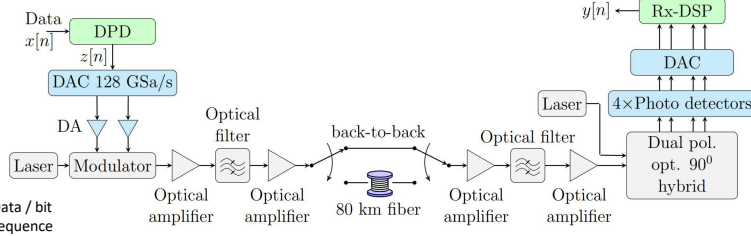
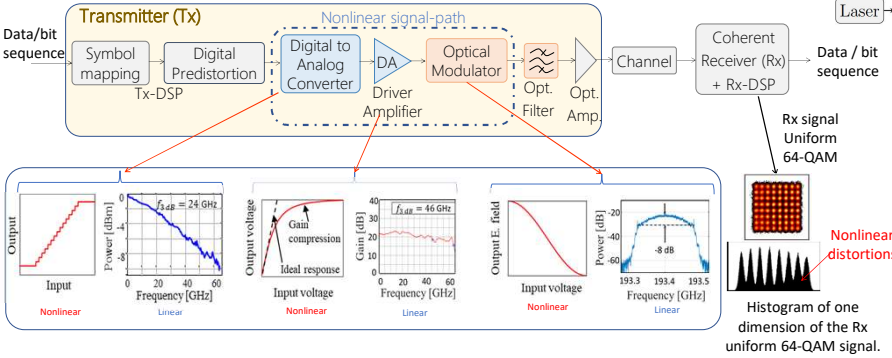


Motivation:

- Optical fiber communication systems carry more than 99% of world's internet traffic.
- The rapid growth in bandwidth hungry services like video-on-demand are driving the need of even higher data transmission capacity systems.
- One of the attractive solution is to speed up the transmitters and the receivers i.e., by operating at > 100 Giga symbol per second.
- Pushing for higher symbol rate is challenging as components are operating at their limits and cause more severe distortions.
- For instance, the signal has to propagate through a nonlinear signal-path at the transmitter.
- A digital pre-distortion technique is required to compensate these distortions at transmitter.

Experiment details:

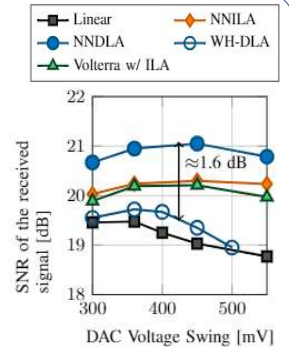
- Investigation of various DPDs on a state-of-the-art experimental setup at Nokia Bell-Labs, Stuttgart.
- Experiments at 128 G symbol/s using two different signals
 - A uniform 64-QAM signal (having 6 bits per symbol).
 - A probabilistic shaped 256-QAM signal (7.5 bits per symbol).
- DAC voltage increased to enhance nonlinear effects and evaluate each DPDs by training at that voltage.



DAC voltage increased to create weak to strong nonlinearity conditions at the transmitter

DPDs evaluated:

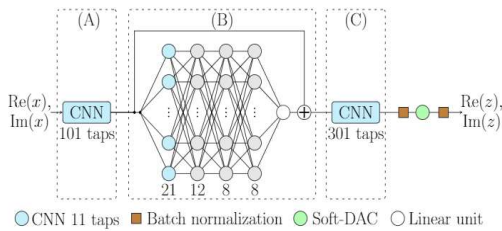
- Linear DPD (with 441 taps)
- Volterra-ILA (Volterra series-based DPD, ILA)
- NN-ILA (NN-based DPD, ILA)
- WH-DLA (Weiner-Hammerstein DPD, DLA)
- NN-DLA (NN-based DPD, DLA)



- Usually, a linear DPD is applied. Because linear effects are relatively more dominant and less difficult to determine.
- In such case, the uncompensated nonlinear distortions degrades quality of transmission.
- For e.g., experimentally obtained trace shows power-dependent noise in the received 64-QAM symbols (6 bits/symbol) in the figure above.

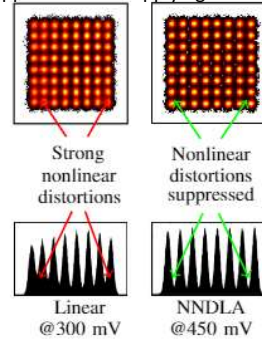
Proposed Solution:

- As the transmitter has complicated chain of linear and nonlinear responses, we investigated neural network (NN)-based DPDs as they are quite promising at modelling nonlinear systems.
- We designed a NN-based architecture by testing over simulation and experiments.
- The NN captures the memory (linear effects) using convolutional layers. The nonlinearity is compensated using fully-connected feed-forward layers.

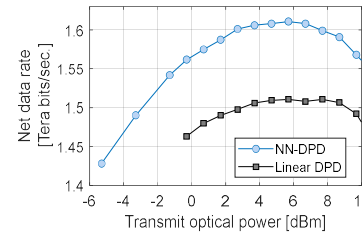


- NN implemented using real values.
- For NN training using DLA, another NN (auxiliary channel) with similar architecture (C → B → A) was used.
- A customized unit Soft-DAC was used to quantize the DPD-NN outputs.
- Training data are the sequence of the transmit and the received symbols obtained from experiments.
- Following two training architectures were used to train DPDs:

Figure shows that nonlinear distortion are suppressed after applying NNDLA-DPD



Net data rate achieved over 80 km of fiber over a single wavelength

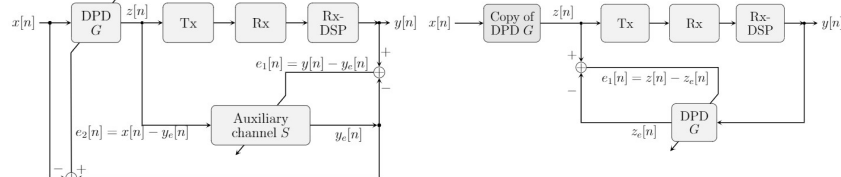


Results & Conclusion:

- The NN-based DPD trained using DLA achieved superior performance among considered DPDs.
- We achieve SNR gains of up to 1.6 dB in comparison to linear DPD.
- A poor performing WH-DPD shows that nonlinearity is mixed with memory, making nonlinear pre-compensation difficult.
- By applying our NN-DLA DPD, we achieve a record data rate of 1.61 Terabits/second over 80 km of optical fiber [3].
- We further show that the complexity of NN-based DPD can be reduced without significant reduction in performance [3].
- In future, we would investigate application NN-based methods to optimize transmitters and receivers.

References:

- D. Rafique et al., "Digital preemphasis in optical communication systems: On the dac requirements for terabit transmission applications", *Journal of Lightwave Technology*, vol. 32, no. 19, pp. 3247–3256, 2014.
- R. Elschner et al., "Improving achievable information rates of 64-gbpdm-64qam by nonlinear transmitter predistortion," in *2018 Optical Fiber Communications Conference (OFC)*, 2018.
- V. Bajaj et al., "Deep Neural Network-Based Digital Pre-distortion for High Baudrate Optical Coherent Transmission," in *Journal of Lightwave Technology*, doi: 10.1109/JLT.2021.3122161.



Direct learning architecture (DLA):

- Finds pre-inverse in two steps
 - Minimize e_1
 - Minimize e_2 (keep S fixed)..

Indirect learning architecture (ILA):

- Finds post-inverse in single step
- Uses its copy as pre-inverse.