Two methods for frequency down extrapolation

Rongrong Wang & Felix J. Herrmann

SLIM
University of British Columbia

SLIM

Outline

1. Motivation and extrapolation workflow

2. TV norm minimization

3. Lq norm minimization

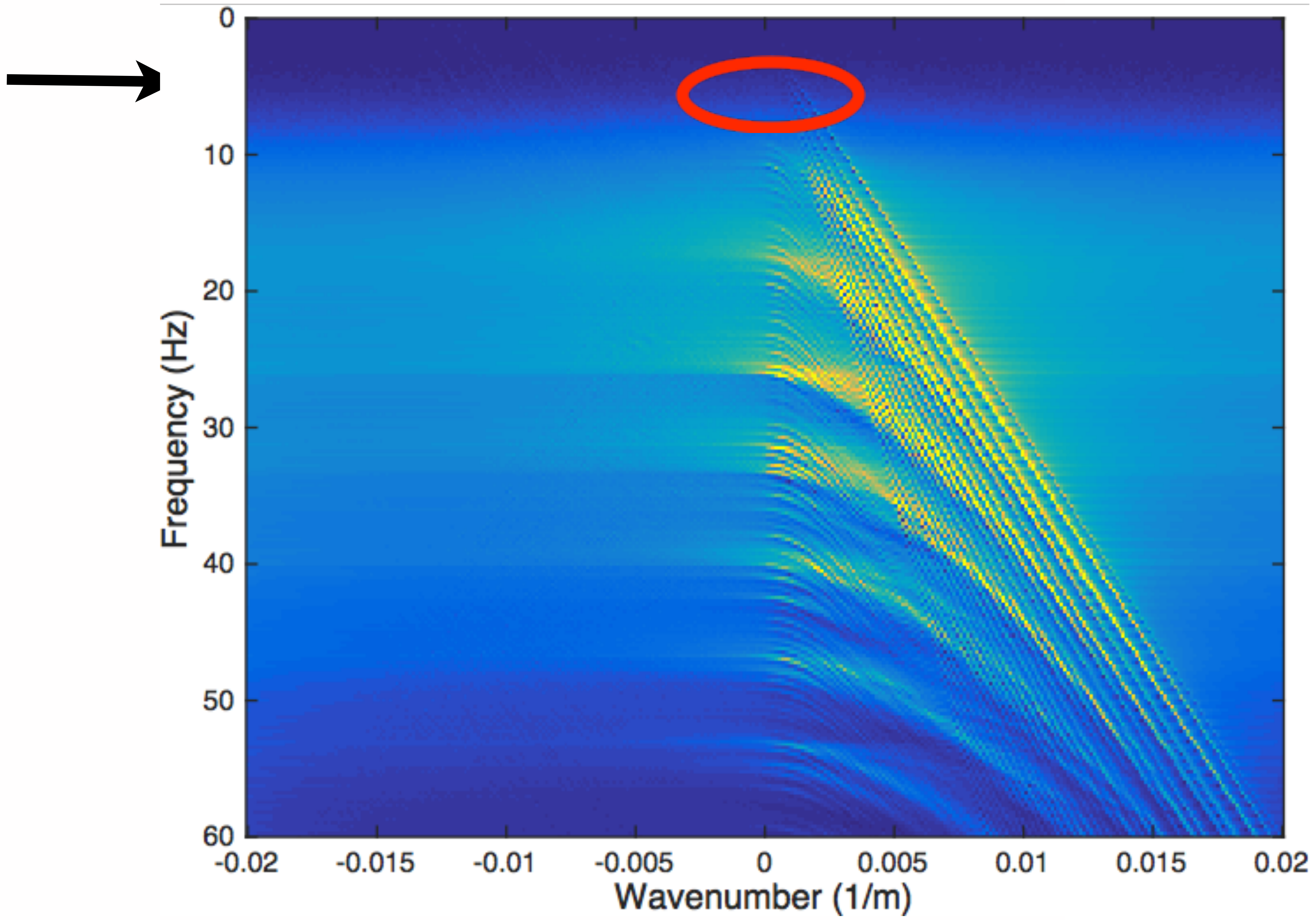
4. Numerical results

Motivation

Challenges in FWI:

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

A shot gather in the f-k domain Chevron (2014)



Our goal:

use mid-band data to extrapolate towards low frequencies

Convolutional model

Near offset trace

Source

$$\text{Trace} \leftarrow \mathbf{d}(t) = \mathbf{w}(t) * \mathbf{r}(t) \rightarrow \text{Reflectivity series}$$

Assume:

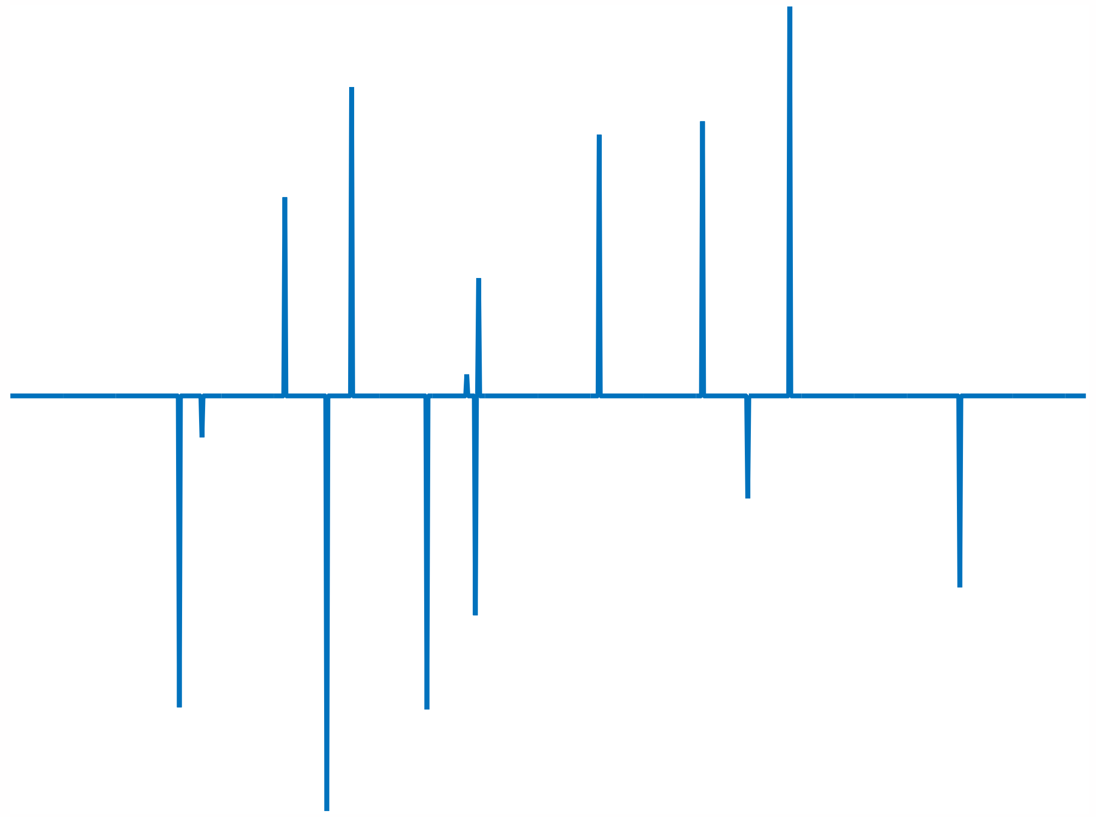
$$\mathbf{r}(t) = \sum_{i=1}^{s} a_i \delta_{t_i}(t)$$

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

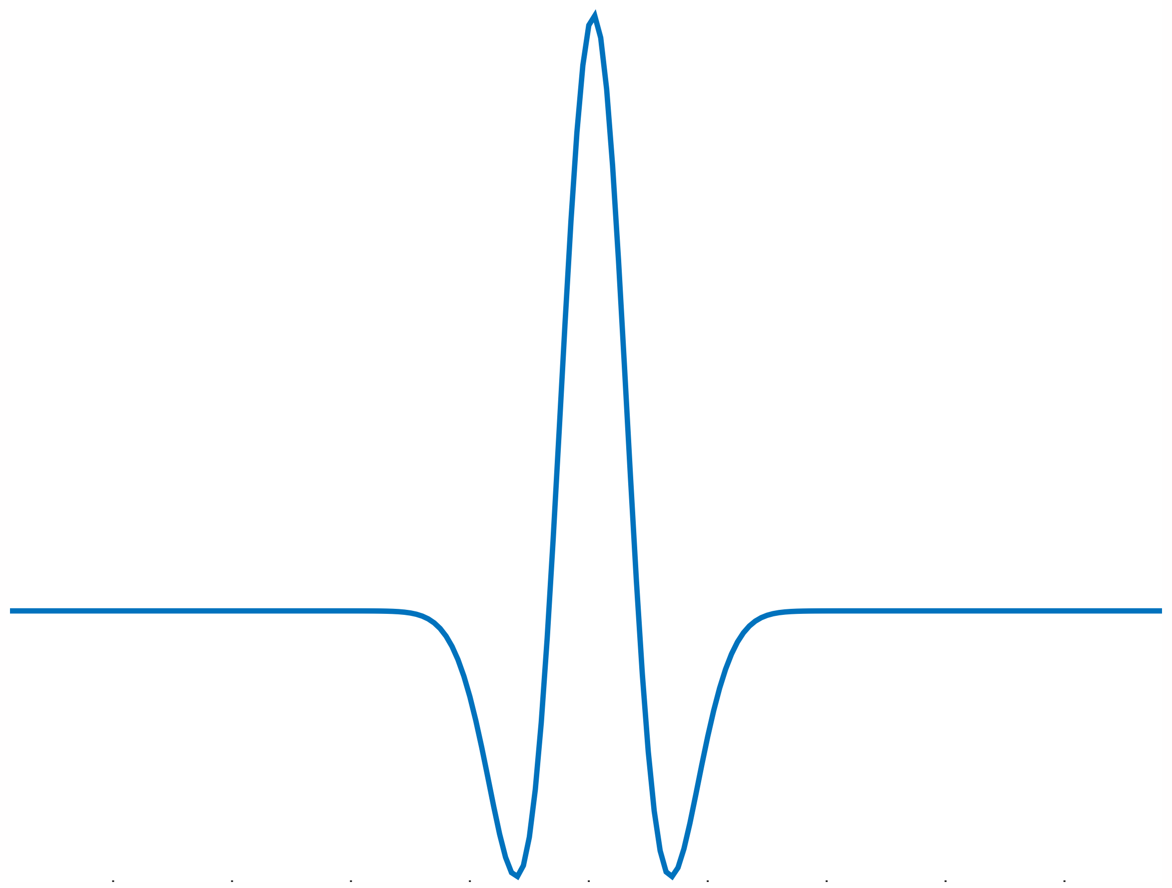
No assumptions on the wavelet

Workflow

Time domain reflectivity

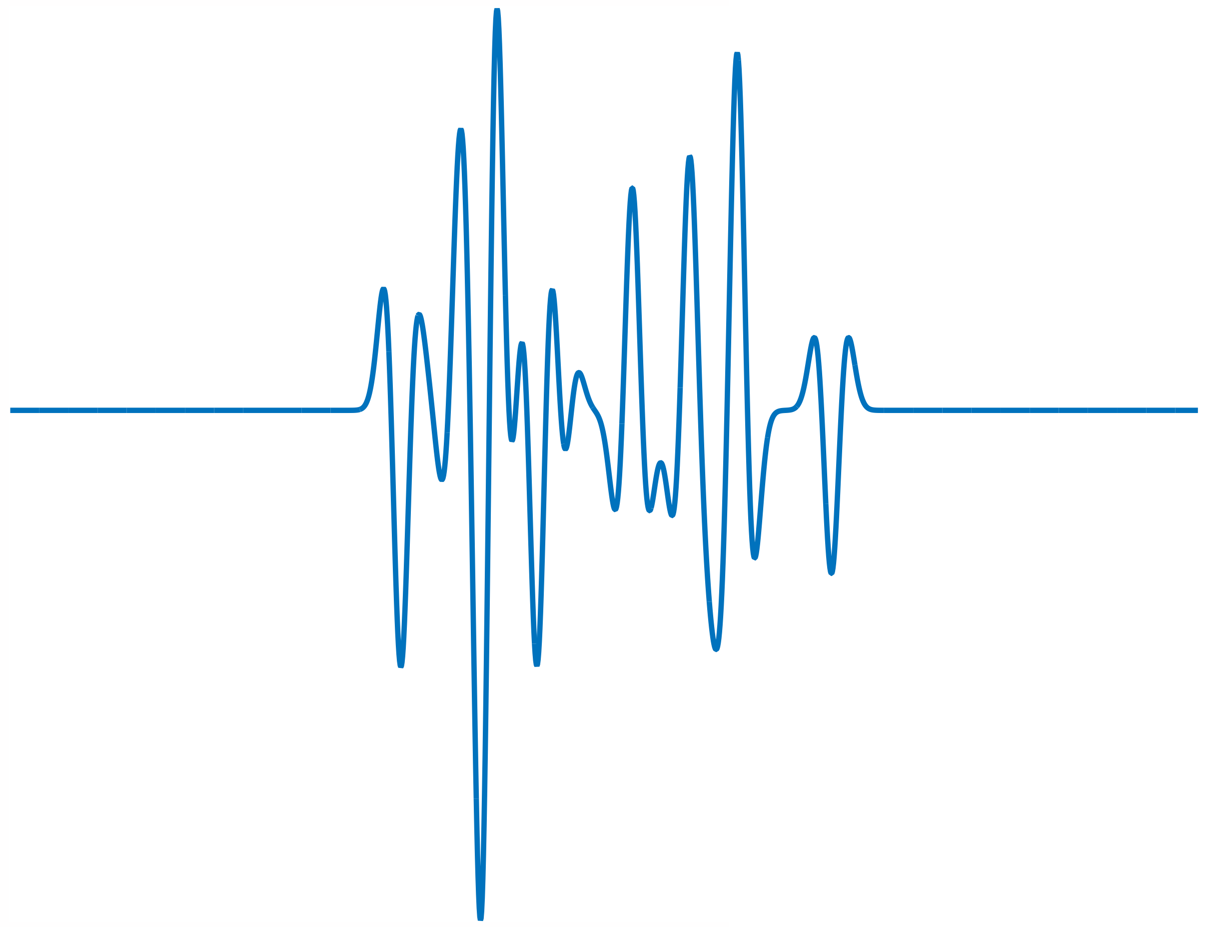


*

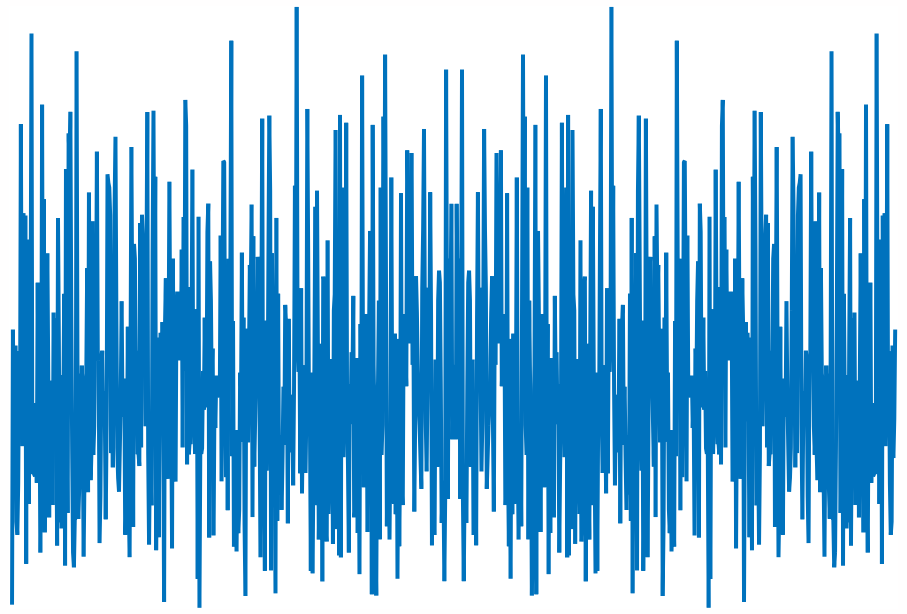


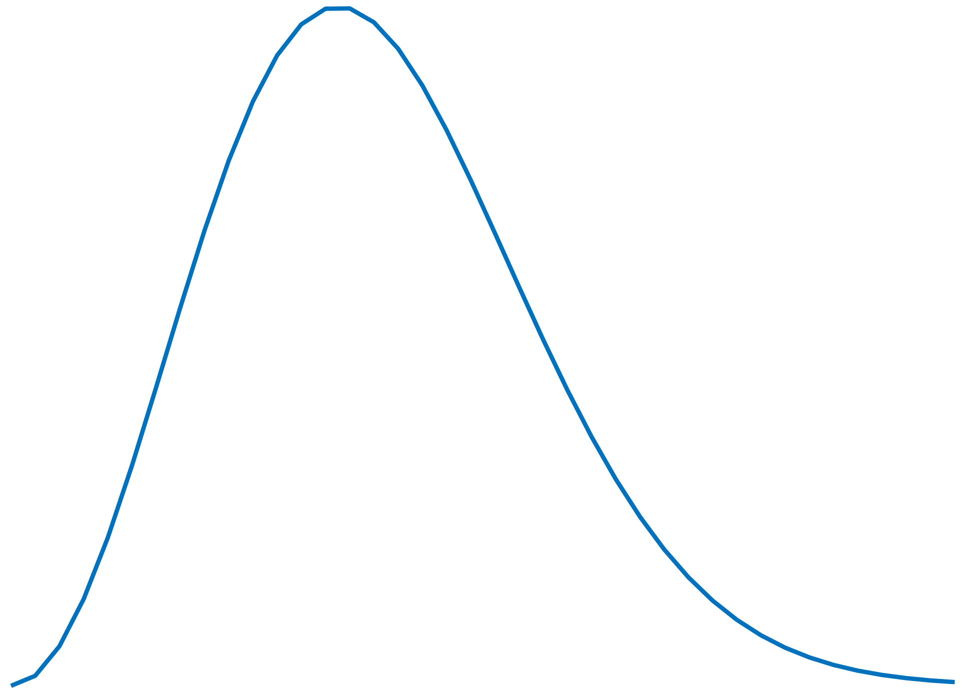
=

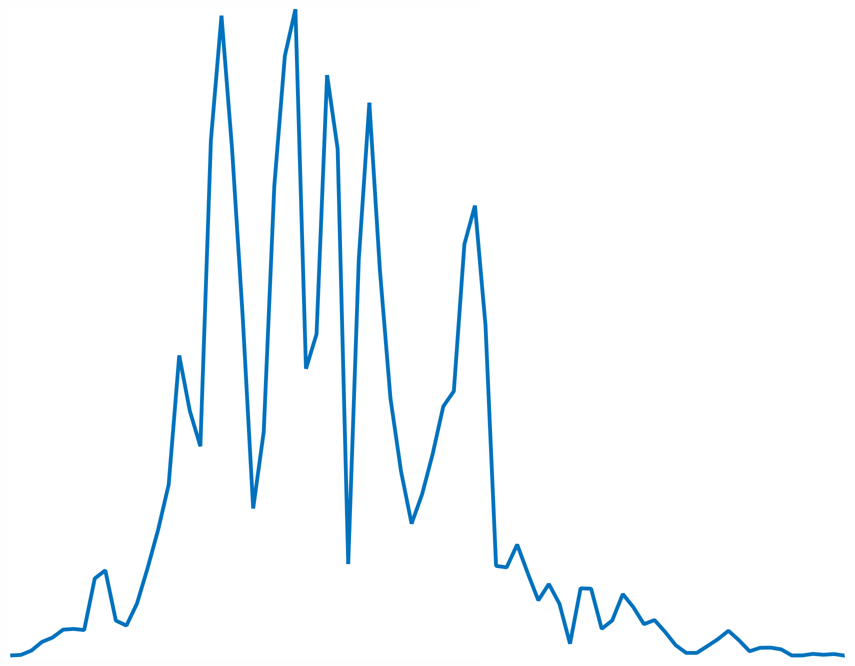
Noise-free data



Spectrum





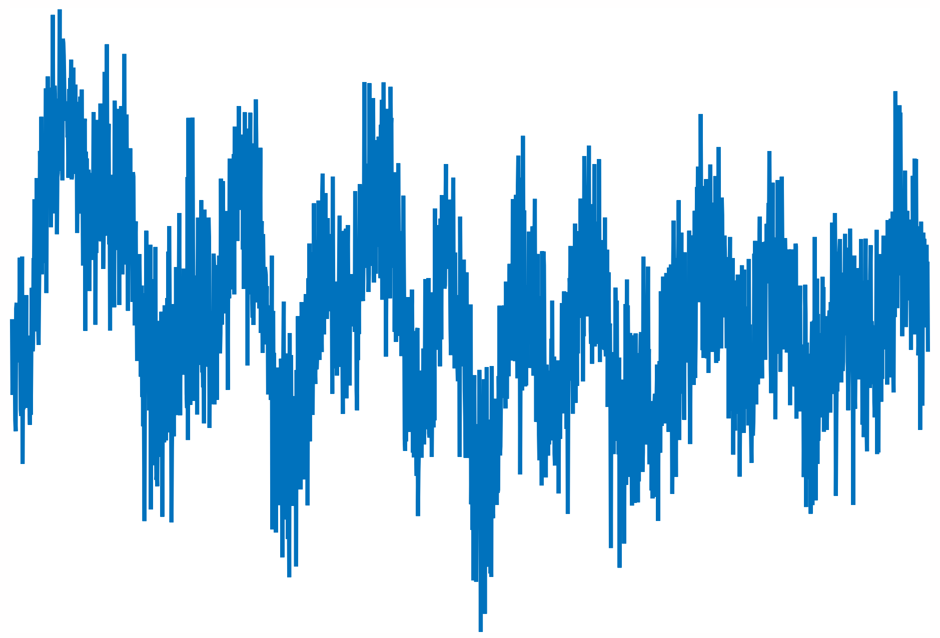


Workflow

Ideal time trace

\+

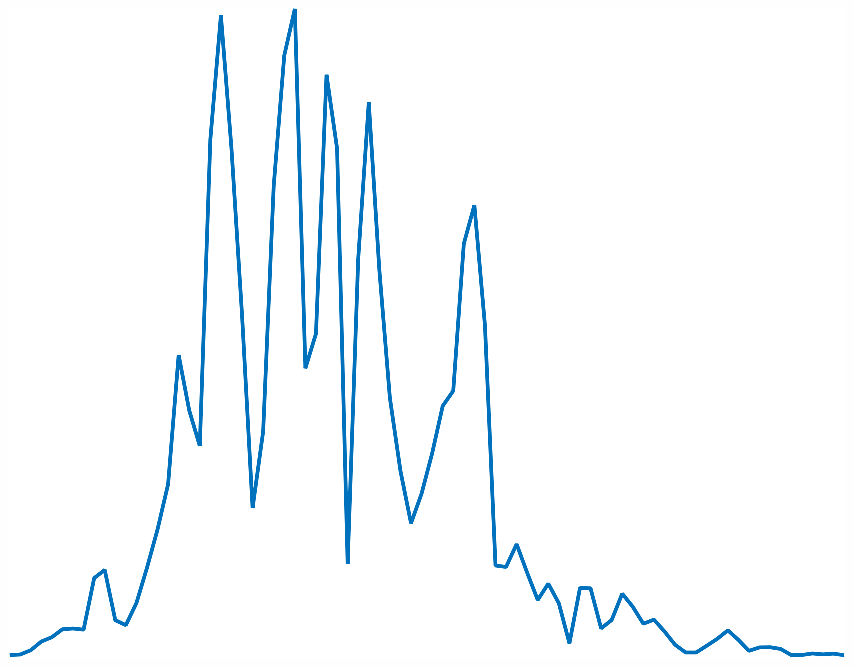
Low frequency contamination

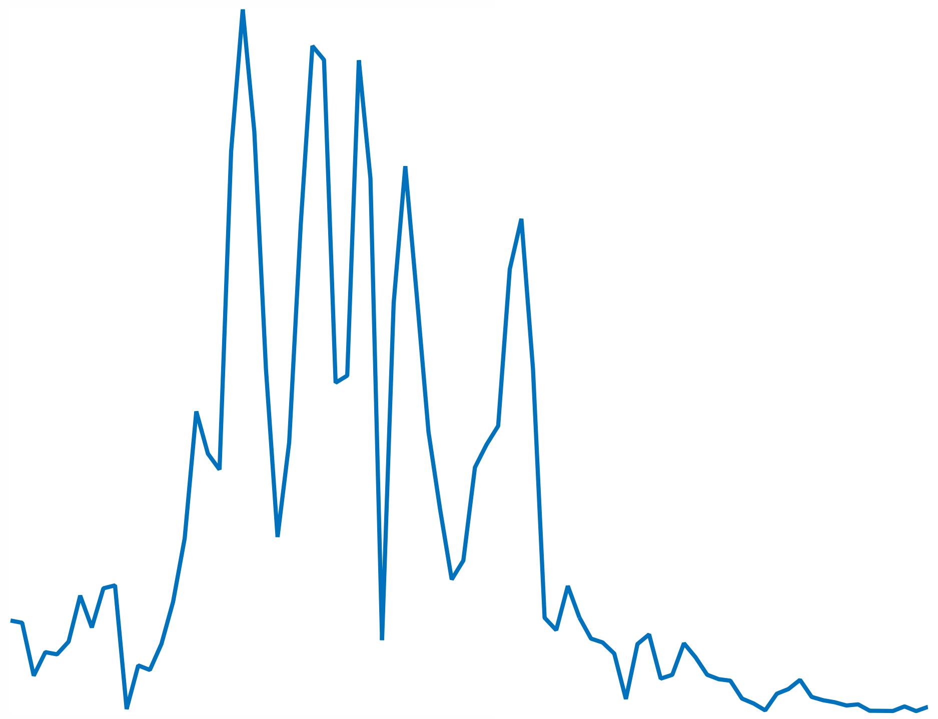


=

Observed data

Spectrum





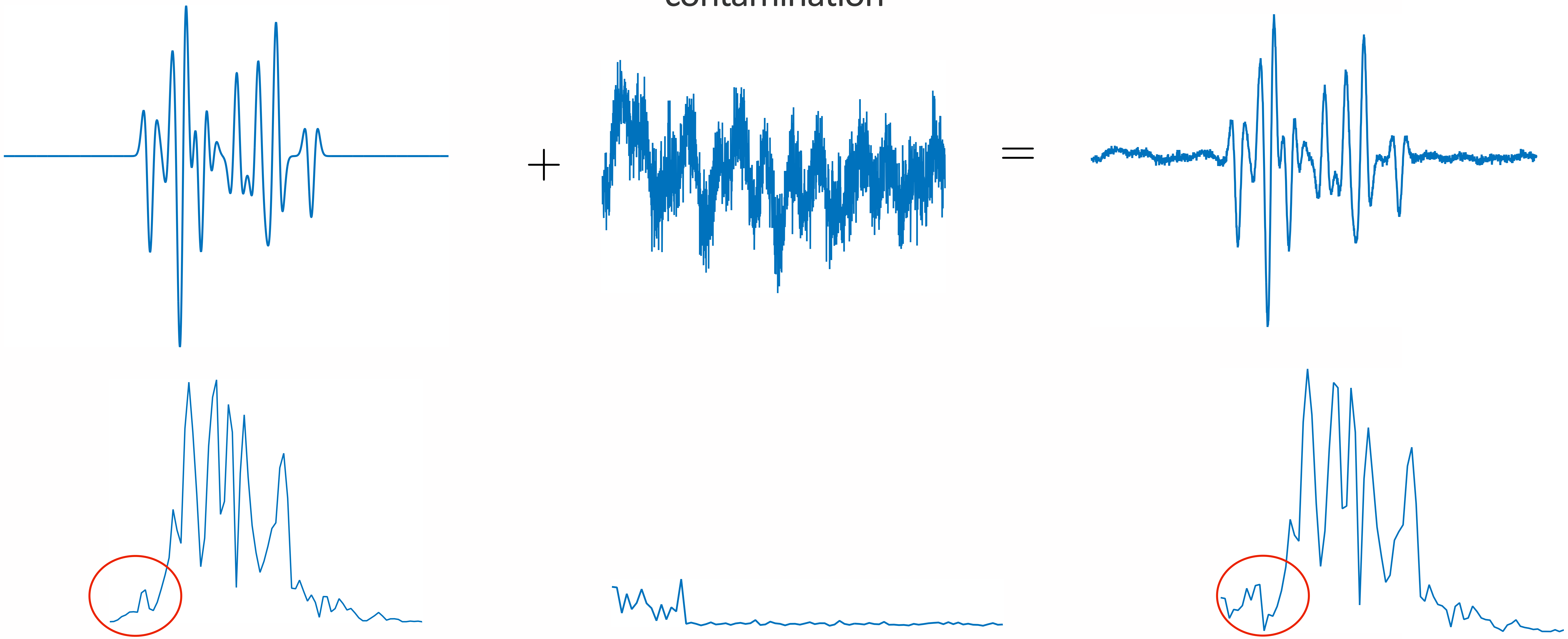
Workflow

Ideal time trace

Low frequency contamination

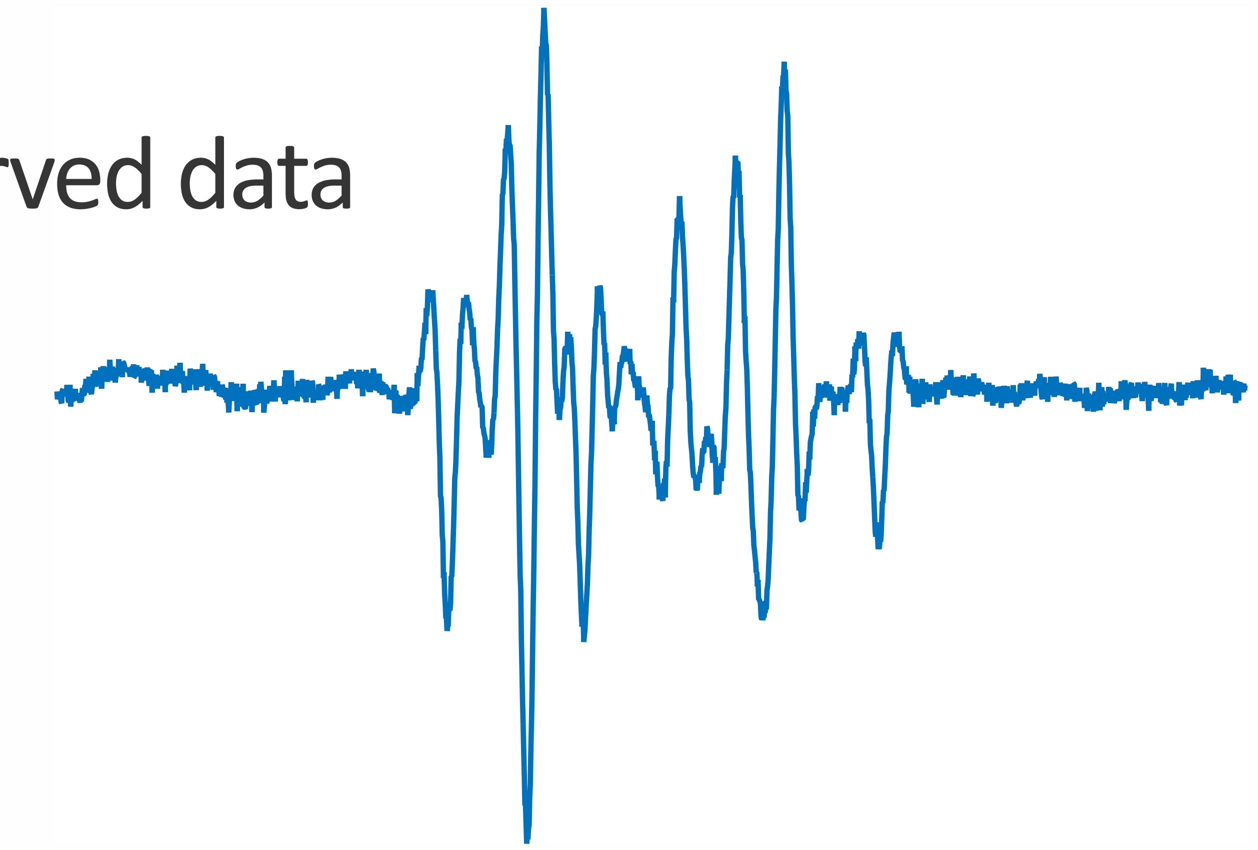
Observed data

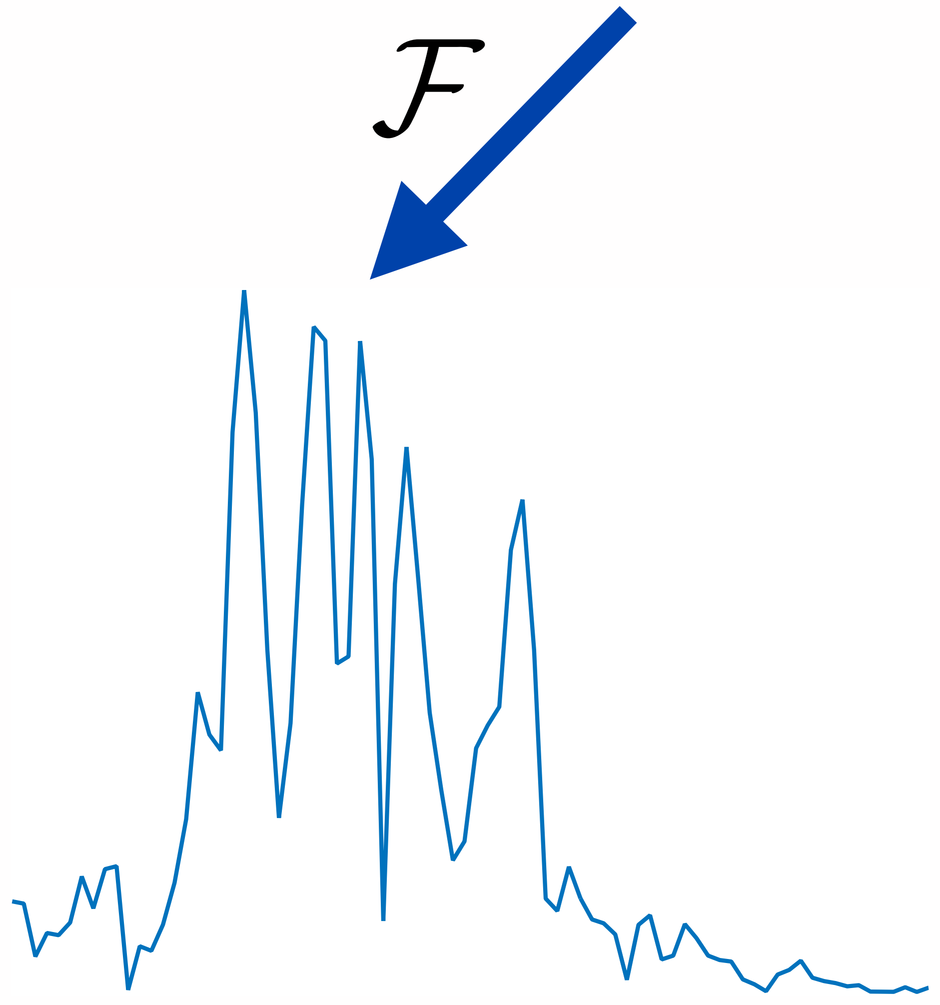
Spectrum



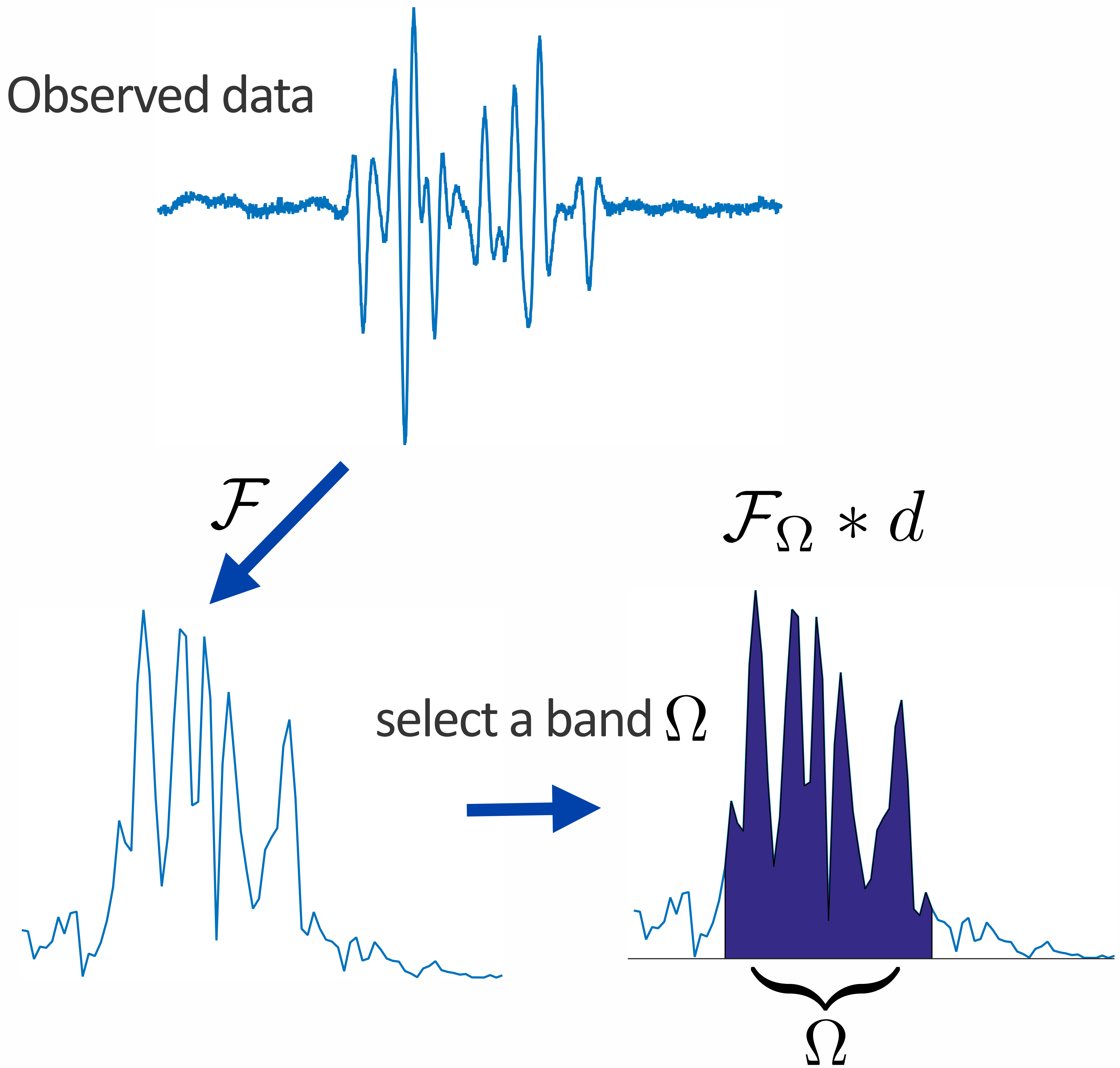
Workflow

Observed data



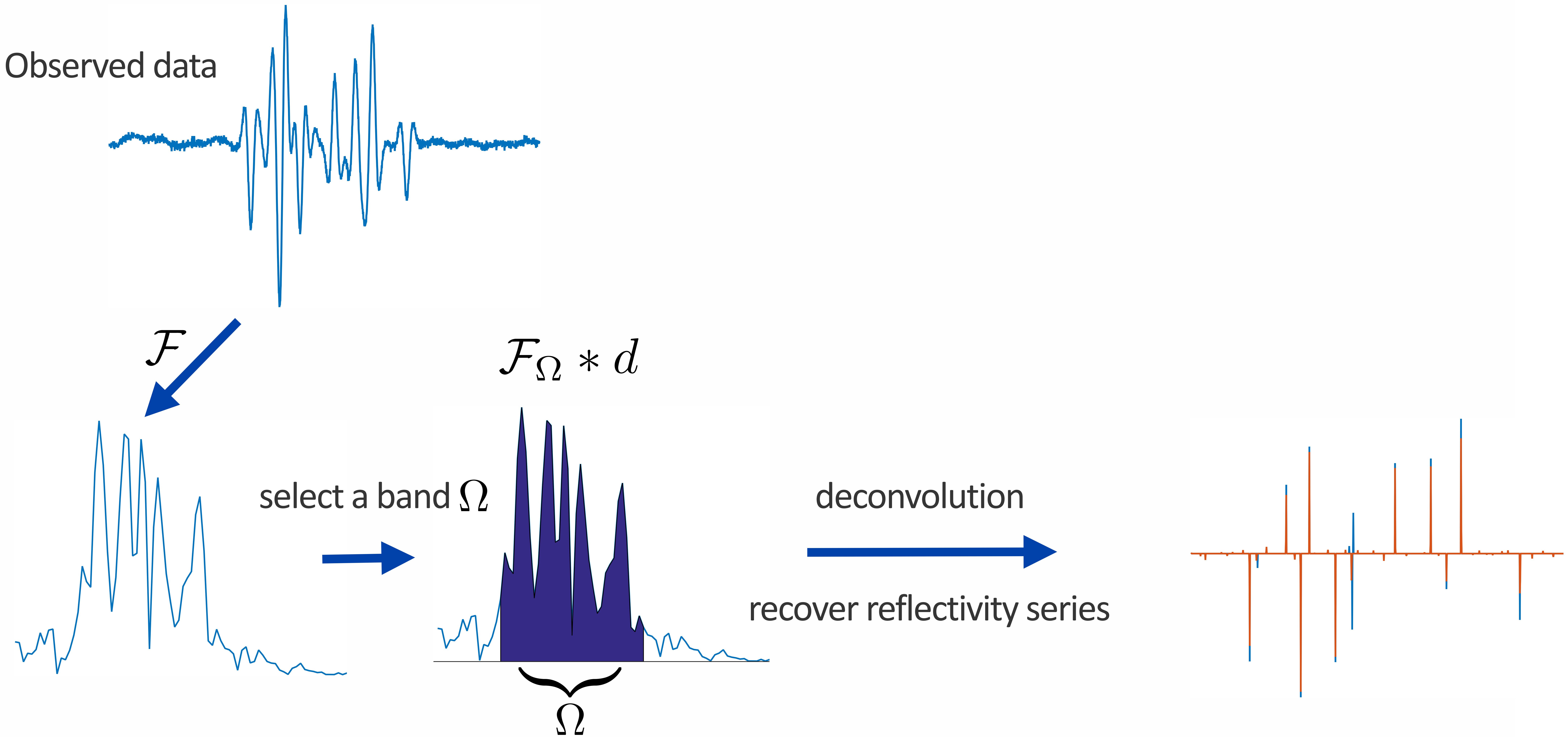


Workflow

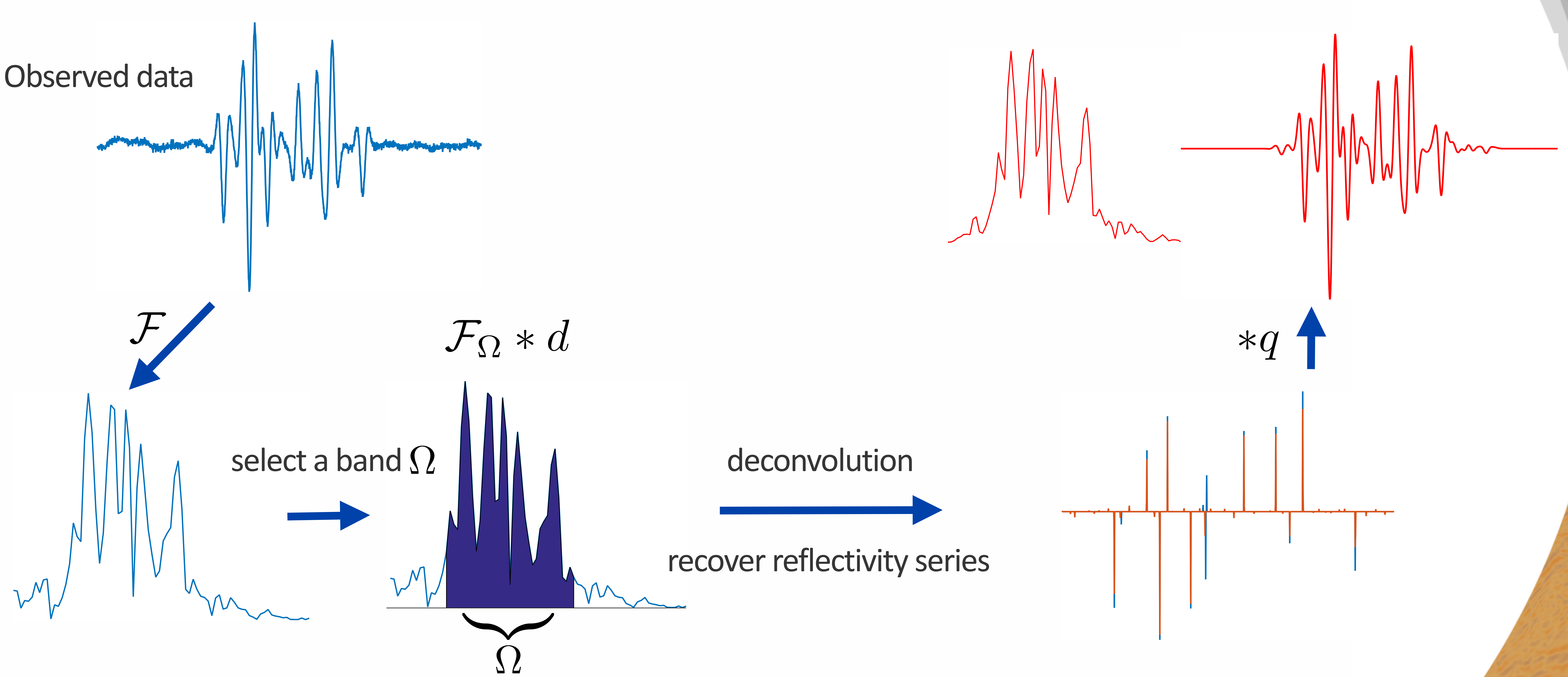


Observed data
$\mathcal{F}$
$\mathcal{F}_\Omega * d$
select a band Ω
Ω

Workflow

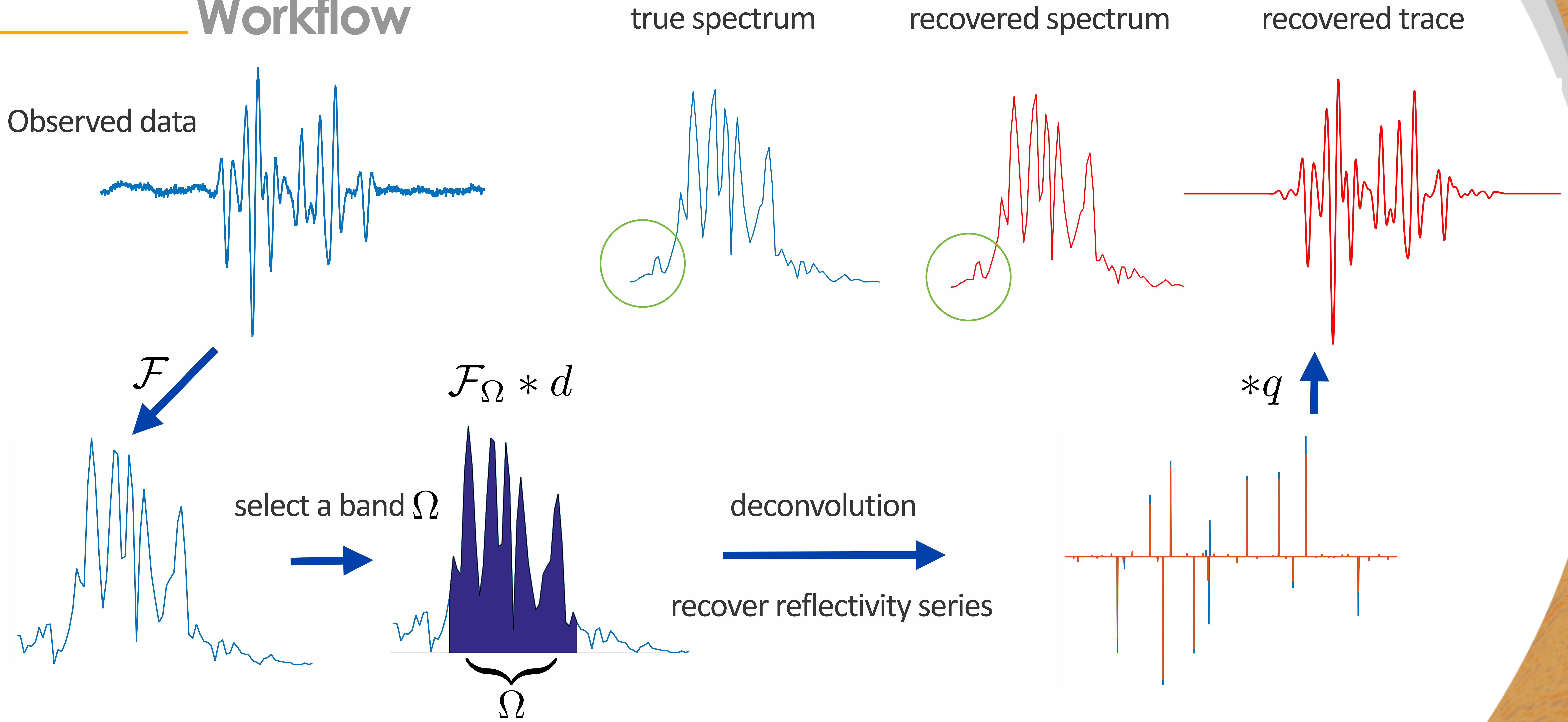

Observed data
$\mathcal{F}$
$\mathcal{F}_\Omega * d$
select a band Ω
deconvolution
recover reflectivity series
Ω

Workflow



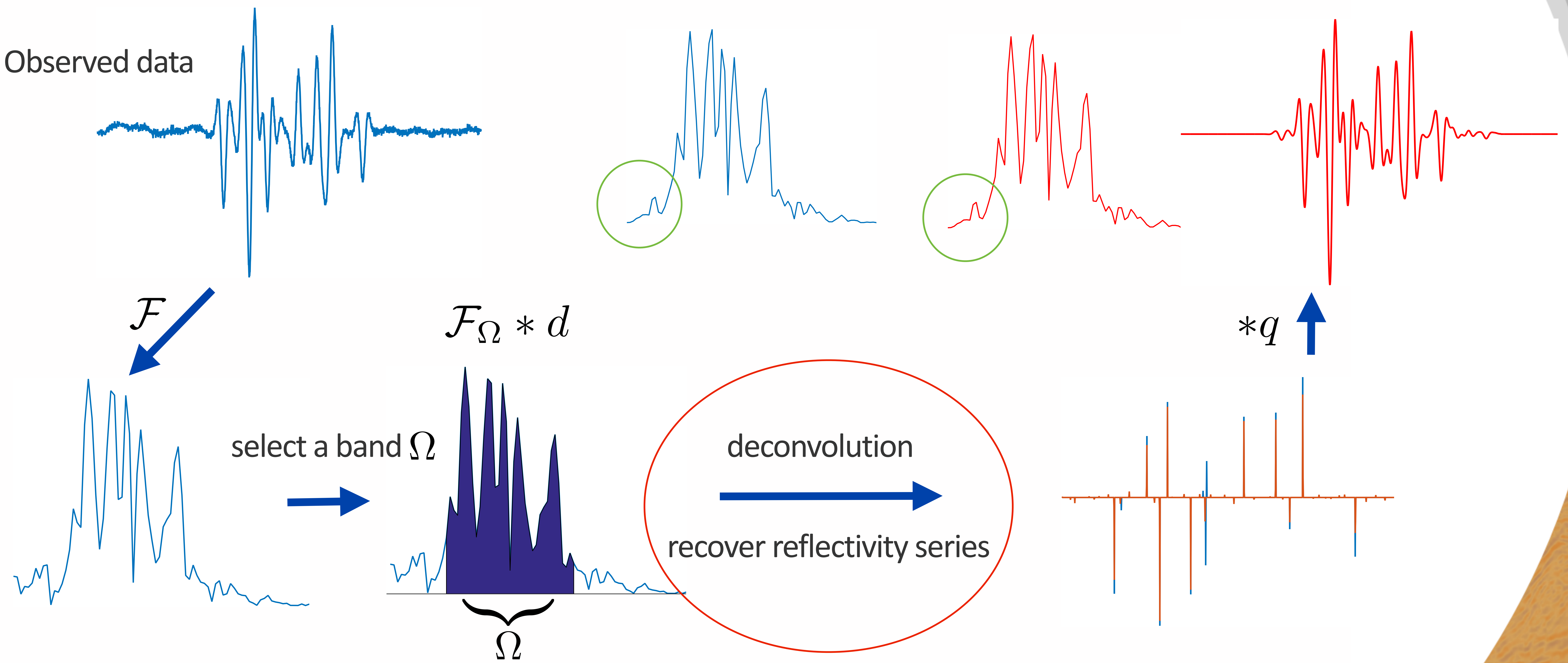
recovered spectrum
recovered trace
Observed data
$\mathcal{F}$
$\mathcal{F}_\Omega * d$
$*q$
select a band Ω
deconvolution
recover reflectivity series
Ω

Workflow



true spectrum
recovered spectrum
recovered trace
Observed data
$\mathcal{F}$
$\mathcal{F}_\Omega * d$
$*q$
select a band Ω
deconvolution
recover reflectivity series
Ω

Workflow


true spectrum
recovered spectrum
recovered trace
Observed data
$\mathcal{F}$
$\mathcal{F}_\Omega * d$
$*q$
select a band Ω
deconvolution
recover reflectivity series
Ω

[1] Schmidt, R.O, "Multiple Emitter Location and Signal Parameter Estimation," IEEE Trans. Antennas Propagation, Vol. AP-34 (March 1986), pp.276-280.

[2] H .L. Taylor, S. C. Banks, J. F. McCoy, "Deconvolution with the L1 norm." Geophysics (1979): 39-52.

[3] Candès, E. J., & Fernandez-Granda, C. (2013). Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, *19*(6), 1229-1254.

[4] Candès, Emmanuel J., and Carlos Fernandez−Granda. "Towards a Mathematical Theory of Super−resolution." *Communications on Pure and Applied Mathematics* 67.6 (2014): 906-956.

Two approaches for 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 spikes

TV norm minimization: a stabilized version of L1minimization

SLIM

L1 minimization

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

$$\text{subject to } \mathcal{F}_\Omega(\mathbf{w} * \mathbf{r}) = \mathcal{F}_\Omega \mathbf{d}$$

$\mathbf{d}$: data

$\mathbf{w}$: wavelet

$\mathbf{r}$: reflectivity series

$\mathcal{F}_\Omega$: bandpass filter

$\Omega = [f_L, f_H]$: pass bands — user defined

[1] J.F. Clear bout and F.Muir. “Robust modelling of erratic data“, Geophysics, (1973), 826-844
[2] H .L. Taylor, S. C. Banks, J. F. McCoy, “Deconvolution with the L1 norm.” Geophysics (1979): 39-52.
[3] M. Rudelson and R. Vershynin. “On sparse reconstruction from Fourier and Gaussian measurements.” Communications on Pure and Applied Mathematics 61.8 (2008), 1025-1045.
[4] Candès, Emmanuel J., and Carlos Fernandez−Granda. "Towards a Mathematical Theory of Super−resolution." *Communications on Pure and Applied Mathematics* 67.6 (2014): 906-956.
[5] C. Dossal and S. Mallat. “Spare spike deconvolution with minimum scale”, Proc. SPASSR (2005)), 123-126

SLIM

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]
- Candes 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] G(t) t∈[0,1] 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 G(t) is proportional to the energy of the noise.

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

Theoretical foundation for low-frequency extrapolation

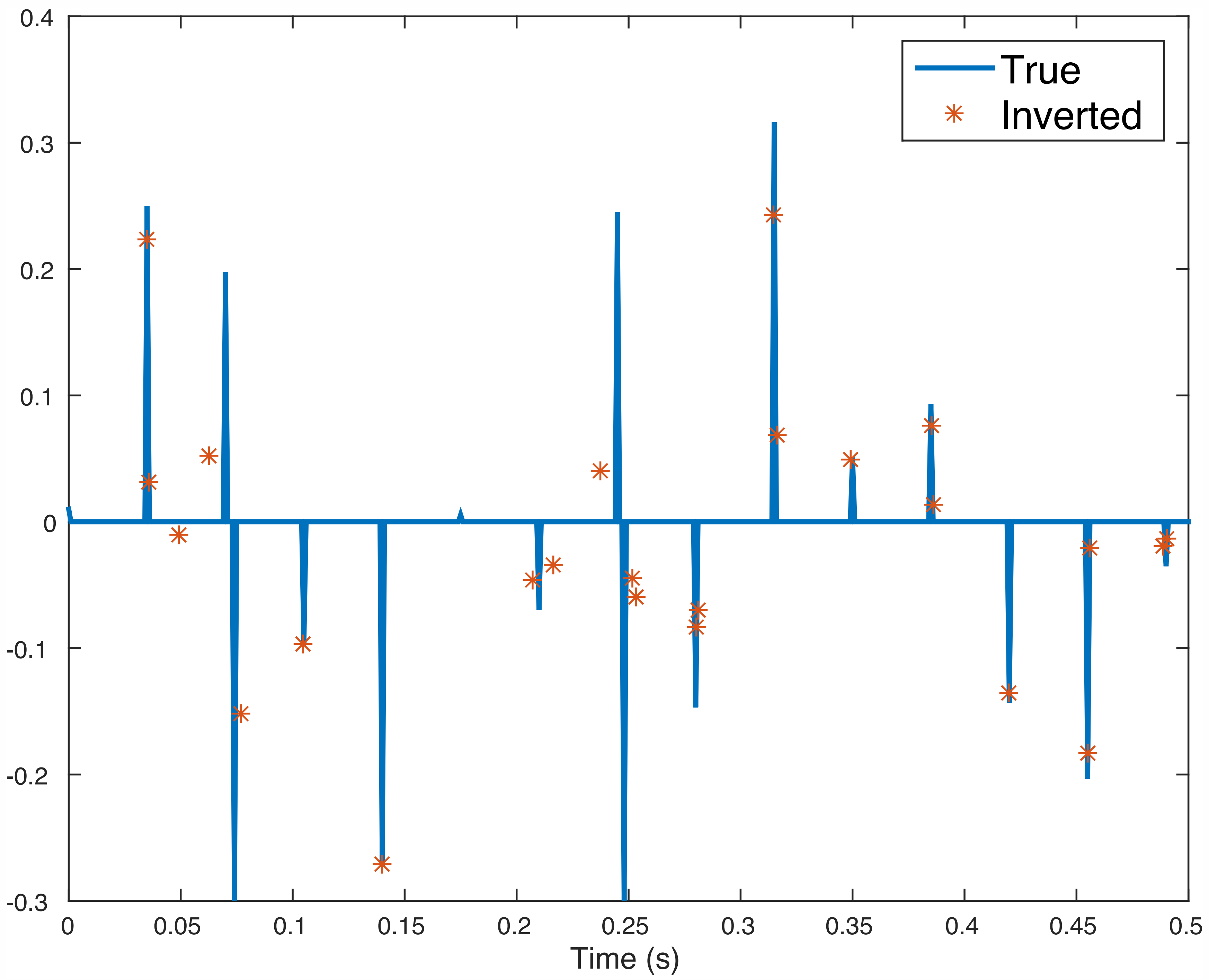
Theorem [RW,2016] G(t) t∈[0,1] can be exactly recovered by L1 minimization if the spikes are separated by 3.5 wavelengths* and the available bandwidth is greater than 60Hz. If noise exists, then the error in the estimate of G(t) is proportional to the energy of the noise.

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

*From numerical experiments: 1.5 wavelength is sufficient

When the minimal distance condition is not satisfied …

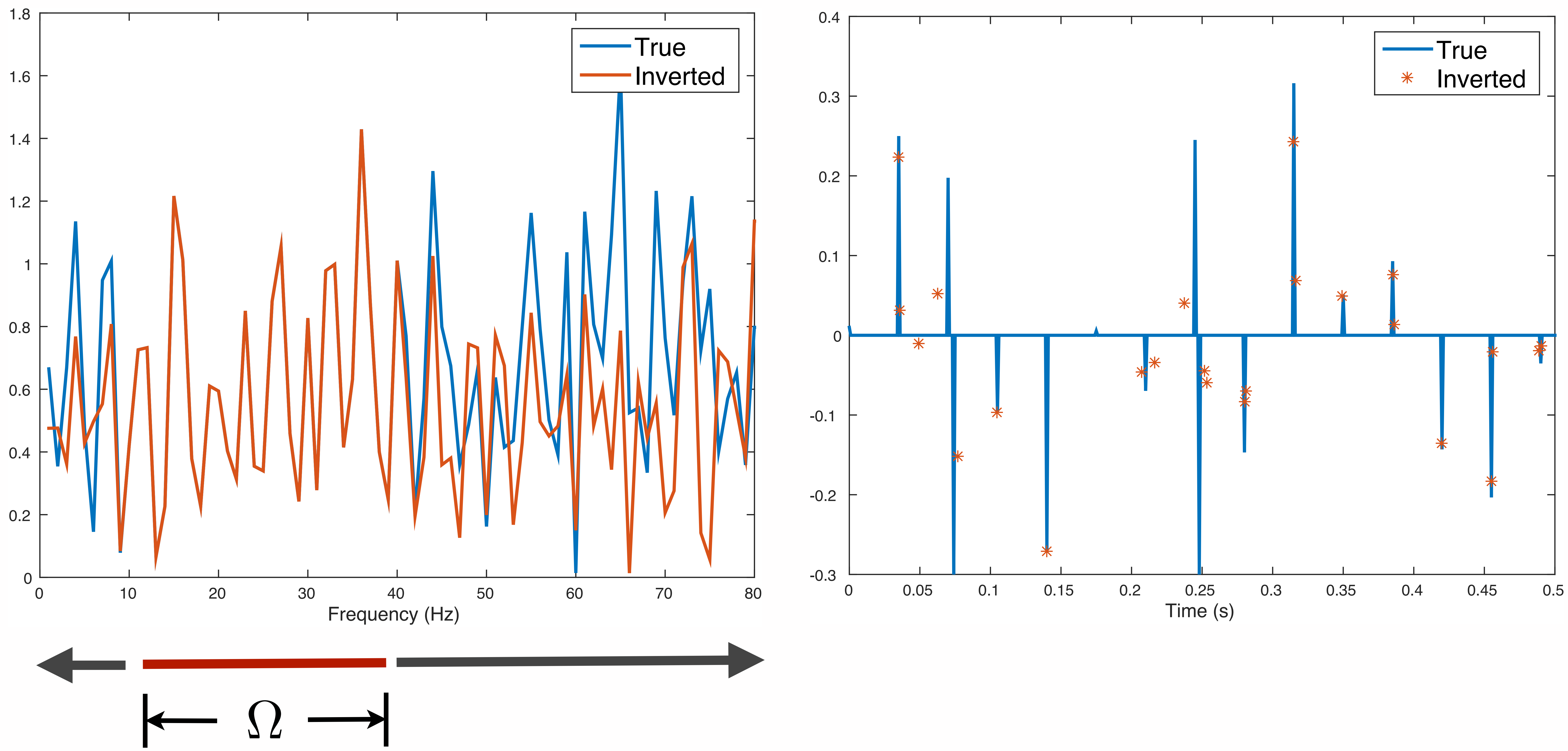
Reconstructions are affected



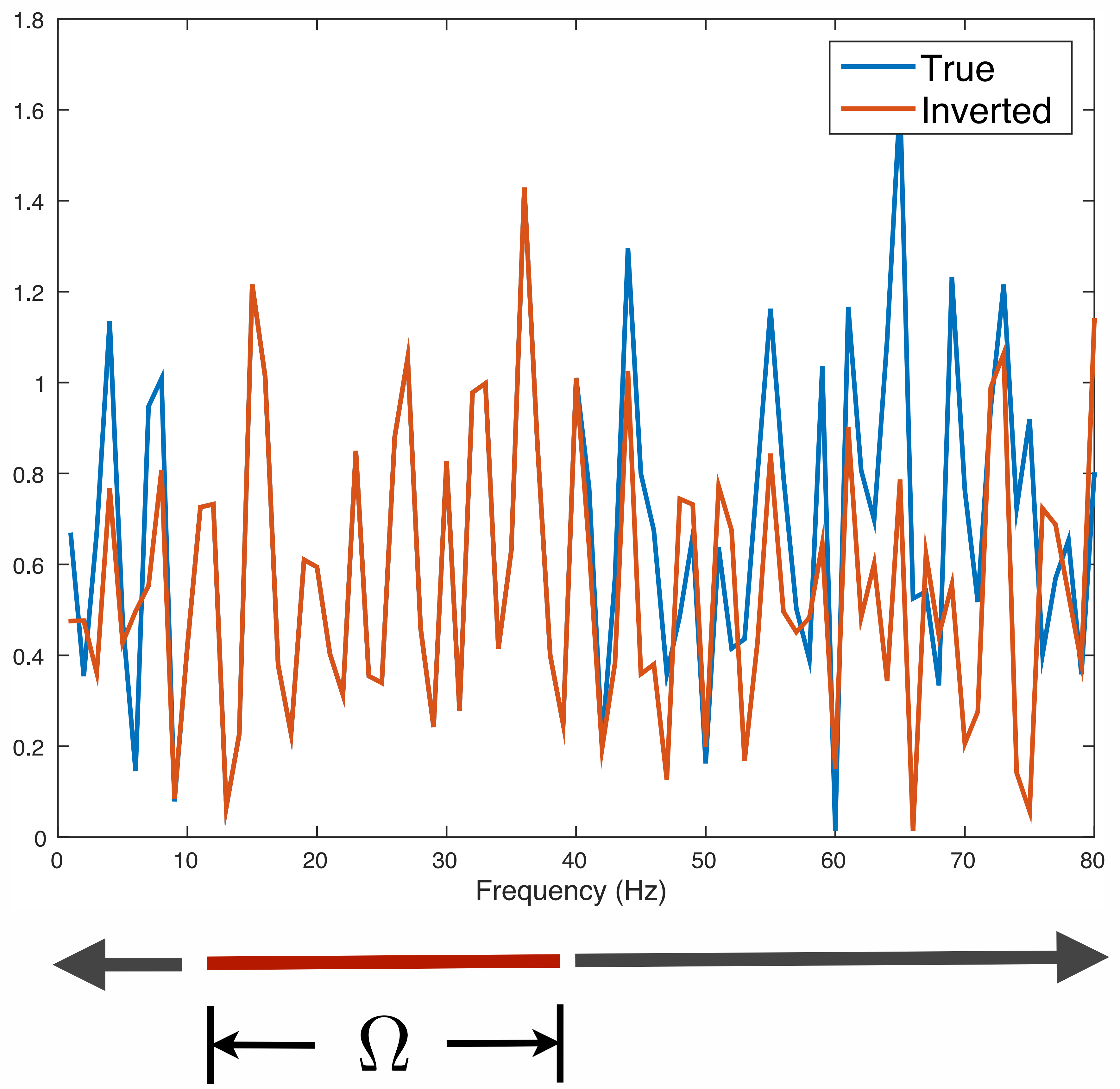
Left: L1 reconstruction of reflectivity series using frequencies $\Omega = \{10,\ldots,40\}$Hz

Deconvolution result is independent of wavelet choice

Error is proportional to distance to Ω



Error is proportional to distance to Ω



It can be shown that

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

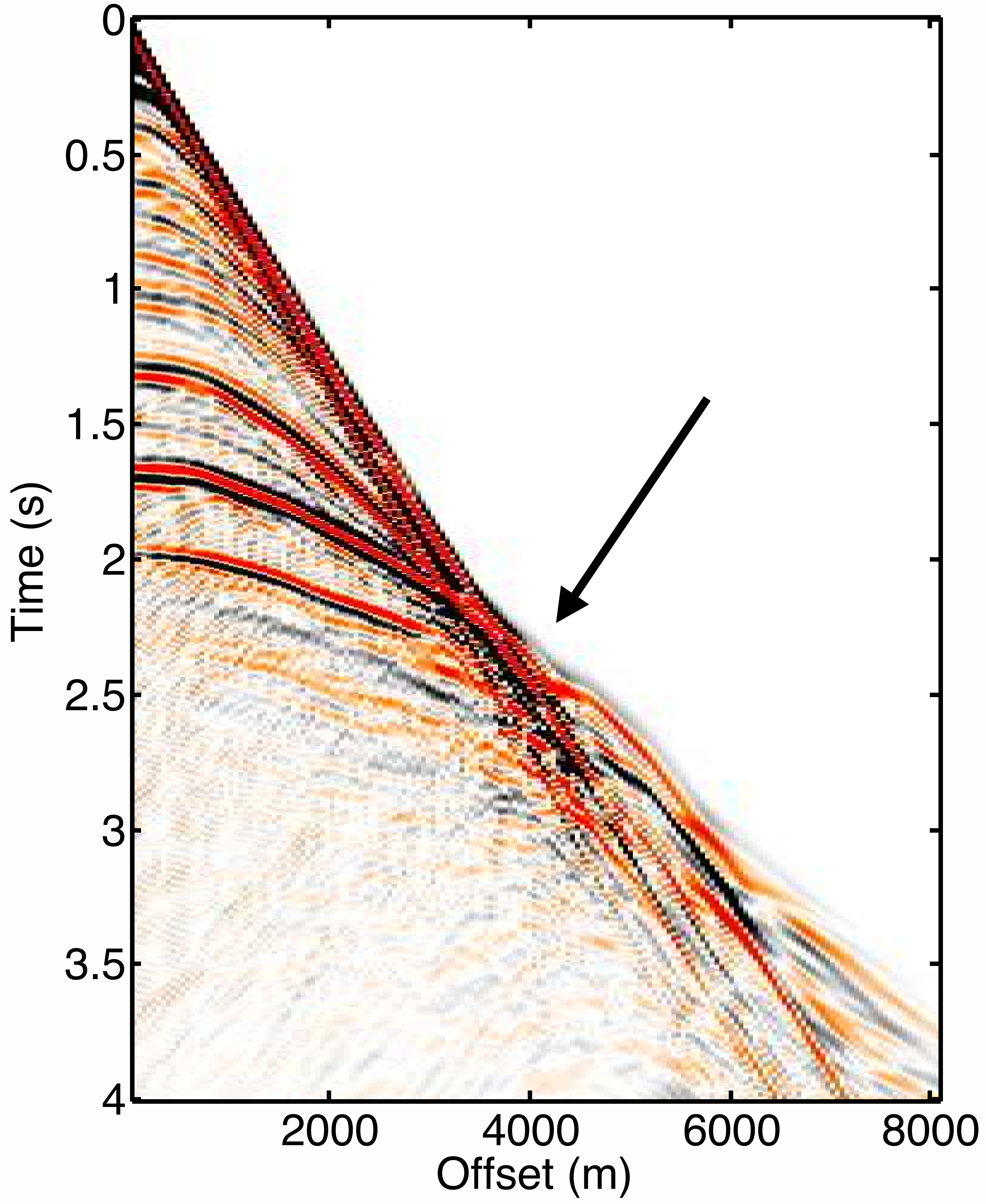
Difficulties in extrapolation

- 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

} noise

→ causes L1 to fail

The L1 minimization w/ TV-norm stabilizer

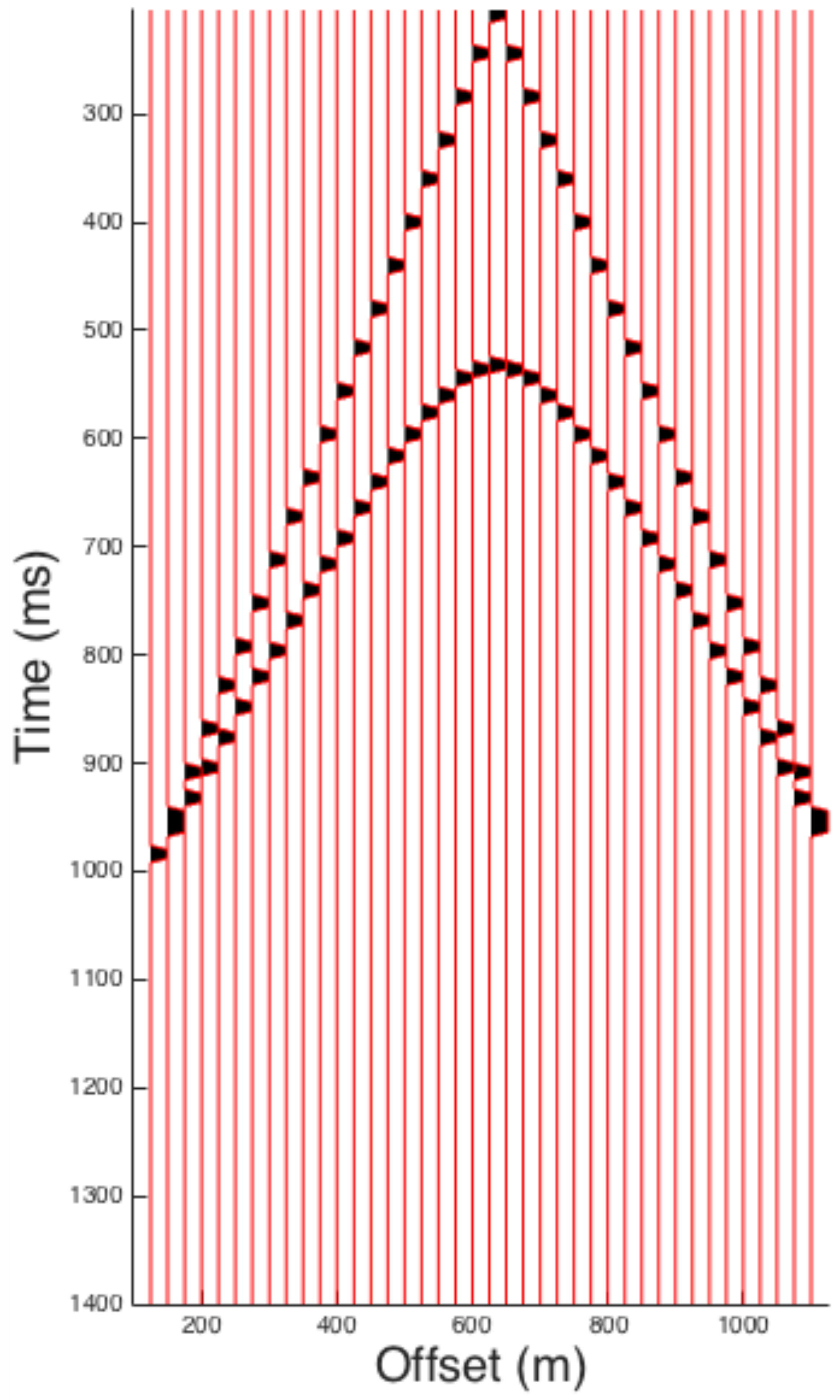


Conflicting events generate very close spikes

L1 minimization has trouble in the pointed region

Goal: utilize spatial correlations

Spatial similarity between adjacent traces



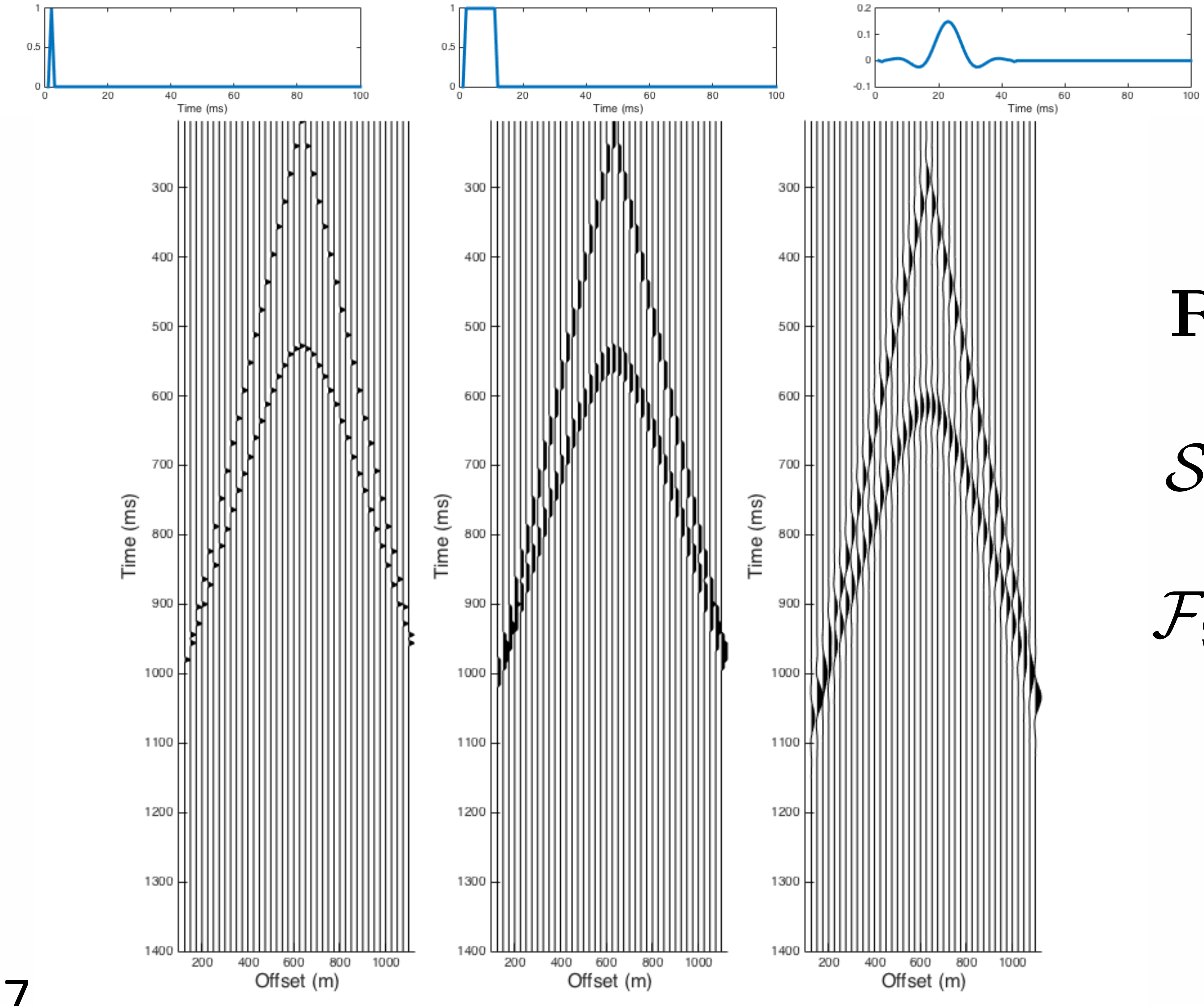
Define spatial similarity by

$$\mathcal{S}(\mathbf{R}) = \frac{\sum_{j=1}^{m-1} (\|\mathbf{R}_j\|_2^2 + \|\mathbf{R}_{j+1}\|_2^2)}{\sum_{j=1}^{m-1} \|\mathbf{R}_j - \mathbf{R}_{j+1}\|_2^2}$$

$\mathbf{R}$ has no spatial similarity when $\mathcal{S}(\mathbf{R}) \leq 1$

Left: visually continuous but has no spatial similarity

Increase spatial similarity by filtering



$$\mathbf{R}_m = \mathbf{f} * \mathbf{R}, \quad \text{where} \quad \mathbf{f} = (1, ..., 1, 0, ..., 0)$$

$$\mathcal{S}(\mathbf{R}_m) \geq \mathcal{S}(\mathbf{R})$$

$$\mathcal{F}_\Omega(\mathbf{R}_m) = \mathcal{F}_\Omega(\mathbf{R}) * \mathcal{F}_\Omega(\mathbf{f}) = \mathbf{d} * \mathcal{F}_\Omega(\mathbf{f})$$

TV norm minimization for Rm

NESTA is used to solve the following optimization problem

$$\mathbf{R}_{m,\text{est}} = \arg\min_{\mathbf{R}_m} \|\mathbf{R}_m\|_{TV}^{\alpha,\beta},$$

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

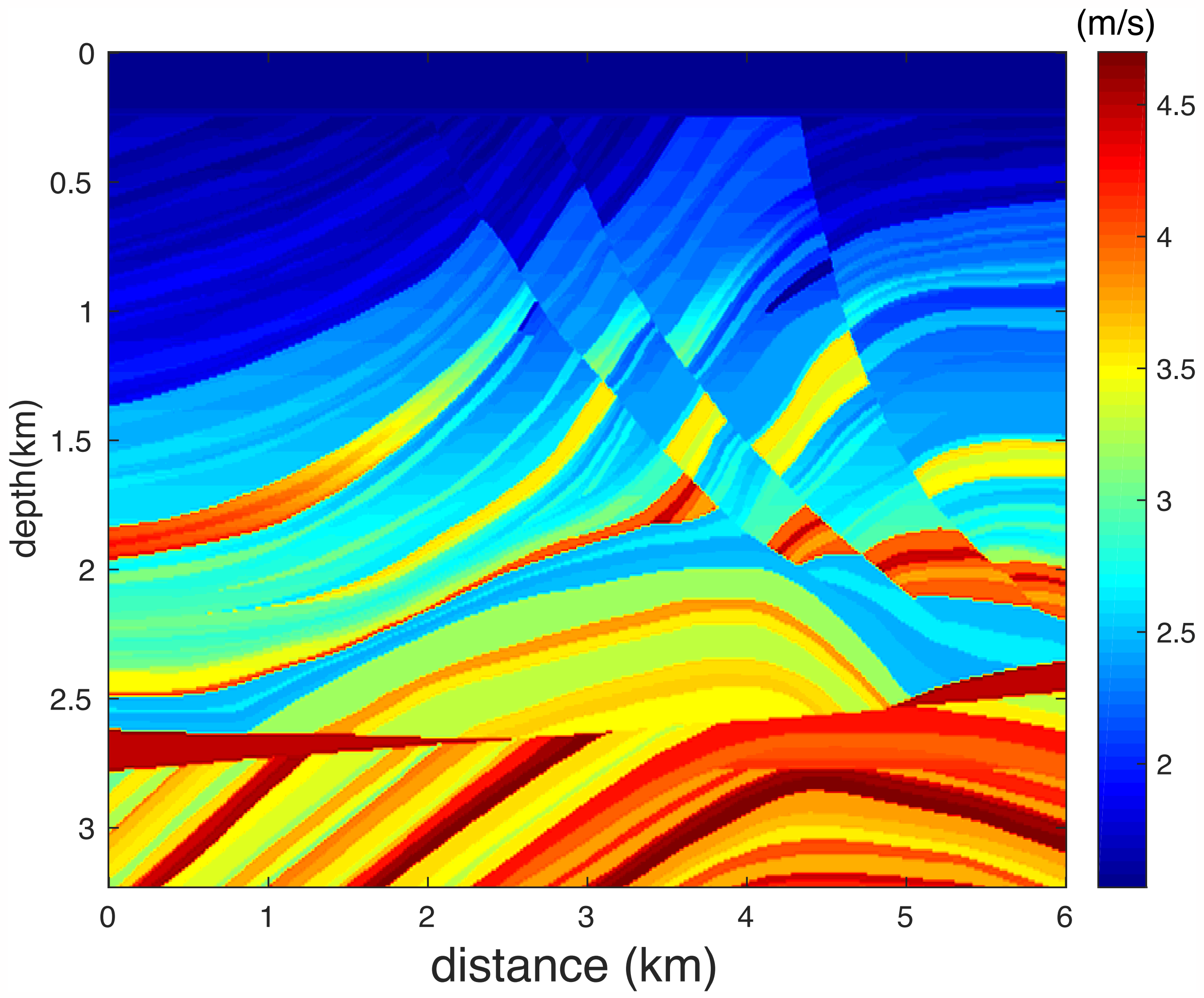
where

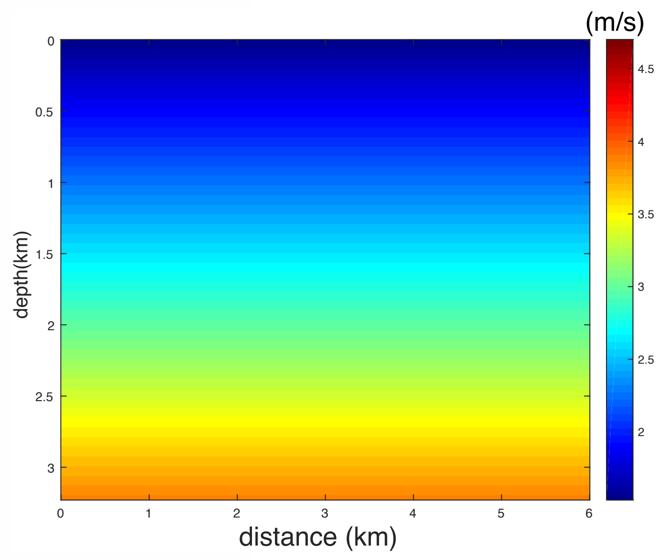
$$\|\mathbf{R}_m\|_{TV}^{\alpha,\beta} := \sum_{i,j} \|\nabla_{\alpha,\beta}\mathbf{R}_m(i,j))\|_{\ell_1}$$

$$\nabla_{\alpha,\beta}\mathbf{R}_m(i,j) = \begin{bmatrix} \alpha(\mathbf{R}_m(i,j) - \mathbf{R}_m(i,j+1)) \\ \beta(\mathbf{R}_m(i,j) - \mathbf{R}_m(i+1,j)) \end{bmatrix}.$$

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
- Ω = {5,...,15}Hz
- mute direct waves





Synthetic data - Non-inversion crime

Ω = {5,...,15}Hz

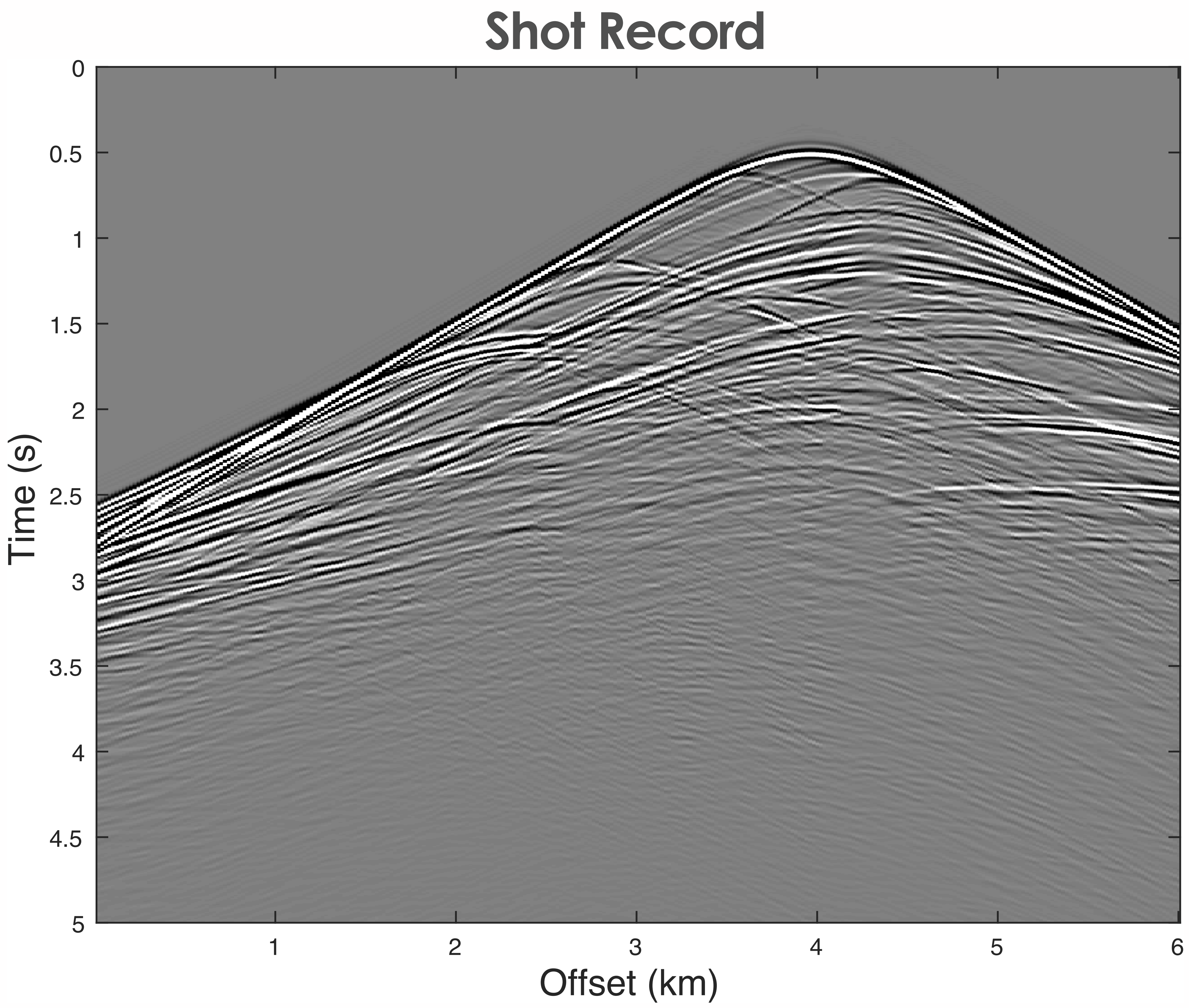
Direct inversion

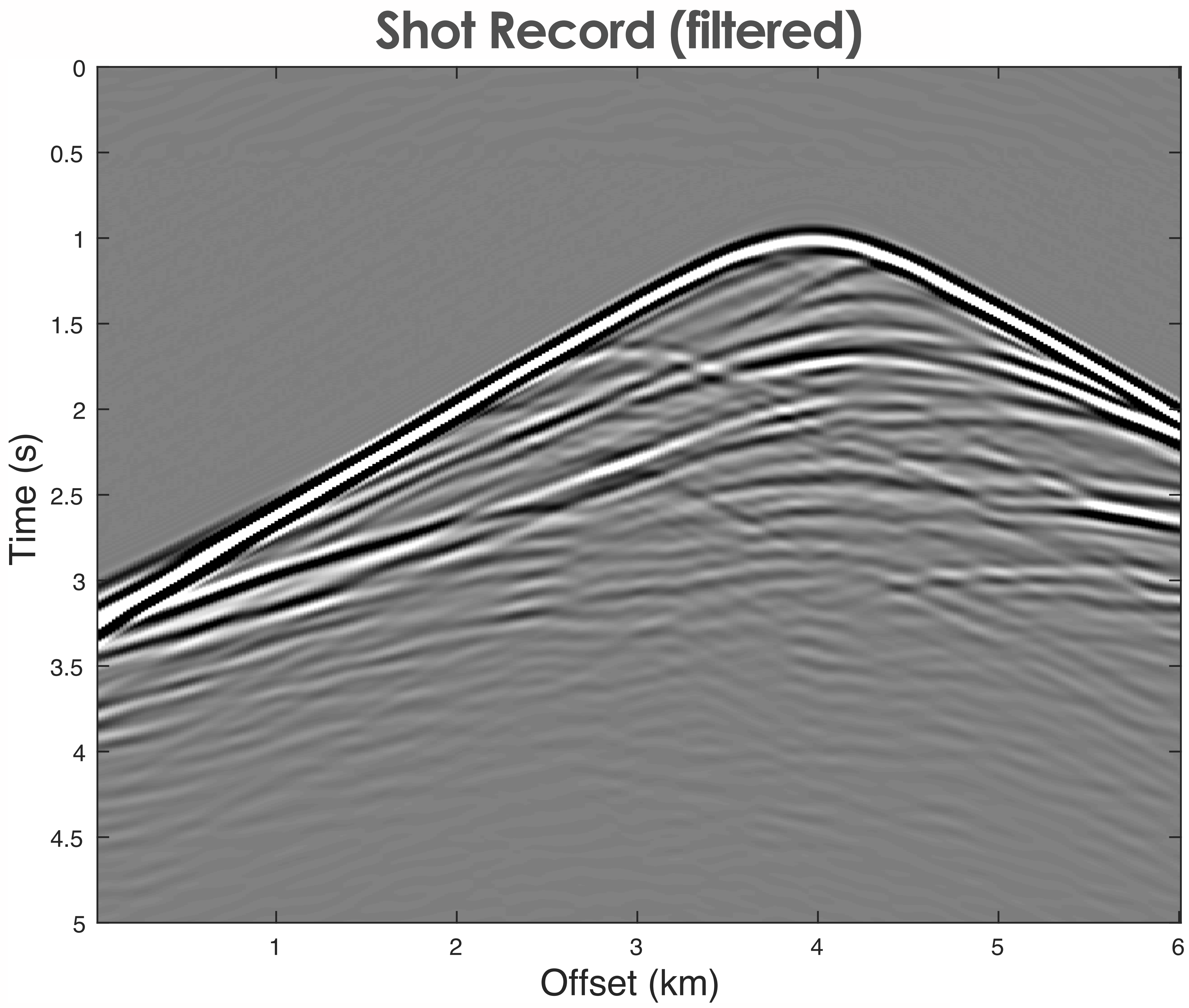
- frequency continuation with batches [5,5.25]Hz,[5.5,5.75]Hz, [6,6.25]Hz....[15,15.25]Hz
- perform the previous step two more sweeps

Inversion with extrapolation

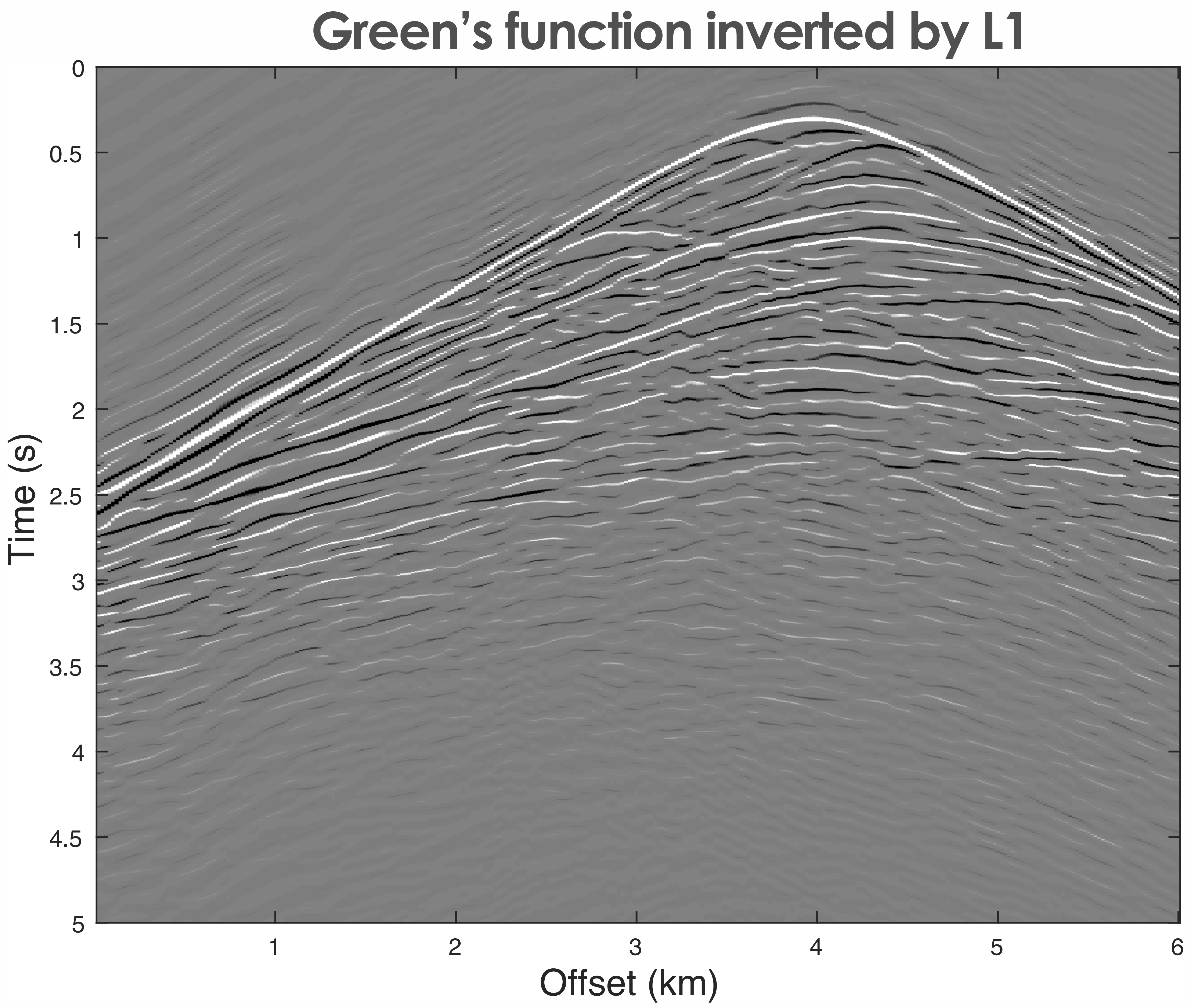
- extrapolation from {5,...,15}Hz to {1,...5}Hz
- frequency continuation with batches [1,1.25]Hz,[1.5,1.75]Hz,....[4.5,4.75]Hz
- frequency continuation with batches [5,5.25]Hz,[5.5,5.75]Hz, [6,6.25]Hz.... [15,15.25]Hz
- perform the previous step two more sweeps

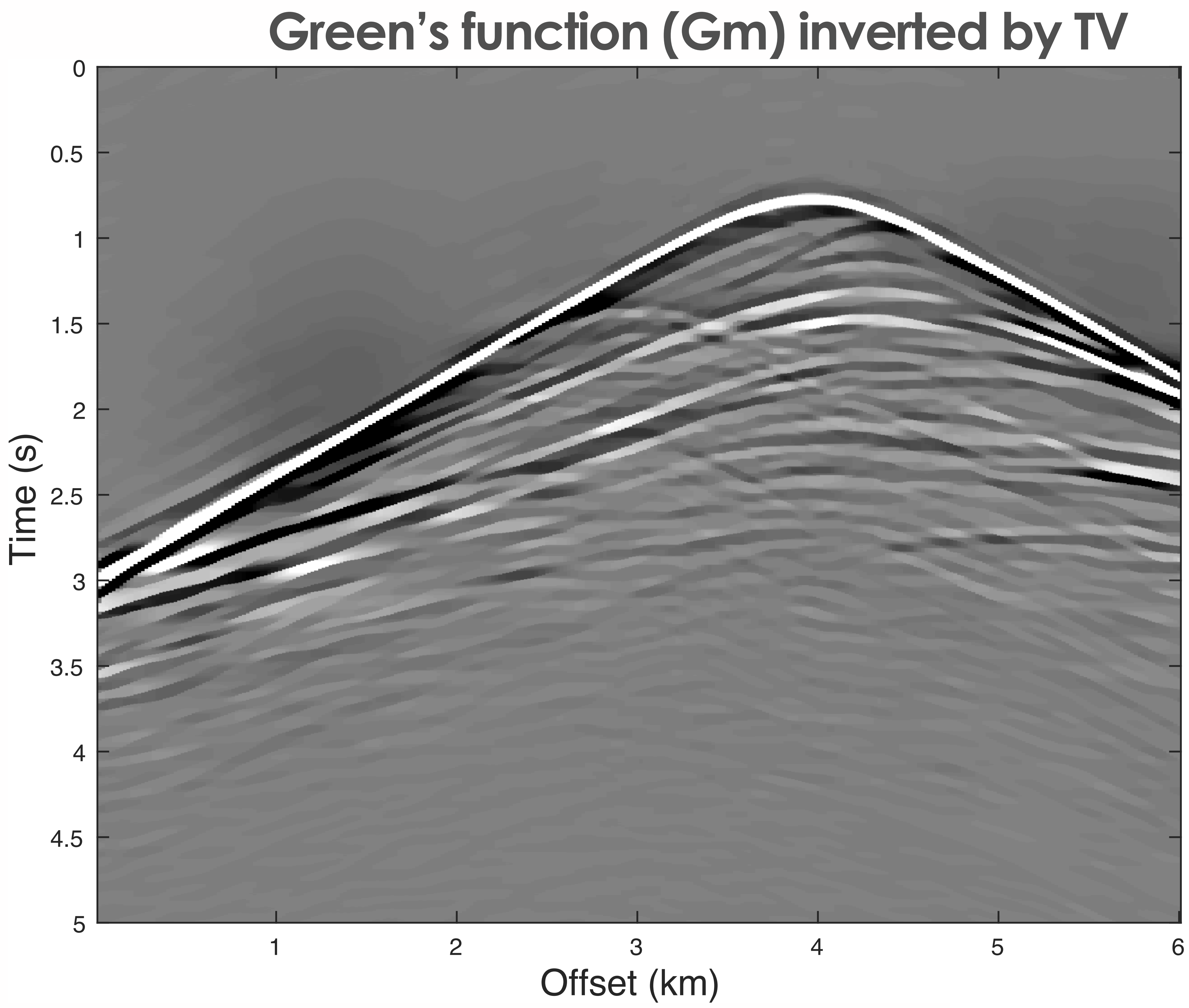
The effect of filtering



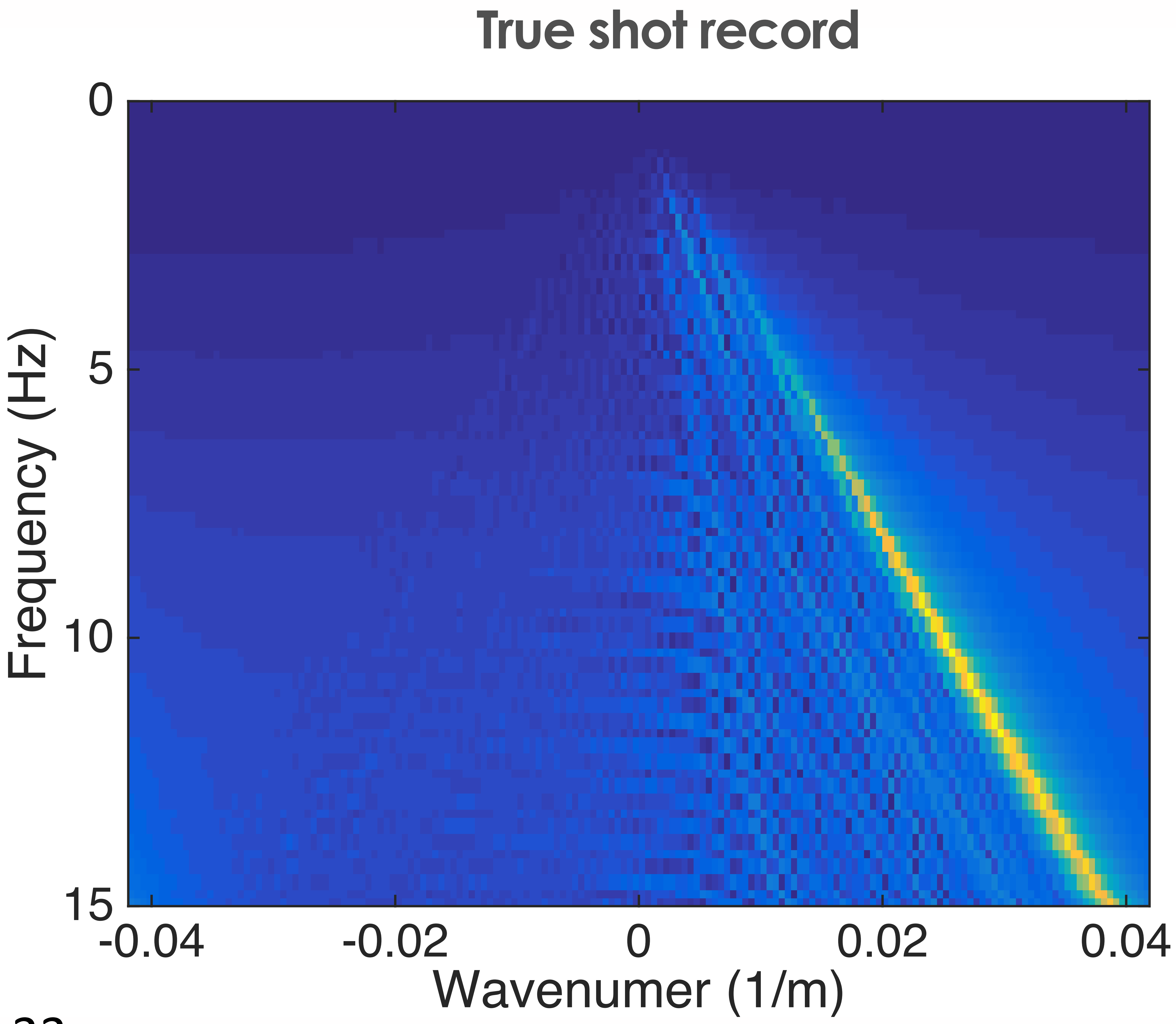


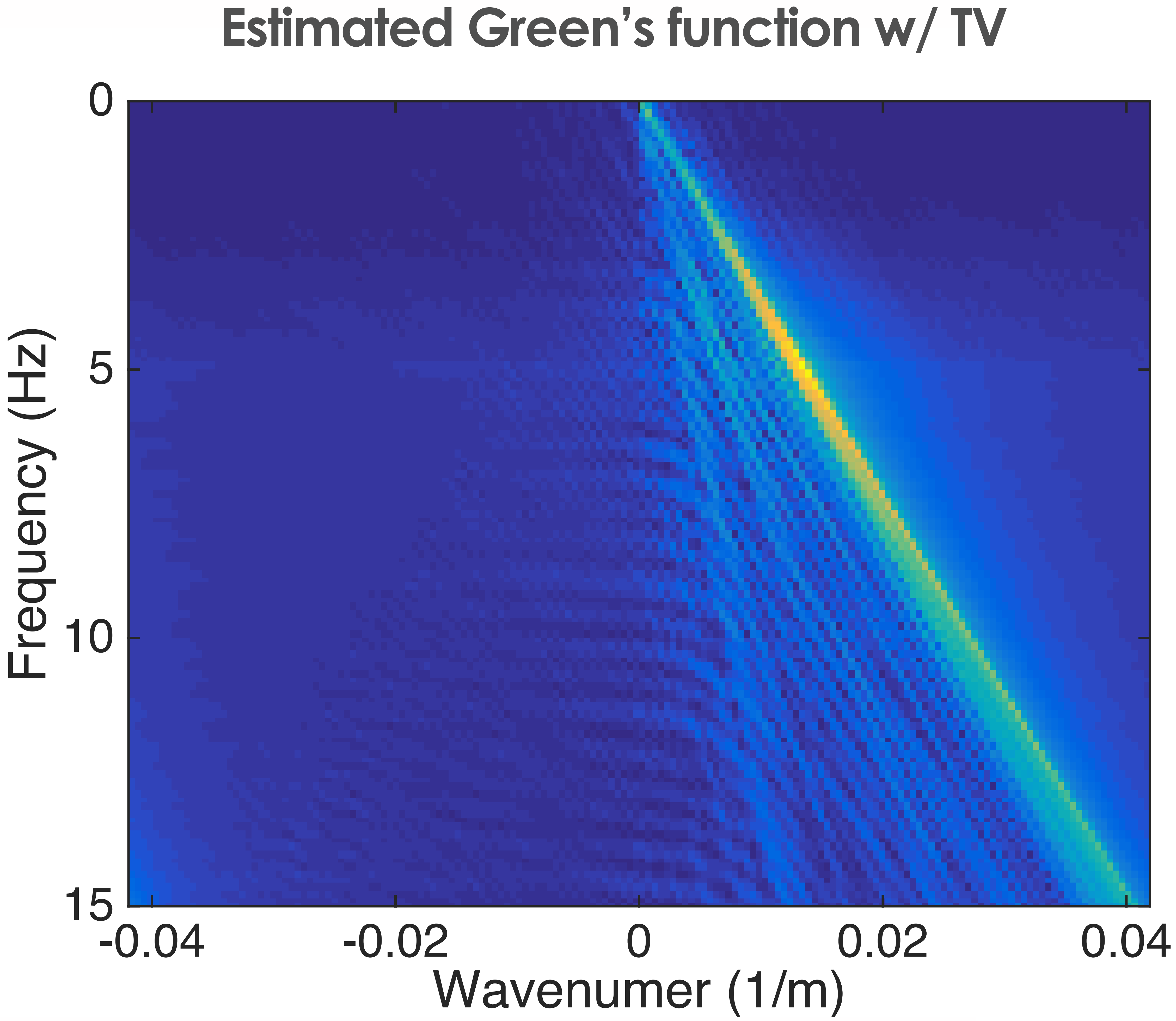
Deconvolution result using 5-15Hz data





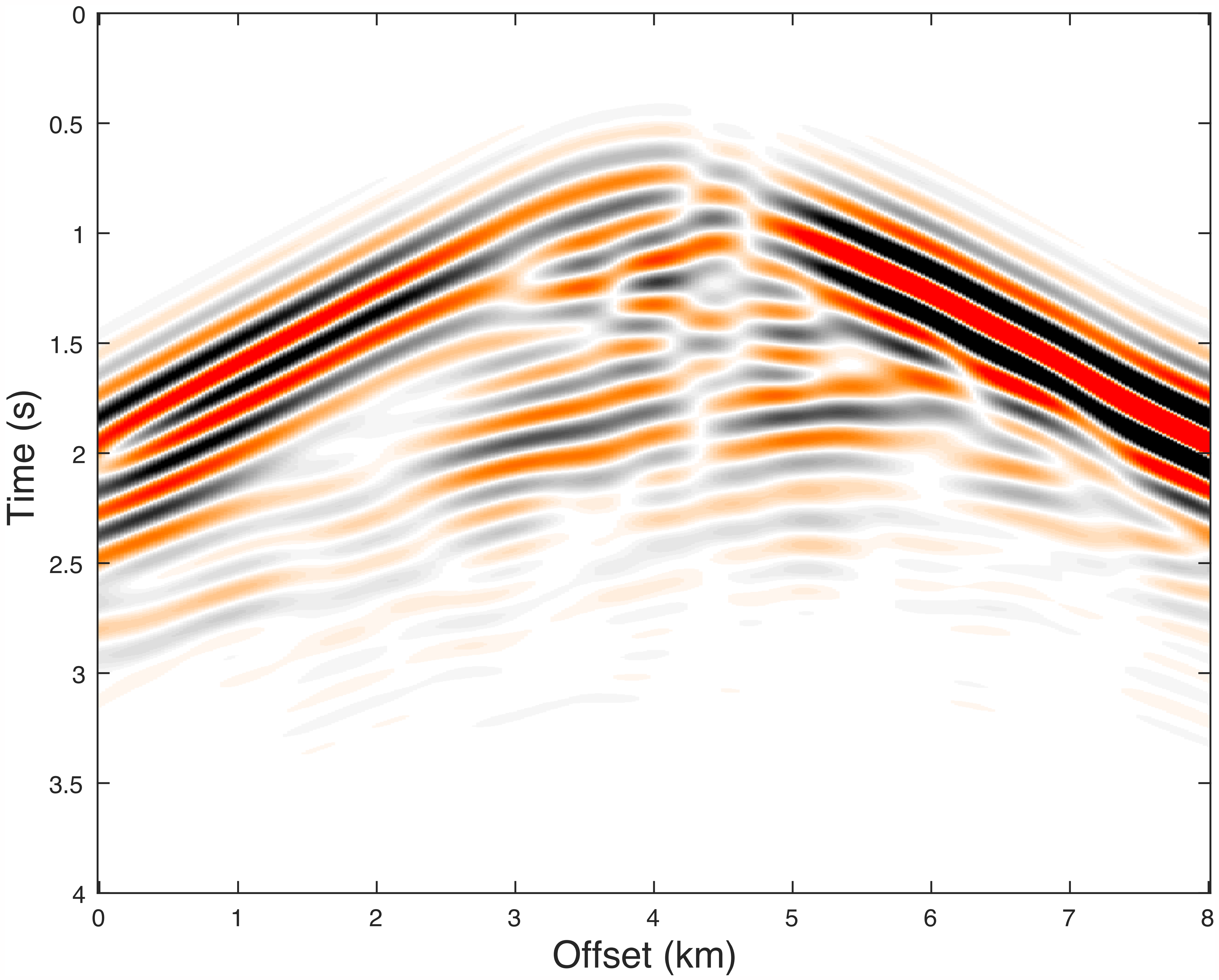
Deconvolution result from 5-15Hz data



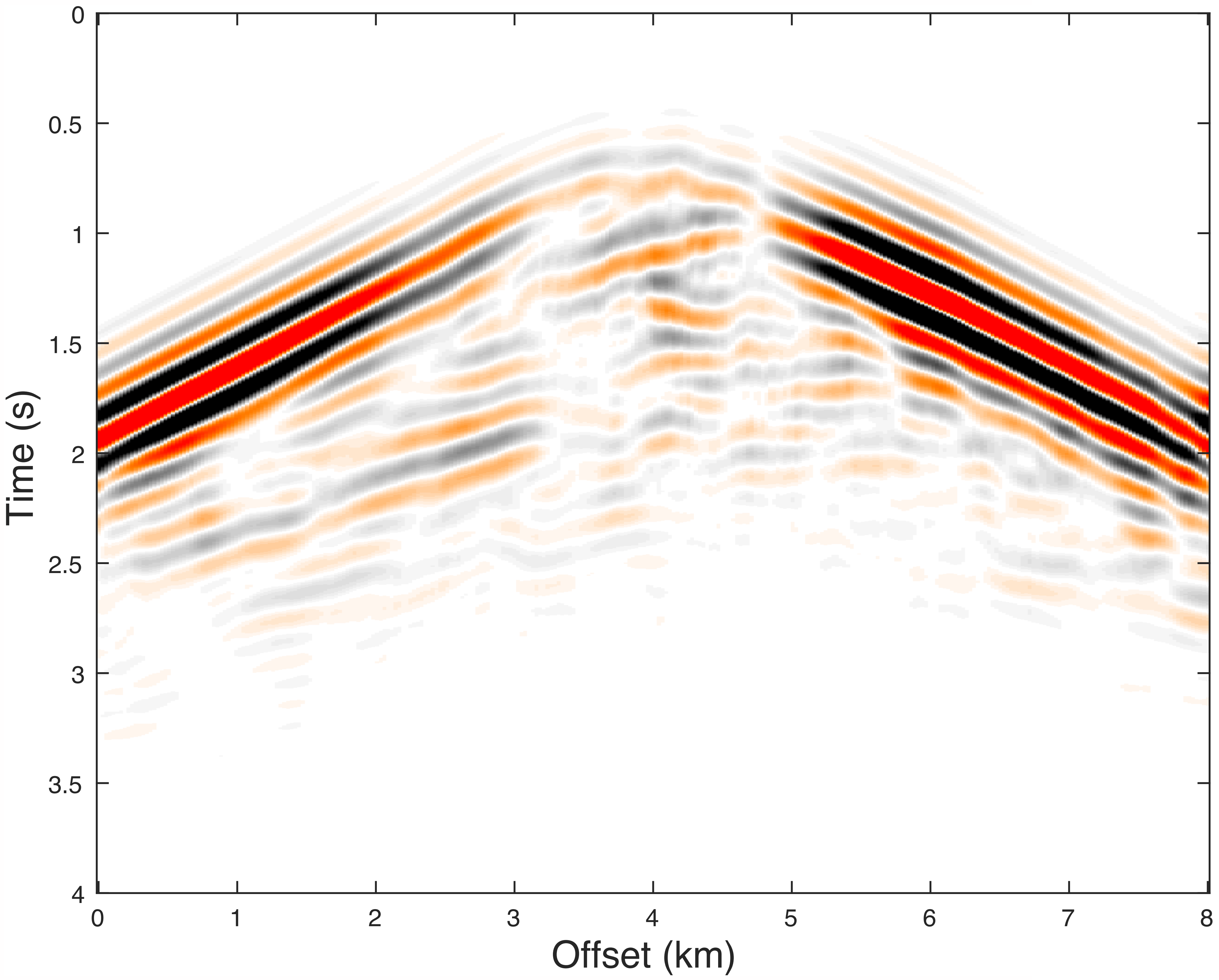


Deconvolution result from 5-15Hz data

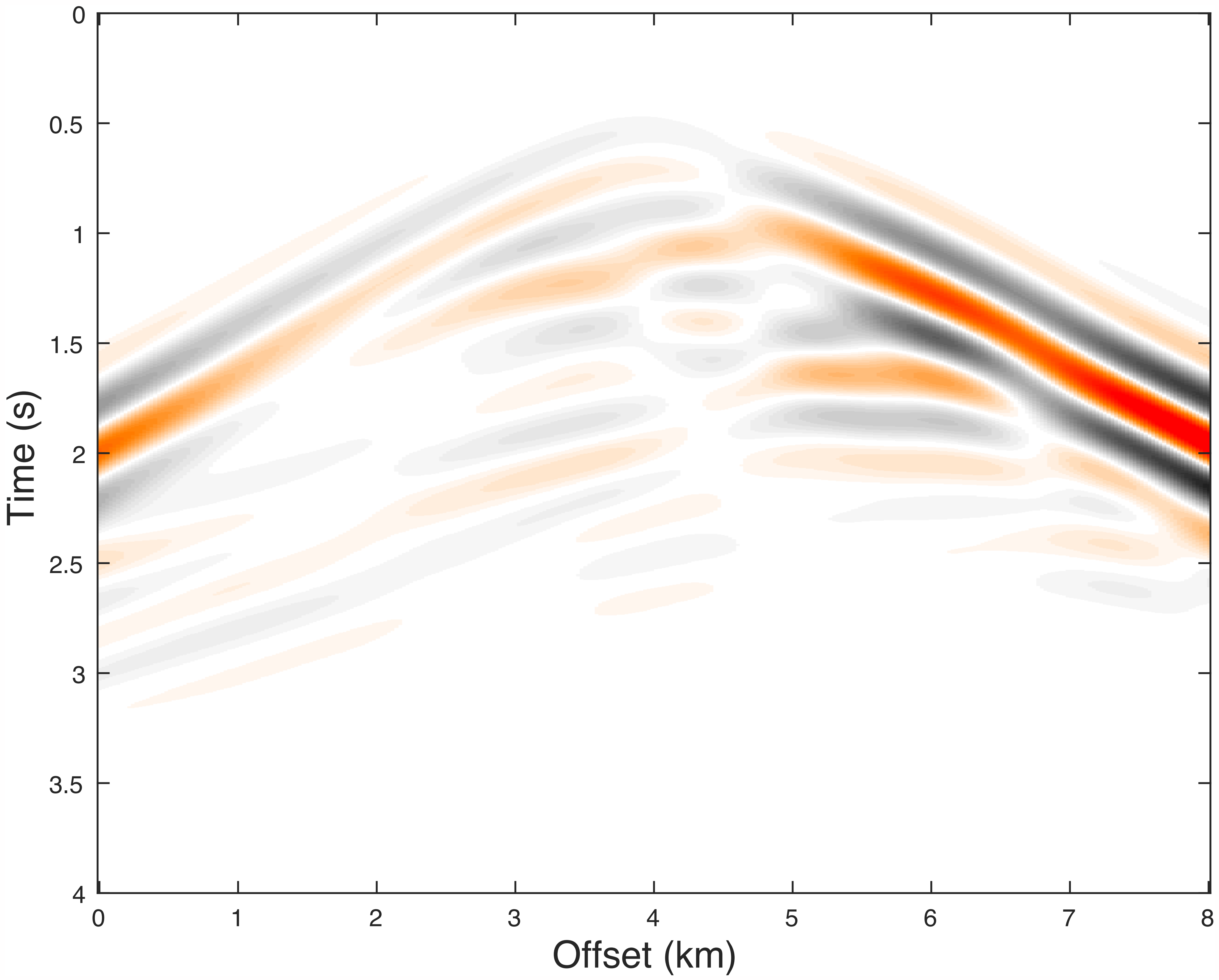
True data 0-4Hz

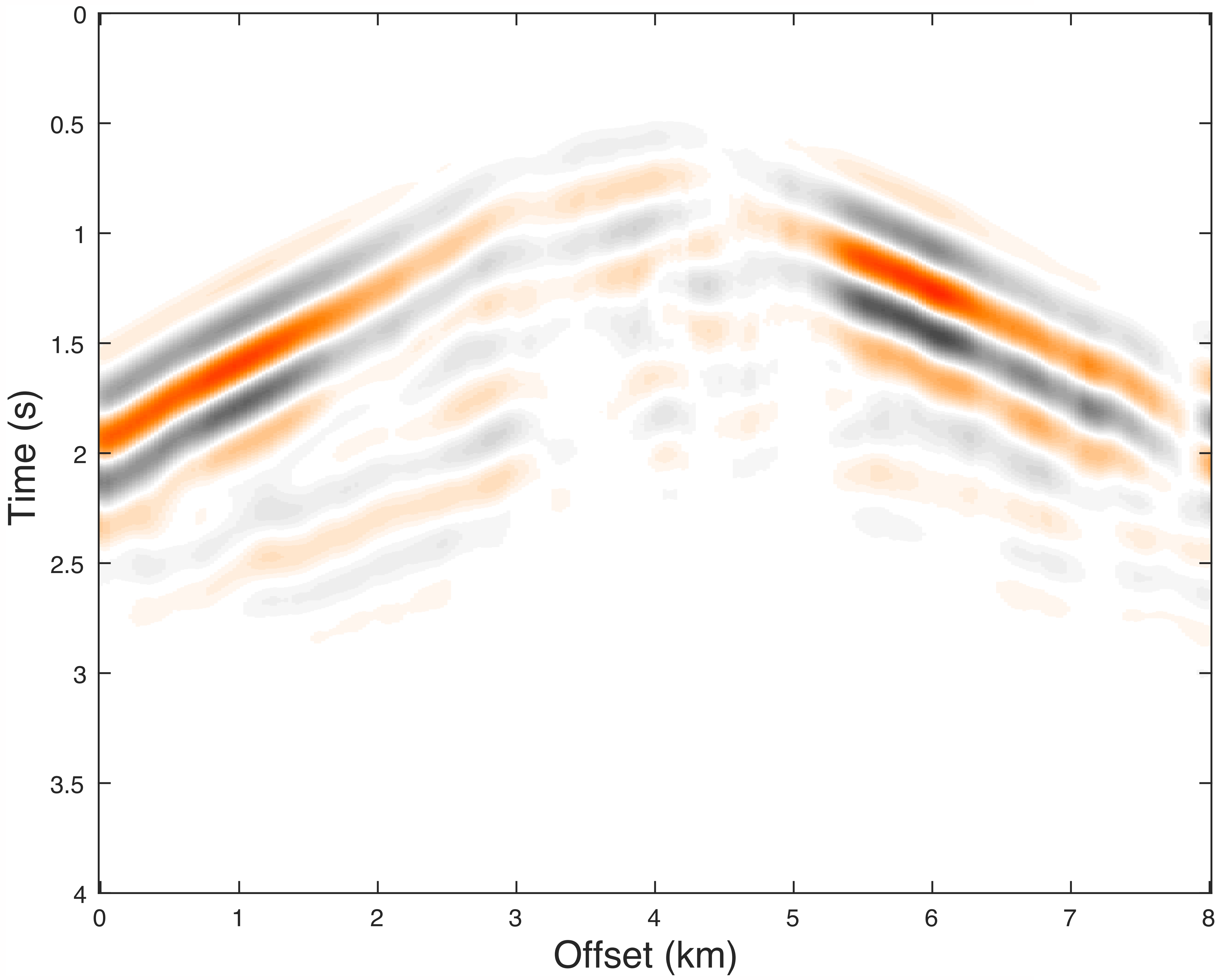


Extrapolated data 0-4Hz

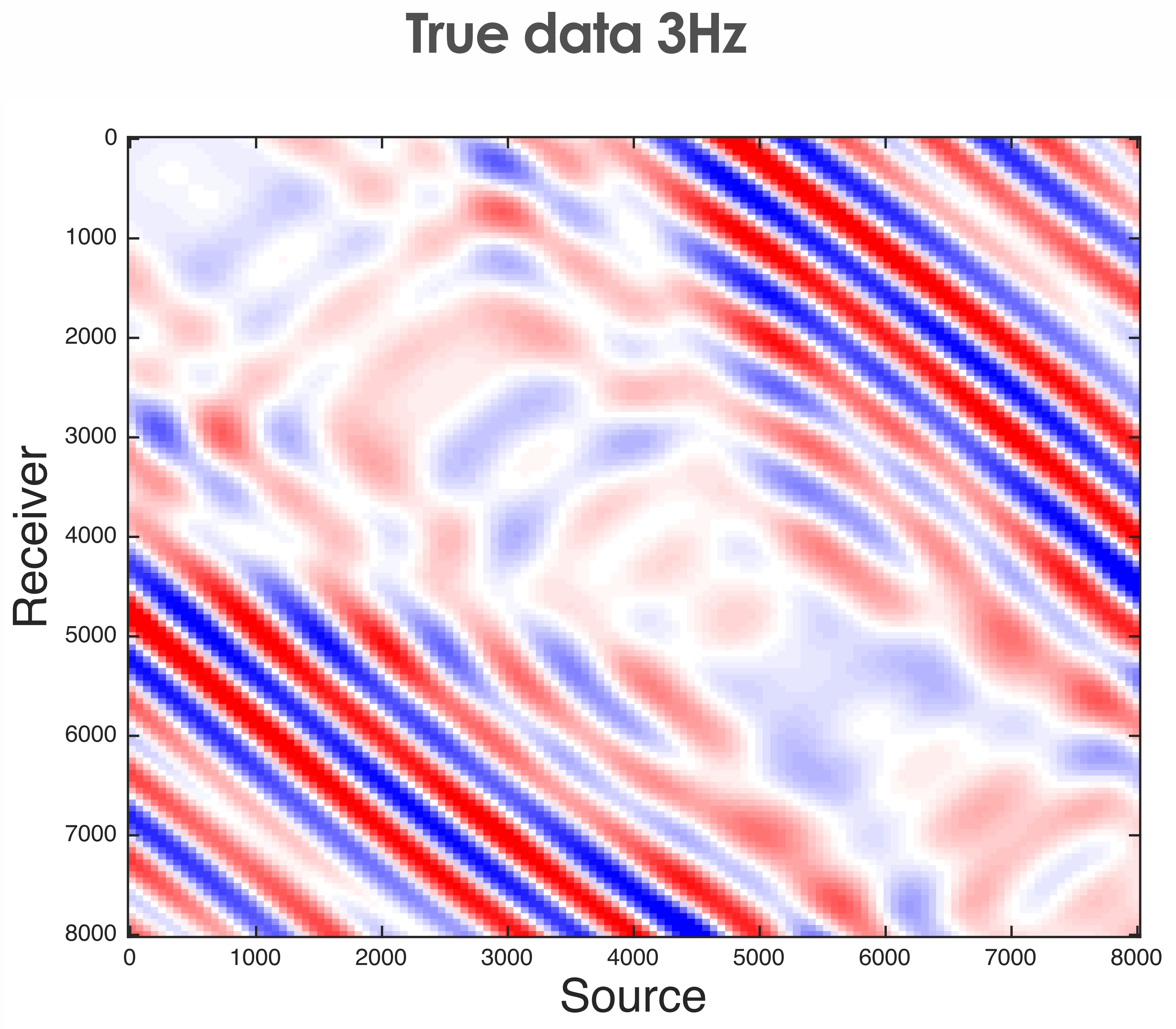


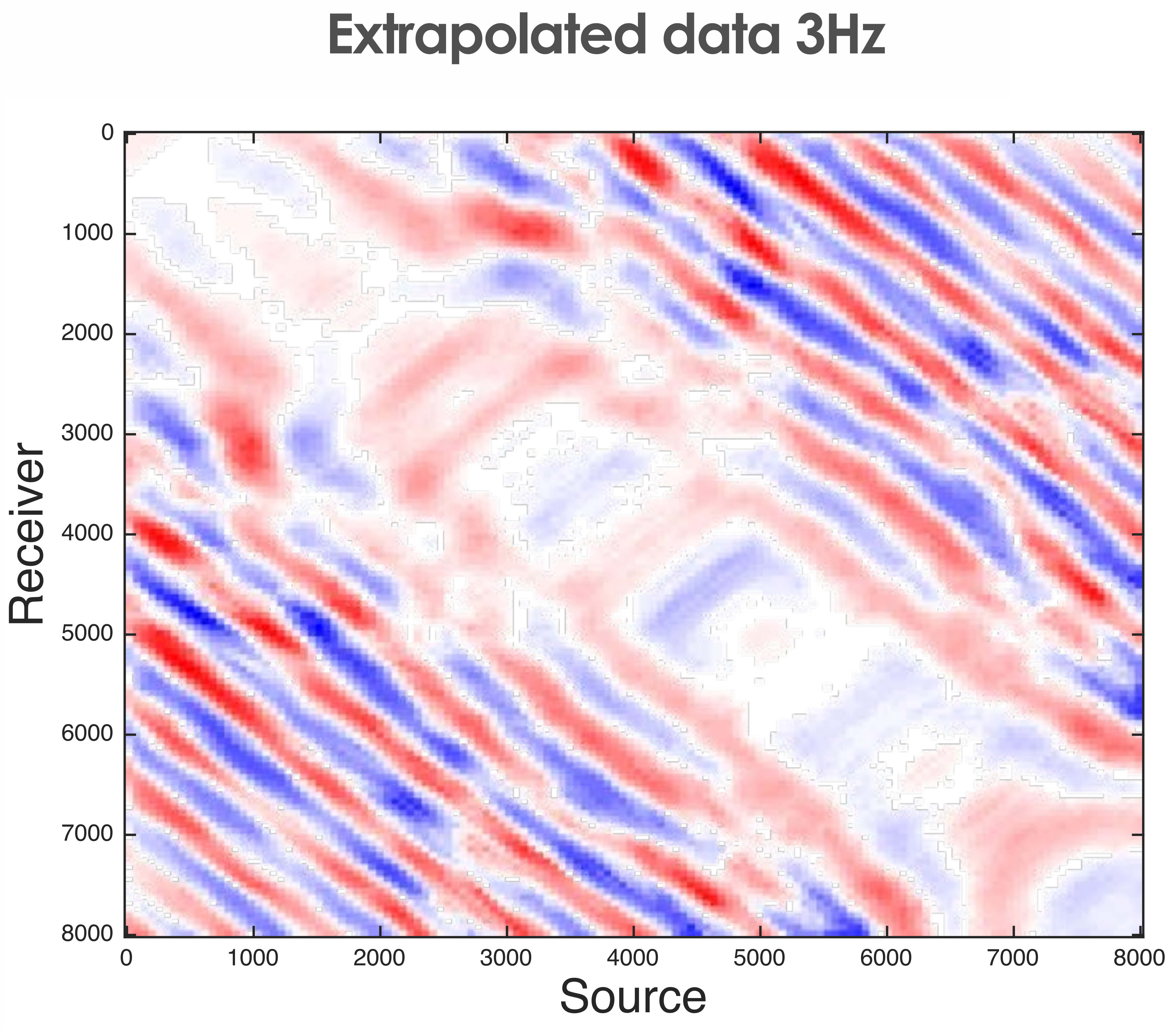
Deconvolution result from 5-15Hz data



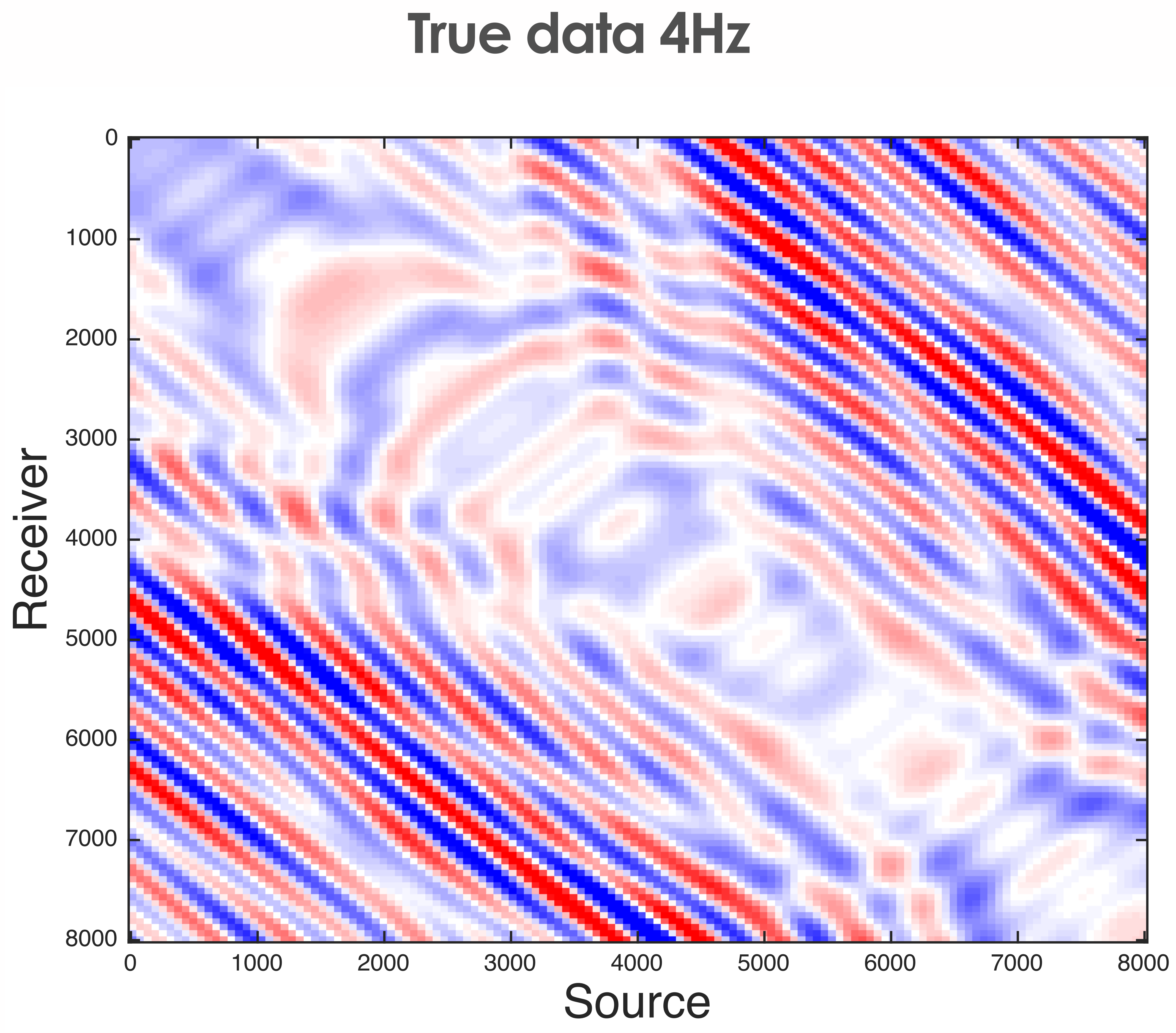


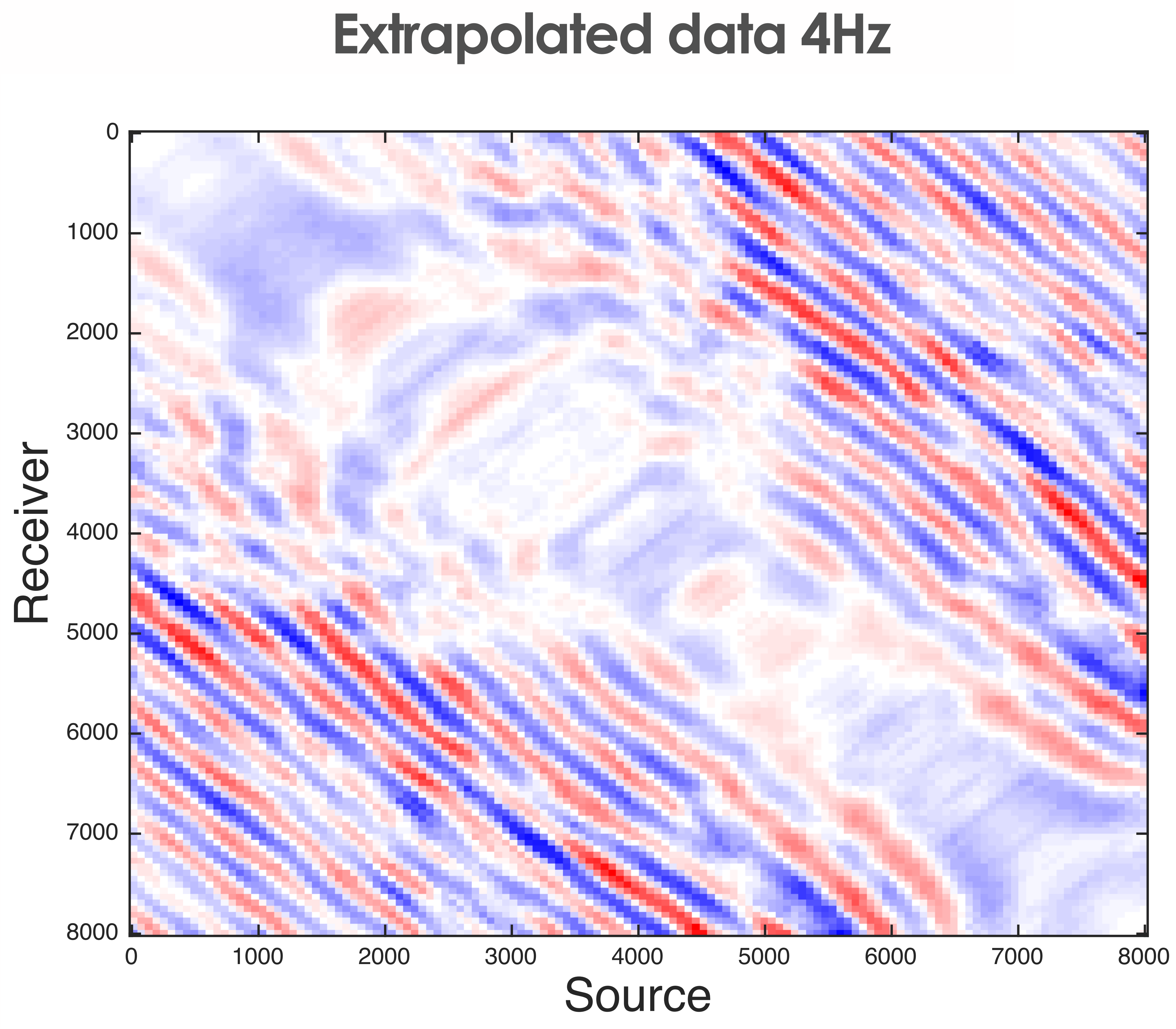
Deconvolution result from 5-15Hz data



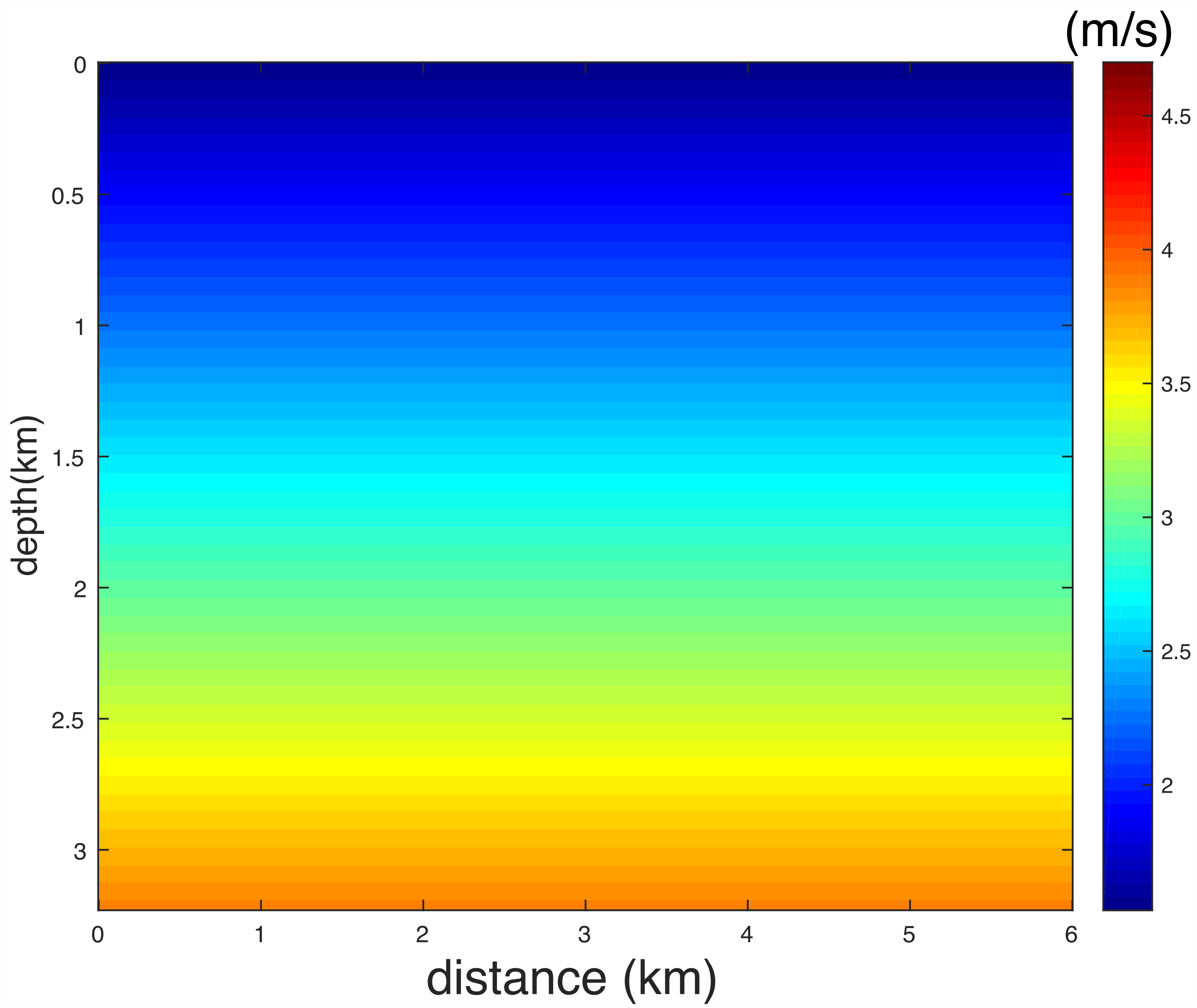


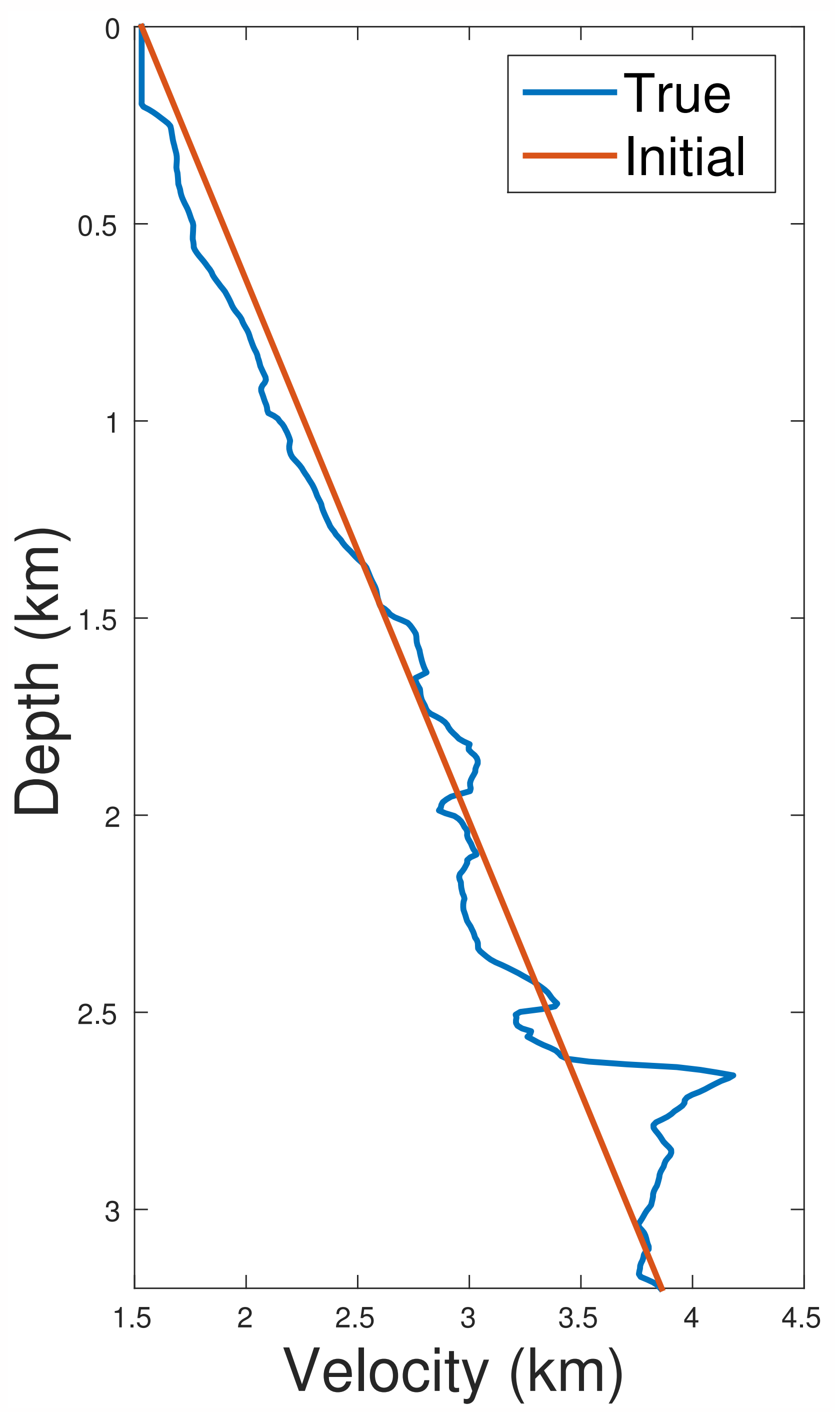
Deconvolution result from 5-15Hz data



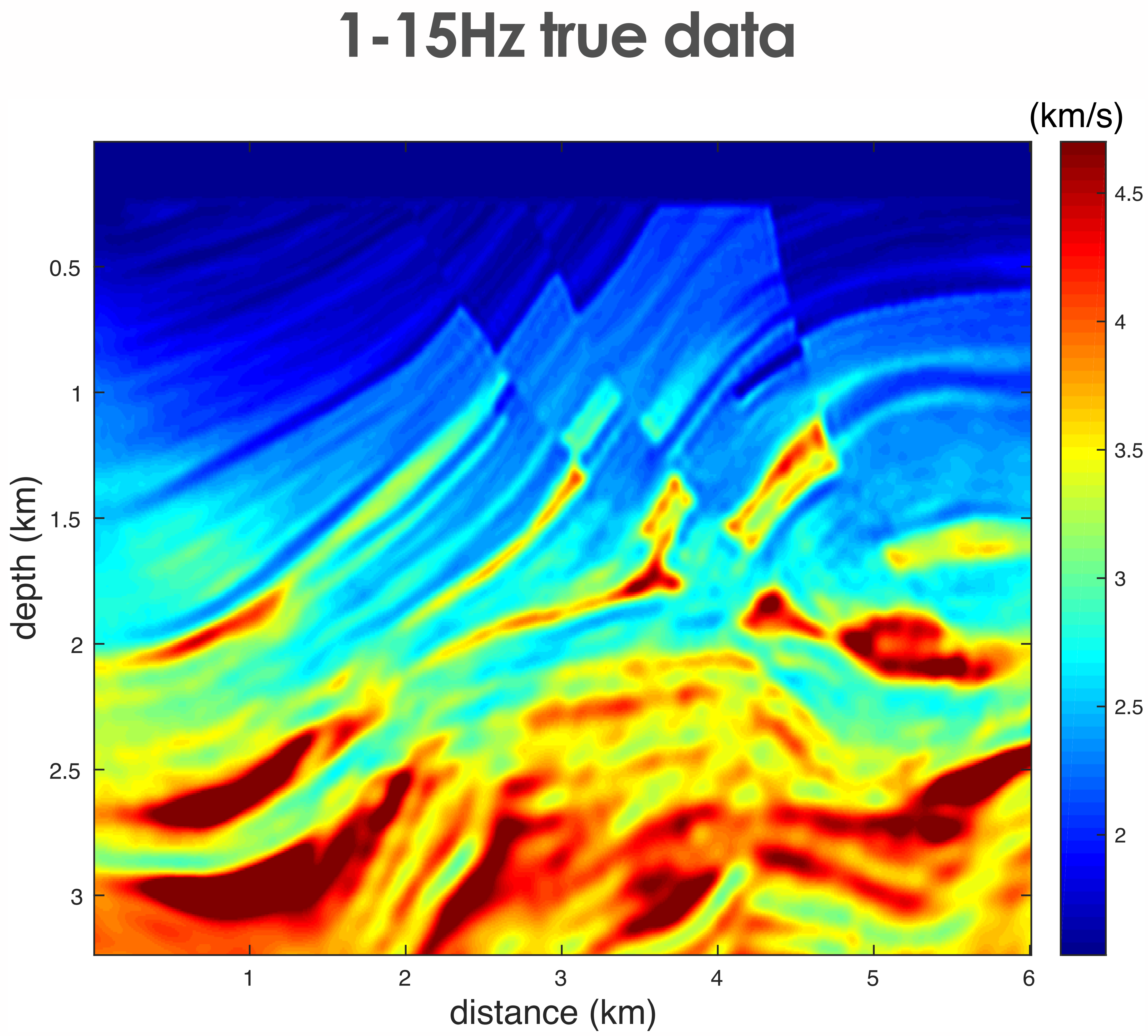


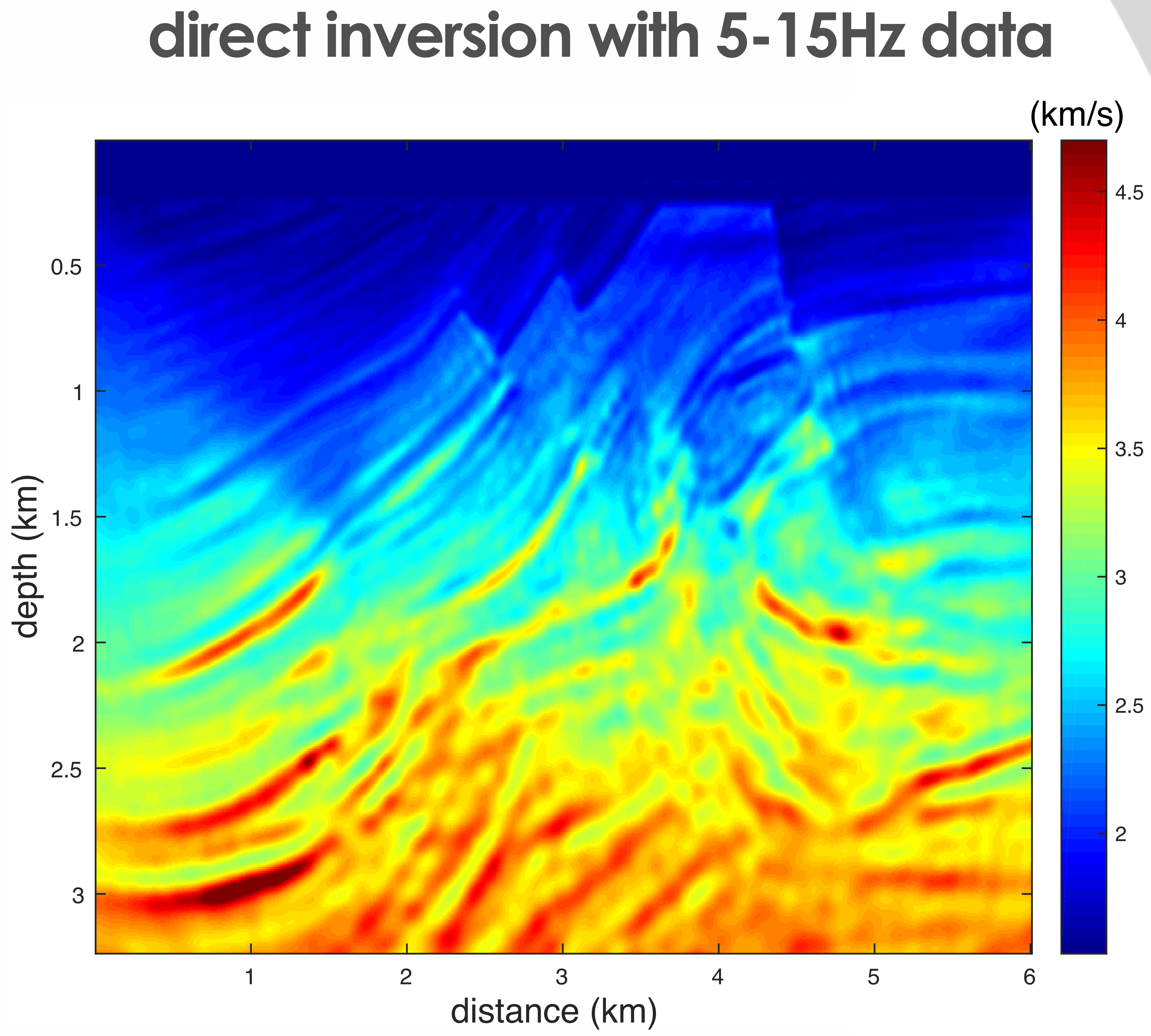
Initial guess





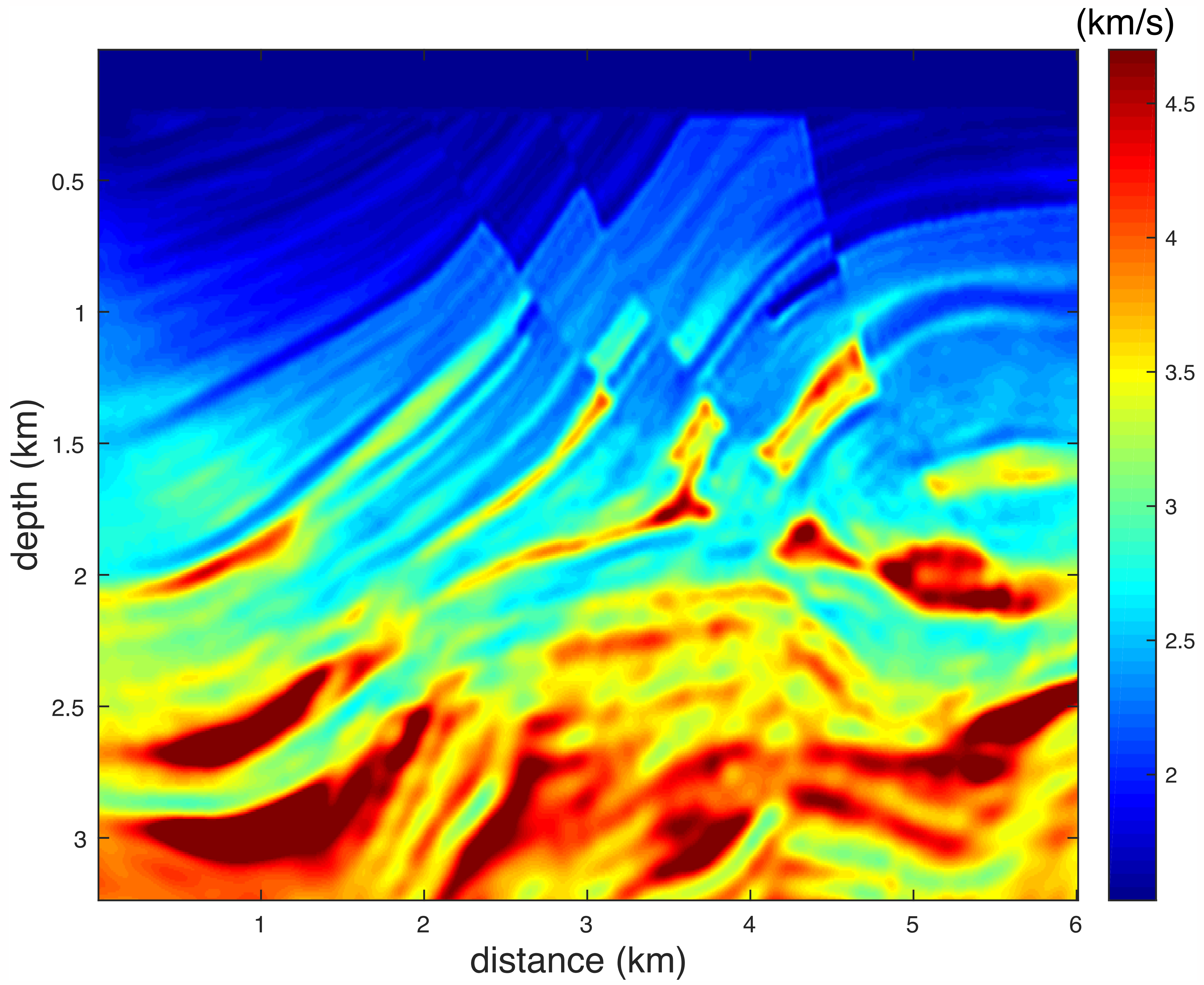
Recovery of FWI



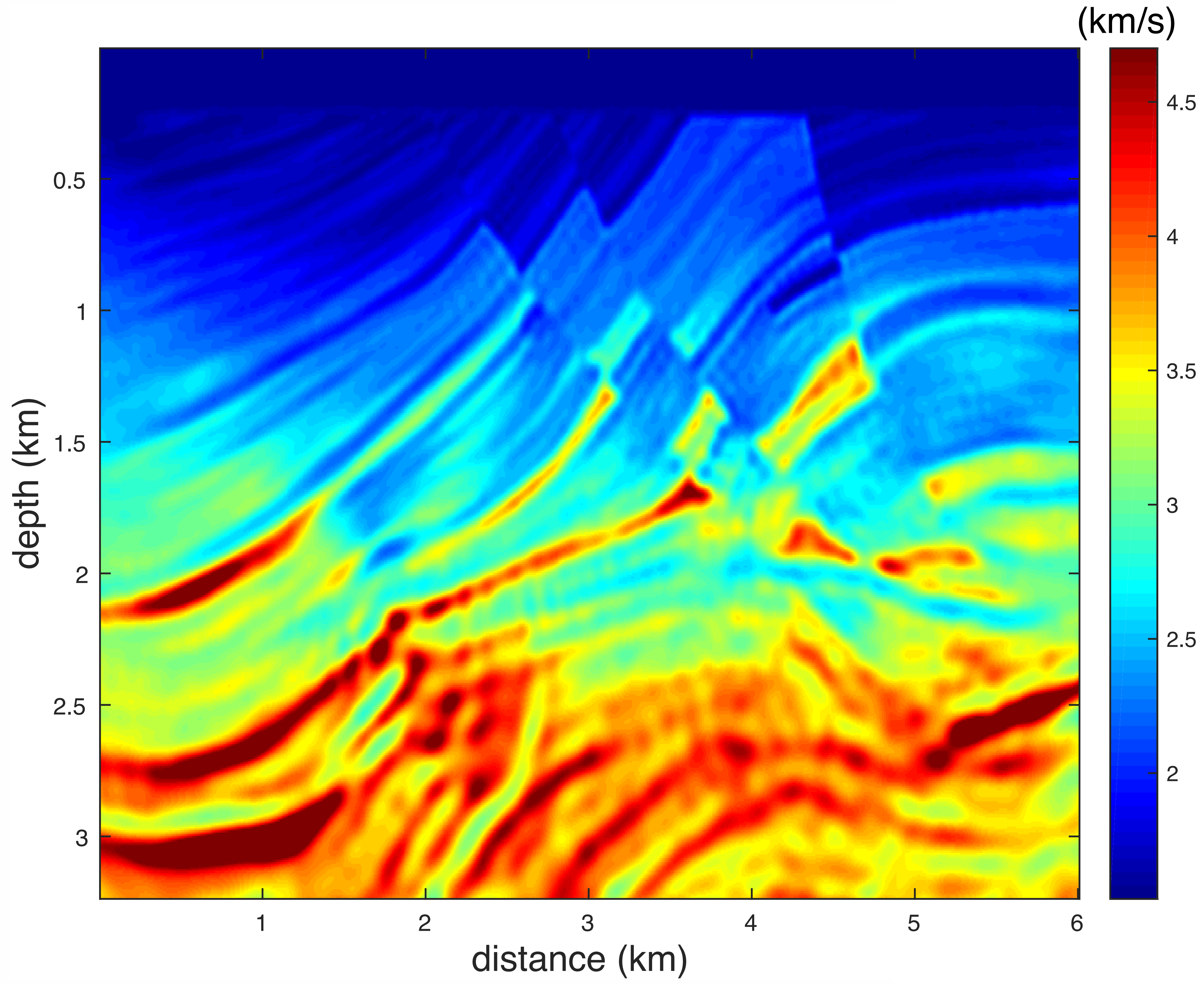


Inversion result via FWI

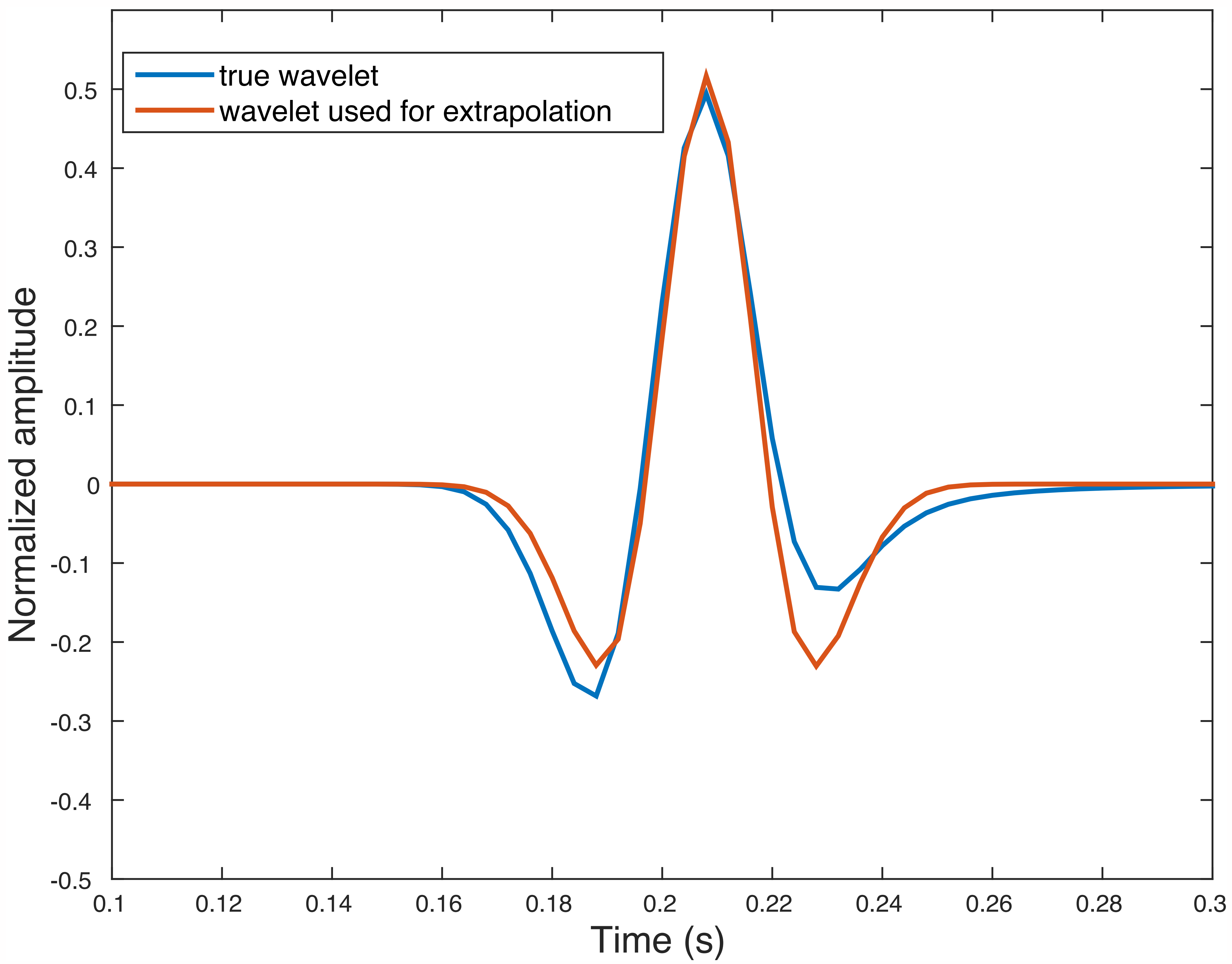
1-15Hz true data

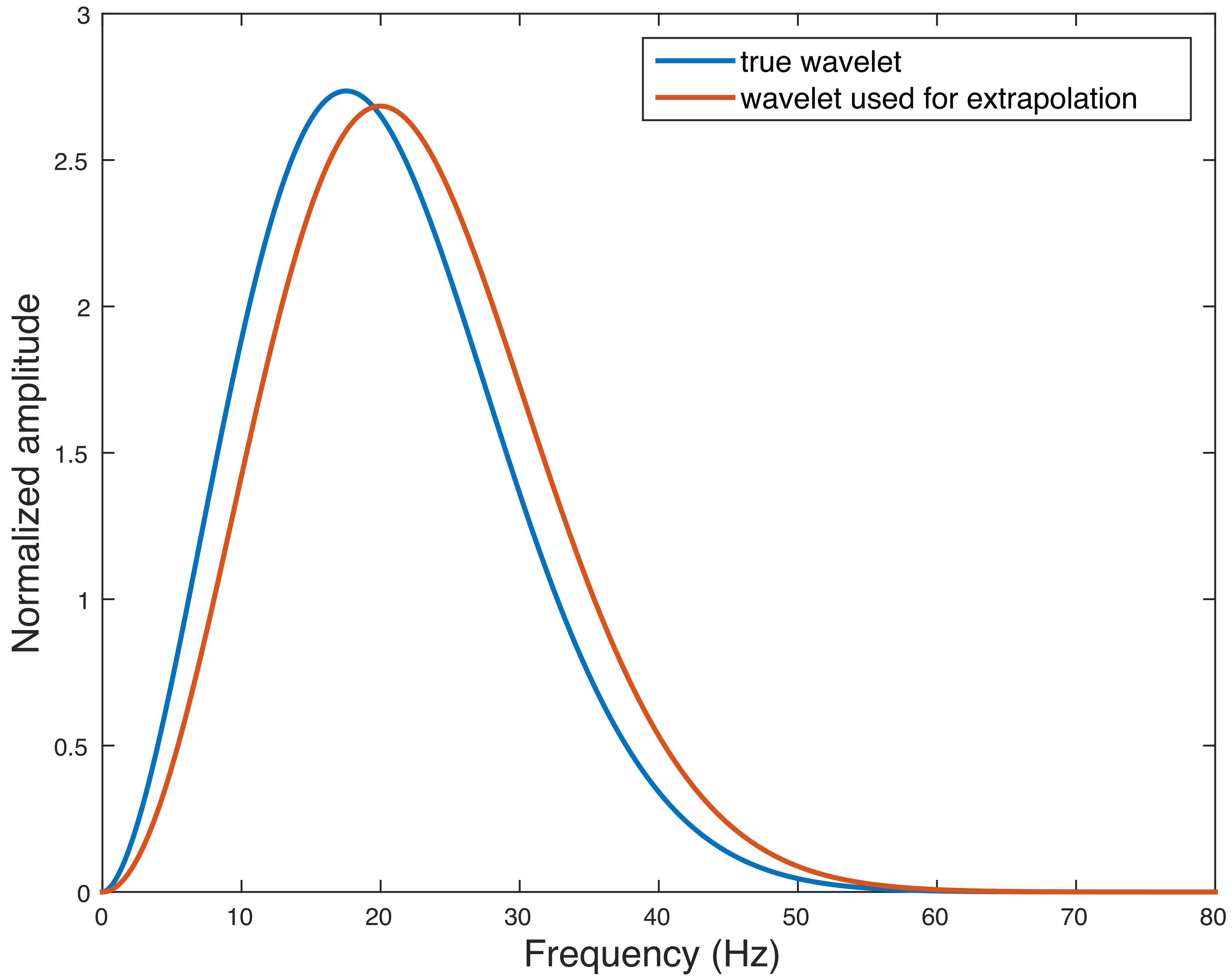


5-15Hz data with extrapolation

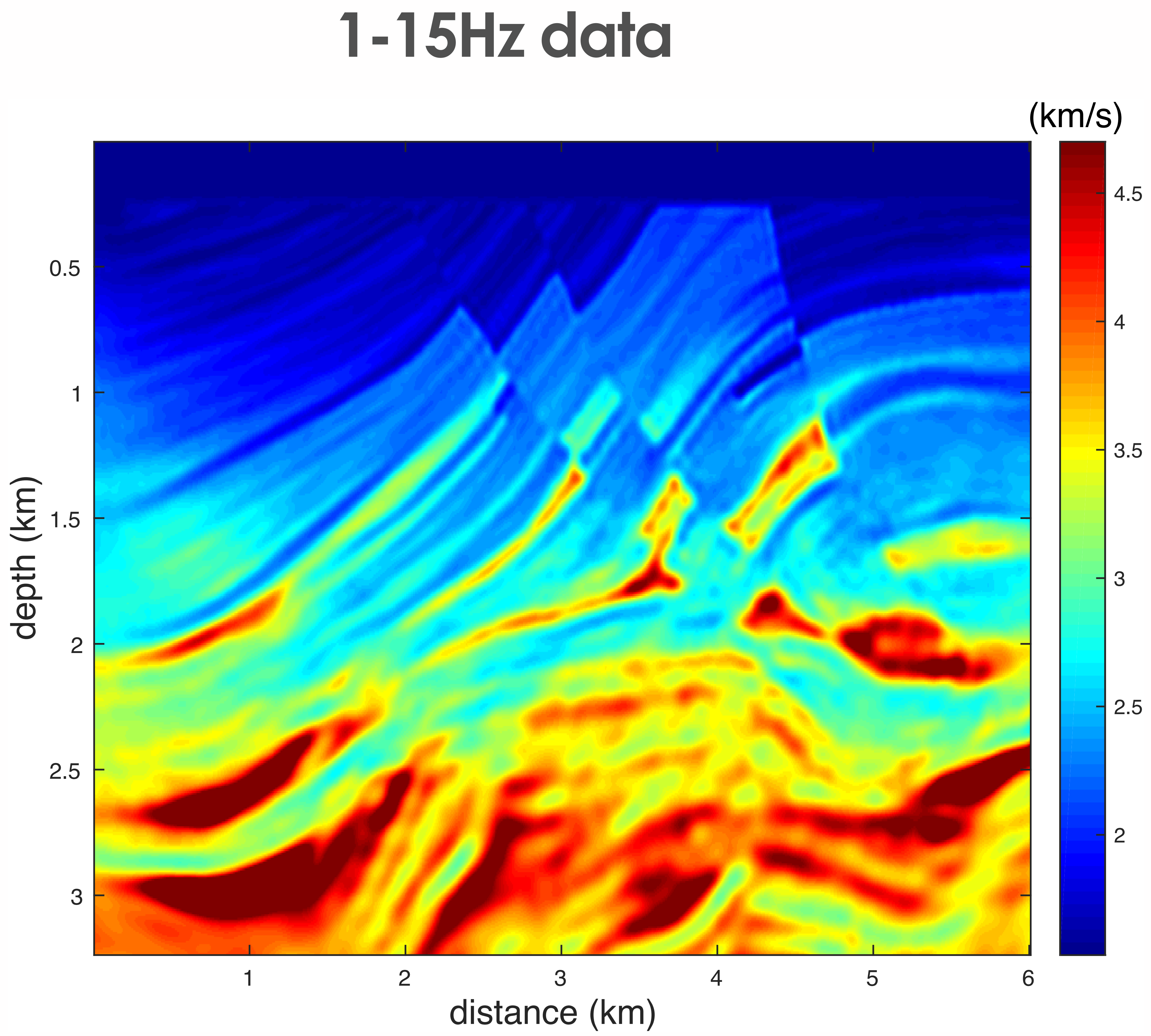


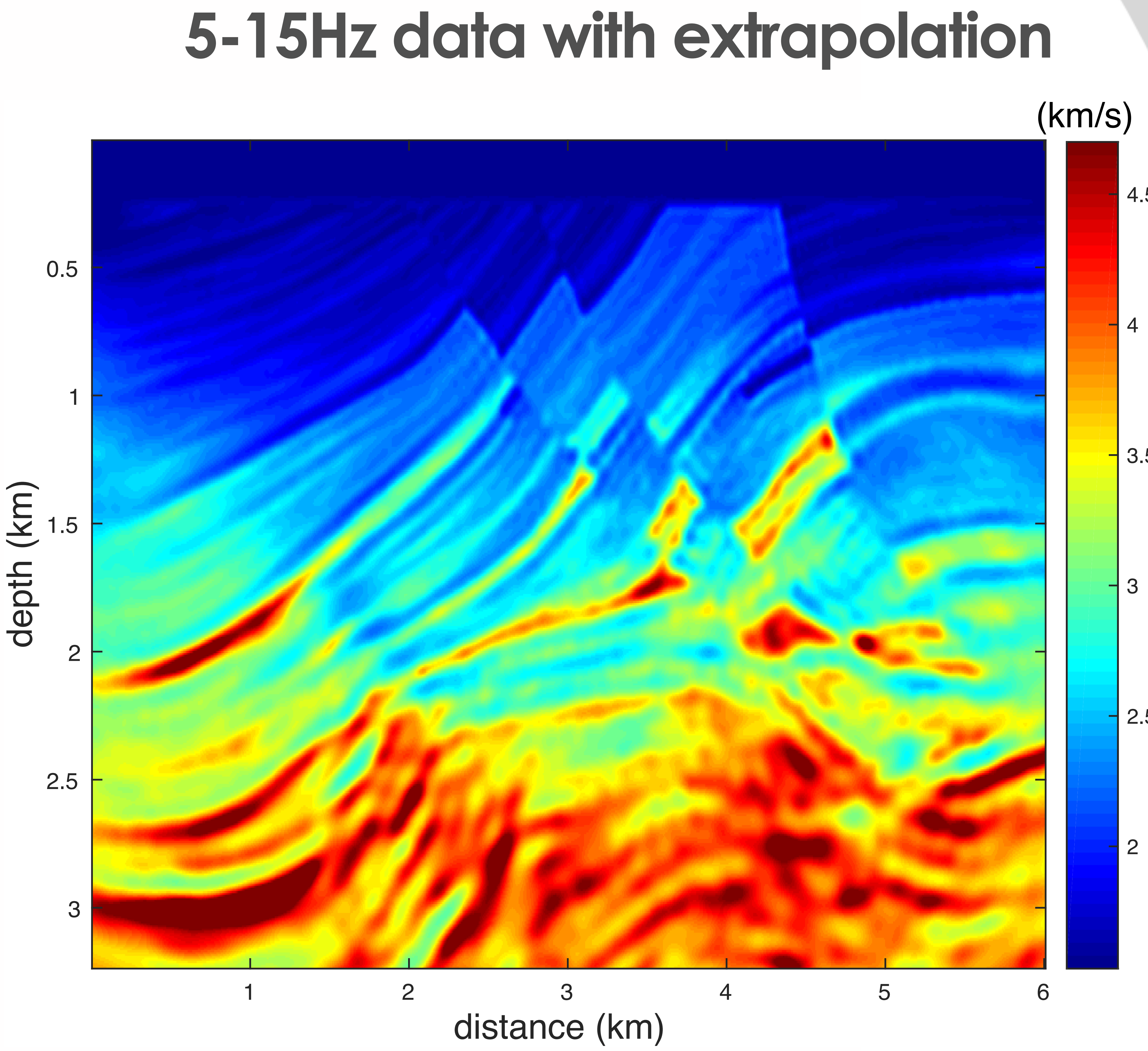
Stability test





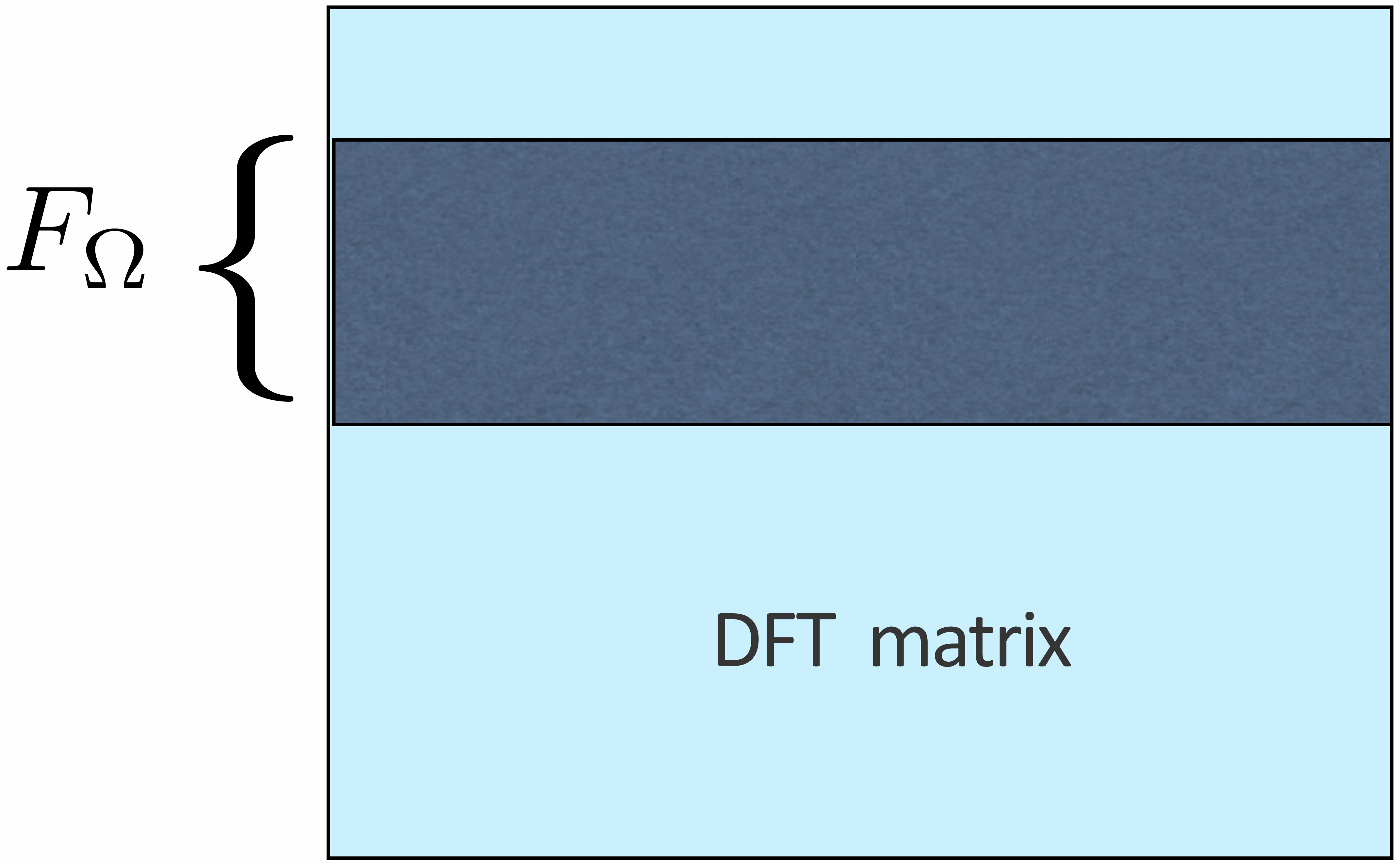
Inversion result via FWI





Lq norm minimization: to overcome the minimal distance barrier

Discrete setting



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

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

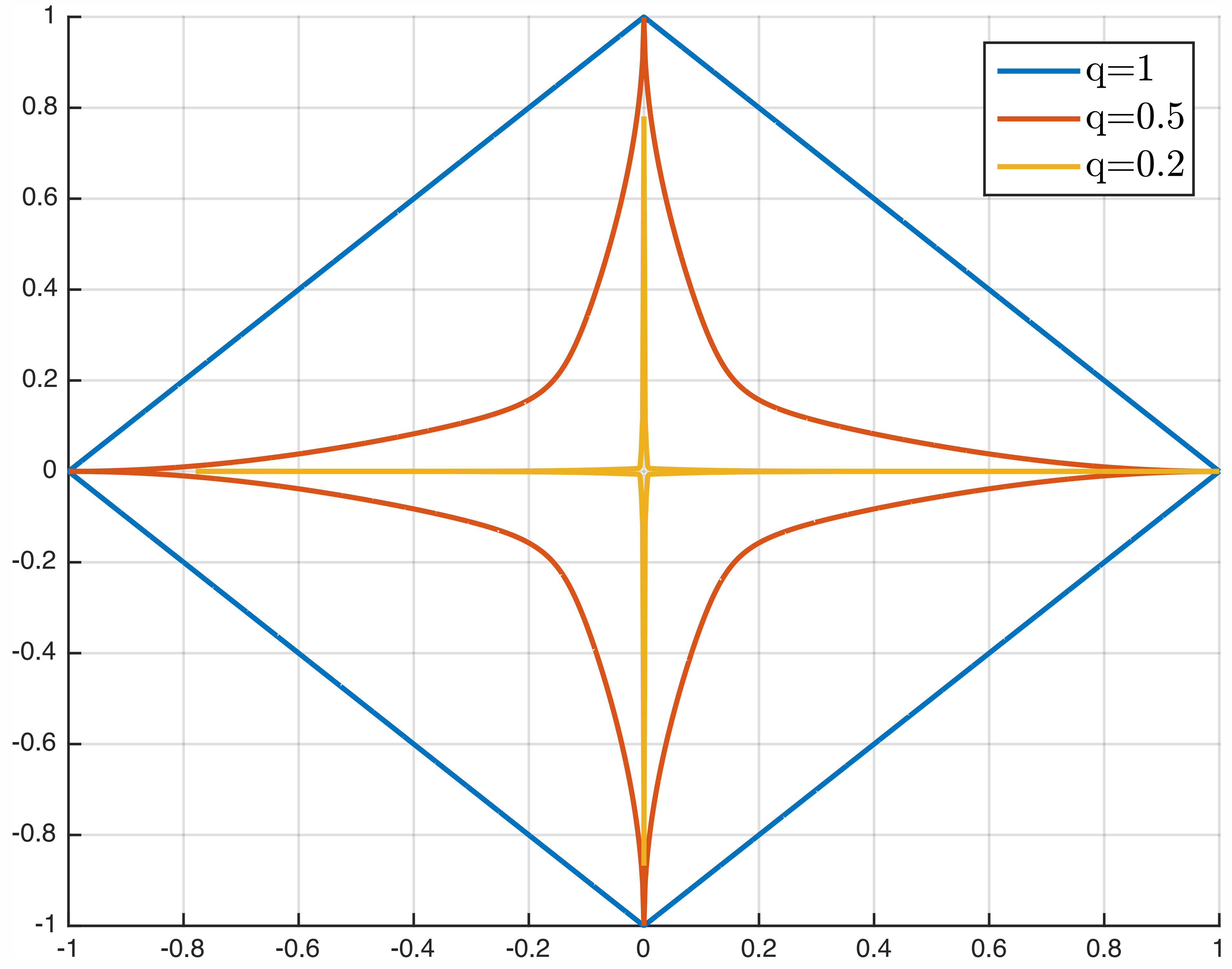
Bands in use: $\Omega = \{m_1, ..., m_2\}$

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

The Lq norm

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

Lq unit ball



$q \rightarrow 0$

more spiky, less convex

Lq objective

Lq norm minimization

$$\min_{\mathbf{r}} \|\mathbf{r}\|_q \qquad 0 < q < 1$$

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

$\mathbf{d}$: data
$\mathbf{w}$: wavelet
$\mathbf{r}$: reflectivity series
F_Ω: DFT matrix restricted to $\Omega = [m_1, m_2]$

Lq objective

Lq norm minimization

$$\min_{\mathbf{r}} \|\mathbf{r}\|_q \qquad 0 < q < 1$$

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

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

$$\min_{\mathbf{r}} \|\mathbf{r}\|_1 + \lambda\|\mathbf{r}\|_q^q$$

$$\text{subject to } F_\Omega(\mathbf{w} * \mathbf{r}) = F_\Omega \mathbf{d} \qquad 0 < q < 1$$

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

$$q \leq \frac{C}{\log N} \text{ and } s \leq c(m_2 - m_1)$$

Solver: Reweighted L1

To solve

$$\min_{\mathbf{x}} \|\mathbf{x}\|_q^q$$

$$\text{subject to } \mathbf{A}\mathbf{x} = \mathbf{b}$$

Write it as

$$\min_{\mathbf{x}} \|\mathbf{x}^{q-1} \odot \mathbf{x}\|_1$$

$$\text{subject to } \mathbf{A}\mathbf{x} = \mathbf{b}$$

RLS iterates as

$$\mathbf{x}_{k+1} = \min_{\mathbf{x}} \|\mathbf{x_k}^{q-1} \odot \mathbf{x}\|_1$$

$$\text{subject to } \mathbf{A}\mathbf{x} = \mathbf{b}$$

Solver: Reweighted L1

$$\mathbf{x}_{k+1} = \min_{\mathbf{x}} \|\mathbf{x_k}^{q-1} \odot \mathbf{x}\|_1$$
$$\text{subject to } \mathbf{Ax} = \mathbf{b}$$

can be written more formally as

$$\mathbf{x}_{k+1} = \min_{\mathbf{x}} \|\mathbf{a}_k \odot \mathbf{x}\|_1$$
$$\text{subject to } \mathbf{Ax} = \mathbf{b} \qquad \text{(RLS)}$$
$$\mathbf{a}_{k+1} = (|\mathbf{x}_k| + \epsilon)^{q-1}$$

RL1 for L1+Lq

Initialize weights: $\mathbf{a}_0 = [1, ..., 1]$

For $k = 1, ..., K$**:**

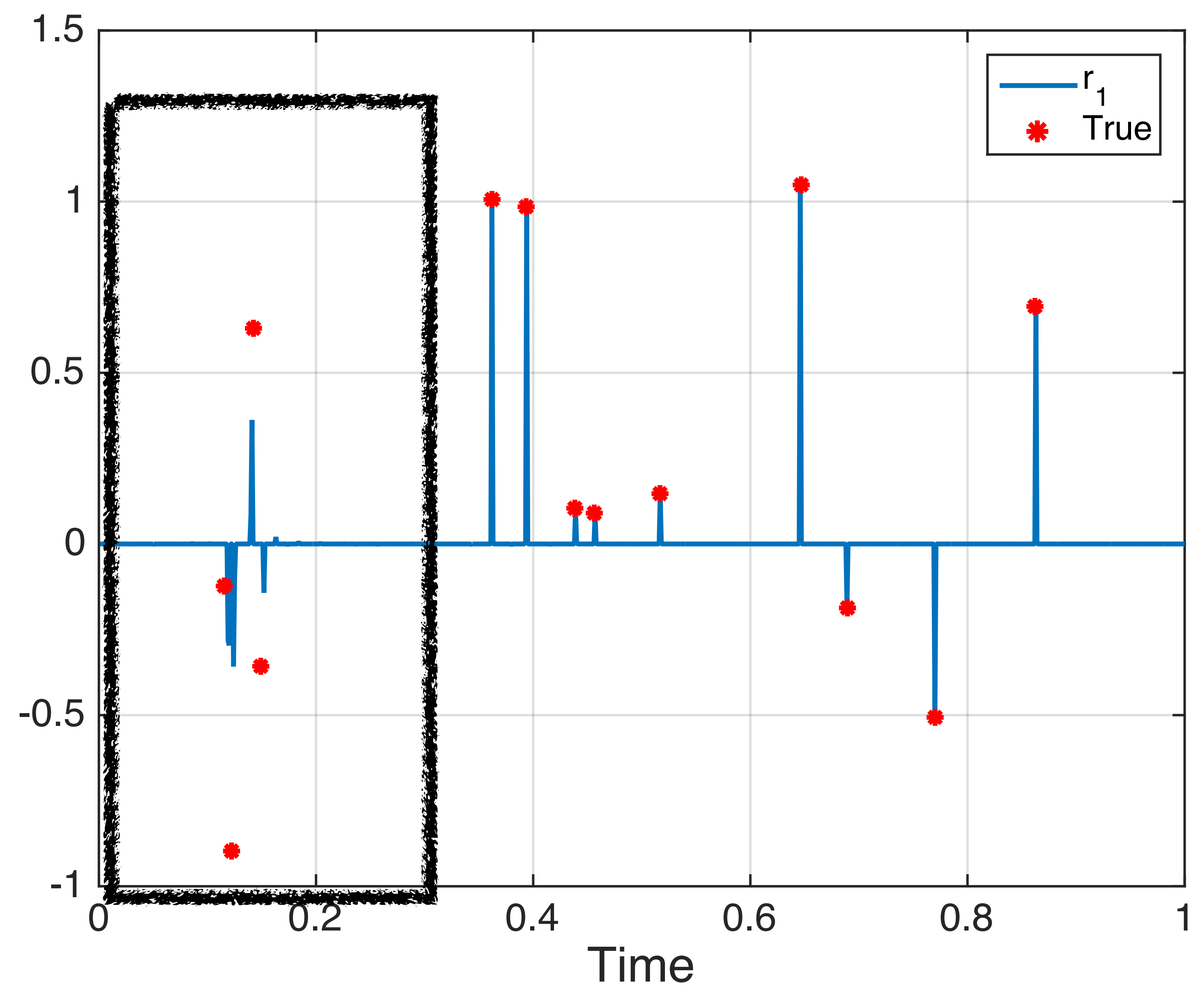
Define weight: $\mathbf{a}_k = [\mathbf{a}_k(1), ..., \mathbf{a}_k(n)], \quad \mathbf{a}_k(i) = \dfrac{1}{\epsilon + |\mathbf{r}_{k-1}(i)|^{1-q}}$

Update $\mathbf{r}_k$: $\quad \mathbf{r}_{k+1} = \min_{\mathbf{r}} \|\mathbf{r}\|_1 + \lambda \|\mathbf{a}_k \otimes \mathbf{r}\|_1$

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

End

Why RL1 does not work for our case?

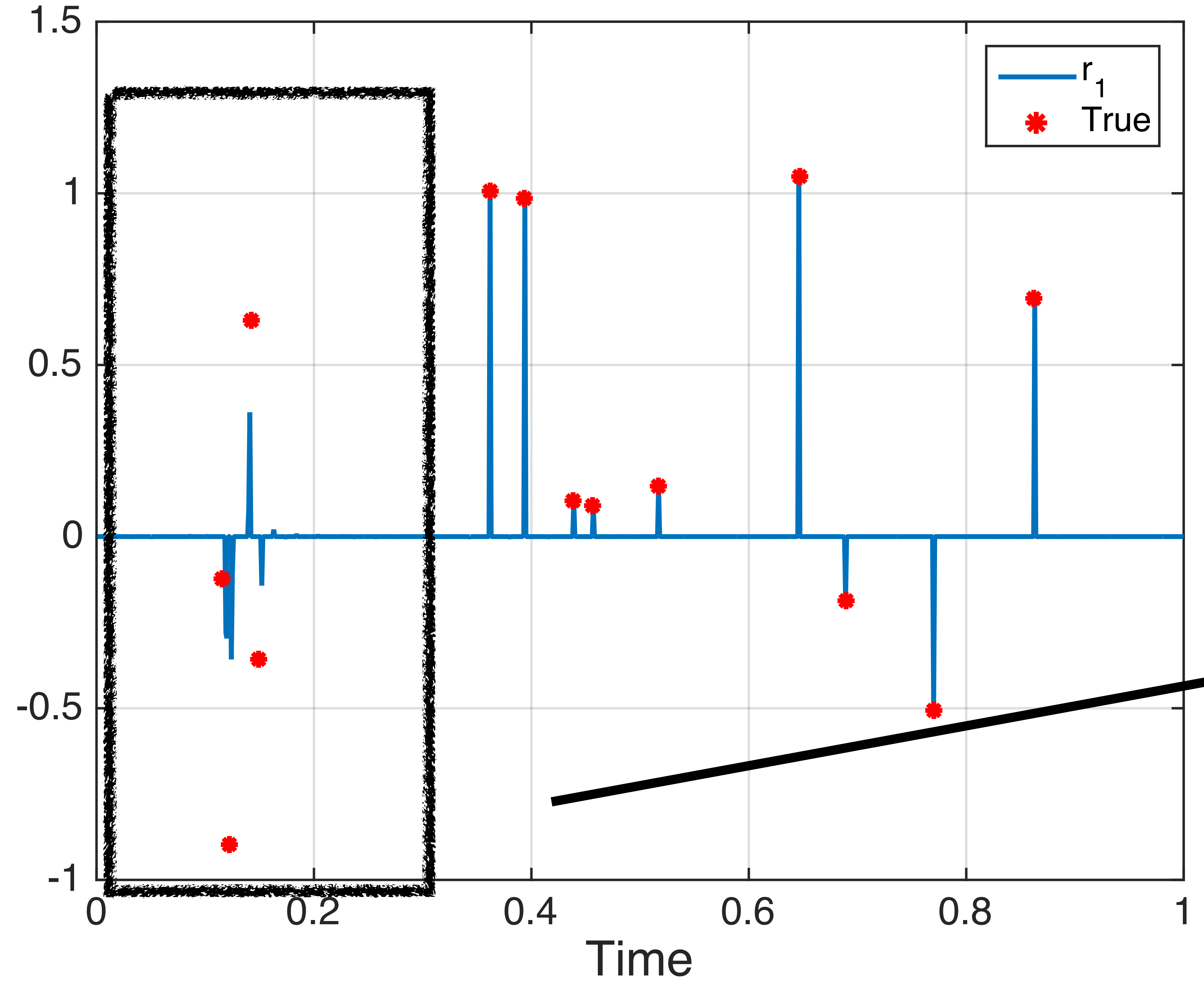


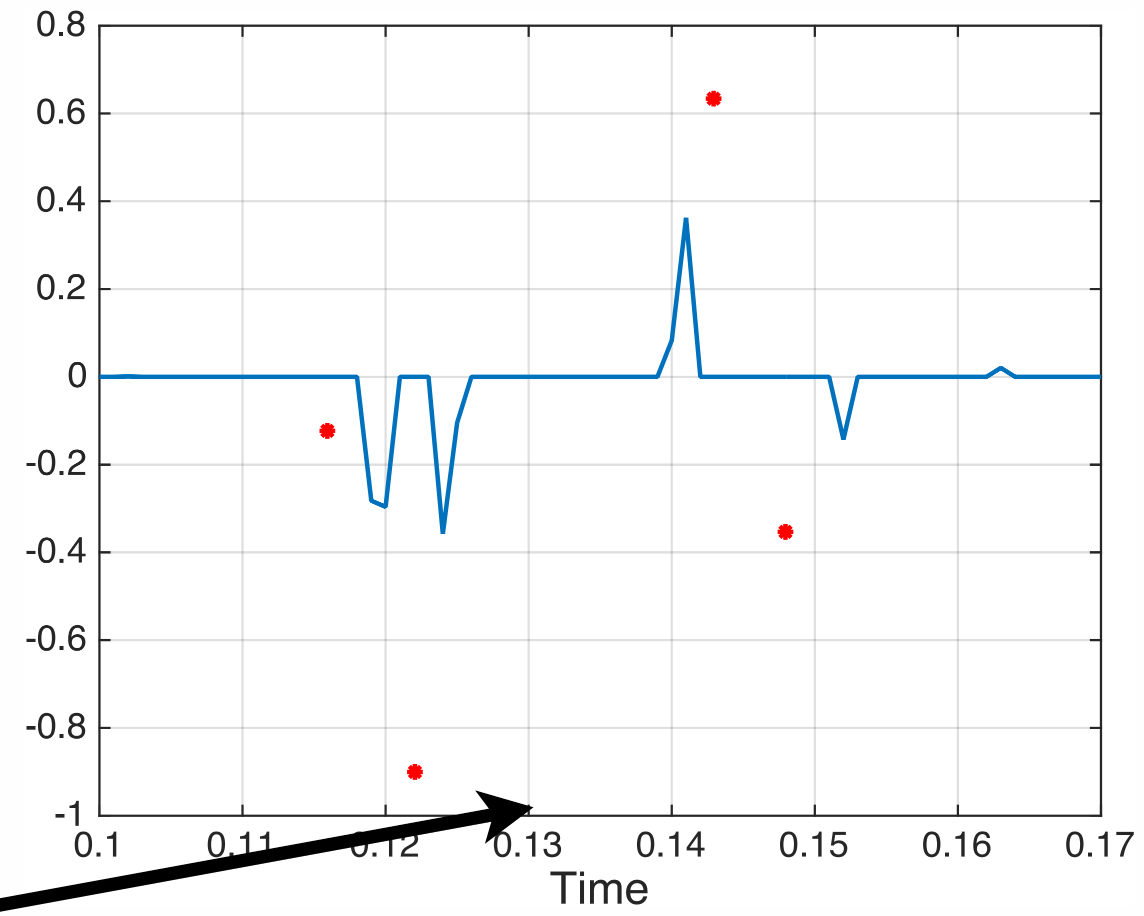
$$N = 1000$$

$$\Omega = [m_1, m_2] = [10, 50]$$

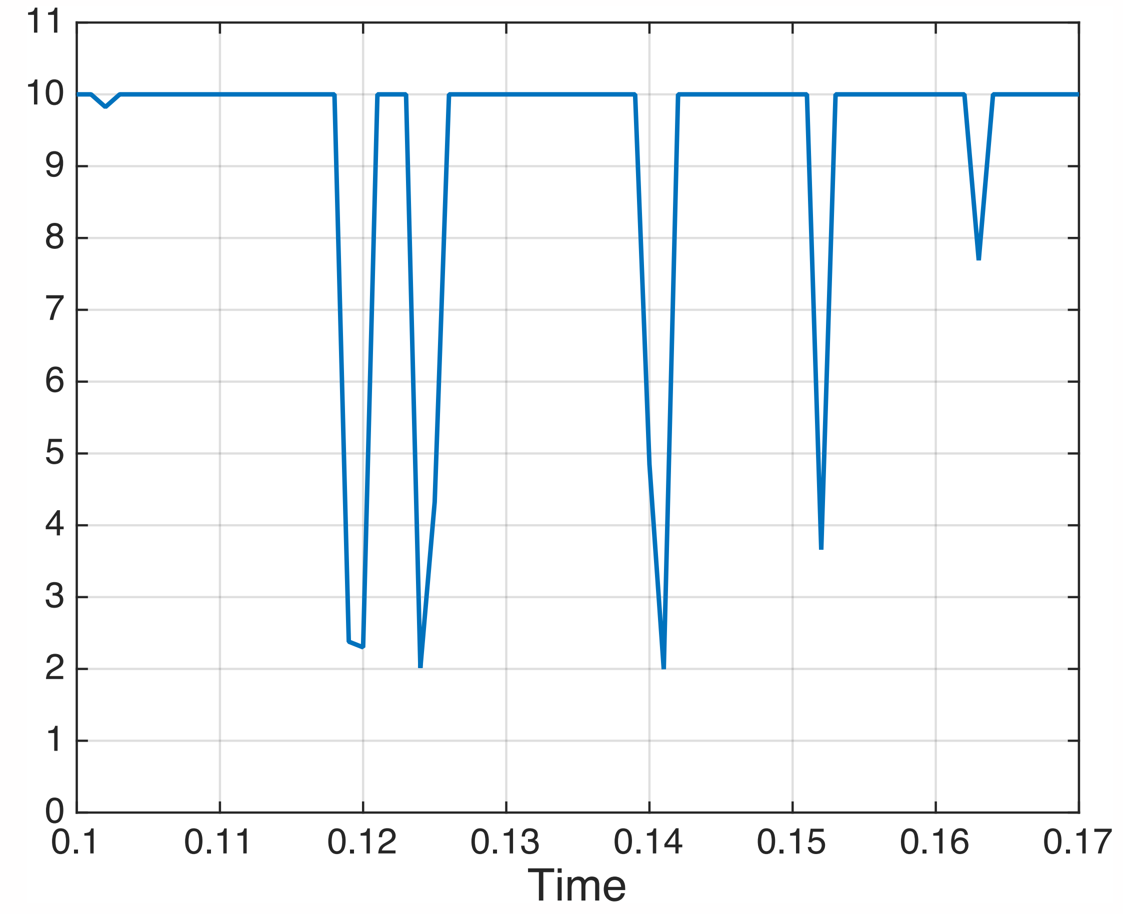
$\mathbf{r}_1$: estimated reflectivity series after the first step of RLS

Why RL1 does not work for our case?



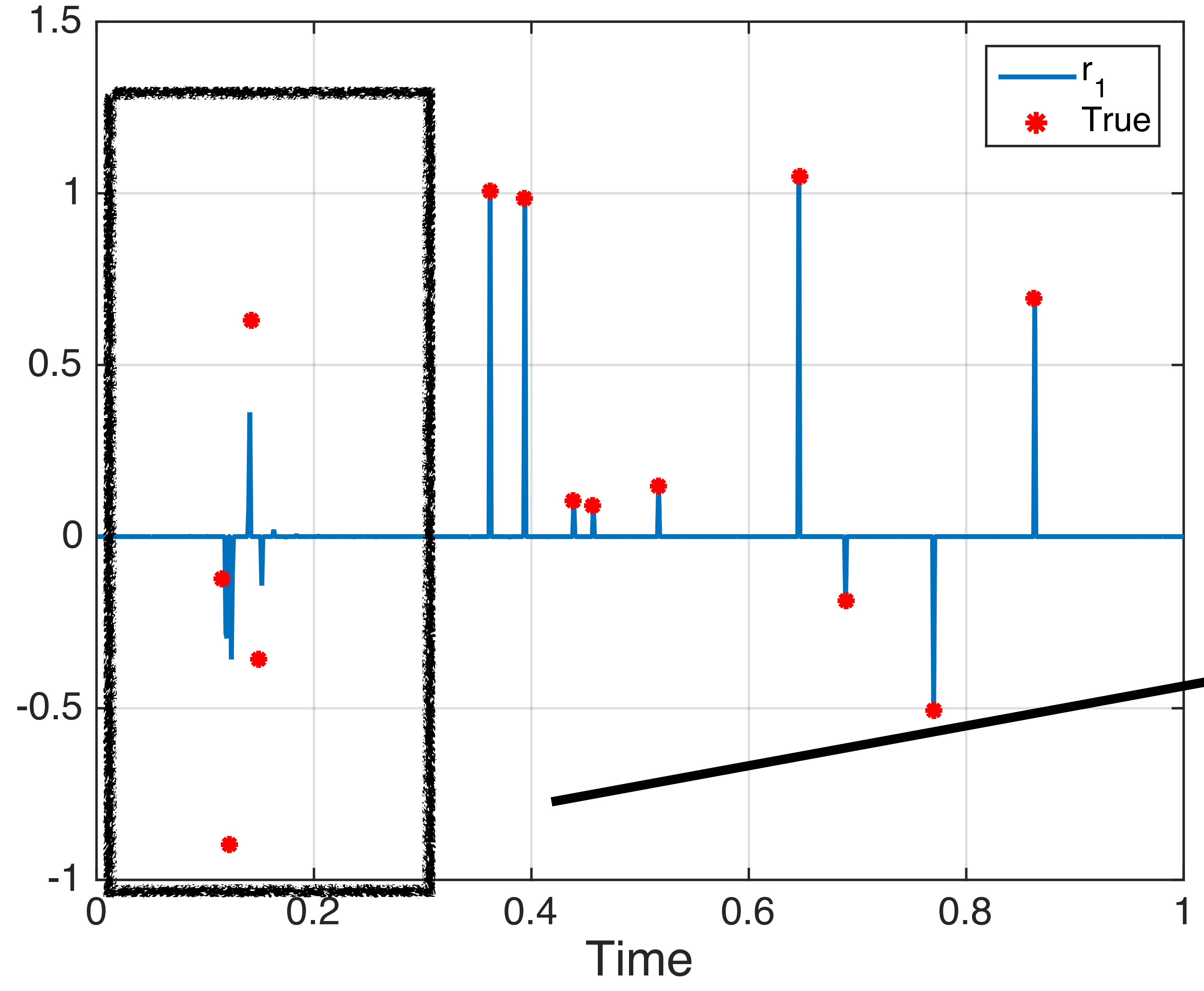


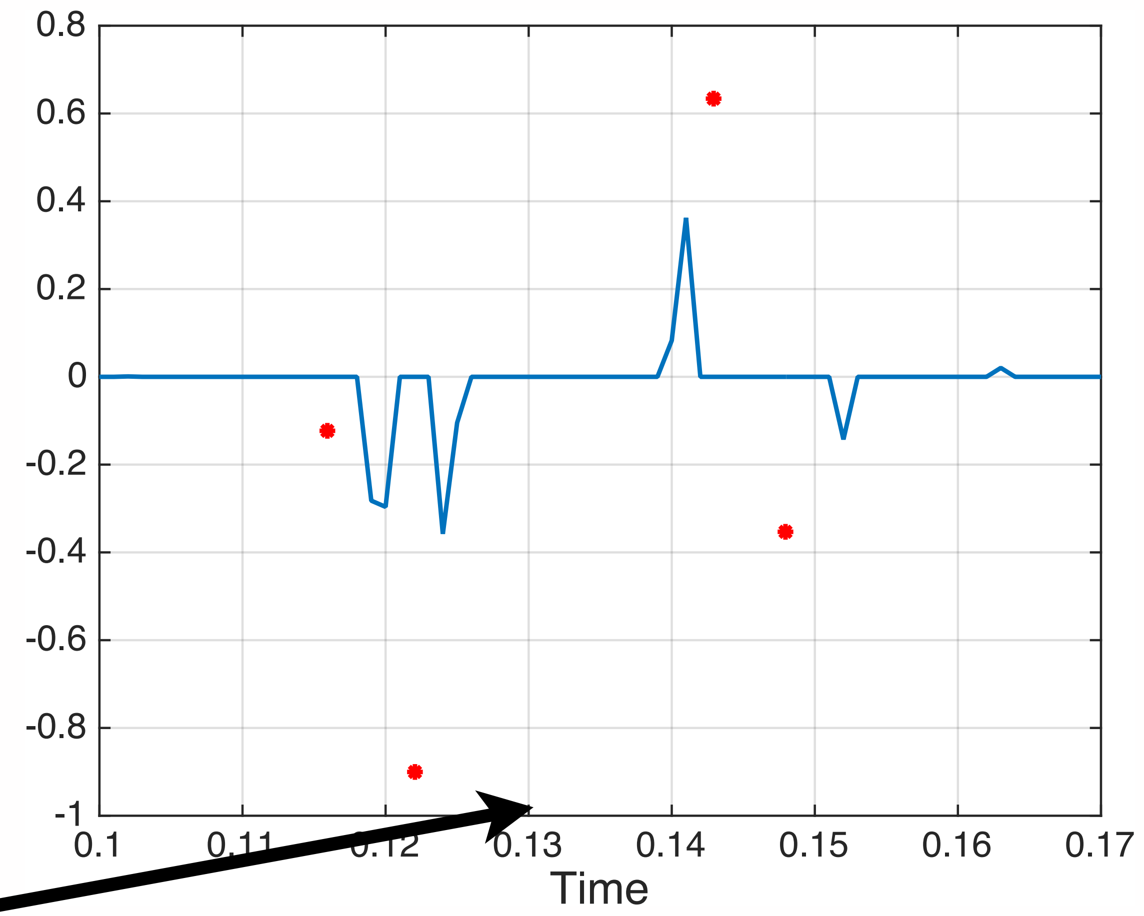
$\mathbf{r}_1$



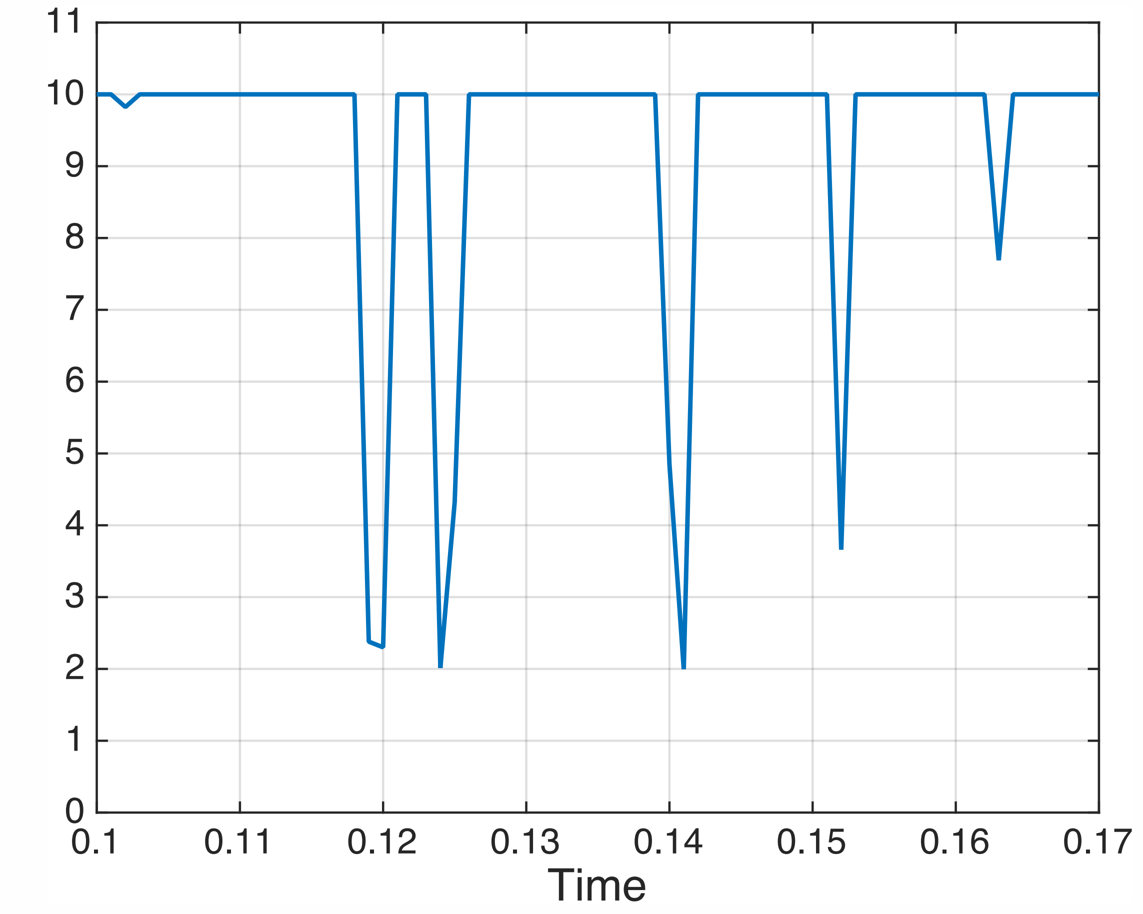
$\mathbf{a}_2$

Why RL1 does not work for our case?





$\mathbf{r}_1$

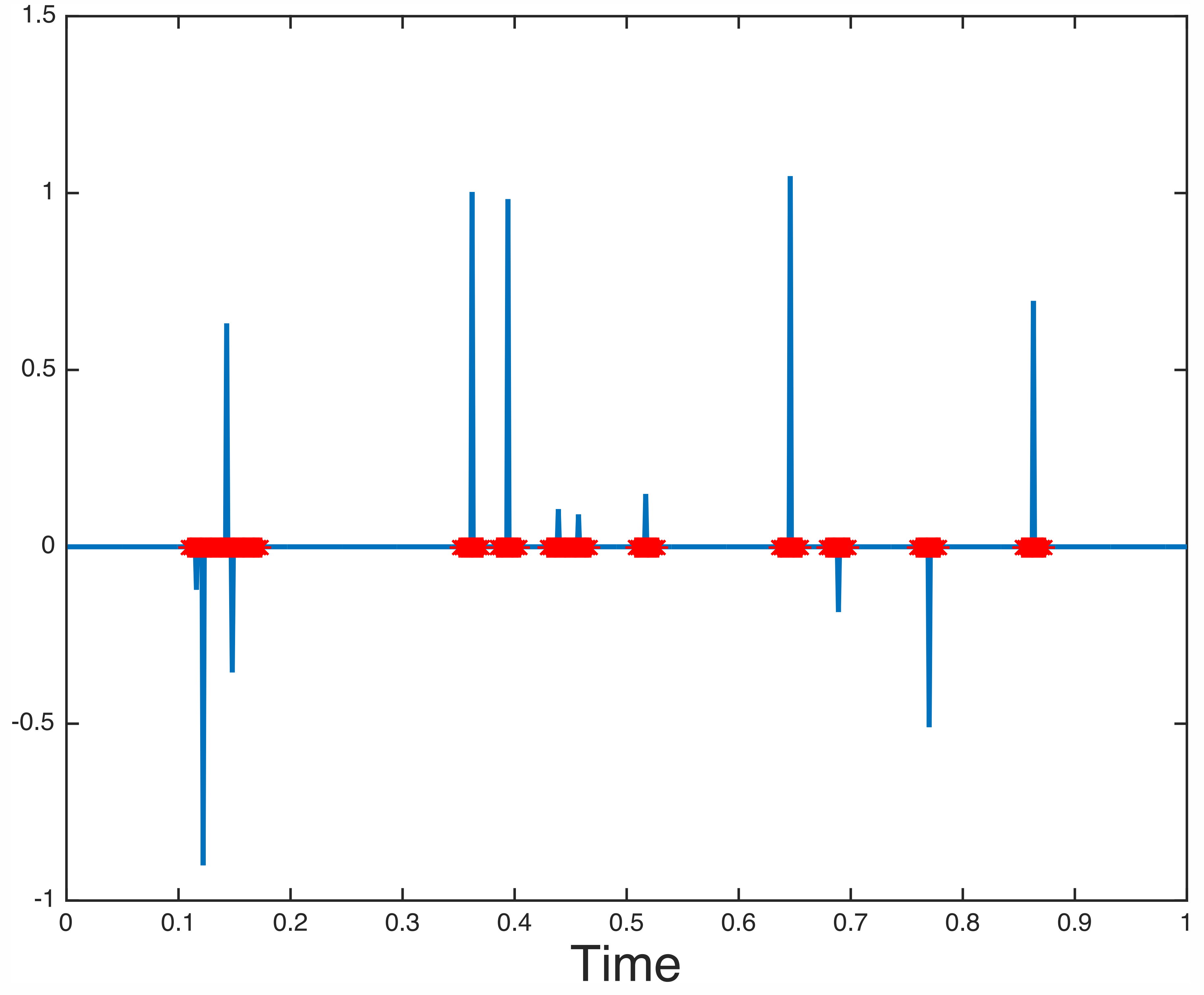


$\mathbf{a}_2$

Weights only promotes previously detected supports

Solution

The true support is within 1.5 wavelength of the nonzeros of $\mathbf{r}_1$



Modified Reweighted L1

For $l = 1, ..., L$**:**

$$\text{size}(\mathcal{N}_i) = \frac{1.5N}{m_1 - m_2}\frac{1}{l}, \text{ for } i = 1, ..., N$$

For $k = 1, ..., K$**:**

$$\text{Define weight: } \mathbf{a}_k = [\mathbf{a}_k(1), ..., \mathbf{a}_k(n)], \quad \mathbf{a}_k(i) = \frac{1}{\epsilon + \min_{j \in \mathcal{N}_i} |\mathbf{r}_k(j)|^{1-q}}$$

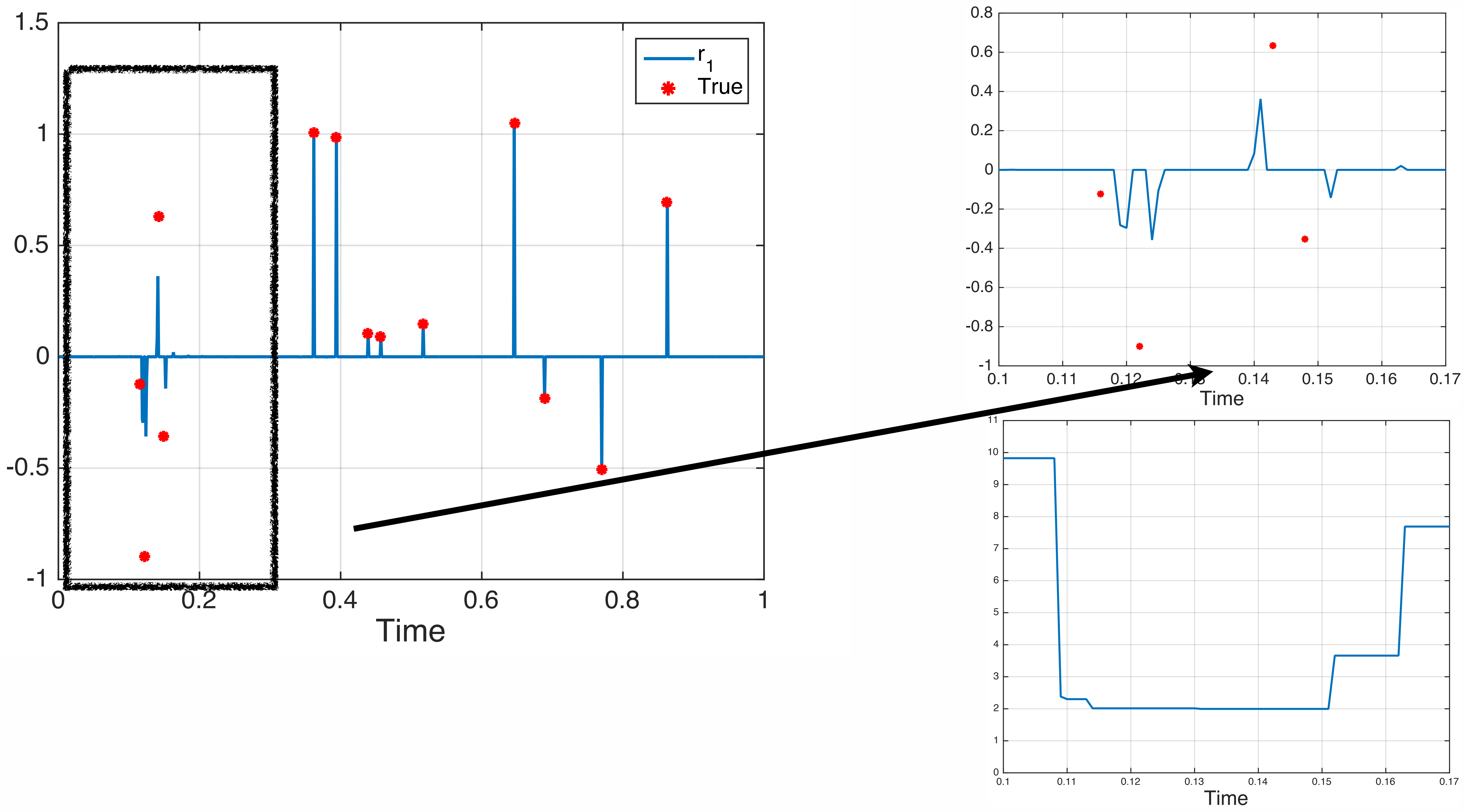
$$\text{Update } \mathbf{r}_k: \quad \mathbf{r}_{k+1} = \min_{\mathbf{r}} \|\mathbf{r}\|_1 + \lambda \|\mathbf{a}_k \odot \mathbf{r}\|_1$$

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

End

End

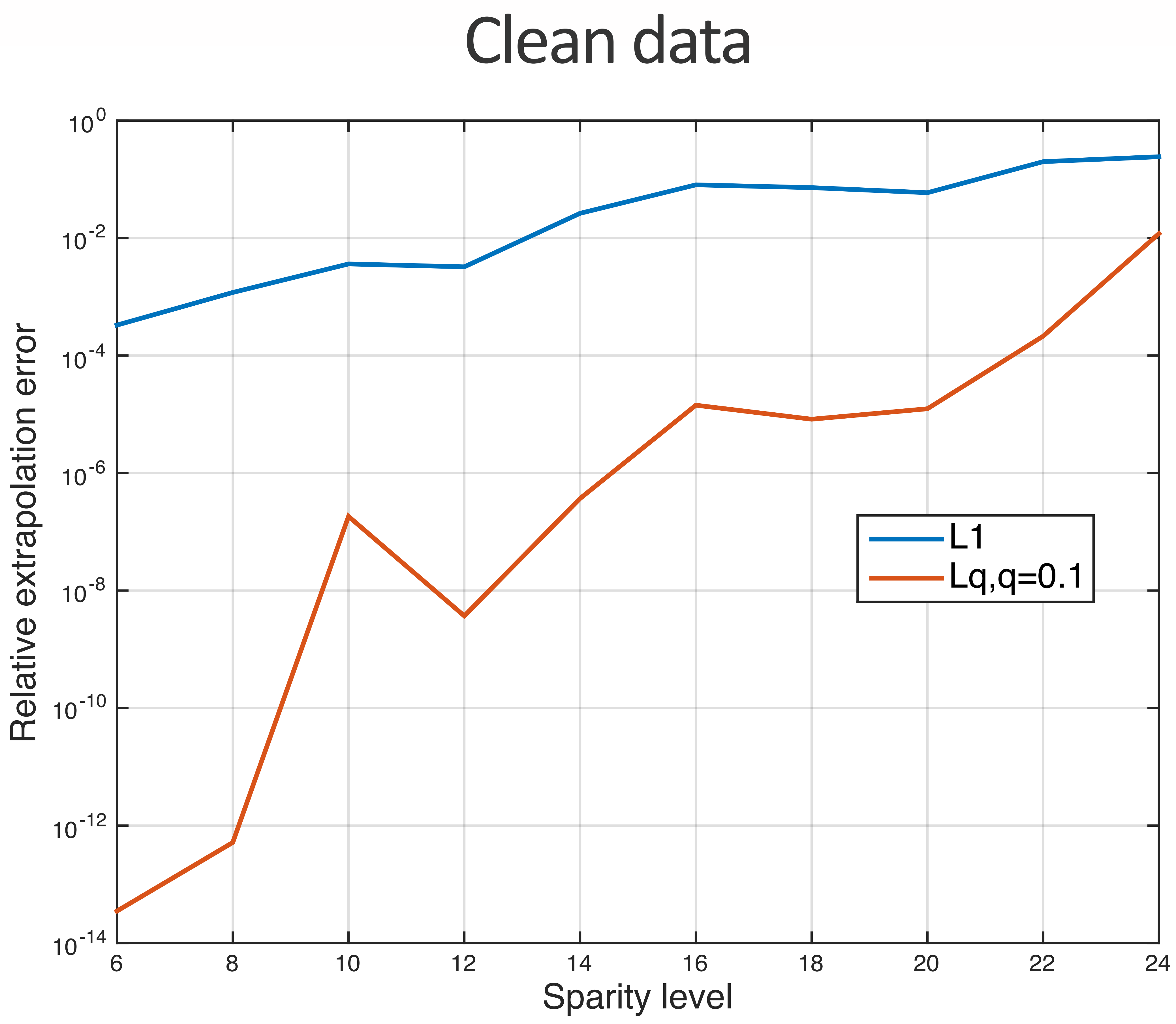
Back to previous example



$\mathbf{r}_1$

$\mathbf{a}_2$

Numerical test



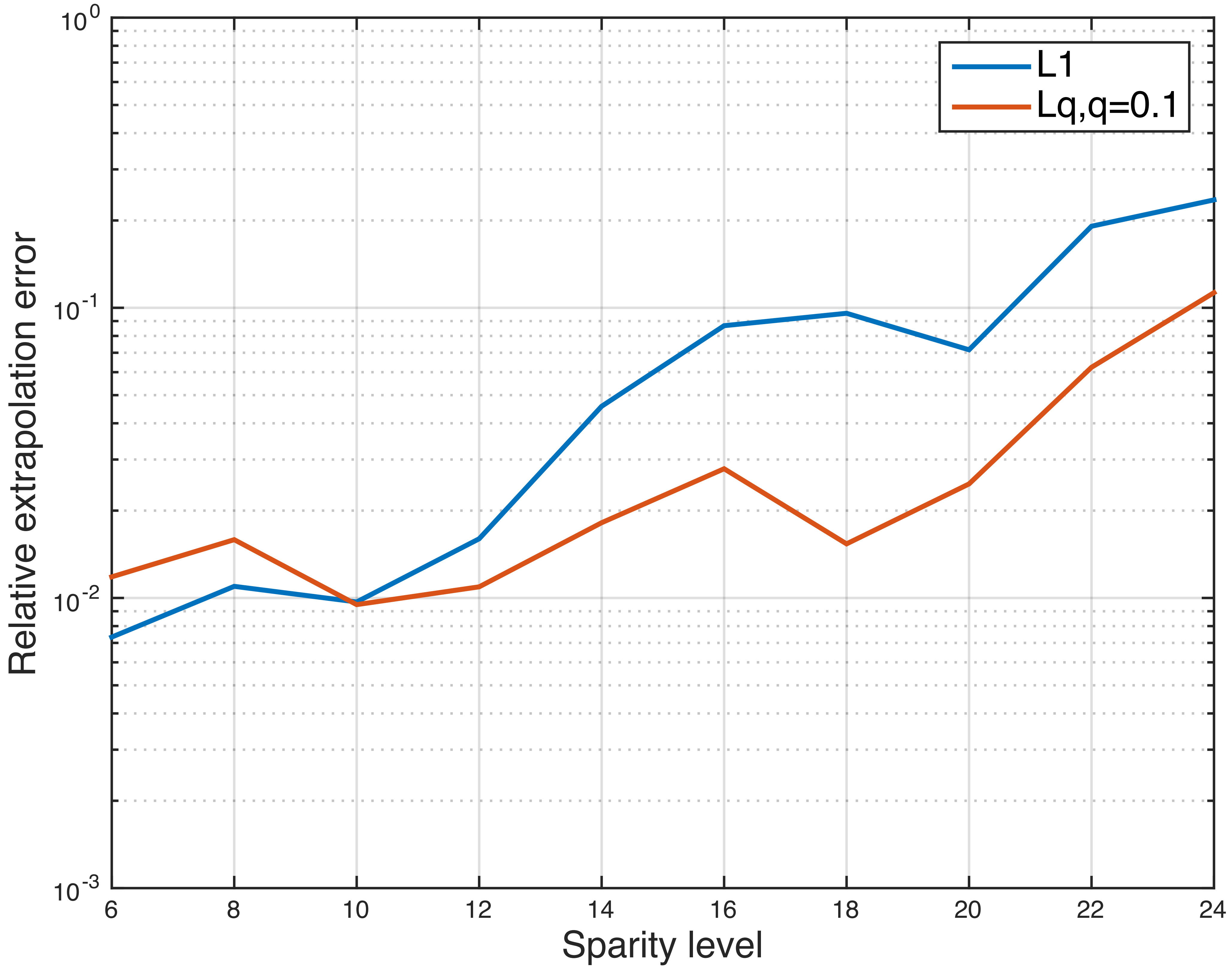
Extrapolation from {10,..,50}Hz towards {1,...,10}Hz

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

N=1000, m1=10, m2=50

Numerical test

SNR=20



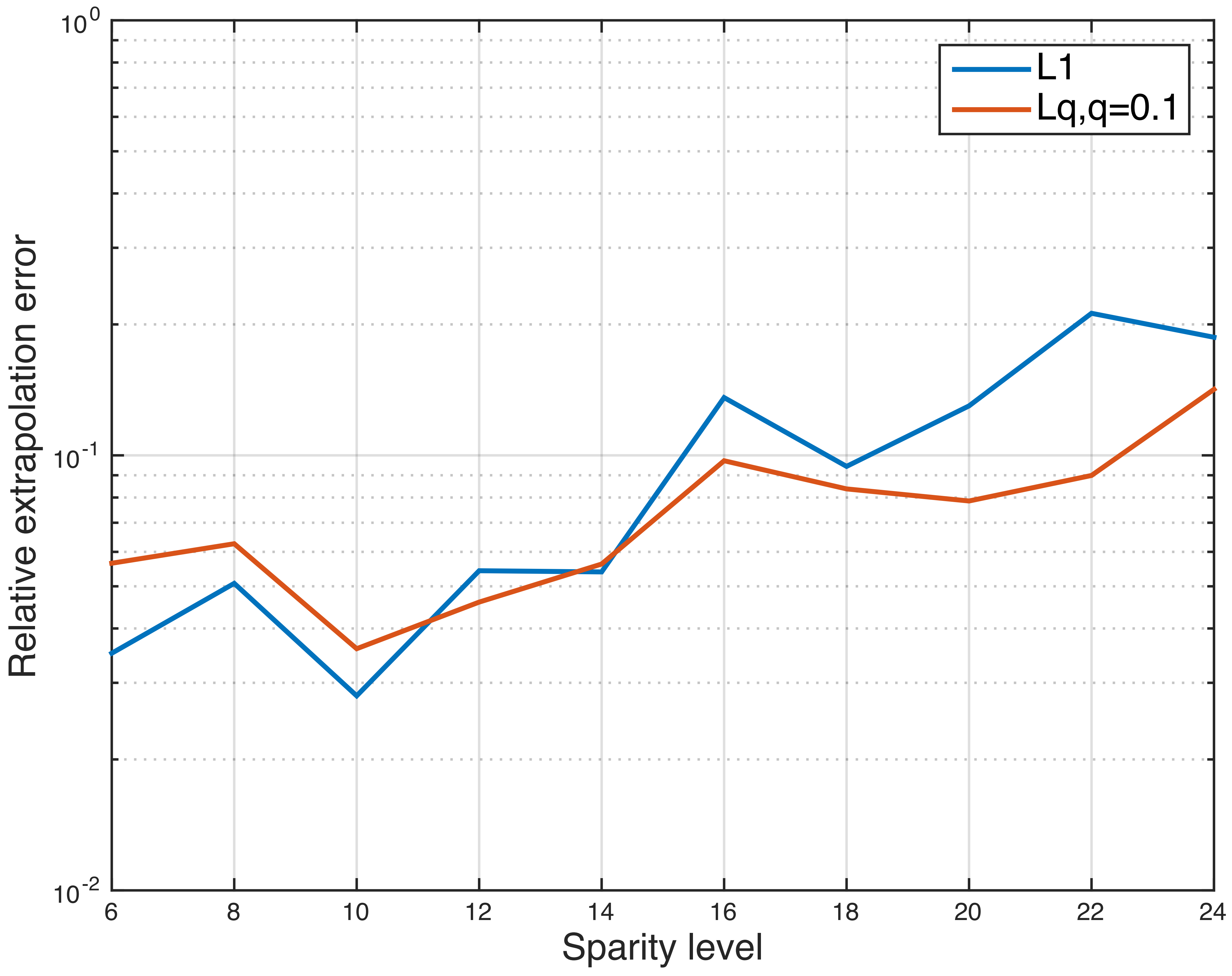
Extrapolation from {10,..,50}Hz towards {1,...,10}Hz

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

N=1000, m1=10, m2=50

Numerical test

SNR=13



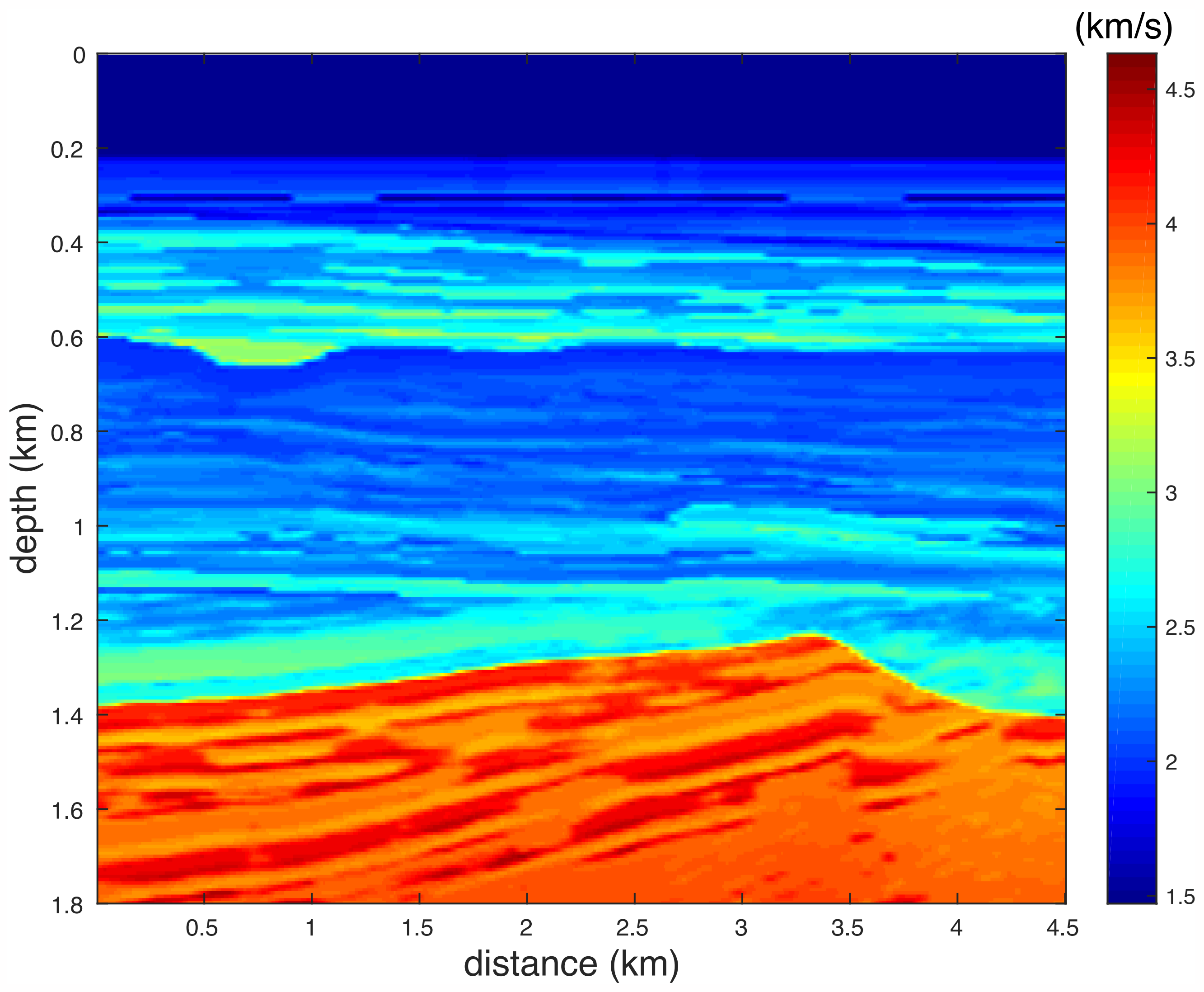
Extrapolation from {10,..,50}Hz towards {1,...,10}Hz

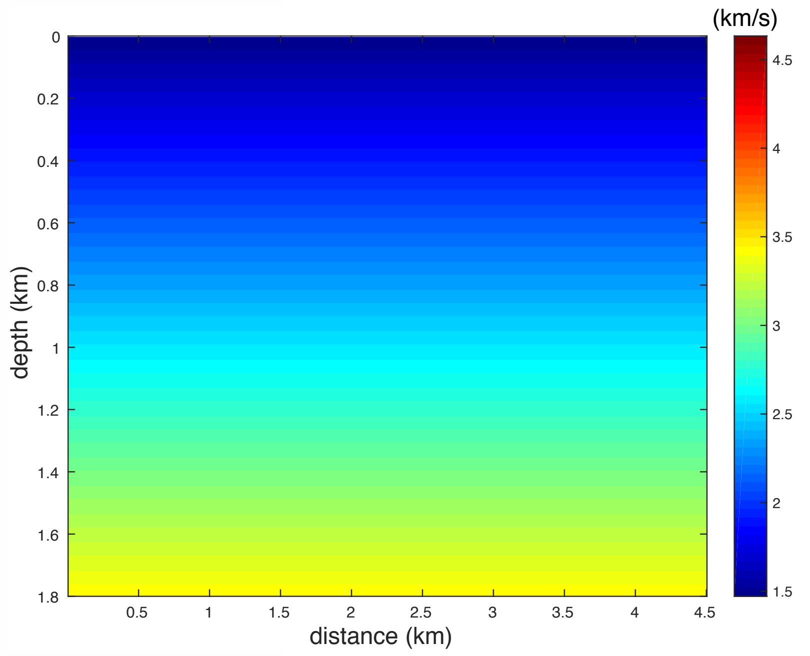
Each point is the mean error on 50 random draws of sparse signals

N=1000, m1=10, m2=50

Synthetic data - Non-inversion crime

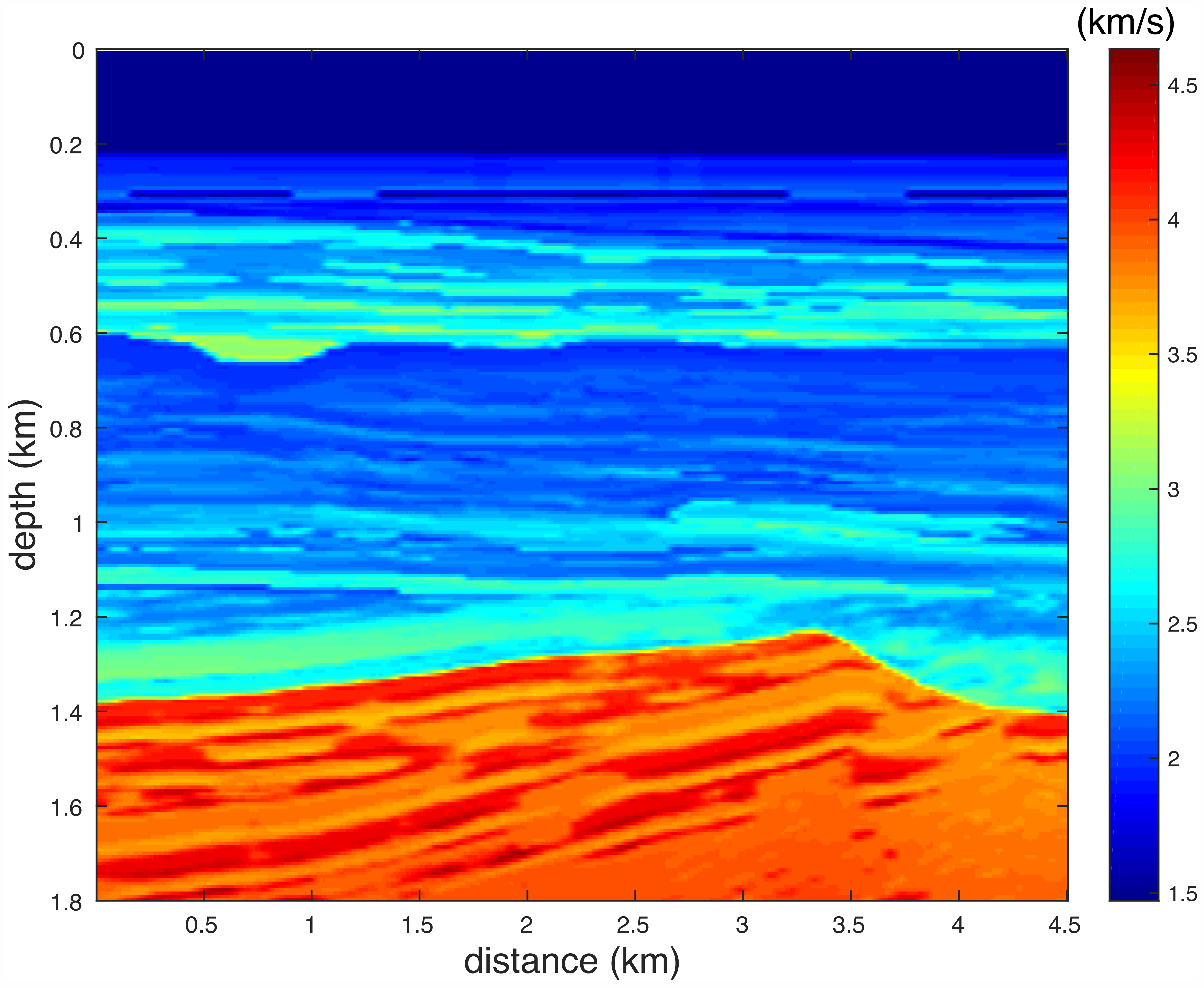
- IWAVE generated data
- inversion using time harmonics
- 3 frequency sweeps, 20 lbfgs-iters for each batch
- 20Hz Ricker wavelet
- source spacing : 0.2km
- receiver spacing : 20m
- maximum offset : 1km
- model size : 1.8km × 4.5km



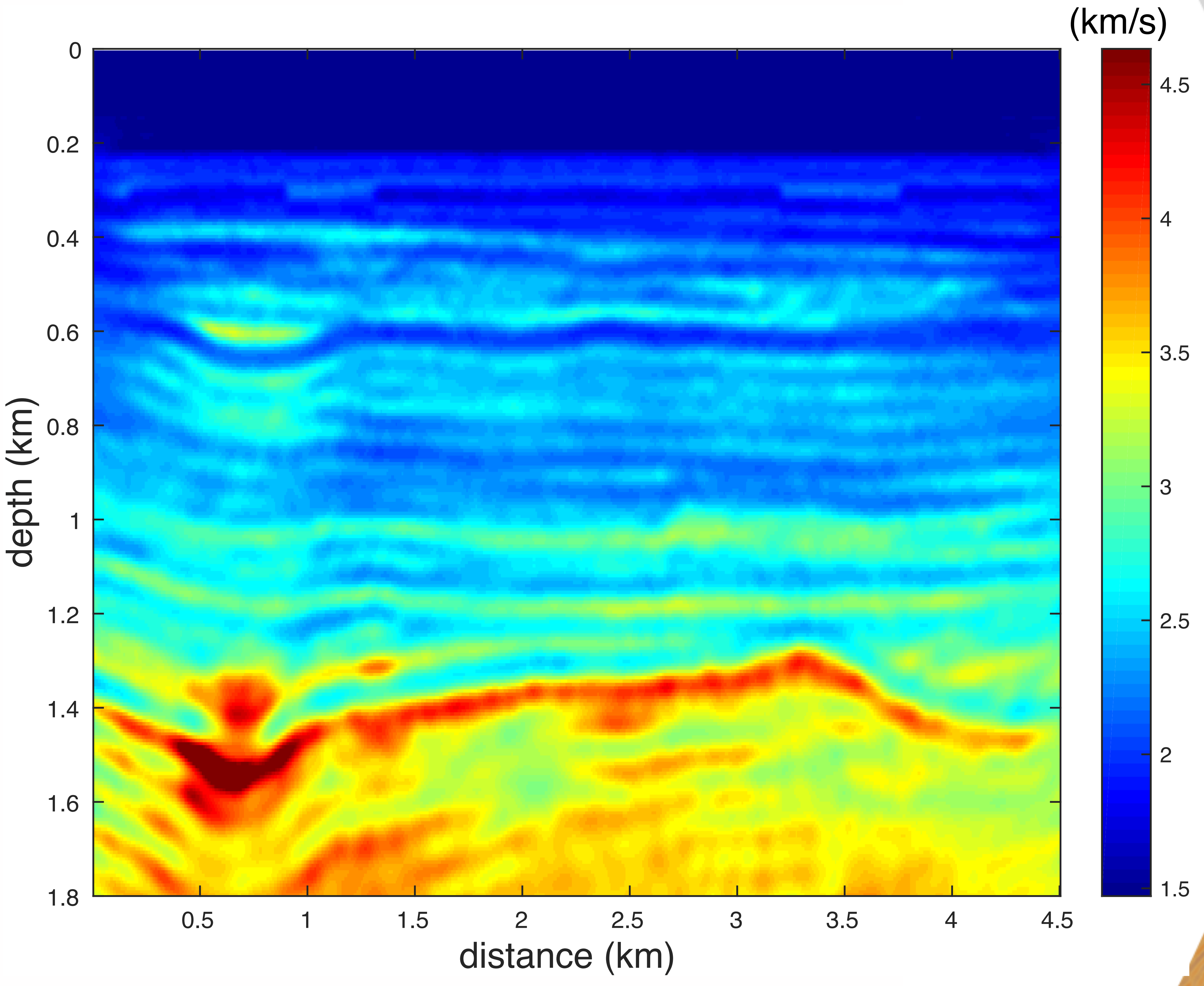


Recovery of FWI

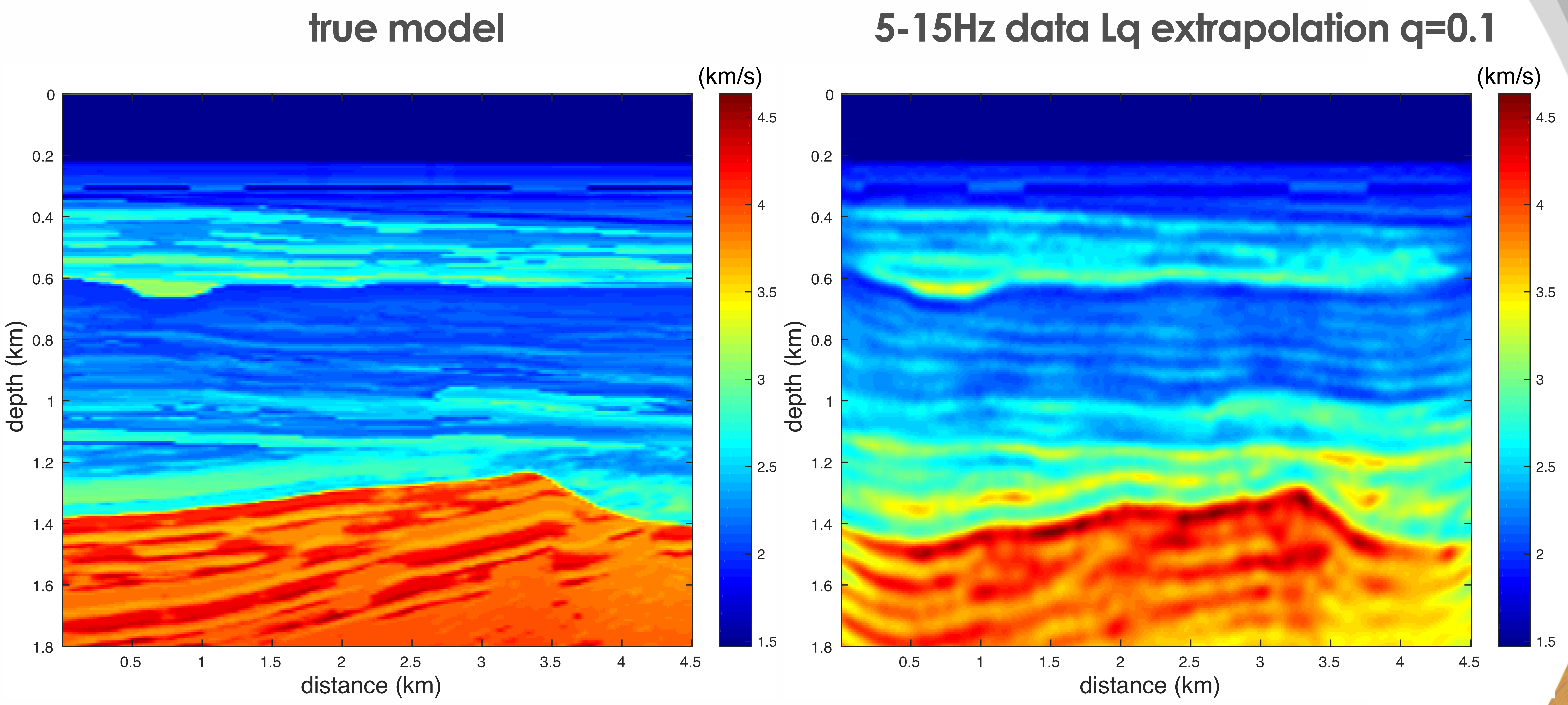
true model



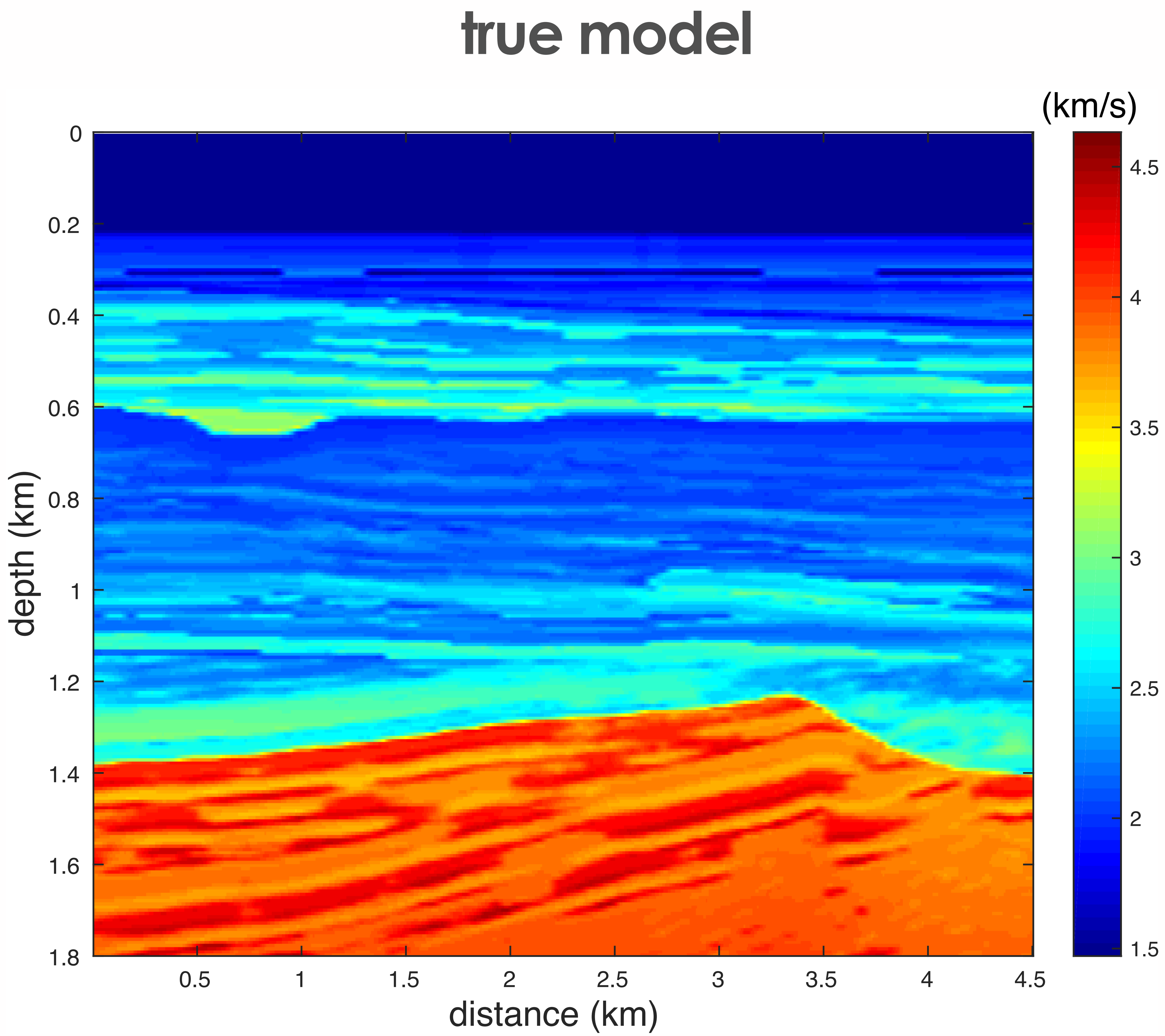
5-15Hz data w/o extrapolation

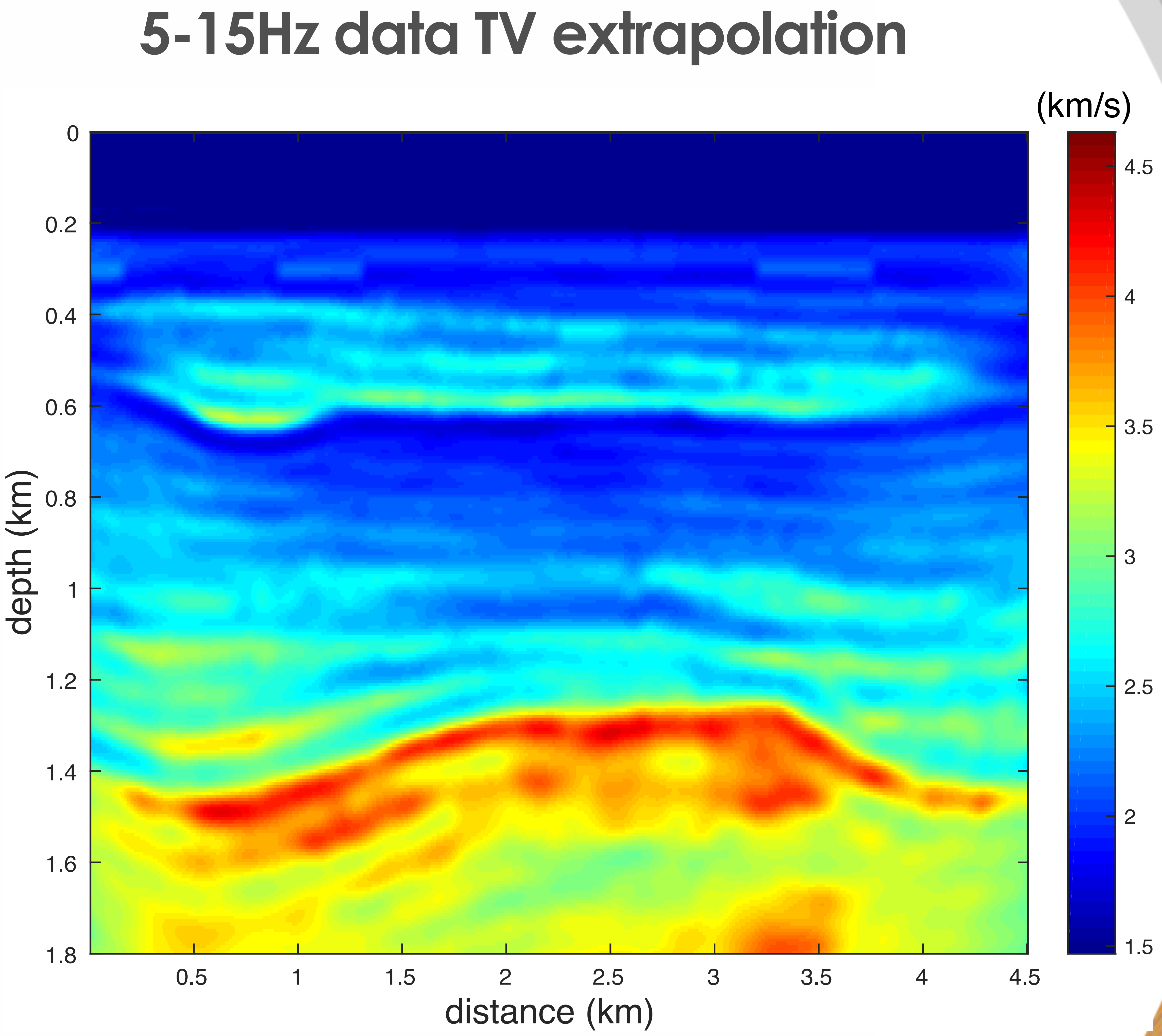


Recovery of FWI

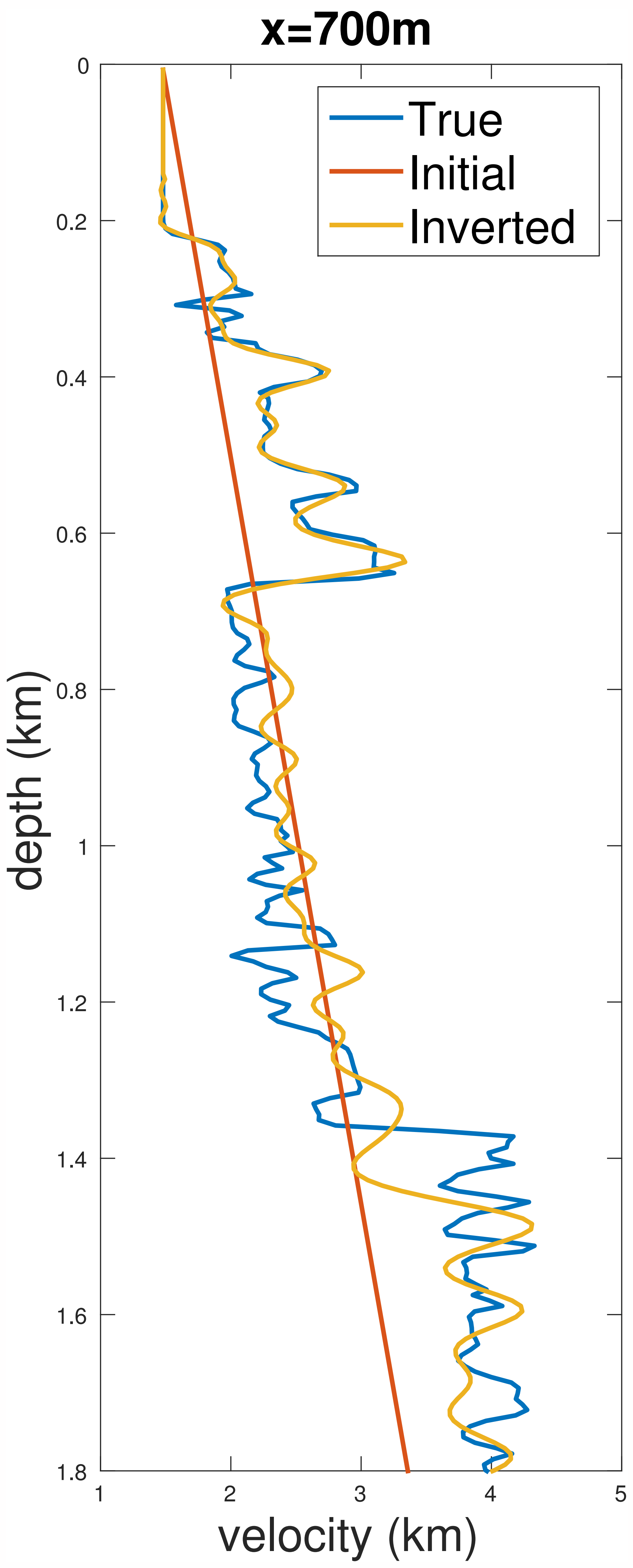


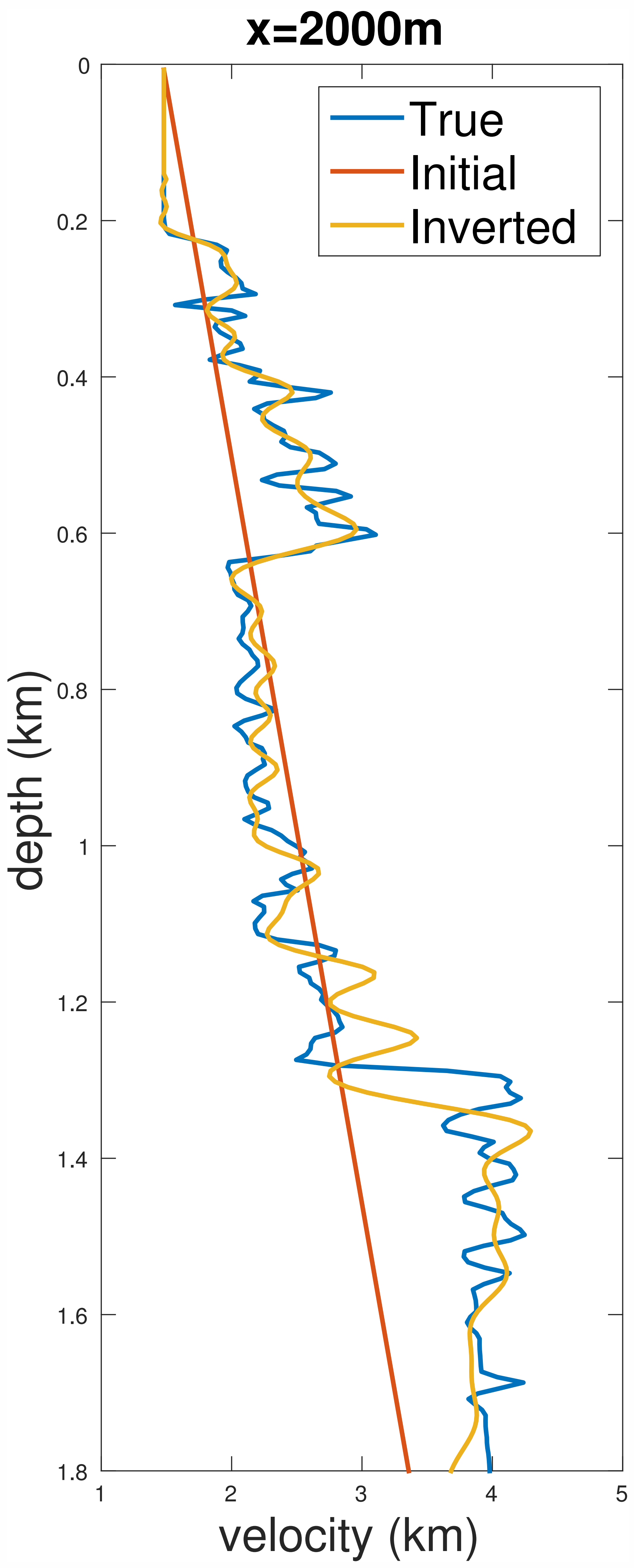
Recovery of FWI

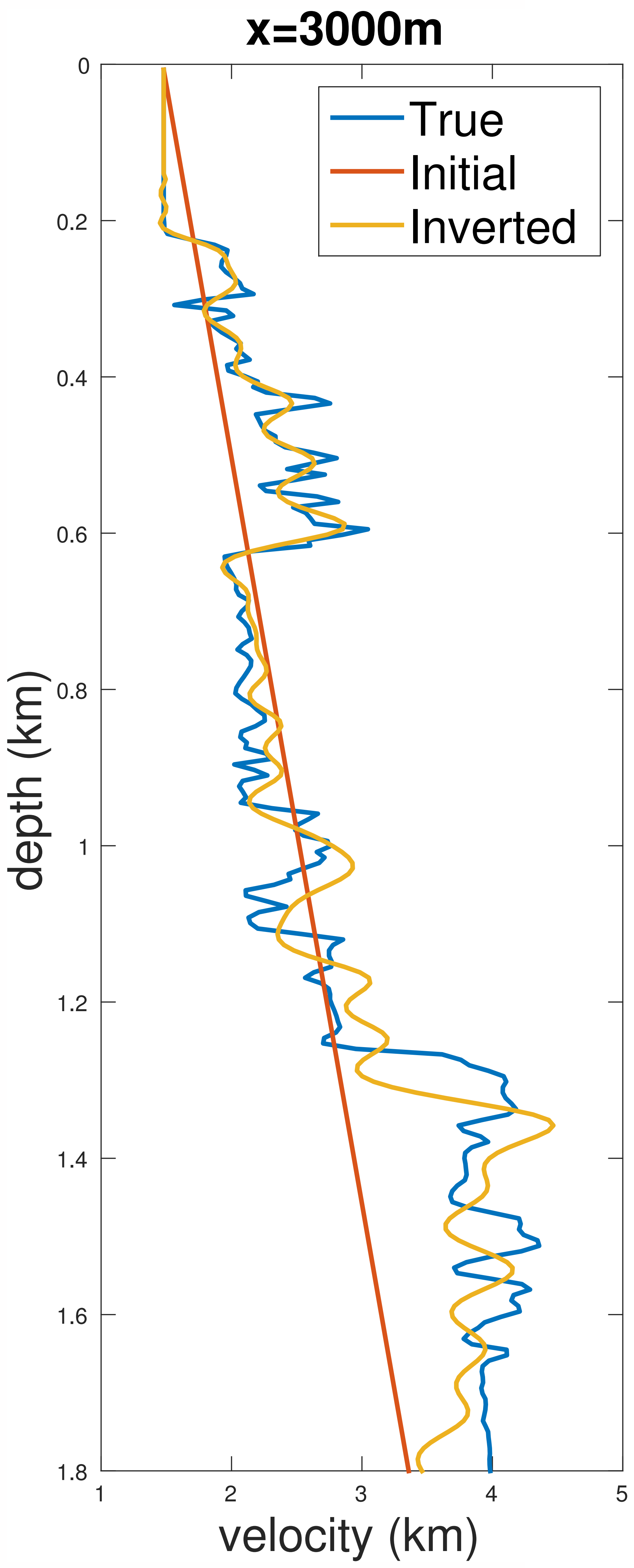




Cross-sections for the inverted model by Lq







Conclusion

- We proposed two methods for frequency down-extrapolation
- Both are stable with respect to additive noise and dispersion
- The TV norm minimization is more stable
- The Lq+L1 norm minimization is better at resolving close by spikes

Future work

- Incorporate TV regularizer to the Lq minimization
- Test the methods on real data.

Acknowledgements

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