OLS-Based RBF Neural Network for Nonlinear and Linear Impairments Compensation in the CO-OFDM System

Volume 10, Number 2, April 2018

Ling Liu
Meihua Bi
Shilin Xiao
Jiafei Fang
Tiancheng Huang
Weisheng Hu

DOI: 10.1109/JPHOT.2018.2808919
1943-0655 © 2018 IEEE
Abstract: For the first time, we propose and experimentally verify a novel low-complexity orthogonal least square (OLS) based radial basis function (RBF) neural network (NN) nonlinear equalizer for high-speed coherent optical orthogonal frequency division multiplexing (CO-OFDM) system. Its ability to compensate nonlinear impairments as well as linear impairments is comprehensively evaluated and compared with the linear equalizer. The impact of training overhead on system performance is also investigated. Results show that in a single-channel 40-Gb/s 16-quadrature amplitude modulation CO-OFDM system, with the training overhead of 4%, the maximum transmission distance is extended to 800 km at Q threshold of 8.7 dB, and RBF-NN outperforms the linear equalizer by 2.8 and 5.6 dB Q-factors after 800 and 600 km transmissions, respectively.

Index Terms: Coherent optical orthogonal frequency division multiplexing (CO-OFDM), nonlinear impairment, neural network (NN).

1. Introduction

To adapt to the ever-increasing demand of transmission capacity and data rate in the fiber-optic community [1], the coherent optical orthogonal frequency division multiplexing (CO-OFDM) has attracted massive interest due to its high spectral efficiency, dispersion resilience, superior receiver sensitivity and convenient digital signal processing (DSP) [2]. However, in high-speed long-haul CO-OFDM system, there exists linear and nonlinear impairments. The linear impairment mainly includes cumulative dispersion and laser phase noise [3]. The nonlinear impairment mainly contains Kerr-induced fiber nonlinearity [4]. The dispersion can be simply compensated by linear equalization. While coherent detection requires the precise tracking of phase and frequency between transmitted signal and the local oscillator (LO) output, which will be disturbed by laser phase noise [6]. In addition, the high peak-to-average power ratio of OFDM signal makes it vulnerable to fiber nonlinearity, especially for coherent system where long transmission distance requires relatively-high fiber launch power [5]. To mitigate the laser phase noise and fiber nonlinearity in
CO-OFDM system, a number of DSP algorithms have been proposed [7]–[9]. Recently, the most attractive innovative approach is neural network (NN) [10], [11]. NN maps low-dimensional data to high-dimensional space so that linearly inseparable problem in low-dimensional space is linearly separable in high-dimensional space, that is, provides nonlinear decision boundary [11] for impairment equalization which can be regarded as classification; while the tradition linear equalization only provides linear decision boundary. Owing to the fact that the channel of CO-OFDM system is quasi-static [12], feedforward artificial neural network (ANN) is the suitable method [13]. There are two typical ANNs, multi-layer perceptron (MLP)-NN and radial basis function (RBF)-NN [14], [15]. Recently, some researches have proposed the use of MLP based ANN for nonlinearity mitigation [16], [17]. However, compared with RBF-NN, MLP-NN has several disadvantages [18]–[20]. First, MLP-NN often uses the sigmoid function as the activation function and has a large input visible field. While in the RBF-NN, when the input value deviates from the center of the basis function, the output gradually decreases and tends to zero quickly, which is more consistent with the feature that the response of neurons is based on the receptive field and converges faster. Second, MLP-NN approaches the nonlinear mapping globally, while RBF-NN uses the locally exponential decaying nonlinear function to locally approach, the required parameters of RBF-NN are less than MLP-NN when achieving the same precision, which massively reduces the complexity. For RBF-NN, to determine the network structure's parameters, the tradition learning algorithms are K-means clustering [20], density peaks clustering and evolutionary computation, which require a large number of matrix inversion operations, resulting in large computation and low computational efficiency.

In this paper, for the first time, we demonstrate a novel low-complexity orthogonal least square (OLS) based RBF-NN equalizer for high-speed CO-OFDM system, its ability to compensate non-linear and linear impairments are evaluated by experiment. The OLS algorithm does not require a large number of sample matrix's inversion calculation, so it has lower computation complexity and is easier to operate than the tradition learning algorithms [21]–[23]. Since it is well known that when the transmission rate is 40 Gb/s or larger, the major nonlinear penalties are from intra-channel nonlinearities rather than inter-channel nonlinearities [23], the experiment of single-channel 40-Gb/s 16-quadrature amplitude modulation (QAM) CO-OFDM system is conducted to evaluate the performance of RBF-NN equalizer under intra-channel nonlinear and linear impairments. Results show that the OLS based RBF-NN equalizer can jointly mitigate linear and nonlinear impairments. Compared to the linear equalizer with the training overhead of 4\%, the maximum transmission distance is extended to 800 km at Q threshold of 8.7 dB, Q-factor improvements of 2.8 dB and 5.6 dB are obtained by RBF-NN for 800 km and 600 km transmissions respectively. Compared to the MLP-NN equalizer, Q-factor improvement of 2.1 dB is obtained by RBF-NN for 600 km transmission. The remainder of this paper is organized as follows. Section 2 describes the OFDM generation/post-processing procedures with detailed principle of OLS based RBF-NN equalizer. Section 3 first presents the experiment setup, then the impacts of training overhead on the performance of linear and RBF-NN equalizers are analyzed. After the training overhead is selected, the ability of RBF-NN equalizer to compensate linear impairment, fiber nonlinearity and laser phase noise are evaluated respectively. Conclusions are given in Section 4.

2. Principle of OLS Based RBF-NN Equalizer in CO-OFDM System

Fig. 1(a) depicts the configuration of the single-channel CO-OFDM system. The procedures of its electrical parts, OFDM generation and post-processing, are described as follows. In the OFDM generation module at the transmitter, the user's binary data is converted from serial to parallel and mapped to the M-QAM constellation every $m$ bits as a group ($m = \log_2 M$). For OFDM modulation, the number of total subcarriers is $N_{sc}$ of which $N_d$ subcarriers carry data and others are mapped with zeros, the number of OFDM symbols is $N_s$. The M-QAM data is symmetric mapping to the data subcarriers. Extra $N_p$ symbols are added in the beginning of the frame as the training sequences for channel equalization. After inverse fast Fourier transform (IFFT), cyclic prefix (CP) insertion and parallel to serial conversion, the pseudo-noise (PN) sequence is added for synchronization at
Fig. 1. (a) Configuration of single-channel CO-OFDM system. DAC: digital to analog convertor, I/Q modulator: in-phase/quadrature modulator, EDFA: erbium doped fiber amplifier, SMF: single mode fiber, LO: local oscillator, BPD: balanced photo diode, ADC: analog to digital convertor. (b) Schematic diagram of OFDM processing module. (c) Structure of the real part of the OLS based RBF-NN equalizer.

receiver, similar to our previous work [25]. Finally, the in-phase (I) and quadrature (Q) components are separately transmitted.

At the receiver, the procedures of the OFDM processing module are shown in Fig. 1(b). The received I and Q signals are first filtered by the low-pass filters (LPF) respectively and form the complex OFDM signal. After the OFDM symbol synchronization [21], serial to parallel conversion, CP removal, training sequences removal, fast Fourier transform (FFT), subcarrier demapping and linear frequency-averaging channel equalization [26], the frequency-domain linear-equalized OFDM signal is obtained, labeled as $d(n)$. The transmitted and received training symbols are labeled as $X_{p,b}$, respectively.

In the subsequent RBF-NN equalizer, there exists two sub-neural networks, real and imaginary sub-neural networks. Note that, the training data in RBF-NN equalizer are the same as those used in the linear equalizer. Since the principles of real and imaginary sub-neural networks are similar, only the principle of the real sub-neural network is given. As show in Fig. 1(c), the RBF-NN comprises three layers: the input layer, the hidden layer and the output layer. The input data and output data have the same dimension. There are two stages, the training stage and testing stage. In the training stage, $X_i$ are input to the input layer. There are four parameters that need to learn and train by OLS algorithm: the number of neurons in hidden layer, the standard deviation of basis function, and the weights between the hidden layer and output layer [14]. The RBF-NN can be considered as a special case of linear regression model [21]–[23],

$$d(n) = \sum_{h=1}^{H} \rho_h(n) w_h(n) + E(n), \quad n = 1, 2, \ldots, N$$

where $N$ is the number of data in the input layer, $H$ is the number of neurons in hidden layer, $w_h(n)$ is the weight from the $h$-th hidden neuron to the output layer, $d(n)$ is the expected output, $E(n)$ is the error between the actual output of RBF-NN and expected output, $\rho_h(n)$ is the regressors which is the response of RBF-NN under certain basis function. For the input $X_i$, the expected output $d(n)$ is $Y_i$, $N$ equals to $N_d \cdot N_p$. The Gaussian function is chosen as the basis function $\phi$, then the output
of h-th hidden neuron is
\[ p_h(n) = \phi(\|X_{in} - \mu_h\|) = \exp\left(\frac{-1}{2\sigma^2}\|X_{in} - \mu_h\|^2\right) \] (2)

where \(\|\cdot\|\) is the Euclidean norm, \(\mu_h\) is the h-th center of basis function, \(\sigma\) is the standard deviation which is set to same for all the neurons and needs selecting based on the channel characteristic. Eq. (1) can be written in the form of matrix,
\[ d = Pw + E \] (3)

where \(d = [d_1, d_2, d_3, \ldots, d_n]^T\), \(w = [w_1, w_2, w_3, \ldots, w_H]^T\), \(P = [p_1, p_2, p_3, \ldots, p_H]^T\), \(p_h = [p_{h1}, p_{h2}, p_{h3}, \ldots, p_{HN}]^T\), \(E = [e_1, e_2, e_3, \ldots, e_N]^T\). The aim of the OLS algorithm is to select the suitable regressor vectors \(p_h\) by learning. Since one regressor \(p_h\) is corresponding to one center \(\mu_h\), the number of the suitable regressor vectors is exactly the number of neurons in hidden layer. The basic idea of OLS is analyzing the contribution of \(p_h\) to reducing the error \(E\) by orthogonalizing \(p_h\), based on certain criterion, the suitable regressor vectors \(p_h\) are chosen. Applying the orthogonal triangular decomposition to the regressor matrix \(P\), we can get
\[ P = RA \] (4)

Here, \(R\) is an \(N\)-by-\(H\) matrix whose columns are orthogonal and satisfies \(R^TR = H\) where \(H\) is a diagonal matrix. \(A\) is an \(H\)-by-\(H\) upper triangular matrix whose elements of the principal diagonal are ones. Due to the fact that the regressors are linearly independent after OLS turns \(P\) orthogonal, their mutual interaction is small, leading to superior performance and lower computation time in function approximation than K-means clustering algorithm [20]. Eq. (3) can be written as
\[ d = RAw + E = RG \] (5)

Left-multiply \(R^T\) in (5),
\[ R^Td = R^TRG = HG \] (6)

we can get
\[ G = H^{-1}R^Td \text{ or } G_h = \frac{R_h^Td}{R_h^TR_h} \quad (1 \leq h \leq H) \] (7)

The contribution of orthogonal vector \(R_h\) to reducing the error \(E\) is defined as
\[ \xi_h = \frac{G_h^2R_h^TR_h}{d^Td} \] (8)

\(h\) is increased from 1 to \(N_h\), and the termination criterion for increasing \(h\) is
\[ 1 - \sum_{h=1}^{H} \xi_h = E_T \] (9)

where \(E_T\) is the error threshold. When the iteration terminates, the \(h\) value at this moment is set to the number of neurons in hidden layer, that is, \(H\) is updated from \(N_h\) to \(h\). After the number of neurons in hidden layer is trained, center \(\mu_h\) can be obtained based on regressor vectors \(p_h\). Meanwhile, orthogonal matrix \(R\) and triangular matrix \(A\) are obtained, then \(w\) can be worked out based on (5). So far, RBF-NN is trained. To sum up, the procedures of training stage are listed as follows. Step 1: Initialize two sub-neural networks for the real data \((\cdot)_p\) and imaginary data \((\cdot)_q\), respectively. Step 2: Present the input vector \(X_p, X_q\), and the expected output vector \(Y_p, Y_q\) for real and imaginary sub-neural networks, respectively. Step 3: Set the maximum number of neurons in hidden layer to \(N_h\) \((N_h < N)\). Initialize the current neurons’ number in hidden layer \(H\) to 1. Increase and update \(H\) from 1 to \(N_h\). Validate the termination criterion after each increment of \(H\) using (8) and (9). When the criterion is satisfied, \(H\) stops increasing. If criterion is not satisfied even when \(H\) is up to \(N_h\), then \(H\) is set to \(N_h\). Here, the number of neurons in hidden layer is trained. Step 4: The centers of basis
function can be obtained from suitable regressor vectors. Step 5: The weights are calculated based on (5). Next, in the testing stage, $x_I$ and $x_Q$ are input into the input layers of the trained real and imaginary sub-neural networks respectively, $x_{eI}$ and $x_{eQ}$ are output respectively. The complex signal $x_e = x_{eI} + i \cdot x_{eQ}$ then performs M-QAM hard-decision demodulation to get received binary data.

3. Experimental Setup and Results Analysis

Following the system configuration in Fig. 1(a), the experiment is conducted. At the transmitter, the output of a continuous-wave laser (Koheras AdjustiK-E15) at wavelength of 1550 nm passes through a polarization controller (PC) which is precisely adjusted before injecting into the I/Q modulator (FUJITSU FTM7960EX) to get the best modulation performance. The I and Q components of the 16-QAM OFDM signal, which are generated offline in MATLAB, are loaded into the arbitrary wave generator (AWG, Tektronix AWG7122C) with 20-GS/s sampling rate. Two output channels of AWG drive the I/Q modulator biased at the null point to achieve the linear radio frequency (RF)-to-optical conversion and eliminate the modulator nonlinearity. The output optical CO-OFDM signal from the I/Q modulator is first amplified by an erbium doped fiber amplifier (EDFA, WXZTE-WZEDFA) and then transmits through a loop link whose span consists of the 100-km standard single mode fiber (SMF) and an EDFA. At the receiver, the received signal first passes through a PC and then injects into a 90° optical hybrid (Kylia COH28-X), together with the output of the LO (Koheras AdjustiK-E15). Note that, laser at the transmitter and LO at the receiver both have the same claimed linewidth below 10 kHz, which is a factor to the success of coherent detection. Four outputs of the hybrid are detected by a pair of balanced photodetectors (BPDs, Discovery DSC-R412) to implement linear optical-to-RF conversion. Two outputs of BPDs are sampled by a real-time oscilloscope (LeCroy SDA830Zi-A) with 20-GS/s sampling rate and then offline processed in MATLAB. The parameter settings of OFDM are as follows. The number of total subcarriers $N_{sc}$ and data subcarriers $N_d$ are 128 and 64 respectively. Two subcarriers from the zero frequency are mapped with zeros to spectrally separate the baseband from the ground noise and high frequency aliasing products [4]. The CP length is 16. The total number of measured OFDM symbol $N_s$ is 100000. Thus the data rate is 40 Gb/s. The parameter settings of RBF-NN are: the error threshold $E_T$ is set to 0.000001, the optimal standard deviation $\sigma$ is 0.11.

There is no doubt that larger training data can bring more accurate estimation for channel model, while it will introduce larger computational complexity which should be reduced in practical use. Thus, firstly, the impact of training overhead on system performance is investigated in the linear domain of the CO-OFDM system. Taking 40-Gb/s 600-km CO-OFDM system with the optimum fiber launch power applied as an example, Fig. 2 demonstrates the Q-factor versus training overhead when only linear equalization (denoted as w/o NN) is applied and both RBF-NN and linear
Fig. 3. Experiment results of (a) Q-factor and (b) Q-factor improvement by RBF-NN and MLP-NN versus fiber launch power in 40-Gb/s CO-OFDM system over different transmission distances. (c) and (d) are received 16-QAM constellation diagrams at 300 km and 600 km of transmissions when their optimum fiber launch powers are applied respectively (Blue are before the RBF-NN equalizer, green are after the RBF-NN equalizer).

equalizations are applied (denoted as w/ RBF-NN), respectively. The system performance is evaluated by Q-factor in dB, measured from the bit error rate (BER) using $Q = 20 \log_{10}( \frac{1}{2} \text{erfc}^{-1}(2BER))$. And we define the Q threshold as the corresponding Q-factor to the 7% forward error correction (FEC) threshold of $1 \times 10^{-3}$, which is 8.7 dB [27]. It is seen that Q-factor increases with the training overhead, the phenomenon is more obvious when both RBF-NN and linear equalizations (w/ RBF-NN) are used in comparison with the case where only linear equalization (w/o NN) is used. When only linear equalizer is applied, the Q-factor stays almost the same when the training overhead is larger than 4% and is smaller than the Q threshold, showing the limitation of linear equalization even when only linear impairment exists. While with the help of NN equalizer, the Q-factor significantly increases and is larger than the Q threshold, revealing the excellent resistance of NN equalizer to linear impairment. Specifically, the Q-factors are enhanced by 3.8 dB, 5.3 dB, 5.6 dB, 6 dB, 6.1 dB and 6.2 dB when the training overheads are 2%, 3%, 4%, 10%, 20% and 30%, respectively. It can be observed that the Q-factor increment is less sharp when the training overhead is increased to 4% or larger. Thus, trading off between the computational complexity and system performance, the training overhead is set to 4% in the following content.

Then, the ability of RBF-NN equalizer to combat fiber nonlinearity and laser phase noise are separately evaluated. Since the laser linewidths are directly proportional to laser phase noises [4], and the linewidths of continuous-wave laser and the LO used in the experiment are quite small, the laser phase noise-induced nonlinear impairment can be ignored in our experiment. Thus, firstly, the capability of NN equalizer to compensate fiber nonlinearity is discussed. Fig. 3(a) presents the experimental results of Q-factor versus fiber launch power in 40-Gb/s CO-OFDM system over different transmission distances respectively. The maximum transmission distance is extended to 800 km by RBF-NN since its Q-factor just reaches the Q threshold at the optimum fiber launch power of $-7.8$ dBm. The performances of 800 km and 600 km transmissions are demonstrated as representatives. Fig. 3(b) shows the corresponding Q-factor improvement by NN as a function of fiber launch power. The results of the case when both MLP-NN and linear equalizations are applied (denoted as w/ MLP-NN) are given as well. For a fair comparison, MLP-NN has the same training overhead as RBF-NN. At the launch power of $-7.8$ dBm, RBF-NN outperforms the linear equalizer by 2.8 dB and 5.6 dB for 800 km and 600 km transmissions respectively, RBF-NN outperforms MLP-NN by 2.1 dB for 600 km transmission case. The system undergoes nonlinear domain when the fiber launch powers are higher than $-7.8$ dBm. In the nonlinear domain, for instance, the Q-factors are enhanced by 1.96 dB and 4.6 dB at the launch power of $-3.8$ dBm for 800 km and 600 km transmissions, respectively. To visually demonstrate the superiority of the RBF-NN equalizer, the received 16-QAM constellation diagrams after 300 km and 600 km transmissions with their optimum fiber launch powers applied are given in Fig. 3(c) and (d) respectively. Blue dots
represent the signal with only linear equalization, green dots represent the signal with both linear and RBF-NN equalizations. Seen from the constellation points with reduced error vector magnitude, it is obvious that RBF-NN equalizer yields significant performance improvement in both nonlinear and linear domains thanks to its nonlinear decision boundary.

Secondly, the robustness of RBF-NN equalizer to laser phase noise is discussed by simulation. The Monte-Carlo simulation is conducted in commercial software OptiSystem 13.0 combined with MATLAB, following the same scheme and parameters as the experiment. The impact of fiber nonlinearity is minimized by adopting the optimum fiber launch power. The laser linewidth is varied from 0 MHz to 3 MHz to introduce various amounts of laser phase noise. Note that, the LO linewidth at receiver stays consistent with the laser linewidth at transmitter in each case. Taking 40-Gb/s 600-km CO-OFDM system as an example, Fig. 4 shows the simulation results of Q-factor as a function of laser linewidth. With the help of NN equalizer, the Q-factors are enhanced by 5.3 dB, 5.17 dB, 4.9 dB, 4.48 dB and 2.46 dB when the laser linewidths are 0 MHz, 0.75 MHz, 1.5 MHz, 2.25 MHz and 3 MHz respectively, proving that the RBF-NN equalizer effectively mitigates the laser phase noise-induced nonlinear impairment. Thus, it can be concluded that nonlinear and linear impairments can be both compensated by RBF-NN equalizer.

4. Conclusion

In this paper, a novel low-complexity OLS based RBF-NN nonlinear equalizer is proposed and experimentally verified for 40-Gb/s 16-QAM CO-OFDM system. Results show that OLS based RBF-NN can simultaneously mitigate linear and nonlinear impairments. The training overhead is set to 4% after balancing the complexity and performance. When comparing RBF-NN to linear equalizer, the maximum transmission distance is extended to 800 km at Q threshold of 8.7 dB, the maximum Q-factor improvements of 2.8 dB and 5.6 dB are obtained under 0.01-MHz laser linewidth for 800 km and 600 km transmissions respectively, and the Q-factor is enhanced by 2.46 dB even when the laser linewidths are both up to 3 MHz in 600 km transmission aided by RBF-NN. Also, RBF-NN outperforms MLP-NN by 2.1 dB for 600 km transmission case. Thus, OLS based RBF-NN is an applicable option for compensating linear and nonlinear impairments in high-speed long-haul CO-OFDM.

References


