A practical workflow for land seismic wavefield recovery with weighted matrix factorization

Yijun Zhang and Felix J. Herrmann

SEG 2021

SLIM
Georgia Institute of Technology

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.
Chiu, Stephen K. "3D attenuation of aliased ground roll on randomly undersampled data." *SEG International Exposition and Annual Meeting*. OnePetro, 2019.
Zhang, Y. et al. "High-frequency wavefield recovery with weighted matrix factorizations." In SEG Technical Program Expanded Abstracts, 2019.

SLIM

Motivation

Fully sample data is a prerequisite to

- multiple removal, migration & FWI

Seismic data is collected randomly along the spatial coordinate to

- shorten the acquisition time, reduces cost
- shift the burden from field acquisition to data processing

Matrix completion for low-to-mid frequencies

- exploits low-rank structure for recovery, simple & computationally cheap

Weighted matrix completion for higher frequencies

- gives good results for higher frequency slices on marine data
- **application to land data is hampered by ground roll**

Van De Coevering, Norbert, Klaas Koster, and Rob Holt. "A sceptic's view of VVAz and AVAz." *SEG Technical Program Expanded Abstracts 2019*.
Tan, Justin, et al. "SEAM Phase II Barrett model classic data study: Processing, imaging, and attributes analysis." *SEG Technical Program Expanded Abstracts 2019.*

Blind study on 3D SEAM Barrett dataset

Data dimension: 80 x 21 x 641 x 641 x 667

$(n_{sx} \times n_{sy} \times n_{rx} \times n_{ry} \times n_t)$

Receiver sampling interval: 12.5 m

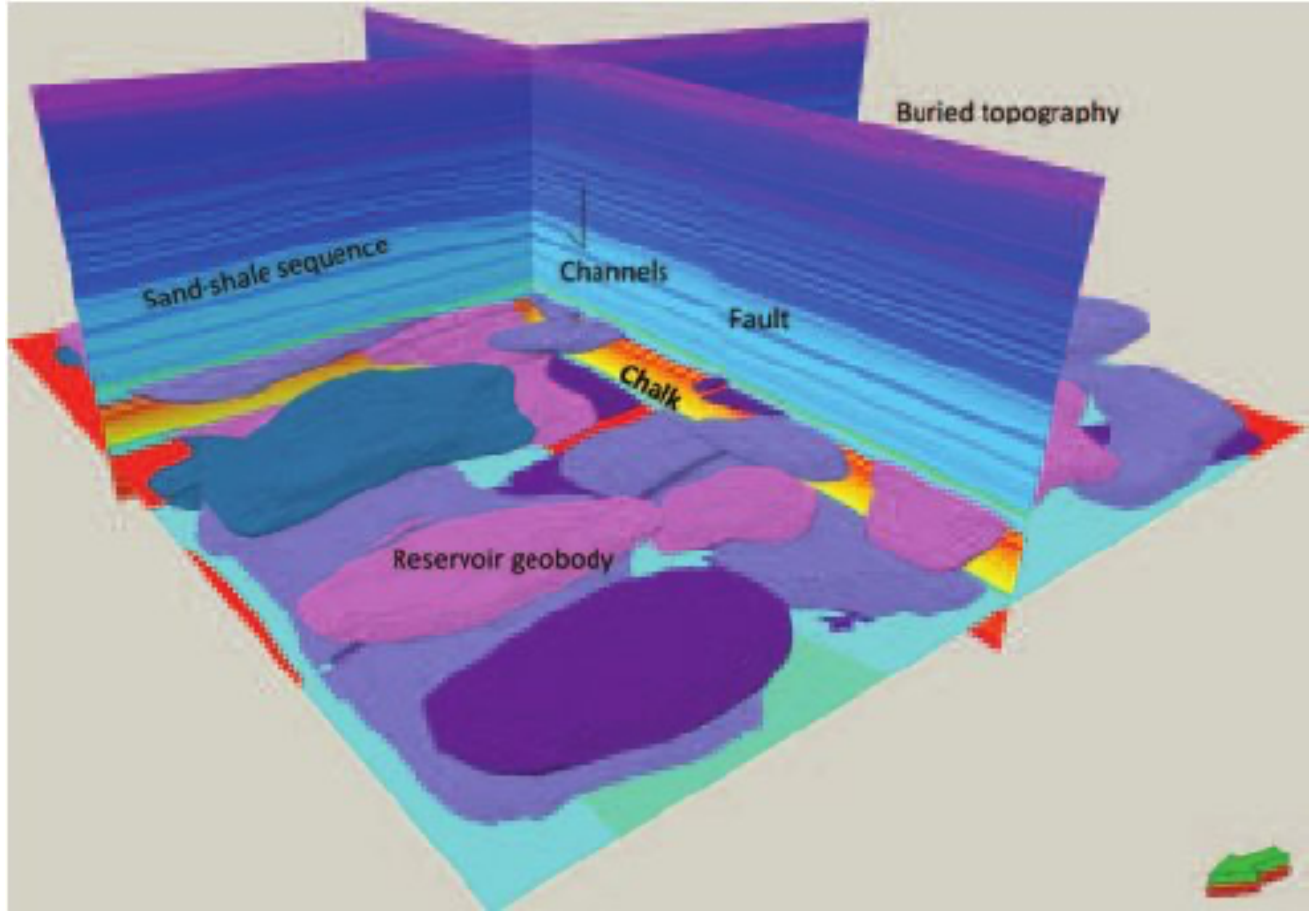
Source sampling interval: 25 m in the shot line direction and 100 m in the perpendicular direction

Time sampling interval: 0.006 s

Subset of dataset

- benchmark for land data
- contains realistic surface waves

https://www.researchgate.net/publication/



Acquisition geometry w/ $\sim 75\,\%$ missing receivers, 21 source lines

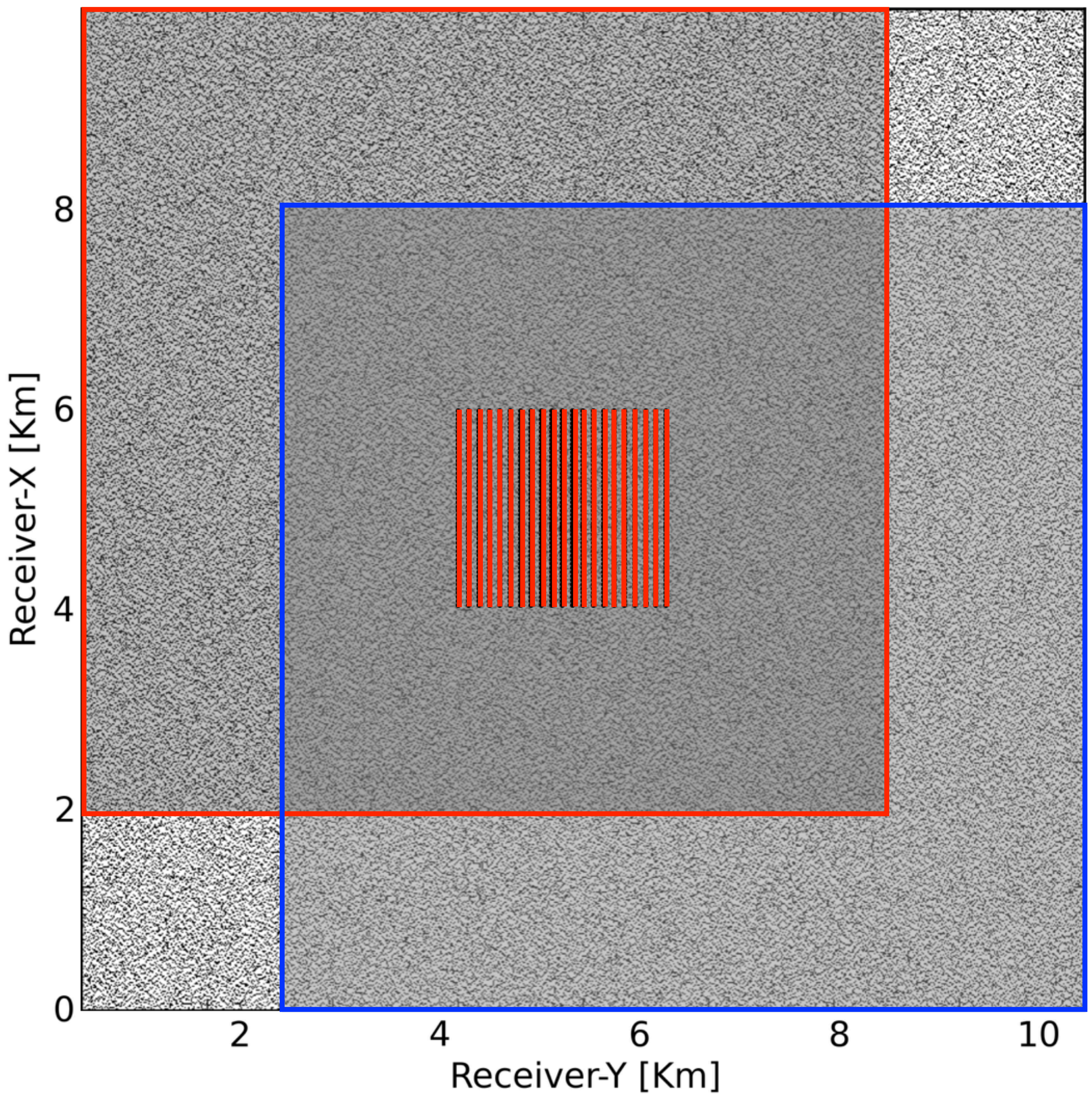
A subset consists of 21 source lines (red lines in the center area)

Each source line contains 80 sources

The 8×8 km receiver aperture moves with the source location

- neighboring shots share most randomly sampled receivers (black dots in the figure)
- some drop-off and others add (red and blue rectangles in the figure)

Acquisition geometry for the observed dataset



Liu, Xuewei. "Ground roll supression using the Karhunen-Loeve transform." *Geophysics* 64.2 (1999): 564-566.

SLIM

Impact of ground roll

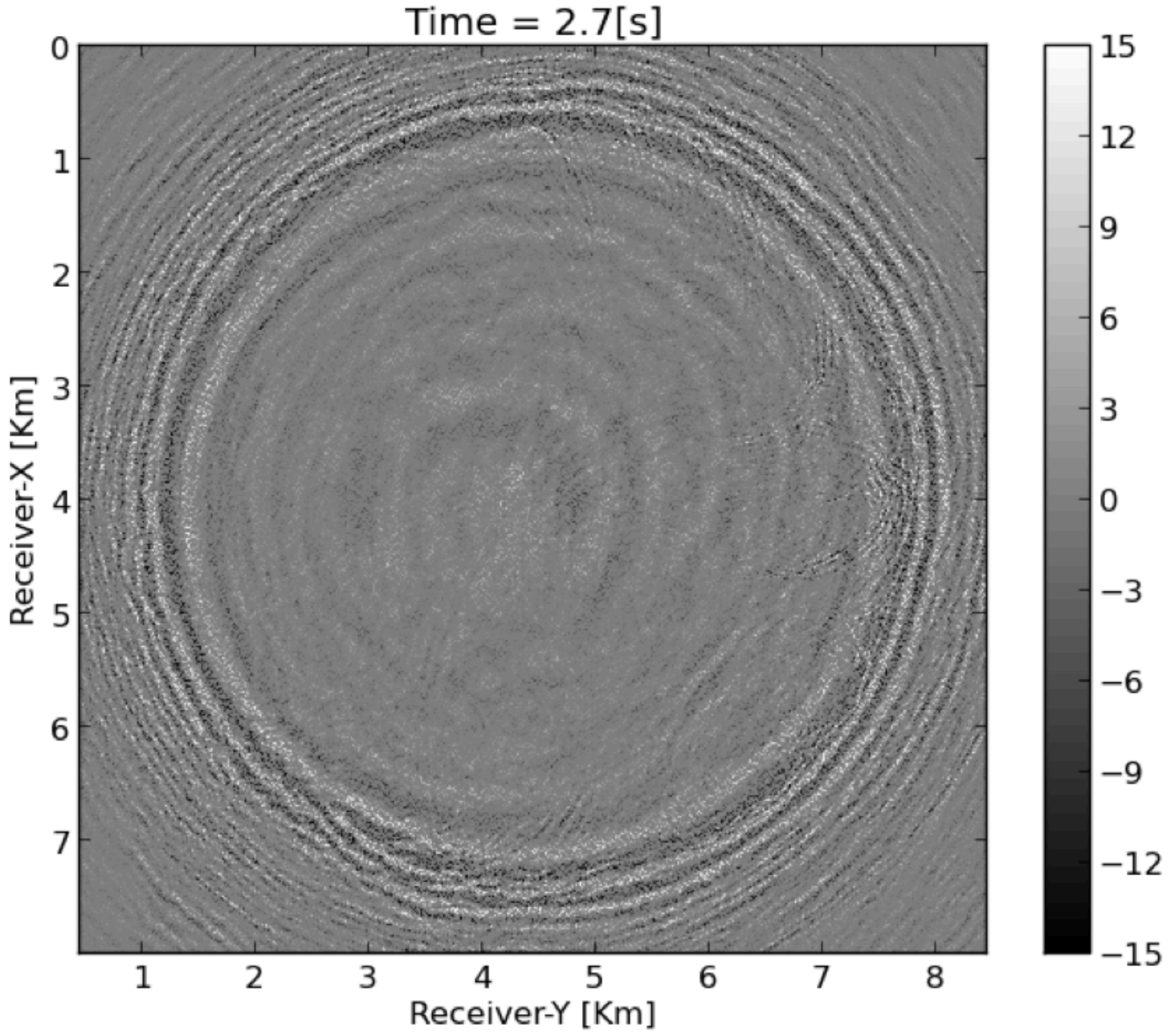
Ground roll corresponds to Rayleigh-type surface waves

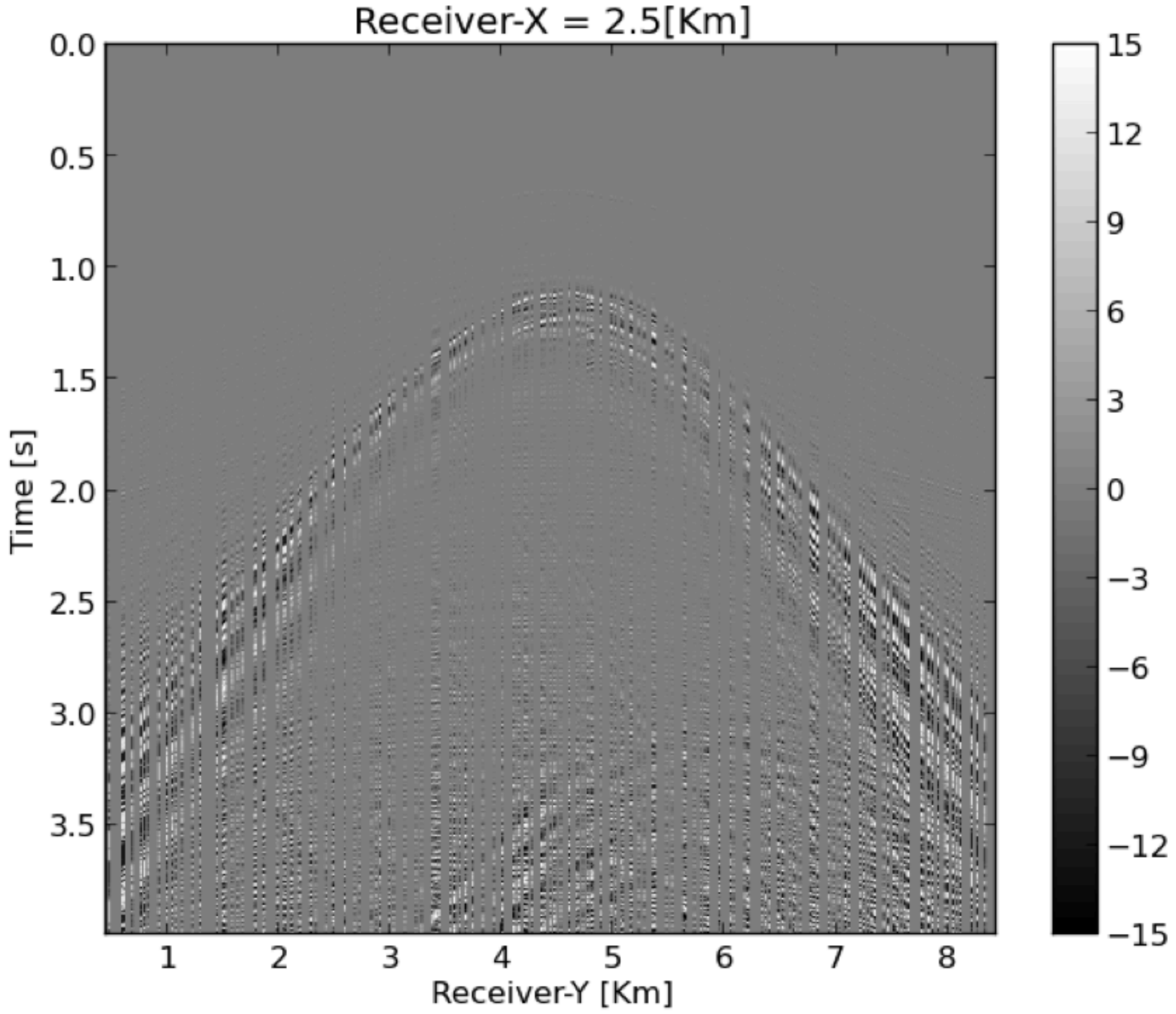
- slow & aliased
- strong amplitude

Dominate the reconstruction at the expense of weaker body waves.

Observed data in time domain (one shot)

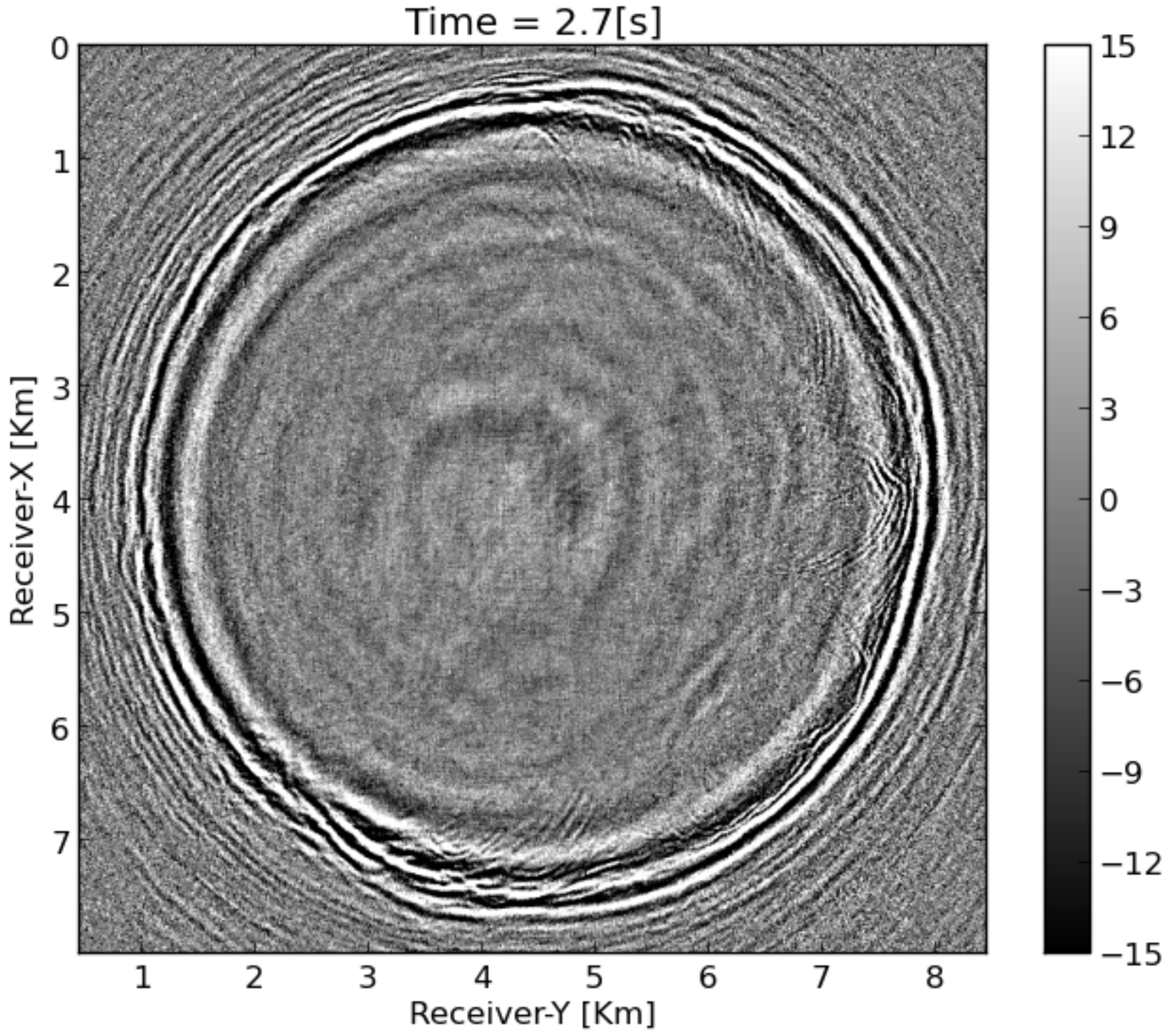
w/ $\sim 75\ \%$ missing receivers

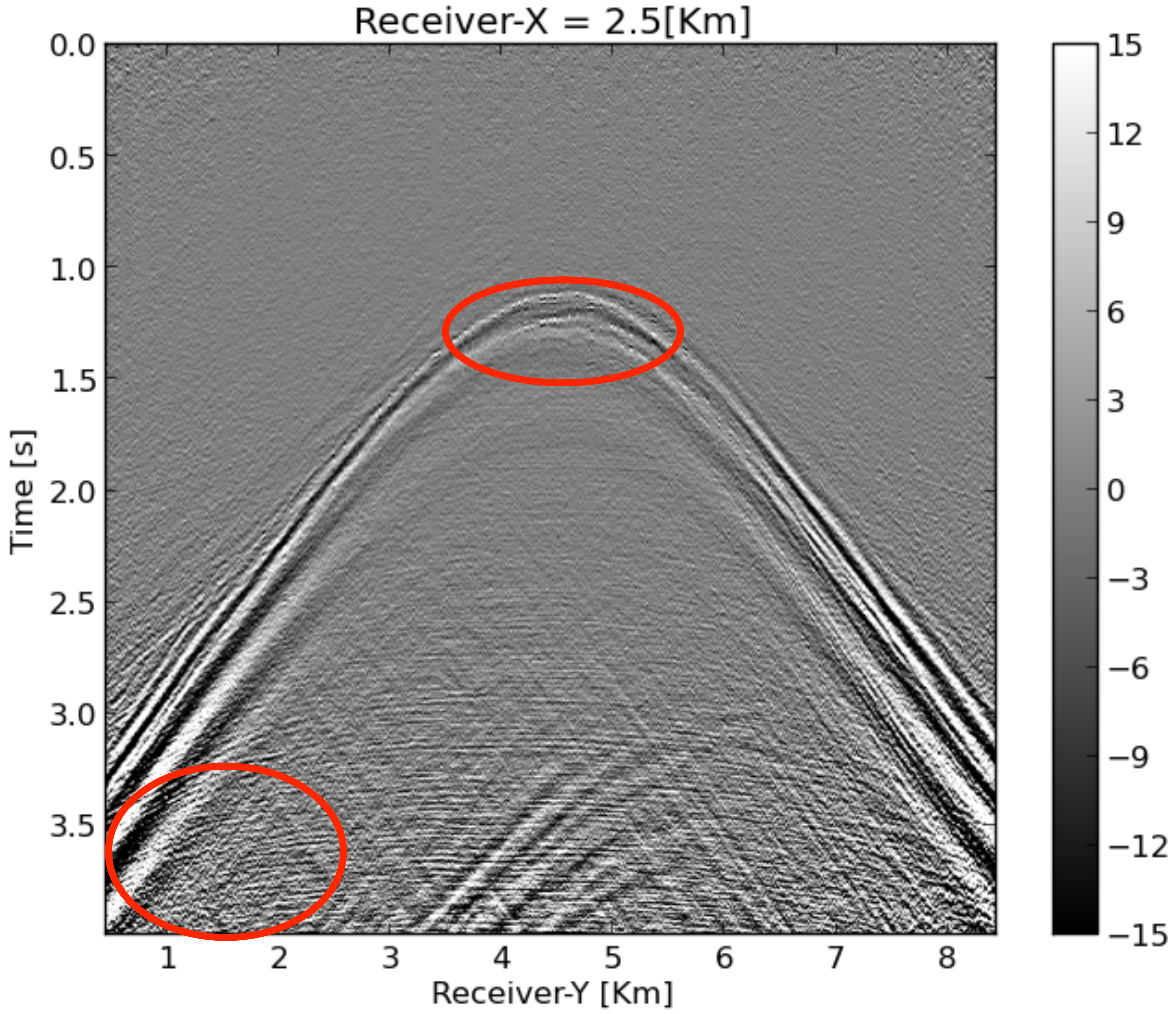




Reconstructed data in time domain (one shot)

via weighted matrix completion w/o proposed workflow





Main research questions

How can we use weighted matrix factorization on land data while avoiding the impact of ground roll?

Answer:

- ***Reconstruct the body and surface (ground roll) waves separately.***

Why separation?

Answer:

- ***Reduce the effect of strong aliased ground roll on body wave reconstruction.***
- ***The ground roll can be separated from body waves, at least in an approximate sense.***

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

Wavefield reconstruction via matrix completion

Successful reconstruction schemes

- 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)

Weighted matrix completion

- further improves the wavefield recovery at higher frequencies by introducing matrix 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.

Weighted matrix completion

Variational formulation:

Expensive for large scale

$$\underset{\mathbf{X}}{\text{minimize}} \quad \overbrace{\|\mathbf{QXW}\|_*}^{\text{Sum of singular values}} \quad \text{subject to} \quad \underbrace{\|\mathscr{A}(\mathbf{X}) - \mathbf{B}\|_F}_{\text{Frobenius norm}} \leq \epsilon$$

where

$$\mathbf{Q} = w_1\mathbf{U}\mathbf{U}^H + \mathbf{U}^{\perp}\mathbf{U}^{\perp H}, \;\; \mathbf{W} = w_2\mathbf{V}\mathbf{V}^H + \mathbf{V}^{\perp}\mathbf{V}^{\perp H}, \;\; \mathbf{B}, \mathbf{X} \in \mathbb{C}^{(N_{sx}\times N_{rx})\times(N_{sy}\times N_{ry})}.$$

- $\mathbf{U}, \mathbf{V}$ column and row subspaces (prior information) derived from neighboring frequencies
- $w_1, w_2 \in (0,1]$ similarity of the prior information and to-be-recovered data
- Small values for scalars indicate more confidence in the 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.

SLIM

Weighted matrix factorization

Weighted matrix factorization:

Computation expensive

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

where

$$\mathbf{X} = \mathbf{LR}^H, \;\; \mathbf{L} \in \mathbb{C}^{(N_{sx}\times N_{rx})\times r}, \;\; \mathbf{R} \in \mathbb{C}^{(N_{sy}\times N_{ry})\times r}.$$

Incorporating weighting matrices into the objective function

- complicates optimization
- increases the computational cost of minimization

Zhang, Yijun, et al. "Wavefield recovery with limited-subspace weighted matrix factorizations." SEG International Exposition and Annual Meeting. OnePetro, 2020.

SLIM

Computationally efficient formulation

Easier to compute by moving weighting matrices to the data misfit constraint

Variational formulation:

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

where

$$\bar{\mathbf{X}} = \mathbf{QXW}, \mathbf{Q}^{-1} = \frac{1}{w_1}\mathbf{UU}^H + \mathbf{U}^{\perp}\mathbf{U}^{\perp H}, \ \mathbf{W}^{-1} = \frac{1}{w_2}\mathbf{VV}^H + \mathbf{V}^{\perp}\mathbf{V}^{\perp H}.$$

Weighted matrix factorization:

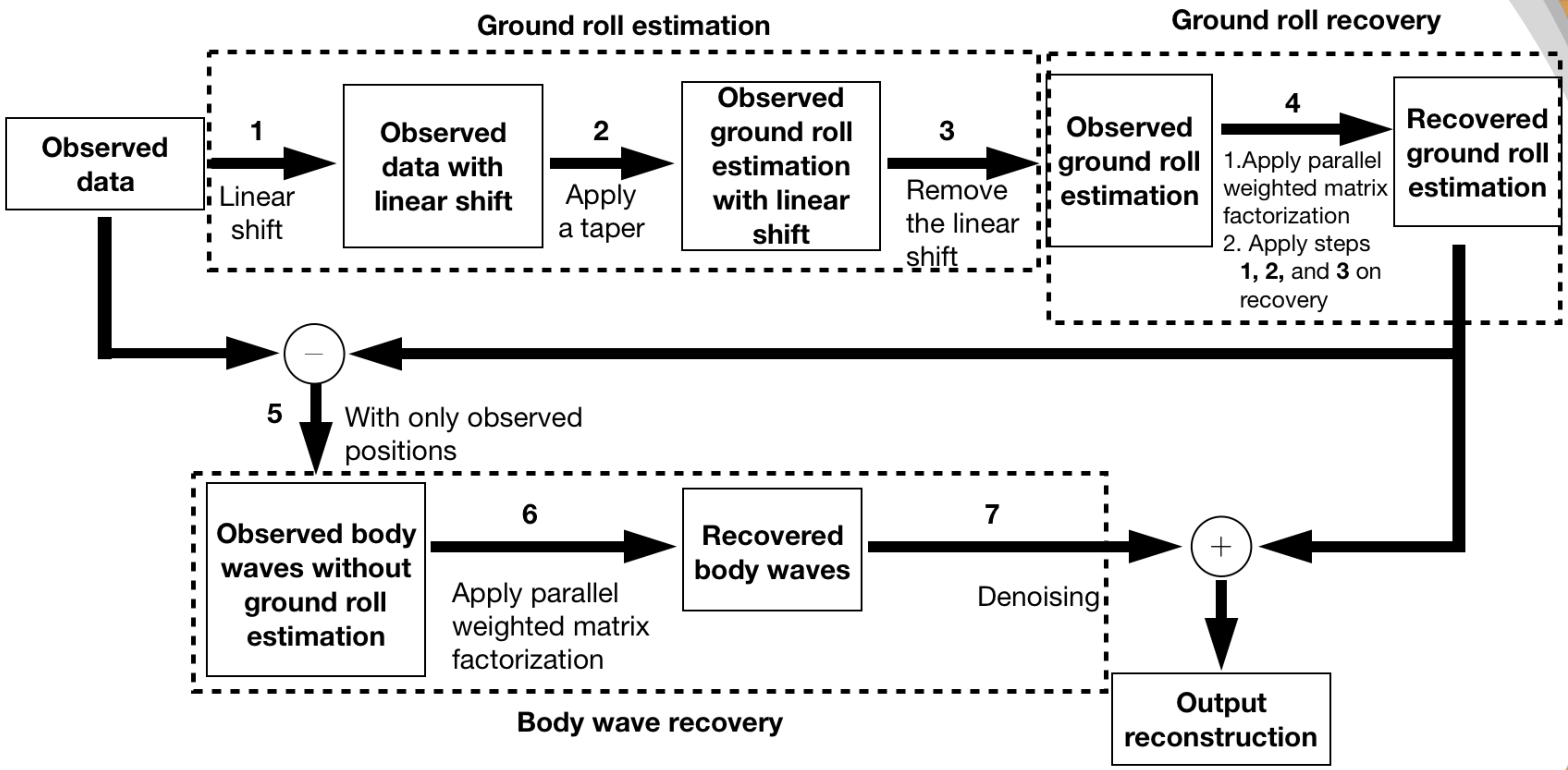
$$\underset{\bar{\mathbf{L}},\bar{\mathbf{R}}}{\text{minimize}} \quad \frac{1}{2}\left\| \begin{bmatrix} \bar{\mathbf{L}} \\ \bar{\mathbf{R}} \end{bmatrix} \right\|_F^2 \quad \text{subject to} \quad \| \mathcal{A}(\mathbf{Q}^{-1}\bar{\mathbf{L}}\bar{\mathbf{R}}^{\mathbf{H}}\mathbf{W}^{-1}) - \mathbf{B}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

where

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

- The original solution can be recovered by $\mathbf{X} = \mathbf{Q}^{-1}\bar{\mathbf{X}}\mathbf{W}^{-1}$.

Proposed workflow



Ground roll estimation

A. Observed data

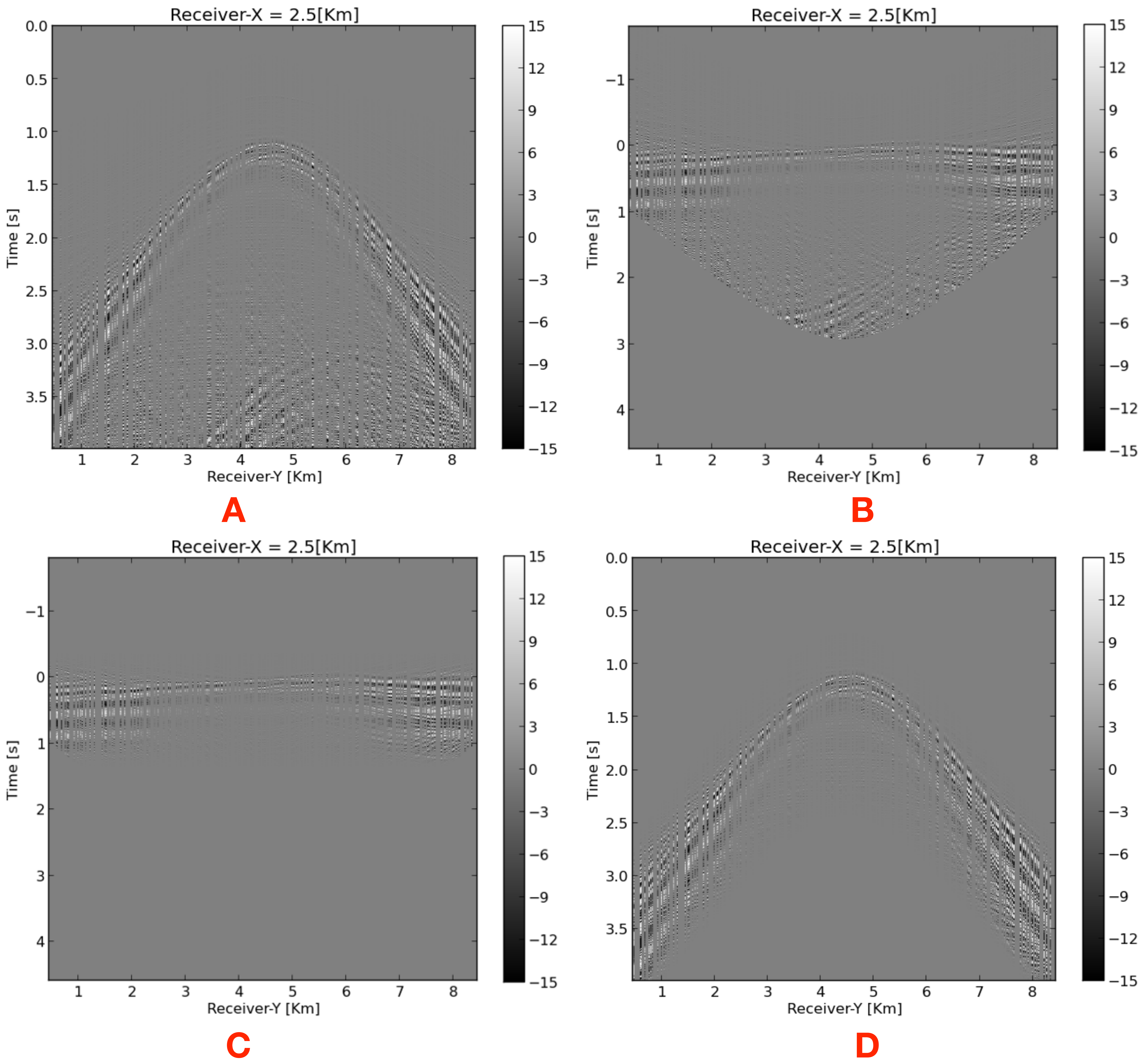
B. The output after aligning the ground roll

- by using a linear shift

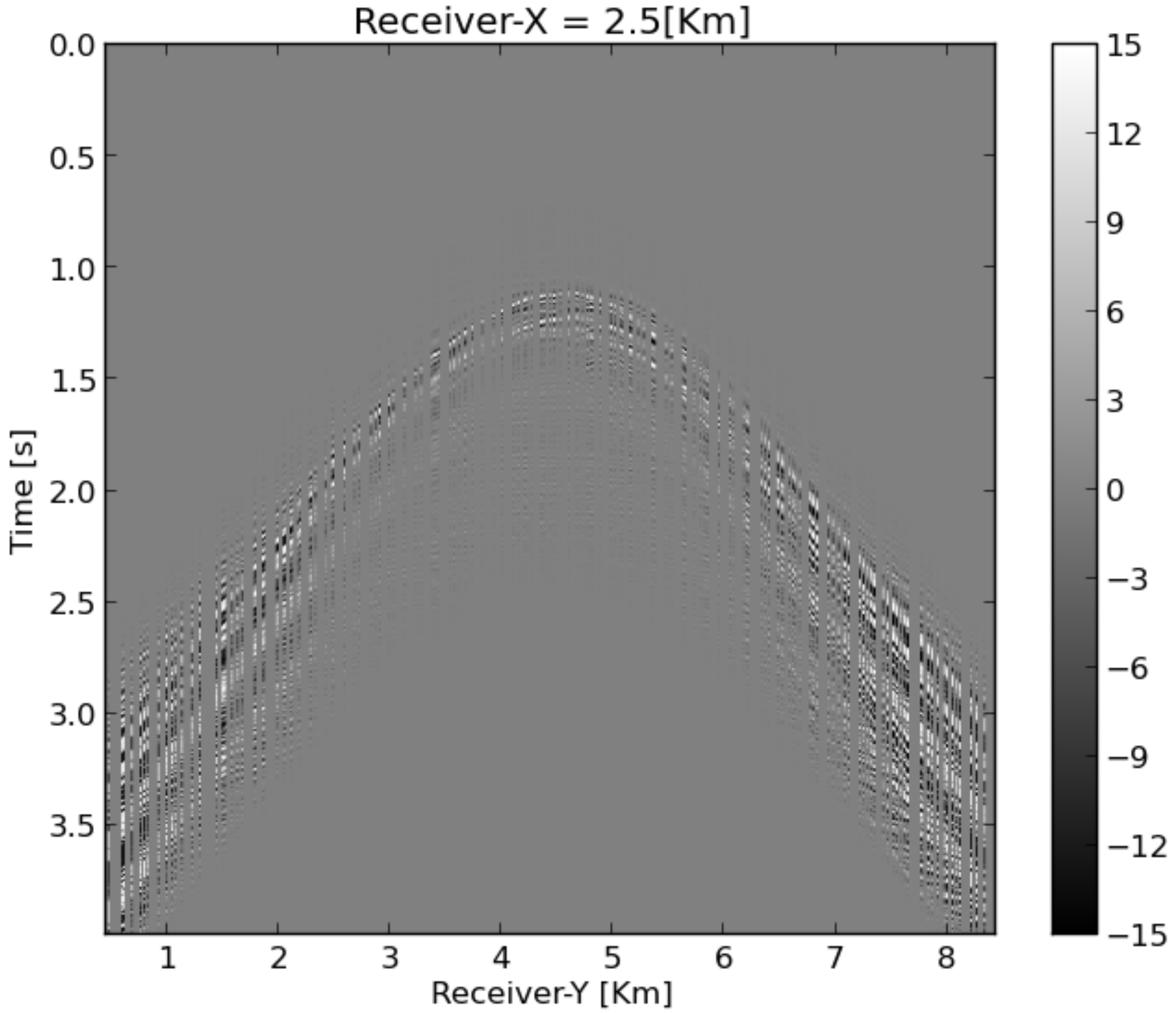
C. The output after applying a smooth taper

D. The ground roll estimation

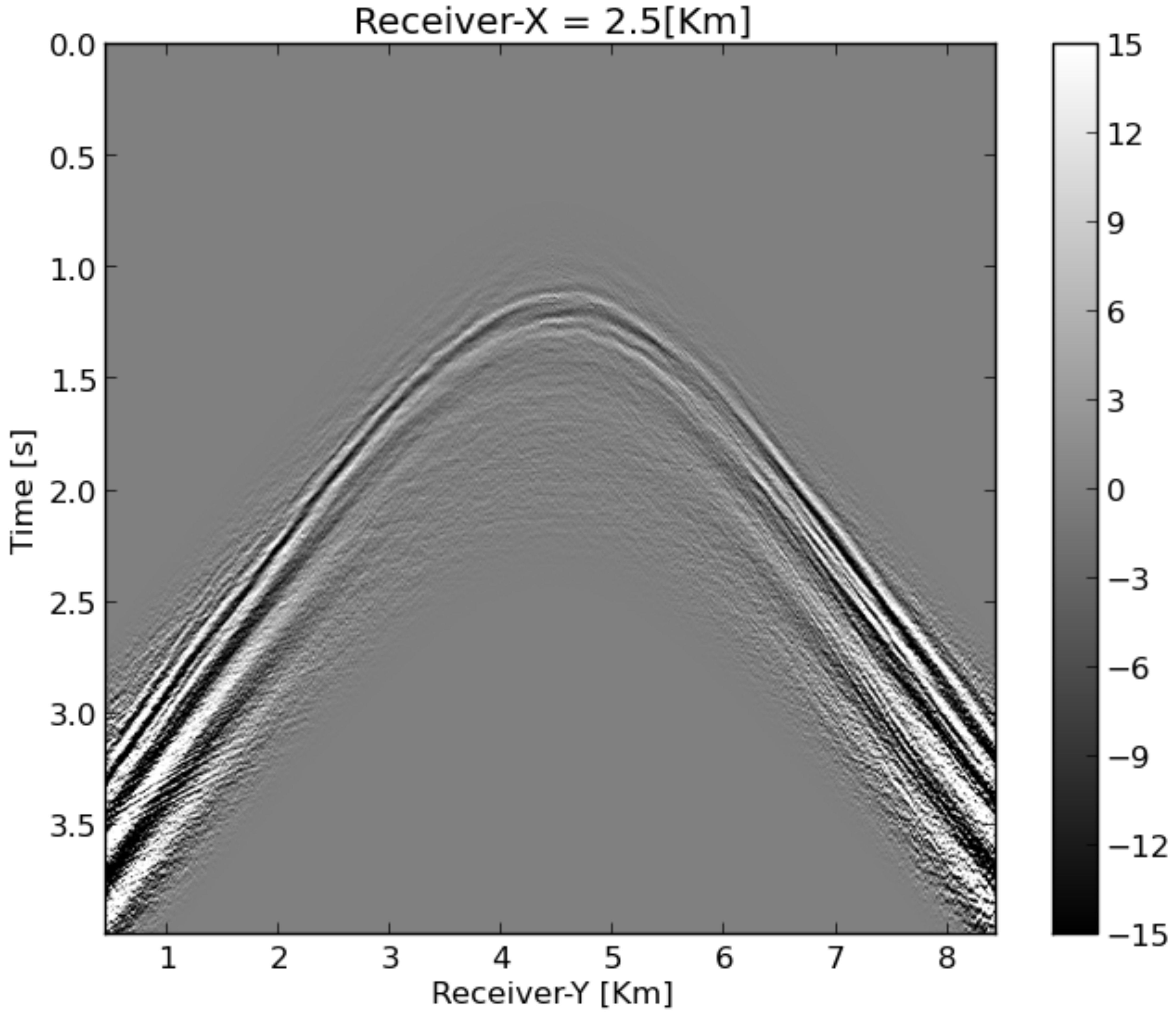
- by undoing the linear shift



Ground roll recovery

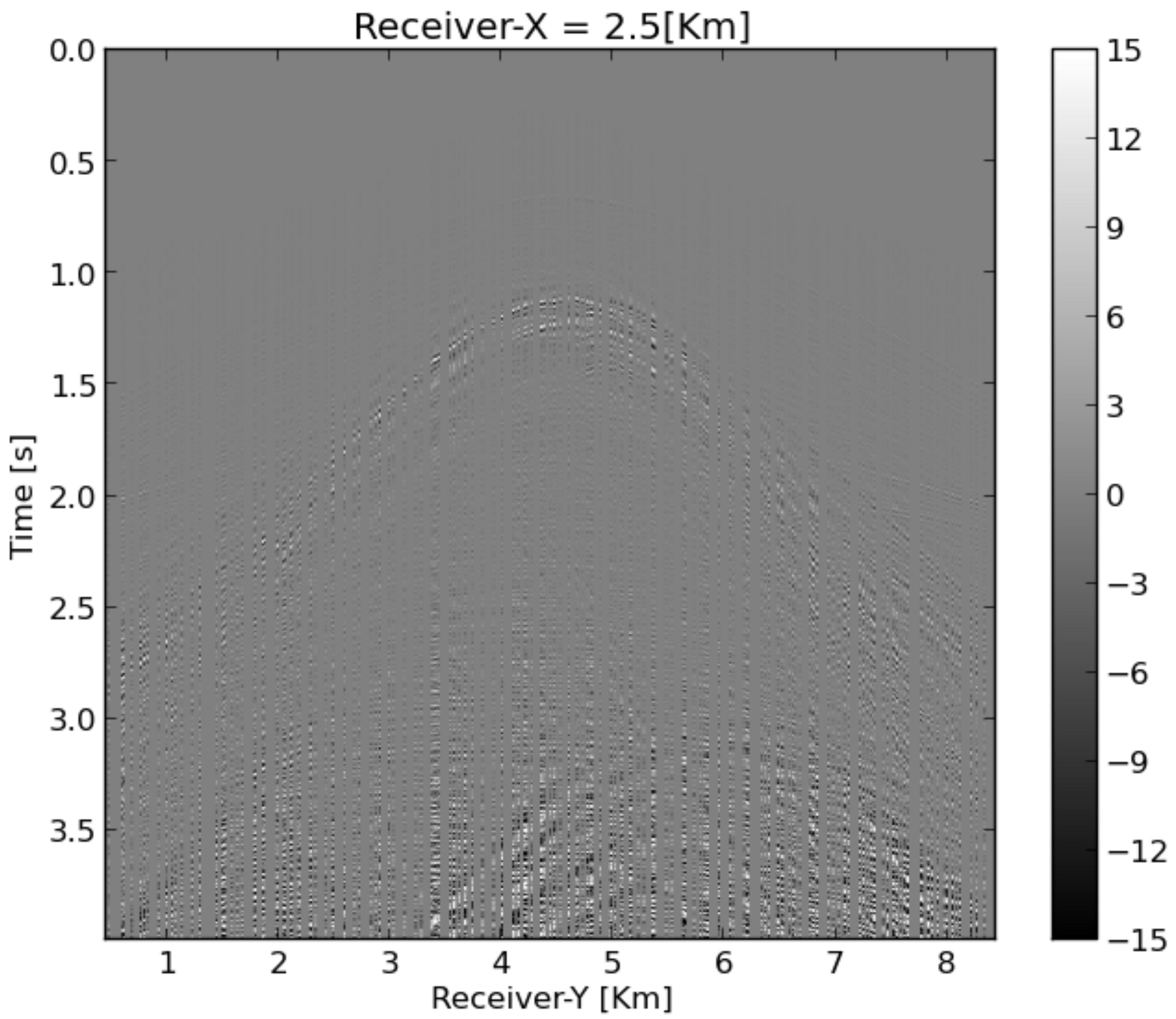


Observed ground roll

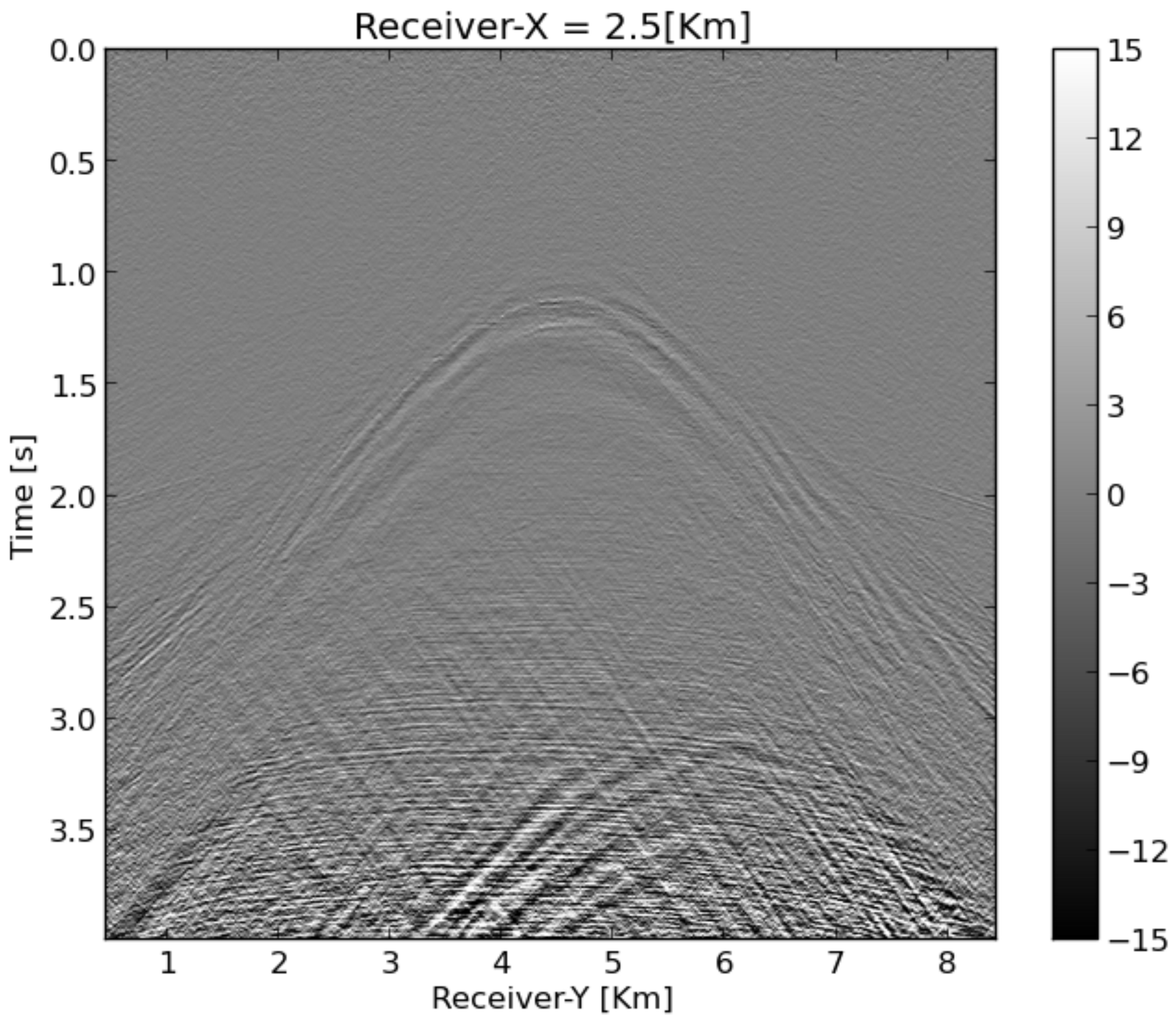


Recovered ground roll

Body wave recovery

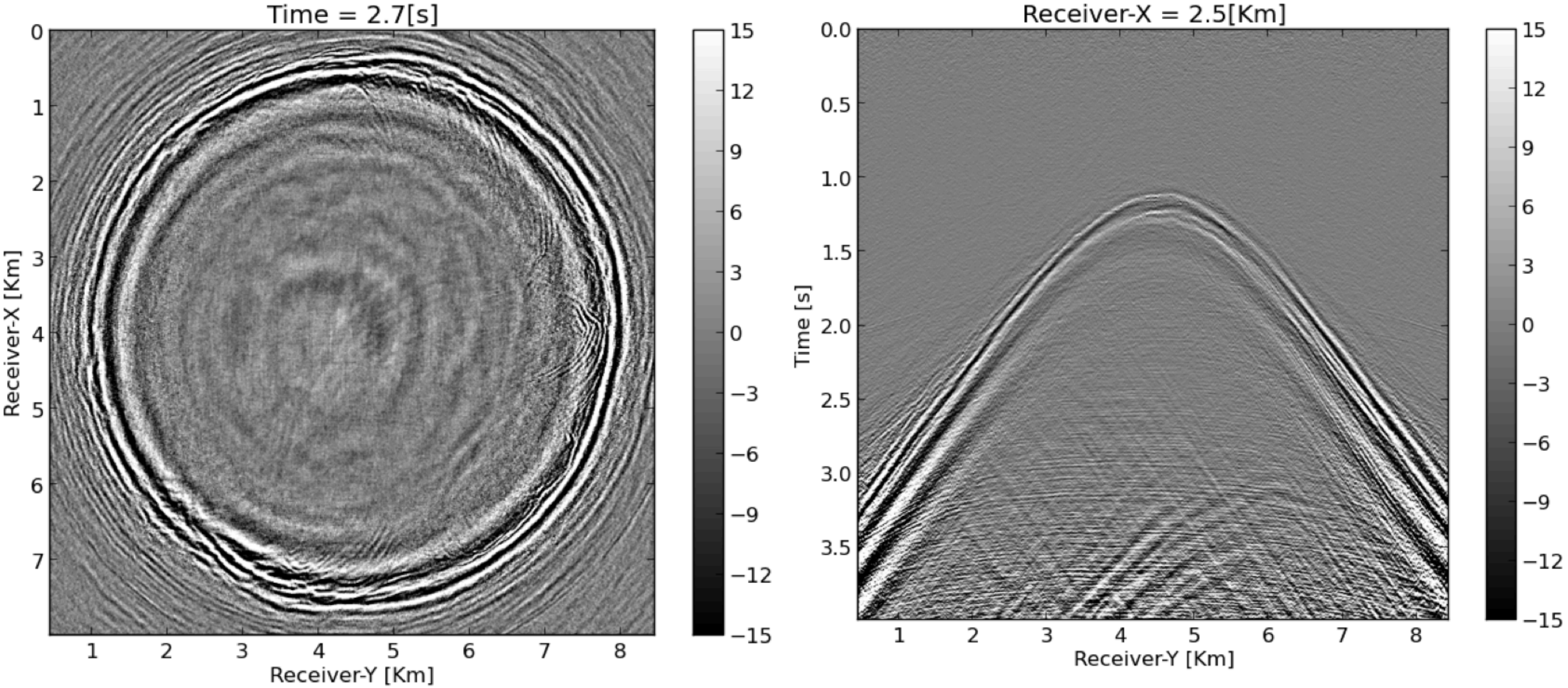


Observed body wave

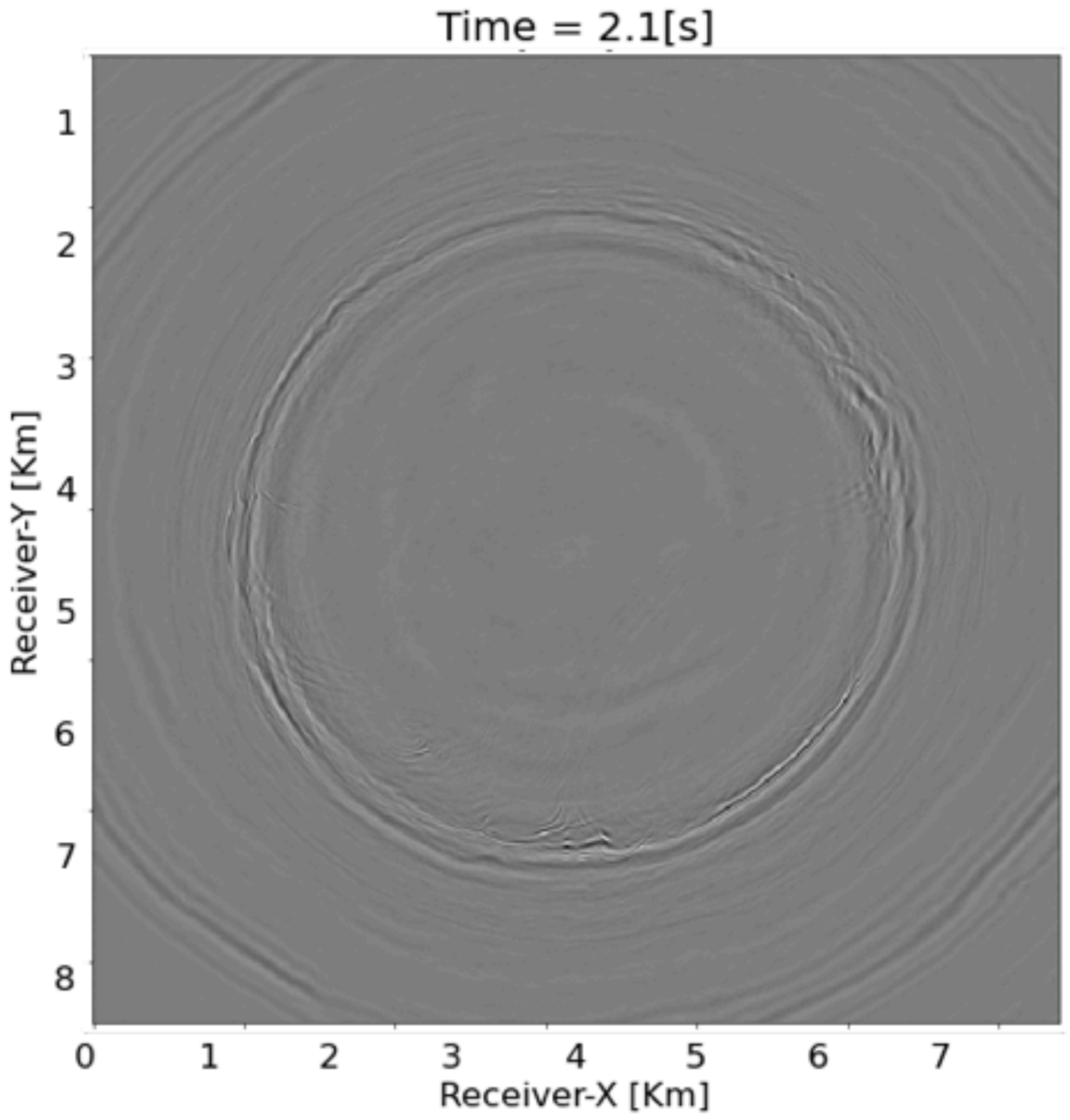


Recovered body wave

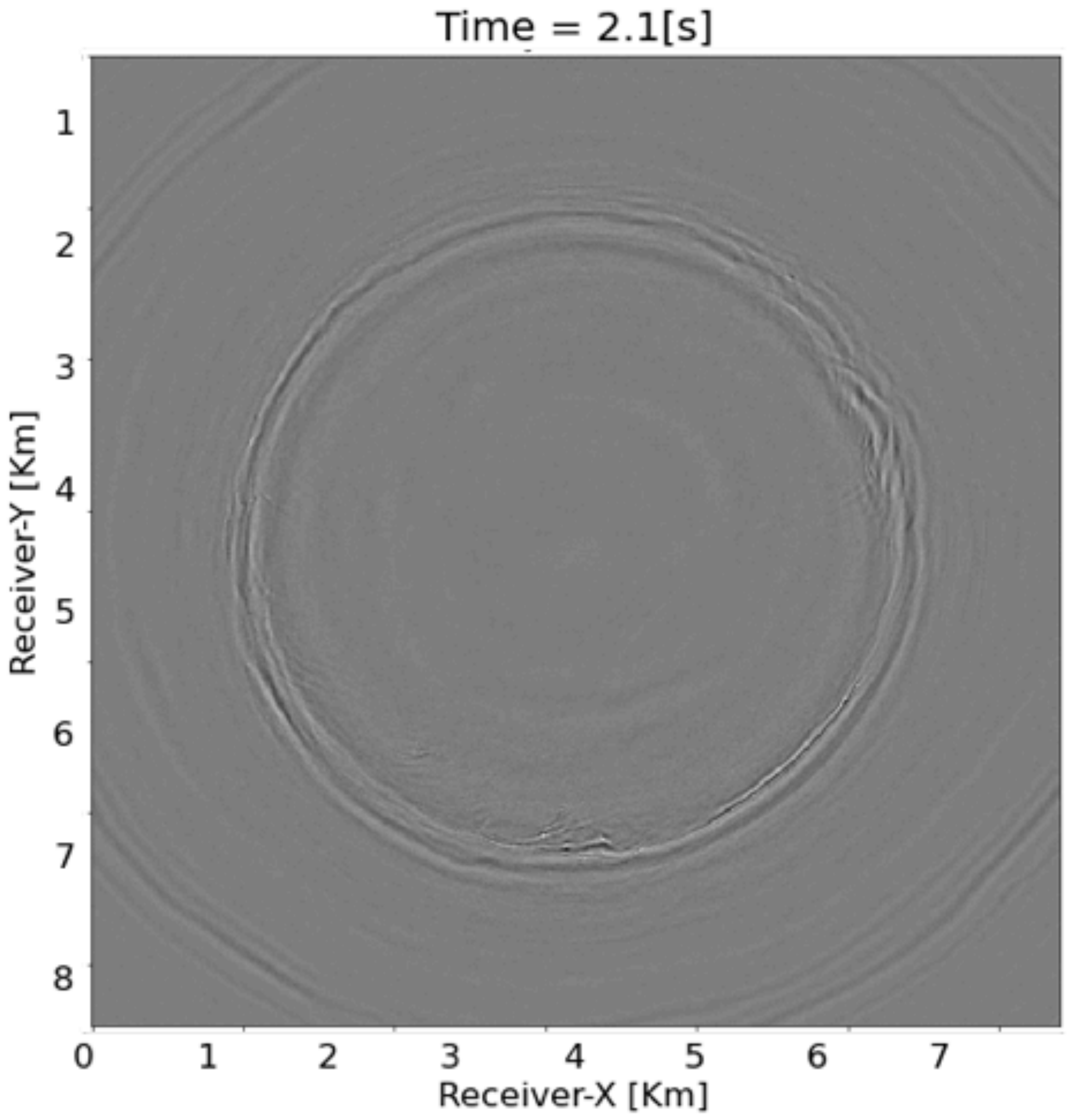
Final reconstructed result in time domain



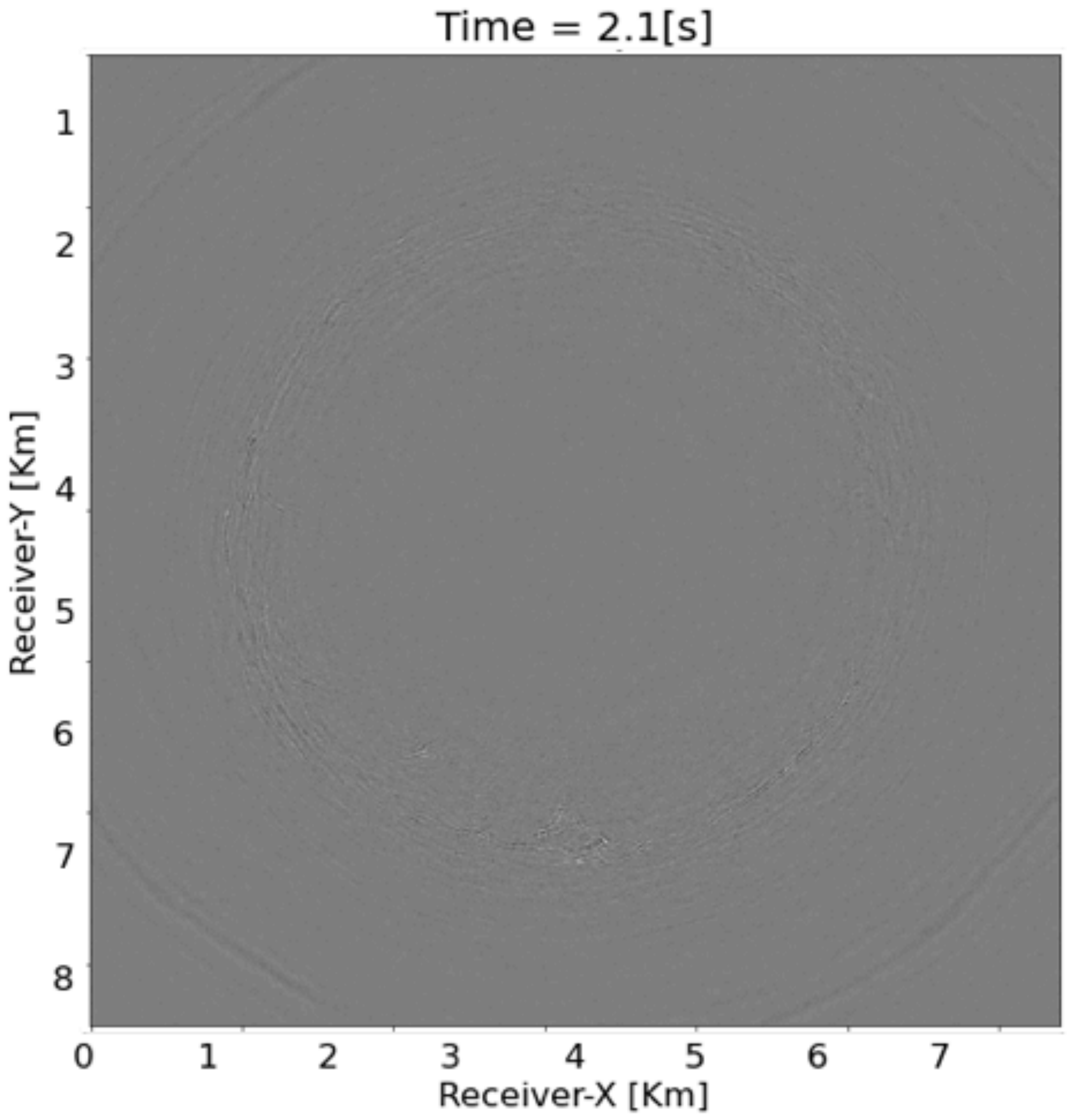
QC



Ground truth



Recovered data



Difference

Conclusions

We mitigate the effects of strongly aliased ground roll by employing the proposed separation.

Furthermore, the proposed workflow successfully recovers body waves (reflections and diffractions).

Acknowledgement

This research was carried out with the support of Georgia Research Alliance and partners of the ML4Seismic Center.

We would like to express our gratitude to Georgia Institute of Technology for funding this study.

We would also like to thank Occidental for providing the dataset, as well as Klaas Koster for assisting us with the blind test.

Thank you for your attention!!