Preconditioning Seismic Inverse Problems with AutoEncoders

M. Ravasi

Outline

- A refresher on Deep Preconditioners
- Strategies for robust training
- Wavefield separation example
- Field data
- Conclusions

Outline

- **A refresher on Deep Preconditioners**
- Strategies for robust training
- Wavefield separation example
- Field data
- Conclusions

A refresher on Deep Preconditioners

Regularization

$$J = \left\| y - Gx \right\|_2 + \varepsilon \left\| Dx \right\|_p$$

D :

- Identity (aka Tikhonov)
- First/second derivative
- Sparsity transforms

Preconditioning

$$J = \left\| y - GPz \right\|_2 + \varepsilon \left\| z \right\|_p$$

P :

- Smoothing
- Sparsity transforms
- PCA (aka dim reduction)

A refresher on Deep Preconditioners

Deep Regularization – Obman et al., 2020

$$J = \|y - Gx\|_2 + \varepsilon \|x - d_\theta(e_\theta(x))\|_p$$

Latent space 'discovery'

$$J_{AE} = \|x - d_\theta(e_\theta(x))\|_p$$

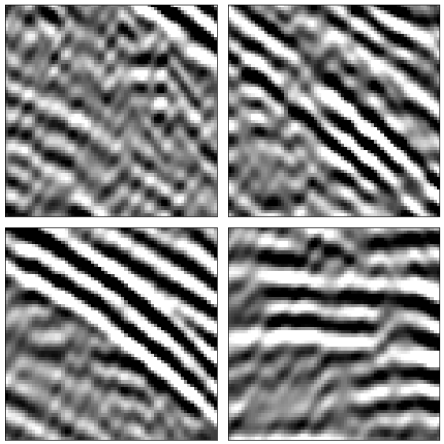
Deep Preconditioning – Ravasi, 2021

$$J = \|y - Gd_\theta(z)\|_2 + \varepsilon \|z\|_p$$

A refresher on Deep Preconditioners

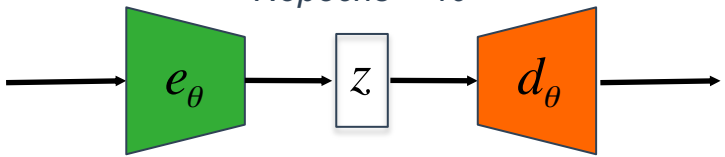
Training

Input



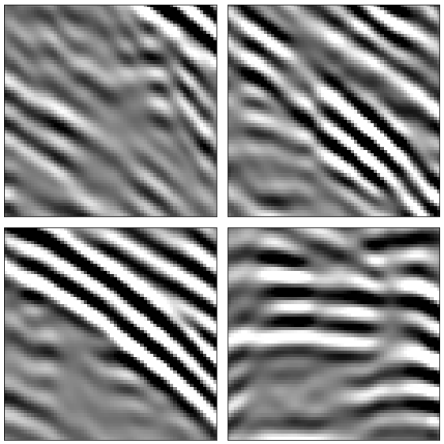
3 level CCN Enc/Dec

Nepochs = 40



$$J_{AE} = \|d_\theta(e_\theta(x)) - x\|_2$$

Prediction

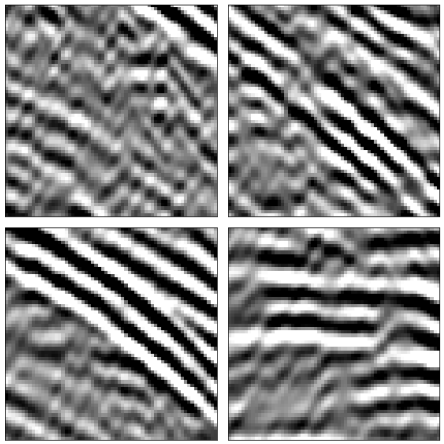


Valid MSE: 0.63

A refresher on Deep Preconditioners

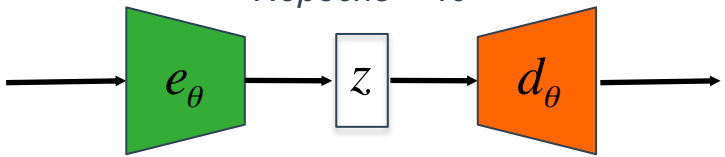
Training

Input



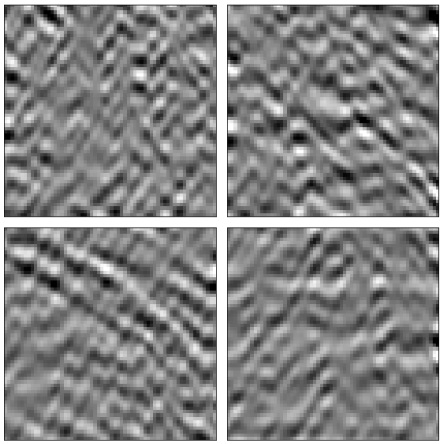
3 level CCN Enc/Dec

Nepochs = 40



$$J_{AE} = \|d_\theta(e_\theta(x)) - x\|_2$$

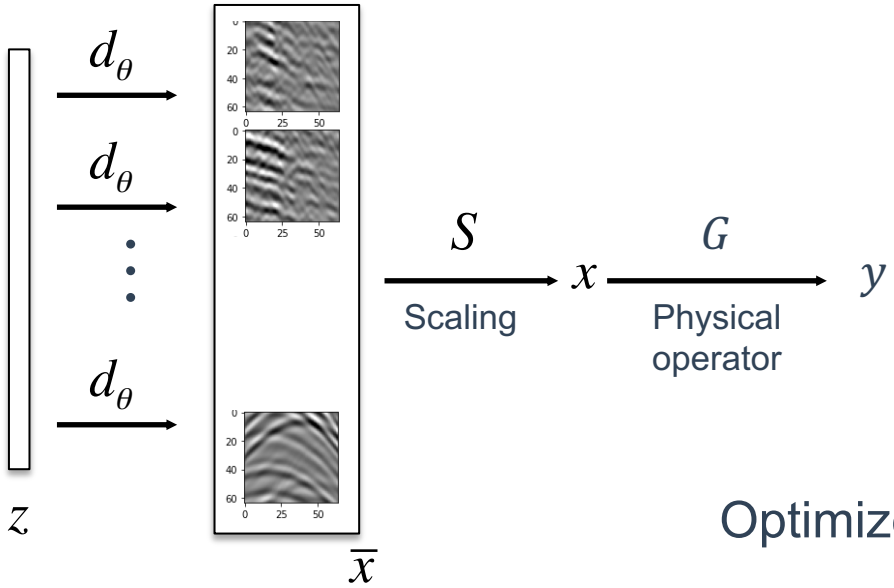
Error



Valid MSE: 0.63

A refresher on Deep Preconditioners

Inversion



Optimize (e.g., L-BFGS):

$$J = \|y - GSd_\theta(z)\|_2 + \varepsilon\|z\|_p$$

A refresher on Deep Preconditioners

Inversion

General: $y = Gx$

Deghosting: $p = R(I + \Phi)p^-$

Initialization of z is crucial:

$z_0 = e_\theta(S^{-1}(I + \Phi)^H R^H y)$

Wavefield separation: $\begin{bmatrix} p \\ v_z \end{bmatrix} = \begin{bmatrix} I & I \\ W^+ & W^- \end{bmatrix} \begin{bmatrix} p^+ \\ p^- \end{bmatrix}$

$d = RWp^{\pm}$

Initialization of z is crucial:

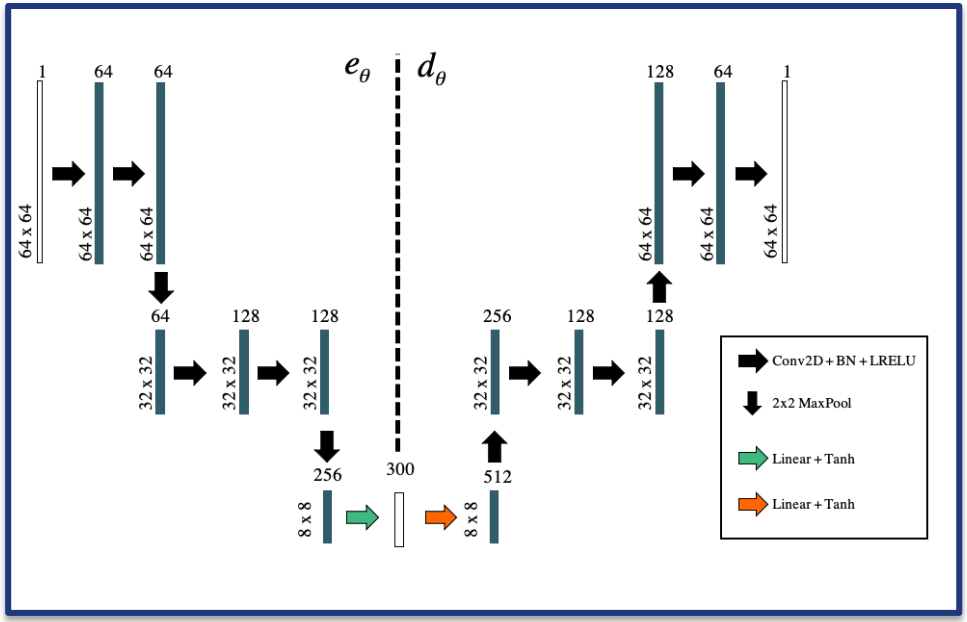
$z_0^+ = e_\theta(S^{-1}(p + v_z))$

$z_0^- = e_\theta(S^{-1}(p - v_z))$

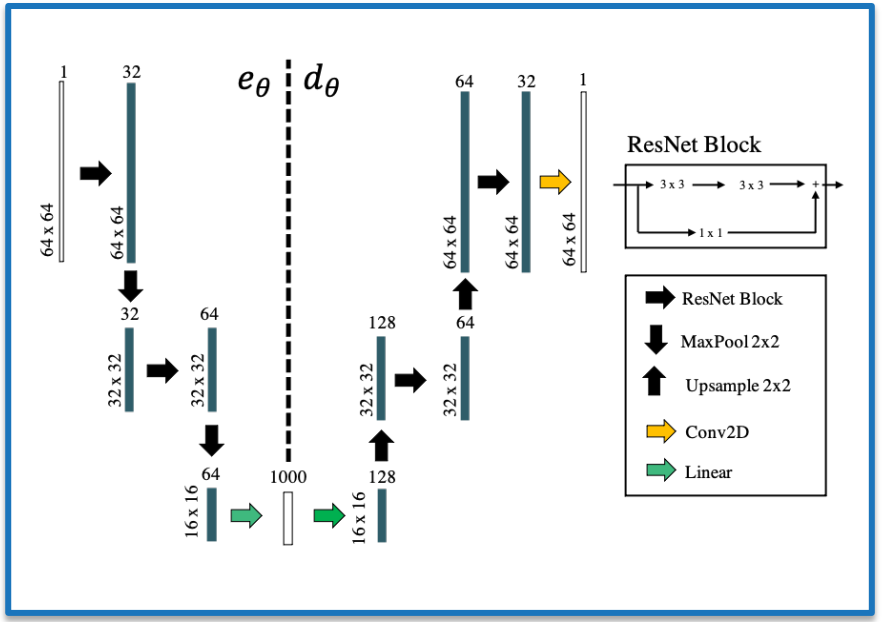
Outline

- A refresher on Deep Preconditioners
- **Strategies for robust training**
- Wavefield separation example
- Field data
- Conclusions

Strategies for robust training - Architecture



Convolutional Network

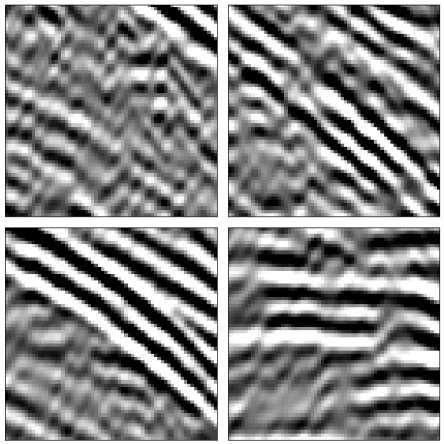


ResNet Network

Strategies for robust training - Architecture

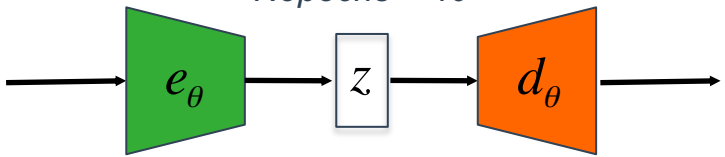
Training

Input



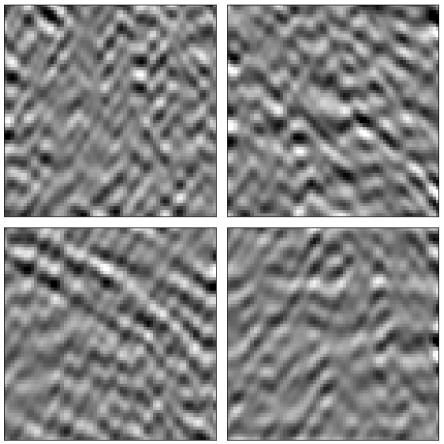
3 level CCN Enc/Dec

Nepochs = 40



$$J_{AE} = \|d_\theta(e_\theta(x)) - x\|_2$$

Error

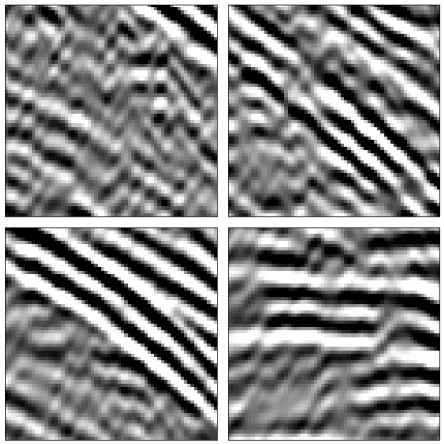


Valid MSE: 0.63

Strategies for robust training - Architecture

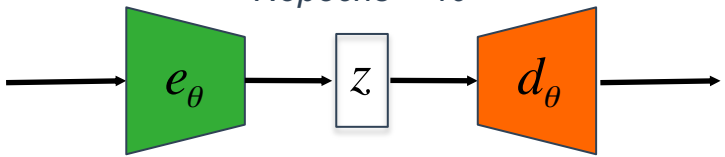
Training

Input



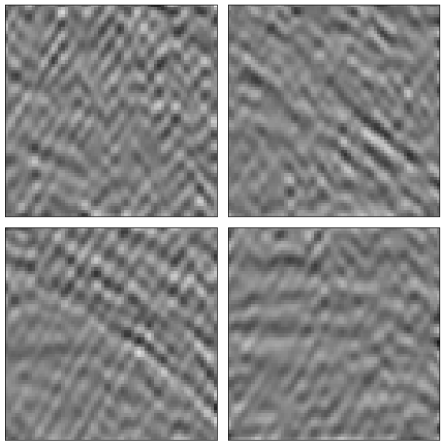
3 level CCN
ResNet Enc/Dec

Nepochs = 40



$$J_{AE} = \|d_\theta(e_\theta(x)) - x\|_2$$

Error



Valid MSE: 0.34

Strategies for robust training – Multi-term loss

Learnable weights

Compensation term (to avoid learning $\sigma_1 = \sigma_2 = 0$)

$$J(\theta, \sigma_1, \sigma_2) = \frac{1}{2\sigma_1^2}\mathcal{L}_1(\theta) + \frac{1}{2\sigma_2^2}\mathcal{L}_2(\theta) + \log(\sigma_1\sigma_2)$$

Multiple losses focusing on different properties of the data, *eg:*

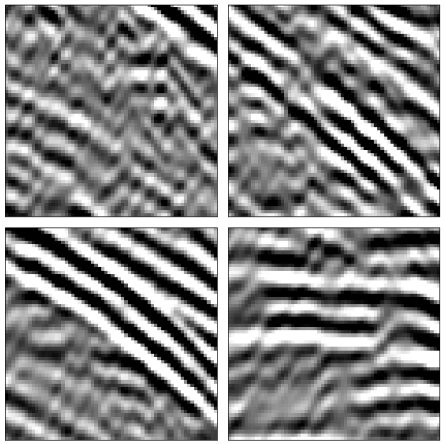
- ***Local**: MSE, MAE* - ***Semi-local**: SSIM* - ***Global**: R^2, CCC*

Inspired by: *Multi-Task Learning Using Uncertainty to Weight Losses, Kendall et al 2017.*

Strategies for robust training – Multi-term loss

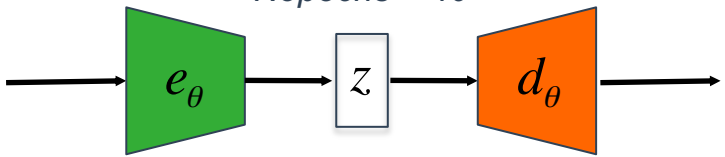
Training

Input



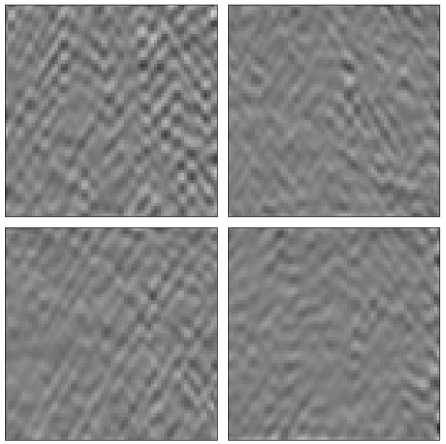
3 level CCN
ResNet Enc/Dec

Nepochs = 40



$$J_{AE} = \frac{1}{2\sigma_1^2}\|d_\theta(e_\theta(x)) - x\|_2$$
$$+ \quad \frac{1}{2\sigma_2^2} R^2\left(d_\theta\left(e_\theta(x)\right), x\right)$$
$$+ \quad \log(\sigma_1\sigma_2)$$

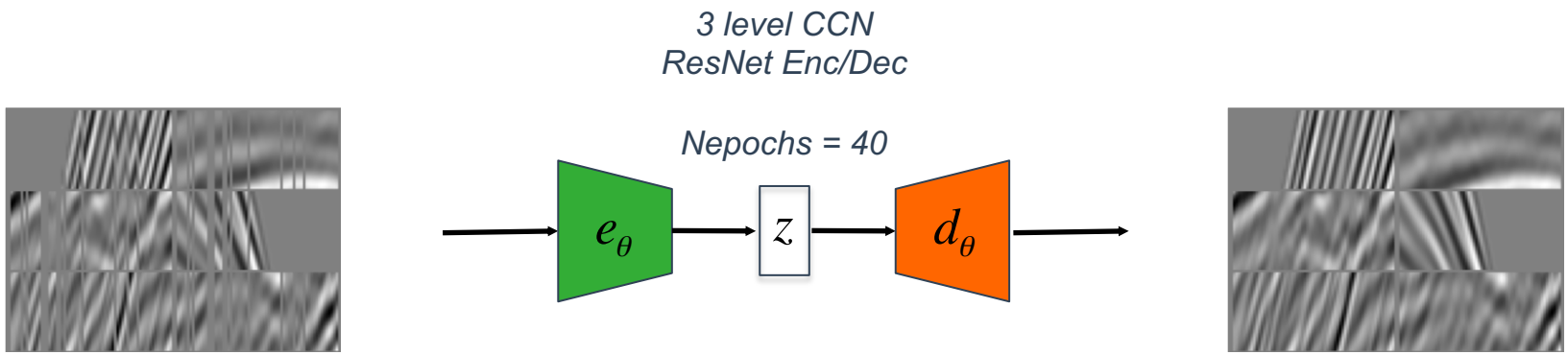
Error



Valid MSE: 0.22

Strategies for robust training – Gaps in inputs

Training



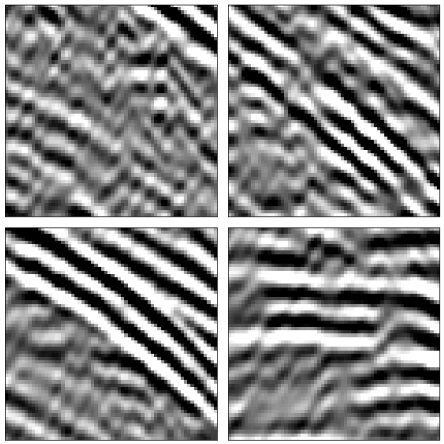
Randomly masking inputs: ***regularizer*** *(similar to noise injection)*

*When data have missing traces, they are naturally filled in (****AE act as denoisers****)*

Strategies for robust training – Gaps in inputs

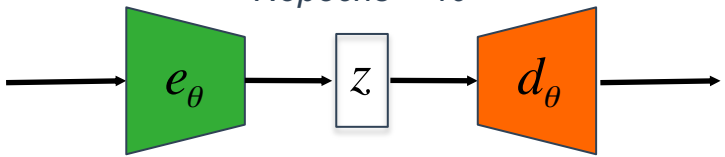
Training

Input

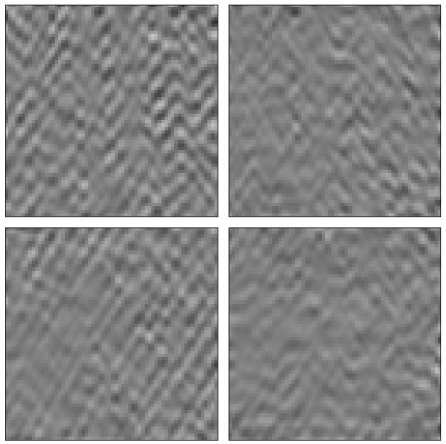


3 level CCN
ResNet Enc/Dec

Nepochs = 40



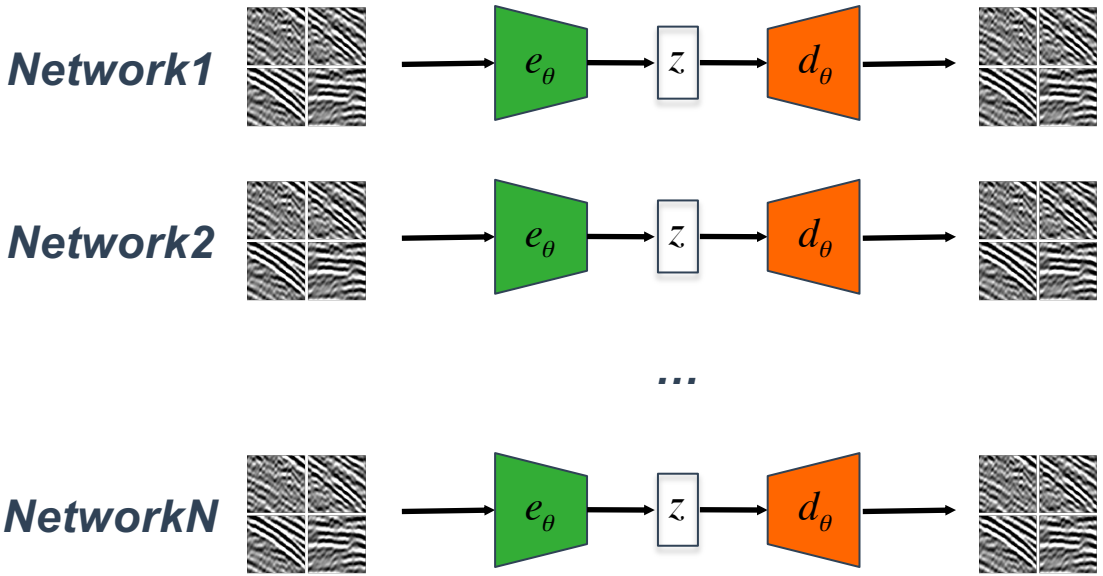
Error



Valid MSE: 0.24

Strategies for robust training – Ensembling

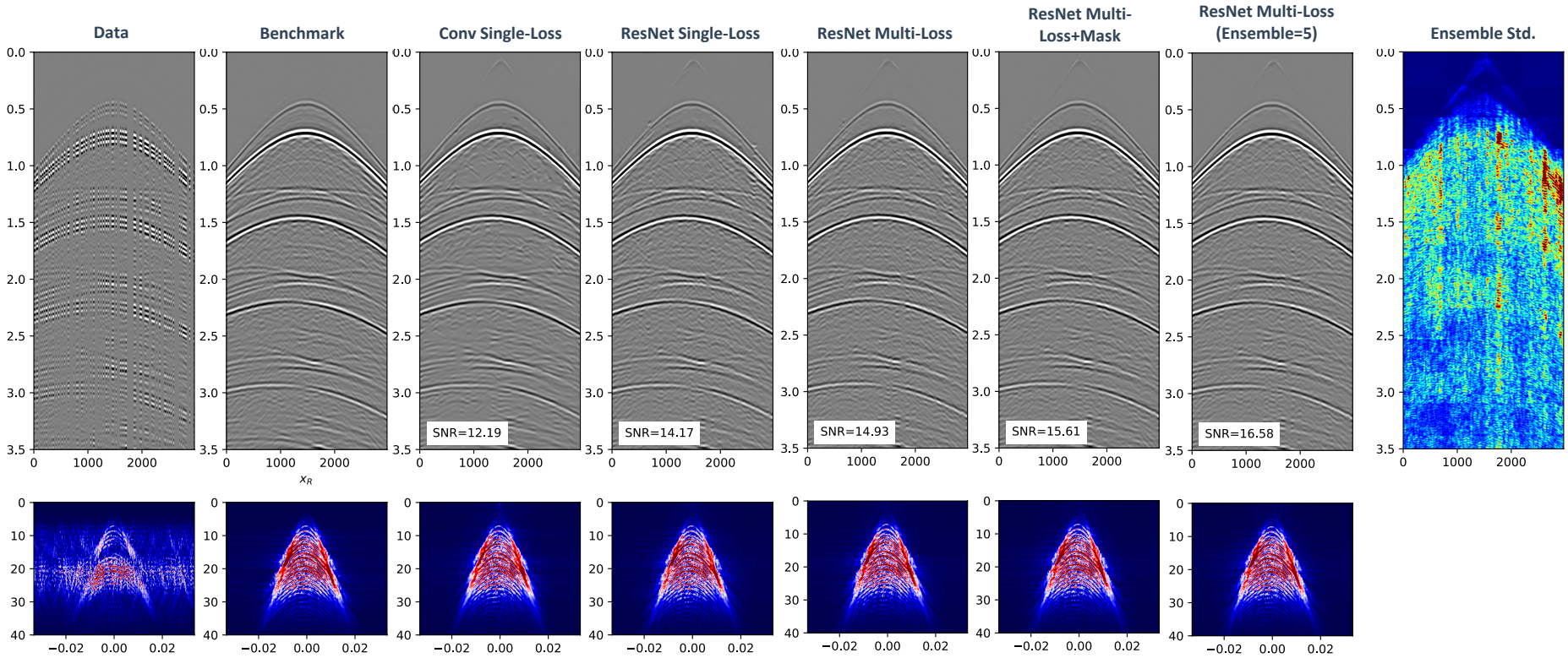
Training



… with different weight initialization

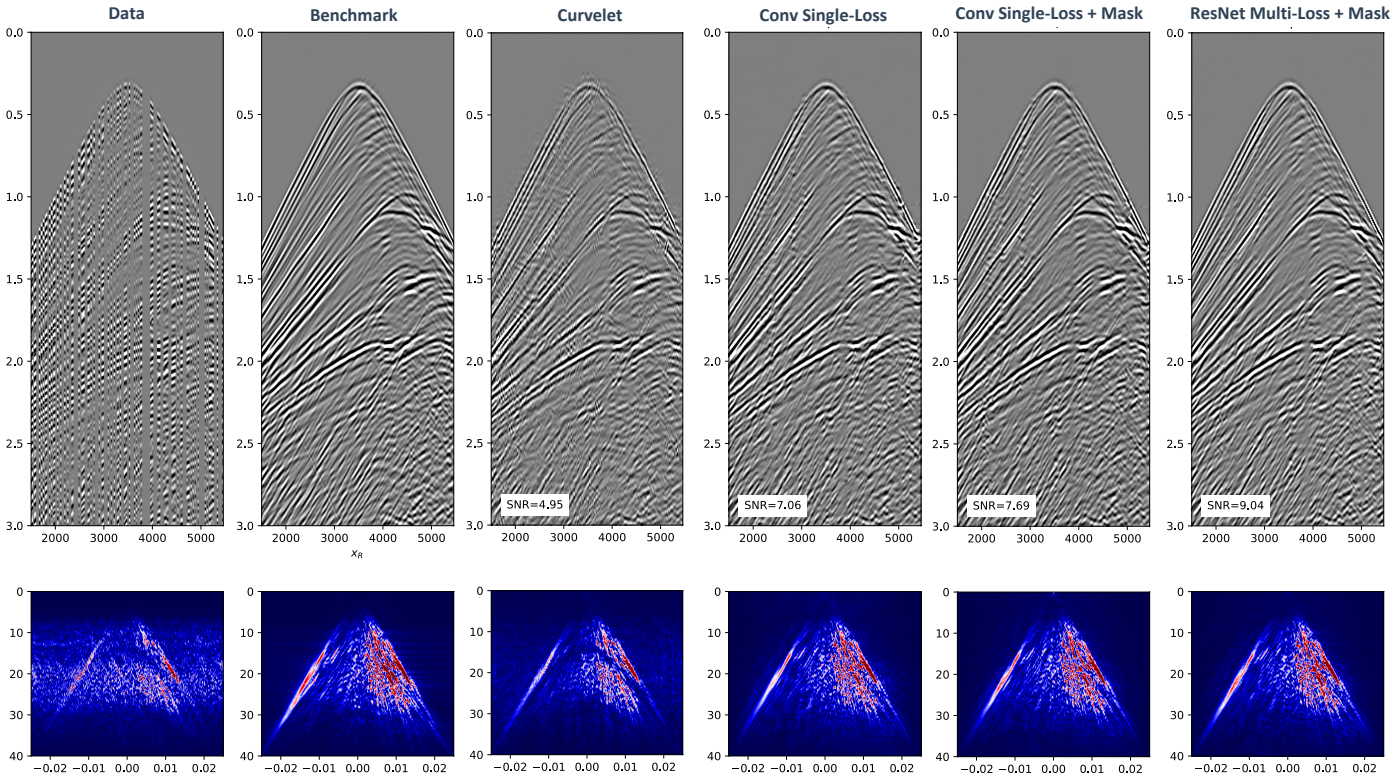
Strategies for robust training – Synthetic (30% irreg)

Inversion



Strategies for robust training – Marmousi (40% irreg)

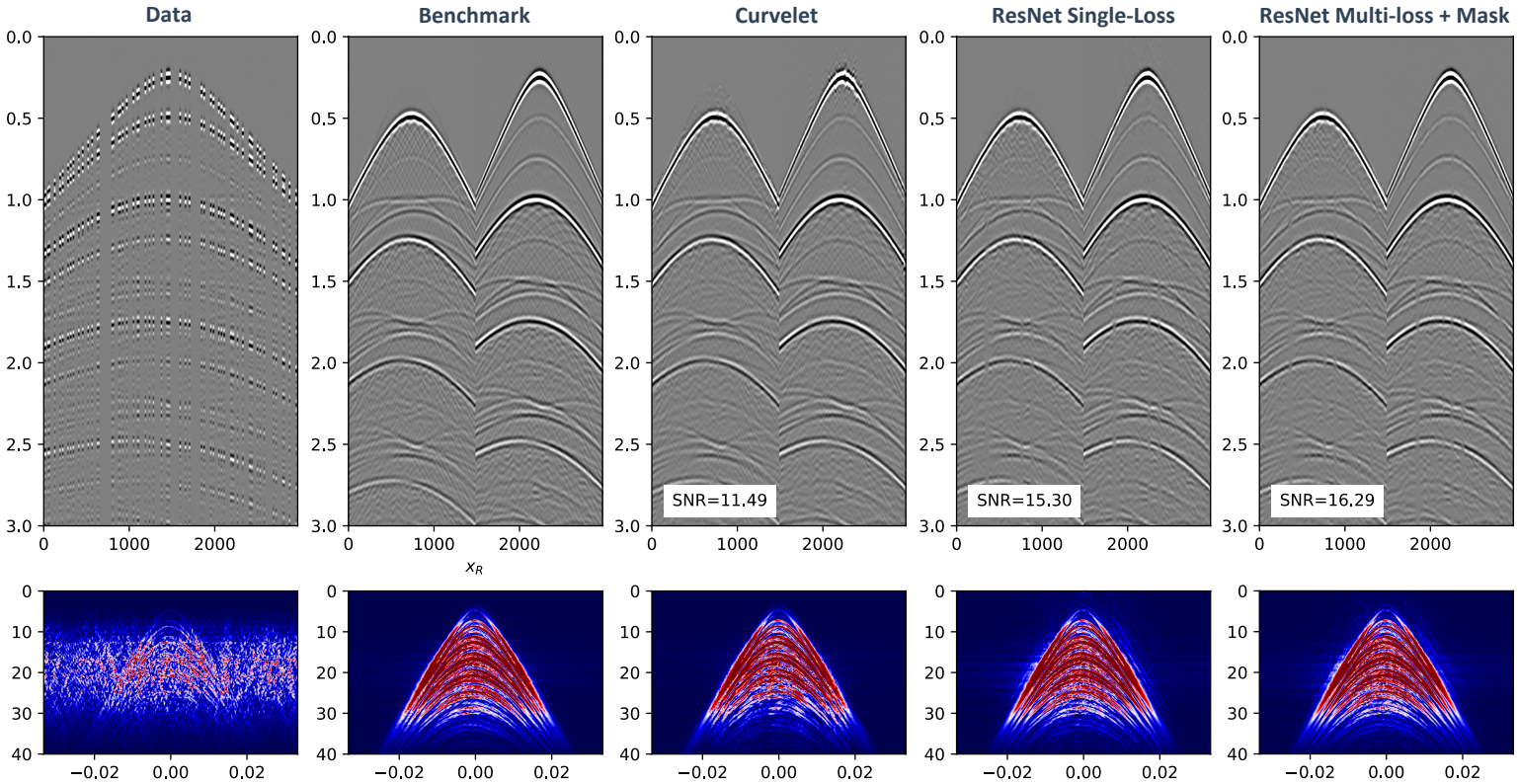
Inversion



Outline

- A refresher on Deep Preconditioners
- Strategies for robust training
- **Wavefield separation example**
- Field data
- Conclusions

Wavefield separation example

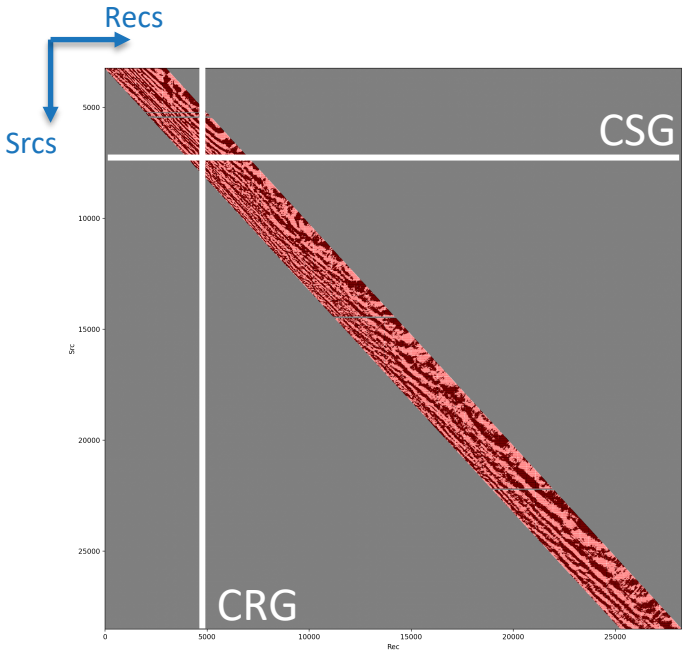


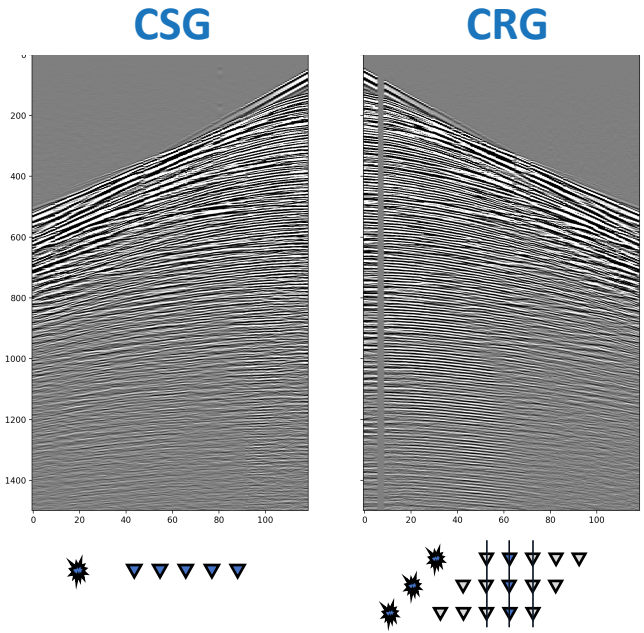
Outline

- A refresher on Deep Preconditioners
- Strategies for robust training
- Wavefield separation example
- **Field data**
- Conclusions

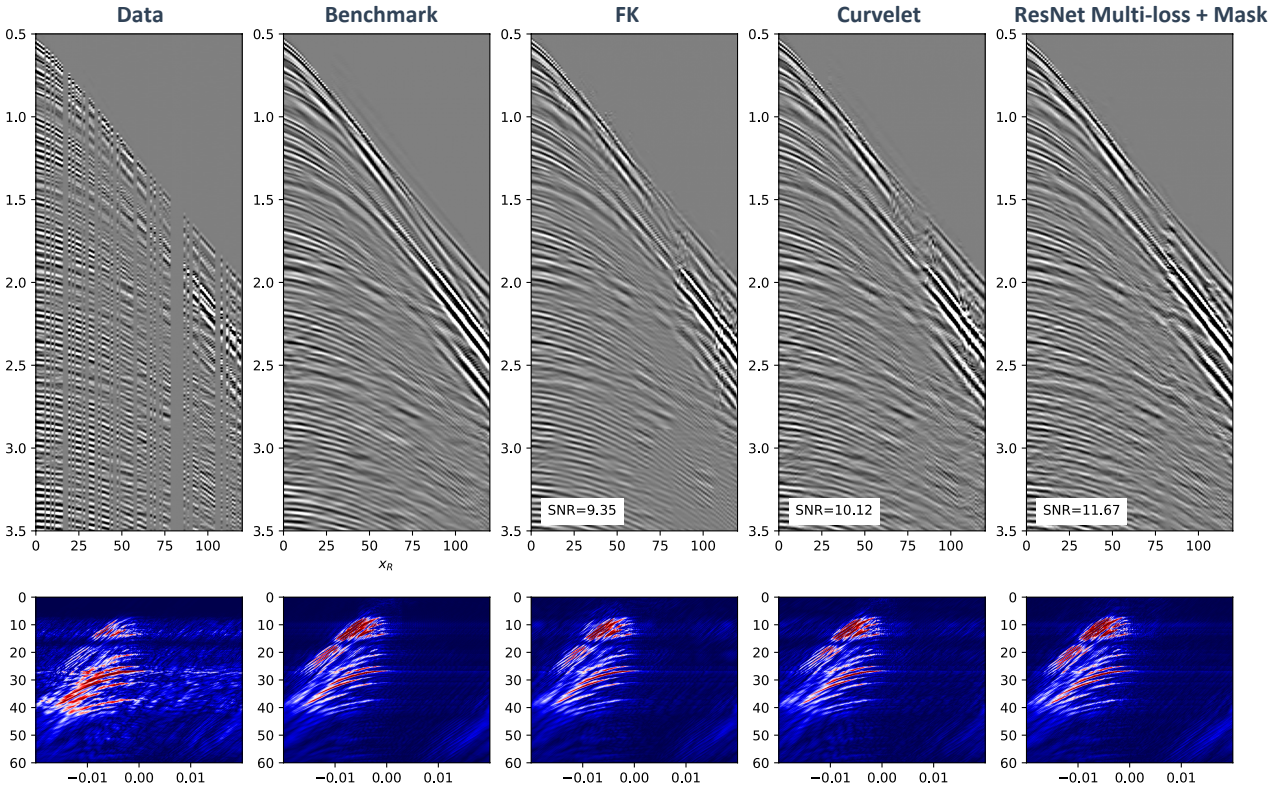
Field data – preliminary results

Mobil AVO viking graben line 12 – streamer data

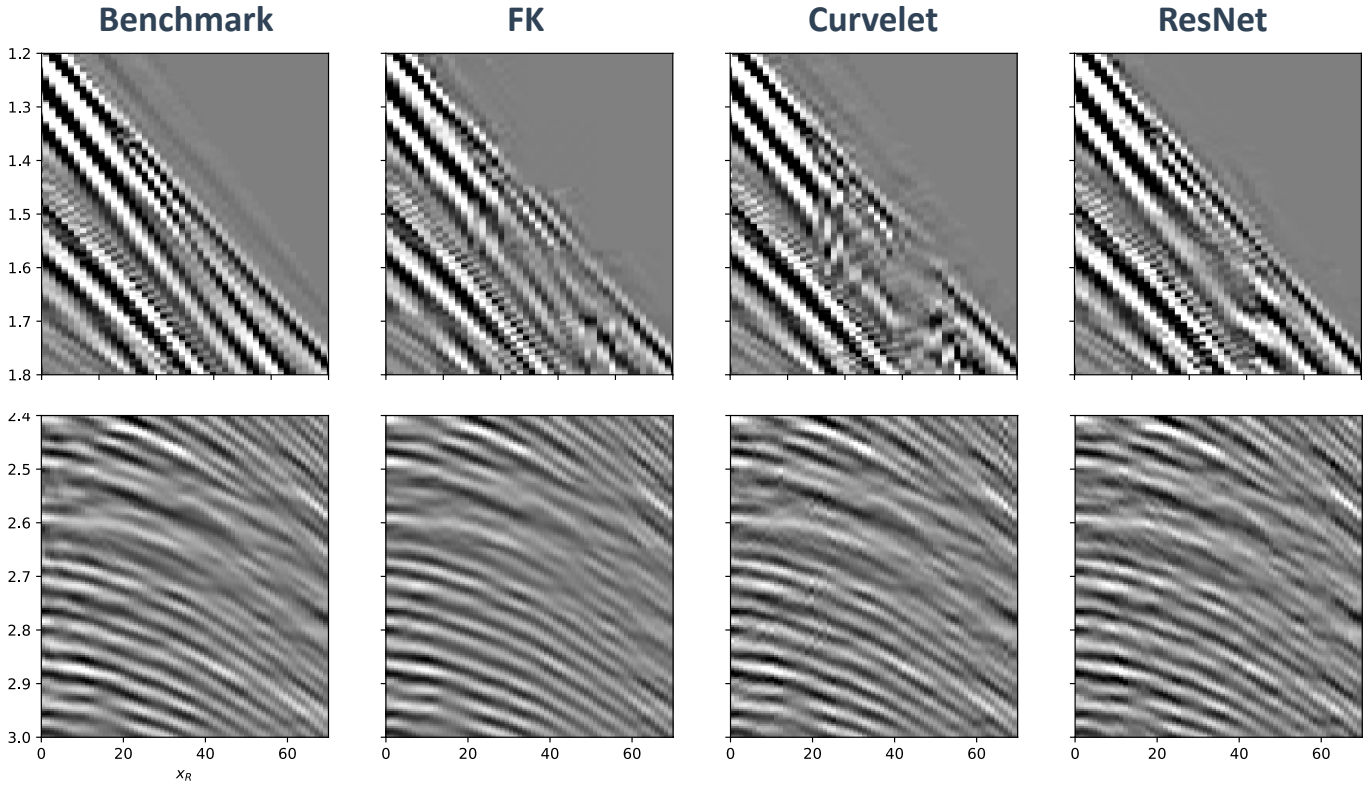




Field data – preliminary results



Field data – preliminary results



Conclusions

Deep Preconditioners *are an effective tool for ill-posed (geophysical) inverse problems*

✓ No need for training samples (beyond the data itself)

✗ Skip connections cannot be used → **robust training strategies**

? Can we push the AE to naturally fill gaps *(streamer case)*

Thank you for listening!

01. 12. 2021