



112-Gb/s SSB 16-QAM signal transmission over 120-km SMF with direct detection using a MIMO-ANN nonlinear equalizer

SHAOHUA AN, QINGMING ZHU, JINGCHI LI, YUYE LING, AND YIKAI SU*

State Key Lab of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

*yikaisu@sjtu.edu.cn

Abstract: We propose and experimentally demonstrate a multiple input multiple output - artificial neural network (MIMO-ANN) nonlinear equalizer (NLE) to process the complex quadrature amplitude modulation (QAM) signal in a single-sideband (SSB) self-coherent detection (SCD) system. In the proposed scheme, a 2-by-2 MIMO structure with two ANNs is employed to effectively mitigate the signal distortions induced by in-phase and quadrature (IQ) imbalance and fiber nonlinear effects. By using the proposed MIMO-ANN NLE, we successfully transmit a 112-Gb/s SSB 16-QAM signal over a single-span 120-km single mode fiber (SMF) in a direct detection (DD) system with a bit error rate (BER) lower than 3.8×10^{-3} . We also conduct a comparative study between the proposed MIMO-ANN NLE, a feedforward equalizer (FFE), a NLE consisting of two independent real-valued Volterra filters, and a MIMO-Volterra filter. The proposed MIMO-ANN NLE outperforms other equalizers with the longer fiber length and thus stronger nonlinearities, since it can easily approximate a complicated nonlinear function. To the best of our knowledge, this is the first experimental demonstration of an ANN-based equalizer in an SSB SCD system.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

With the tremendous increase of broadband applications such as cloud computing and high-definition video, capacity demands in metro and data center interconnections are growing rapidly [1]. In the scenarios of ~100-km reach, cost-effective transceivers with data rates beyond 100 Gb/s are highly desired. Coherent transceivers exhibit high spectral efficiencies, but they face challenges of high costs and power consumptions. Self-coherent optical orthogonal frequency-division multiplexing (SCO-OFDM) technology with a chip-based stimulated Brillouin scattering (SBS) filter provides an interesting solution [2]. Alternatively, direct detection (DD) systems have received considerable attention due to their lower costs, higher power efficiencies, and easier implementations when compared with coherent systems. Recently, various DD systems with advanced digital signal processing (DSP) technologies have been proposed and experimentally demonstrated [3–5]. Among these DD systems, single-sideband (SSB) self-coherent detection (SCD) system based on Kramers-Kronig (KK) algorithm is a promising solution, due to its high spectral efficiency and excellent optical field reconstruction capability [6]. In an SSB SCD system, high-order quadrature amplitude modulation (QAM) can achieve a high spectral efficiency, and KK algorithm may allow an accurate optical field reconstruction [7]. Thus, linear distortions including chromatic dispersion (CD) can be compensated at the receiver [8]. With the increased fiber length, a higher launch power is required to improve the optical signal-to-noise ratio (OSNR) at the receiver. Stronger fiber nonlinear effects will therefore be introduced. Different from a very short-reach link, both the nonlinear impairments and the degraded OSNR may limit the achievable performance in an optimized system with a reach above tens of kilometers [9]. Moreover, in-phase (I) and quadrature (Q) imbalance, crosstalk between the I and Q

components and other nonlinearities from components further limit the system performance [10]. In order to mitigate these distortions, nonlinear equalizers (NLEs) can be used in the receiver DSP. Conventional NLEs, such as a real-valued Volterra filter (VF), have been employed in intensity modulation-direct detection (IM-DD) systems. But they cannot be directly implemented in the SSB SCD system, since a complex-valued QAM signal is the input instead of a real-valued signal. To address this issue, a sparse IQ VF was proposed to alleviate nonlinear distortions in an SSB SCD system [10], enabling a better performance than a NLE with two independent real-valued VFs. However, a VF-based NLE shows a limited performance in the scenario with strong nonlinearities [11].

Recently, machine learning algorithms have shown remarkable performances when applied in optical communication systems [12,13]. Among these algorithms, artificial neural network (ANN) is considered as a promising equalization tool. Various ANN-based NLEs have been proposed and experimentally demonstrated both in long-haul coherent systems and short-reach IM-DD systems [11,14,15]. Compared to VF-based NLEs, ANN-based NLEs could bring more effective improvements, since it can easily approximate a complicated nonlinear function [16]. Nonetheless, there have been no efforts to mitigate nonlinear distortions using machine learning algorithms in an SSB SCD system, which are more serious due to the complex-valued QAM signal compared to an IM-DD system.

In this paper, we propose a multiple input multiple output (MIMO) - ANN NLE to process the complex QAM signal in an SSB SCD system. In our proposed scheme, two ANNs in a 2-by-2 MIMO structure constitute the real and the imaginary parts of a desired QAM signal, respectively. For each ANN, both the I and Q components of a received QAM signal are sent into the input layer, to mitigate the IQ imbalance and crosstalk. The training processes of the two ANNs are independent, due to the independence between the real and the imaginary parts of a QAM signal. Based on the proposed MIMO-ANN NLE, we experimentally demonstrate an SSB-SCD 112-Gb/s 16-QAM signal transmission over a single-span 120-km single mode fiber (SMF) with a bit error rate (BER) below 3.8×10^{-3} . Moreover, we compare the performances of the proposed MIMO-ANN NLE, a feedforward equalizer (FFE), a NLE with two independent real-valued VFs (2 real VFs), and a MIMO-VF. The results show that the proposed MIMO-ANN NLE outperforms other equalizers in the scenarios where the nonlinearities are stronger with the increases of the fiber length and the signal launch power. If the single-span length is extended from 120 km to 145 km, only the proposed MIMO-ANN NLE can reach the BER threshold of 2.4×10^{-2} in the SSB SCD system. To the best of our knowledge, this is the first experimental demonstration of an ANN-based equalizer in an SSB SCD system.

2. Operation principle

Figure 1(a) depicts the block diagram of the proposed MIMO-ANN NLE. The received complex-valued QAM signal, which can be divided into an I component x_i and a Q component x_q , is simultaneously fed into two ANNs. In each ANN, the I and Q components are linked with each other and processed by nonlinear operations to effectively suppress the nonlinearities. Two independent training processes on the ANNs are performed to obtain two output signals y_i and y_q , which are combined to form the equalized QAM signal.

Here we use a 2-by-2 MIMO structure to process the complex-valued QAM signal, since it can effectively mitigate the interference between two orthogonal signals. Previous applications of the 2-by-2 MIMO construction included polarization demultiplexing in the coherent system, and IQ interference cancellation of a carrier-less amplitude and phase modulation (CAP) signal, with excellent performances [17,18]. It is worth noting that the proposed scheme is different from the previous scheme in [13], which utilizes one ANN to form both the real and the imaginary parts of a desired QAM signal in a coherent system. In the proposed NLE, we use the 2-by-2 MIMO structure for the QAM signal to recover the two quadrature components.

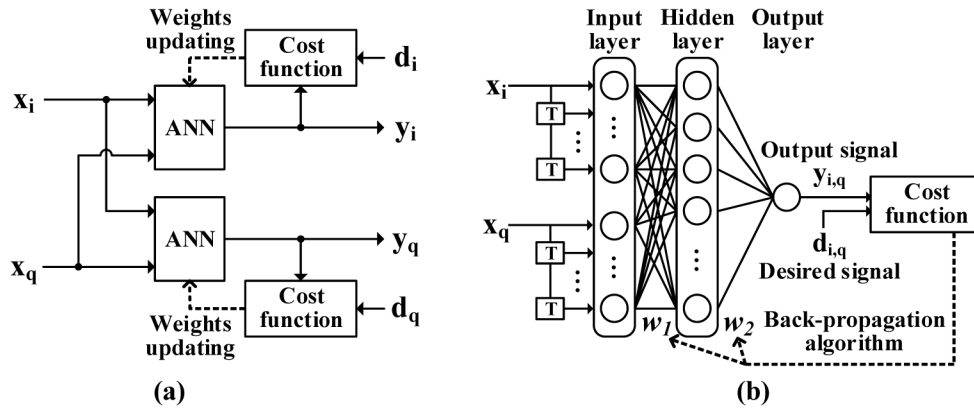


Fig. 1. (a) Block diagram of the proposed MIMO-ANN NLE. (b) Block diagram of each ANN in the proposed MIMO-ANN NLE.

Figure 1(b) shows the block diagram of each ANN in the proposed MIMO-ANN NLE. For a proof-of-concept demonstration, we design the equalizer with a fully connected two-layer neural network with one hidden layer. A better performance can be expected by using a deep neural network (DNN)-based nonlinear equalizer, at the cost of an increased complexity. Each circle in Fig. 1(b) denotes a neuron, which computes the output using an activation function [19]:

$$y_m = f\left(\sum_k w_{mk} x_k\right), \quad (1)$$

where y_m is the output of the m -th neuron in the current layer, x_k is the output of the k -th neuron in the previous layer and also the input of the m -th neuron in the current layer, w_{mk} is the corresponding weight and $f(\cdot)$ is an activation function. A nonlinear activation function is usually employed in the hidden layer, which enhances the expression ability of the neural network. For simplicity, we have omitted bias terms.

As shown in Fig. 1(b), two samples of the I and Q components with their delayed copies are sent into the input layer. After computed by a linear activation function, the outputs of the input layer are delivered to the hidden layer, where the rectified linear units (ReLU) function is applied as the activation function [19]:

$$\text{ReLU}(x) = \max(0, x). \quad (2)$$

Here the ReLU function is used instead of a sigmoid function, due to its faster training process [20]. Subsequently, the outputs of the hidden layer are delivered to the output layer, which consists of one neuron with a linear activation function. In the training process, the mean square error (MSE) function is employed as the cost function [15]:

$$J(\mathbf{w}) = |d(n) - y(n)|^2, \quad (3)$$

where \mathbf{w} is the weight matrix, $d(n)$ is the n -th sample of the desired signal, $y(n)$ is the output of the ANN at the n -th training iteration. In order to minimize the cost function, we employ the back-propagation (BP) algorithm [21] where the weights-updating process between the hidden and the output layers can be expressed as follows:

$$\Delta \mathbf{w}_2(n) = -\mu \frac{\partial J(\mathbf{w}_2)}{\partial \mathbf{w}_2(n)} = 2\mu (d(n) - y(n)) \mathbf{y}_1^T(n), \quad (4)$$

$$\mathbf{w}_2(n+1) = \mathbf{w}_2(n) + \Delta \mathbf{w}_2(n), \quad (5)$$

where $\mathbf{w}_2(n)$ is the weight matrix linking the hidden and the output layers at the n -th training iteration, $\mathbf{y}_1(n)$ is the output vector of the hidden layer at the n -th training iteration, μ is the step size, and $(\cdot)^T$ denotes the transposing operation. After updated, $\mathbf{w}_2(n)$ back-propagates to update $\mathbf{w}_1(n)$, which is the weight matrix linking the input and the hidden layers at the n -th training iteration. The weights-updating process can be described as follows [21]:

$$\Delta \mathbf{w}_1(n) = [(\mathbf{v}_1(n) > 0) \mathbf{w}_2^T(n)(d(n) - y(n))] * \mathbf{x}^T(n), \quad (6)$$

$$\mathbf{w}_1(n+1) = \mathbf{w}_1(n) + \Delta \mathbf{w}_1(n), \quad (7)$$

where $\mathbf{v}_1(n)$ is the vector of the weighted sums in the hidden layer at the n -th training iteration, $\mathbf{x}(n)$ is the vector of the input samples of the input layer at the n -th training iteration, and $*$ denotes the multiplication between matrices. After the training processes, the optimized ANNs are used to equalize the received data. In the experiment, a data set consisting of 204,800 QAM symbols is employed, which is the largest data set we can generate. Since an ANN training process takes a large amount of data, we use 80% symbols for the training and the remaining 20% symbols for the testing, respectively. In practice, the training symbols are only used in the beginning to obtain the optimum weights of the neural network. After the training process, the neural network can be used to equalize the received data and no more training symbols are needed. Thus, the training symbols are not an overhead for a transmitted data, and they do not reduce the system capacity. No validation set is used in the experiment, as the data set used in the experiment is large enough. With a large data set, a neural network tested on one unknown data set can already represent a good generalization ability. This has been supported by previous experiments using only training/testing split process [15,22]. Moreover, validation schemes such as a multi-fold cross validation would be hard to apply in practice, due to the limited computational resource.

3. Experimental setup

We perform an experiment to verify the feasibility of the proposed MIMO-ANN NLE in an SSB SCD system. The experimental setup is illustrated in Fig. 2. At the transmitter, an arbitrary waveform generator (AWG) (Keysight M8195A) is used to generate a 28-GBaud Nyquist 16-QAM signal, with a sampling rate of 65 GSa/s. After amplified by two electrical amplifiers (EAs), the I and Q components of the electrical 16-QAM signal drive a 22-GHz IQ modulator (IQM), which is biased at its transmission null with a half-wave voltage of ~ 8.4 V. Two external cavity lasers (ECLs) with ~ 15 -KHz linewidths are employed as the light sources. Here, we use a second laser source instead of a virtual carrier to generate an SSB signal, in order to achieve a better performance without sacrificing the dynamic range of the digital-to-analog converter (DAC) [23]. A continuous wave (CW) light from the ECL1 at 1550.298 nm is fed into the IQM. After electrical-to-optical (E/O) conversion, the optical 16-QAM signal is combined with another optical CW tone located at the edge of the signal spectrum, producing an optical carrier-assisted SSB signal. The optical CW tone is emitted from the ECL2 at 1550.412 nm, followed by a polarization controller (PC) to align the polarization states between the optical 16-QAM signal and the optical CW tone. After the polarization alignment, a variable optical attenuator (VOA) is inserted to vary the carrier-to-signal power ratio (CSPR) of the generated optical SSB signal, which is then boosted by an erbium-doped fiber amplifier (EDFA) before launched into a single-span SMF. At the receiver, the received optical SSB signal is firstly amplified by a second EDFA and then filtered by a 1-nm optical bandpass filter (OBPF) to suppress the amplified spontaneous emission (ASE) noise. A following 40-GHz photodetector (PD) is used to detect the optical signal. After optical-to-electrical (O/E) conversion, the electrical signal is captured by an 80-GSa/s digital storage oscilloscope (DSO) (LeCroy 36Zi-A), followed by the offline DSP.

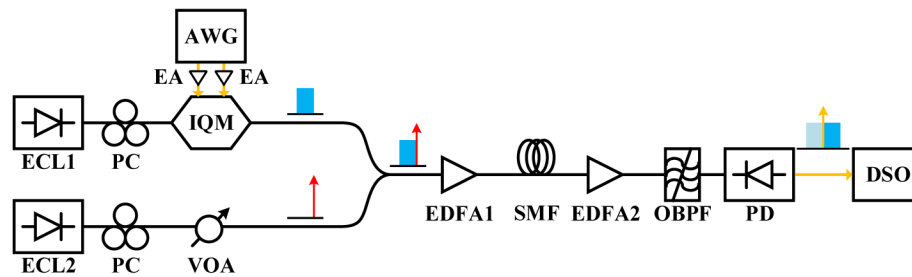


Fig. 2. Experimental setup.

The DSP flow charts are shown in Fig. 3(a). At the transmitter, a random binary string is firstly mapped to a 16-QAM symbol string. Here we use the random data instead of a pseudo-random bit sequence (PRBS) to avoid the overestimation effect of the ANN. The random binary string of 819,200 bits is generated by using the Mersenne Twister algorithm in Matlab. Recent studies have shown a risk of overestimating the performance gain when applying the ANN in a system with a PRBS [11,16]. By learning the generation method of the PRBS rather than the channel characteristics, the ANN shows a superior performance but has a severe overestimation effect. However, using a random binary string as the training data can avoid the risk of overestimation, thus evaluate the performance of the ANN more accurately [11]. After the 16-QAM mapping, a 512-point Zadoff-Chu sequence is added for synchronization due to its excellent correlation property. Both the synchronization sequence and the 16-QAM symbol string are then up-sampled by 65 times, followed by a root raise cosine (RRC) filter with a roll-off factor of 0.01. Previous studies have shown a significant BER reduction with the increase of the roll-off factor, but at a cost of a lower spectral efficiency [24]. Here we use a small roll-off factor to achieve a high spectral efficiency, considering the practical wavelength division multiplexing (WDM) applications [4]. After the RRC filtering, the signal is down-sampled by 28 times to align with the sampling rate of the AWG and then sent into the AWG. In practice, an up-sampling rate of 2 without down-sampling is enough with the very small roll-off factor of 0.01 we employed. At the receiver, the received signal is firstly resampled to a sampling rate of 112 GSa/s, corresponding to 4 samples per symbol (SPS). After adding a direct current (DC) item, the KK algorithm is used to reconstruct the optical field, followed by the CD compensation. Due to the frequency offset between the two ECLs, the signal is down-converted by 14.28 GHz (0.51×28 GBaud). The residual frequency offset is removed by the carrier frequency recovery algorithm [25]. After the matched filtering and synchronization, different linear and nonlinear equalizers are implemented. In the FFE with T_s as the symbol duration, a $T_s/4$ -spaced finite impulse response (FIR) filter with the decision-directed least mean square (LMS) algorithm is used to mitigate linear distortions including the residual CD. The output sampling rate of the FFE is 1 SPS. As for the nonlinear equalization, we investigate three symbol-spaced NLEs including a NLE with 2 real VFs, a MIMO-VF, and the proposed MIMO-ANN NLE. The block diagrams of the NLE with 2 real VFs and the MIMO-VF are depicted in Fig. 3(b) and 3(c), respectively [10,18]. LMS algorithms based on the training sequences are employed to update the weights in the two VF-based NLEs. All the three investigated NLEs are cascaded with a $T_s/4$ -spaced FFE, for fast convergences and fair comparisons. In the experiment, no phase recovery algorithm was employed to compensate the phase noise, thanks to the narrow linewidths of the ECLs. After the equalization, QAM de-mapping and BER calculation are performed to evaluate the system performance.

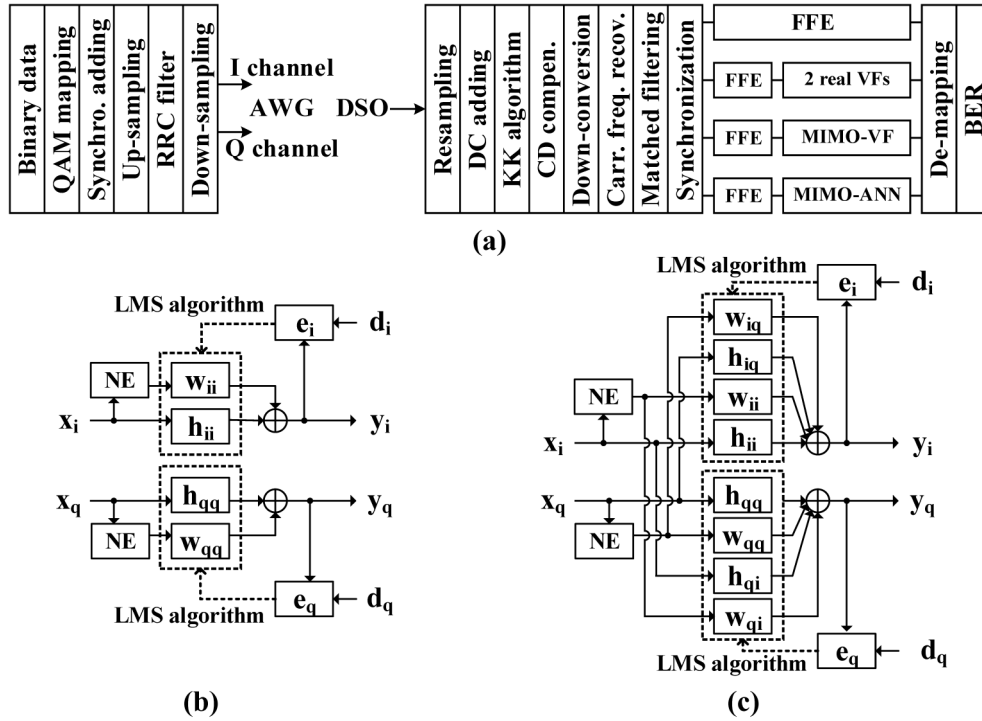


Fig. 3. (a) DSP flow charts. (b) Block diagram of the NLE with 2 real VFs. (c) Block diagram of the MIMO-VF. NE: construction of nonlinear items.

4. Experimental results and discussion

4.1 Optical back-to-back transmission

Figure 4 shows the optical spectra measured after different transmission lengths measured by an optical spectrum analyzer (OSA) (APEX AP2040C) with a 1.12-pm resolution. It can be seen that the optical carrier is located at the edge of the optical QAM signal, generating an optical SSB signal. In our system, no guard band is used to avoid sacrificing the spectral efficiency. However, the signal-to-signal beating interference (SSBI) induced by DD falls into the signal spectrum and degrades the performance after the O/E conversion [23]. In order to remove the SSBI, the KK algorithm is used in the receiver DSP to reconstruct the optical field. After the CD compensation, the linear equalizer followed by the proposed MIMO-ANN NLE is employed to improve the system performance.

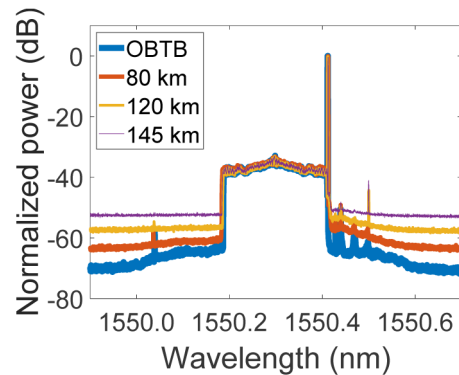


Fig. 4. Optical spectra measured after different transmission distances.

In the optical back-to-back (OBTB) case, we firstly optimize the delay length in the input layer and the neuron number in the hidden layer for each ANN. The same parameters are used in both the two ANNs. As shown in Fig. 5(a), BER converges with the increase of the delay length in the input layer, when the neuron number in the hidden layer is set to 14. If the delay length is larger than 51, no further BER reduction is observed. Thus, the delay length of 51 is adopted in the following optimization, corresponding to 102 neurons in the input layer. Figure 5(b) shows the measured BER versus the neuron number in the hidden layer, while the delay length in the input layer is set to 51. The BER performance exhibits an independence on the neuron number in the hidden layer, since the nonlinearities in the OBTB case are not strong. The slight BER fluctuation is due to the change of the network structure, which obtains different optimized weights. The neuron number of 14 in the hidden layer is employed in the following experiment.

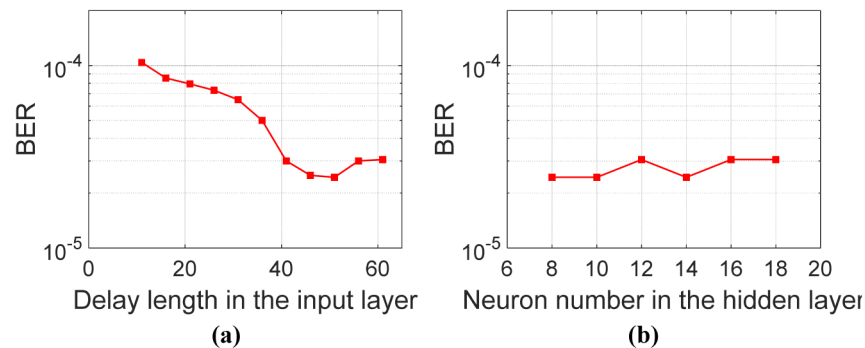


Fig. 5. (a) BER versus delay length in the input layer. (b) BER versus neuron number in the hidden layer.

We then study the tolerance performance of the MIMO-ANN NLE to the IQ imbalance. By adjusting the bias voltage of the phase shifter in the IQM, the IQ imbalance can be deliberately introduced, thus distorting the constellations. In the OBTB case, the $T_s/4$ -spaced FFE tap number is set to 201. The 1st, 2nd, and 3rd memory lengths of the two VF-based NLEs are both optimized to be 51, 9, and 1, respectively. As seen in Fig. 6, the FFE and the NLE with 2 real VFs both show increased BERs when the bias voltage is far from the optimum point, exhibiting low tolerances to the IQ imbalance. However, the MIMO-VF and the proposed MIMO-ANN NLE both show high tolerances to the IQ imbalance, due to the employed 2-by-2 MIMO structures. The results verify the excellent performance of the MIMO scheme to mitigate the interference between two orthogonal signals. The recovered

constellations at the 6.9-V bias of the phase shifter in the IQM for the NLE with 2 real VFs and the proposed MIMO-ANN NLE are shown in the insets (i) and (ii), respectively.

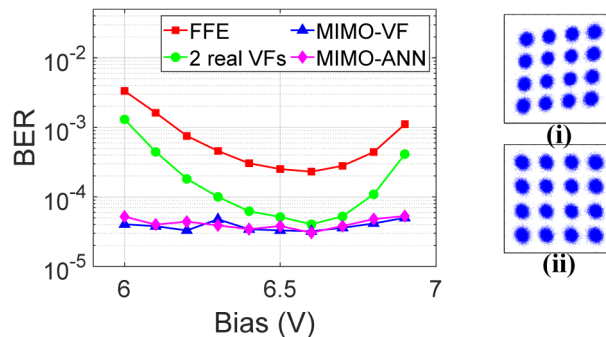


Fig. 6. BER versus bias voltage of the phase shifter in the IQM. Insets (i-ii): constellations at the 6.9-V bias of the phase shifter in the IQM for the NLE with 2 real VFs, and the proposed MIMO-ANN NLE, respectively.

4.2 Single-span fiber transmission

In an SSB SCD system, CSPR is a critical parameter to be optimized. Decreasing the CSPR can improve the system sensitivity, but degrade the performance of the KK algorithm due to the violation of the minimum phase condition (MPC) [23]. Thus, an optimum CSPR value needs to be estimated to achieve the best transmission performance [26]. The optimization results of the CSPR and the launch power after single-span 120-km and 145-km transmissions are shown in Fig. 7(a) and 7(b), respectively. Here we use a relatively high launch power to find the longest single-span length. In the proposed MIMO-ANN NLE, the neuron numbers in the input and the hidden layers for each ANN are set to 102 and 14 respectively, since no BER reduction is observed when further increasing the neuron numbers. After the 120-km fiber transmission, the optimal launch powers for the signals with 12-, 13-, and 14-dB CSPRs are 7, 8, and 9 dBm, respectively. It can be observed that a larger CSPR requires a higher optimum launch power, since a signal with a larger CSPR means less signal component ratio in the total optical power, resulting in a lower receiver sensitivity [26]. From Fig. 7(a), the CSPR of 13 dB and the launch power of 8 dBm can achieve the minimum BER, thus they are adopted at the 120-km distance. Similarly, the optimum 14-dB CSPR and 11-dBm launch power are observed after the 145-km transmission, as shown in Fig. 7(b). The increased optimum launch power in the 145-km case is needed to ensure the same OSNR with the longer fiber length and thus the higher loss. The higher optimum CSPR is used to avoid the violation of MPC caused by stronger nonlinear effects and the increased peak-to-average power ratio (PAPR) through the 145-km transmission.

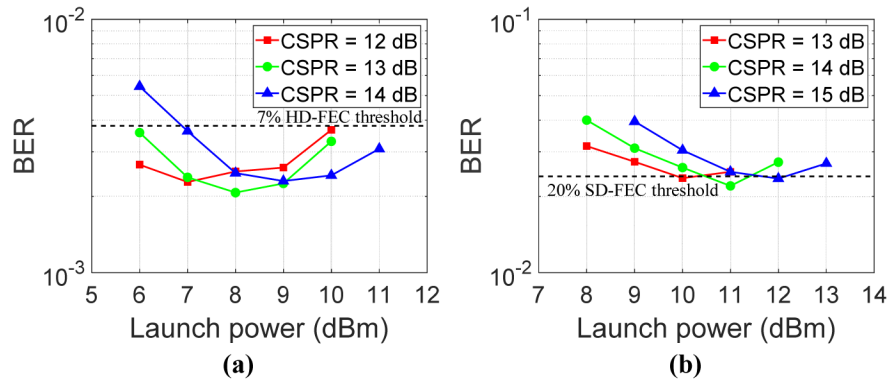


Fig. 7. BERs versus launch powers at different CSPRs after the single-span 120-km (a), and 145-km (b) transmissions, respectively.

With the optimum CSPRs, the performances of the four equalizers under study are investigated at two distances and shown in Fig. 8(a) and 8(b) respectively. In the experiment, we carefully adjusted the bias voltage of the phase shifter in the IQM for each launch power value, thus minimizing the IQ imbalance. After optimization, the $T_s/4$ -spaced FFE tap number is set to 201. The 1st, 2nd, and 3rd memory lengths of the two VF-based NLEs are both optimized to 51, 9, and 3, respectively. Therefore, 51, 45, and 10 kernels are employed for the in-phase or quadrature equalization in the NLE with 2 real VFs, leading to 212 kernels in total. In the MIMO-VF, 102, 90, and 20 kernels are employed for the in-phase or quadrature equalization, resulting in 424 kernels in total. Compared to the OBTB case, only the 3rd memory length is increased, since the CD and the SSBI are effectively compensated and cancelled in the receiver DSP, respectively. In Fig. 8(a), the conventional FFE cannot reach the BER of the 7% hard-decision forward error correction (HD-FEC) threshold after the 120-km transmission, since only the linear distortions are suppressed. In order to mitigate the nonlinear distortions, the NLE with 2 real VFs, the MIMO-VF, and the proposed MIMO-ANN NLE can be employed, exhibiting significant BER reductions. The three NLEs show similar performances at the 120-km distance. After the 145-km transmission, stronger nonlinearities are induced by the increased CSPR and higher launch power. The proposed MIMO-ANN NLE shows a better BER performance than other equalizers, since it can easily approximate a complicated nonlinear function. As shown in Fig. 8(b), only the proposed MIMO-ANN NLE can reach the BER of the 20% soft-decision forward error correction (SD-FEC) threshold at the 145-km distance.

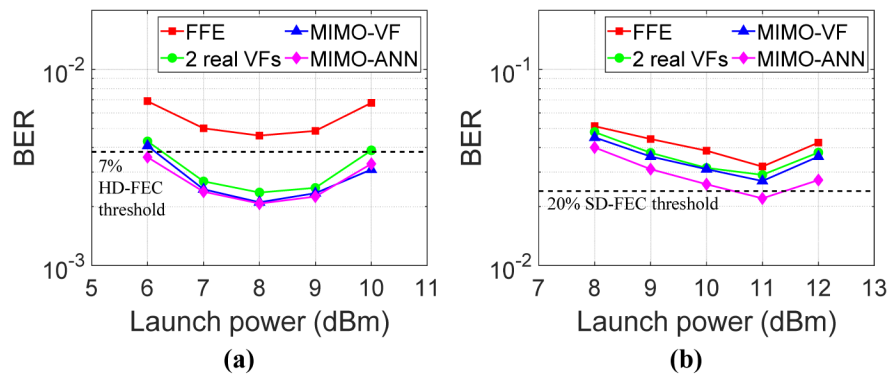


Fig. 8. BERs versus launch powers with different equalizers after the single-span 120-km (a), and 145-km (b) transmissions, respectively.

In addition, we provide the measured BER performance versus the frame length after the 120-km transmission, as shown in Fig. 9. When implemented in image classification, an ANN usually needs a large data set with up to one million images to train the network [20]. Therefore, it is interesting to investigate how many data symbols are needed for the equalization applications in the optical communication systems. Limited by the memory of the AWG, we employ 204,800 16-QAM symbols as the data set, in which 80% symbols are used for the training and the remaining 20% symbols are used for the testing, respectively. In Fig. 9, we decrease the frame length step by step and measure the BER accordingly, while the length ratio between the training data and the testing data remains the same. A slight BER fluctuation is observed if the frame length is longer than 50,000, which verifies that the frame length used in the experiment is sufficient for the proposed MIMO-ANN NLE to learn the channel characteristics.

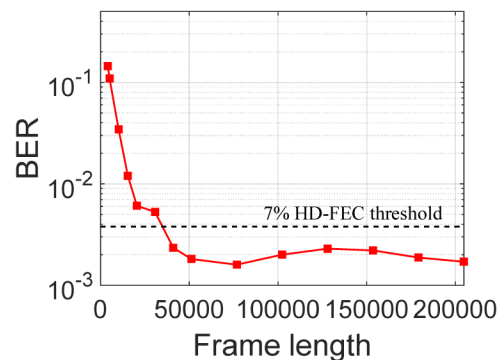


Fig. 9. BER versus frame length.

We also compare the computational complexities of the four investigated equalizers, as shown in Table 1. The number of real-valued multiplications required by an equalizer to output a symbol is used to characterize the complexity. In the $T_s/4$ -spaced FFE with 201 taps, each output symbol needs 201 complex-valued multiplications in the weighting process. Since one complex-valued multiplication can be simply realized by using four real-valued multiplications, 804 real-valued multiplications are needed in the FFE in total. As for a real-valued VF, the 1st, 2nd, and 3rd memory lengths of 51, 9, and 3 can generate the corresponding kernels of 51, 45, and 10, respectively. Both the kernel generating process and the weighting process require multiplications in a VF-based NLE. In the NLE with 2 VFs, 130 ($2 \cdot (45 + 2 \cdot 10)$) and 212 ($2 \cdot (51 + 45 + 10)$) real-valued multiplications are performed during the kernel generating process and the weighting process, respectively. Thus, the number of the real-valued multiplications is 342 in the NLE with 2 VFs, leading to 1146 real-valued multiplications in the scheme of FFE + 2 real VFs. Similarly, the MIMO-VF requires 130 ($2 \cdot (45 + 2 \cdot 10)$) and 424 ($4 \cdot (51 + 45 + 10)$) real-valued multiplications in the kernel generating process and the weighting process, respectively. Therefore, 1358 real-valued multiplications are performed in the scheme of FFE + MIMO-VF. In the proposed MIMO-ANN NLE, the number of real-valued multiplications is 2884 ($2 \cdot (102 \cdot 14 + 14 \cdot 1)$), considering the 102, 14 and 1 neurons in the three layers respectively. As a result, the total number of real-valued multiplications is 3688 in the scheme of FFE + MIMO-ANN. As shown in Table 1, the ANN-based NLEs need more multiplications than the conventional FFE and VF-based NLEs, since an ANN has a more complicated structure than a transversal filter. Compared with the scheme in [15] as a benchmark, the proposed scheme has a slightly higher complexity. This is mainly attributed to the 2-by-2 MIMO structure we employed to process the complex-valued QAM signal, instead of the real-valued signal in [15]. In order to reduce the complexity of the proposed MIMO-ANN NLE, some pruning methods and simplified network structures can be exploited.

Table 1. Computational complexity comparisons

Equalizer	FFE	FFE + 2 real VFs	FFE + MIMO-VF	FFE + MIMO-ANN	FFE + Conv. ANN [15]
Multiplication number	804	1146	1358	3688	2896

5. Conclusion

We have proposed and experimentally demonstrated a MIMO-ANN NLE in an SSB SCD system. In our scheme, a 2-by-2 MIMO structure with two ANNs is employed to process the complex QAM signal, thus effectively mitigating the signal distortions including IQ imbalance and fiber nonlinear effects. Based on the proposed MIMO-ANN NLE, we experimentally demonstrated a 112-Gb/s SSB-SCD 16-QAM signal transmission over a single-span 120-km SMF with a BER lower than 3.8×10^{-3} . We also compared the performances of the proposed MIMO-ANN NLE, the FFE, the NLE with 2 real VFs, and the MIMO-VF. The results show that the proposed MIMO-ANN NLE has a better performance than other equalizers in the scenarios with strong nonlinearities, induced by the long fiber transmission length and the high signal launch power. After the single-span 145-km transmission, only the proposed MIMO-ANN NLE can reach the BER below 2.4×10^{-2} . To the best of our knowledge, this is the first experimental demonstration of an ANN-based equalizer in an SSB SCD system.

Funding

National Natural Science Foundation of China (NSFC) (61835008/61860206001).

References

1. Cisco, "Cisco Global Cloud Index: Forecast and Methodology, 2016–2021 White Paper", 2018, <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf>
2. E. Giacomidis, A. Choudhary, E. Magi, D. Marpaung, K. Vu, P. Ma, D.-Y. Choi, S. Madden, B. Corcoran, M. Pelusi, and B. J. Eggleton, "Chip-based Brillouin processing for carrier recovery in self-coherent optical communications," *Optica* **5**(10), 1191–1199 (2018).
3. D. Che and W. Shieh, "Polarization demultiplexing for Stokes vector direct detection," *J. Lightwave Technol.* **34**(2), 754–760 (2016).
4. X. Chen, C. Antonelli, S. Chandrasekhar, G. Raybon, A. Mecozzi, M. Shtaif, and P. Winzer, "Kramers–Kronig Receivers for 100-km Datacenter Interconnects," *J. Lightwave Technol.* **36**(1), 79–89 (2018).
5. G. N. Liu, L. Zhang, T. Zuo, and Q. Zhang, "IM/DD transmission techniques for emerging 5G Fronthaul, DCI, and metro applications," *J. Lightwave Technol.* **36**(2), 560–567 (2018).
6. X. Chen, S. Chandrasekhar, and P. Winzer, "Self-Coherent Systems for Short Reach Transmission," in *European Conference on Optical Communication*, (IEEE, 2018), 1–3.
7. A. Mecozzi, C. Antonelli, and M. Shtaif, "Kramers–Kronig coherent receiver," *Optica* **3**(11), 1220–1227 (2016).
8. Z. Li, M. S. Erkılınç, K. Shi, E. Sillekens, L. Galdino, B. C. Thomsen, P. Bayvel, and R. I. Killey, "SSBI mitigation and the Kramers–Kronig scheme in single-sideband direct-detection transmission with receiver-based electronic dispersion compensation," *J. Lightwave Technol.* **35**(10), 1887–1893 (2017).
9. Y. Su, G. Raybon, L. K. Wickham, R.-J. Essiambre, S. Chandrasekhar, S. Radic, L. E. Nelson, L. Gruner-Nielsen, and B. J. Eggleton, "40-Gb/s transmission over 2000 km of nonzero-dispersion fiber using 100-km amplifier spacing," in *Optical Fiber Communications Conference*, (Optical Society of America, 2002), ThFF3.
10. L. Shu, J. Li, Z. Wan, Z. Yu, X. Li, M. Luo, S. Fu, and K. Xu, "Single-photodiode 112-Gbit/s 16-QAM transmission over 960-km SSMF enabled by Kramers-Kronig detection and sparse I/Q Volterra filter," *Opt. Express* **26**(19), 24564–24576 (2018).
11. L. Yi, T. Liao, L. Huang, L. Xue, P. Li, and W. Hu, "Machine learning for 100Gb/s/λ passive optical network," *J. Lightwave Technol.* **1**, 1 (2018).
12. F. N. Khan, C. Lu, and A. P. T. Lau, "Machine Learning Methods for Optical Communication Systems," in *Advanced Photonics (IPR, NOMA, Sensors, Networks, SPPCom, PS)*, (Optical Society of America, 2017), SpW2F.3.
13. O. Sidelnikov, A. Redyuk, and S. Sygletos, "Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems," *Opt. Express* **26**(25), 32765–32776 (2018).
14. M. A. Jarajreh, E. Giacomidis, I. Aldaya, S. T. Le, A. Tsokanos, Z. Ghassemlooy, N. J. Doran and S. T. Lee, "Artificial neural network nonlinear equalizer for coherent optical OFDM," *IEEE Photonics Technol. Lett.* **27**(4), 387–390 (2015).

15. Z. Wan, J. Li, L. Shu, M. Luo, X. Li, S. Fu, and K. Xu, "Nonlinear equalization based on pruned artificial neural networks for 112-Gb/s SSB-PAM4 transmission over 80-km SSMF," *Opt. Express* **26**(8), 10631–10642 (2018).
16. T. A. Eriksson, H. Bülow, and A. Leven, "Applying neural networks in optical communication systems: possible pitfalls," *IEEE Photonics Technol. Lett.* **29**(23), 2091–2094 (2017).
17. F. Zhang, D. Wang, R. Ding, and Z. Chen, "Terabit Nyquist PDM-32QAM signal transmission with training sequence based time domain channel estimation," *Opt. Express* **22**(19), 23415–23426 (2014).
18. J. Shi, J. Zhang, X. Li, N. Chi, G.-K. Chang, and J. Yu, "112 Gb/s/λ CAP Signals Transmission over 480 km in IM-DD System," in *Optical Fiber Communication Conference*, (Optical Society of America, 2018), W1J.5.
19. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**(7553), 436–444 (2015).
20. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, (NIPS, 2012), 1097–1105.
21. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature* **323**(6088), 533–536 (1986).
22. Z. Yang, F. Gao, S. Fu, X. Li, L. Deng, Z. He, M. Tang, and D. Liu, "Radial basis function neural network enabled C-band 4 × 50 Gb/s PAM-4 transmission over 80 km SSMF," *Opt. Lett.* **43**(15), 3542–3545 (2018).
23. S. T. Le, K. Schuh, M. Chagnon, F. Buchali, R. Dischler, V. Aref, H. Buelow, and K. M. Engenhardt, "1.72-Tb/s virtual-carrier-assisted direct-detection transmission over 200 km," *J. Lightwave Technol.* **36**(6), 1347–1353 (2018).
24. X. Tang, Y. Qiao, J. Zhou, M. Guo, J. Qi, S. Liu, X. Xu, and Y. Lu, "Equalization scheme of C-band PAM4 signal for optical amplified 50-Gb/s PON," *Opt. Express* **26**(25), 33418–33427 (2018).
25. X. Zhou, "an improved feed-forward carrier recovery algorithm for coherent receivers with M-QAM modulation format," *IEEE Photonics Technol. Lett.* **22**(14), 1051–1053 (2010).
26. Y. Zhu, K. Zou, Z. Chen, and F. Zhang, "224 Gb/s optical carrier-assisted Nyquist 16-QAM half-cycle single-sideband direct detection transmission over 160 km SSMF," *J. Lightwave Technol.* **35**(9), 1557–1565 (2017).