Transfer learning in large-scale ocean bottom seismic wavefield reconstruction

Mi Zhang[1, 2], Ali Siahkoohi[1], and Felix J. Herrmann[1]

[1] Georgia Institute of Technology
[2] China University of Petroleum - Beijing

SLIM

SEG20 11-16 OCT 2020

Problem setup

Ocean bottom node (OBN) geometry:

- desirable source sampling
- on the grid sparse receivers on Ocean bottom

SLIM

Problem setup

Ocean bottom node (OBN) geometry:

- desirable source sampling
- on the grid sparse receivers on Ocean bottom

Objective:

- Recover missing receivers from severe subsampling
- Improve the reconstruction quality & computational efficiency

Previous works

Deep-learning-based wavefield reconstruction:

- ► training only relies on acquired data
- ► implicit deep “factorization” via a **nonlinear** neural net
- ► recover **randomly** or **periodically** from missing receivers
- ► **high** missing rates (~90%)

Main contribution

Transfer learning to accelerate the wavefield reconstruction.

Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Seismic data in a 3D survey

Seismic data is 5D: $(t$, Src x, Src y, Rec x, Rec $y)$

Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Seismic data in a 3D survey

Seismic data is 5D: $(t$, Src x, Src y, Rec x, Rec $y)$

$\Downarrow$

Taking Fourier transfer w.r.t. time: $(\omega$, Src x, Src y, Rec x, Rec $y)$

Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Seismic data in a 3D survey

Seismic data is 5D: $(t$, Src x, Src y, Rec x, Rec $y)$

$\Downarrow$

Taking Fourier transfer w.r.t. time: $(\omega$, Src x, Src y, Rec x, Rec $y)$

$\Downarrow$

Monochromatic seismic data is 4D: (Src x, Src y, Rec x, Rec y)

Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Matricization of monochromatic seismic data

Our framework operates on *monochromatic* frequency slices

(Src x, Src y, Rec x, Rec y)

Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Matricization of monochromatic seismic data

Our framework operates on *monochromatic* frequency slices

(Src x, Src y, Rec x, Rec y)

Two choices for matricization:

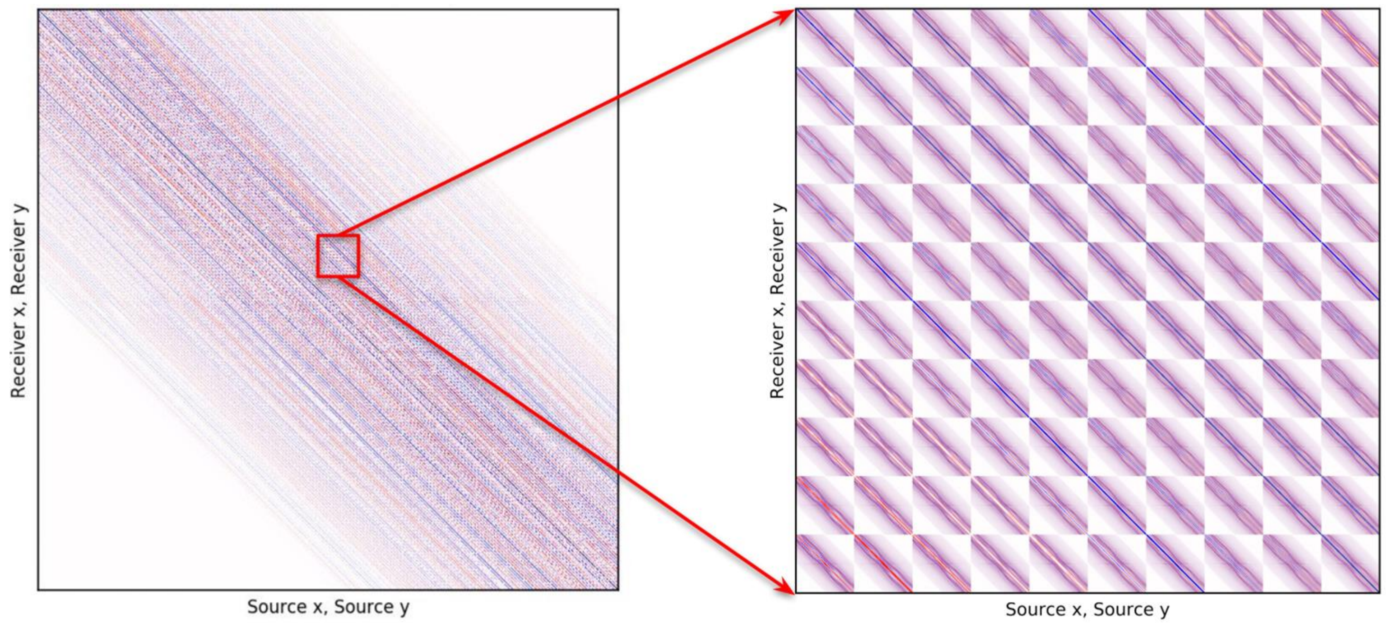
- ► (Rec x, Rec y) × (Src x, Src y) domain (canonical)
- ► **(Rec y, Src y) × (Rec x, Src x) domain (non-canonical)**

Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Matricization of monochromatic seismic data

- (Rec x, Rec y) × (Src x, Src y) domain (canonical)
- (Rec y, Src y) × (Rec x, Src x) domain (non-canonical)

fully-sampled data

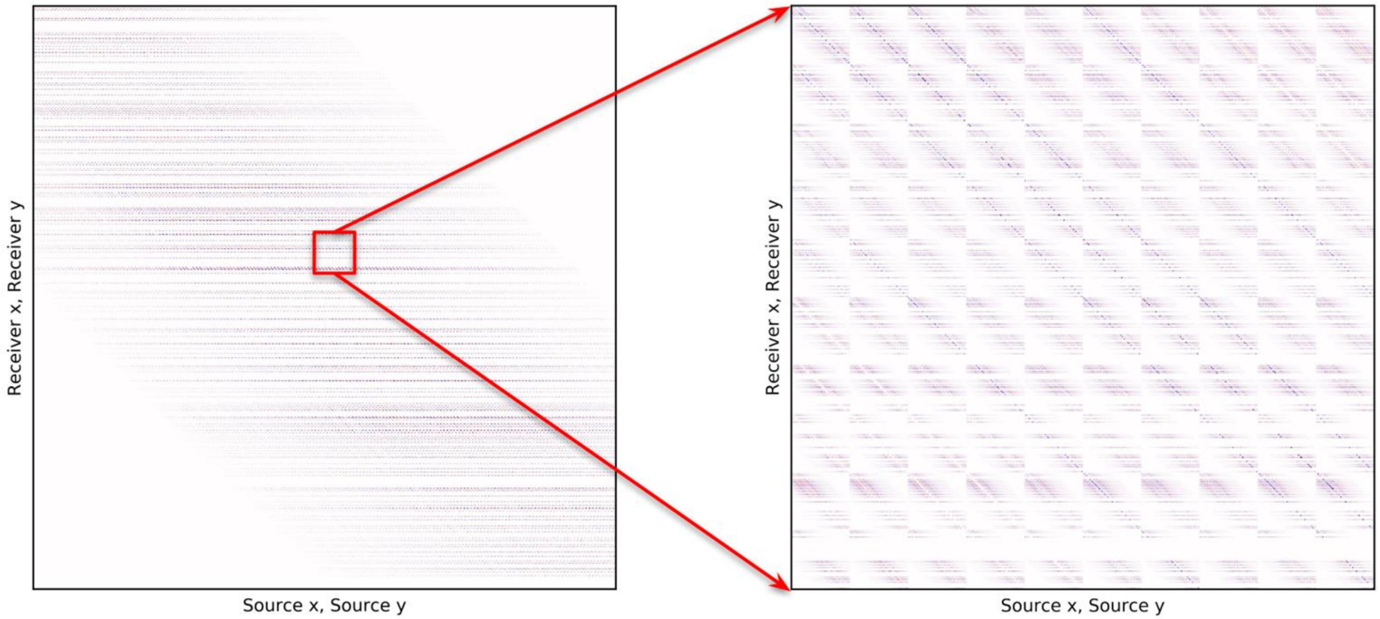


Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Matricization of monochromatic seismic data

- (Rec x, Rec y) × (Src x, Src y) domain (canonical)
- (Rec y, Src y) × (Rec x, Src x) domain (non-canonical)

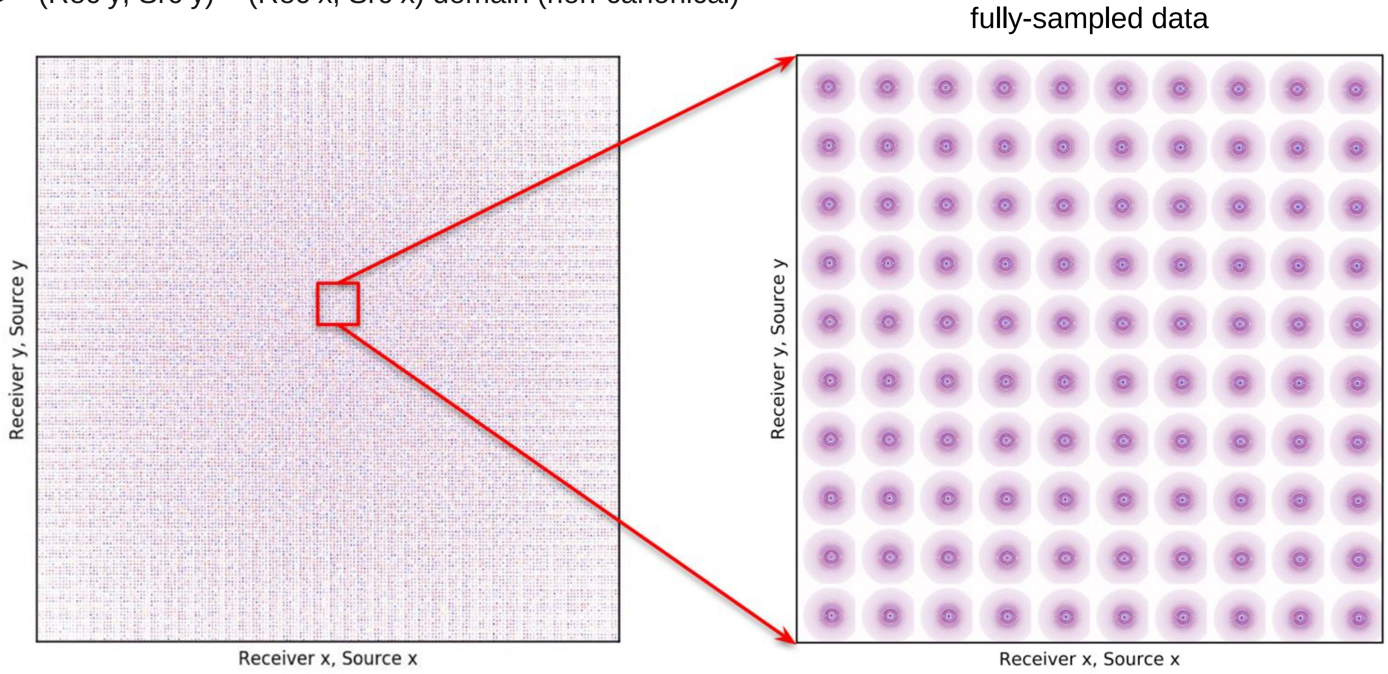
observed data – sampling rate 10%



Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Matricization of monochromatic seismic data

- (Rec x, Rec y) × (Src x, Src y) domain (canonical)
- (Rec y, Src y) × (Rec x, Src x) domain (non-canonical)

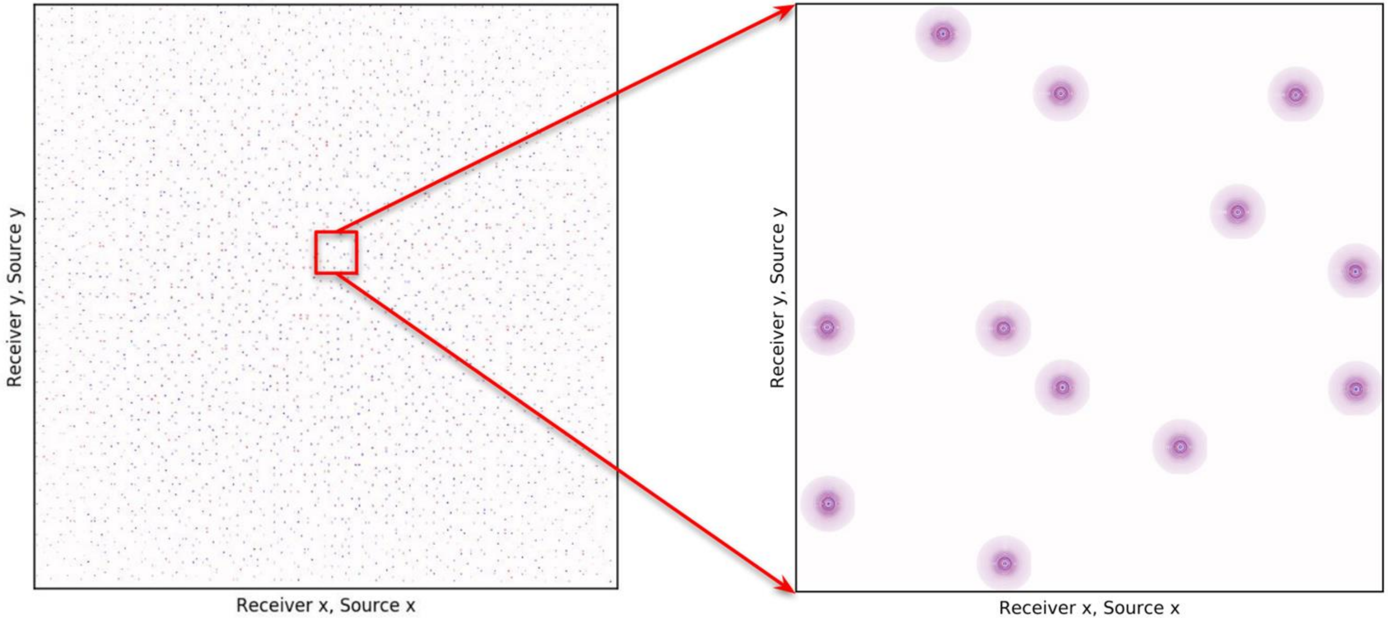


Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Matricization of monochromatic seismic data

- (Rec x, Rec y) × (Src x, Src y) domain (canonical)
- (Rec y, Src y) × (Rec x, Src x) domain (non-canonical)

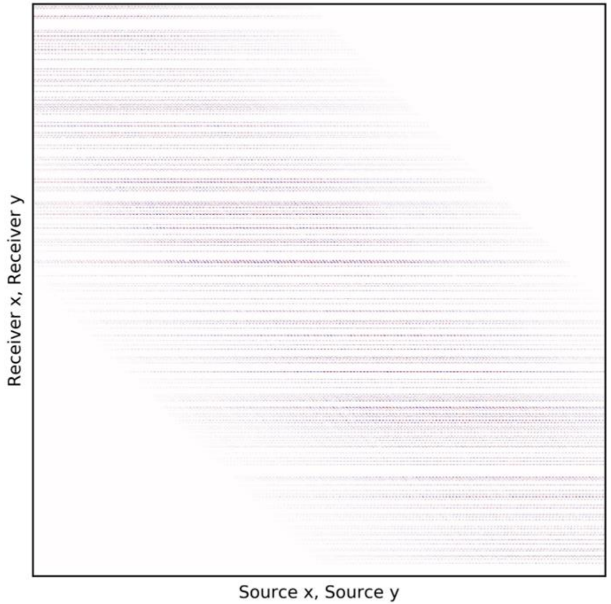
observed data – sampling rate 10%



Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

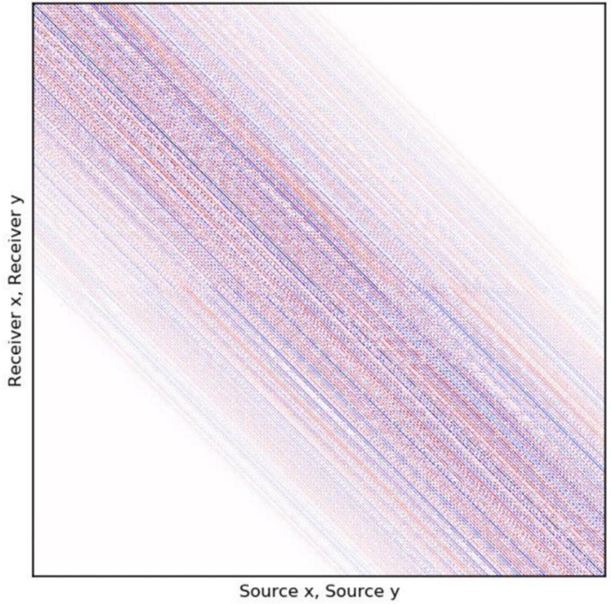
Objective 1: Recover missing receivers

observed monochromatic data



?

desired monochromatic seismic data

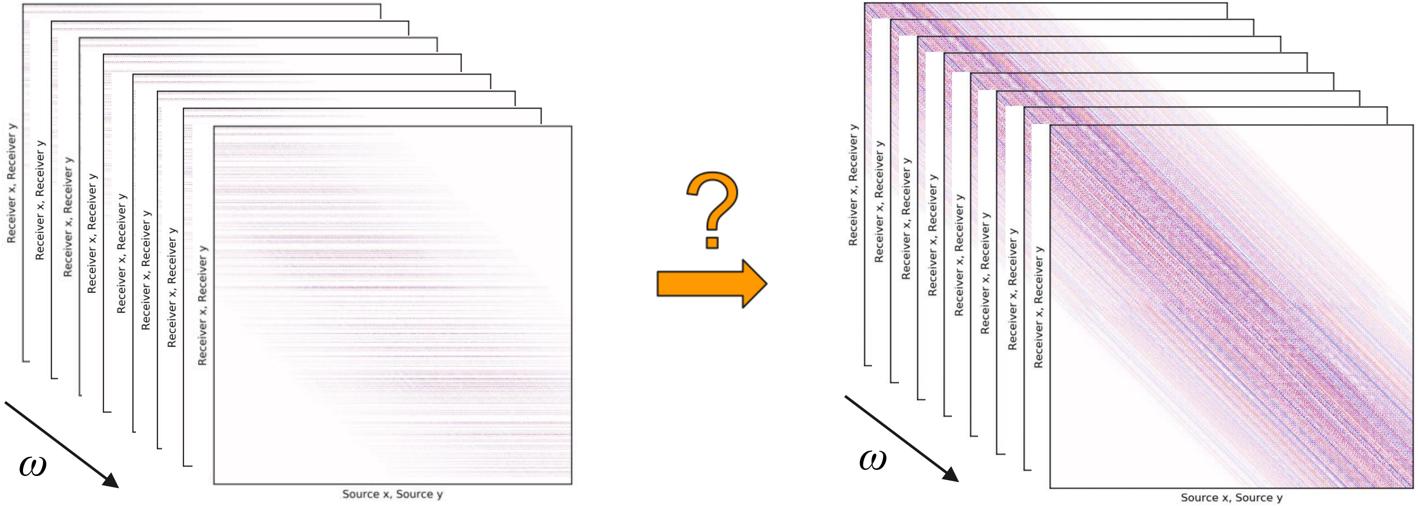


Demanet, L., 2006, Curvelets, wave atoms, and wave equations: PhD thesis, California Institute of Technology.
Silva, C. D., and F. J. Herrmann, 2013, Hierarchical Tucker tensor optimization - applications to tensor completion: Presented at the Sampling Theory and Applications conference.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Objective 2: Speed up the reconstruction

observed all frequency slices

desired all frequency slices



Mao X, Li Q, Xie H, Lau RY, Wang Z, Paul Smolley S. Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision 2017, pages 2794-2802.
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5967–5976, July 2017.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Training framework: Generative Adversarial Network (GAN)

$$\min_{\theta} \mathop{\mathbb{E}}_{\mathbf{X} \sim p(\mathbf{X})} \left[\left(1 - \mathscr{D}_{\phi}\left(\mathscr{G}_{\theta}(\mathbf{M} \odot \mathbf{X})\right)\right)^2 + \lambda \left\| \mathscr{G}_{\theta}(\mathbf{M} \odot \mathbf{X}) - \mathbf{X} \right\|_1 \right],$$

$$\min_{\phi} \mathop{\mathbb{E}}_{\mathbf{X} \sim p(\mathbf{X})} \left[\left(\mathscr{D}_{\phi}\left(\mathscr{G}_{\theta}(\mathbf{M} \odot \mathbf{X})\right)\right)^2 + \left(1 - \mathscr{D}_{\phi}(\mathbf{X})\right)^2 \right],$$

$\mathbf{X}$: input ground truth drawn from the probability distributions $p(\mathbf{X})$

$\mathscr{G}_{\theta}$: generator $\mathscr{D}_{\phi}$: discriminator $\mathbf{M}$: training mask

λ : ensures that each realization of discriminator maps to a particular input, i.e., rather than solely fooling the discriminator.

Mao X, Li Q, Xie H, Lau RY, Wang Z, Paul Smolley S. Least squares generative adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision 2017, pages 2794-2802.
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5967–5976, July 2017.
Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Training framework: Generative Adversarial Network (GAN)

$$\min_{\theta} \mathop{\mathbb{E}}_{\mathbf{X}\sim p(\mathbf{X})} \left[\left(1-\mathscr{D}_{\phi}\left(\mathscr{G}_{\theta}(\mathbf{M}\odot\mathbf{X})\right)\right)^2 + \lambda \left\|\mathscr{G}_{\theta}(\mathbf{M}\odot\mathbf{X}) - \mathbf{X}\right\|_1\right],$$

$$\min_{\phi} \mathop{\mathbb{E}}_{\mathbf{X}\sim p(\mathbf{X})} \left[\left(\mathscr{D}_{\phi}\left(\mathscr{G}_{\theta}(\mathbf{M}\odot\mathbf{X})\right)\right)^2 + \left(1-\mathscr{D}_{\phi}(\mathbf{X})\right)^2\right],$$

Training pair: $(\mathbf{M}\odot\mathbf{X}, \mathbf{X})$

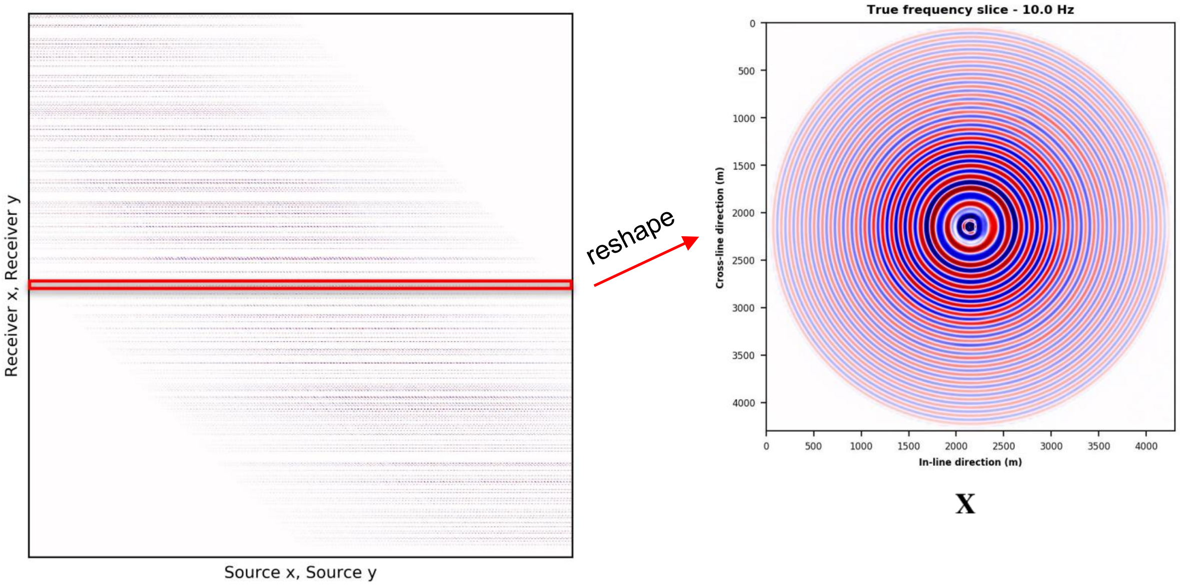
$\mathbf{X}$: input ground truth drawn from the probability distributions $p(\mathbf{X})$

$\mathscr{G}_{\theta}$: generator $\mathscr{D}_{\phi}$: discriminator $\mathbf{M}$: training mask

λ : ensures that each realization of discriminator maps to a particular input, i.e., rather than solely fooling the discriminator.

Extract training pairs for training

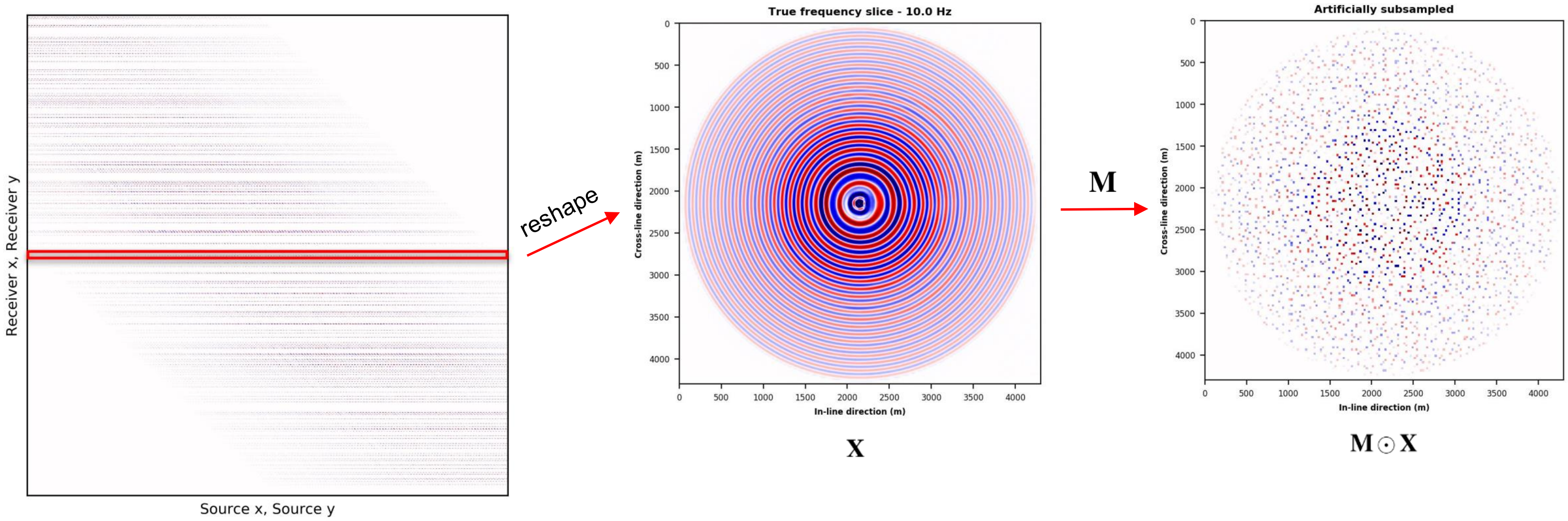
- (Rec x, Rec y) × (Src x, Src y) domain



X

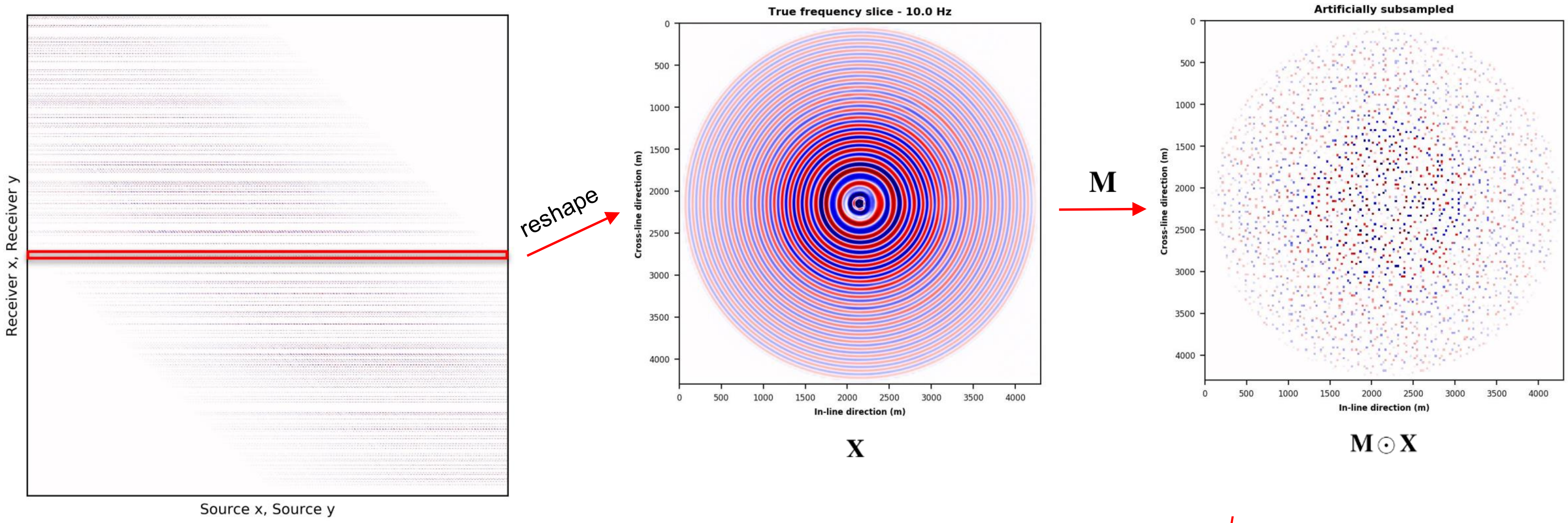
Extract training pairs for training

- (Rec x, Rec y) × (Src x, Src y) domain



Extract training pairs for training

- (Rec x, Rec y) × (Src x, Src y) domain



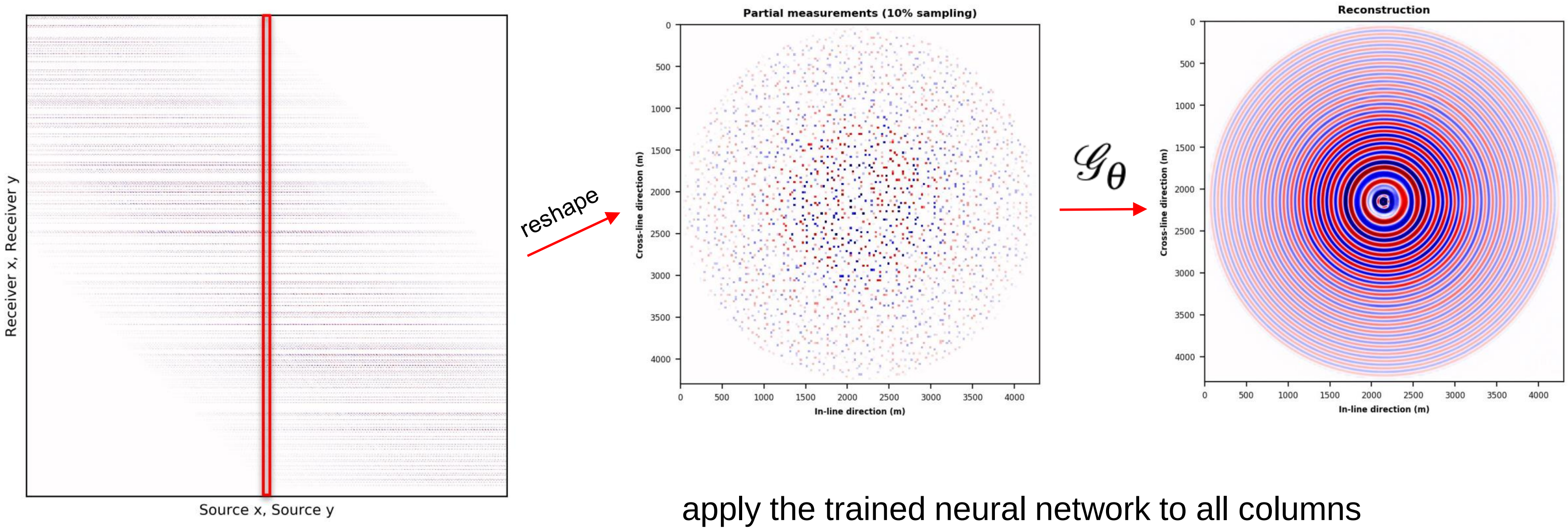
Training pair: $(\mathbf{M} \odot \mathbf{X}, \mathbf{X}) \longrightarrow \mathscr{G}_\theta$ $\surd$

Testing Stage: reconstruction

$\mathcal{G}_\theta$ √

- (Rec x, Rec y) × (Src x, Src y) domain

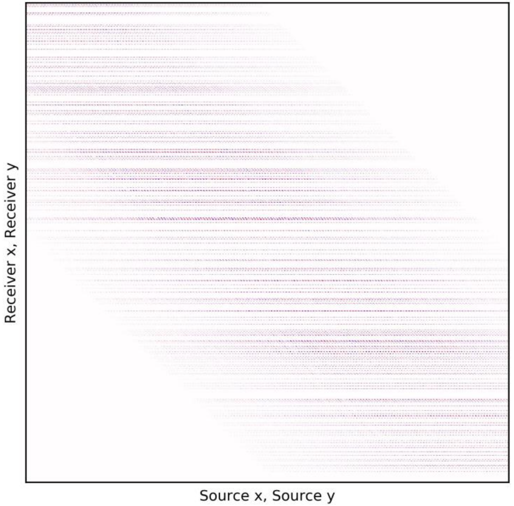
source-receiver reciprocity



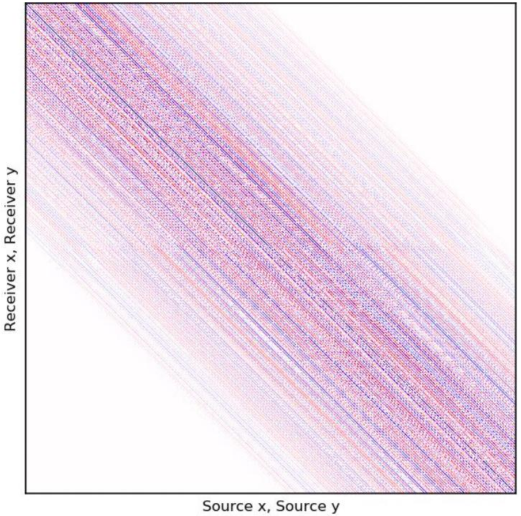
apply the trained neural network to all columns

The problem of the deep-learning-based method

- (Rec x, Rec y) × (Src x, Src y) domain



train from scratch



Problem: each frequency slice is treated independently → high training costs

Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.
Zhang, Y., S. Sharan, and F. J. Herrmann, 2019, High-frequency wavefield recovery with weighted matrix factorizations: SEG Technical Program Expanded Abstracts 2019.

SLIM

Transfer learning to speed up the reconstruction

Problem: each frequency slice is treated independently → high training costs

Solution: train from scratch

↓

exploit frequency-to-frequency similarities

Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.
Zhang, Y., S. Sharan, and F. J. Herrmann, 2019, High-frequency wavefield recovery with weighted matrix factorizations: SEG Technical Program Expanded Abstracts 2019.

SLIM

Transfer learning to speed up the reconstruction

Problem: each frequency slice is treated independently → high training costs

Solution: train from scratch

↓

exploit frequency-to-frequency similarities

↓

train from transferred weights from neighboring frequencies
(do recursively for all frequencies from low to high)

Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.
Zhang, Y., S. Sharan, and F. J. Herrmann, 2019, High-frequency wavefield recovery with weighted matrix factorizations: SEG Technical Program Expanded Abstracts 2019.

Transfer learning to speed up the reconstruction

Problem: each frequency slice is treated independently → high training costs

Solution: train from scratch

↓

exploit frequency-to-frequency similarities

↓

train from transferred weights from neighboring frequencies (do recursively for all frequencies from low to high) → lower training costs

SLIM

Numerical experiments

Numerically simulated data on 3D BG Compass model

- ► 172×172 2D periodic source grid
- ► 172×172 2D periodic receiver grid
- ► spatial subsampling in both horizontal directions
- ► complex velocity model w/ strong vertical & lateral variations

Numerical experiments

We know the missing pattern of receivers

Training mask:

- **previous:** training mask equal to the missing receiver sampling mask

Siahkoohi, A., R. Kumar, and F. J. Herrmann, 2019, Deep-learning based ocean bottom seismic wave field recovery: SEG Technical Program Expanded Abstracts 2019.

Numerical experiments

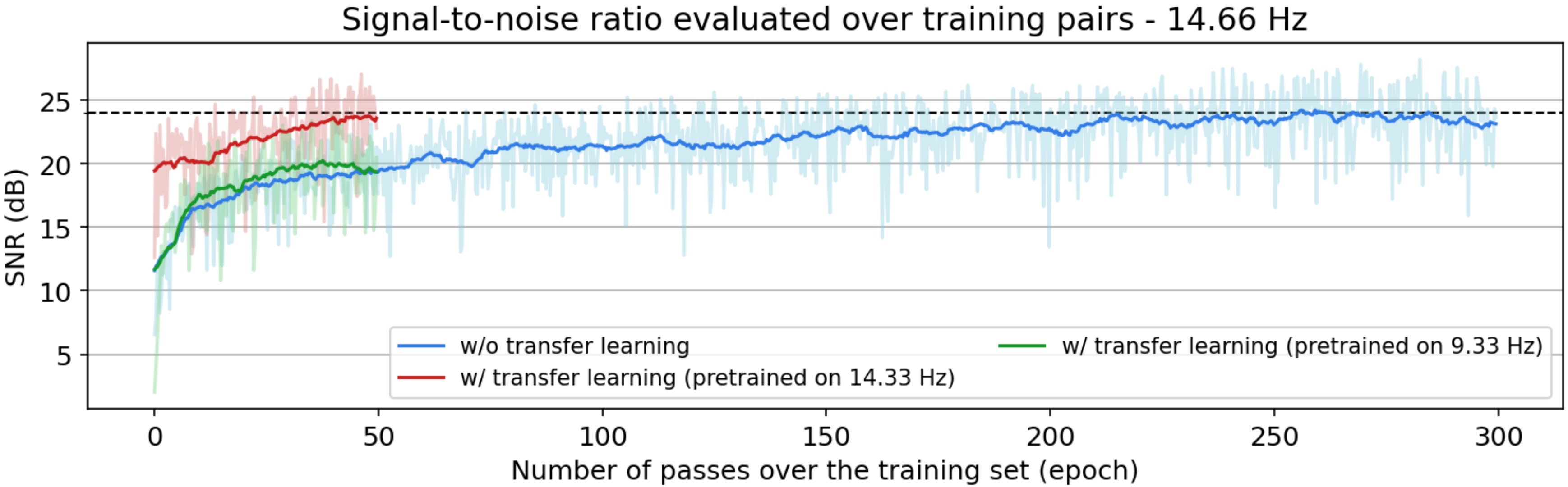
We know the missing pattern of receivers

Training mask:

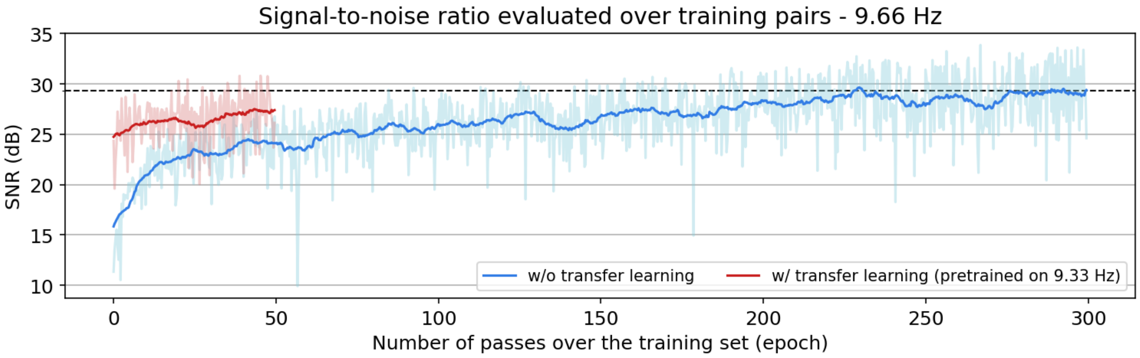
- **previous:** training mask equal to the missing receiver sampling mask
- **now:** change the training mask (with a same missing rate) at every epoch

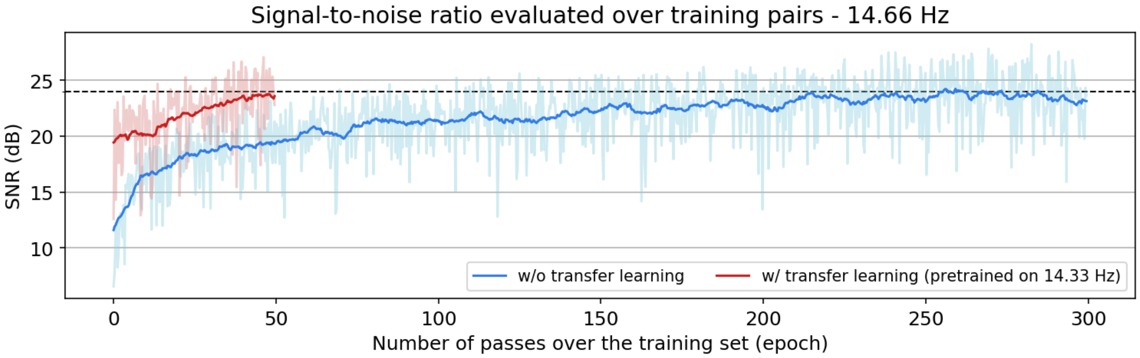
Similarity: neighboring frequency slices > non-neighboring frequency slices

(14.33 Hz -> 14.66 Hz) (9.33 Hz -> 14.66 Hz)



Similarity: (14.33 Hz -> 14.66 Hz) > (9.33 Hz -> 9.66 Hz)

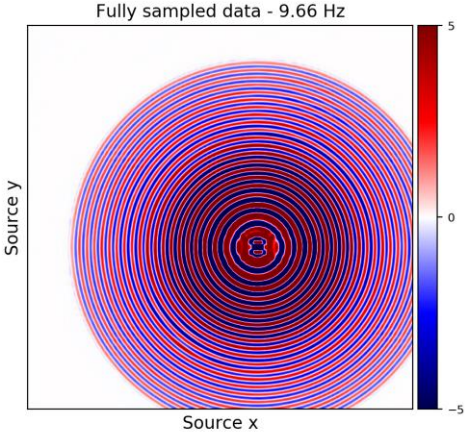


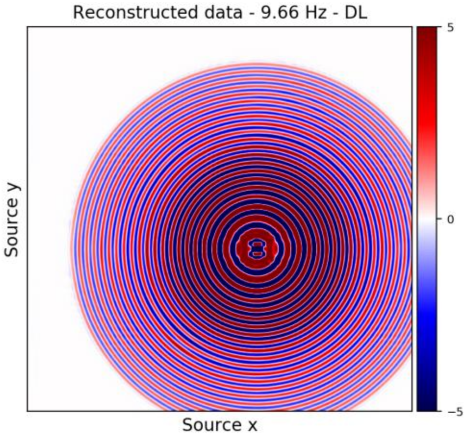


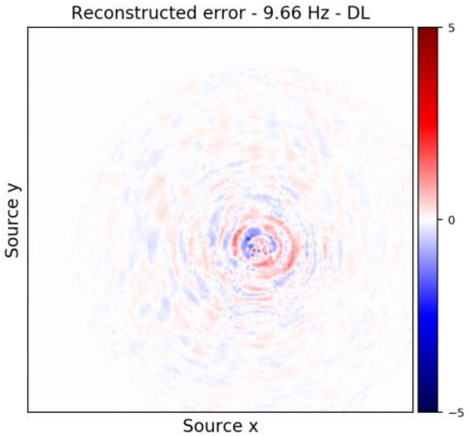
~ six fold speedup

Reconstruction performance

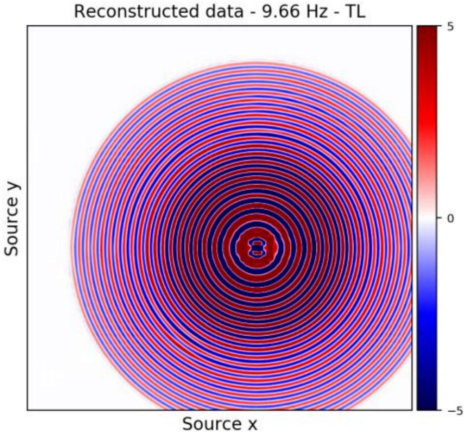
Frequency: 9.66 Hz
Neighboring: 9.33 Hz

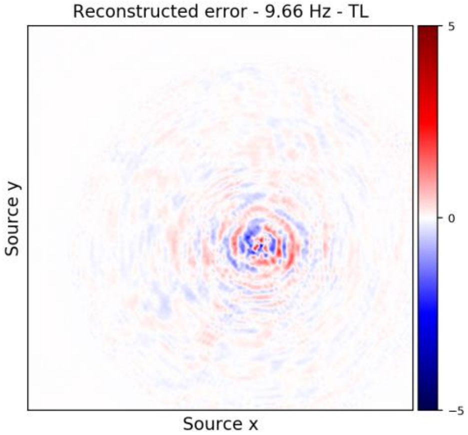






Training from scratch: **300 epochs**



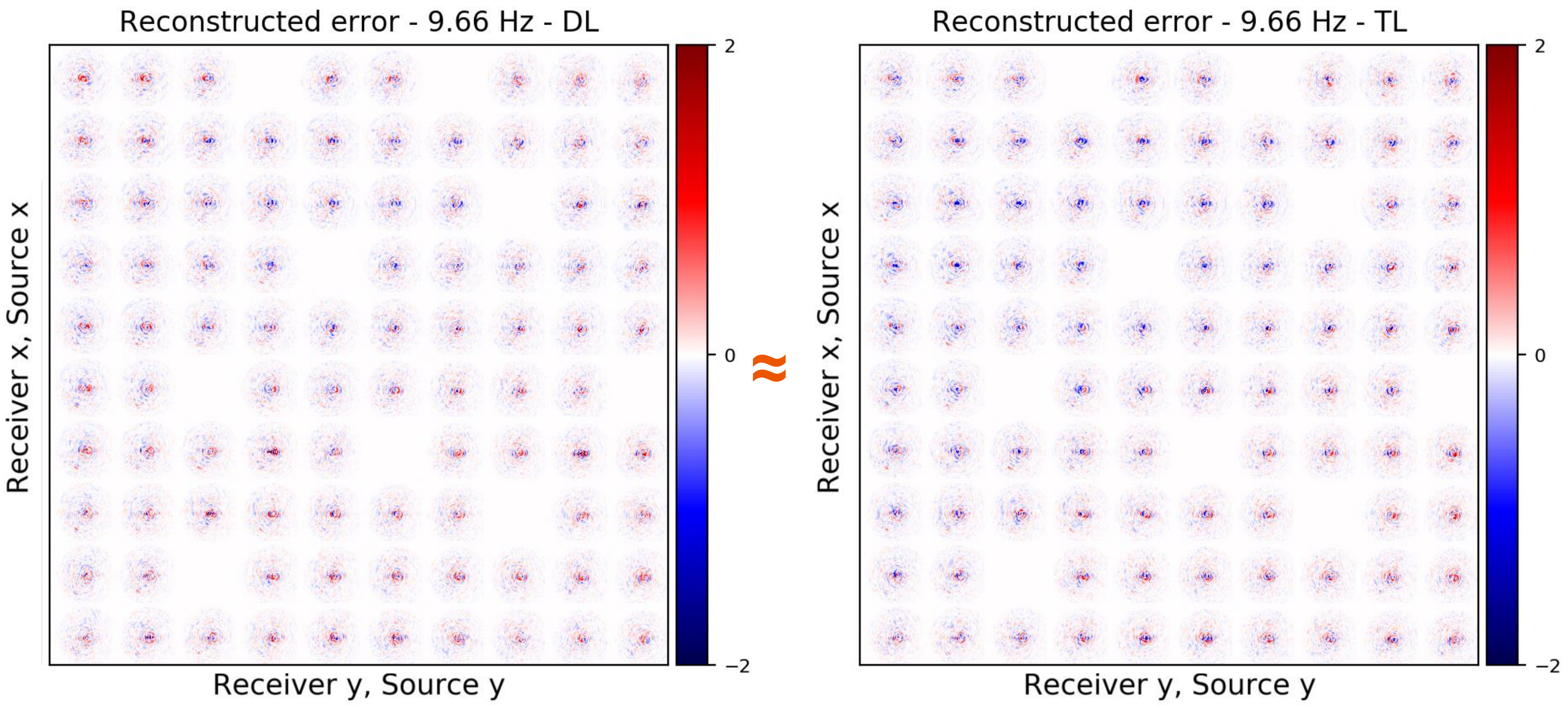


Transfer learning from 9.33 Hz: **50 epochs**

Reconstruction performance

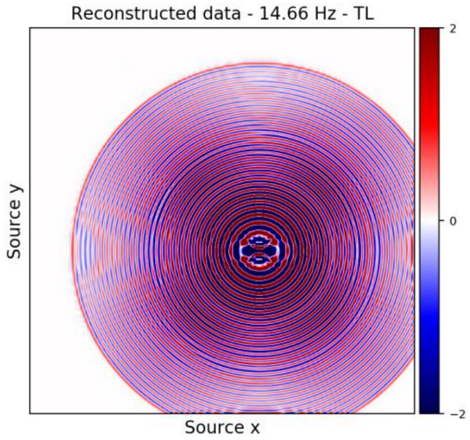
- (Rec y, Src y) × (Rec x, Src x) domain (non-canonical)

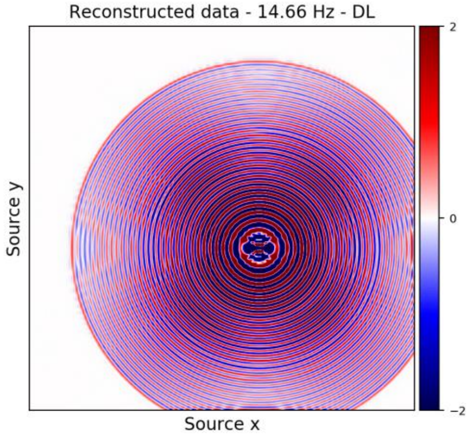
Training from scratch: **300 epochs** **>** Transfer learning from 9.33 Hz: **50 epochs**

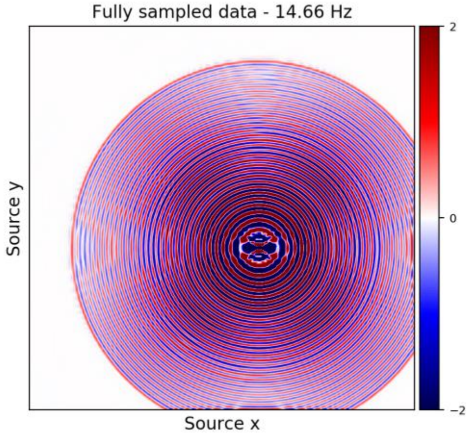


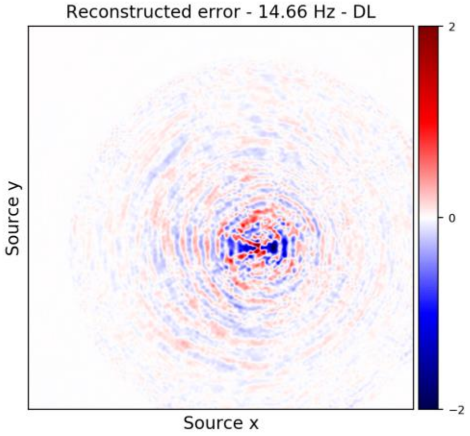
Reconstruction performance

Frequency: 14.66 Hz
Neighboring: 14.33 Hz



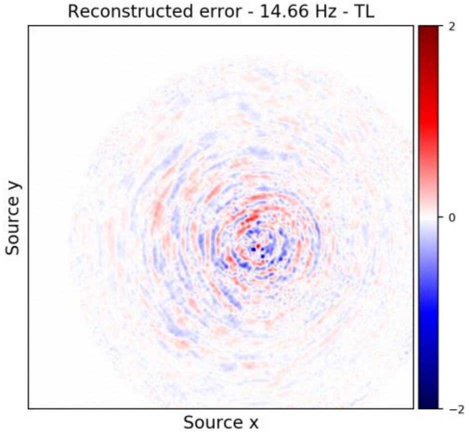






Training from scratch: **300 epochs**

SNR: 17.18 dB



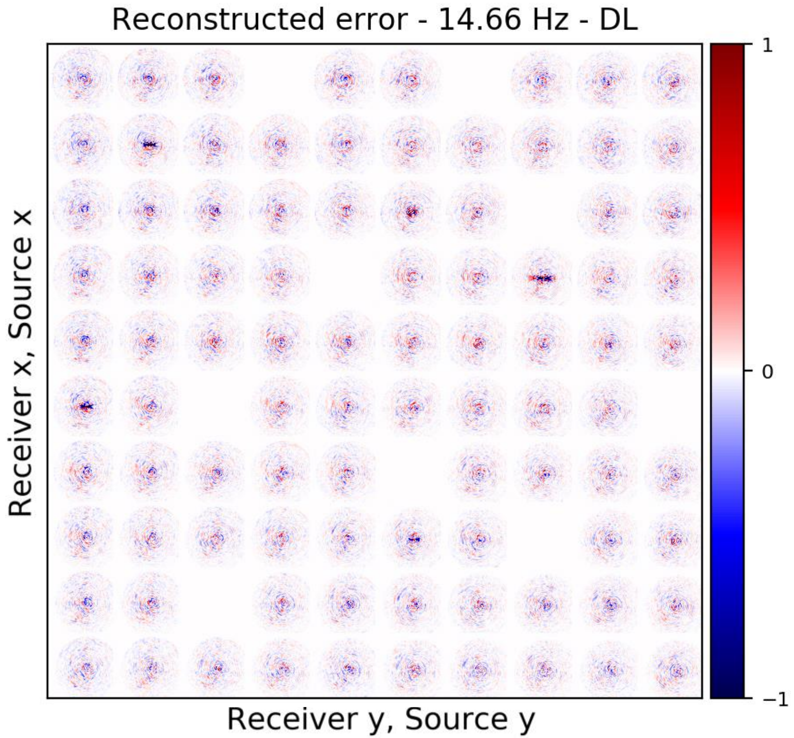
Transfer learning from 14.33 Hz: **50 epochs**

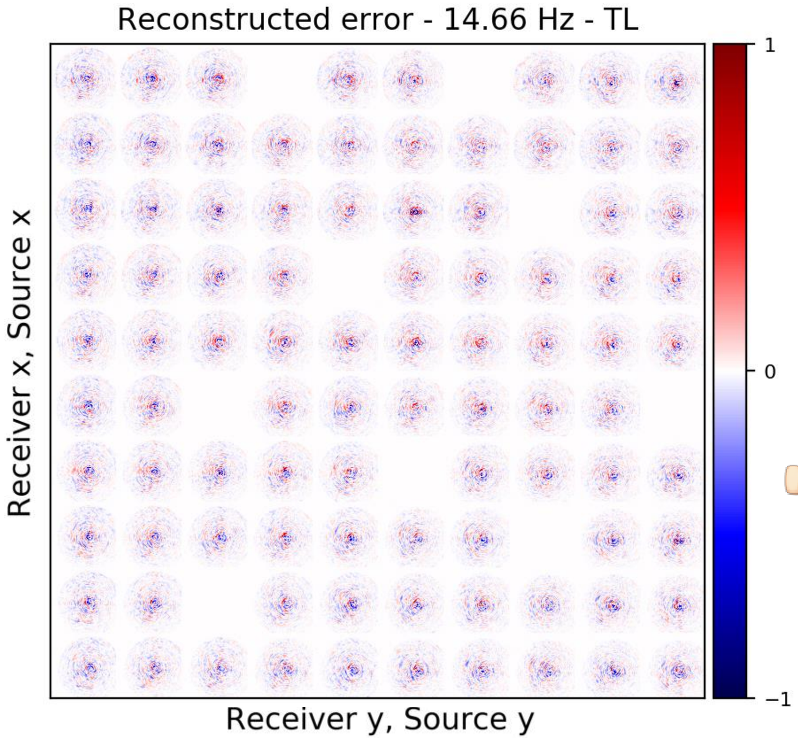
SNR: 18.09 dB

Reconstruction performance

- (Rec y, Src y) × (Rec x, Src x) domain (non-canonical)

Training from scratch: **300 epochs** **>** Transfer learning from 14.33 Hz: **50 epochs**





Conclusions

The method does not need any external training data, assuming:

- desirable source sampling
- source-receiver reciprocity

Transfer learning (recursively for all frequencies) can significantly **speed up** in the training, specially at consecutive frequency slices with relatively **high correlation**.

Conclusions

The method does not need any external training data, assuming:

- desirable source sampling
- source-receiver reciprocity

Transfer learning (recursively for all frequencies) can significantly **speed up** in the training, specially at consecutive frequency slices with relatively **high correlation.**

Future work: improve the reconstruction accuracy of high-frequency slices

Thank you for your attention!

GitHub https://github.com/slimgroup/Software.SEG2020