Frequency down extrapolation with TV norm minimization

Rongrong Wang & Felix J. Herrmann

University of British Columbia
Outline

1. Motivation and extrapolation workflow
2. Theoretical understanding of L1 minimization
3. TV minimization to enhance stability of L1
4. Numerical results
Motivation

Challenges in FWI:

- High frequency data introduces abundant local minima.
- Field data lacks
  - low frequencies
  - or low frequencies are noisy

Our goal:

use mid-band data to extrapolate towards low frequencies
Convolutional model

Near offset trace

\[ d(t) = w(t) * r(t) \]

Assume

\[ r(t) = \sum_{i=1}^{s} a_i \delta_{t_i}(t) \]

\[ \hat{r}(\omega) = \sum_{i=1}^{s} a_i e^{\pi i t_i \omega} \]

No assumptions on the wavelet
Workflow

Time domain reflectivity

Spectrum

Noise-free data

Wavelet Noise-free data

Spectrum

Time domain reflectivity

Workflow
Workflow

Ideal time trace

Contamination

Observed data
Workflow

Ideal time trace

Contamination

Observed data

Spectrum
Workflow

Observed data

$\mathcal{F}$
Workflow

Observed data

\[ \mathcal{F} \]

select a band \( \Omega \)

\[ \mathcal{F}_\Omega \ast d \]
Workflow

Observed data

\[ F \]

select a band \( \Omega \)

\[ F_\Omega \ast d \]

deconvolution

recover reflectivity series
Workflow

Observed data

\( F \)

select a band \( \Omega \)

\( F_\Omega \ast d \)

deconvolution

recover reflectivity series

\( \ast q \)

recovered spectrum

recovered trace
**Workflow**

Observed data

true spectrum

recovered spectrum

recovered trace

select a band $\Omega$

$\mathcal{F}_\Omega * d$

deconvolution

recover reflectivity series

$\mathcal{F}$

$*q$
Workflow

Observed data

true spectrum

recovered spectrum

recovered trace

select a band $\Omega$

$F \Omega \ast d$

$F$ deconvolution

recover reflectivity series

*\(q\)
Two approaches to deconvolution

- **MUltiple SIgnal Classification (MUSIC)**
  - needs only 2s+1 measurements
  - needs prior information on the number of events
  - has some stability w.r.t. noise

- **L1 minimization (Linear Programming)**
  - has greater stability
  - fits data exactly
  - needs constraint on minimal distance between opposite spikes

---

L1 minimization

\[
\min_r \|r\|_1 \\
\text{subject to } \mathcal{F}_\Omega(w \ast r) = \mathcal{F}_\Omega d
\]

d: data
w: wavelet
r: reflectivity series
\(\mathcal{F}_\Omega\): bandpass filter
\(\Omega = [f_L, f_H]\): pass bands
Prior art

- L1 based deconvolution first appeared in geophysics literature in 1970s [1,2]
- Compressed Sensing provided theoretical support for randomly selected Fourier coefficients [3]
- Candès et al. established a super-resolution theory assuming high frequency is missing [4]
- Dossal et al. introduced the minimal scale condition [5]
Theoretical foundation for low-frequency extrapolation

Theorem [RW,2016] \( r(t) \) can be exactly recovered by L1 minimization if the spikes are separated by 3.5 wavelengths* and the available bandwidth is greater than 60Hz. If noise exists, then the error in the estimate of \( r(t) \) is proportional to the energy of the noise.

* wavelength: \( \lambda_c = \frac{1}{f_H - f_L} \)
Theoretical foundation for low-frequency extrapolation

Theorem [RW,2016] $r(t)$ can be exactly recovered by L1 minimization if the spikes are separated by 3.5 wavelengths* and the available bandwidth is greater than 60Hz. If noise exists, then the error in the estimate of $r(t)$ is proportional to the energy of the noise.

*wavelength: $\lambda_c = \frac{1}{f_H - f_L}$

*From numerical experiments: 1.5 wavelength is sufficient
When the minimal distance condition is not satisfied …
Reconstructions are affected

Left: L1 reconstruction of reflectivity series using frequencies $\Omega = \{10,\ldots,40\} \text{Hz}$

Deconvolution result is independent of wavelet choice
Error is proportional to distance to $\Omega$
It can be shown that

- As distance from $\Omega$ ↑, error ↑
- extrapolation towards low frequencies is more stable than towards high frequencies
Difficulties in extrapolation

- FWI requires high accuracy in both phase and amplitude of low frequency data
- Wavelet estimation is not accurate
- Existence of dispersion
- 2D modeling: reflectivity series do not contain perfect spikes
- Existence of very close spikes at crossing of events

\[
\text{noise} \quad \Rightarrow \quad \text{causes L1 to fail}
\]
The L1 minimization w/ TV-norm stabilizer

Conflicting events generate very close spikes

L1 minimization has trouble in the indicated region

Goal: utilize spatial correlations
Spatial similarity between adjacent traces

Define spatial similarity by

$$S(R) = \frac{\sum_{j=1}^{m-1} (\| R_j \|_2^2 + \| R_{j+1} \|_2^2)}{\sum_{j=1}^{m-1} \| R_j - R_{j+1} \|_2^2}$$

G has no spatial similarity when $S(R_m) \leq 1$

Left: visually continuous but has no spatial similarity
Increase spatial similarity by low-pass filtering

\[ R_m = f \ast R, \quad \text{where} \quad f = (1, \ldots, 1, 0, \ldots, 0) \]

\[ S(R_m) \geq S(R) \]

\[ \mathcal{F}_\Omega(w \ast R_m) = \mathcal{F}_\Omega(w \ast R) \ast \mathcal{F}_\Omega(f) = d \ast \mathcal{F}_\Omega(f) \]
TV norm minimization for Gm

NESTA is used to solve the following optimization problem

\[
\mathbf{R}_{m, \text{est}} = \arg \min_{\mathbf{R}_m} \| \mathbf{R}_m \|_{T^\alpha \! V}^\alpha, \beta,
\]

subject to \( \mathbf{R}_m(\omega) = \mathbf{R}(\omega) \hat{f}(\omega) = \hat{d}_\Omega(\omega) \hat{f}(\omega), \omega \in \Omega, \)

where

\[
\| \mathbf{R}_m \|_{T^\alpha \! V}^\alpha, \beta := \sum_{i,j} \| \nabla_{\alpha, \beta \mathbf{R}_m(i,j)} \|_1
\]

\[
\nabla_{\alpha, \beta \mathbf{R}_m(i,j)} = \begin{bmatrix}
\alpha(\mathbf{R}_m(i,j) - \mathbf{R}_m(i,j + 1)) \\
\beta(\mathbf{R}_m(i,j) - \mathbf{R}_m(i + 1,j))
\end{bmatrix}.
\]
A stylized example

Data

Reflectivity gather

- Three events
- $\Omega = 5$-15Hz
- Receiver spacing: 10m
- Maximum offset: 1km
A stylized example

- True reflectivity gather
- Inverted with L1
- Filtered reflectivity gather
- Inverted with TV
Synthetic data - Non-inversion crime

- IWAVE generated data
- inversion using time harmonics
- 3 frequency sweeps, 40 l-BFGS-iterations for each batch
- 20Hz Ricker wavelet
- source spacing: 0.2km
- receiver spacing: 20m
- maximum offset: 2km
- model size: 3.2 × 6km
- $\Omega = 5-15$Hz
Synthetic data - Non-inversion crime

\( \Omega = 5 - 15 \text{Hz} \)

Direct inversion
- frequency continuation with batches \([5,5.25] \text{Hz}, [5.5,5.75] \text{Hz}, [6,6.25] \text{Hz} \ldots [15,15.25] \text{Hz}\)
- perform the previous step two more sweeps

Inversion with extrapolation
- extrapolation from \(\{5, \ldots, 15\} \text{Hz}\) to \(\{1, \ldots, 5\} \text{Hz}\)
- frequency continuation with batches \([1,1.25] \text{Hz}, [1.5,1.75] \text{Hz}, \ldots [4.5,4.75] \text{Hz}\)
- frequency continuation with batches \([5,5.25] \text{Hz}, [5.5,5.75] \text{Hz}, [6,6.25] \text{Hz} \ldots [15,15.25] \text{Hz}\)
- perform the previous step two more sweeps
The effect of filtering

Shot Record

Shot Record (filtered)
Deconvolution result using 5-15Hz data

Green's function inverted by L1

Green's function (Gm) inverted by TV
Deconvolution result from 5-15Hz data

True shot record

Estimated Green’s function w/ TV
Deconvolution result from 5-15Hz data

True data 0-4Hz

Extrapolated data 0-4Hz
Deconvolution result from 5-15Hz data

True data 0-2Hz

Extrapolated data 0-2Hz
Deconvolution result from 5-15Hz data

True data 3Hz

Extrapolated data 3Hz
Deconvolution result from 5-15Hz data

True data 4Hz

Extrapolated data 4Hz
Initial guess

![Graph showing depth and distance with velocity](image)

- Distance (km): 0, 1, 2, 3, 4, 5, 6
- Depth (km): 0, 0.5, 1, 1.5, 2, 2.5, 3
- Velocity (km/s): 1.5, 2, 2.5, 3, 3.5, 4, 4.5

True vs. Initial graph
Recovery of FWI

1-15Hz true data

direct inversion with 5-15Hz data
Inversion result via FWI

1-15Hz true data

5-15Hz data with extrapolation
Stability test

![Time Domain Graph](image1)

![Frequency Domain Graph](image2)
Inversion result via FWI

1-15Hz data

5-15Hz data with extrapolation
Lq norm minimization: to overcome the minimal distance barrier
Discrete setting

Signal: $x \in \mathbb{R}^N$

Wavelet: $w \in \mathbb{R}^N$

Bands in use: $\Omega = \{m_1, ..., m_2\}$

Data: $y = F_\Omega^* F_\Omega (w * x) \in \mathbb{C}^N$

discrete filter
The $L^q$ norm

$$\| x \|_q = \left( x_1^q + x_2^q + \cdots + x_n^q \right)^{1/q} \quad q \leq 1$$

$L^q$ unit ball

$q \to 0$

more spiky, less convex
Lq objective

Lq norm minimization

$$\min_r \|r\|_q \quad 0 < q < 1$$

subject to $F_\Omega(w \ast r) = F_\Omega d$

$d$: data
$w$: wavelet
$r$: reflectivity series
$F_\Omega$: DFT matrix restricted to $\Omega = [m_1, m_2]$
Lq objective

Lq norm minimization

\[
\min_r \|r\|_q \quad 0 < q < 1
\]

subject to \( F_{\Omega}(w * r) = F_{\Omega}d \)

Exact recovery is guaranteed if

\[
q \leq \frac{C}{s \log(N/(m_2 - m_1))}
\]

The algorithm is less stable as q gets smaller
L1+Lq

Weighted Lq

\[
\min_{\mathbf{r}} \|\mathbf{r}\|_1 + \lambda \|\mathbf{r}\|_q^q \\
\text{subject to } F_\Omega(\mathbf{w} \ast \mathbf{r}) = F_\Omega \mathbf{d} \quad 0 < q < 1
\]

Theoretical analysis suggested choice

\[
\lambda \in \left[ c_q \frac{\|x\|_1}{\|x\|_q^q}, C_q \frac{\|x\|_1}{\|x\|_q^q} \right]
\]

Approximate recovery guaranteed by

\[
q \leq \frac{C}{\log N} \quad \text{and} \quad s \leq c(m_2 - m_1)
\]
Solver: Reweighted Least Squares

To solve
\[ \min_x \|x\|_q^q \quad \text{subject to } Ax = b \]

Write it as
\[ \min_x \|x^{q-1} \odot x\|_1 \quad \text{subject to } Ax = b \]

RLS iterates as
\[ x_{k+1} = \min_x \|x_k^{q-1} \odot x\|_1 \quad \text{subject to } Ax = b \]
Solver: Reweighted Least Squares

\[ x_{k+1} = \min_x \| x_k^{q-1} \odot x \|_1 \]

subject to \( Ax = b \)

can be written more formally as

\[ x_{k+1} = \min_x \| a_k \odot x \|_1 \]

subject to \( Ax = b \)

\[ a_{k+1} = (|x_k| + \epsilon)^{q-1} \]
RLS for L1+Lq

Initialize weights: \( a_0 = [1, ..., 1] \)

For \( k = 1, ..., K \):

Define weight: \( a_k = [a_k(1), ..., a_k(n)], \ a_k(i) = \frac{1}{\epsilon + |r_{k-1}(i)|^{1-p}} \)

Update \( r_k \): \( r_{k+1} = \min_r \|r\|_1 + \lambda \|a_k \otimes r\|_1 \)

subject to \( F_\Omega(w * r) = F_\Omega d \)

End
Why RLS does not work for our case?

\[ N = 1000 \]
\[ \Omega = [m_1, m_2] = [10, 50] \]

r1: estimated reflectivity series after the first step of RLS
Why RLS does not work for our case?

---

Why RLS does not work for our case?
Why RLS does not work for our case?

Weights promotes the previously detected wrong supports
Solution

The true support is within 1.5 wavelength of the nonzeros of $r_1$
Modified Reweighted Least Squares

For $l = 1, \ldots, L$:

\[ \text{size}(\mathcal{N}_i) = \frac{1.5N}{m_1 - m_2} \cdot \frac{1}{l}, \quad \text{for } i = 1, \ldots, N \]

For $k = 1, \ldots, K$:

Define weight: $a_k = [a_k(1), \ldots, a_k(n)]$, 

\[ a_k(i) = \frac{1}{\epsilon + \min_{j \in \mathcal{N}_i} |r_k(j)|^{1-p}} \]

Update $r_k$: 

\[ r_{k+1} = \min_r \|r\|_1 + \lambda \|a_k \odot r\|_1 \]

subject to $F_\Omega(w \ast r) = F_\Omega d$

End

End
Numerical test

Clean data

Extrapolation from \{10,..,50\}Hz towards \{1,...,10\}Hz

Each point is the mean error on 50 random draws of sparse signals

N=1000, m1=10, m2=50
Numerical test

SNR=20

Extrapolation from \{10,\ldots,50\}Hz towards \{1,\ldots,10\}Hz

Each point is the mean error on 50 random draws of sparse signals

N=1000, m1=10, m2=50
Numerical test

SNR=13

Extrapolation from \{10,\ldots,50\}Hz towards \{1,\ldots,10\}Hz

Each point is the mean error on 50 random draws of sparse signals

N=1000, m1=10, m2=50
Synthetic data - Non-inversion crime

- IWAVE generated data
- Inversion using time harmonics
- 3 frequency sweeps, 20 lbfgs-ITers for each batch
- 20Hz Ricker wavelet
- Source spacing: 0.2km
- Receiver spacing: 20m
- Maximum offset: 1km
- Model size: 1.8km × 4.5km
Recovery of FWI

true model

5-15Hz data w/o extrapolation

(distance (km)

(depth (km))

(km/s)
Recovery of FWI

true model

5-15Hz data Lq extrapolation $q=0.1$
Recovery of FWI

true model

5-15Hz data TV extrapolation
Cross-sections for the inverted model by Lq
Conclusions

• We provide theoretical understanding of the classical L1 minimization.
• We use the L1 based approach for low frequency extrapolation.
• We used TV norm regularization to enhance stability of the classical L1 minimization.
• The efficacy of the method is still limited by the minimal distance of opposite events.
• The method is stable with respect to additive noise and dispersion.
Acknowledgements

This research was carried out as part of the SINBAD project with the support of the member organizations of the SINBAD Consortium.