Time domain sparsity promoting LSRTM with source estimation

Mengmeng Yang, Philipp Witte, Zhilong Fang & Felix J. Herrmann

SLIM
University of British Columbia

Motivation

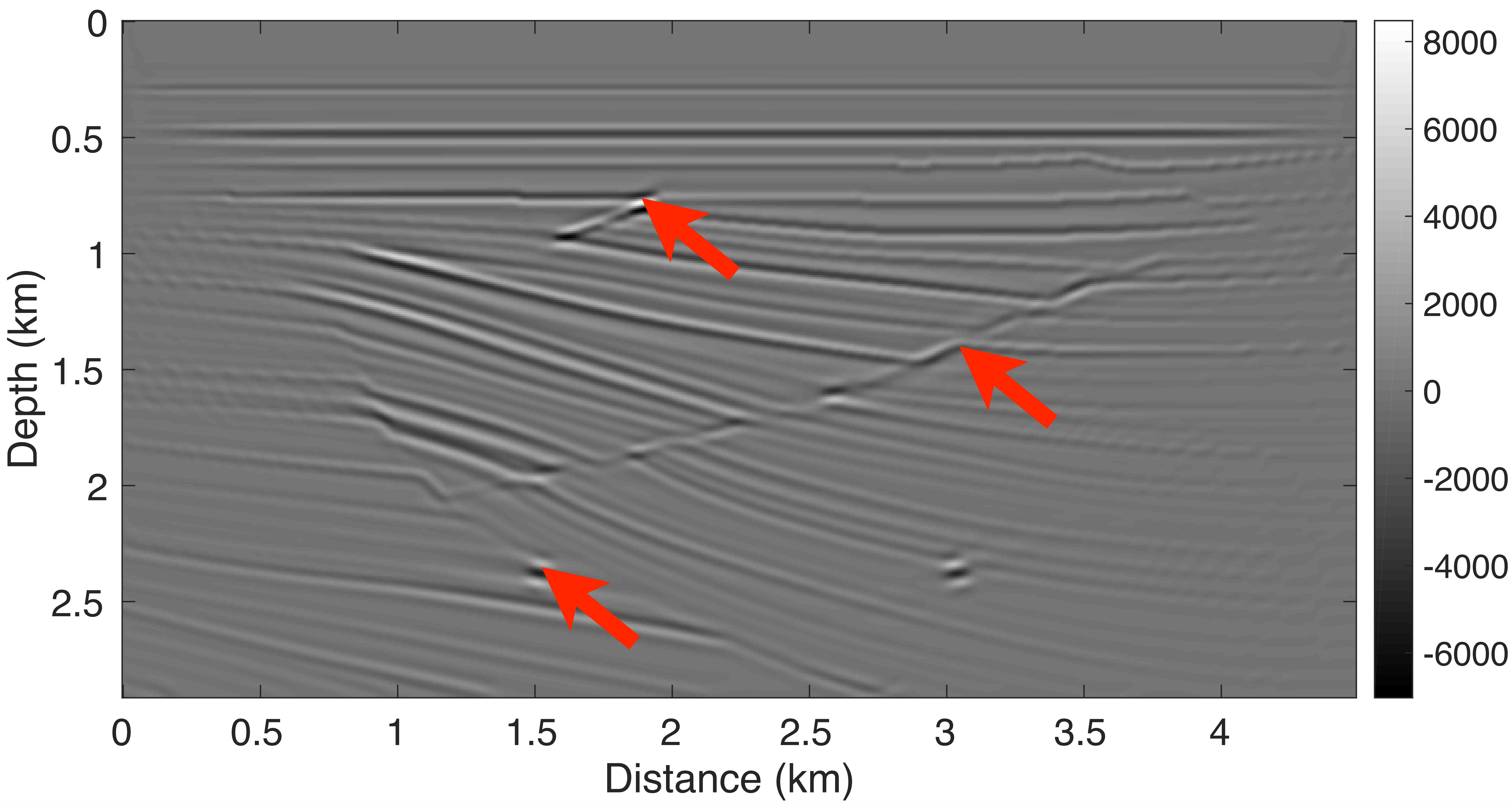
Features of RTM:

- pros
  - no dip limitation
  - strong lateral velocity variations
- cons
  - inaccurate amplitudes & low resolution

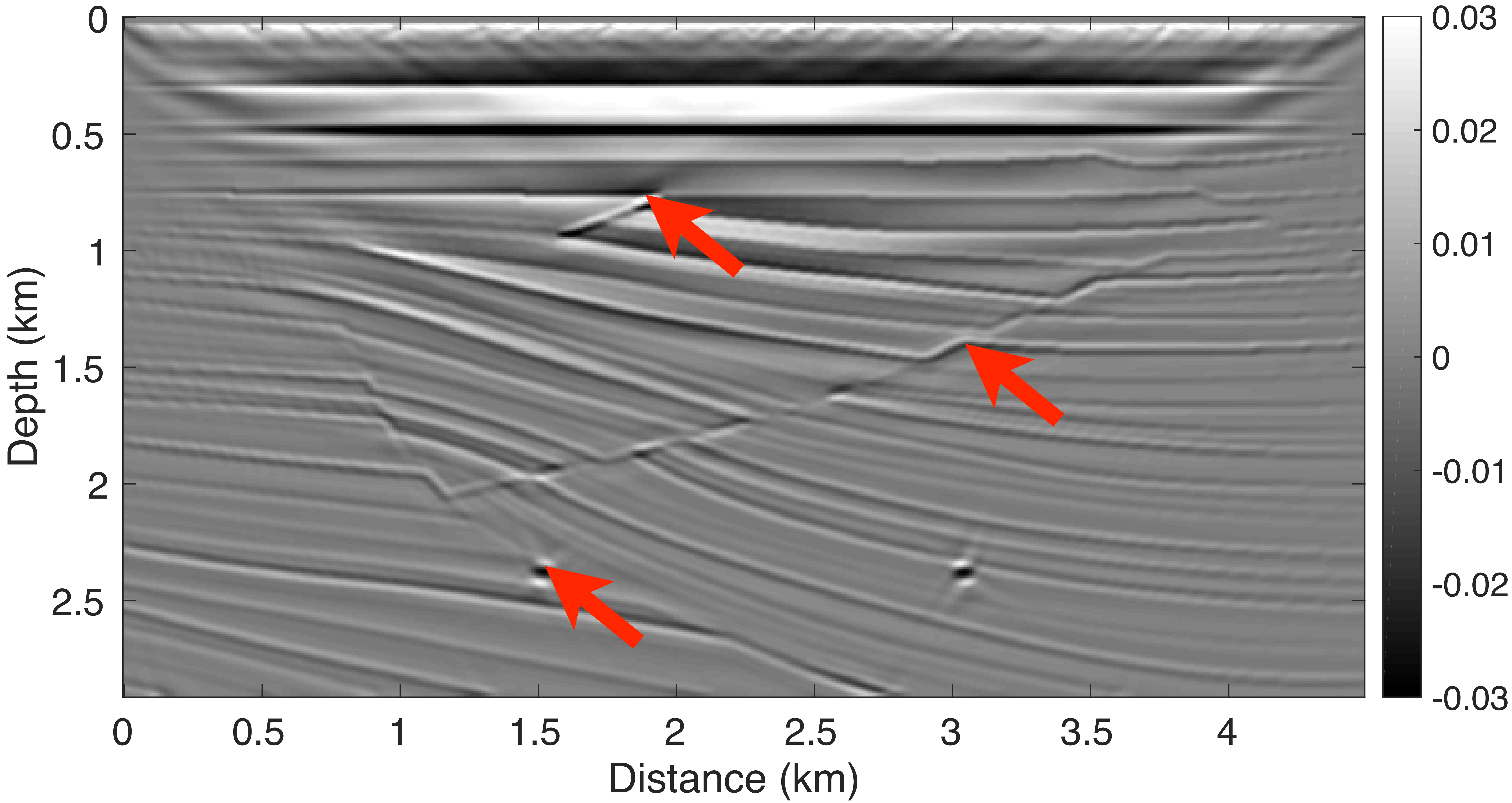
Problems of LS-RTM:

- iterations that touch all shots are too expensive
- data can be overfitted

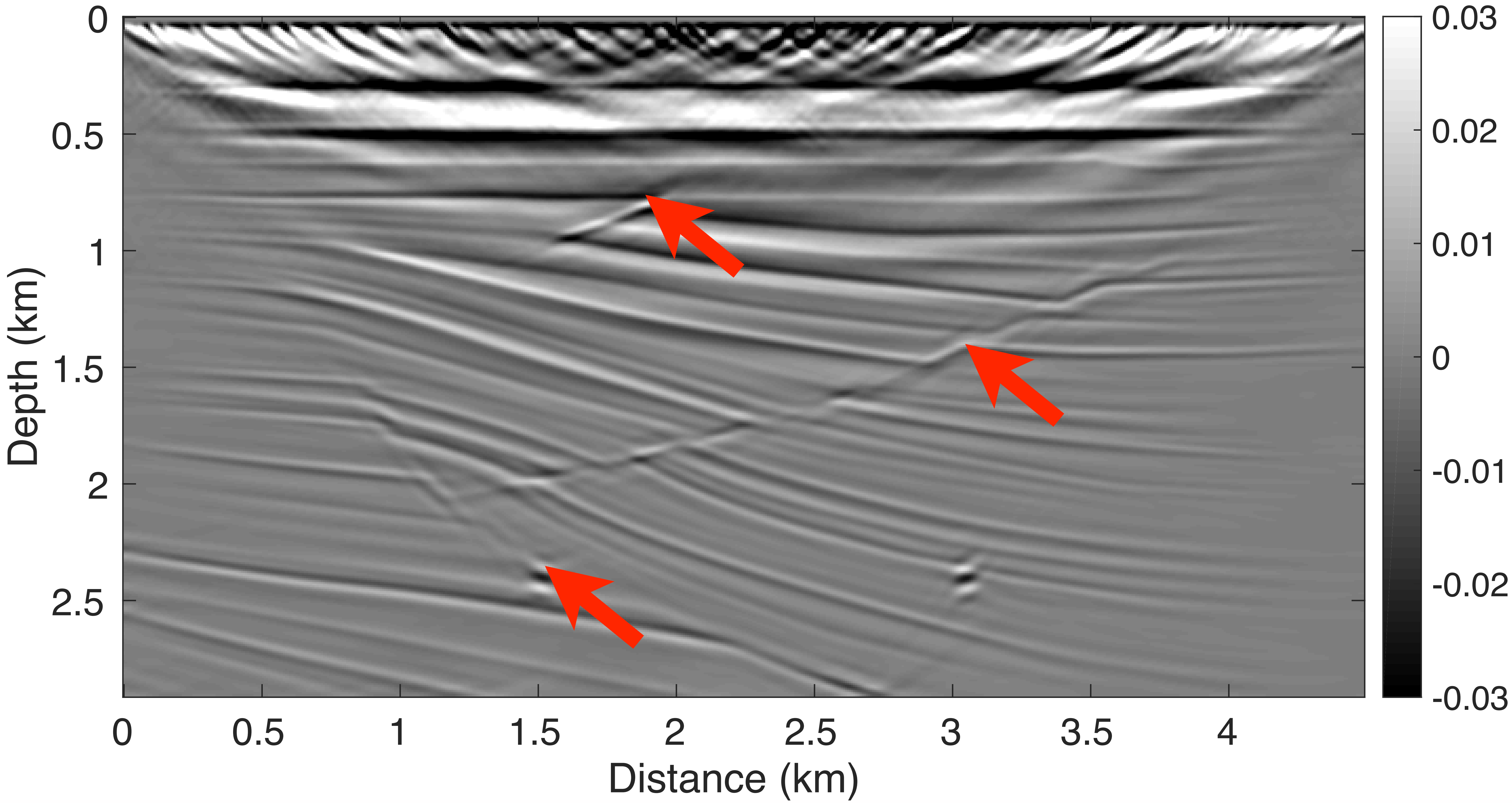
RTM w/ correct wavelet



Sparsity promoting LS-RTM w/ correct wavelet



Sparsity promoting LS-RTM w/ wrong wavelet



SLIM

LS-RTM

$$\min_{\delta \mathbf{m}} \sum_{i=1}^{n_s} \|\mathbf{J}_i[\mathbf{m}_0, \mathbf{q}_i]\delta \mathbf{m} - \mathbf{b}_i\|^2$$

$\mathbf{m}_0$: background model

$\mathbf{J}_i$: Born modelling operator for i^{th} shot

$\delta \mathbf{m}$: model perturbation

$\mathbf{q}_i$: source wavelet for i^{th} shot

$\mathbf{b}_i$: vectorized reflections for i^{th} shot

SLIM

Sparsity promoting inversion

$$\min_{\mathbf{x}} \ \|\mathbf{x}\|_1$$

$$\text{s.t.} \ \sum_{i=1}^{ns} \| \underbrace{\mathbf{J}_i[\mathbf{m}_0, \mathbf{q}_i]\mathbf{C}^*}_{\hat{\mathbf{J}}} \mathbf{x} - \underbrace{\mathbf{b}_i}_{\mathbf{b}} \|_2 \leq \sigma$$

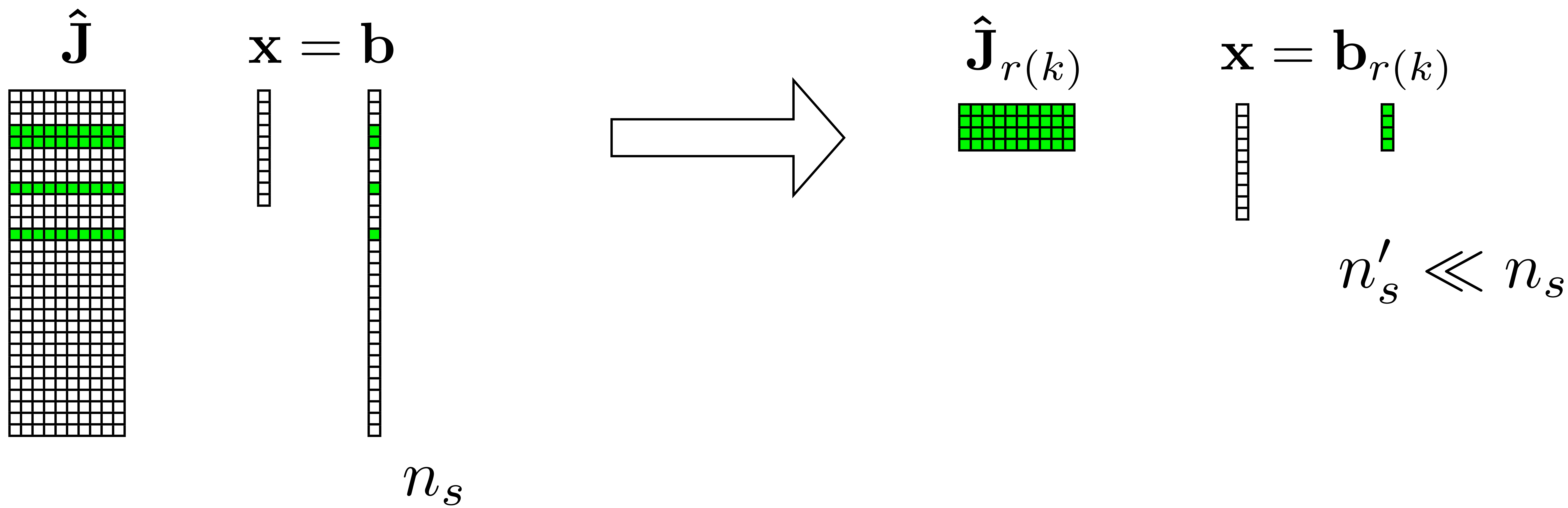
$\mathbf{C}^*$: the transpose of Curvelet transform

$\mathbf{x}$: Curvelet coefficients

σ : tolerance for noise or modelling error

Felix J. Herrmann,Ning Tu and Ernie Esser, “Fast ‘online’ migration with Compressive Sensing”, EAGE Annual Conference Proceeding, 2015, vol. 60, p. 696-712, 2012
Lorenz, Dirk A.; Wenger, Stephan; A sparse Kaczmarz solver and a linearized Bregman method for online compressed sensing. arXiv:1403.7543

Randomized subsampling



W, Yin. Analysis and generalizations of the linearized Bregman method. SIAM J. Imaging Sci., 3(4):856–877, 2010.
Herrmann F J, Tu N, Esser E. Fast "online" migration with Compressive Sensing[J].

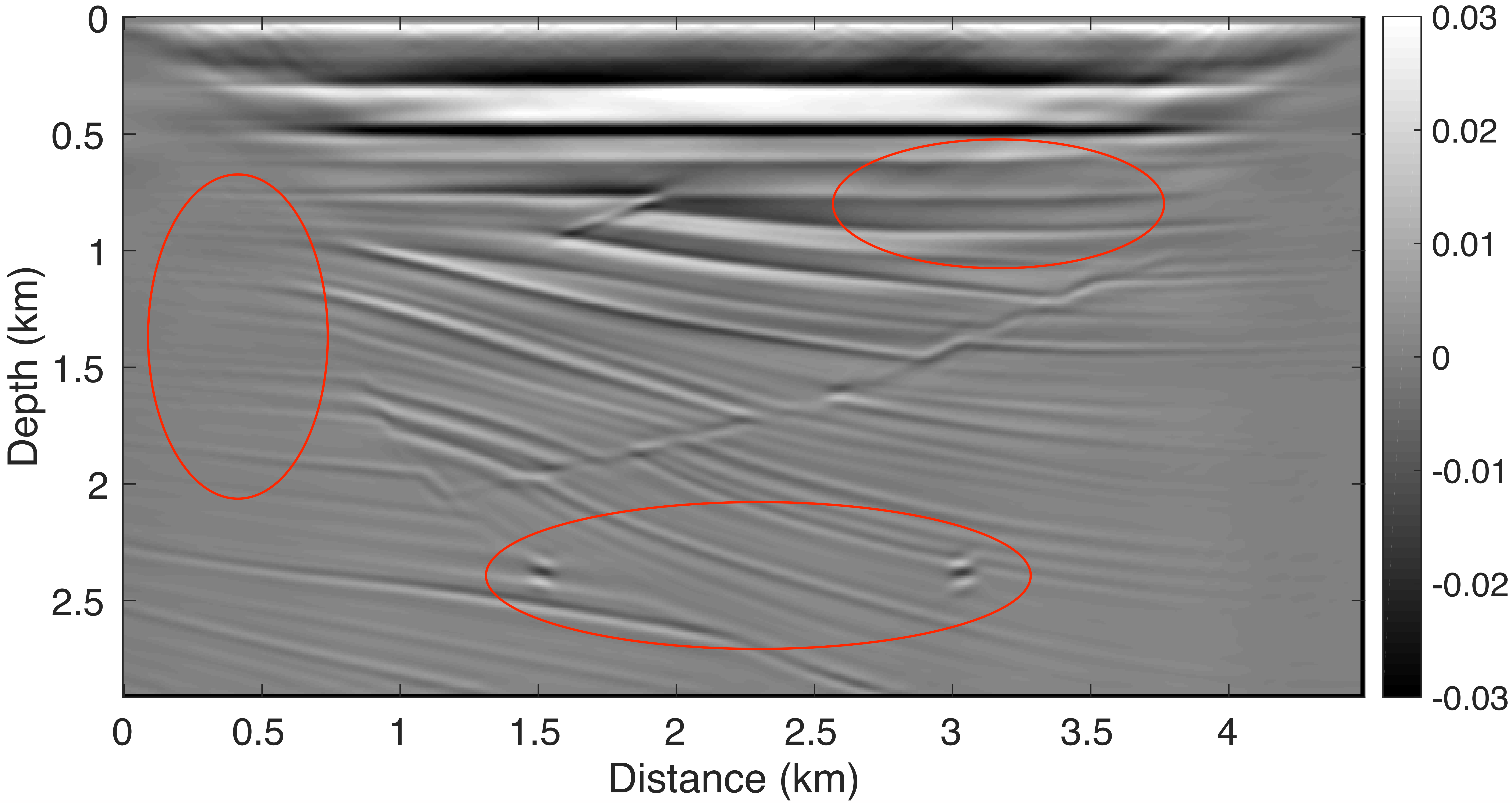
SLIM

Solvers for sparsity promoting inversion

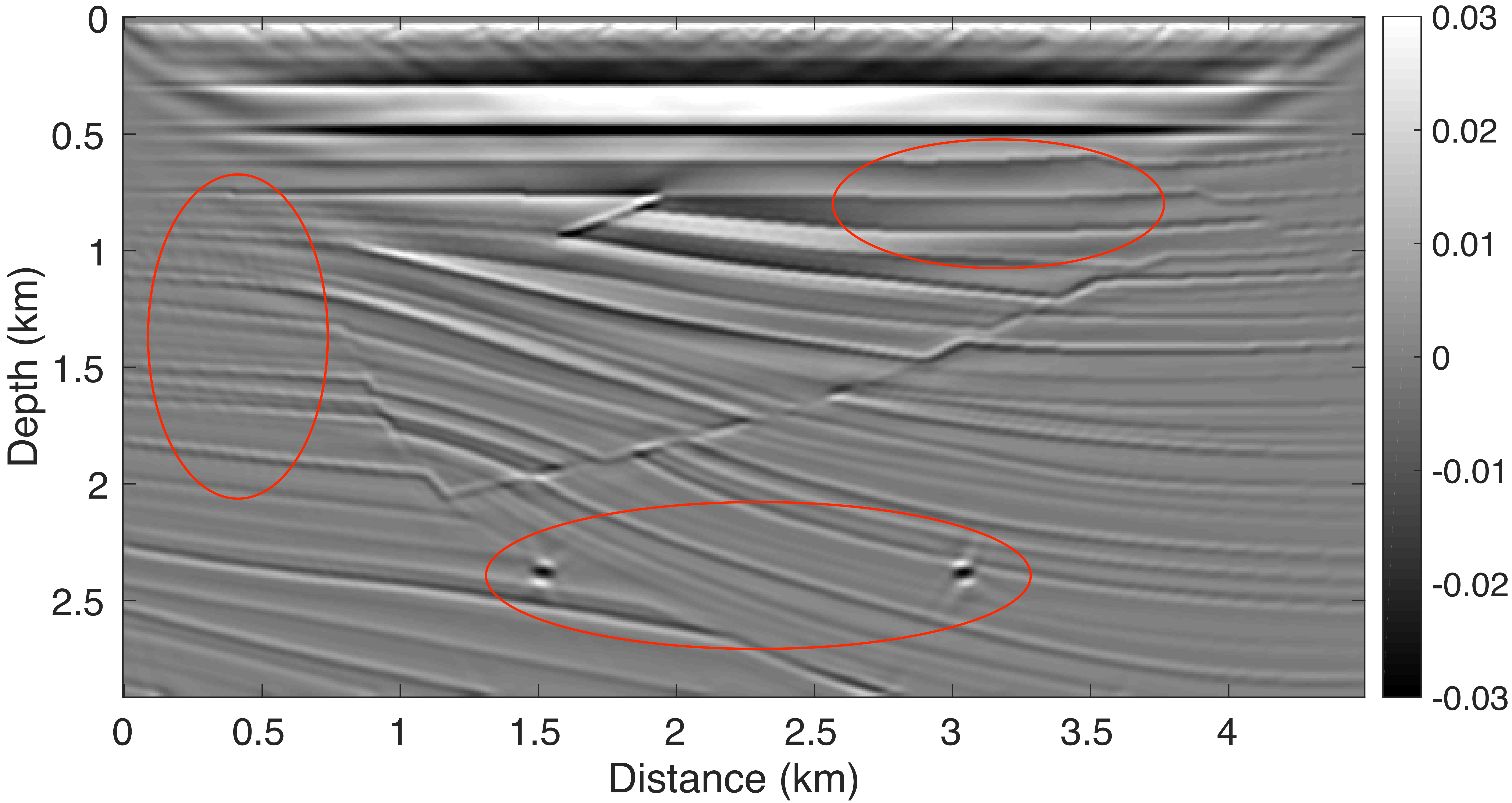
Many solvers for sparse. inversion:

- Iterative soft thresholding
  (simple, but slow convergence, cooling of threshold ...)
- Spectral projected gradients w/ L1 constraint – SPGL1
  (expensive, difficult to implement, slow convergence)
- Linearized Bregman (LB)
  (easy to implement, proven convergence w/ subsampling)

Sparsity promoting LS-RTM
w/ correct wavelet & SPGL1



Sparsity promoting LS-RTM
w/ correct wavelet & LB



W, Yin. Analysis and generalizations of the linearized Bregman method. SIAM J. Imaging Sci., 3(4):856–877, 2010.
Herrmann F J, Tu N, Esser E. Fast “online” migration with Compressive Sensing[J].

SLIM

Modification

$$\underset{\mathbf{x}}{\text{minimize}} \quad \lambda\|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|^2$$
$$\text{s.t.} \quad \|\hat{\mathbf{J}}\mathbf{x} - \mathbf{b}\|_2 \leq \sigma$$

- strongly convex objective function because of additional 2-norm term
- for big enough λ solves BP problem

Workflow for LB

$$\underset{\mathbf{x}}{\text{minimize}} \quad \lambda\|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|^2$$

$$\text{s.t.} \quad \|\hat{\mathbf{J}}\mathbf{x} - \mathbf{b}\|_2 \leq \sigma$$

1. **Initialize** $\mathbf{x}_0 = \mathbf{0}$, $\mathbf{z}_0 = \mathbf{0}$, q, λ, batchsize $n_s' \ll n_s$
2. **for** $k = 0, 1, \cdots$
3. **Randomly choose shot subsets** $\mathcal{I} \in [1 \cdots n_s]$, $|\mathcal{I}| = n_s'$
4. $\hat{\mathbf{J}}_k = \{\mathbf{J}_i(\mathbf{m}_0, q_i)\mathbf{C}^*\}_{i\in\mathcal{I}}$
5. $\mathbf{b}_k = \{\mathbf{b}_i\}_{i\in\mathcal{I}}$
6. $\mathbf{z}_{k+1} = \mathbf{z}_k - t_k \hat{\mathbf{J}}_k^T P_\sigma(\hat{\mathbf{J}}_k\mathbf{x}_k - \mathbf{b}_k)$
7. $\mathbf{x}_{k+1} = S_\lambda(\mathbf{z}_{k+1})$
8. **end**

note: $S_\lambda(\mathbf{z}_{k+1}) = \text{sign}(\mathbf{z}_{k+1}) \max\{0, |\mathbf{z}_{k+1}| - \lambda\}$

$\mathcal{P}_\sigma(\hat{\mathbf{J}}_k\mathbf{x}_k - \mathbf{b}_k) = \max\{0, 1 - \frac{\sigma}{\|\hat{\mathbf{J}}_k\mathbf{x}_k - \mathbf{b}_k\|}\} \cdot (\hat{\mathbf{J}}_k\mathbf{x}_k - \mathbf{b}_k)$

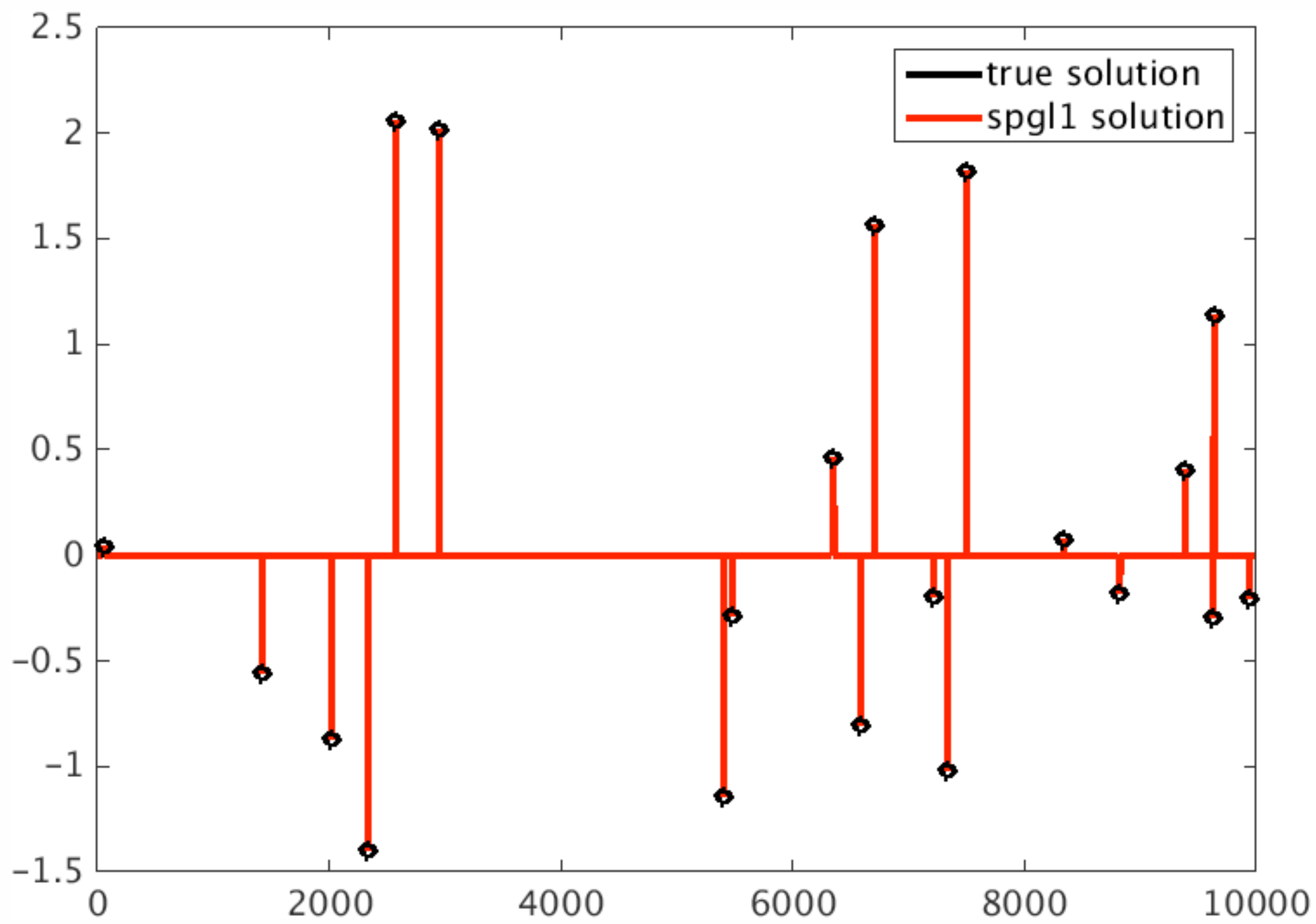
Toy example

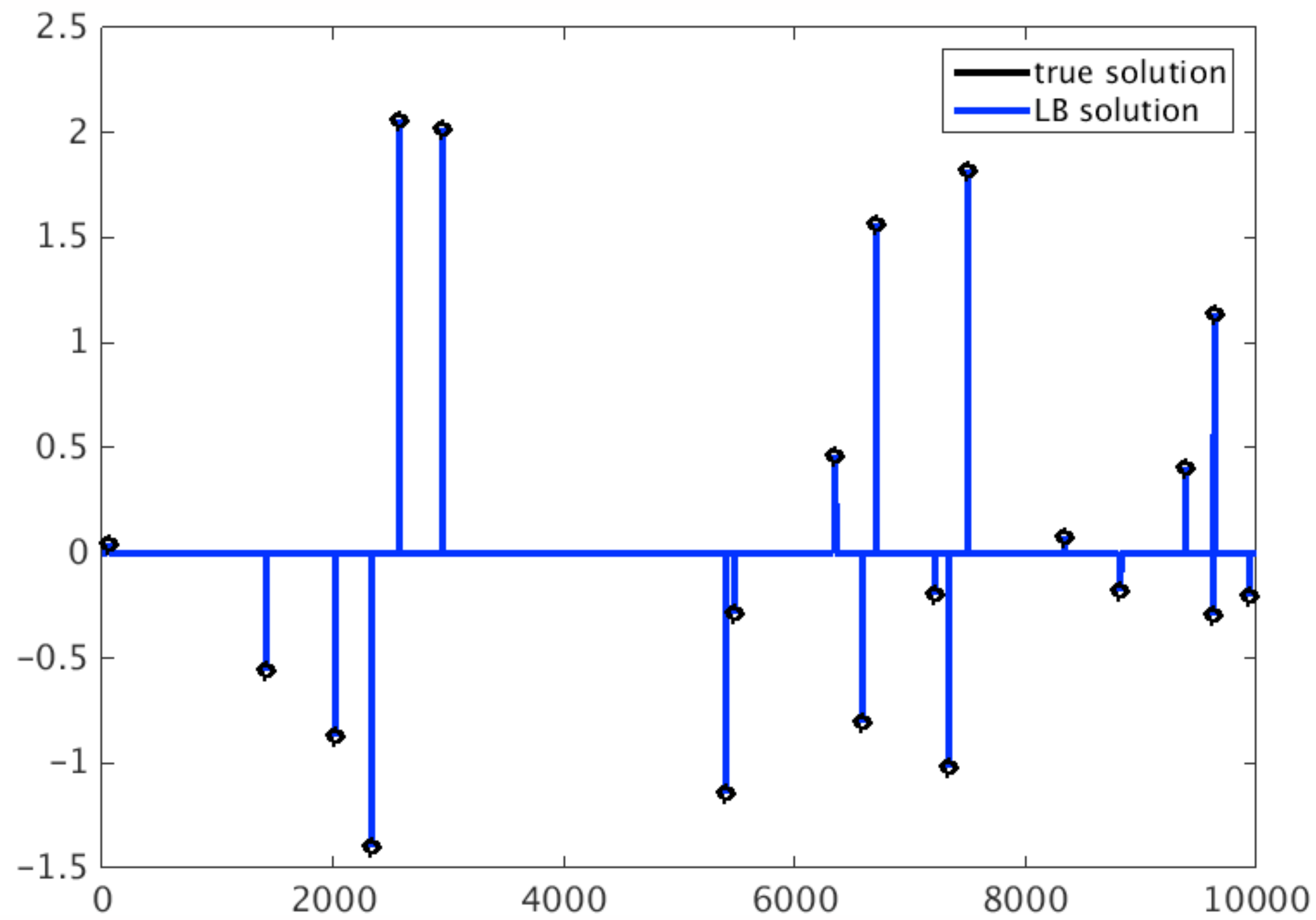
Sparsity recovery with tall ill-conditioned matrix

A: 20000 X 10000, with Rank 500

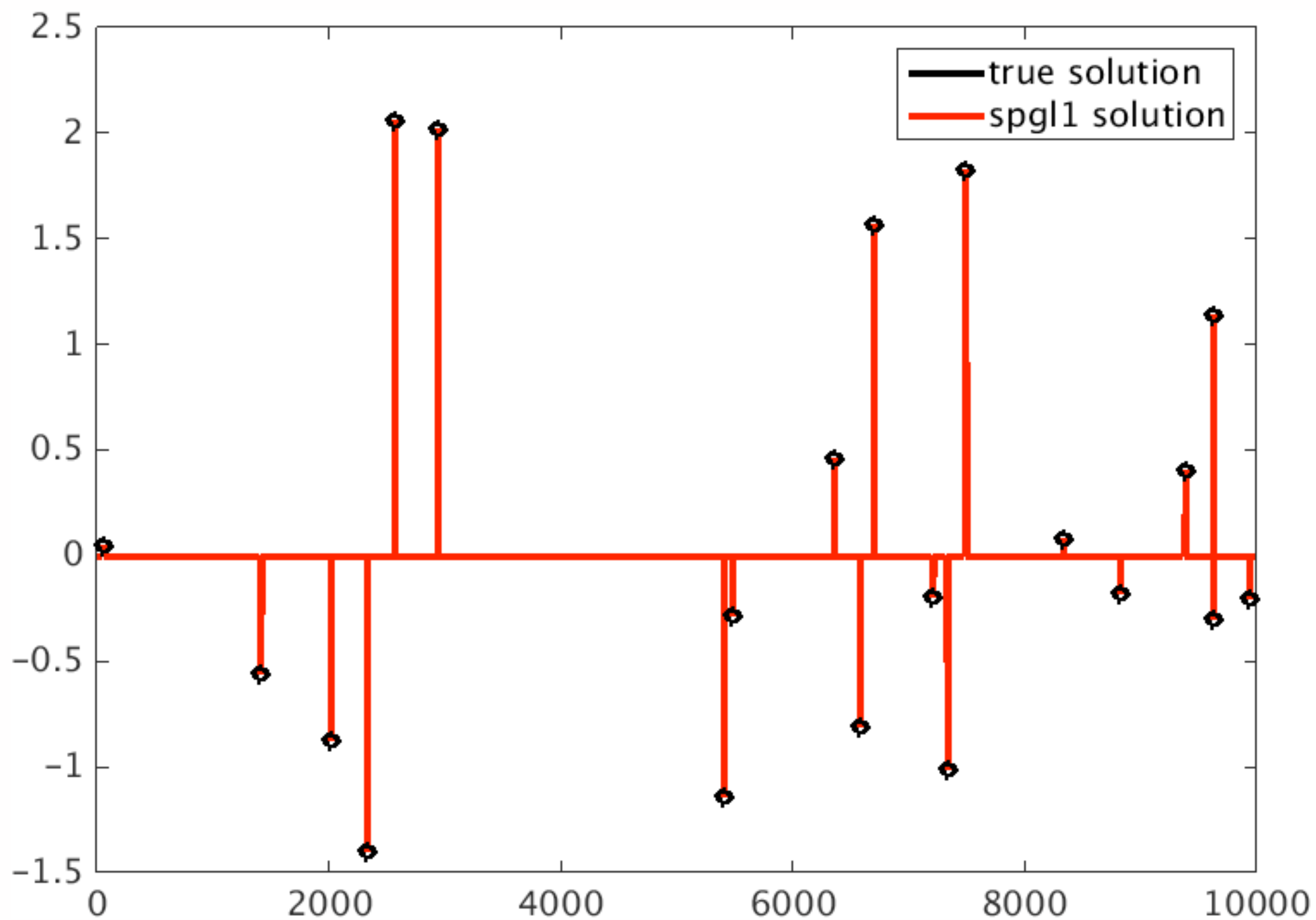
x: 10000 X 1, with 20 non-zeros

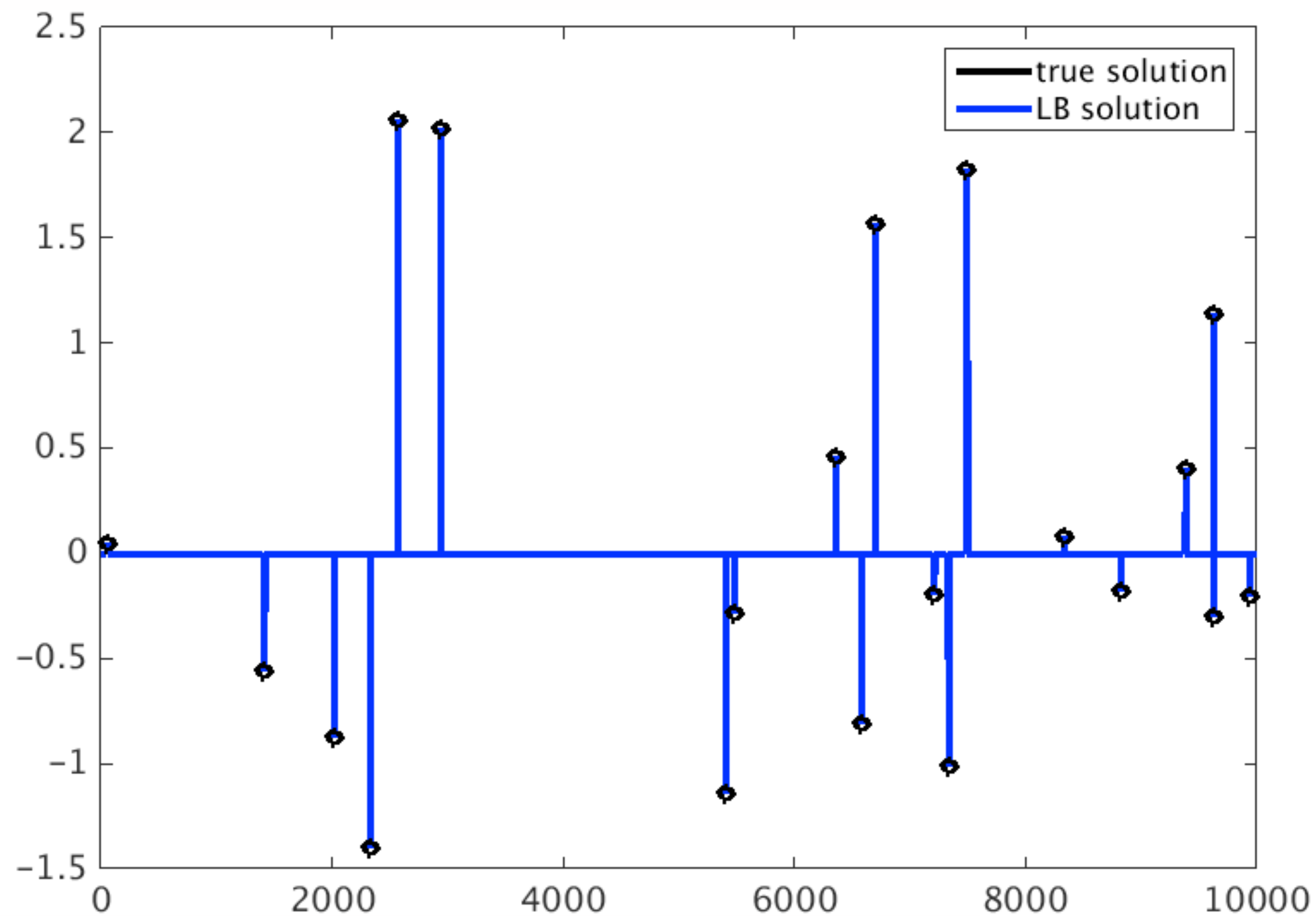
SPGL1 vs LB
no subsampling



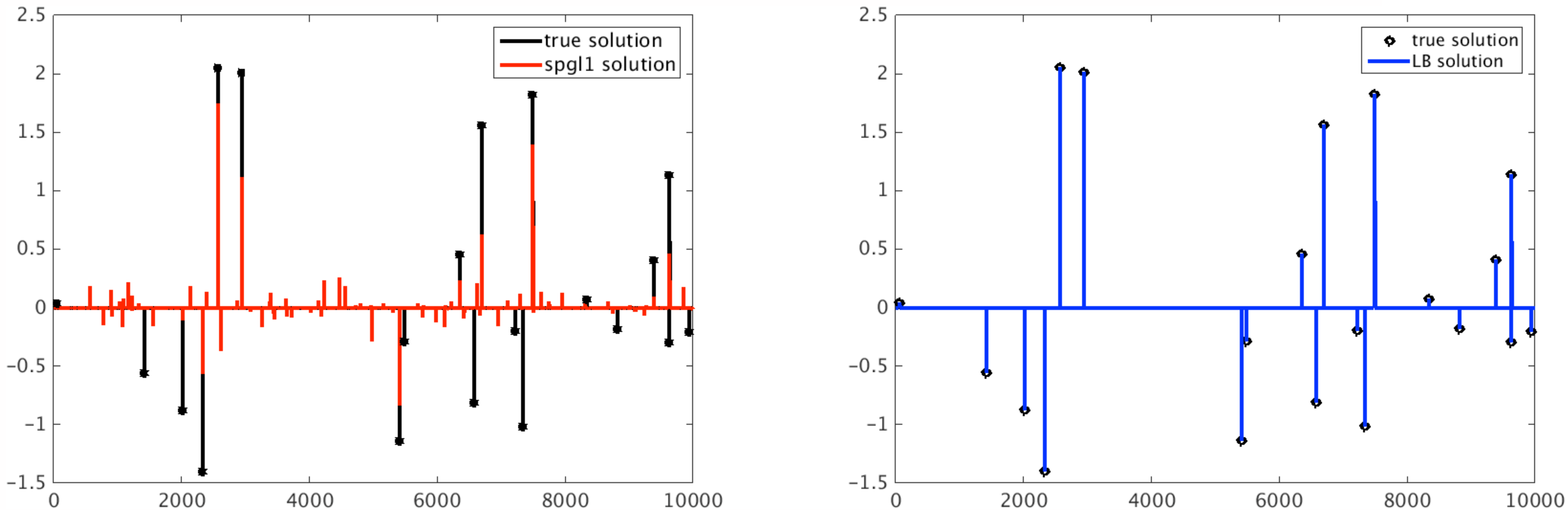


SPGL1 vs LB
50% subsampling

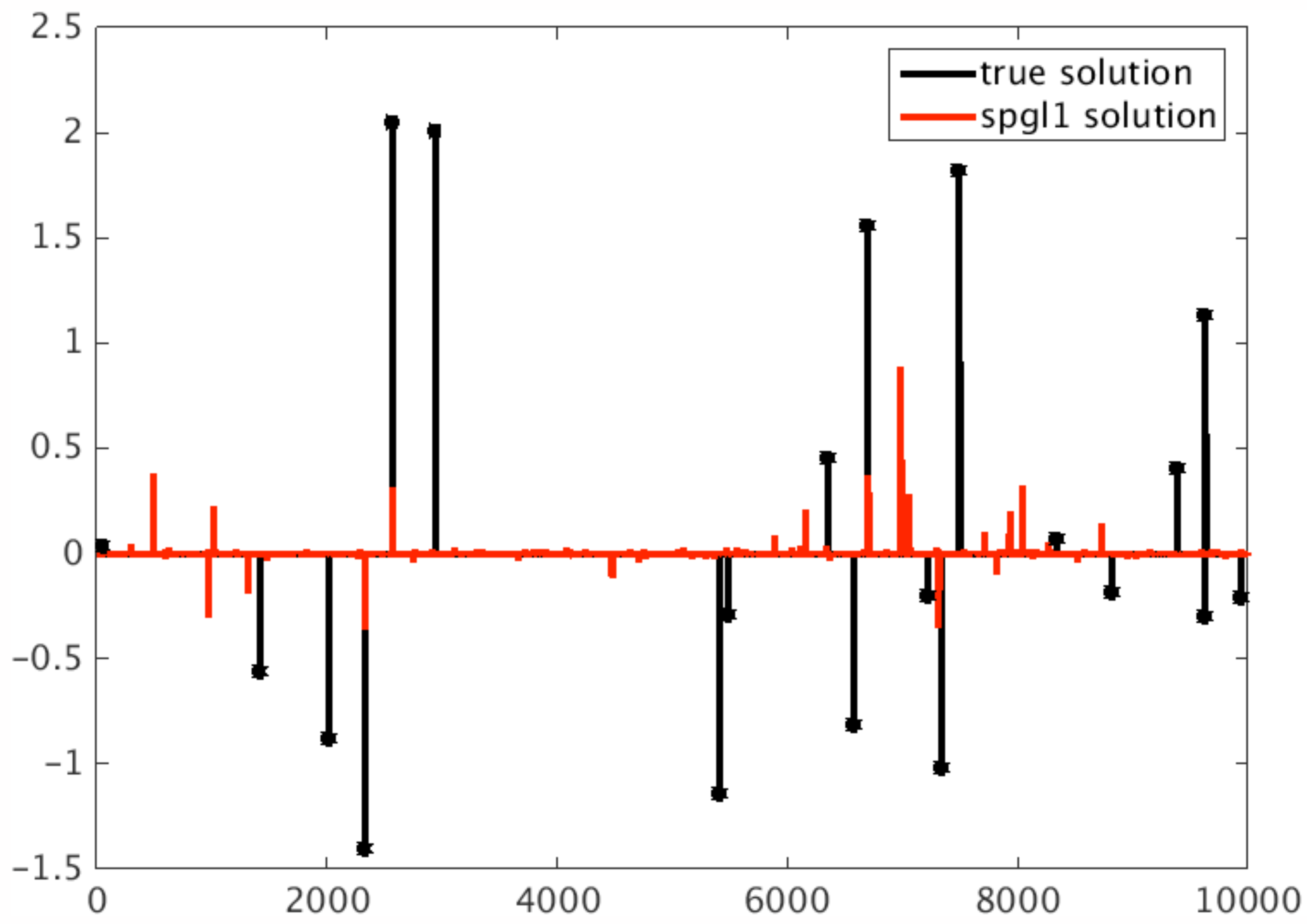


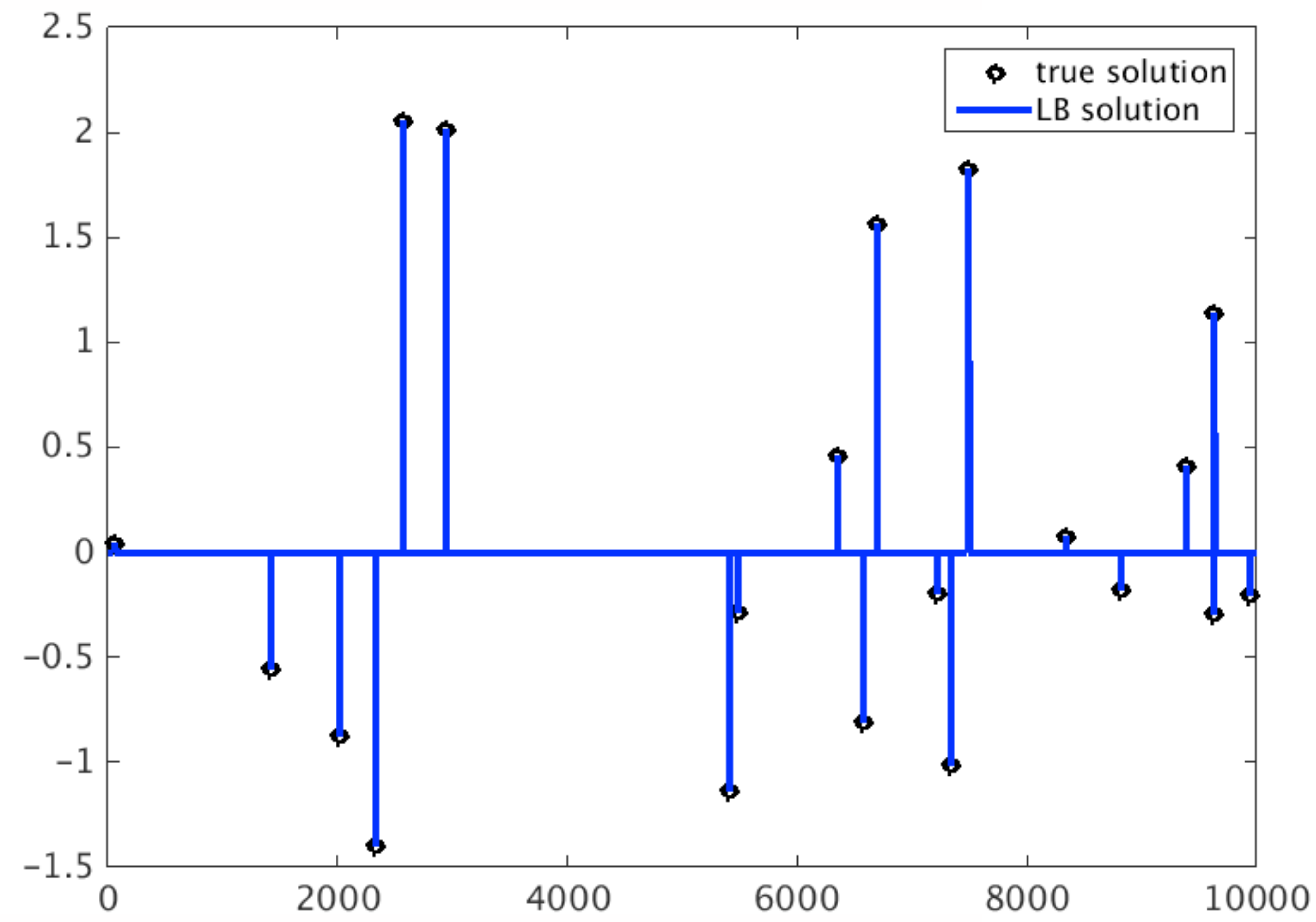


SPGL1 vs LB
80% subsampling



SPGL1 vs LB
90% subsampling





SLIM

What if source signature is unknown?

Estimate source by solving least-square problem:

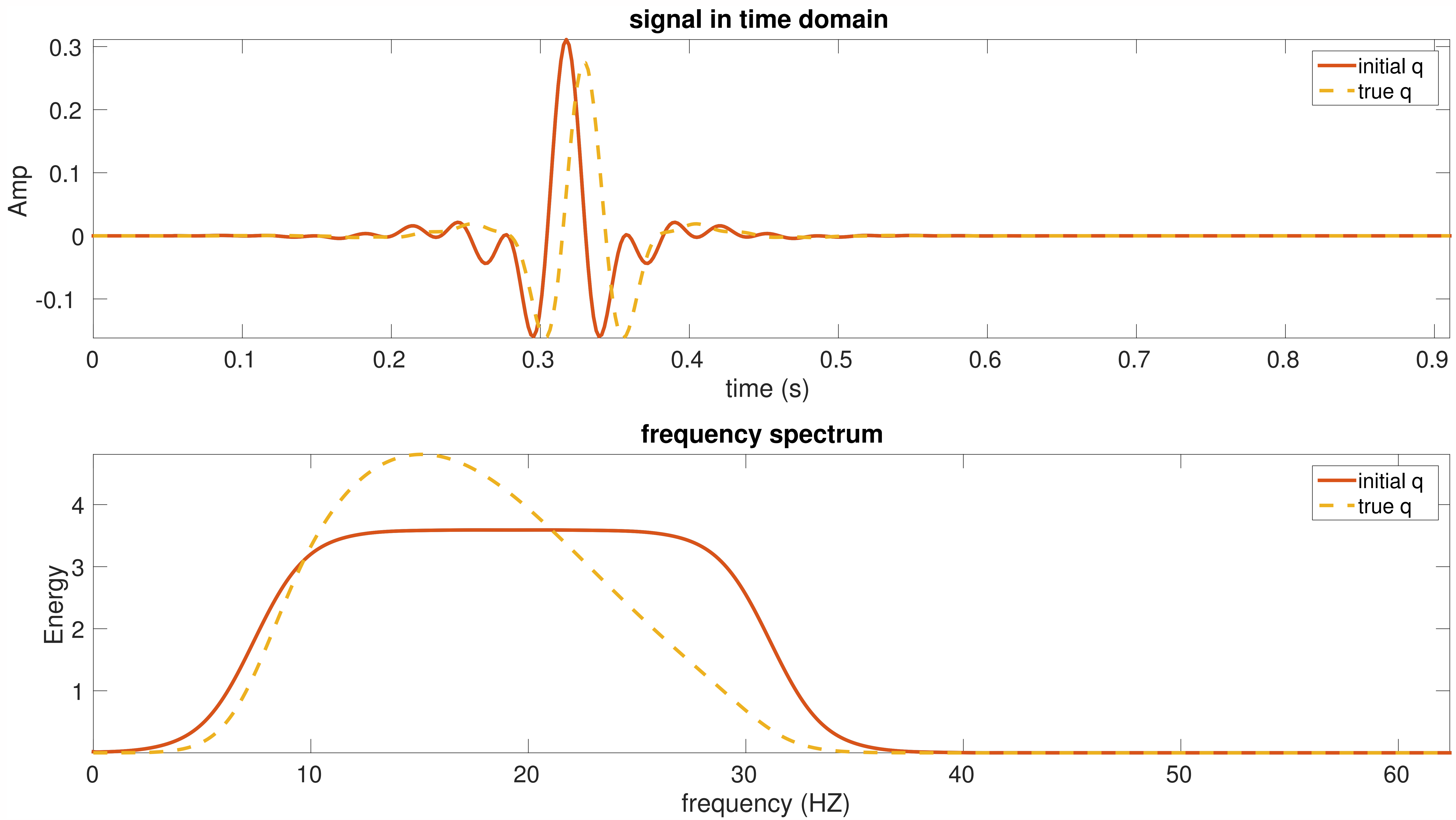
$$\min_{\mathbf{w}} \sum_j^{Ntr} \|\mathbf{w} * \tilde{\mathbf{b}}_j - \mathbf{b}_j\|^2 \ \ \mathbf{s.t.} \|\mathbf{w}\|^2 = 1$$

where $\widetilde{\mathbf{b}} = \mathbf{J}(\mathbf{m_0}, \mathbf{q_0})\widetilde{\delta\mathbf{m}}$

Suppose that $\mathbf{q} = \mathbf{w} * \mathbf{q}_0$, and $\mathbf{q}$ is the same for all shots

$\mathbf{q}_0$ is the initial guess of $\mathbf{q}$

Initial wavelet setting

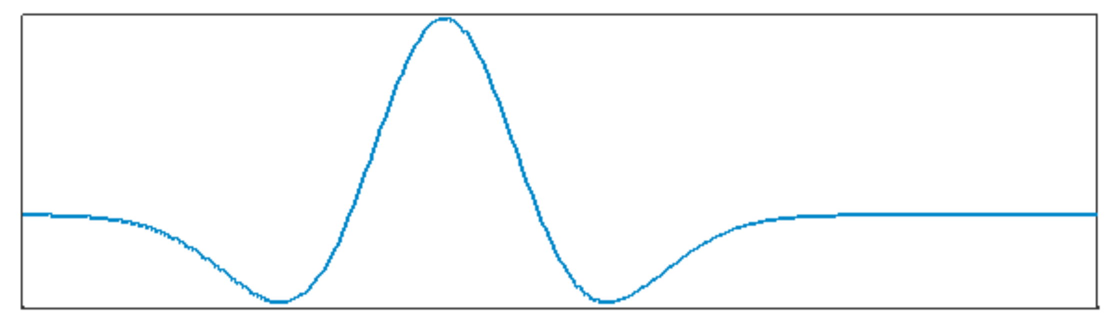


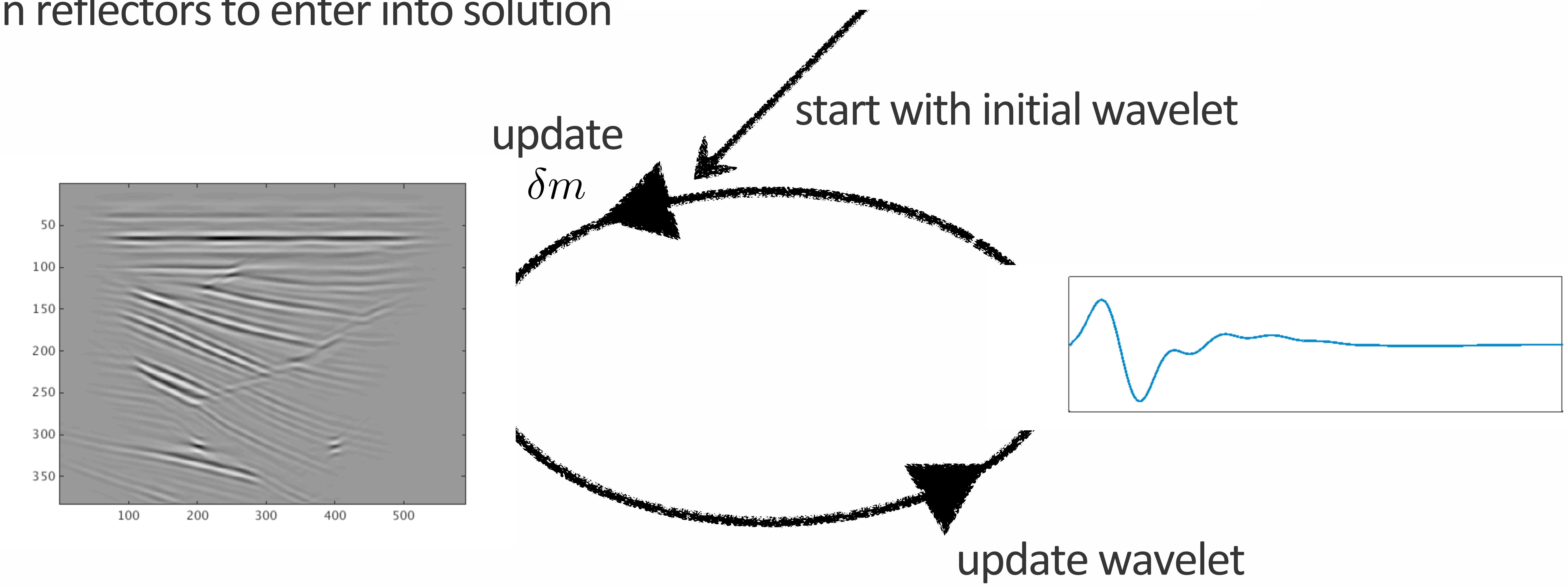
approximate duration

frequency bandwidth wider due to factorization

Combine the image inversion & source estimation

start w/ sufficiently small threshold
to allow main reflectors to enter into solution





Workflow for sparsity-promoting LS-RTM w/ source estimation

1. **Initialize** $\mathbf{x}_0 = \mathbf{0}$, $\mathbf{z}_0 = \mathbf{0}$, $\mathbf{q}_0$, λ, λ_2, **batchsize** $n_s' \ll n_s$, **weights** r
2. **for** $k = 0, 1, \cdots$
3. **Randomly choose shot subsets** $\mathcal{I} \in [1 \cdots n_s], |\mathcal{I}| = n_s'$
4. $\hat{\mathbf{J}}_k = \{\mathbf{J}_i(\mathbf{m}_0, q_0)\mathbf{C}^*\}_{i\in\mathcal{I}}$
5. $\mathbf{b}_k = \{\mathbf{b}_i\}_{i\in\mathcal{I}}$
6. $\tilde{\mathbf{b}}_k = \hat{\mathbf{J}}_k\mathbf{x}_k$
7. $\mathbf{w}_k = \arg\min_{\mathbf{w}} \sum_{\mathcal{I}} \|\mathbf{w} * \tilde{\mathbf{b}}_k - \mathbf{b}_k\|^2 + \|r(\mathbf{w} * \mathbf{q}_0)\|^2 + \lambda_2\|\mathbf{w} * \mathbf{q}_0\|^2$
8. $\mathbf{z}_{k+1} = \mathbf{z}_k - t_k\hat{\mathbf{J}}_k^*\Big(\mathbf{w}_k \star P_\sigma(\mathbf{w}_k * \tilde{\mathbf{b}}_k - \mathbf{b}_k)\Big)$
9. $\mathbf{x}_{k+1} = S_\lambda(\mathbf{z}_{k+1})$
10. **end**

Experiments

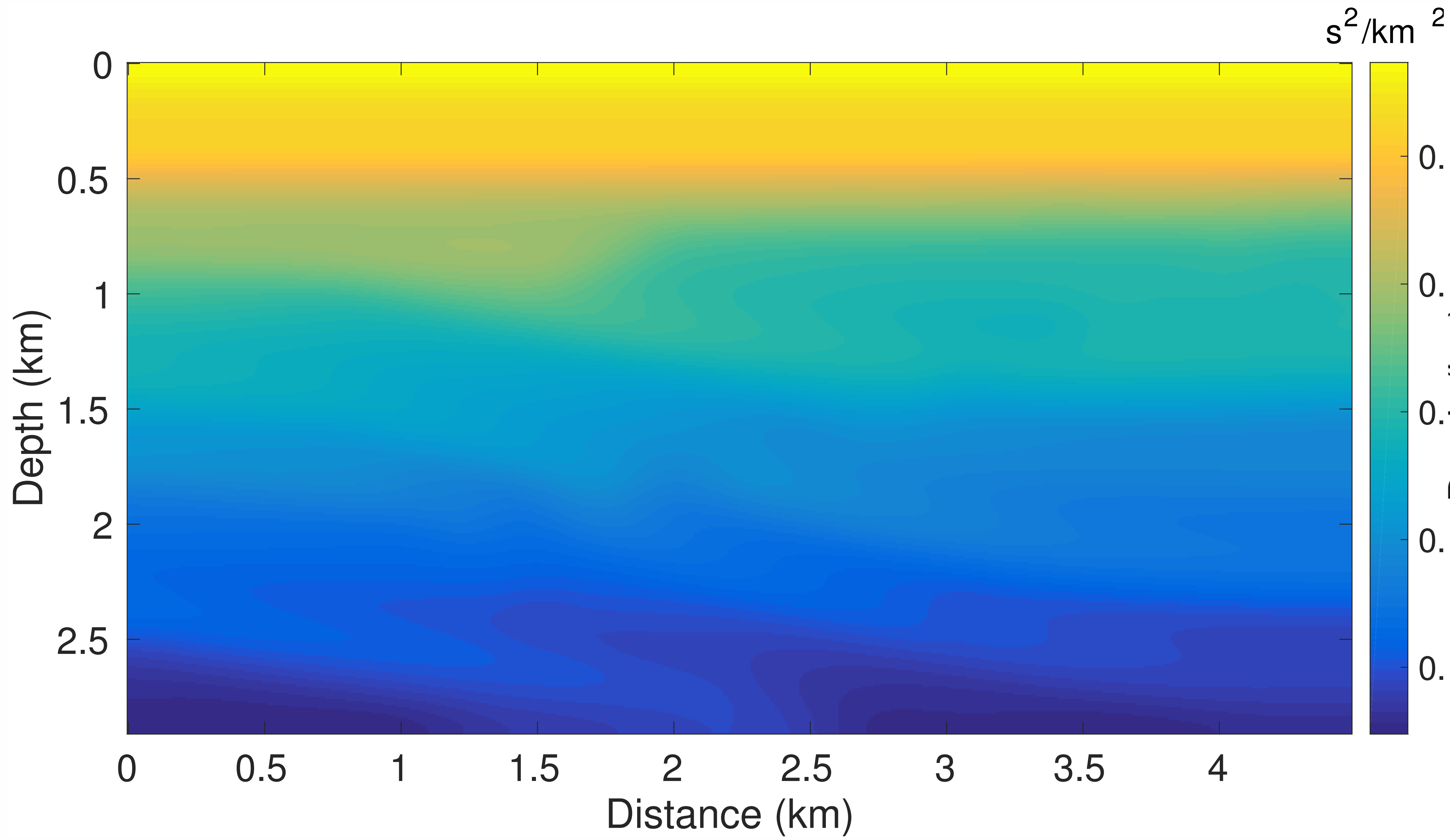
Data:

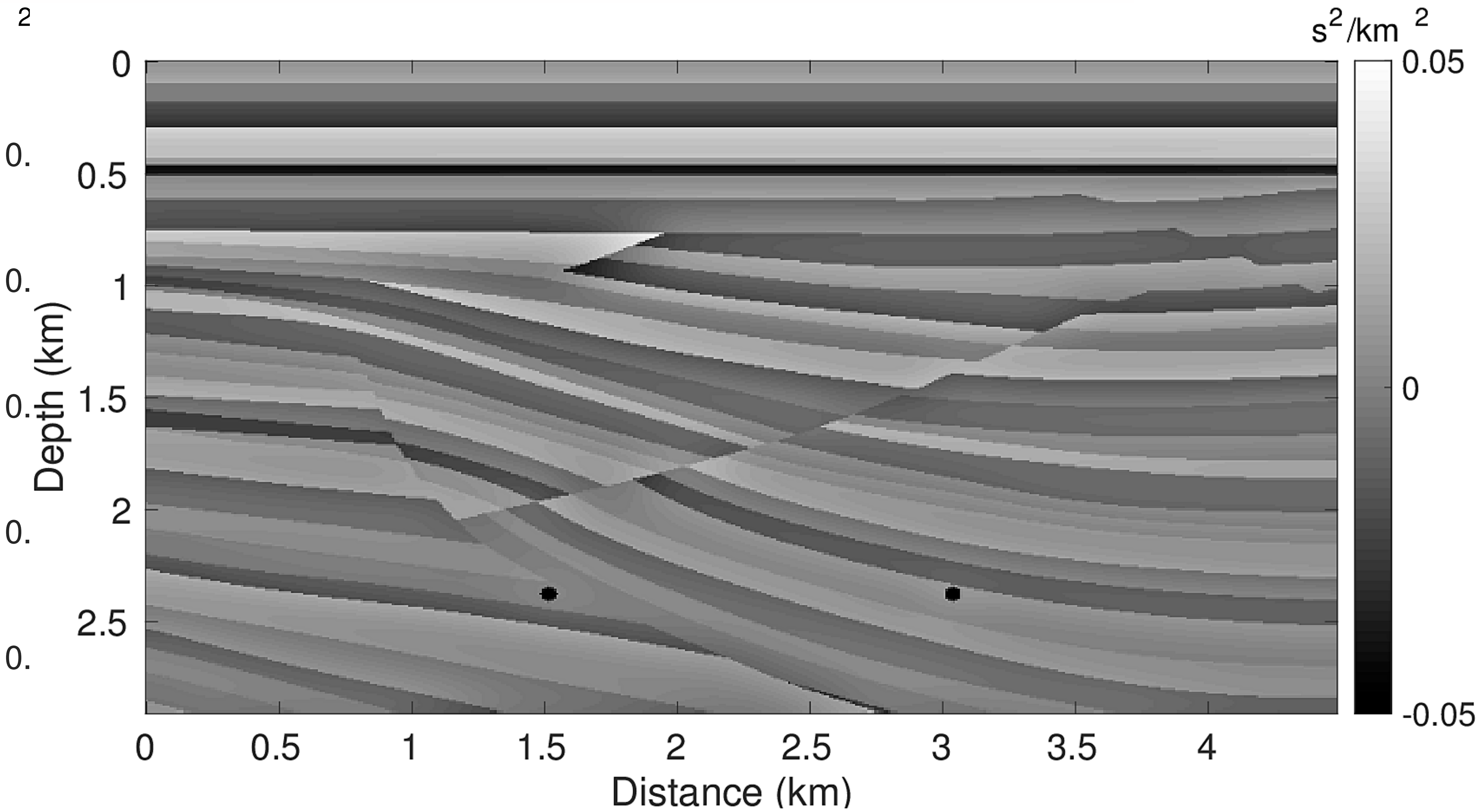
- 295 shots with shot interval 15m
- 295 receivers with receiver interval 15m
- 4s record, 15Hz peak frequency designed wavelet
- synthetic linearized data

Experiments:

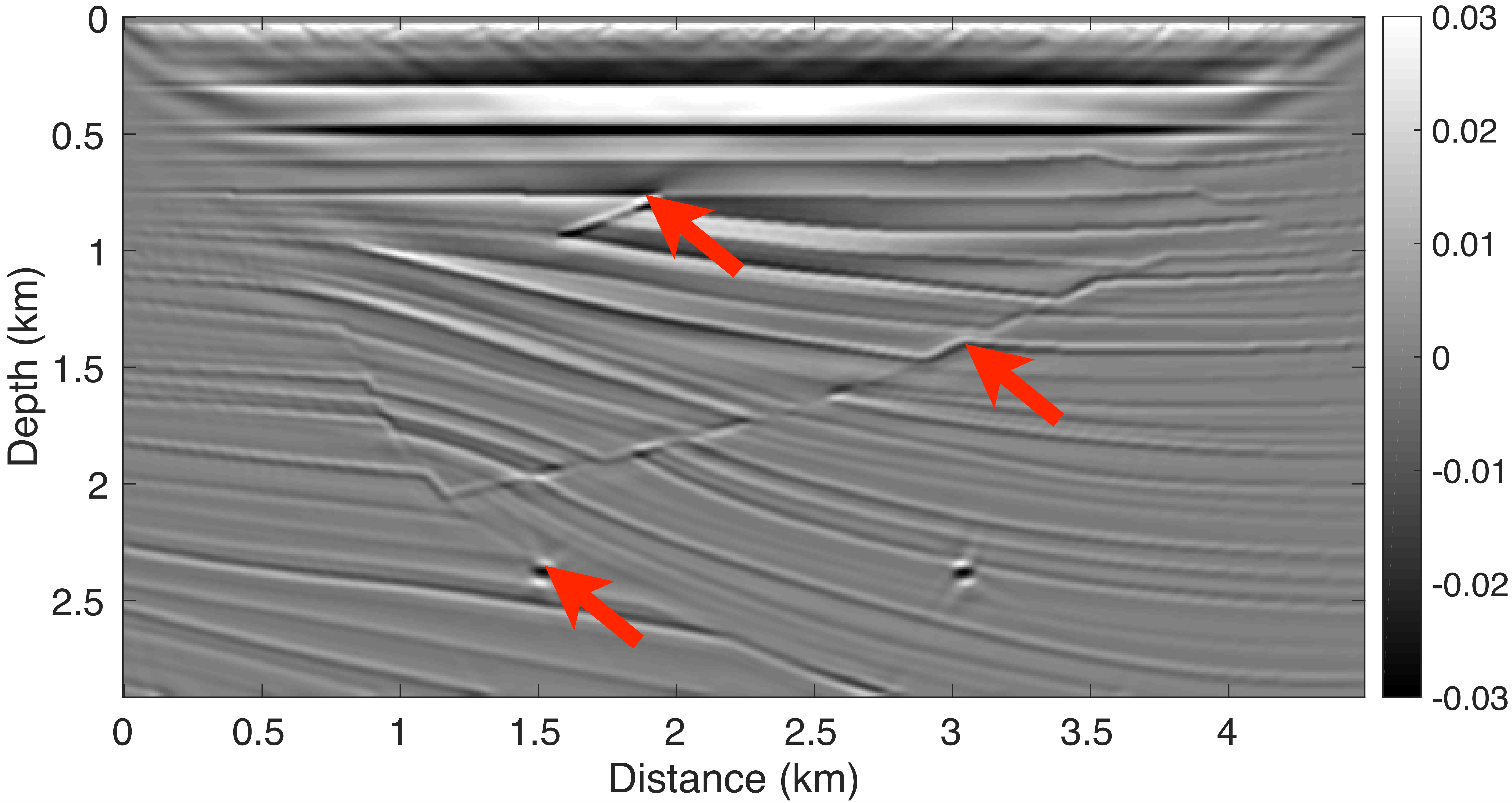
- one pass through the data with batch sizes 2.5% data
- randomized subset of shots
- normalized true source wavelet & initial guessed wavelet

background model and model perturbation

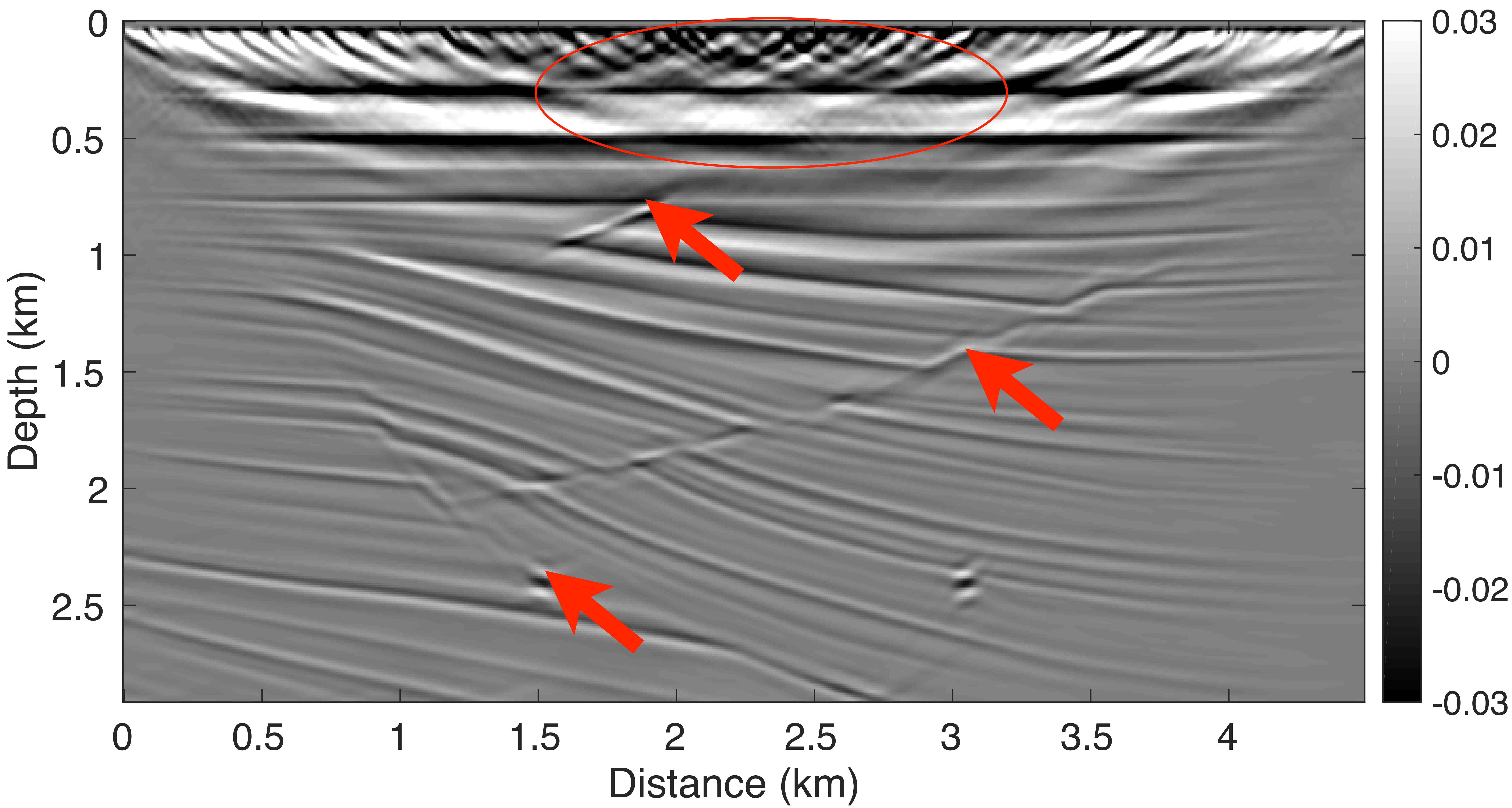




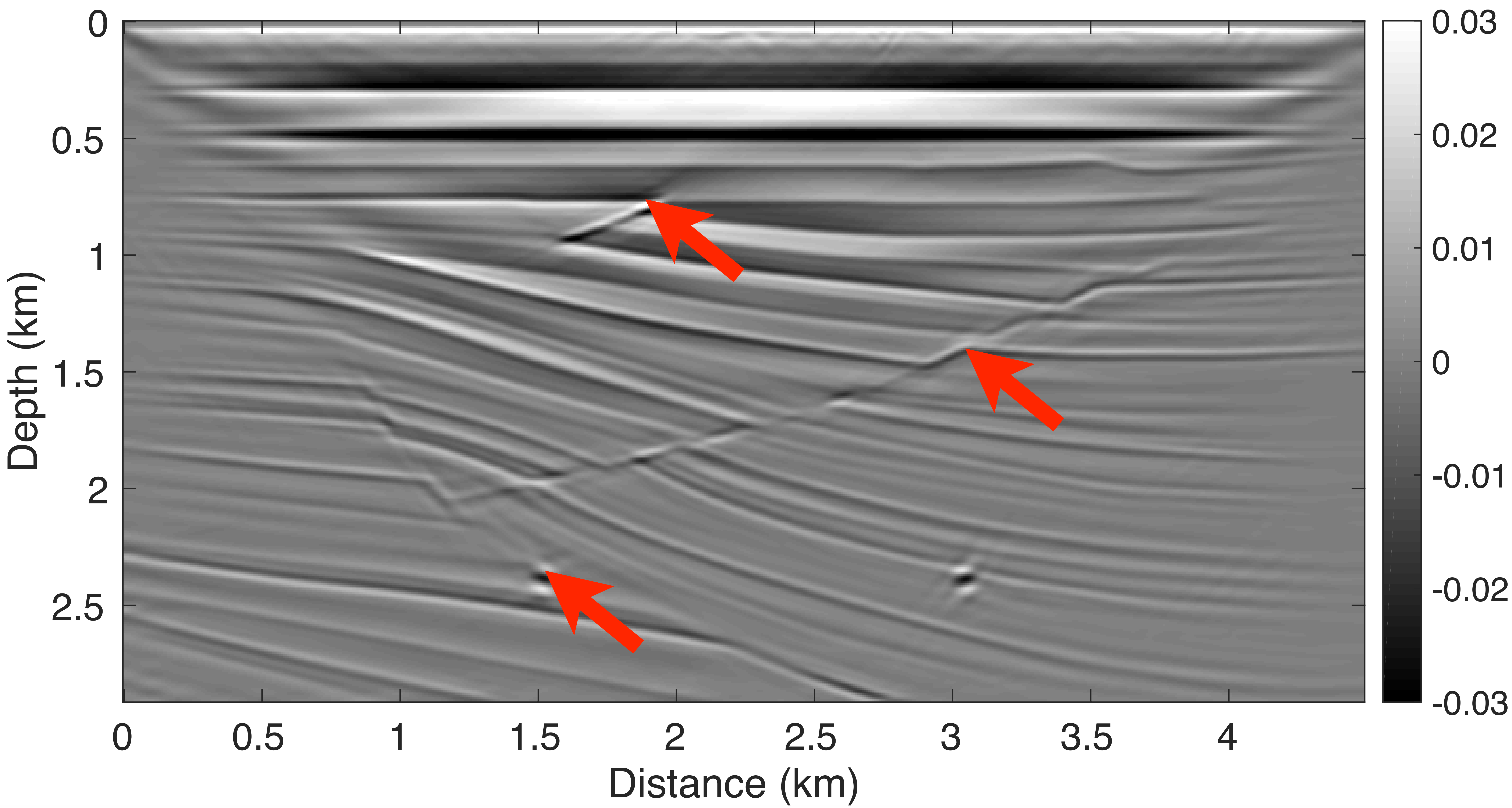
Sparsity promoting LS-RTM
w/ correct wavelet & LB



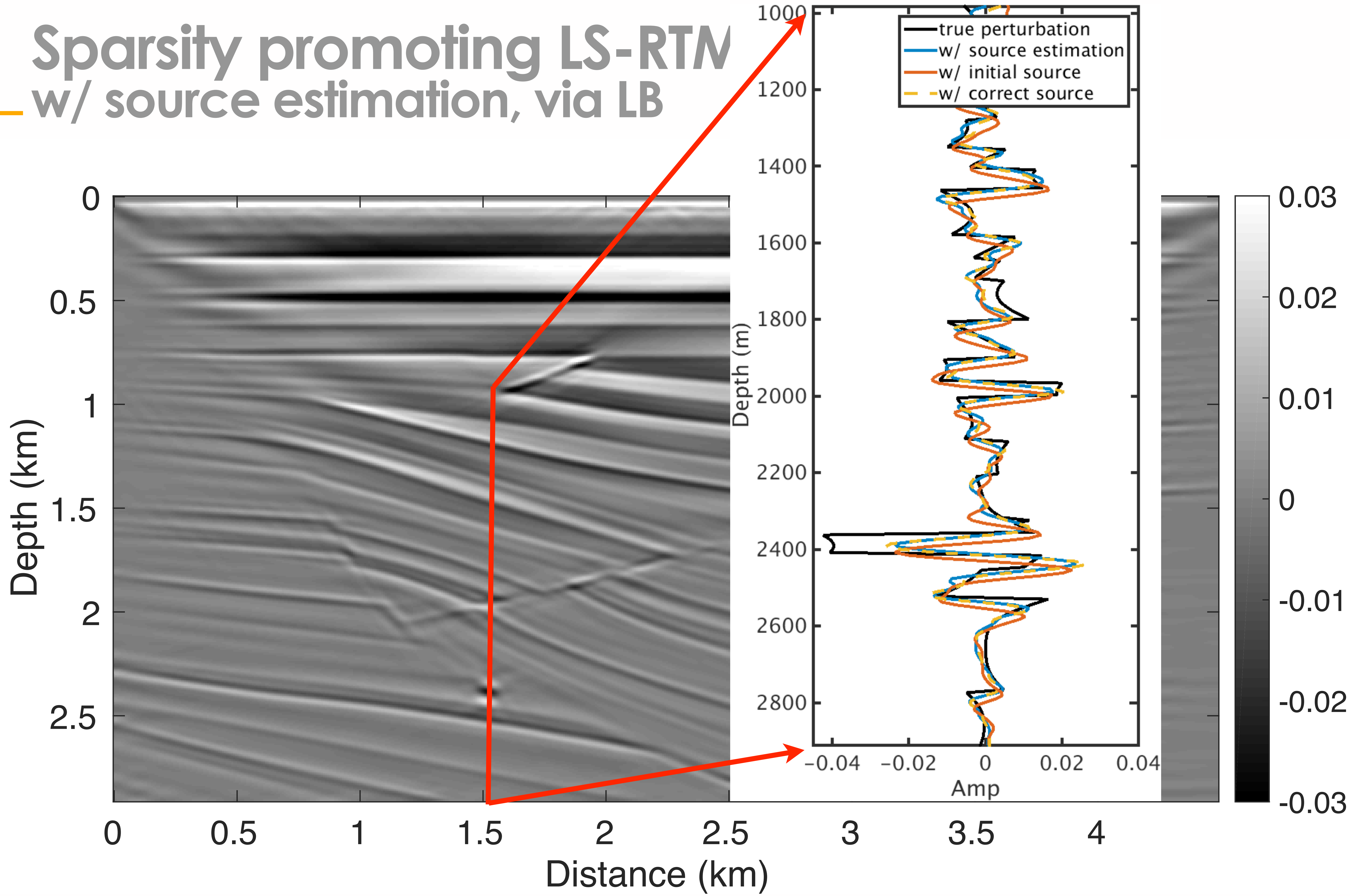
Sparsity promoting LS-RTM
w/ wrong wavelet & LB



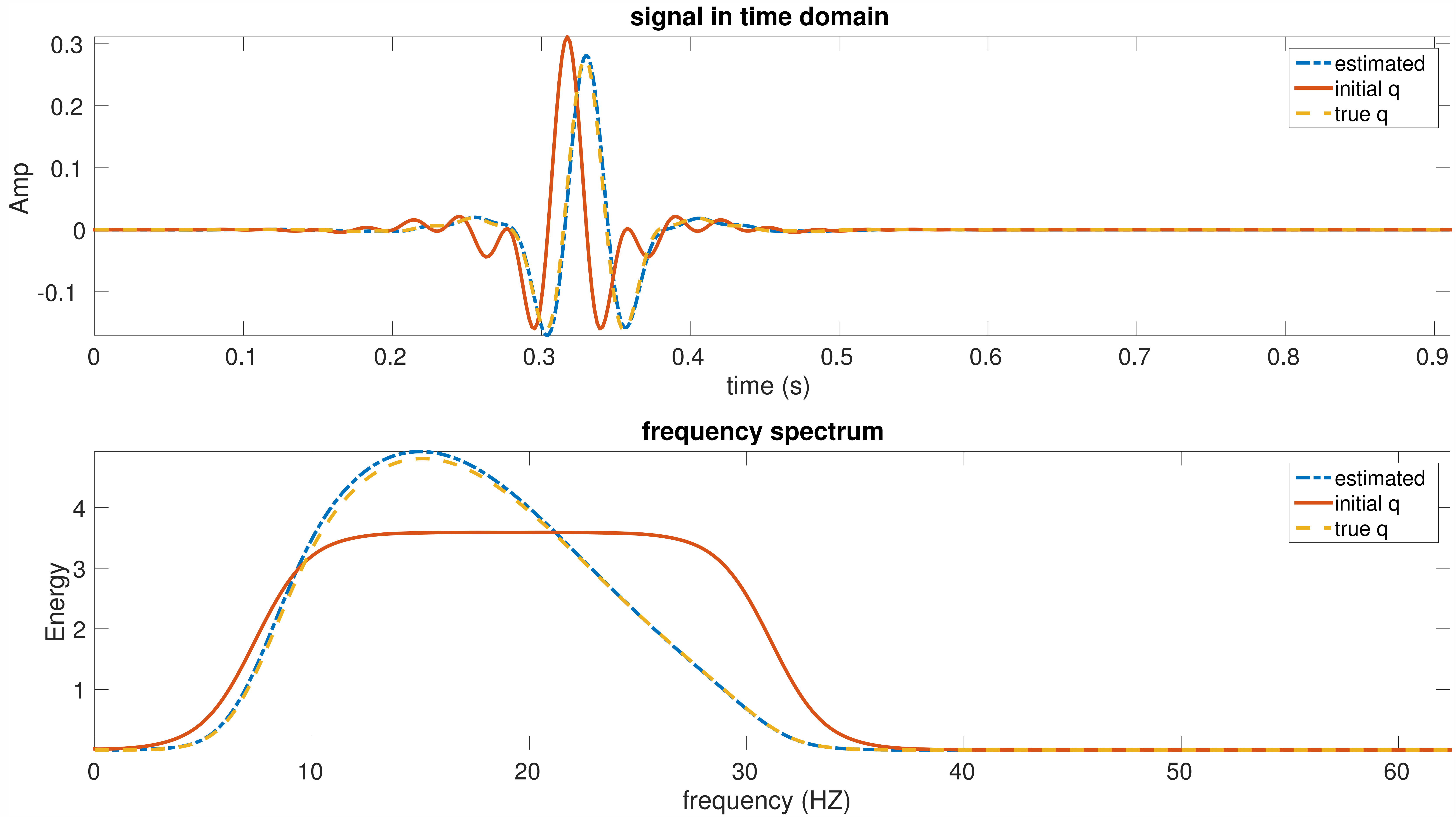
Sparsity promoting LS-RTM
w/ source estimation w/ LB



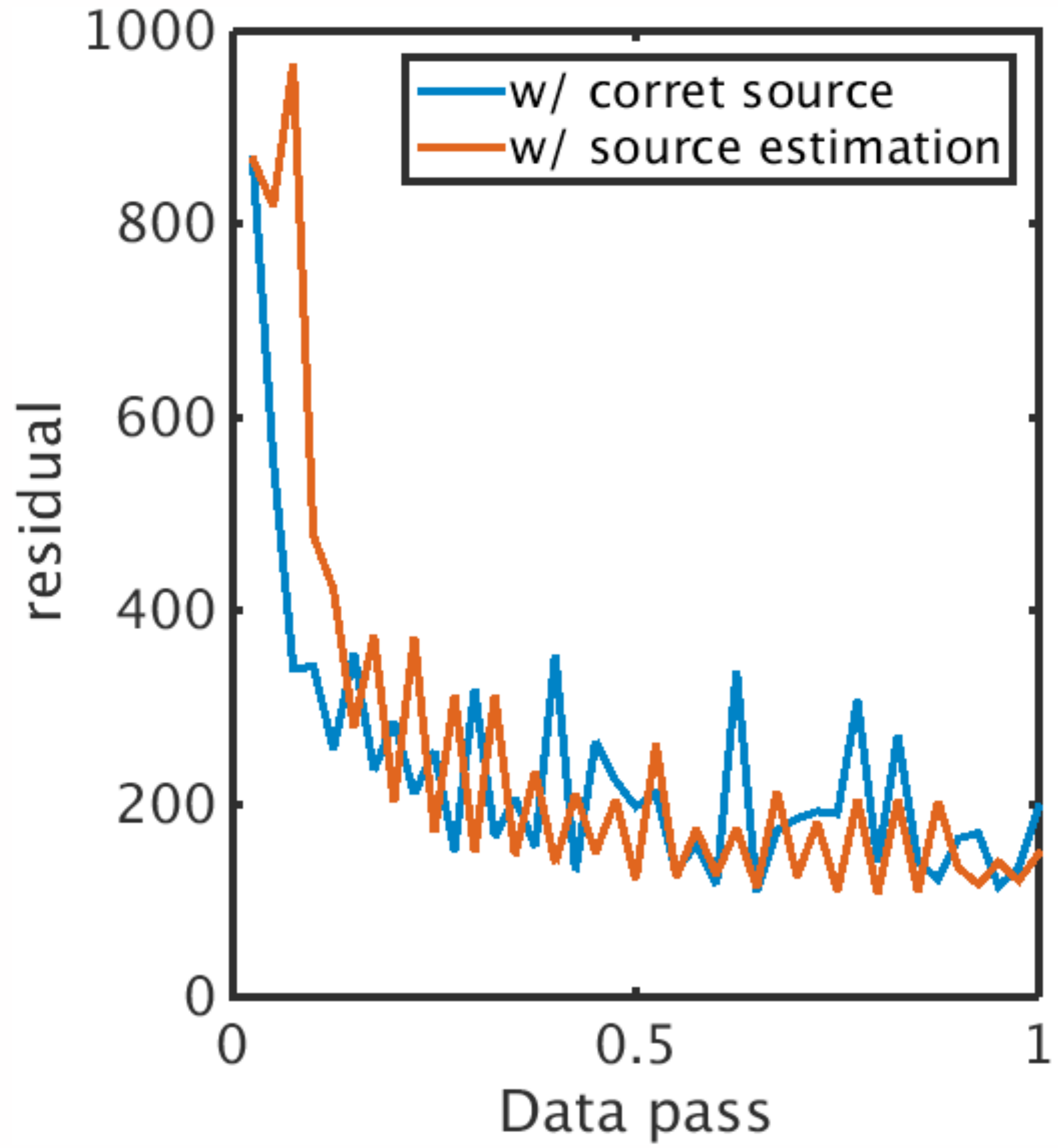
Sparsity promoting LS-RTM
w/ source estimation, via LB



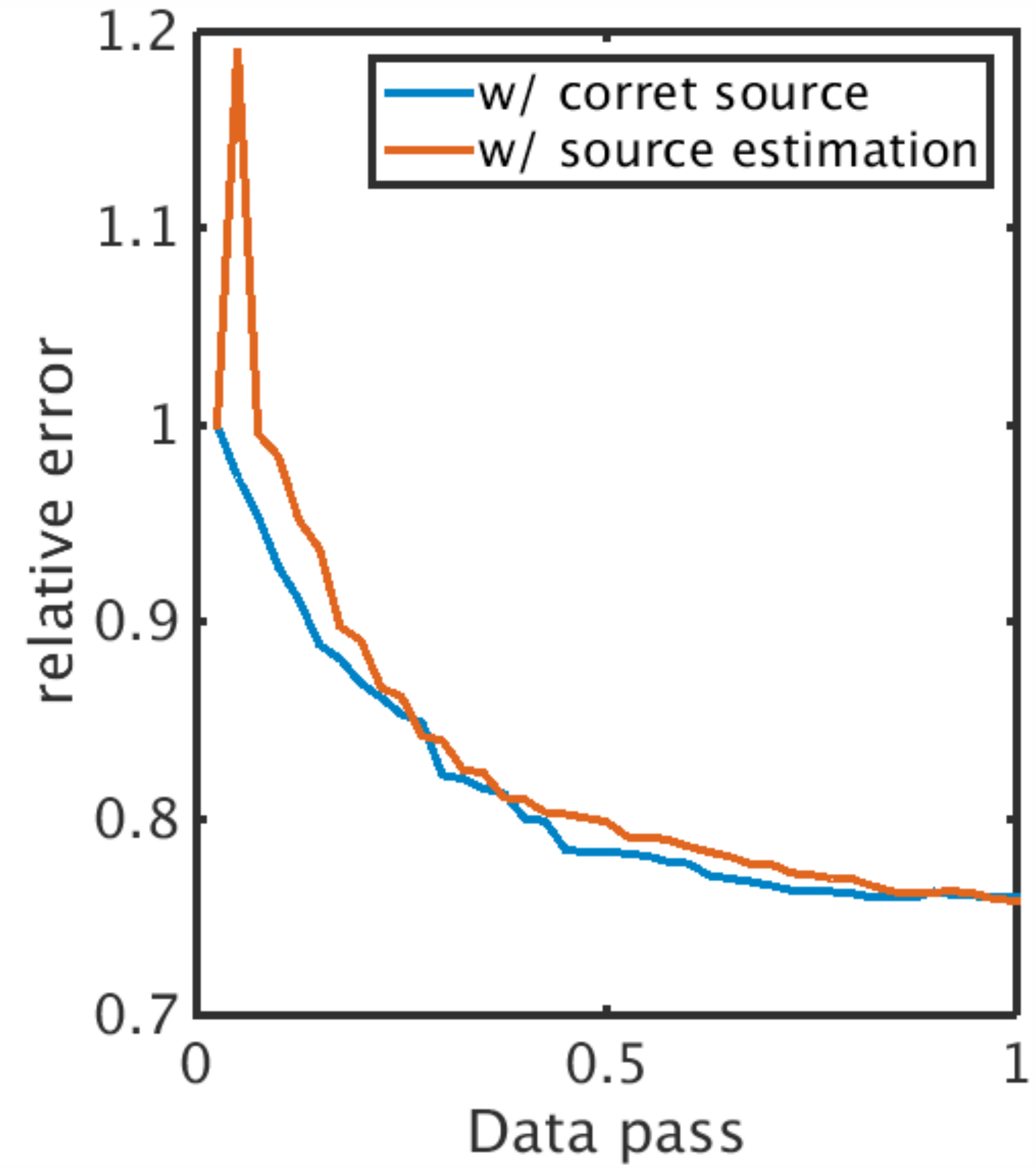
Sparsity promoting LS-RTM
w/ source estimation



Residual & model error

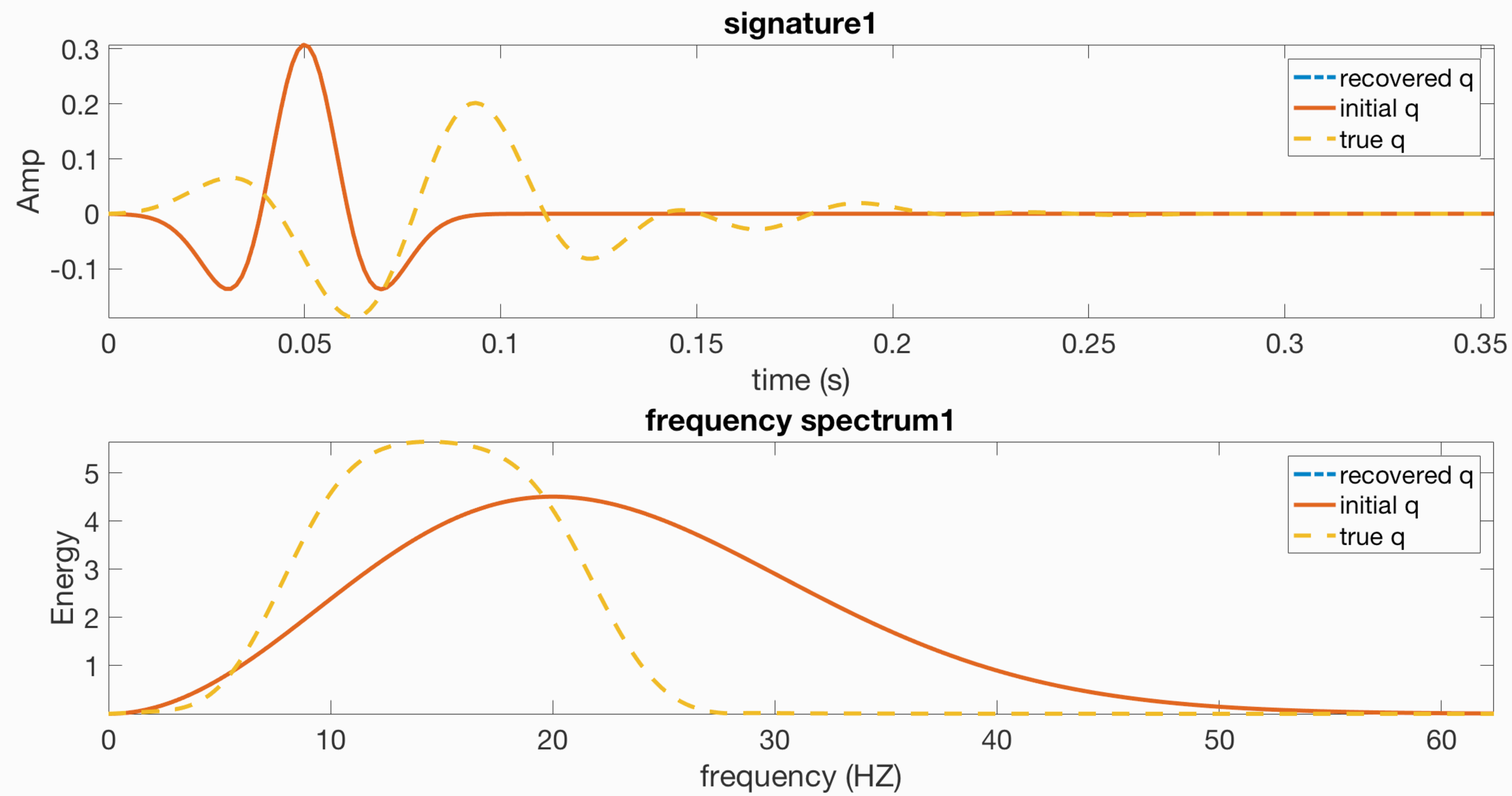


Residuals

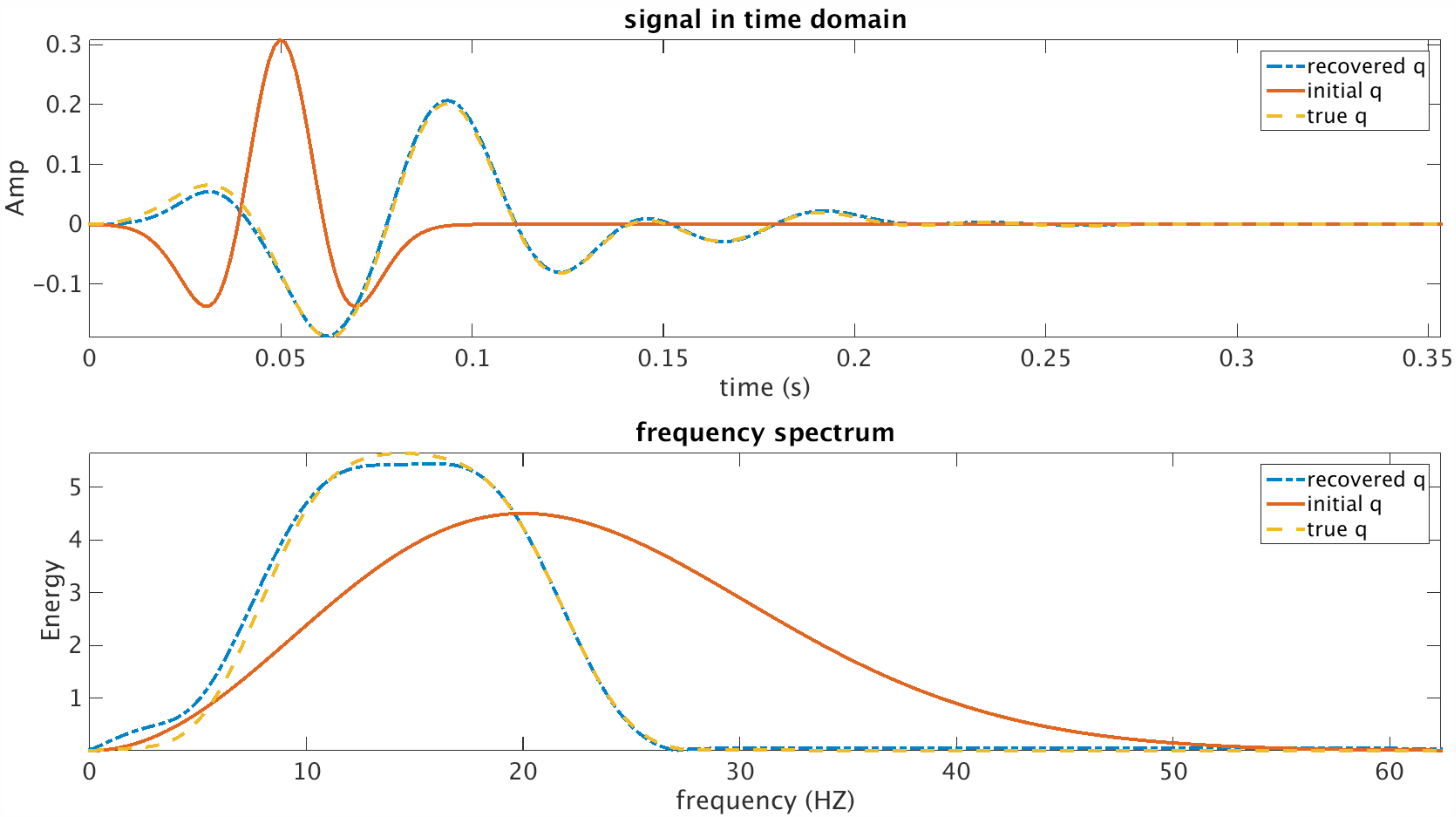


Relative model error

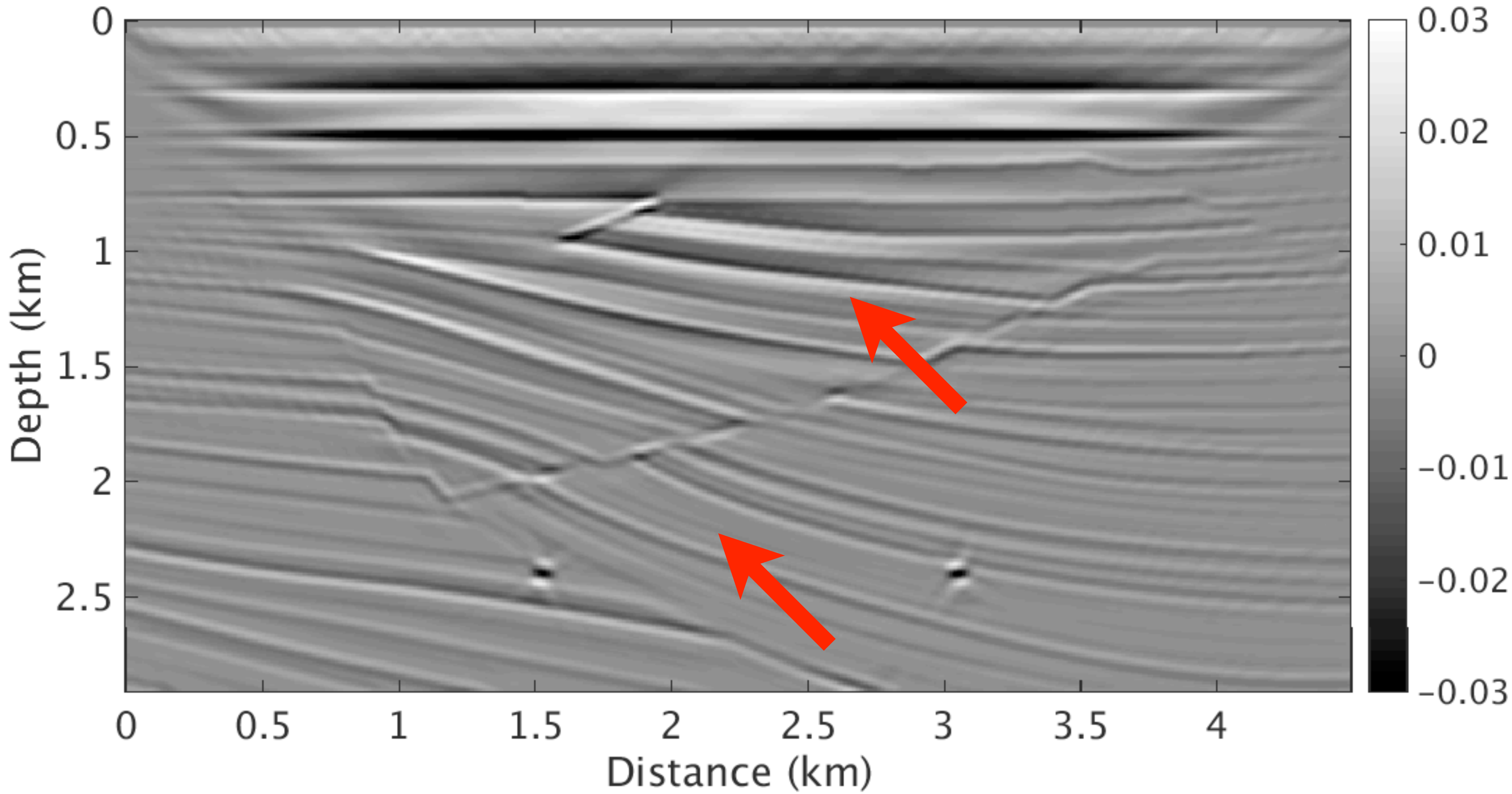
Robustness of source estimation
starting w/ zero-phase wavelet



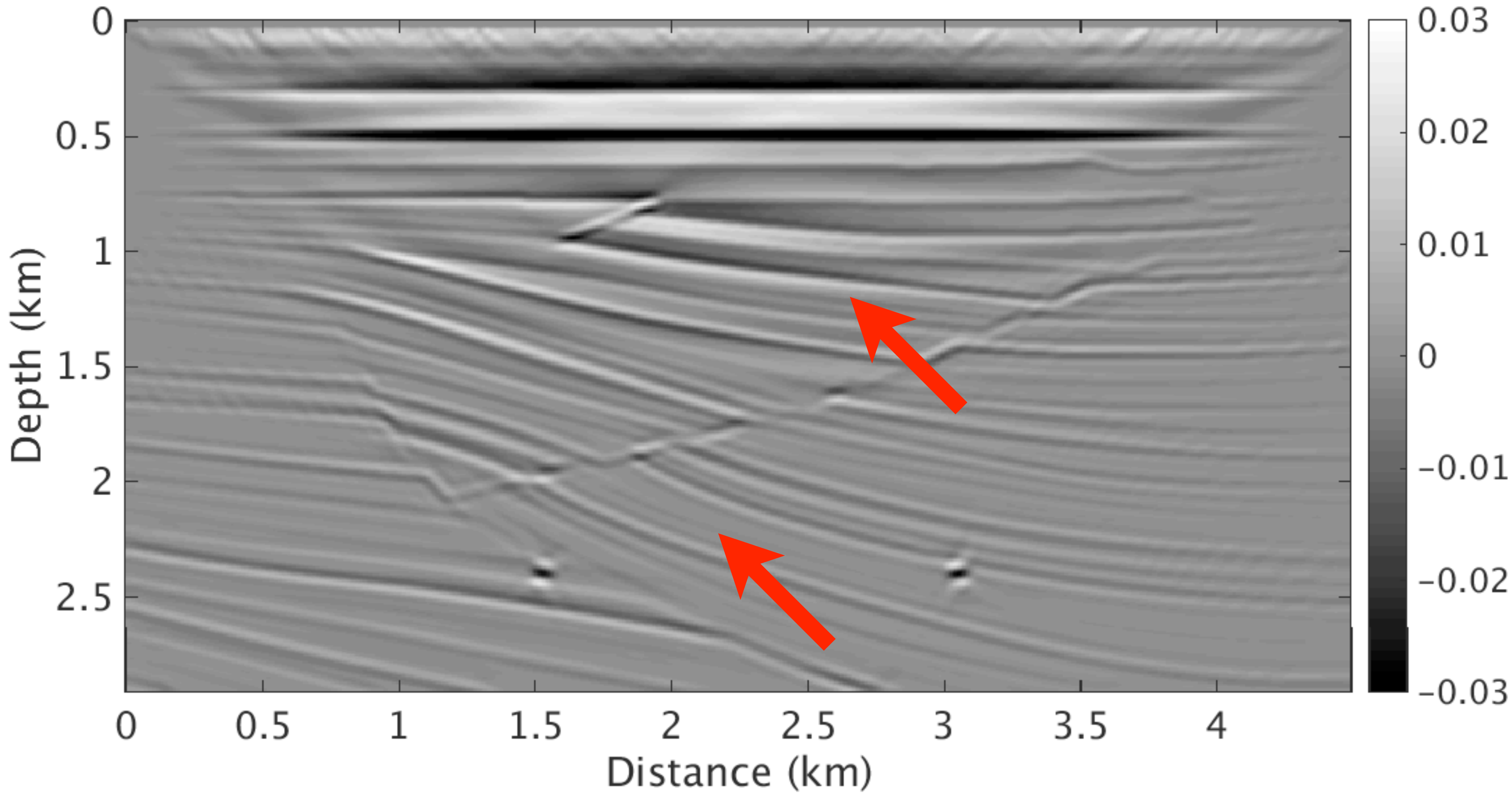
Robustness of source estimation
starting w/ zero-phase wavelet



Sparsity promoting LS-RTM
w/ correct wavelet & LB



Sparsity promoting LS-RTM
w/ source estimation & LB



Conclusions

- LB with correct source signature gives image with sharp interfaces w/ correct amplitudes
- Computational complexity is controlled to ~1 RTM w/ randomized source subsampling
- LB improves inversion results compared to other one-norm solvers
- LB can be combined w/ on-the-fly source estimation w/o a large computational overhead

SLIM

Future work

Accelerate LB algorithm with faster decades on dual variables and test the performance especially on salt models

Acknowledgements

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