Generalization Capabilities of Machine Learning-based PDM Equalization

Samuel Lennard*, Fabio A. Barbosa, Filipe M. Ferreira
Optical Networks Group, Dept. Electronic & Electrical Eng., University College London (UCL), UK
**s.lennard@ucl.ac.uk*

Abstract: We investigate the generalization capabilities of a novel machine learning-based receiver for recovering PDM 16-QAM symbols over unseen chromatic dispersion, Kerr nonlinear distortion, and stochastic polarization evolution.

1. Introduction

As demand for fiber transmission capacity increases, spatial division multiplexing (SDM) has risen to prominence as the next major step in optical fiber research [1]. However, this comes with multiple challenges, including developing robust DSP to effectively recover spatially multiplexed data with varying spatial mode coupling [2]. In this paper, we introduce *GEqNet* (Generalized Equalization Network), a novel machine learning model for coherent receivers currently able to replace the DSP blocks for chromatic dispersion (CD) compensation, channel estimation, and equalization. *GEqNet* does not require precise knowledge of the fiber characteristics, being able to generalize from a sufficiently broad offline training stage. Here, we aim to develop a proof of concept using a polarization-division multiplexing (PDM) system in a single-mode fiber (SMF) with varying polarization-mode coupling strength, CD, and nonlinearities.

2. System Model

The optical transmission system setup considers 100 km transmission over a SMF channel using an ideal transmitter. A default CD of 16.75 ps/(nm·km), dispersion slope 0.06 ps/(nm^2·km), nonlinear coefficient $\gamma = 1.2$ /W/km, and attenuation 0.152 dB/km at 1550 nm are considered, with differential group delay and phase noise neglected. The transmission signal consist of a single-channel with 30 Gbaud/s PDM 16-QAM, and is root raised cosine shaped with a 0.01 roll-off factor.

The optical fiber is modelled using the single-mode split-step Fourier method including the effects of CD and Kerr nonlinearity [3] as well as stochastic polarization evolution with a given polarization correlation length [4]. A stochastic polarization evolution is introduced in Jones space every k-th step using a unitary mode coupling matrix given by $\exp(\mathbb{I}_{2x2} + M_{sh})$, where perturbation M_{sh} is a skew-Hermitian matrix whose elements m_{ij} follow $\Re(m_{ij}), \Im(m_{ij}) \sim \mathcal{N}(0, \sigma^2)$ and $m_{ij} = -\overline{m_{ji}}$; an approach similar to the one in [5]. The crosstalk strength (XT) can be controlled via the perturbation variance of M_{sh}. Before detection, additive white Gaussian noise is introduced such that an optical signal-to-noise ratio of 11.6 dB/0.1nm is obtained for a launching power of 0 dBm. After homodyne detection, the baseband electrical signals are sampled at 4 samples/symbol. The coherently received signals are then compensated for all fiber impairments using *GEqNet*.

3. Learned DSP for PDM

ML approaches that utilize online learning can be costly to deploy in receivers, and we therefore look to mitigate this by generalizing over unseen fibers in an offline training stage. Equalization can then be applied by exposing a received training and data sequence to the ML model to predict the transmitted sequence.

GEqNet uses 8 convolutional neural network (CNN) channels to encode its input, corresponding to two polarizations made up of a real and imaginary part, and with a training and data sequence. The input is normalized, passed through multiple CNN layers with the hyperbolic tangent activation function, flattened, and passed through a dense feedforward neural network with a ReLU activation and linear activation in its respective layers to produce the real and imaginary parts of the predicted transmitted symbols. The network architecture is shown in Fig. 1.

To showcase the generalization capabilities of *GEqNet*, two models are trained with different procedures, with one aiming to generalize over signal transmit powers, and another generalizing over dispersion coefficients. In the first case, 150 different realizations of fibers are generated with random polarization mode coupling, with average crosstalk varying between -40dB/km and -10dB/km. On each fiber, 24 random signals of 128 16QAM symbols are sampled at a random transmit power drawn from the distribution $P \sim \mathcal{N}(\mu = 9, \sigma = 4)$ dBm for both polarizations and launched down the fiber, as well as the pre-determined training sequence. In the second case, fibers are generated in the same way but also with random dispersion coefficients between 10 and 20 ps/(nm·km), while transmit power is kept to 9dBm. The mean-square error between the predicted and actual symbols transmitted are then optimized using the default Adam optimizer provided by Pytorch for 500 epochs with a batch size of 1024.

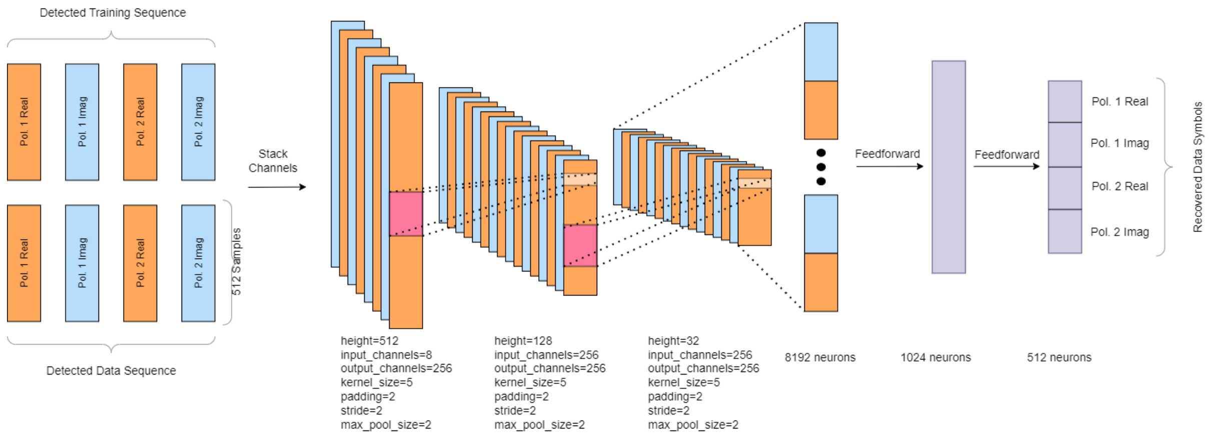

Fig. 1: Diagram of *GEqNet*, comprising of a convolutional component and a feedforward component.

4. Numerical Results

To test the ability of the model to generalize to unseen fibers and signals, it is evaluated on realizations of fibers that it has not been trained on. These fibers have a range of average crosstalks (i.e., polarization correlation length [4]), and for each fiber 100 random 16-QAM signals are generated and launched with the known training sequence. The model then predicts the transmitted symbols, and SNR is calculated from the resulting constellation.

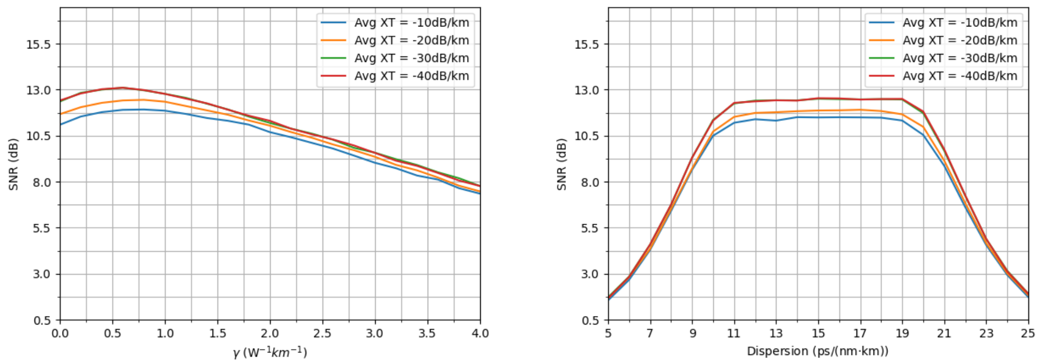

Fig. 2: SNR performance of *GEqNet*: (a) SNR vs gamma with *GEqNet* trained for a range of transmit powers and constant dispersion, (b) SNR vs CD with *GEqNet* trained on constant transmit power and a range of dispersions.

In Fig. 2 (a) it can be seen that *GEqNet* is able to achieve generalization over a range of nonlinear coefficient values. It achieves high SNR for $\gamma \in [0, 1.5]/W/km$, before decreasing steadily as γ increases. Meanwhile, Fig 2 (b) shows that *GEqNet* is able to generalize over different values for chromatic dispersion when the training set is sufficiently varied. However, it should be noted that with the current model architecture, generalizing over both at the same time results in an approximate loss of 1.25dB of SNR at optimal launch power. In addition, it is observed that *GEqNet* suffers from poor out-of-distribution performance – in order to generalize to other fibers, the testing distribution must be a subset of the training distribution.

5. Conclusion and Future Work

In this paper we introduced a novel machine learning approach to equalization in coherent receivers that can generalize to unseen fibers, eliminating the need for costly online training. While the realization of *GEqNet* shown here is a proof of concept, it shows that the possibility of mitigating online training for machine learning-based DSP is plausible, whilst presenting a framework that can be extended to SDM. Further work may include applying *GEqNet* so that it can generalize to unseen *real fiber channels*, and extending it to multiple modes.

Acknowledgements: This work was supported by a UKRI Future Leaders Fellowship under grant no. MR/T041218/1. Underlying data at doi.org/10.5522/04/24428044 and doi.org/10.5522/04/24461179.

References

1. D. Richardson, *et al.*, "Space-division multiplexing in optical fibers," Nature Photonics **7**(5), 354-362 (2013).
2. F. Barbosa, *et al.*, "On a Scalable Path for Multimode SDM Transmission," *J. Lightwave Technol.*, *early access* (2023).
3. E. Ip, *et al.*, "Compensation of Dispersion and Nonlinear Impairments Using Digital Backpropagation," *J. Lightwave Technol.* **26**(20), 3416-3425 (2008).
4. C. Poole, *et al.*, "Chapter 6 - Polarization Effects in Lightwave Systems," Optical Fiber Telecommunications IIIA (Third Edition), Academic Press, 114-161 (1997).
5. C. Czegledi, *et al.*, "Polarization Drift Channel Model for Coherent Fibre-Optic Systems," Scientific Reports **6** (2016).