A linearized Bregman method for sparsity-promoting full-waveform inversion

Xintao Chai[1, 2], Mengmeng Yang[1], Philip A. Witte[1], Rongrong Wang[1], Zhilong Fang[1], and Felix J. Herrmann[1]

[1]Seismic Laboratory for Imaging and Modeling (SLIM), University of British Columbia (UBC), Vancouver, Canada.

[2]China University of Petroleum (Beijing).

Outline

- *Introduction*
- Theory
- Numerical experiments
- Conclusions

Introduction

FWI can be formulated as a least-squares (LS) problem

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) := \frac{1}{2}\left\|\mathbf{D} - \mathbf{F}(\mathbf{m},\mathbf{S})\right\|_2^2$$

- **F** : wave-equation based modeling operator
- **m**: model
- **S :** source
- **D :** observed data

◆The Gauss-Newton (GN) method (Pratt et al., 1998) solves the FWI problem without explicit computation of the Hessian.

◆Each iteration for GN subproblem requires a large number of PDE solves.

◆To reduce the computational cost, source encoding or subsampling…

◆Dimension reduction introduces source crosstalks.

Introduction

Compressed sensing theory:
If the signals' energy is concentrated in a few large coefficients, these signals can be reconstructed from small amounts of data by solving sparsity-promoting problems.

Li et al. (2012) presented a modified GNFWI method,

$$\min_{\mathbf{x}} \frac{1}{2}\left\|\underline{\delta\mathbf{d}} - \nabla\mathbf{F}\left(\mathbf{m}_0, \underline{\mathbf{S}}\right)\mathbf{C}^*\mathbf{x}\right\|_2^2 \quad \text{subject to} \left\|\mathbf{x}\right\|_1 \le \tau$$

Li et al. (2012) solve the above problem with a spectral-gradient method, SPGℓ_1.

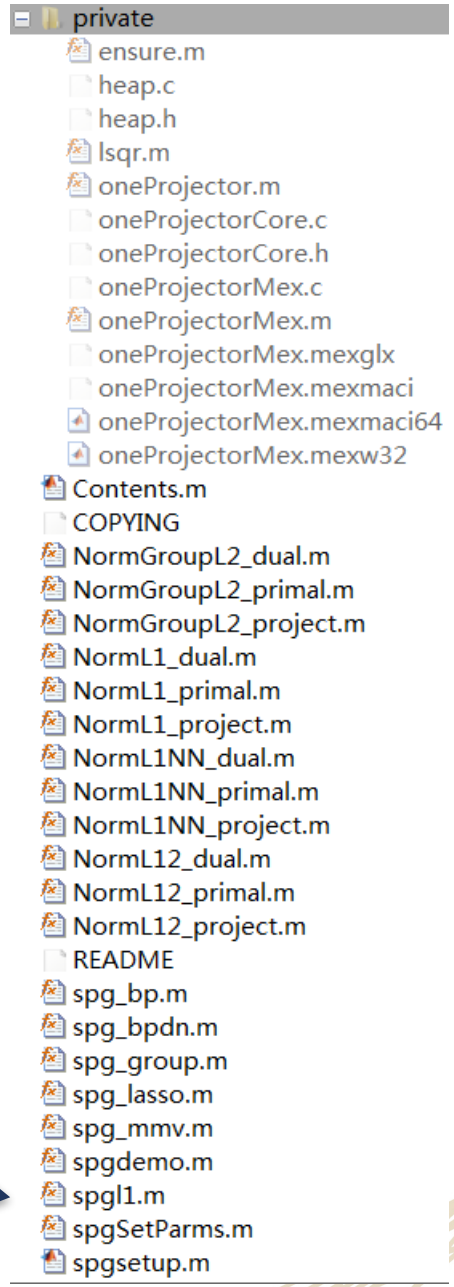
The complexity of SPGℓ_1 hinders its industry implementation.

Complexity of SPGl_1

SPGℓ_1: Multi parameters

```
fprintf(' Default parameters for l1Set.m:\n');
fprintf('        fid : [ positive integer        |     1 ]\n');
fprintf('  verbosity : [ integer: 1, 2, or 3     |     3 ]\n');
fprintf(' iterations : [ positive integer        |  10*m ]\n');
fprintf('  nPrevVals : [ positive integer        |    10 ]\n');
fprintf('      bpTol : [ positive scalar         | 1e-06 ]\n');
fprintf('      lsTol : [ positive scalar         | 1e-06 ]\n');
fprintf('     optTol : [ positive scalar         | 1e-04 ]\n');
fprintf('     decTol : [ positive scalar         | 1e-04 ]\n');
fprintf('    stepMin : [ positive scalar         | 1e-16 ]\n');
fprintf('    stepMax : [ positive scalar         | 1e+05 ]\n');
fprintf' rootMethod : [ 1=linear, 2=quadratic    |     2 ]\n');
fprintf('activeSetIt : [ positive integer        |   Inf ]\n');
fprintf('subspaceMin : [ 0=no, 1=yes             |     0 ]\n');
fprintf('  iscomplex : [ 0=no, 1=yes, NaN=auto   |   NaN ]\n');
fprintf('  maxMatvec : [ positive integer        |   Inf ]\n');
fprintf('    weights : [ vector                  |     1 ]\n');
fprintf('    project : [ projection function     |    @()]\n');
fprintf('primal_norm : [ primal norm eval fun    |    @()]\n');
fprintf('  dual_norm : [ dual norm eval fun      |    @()]\n');
```

SPGℓ_1: Hundreds of lines of code



Outline

- Introduction
- *Theory*
- Numerical experiments
- Conclusions

Theory

A simple linearized Bregman method (Yin, 2010; Lorenz et al., 2013) for solving the large-scale sparsity-promoting GNFWI subproblem.

The linearized Bregman method solves the regularized BPDN problem

$$\min_{\mathbf{x}} \lambda \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \quad \text{subject to} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \leq \sigma^2$$

Simplicity of the linearized Bregman (LB) method

Algorithm 1 Linearized Bregman iterations for CS

1. **Input data and parameters:**
   $\mathbf{A}$, $\mathbf{b}$, threshold λ, iteration limit $k^{\max}$, noise factor σ
2. Initialize iteration index $k = 0$, $\mathbf{x}_k = \mathbf{g}_k = \mathbf{0}$
3. **while** stopping criteria not satisfied **do**
4. $\quad \mathbf{g}_{k+1} = \mathbf{g}_k - t_k \mathbf{A}^* \Pi_\sigma(\mathbf{A}\mathbf{x}_k - \mathbf{b})$
5. $\quad \mathbf{x}_{k+1} = \mathrm{shrink}(\mathbf{g}_{k+1}, \lambda)$
6. $\quad k \leftarrow k + 1$
7. **end while**
8. **Output:** Sparse solution $\mathbf{x}$

Dynamic step-size t

$$t_{k_i} = \left\| \mathbf{A}_{k_i} \mathbf{x}_{k_i} - \mathbf{b}_{k_i} \right\|_2^2 / \left\| \mathbf{A}_{k_i}^* \left(\mathbf{A}_{k_i} \mathbf{x}_{k_i} - \mathbf{b}_{k_i} \right) \right\|_2^2$$

Soft shrinkage function

$$\mathrm{shrink}(x, \lambda) = \max(|x| - \lambda, 0)\ \mathrm{sign}(x)$$

Algorithm 2 A linearized Bregman approach for sparsity-promoting FWI

1. **Input data:** $\mathbf{d}_{\text{obs}}$, source wavelet, initial model $\mathbf{m}_0$
2. **Inversion parameters:**
   number of source-encoding experiments, threshold λ
   noise factor σ, outer iteration limit $k^{\max}_{\text{outerloop}}$,
   inner iteration limit $k^{\max}_{\text{innerloop}}$
3. **while** stopping criteria not satisfied **do** //main loop
4. Initialize outer loop iteration index $k_{\text{o}} = 0$
5. **while** $k_{\text{o}} < k^{\max}_{\text{outerloop}}$ **do** //outer loop
6. Initialize inner loop iteration index $k_{\text{i}} = 0$,
   $\mathbf{x}_{k_{\text{i}}} = \mathbf{g}_{k_{\text{i}}} = \mathbf{0}$, $\delta\mathbf{d} = \mathbf{d}_{\text{obs}} - \boldsymbol{\mathcal{F}}(\mathbf{m}_0)$
7. **while** $k_{\text{i}} < k^{\max}_{\text{innerloop}}$ **do** //inner loop
8. new random draw source-encoding operator,
   encoding and subsampling the measurements
   $\mathbf{A}_{k_{\text{i}}} = \underline{\mathbf{J}(\mathbf{m}_0)}\mathcal{S}^*$, $\mathbf{b}_{k_{\text{i}}} = \underline{\delta\mathbf{d}}$
9. $\mathbf{g}_{k_{\text{i}}+1} = \mathbf{g}_{k_{\text{i}}} - s_{k_{\text{i}}}\mathbf{A}^*_{k_{\text{i}}}\Pi_\sigma(\mathbf{A}_{k_{\text{i}}}\mathbf{x}_{k_{\text{i}}} - \mathbf{b}_{k_{\text{i}}})$
10. $\mathbf{x}_{k_{\text{i}}+1} = \text{shrink}(\mathbf{g}_{k_{\text{i}}+1}, \lambda)$
11. $k_{\text{i}} \leftarrow k_{\text{i}} + 1$
12. **end while** //end inner loop
13. $\delta\mathbf{m} = \mathcal{S}^*\mathbf{x}$, $\mathbf{m}_0 \leftarrow \mathbf{m}_0 + \delta\mathbf{m}$, $k_{\text{o}} \leftarrow k_{\text{o}} + 1$
14. **end while** //end outer loop
15. **end while** //end main loop
16. **Output:** the inverted model

the linearized Bregman iterations

Outline

- Introduction
- Theory
- *Numerical experiments*
- Conclusions

To demonstrate the superiority of sparsity-promoting by ℓ_1-norm,
we also use the LSQR algorithm to solve a LS problem to get the model update.

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{x}\|_2^2 \quad \text{subject to} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2 \le \sigma^2$$

To evaluate the quality of the inverted model,
(1) Relative least-squares error (RLSE)

$$\mathrm{RLSE} = \|\mathbf{m}_{\mathrm{inv}} - \mathbf{m}_{\mathrm{true}}\|_2^2 \,/\, \|\mathbf{m}_{\mathrm{true}}\|_2^2$$

The smaller the RLSE, the better the results.

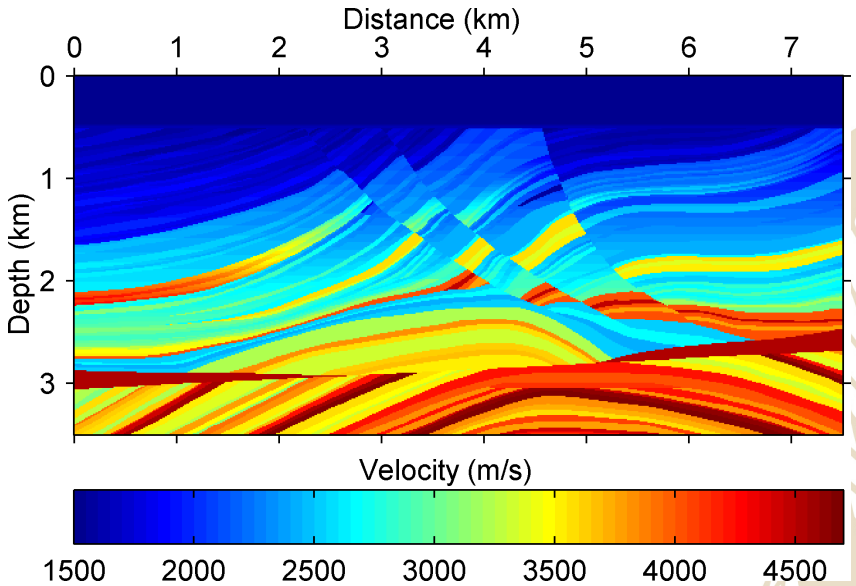
(2) SNR

$$\mathrm{SNR(dB)} = -20\log_{10}\frac{\|\mathbf{m}_{\mathrm{true}} - \mathbf{m}_{\mathrm{inv}}\|_2}{\|\mathbf{m}_{\mathrm{true}}\|_2}$$

Marmousi II model

Experiment setup

- model size: 281×601
- grid spacing: $d_x = d_z = 12.5$ m
- source: 10 Hz Ricker wavelet, phase shift of 0.1 s
- 26 frequencies in the range [3 15.5] Hz
- 301 shots and 601 receivers



Inversion settings

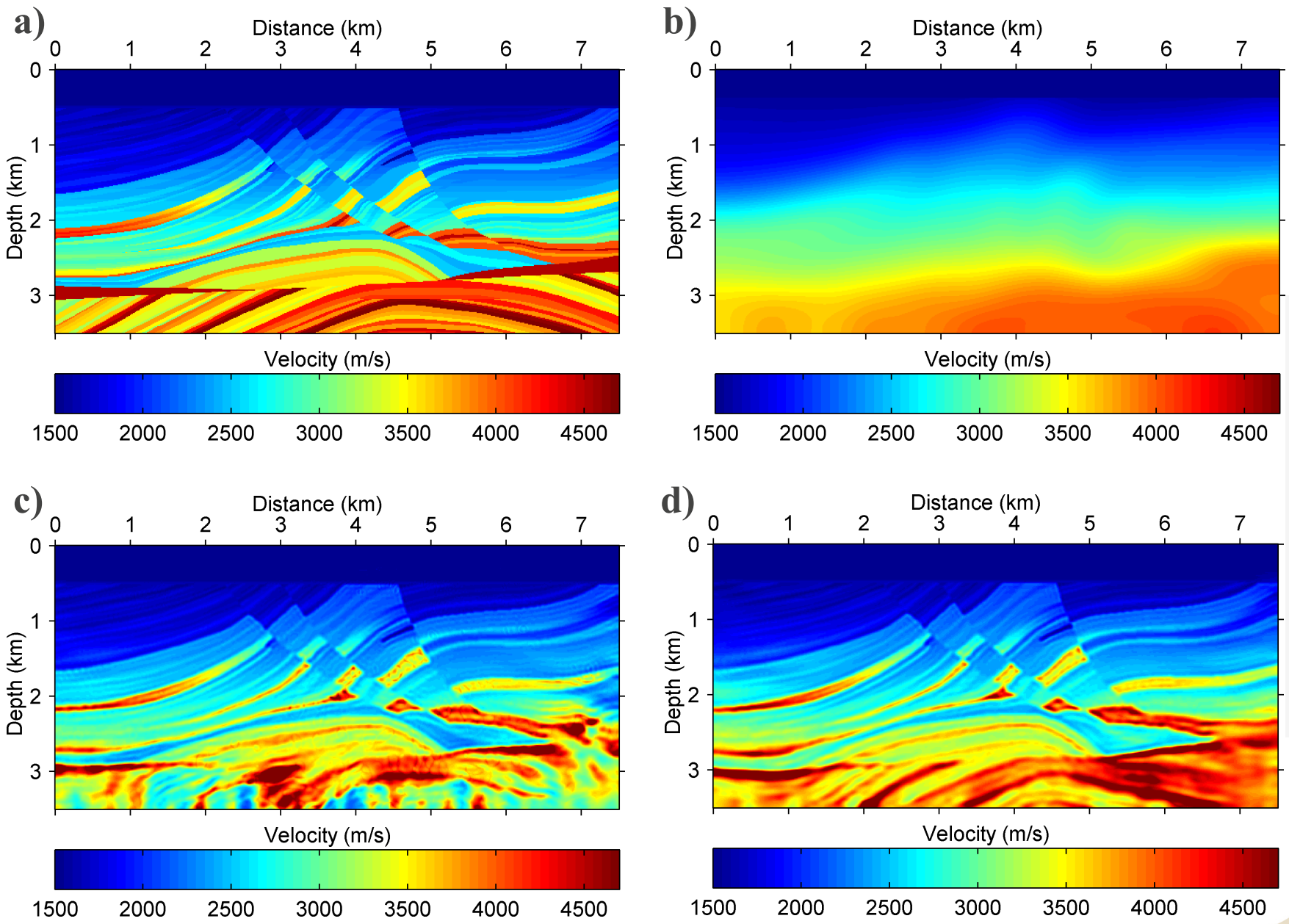
✓6 frequencies

✓10 simultaneous shots

✓moving from low to high frequencies in overlapping batches of 3

✓10 GN outer iterations for each frequency batch

✓each GN subproblem, 20 inner iterations of LSQR and LB



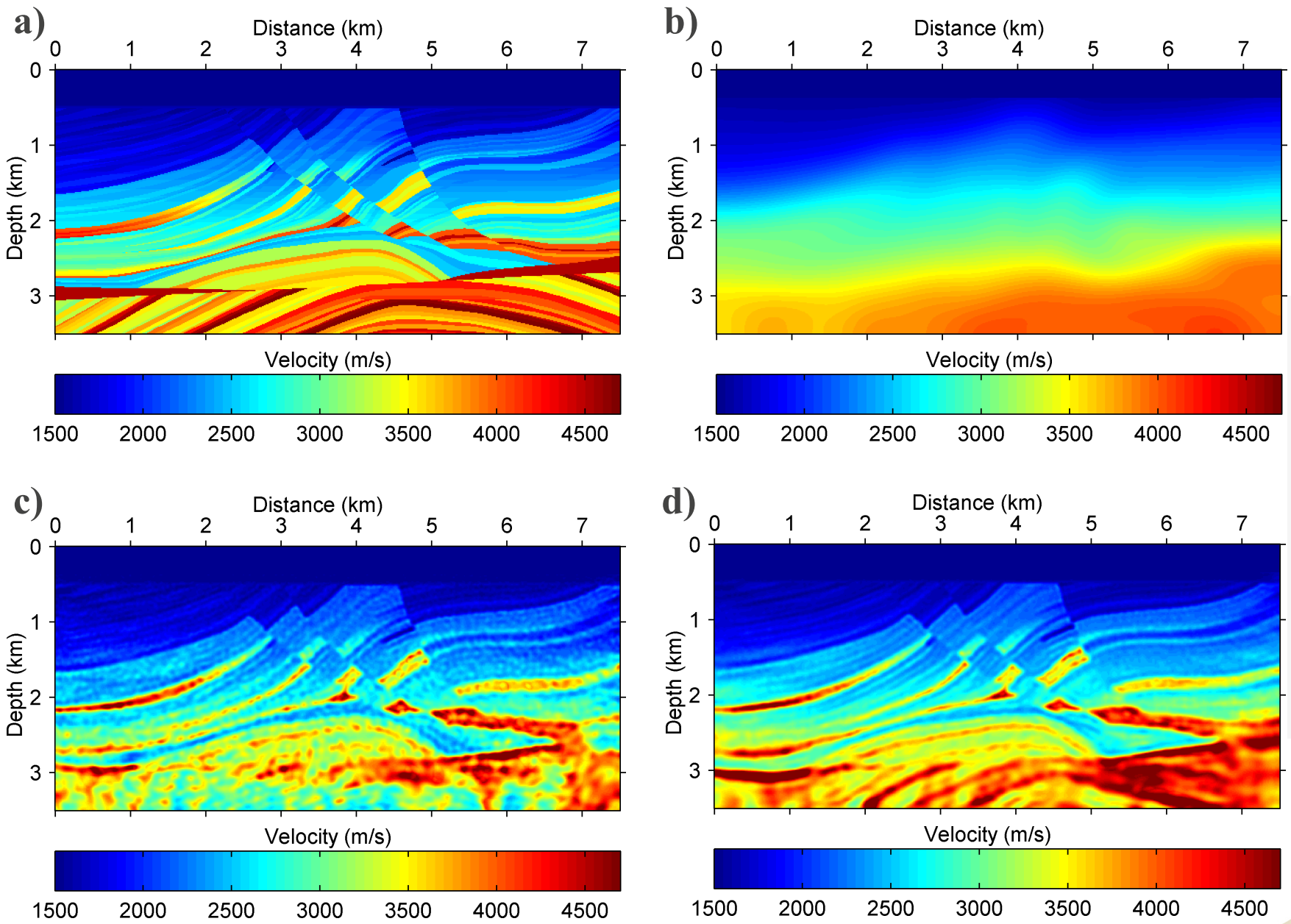
(a) True model

(b) Initial model

(c) Inverted by LSQR, SNR = 17.5233 (dB)

(d) Inverted by LB, SNR = 18.8012 (dB)

Results for noise-free data.



(a) True model

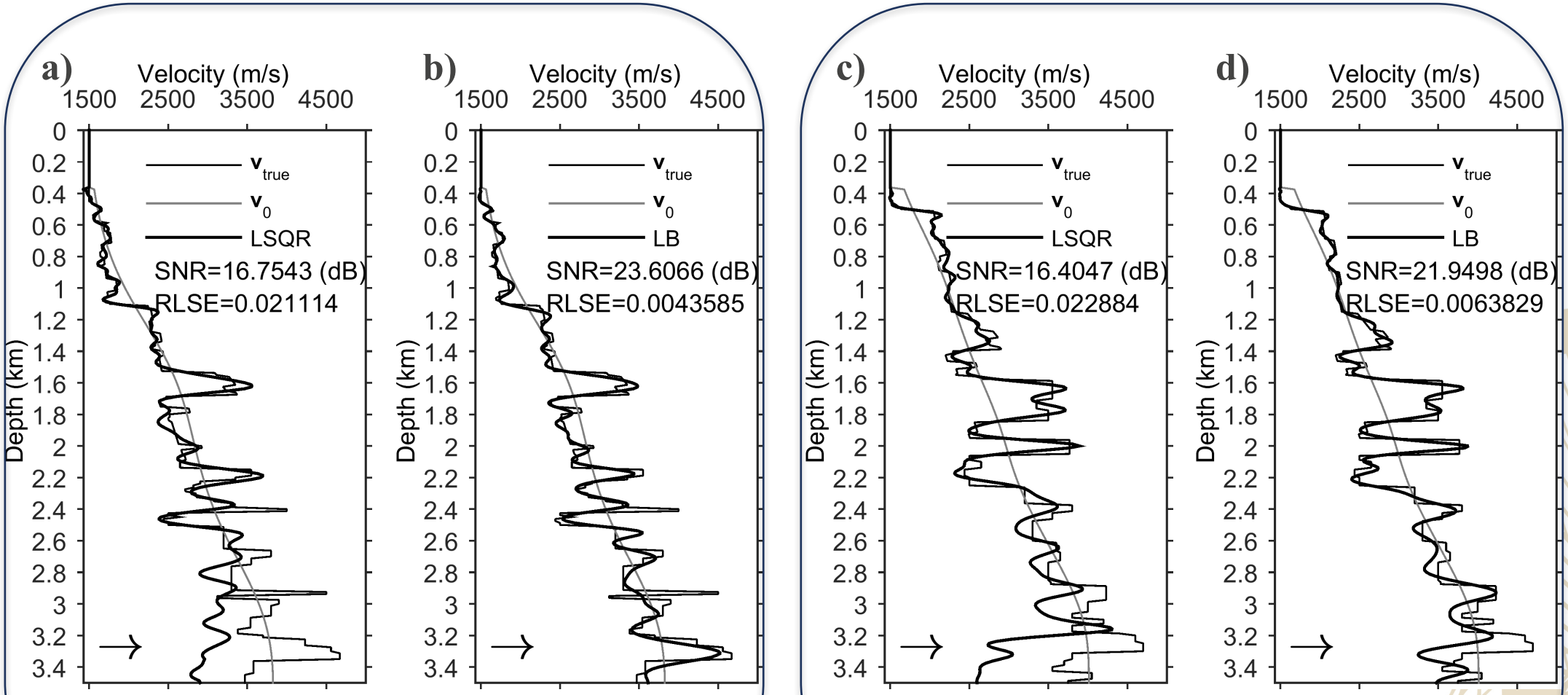
(b) Initial model

(c) Inverted by LSQR, SNR = 17.2728 (dB)

(d) Inverted by LB, SNR = 18.3811 (dB)

Results for 20% random noise added data.

Well data comparison for noisy data set



Well comparison at 2.25 km for (a) LSQR and (b) LB.

Well comparison at 3.95 km for (c) LSQR and (d) LB.

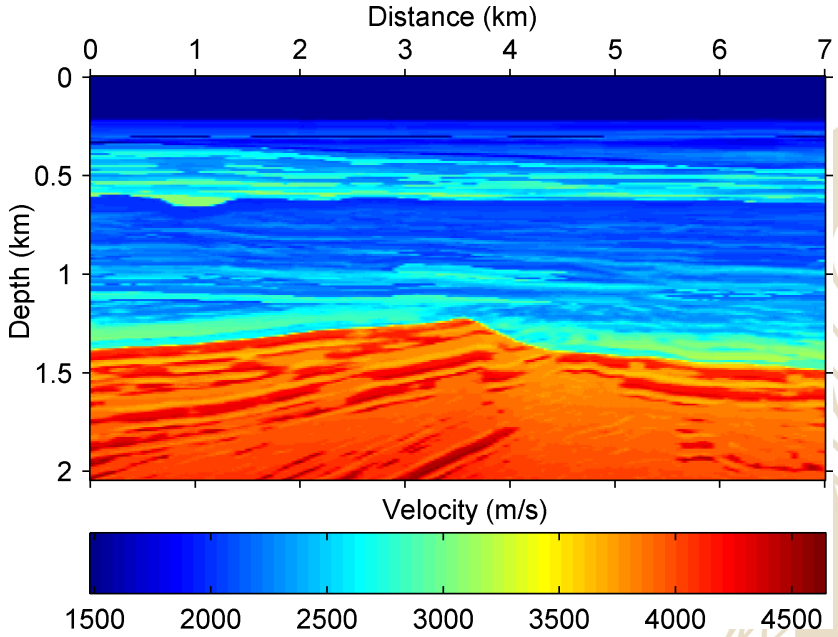
Thin black lines: True model
Thin gray lines: Initial model
Bold black lines: Inverted model

BG Compass model

Experiment setup

➢Same input data set and inversion parameter settings as Li et al. (2012)

- model size: 205×701
- grid spacing: $d_x = d_z = 10$ m
- 58 frequencies in the range [2.9 22.5] Hz
- 351 shots and 701 receivers



Inversion settings

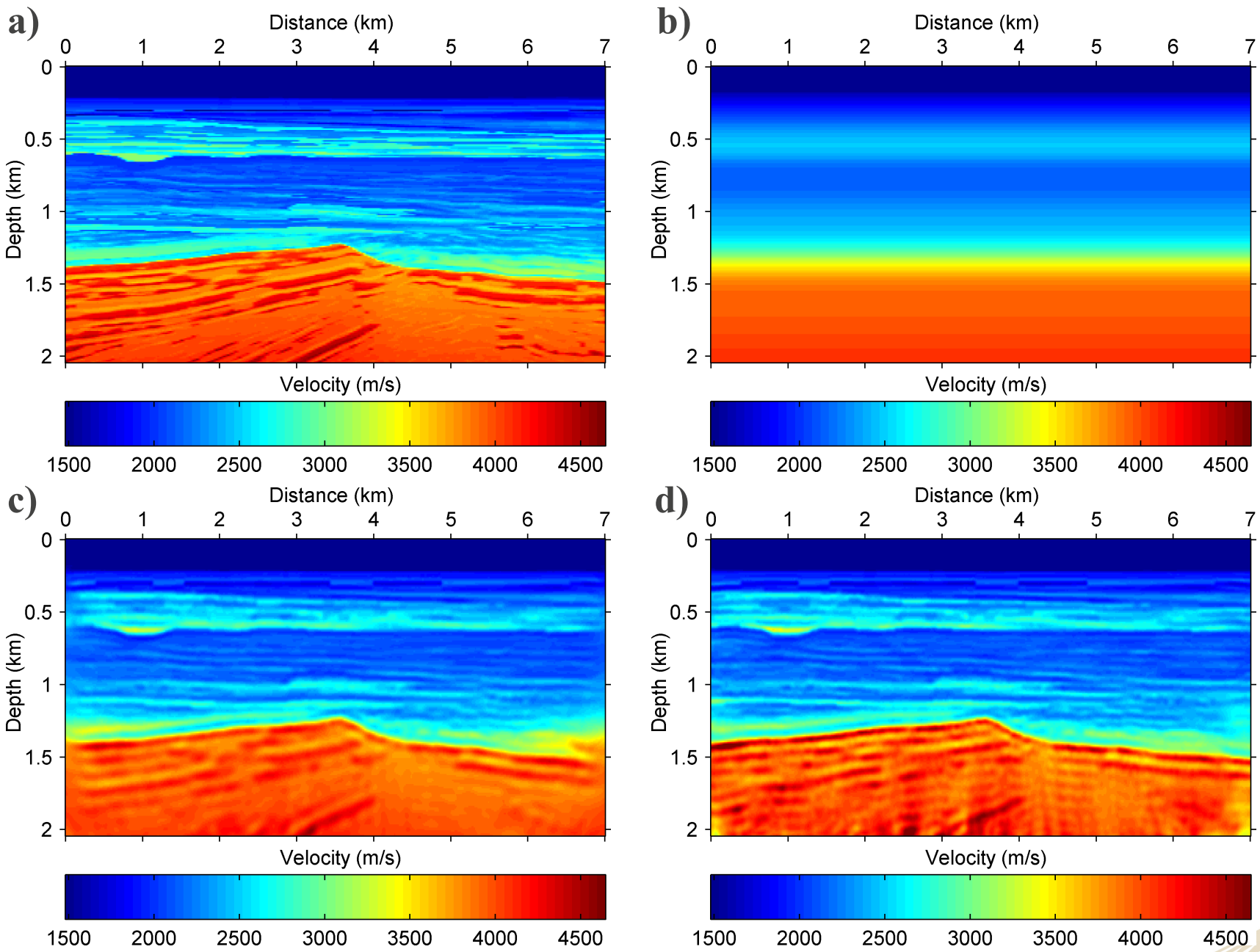
✓10 overlapping frequency bands

✓7 simultaneous shots

✓10 selected frequencies

✓10 GN iterations for each frequency batch

✓each GN subproblem, 20 inner iterations of SPGl_1 and LB



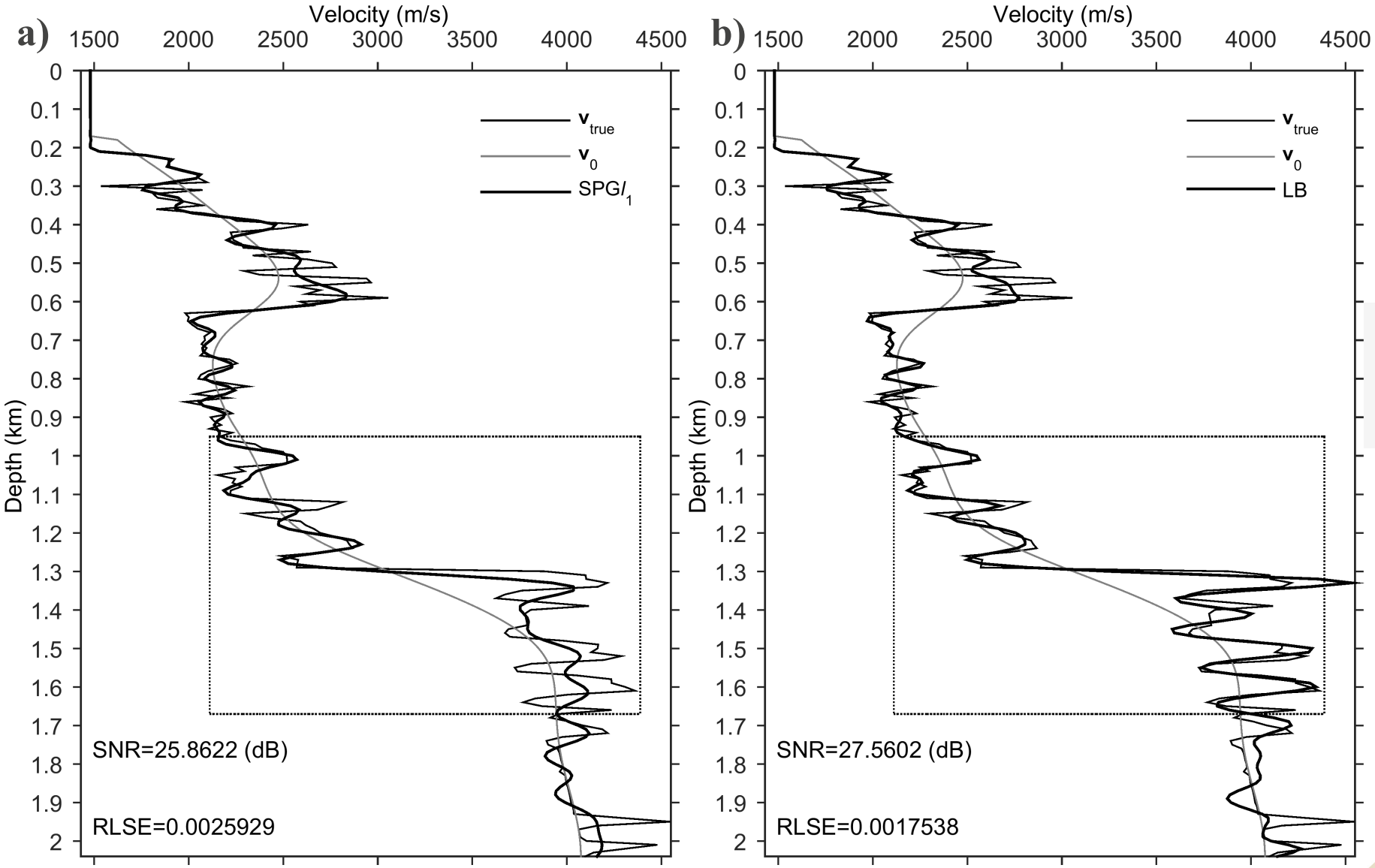
(a) True model

(b) Initial model

(c) Inverted by SPGl_1

(d) Inverted by LB

BG compass model: Well data comparison



Thin black lines: True model
Thin gray lines: Initial model
Bold black lines: Inverted model

Well comparison at 2 km (a) SPGl_1 and (b) LB.

Outline

- Introduction
- Theory
- Numerical experiments
- *Conclusions*

Conclusions

➢A simple linearized Bregman method for sparsity-promoting GNFWI.

1. the much simpler linearized Bregman method does an excellent job compared to the complicated SPGℓ_1;

2. the linearized Bregman method is more efficient than SPGℓ_1;

3. GNFWI result with l_1-norm constraint on the updates better than with l_2-norm constraint on the updates.

Acknowledgments

This work was financially supported in part by the Natural Sciences and Engineering Research Council of Canada Collaborative Research and Development Grant DNOISE II (CDRP J 375142-08). This research was carried out as part of the SINBAD II project with the support of the member organizations of the SINBAD Consortium. We are grateful to Charles Jones from BG Group for providing us with the BG Compass velocity model.

Selected references

1. Herrmann, F. J., N. Tu, and E. Esser, 2015, Fast “online” migration with Compressive Sensing: EAGE Annual Conference Proceedings, doi: 10.3997/2214-4609.201412942.

2. Li, X., A. Y. Aravkin, T. van Leeuwen, and F. J. Herrmann, 2012, Fast randomized full-waveform inversion with compressive sensing: Geophysics, 77, no. 3, A13–A17, doi: 10.1190/GEO2011-0410.1.

3. Li, X., E. Esser, and F. J. Herrmann, 2016, Modified Gauss-Newton full-waveform inversion explained–-why sparsity-promoting updates do matter: Geophysics.

4. Lorenz, D. A., S. Wenger, F. Schöpfer, and M. Magnor, 2014, A sparse Kaczmarz solver and a linearized Bregman method for online compressed sensing.

5. Pratt, R. G., C. Shin, and G. J. Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, 133, 341-362, doi: 10.1046/j.1365-246X.1998.00498.x.

6. van den Berg, E., and M. P. Friedlander, 2008, Probing the Pareto frontier for basis pursuit solutions: SIAM Journal on Scientific Computing, 31, 890–912, doi: 10.1137/080714488.

SEG
INTERNATIONAL EXPOSITION AND 86TH
ANNUAL MEETING
DALLAS 2016

Thank you.