21 Gbaud DP-8QAM TRANSMISSION WITH ARTIFICIAL INTELLIGENCE BASED NONLINEAR COMPENSATION OVER A 13212 KM LINK

Yuliang Gao, Ziad A. El-Sahn, Demin Yao , Han Sun , Pierre Mertz, Kuang-Tsan Wu (Infinera)
Email: ygao@infinera.com

Infinera Canada, 555 Legget Dr., Kanata, ON K2K 2X3, Canada.

Abstract: We demonstrate the effectiveness of machine learning and artificial intelligence (AI) to compensate for fiber nonlinearity over transoceanic distances for 21 Gbaud DP-8QAM. We show that using a deep neural network (DNN) with three hidden layers and a single output layer we can achieve 0.24 dB improvement in Q factor compared to the case without fiber nonlinearity compensation (NLC). DP-8QAM transmission was performed over a 13212 km field-deployed uncompensated ultra-large effective area fiber link with 21.66 ps/nm/km dispersion, 56.46 km average span length, and 4.5 dB noise figure erbium-doped fiber amplifiers (EDFAs).

To simplify the design of the AI-NLC engine, we use a 4-subcarrier signal and asymmetrical dispersion compensation with 40% of the dispersion compensated at the transmitter and 60% at the receiver. Compared to previous implementations, the neural network is only fed with a window of 101 symbols, with no nonlinearity perturbation triplets at its input. Moreover, the DNN is inserted in the forward-error correcting (FEC) decoding loop, for accurate estimation of the nonlinear perturbation terms achieving 0.86 dB Q improvement. In this paper, we also confirm that the DNN is not learning the sequence of received symbols, by examining different sets of data (other than the set used for training).

1. INTRODUCTION

Fiber nonlinearity mitigation in subsea and long-haul systems is becoming crucial to boost the link capacity [1]. However, using traditional techniques such as digital back-propagation (DBP), perturbative approaches, and Volterra equalization may not be well suited for practical implementation [2]-[4]. Alternatively, the use of machine learning techniques and artificial intelligence (AI) for fiber nonlinearity compensation (NLC) has recently attracted many researchers from the industry and academia sectors [5]-[10].

Unlike initial research efforts in using deep neural networks (DNNs) for combating fiber nonlinearity effects [5], [6], recent demonstrations of AI-NLC [7]-[10] added intra-channel cross-phase modulation and intra-channel four-wave mixing triplets as learning features to their simplified DNN models. However, calculating these triplets introduce extra computation complexity, and thus, is hard to be implemented in real-time digital signal processing (DSP) application specific integrated circuits (ASICs). In [7], the DNN was used for post-compensation and was fed with a window of soft received symbols and the corresponding triplets from both polarizations. The same authors in [7], proposed a smarter way [8] by moving the AI-NLC at the transmitter side to pre-distort the transmitted symbols similar to traditional perturbation technique by only feeding triplets to the DNN. In that scenario, a look-up table can be used to form the triplets to reduce the implementation complexity.

More recently, attempts have been made to balance the trade-off between the DNN size

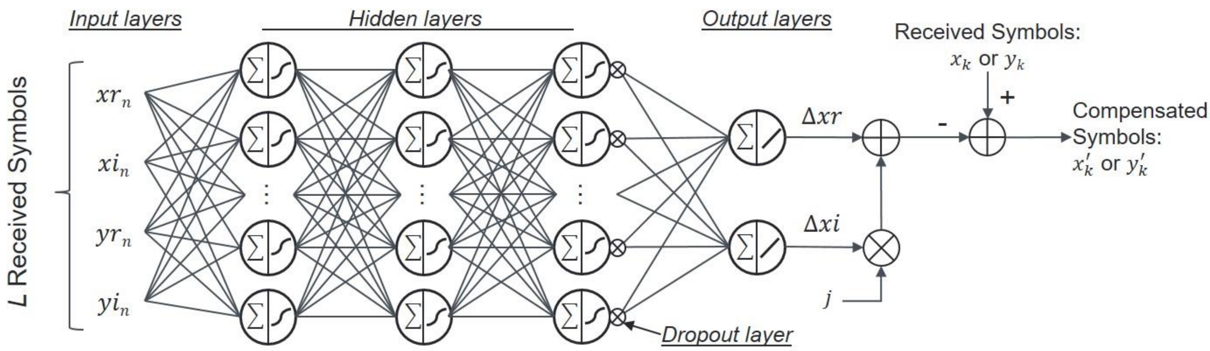


Figure 1: The proposed AI-NLC algorithm. xr_n, xi_n, yr_n, yi_n are real and imaginary parts of received symbols on x and y polarizations. n is the symbol index $k-(L-1)/2 \leq n \leq k+(L-1)/2$. k is the current symbol of interest. Δxr and Δxi are the real and imaginary parts of the perturbation estimated by DNN. x'_k and y'_k are the compensated

(i.e., number of layers/nodes) and the number of learning features (i.e., inputs to DNN), while preserving the performance. In [9], we have proposed to use the principal component analysis (PCA) to reduce the number of triplets feeding a DNN with the same architecture as in [8]. Whereas in [10], the authors demonstrated nonlinear gains and superiority over DBP using AI-NLC without input triplets and a DNN with 2 hidden layers formed from 16 nodes each. However, the transmission distance was limited to less than 2500 km for 32 Gbaud DP-16QAM (based on simulations). Moreover, a window of ~50 symbols was used by the DNN to predict the nonlinear perturbation associated to a given symbol.

In this paper, we propose a simplified AI-NLC algorithm based on DNNs consisting of three densely connected layers for fiber nonlinearity compensation and without using triplets. Unlike [10], we use symbols from both polarizations to form the input to the DNN. Moreover, the proposed AI-NLC engine is evaluated experimentally over a deployed transoceanic link. To enable receiver side AI-NLC and to facilitate the learning process about fiber nonlinearity in long-haul systems, we introduce a novel pattern averaging technique that partially mitigates the amplified spontaneous emission (ASE) noise during the training stage. In the end, for an open discussion, we discuss the possibility of implementing the proposed AI-NLC algorithm either at the transmitter side or embedded at the receiver side within the forward error correction (FEC) iterations similar to [11], to achieve even better Q improvement.

2. PROPOSED AI-NLC ALGORITHM

We built a DNN model and training strategy based on merely first order soft decoded symbols and with much less features (without nonlinear perturbation triplets) as shown in Figure 1.

After the carrier recovery at the receiver DSP, the soft decoded symbols are separated into in-phase and quadrature components on both x and y polarizations. L neighboring symbols, i.e. $4L$ real components, are fed into the DNN in a sliding window manner. Each layer can be described as follows:

$$\boldsymbol{a}^l = f_{NL}^l(\boldsymbol{H}^l \boldsymbol{a}^{l-1} + b^l),$$

where l is the layer index ranging from 1 to 4, $\boldsymbol{H}^l$ is the weights matrix at layer l, $\boldsymbol{b}^l$ are biases at layer l, f_{NL}^l is defined as nonlinear hyperbolic tangent activation function in the hidden layers and straight forward linear activation function in the output layer, and $\boldsymbol{a}^l$ is the output vector from layer l.

Since the proposed DNN does not rely on the assistance of pre-calculated triplets, one additional hidden layer has been added, as

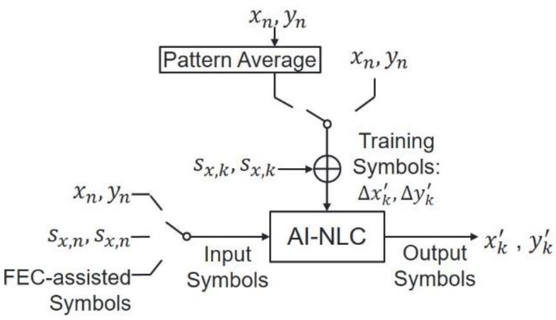
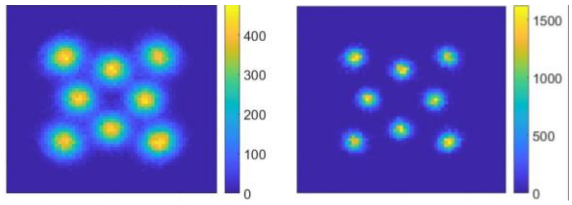


Figure 2: the proposed AI-NLC at either Tx-side or Rx-side. The input can be received symbols (xn, yn), transmitted symbols (sx,n, sy,n), or FEC assisted symbols. The training symbols are from the receiver side with or without pattern

compared to [7]-[9], to better model the nonlinearity impairments. The three densely connected hidden layers each has 10 nodes. A dropout layer with probability of 50% is used after the third hidden layer to avoid over-fitting. The output layer is simply formed from 2 nodes representing the in-phase and quadrature phase components of the estimated nonlinear perturbation term.

If the AI-NLC is implemented in the receiver, the input can be soft decoded symbols (x_n, y_n) or FEC assisted symbols [11] as shown in Figure 2. At transmitter, reliable transmitted symbols are available as the input feature while the training signal could be looped back from the receiver. Due to extremely long transmission distances in the transoceanic links, the received signal is usually at low signal-to-noise ratio (SNR), preventing DNNs to converge effectively. To mitigate the ASE impact, identical training patterns were sent in the training stage and averaged at the receiver. The training signal was then defined as the difference between the transmitted symbols ($s_{x,n}$, $s_{y,n}$) and the averaged received data patterns (symbols), such that the training signals of DNNs are dominated by nonlinear fiber impairments. In our experiment, 14 of the received patterns each with a pattern length of 2^{14} were averaged together, improving SNR by 11.5 dB while leaving the fiber nonlinear impairments untouched as shown in Figure 3.



(a) (b)

Figure 3: training signals at the receiver after a deployed 13,212 km transoceanic link without (a) and with (b) pattern averaging.

The cost is calculated as mean square error (MSE) between the transmitted symbols and averaged patterns in the training phase. ADAM optimizer is utilized to optimize the parameters of the DNN for more reliable and faster convergence [12]. In the end, the performance of the trained DNN is evaluated using another set of data pattern with the dropout layer by-passed. The received symbols are then compensated using the trained perturbations terms from the AI-NLC. Finally, the Q factor is evaluated to assess the performance.

3. DP-8QAM TRANSMISSION OVER A DEPLOYED FIBER LINK

For the experimental investigation, a 4-subcarrier multiplexed DP-8QAM signal at 21 Gbaud is transmitted over a 13,212 km deployed fiber link. Asymmetric chromatic dispersion (CD) compensation is used, with 40% of the total dispersion pre-compensated at the transmitter, and the remaining 60% post-compensated at the receiver. Following the receiver CD compensation, the DSP include timing recovery, polarization mode de-multiplexing, and carrier phase recovery. Following the carrier recovery, the soft decoded symbols are separated into in-phase and quadrature components to be fed to 8 standalone DNNs (that's for the 4 subcarriers and the 2 polarizations) for fiber nonlinearity compensation.

In the experiment, the eight DNNs are trained independently using 229376 symbols from x-polarization and another 229376 symbols

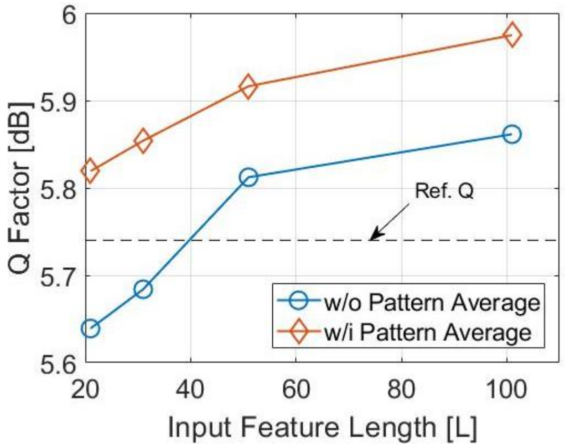


Figure 4: Q factor versus input feature length with soft decoded inputs for AI-NLC nonlinearity compensation. Without AI-NLC, reference Q is 5.74 dB.

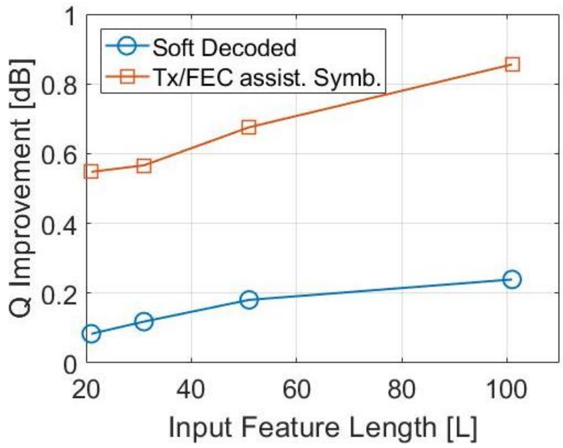


Figure 5: Q factor improvement of DNN trained on transmitter symbols versus soft decoded symbols. Without AI-NLC, the reference Q is 5.74 dB.

from y-polarization. These symbols are formed from 14 repetitive data patterns each having 2^{14} symbols. The averaged received symbols are then subtracted from the transmitter side symbols to provide reliable nonlinear impairments for DNN as reference signal while training. Figure 3 shows the SNR improvement with/without the pattern averaging.

In order for the DNN to learn about fiber nonlinearities in transoceanic transmission, considerable input features are needed by the DNN. As shown in Figure 4, higher Q-factors are achieved with increased number of input features. In Figure 4, 101 symbols, or 404 input features (*xi*, *xr*, *yi* and *yr*) on each subcarrier are enough to achieve 5.97dB Q with pattern averaging enabled. Further increment in the input features will saturate the benefit of the AI-NLC. Due to less channel memory length, we were able to use much less input features on each subcarrier compared with other literatures reported with similar transmission links [8]. This is due to the fact that 40% CD was pre-compensated and the nature of the lower baud rate in the subcarrier systems. Without pattern averaging for ASE mitigation during training, only 0.13 dB Q improvement has been observed with respect to the reference Q of 5.74dB (which confirms that the excess ASE alters the learning process of DNNs about fiber nonlinearity). With input length of 21 and 31, it even introduces more penalty.

Similar to perturbation pre-compensation (PPD), DNN is more effective to be implemented at the transmitter side than the receiver side, but requires a feedback loop to provide a cost function signal from the receiver side for training the DNNs. Another option would be to embed the DNN within the FEC decoder. During the FEC iterations, the hard decoded signal with lower BER can be used as DNN features to further remove the nonlinearity impact. As an open discussion, we compare the DNN trained on the transmitter signal with the case where only receiver signal is available and pattern averaging is switched on. In Figure 5, Q improvement of the receiver-side AI-NLC saturates at 0.24 dB when input length is 101. Without the impact of ASE noise, the Tx symbols assisted AI-NLC provides 0.55 dB Q improvement even with an input length of L = 21. With same 404 input features, the improvement increases to 0.86 dB.

Note that inter-subcarrier nonlinearity could also be estimated by feeding neighbouring subcarrier signals into the DNN. However, it is out of the scope of this paper and can be investigated in future.

4. CONCLUSIONS

We proposed a low complexity pattern averaging assisted AI-NLC technique for fiber nonlinearity compensation using only singlets (no triplets fed to the DNN). The algorithm was experimentally tested using a 21 Gbaud 4-subcarrier multiplexed signal transmitted over a deployed 13,212 km transoceanic fiber link. Results show that with 101 input symbol memory length, 0.24 dB Q-factor improvement can be achieved and 0.86 dB Q improvement can be obtained if transmitter symbols are available (e.g., pilot or training symbols). With pattern averaging for ASE mitigation during the training phase, asymmetric CD compensation and subcarrier multiplexing, the number of input features to the DNN has been greatly reduced. To our best knowledge, this is the longest deployed transmission link with AI-NLC for fiber nonlinearity compensation.

5. REFERENCES

[1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," IEEE J. Lightwave Technology vol. 28, no. 4, pp. 662-701, February 2010.

[2] E. Ip, and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," IEEE J. Lightwave Technology vol. 26, no. 20, pp. 3416-3425, October 2008.

[3] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussenet, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," IEEE J. Lightwave Technology vol. 29, no. 17, pp. 2570-2576, September 2011.

[4] K. V. Peddanarappagari, and M. Brandt-Pearce, "Volterra series transfer function of single-mode fibers," IEEE J. Lightwave Technology vol. 15, no. 12, pp. 2232-2241, December 1997.

[5] D. Zibar, O. Winther, N. Franceschi, R. Borkowski, A. Caballero, V. Arlunno, M. N. Schmidt, N. G. Gonzales, B. Mao, Y. Ye, K. J. Larsen, and I. T. Monroy, "Nonlinear impairment compensation using expectation maximization for dispersion managed and unmanaged PDM 16-QAM transmission," OSA Optics Express vol. 20, no. 26, pp. B181-B196, December 2012.

[6] T. S. R. Shen, and A. P. T. Lau, "Fiber nonlinearity compensation using extreme learning machine for DSP-based coherent communication systems," Proc. Opto-Electronics and Communications Conference (OECC), Kaohsiung, Taiwan, pp. 816-817, July 2011.

[7] V. Kamalov, L. Jovanovski, V. Vusirikala, S. Zhang, F. Yaman, K. Nakamura, T. Inoue, E. Mateo, and Y. Inada, "Evolution from 8QAM live traffic to PS 64-QAM with neural-network based nonlinearity compensation on 11000 km open subsea cable," Proc. Optical Fiber Communication (OFC) Conference, San Diego, CA, Paper Th4D.5, March 2018.

[8] S. Zhang, F. Yaman, E. Mateo, and Y. Inada, "Neuron-network-based nonlinearity compensation algorithm," Proc. European Conference on Optical Communications (ECOC), Roma, Italy, Paper Tu1F.5, September 2018.

[9] Y. Gao, Z. A. El-Sahn, A. Awadalla, D. Yao, H. Sun, P. Mertz, and K.-T. Wu, "Reduced complexity nonlinearity compensation via principal component analysis and deep neural networks," Proc. Optical Fiber Communication (OFC) Conference, San Diego, CA, Paper Th2A.49, March 2019.

[10] O. Sidelnikov, A. Redyuk, and S. Sygletos, "Nonlinear equalization in long haul transmission systems using dynamic multi-layer perceptron networks," Proc. European Conference on Optical Communications (ECOC), Roma, Italy, Paper Tu4F.3, September 2018.

[11] E. P. da Silva, M. P. Yankov, T. Morioka, and L. K. Oxenlowe, “FEC-assisted perturbation-based nonlinear compensation for WDM sytems,” Proc. Optical Fiber Communication (OFC) Conference, San Diego, CA, Paper W3A.3, March 2018.
[12] Y. Bengio, “Practical recommendations for gradient-based training of deep architectures,” Neural Networks: Tricks of the Trade, K.-R. Muller, G. Montavon, and G. B. Orr, eds., Springer 2013.

