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 $\iff$ 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

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

- second order hyperbolic PDE
- interested in the singularities of

\[ m = c - \bar{c} \]
Inverse problem

Minimization:

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

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:

\[
(K^T d)(x) = (K^T K m)(x) + (K^T n)(x) \\
y(x) = (\Psi m)(x) + e(x)
\]

\( \Psi \) 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

“Sparsety- 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 Gramm 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 $\Psi$, 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


Alternative to expensive least-squares migration. In high-frequency limit \( \Psi \) 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 \( \overline{c} \)

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

Theorem 1. The following estimate for the error holds

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

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

Allows for the decomposition

$$(\Psi\varphi_\mu)(x) \simeq (C'^T D\Psi C\varphi_\mu)(x)$$

$$= (AA^T \varphi_\mu)(x)$$

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

Solve

\[
P : \begin{cases}
\min_x J(x) \quad \text{subject to} \quad \|y - Ax\|_2 \leq \epsilon \\
\tilde{m} = (A^H)^\dagger \tilde{x}
\end{cases}
\]

with

\[
J(x) = \alpha \|x\|_1 + \beta \|\Lambda^{1/2} (A^H)^\dagger x\|_p .
\]

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

- two-way reverse time wave-equation migration with checkpointing [Symes ‘07]
- adjoint state method with 8000 time steps
- evaluation $K^T$ 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.,

\[ u = T_\tau v \]
\[ = e^{-\tau D} v \]
\[ = L e^{-j\tau \Omega} L^H v \]

where

\[ D = L \Omega L^H \]
\[ L = \text{The Fourier Transform} \]

- Eigen modes <=> Fourier transform.
- Can this operation be compressed by compressive sampling?
Operators on spikes

[Candes et. al, Donoho]

Calculate instead

\[ \begin{align*}
    y' & = R e^{j \Omega \tau} \mathcal{F} v \\
    A & = R \mathcal{F} \\
    \tilde{u} & = \arg \min_u \| u \|_1 \quad \text{s.t.} \quad Au = y'
\end{align*} \]

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

\[ R \in \mathbb{R}^{m \times N} \quad \text{with} \quad m \ll N \]

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

<table>
<thead>
<tr>
<th></th>
<th>diagonalization operator</th>
<th>parsimony wavefield</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD/Lanczos/modal</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>curvelets</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>
## Different representations

<table>
<thead>
<tr>
<th></th>
<th>diagonalization operator</th>
<th>parsimony wavefield</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD/Lanczos/modal</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>curvelets</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

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 \( \mathbf{H}_2 \)
  \[ \mathbf{H}_2 = \begin{bmatrix} \left( \frac{\omega}{c_1} \right)^2 & 0 & \cdots & 0 \\ 0 & \left( \frac{\omega}{c_2} \right)^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left( \frac{\omega}{c_{n_1}} \right)^2 \end{bmatrix} + \mathbf{D}_2 \]

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

(Claerbout, 1971; Wapenaar and Berkhout, 1989)
Modal transform

- Solve eigenproblem & take square root
  \[
  H_1 = L \Lambda^{1/2} L^H
  \]
- \( L \) is orthonormal & defines the modal transform that diagonalizes one-way wavefield extrapolation
- Eigenvalues play role of vertical wavenumbers
- Extrapolation operator is diagonalized

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

- Radiating and guided modes

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

Recorded Data

Original events

Reconstruct point scatterers from recorded data ....

\[ \mathbf{u} = \mathbf{L} e^{-j \frac{1}{2} \Lambda \Delta x_3} \mathbf{L}^H \mathbf{v} \]
Compressed wavefield extrapolation

\[
\begin{align*}
    y &= RL^H u \\
    A &= Re^j \Lambda^{1/2} \Delta x_3 L^H \\
    \tilde{x} &= \arg\min_x \|x\|_1 \quad \text{s.t.} \quad Ax = y \\
    \tilde{v} &= \tilde{x}
\end{align*}
\]

- 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 $\iff$ Fourier recovery
Compressed wavefield extrapolation

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 ...
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:**

  \[ W_1 : \begin{cases} 
  y' = \text{Re}^j \Lambda^{1/2} \Delta x_3 L^H \\
  A := RL^H FC^T \\
  \tilde{x} = \text{arg min}_x \|x\|_1 \quad \text{s.t.} \quad Ax = y' \\
  \tilde{u} = C^T \tilde{x},
  \end{cases} \]

- **Inverse extrapolation:**

  \[ F_1 : \begin{cases} 
  y = RL^H FC u \\
  A' = \text{Re}^j \Lambda^{1/2} \Delta x_3 L^H C^H \\
  \tilde{x} = \text{arg min}_x \|x\|_1 \quad \text{s.t.} \quad A'x = y \\
  \tilde{v} = C^T \tilde{x}.
  \end{cases} \]
Forward Extrapolation

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

Figure 11: Compressed forward extrapolation according to W1 wcf. Eq. w42xx for restrictions. Observe that the forward propagated wavefield is largely recovered for the restriction in (c).

Herrmann et.al. – 65
Inverse Extrapolation

Figure 13: Compressed inverse extrapolation according to Eq. 45 for different parameters. Observe that the recovery for the velocity high is slightly better.

- (a) $p = 0.04$
- (b) $p = 0.16$, (c) $p_f=0.4$, $p_x=0.4$
Evanescent Recovery

(a) is downward extrapolated wavefield
(b) is matched filter
(c) is “compressed” inverse extrapolation
Velocity model

Figure 13: Lateral velocity profile for the overthrust examples. Herrmann et al. –
Compressed inverse extrapolation

Overthrust exploding reflector

Full forward extrapolation

Matched filter

Recovered from p=0.25
Propose a scheme motivated by extensions of CS

$$F_j^1 : \begin{cases} 
  y_j = R_j M_j u \\
  A'_j := R_j M'_j C_j^T \\
  \tilde{x}_j = \text{arg min}_{x_j} \| x_j \|_1 \quad \text{s.t.} \quad A'_j x_j = y_j \\
  \tilde{v} = \sum_j C_j^T \tilde{x}_j,
\end{cases}$$

with $j = \{j, l\}$ 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 focusssing

Solve

\[
P_\epsilon : \begin{cases} 
\tilde{x} = \arg \min_x \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \epsilon \\
\tilde{f} = S^T \tilde{x}
\end{cases}
\]

with

\[
A := R \Delta PC^T \\
S^T := \Delta PC^T \\
y = RP(\cdot) \\
R = \text{picking operator.}
\]
Recovery with focusing

Solve

\[ P_\epsilon : \begin{cases} \tilde{x} = \arg \min_x \|x\|_1 \quad \text{s.t.} \quad \|Ax - y\|_2 \leq \epsilon \\ \tilde{f} = S^T\tilde{x} \end{cases} \]

with

\[ A := R\Delta PC^T \]
\[ S^T := \Delta PC^T \]
\[ y = RP(\cdot) \]
\[ R = \text{picking operator.} \]
Green's function $\Delta P$
Focused curvelet recovery

\[ A := R \Delta P C^T \]
The reconstruction of seismic wavefields from regularly sampled data with missing traces is a significant problem. Comparing the 'ground truth' in Figure \(1\) with the recovered data in Figure \(2\), we observe a remarkable match. The mathematical formulation for this recovery is given by:

\[ A := RC^T \]

This equation encapsulates the process of recovering the seismic wavefield from incomplete data.
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 focused
  - 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 $\iff$ curvelets
    - radiation vs guided modes $\iff$ 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.