Centre for Oil and Gas - DTU
The Danish Hydrocarbon Research and Technology Centre

Information Theory Considerations in Patch-Based Training of Deep Neural Networks on Seismic Time-Series

J.S. Dramsch[1]*, M. Lüthje[1]
[1] Technical University Of Denmark

Summary
Recent advances in machine learning rely on convolutional deep neural networks. These are commonly trained on cropped image patches. Pertaining to non-stationary seismic signals this may introduce low frequency noise and non-generalizability. While the Nyquist-Shannon sampling theorem applies to high frequency aliasing, we show that real-valued operations on seismic data can introduce a DC- or low frequency aliasing.

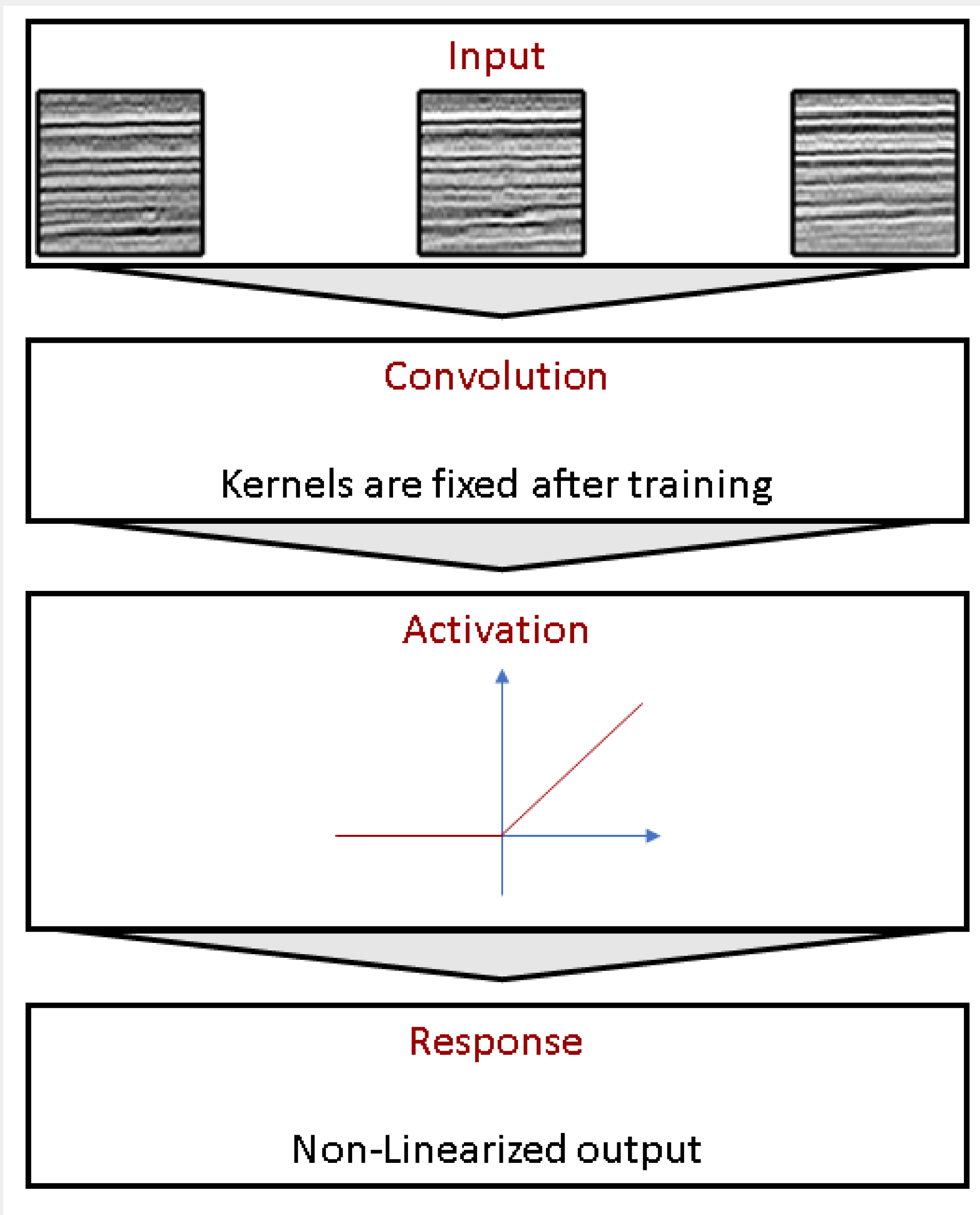
Sensitive Activations

Neurons in neural networks are described by the activation $\sigma(w \cdot x + b)$, where w is the network weights, x is the input data, b is the network bias, and sigma is a non-linear activation function. A common non-linear activation is the rectified linear unit (RELU) $\sigma(x) = max(0, x)$. Considering the inference stage, the network weights w and biases b are fixed, x is the only variable parameter.

Learning on a mean-shift of q of an arbitrary distribution over x leads to $\sigma(w \cdot (x + q) + b)$, which changes the neuron response by q, weighted by w.

Convolutions that take phase information into account would be the cleanest solution. This could be accomplished with complex convolutions.

In modern neural networks, batch normalization is often used, it shifts the mean of the activation response to zero and unit variance. This does not correct for the aliasing around the zero frequency but fixes the mean shift problem.



We observed that image patches of seismic data vary in lightness. We investigated, whether windowing and subsequent real-valued convolution results in different learnt distributions. This is known as mean shift in ML.

Equivalent Concepts

- Frequency: DC offset
- Image: Alpha Value
- ML: Mean Shift

Proposed Solutions

- Complex-valued Convolution
- Batch Normalization

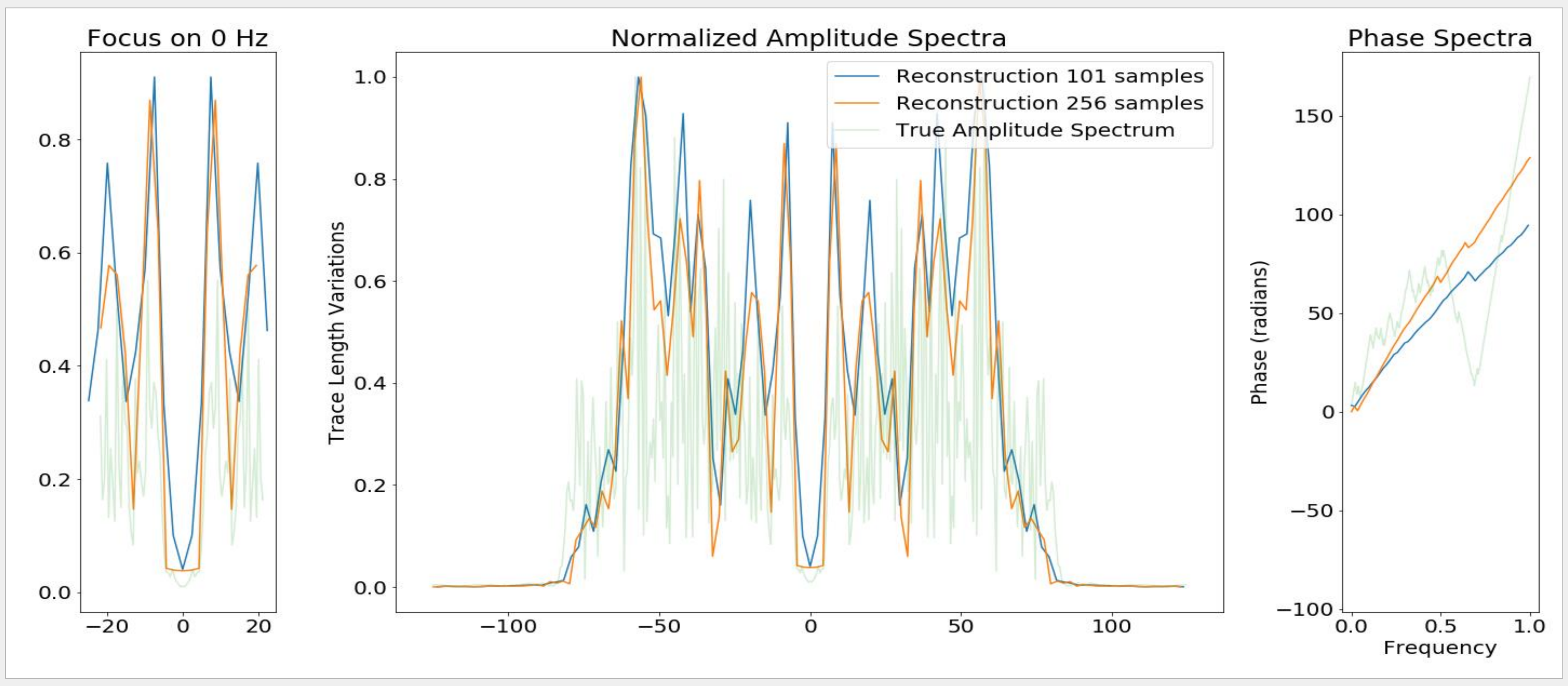


Figure 2:Effect of window size on low-frequency aliasing in a "problematic" area.
We observe that the aliasing effect of smaller window-sizes is stronger for real-valued operations. The phase spectra are visibly different.

PES GB

*

EAGE