Classification Weighted Deep Neural Network Based Channel Equalization for Massive MIMO-OFDM Systems

Lijun GE [1], Changcheng QI [1], Yuchuan GUO [1], Lei QIAN [1], Jun TONG [2], Peng WEI [3]

[1]School of Electronics and Information Engineering, Tianjin Key Laboratory of Optoelectronic Detection Technology and System, Tiangong University, Tianjin, 300160, China
[2] Electrial, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, 2522, Australia
[3]Department of Electronic Engineering, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing, 100084, China

gelj_001@hotmail.com, 18322285826@163.com, 1310433616@qq.com,
qianlei@tiangong.edu.cn, eejtong@gmail.com, wpwwwhttp@163.com

Submitted March 12, 2022 / Accepted July 9, 2022 / Online first August 3, 2022

Abstract. *Massive multi-input multi-output (MIMO) has attracted significant interest in academia and industry, which can efficiently increase the transmission rate. However, the error rate of conventional channel equalizations in massive MIMO systems may be high owing to the dynamic channel states in practical conditions. To solve this problem, in this paper, we propose an improved channel equalization framework based on the deep neural network (DNN). Based on the analyzed relationship between the input and output of the DNN, the data can be recovered without the channel state information. Furthermore, aiming at reducing the convergence time and enhancing the learning ability of the DNN, a classification weighted algorithm is proposed to optimize the cost function of the DNN, which is named as classification weighted deep neural network (CW-DNN). Simulation results demonstrate that compared to conventional counterparts, the proposed CW-DNN based equalizer can achieve a better normalized mean square error (NMSE). Upon approximating the optimal neural network parameters with the significantly improved convergence speed and reduced training time of the network, under the condition of the fixed learning rate.*

Keywords

Channel equalization, classification weighted, deep neural network, massive MIMO, optimization algorithm

1. Introduction

Massive multi-input multi-output (MIMO) systems with tens or hundreds of antennas are one of the most promising transmission techniques in wireless communication systems, and they can considerably enhance spatial multiplexing gain and system capacity [1], [2]. However, due to the dynamic characteristics of the channel state information (CSI) and the user scheduling, limited radio frequency (RF) resources, and the pilot contamination [2–4], efficient channel equalization is required in wireless communication systems.

Channel equalization strategies for massive MIMO orthogonal frequency division multiplexing (OFDM) systems have been extensively researched in [5–11], with linear and nonlinear methods being distinguished. The former contains equalizers with zero forcing (ZF) and minimum mean square error (MMSE) ,while the latter includes the maximum likelihood (ML) and lattice reduction-aided (LRA) equalizers. Compared to the nonlinear methods, the linear equalization methods have low complexity at the cost of degraded error performance. Therefore, significant efforts are dedicated to achieving a reasonable compromise between error performance and complexity. To improve the bit error rate (BER) performance for multi-user (MU) MIMO-OFDM systems, linear equalizers based on the least mean square (LMS) and recursive least square (RLS) algorithms were proposed by the authors of [5]. Besides, low-complexity algorithms for implementing the ML detector are considered in [6] and [7].

Full noun	Abbreviation
channel state information	CSI
multi-input multi-output	MIMO
normalized mean square eerror	NMSE
fast fourier transform	FFT
artificial neural network	ANN
multilayer perceptron	MLP
radial basis function	RBF
back propagation neural network	BPNN
deep neural network	DNN
uniform linear array	ULA
base station	BS
orthogonal frequency division multiplexing	OFDM
least mean square	LMS
recursive least square	RLS

Tab. 1. Abbreviations used in this article.

DOI: 10.13164/re.2022.0346

However, for massive MIMO systems, since the number of antennas is tens or hundreds times of that in MIMO systems, the above algorithms still have high computation. To improve the efficiency of channel equalization for massive MIMO systems, an adaptive equalizer combined with a channel tracking method is proposed for the multi carrier-code division multiple access (MC-CDMA) system over the rapidly fading channel under colored noise in [8]. Furthermore, to enhance the stability of the adaptive channel equalizer, using adaptive generalized decision feedback equalization and ordered consecutive interference cancellation, the authors of [9] suggested an adaptive MIMO channel equalizer. In general, the above methods are classified as the model-driven channel equalization [10], which may suffer from severe error performance degradation in practical communication scenarios with unknown essential parameters, such as CSI, bandwidth, and pilot information.

To break the model-driven frame, the artificial neural network (ANN) has been paid significant attention for channel equalization, owing to its strong capability of nonlinear mapping and the characteristic learning in classification or recognition [11]. In the ANN, multiple training sets of data are introduced to optimize the weights and bias until a stable network state [12]. In the last decade, various neural network models have been developed for machine learning-aided communication systems. The conventional multilayer perceptron (MLP) has been used to mitigate the inter-channel interference in OFDM systems in [13]. A conventional radial basis function (RBF) based neural network is developed in [14–16] to enhance channel estimation in pilot-aided OFDM systems. Although the RBF-based neural network has a stronger learning ability than the MLP based neural network, the radial base center in RBF needs to be determined artificially. Furthermore, the back propagation neural network (BPNN) is adopted in the space-time coded MIMO-OFDM systems for effectively estimating the channel correlation coefficients in [17]. Generally, the ANN-based channel equalization has good performance for systems with a few antennas. However,with the increase of the number of antennas, its performance decreases obviously. Therefore, the ANN-based channel equalization method is not suitable for massive MIMO systems.

The deep neural network (DNN) with more layers and more neurons is significantly studied based on deep learning (DL) to obtain better feature extraction and generalization from a large amount of data [2, 18–20]. The authors of [2] proposed a novel channel prediction framework that integrates the imperfect channel estimation of the massive MIMO-OFDM into the DNN scheme and proves that DNN is an effective method. The work of this article is to perform channel estimation, which uses DNN to directly restore data without channel state information. Compared with [2], the number of layers of the DNN used, the number of neurons and the data processing method are very different. To solve the problems of channel equalization and signal detection in wireless communication systems, the DNN is leveraged in [21–25]. The authors of [21] proposed a nonlinear signal detector based on a DNN for the OFDM system, which achieves a good error performance with high complexity. Then, to reduce the complexity, a DNN based MIMO detector was designed by the gradient descent method in [23]. Furthermore, to obtain the error performance close to the ML detection, the authors of [24] proposed a dense-layer DNN for a single-path MIMO system. Aiming at improving the high error performance in a MIMO system, the cross-entropy modelled loss function of the DNN was optimized for the softly detected MIMO signal [26]. It shows that the selection of cost function has an important effect on DNN training. However, current signal detection methods cannot be easily adapted to DNN-based signal detection methods. In addition, there are also many researches focusing on the combination of DNN and current signal detection methods to further improve the error performance. For example, the iterative signal detection was projected into the DNN structure and realized by the DL method in [27–29] to improve the MIMO detection performance and robustness. Besides, an MU-SIMO channel detection scheme was proposed in [30], which combines the advantages of the feed-forward DNN and parallel interference cancellation to effectively eliminate the co-channel interference. Additionally, according to the idea of Residual Network (ResNet), a deeper DNN model was built to improve the performance of the channel estimation and equalization in [31]. Channel equalization first needs to obtain channel state information through channel estimation. The proposed CLMMSE algorithm [1] calculates the channel autocorrelation matrix by investigating the channel prior information based on compressive sensing (CS) theory. The channel state information is estimated by the algorithm, but it brings high complexity.

Motivated by this problem, an optimized channel equalization method is proposed based on the classification weighted DNN (CW-DNN). Furthermore, the performance of the proposed method is verified by using a signal detection of the minimum distance-based symbolic slicer. We formulate the equalization of the massive MIMO channel as a classification problem using the DNN. In this paper, to solve the above problems, an optimized channel equalization method is proposed based on the classification weighted DNN (CW-DNN). The main contributions are summarized as follows.

- In a typical massive MIMO-OFDM system, the input signal is classified into different categories according to the one-hot mapping for the constellation. We use a DNN to directly recover the data without channel state information and then recover the corresponding classes of transmitted symbols from the received signal.
- Furthermore, in order to improve the performance of channel equalization, we specially design a new loss function. The classification weighting is adopted to optimize the cost function of the DNN. Without adjusting the learning rate, the proposed CW-DNN is capable of

enhancing the training speed of the DNN, while preventing the solutions from falling into the saddle point or the local minimum point, so as to obtain the optimal neural network parameters.

- Finally, simulation results show that the proposed method using the classification weighted optimization of the cost function is an effective alternative compared to the methods using the classical cost function.

Notation and symbols: Abbreviations used in this article in Tab. 1. Notation and symbols used in this article are shown in Tab. 2.

The following is a summary of the remainder of this paper: Massive MIMO-OFDM system is presented in Sec. 2. Deep learning and optimization algorithm-based channel equalizations are analyzed in Sec. 3. Simulation results are discussed in Sec. 4, and a brief conclusion is summarized in Sec. 5.

Symbol	Definition/Explanation
N_t	Number of transmit antennas
N_r	Number of receive antennas
$(\cdot)^\mathrm{T}$	Transpose operator
$(\cdot)^\mathrm{H}$	Hermitian operator
$\|\cdot\|$	Frobenius norm operator
$\mathbb{E}[\cdot]$	Expectation operator
$(\cdot)^{-1}$	Inverse operator
$\otimes$	Kronecker product
Vec	Vectorization operator

Tab. 2. Notation and symbols used in this article

2. Massive MIMO-OFDM System

As shown in Fig. 1, the transmitter sends an information bit stream. The number of sub-carriers is N, after serial-to-parallel conversion, it becomes a group of N parallel data. And then IFFT is performed, cyclic prefix is added, after quadrature amplitude modulation (16/32QAM) a set of parallel data after constellation mapping, modulate N subcarriers to obtain OFDM symbols. Then mix white Gaussian noise to the receiving end for demodulation. The model of a massive MIMO-OFDM system with N_t transmit antennas and N_r receive antennas. The frequency-domain symbols $X_i(k)$ on the subcarrier k at the i-th transmit antenna are modulated on N subcarriers with $k = 0, 1, \dots, N-1$. Thus, the time-domain signal at the i-th transmit antenna in time n is represented as

$$x_i(n) = \sum_{k=0}^{N-1} X_i(k)\mathrm{e}^{\mathrm{j}(2\pi kn/N)}, n = 0, 1, \dots, N-1. \quad (1)$$

A uniform linear array (ULA) of N_t antennas is assumed to be installed on base station (BS). The Rayleigh channel fading model can be given as

$$h(t) = \sum_{L} a_L \exp\left[\mathrm{j}\left(\phi_L + 2\pi f \cos\left(\beta_L\right) t\right)\right] \quad (2)$$

where a_L is the amplitude of the L-th propagation path from the BS to the receivers. ϕ_L and β_L are, respectively, the angle of arrival and random phase of the L-th path.

The data stream at each transmit antenna can be expressed as $x(N_i)$. At the receiving end, the data stream of the massive MIMO-OFDM system can be expressed as

$$y_j(n) = \sum_{N_i=1}^{N_\mathrm{t}} \sum_{l=0}^{L-1} h_{jN_i}(n,l) x_{N_i}(n-l) + w_j(n). \quad (3)$$

Converting the signals from time domain to frequency domain by fast Fourier transform (FFT) operations, and then the obtained signals can be expressed as

$$Y_j(k) = \sum_{N_i=1}^{N_\mathrm{t}} \sum_{l=0}^{L-1} H_j N_i(n,l) X_{N_i}(n-l) + W_j(n). \quad (4)$$

We assume that in a time-invariant MIMO channel, the channel impulse response remains constant over the coherence time. Therefore, the received frequency-domain signal $\mathbf{Y}(k) = [Y_1(k), Y_2(k), \dots, Y_{N_\mathrm{r}}(k)]^\mathrm{T}$ at the k-th subcarrier is expressed as

$$\mathbf{Y}(k) = \mathbf{H}(k)\mathbf{X}(k) + \mathbf{W}(k) \quad (5)$$

where $\mathbf{W}(k)$ is the additive white Gaussian noise (AWGN).

In most scenarios, $\mathbf{H}(k)$ is often an underdetermined factor and $\mathbf{X}(k)$ can be set as certain sequence which is known at both transmitting and receiving sides. Thus, when apply channel estimation algorithms to get the channel state information, (5) can be rewritten as

$$\mathbf{Y}(k) = \left(\mathbf{X}(k)^\mathrm{T} \otimes I\right) \mathrm{Vec}(\mathbf{H}(k)) + \mathbf{W}(k). \quad (6)$$

Here, $\otimes$ represents Kronecker product, Vec is the vectorization operation which stacks the column vector by column of the matrix.

In (6), we assume that the channel gain $\mathbf{H}(k)$ and the constellation of the data $\mathbf{X}(k)$ are unknown. Thus, the channel equalization aims at recovering $\mathbf{X}(k)$ from the received signal $\mathbf{Y}(k)$ at the receiver. Additionally, convert this system equation into a matrix form, the matrix form of (4) can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (7)$$

with the transmit data vector $\mathbf{x}$, the massive MIMO channel impulse matrix $\mathbf{H}$, the additive white Gaussian noise (AWGN) $\mathbf{n}$, and the received signal vector $\mathbf{y}$.

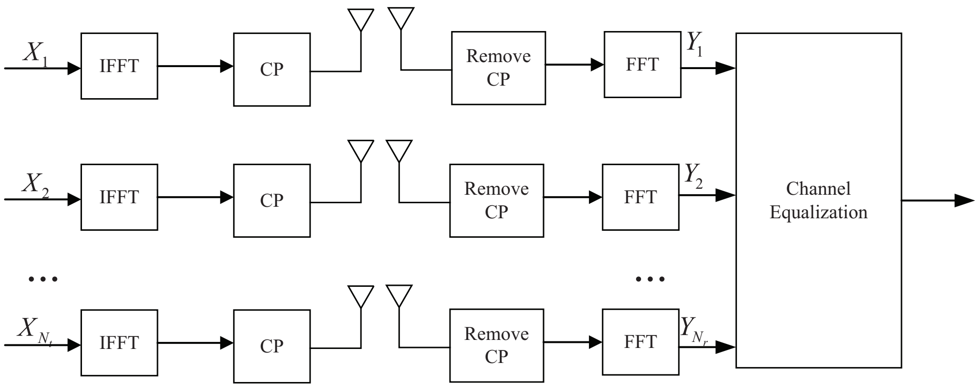

Fig. 1. The block diagram of the massive MIMO-OFDM system.

3. Deep Learning and Optimization Algorithm Based Channel Equalizations

The massive MIMO channel equalization based on DNN can be formulated as a classification problem, where the transmitted signal is classified into different groups according to the constellation. The DNN based equalizer solves the classification problem to recover the transmitted signal. The pilot symbols are used to train the DNN, and the variation of the pilot value is used to reflect the variation of the channel.

3.1 Algorithmically for Deep Learning

DNN has the same structure as a traditional neural network, including hidden layers and neurons. Furthermore, its deep design facilitates the establishment of models, particularly in highly nonlinear contexts where processing is challenging. DNN learning is divided into two phases: training and testing. The network model is initially trained using the gradient descent approach to minimize the error between the output value and the real value before implementing the effective channel parameter estimation. Afterwards when, using partial derivatives of the cost function, the network weights and biases are maintained in real time.

The proposed DNN-based equalizer uses a fully connected DNN with L layers, comprising one input layer, $(L-2)$ hidden layers, and one output layer, as illustrated in Fig. 2. Let w_{ij}^l represent the weight from the i-th neuron of the $(l-1)$-th layer to the j-th neuron to the l-th layer, and b_i^l represent the bias unit. Thus, the layer pre-activation is delivered by

$$z_i^l = \sum_j w_{ij}^l a_j^{l-1} + b_i^l. \tag{8}$$

The activation of neuron output may then be rewritten as

$$a_i^l = f(z_i^l) = f\left(\sum_j w_{ij}^l a_j^{l-1} + b_i^l\right). \tag{9}$$

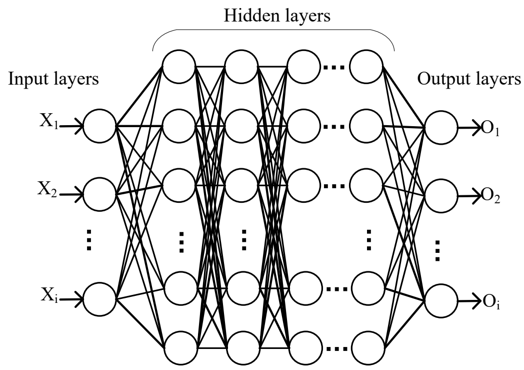

Fig. 2. The illustration of the DNN model.

3.2 DNN Based Channel Equalization Framework

This subsection introduces the massive MIMO channel equalization based on the DNN, where the non-sparse Rayleigh channel is considered. As shown in Fig. 3, different from the CSI requirement in conventional equalizer, the DNN equalizer removes the channel estimation and directly recovers the transmitted data.

In the DNN model, for the transmitted data $\mathbf{X} = \left[X_1, \ldots, X_{N_t}\right]^{\mathrm{T}}$ and the received data $\mathbf{Y} = \left[Y_1, \ldots, Y_{N_r}\right]^{\mathrm{T}}$, we can rewrite them as

$$\begin{cases} \mathbf{X} \triangleq \begin{bmatrix} \Re(\tilde{\mathbf{X}}) \\ \Im(\tilde{\mathbf{X}}) \end{bmatrix} \\ \mathbf{Y} \triangleq \begin{bmatrix} \Re(\tilde{\mathbf{Y}}) \\ \Im(\tilde{\mathbf{Y}}) \end{bmatrix} \end{cases} \tag{10}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts, respectively, therefore, $\mathbf{X} \in R^{2N_t}$, and $\mathbf{Y} \in R^{2N_r}$. In this paper, we adopt the one-hot mapping to Re-parameterize the discrete constellations $\mathbf{S} = \{s_1, s_2, \ldots, s_M\}$, where M denotes digital modulation order. Corresponding to each s_m for $m = 1, \ldots, M$, we define a unit vector $\mathbf{u}_i \in R^M$. The four-dimensional one-hot mapping of the actual component of 16-QAM constellations, is depicted as

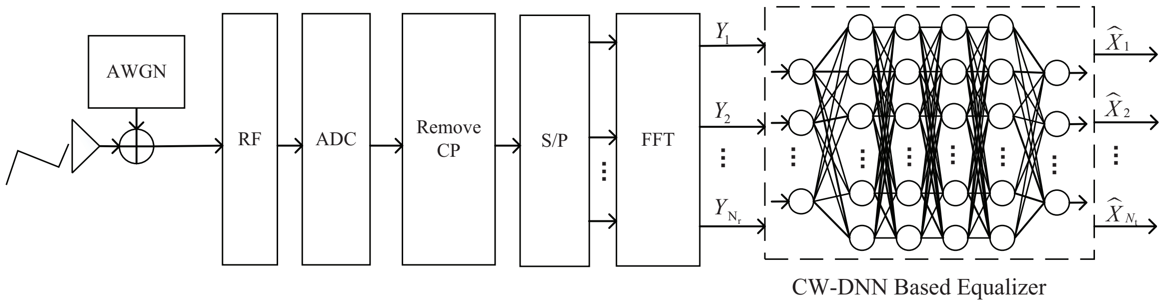

Fig. 3. Block diagram of the CW-DNN based receiver in the massive MIMO system.

$$\begin{aligned} \mathbf{S}_1 &= -3 \leftrightarrow \mathbf{u}_1 = [1, 0, 0, 0], \\ \mathbf{S}_2 &= -1 \leftrightarrow \mathbf{u}_2 = [0, 1, 0, 0], \\ \mathbf{S}_3 &= \ \ \ 1 \leftrightarrow \mathbf{u}_3 = [0, 0, 1, 0], \\ \mathbf{S}_4 &= \ \ \ 3 \leftrightarrow \mathbf{u}_4 = [0, 0, 0, 1]. \end{aligned} \tag{11}$$

Then, we define this mapping as the function $s = f_{\text{oh}}(\mathbf{u})$, so that $s_i = f_{\text{oh}}(\mathbf{u}_i)$ for $i = 1, \ldots, M$. Especially, when the inputs are not unit vectors, the function is defined as

$$x = f_{\text{oh}}(x_{\text{oh}}) = \sum_{i=1}^{M} s_i [x_{\text{oh}}]_i . \tag{12}$$

The MIMO model involves a vector of $2N_t$ symbols which is handled by stacking the one-hot mapping of each elements. In other words, a vector $x_{\text{oh}} \in \{0, 1\}^{2N_t M}$ is mapped to $x \in M^{2N_t}$.

In the training samples, the training set, the test set, and the validation set occupy 70%, 15%, and 15%, respectively. The input of the DNN based equalizer is $\mathbf{Y}_{\text{int}} = [y_{1,1}, \ldots, y_{k,n_r}]^{\text{T}}$, where y_{k,n_r} represents the data on the k-th subcarrier at the n_r-th transmit antenna. The output of the DNN based equalizer is $\widehat{\mathbf{X}}_{\text{out}} = [\widehat{x}_{1,1}, \ldots, \widehat{x}_{k,n_t}]^{\text{T}}$, where $\widehat{x}_{k,n_t}$ represents the estimated data on the k-th subcarrier from the n_r-th transmit antenna.

First, the DNN is trained with the known Y and X to minimize the error between the input and output data. After the DNN based equalizer is trained, this paper uses the gradient descent algorithm to update the weight values and the hidden layer bais values. The symbolic slicer is a signal detection method based on minimum distance, divides the network output data through the DNN based equalizer into the correct class, and then restore the original signal.

Let the output signal vector of the j-th estimation be $\hat{y}_j$ and the transmit signal vector of the j-th sample be x_j, the cost function of (20) is expressed as

$$E = \frac{1}{2} \sum_{j=1}^{P} \left(x_j - \hat{y}_j \right)^2 \tag{13}$$

where P denotes the total number of samples.

The goal of the proposed DNN based estimation model is computing and updating the partial derivatives of weights and bias in DNN. In this paper, mini-Batch Gradient Descent (MBGD) is used, taking N $(1 \leq N \leq P)$ samples in each training batch, the cost function can be approximated as

$$E = \frac{1}{N} \sum_{j=1}^{N} \left(x_j - \hat{y}_j \right)^2 . \tag{14}$$

The DNN updates the parameters from the output of each mini-batch, which can be expressed as

$$W^l = W^l - u \sum_{i=1}^{N} \frac{\partial E_i}{\partial W^l}, \tag{15}$$

$$b^l = b^l - u \sum_{i=1}^{N} \frac{\partial E_i}{\partial b^l} \tag{16}$$

where u denotes the learning rate of the DNN.

The corresponding input and output training data of the DNN based channel equalization is indexed by received data (include data symbols and pilot symbols) at the receiver and one-hot mapping of the original constellation symbols at the BS, respectively. The pilot symbols are uniformly distributed on the subcarrier of each OFDM symbol, and the pilot positions are fixed. The input of the DNN includes the received data symbols and the received pilot symbols. Since the change of the received pilot symbols value reflects the time-varying characteristics of the channel, the implicit estimation of the channel can be taken.

3.3 CW-DNN Based Channel Equalization and Signal Dection

Due to the diversity and complexity of the antenna array, the DNN based equalizer requires the longer processing time. Furthermore, the convergence rate during the initial training process is slow. During the training of the DNN, it is difficult to find an appropriate learning rate, and the parameters of the weight and bias always approach to a local optimum instead of a global optimum. To solve the problem, this section analyzes the relationship between the input pattern sent by the known pilot and the network output. Then, aiming at reducing the convergence time and enhacing the learning ability of the neural network, a classification weighted algorithm is proposed by optimizing the cost function of the DNN, referred to as CW-DNN. The CW-DNN based equalizer is shown in Fig. 3. The new cost function by using classification weighted is proposed in this part. Finally, at a fixed learning rate, the original gradient is increased or decreased by changing the cost function to obtain better neural network parameters.

In multipath channels, the transmitted signal reaches the receiver through different paths, which may enhance or attenuate. When the signal bandwidth is larger than the coherent bandwidth of the channel, the signal will be severely distorted. On the contrary, the distortion of the signal is very small. The purpose of equalization is to eliminate the distortion caused by the channel, and the elimination of inter-symbol interference is the result of equalization. The received signal in the frequency domain is the product of the transmitted signal and the frequency domain response of the channel. The frequency domain response of the channel can be directly obtained by the pilot inserted in each symbol, so that the equalization of the multi-carrier signal can be completed by a simple single-point equalizer.

In this paper, DNN-based channel equalization is frequency-domain equalization. x_j and $\hat{y}_j$ are mapped to -1 and 1. The scaling factor c is set by judging the positive or negative of the product of the transmitted data and the received data, and the cost function is improved to ensure the gradient of the cost function. It must be large enough and predictable enough to provide good guidance for the learning algorithm, maintain convergence and continue to search for the global optimum.The neural network can increase the training speed of the DNN, while preventing the solution from falling into a saddle point or a local minimum, so as to obtain the optimal neural network parameters. Then use the symbol limiter to restore the equalized signal to the corresponding constellation point.

The output layer of the CW-DNN uses the tanh function instead of the traditional linear function. The output of the network is a one-dimensional real array distributed in $[-1, 1]$. The output sample $\hat{y}_j$ need to be transformed to $\hat{y}_j/c$, where $\hat{y}_j$ denotes the target data that the network tries to approximate. c ($c > 0$) is a scaling constant to make the range of all target data match the tanh activation function of the output layer. According to the characteristics of the pilot symbols combined with the network output, when the output value is positive, it is regarded as a mapping of 1. On the contrary, when the output data is negative, the approximated value is mapped to -1.

If the digital modulation at the pilot frequency is BPSK, the input value of the CW-DNN will be a sequence of ± 1. CW-DNN determines whether the received data is same as the original data by judging the $x_j \hat{y}_j$. The CW-DNN uses the characteristics of the symbolic function $\operatorname{sign}(x)$: if $x \geq 0$, returns 1; if $x < 0$, returns -1. Set two adjustable weight values c_1 and c_2 respectively. When $x_j \hat{y}_j \geq 0$, it indicates that the estimated data symbol is the same as the real sample symbol. For example, if the transmitted data is 1, the received data is also 1, or the transmitted data is -1, and the received data is also -1. At this time $\operatorname{sign}(x_j \hat{y}_j) = 1$, setting c_2 is relatively small, and the convergence can be maintained when the known pilot symbols and the output data through the CW-DNN are the same, and continue to search for the global best. When $x_j \hat{y}_j < 0$, it indicates the symbols are different. At this time $\operatorname{sign}(x_j \hat{y}_j) = -1$, setting c_1 is relatively large, the penalty can be increased when the known pilot symbols and the output data through the CW-DNN are different.

The CW-DNN based equalization can improve the convergence speed in the early stage of network training and continue to converge in the later stage of training, greatly reducing the training time of network parameters and improving learning efficiency. The original cost function of the DNN changes to (18) of the CW-DNN, the gradient descent algorithm is used to obtain partial derivatives of weights and offset values, the formula can be written as

$$E^*(x_j, \hat{y}_j) = \frac{1}{N} \sum_{j=1}^{N} \left[c_j (x_j - \hat{y}_j)^2 \right], \tag{17}$$

$$c_j = \begin{cases} c_1, & \operatorname{sign}(x_j \hat{y}_j) = -1, \\ c_2, & \operatorname{sign}(x_j \hat{y}_j) = 1. \end{cases} \tag{18}$$

Then use the symbol limiter to restore the equalized signal to the corresponding constellation point. Minimum-distance based signal detection is used to find the symbol having the minimum Euclidean distance between the constellation alphabet and the equalized symbol $\tilde{Y}_j$. The corresponding symbol is taken as the hard detection result $\tilde{X}_j$, which can be expressed as

$$\tilde{X}_j = \arg\min_{X \in A} \left| \tilde{Y}_j - X \right| \tag{19}$$

where A is the constellation set, such as the 4-QAM modulation.

In order to evaluate the performance of the proposed CW-DNN based channel equalization scheme, the normalized mean square error (NMSE) performance is used, which is defined as

$$\mathrm{NMSE} = E \left\{ \sum_{k=1}^{N} \frac{\|x(k) - \hat{x}(k)\|_2^2}{\sum_{k=1}^{N} \|x(k)\|_2^2} \right\} \tag{20}$$

where N is the total number of the transmit data bits, $x(k)$ and $\hat{x}(k)$ are the transmit signal and the recovery signal, respectively.

The proposed channel equalization method based on CW-DNN collects pilot position data through the massive MIMO channel and adopts offline model training to provide the better NMSE performance. Besides, the proposed CW-DNN algorithm contains more hidden layers and the number of neurons, which can adjust the cost function adaptively, so as to learn channel data information better. Via combining the characteristics of the pilot signal and the output data, the DNN based on the optimization of the classification weighting method is used. The specific process of the algorithm is shown in Apendix A.

4. Simulation Results

In this section, the performance of the proposed CW-DNN based channel equalization algorithm is simulated by the Monte Carlo method. For comparison, the ZF and BPNN based equalization algorithms are also simulated. System parameters are listed in Tab. 3.

The channel adopts a Rayleigh distribution channel. The transmitted signal is multiplied by the Rayleigh fading coefficient, and then a Gaussian white noise signal with a different signal-to-noise ratio power is added to the signal. The multipath amplitude follows the Rayleigh distribution and the multipath follows a random distribution. Pilot insertion method is comb pilot insertion. In the process of channel data generation, the original channel data is updated by multiplying time expansion and multipath fading to obtain the original channel data at the current moment. The transmission and reception antenna expansion is used to convert the original channel data into a channel matrix. Then, the channel matrices on the path are combined to form a channel matrix on the distinguishable delay path. The simulation data is generated in MATLAB R2019B. Trainings are performed using TensorFlow 2.0, Keras and Pyhton. The experiments are performed on a computer with Intel Core i7 CPU (1.5 GHz).

The CW-DNN pilot insertion rate of the CW-DNN channel equalization is set to 20% and the training rate is set to 0.001. In addition, 500000 data sample sets are randomly allocated through the massive MIMO channel, 70% of which are considered as the training set, 15% of which are considered as the test set, and 15% of which are as the verification set for training the neural network. To optimize the cost function, the weighted factor c_1 and c_2 are defined as 2 and 0.5, respectively. In addition, the entire CW-DNN is composed of 5 layers, including an input layer, 3 hidden layers and an output layer, where the hidden layer activation function is defined as the ReLU function, and the output layer activation function is the tanh function.

Parameters	Values
Channel type	AWGN, multi-path fading channels
Number of transmit antennas	64
Number of receive antennas	64
Number of subcarriers	256
FFT size	256
Modulation	16, 32QAM
Length of CP	64
Subcarrier spacing	15 KHz

Tab. 3. Simulation parameters of the massive MIMO-OFDM system.

The performance of BER under various modulation schemes is compared in Fig. 4. As predicted, the BER curve in Fig. 4 continues to have a clear declining trend as the SNR increases. In the simulated massive MIMO system, there is no FEC under the Rayleigh channel and multipath fading channels. The data transmission is directly resumed without channel state information obtained. As shown in figure 4, the system BER performance is affected.

In Fig. 5, with various SNR levels, the relationship between the number of hidden layer neurons and BER performance in the massive MIMO-OFDM system is given. The SNR levels have been set at 15 dB, 20 dB, and 25 dB, respectively. The number of neurons in the hidden layer grows by 150 in each step, starting with 100.

It can be seen in Fig. 5, the BER performance is not directly proportional to the number of neurons in the hidden layers. When the number of neurons in the hidden layers increases to 850, the BER performance improvement becomes small. While, the complexity of the algorithm also increases. Furthermore, when the number of neurons in the hidden layer increases to 1150, the BER performance is degraded, due to the phenomenon of "Overfitting".

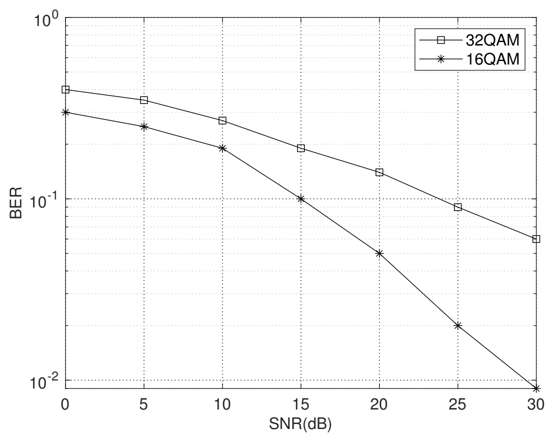

Fig. 4. BER performance of the CW-DNN under different modulation methods.

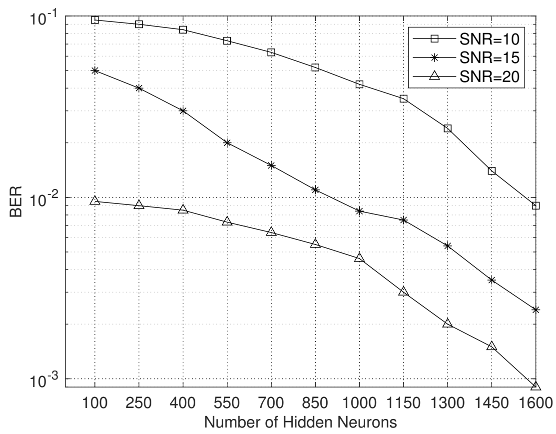

Fig. 5. BER performance versus the number of hidden neurons with varying transmitted SNRs.

Figure 6 provides the NMSE comparison of the proposed CW-DNN based scheme with batch size = 10, 30, and 50. It can be observed from Fig. 6 that as the SNR increases, the result of the NMSE value decreases. Furthermore, the NMSE performance is severely degraded by the small batch size. For example, when SNR = 25 dB, the NMSEs are 0.5×10^{-3}, 1.1×10^{-3}, 1.5×10^{-3} and 2.5×10^{-3} for batch size = 10, 30, 50, and 100, respectively. Besides, combined with the results in Fig. 7, it can be inferred that the small batch size can effectively alleviate the problem of the gradient dispersion.

Figure 7 shows the NMSE of the DNN using the MSE cost function and the proposed classification weighted optimized cost function with the varying number of the training iterations. It can be seen from Fig. 7 that the proposed system using the CW-DNN converges faster at the initial stage of training, compared to conventional DNN. As the number of the training iterations increases, the NMSE of the proposed method will converge to a fixed value.

We compared the DNN structure used in the literature [32]. Figure 8 and Figure 9 are the NMSE performance of the conventional ZF, BPNN, DNN [32], and the proposed CW-DNN based channel equalizations, with 16QAM and 32QAM, respectively. The number of the DNN training iterations is 1000, the number of hidden layer neurons is 850, and the batch size is 10. It can be seen from the Fig. 8 and Fig. 9 that with the increased value of the SNR, the NMSE performance of all equalization algorithms will be improved, where the performance of the proposed CW-DNN method is the best.

Figure 10 shows the effect of different numbers of channel paths on BER performance. We used 16QAM modulation. It can be seen from Fig. 10 that as the number of channel paths increases, the BER performance does not improve significantly. Thus, the performance of the proposed method is independent of the number of paths in the multipath channel.

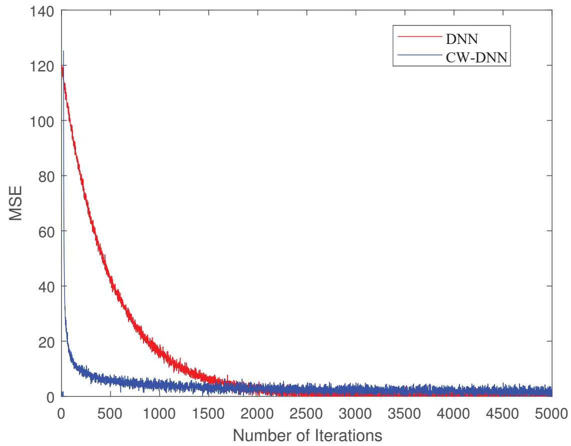

Fig. 7. Performance of a DNN-based equalizer in terms of convergence under various cost functions.

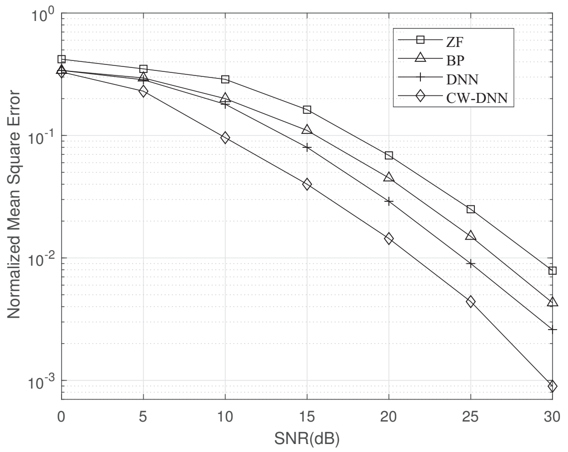

Fig. 8. NMSE performance when digital modulation is 16 QAM under different equalizers.

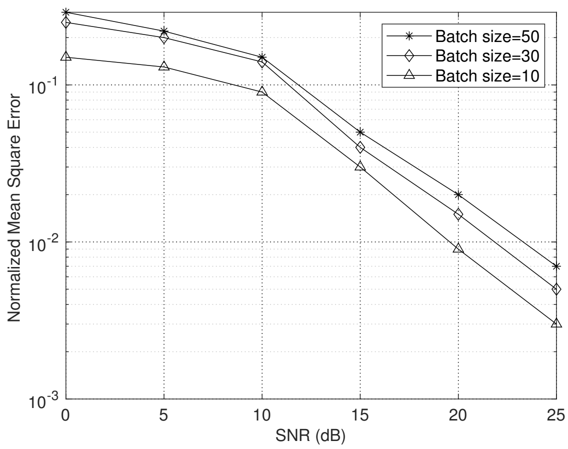

Fig. 6. The CW-DNN based scheme's NMSE performance with batch sizes of 10, 30, and 50.

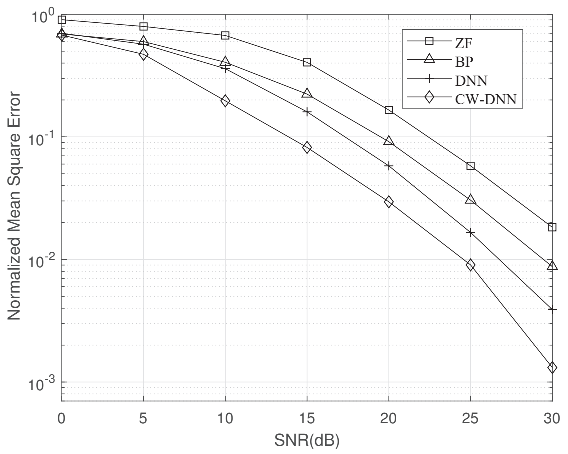

Fig. 9. NMSE performance when digital modulation is 32 QAM under different equalizers.

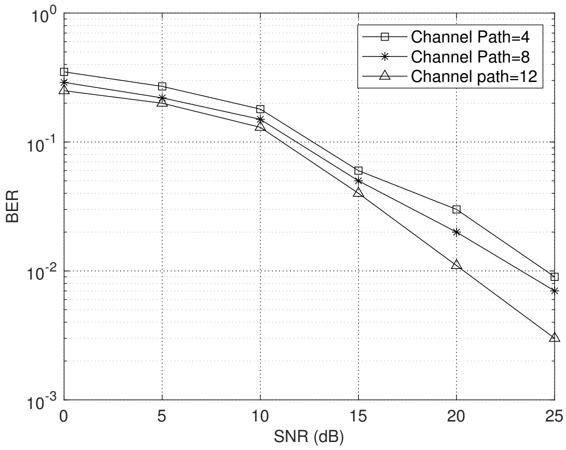

Fig. 10. The BER performance of the CW-DNN based channel equalization in different channel paths.

Operation	Complexity
MMSE	$O(N_r^3)$
ML	$O(MN_t)$
CW-DNN	$O(3MN_t)$

Tab. 4. Computational complexity of schemes.

We compare the computation complexities of the proposed DNN-based channel Equalizations scheme and the conventional channel Equalizations scheme in Tab. 4. Assuming that the modulation method used in communication is M-QAM, the modulated signal reaches the receiving end through the channel, and the constellation diagram of the modulated signal is divided into M types. Note that, in our proposed scheme, we train a neural network offline using channel simulation values generated from specific channel scenarios. In the case of offline learning, computational complexity is a lesser concern, as the time required is usually not strictly limited. The computational complexity of the scheme in the test phase includes three fully connected operations.

5. Conclusions

In this paper, we propose a DNN based improved channel equalization framework, which is formulated as a classification problem. Aiming at reducing the convergence time and enhacing the learning ability of the neural network, we also propose a classification weighted algorithm to optimize the cost function of the DNN. Simulation results demonstrate that under the condition of the fixed learning rate, the proposed CW-DNN based equalizer can obtain the approximated optimal neural network parameters with the significantly improved convergence speed and reduced training time. Compared to the conventional channel equalization methods of ZF and BPNN, the CW-DNN based channel equalization algorithm can provide better NMSE performance.

Acknowledgments

This work has been supported in part by the national natural science foundation of China (61302062, 61601325, 61901298), in part by the natural science foundation of Tianjin (13JCQNJC00900, 20JCQNJC00300) in China, and in part by Tianjin Key Laboratory of Optoelectronic Sensor and Sensing Network Technology.

References

[1] GE, L., ZHANG, Y., CHEN, G., et al. Compression-based LMMSE channel estimation with adaptive sparsity for massive MIMO in 5G systems. *IEEE Systems Journal*, 2019, vol. 13, no. 4, p. 3847–3857. DOI: 10.1109/JSYST.2019.2897862

[2] GE, L., GOU, Y., ZHANG, Y., et al. Deep neural network based channel estimation for massive MIMO-OFDM systems with imperfect channel state information. *IEEE Systems Journal*, 2021, p. 1–11. DOI: 10.1109/JSYST.2021.3114229

[3] MARZETTA, T. L. Non-cooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 2010, vol. 9, no. 11, p. 3590–3600. DOI: 10.1109/TWC.2010.092810.091092

[4] LUO, H., ZHANG, Y., HUANG, L., et al. A closed-loop reciprocity calibration method for massive MIMO in terrestrial broadcasting systems. *IEEE Transactions on Broadcasting*, 2017, vol. 63, no. 1, p. 11–19. DOI: 10.1109/TBC.2016.2606890

[5] ATAPATTU, L., ARACHCHIGE, G. M., ZIRI-CASTRO, K., et al. Linear adaptive channel equalization for multiuser MIMO-OFDM systems. In *Australasian Telecommunication Networks and Applications Conference (ATNAC)*. Brisbane (Australia), 2012, p. 1–5. DOI: 10.1109/ATNAC.2012.6398080

[6] HIGUCHI, K., KAWAI, H., MAEDA, N., et al. Likelihood function for QRM-MLD suitable for soft-decision turbo decoding and its performance for OFCDM multiplexing in multipath fading channel. In *IEEE 15th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. Barcelona (Spain), 2004, p. 1142–1148. DOI: 10.1109/PIMRC.2004.1373877

[7] SHIN, W., KIM, H., SON, M., et al. An improved LLR computation for QRM-MLD in coded MIMO systems. In *IEEE 66th Vehicular Technology Conference*. Baltimore (USA), 2007, p. 447–451. DOI: 10.1109/VETECF.2007.105

[8] YANG, C. Y., CHEN, B. S. Adaptive channel-tracking method and equalization for MC-CDMA systems over rapidly fading channel under colored noise. *EURASIP Journal on Advances in Signal Processing*, 2010, article no. 871250, p. 1–12. DOI:10.1155/2010/871650

[9] ARABLOUEI, R., DOĜANÇAY K. Low-complexity adaptive decision-feedback equalization of MIMO channels. *Signal Processing*, 2012, vol. 92, no. 6, p. 1515–1524. DOI: 10.1016/j.sigpro.2011.12.012

[10] CHEN, S., ZHU C. ICI and ISI analysis and mitigation for OFDM systems with insufficient cyclic prefix in time-varying channels. *IEEE Transactions on Consumer Electronics*, 2004, vol. 50, no. 1, p. 78–83. DOI: 10.1109/TCE.2004.1355879

[11] YANG, L., NIE, M., ZHONG, Z. Channel equalization of MIMO-OFDM system based on extreme learning machine. *Applied Mechanics and Materials*, 2014, vol. 536–537, p. 1751–1757. DOI:10.4028/www.scientific.net/AMM.536-537.1751

[12] NAWAZ, S. J., MOHSIN, S., IKARAM, A. A. Neural network based MIMO-OFDM channel equalizer using comb-type pilot arrangement. In *International Conference on Future Computer and Communication*. Kuala Lumpar (Malaysia), 2009, p. 36–41. DOI: 10.1109/ICFCC.2009.136

[13] BURSE, K., YADAV, R. N., SHRIVASTAVA, S. C. Channel equalization using neural networks: A review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2010, vol. 40, no. 3, p. 352–357. DOI: 10.1109/TSMCC.2009.2038279

[14] HAO, J., YANG, L. Semi-blind channel estimation of MIMO-OFDM systems based on RBF network. In *IET International Communication Conference on Wireless Mobile and Computing (CCWMC 2011)*. Shanghai (China), 2011, p. 187–191. DOI: 10.1049/cp.2011.0872

[15] HA, C. B., YOU, Y. H., SONG, H. K. Machine learning model for adaptive modulation of multi-stream in MIMO-OFDM system. *IEEE Access*, 2018, vol. 7, p. 5141–5152. DOI: 10.1109/ACCESS.2018.2889076

[16] SIMSIR, S., TASPINAR, N. Channel estimation using neural network in orthogonal frequency division multiplexing-interleave division multiple access (OFDM-IDMA) system. In *International Telecommunications Symposium (ITS)*. Sao Paulo (Brazil), 2014, p. 1–5. DOI: 10.1109/ITS.2014.6947977

[17] TAŞPINAR, N., SEYMAN, M. N. Back propagation neural network approach for channel estimation in OFDM system. In *IEEE International Conference on Wireless Communications, Networking and Information Security*. Beijing (China), 2010, p. 265–268. DOI: 10.1109/WCINS.2010.5541934

[18] BOCCARDI, F., HEATH, R. W., LOZANO, A., et al. Five disruptive technology directions for 5G. *IEEE Communications Magazine*, 2014, vol. 52, no. 2, p. 74–80. DOI: 10.1109/MCOM.2014.6736746

[19] SEYMAN, M. N., TAŞPINAR, N. Channel estimation based on neural network in space time block coded MIMO-OFDM system. *Digital Signal Processing*, 2013, vol. 23, no. 1, p. 275–280. DOI:10.1016/j.dsp.2012.08.003

[20] WANG, T., WEN, C., WANG, H., et al. Deep learning for wireless physical layer: Opportunities and challenges. *China Communications*, 2017, vol. 14, no. 11, p. 92–111. DOI: 10.1109/CC.2017.8233654

[21] YE, H., LI, G. Y., TUANG, B. H. Power of deep learning for channel estimation and signal detection in OFDM systems. *IEEE Wireless Communications Letters*, 2018, vol. 7, no. 1, p. 114–117. DOI: 10.1109/LWC.2017.2757490

[22] WANG, X., HAU, H., XU, Y. Pilot-assisted channel estimation and signal detection in uplink multi-user MIMO systems with deep learning. *IEEE Access*, 2020, vol. 8, p. 44936–44946. DOI: 10.1109/ACCESS.2020.2978253

[23] SAMUEL, N., DISKIN, T., WIESEL, A. Learning to Detect. *IEEE Transactions on Signal Processing*, 2019, vol. 67, no. 10, p. 2554–2564. DOI: 10.1109/TSP.2019.2899805

[24] BAEK, M., KWAK, S., JUNG, J., et al. Implementation methodologies of deep learning-based signal detection for conventional MIMO transmitters. *IEEE Transactions on Broadcasting*, 2019, vol. 65, no. 3, p. 636–642. DOI: 10.1109/TBC.2019.2891051

[25] HUANG, Q., ZHAO, C., JIANG, M., et al. A novel OFDM equalizer for large doppler shift channel through deep learning. In *IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. Honolulu (USA), 2019, p. 1–5. DOI: 10.1109/VTCFall.2019.8891326

[26] LI, Q., ZHANG, A., LI J., et al. Soft decision signal detection of MIMO system based on deep neural network. In *5th International Conference on Computer and Communication Systems (ICCCS)*. Shanghai (China), 2020, p. 665–669. DOI: 10.1109/ICCCS49078.2020.9118425

[27] HE, H., WEN, C., JIN, S., et al. Model-driven deep learning for MIMO detection. *IEEE Transactions on Signal Processing*, 2020, vol. 68, p. 1702–1715. DOI: 10.1109/TSP.2020.2976585

[28] TAN, X., XU, W., SUN, K., et al. Improving massive MIMO message passing detectors with deep neural network. *IEEE Transactions on Vehicular Technology*, 2020, vol. 69, no. 2, p. 1267–1280. DOI: 10.1109/TVT.2019.2960763

[29] LIAO, J., ZHAO, J., GAO, F., et al. A Model-driven deep learning method for massive MIMO detection. *IEEE Communications Letters*, 2020, vol. 24, no. 8, p. 1724–1728. DOI: 10.1109/LCOMM.2020.2989672

[30] XUE, S., MA, Y., LI, A., et al. On unsupervised deep learning solutions for coherent MU-SIMO detection in fading channels. In *IEEE International Conference on Communications (ICC)*. Shanghai (China), 2019, p. 1–6. DOI: 10.1109/ICC.2019.8761999

[31] CHENG, X., LIU, D., ZHU, Z., et al. A ResNet-DNN based channel estimation and equalization scheme in FBMC/OQAM systems. In *10th International Conference on Wireless Communications and Signal Processing (WCSP)*. Hangzhou (China), 2018, p. 1–5. DOI: 10.1109/WCSP.2018.8555649

[32] MIAO, P., YIN, W., PENG, H., et al. Deep learning based nonlinear equalization for DCO-OFDM systems. In *IEEE International Conference on Electrical Engineering and Mechatronics Technology (ICEEMT)*. Qingdao (China), 2021, p. 699–703, DOI: 10.1109/ICEEMT52412.2021.9602129

About the Authors . . .

Lijun GE received the B.E. degree in the major of Electronics Science and Technology in 2006 from Nankai University, Tianjin, China, where subsequently, he did graduate study in the major of Communication and Information Systems, and the Ph.D. degree in the major of Communication and Information Systems after five-year graduate study from Nankai University, in 2011. From 2008 to 2010, he was a Teaching Assistant with Nankai University, teaching communicationrelated experiment courses. In 2011, he joined the Department of Communication Engineering, School of Electrical and Electronic Engineering, Tiangong University, Tianjin, China, as a Lecturer in Communication and Information Systems. Since 2013, he has been an Associate Professor in the same academic field, and he has been a Professor and Doctoral Supervisor since 2021. His research interests include OFDM, UWB, massive MIMO wireless communication technologies and wearable Internet of Things, FPGA technologies and applications, and development of communication and information systems. During the past years, he was in charge of eight projects supported by the nation or the city or some companies, and published more than 40 academic papers.

Changcheng QI received the B.S. degree in Communication Engineering in 2020 from the School of Electrical and Electronic Engineering, Tiangong University, Tianjin, China. He is currently studying toward the M.S. degree in the major of Information and Communication Engineering from the School of Electrical and Electronic Engineering, Tiangong

University, Tianjin, China, since 2020. His research interests include massive MIMO systems, deep learning based wireless communication algorithms, and simulation of communication systems.

Yuchuan GUO received the B.S. degree in Communication Engineering in 2019 from the School of Electrical and Electronic Engineering, Tiangong University, Tianjin, China. He is currently studying toward the M.S. degree in the major of Information and Communication Engineering from the School of Electrical and Electronic Engineering, Tiangong University, Tianjin, China, since 2019. His research interests include massive MIMO systems, deep learning based wireless communication algorithms, and simulation of communication systems.

Lei QIAN (corresponding author) received the B.Eng. and Ph.D. degrees from the College of Communications Engineering, Jilin University, Changchun, China, in 2016 and 2021, respectively. Now she is with the Tianjin Key Laboratory of Optoelectronic Detection Technology and System, School of Electrical and Electronic Engineering, Tiangong University, as a Lecturer. She was a visiting Ph.D. student with the School of Engineering, University of British Columbia, Canada, from 2019 to 2020, sponsored by the Chinese Scholarship Council. Her current research interests include delay QoS guarantee, effective capacity, visible light communications, resource allocation, physical layer security and channel estimation.

Jun TONG (corresponding author) received the Ph.D. degree in Electronic Engineering from the City University of Hong Kong, Hong Kong SAR, China, in 2009. He is currently a Senior Lecturer with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia. His research interests include signal processing and its applications to wireless communications and instrumentation.

Peng WEI received the Ph.D. degree in Communication and Information Systems from the University of Electronic Science and Technology of China (UESTC) in 2017. He was also a visiting student in the department of Electrical and Computer Engineering at the University of Delaware (UD) from 2014 to 2016. He has been a Lecturer of the School of Electronics and Information Engineering at Tiangong University (TGU) from 2017 to 2020. He is now a Post Doctor with the Department of Electronic Engineering, Tsinghua University, from 2020. His research interests are in wireless communication network, multicarrier system, and signal processing.

Appendix A: Algorithm 1

Algorithm 1. CW-DNN Based Algorithm for Channel Equalization and Signal Detection.

Input: The pilot subcarriers $x\left(k_p\right)$ and data subcarriers $x\left(k_d\right)$ of the received signal in the training set and testing set.
Output: The equalized signal $\hat{x}_(k_d)$.
Initialization: The weights W and bias b of the hidden layers and the output layer are randomly initialized.
Training:

1: Input the massive MIMO frequency domain data sets $\left\{\left(x_1, x_2 \ldots, x_{k_p+k_d}\right), \left(o_1, o_2 \ldots, o_{k_d}\right)\right\}$ into the CW-DNN.
2: Calculate the outputs of the CW-DNN.
3: Calculate the value of the cost function E^* according to (17) and (18).
4: Calculate and update the weights and biases according to (15) and (16).
5: Repeat steps 1-4, until the stopping criterion is satisfied (the error difference between adjacent two times is very small or limit the number of iterations directly).
6: Return trained CW-DNN models with optimal weights and bias.

Testing:

7: Input the testing set into the CW-DNN, and calculate the outputs of the CW-DNN, get the equalized signal.
8: Using the signal detection method based on symbolic slicer , restore the equalized signal to the corresponding constellation point, get the detection signal.
9: Calculate the NMSE to evaluate the performance of the proposed CW-DNN based algorithm.

End