DSP optimization technique based on deep learning to account for the effects of fiber nonlinearity

Jing Yang *

Shanghai University, Shanghai, China

* Corresponding Author Email: j417290076@shu.edu.cn

Abstract. The first section of this work is the classification of digital signal processing (DSP) techniques for adjusting, minimizing, and utilizing fiber nonlinearities in coherent fiber optic transmission systems. Then, different experimental optimization methods based on the Volterra series of Deep Neural Networks (DNN) and the Digital Back Propagation (DBP) methods are contrasted with the aid of specific instances. Concrete examples reinforce the techniques. The findings demonstrate the viability of a DSP strategy using deep learning for inexpensive optical comms, resulting in significant performance gains and computational efficiencies. The optimal DSP, it is further determined, should balance minimizing transmission flaws with minimizing additional distortions brought on by ineffective linear or nonlinear correction steps as a result of in-line amplifier noise. It is demonstrated that, in addition to improving system performance, machine learning (ML) improves analytical reasoning, provides deeper mathematical comprehension of fiber nonlinearity correction, and advances theoretical knowledge in related fields.

Keywords: DSP, nonlinearities, DNN, optimization.

1. Introduction

This thesis focuses on intra- and inter-channel nonlinearities in optical channels, a long-standing basic bottleneck in optical communications, and the urgency of research breakthroughs in optical communications to fulfill the current and future rapidly expanding connectivity needs. For a given level of improvement, multiple strategies have been developed to simplify nonlinear compensation (NLC) algorithms, some of which, like filtering DBPs [1], do so by several orders of magnitude. A new strategy has recently been tried using machine learning algorithms, but these efforts are still in their early stages. Research and development (R&D) in deep learning, a branch of machine learning (ML) that employs neural networks with biologically influenced architectures to carry out learning tasks, is responsible for the emergence of AI since 2012 and its technological advancements [2]. Deep neural networks (DNNs) have recently set records for performance in various functions of communication systems, including speech and object recognition [3]. Jack et al [4] contend that some hierarchical or Markovian process produces the majority of data of practical significance, typically adhering to physical criteria like locality and symmetry, to explain this success. Based on this, DNNs have been used in a number of papers to solve fiber nonlinearities in long-range coherent links and O/E characterization [5-7] components of short-range IM/DD systems [8-10]. Deep learning techniques [11] are regarded as a competitive alternative to traditional digital signal processing from a physical layer viewpoint (DSP) [12, 13]. Many machine learning (ML) solutions in fiber section are amazing "black-box" data-driven models with unmatched properties, or, they don't really shed any new light on the problem.

Recently people have witnessed a huge peak in interest in deep learning and deep neural networks (NNs). This summary paper outlines recent work on using deep neural networks for fiber optic communications. This paper focuses on ML. Firstly, it generates easy-to-implement algorithms to improve system performance, and secondly, it complements analytical thinking by providing deeper mathematical insights into fiber optic nonlinear compensation, which in turn helps to advance theoretical understanding in related fields. In the first section here, digital signal processing methods that work with fiber optic nonlinearities are checked. Then, two distinct digital signal processing (DSP) strategies based on deep neural networks (DNNs) to account for nonlinearities for long- and

short-haul fiber optic communications are presented and contrasted, with specific examples to highlight machine learning as a potent tool to advance the field of optics among other disciplines.

2. Physics-Based Deep-Learning Methods

The standard configuration of feedforward NNs is briefly reviewed at the beginning of this section. It is crucial to stress that in our work, instead, the fundamental driving factor behind the strategy described in the following section is their underlying mathematical structure.

2.1. Deep Neural Networks

Multiple levels of artificial neurons are used to construct the computing structures of DNNs [14]. As seen in Fig. 1, the processing unit known as a single artificial neuron has many different inputs and a single output.

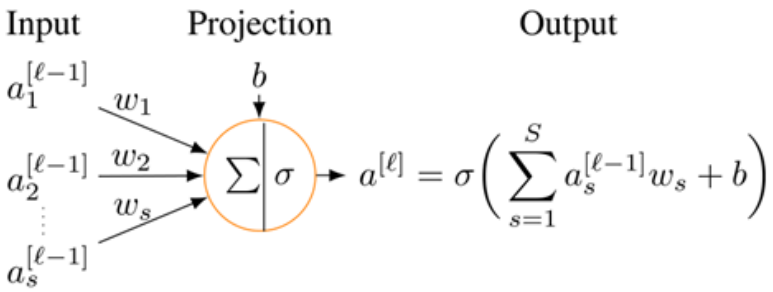


Figure 1. Single artificial neuron.

The parametric mapping $y = f_\theta(x)$, which is a parametric vector, is defined by a feedforward neural network with l levels. Additionally, the first layer's input, $x \in X$, could mapped tthe $y \in Y$ by a succession as fine transformations and point-wise nonlinear mappings [2] [4].

$$\tilde{x}^{(i)} = W^{(i)} x^{(i-1)} + b^{(i)}, i = 1,2,\ldots,\ell \quad (1)$$

where $x \triangleq x(0)$ is the first layer's input, $W^{(i)}$ is mapped to the output, and $b^{(i)}$ via a succession of fine transformations and point-wise nonlinear mappings $x^{(i)} = \sigma(\tilde{x}^{(i)})$ and σ refers to applying some nonlinear activation function σ. Nonlinear activation functions that are frequently used include $\sigma(z)=\max$ [15], $\sigma(z)=1/(1+e^{-z})$, and $\sigma(z)=\tanh(z)$.

2.2. Volterra Equalizer & Digital Back Propagation

The main cause of nonlinearity in short-range fiber optic transmission is electrical components. Although the construction of Volterra equalizers is typically highly complex, the Volterra family is a well-liked countermeasure for balancing the damage of nonlinear components and their memory effects on the receiver side [16]. Nonlinear equalization can be replaced by deep neural network (DNN) topologies.

Fig.2 and Fig.3 show a previously published concept [17] regarding the above, namely the one-to-one replacement of VNLE with DNNE.

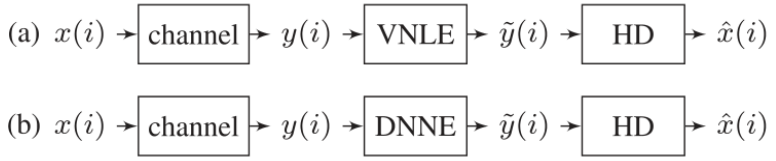


Figure 2. Hardecision (HD) demapper and (a) VNLE or (b) DNNE.

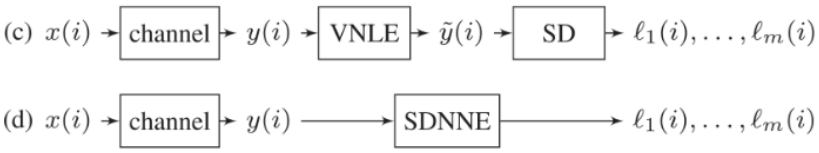


Figure 3. Soft Deep Neural Network Amplifier in Combination with VNLE; or (d) Soft Decision (SD) Demapper (SDNNE).

The requirement to instantly invert the nonlinear Schrödinger equation (NLSE), also known as Digital Back Propagation, is a crucial challenge in fiber optic communication (DBP). Intriguingly, the widely used split-step Fourier method (SSFM) produces a constructed graph resemble deep neural networks (DNNs). As a result, the linear steps can be thought of as general linear functions by parameterizing the SSM [18], which could seen as the weight matrix of a neural network. This is so because a deep multilayer neural system's functional form and the numerical answer to the nonlinear Schrödinger equation are very similar. Deep learning is used at each step to jointly optimize all filters, using physically based models and trainable filters.

Several studies make the assumption that the signal x(t) is discharged and moves along it following the NLSE, which seen in Fig. 4. The constants β_2 and γ are stand in for the fiber's nonlinear Kerr parameter and chromatic dispersion coefficient. For the sake of simplicity, the loss term $\alpha u/2$ is neglected. The signal x(t) or a sampled representation of it is successfully recovered from $\{y_k\}_{k\in Z}$.

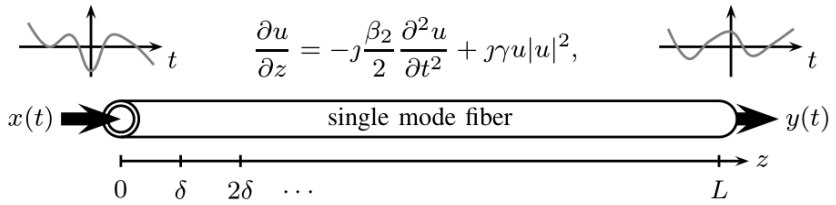


Figure 4. Evolution of the signal conceptually using single-mode fibers. The output signal y(t) = u(z = L,t) as well as the input signal x(t) = u(z = 0,t) are both explicable by the Schrödinger equation.

2.3. The Split-Step Method

By accurately parameterizing the SSM and utilizing the resulting model but a general NN, the main objective is to make advantage of SSFM. Consider the NLSE-propagated complex baseband signal x(t).[19].

$$\frac{\partial u(z,t)}{\partial z} = -\frac{\alpha}{2}u(z,t) - \int\frac{\beta_2}{2}\frac{\partial^2}{\partial t^2}u(z,t) + \int\gamma|u(z,t)|^2u(z,t) \quad (2)$$

where the loss parameter is α in u(z=0,t)≙x(t), the CD coefficient is β_2, and the nonlinear Kerr parameter is γ. The signal is represented as y $y(t) \triangleq u(z = L, t)$ after propagation distance.

Generally speaking, closed-form solutions are not allowed by (2). SSM, one of the most popular methods, works by processing sampled waveforms (i.e., discrete-time waveforms). Considering that x(t) and y(t) signals have sampling point at a rate of k/f_s, the sample sequences $\{x_k\}_{k\in Z}$ and $\{y_k\}_{k\in Z}$ respectively, are further mapped to another vector $x = (x_1, \ldots, x_n)^\top$, $y = (y_1, \ldots, y_n)^\top$, which $x, y \in C^n$, and the time-accurate NLSE is considered in this study[20].

$$\frac{du(z)}{dz} = -\frac{\alpha}{2}u(z) + Au(z) + \int\gamma\rho\big(u(z)\big) \quad (3)$$

where $A = F^{-1}\mathrm{diag}(H_1, \ldots, H_n)F$. The kth DFT corner frequency is $H_k = \int\frac{\beta_2}{2}\omega_k^2$ and $\omega_k = 2\pi\varphi_k$(i.e.,if $k < n/2, f_k/f_s = \frac{(k-1)}{n}$, if $k \geq n/2$, $f_k/f_s = (k-1-n)/n$). $\rho: C_n \to C_n$ is defined

by using the formula $\rho(x) = x|x|^2$. If $\gamma = 0$, then (3) is linear and the answer is $u(z) = A_z u_0$, where $A_z \triangleq e^{-\frac{\alpha}{2}z} e^{zA}$ and $e^{zA} = F^{-1}\mathrm{diag}(e^{zH1}, \ldots, e^{zH1})F$. It is possible to demonstrate that the approach is $u(z) = \sigma_z(u_0)$ for $\alpha = 0$ and $\beta_2 = 0$, where $\sigma_z: C_n \to C_n$ is the element-by-element application of $\sigma_z(x) = xe^{\int \gamma z|x|^2}$.

SSM can be obtained by alternating between these two operators.

$$u_i = \sigma_{\delta_i}(A_{\delta_i} u_{i-1}), i = 1,2,\ldots,M, \tag{4}$$

where the estimate u_i is given as $u_i \approx u(\delta 1 + \ldots + \delta i)$. In reality, sufficient temporal and spatial resolution must be guaranteed by f_s and M.

3. Discussion of DNN-based DSP

3.1. DSP of Fiber Nonlinearities

First, this paper reviews five well-known digital signal processing techniques for coherent fiber optic transmission's fiber nonlinearity in an article.

The pre-compensation technique can be implemented with a sample of symbols and describes the fiber nonlinearity as a first-order perturbation term [21-23]. Fig.5 illustrates the bit error rate (BER) dependence in the transmit wattage. RRC-SEDC stands for cosine pulse shaping based and precompensated symmetric linear pre-summed post-compensated dispersion based on perturbation. LC stands for linear post-compensated decentralization, LC-SEDC, and LC-SEDC for symmetric linear precompensated post-compensated dispersion [24]. The outcomes demonstrate that the perturbation-based technical formulation can reduce execution complexity and not damage to sacrificing performance.

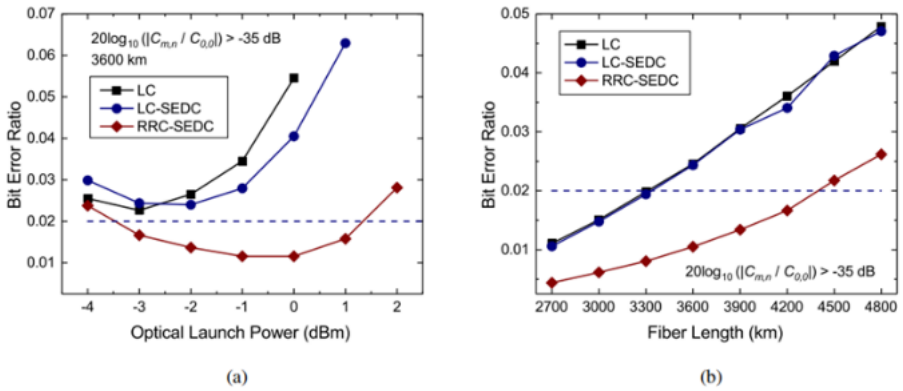


Figure 5. Fiber length dependence of BER at optimal launch power

DBP is the most widely used DSP skill for flaws compensation. Signal-signal nonlinear interference (NNLI) effects can be fully determined by this approach, which makes it effective. As further enhancing the NLC bandwidth does not seem feasible at present. Analytical techniques reveal a decline in the performance of multi-channel DBP.

To accurately characterize the functionality of DBP in broadband transmission, it is fair to use a closed-form formulation known as the improved GN model (EGN). The relationship from SNR to transmitted power is depicted in Fig. 6. In comparison, the full-field DBP gain has a mild reliance on ASE noise ($P_{ASE}^{-\frac{1}{3}}$) and transmission bandwidth ($\eta^{-1/6}$) and significant dependence on transmission distance ($N_S^{-\frac{1}{2}}$).

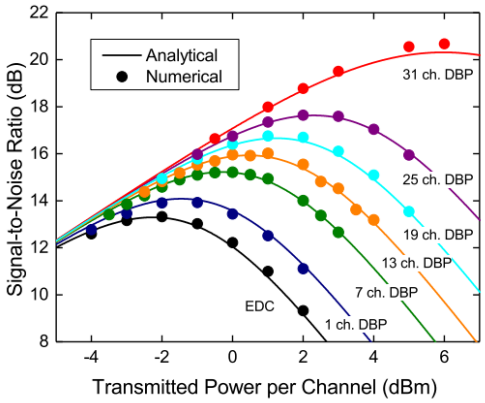


Figure 6. DBP SNR performance in 31×32G baud PM-16QAM channels.

In conclusion, this work demonstrates that the currently known closed-form expressions could precisely forecast the receiver SNR adopting multichannel DBPt to correct for intra- and interchannel NLI. This is done by comparison with SSFM results.

The NLSE problem was solved in the frequency domain using VSTF, a tool modeling and compensating nonlinear dynamic processes, to derive a set of nonlinear transfer functions for discharging optical fibers [25]. Examples are based on [26, 27] proposals for a streamlined Volterra series nonlinear equalizer (VSNE) with a precise reconstruction of the K3 kernel that reduces computational complexity by 3 times without sacrificing performance [27], the results of which are displayed in Fig. 7.

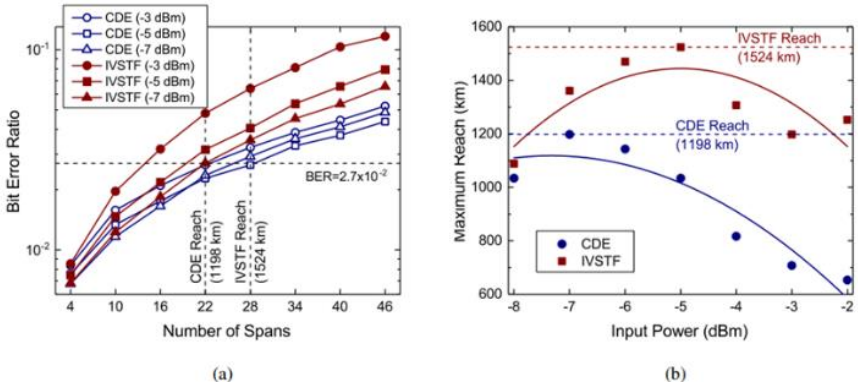


Figure 7. Maximum signal range and BER performance (a) BER versus figure of spans at various transmit power channels; (b) maximum arrival rate versus transmit power.

By evaluating the possible information rate (AIR), advanced modulation formats influence the performance of systems. The signal-to-noise & interference ratio is grew through this method's constellation rule for reducing nonlinear interference noise (NLIN) (SNR). Still, because of its design restrictions, the entropy H(XK1) is also reduced, creating a conflict between geometric and probabilistic design. It is demonstrated that the increased dimensionality makes it difficult to optimize the PMF.

The signal nonlinear Fourier spectrum is used by mathematical foundation underlying the nonlinear Fourier transform to encode information (additionally known as the IST (inverse scattering transformer)). It has been shown in numerous recent articles [28-31] that nonlinearities can greatly reduce the effectiveness of traditional transmission techniques. Even though these NFT-based systems have a definite promise for generating reasonably significant spectrum efficiencies and enabling transmission with higher transmit power than current techniques, there is still a lot of work to be done before they can be considered competitive in practice.

3.2. DNN-based Volterra Equalizer

This part, which follows the preceding one, provides an experimental illustration of traditional nonlinear Volterra equalization utilizing feedforward or modern neural network optimization. Hence, it is demonstrated that DSP-based deep learning can achieve enhanced experimental performance and system efficiency.

Compared to conventional application models, end-to-end machine learning could be a practical DSP solution for short-range fiber optic communication [12,13]. The SBRNN autoencoder experimental study is extended in this work with a simple approach that looks into further performing better when transmission link parameters are learned on the channel model. Deep learning is helpful for low-cost optical section, the study's findings suggest, as it can boost arrival rates or enhance data transmission rates over shorter distances.

The experimental setup is based on an IM/DD testbed sending T=5152 random messages of input, $m_{i,j} \in M$ ($M = |M| = 64$) , in Z=800 distinct sequences. An example message carry of 2M=6 bits is encoded using the BRNN transmitter. The tested data set and methods to optimize the test bed parameters are shown in Fig. 8.

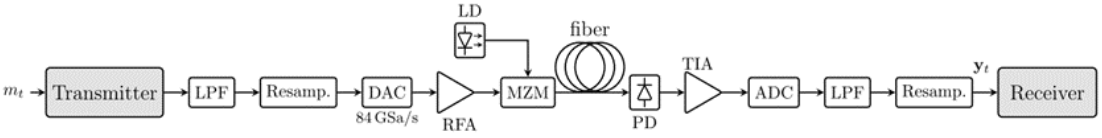


Figure 8. Experimental optical IM/DD transmission test-bed designs that were employed

This study focuses on the experimental performance of the receiver-only application. For the optimization and testing of the receiver, the experiment uses a sequence with $i \in \{1,\ldots,Z\}$ and $j \in \{1,\ldots,T\}$. The received symbol sequences $y_{i,j}$ are used to optimize and test DSP parameters together with the labels $m_{i,j}$ after transmission across the test bed link. The training set is composed of the sequences $i \in \{1,\ldots,90\}$.

$$D^{PAM} := \begin{bmatrix} y_{1,1} & \cdots & y_{1,T/2} \\ y_{1,T/2+1} & \cdots & y_{1,T} \\ \cdots & \cdots & \cdots \\ y_{90,1} & \cdots & y_{90,T/2} \\ y_{90,T/2+1} & \cdots & y_{90,T} \end{bmatrix} \tag{5}$$

$$L^{PAM} := \begin{bmatrix} m_{1,1} & \cdots & m_{1,T/2} \\ m_{1,T/2+1} & \cdots & m_{1,T} \\ \cdots & \cdots & \cdots \\ m_{90,1} & \cdots & m_{90,T/2} \\ m_{90,T/2+1} & \cdots & m_{90,T} \end{bmatrix} \tag{6}$$

where $D^{PAM} \in R^{180\times T*nc}$ makes up data elements $y_{i,j}$ and $L^{AE} \in R^{180\times 4*T}$ contains a large quantity of elements $m_{i,j}$.

The communication channel is nonlinear and requires processing to restore the data sequence to the interaction of the IM/DD transmission's dispersion and the photodiode's square law detection [32]. Hence, both traditional DSP and deep learning-based solutions are examined in this section.

The BRNN autoencoder is an entire technology chain that implements an end-to-end deep ANN, concerning the channel, the transmitter, and the receiver [33]. Fig.9 (a) shows a schematic of the BRNN building blocks of the transmitter and receiver, while the end-to-end deep learning processing and data sequence estimates are schematically in Fig. 9(b). The training performance of the aforementioned system is displayed in Fig.10 with a sliding window with a fixed size of W=10. The training work utilizes the sliding window BRNN PAM receiver as well as the sliding window FFNN

PAM receiver. Experiments and simulations [34] show that the benefit from subsequent increases in W decreases with time.

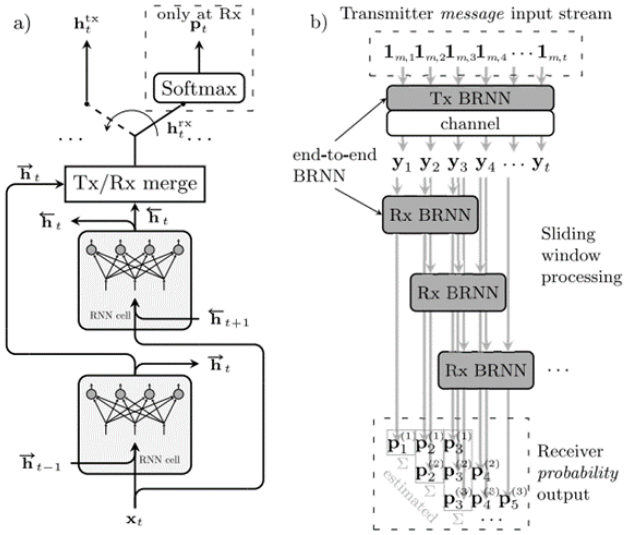


Figure 9. a) A schematic diagram depicts the two-way RNN, b) the sliding window series estimation algorithm.

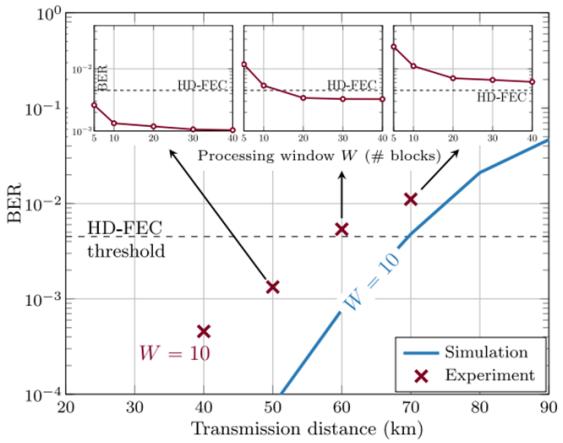


Figure 10. SBRNN autoencoder trained on channel model: BER against distance

As shown in Fig. 11, the Volterra equalizer, which is popular for PAM-based IM/DD links and is comparable to other PAM algorithms, is utilized in this research. This work used SFFNN in an 84 Gb/ss system with PAM4 and trained various distance and multi-distance learning techniques for DSP optimization.

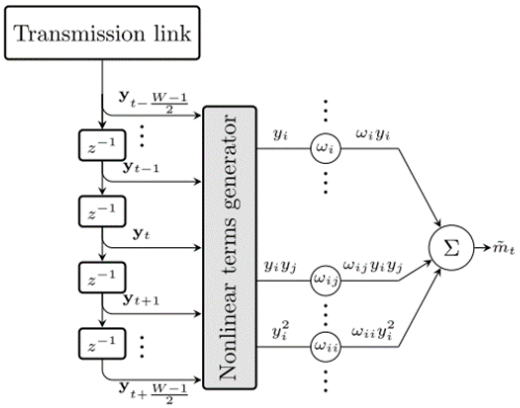


Figure 11. The nonlinear Volterra equalizer for PAM symbols is depicted in a schematic figure.

3.3. DNN-based DBP

This section is on the basis of the Digital Back Propagation (DBP) algorithm, currently the best as well as most popular DSP for nonlinear compensation in fiber optic communication [1,35,36]. It indicates that machine learning (ML) could create simple and easy algorithms to improve the system's performance (previous section) and provide insightful mathematical insights for nonlinear compensation.

The DNN-based DBP architecture presented in this research is a novel example for long-haul optical communication systems. The implementation of low-complexity for DNN-based DBP is found to have useful gain in a single-channel 28-G Baud 16-QAM system, while the best DBP performance is of arbitrary complexity. The Q-factor gains of DNN-based low-complexity DBP and arbitrary complexity DBP in single-channel PDM and WDM-PDM systems, on the other hand, are 0.6 dB and 0.25 dB, suggests that the ideal DSP should balance mitigating for transmission impairment with the additional distortion caused by DBP. The ideal parameter configuration examines how CD, nonlinearity, and noise interact and theorizes that transceiver DSPs shouldn't totally subvert the linear and nonlinear fiber propagation effects. This study is an important step toward bringing nonlinear compensation DSPs to realistic WDM systems.

Contrary to the usual DBP implementation, the DBP's linear and nonlinear interleaving stages can be considered, Fig.12 illustrates the the multilayer neural network's linear and nonlinear operations. The nonlinear operation in the instance of the figure primarily refers to nonlinear phase inference, where all filters are made up of parameters and are amenable to optimization by machine learning methods.

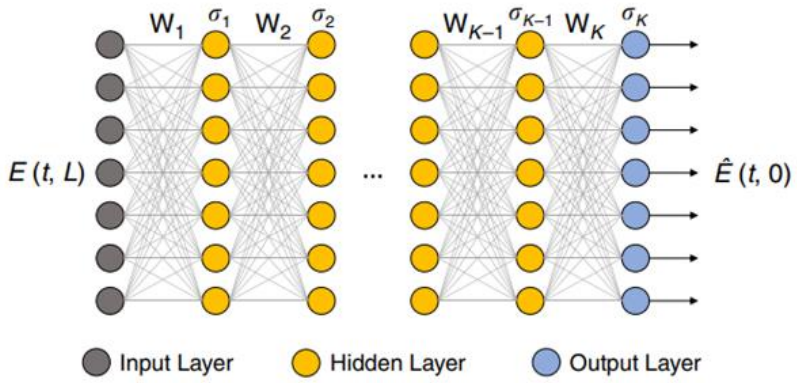


Figure 12. DNN as DBP

The DNN-based DBP needs a number of adjustments, which this study suggests.

1) The training data were split into the initial 32,768 symbols and the remaining 32,768 symbols in order to free for duplicating the test set for the training data (25 copies).

2) Use overlapping preservation in the the execution while concerning the time series as well as dispersive character for the incoming signals [37].

3) The neural networks' input blocks were created from the 65,536 training samples. The length of the linear filter, initial or final study rates of the Ada Bound optimizer, and the input vector size were all modified together.

4) To properly account for impulse augmentation effects in the overlap and preservation methods, 60 adjacencies were added to the ends of each input.

5) Changing the computation to real multiplication and addition in order to enable the optimization objective to be handled by a widespread deep learning framework like Tensor flow [38].

Fig.13 depicts the structure of additional DSP blocks in a coherent digital receiver along with the testing and training phases of the DNN-based DBP.

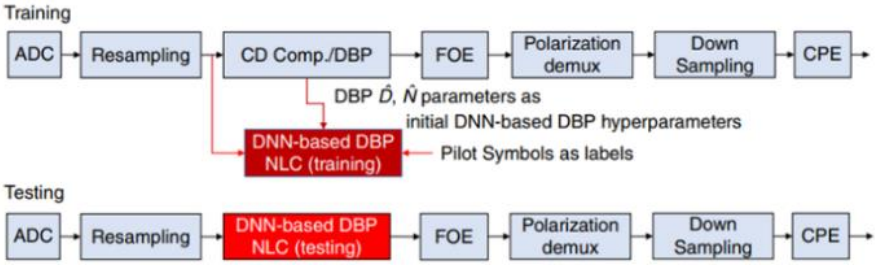


Figure 13. Comprehensive DBP structure based on DNN, a frequency offset estimate is a FOE, while a carrier phase estimate is a CPE.

The neural network model from Häger and Pfister's15 study is successfully expanded in this study to PDM systems for PDM and WDM tests. This is accomplished by fusing methods for ML laser and polarization damage compensation with neural networks. Fig.14 compares the Q-factor for 28-G Baud PDM 16-QAM transmission over an 815 km distance utilizing pure CDC, DBP based on DNN and DBP with different step sizes. A linear filter with the same quadratic phase diagram for the CD compensated filter is shown in b and c to have "M" and "U" shaped amplitude response characteristics.

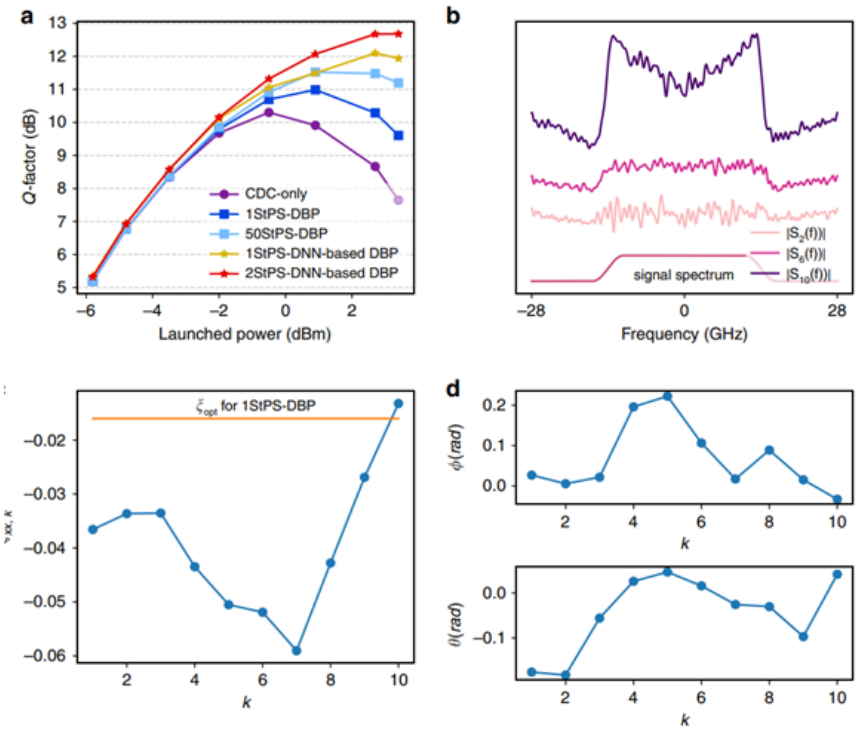


Figure 14. For single-channel PDM transmissions, the results

3.4. Fiber optic communications DSP optimization for nonlinearities comparison

Five digital signal processing approaches were employed in the first review stated above in this study to account for and take advantage of the effects of fiber nonlinearity in coherent fiber transmission systems. Table 1 compares these techniques, highlighting the methods used, the issues resolved, the benefits of the technique, and the areas that could use further development.

It is important to note that the two research publications described above used the Volterra equalization and Digital Back Propagation (DBP) methodologies. Volterra sequence is a notable DSP technique for modeling and compensating nonlinear dynamic processes, while Digital Back Propagation (DBP) is now the most well-liked DSP technique.

Table 1. Performance Comparison of DSP Technologies

DSP	Perturbation	DBP	Volterra	Advanced modulation	Fourier encoding
Method	Quantification	Closed expressions	VSTF	AIR	IST
Trouble	cross-channel nonlinearity	Linear transmission defects	NLSE problems		Nonlinear Fourieris spectrum
Advantage	Further reducted calculations	prediction of receiver SNR	Same homogeneity Lower computation	Improved tolerance Allowed equalization	Greater spectral efficiency and higher transmit power
Optimization	Extended Algorithm		Scaling algorithms Scaling algorithms	optimized constellation properties	Interoperability

The experimental and theoretical comparisons between the two research articles are shown in Table 2. Included are the methodologies utilized in the study, the scenarios applied to the study, the employed DSP techniques, the experiments carried out, and the experimental conclusions. " Short-Range Fiber Optic Communication Using Digital Signal Processing" is number 1 and number 2 has a topic of machine learning is being used to advance theoretical understanding and practical performance of signal processing for nonlinear optical communications.

Paper 1 focused on the experimental performance enhancement method of the RNN autoencoder. It is thoroughly studied, focusing on the performance enhancement of sliding window sequence estimation (SBRNN) autoencoder, and the research process is more explicit in Paper 1. More importantly, the neural network part of the paper used feedforward or current neural networks, the DSP algorithm used a classical nonlinear Volterra equalizer. The objective of Paper 2 was to optimize the standard DBP to DNN-DBP in an idealized simulation environment for fiber nonlinear compensation, interleaving linear and nonlinear operations. In contrast, Paper 2 focused more on theoretical explanation and included experimental comparison. Later, in Paper 2, this idea is expanded to include the single channel, polarization division multiplexing, and wavelength division multiplexing tests in Table 2, to provide a mathematically-based theoretical justification for the ML approach to fiber optic communication nonlinearity.

The two pieces of work share a lot in common because they both attempt to use deep learning to optimize DSP techniques to solve the fiber optic transmission nonlinearity issue. These studies demonstrate how machine learning is a promising research topic and helps to enhance theoretical knowledge and technological guidance in fields like optics.

Table 2. Performance comparison of DSP technologies

Paper number	1	2
Scenario	Short-range fiber optic communication	Long-range fiber optic communication
Method	Nonlinear Volterra equalization	Optimized DBP
Experiment	SBRNN fiber autoencoder Single channel	polarization division multiplexing and wavelength division multiplexing
Conclusion	Achieved improved BER performance	Optimal DSP should strike a balance between suppressing distortion and overcoming fiber propagation effects
	Allows to increase the total capacity of the system	This distortion is transferred in a quantifiable way in different DBP stages

4. Conclusion

This paper compares Volterraequalization for deep learning of experimental performance in short-haul fiber optic cables with that for receivers only, based on a review of DSPs for compensating fiber nonlinearities. This comparison shows that, for low-cost optical communications, deep learning can be a useful DSP solution to increase reach or data transfer speeds over shorter distances. In long-haul fiber optic communications, Examining the experimental results of configurations combining digital backpropagation algorithms with deep neural networks' linear and nonlinear operators, demonstrating the network parameter optimization capabilities of machine learning techniques together with the use of DNN-based DBP and PDM-WDM systems. As a result, properties is significantly improved, and calculation costs are reduced, enabling the introduction of nonlinearly compensated DSP into realistic long-haul fiber optic communications. Notably, the results show that the DSP should balance mitigating transmission flaws with reducing further distortions brought on by ineffective linear and nonlinear correction steps brought on by in-line amplifier noise. It is shown that machine learning techniques can give a means for real systems to intelligent solutions, move beyond traditional thinking, and create algorithms that go beyond the most recent performance for nonlinear fiber optic transmission. Many fields might benefit greatly from this area of inquiry.

References

[1] B. D. Liang and A. J. Lowery, "Improved single channel backpropagation for intra-channel fiber nonlinearity compensation in long-haul optical communication systems," Optics Express, vol. 18, no. 16, pp. 17075-88, 2010.

[2] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, p. 436, 2015.

[3] S. Lefkimmiatis, "Non-local Color Image Denoising with Convolutional Neural Networks," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[4] H. W. Lin, M. Tegmark, and D. Rolnick, "Why does deep and cheap learning work so well?," 2016.

[5] V. Kamalov, L. Jovanovski, V. Vusirikala, S. Zhang, and Y. Inada, "Evolution from 8QAM live traffic to PS 64-QAM with Neural-Network Based Nonlinearity Compensation on 11000 km Open Subsea Cable," in Optical Fiber Communication Conference, 2018.

[6] T. Koike-Akino, D. S. Millar, K. Parsons, and K. Kojima, "Fiber Nonlinearity Equalization with Multi-Label Deep Learning Scalable to High-Order DP-QAM," 2018.

[7] S. Zhang et al., "Field and lab experimental demonstration of nonlinear impairment compensation using neural networks," Nature Communications, vol. 10, no. 1, p. 3033, 2019/07/10 2019, doi: 10.1038/s41467-019-10911-9.

[8] A. G. Reza and J. Rhee, "Nonlinear Equalizer Based on Neural Networks for PAM-4 Signal Transmission Using DML," IEEE Photonics Technology Letters, 2018.

[9] R. Rios-Müller, J. M. Estaran, and J. Renaudier, "Experimental estimation of optical nonlinear memory channel conditional distribution using deep neural networks," in Optical Fiber Communications Conference & Exhibition, 2017.

[10] P. Li, L. Yi, L. Xue, and W. Hu, "100Gbps IM/DD Transmission over 25km SSMF using 20G-class DML and PIN Enabled by Machine Learning," in Optical Fiber Communication Conference, 2018.

[11] I. Goodfellow, Y. Bengio, and A. Courville, "Regularization for Deep Learning," 2016.

[12] O. Simeone, "A Very Brief Introduction to Machine Learning With Applications to Communication Systems," ed: IEEE, 2018, pp. 648-664.

[13] F. N. Khan, Q. Fan, C. Lu, and A. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," Journal of Lightwave Technology, vol. PP, no. 99, pp. 1-1, 2019.

[14] C. M. Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics). Pattern Recognition and Machine Learning (Information Science and Statistics), 2006.

[15] H. Alva-Sanchez et al., "A Small-Animal PET System Based on LYSO Crystal Arrays, PS-PMTs and a PCI DAQ Board," IEEE Transactions on Nuclear Science, vol. 57, no. 1, pp. 85-93, 2010.

[16] M. Schaedler, G. Bcherer, and S. Pachnicke, "Soft-Demapping for Short Reach Optical Communication: A Comparison of Deep Neural Networks and Volterra Series," Journal of Lightwave Technology, vol. PP, no. 99, 2021.

[17] C. Bluemm, M. Schaedler, S. Calabrò, G. Charlet, and M. Kuschnerov, "Equalizing nonlinearities with memory effects: Volterra series vs. deep neural networks," in 45th European Conference on Optical Communication (ECOC 2019), 2019.

[18] C. Hger and H. D. Pfister, "Deep Learning of the Nonlinear Schr"odinger Equation in Fiber-Optic Communications," 2018.

[19] G. P. Agrawal, Nonlinear Fiber Optics (Fourth Edition). Nonlinear Fiber Optics (Fourth Edition), 2006.

[20] C. Hger and H. D. Pfister, "Physics-Based Deep Learning for Fiber-Optic Communication Systems," IEEE Journal on Selected Areas in Communications, vol. PP, no. 99, pp. 1-1, 2020.

[21] Mecozzi, A., Essiambre, and R., "Nonlinear Shannon Limit in Pseudolinear Coherent Systems," Lightwave Technology, Journal of, 2012.

[22] A. Mecozzi, C. B. Clausen, and M. Shtaif, "Analysis of intrachannel nonlinear effects in highly dispersed optical pulse transmission," Photonics Technology Letters IEEE, vol. 12, no. 4, pp. 392-394, 2000.

[23] D. Liang, Z. Tao, L. Lei, W. Yan, and J. C. Rasmussen, "A Low Complexity Pre-Distortion Method for Intra-channel Nonlinearity," in 2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference.

[24] Y. Gao et al., "Reducing the complexity of perturbation based nonlinearity pre-compensation using symmetric EDC and pulse shaping," Optics Express, vol. 22, no. 2, pp. 1209-19, 2014.

[25] M. Schetzen, The Volterra And Wiener Theories of Nonlinear Systems. The Volterra And Wiener Theories of Nonlinear Systems, 1980.

[26] L. Liu et al., "Intrachannel Nonlinearity Compensation by Inverse Volterra Series Transfer Function," Journal of Lightwave Technology, vol. 30, no. 3, pp. 310-316, 2012.

[27] F. P. Guiomar and A. N. Pinto, "Simplified Volterra Series Nonlinear Equalizer for Polarization-Multiplexed Coherent Optical Systems," Journal of Lightwave Technology, vol. 31, no. 23, pp. 3879-3891, 2013.

[28] J. E. Prilepsky, S. A. Derevyanko, and S. K. Turitsyn, “Nonlinear Spectral Management: Linearization of the lossless fiber channel,” Optics Express, vol. 21, no. 20, p. 24344, 2013.

[29] J. E. Prilepsky, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn, "Nonlinear Inverse Synthesis and Eigenvalue Division Multiplexing in Optical Fiber Channels," Physical Review Letters, vol. 113, no. 1, p. 013901, 2014.

[30] S. Hari, F. Kschischang, and M. Yousefi, "Multi-eigenvalue communication via the nonlinear Fourier transform," in Communications, 2014.

[31] I. Tavakkolnia and M. Safari, "Signalling over nonlinear fibre-optic channels by utilizing both solitonic and radiative spectra," in European Conference on Networks & Communications, 2015.

[32] G. P. Agrawal, "Fiber-Optic Communication Systems, 4th Edition," 2010.

[33] T. Oshea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," IEEE Transactions on Cognitive Communications & Networking, vol. 3, no. 4, pp. 563-575, 2017.

[34] B. Karanov, D. Lavery, P. Bayvel, and L. Schmalen, "End-to-End Optimized Transmission over Dispersive Intensity-Modulated Channels Using Bidirectional Recurrent Neural Networks," Optics Express, vol. 27, no. 14, p. 19650, 2019.

[35] D. Rafique, M. Mussolin, M. Forzati, J. Mrtensson, and A. D. Ellis, "Compensation of intra-channel nonlinear fibre impairments using simplified digital back-propagation algorithm," Optics Express, vol. 19, no. 10, pp. 9453-9460, 2011.

[36] N. V. Irukulapati, H. Wymeersch, P. Johannisson, and E. Agrell, "Stochastic Digital Backpropagation," IEEE Transactions on Communications, vol. 62, no. 11, pp. 3956-3968, 2014.

[37] C. Fougstedt, M. Mazur, L. Svensson, H. Eliasson, and P. Larsson-Edefors, "Time-Domain Digital Back Propagation: Algorithm and Finite-Precision Implementation Aspects," in Optical Fiber Communication Conference, 2017.

[38] M. Abadi, P. Barham, J. Chen, Z. Chen, and X. Zhang, "TensorFlow: A system for large-scale machine learning," USENIX Association, 2016.