Large scale high-frequency seismic wavefield reconstruction, acquisition via rank minimization and sparsity-promoting source estimation

Shashin Sharan
Ph.D. Defense of Dissertation
October 27, 2020
Seismic data acquisition

[Caldwell and Walker, '11]
Seismic data acquisition

Objective

- Acquire Dense seismic data
- for high-resolution imaging

[Caldwell and Walker,'11]
Seismic data acquisition

Objective

- Acquire Dense seismic data
- for high-resolution imaging

Challenge

- Operationally complex
- to acquire dense data

[Caldwell and Walker,’11]
Seismic data acquisition

Objective

- Acquire Dense seismic data
- for high-resolution imaging

Challenge

- Operationally complex
- to acquire dense data

Solution

- Acquisition on coarse grid
- followed by data reconstruction
- on periodic fine grid

[Caldwell and Walker, '11]
Common Shot Gather Visualization

![Common Shot Gather Visualization Diagrams](image-url)
Seismic data reconstruction w/ Low-rank

Advantages

› Scalable for large scale 3D data
› Performs well at lower frequencies

Limitations

› Performs poorly at higher frequencies
Seismic data reconstruction with Low-rank

60 Hz Frequency slice  data w/ 75% missing sources

Advantages

- Scalable for large scale 3D data
- Performs well at lower frequencies

Limitations

- Performs poorly at higher frequencies

Reconstructed data, 2.83 dB  Data residual

[Kumar et al., ’16]
Key Contributions: Chapters 2 & 3

Recursively Weighted Matrix Completion Framework

- Improved data reconstruction at high frequency
- Scalable for large scale 3D data
- Computationally faster weighted method in comparison to traditional weighted method

5D Time-Jittered Marine Acquisition

- Simultaneous separation and reconstruction of sources from blended data
- Scalable for large scale 3D data
Successful reconstruction scheme

- **exploit structure**
  - low-rank / fast decay of singular values

- **sampling**
  - randomness increases rank in “transform domain”

- **optimization**
  - via rank-minimization (nuclear norm-minimization)

Matrix completion

[Candès and Plan,’09; Kumar et al.,’16]
Nuclear-norm minimization

\[ \begin{align*} 
\text{minimize} & \quad \|X\|_* \quad \text{subject to} \quad \|A(X) - B\|_F \leq \epsilon \\
X \in \mathbb{C}^{m \times n} & \end{align*} \]

**Sum of singular values of X**

*where \(\|\cdot\|_F\) is the Frobenius norm
\(A\) is the Measurement operator
\(B\) is the observed data
\(\epsilon\) is the noise level*
Nuclear-norm minimization

\[
\begin{align*}
\text{minimize} & \quad \|X\|_* \quad \text{subject to} \quad \|A(X) - B\|_F \leq \epsilon \\
\text{X} & \in \mathbb{C}^{m \times n}
\end{align*}
\]

*where \(\|\cdot\|_F\) is the Frobenius norm
\(A\) is the Measurement operator
\(B\) is the observed data
\(\epsilon\) is the noise level

Sum of singular values of \(X\)

\[
\text{Convex relaxation of rank-minimization}
\]

Weighted Nuclear-norm minimization

\[
\begin{align*}
\text{minimize} & \quad \|QXW\|_* \quad \text{subject to} \quad \|A(X) - B\|_F \leq \epsilon \\
\text{X} & \in \mathbb{C}^{m \times n}
\end{align*}
\]

*where \(Q = w_1UU^H + U^\perp U^{\perp H}\) and \(W = w_2VV^H + V^\perp V^{\perp H}\)

\(U, V\) are row and column subspaces of adjacent frequency slice
scalars \(w_1, w_2 \in (0, 1]\) are weights

[Recht et al.,'16; Aravkin et al.,'14; Eftekhar et al.,'18]
Nuclear-norm minimization

\[
\text{minimize } \|X\|_* \quad \text{subject to } \|A(X) - B\|_F \leq \epsilon
\]

\text{Sum of singular values of } X

*where \(\|\cdot\|_F\) is the Frobenius norm
\(A\) is the Measurement operator
\(B\) is the observed data
\(\epsilon\) is the noise level

Weighted Nuclear-norm minimization

\[
\text{minimize } \|QXW\|_* \quad \text{subject to } \|A(X) - B\|_F \leq \epsilon
\]

\*where \(Q = w_1UU^H + U^\perp U^\perp H\) and \(W = w_2VV^H + V^\perp V^\perp H\)

\(U, V\) are row and column subspaces of adjacent frequency slice
\(w_1, w_2 \in (0, 1)\) are weights
Efficient Weighted Matrix Completion

\[
\begin{align*}
\text{minimize} & \quad \|\bar{X}\|_* \quad \text{subject to} \quad \|A(Q^{-1}\bar{X}W^{-1}) - B\|_F \leq \epsilon \\
\text{where} & \quad \bar{X} = QXW \\
\text{and} & \quad X = Q^{-1}\bar{X}W^{-1}
\end{align*}
\]
Efficient Weighted Matrix Completion

\[
\begin{align*}
\text{minimize} \quad & \|\bar{X}\|_* \quad \text{subject to} \quad \|A(Q^{-1}\bar{X}W^{-1}) - B\|_F \leq \epsilon \\
\text{where} \quad & \bar{X} = QXW \\
\text{and} \quad & X = Q^{-1}\bar{X}W^{-1}
\end{align*}
\]

\[
\begin{align*}
\text{minimize} \quad & \frac{1}{2} \left\| \begin{bmatrix} \bar{L} \\ \bar{R} \end{bmatrix} \right\|_F^2 \\
\text{subject to} \quad & \|A(Q^{-1}\bar{L}\bar{R}^H W^{-1}) - B\|_F \leq \epsilon
\end{align*}
\]
Efficient Weighted Matrix Completion

\[
\begin{align*}
\text{minimize} \quad & \| \tilde{X} \|_* \quad \text{subject to} \quad \| A(Q^{-1}\tilde{X}W^{-1}) - B \|_F \leq \epsilon \\
\text{where} \quad & \tilde{X} = QXW \\
\text{and} \quad & X = Q^{-1}\tilde{X}W^{-1}
\end{align*}
\]

*where \( \tilde{X} = QXW \)
and \( X = Q^{-1}\tilde{X}W^{-1} \)

Factorized Form

\[
\begin{align*}
\text{minimize} \quad & \frac{1}{2} \left\| \begin{bmatrix} \bar{L} \\ \bar{R} \end{bmatrix} \right\|_F^2 \\
\text{subject to} \quad & \| A(Q^{-1}\bar{L}\bar{R}^HW^{-1}) - B \|_F \leq \epsilon
\end{align*}
\]

*where \( \tilde{X} = \bar{L}\bar{R}^H \)
with \( \bar{L} \in \mathbb{C}^{m \times k} \) and \( \bar{R} \in \mathbb{C}^{n \times k} \)
and \( k << m, n \)
and \( L = Q^{-1} \bar{L}, \ R = W^{-1} \bar{R} \)
Efficient Weighted Matrix Completion

\[
\begin{align*}
\text{minimize} & \quad \| \tilde{X} \|_* \quad \text{subject to} \quad \| A(Q^{-1}\tilde{X}W^{-1}) - B \|_F \leq \epsilon \\
\text{where} & \quad \tilde{X} = QXW \\
\text{and} & \quad X = Q^{-1}\tilde{X}W^{-1}
\end{align*}
\]

*where \( \tilde{X} = L\tilde{R}^H \)

with \( \tilde{L} \in \mathbb{C}^{m \times k} \) and \( \tilde{R} \in \mathbb{C}^{n \times k} \)

and \( k \ll m, n \)

and \( L = Q^{-1}\tilde{L}, R = W^{-1}\tilde{R} \)

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} \left\| \begin{bmatrix} L \\ R \end{bmatrix} \right\|^2_F \\
\text{subject to} & \quad \| A(Q^{-1}L\tilde{R}^HW^{-1}) - B \|_F \leq \epsilon
\end{align*}
\]

**Factorized Form**

**Runtime Comparison**
Recursively Weighted vs Pair Weighted

Reconstructed data

Data residual

Pair Weighted, SNR = 5.08 dB
Recursively Weighted vs Pair Weighted

**Reconstructed data**

**Data residual**

Pair Weighted, SNR = 5.08 dB

Recursively Weighted, SNR = 8.72 dB
Parallel Implementation

Recursively weighted method challenges
- need to wait for reconstruction of previous frequencies
- computationally demanding for large scale 3D data

Alternating minimization and decoupling strategies enable
- recursively framework scalable for large 3D data
- parallelized framework

[Recht et al.,’13; Lopez et al.,’15]
Parallelization w/o weights

\[ \mathbf{R}(l_1, :)^H := \arg \min_{\mathbf{v}} \frac{1}{2} \| \mathbf{v} \|^2 \ \text{subject to} \| \mathbf{A}_{l_1}(\mathbf{L}\mathbf{v}) - \mathbf{B}(::, l_1) \| \leq \gamma \]

for \( l_1 = 1, 2, \cdots, n \)

\[ \mathbf{L}(l_2, :)^H := \arg \min_{\mathbf{u}} \frac{1}{2} \| \mathbf{u} \|^2 \ \text{subject to} \| \mathbf{A}_{l_2}((\mathbf{R}\mathbf{u})^H) - \mathbf{B}(l_2, :) \| \leq \gamma \]

for \( l_2 = 1, 2, \cdots, m \)

* where \( \mathbf{A}_{l_1} \) is acquisition mask for \( l_1^{th} \) row of \( \mathbf{B} \)

and \( \mathbf{A}_{l_2} \) is acquisition mask for \( l_2^{th} \) column of \( \mathbf{B} \)

Inclusion of weight matrices poses challenge in parallelizing
Parallelization w/ weights

For large weights:

\[ QA(Q^{-1}\bar{X}W^{-1}) \approx A(\bar{X}W^{-1}) \quad \text{and} \quad A(Q^{-1}\bar{X}W^{-1})W \approx A(Q^{-1}\bar{X}) \]
Parallelization w/ weights

For large weights:

\[ Q_A(Q^{-1}\tilde{X}W^{-1}) \approx A(\tilde{X}W^{-1}) \quad \text{and} \quad A(Q^{-1}\tilde{X}W^{-1})W \approx A(Q^{-1}\tilde{X}) \]

This commutation property allows parallelization

\[ \bar{R}(l_1,:)^H := \arg \min_{\bar{v}} \quad \frac{1}{2} \|\bar{v}\|^2 \quad \text{subject to} \quad \|A_{l_1}(\hat{Q}\bar{L}\bar{v}) - w_1w_2B(:,l_1)\| \leq w_1w_2\gamma \]

\[ \bar{L}(l_2,:)^H := \arg \min_{\bar{u}} \quad \frac{1}{2} \|\bar{u}\|^2 \quad \text{subject to} \quad \|A_{l_2}((\hat{R}\bar{u})^H\hat{W}) - w_1w_2B(l_2,:)\| \leq w_1w_2\gamma \]

*where \( \hat{Q} = UU^H + w_1V^\perp V^\perp H = w_1Q^{-1} \)

and \( \hat{W} = VV^H + w_2V^\perp V^\perp H = w_2W^{-1} \)
Case Study: BG Synthetic 3D Data

**Data dimension:** 501 x 201 x 201 x 41 x 41 (nt x nrx x nry x nsx x nsy)

**Time sampling interval:** 10 ms

**Source sampling interval (x and y):** 150 m

**Receiver sampling interval (x and y):** 25 m

**Velocity model:** BG Compass

**Observed data:** 90 % missing receivers
Optimization Information

Number of alternations per frequency: 4

Number of inner iterations per alternation: 40

Rank parameter: 228

Weight: 0.75

Computational resource: AWS cloud

Time per frequency: 8 minutes

Final volume: 7 GB (95% compressed)
15 Hz Frequency Slice

Ground Truth

Observed w/ 90% missing receivers
15 Hz Frequency Slice

- **Ground Truth**
- **Reconstruction w/ conventional, SNR = 3.7 dB**
- **Observed w/ 90% missing receivers**
- **Data Residual w/ conventional**
15 Hz Frequency Slice

Ground Truth

Reconstruction w/ conventional, SNR = 3.7 dB

Reconstruction w/ recursively weighted, SNR = 12.5 dB

Observed w/ 90% missing receivers

Data Residual w/ conventional

Data Residual w/ recursively weighted
Time Domain Results

Ground Truth

Observed Data w/ 90% missing receivers
Time Domain Results

Reconstruction w/ conventional Data residual
Time Domain Results

Reconstruction w/ recursively weighted

Data residual
Signal to noise ratio comparison

![Signal to noise ratio comparison graph]

- **Conventional**
- **Recursively Weighted**

**Y-axis**: Signal to noise ratio (dB)

**X-axis**: Frequency (Hz)
5D Time-Jittered marine acquisition
Objective

- Acquire blended seismic data using multiple sources
- Simultaneous separation and reconstruction of sources on dense grid
5D Time-Jittered marine acquisition

Objective

‣ Acquire blended seismic data using multiple sources
‣ Simultaneous separation and reconstruction of sources on dense grid

Benefits

‣ Reduction in overall cost of acquiring dense seismic data
Low-rank formulation

Restriction operator is non-separable
  ▸ combination of time-shifting and shot-jittered operator
Low-rank formulation

Restriction operator is non-separable
  ▸ combination of time-shifting and shot-jittered operator

\[ \mathcal{A}(.) = MF^H S^H(.) \]

\[ M \quad \text{time-jittered operator} \]
\[ F^H \quad \text{inverse Fourier transform along frequency axis} \]
\[ S \quad \text{rank-revealing transform} \]
Low-rank formulation

Restriction operator is non-separable
- combination of time-shifting and shot-jittered operator

\[ \mathcal{A}(.) = MF^H S^H(.) \]

- \( M \) time-jittered operator
- \( F^H \) inverse Fourier transform along frequency axis
- \( S \) rank-revealing transform

Can’t perform matrix-completion over independent frequencies
- reformulate low-rank factorization over temporal-frequency domain
Low-rank formulation

Restriction operator is non-separable
  ▸ combination of time-shifting and shot-jittered operator

\[ \mathcal{A}(.) = MF^H S^H(.) \]

- \( M \) time-jittered operator
- \( F^H \) inverse Fourier transform along frequency axis
- \( S \) rank-revealing transform

Can’t perform matrix-completion over independent frequencies
  ▸ reformulate low-rank factorization over temporal-frequency domain

\[
\min_{L, R} \sum_{j}^{n_f} \frac{1}{2} \left\| \begin{bmatrix} L_j \\ R_j \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \| \mathcal{A}(LR^H) - b \|_2 \leq \epsilon
\]
Case Study: 3D BG Compass model

Temporal length
  ▪ 65 minutes

25 m flip-flop shooting
  ▪ source-sampling ranges from 25 m to 175 m
  ▪ acquired 400 sources

101 x 101 receivers (nrx x nry)

Ricker wavelet with central frequency 15 Hz

Dimensions of deblended/interpolated data volume on 6.25 m grid
  ▪ 2501 x 101 x 101 x 40 x 40 (nt x nrx x nry x nsx x nsy)
Optimization Information

Computational environment
- SENAI Yemoja cluster

Parallelized over
- receivers and frequencies

Computational information
- 200 iterations, 42 hours

Final volume
- 13 GB (98% compression)
5D Time-Jittered marine acquisition

Blended data @ 25 m flip-flop (overlapping & missing shots)

Separation + Interpolation (recovered grid @ 6.25m)

Recovery

SNR = 21 dB
Conclusions

Low-rank matrix factorization based wavefield reconstruction method

- performs poorly at higher frequencies
- recursively weighted method improves reconstruction at higher frequencies
- by including prior information from lower frequencies
Conclusions

Low-rank matrix factorization based wavefield reconstruction method

- performs poorly at higher frequencies
- recursively weighted method improves reconstruction at higher frequencies
- by including prior information from lower frequencies

Scaling for full azimuth industry-size data is achieved via

- shifting the weights from objective to data-misfit to avoid expensive projections
- using strategies of alternating minimization and decoupling
- parallelizing over rows of low-rank factors
Conclusions

Low-rank matrix factorization based wavefield reconstruction method

- performs poorly at higher frequencies
- recursively weighted method improves reconstruction at higher frequencies
- by including prior information from lower frequencies

Scaling for full azimuth industry-size data is achieved via

- shifting the weights from objective to data-misfit to avoid expensive projections
- using strategies of alternating minimization and decoupling
- parallelizing over rows of low-rank factors

Factorization based time-jittered acquisition

- scalable for large scale 5D data
- by using parallel computation
- achieves data compression by saving low-rank factors
Future work

In recursively weighted framework
  ▶ include smaller weights
  ▶ for large scale data
Future work

In recursively weighted framework
- include smaller weights
- for large scale data

In time-jittered acquisition
- include weights
- to further improve reconstruction quality
Sparsity-promoting source estimation
Unconventional Reservoir
Unconventional Reservoir
Unconventional Reservoir
Unconventional Reservoir
Unconventional Reservoir
Unconventional Reservoir

Objectives

- detection of microseismic events in space and time
- estimation of source-time function
Key Contributions: Chapters 4 & 5

Sparsity-promoting microseismic estimation

- Detection of closely spaced microseismic sources from noisy data
- Estimation of associated source-time function
Proposed method w/ sparsity promotion
Proposed method w/ sparsity promotion
Proposed method w/ sparsity promotion

Assumptions
- localized in space
- finite energy along time
Proposed method w/ sparsity promotion

\[
\begin{align*}
\text{minimize} & \quad \| Q \|_{2,1} \\
\text{subject to} & \quad \| \mathcal{F}[m](Q) - d \|_2 \leq \epsilon
\end{align*}
\]

Source field

Forward modeling operator

Slowness square

Noise level

Observed data

\( Q \in \mathbb{R}^{n_x \times n_t} \)

\( n_x \): number of grid points

\( n_t \): number of time samples
Proposed method w/ sparsity promotion

\[
\minimize_{Q} \|Q\|_{2,1} \quad \text{subject to} \quad \|F[m](Q) - d\|_{2} \leq \epsilon
\]

Source field

Noise level

Forward modeling operator

Slowness square

Observed data

\( Q \in \mathbb{R}^{n_x \times n_t} \)

\( n_x \): number of grid points

\( n_t \): number of time samples

Similar to classic Basis pursuit denoising (BPDN)

[Van Den Berg et al.,'08]
Solving w/ Linearized Bregman

\[
\begin{align*}
\text{minimize} & \quad \|Q\|_{2,1} + \frac{1}{2\mu} \|Q\|_F^2 \\
\text{subject to} & \quad \|F[m](Q) - d\|_2 \leq \epsilon
\end{align*}
\]

*where \( \|.\|_F \) is the Frobenius norm

- Recent successful application to seismic imaging problem
- Three-step algorithm simple to implement
- Choice of \( \mu \) controls the trade off between sparsity and the Frobenius norm
- \( \mu \uparrow \infty \) corresponds to solving original BPDN problem
Linearized Bregman algorithm

1. **Data** $d$, **slowness square** $m$  //Input
2. **for** $k = 0, 1, \ldots$
3. $V_k = \mathcal{F}^\top [m] (\Pi_\epsilon (\mathcal{F}[m](Q_k) - d))$  //adjoint solve
4. $Z_{k+1} = Z_k - t_k V_k$  //auxiliary variable update
5. $Q_{k+1} = \text{Prox}_{\mu \ell_2,1}(Z_{k+1})$  //sparsity promotion
6. **end**
7. $I(x) = \sum_t |Q(x, t)|$  //Intensity plot
Linearized Bregman algorithm

1. **Data d, slowness square m**  //Input
2. **for** $k = 0, 1, \ldots$
3. \[ V_k = \mathcal{F}^\top [m] (\Pi_{\epsilon} (\mathcal{F}[m](Q_k) - d)) \]  //adjoint solve
4. \[ Z_{k+1} = Z_k - t_k V_k \]  //auxiliary variable update
5. \[ Q_{k+1} = \text{Prox}_{\mu \ell_2, 1} (Z_{k+1}) \]  //sparsity promotion
6. **end**
7. **I(x) = \sum_t |Q(x, t)|**  //Intensity plot

* $\Pi_{\epsilon}(x) = \max\{0, 1 - \epsilon/\|x\|\}, (x)$ the projection on to $\ell_2$ norm ball
Linearized Bregman algorithm

1. **Data** \( d \), slowness square \( m \)  //Input  
2. **for** \( k = 0, 1, \cdots \)  
3. \( V_k = \mathcal{F}^\top (m) (\Pi_\epsilon (\mathcal{F}[m](Q_k) - d)) \)  //adjoint solve  
4. \( Z_{k+1} = Z_k - t_k V_k \)  //auxiliary variable update  
5. \( Q_{k+1} = \text{Prox}_{\mu \ell_2,1} (Z_{k+1}) \)  //sparsity promotion  
6. **end**  
7. \( I(x) = \sum_t |Q(x, t)| \)  //Intensity plot

* \( \Pi_\epsilon (x) = \max\{0, 1 - \frac{\epsilon}{\|x\|}\}, (x) \) the projection on to \( \ell_2 \) norm ball  

*where \( t_k = \frac{\|\mathcal{F}[m](Q_k) - d\|^2}{\|\mathcal{F}[m](\mathcal{F}[m](Q_k) - d)\|^2} \) is the dynamic step length
**Linearized Bregman algorithm**

1. **Data d, slowness square m**  //Input
2. **for** $k = 0, 1, \cdots$
3. \[ V_k = \mathcal{F}^\top [m] (\Pi_\varepsilon (\mathcal{F}[m](Q_k) - d)) \]  //adjoint solve
4. \[ Z_{k+1} = Z_k - t_k V_k \]  //auxiliary variable update
5. \[ Q_{k+1} = \text{Prox}_{\mu \ell_2,1}(Z_{k+1}) \]  //sparsity promotion
6. **end**
7. \[ I(x) = \sum_t |Q(x, t)| \]  //Intensity plot

* $\Pi_\varepsilon(x) = \max\{0, 1 - \frac{\varepsilon}{\|x\|}\}(x)$ the projection on to $\ell_2$ norm ball

*where $t_k = \frac{\|\mathcal{F}[m](Q_k) - d\|^2}{\|\mathcal{F}[m](\mathcal{F}[m](Q_k) - d)\|^2}$ is the dynamic step length

* $\text{Prox}_{\mu \ell_2,1}(C) := \arg \min_B \|B\|_{2,1} + \frac{1}{2\mu} \|C - B\|_F^2$ is the proximal mapping of the $\ell_{2,1}$ norm
Linearized Bregman algorithm

1. **Data** \( d \), **slowness square** \( m \)  //Input
2. **for** \( k = 0, 1, \cdots \)
3. \( V_k = \mathcal{F}^T[m](\Pi_\varepsilon(\mathcal{F}[m](Q_k) - d)) \)  //adjoint solve
4. \( Z_{k+1} = Z_k - t_k V_k \)  //auxiliary variable update
5. \( Q_{k+1} = \text{Prox}_{\mu \ell_{2,1}}(Z_{k+1}) \)  //sparsity promotion
6. **end**
7. \( I(x) = \sum_t |Q(x, t)| \)  //Intensity plot

* \( \Pi_\varepsilon(x) = \max\{0, 1 - \frac{\varepsilon}{\|x\|}\} \) \( \varepsilon \) the projection on to \( \ell_2 \) norm ball

*where \( t_k = \frac{\|F[m][Q_k]-d\|^2}{\|F[m][F[m][Q_k]-d]\|^2} \) is the dynamic step length

* \( \text{Prox}_{\mu \ell_{2,1}}(C) := \arg\min_B \|B\|_{2,1} + \frac{1}{2\mu}\|C - B\|_F^2 \) is the proximal mapping of the \( \ell_{2,1} \) norm

- **Source location**: estimated as outlier in intensity plot
- **Source-time function**: temporal variation of wavefield at estimated source location
$\mathbf{V}_1 = \mathcal{F}^\dagger [\mathbf{m}] (\Pi_\epsilon (\mathcal{F} [\mathbf{m}] (\mathbf{Q}_0 - \mathbf{d})))$

**Adjoint solve**
\[ V_1 = \mathcal{F}^\dagger [m](\mathbb{II}_\varepsilon(\mathcal{F}[m](Q_0) - d)) \]

Adjoint solve
\[ V_1 = \mathcal{F}^\dagger [m][\Pi_\xi(\mathcal{F}[m](Q_0 - d))] \]

**Adjoint solve**

**Auxiliary variable update**

\[ Z_1 = Z_0 - t_1 V_1 \]
\( \mathbf{V}_1 = \mathcal{F}^\dagger [\mathbf{m}] (\Pi_\epsilon (\mathcal{F}[\mathbf{m}](\mathbf{Q}_0) - \mathbf{d})) \)

Adjoint solve

\[ \mathbf{Z}_1 = \mathbf{Z}_0 - t_1 \mathbf{V}_1 \]
\[
V_1 = F^\dagger m (\Pi_\varepsilon (F m (Q_0 - d)))
\]

**Adjoint solve**

\[
Z_1 = Z_0 - t_1 V_1
\]

**Auxiliary variable update**

\[
Q_1 = \text{Prox}_{\mu \ell_2,1}(Z_1)
\]

**Sparsity promotion**
\[ V_1 = \mathcal{F}^\dagger [m](\Pi_{\ell}(\mathcal{F}[m](Q_0) - d)) \]

**Adjoint solve**

\[ Z_1 = Z_0 - t_1 V_1 \]

**Auxiliary variable update**

\[ Q_1 = \text{Prox}_{\mu \ell_{2,1}}(Z_1) \]

**Sparsity promotion**

\[ I(x) = \sum_t |Q_1(x, t)| \]
\[ V_1 = F^\dagger [m] (\Pi_\epsilon (F[m] (Q_0) - d)) \]

**Adjoint solve**

\[ Z_1 = Z_0 - t_1 V_1 \]

**Auxiliary variable update**

**Sparsity promotion**

\[ Q_1 = \text{Prox}_{\mu \ell_{2,1}} (Z_1) \]
Modeling information:

Model size: 0.7 km x 0.7 km
Grid spacing: 5m
Receiver spacing: 10m
Receiver depth: 20m
Fixed spread: 0.69 km
Sampling interval: 2 ms
Recording length: 1 s
Peak frequency: 30 Hz
Dominant wavelength: 46 m
Source separation: 22 m
Results w/ different threshold parameters

\[ \mu = 8 \times 10^{-5}, 200 \text{ iterations} \]

\[ \mu = 8 \times 10^{-4}, 600 \text{ iterations} \]

\[ \mu = 8 \times 10^{-3}, 4900 \text{ iterations} \]
Acceleration with quasi-Newton: Algorithm

1. Data \( d \), slowness square \( m \), number of iterations \( k \)  //Input
2. Initialize dual variable \( y = 10^{-3}d \)
3. \( \hat{y} = \text{L-BFGS}(f(y), g(y), y, k) \)  //Dual solution
   where \( f(y) = \Psi(y) - \epsilon\|y\|_2 \)  //L-BFGS objective
   and \( g(y) = \Psi'(y) - \epsilon y/\|y\|_2 \)  //L-BFGS gradient
4. \( \hat{Q} = \text{Prox}_{\mu \ell_{2,1}}(\mu F[m]^\top(\hat{y})) \)  //Primal solution
5. \( I(x) = \sum_t |\hat{Q}(x, t)| \)  //Intensity plot

*where \( \Psi(y) = \min_Q \|Q\|_{2,1} + \frac{1}{2\mu} \|Q\|_F - y^\top (F[m](Q) - d) \)

* \( \Psi'(y) = d - F[m](\text{Prox}_{\mu \ell_{2,1}}(\mu F[m]^\top(y))) \) is the gradient of \( \Psi(y) \)
Acceleration with quasi-Newton: Algorithm

1. Data $d$, slowness square $m$, number of iterations $k$  //Input
2. Initialize dual variable $\hat{y} = 10^{-3}d$
3. $\hat{y} = \text{L-BFGS}(f(y), g(y), y, k)$  //Dual solution
   where $f(y) = \Psi(y) - \epsilon\|y\|_2$  //L-BFGS objective
   and $g(y) = \Psi'(y) - \epsilon y/\|y\|_2$  //L-BFGS gradient
4. $\hat{Q} = \text{Prox}_{\mu\ell_2,1}(\mu\mathcal{F}[m]^\top(\hat{y}))$  //Primal solution
5. $I(x) = \sum_t |\hat{Q}(x, t)|$  //Intensity plot

*where $\Psi(y) = \min_Q \|Q\|_{2,1} + \frac{1}{2\mu} \|Q\|_F - y^\top (\mathcal{F}[m](Q) - d)$

* $\Psi'(y) = d - \mathcal{F}[m](\text{Prox}_{\mu\ell_2,1}(\mu\mathcal{F}[m]^\top(y)))$ is the gradient of $\Psi(y)$
Further acceleration w/ 2D Preconditioning

\[
\begin{align*}
\text{minimize} & \quad \|Q\|_{2,1} + \frac{1}{2\mu} \|Q\|_F^2 \\
\text{subject to} & \quad \|\mathcal{M}_L \mathcal{F}[m](Q) - \mathcal{M}_L d\|_2 \leq \gamma
\end{align*}
\]

*with \( \mathcal{M}_L := \partial_{|t|}^{1/2} \) is the half differentiation correction

*where \( \partial_{|t|}^{1/2} = \mathbf{F}^{-1}|\omega|^{1/2}\mathbf{F} \)

*\( \mathbf{F} \) is the Fourier transform and \( \omega \) is the frequency

*\( \gamma \) is the noise level
Estimated location
w/ $\mu = 8e-2$ and 10 iterations
Convergence comparison: LBR vs L-BFGS

Convergence comparison
- Using same value of $\mu$

Improvement in convergence with
- Dual formulation and
- 2D Preconditioning
Multiple source cluster experiment in Marmousi model

Modeling information:

- **Model size:** 3.15 km x 1.08 km
- **Grid spacing:** 5 m
- **Total number of sources:** 7
- **Peak frequency:** 22 Hz, 25 Hz & 30 Hz
- **Receiver spacing:** 10m
- **Receiver depth:** 20m
- **Sampling interval:** 0.5 ms
- **Recording length:** 1 s
- **Free surface:** No
Multiple source cluster experiment in Marmousi model

Modeling information:

- **Model size**: 3.15 km x 1.08 km
- **Grid spacing**: 5 m
- **Total number of sources**: 7
- **Peak frequency**: 22 Hz, 25 Hz & 30 Hz
- **Receiver spacing**: 10m
- **Receiver depth**: 20m
- **Sampling interval**: 0.5 ms
- **Recording length**: 1 s
- **Free surface**: No

- Contaminated with 5 to 45 Hz random noise
- \(\text{SNR} = 3.5\) dB
Estimated location
w/$\mu = 9e-4$ and 10 iterations
Estimated location w/ $\mu = 9e-3$ and 30 iterations
Wavelet comparison

![Wavelet comparison](image)
Wavelet comparison
Wavelet comparison
Wavelet comparison

In chapter 5
- Debiasing step to correct amplitude and
- Detection of microseismic sources from noisy data
Key Contributions: Chapter 6

Sparsity-promoting photo acoustic imaging

- Simultaneous imaging of absorption map and source estimation
- With reduced number of transducers and with smooth velocity model
Photoacoustic Imaging

Pulsed laser excitation → Ultrasonic emission → Ultrasonic detection

Laser/RF pulse → Absorption → Thermal expansion → Acoustic waves → Ultrasonic detection → Image formation
Photoacoustic Imaging

Objectives

- detection of photoabsorbers
- estimation of associated source-time function
Objectives

- detection of photoabsorbers
- estimation of associated source-time function

Challenges

- Time-reversal methods do not estimate source-time function
- Require dense transducer coverage
Solving w/ Linearized Bregman

\[
\text{minimize} \quad \|Q\|_{2,1} + \frac{1}{2\mu} \|Q\|_F^2
\]

subject to \[\|\mathcal{F}[\mathbf{m}](Q) - \mathbf{d}\|_2 \leq \epsilon\]

*where \(T_\tau = \{Q|Q(:,t) = 0, t > \tau\}\)

*where \(\tau\) is the user defined duration
Case Study: Blood vessel phantom

True Image

Estimated Image

Wavelet Comparison
Effect of transducer sampling

**Sparsity-Promoting**

Transducers every 2 degrees

Transducers every 6 degrees

**Time reversal w/ K-wave**
Conclusions

Sparsity promotion based method

- can simultaneously estimate multiple source locations & source-time functions

- can provide locations of fractures by resolving microseismic sources within half a wavelength

- works w/ sources of different frequencies & origin times
Conclusions

Sparsity promotion based method
- can simultaneously estimate multiple source locations & source-time functions
- can provide locations of fractures by resolving microseismic sources within half a wavelength
- works w/ sources of different frequencies & origin times

Dual formulation provides a computationally efficient scheme
Conclusions

Sparsity promotion based method

- can simultaneously estimate multiple source locations & source-time functions
- can provide locations of fractures by resolving microseismic sources within half a wavelength
- works w/ sources of different frequencies & origin times

Dual formulation provides a computationally efficient scheme

Wavelet scaling can be corrected using debiasing method

[Sharan et al.,'18]
Future work

Checkpointing strategy to avoid storage of complete source-wavefield

Inclusion of TV-norm instead of $\ell_1$ norm in space
  - for sources along a plane
  - with sharp boundaries
1. S. Sharan, Y. Zhang, O. Lopez and F.J. Herrmann, 2020, *Large scale high-frequency wavefield reconstruction with recursively weighted matrix factorizations*, Submitted to *Geophysics*


Acknowledgement

Many thanks to

- my advisor Prof Felix J. Herrmann
- PhD advisory committee at Georgia Tech and UBC
- PhD examination committee
- My colleagues at SLIM Lab
- Georgia Research Alliance, NSERC, SINBAD consortium
- Nick Moldoveanu from Schlumberger
- Developers of open source software packages (Devito, JUDI)