Neural Nets to Approach Optimal Receivers for High Speed Optical Communication

Sai Chandra Kumari Kalla, Rizan Homayoun Nejad, Sasan Zhalehpour, and Leslie A. Rusch

IEEE/OSA Conference on Lasers and Electro-Optics (CLEO 2020)

Neural Nets to Approach Optimal Receivers for High Speed Optical Communication

Sai Chandra Kumari Kalla, Rizan Homayoun Nejad, Sasan Zhalehpour and Leslie A. Rusch
Centre for Optics, Photonics and Lasers (COPL), ECE Department, Université Laval, Quebec, Canada
rusch@gel.ulaval.ca

Abstract: We show via simulation that two nonlinear neural network solutions can achieve performance close to that of the maximum likelihood sequence estimator for experimental frequency responses of a silicon photonic modulator and coherent detection.

1. INTRODUCTION

Inter-symbol interference (ISI) is a principal impairment in high speed optical communications with bandlimited components. Minimum mean squared error (MMSE) filtering combats ISI and is the optimal linear one-shot receiver for Gaussian noise channels. The maximum likelihood sequence estimator (MLSE) examines all possible sequences to provide an optimal nonlinear filter. However, MMSE performance deteriorates with frequency band attenuation, and MLSE requires complete channel state information (CSI). Furthermore, the complexity of MLSE receiver increases exponentially with modulation order and ISI memory length. We examine the performance of neural networks vis-à-vis MMSE and MLSE solutions for severely bandlimited channels. The neural networks (NN) hold the potential to learn the channel indirectly during the training stage without any explicit need of CSI.

In [1], a deep neural network (DNN) is implemented for dispersion mitigation in optical fibers, however, they do not address bandlimited channels. In [2], a recurrent neural network (RNN) architecture is used to combat ISI induced by a Poisson channel with performance approaching MLSE; however, the Poisson channel is a poor model for high bit rate optical communications. We examine two distinct nonlinear neural network architectures for ISI mitigation: a nonlinear functional neural network (NLFNN) and a long-short term memory RNN (LSTM-RNN). Both solutions are able to approach the MLSE equalizer performance with small power penalty. A third neural network without nonlinear activation, in the form of a Linear Multilayer Perceptron (LMLP), is presented for contrast. It can achieve MMSE performance, but not MLSE. We investigate via simulation, channels where large performance gaps exist between MMSE and MLSE: both a highly studied synthetic channel [3], and a measured optical channel at 100 Gbaud using a silicon photonic (SiP) modulator [4].

2. NEURAL NETWORKS FOR ISI MITIGATION

Linear Multilayer Perceptron (LMLP): This simple linear NN provides a baseline, and has input and output layers and one hidden layer with 70 linear neurons. The input layer has 58 features: in-phase (I) and quadrature (Q) components of targeted symbol, 14 past and 14 future symbols. The weights of all NN models we examine are adapted to minimise the mean squared error of training data using a stochastic-gradient-descent algorithm. The input data is convoluted with the final LMLP weights to give the equalized channel outputs. We will see the LMLP with no nonlinear operations and no feedback path, unsurprisingly achieves performance equal to that of the classic linear MMSE equalizer. We move to two new NN structures to target MLSE performance.

Long Short Term Memory-Recurrent Neural Network (LSTM-RNN): The optimal MLSE performance is achieved by examining a long sequence of symbol decisions; the RNN architecture uses a feedback path to achieve similar visibility. RNNs, in contrast to feed-forward NNs (like LMLP) have additional memory cells in each hidden layer neuron that stores past decision information. This structure combines input data with past information to process sequences rather than single data points, an approach similar to MLSE. We investigate LSTM-RNNs as they can also focus on long term sequence dependencies during training. We have 18 input features (4 past/future), an output layer and one hidden layer with 70 LSTM units [5]. In addition to standard memory cell in RNN neurons, each LSTM unit has 3 gates: input, output and forget gate. These gates regulate the flow of information to/from memory and result in capturing long term effects; we will see performance approaching that of MLSE receivers.

Nonlinear Functional Neural Network (NLFNN): MLSE achieves nonlinear behaviour via feedback on previous decisions. Our second NN focuses on non-linearity without feedback. Our NLFNN has an input layer with 18 features, an output layer and two hidden layers. Each hidden layer has 50 neurons with *tanh* as the nonlinear activation function. In contrast to our LMLP, NLFNN has 2 hidden layers with fewer neurons in each layer. The concatenated nonlinearities in the two NLFNN hidden layers are meant to capture the nonlinear nature of MLSE.

3. SIMULATION, RESULTS AND DISCUSSION

We compare our NN equalizers with optimal receivers for 4-QAM transmission for two channel responses: (i) Proakis synthetic channel [3] (multipath response: $[0.41, 0.81, 0.41]$), (ii) 100 Gbaud optical channel estimated from the experimental setup in [4]. Fig. 1a shows the frequency responses. The 513 tap estimated channel is approximated by $[0.25+0.38i, 0.75+0.84i, 0.42+0.09i]$ in our simulations. Both channels are severely attenuated at higher frequencies (i.e., bandlimited). The 100 Gbaud system has a SiP modulator with 35 GHz bandwidth and complex response is due to a nonideal modulator. In Fig. 1b and Fig. 1c, we report Monte Carlo simulations of bit error rate (BER) vs. Signal-to-noise-ratio (SNR) estimated over 10^5 symbols. The BER curves for no equalization and for a channel with infinite bandwidth (no ISI) are provided as benchmarks. The performance of MMSE and MLSE receivers are plotted with dashed lines; they exhibit a large gap as the channels are bandlimited. For both channels, our LMLP (diamond markers) attains the performance of the MMSE equalizer. Both nonlinear NN equalizers, NLFNN (circle markers) and LSTM-RNN (red line), achieve very similar performance. At 3.8e-3 BER (7% forward error correction threshold) they have a small power penalty vis-à-vis the MLSE equalizer: 1.2 dB and 1.9 dB for the optical and Proakis channels, respectively. For both channels, the nonlinear NN equalizers outperform the MMSE receiver. The MLSE without ideal CSI (used in our results) will have lower performance, while our NN results reflect the inferred CSI. In Proakis channel, MMSE reach 3.8e-3 BER at 25.6 dB SNR, for a gap of $\sim$ 14 dB with MLSE, vs. a gap of 6 dB for the optical channel. The nonlinear NN equalizers cover a large portion of the performance gap between MMSE and MLSE. Interestingly, despite the much larger gap for one, the channels have similar SNR penalty when using the nonlinear NN equalizers.

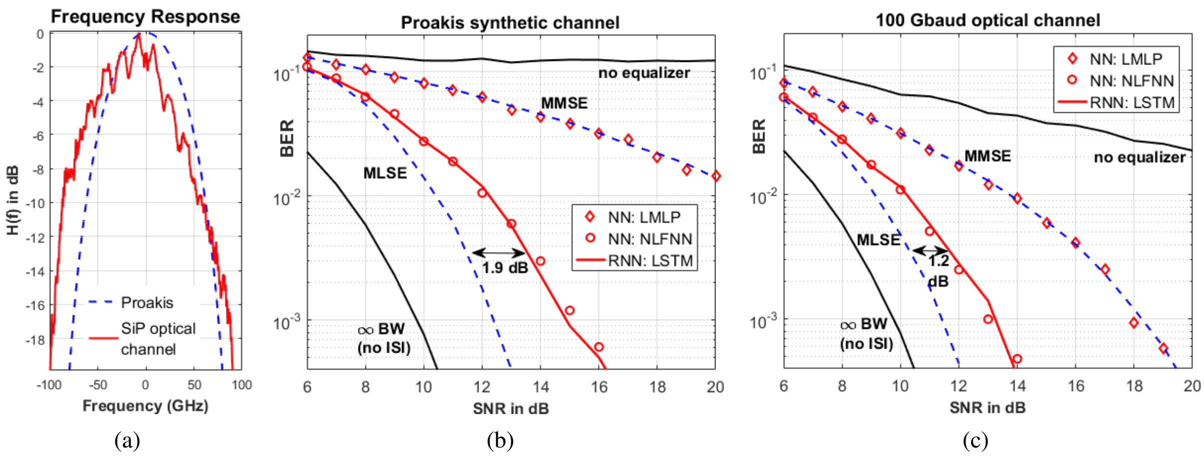

Fig. 1: (a) Frequency responses & BER performance for (b) Proakis, and (c) 100 Gbaud SiP channel

4. CONCLUSION

We demonstrated nonlinear NN equalizers approach MLSE equalizer performance with small (<2 dB) power penalty, whereas LMLP achieves only MMSE performance. The MLSE and nonlinear NN equalizers have very different architectures. The MLSE complexity scales very poorly with QAM modulation level and memory depth. As we have two NN equalizer architectures that are effective, they may offer an opportunity for better scaling.

Acknowledgment: This research was supported by the Huawei/NSERC Industrial Research Chair, 537311-18.

References

1. Yung Hsu, et al., “Implementing Deep Neural Network for Signal Transmission Distortion Mitigation of PAM-4 Generated by Silicon Mach-Zehnder Modulator,” OECC/PSC, WP4-B10 (2019).
2. Nariman Farsad, et al., “Neural Network Detection of Data Sequences in Communication Systems,” *IEEE Trans. on Signal Processing* **66**, pp. 5663-5678 (2018).
3. John G. Proakis, *Digital Communications* (McGraw-Hill Education, 2000), Chap. 10.
4. Sasan Zhalehpour, et al., “All-Silicon IQ Modulator for 100 GBaud 32QAM Transmissions,” OFC Post-deadline Paper, Th4A.5 (2019).
5. Jesús Gonzalez, et al., “Non-linear system modeling using LSTM neural networks,” IFAC-PapersOnLine **51**, pp. 485–489 (2018).