

# After FEC Performance of 50G PON with Deep Neural Network Equalization

Amitkumar Mahadevan, Vincent Houtsma, Noriaki Kaneda, and Doutje van Veen

Nokia Bell Labs, 600-700 Mountain Ave., Murray Hill, NJ, 07974, USA.  
amitkumar.mahadevan@nokia-bell-labs.com

**Abstract** Performance of Deep Neural Network receiver equalization is investigated after LDPC FEC. It is shown that DNN equalization results in a lower level of error clustering and higher performing equalization relative to more traditional equalization for bandwidth-limited and dispersion-limited channels for 50G PON.

## Introduction

The emergence of new high-bandwidth services continues to drive the demand for progressively higher speed optical access networks; the recent paradigm shift towards telepresence is only expected to further fuel this demand. The IEEE 802.3ca standard for 25 and 50 Gbps (2 wavelengths) TDM-EPON was recently published<sup>[1]</sup>; in the ITU-T, standardization of a single wavelength 50 Gbps TDM-PON is under active development in project G.hsp<sup>[2]</sup>. Efforts have mostly been focused on increasing line-rate of PON networks while enabling cost-effective limited bandwidth optical components through adoption of high gain forward error correction (FEC) and equalization. Hard-decision input low-density parity-check (LDPC) codes have been adopted for PON<sup>[1]</sup>. For 50 Gbps non-return to zero (NRZ) IM-DD PON systems, complex receiver equalization schemes are required to combat the inter-symbol interference (ISI) introduced by the limited bandwidth components as well as reduced tolerance to chromatic dispersion (CD)<sup>[3]</sup>. However, equalization produces correlated errors due to the inherent filtering operations; these are exacerbated by error propagation when feedback structures are involved. The impact of these errors is even more significant in FEC-based systems involving binary LDPC codes.

In<sup>[4]</sup> the impact of error clustering on the after FEC performance for 50G PON has been studied for traditional equalizers (FFE+DFE and MLSE) under transceiver bandwidth limitations and dispersion. These error clusters become especially pronounced in the DFE when the magnitude of the feedback tap is more than half

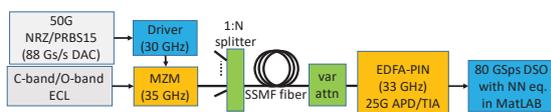
the main feedforward tap (cursor)<sup>[5]</sup>.

Equalization based on Neural Networks (NN) has recently been proposed in research for the PON optical physical layer due to its excellent performance<sup>[6]</sup>. Also, to validate feasibility, a fixed-point deep NN based equalizer was implemented in FPGA and shown to outperform conventional linear and non-linear equalizers for the 50 Gbps PON downstream<sup>[7]</sup>, at the typically considered FEC input bit-error rate (BER) threshold of 1E-2. However, the impact of error clustering after NN equalization and its influence on the resulting FEC performance for 50G PON has not been investigated yet.

In this paper, we study error clustering due to NN equalization and present after-FEC performance for 50G PON based on transmission experiments. To separately evaluate the impact of ISI due to bandwidth limitations and chromatic dispersion on 50G PON after-FEC performance with NN equalization, we evaluated error clustering at 50G in C-band ( $\lambda=1565$  nm) using a full bandwidth receiver as well as in the O-band ( $\lambda=1342$  nm) with limited bandwidth receivers after 20 and 30 km of standard single mode fiber (SSMF), respectively. While using an O-band wavelength has the advantage to be near zero dispersion for the SSMF, the dispersion at 1565 nm is considerable.

## Neural Network implementation

We chose a simple feed forward deep Neural Network (DNN) architecture as the receiver side equalizer with 1 symbol output. Among NN architectures, the feed forward DNN architecture has the lowest complexity and is amenable to easier hardware implementation<sup>[7]</sup>. Two different DNN sizes were used, one DNN optimized for performance at 50 Gbps under high dispersive conditions (total CD=365 ps/nm) and one for limited bandwidth reception at more modest dispersion (total CD = 83 ps/nm). The minimum required number of layers and number of neurons for each NN depends on the expected impulse



**Fig. 1:** Experimental setup of 50Gb/s downstream PON with NN equalization

response of the inverse channel to deal with the dispersive/bandwidth limited channel and is determined by a parameter search.

The first DNN (DNN-C) designed for C-band operation under full bandwidth reception has 15 inputs, 1 output and 4 hidden layers with 15, 8, 4, 2 neurons totalling 417 weights including biases. The second DNN (DNN-O) for O-band and limited bandwidth reception has 11 inputs, 1 output, and 2 hidden layers with 33 and 14 neurons with a total of 887 weights including biases. The weights and biases are obtained by the backward propagation of DNN using training data from optical transmission experiments.

### Experimental Setup

Figure 1 shows the experimental setup. A 50 Gbps NRZ signal generated using a 88 GSa/s (1.75x oversampling) digital-to-analog converter (DAC) was modulated using a Mach Zehnder Modulator (MZM) on either a C- or O-band wavelength generated by a tunable external cavity laser (ECL). The wavelength of  $\lambda=1565$  nm was chosen in the C-band for maximum dispersion, while  $\lambda=1342$  nm, the agreed 50G downstream wavelength by the ITU for G.hsp<sup>[2]</sup>, was chosen in the O-band. The optical signal was transmitted over SSMF and attenuated prior to reception. The O-band receiver was a limited bandwidth 25 Gb/s class avalanche photo diode (APD) with integrated linear transimpedance amplifier (TIA); the C-band receiver was a full bandwidth (33 GHz at -3 dB) PIN with a constant 18 dB optical pre-amplification using an Erbium Doped Fiber Amplifier (EDFA) without an optical noise filter. The waveforms were captured using an 80 GSa/s oscilloscope, resampled to 1 Sa/symbol, and fed to the DNN for equalization. The sampling point was chosen at the point of minimum BER for direct-detection (DD).

### Experimental Results

Fig. 2 shows 50 Gbps NRZ experimental results. For NN equalization at an output BER =  $1E-2$ , we achieved -28.4 dBm sensitivity for b2b and -24.5 dBm after 20 km of fiber at  $\lambda=1565$  nm (CD of 365 ps/nm). A sensitivity of -24.2 dBm (b2b) and -23.7 dBm after 30 km of fiber was achieved at  $\lambda=1342$  nm (CD of 83 ps/nm), which covers the worst dispersion for 20 km SSMF at 1342 nm. The sensitivity performance of equalization schemes considered in [4] (CTLE+DFE and MLSE) as well as that of a 11-tap FFE, for 30km at 1342 nm is also shown in Fig. 2 for reference. These results show the promise of using DNN-based receiver-side equalization for either bandwidth limited systems or high dispersion values to support 50G-PON downstream transmission over 20-km

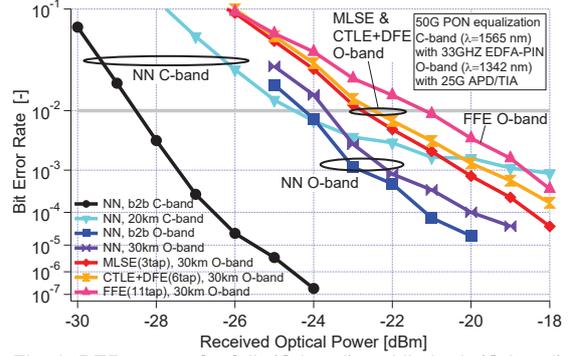


Fig. 2: BER-curves for full- (C-band) and limited- (O-band) BW optical receivers for b2b and after 20 km (C-band 365 ps/nm) and 30 km (O-band 83 ps/nm) of fiber with NN

SSMF fiber with 29 dB link budget at the FEC threshold of  $1E-2$ . However, for overall system performance, the sensitivity at FEC decoder output BER of  $1E-12$  is important as well. The optical power penalty due to error clustering after NN equalization should be included in the receiver sensitivity when comparing equalization schemes that introduce correlated errors at the LDPC decoder input.

The error correlation at the equalizer output is characterized using the Error Cluster Distribution (ECD)<sup>[4]</sup>, i.e., the distribution of the relative probability of consecutive error events. Figure 3(a) compares the ECD (normalized to the 1-error event probability) as a function of the cluster length for NN equalization against those of other equalization schemes considered in [4], as well as the FFE with 11 input taps – the same number as the O-band NN. Solid lines with filled markers are measurement data for points in Fig. 2 close to the  $1E-2$  BER threshold for all cases. The simulated ECD for the binary symmetric channel (BSC) with average BER =  $1E-2$  is also shown for reference. A larger deviation to the right of the ECD of the BSC implies stronger error correlation at the equalizer output. The CTLE+DFE shows the highest error correlation, which is primarily due to the error-propagation induced by the feedback structure. On the other hand, the 11-tap FFE shows the least correlation; such behavior

**Tab. 1:** LDPC performance for output BER of  $1E-12$  for different equalized 50 Gbps NRZ PON transmission cases. O-band results are for 83 ps/nm CD with 25G class APD; C-band results are for 365 ps/nm CD with full bandwidth (33 GHz) EDFA/PIN.

Equalizer Case	Reqd. input BER	Elect. penalty (dB)	Optical penalty (dB)
BSC	$1.03E-2$	-	-
DNN-O	$8.9E-3$	0.20	0.08
DNN-C	$9.0E-3$	0.19	0.11
CTLE+DFE -O	$6.8E-3$	0.55	0.6
MLSE-O	$8.3E-3$	0.3	0.3

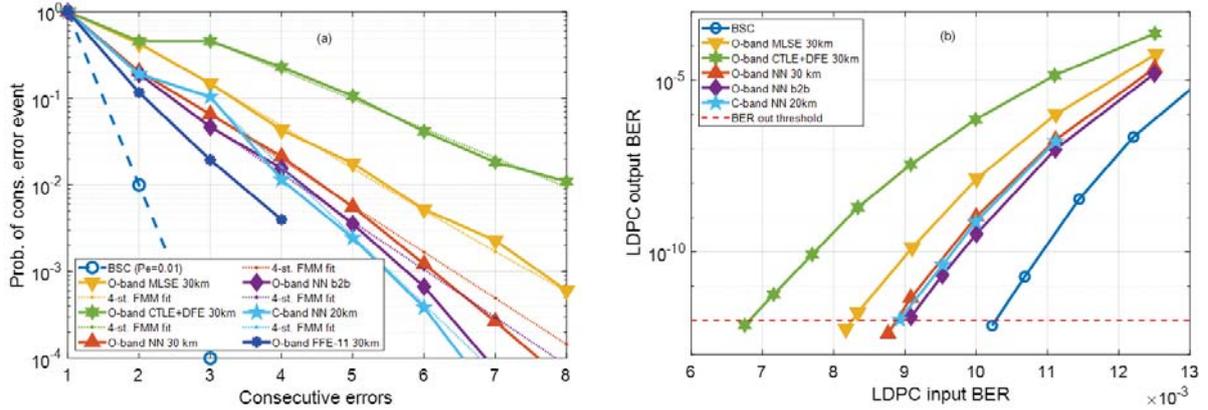


Fig 3(a). Probability of consecutive error events (normalized to 1-error event probability) vs. consecutive error length: measured results are shown with solid lines, while 4-state FMM fits are shown with dotted lines. (b) Information bit BER after LDPC decoding vs. BER at decoder input with errors generated using the 4-state FMM.

may be attributed to its feedforward and linear structure. The feedforward structure of the DNN also results in low error correlation, but its non-linear structure is likely responsible for the higher correlation as compared to the FFE. Characterization of the MLSE ECD is not as straightforward given that an error event during the Viterbi decoding process involves divergence from the intended path in the trellis followed by reconvergence after a minimum of three stages for a 4-state (3-tap) MLSE.

The 4-state Fritchman's Markov model (FMM) is used to fit to the ECD<sup>[4, 8]</sup>; these fits are shown in Fig. 3(a) using dotted lines with like colors. The fitting parameters are chosen to accurately capture the relative behavior of error clusters of length 1 to 3, and then the asymptotic slope for longer cluster lengths while ignoring points with poor confidence due to low event counts. The fitted FMM is then used as a generative model to inject correlated errors prior to LDPC decoding. Figure 3(b) shows the LDPC decoder output BER vs. input BER (equalizer output BER) for the IEEE 802.3ca quasi-cyclic LDPC mother code<sup>[1]</sup> with the last two parity columns punctured (code rate of ~0.85). Table 1 shows the required input BER to achieve the target BER = 1E-12 at the LDPC decoder output. For the BSC channel, the input BER is 1.03E-2. For the 4-state fit of the DNN-O measurements after 30 km, the input BER is 8.9E-3. The corresponding electrical penalty (EP) reported in Tab. 1 as 0.2 dB is directly calculated from the BER values as  $20 \log_{10}(Q^{-1}(\text{BER})/Q^{-1}(\text{BER}_{\text{BSC}}))$  dB, where  $Q(x) = \int_x^\infty e^{-t^2/2} dt$ . Meanwhile, the optical penalty (OP) due to reduction in input BER with respect to the BSC to achieve an output BER of 1E-12 is derived via interpolation from the BER vs. ROP curve in Fig. 2 and reveals an OP of 0.08 dB due to correlated errors at the output of the DNN-O. For the C-band measurements, the OP

due to error clustering after 20 km of fiber transmission is 0.11 dB. In comparison, the OP due to error clustering is 0.3 dB for the 3-tap MLSE and 0.6 dB for CTLE+DFE equalization in O-band after 30 km of fiber<sup>[4]</sup>. In addition to their lower EP, the DNNs also exhibit a lower OP/EP ratio as compared to the traditional equalizer schemes due to the steeper slope of their BER vs. ROP curves near the BER of 1E-2.

For 83 ps/nm dispersion, the DNN-O yields a 1.0 dB gain over the CTLE+DFE if compared at an equalizer output BER of 1E-2 (Fig. 2). Moreover, when correlated errors are considered and the comparison is performed at the LDPC output BER of 1E-12, the gain increases to 1.5 dB, since the NN equalizer also results in lower error correlation as compared to CTLE+DFE. Similarly, the DNN-O gain over the MLSE of 0.7 dB at equalizer output BER of 1E-2 increases to 0.9 dB at LDPC output BER of 1E-12.

## Conclusions

Deep Neural Network receiver side equalization is a promising alternative for equalization of bandwidth-limited as well as dispersion-limited channels for 50G PON. DNN equalization outperforms traditional equalization not only in terms of resulting optical receiver sensitivity at the typical FEC input BER=1E-2, but it also results in lesser error correlation, which leads to even higher gain at the target LDPC output BER = 1E-12.

To our knowledge, this is the first time that consecutive error behaviour after Neural Network equalization for 50G PON transmission experiments with transceiver bandwidth limitation and dispersion has been studied.

## Acknowledgements

The authors thank Ed Harstead and Jochen Maes for their inputs.

## References

- [1] IEEE Std 802.3ca-2020, "IEEE Standard for Ethernet Amendment 9: Physical Layer Specifications and Management Parameters for 25 Gb/s and 50 Gb/s Passive Optical Networks", 3 July 2020.
- [2] "ITU-T G.hsp recommendation: 50-Gigabit-capable passive optical networks (50G-PON): Physical media dependent (PMD) layer specification," 2020.
- [3] M. Tao, J. Zheng, X. Dong, K. Zhang, L. Zhou, H. Zeng, Y. Luo, S. Li and X. Liu "Improved dispersion tolerance for 50G-PON downstream transmission via receiver-side equalization," Proc. OFC '19, M2B.3, 2019.
- [4] A. Mahadevan, D. van Veen, N. Kaneda, A. Duque, A. de Lind van Wijngaarden, and V. Houtsma "50G PON FEC Evaluation with Error Models for Advanced Equalization", Proc. OFC '20, Th1B.6, 2020.
- [5] R. Narasimha, N. Warke, and N. Shanbhag, "Impact of DFE error propagation on FEC-based high-speed I/O links," Proc. IEEE Global Telecommun. Conf., Honolulu, HI, USA, 2009.
- [6] V. Houtsma, E. Chou, and D. van Veen, "92 and 50 Gbps TDM-PON using neural network enabled receiver equalization specialized for PON," in Proc. OFC '19, M2B.6, 2019.
- [7] N. Kaneda, Z. Zhu, C-Y Chuang, A. Mahadevan, B. Farah, K. Bergman, D. Van Veen, and V. Houtsma, "FPGA Implementation of Deep Neural Network Based Equalizers for High-Speed PON" , Proc. OFC '20, T4D.2, 2020.
- [8] B. Fritchman, "A binary channel characterization using partitioned Markov chains," IEEE Trans. Inf. Theory, vol. 13, no. 2, Apr. 1967.