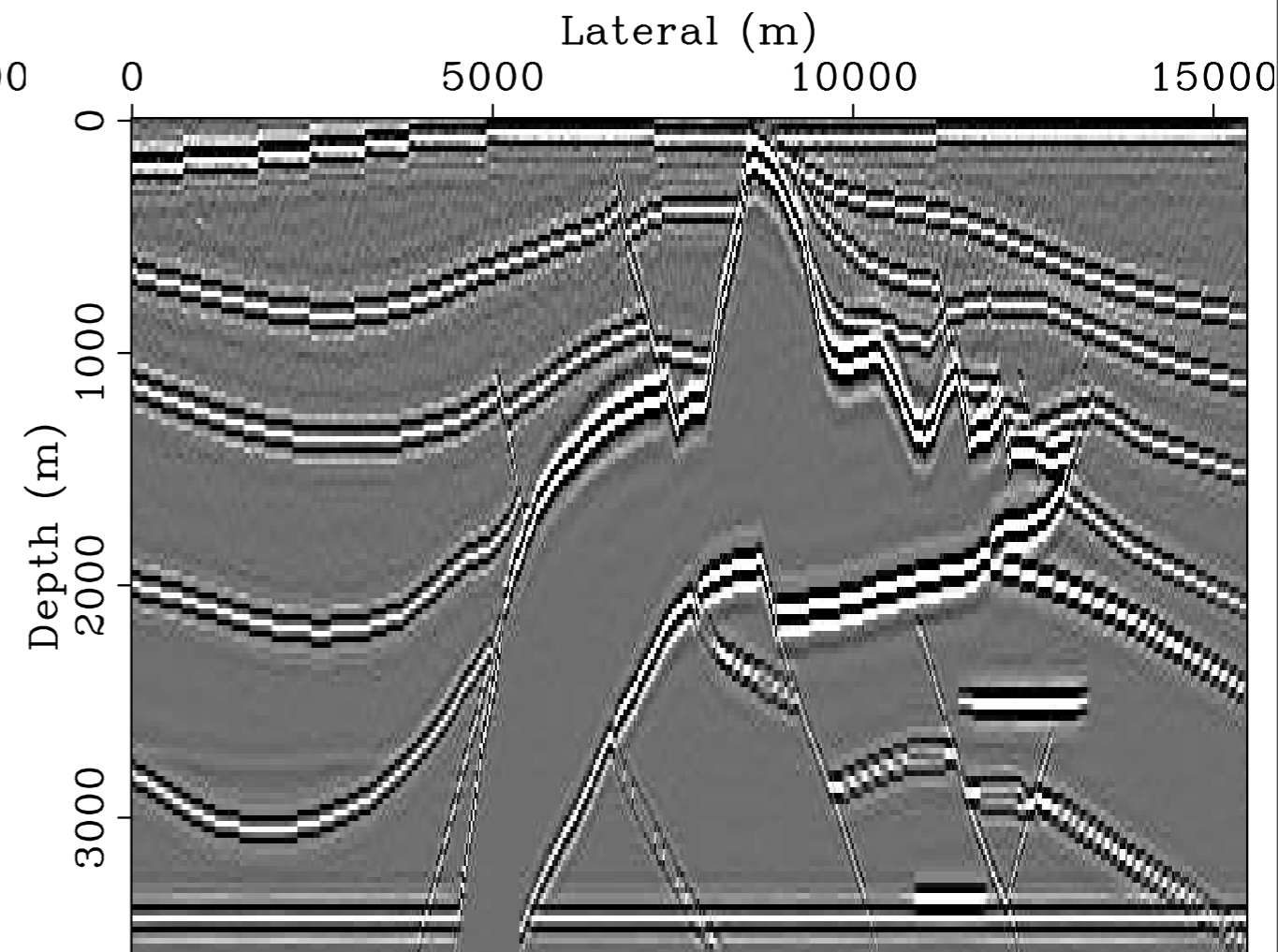
$$J(\mathbf{x}) = \overbrace{\alpha \|\mathbf{x}\|_1}^{\text{sparsity}} + \beta \underbrace{\left\| \mathbf{\Lambda}^{1/2} \left(\mathbf{A}^H \right)^\dagger \mathbf{x} \right\|_p}_{\text{continuity}}.$$

- uses curvelet sparsity on the model
- employs curvelet invariance under the Gram operator
- removes the curvelet frame ambiguity
- removes artifacts by anisotropic diffusion
- does **not** really incorporate ideas from compressive

Imaging example



Migrated data

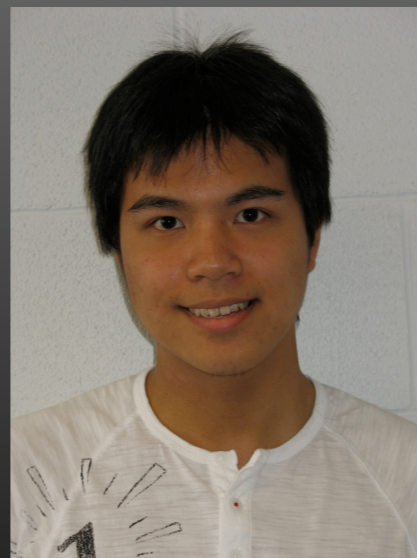


Amplitude-corrected & denoised migrated data

- two-way reverse time wave-equation migration with checkpointing [Symes '07]
- adjoint state method with 8000 time steps
- evaluation \mathbf{K}^T takes 6 h on 60 CPU's

Compressed wavefield extrapolation

joint work with Tim Lin



“Compressed wavefield extrapolation” to appear in Geophysics

Motivation

Synthesis of the discretized operators form bottle neck of imaging

Operators have to be applied to multiple right-hand sides

Explicit operators are feasible in 2-D and lead to an order-of-magnitude performance increase

Extension towards 3-D problematic

- storage of the explicit operators
- convergence of implicit time-harmonic approaches

First go at the problem using CS techniques to compress the operator ...

Related work

Curvelet-domain diagonalization of FIO's

- The Curvelet Representation of Wave Propagators is Optimally Sparse (Candes & Demanet '05)
- Seismic imaging in the curvelet domain and its implications for the curvelet design (Chauris '06)
- Leading-order seismic imaging using curvelets (Douma & de Hoop '06)

Explicit time harmonic methods

- Modal expansion of one-way operators in laterally varying media (Grimbergen et. al. '98)
- A new iterative solver for the time-harmonic wave equation (Riyanti '06)

Fourier restriction

- How to choose a subset of frequencies in frequency-domain finite-difference migration (Mulder & Plessix '04)

Inspiration

Suppose we want to shift a sparse spike train, i.e.,

$$\begin{aligned}\mathbf{u} &= \mathbf{T}_\tau \mathbf{v} \\ &= e^{-\tau} \mathbf{D} \mathbf{v} \\ &= \mathbf{L} e^{-j\tau \boldsymbol{\Omega}} \mathbf{L}^H \mathbf{v}\end{aligned}$$

where

$$\mathbf{D} = \mathbf{L} \boldsymbol{\Omega} \mathbf{L}^H$$

$$\mathbf{L} = \text{The Fourier Transform}$$

- Eigen modes \Leftrightarrow Fourier transform.
- Can this operation be compressed by compressive sampling?

Operators on spikes

[Candes et. al, Donoho]

Calculate instead

$$\begin{cases} \mathbf{y}' &= \mathbf{R} e^{j\Omega\tau} \mathcal{F} \mathbf{v} \\ \mathbf{A} &= \mathbf{R} \mathcal{F} \\ \tilde{\mathbf{u}} &= \arg \min_{\mathbf{u}} \|\mathbf{u}\|_1 \quad \text{s.t.} \quad \mathbf{A} \mathbf{u} = \mathbf{y}' \end{cases}$$

- Take compressed measurements in Fourier space.
- Recover with sparsity promotion
- Shift operator is compressed by the restriction

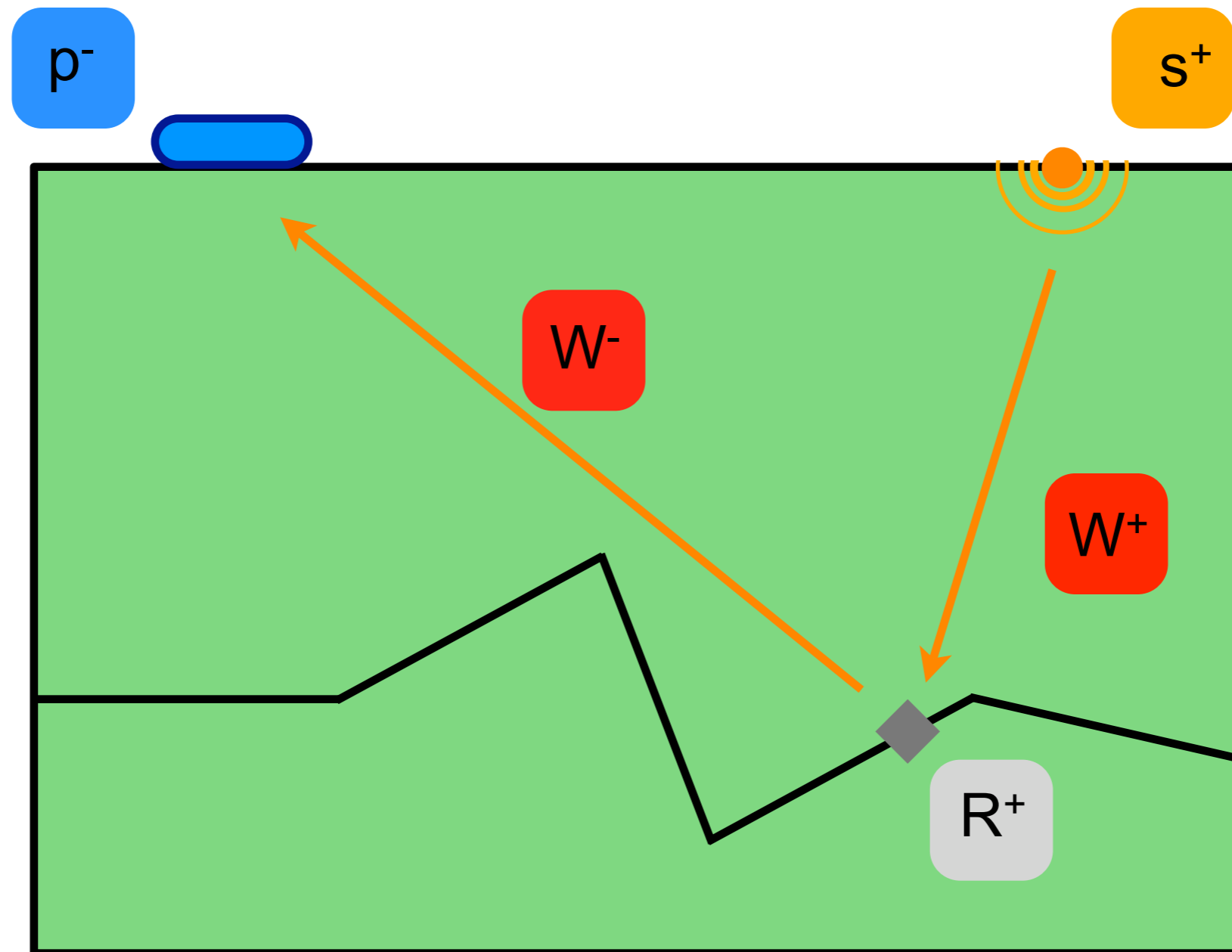
$$\mathbf{R} \in \mathbb{R}^{m \times N} \quad \text{with } m \ll N$$

yielding compressed rectangular operators.

- **Extend this idea to wavefield extrapolation?**

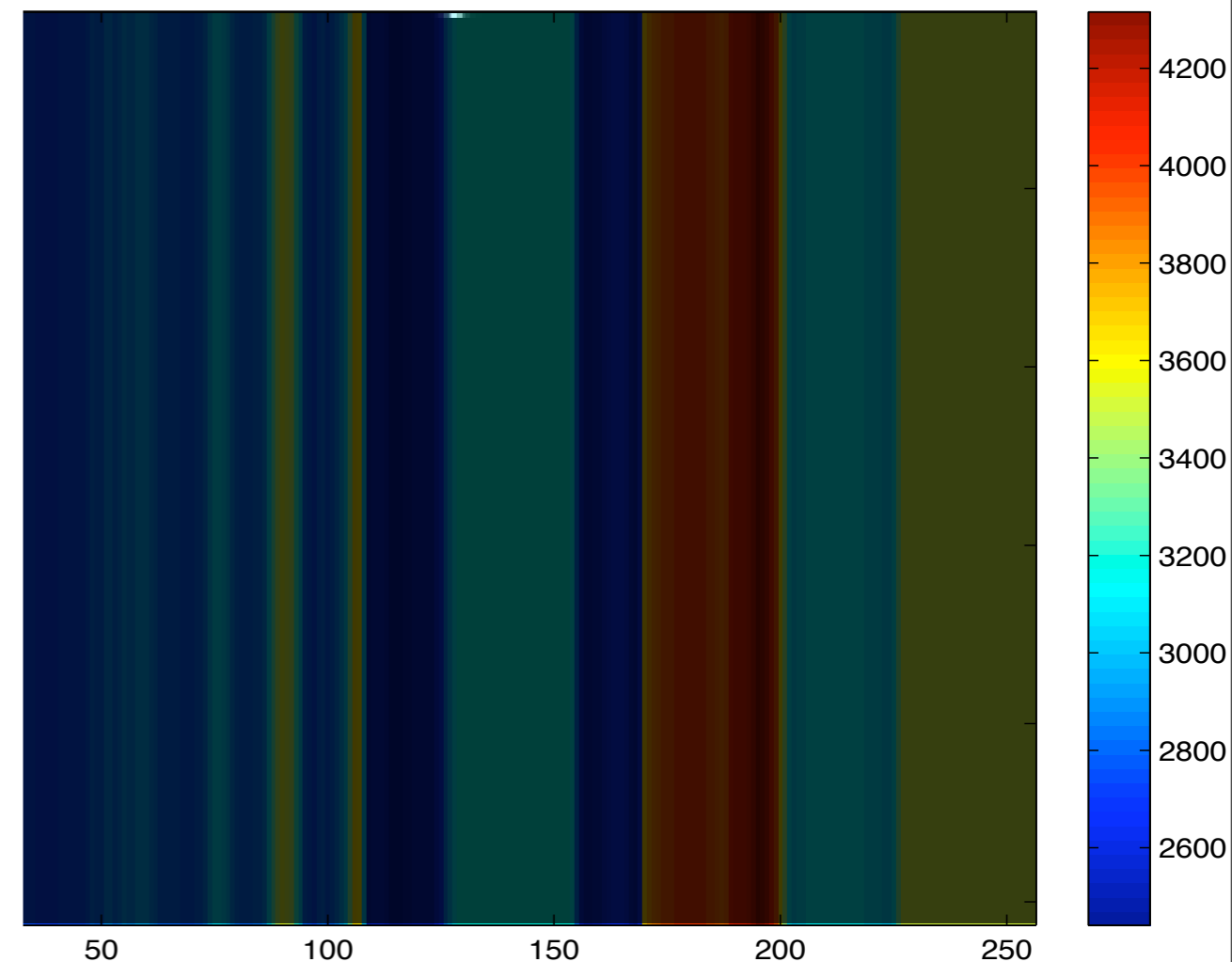
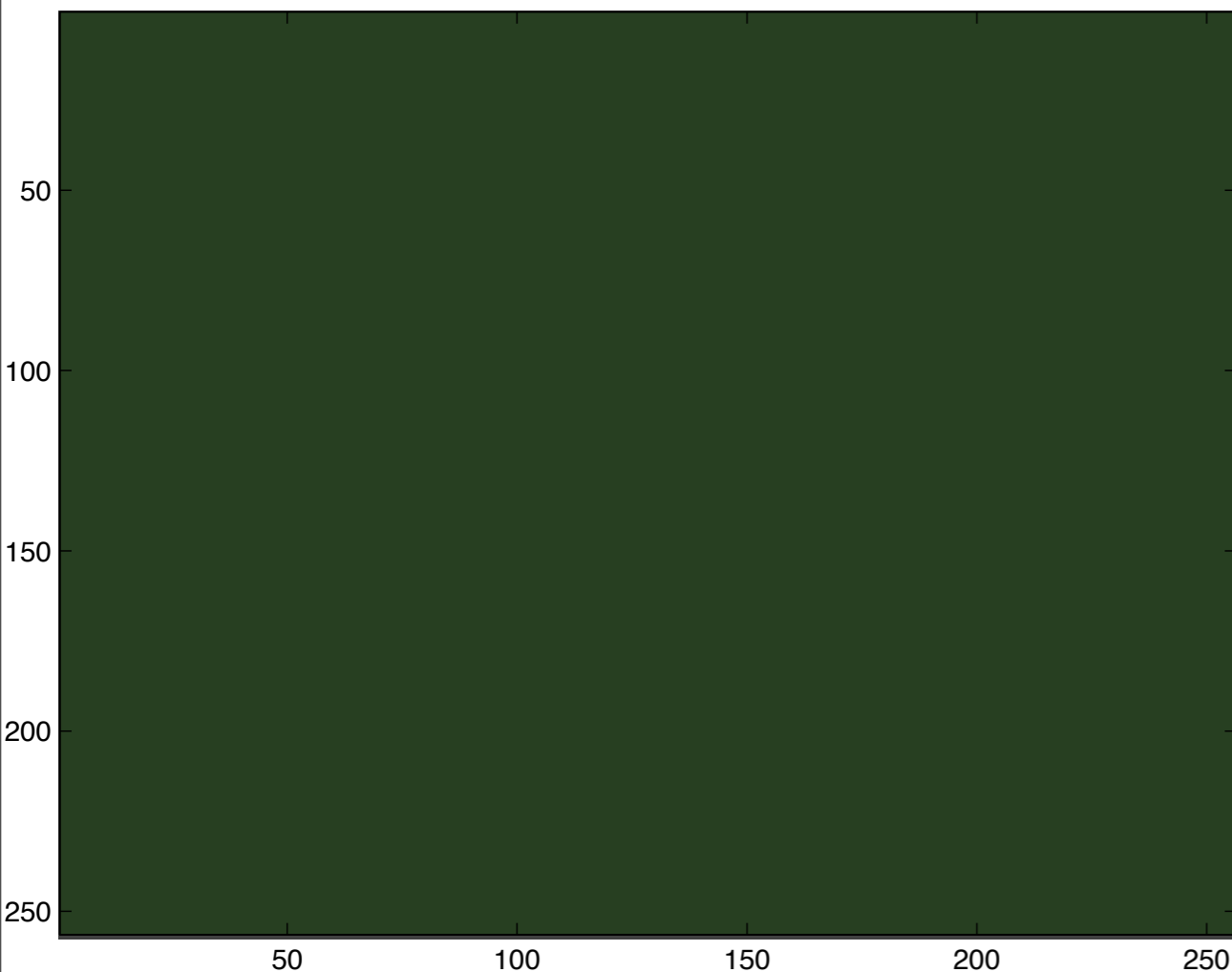
Representation for seismic data

[Berkhout]



One-way forward & inverse wavefield extrapolation

[Claerbout, 1971: Wapenaar and Berkhout, 1989]



Different representations

	diagonalization operator	parsimony wavefield
SVD/Lanczos/modal	✓	✗
curvelets	✗	✓

Different representations

	diagonalization operator	parsimony wavefield
SVD/Lanczos/modal	✓	×
curvelets	×	✓

If incoherent this may actually work

One-Way Wave Operator

- Solution of the one-way wave equation

$$\mathcal{W}(x_3; x'_3) = \exp(-j(x_3 - x'_3)\mathcal{H}_1)$$

- After discretization solve eigenproblem on \mathbf{H}_2

$$\mathbf{H}_2 = \begin{bmatrix} \left(\frac{\omega}{\bar{c}_1}\right)^2 & 0 & \dots & 0 \\ 0 & \left(\frac{\omega}{\bar{c}_2}\right)^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \left(\frac{\omega}{\bar{c}_{n_1}}\right)^2 \end{bmatrix} + \mathbf{D}_2$$

- Helmholtz operator is Hermitian
- monochromatic
- velocity \bar{c} varies laterally

(Claerbout, 1971; Wapenaar and Berkhout, 1989)

Modal transform

- Solve eigenproblem & take square root

$$\mathbf{H}_1 = \mathbf{L}\mathbf{\Lambda}^{1/2}\mathbf{L}^H$$

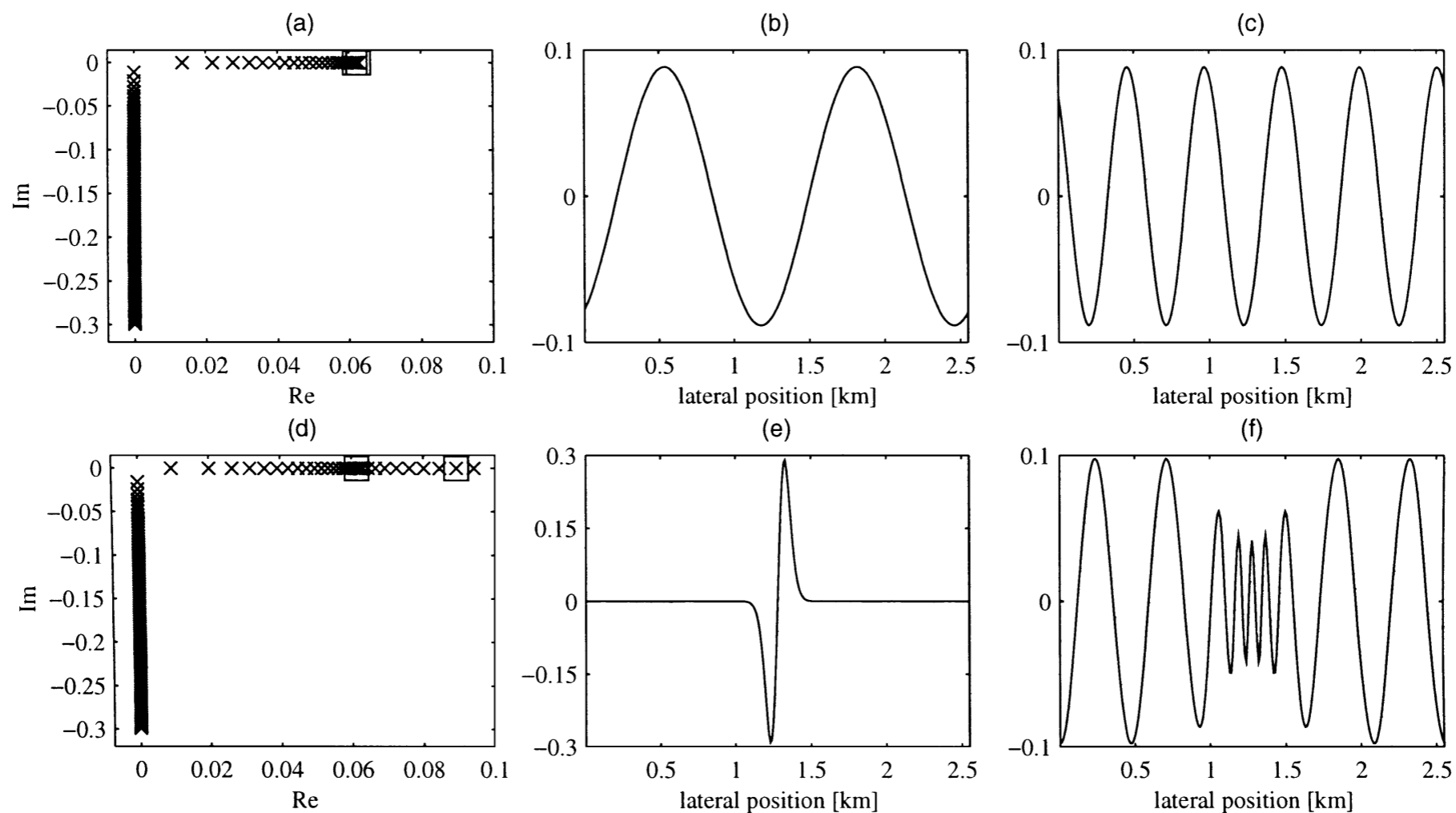
- \mathbf{L} is orthonormal & defines the modal transform that diagonalizes one-way wavefield extrapolation
- Eigenvalues play role of vertical wavenumbers
- Extrapolation operator is diagonalized

$$\mathbf{W} = \mathcal{F}^H \mathbf{L} e^{-j\mathbf{\Lambda}^{1/2}(x_3 - x'_3)} \mathbf{L}^H \mathcal{F}$$

Eigenfunctions

□ Radiating and guided modes

(from Grimbergen '98)

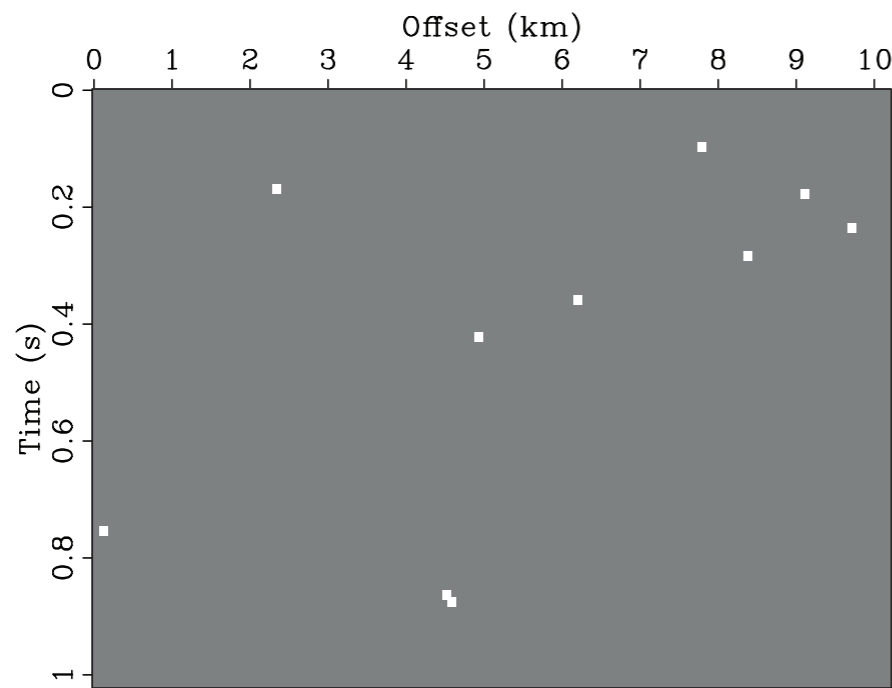


- Eigenmodes form a “complete” & orthonormal basis
- Evanescent eigenmodes decay exponentially

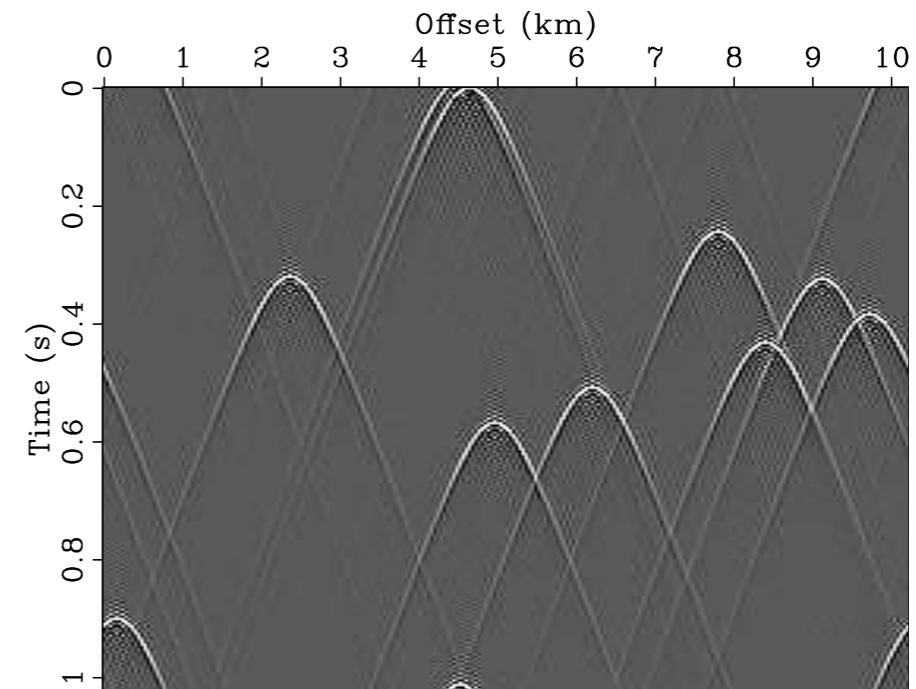
Compressed wavefield extrapolation

Forward model

$$\mathbf{u} = \mathbf{L} e^{-j\mathbf{\Lambda}^{1/2} \Delta x_3} \mathbf{L}^H \mathbf{v}$$



Original events



Recorded Data

Reconstruct point scatterers from recorded data

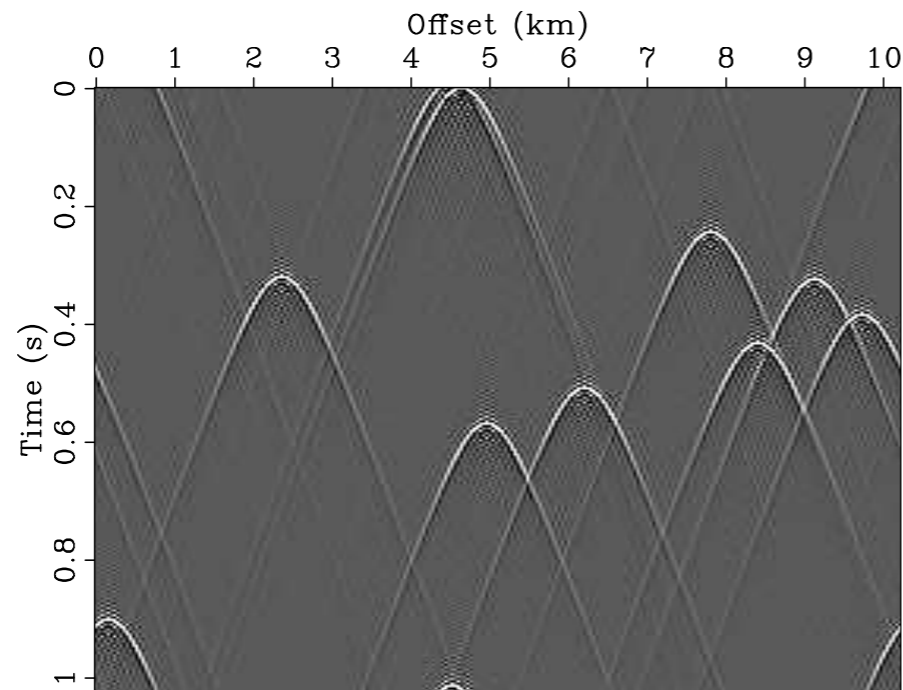
Compressed wavefield extrapolation

$$\begin{cases} \mathbf{y} &= \mathbf{R}\mathbf{L}^H \mathbf{u} \\ \mathbf{A} &= \mathbf{R}e^{j\mathbf{\Lambda}^{1/2} \Delta x_3} \mathbf{L}^H \\ \tilde{\mathbf{x}} &= \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{y} \\ \tilde{\mathbf{v}} &= \tilde{\mathbf{x}} \end{cases}$$

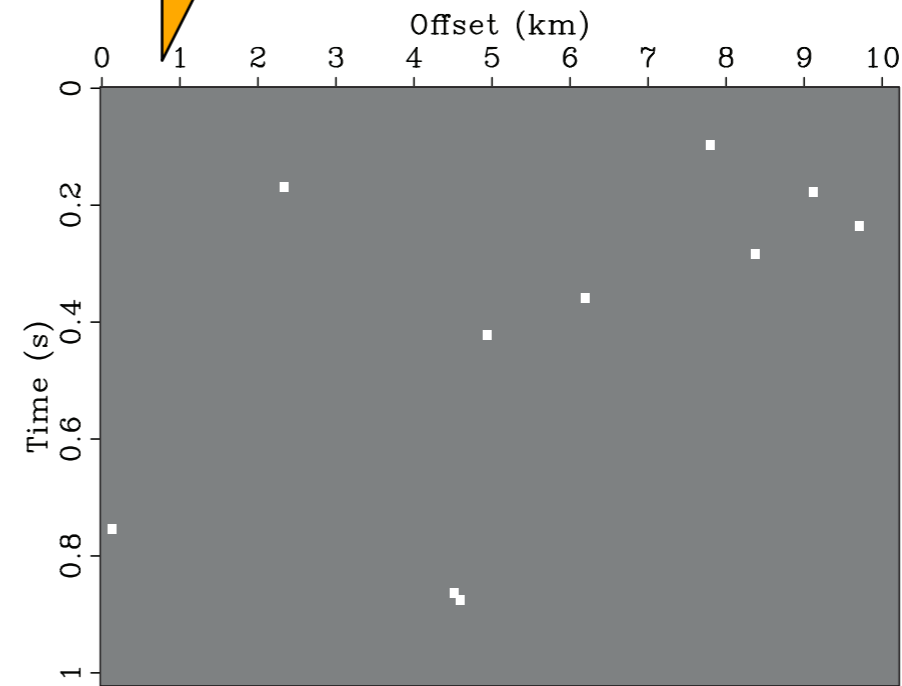
- Randomly subsample & phase rotate in Modal domain
- Recover by norm-one minimization
- Capitalize on
 - the incoherence modal functions and point scatterers
 - reduced explicit matrix size
 - constant velocity \Leftrightarrow Fourier recovery

Compressed wavefield extrapolation

Reconstruction



Recorded Data

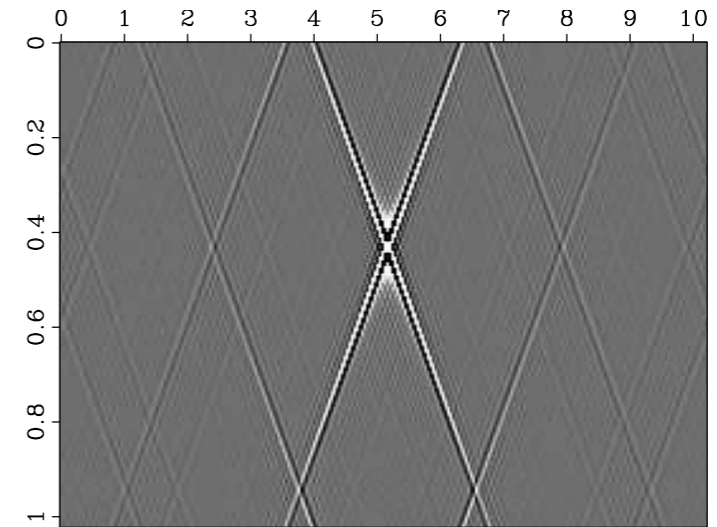


Reconstructed events

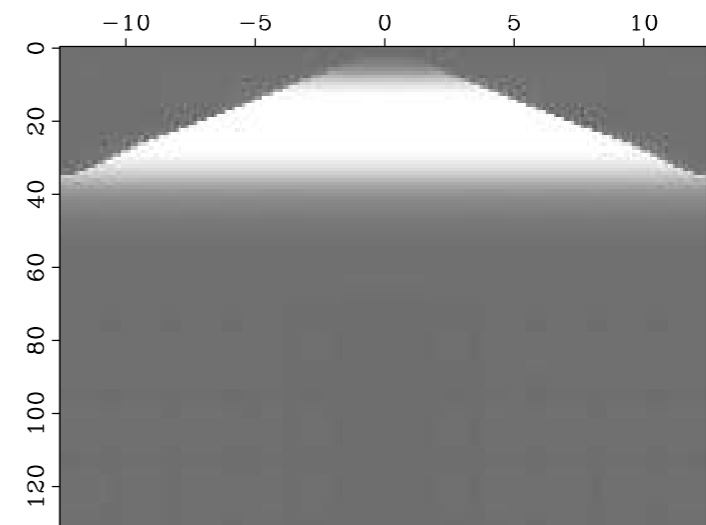
Only 1 % of original modes were used ...

Observations

- Despite the existence of evanescent (exponentially decaying) waves modes recovery is successful
- If you are looking for point-scatterers, we have a proof of concept that is fast
- Earth is more complex ...

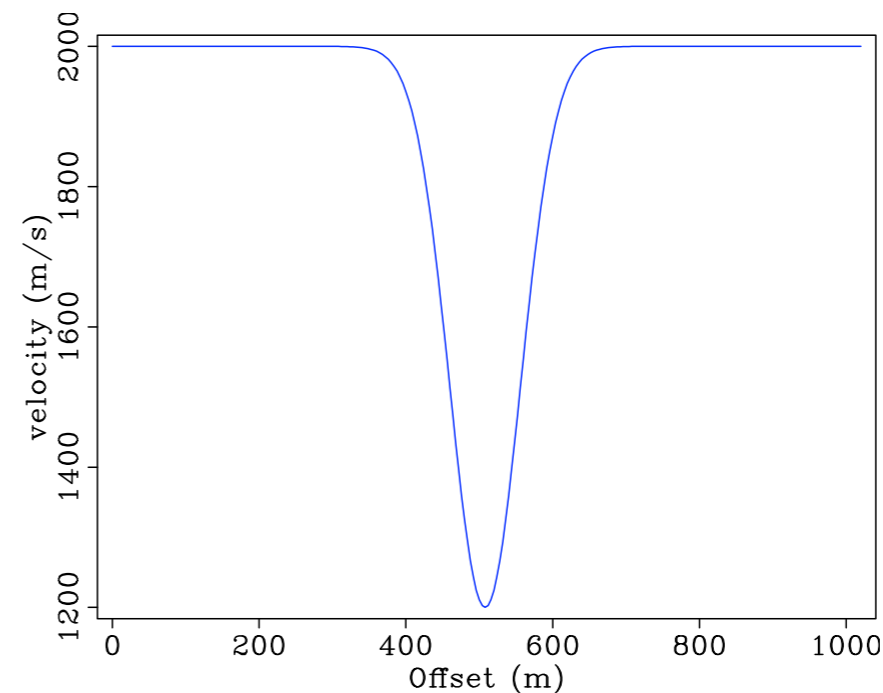
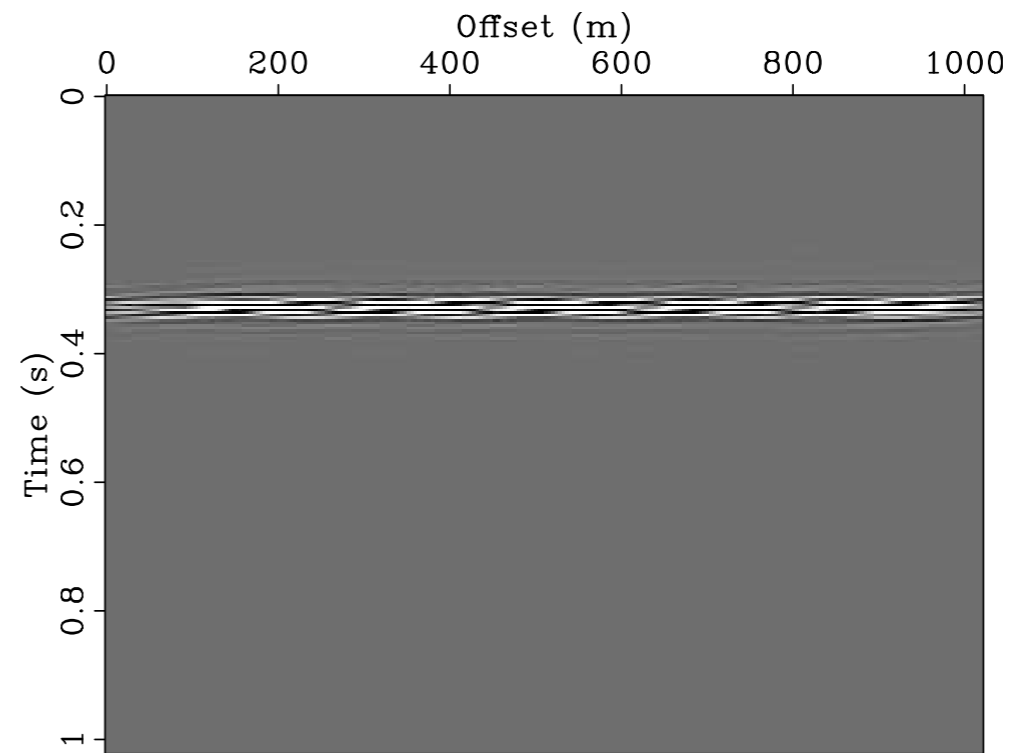


(c)



Compressed wavefield extrapolation

- Extend to general wavefields
- Use curvelets as the sparsity representation
- Use the full & compressed forward operator operator
- Compressively extrapolate back 600m to the source



Restriction & sparsity strategies

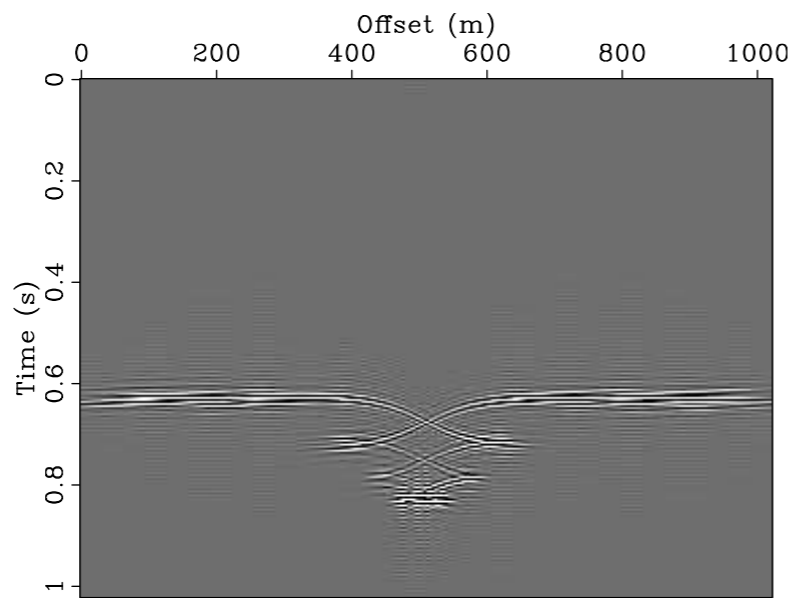
- **Forward extrapolation:**

$$\mathbf{W}_1 : \begin{cases} \mathbf{y}' = \mathbf{R} e^{j\Lambda^{1/2} \Delta x_3} \mathbf{L}^H \\ \mathbf{A} := \mathbf{R} \mathbf{L}^H \mathcal{F} \mathbf{C}^T \\ \tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A} \mathbf{x} = \mathbf{y}' \\ \tilde{\mathbf{u}} = \mathbf{C}^T \tilde{\mathbf{x}}, \end{cases}$$

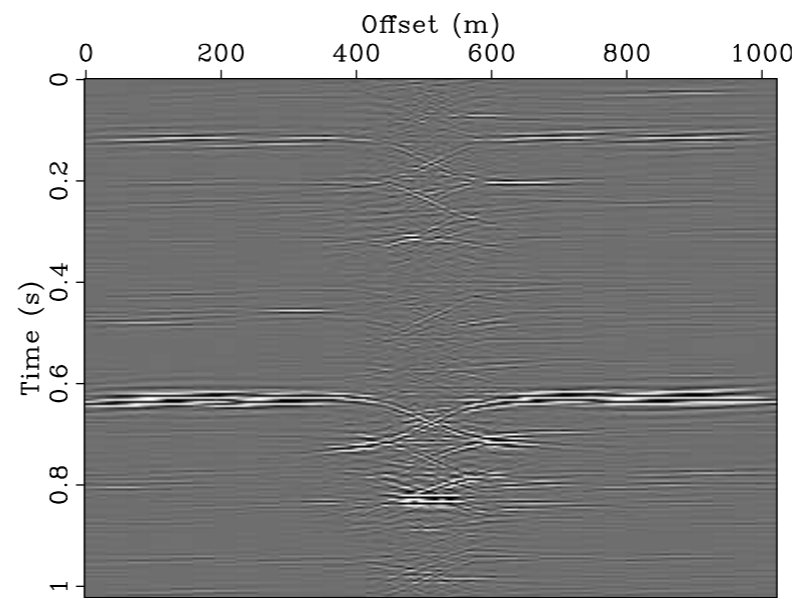
- **Inverse extrapolation:**

$$\mathbf{F}_1 : \begin{cases} \mathbf{y} = \mathbf{R} \mathbf{L}^H \mathcal{F} \mathbf{u} \\ \mathbf{A}' = \mathbf{R} e^{j\Lambda^{1/2} \Delta x_3} \mathbf{L}^H \mathbf{C}^H \\ \tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{A}' \mathbf{x} = \mathbf{y} \\ \tilde{\mathbf{v}} = \mathbf{C}^T \tilde{\mathbf{x}}. \end{cases}$$

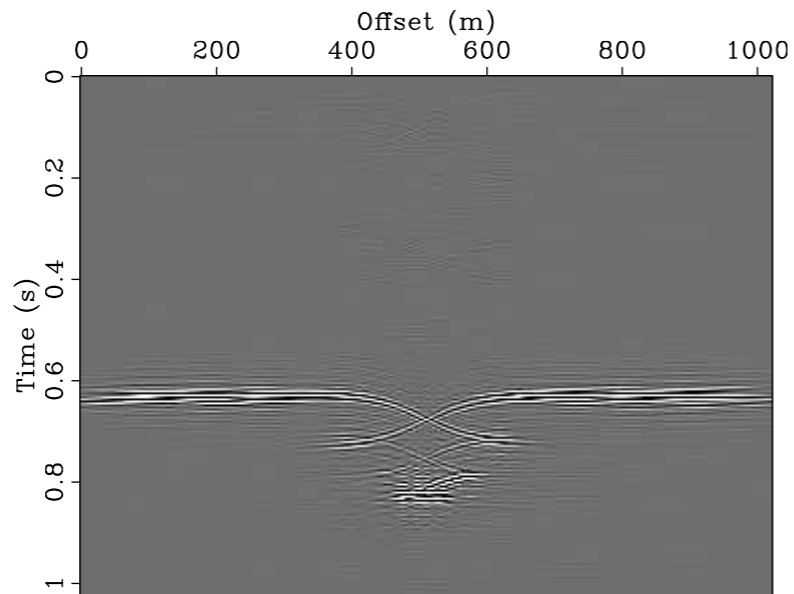
Forward Extrapolation



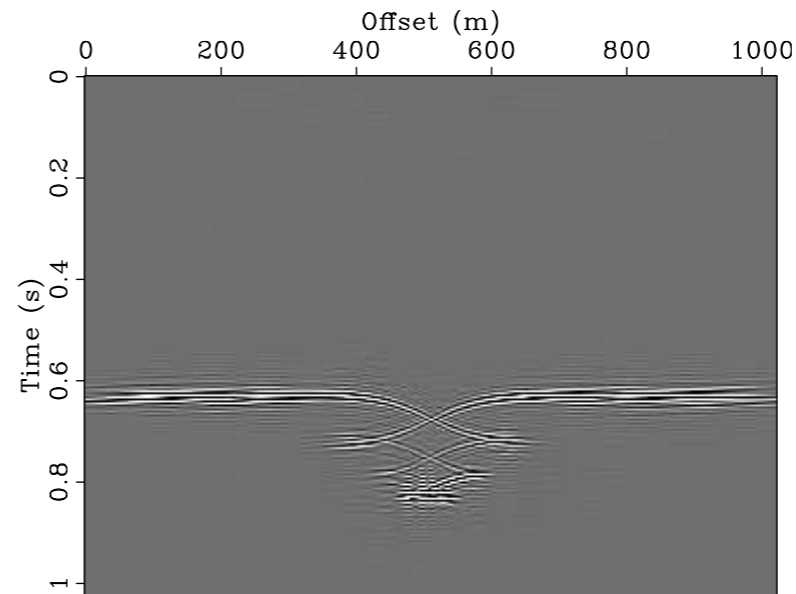
(a)



(b)



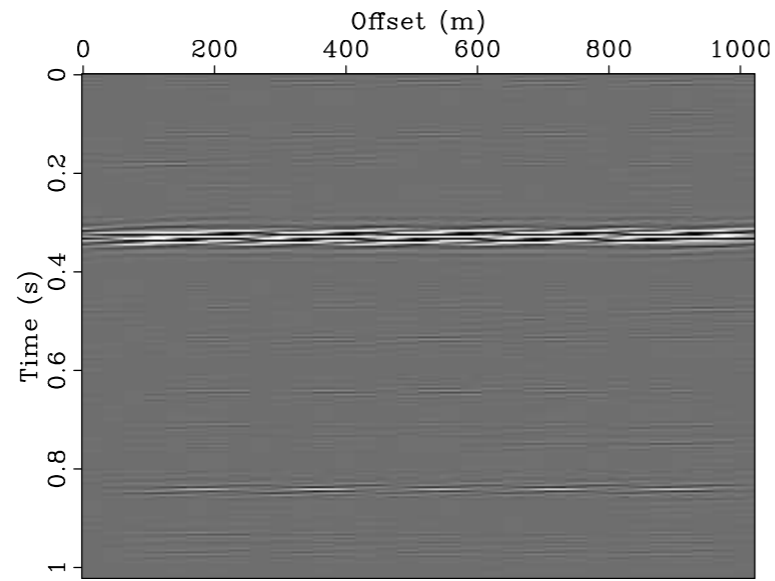
(c)



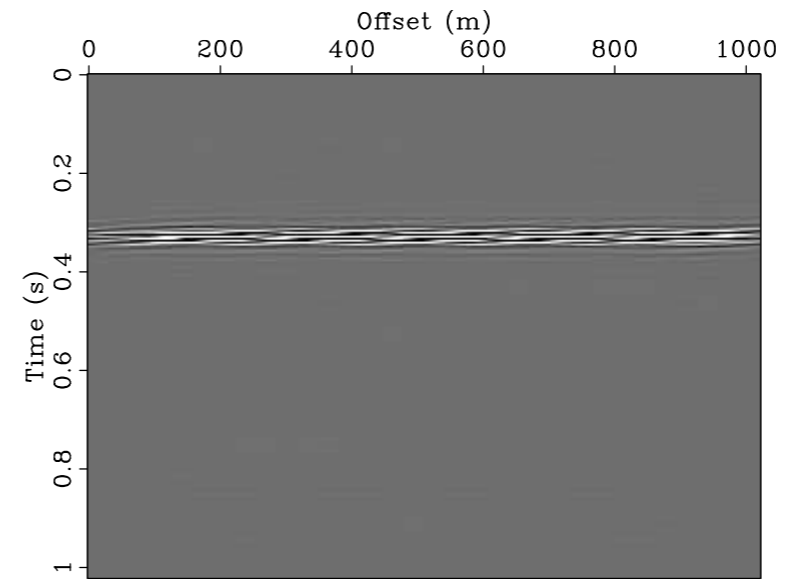
(d)

- (a) is Full extrapolation
- (b)-(d) is compressed extrapolation, (b) $p = 0.04$, (c) $p = 0.16$, (d) $p = 0.24$

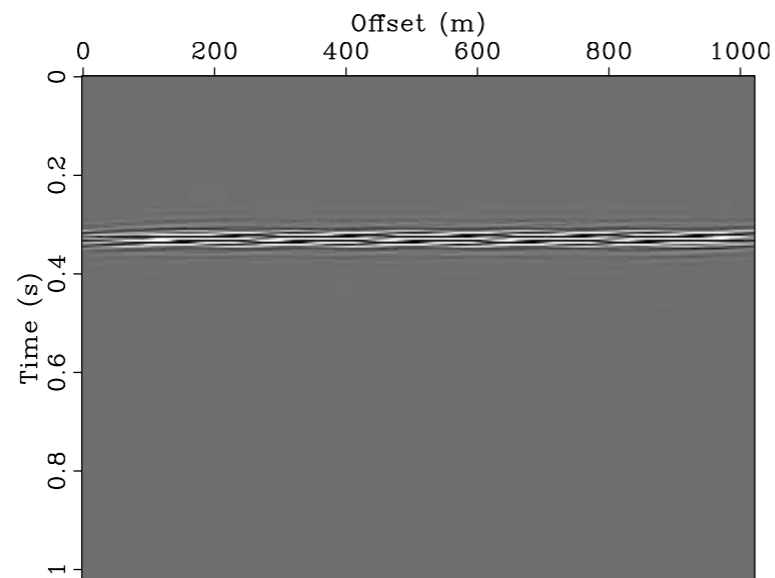
Inverse Extrapolation



(a)



(b)

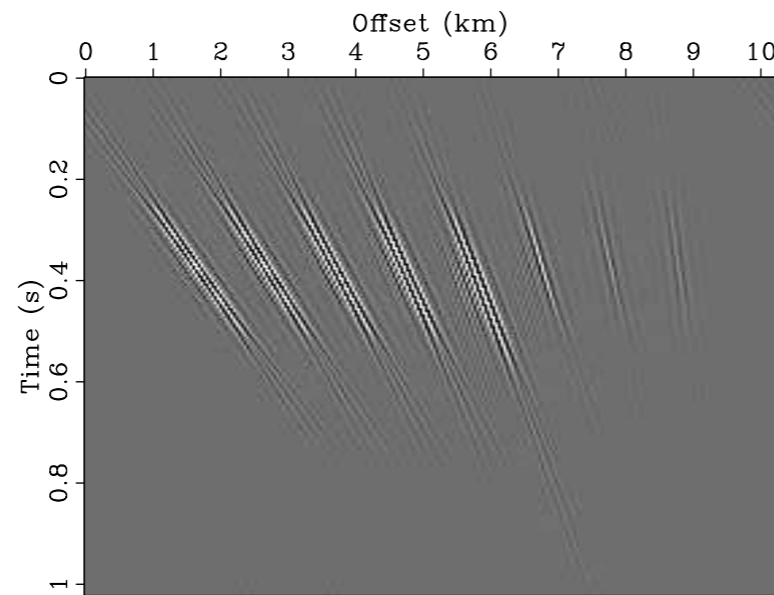


(c)

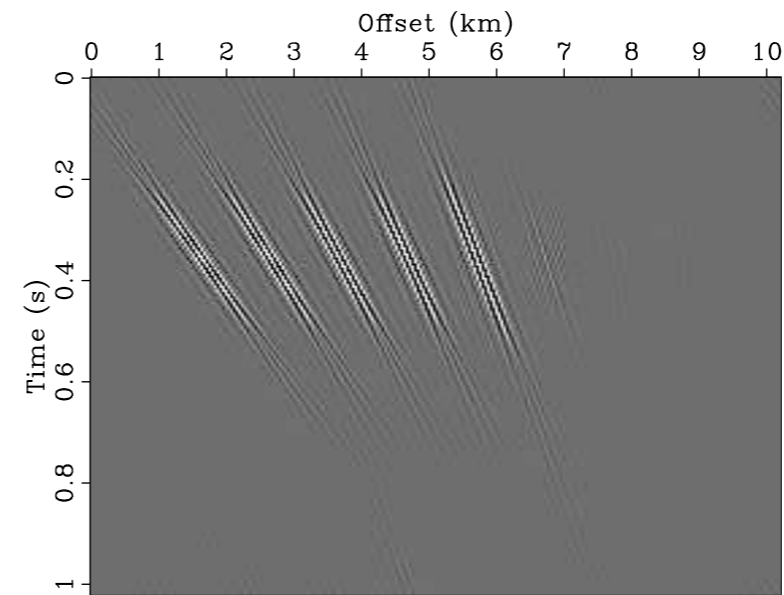
□ (a) $p = 0.04$

□ (b) $p = 0.16$, (c) $pf=0.4$, $px=0.4$

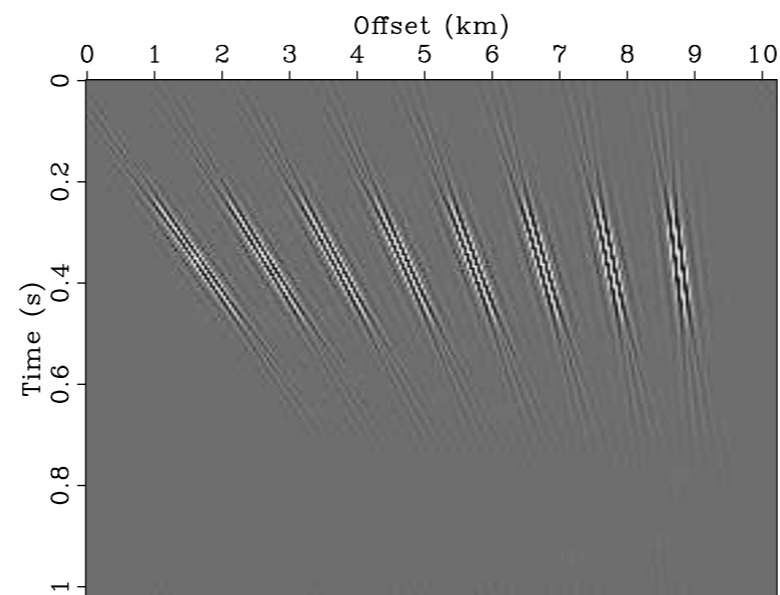
Evanescent Recovery



(a)



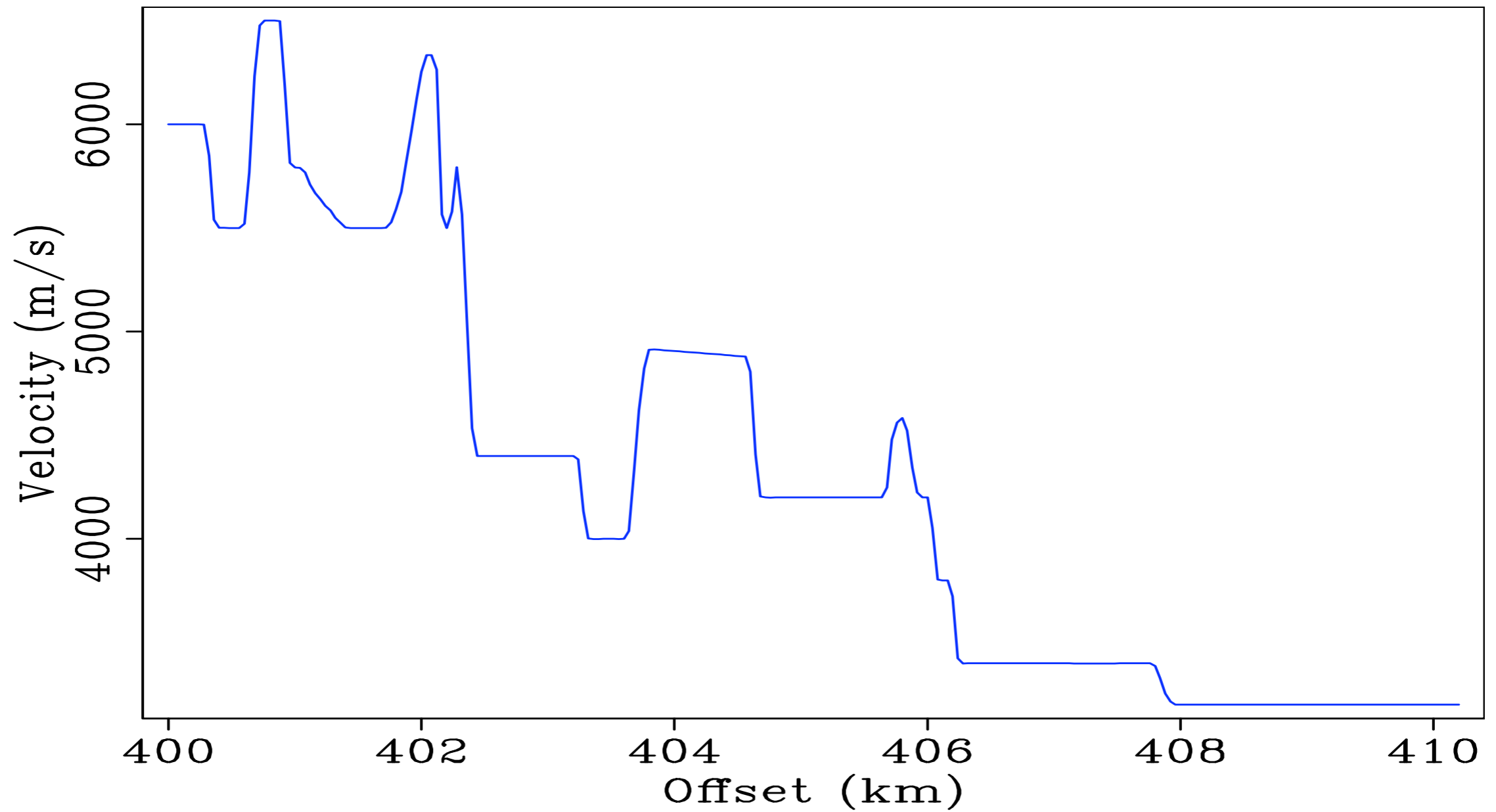
(b)



(c)

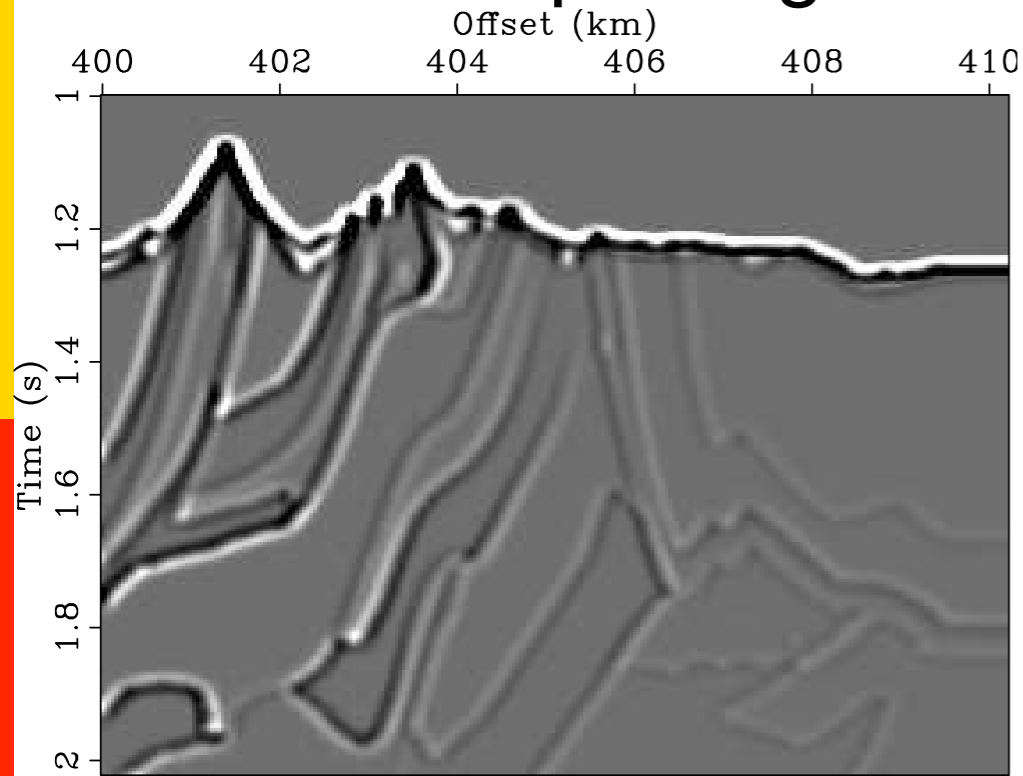
- (a) is downward extrapolated wavefield
- (b) is matched filter
- (c) is "compressed" inverse extrapolation

Velocity model

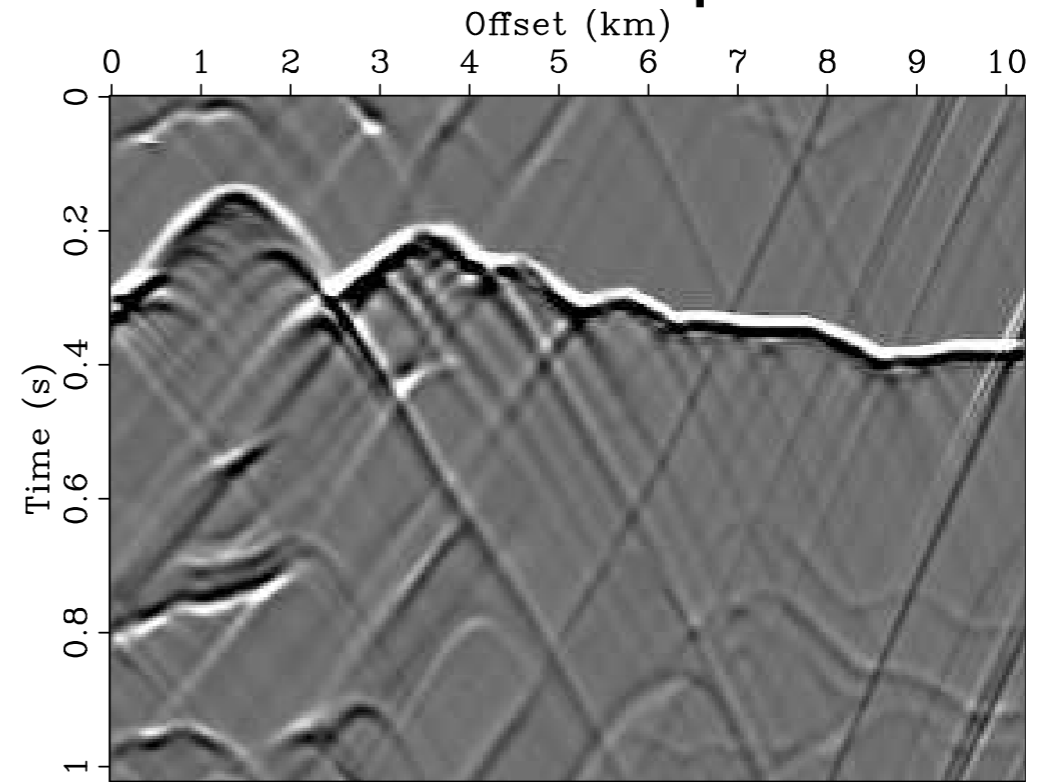


Compressed inverse extrapolation

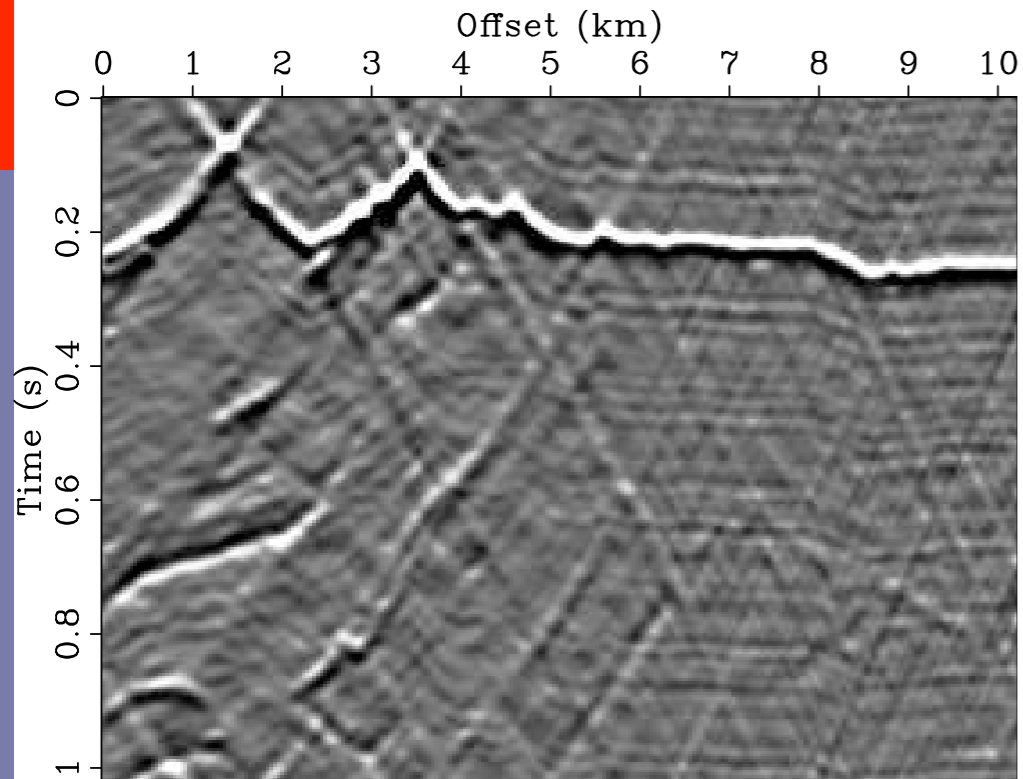
Overthrust exploding reflector



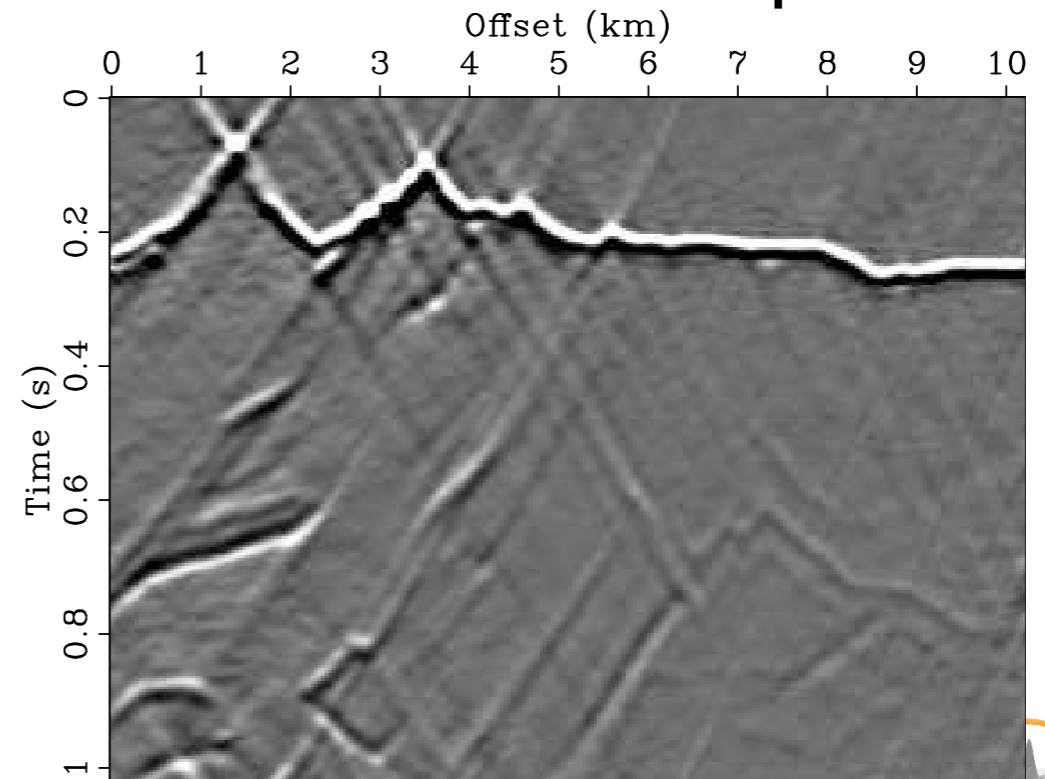
Full forward extrapolation



Matched filter



Recovered from $p=0.25$



Multiscale and angular compressed wavefield extrapolation

- Propose a scheme motivated by extensions of CS

(Tsaig and Donoho '06)

$$\mathbf{F}_1^{\mathbf{j}} : \begin{cases} \mathbf{y}_{\mathbf{j}} = \mathbf{R}_{\mathbf{j}} \mathbf{M}_{\mathbf{j}} \mathbf{u} \\ \mathbf{A}'_{\mathbf{j}} := \mathbf{R}_{\mathbf{j}} \mathbf{M}'_{\mathbf{j}} \mathbf{C}_{\mathbf{j}}^T \\ \tilde{\mathbf{x}}_{\mathbf{j}} = \arg \min_{\mathbf{x}_{\mathbf{j}}} \|\mathbf{x}_{\mathbf{j}}\|_1 \quad \text{s.t.} \quad \mathbf{A}'_{\mathbf{j}} \mathbf{x}_{\mathbf{j}} = \mathbf{y}_{\mathbf{j}} \\ \tilde{\mathbf{v}} = \sum_{\mathbf{j}} \mathbf{C}_{\mathbf{j}}^T \tilde{\mathbf{x}}_{\mathbf{j}}, \end{cases}$$

with $\mathbf{j} = \{j, l\}$ the scale and angle.

- adapt discretization & restriction
- parallel implementation

Compressed focusing with curvelets

joint work with Deli Wang (visitor
from Jilin university) and Gilles
Hennenfent



Related work

Focusing:

Focal transformation, an imaging concept for signal restoration and noise removal (Berkhout & Verschuur '06)

- mapping of multiples => primaries
- incorporation of *prior* information on the Green's function in the recovery

Acquisition restriction in migration:

A quasi-Monte Carlo approach to 3-D migration: Theory (Sun et. al. '97)

Recovery with focussing

Solve

$$\mathbf{P}_\epsilon : \begin{cases} \tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 & \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon \\ \tilde{\mathbf{f}} = \mathbf{S}^T \tilde{\mathbf{x}} \end{cases}$$

with

$$\mathbf{A} := \mathbf{R}\Delta\mathbf{P}\mathbf{C}^T$$

$$\mathbf{S}^T := \Delta\mathbf{P}\mathbf{C}^T$$

$$\mathbf{y} = \mathbf{R}\mathbf{P}(\cdot)$$

$$\mathbf{R} = \text{picking operator.}$$

Recovery with focussing

Solve

$$\mathbf{P}_\epsilon : \begin{cases} \tilde{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 & \text{s.t.} \quad \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon \\ \tilde{\mathbf{f}} = \mathbf{S}^T \tilde{\mathbf{x}} \end{cases}$$

with

$$\mathbf{A} := \mathbf{R}\Delta\mathbf{P}\mathbf{C}^T$$

$$\mathbf{S}^T := \Delta\mathbf{P}\mathbf{C}^T$$

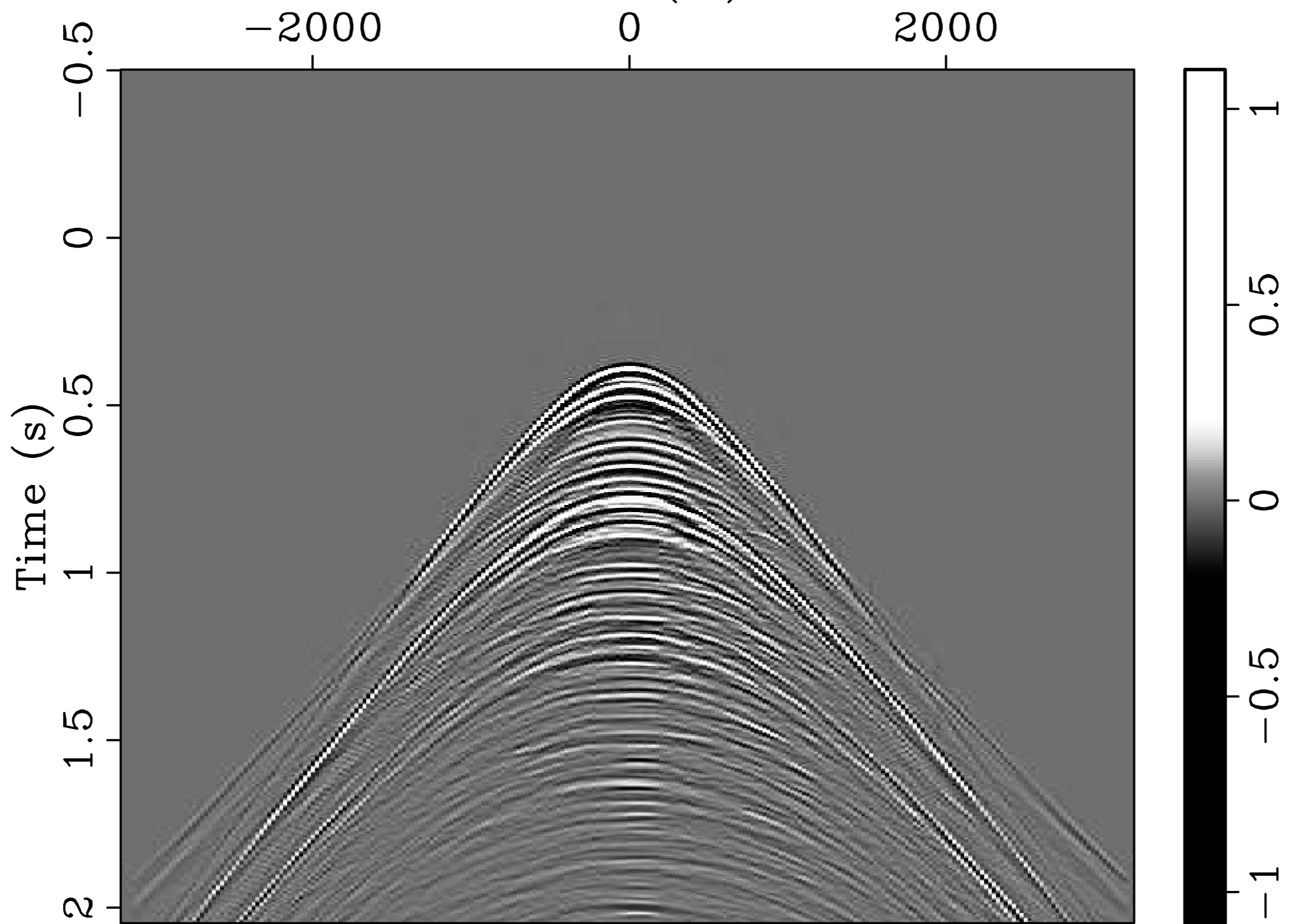
$$\mathbf{y} = \mathbf{R}\mathbf{P}(\cdot)$$

$$\mathbf{R} = \text{picking operator.}$$

compression
of the operator

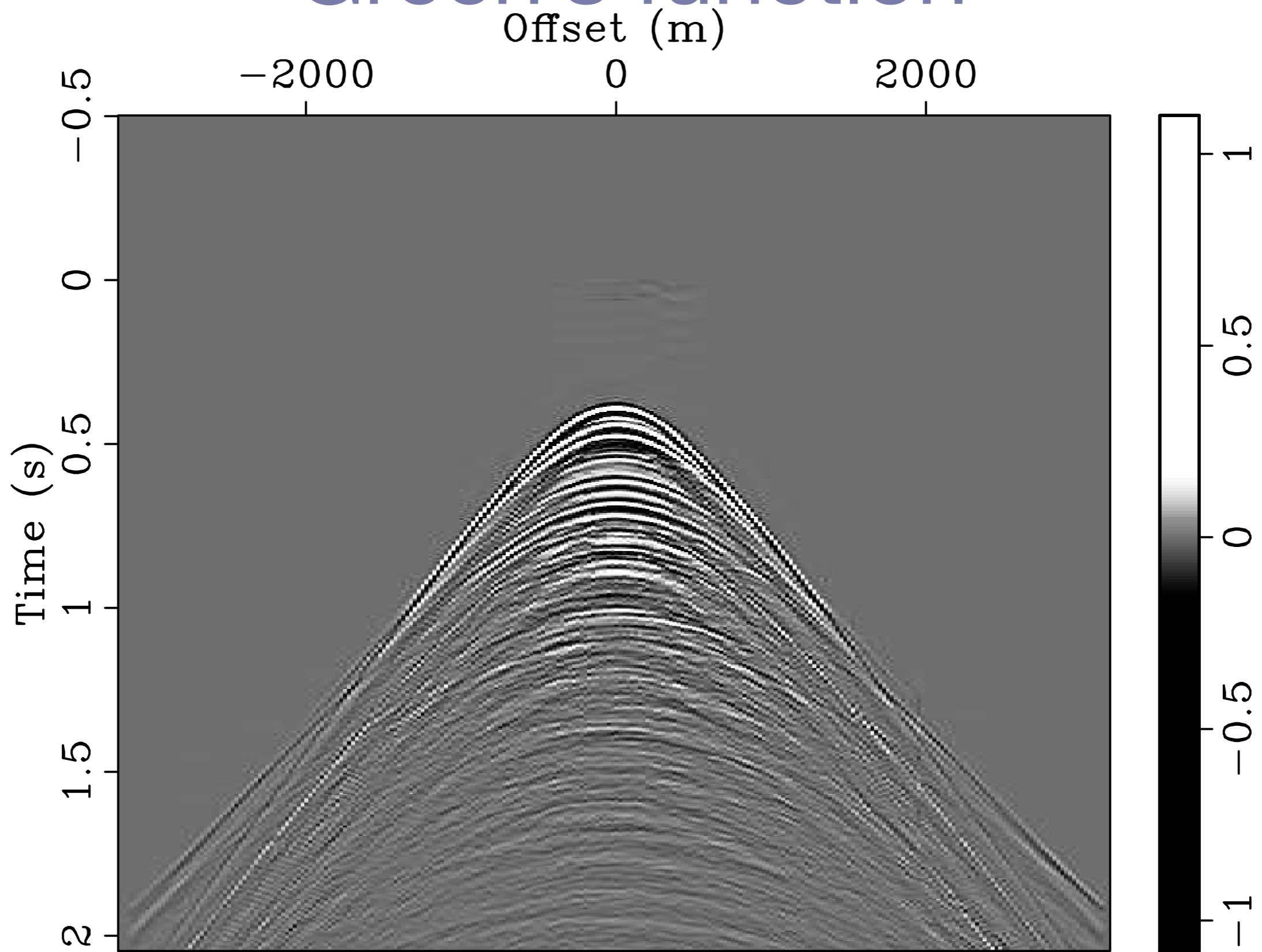
Shot

Offset (m)

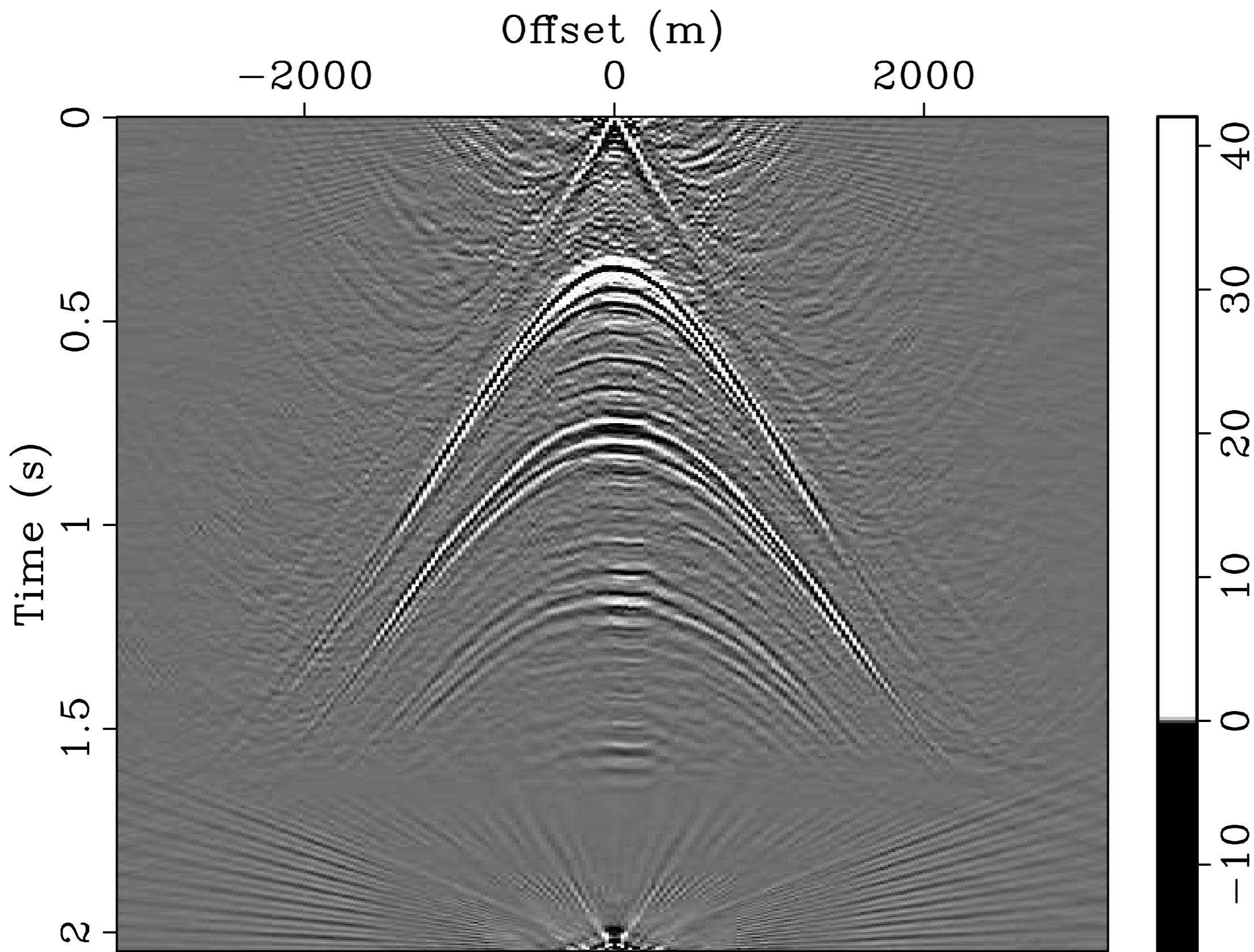


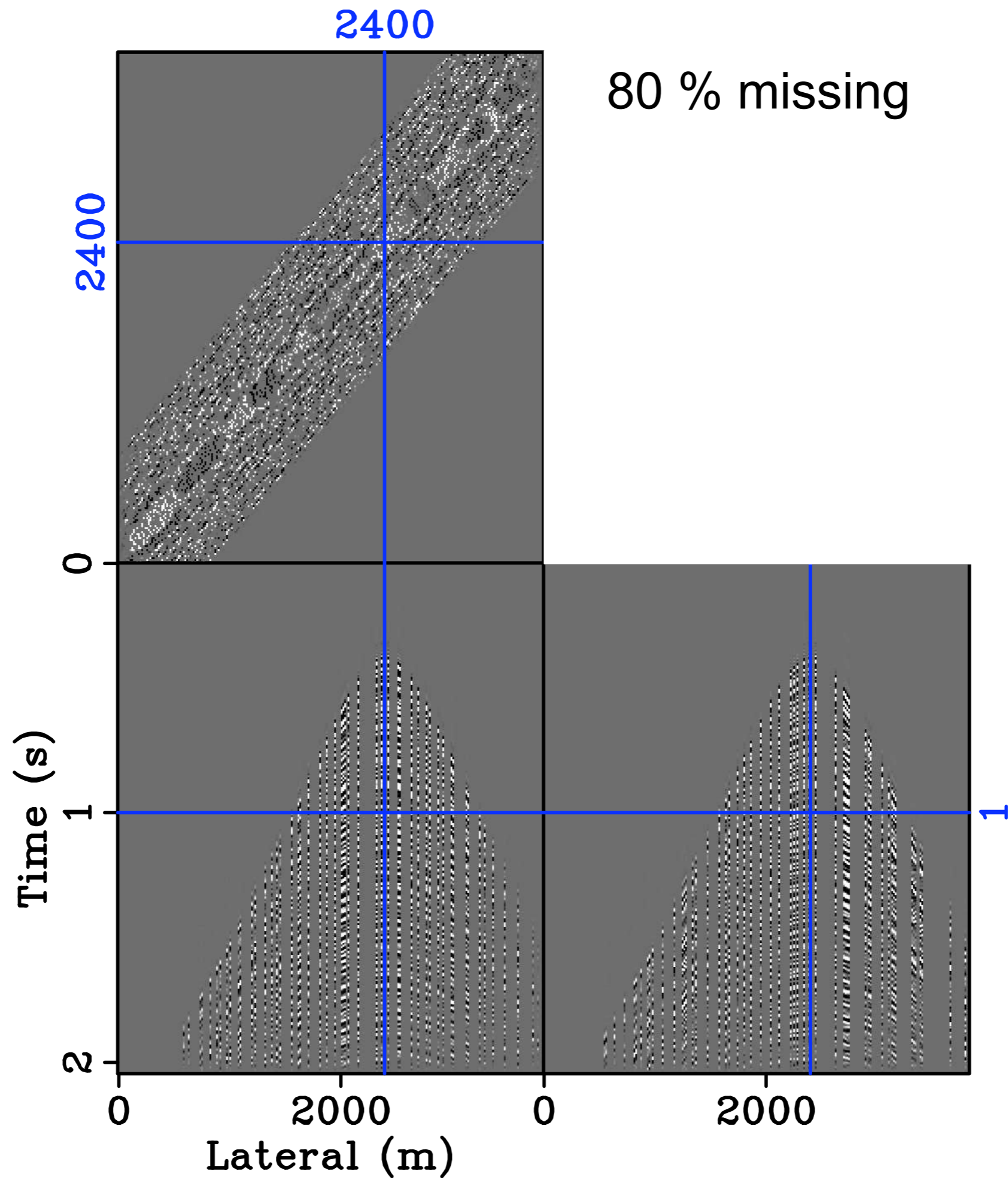
Original data

Green's function ΔP



SRME primaries

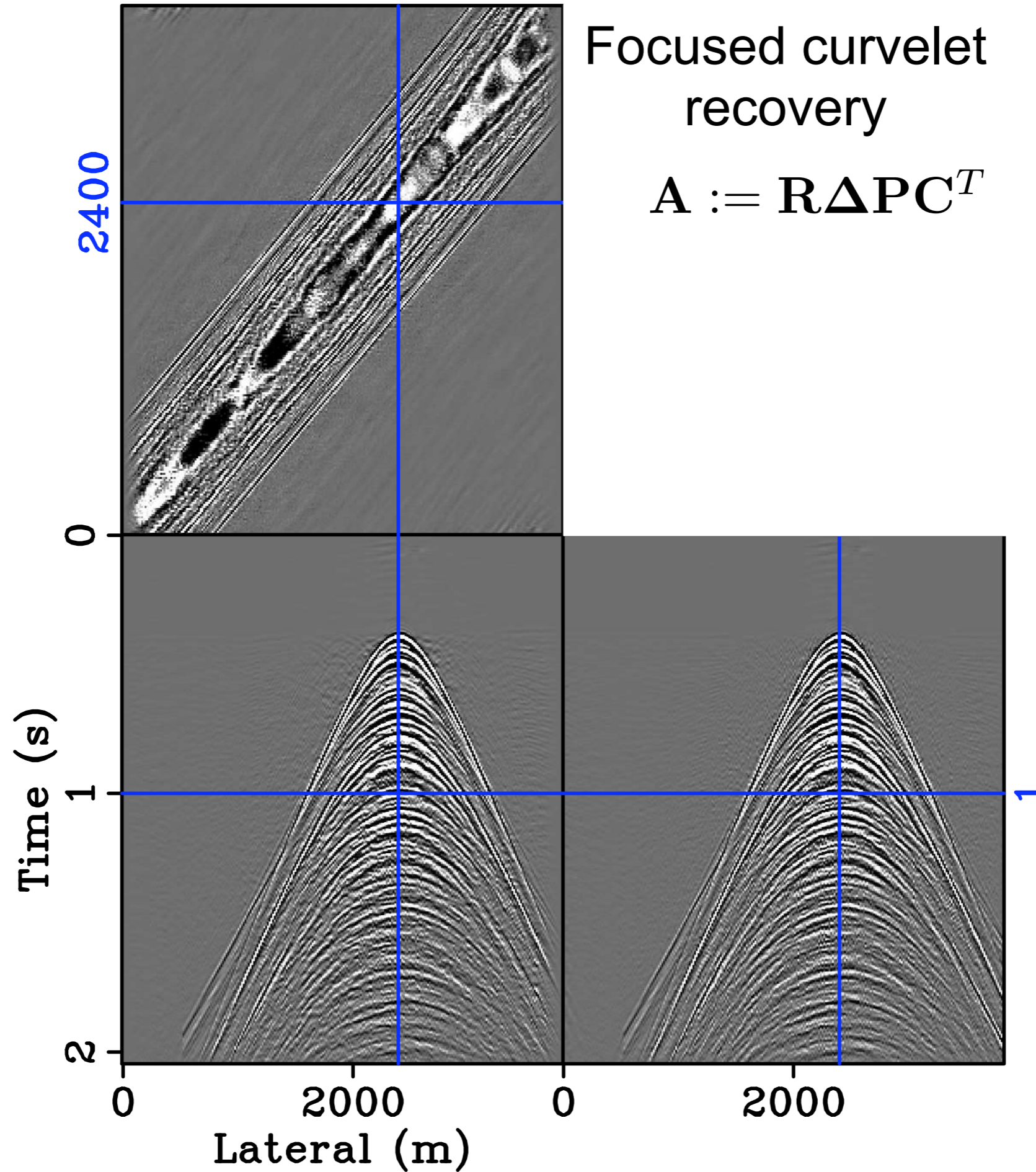


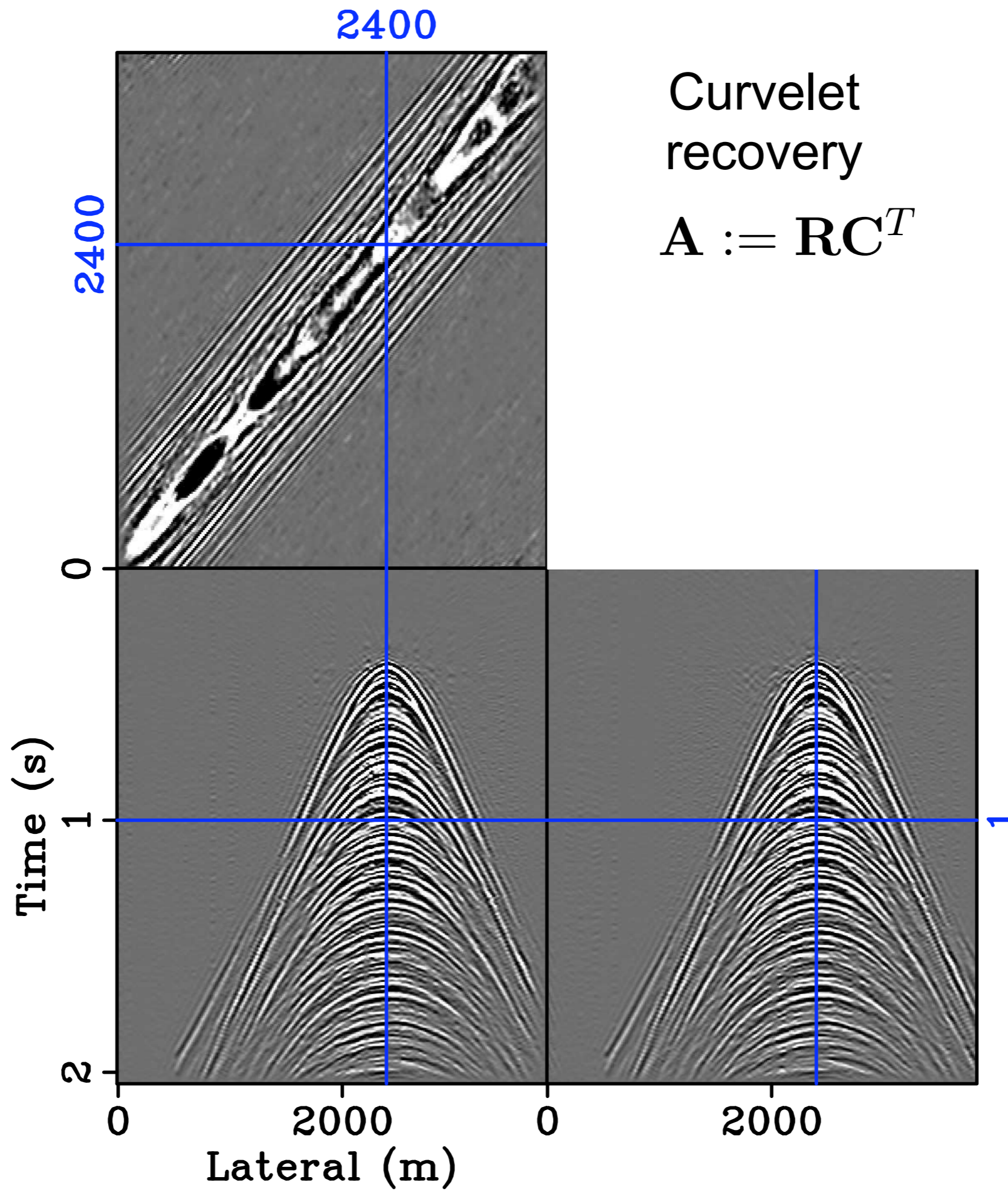


2400

Focused curvelet
recovery

$$\mathbf{A} := \mathbf{R}\mathbf{\Delta}\mathbf{P}\mathbf{C}^T$$





Curvelet
recovery

$$\mathbf{A} := \mathbf{RC}^T$$

Conclusions

- Curvelets sparsity on the model and near diagonalization yields stable inversion Gramm matrix
- Compressed wavefield extrapolation
 - reduction in synthesis cost
 - inverse extrapolation works well when focussed
 - mutual coherence curvelets and modes
 - performance of norm-one solver
- Double-role CS matrix is cool ... upscale to “real-life” will be a challenge
- Focusing in combination with curvelets leads to better recovery
- That is good seismic because **imaging = focusing**

Open problems

- What deeper insights can CS give?
 - CS principles and near unitary operators
 - Coherence generalized to frames to study
 - cols modeling operator \Leftrightarrow curvelets
 - radiation vs guided modes \Leftrightarrow curvelets

- Norm-one solver for reduced system as fast a LSQR on the full system

- Fast random eigenvalue solver does not exist

- Many more ...

Acknowledgments

The audience for listening and the organizers for putting this great program together

The authors of CurveLab (Demanet, Ying, Candes, Donoho)

Dr. W.W. Symes for his reverse-time migration code

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE (334810-05) of F.J.H. This research was carried out as part of the SINBAD project with support, secured through ITF (the Industry Technology Facilitator), from the following organizations: BG Group, BP, Chevron, ExxonMobil and Shell.