135-GHz D-Band 60-Gbps PAM-8 Wireless Transmission Employing a Joint DNN Equalizer With BP and CMMA

Wen Zhou, Li Zhao, Jiao Zhang, Kaihui Wang, Jianjun Yu, *Fellow, IEEE, Fellow, OSA*, You-Wei Chen, Shuyi Shen, Run-Kai Shiu, and Gee-Kung Chang, *Fellow, IEEE, Fellow, OSA*

***Abstract*—In this article, a novel scheme to effectively mitigate the nonlinear impairments in a PAM-8 radio-over-fiber (ROF) delivery is proposed by a joint deep neuron network (J-DNN) equalizer, which has more superiority in terms of good training accuracy, satisfactory tracking speed, and over-fitting suppression compared with a typical deep neuron network (DNN) equalizer. Our proposed J-DNN equalization scheme is mainly based upon back-propagation (BP) algorithm and blind cascaded multi-modulus algorithm (CMMA), which can be trained via two steps including DNN initialization and DNN optimization. By using the proposed J-DNN equalizer, 60-Gbps PAM-8 signal generation and transmission over 10-km SMF and 3-m wireless link at 135-GHz can be achieved. For the digital signal processing (DSP) at receiver, comparisons between CMMA equalizer, DNN equalizer, and J-DNN equalizer are demonstrated. The results indicate that J-DNN equalizer has a much better BER performance in receiver sensitivity than the traditional CMMA, and an improvement of receiver sensitivity can be achieved as much as 1 dB compared with a DNN equalizer at the BER of 3.8×10^{-3}. To the best of our knowledge, this is the first time to propose a novel joint DNN equalizer, which is promising for the development in integrated microwave photonics and microwave/millimeter-wave photonics for 5G applications and beyond.**

***Index Terms*—D-band, J-DNN equalizer, nonlinear effect, PAM-8.**

Manuscript received November 26, 2019; revised February 8, 2020; accepted March 4, 2020. Date of publication March 9, 2020; date of current version July 20, 2020. This work was supported in part by the National Key R&D Program of China under Grant 2018YFB1801004, in part by the NNSF of China under Grant 61527801, Grant 61805043, Grant 61835002, Grant 61675048, and Grant 61720106015, and in part by International Postdoctoral Exchange Fellowship Program from Chinese Postdoctoral Council (No. 20180097). *(Corresponding author: Wen Zhou.)*

Wen Zhou is with the Key Laboratory for Information Science of Electromagnetic Waves, Shanghai Institute for Advanced Communication and Data Science, Fudan University, Shanghai 200433, China, and also with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: zwen@fudan.edu.cn).

Li Zhao, Jiao Zhang, Kaihui Wang, and Jianjun Yu are with the Key Laboratory for Information Science of Electromagnetic Waves, Shanghai Institute for Advanced Communication and Data Science, Fudan University, Shanghai 200433, China (e-mail: 16110720097@fudan.edu.cn; 17110720046@fudan.edu.cn; 16110720021@fudan.edu.cn; jianjun@fudan.edu.cn).

You-Wei Chen, Shuyi Shen, Run-Kai Shiu, and Gee-Kung Chang are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: yu-wei.chen@ece.gatech.edu; ssyzoe@gatech.edu; rshiu6@gatech.edu; geekung.chang@ece.gatech.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2020.2979070

I. Introduction

With the advent of data explosion, the growth of global mobile traffic will be over 40 thousand times from 2010 to 2030 [1]. In order to cope with the future mobile data traffic, the upcoming 5G opens up avenues to new industrial applications, including 3D video, UHD screen, cloud office, virtual reality (VR)/ augmented reality (AR), smart factory, internet of vehicles and 5G is also a critical element to build up a widely connected "smart city". Enhanced mobile broadband (eMBB) is a main challenge for 5G radio network to increase the capacity to 1 Gbit/s, and the peak data rate up to 10 Gbit/s. Millimeter wave (Mm-wave) technology is a promising candidate for providing large capacity [2]–[8], since it occupies wide bandwidth from 30 GHz to 300 GHz. Instead of expensive electronics operating at these high carrier frequencies with a bandwidth limitation, photonics-aided mm-wave technology has been widely applied for the generation and processing of mm-wave [9]–[11]. The millimeter ROF delivery based on more spectrum-efficient modulation will be a promising direction in the development of larger capacity links. However, as for the spectrum-efficient modulation, such as QPSK, 8QAM, 16QAM and 64QAM, it is not a low cost solution because an expensive IQ modulator should be employed. Instead, a more cost-effective MZM is employed to carry PAM-8 data. Moreover, PAM has a better phase noise tolerance than QAM, thus, it can reduce the system complexity effectively. Finally, compared to QAM signals with both I and Q components, PAM modulation format carries only amplitude information, so only the amplitude information needs to be equalized, making it superior in the requirement of equalizer simplicity. Pulse amplitude modulation with 8 levels (PAM-8) format has been extensively investigated due to its high spectral efficiency, lower power consumption and simple configuration [12], [13], and thus it is an effective solution to address the receiver sensitivity problem in radio-over-fiber (ROF) links.

Photonics enabled high frequency transmission work has been massively researched, like [14]–[19], and several nonlinear equalization methods, like Volterra [20], [21], MLSE [22], Kernel [23] has been also employed. References [12], [13] are research work in VLC and DCI. Recently, some reported achievements demonstrated that the D-band (110–170 GHz) transmission has great potentialities for large-capacity line of sight (LOS) indoor applications [24], [25]. However, D-band

is often underutilized due to a larger wireless transmission loss caused by atmospheric absorption [26], [27], and the lack of D-band amplifiers aggravates this problem. Therefore, how to amplify D-band signals and improve the receiver sensitivity over a wireless delivery lead to an interesting issue. At present, intensity modulation coherent detection for m-PAM signal transmission offers cost-effective modulation, high receiver sensitivity and linear detection enabling channel impairment compensation in the digital domain [28]–[30]. At receiver side, PAM-8 signals can be received via coherent detection and processed by the carrier recovery technologies in order to improve the capability of noise resistance such as phase and frequency noises.

In general, nonlinearity in the wireless system is originated in the nonlinear devices including mixers, power amplifiers in the wireless transmitter, and low-noise amplifiers in the receiver [31]. But beyond that, a photonics-aided mm-wave ROF system also suffers from the nonlinear impairment caused by the electro-optical component, photoelectric conversion, as well as the laser and fiber nonlinearity. To address the nonlinearity issue, lots of efforts have been made, i.e., look-up-table pre-distortion [32] has been employed to mitigate the amplitude distortion in the transmitter. But it increases the computation burden at Tx-side and has to collaborate with the other equalization algorithms at Rx-side [33], [34]. In addition, blind equalization is a technology that realizes adaptive channel equalization without the aid of training sequence. In particular, constant modulus algorithm (CMA) has become a favorite of practitioners [35]. However, CMA has a large residual error after convergence and it is not suitable for the nonlinear channel balance.

At present, the most effective approaches to modeling the nonlinear wireless behavior are those based on neuron network equalizers including artificial neuron network (ANN) [36], [37] and deep neuron network (DNN) [38], [39]. In order to obtain the optimum effect of neural network training, a large quantity of training data is required. In Ref. [40], the result has validated that the stable safety neuron network can be implemented when the 10% of data is used for testing while the 90% is for training. However, in a real wireless communication system, a lot of training data means that an effective channel capacity is reduced, and a long training period is undesirable especially for the time varying channel tracking. To reduce the training time period, *Liu et al.* [41] experimentally proposed a 5-Gbps BPSK signal ROF link at 60-GHz based on an adaptive ANN equalizer. The proposed ANN equalizer with an adjustable parameter α can be optimized to achieve a faster learning rate. Besides, he also demonstrated that the equalization performance is improved with the increasing neuron numbers in the hidden layer [42], but the improvement is limited since over-fitting effect is introduced. Besides, although the implementation of DNN involves a larger number of matrix and vector calculations compared with ANN, a more improvement of BER decision accuracy can be achieved, which is more helpful to handle the receiver sensitivity issue in D-band wireless links.

In this paper, we proposed a novel joint DNN (J-DNN) equalizer for the combination of two error functions, which is based upon back-propagation (BP) algorithm and blind cascaded multi-modulus algorithm (CMMA). In our proposed scheme,

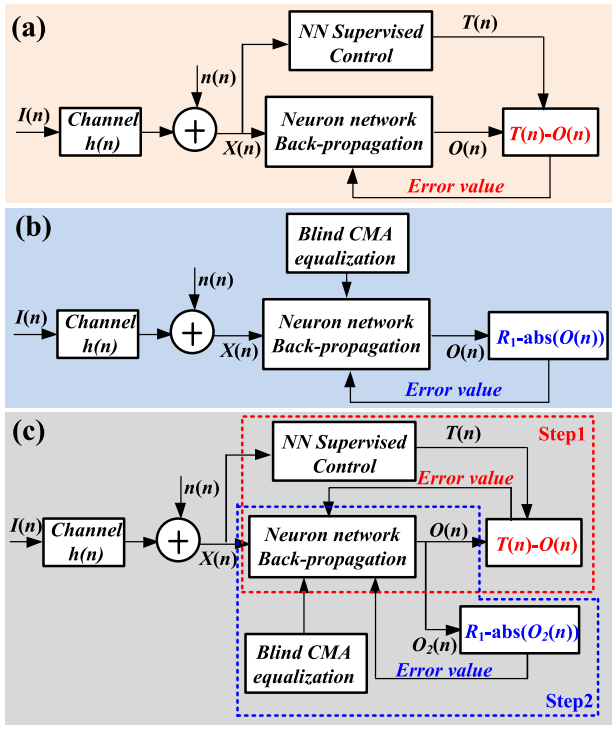

Fig. 1. Schematic diagrams of NN equalizers (a) the adaptive NN equalization, (b) CMA blind equalization based on NN, and (c) our proposed J-DNN equalizer.

DNN network can be established via two steps. Firstly, it can be initialized with the aid of the training sequence, and then the weight value of which can be further optimized by employing the error function of CMMA. Adopting our proposed J-DNN equalizer, both of the length of the training data and the training time can be effectively reduced, as well as the over-fitting effect. Based on our proposed scheme, we experimentally demonstrated a 60 Gbit/s PAM-8 signal generation and 3-m wireless transmission at 135 GHz with BER below 3.8×10^{-3}. The results show that J-DNN equalizer can improve the performance of receiver sensitivity greatly with the comparison of the traditional CMMA, and an improvement of receiver sensitivity can be achieved as much as 1 dB compared with a DNN equalizer at the BER of 3.8×10^{-3}.

II. Operation Principle for a J-DNN Equalizer

Conventionally, NN consists of one input layer, multi hidden layers and one output layer, and the connection of which can be held through the nodes of each layer. There are two main classification methods for NN equalization algorithm such as blind NN equalization and adaptive NN equalization, as shown in Fig. 1. Fig. 1(a) corresponds to the adaptive NN equalization [37], [41], [42], the cost function of which is $J_n = \frac{1}{2}\sum_{n=1}^{S}(T_n - O_n)^2$, where T_n is the expected value and O_n is the obtained value after NN adaptive equalizer. But it also has severe drawbacks such as a large size of training sequence, long training period, over-fitting effect and slow convergence. Fig. 1(b) gives the schematic diagram of blind equalization in neural network, which is a self-recovering equalization method without the aid

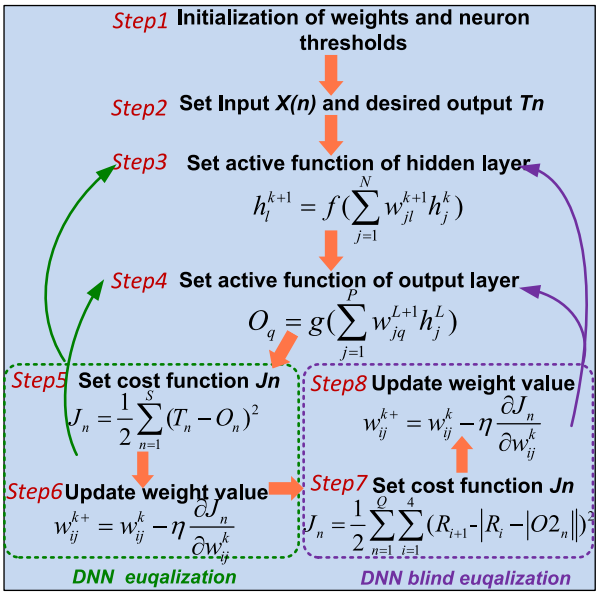

Fig. 2. Framework of the proposed J-DNN equalizer.

of training sequence and enhances the bandwidth effectively. For example, CMA blind equalization can be combined with NN algorithm [43], the cost function of which is defined as $J_{n\_CMA} = \frac{1}{2}\sum_{n=1}^{Q}(R_1 - |O_n|)^2$, where R_1 is the constant module of signals (i.e., OOK and QPSK). However, its higher residual mean square error (MSE) is a major challenge for equalizing nonlinear channels. So, we propose a new joint DNN (J-DNN) equalizer to conquer the defects of both blind NN equalization and adaptive NN equalization methods, as shown in Fig. 1(c).

Fig. 2 shows the flow chart of our proposed J-DNN equalizer. Each link between one input layer, L hidden layers and one output layer in J-DNN is associated with the weight value w_{ij}^k, where k denotes the *k-th* layer (i.e., 0-*th* for the input layer, (L + 1)-*th* for the output layer), i and j represent the *i-th* node in the former hidden layer and the *j-th* node in the current layer, respectively. From the first step, J-DNN can be initialized via the weight value randomization ($w_{ij}^k \in [-1, 1]$), learning rate and iterative learning epoch setting. Then the input samples $X(n)$ are randomly selected, and the corresponding target output data T_n is decided. Here, $X(n) = [x(n), x(n-1) \ldots, x(n-M+1)]^{\mathbf{T}}$ is defined as the input vector with M memory length, which is multiplied by $w_{ij}^1 (i = 1 \ldots M)$ in the first hidden layer and after which, the output of different nonlinear neurons are summed as

$$h_j^1 = f\left(\sum_{i=1}^{M} w_{ij}^1 x_i\right) \tag{1}$$

Where f (.) denotes the nonlinear active function between multiple hidden layers. Selection of a good nonlinear activation function is an important part of DNN network construction. Some typical active functions have been extensively deployed such as tanh and sigmoid. Since sigmoid nonlinear function has non zero-centered output with a value ranging from 0 to 1, a slower learning would be caused from "Zigzag" reverse updating of the weights [44]. So, we use tanh function with a zero-centered output of [−1, 1] to avoid this issue, which can be depicted as

$$f(x) = \tanh(x) \tag{2}$$

The derivate function of Eq. (2) is

$$f'(x) = 1 - \tanh^2(x) \tag{3}$$

In the training feed-forward process, the output of the *l-th* neuron in the other (k + 1)-*th* layer can be calculated as

$$h_l^{k+1} = f\left(\sum_{j=1}^{N} w_{jl}^{k+1} h_j^k\right) \tag{4}$$

where h_j^k is the output of the *j-th* neuron cell in the previous *k-th* hidden layer, and the total number of cells in this layer is N. When k is increased as L, we define h_l^{L+1}as the final output signal calculated from the output layer. It is worth noting that the transformation from input layer to hidden layer is nonlinear but that from hidden layer to output layer is linear. Thus, in Step 4, the function g (.) denotes linear function "purelin" employed in output layer.

Different from steps 1–4 illustrated as forward propagation, the steps 5–8 are used to update the weight value based on BP algorithm. According to the regular minimum mean square error (MMSE) algorithm, the cost function can be defined as

$$J_n = \frac{1}{2}\sum_{n=1}^{S}(T_n - O_n)^2 \tag{5}$$

where S is the length of the training data $X(n)$. Compared the obtained value O_n with corresponding expected value T_n, the error $e_n = T_n - O_n$ is sent to network with reverse learning algorithm. So that the connected weight vector w_{ij}^{k+} can be iteratively updated until the desired epoch or error value is reached, which can be represented as

$$w_{ij}^{k+} = w_{ij}^k - \Delta w_{ij}^k = w_{ij}^k - \eta \frac{\partial J_n}{\partial w_{ij}^k} \tag{6}$$

where

$$\frac{\partial J_n}{\partial w_{ij}^k} = e_n \cdot \frac{\partial O_n}{\partial w_{ij}^k} = e_n \cdot \frac{\partial O_n}{\partial h_j^L} \cdot \frac{\partial h_j^L}{\partial h_j^{L-1}} \cdots \frac{\partial h_j^k}{\partial w_{ij}^k} \tag{7}$$

Here, η is the learning rate and Δ symbol represents the gradient operation. In order to update the weight value of every nonlinear node accurately, we have to calculate the gradient of the whole training sequence. That is the reason a very slow convergence is induced in DNN network, especially dealing with a large dataset. If the error function of blind equalization is introduced into DNN, we will decrease the size of the training set and improve the computation speed effectively.

As we know, CMA is a typical blind adaptive equalizer for OOK or QPSK modulation format with a single reference radius, but it is less effective for higher order modulation format like PAM-8 with four symmetrical levels of amplitude. Therefore, CMMA has been proposed to address this issue since its error function employs multilevel decision, as shown in Fig. 3.

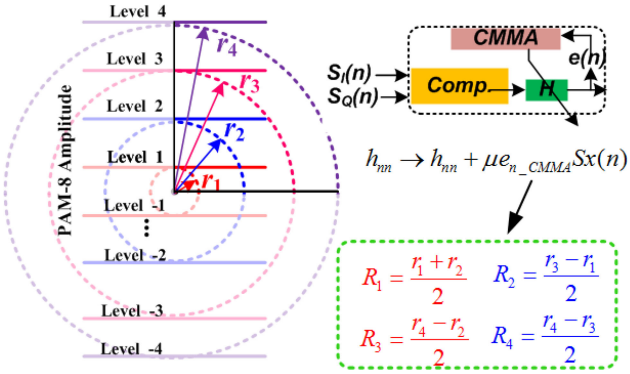

Fig. 3. CMMA equalization algorithm for PAM-8 signal.

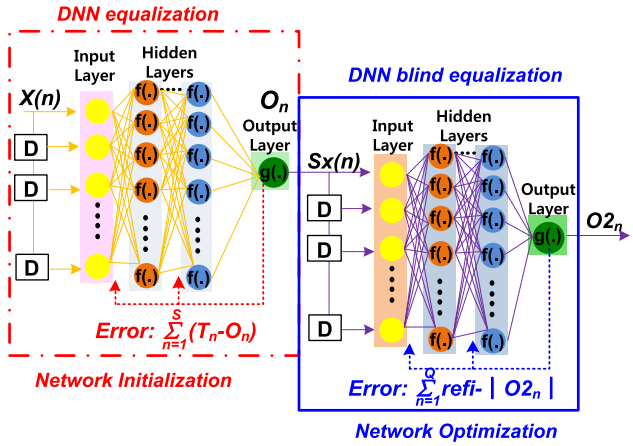

Fig. 4. Architecture of the proposed delay-tap joint DNN equalizer.

For received PAM-8 signal $Sx(n)$ equalization, μ is defined as the CMMA convergence parameter, so the corresponding filter tap weight h_{nn} updates as follows

$$h_{nn} \to h_{nn} + \mu e_{n\_CMMA} Sx(n). \tag{8}$$

Where the error function e_{n_CMMA}can be defined as

$$\begin{aligned} e_{n\_CMMA} = & \left(R_4 - |R_3 - |R_2 - |R_1 - |O2_n||||\right) \\ & \cdot sign\left(R_3 - |R_2 - |R_1 - |O2_n|||\right) \ldots \\ & \cdot sign\left(R_2 - |R_1 - |O2_n||\right) \\ & \cdot sign\left(R_1 - |O2_n|\right) \cdot sign(O2_n). \end{aligned} \tag{9}$$

Here, $R_1 = \frac{r_1+r_2}{2}$, $R_2 = \frac{r_3-r_1}{2}$, $R_3 = \frac{r_4-r_2}{2}$ and $R_4 = \frac{r_4-r_3}{2}$. r_1, r_2, r_3 and r_4 represent the absolute value of four reference levels respectively, as shown in Fig. 2. $O2_n$ is the output of CMMA equalizer. In practical, we can fix the optimum μ and tap number to ensure better channel equalization effect.

Unlike CMMA, our proposed J-DNN just utilizes the error function in Eq. (9), and then substitutes it into the cost function based on blind equalization algorithm, which is depicted as

$$J_{n\_CMMA} = \frac{1}{2} \sum_{n=1}^{Q} \left(R_4 - |R_3 - |R_2 - |R_1 - |O2_n||||\right)^2 \tag{10}$$

where $O2_n$is the output from J-DNN network, and the input data into J-DNN optimization network is the test signal Sx_n with a size of Q. So the weight value can be further optimized with the aid of blind equalization and BP methods in Step 8, which can be given as

$$w_{ij}^{k+} = w_{ij}^{k} - \Delta w_{ij}^{k} = w_{ij}^{k} - \eta \frac{\partial J_{n\_CMMA}}{\partial w_{ij}^{k}} \tag{11}$$

where

$$\begin{aligned} \frac{\partial J_{n\_CMMA}}{\partial w_{ij}^{k}} &= e_{n\_CMMA} \cdot \frac{\partial O2_n}{\partial w_{ij}^{k}} \\ &= e_{n\_CMMA} \cdot \frac{\partial O2_n}{\partial h_j^L} \cdot \frac{\partial h_j^L}{\partial h_j^{L-1}} \cdots \frac{\partial h_j^k}{\partial w_{ij}^{k}} \end{aligned} \tag{12}$$

Thus, our proposed J-DNN is composed of two parts including DNN equalization and DNN blind equalization. The former updating of the weight value is based on traditional BP algorithm in Steps 5 and 6, while the latter one deploys the blind CMMA equalization algorithm to further optimize the network in Steps 7 and 8. The total number of multiplications in a traditional 3-layer (input, hidden and output layer) NN is $4\sum_{i=0}^{1} n_i n_{i+1} + 3\sum_{i=1}^{2} n_i - n_0 n_1 + 2n_2$ [45], where n_0 is the number of neurons in input layer, n_1 represents the number of neurons in input layer, and n_2 is the number of neurons in output layer. When the number of neurons in output layer is fixed as $1(n_2 = 1)$, the total multiplications is $3n_0 n_1 + 7n_1 + 5$ in one iteration. Since our proposed J-DNN can be processed via two steps including DNN equalization and DNN blind equalization, the total number of multiplications is the sum of the two steps. Assumed the number iterations in step 1 and step 2 is i_1 and i_2, respectively, so the number of multiplications in J-DNN algorithm equals to $(i_1 + i_2)(3n_0 n_1 + 7n_1 + 5)$.

Fig. 4 shows the architecture of the proposed delay-tap joint DNN equalizer. Our proposed J-DNN is a focused time delay neural network (FTDNN), the output of which depends not only on the current input to the network, but also on the current or previous inputs of the network. So 'delay-tap' means that the input weight has a tap delay line associated with it. This allows the network to have a finite dynamic response to time series input data, e.g., $X(n) = [x(n), x(n-1) \ldots, x(n-M+1)]^{\mathrm{T}}$ is defined as the input vector with M memory length in our paper. In summary, the proposed J-DNN is processed via two steps. The first training procedure shown as the dotted red block is used for the network initialization, where $X(n)$ is the training set and O_n presents the approximated output after DNN equalization. Here, the difference between the target output T_n and the training output O_n is utilized as the error function. Then with the initialized DNN, we can send the testing set $Sx(n)$ into the network, and take an action according to the blind CMMA approach to achieve a better equalization effect. Actually, the network shown in blue block is our finally trained DNN network. There is only one DNN in our proposed J-DNN. In DNN blind equalization, the weight value of the same DNN in DNN equalization process can be optimized according to the cost function based on blind equalization algorithm. Our proposed J-DNN can enhance the

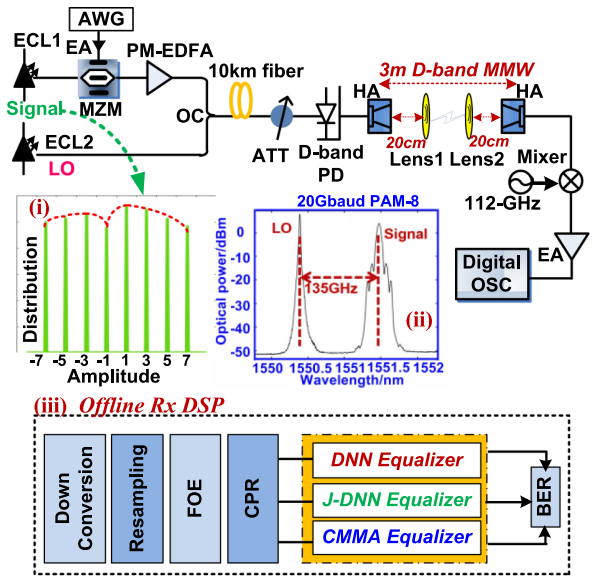

Fig. 5. The ROF experimental setup for 20 Gbaud PAM-8 wireless transmission at D-band with detailed DSP blocks such as DNN equalizer, J-DNN equalizer and CMMA equalizer, respectively. ECL: external cavity laser, EA: electrical amplifier, MZM: Mach–Zehnder modulator, PM-EDFA: polarization-maintaining Erbium-doped fiber amplifier, OC: optical coupler, ATT: attenuator, PD: photodiode, HA: horn antenna, AWG: arbitrary wave generator, OSC: oscilloscope. Insets (i) Transmitted PAM-8 amplitude distribution, (ii) the optical spectrum of 20 Gbaud PAM-8 signal after the optical coupler, (iii) the detailed DSP block at the receiver end.

adaptive ability to the channel variance of the wireless system without the aid of training samples.

III. Experimental Setup

Fig. 5 depicts the experimental setup of our demonstrated D-band PAM-8 delivery over 10 km single mode fiber (SMF) and 3 m wireless distance, where the transmitted PAM-8 signal distribution $X(k)\in[\pm7, \pm5, \pm3, \pm1]$ with red dotted envelope is illustrated in Inset (i). Here, we employ two individual external cavity lasers (ECL1 and ECL2) to generate D-band mm-wave signals although the phases of the two optical waves are not correlated. Techniques have been used to reduce phase noise, i.e., (a) Optical Phase Locked Loops (OPLL) and (b) Optical Injection Locking (OIL), but both OPLL and OIL methods need stable external microwave signal sources, which leads to cost increase [46], [47]; (c) Dual-wavelength lasers, but the millimeter wave signals continuously tuned cannot be generated [48]; (d) Mode-locked-laser (MLL), many devices including MLL, inter-leaver, EDFA, band-pass filter and polarization maintaining 3 dB coupler are required to generate high frequency mm-wave signals, which not only complicates the system but also increases system cost [49]. Moreover, the generated mm-wave frequency $n\triangle f$ depends on the frequency spacing $\triangle f$ of MLL, which is not flexible in the real communication system with fixed generation frequency. So the adoption of two individual lasers is a relatively simple, flexible and cost-effective architecture for D-band mm-wave signal generation.

The baseband PAM-8 signal is digital-to-analog (DAC) converted by the arbitrary wave generator (AWG) with a sampling rate of 80GSa/s. The effective number of bits (ENOB) of the AWG is a little different at different frequency. The ENOB at DC∼5, 5∼10, 10∼15, 15∼20 GHz is 6, 5.5, 5, 4.6bits, respectively. After amplified by a cascaded electrical amplifier (EA), the boosted PAM-8 signal is used to drive the intensity modulator. The external cavity laser (ECL1) at 1551.35 nm with line width of 100 KHz and an average power of 16 dBm is operated as a signal light source, which is modulated via the intensity modulator MZM. Here, the deployed MZM has a 3 dB optical bandwidth of 30-GHz, half-wave voltage of 2.7-V at 1 GHz, and 5-dB insertion loss. ECL2 at the center wavelength of 1550.26 nm works as a local oscillator (LO), which has a frequency space of 135-GHz with the modulated ECL1 light, as shown in Inset (b). This figure shows the optical spectrum after the combination of amplified ECL1 from PM-EDFA and ECL2. The coupled light beam can be delivered over 10 km standard single mode fiber (SSMF) with 17-ps/km/nm CD at 1550 nm, and the optical power of which is adjusted to obtain the optimum input power into D-band photodiode (PD) by an attenuator (ATT). The adopted D-band PD in our experiment is implemented within the frequency range of 10∼170 GHz at −2V DC bias, and the output power of which is −7 dBm.

At the D-band wireless transmitter, the generated 135-GHz signal is emitted from D-band HA with a gain of 25 dBi gain. A paired D-band HA is employed to receive D-band signal. However, the D-band wireless link composed solely of a pair of HAs would not work without appropriate D-band amplifiers, due to low SNR. Here, we employ a pair of identical lens (Lens 1 and Lens 2) between HAs to realize the focusing and amplifying effects of mm-wave signals, the diameter and focal length of which are 10 cm and 20 cm, respectively. The transmitted HA is just placed at focus of Lens 1 and then the wireless mm-wave signal is focused at the position of received D-band HA. According to the power budget of wireless transmission link over free space, the received power can be defined as $P_{R1} = P_T + G_T + G_R + G_{lens} - 20\log(4\pi d/\lambda) - L_m$, where G_T and G_R denote the gain of transmitted and received HAs respectively, P_T is transmitted power,d is the wireless transmission distance, λ is the wavelength of the mm-wave signal, and L_m is the wireless transmission loss caused by atmosphere absorption. We have measured that the power penalty caused by 3.1-m wireless transmission at 135 GHz is ∼7.5 dB compared to the case of no wireless delivery (the HAs are also removed). It means that the received power is $P_{R2} = P_T$ for no wireless transmission (the HAs and lens are removed), and $20\log(4\pi d/\lambda)$ equals to 84.58 dB. L_m is ∼0.0012 dB at 135 GHz over 3 m wireless distance and can be neglected. We can calculate that the gain G_{lens} of a pair of lens is at least 27.08 dB (84.58-25-25-7.5 = 27.08 dB).

At the receiver end, the received signal at 135 GHz is firstly down converted into an intermediate frequency (IF) signal by a commercial D-band mixer with a 9.5 dB conversion loss and a local oscillator source conducted with 112 GHz. And then the IF signal at 23 GHz is boosted by using an electrical amplifier (EA) with 33-dB gain and 14-dBm saturation output power available from DC to 50 GHz frequency band. Finally, the boosted signal is captured by a digital storage oscilloscope

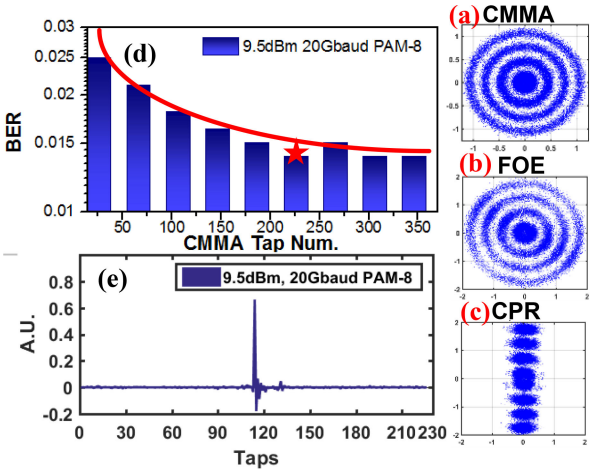

Fig. 6. Constellation diagrams of 20 Gbaud PAM-8 signals after (a) CMMA equalization, (b) FOE, and (c) CPR. (d) BER performance vs CMMA tap number when the input power into PD is 9.5 dBm. (e) The optimum tap distribution.

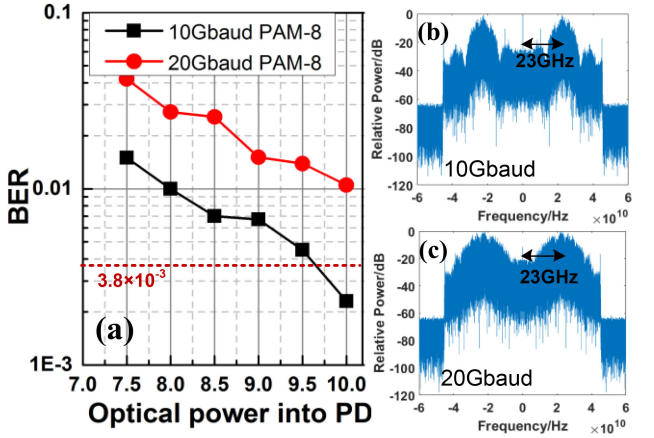

Fig. 7. (a) BER performance vs the optical power into PD when the transmission speed is 10 Gbaud and 20 Gbaud, respectively. Electrical spectrum of IF signal at (b) 10Gbaud, (c) 20 Gbaud.

(OSC). The deployed OSC has a sampling rate of 120 GSa/s and an electrical bandwidth of 45 GHz. As Inset (iii) shown in Fig. 5, the captured signal is offline processed by DSP steps including down conversion into base band, resampling, frequency offset estimation (FOE) and carrier phase recovery (CPR). In particular, we also compare the BER performance between CMMA equalizer, DNN equalizer and J-DNN equalizer after these DSP steps.

IV. Experimental Results and Discussions

As we know, a well-designed equalizer is good at resolving nonlinear problems and has been successfully applied in radio communications. However, error function is an important factor that affects residual error of equalizer, so if error function is different, the performance of equalizer is also different. We will introduce some equalizers in our proposed system and compare their performance, such as a typical CMMA equalizer with taps, a constructed DNN equalizer and our designed J-DNN equalizer.

A. The Typical CMMA Equalization Method

Figs. 6(a)–(c) show the recovered 20 Gbaud PAM-8 signals after CMMA equalization, FOE, CPR, respectively when the input power into PD is fixed as 9.5 dBm. To ensure a better convergence effect, the number of taps can be adjusted from 25 to 375. Fig. 6(d) corresponds to the comparison result between different taps, from which we can conclude that the increase of taps improves BER performance smoothly and BER is finally stable. As the red star signed in the figure, the optimum tap number is 226 and BER is deceased to 0.0139. The relative 226-tap value distribution is also given in Fig. 6(e).

With the optimized 226-tap, we further discuss BER performance versus the optical power into PD, as shown in Fig. 7(a). It is obviously that BER performance is worsen when the transmission speed increases from 10 Gbaud to 20 Gbaud. BER can achieve under 3.8×10^{-3} threshold for 7% hard-decision forward error correction (HD-FEC) only when the optical power is above 9.5 dBm at 10 Gbaud. Figs. 7(b) and (c) give the relative electrical spectra of received IF signal at 23 GHz when the data rate is 10 Gbaud and 20 Gbaud, respectively. From the experimental result, it can be concluded that the nonlinear mitigation effect of CMMA technique is not satisfying due to its large MSE, which is consist with our previously discussion in Section II.

B. The Comparison Between DNN and J-DNN Equalization

In wireless communication field, the most difficult point is non-linear system identification and mitigation. Some achievements have been achieved in applying deep neural networks to the study of non-linear distortion, but its mitigation accuracy, low training speed, efficiency and over-fitting effect are not ideal enough. Based on our proposed experiment system for 20 Gbaud D-band PAM-8 wireless transmission, we will further compare the performance of DNN and J-DNN equalization in terms of hidden layers, training data size, neuron number in hidden layer, as well as the training time cost.

C. Hidden Layers

Generally, the increase of hidden layers in DNN improves the accuracy effectively, but it also brings a series of problems especially a larger computation burden and time cost during the training process. Fig. 8 illustrates BER of 20 Gbaud PAM-8 signals versus the input optical power into PD. Here, we fix the neuron numbers of input and output are fixed as 171 and 1, respectively. It can be discussed in 4 cases: (1) DNN with a single hidden layer (171-60-1), which means that the number of neurons in the single hidden layer is 60. (2) DNN with two hidden layers (171-60-60-1), which means that the number of neurons in the each hidden layers is 60. (3) DNN with three hidden layers (171-60-60-60-1), which means that the number of neurons in the each hidden layers is 60. (4) There is only one hi&dden layer with 60 neuron cells in our proposed J-DNN

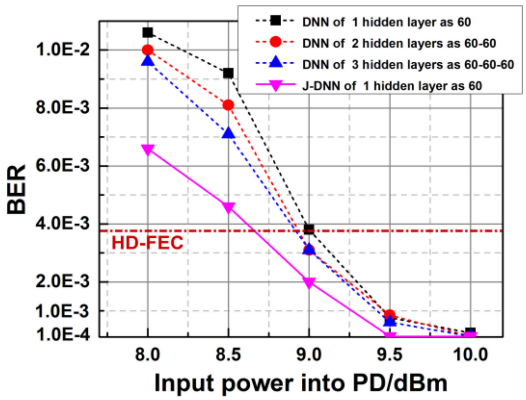

Fig. 8. BER performance vs the optical power into PD when there are 1, 2, 3 hidden layers in DNN, respectively. The pink solid line corresponds to the scenario when there is only one hidden layer in J-DNN.

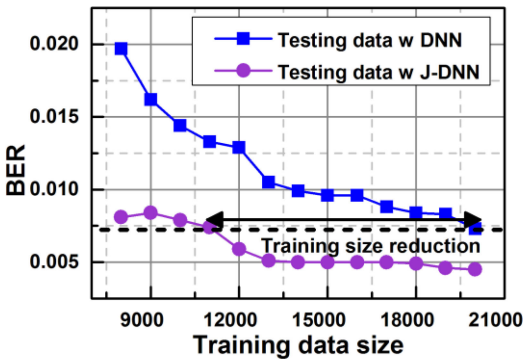

Fig. 9. BER performance of the testing data vs the training data size in a regular DNN equalizer and our proposed J-DNN equalizer, respectively.

structure (171-60-1). For the discussion of these 4 cases, the training data size is 16000, and the iterative epoch is 400. It can be found that BER performance is better when there is 3 hidden layers in DNN compared with the other two cases (case 1 and case 2). However, using our proposed J-DNN equalization method in case 4, BER is reduced significantly in a low SNR region. Besides, from the aspects of complexity and training speed, our proposed J-DNN also performs better since there is only one hidden layer composed of the same amount of nodes.

D. Training Data Size

During our training process, the total received PAM-8 signal is randomly divided into the training data and the testing dat. If a large training dataset is selected, which will not only increase the computation complexity and time delay but also reduce the efficient system capacity. While a shorter length of training data means a lower identification precision. So, the selection of training data scale is a key influential factor of DNN network. Fig. 9 shows the BER performance of testing data changing with the training data size, which corresponds to the scenario when the input power is fixed as 8 dBm. From the result, it is obvious that the precision of the existing DNN equalizer depends on the initial training set, and BER decreases effectively with the increasing of training block length. When there are 20000

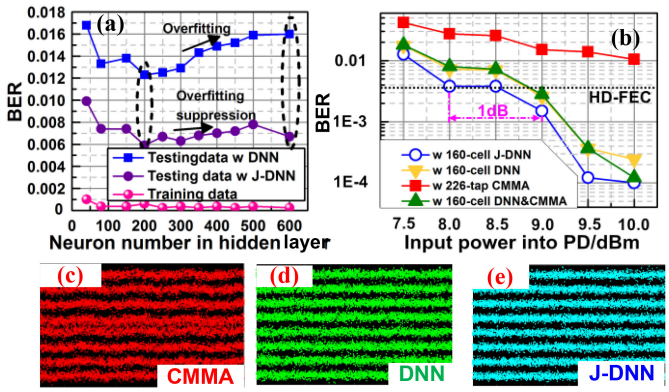

Fig. 10. (a) BER performance vs the neuron number in a hidden layer of DNN and J-DNN network, respectively. (b) BER performance vs the input power into PD. The recovered PAM-8 symbol distribution after (c) CMMA equalizer, (d) DNN equalizer, (e) J-DNN equalizer.

samples used as the training data, BER can be decreased as low as 0.0073. At the same time, using our proposed J-DNN equalizer, BER changes smoothly with the training data size ranging from 7000 to 20000.

Here, the training data is only used for DNN equalization initialization in our proposed J-DNN equalizer since blind DNN equalization and optimization in step 2 is a self-recovering equalization method without the aid of training sequence. It is worth noting that if only 11000 examples as training data are adopted in our proposed J-DNN equalizer, the same BER performance can be obtained as that of a regular DNN equalizer with a training length of 20000 samples. It indicates that J-DNN neural network has good training accuracy and satisfactory tracking speed. With the consideration of accuracy, time cost and complexity, we trained the network using 16000 training samples in the further discussion.

E. Neuron Number in Hidden Layer

The usual criterion of the neuron cells rising in a hidden layer, which is based only on improving training accuracy, often results in an "over-fitting". Although the over-fitting model has outperformance with the training set, it has underperformance with the testing case. Fig. 10(a) gives the relationship between the BER performance and neuron cells in one hidden layer when the input power is 8 dBm. The training error decreases and is approximately zero with the increasing of cell amount. While the testing error in DNN grows when the number of nodes is greater than 200, as a result of the over-fitting effect. Different from DNN, our proposed J-DNN suppress the over-fitting effectively and the testing error remains stable with a large neuron number. That is because DNN blind equalization process in J-DNN avoids the impact of training sequence, and the over parameterization of training set can be reduced in the whole J-DNN network. Therefore, the optimal neuron scale is around 200 for both DNN and J-DNN equalizers. Fig. 10(b) shows BER performance as a function of the input optical power into PD, where it can be found that increasing the optical power can help improve BER performance due to larger SNR. Here, we compare

four kinds of equalizers including J-DNN equalizer with 160 cells in a hidden layer, DNN equalizer with 160 cells, 226-tap CMMA equalizer and 160-cell DNN equalizer combined with CMMA. The yellow triangles overlapped with green triangles indicate that the combination of CMMA and DNN provides a negligible help from the standalone DNN equalizer. In addition, from the comparison between J-DNN and DNN, it can be concluded that the required power under HD-FEC (3.8×10^{-3}) utilizing the J-DNN scheme is close to 8 dBm, which significantly enhances by almost 1 dB in receiver sensitivity compared with the DNN equalization scheme. Figs. 10(c)–(e) depict the recovered PAM-8 symbol distribution after CMMA equalization, DNN equalization and J-DNN equalization, respectively. In Fig. 10(c), it can be found that the amplitude fluctuations and distortion are induced from the nonlinear effect. That is because CMMA has a large residual error after convergence and it is not suitable for the nonlinear channel balance. Our proposed J-DNN network can be established via two steps. Firstly, it can be initialized with the aid of the training sequence, according to the cost function which can be defined as $J_n = \frac{1}{2}\sum_{n=1}^{S}(T_n - O_n)^2$, where T_n is the expected value and O_n is the obtained value after NN adaptive equalizer. It is similar to the conventional DNN equalization, and the nonlinear effect can be mitigated effectively in this step. Secondly, our proposed J-DNN can be further optimized by utilizing the cost function based on blind equalization algorithm as $J_{n\_CMMA} = \frac{1}{2}\sum_{n=1}^{Q}(R_4 - |R_3 - |R_2 - |R_1 - |O2_n||||)^2$. Here, unlike the conventional CMMA, the second CMMA blind equalization step based on NN also has a favorable nonlinear mitigation performance because of the nonlinear functions employed in NN. Therefore, the amplitude distortion issue in Fig. 10(e) can be effectively mitigated by employing our proposed J-DNN equalizer.

F. Training Iterative Epoch

Another key factor that directly influences the learning speed is referred to as the parameter of training iterative epoch. Usually only one epoch iteration is not enough for an optimal updating of weight values. Excessive epochs result in an "over-fitting" while too few epochs lead to an "under-fitting" effect. Fig. 11(a) illustrates MSE versus the epoch value when the optical power is fixed as 8 dBm, where the black line implies that MSE in DNN (171-80-1) decreases with the amount of epoch. Since our proposed J-DNN (171-80-1) is constructed by two steps such as DNN initialization and DNN blind equalization, the epochs of these two steps are summed as the total epoch in the whole J-DNN network. For example, the red line represents MSE of J-DNN after 400 epoch iterations during DNN initialization process. We can conclude that the convergence speed of the DNN is slow, while our proposed J-DNN scheme has a faster convergence speed and achieves a lower MSE floor after more than 100 iterations in DNN blind equalization process. This conclusion can be also observed in Fig. 11(b).

Seen in Fig. 11(b), the corresponding BER of 20 Gbaud PAM-8 signals after DNN equalization reduces with the epoch

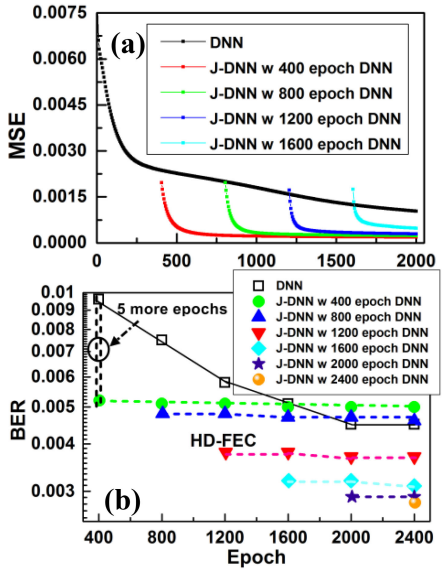

Fig. 11. (a) MSE vs. the epoch iteration for DNN and J-DNN after 400-, 800-, 1200- and 1600-epoch DNN initialization, respectively. (b) BER performance vs the epoch iteration for DNN and J-DNN after 400-, 800-, 1200-, 1600-, 2000- and 2400-epoch DNN initialization, respectively.

increasing and keeps stable when the epoch is improved as 2000, due to the "over-fitting" effect. For the discussion of J-DNN, after DNN initialization with 400 epochs, the network is subsequently optimized with the help of blind DNN equalization, the BER of which is given as the green circles. We also investigate J-DNN cases with 800-, 1200-, 1600-, 2000- and 2400-epoch DNN initialization network, respectively.

It can be found that BER rapidly drops with more than 5 epochs in DNN blind equalization step, and remains constant after totally 405 epoch iterations. There is an obvious epoch reduction between J-DNN with 405 epochs and DNN with 1600 epochs. In particular, BER after 1205 epoch iterations is reduced as low as HD-FEC in our proposed J-DNN scheme, which gradually decreases to 0.0028 when the number of epoch increases to 2405. However, the increase of epoch would be an ineffective method of improving BER performance due to the undesirable over-fitting effect in traditional DNN network.

Fig. 12 shows the BER of the J-DNN equalizer (171-80-1) versus the input power into PD in comparison with DNN equalizer (171-80-1), which correspond to the same scenario when the number of neurons is fixed as 80. It is obvious that there is an under-fitting problem in DNN generalized model with 400 epochs. Meanwhile, with a higher SNR, as much as 2000 epochs cause an over-fitting effect, as the blue solid line going beyond the red solid one in the figure. Differently, the fitting dotted blue and red curves are well separated, which implies that BER of the system deploying J-DNN equalization keeps decreasing when the iteration amount reaches 2005 epochs. Therefore, our suggested J-DNN method can improve the fitting results yielded by the DNN model in the case of over fitting. In case of 1200 epochs, the receiver sensitivity using J-DNN equalization

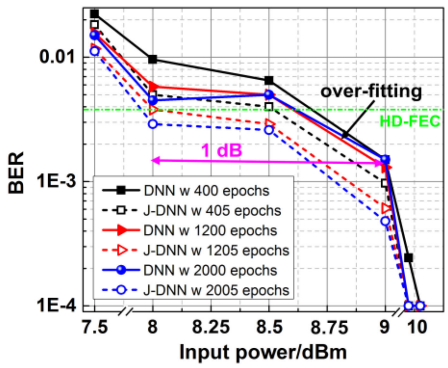

Fig. 12. BER performance vs the optical power into PD for 20 Gbaud PAM-8 signal wireless transmission by employing 400-, 1200- and 2000-epoch DNN equalizer, as well as 405-, 1205- and 2005-epoch J-DNN equalizer, respectively.

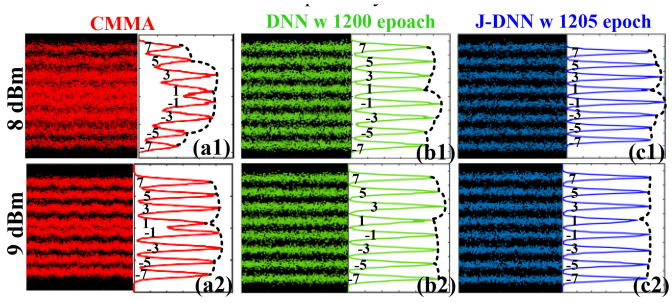

Fig. 13. The recovered PAM-8 symbol pattern with input power of 8 dBm and 9 dBm, respectively after (a1)–(a2) CMMA equalizer, (b1)–(b2) DNN equalizer with 1200 epochs, (c1)–(c2) J-DNN equalizer with 1205 epochs.

can be improved as much as 1 dB compared with the DNN at the BER of 3.8×10^{-3}.

We also give the recovered PAM-8 symbol pattern and distribution after CMMA equalizer, DNN equalizer with 1200 epochs, and J-DNN equalizer with 1205, respectively in Fig. 13. When the optical power is 8 dBm, the random noise dominates the BER performance and introduces serious interferences between lower PAM-8 levels, as shown in Fig. 13(a1). Thanks to the J-DNN equalizer, PAM-8 signal with an average and narrow distribution can be achieved in Fig. 13(c1). With an increasing SNR, the recovered PAM-8 signals presented by the J-DNN equalizer in Fig. 13(c2) are more close to the transmitted data than the traditional DNN results in Fig. 13(b2), the peak distribution of which is closer to that of transmitted data shown as Inset (i) in Fig. 5, and the bottom value is approaching zero.

V. Conclusion

In this paper, a new J-DNN equalizer for the 60-Gbps PAM-8 signal ROF transmission system over 10-km SMF and 3-m wireless link at 135-GHz is experimentally demonstrated, which consists of two steps including DNN initialization and DNN blind equalization. We compare the novel J-DNN equalizer with a classical DNN equalizer in terms of the computation complexity, training data size, learning speed and over-fitting effect. Over-fitting effect is often caused by the training set, while our proposed J-DNN in blind equalization step effectively avoids the training data to reduce the over-fitting effect. The experimental results show that the new method has a good training accuracy, less requirements of training block length and satisfactory tracking speed. The proposed J-DNN converges after approximately 405 iterations, whereas the traditional DNN needs 1600 iterations to achieve the same BER level. Thanks to our proposed J-DNN scheme, an improvement of 1 dB over the J-DNN equalizer in receiver sensitivity at a BER of 3.8×10^{-3} is achieved in comparison with a tradition DNN with a same structure and iteration amount. To the best of our knowledge, this is the first time to deploy a joint DNN equalizer in D-band ROF link. Our proposed J-DNN equalization scheme is promising for the future 5G ROF-based communication application.

References

[1] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, Oct. 2013.
[2] X. Li, J. Zhang, J. Xiao, Z. Zhang, Y. Xu, and J. Yu, "W-band 8QAM vector signal generation by MZM-based photonic frequency octupling," *IEEE Photon. Technol. Lett.*, vol. 27, no. 12, pp. 1257–1260, Mar. 2015.
[3] W. Zhou and C. Qin, "Simultaneous generation of 40, 80 and 120 GHz optical millimeter-wave from one Mach–Zehnder modulator and demonstration of millimeter-wave transmission and down-conversion," *Opt. Commun.*, vol. 398, pp. 101–106, Apr. 2017.
[4] T. Nagatsuma, "Photonic generation of millimeter waves and its applications," in *Proc. Opt. Fiber Commun. Conf.*, Los Angeles, CA, USA, 2012, Paper OM2B.7.
[5] X. Li, J. Yu, J. Xiao, and Y. Xu, "Fiber-wireless-fiber link for 128-Gb/s PDM-16QAM signal transmission at W-band," *IEEE Photon. Technol. Lett.*, vol. 26, no. 19, pp. 1948–1951, Jul. 2014.
[6] H. T. Huang *et al.*, "High spectral efficient W-band OFDM-RoF system with direct-detection by two cascaded single-drive MZMs," *Opt. Express*, vol. 21, no. 14, pp. 16615–16620, Jul. 2013.
[7] C. B. Huang, J. W. Shi, N. W. Chen, H. P. Chuang, J. E. Bowers, and C. L. Pan, "Remotely up-converted 20-Gbit/s error-free wireless on–off-keying data transmission at w-band using an ultra-wideband photonic transmitter-mixer," *IEEE J. Photon.*, vol. 3, no. 2, pp. 209–219, Mar. 2011.
[8] A. Kanno *et al.*, "40 Gb/s W-band (75–110 GHz) 16-QAM radio-over-fiber signal generation and its wireless transmission," *Opt. Express*, vol. 19, no. 26, pp. B56–B63, Sep. 2011.
[9] X. Li, Y. Xu, and J. Yu, "Single-sideband W-band photonic vector millimeter-wave signal generation by one single I/Q modulator," *Opt. Lett.*, vol. 41, no. 18, pp. 4162–4165, Sep. 2016.
[10] X. Li, Y. Xu, and J. Yu, "Over 100-Gb/s V-band single-carrier PDM-64QAM fiber-wireless-integration system," *IEEE J. Photon.*, vol. 8, no. 5, pp. 1–7, Oct. 2016.
[11] J. Zhang *et al.*, "Full-duplex asynchronous quasi-gapless carrier-aggregation using filter-bank multi-carrier in MMW radio-over-fiber heterogeneous mobile access networks," in *Proc. Opt. Fiber Commun. Conf.*, Anaheim, CA, USA, 2016, Paper TU2B. 2.
[12] N. Chi, M. Zhang, J. Shi, and Y. Zhao, "Spectrally efficient multi-band visible light communication system based on Nyquist PAM-8 modulation," *Photon. Res.*, vol. 5, no. 6, pp. 588–597, Dec. 2017.
[13] A. Kakkar *et al.*, "Low complexity timing recovery algorithm for PAM-8 in high speed direct detection short range links," in *Proc. Opt. Fiber Commun. Conf.*, Los Angeles, CA, USA, 2017, Paper W2A.54.
[14] X. Pang *et al.*, "100 Gbit/s hybrid optical fiber-wireless link in the W-band (75–110 GHz)," *Opt. Express*, vol. 19, no. 25, pp. 24944–24945, Dec. 2011.
[15] S. Jia *et al.*, "Integrated dual-DFB laser for 408 GHz carrier generation enabling 131 Gbit/s wireless transmission over 10.7 meters," in *Proc. Opt. Fiber Commun. Conf.*, San Diego, CA, USA, 2019, Paper Th1C.2.
[16] K. Liu *et al.*, "100 Gbit/s THz photonic wireless transmission in the 350-GHz band with extended reach," *IEEE Photon. Technol. Lett.*, vol. 30, no. 11, pp. 1064–1067, Jun. 2018.
[17] X. Li and J. Yu, "Over 100 Gb/s ultra-broadband MIMO wireless signal delivery system at the D-band," *IEEE J. Photon.*, vol. 8, no. 5, Dec. 2016, Art. no. 7906210.

[18] J. J. V. Olmos and I. T. Monroy, "Wireless communications surpassing fiber capacity: Micro-and millimeter-wave solutions up to D-band for 5G systems," in *Proc. 19th Int. Conf. Transparent Opt. Netw.*, Girona, Catalonia, Spain, 2017, pp. 1–4.

[19] X. Li, J. Yu, and G. K. Chang, "Photonics-assisted technologies for extreme broadband 5G wireless communications," *J. Lightw. Technol.*, vol. 37, no. 12, pp. 2851–2865, Jun. 2019.

[20] L. Zhang *et al.*, "Nonlinearity-aware 200 Gbit/s DMT transmission for C-band short-reach optical interconnects with a single packaged electro-absorption modulated laser," *Opt. Lett.*, vol. 43, no. 2, pp. 182–185, Jan. 2018.

[21] H. Yamazaki *et al.*, "Transmission of 400-gbps discrete multi-tone signal using >100-GHz-bandwidth analog multiplexer and InP Mach–Zehnder modulator," in *Proc. Eur. Conf. Opt. Commun.*, Rome, Italy, 2018, pp. 1–3.

[22] N. Stojanovic, C. Prodaniuc, L. Zhang, and J. Wei, "210/225 Gbit/s PAM-6 transmission with BER below KP4-FEC/EFEC and at least 14 dB link budge," in *Proc. Eur. Conf. Opt. Commun.*, Rome, Italy, 2018, pp. 1–3.

[23] L. Zhang *et al.*, "Kernel mapping for mitigating nonlinear impairments in optical short-reach communications," *Opt. Express*, vol. 27, no. 21, pp. 29567–29580, Oct. 2019.

[24] W. Zhou, X. Li, and J. Yu, "Pre-coding assisted generation of a frequency quadrupled optical vector D-band millimeter wave with one Mach–Zehnder modulator," *Opt. Express*, vol. 25, no. 22, pp. 26483–26491, Oct. 2017.

[25] X. Li, J. Yu, L. Zhao, K. Wang, W. Zhou, and J. Xiao, "1-Tb/s photonics-aided vector millimeter-wave signal wireless delivery at D-band," in *Proc. Opt. Fiber Commun. Conf.*, San Diego, CA, USA, 2018, Paper TH4D.1.

[26] R. Puerta, J. Yu, X. Li, Y. Xu, J. J. V. Olmos, and I. T. Monroy, "Single-carrier dual-polarization 328-Gb/s wireless transmission in a D-band millimeter wave 2 × 2 MU-MIMO radio-over-fiber system," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 587–593, Jan. 2018.

[27] T. Schneider, "Ultrahigh-bitrate wireless data communications via THz-links; Possibilities and challenges," *J. Infrared Millimeter Terahertz Waves*, vol. 36, pp. 159–179, Aug. 2014.

[28] J. Zhang *et al.*, "200 Gbit/s/λ PDM-PAM-4 PON system based on intensity modulation and coherent detection," *J. Opt. Commun. Netw.*, vol. 12, no. 1, pp. A1–A8, Jan. 2020.

[29] J. Zhang, J. S. Wey, J. Shi, and J. Yu, "Single-wavelength 100-Gb/s PAM-4 TDM-PON achieving over 32-dB power budget using simplified and phase insensitive coherent detection," in *Proc. Eur. Conf. Opt. Commun.*, Rome, Italy, 2018, pp. 1–3.

[30] C. Xie *et al.*, "Generation and transmission of 100-Gb/s PDM 4-PAM using directly modulated VCSELs and coherent detection," in *Proc. Optical Fiber Commun. Conf.*, San Francisco, CA, USA, 2014, Paper Th3K.2.

[31] K. M. Gharaibeh, *Nonlinear Distortion in Wireless Systems: Modeling and Simulation With Matlab*," Hoboken, NJ, USA: Wiley, 2011.

[32] J. Zhang *et al.*, "PAM-8 IM/DD transmission based on modified lookup table nonlinear predistortion," *IEEE J. Photon.*, vol. 10, no. 3, Jun. 2018, Art. no. 7903709.

[33] W. Zhou, J. Zhang, X. Han, M. Kong, and P. Gou, "PAM-4 delivery based on pre-distortion and CMMA equalization in a ROF system at 40 GHz," *Opt. Commun.*, vol. 418, pp. 61–65, Jan. 2018.

[34] J. Zhang, and J. Yu, "EML-based IM/DD 400G (4x112. 5-Gbit/s) PAM-4 over 80 km SSMF based on linear pre-equalization and nonlinear LUT pre-distortion for inter-DCI applications," in *Proc. Opt. Fiber Commun. Conf.*, Los Angeles, CA, USA, 2017, Paper W4I.4.

[35] I. Fijalkow, C. E. Manlove, and C. R. Johnson, "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 227–231, Jan. 1998.

[36] E. Giacoumidis *et al.*, "Fiber nonlinearity-induced penalty reduction in CO-OFDM by ANN-based nonlinear equalization," *Opt. Lett.*, vol. 40, no. 21, pp. 5113–5116, Nov. 2015.

[37] S. Liu *et al.*, "A multilevel artificial neural network nonlinear equalizer for millimeter-wave mobile fronthaul systems," *J. Lightw. Technol.*, vol. 35, no. 20, pp. 4406–4417, Oct. 2017.

[38] N. Chi, Y. Zhao, M. Shi, P. Zou, and X. Lu, "Gaussian kernel-aided deep neural network equalizer utilized in underwater PAM8 visible light communication system," *Opt. Express*, vol. 26, no. 20, pp. 26700–26712, Sep. 2018.

[39] W. Samek, A. Binder, G. Montavon, S. Lapuschkin, and K. Müller, "Evaluating the visualization of what a deep neural network has learned," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 11, pp. 2660–2673, Nov. 2017.

[40] B. Verónica, P. Diego, A. Amparo, G. Bertha, and S. Noelia, "Scalability analysis of ANN training algorithms with feature selection," in *Proc. Conf. Spanish Assoc. Artif. Intell.*, Berlin, Heidelberg, 2011, pp. 84–93.

[41] S. Liu, X. Wang, W. Zhang, G. Shen, and H. Tian, "An adaptive activated ANN equalizer applied in millimeter-wave RoF transmission system," *IEEE Photon. Technol. Lett.*, vol. 29, no. 22, pp. 1935–1938, Nov. 2017.

[42] S. Liu, Y. M. Alfadhli, S. Shen, M. Xu, H. Tian, and G. Chang, "A novel ANN equalizer to mitigate nonlinear interference in analog-RoF mobile fronthaul," *IEEE Photon. Technol. Lett.*, vol. 30, no. 19, pp. 1675–1678, Oct. 2018.

[43] L. Zhang, "Blind equalization in neural networks: Theory," *Algorithms and Applications[M]*. Berlin, Germany: Walter de Gruyter GmbH & Co KG, 2017.

[44] J. Han and C. Moraga, "The influence of the sigmoid function parameters on the speed of back propagation learning," in *Proc. Int. Workshop Artif. Neural Netw.*, Berlin, Heidelberg, 1995, pp. 195–201.

[45] J. C. Patra, R. N. Pal, R. Baliarsingh, and G. Panda, "Nonlinear channel equalization for QAM signal constellation using artificial neural networks," *IEEE Trans. Syst., Man, Cybern., Part B*, vol. 29, no. 2, pp. 262–271, Apr. 1999.

[46] R. P. Braun, G. Grosskopf, D. Rohde, and F. Schmidt, "Low-phase-noise millimeter-wave generation at 64 GHz and data transmission using optical sideband injection locking," *IEEE Photon. Technol. Lett.*, vol. 10, no. 5, pp. 728–730, May 1998.

[47] A. C. Bordonalli, C. Walton, and A. J. Seeds, "High-performance phase locking of wide linewidth semiconductor lasers by combined use of optical injection locking and optical phase-lock loop," *J. Lightw. Technol.*, vol. 17, no. 2, pp. 328–342, Feb.1999.

[48] A. G. Juárez, I. E. Z. Huerta, J. R. Asomoza, and M. D. R. G. Colín, "Photonic Components for Analog Fiber Links," in *Opt. Commun.*, London, U.K.: IntechOpen, 2012, pp. 132–138.

[49] S. Koenig *et al.*, "100 Gbit/s wireless link with mm-wave photonics," in *Proc. Opt. Fiber Commun. Conf. Expo. Nat. Fiber Opt. Engineers Conf.*, Anaheim, CA, USA, 2013, pp. 1–3.