Wavefield recovery with limited-subspace weighted matrix factorizations

Yijun Zhang, Shashin Sharan, Oscar Lopez, and Felix J. Herrmann

SEG 2020

SLIM
Georgia Institute of Technology

Motivation

Fully sampled data is needed for

- multiple removal
- migration & FWI

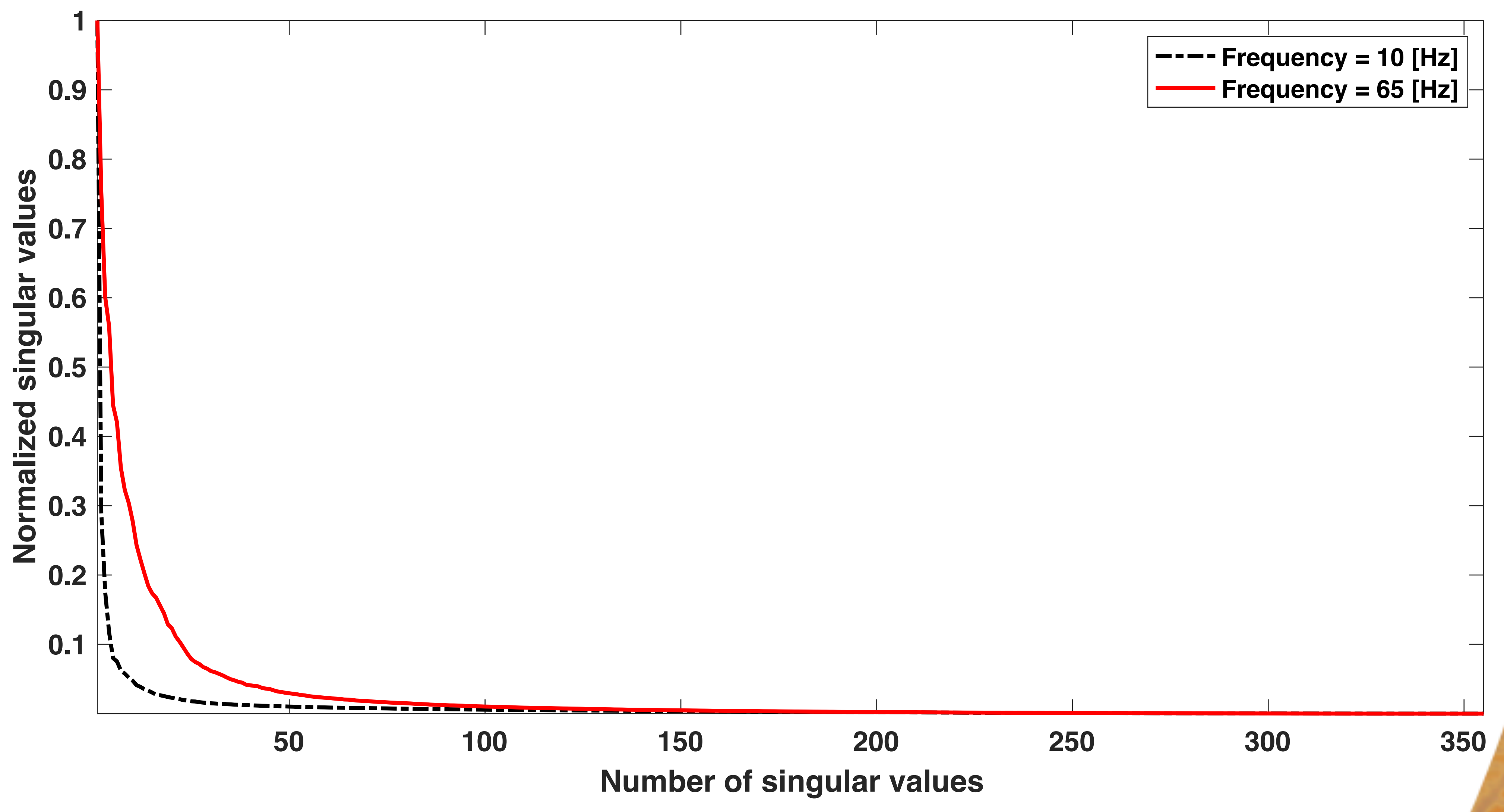
Dense seismic data acquisition

- budget limitations
- operationally challenging

Motivation

Matrix completion

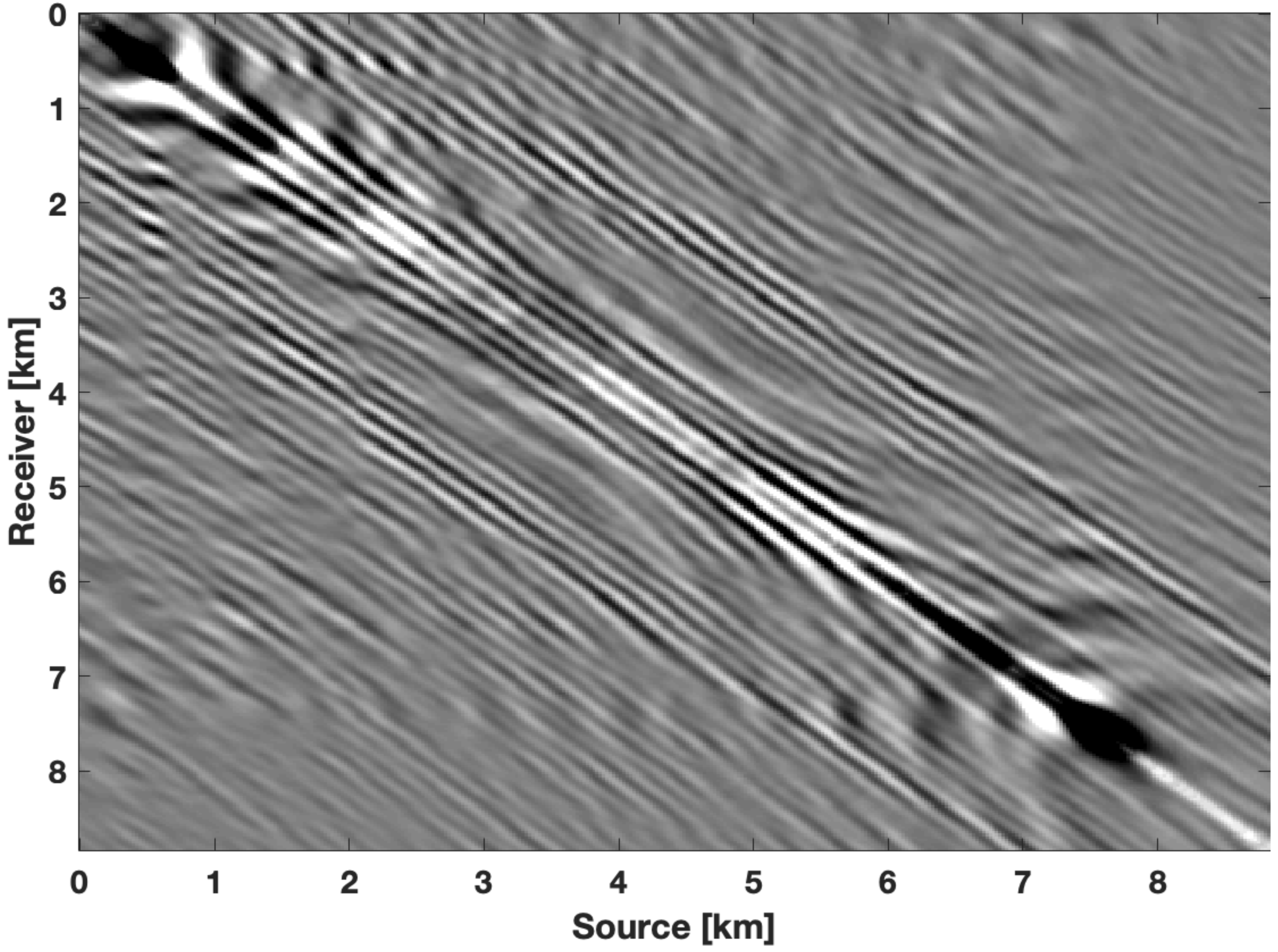
- exploits low-rank structure for recovery
- performance degrades with increasing frequency



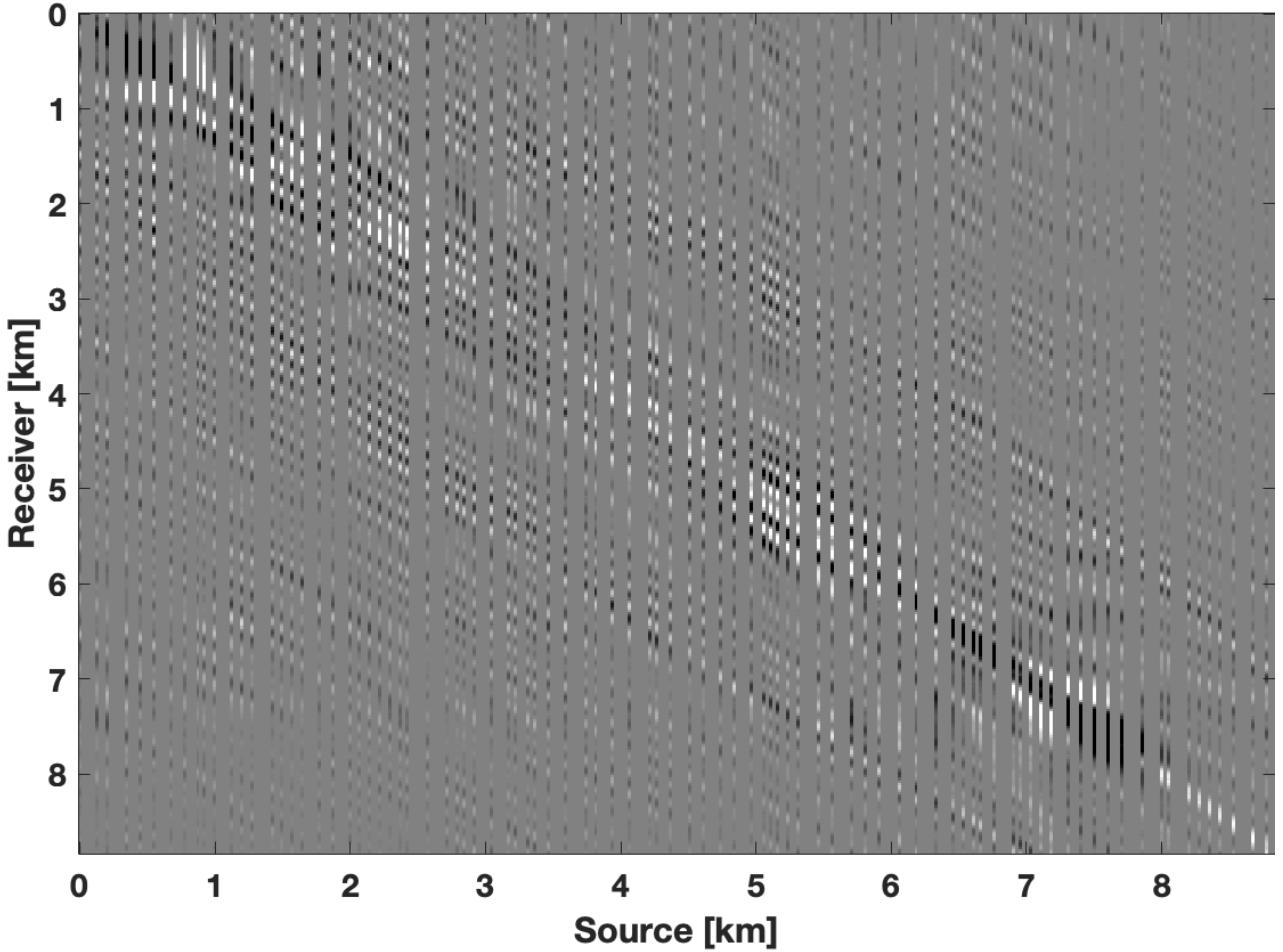
Weighted matrix completion

- gives good results for higher frequency slices
- higher rank needed for higher frequencies
- increases chances of potential overfitting

Fully sampled data: 22 Hz slice

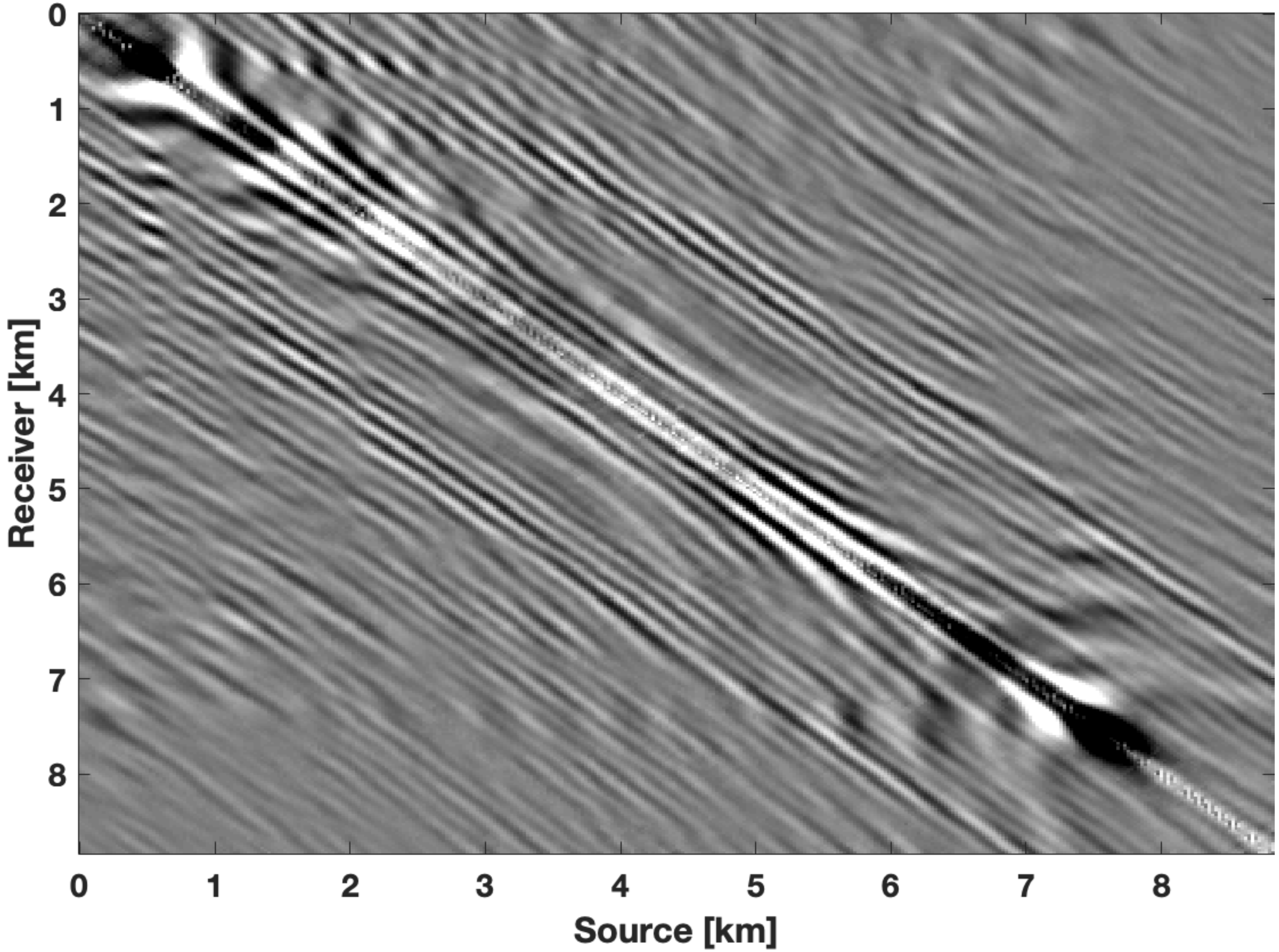


Observed data (75% jittered subsampled)



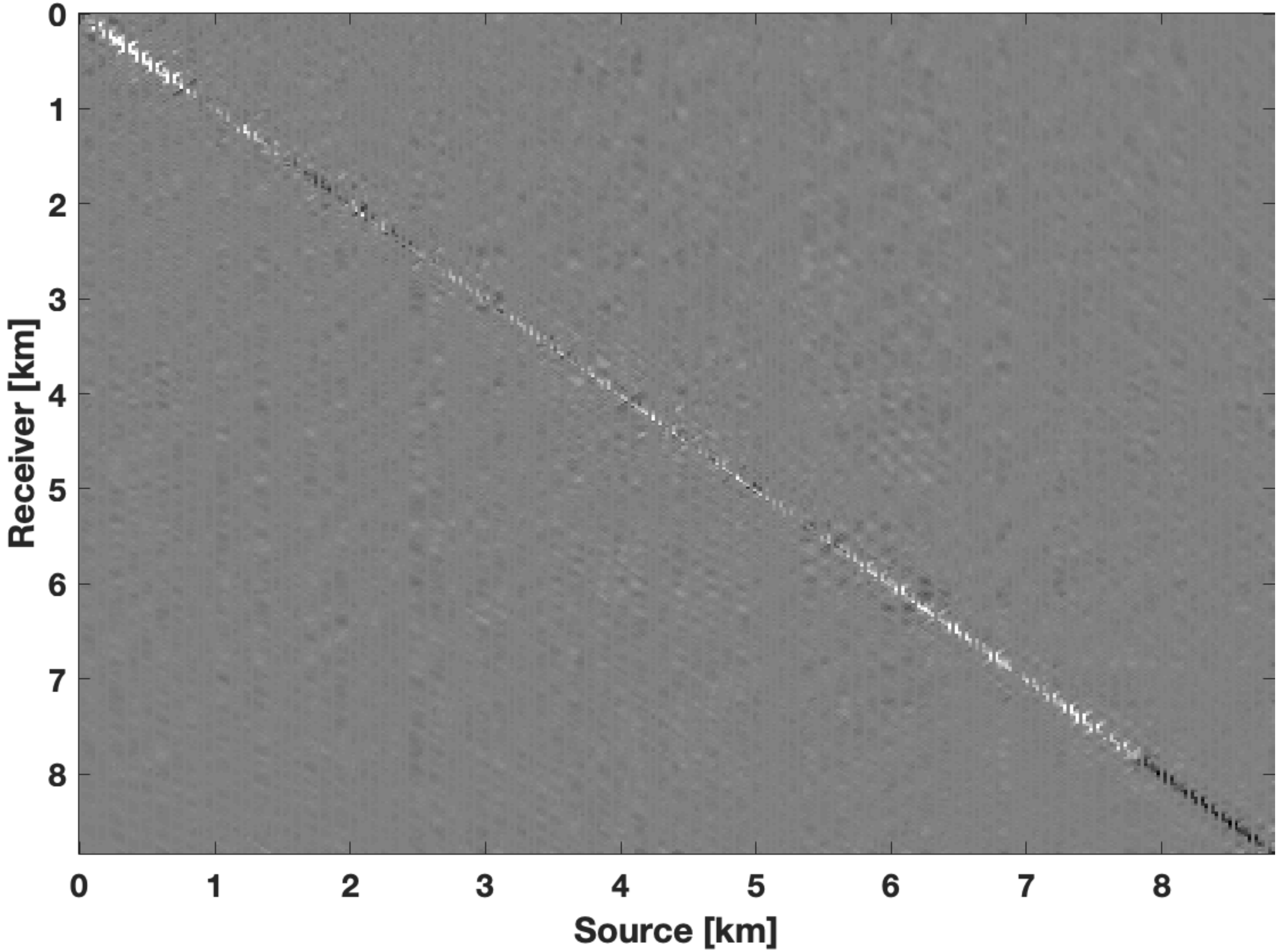
Zhang, Y. et al. "High-frequency wavefield recovery with weighted matrix factorizations." In SEG Technical Program Expanded Abstracts, 2019.

Recovery w/ recursively weighted (rank = 25)



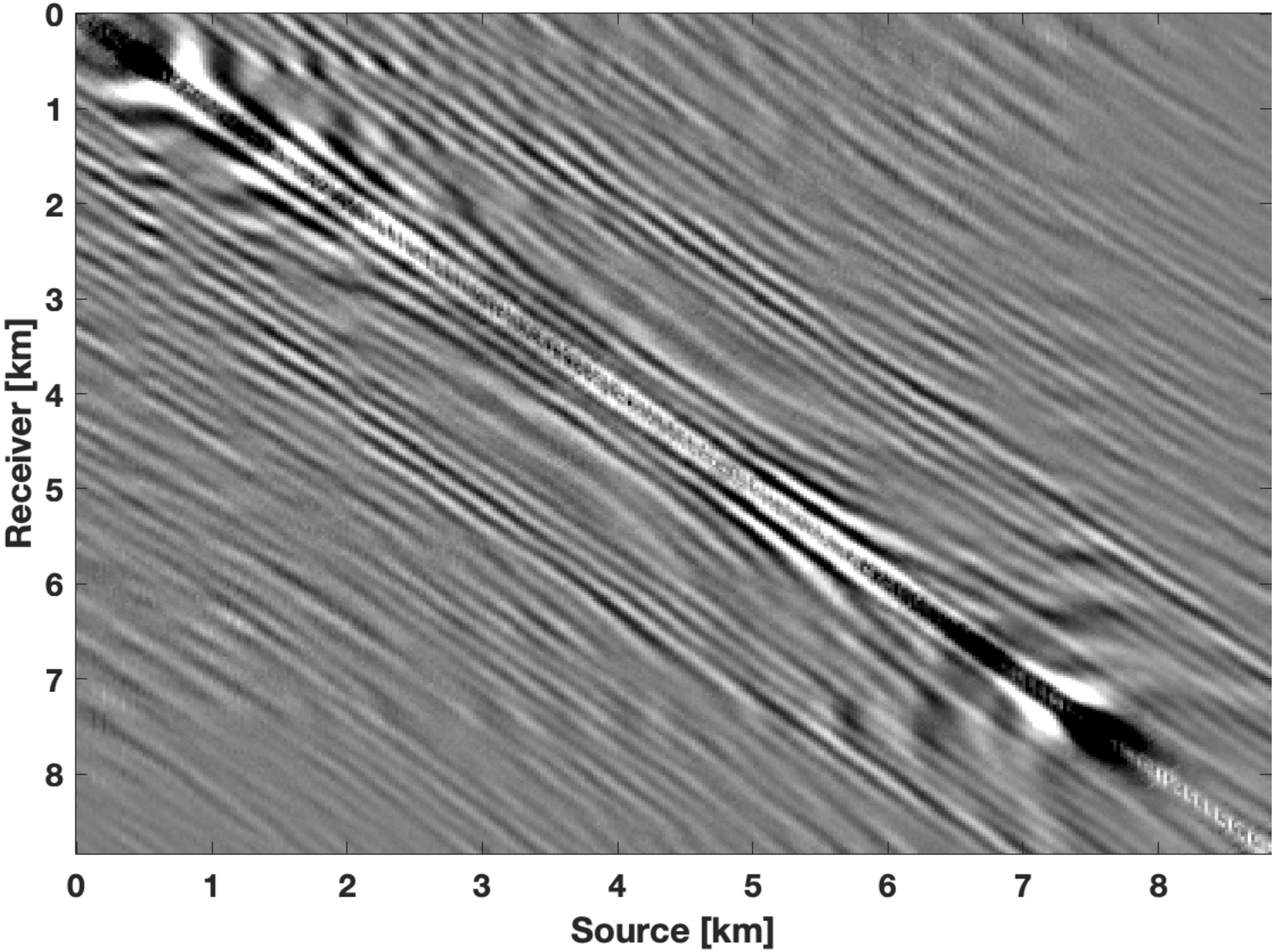
SNR = 15.50 dB
Rank = 25

Difference: True - Recovery w/ recursively weighted (rank = 25)



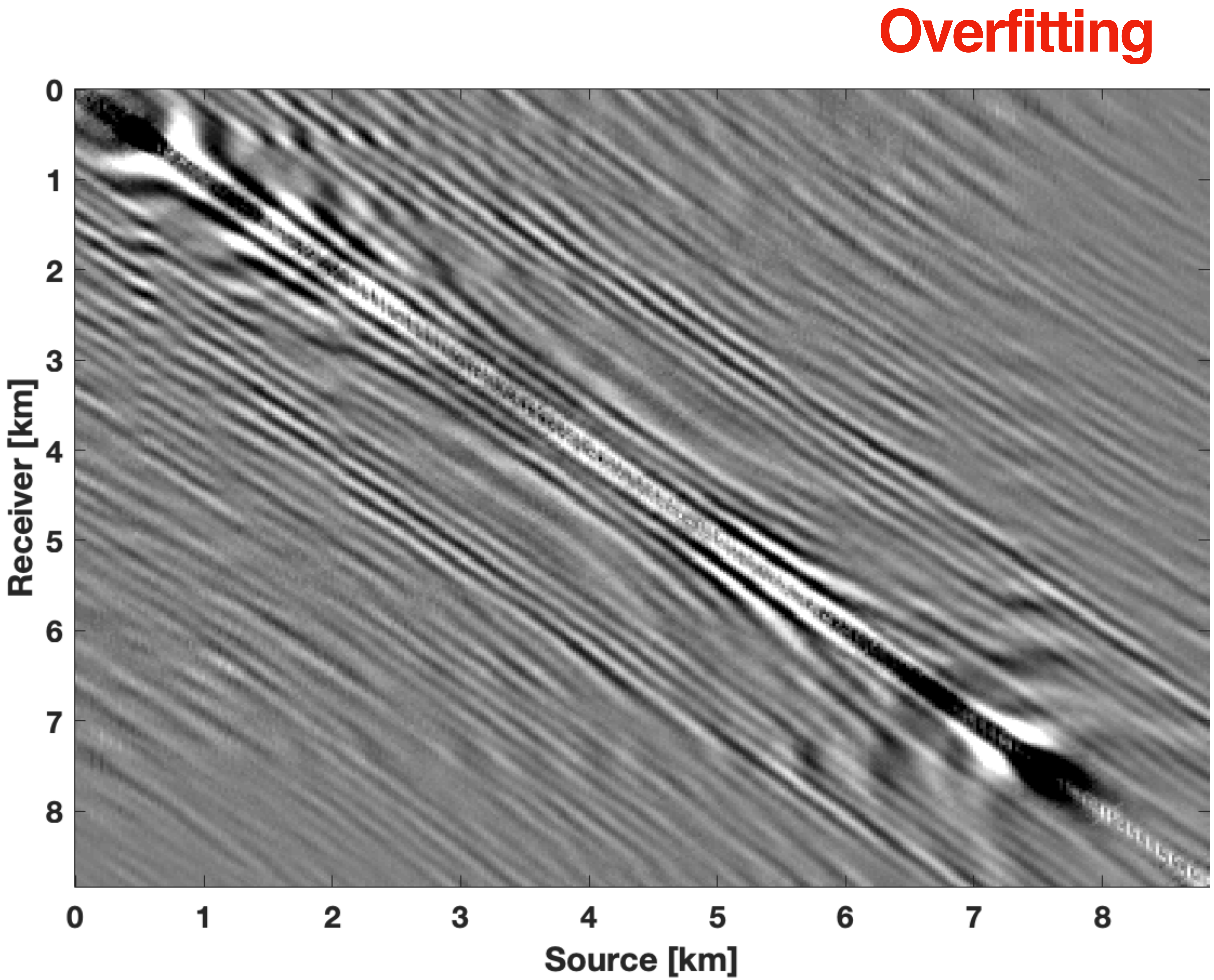
SNR = 15.50 dB
Rank = 25

Recovery w/recursively weighted (rank = 85)



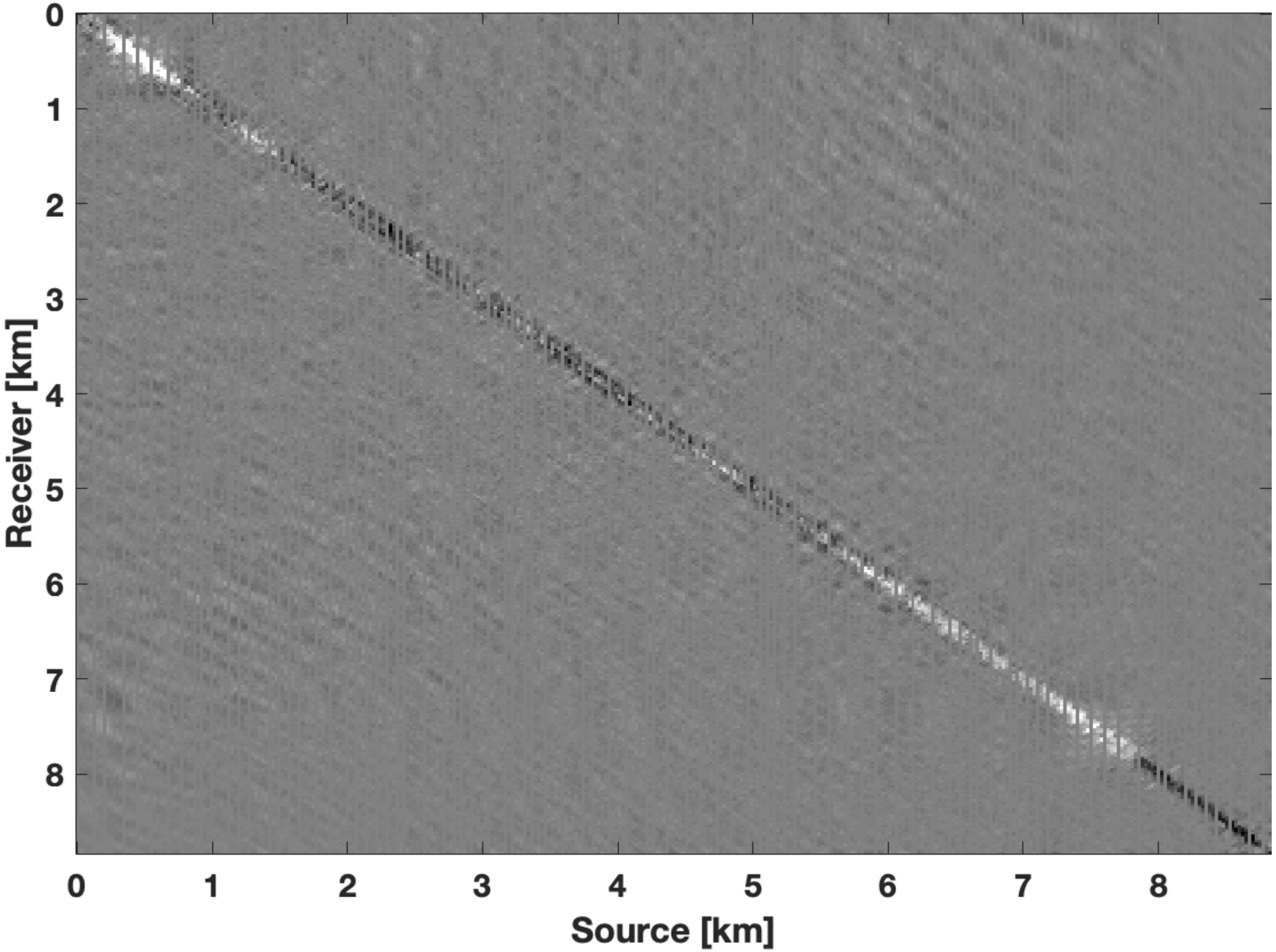
SNR = 13.09 dB
Rank = 85

Recovery w/recursively weighted (rank = 85)



SNR = 13.09 dB
Rank = 85

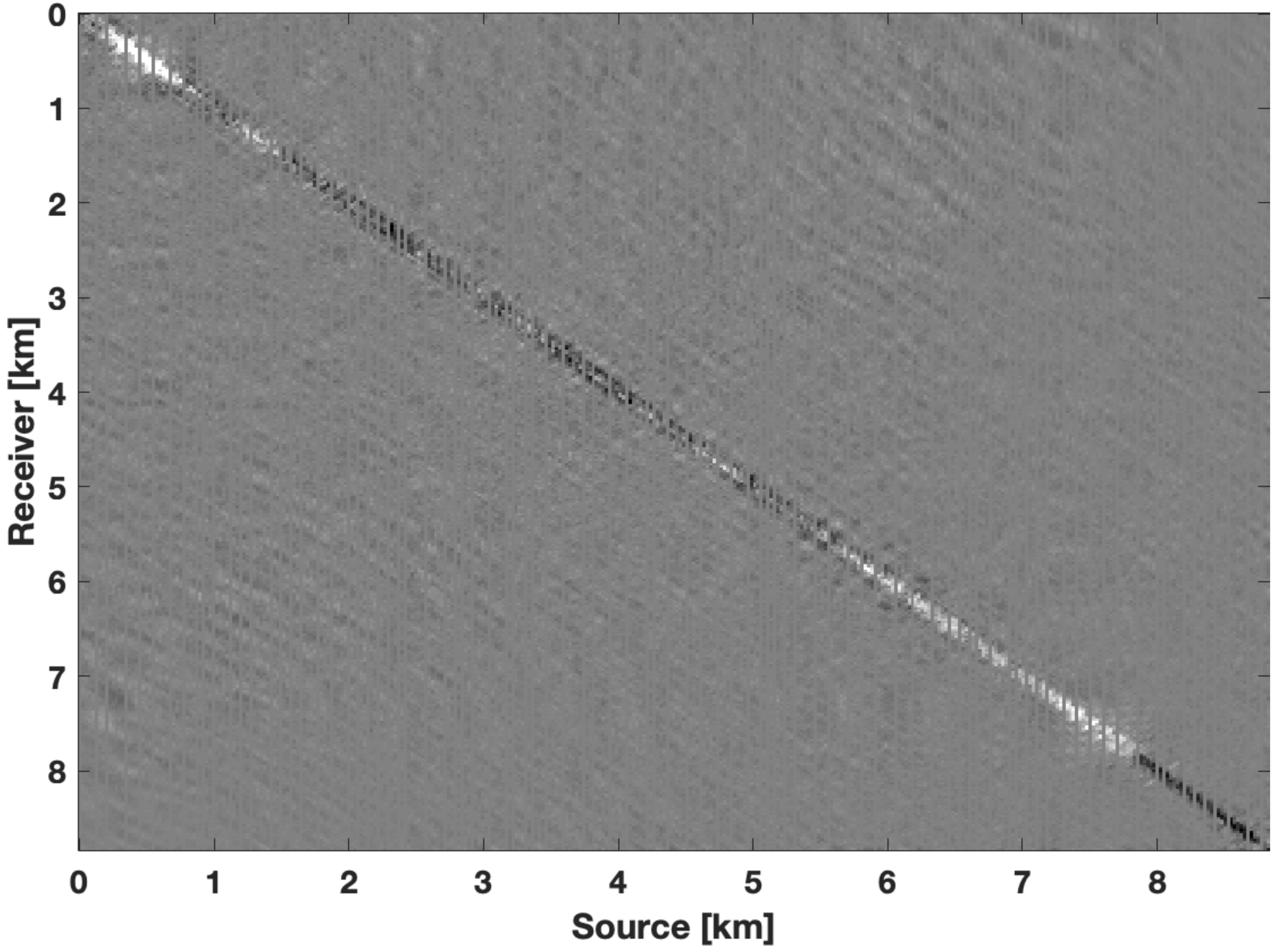
Difference: True - Recovery w/ recursively weighted (rank = 85)



SNR = 13.09 dB
Rank = 85

Difference: True - Recovery w/ recursively weighted (rank = 85)

Overfitting



SNR = 13.09 dB
Rank = 85

Question

How can we use a higher rank for data reconstruction while avoiding the risk of overfitting?

Question

How can we use a higher rank for data reconstruction while avoiding the risk of overfitting?

Answer:

- *Limited-subspace weighted method*

Why limited-subspace weighted method?

Keeps the higher rank for reconstruction

- decreases the approximation errors at higher frequencies

Reduces the rank of prior information

- prevents overfitting
- further improves the wavefield recovery

Matrix completion

Successful reconstruction strategy

- Exploit structure
  - low rank/fast decay of singular values in “transform domain”
- Sample randomly
  - increase rank in “transform domain”
- Optimization
  - via rank-minimization (matrix factorization)

SLIM

Conventional method

Rank minimization

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\text{rank}(\mathbf{X})}_{\textbf{number of singular values of } \mathbf{X}} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \le \epsilon$$

where	$\mathcal{A}$	acquisition mask
	$\mathbf{B}\in\mathbb{C}^{m\times n}$	observed data
	$\|.\|_F$	Frobenius norm

Rajiv Kumar, Curt Da Silva, Okan Akalin, Aleksandr Y. Aravkin, Hassan Mansour, Ben Recht, and Felix J. Herrmann, "Efficient matrix completion for seismic data reconstruction", Geophysics, vol. 80, pp. V97-V114, 2015

SLIM

Conventional method

Rank minimization

Hard to solve

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\text{rank}(\mathbf{X})}_{\textbf{number of singular values of } \mathbf{X}} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

where		
	$\mathcal{A}$	acquisition mask
	$\mathbf{B}\in\mathbb{C}^{m\times n}$	observed data
	$\|.\|_F$	Frobenius norm

Rajiv Kumar, Curt Da Silva, Okan Akalin, Aleksandr Y. Aravkin, Hassan Mansour, Ben Recht, and Felix J. Herrmann, “Efficient matrix completion for seismic data reconstruction”, Geophysics, vol. 80, pp. V97-V114, 2015

Conventional method

Low-rank matrix completion

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\|\mathbf{X}\|_*}_{\textbf{sum of singular values of X}} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{B}\|_F \leq \epsilon$$

Rajiv Kumar, Curt Da Silva, Okan Akalin, Aleksandr Y. Aravkin, Hassan Mansour, Ben Recht, and Felix J. Herrmann, "Efficient matrix completion for seismic data reconstruction", Geophysics, vol. 80, pp. V97-V114, 2015

SLIM

Conventional method

Low-rank matrix completion

Expensive for large scale

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \underbrace{\|\mathbf{X}\|_*}_{\textbf{sum of singular values of } \mathbf{X}} \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

Low-rank matrix factorization

$$\underset{\mathbf{L},\mathbf{R}}{\text{minimize}} \quad \frac{1}{2}\left\|\begin{bmatrix}\mathbf{L}\\ \mathbf{R}\end{bmatrix}\right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{L}\mathbf{R}^H)-\mathbf{B}\|_F \leq \epsilon$$

where

$$\mathbf{X}=\mathbf{L}\mathbf{R}^H$$
$$\mathbf{L}\in\mathbb{C}^{m\times r}$$
$$\mathbf{R}\in\mathbb{C}^{n\times r}$$
$$\mathbf{B}\in\mathbb{C}^{m\times n}$$

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Weighted method

Weighted matrix completion

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \|\mathbf{QXW}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \leq \epsilon$$

where

$$\mathbf{X}_{\text{prior}} \approx \mathbf{U\Sigma V}^H, \;\; \mathbf{U}\in\mathbb{C}^{m\times r}, \;\; \mathbf{V}\in\mathbb{C}^{n\times r}$$

$$\mathbf{Q} = w_1\mathbf{UU}^H + \mathbf{U}^\perp\mathbf{U}^{\perp H}$$

$$\mathbf{W} = w_2\mathbf{VV}^H + \mathbf{V}^\perp\mathbf{V}^{\perp H}$$

Scalars $w_1, w_2 \in (0,1]$ **are weights**

Zhang, Y. et al. "High-frequency wavefield recovery with weighted matrix factorizations." In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. "Weighted matrix completion and recovery with prior subspace information." IEEE Transactions on Information Theory, 2018.

SLIM

Weighted method

Weighted matrix completion

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \|\mathbf{QXW}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{B}\|_F \leq \epsilon$$

where

$$\mathbf{X}_{\text{prior}} \approx \mathbf{U\Sigma V}^H, \;\; \mathbf{U} \in \mathbb{C}^{m\times r}, \;\; \mathbf{V} \in \mathbb{C}^{n\times r}$$

$$\mathbf{Q} = w_1\mathbf{UU}^H + \mathbf{U}^\perp\mathbf{U}^{\perp H}$$

$$\mathbf{W} = w_2\mathbf{VV}^H + \mathbf{V}^\perp\mathbf{V}^{\perp H}$$

Scalars $w_1, w_2 \in (0, 1]$ **are weights**

Weighted matrix factorization

$$\underset{\mathbf{L},\mathbf{R}}{\text{minimize}} \quad \frac{1}{2}\left\|\begin{bmatrix}\mathbf{QL}\\ \mathbf{WR}\end{bmatrix}\right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{LR}^H) - \mathbf{B}\|_F \leq \epsilon$$

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Weighted method

Computation expensive

Higher rank increases potential overfitting

Weighted matrix completion

$$\underset{\mathbf{X}\in\mathbb{C}^{m\times n}}{\text{minimize}} \quad \|\mathbf{QXW}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_F \le \epsilon$$

where

$$\mathbf{X}_{\text{prior}} \approx \mathbf{U\Sigma V}^H, \ \mathbf{U}\in\mathbb{C}^{m\times r}, \ \mathbf{V}\in\mathbb{C}^{n\times r}$$

$$\mathbf{Q} = w_1\mathbf{UU}^H + \mathbf{U}^\perp\mathbf{U}^{\perp H}$$

$$\mathbf{W} = w_2\mathbf{VV}^H + \mathbf{V}^\perp\mathbf{V}^{\perp H}$$

Scalars $w_1, w_2 \in (0,1]$ are weights

Weighted matrix factorization

$$\underset{\mathbf{L},\mathbf{R}}{\text{minimize}} \quad \frac{1}{2}\left\| \begin{bmatrix}\mathbf{QL}\\ \mathbf{WR}\end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{LR}^H)-\mathbf{B}\|_F \le \epsilon$$

Smaller weights correspond to more confidence on prior information (similarity)

Weighted method (Efficient)

SLIM

Weighted matrix completion

$$\underset{\bar{\mathbf{X}} \in \mathbb{C}^{m \times n}}{\text{minimize}} \quad \|\bar{\mathbf{X}}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q^{-1}}\bar{\mathbf{X}}\mathbf{W^{-1}}) - \mathbf{B}\|_F \leq \epsilon$$

where $\bar{\mathbf{X}} = \mathbf{QXW}$

$$\mathbf{Q^{-1}} = \frac{1}{w_1}\mathbf{U}\mathbf{U}^H + \mathbf{U}^{\perp}\mathbf{U}^{\perp H}$$

$$\mathbf{W^{-1}} = \frac{1}{w_2}\mathbf{V}\mathbf{V}^H + \mathbf{V}^{\perp}\mathbf{V}^{\perp H}$$

Weighted matrix factorization

$$\underset{\bar{\mathbf{L}}, \bar{\mathbf{R}}}{\text{minimize}} \quad \frac{1}{2}\left\| \begin{bmatrix} \bar{\mathbf{L}} \\ \bar{\mathbf{R}} \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q^{-1}}\bar{\mathbf{L}}\bar{\mathbf{R}}^{\mathbf{H}}\mathbf{W^{-1}}) - \mathbf{B}\|_F \leq \epsilon$$

where $\bar{\mathbf{X}} = \bar{\mathbf{L}}\bar{\mathbf{R}}^{\mathbf{H}}$

$$\mathbf{X} = \mathbf{Q^{-1}}\bar{\mathbf{X}}\mathbf{W^{-1}}$$

Weighted method (Efficient)

SLIM

Weighted matrix completion

$$\underset{\bar{\mathbf{X}} \in \mathbb{C}^{m \times n}}{\text{minimize}} \quad \|\bar{\mathbf{X}}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

where $\bar{\mathbf{X}} = \mathbf{QXW}$

$$\mathbf{Q}^{-1} = \frac{1}{w_1}\mathbf{U}\mathbf{U}^H + \mathbf{U}^{\perp}\mathbf{U}^{\perp H}$$

$$\mathbf{W}^{-1} = \frac{1}{w_2}\mathbf{V}\mathbf{V}^H + \mathbf{V}^{\perp}\mathbf{V}^{\perp H}$$

Weighted matrix factorization

$$\underset{\bar{\mathbf{L}}, \bar{\mathbf{R}}}{\text{minimize}} \quad \frac{1}{2}\left\| \begin{bmatrix} \bar{\mathbf{L}} \\ \bar{\mathbf{R}} \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{L}}\bar{\mathbf{R}}^{\mathbf{H}}\mathbf{W}^{-1}) - \mathbf{B}\|_F \leq \epsilon$$

where $\bar{\mathbf{X}} = \bar{\mathbf{L}}\bar{\mathbf{R}}^{\mathbf{H}}$

$$\mathbf{X} = \mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}$$

20-25 times faster!

Limited-subspace weighted method

Limited-subspace weighted matrix factorization

$$\underset{\bar{\mathbf{L}},\bar{\mathbf{R}}}{\text{minimize}} \quad \frac{1}{2}\left\|\begin{bmatrix}\bar{\mathbf{L}}\\ \bar{\mathbf{R}}\end{bmatrix}\right\|_F^2 \quad \text{subject to} \quad \left\|\mathcal{A}(\tilde{\mathbf{Q}}^{-1}\bar{\mathbf{L}}\bar{\mathbf{R}}^{\mathbf{H}}\tilde{\mathbf{W}}^{-1}) - \mathbf{B}\right\|_F \leq \epsilon$$

where

$$\tilde{\mathbf{Q}}^{-1} = \frac{1}{w_1}\tilde{\mathbf{U}}\tilde{\mathbf{U}}^H + \tilde{\mathbf{U}}^{\perp}\tilde{\mathbf{U}}^{\perp H}$$

$$\tilde{\mathbf{W}}^{-1} = \frac{1}{w_2}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H + \tilde{\mathbf{V}}^{\perp}\tilde{\mathbf{V}}^{\perp H}$$

$$\tilde{\mathbf{U}} \in \mathbb{C}^{m\times r_s}, \quad \tilde{\mathbf{V}} \in \mathbb{C}^{n\times r_s}$$

$$\mathbf{U} \in \mathbb{C}^{m\times r}, \quad \mathbf{V} \in \mathbb{C}^{n\times r}$$

$r_s < r$, r is the rank used to recovery, r_s is the rank for limited-subspace

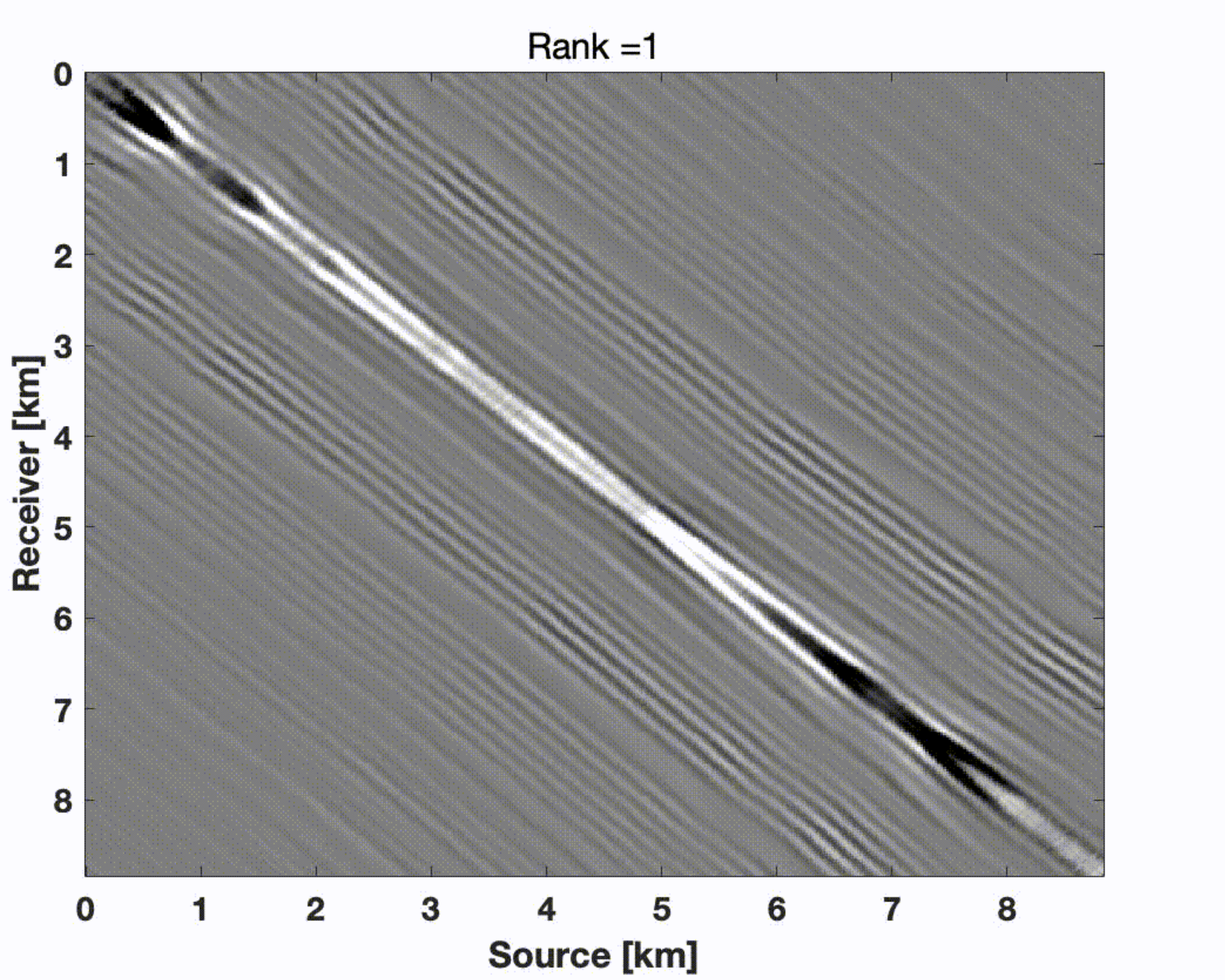
$$\tilde{\mathbf{U}} \subset \mathbf{U}, \quad \tilde{\mathbf{V}} \subset \mathbf{V}$$

Questions

1. Is the limited-subspace enough to represent the prior information?

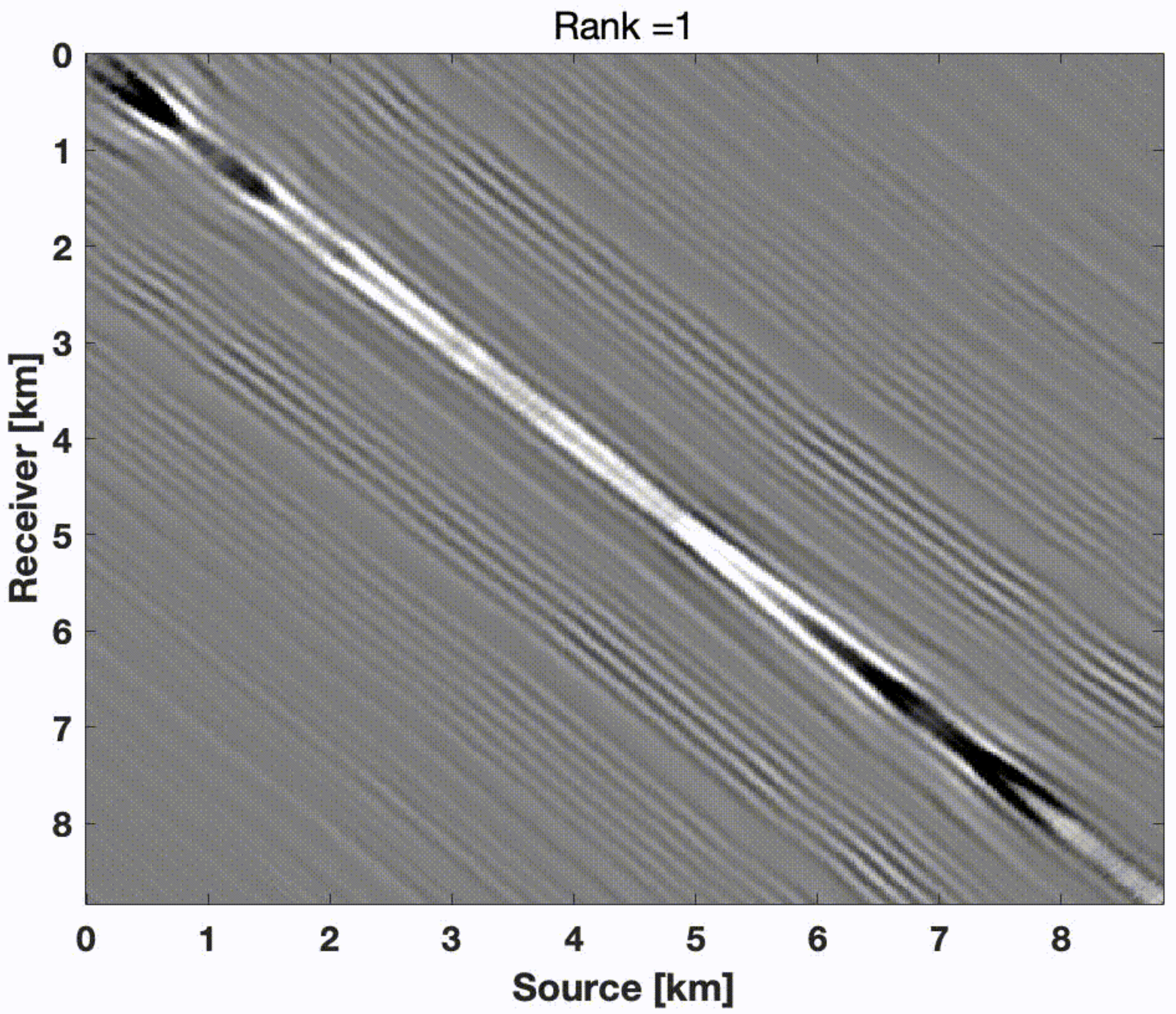
2. Does the limited-subspace break the similarity between the prior information and recovery?

Limited-subspace weighted method



- Higher rank includes lots of null space
- Prior information includes null space increases chances of overfitting
- Limited-subspace weighted method removes null space to prevent overfitting

Limited-subspace weighted method



- Higher rank includes lots of null space
- Prior information includes null space increases chances of overfitting
- Limited-subspace weighted method removes null space to prevent overfitting

Questions

1. Is the limited-subspace enough to represent the prior information?

2. Does the limited-subspace break the similarity between the prior information and recovery?

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

SLIM

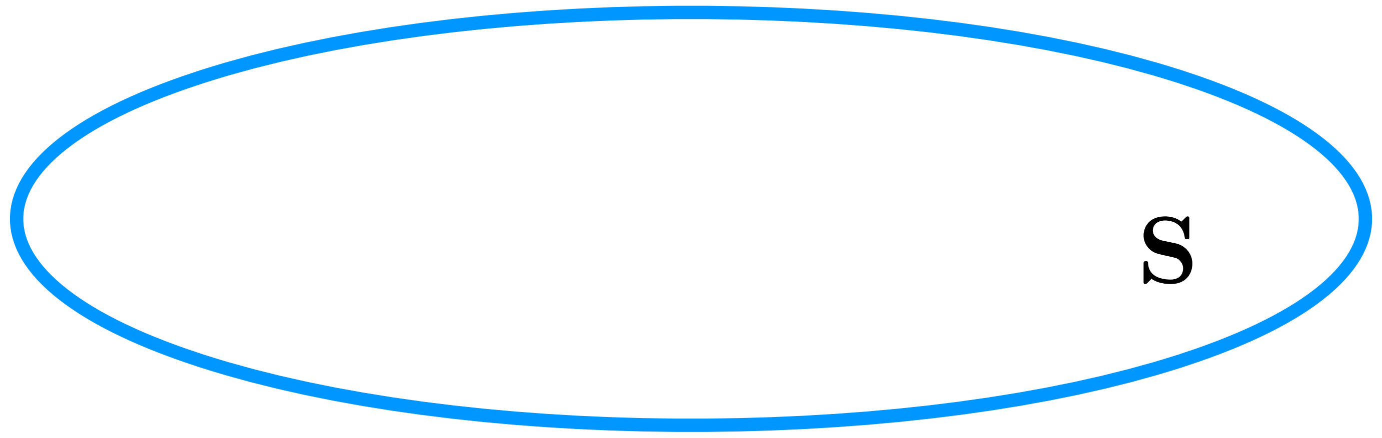
Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

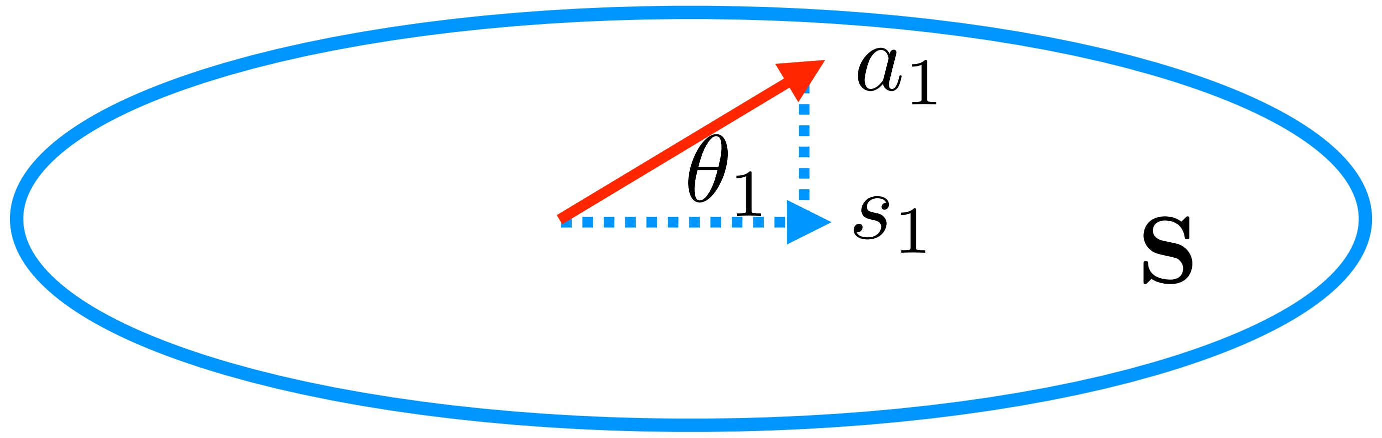
Similarity depends on largest angle between prior information and space of recovery



Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

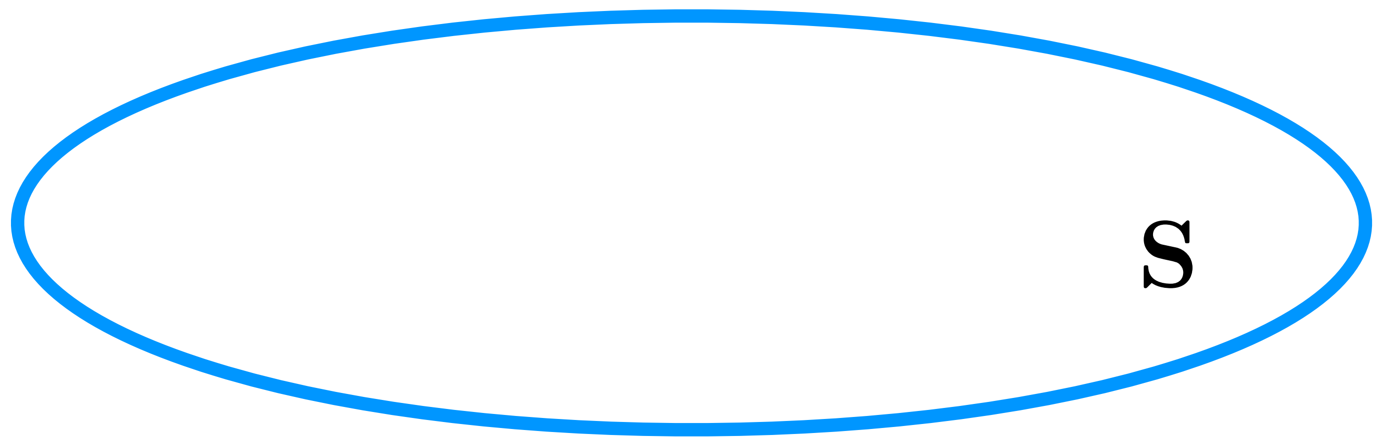
Similarity depends on largest angle between prior information and space of recovery



Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

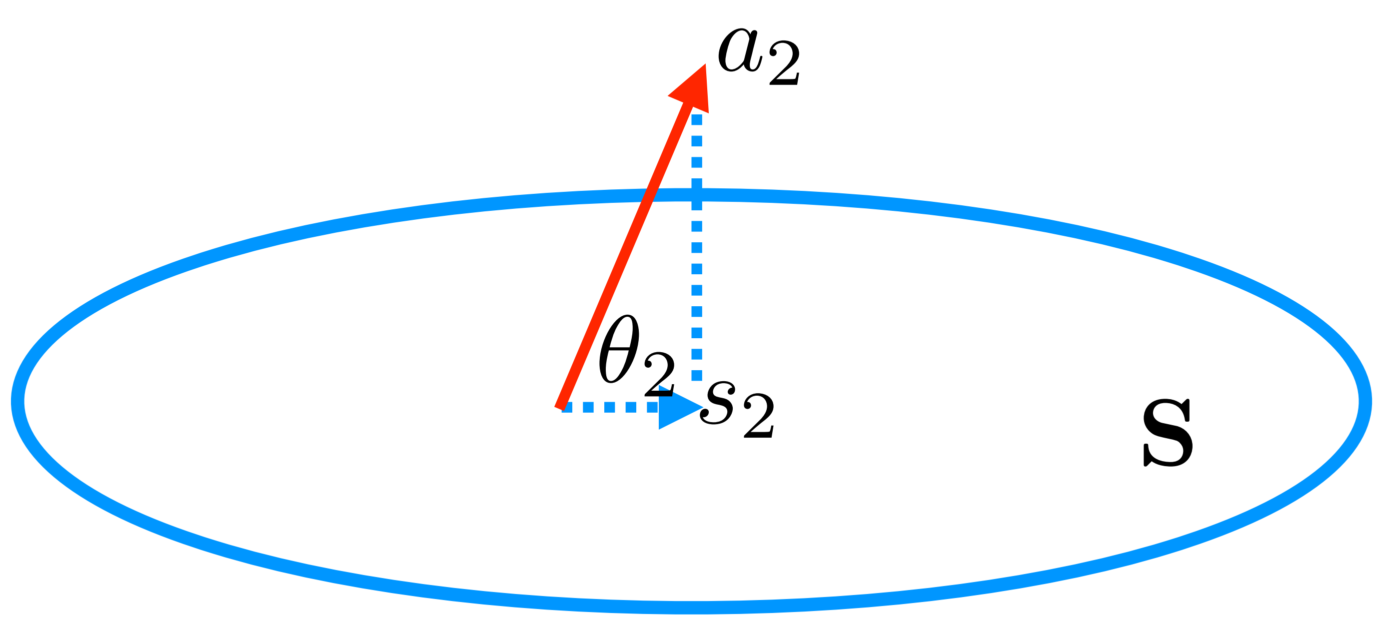
Similarity depends on largest angle between prior information and space of recovery



Zhang, Y. et al. "High-frequency wavefield recovery with weighted matrix factorizations." In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. "Weighted matrix completion and recovery with prior subspace information." IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

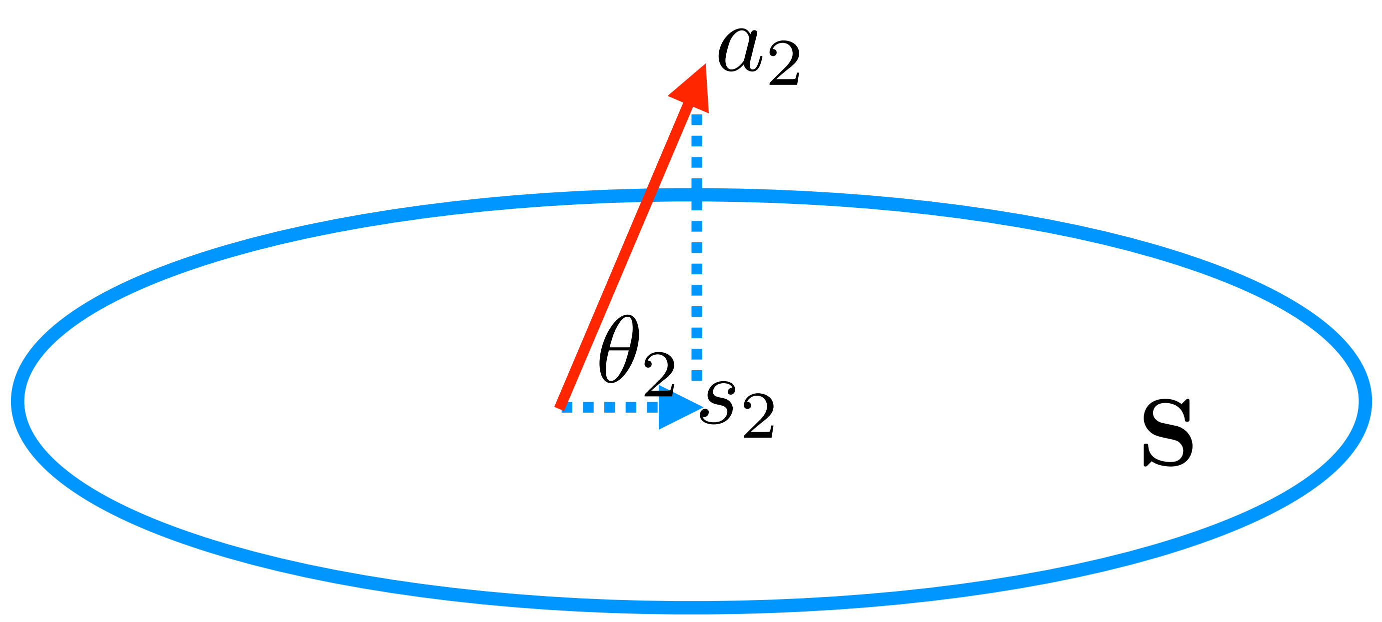
Similarity depends on largest angle between prior information and space of recovery



Zhang, Y. et al. "High-frequency wavefield recovery with weighted matrix factorizations." In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. "Weighted matrix completion and recovery with prior subspace information." IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery

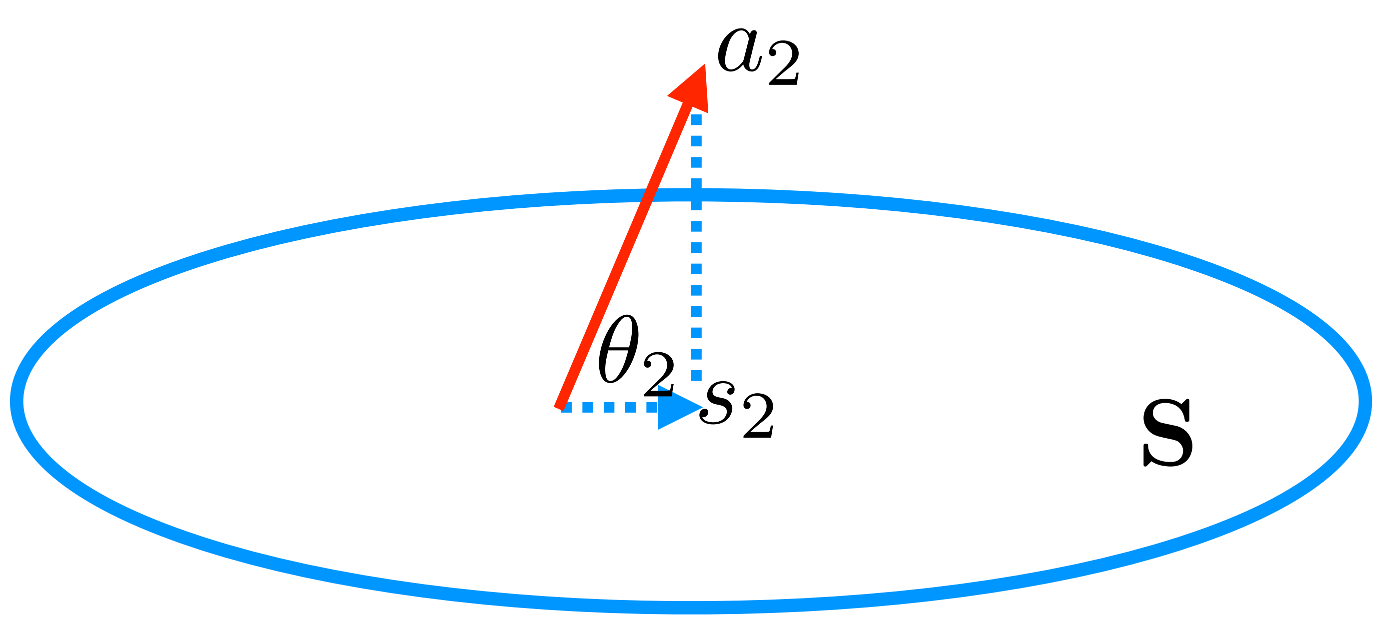


θ_2, a_2 indicates the similarity

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery



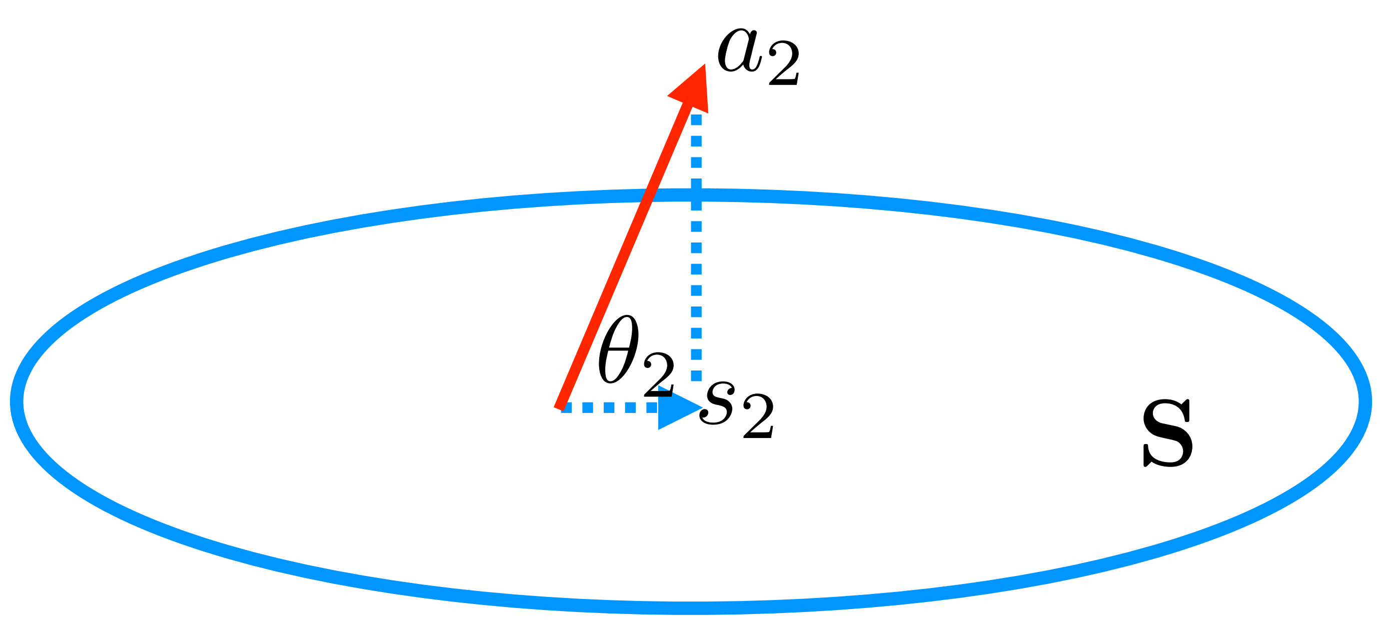
- does not affect the similarity between adjacent frequencies

Full-subspace of prior information

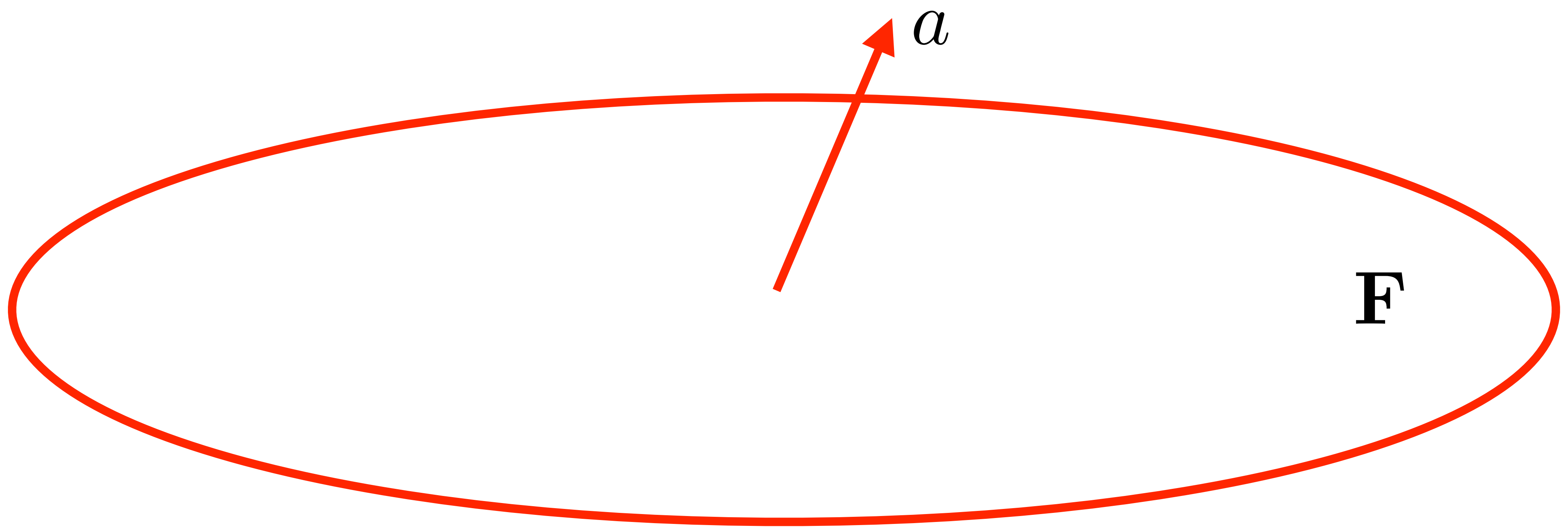
Zhang, Y. et al. "High-frequency wavefield recovery with weighted matrix factorizations." In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. "Weighted matrix completion and recovery with prior subspace information." IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery



θ, a **indicates the similarity**



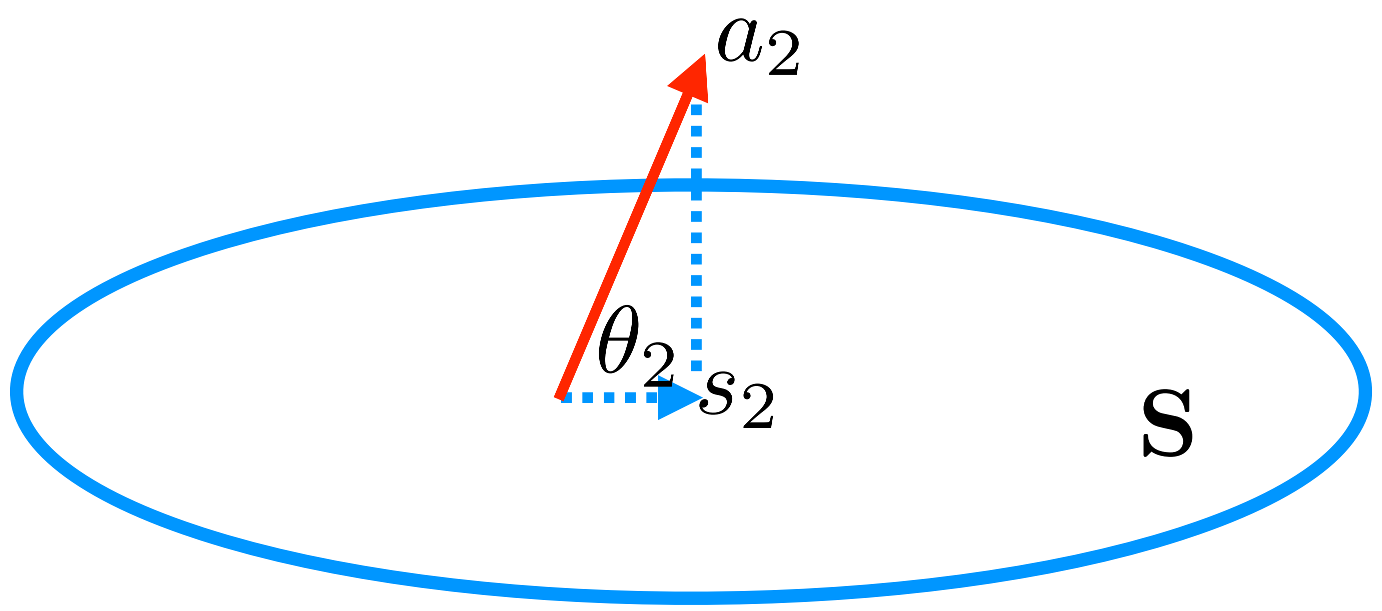
Full-subspace of prior information

- does not affect the similarity between adjacent frequencies

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

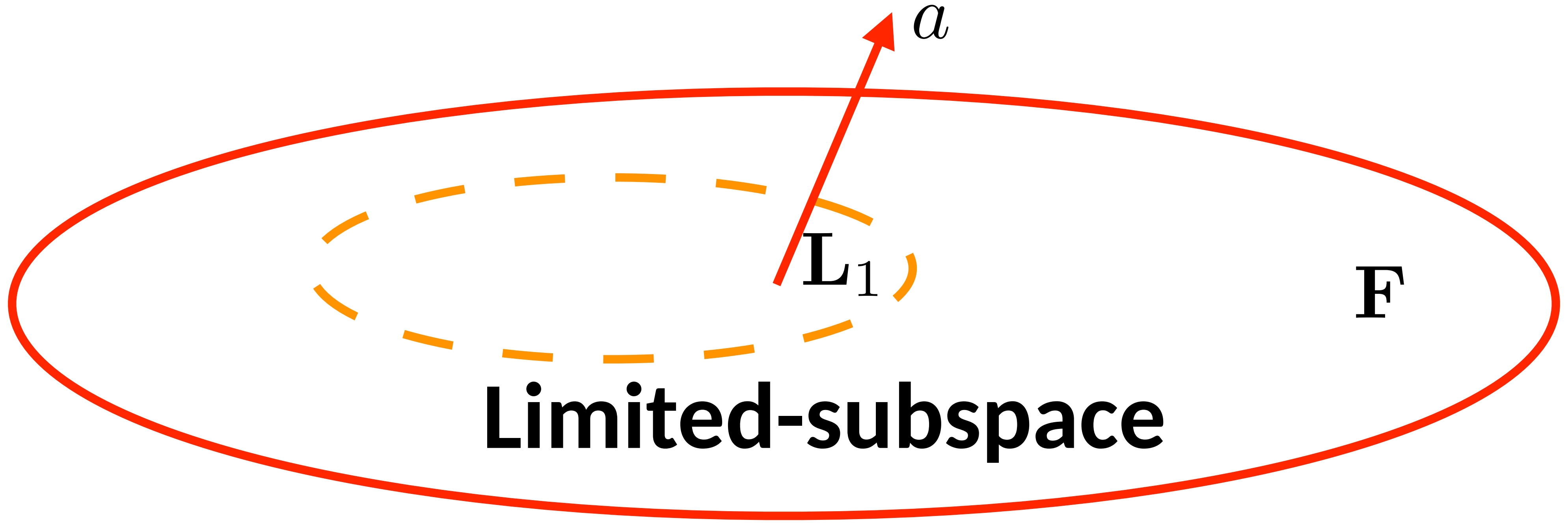
Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery



θ, a **indicates the similarity**

- does not affect the similarity between adjacent frequencies

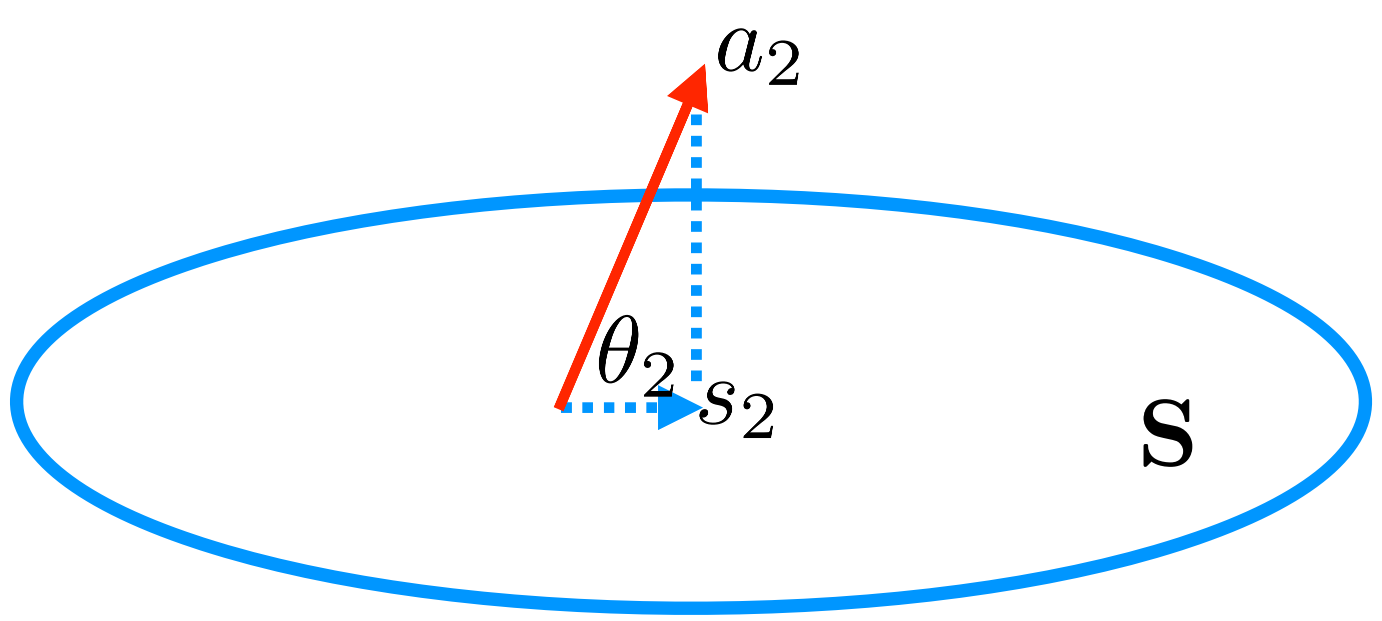


Full-subspace of prior information

Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

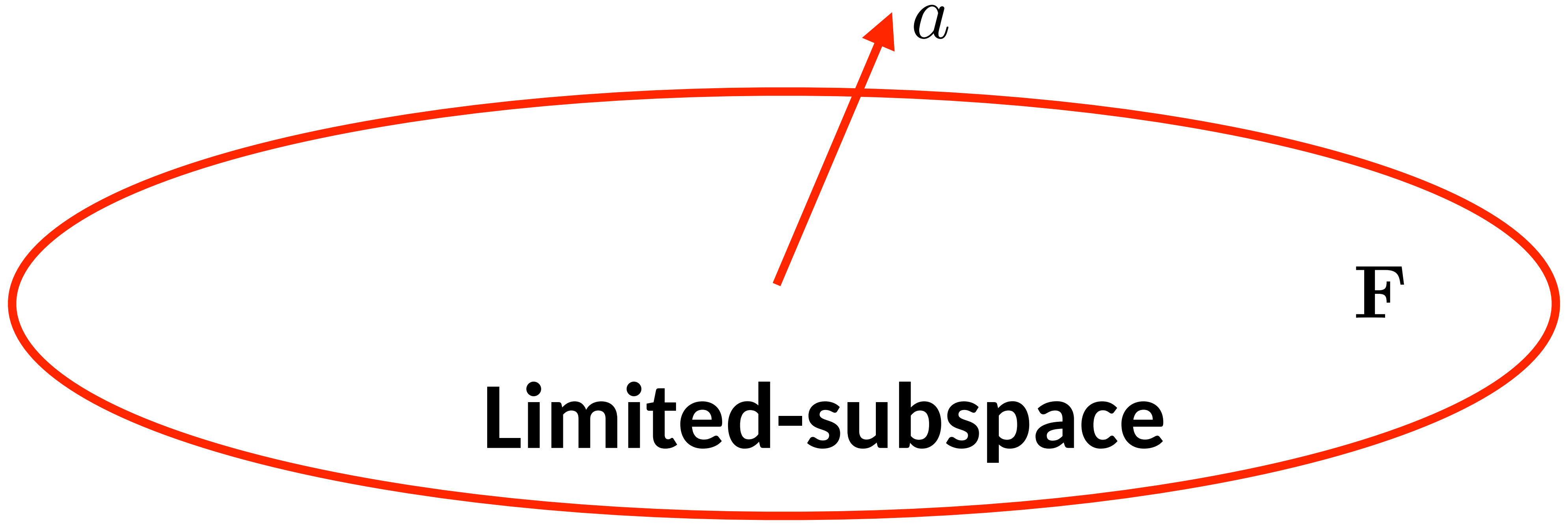
Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery



θ, a **indicates the similarity**

- does not affect the similarity between adjacent frequencies

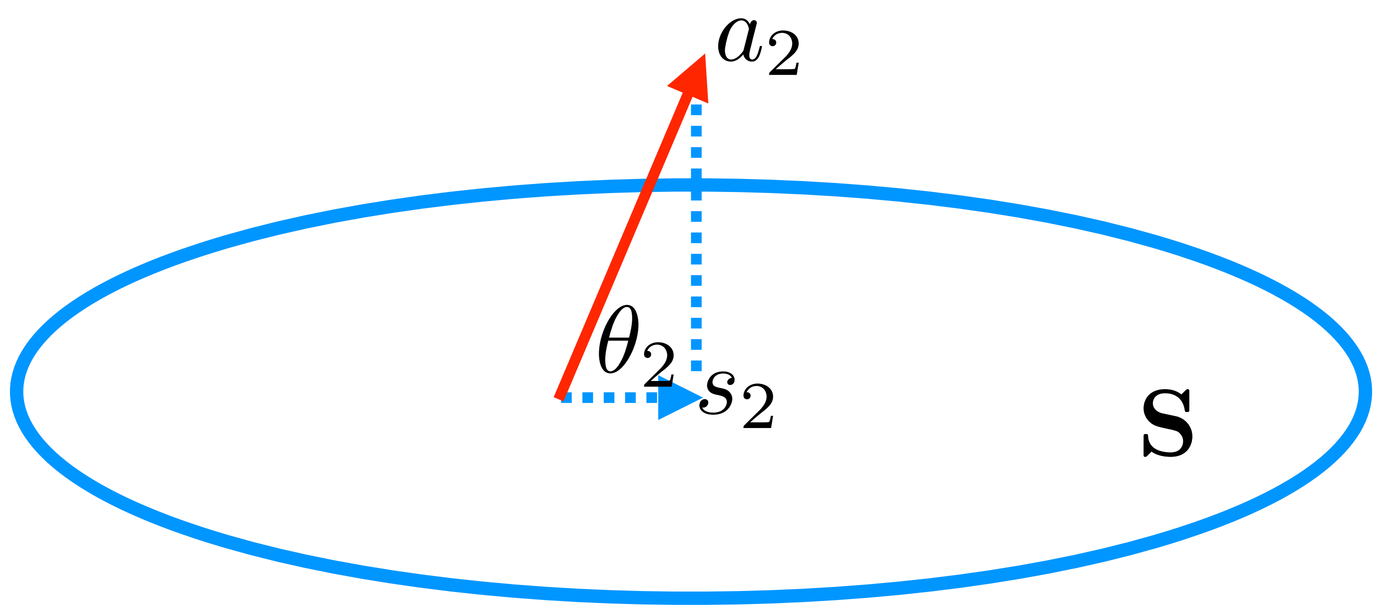


Full-subspace of prior information

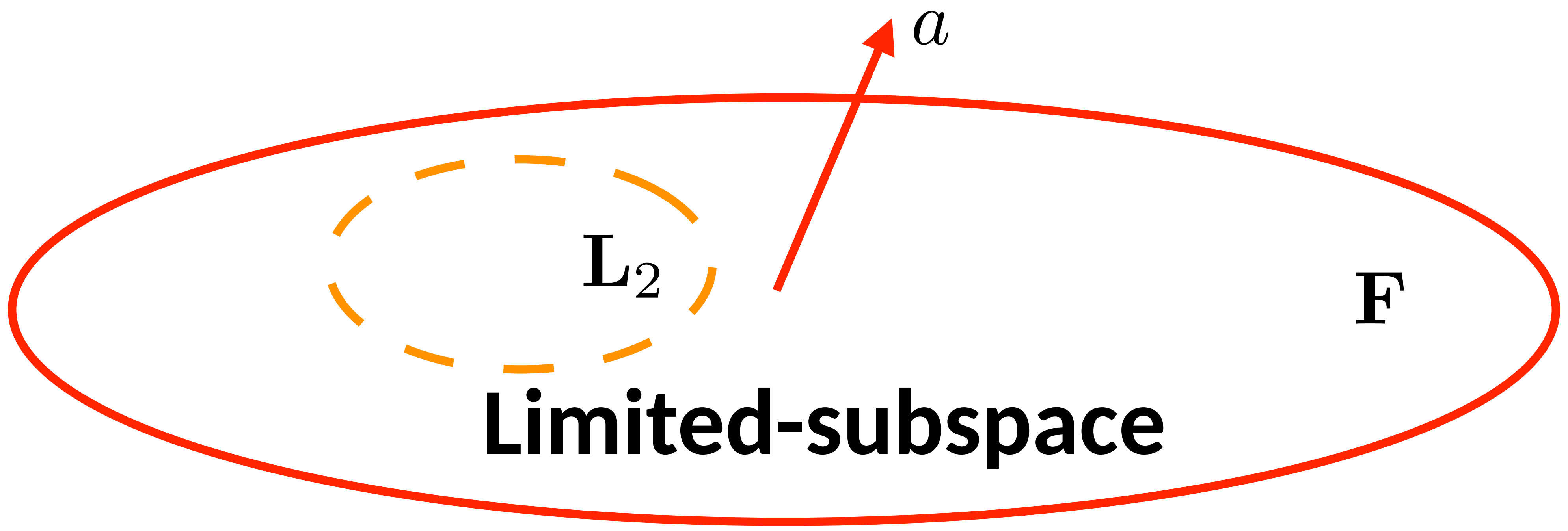
Zhang, Y. et al. “High-frequency wavefield recovery with weighted matrix factorizations.” In SEG Technical Program Expanded Abstracts, 2019.
Eftekhari, A. et al. “Weighted matrix completion and recovery with prior subspace information.” IEEE Transactions on Information Theory, 2018.

Limited-subspace weighted method

Similarity depends on largest angle between prior information and space of recovery



θ, a **indicates the similarity**



Full-subspace of prior information

- does not affect the similarity between adjacent frequencies

2D Field data example: Gulf of Suez

Data acquisition area: Gulf of Suez

Data dimension: 355 x 355 x 1024 (nr x ns x nt)

Dimension of each frequency slice: 355 x 355

Source sampling interval: 25 m

Receiver sampling interval: 25 m

Field data example: Gulf of Suez

Observed data: 75 % missing sources

Result shown for Frequency slice: 22 Hz

Optimization information

Number of iterations per frequency: 150

Scenarios compared

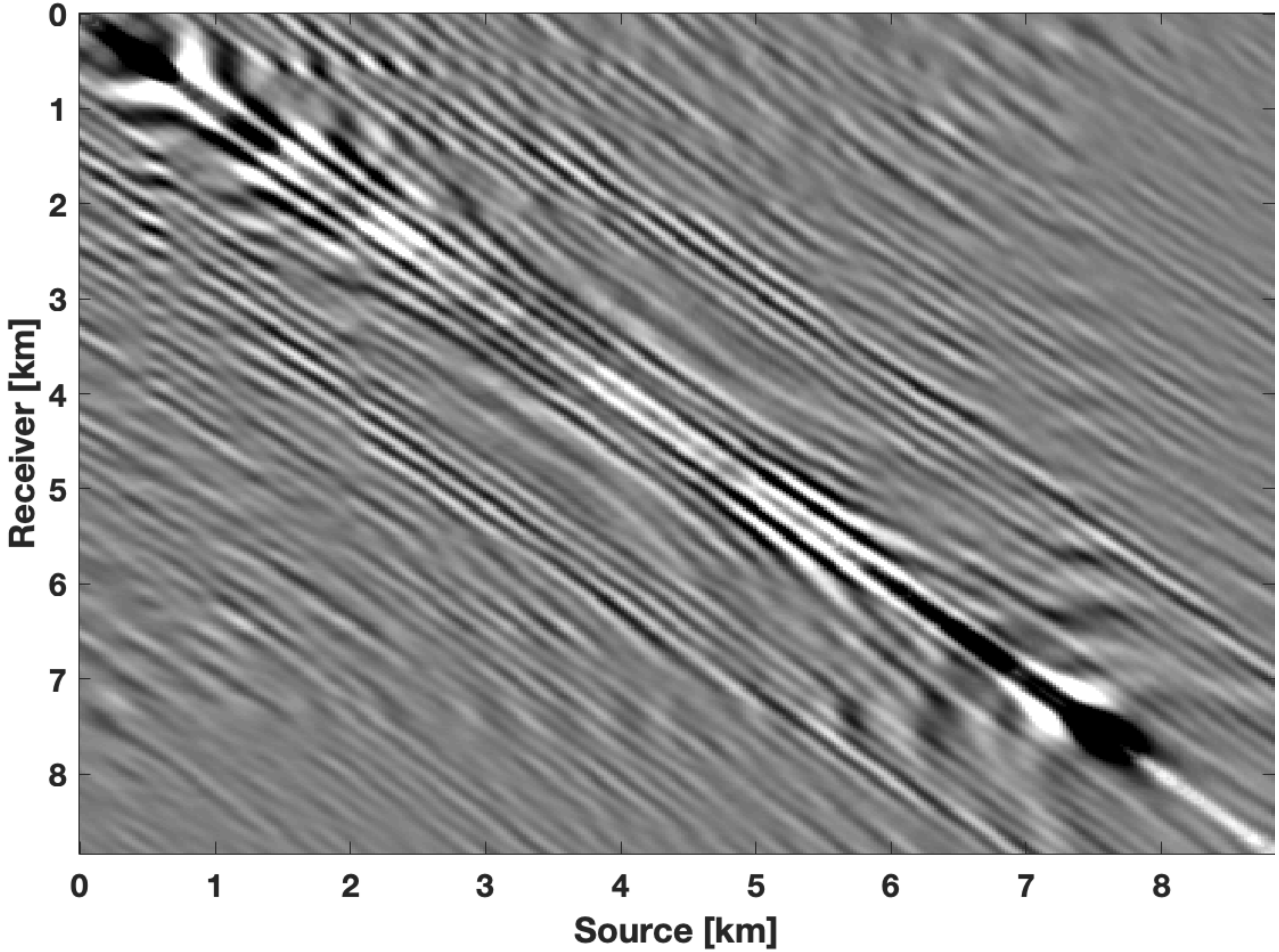
Scenario 1: using recursively weighted matrix completion (rank = 85, subspace rank = 85)

Scenario 2: using recursively weighted matrix completion (rank = 25, subspace rank = 25)

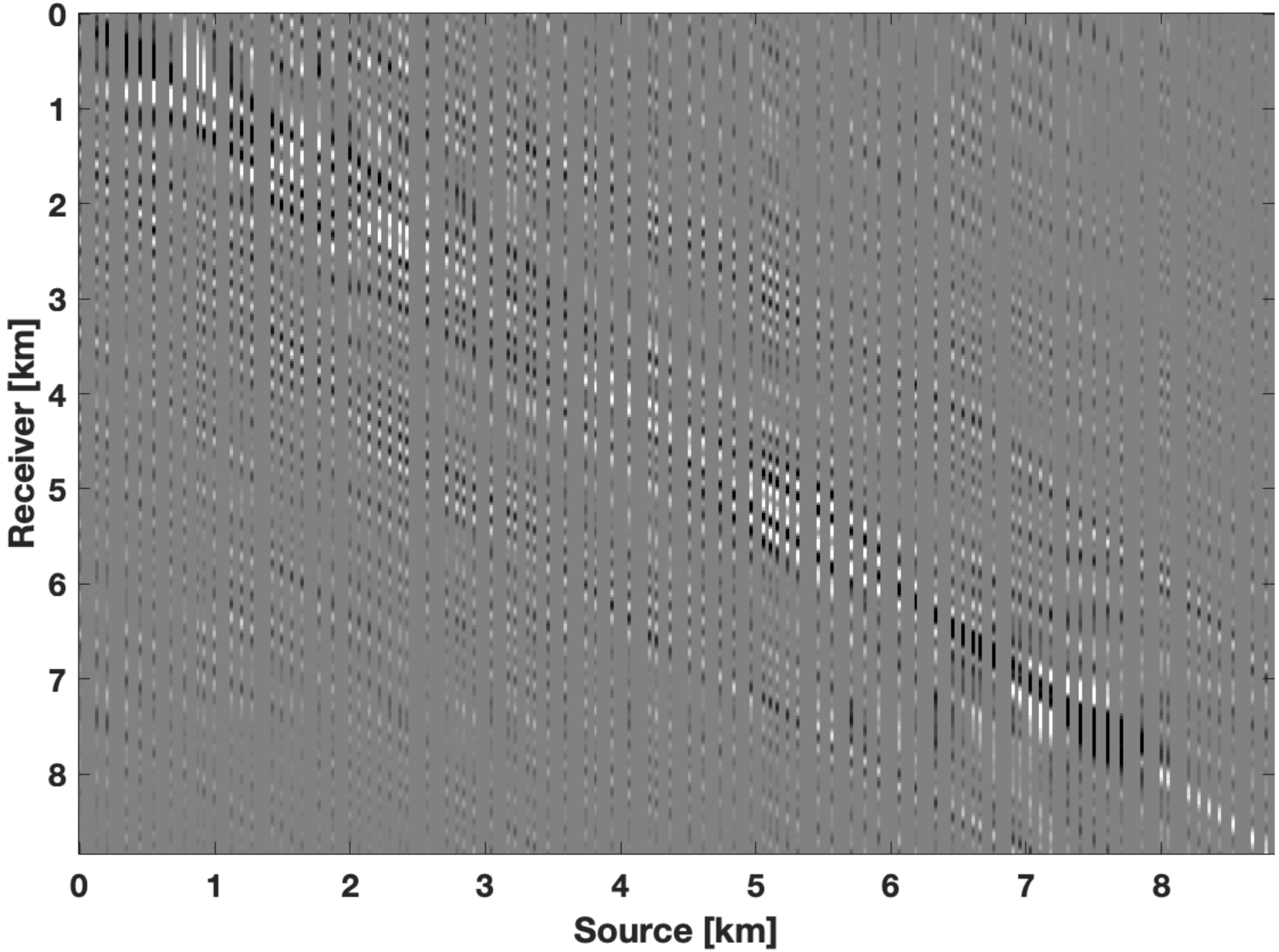
Scenario 3: using recursively limited-subspace weighted matrix completion (rank = 85, subspace rank = 25)

Frequency slice: 22 Hz

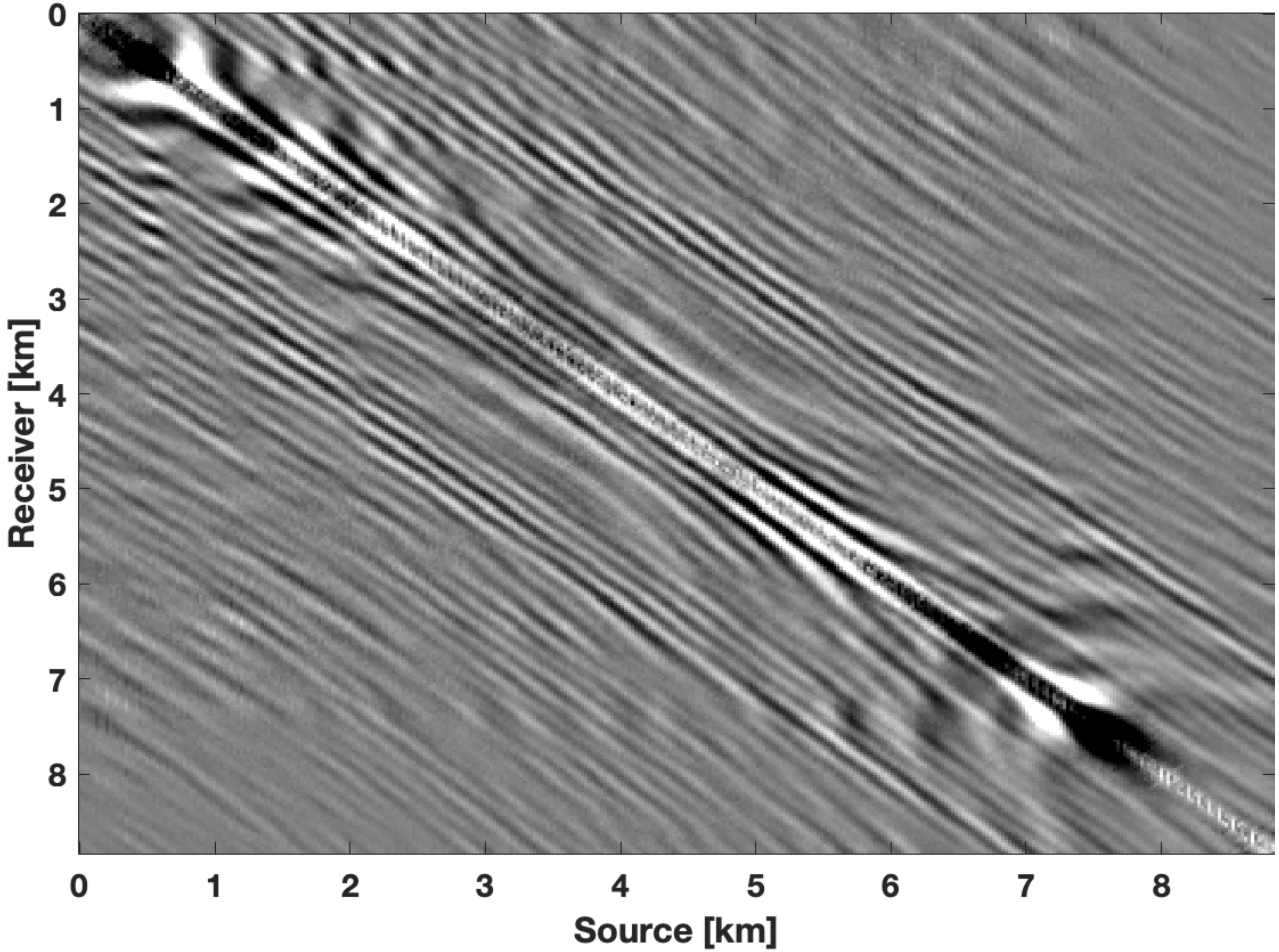
Fully sampled data



Observed data (75% jittered subsampled)

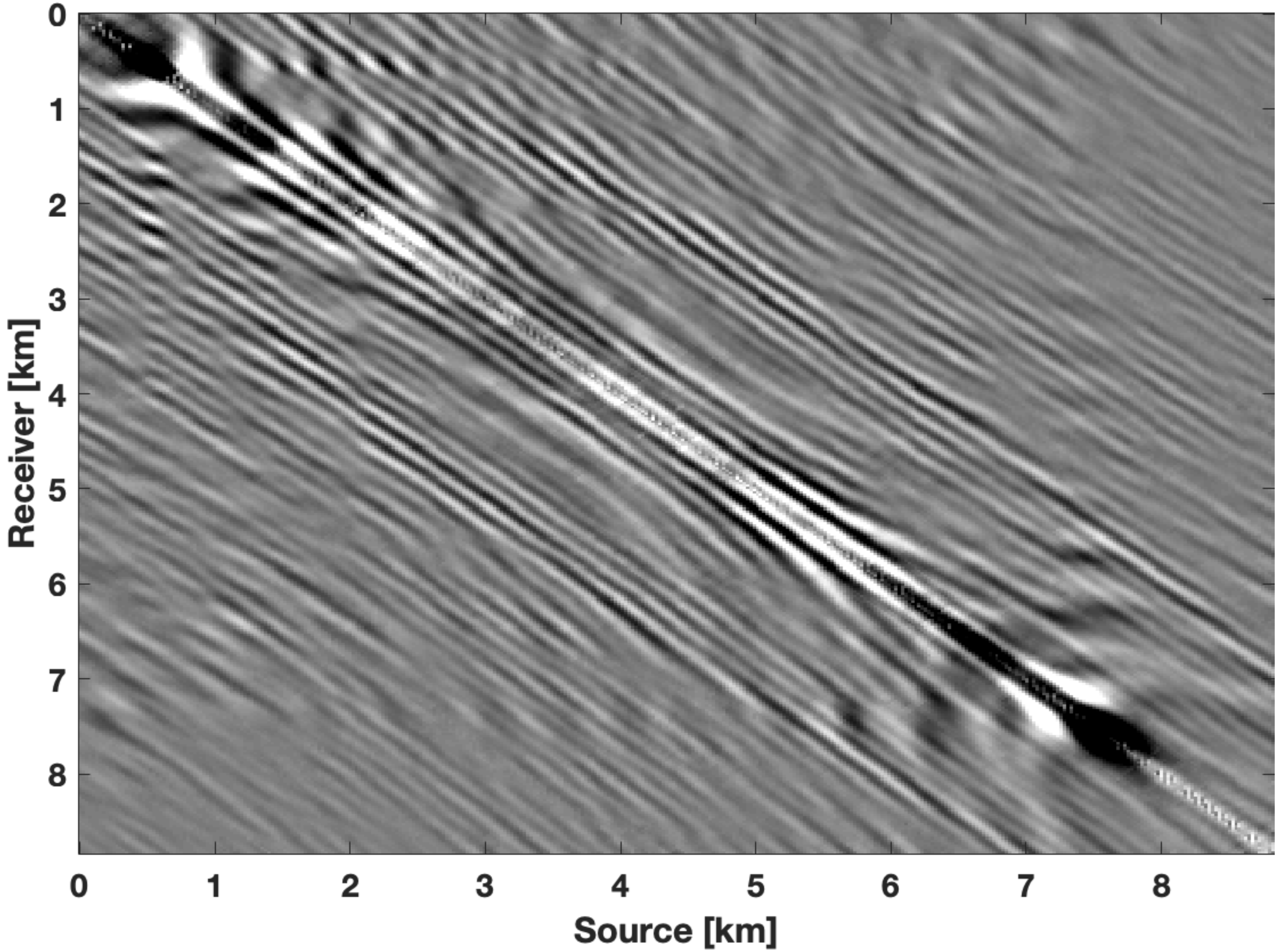


Recovery w/ weighted (rank = 85)



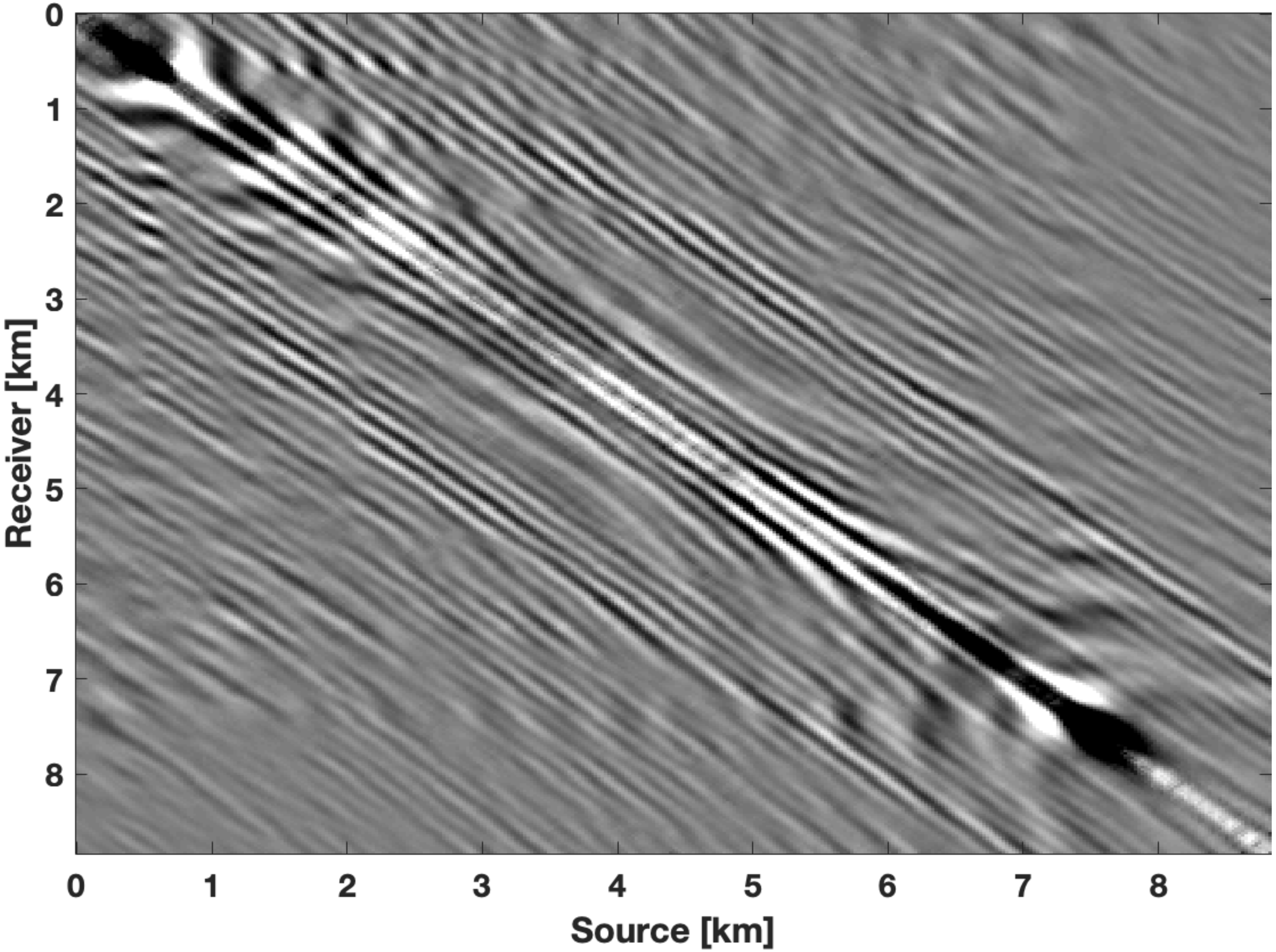
SNR = 13.09 dB
Rank = 85
Subspace rank = 85

Recovery w/ weighted (rank = 25)



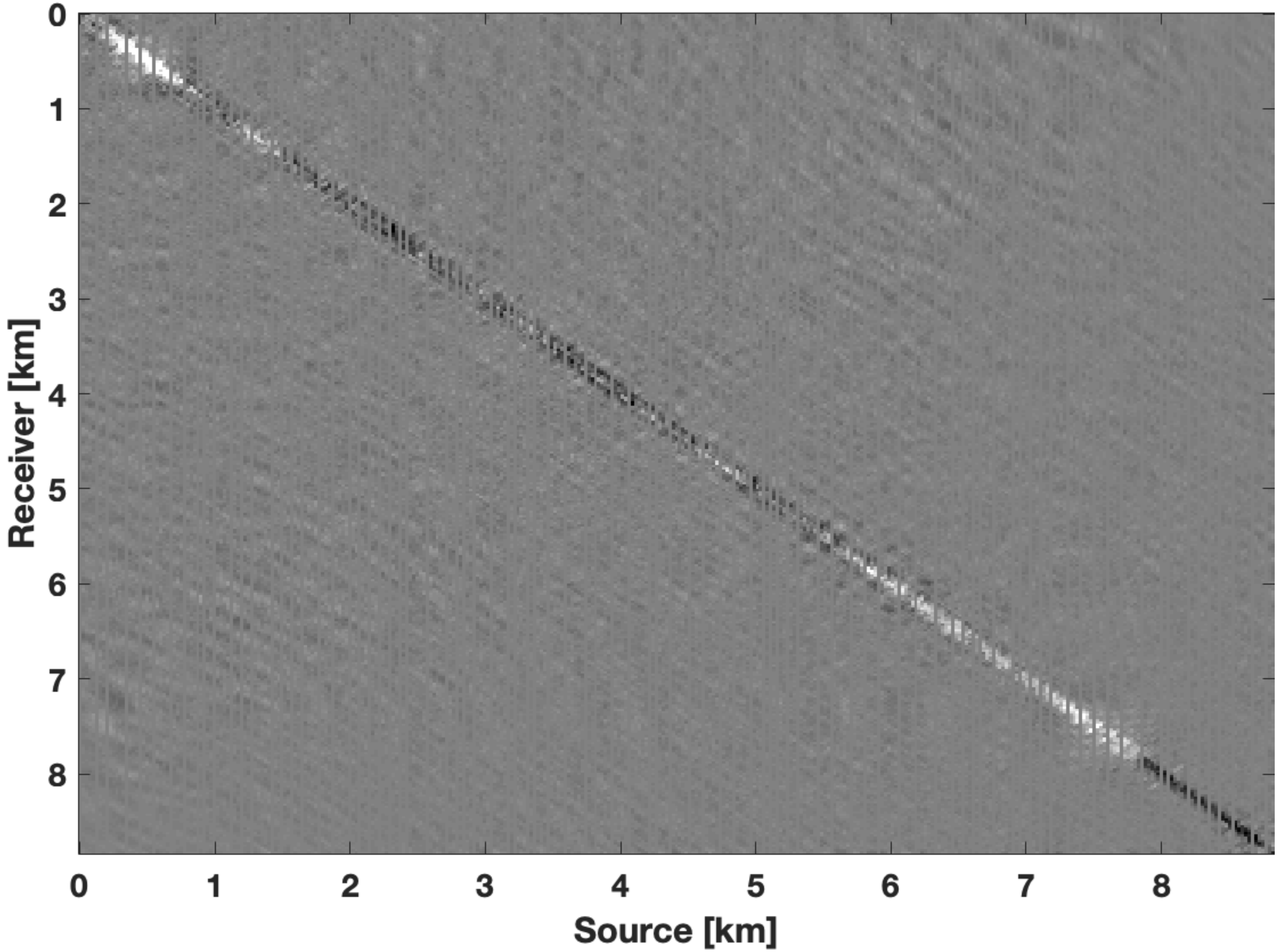
SNR = 15.50 dB
Rank = 25
Subspace rank = 25

Recovery w/ limited subspaces weighted



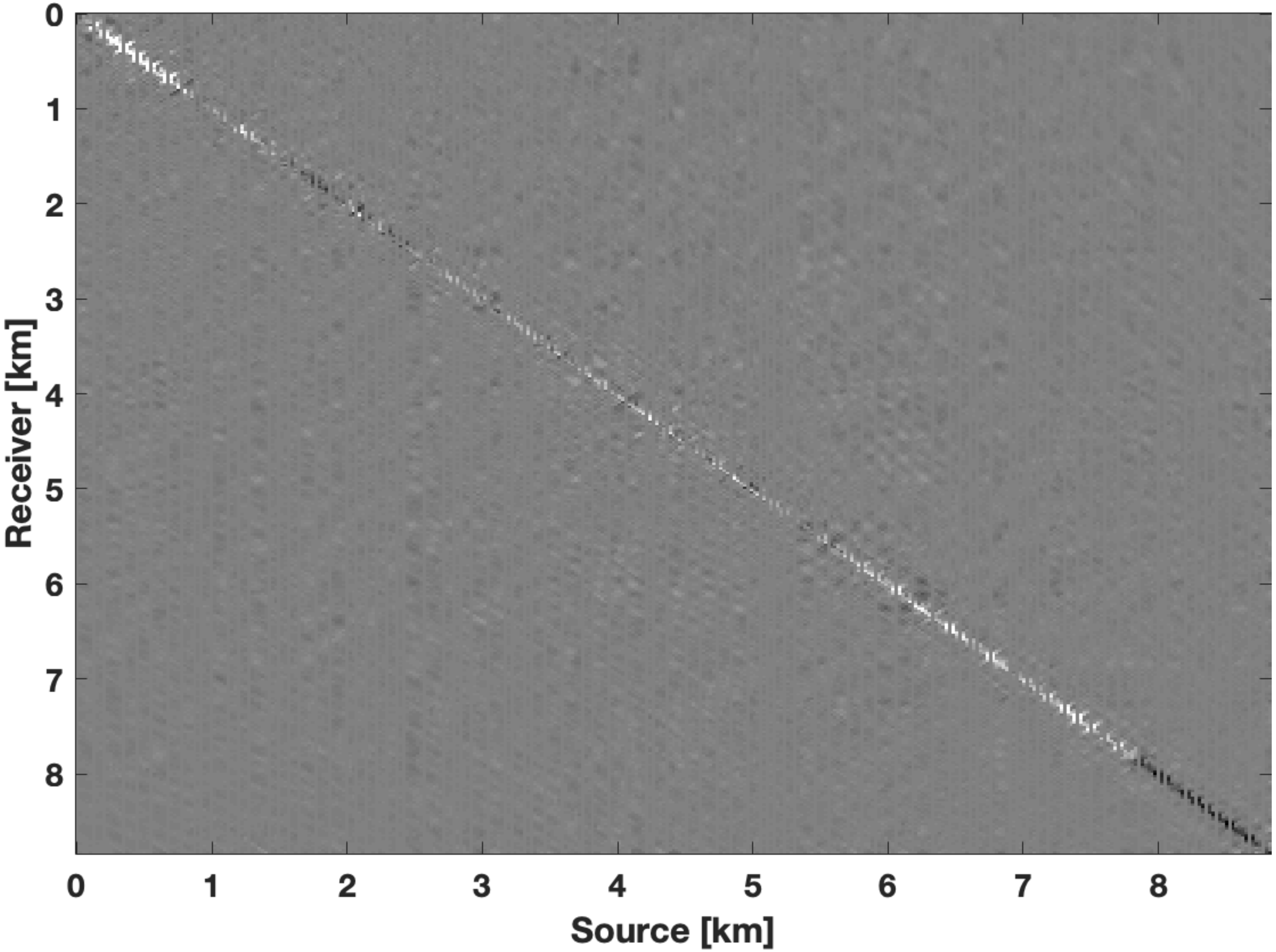
SNR = 19.52 dB
Rank = 85
Subspace rank = 25

Difference: True - Recovery w/ weighted (rank = 85)



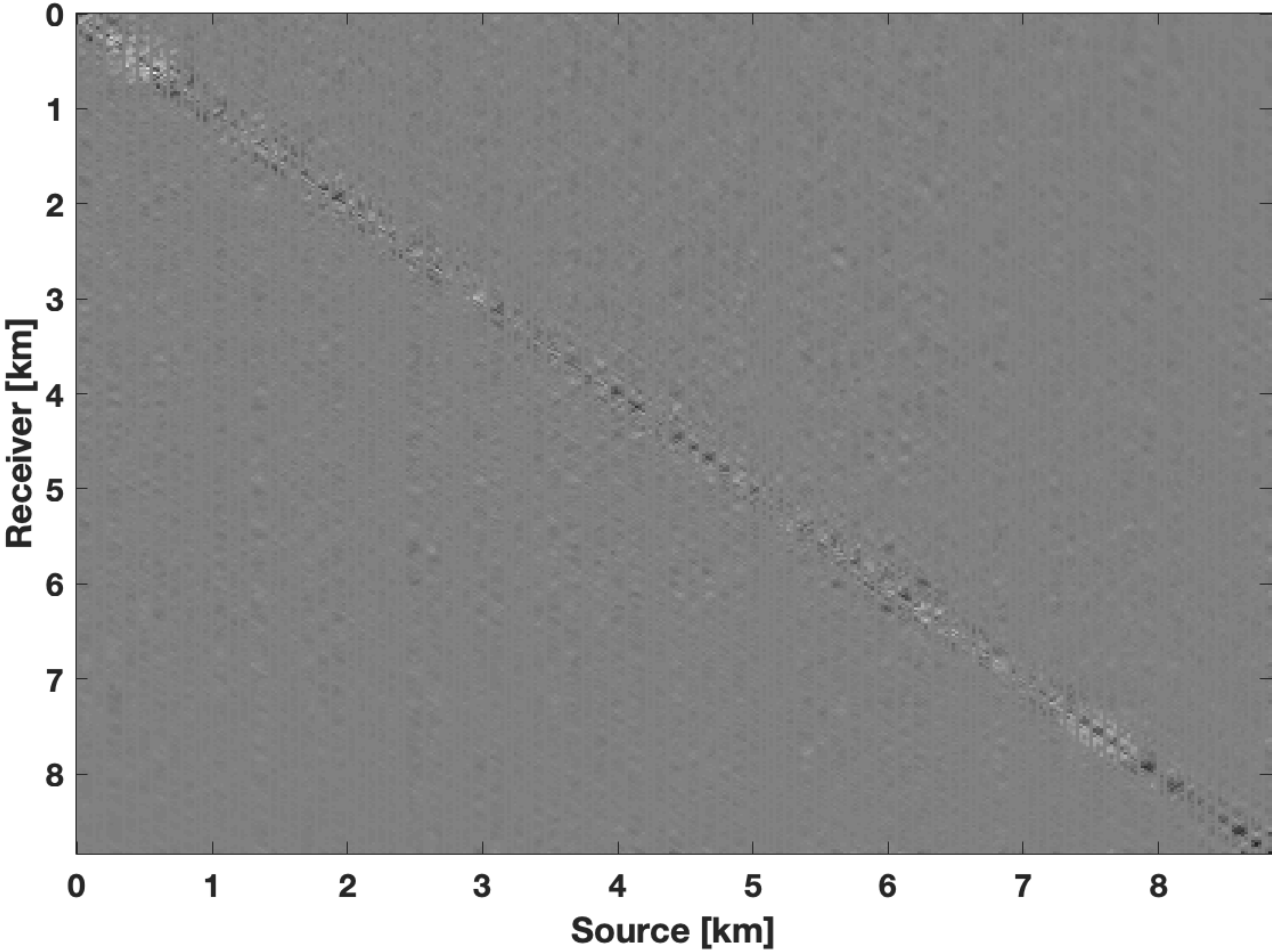
SNR = 13.09 dB
Rank = 85
Subspace rank = 85

Difference: True - Recovery w/ weighted (rank = 25)



SNR = 15.50 dB
Rank = 25
Subspace rank = 25

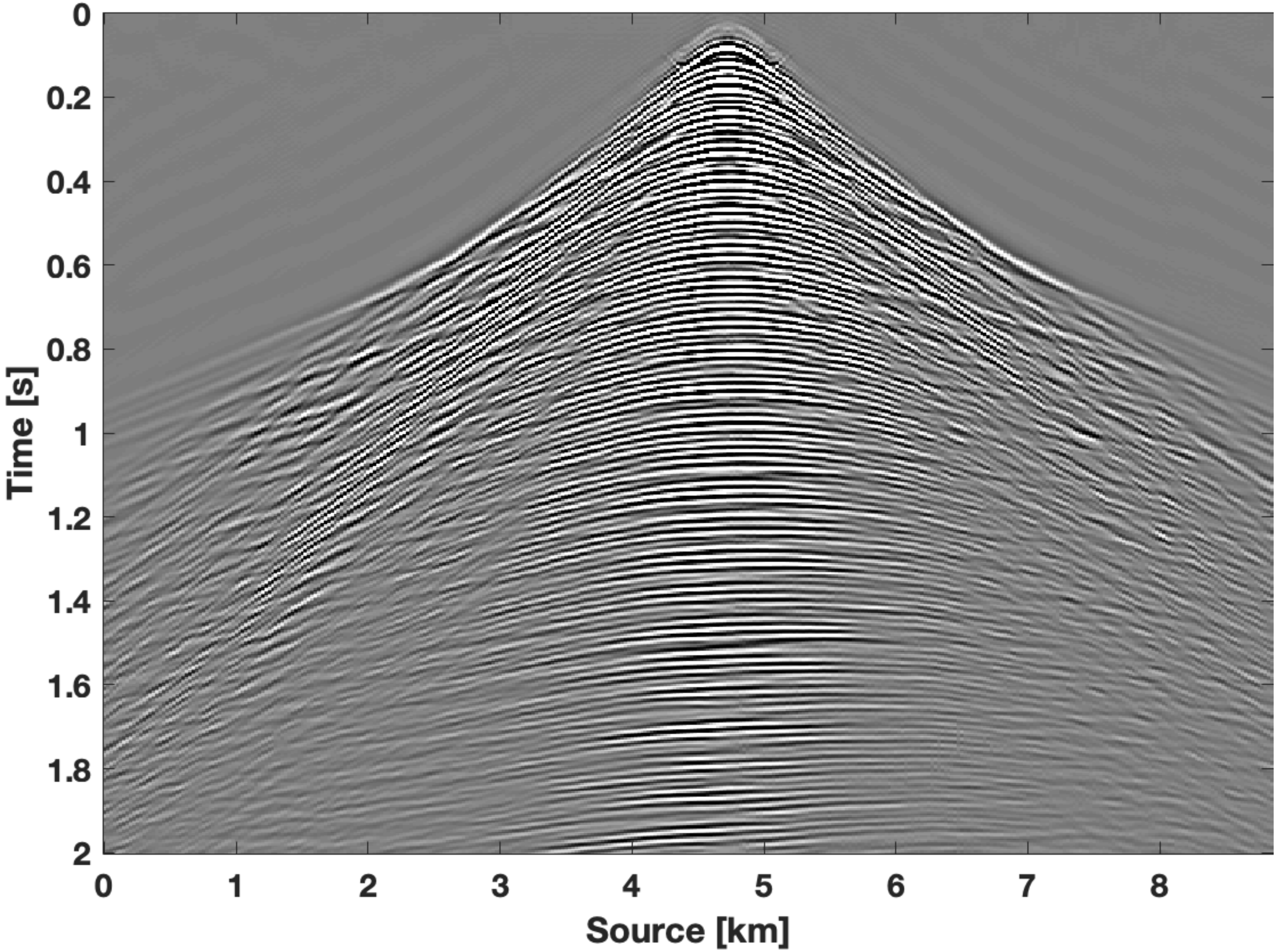
Difference: True - Recovery w/ limited subspaces weighted



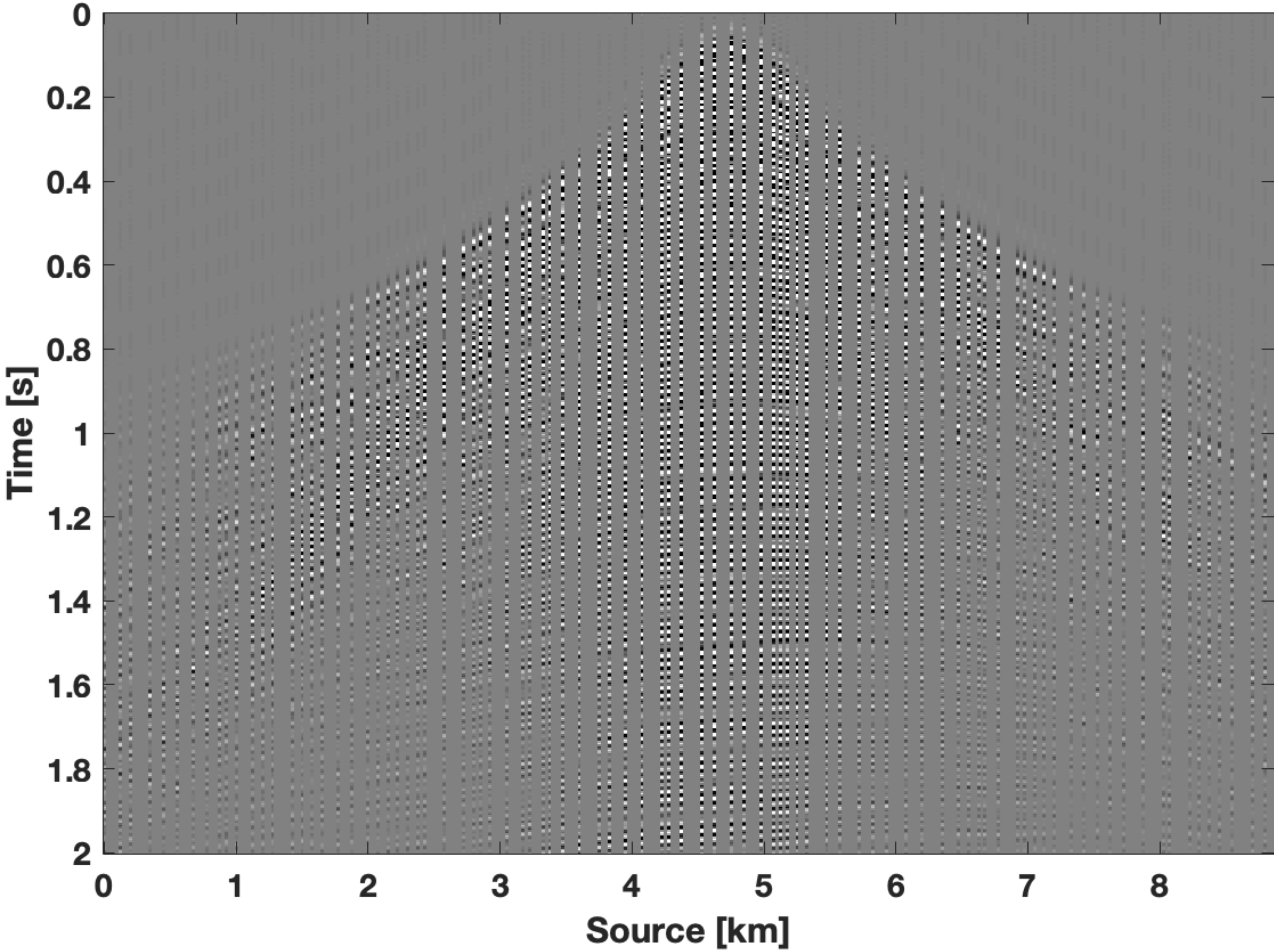
SNR = 19.52 dB
Rank = 85
Subspace rank = 25

Common receiver gather

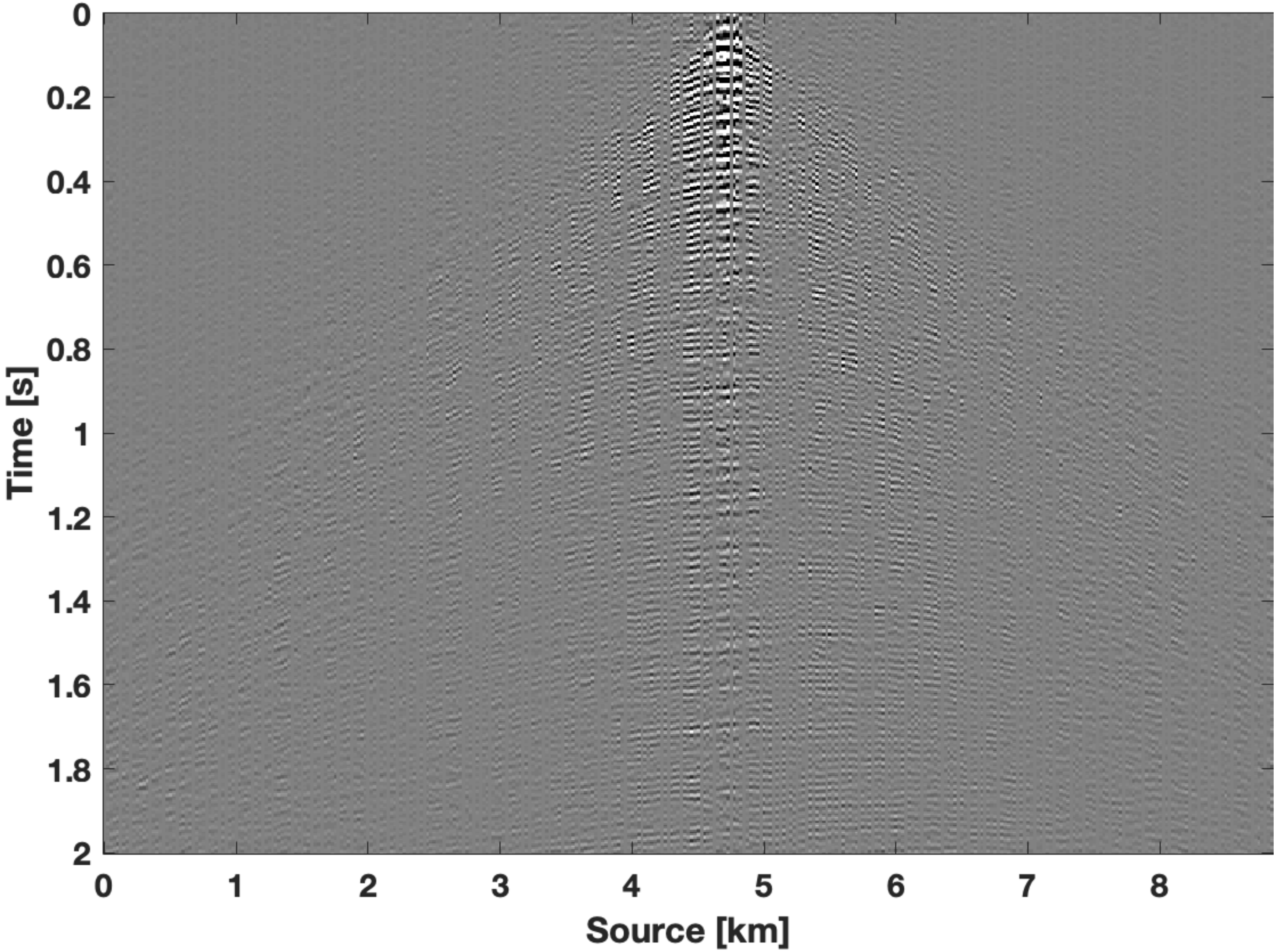
Fully sampled data



Observed data (75% jittered subsampled)

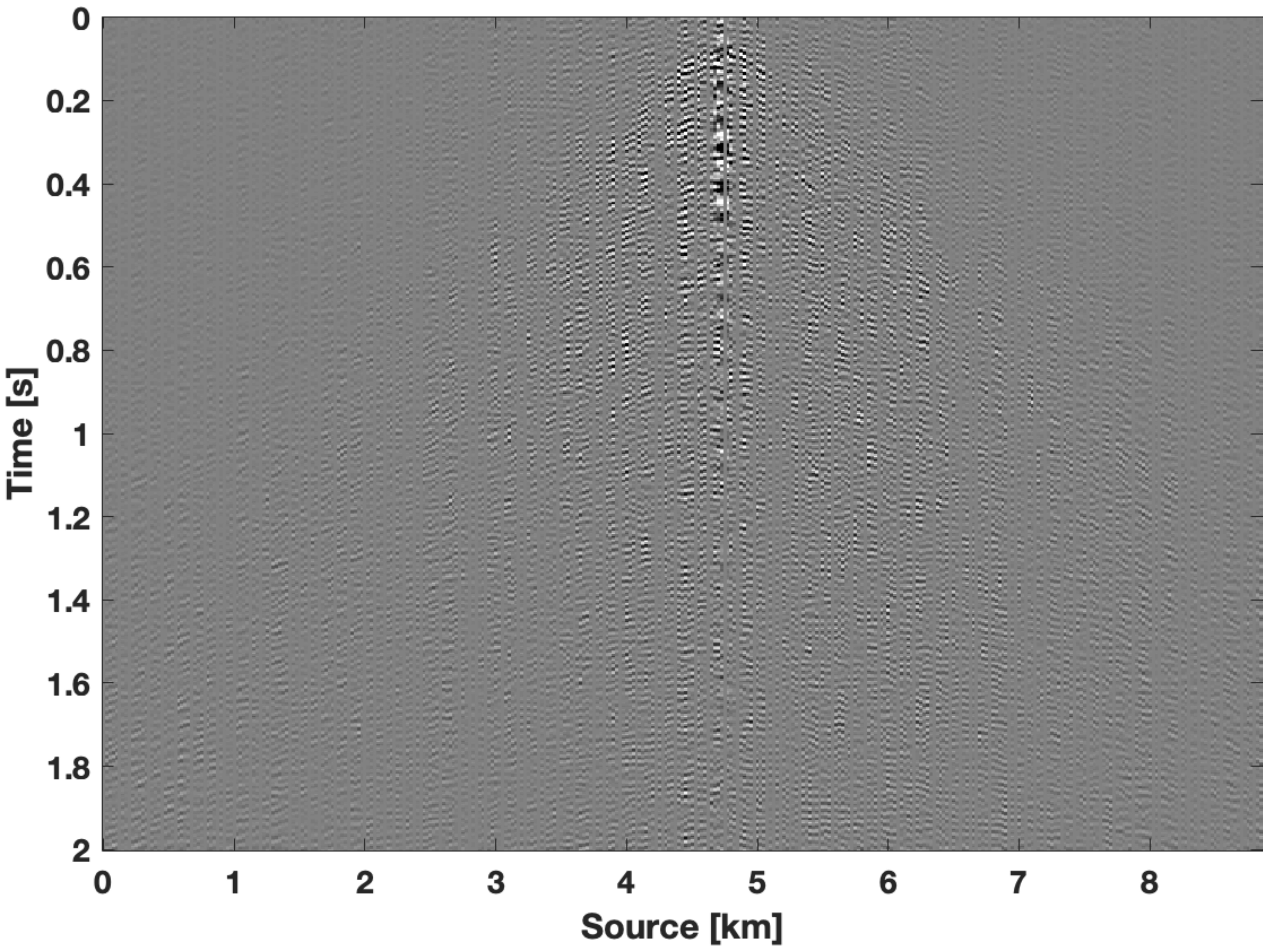


Difference: True - Recovery w/ weighted (rank = 85)



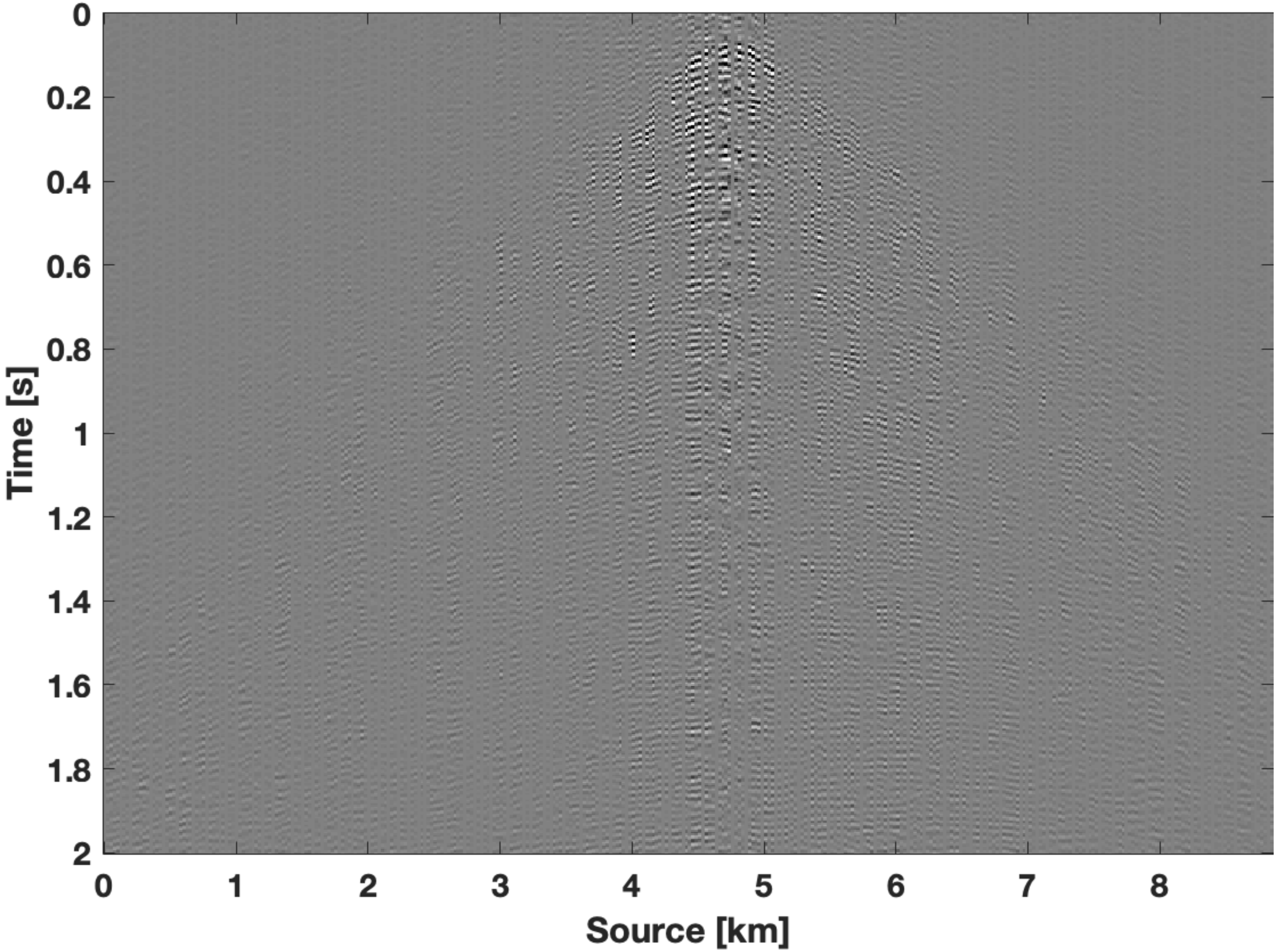
SNR = 10.69 dB
Rank = 85
Subspace rank = 85

Difference: True - Recovery w/ weighted (rank = 25)



SNR = 11.49 dB
Rank = 25
Subspace rank = 25

Difference: True - Recovery w/ limited subspaces weighted



SNR = 13.31 dB
Rank = 85
Subspace rank = 25

Conclusion & future work

Limited weighted strategy

- prevents overfitting at lower frequencies
- improves SNR at higher frequencies

Simultaneous source separation

Acknowledgement

We would like to thank the Georgia Institute of Technology for funding this research

Thank you for your attention!