Article:
Luong, Thien Van, Ko, Youngwook, Matthaiou, Michail et al. (2021) Deep Learning-Aided Multicarrier Systems. IEEE Transactions on Wireless Communications. pp. 2109-2119. ISSN: 1536-1276

https://doi.org/10.1109/TWC.2020.3039180

Deep Learning-Aided Multicarrier Systems

Thien Van Luong, Youngwook Ko, *Senior Member, IEEE,* Michail Matthaiou, *Senior Member, IEEE,* Ngo Anh Vien, Minh-Tuan Le and Vu-Duc Ngo, *Member, IEEE*

***Abstract*—This paper proposes a deep learning (DL)-aided multicarrier (MC) system operating on fading channels, where both modulation and demodulation blocks are modeled by deep neural networks (DNNs), regarded as the encoder and decoder of an autoencoder (AE) architecture, respectively. Unlike existing AE-based systems, which incorporate domain knowledge of a channel equalizer to suppress the effects of wireless channels, the proposed scheme, termed as MC-AE, directly feeds the decoder with the channel state information and received signal, which are then processed in a fully data-driven manner. This new approach enables MC-AE to jointly learn the encoder and decoder to optimize the diversity and coding gains over fading channels. In particular, the block error rate of MC-AE is analyzed to show its higher performance gains than existing hand-crafted baselines, such as various recent index modulation-based MC schemes. We then extend MC-AE to multiuser scenarios, wherein the resultant system is termed as MU-MC-AE. Accordingly, two novel DNN structures for uplink and downlink MU-MC-AE transmissions are proposed, along with a novel cost function that ensures a fast training convergence and fairness among users. Finally, simulation results are provided to show the superiority of the proposed DL-based schemes over current baselines, in terms of both the error performance and receiver complexity.**

***Index Terms*—Autoencoder, deep learning, deep neural network, DNN, fading channels, MC-AE, multicarrier systems.**

I. Introduction

Multicarrier modulation (MCM) has been widely adopted in various wireless systems, in which orthogonal frequency division multiplexing (OFDM) [1] is the most common MC scheme that has been included in a wide range of wireless standards, such as IEEE 802.11, IEEE 802.16, 3GPP-LTE and LTE-Advanced. Particularly, MC systems divide the transmitted data stream into many substreams which are sent via multiple parallel narrowband subchannels in order to make them experience relatively flat fading. Hence, the adverse effects of multipath fading, such as intersymbol interference (ISI) and delay spreading, can be effectively combated with simple receivers, which makes OFDM-based MCM a key technique for current and next-generation wireless networks.

In recent years, a range of advanced MC schemes based on OFDM have been explored, aiming to enhance the reliability and spectral or energy efficiency. For example, in [2], OFDM with index modulation (OFDM-IM) was proposed, which activates only a subset of sub-carriers to convey data bits via active indices in addition to the M-ary symbols, leading to higher reliability and energy efficiency than classical OFDM. The error performance of OFDM-IM was investigated in [3], to show that its diversity order is limited to one as in classical OFDM. Various IM-based schemes with enhanced transmit diversity have also been proposed, such as the coordinate interleaved OFDM-IM (CI-OFDM-IM) [4] and a repetition code [5]. Especially, in [6], a spreading code was applied to OFDM-IM to maximize the diversity gain, wherein the resulting scheme is termed as spread OFDM-IM (S-OFDM-IM). Prior to this, the spread OFDM (S-OFDM) which uses the rotated Walsh-Hadamard spreading matrix was introduced in [7]. It is shown in [6] that S-OFDM-IM yields better performance than S-OFDM when low-complexity detection schemes are employed, such as minimum mean squared error (MMSE)-based detectors. On a similar note, to improve the spectral efficiency (SE), the dual model OFDM (DM-OFDM) that employs multiple distinguishable signal constellations was proposed in [8]. It is worth noting that the performance improvements of the above mentioned schemes come at the cost of increased receiver complexity. Moreover, they are all based on hand-crafted designs, thus not guaranteed to achieve an optimal performance for each specific channel. These fundamental issues will be addressed by deep learning (DL) in this work.

A multiuser version of OFDM is known as orthogonal frequency-division multiple access (OFDMA) [1], in which each user is assigned different orthogonal sub-carriers. Unlike OFDMA, multicarrier code-division multiple access (MC-CDMA) [9] that combines MCM and CDMA to spread data symbols of multiple users over the same set of sub-carriers, provides improved diversity compared to OFDMA. For practical implementations, a wide range of linear detection schemes are designed for MC-CDMA, such as MMSE, zero-forcing (ZF) and maximum-ratio combining (MRC). Recently, MC-CDMA was combined with IM [10] to result in the IM-MC-CDMA scheme, which exploits the indices of spreading codes to carry data bits. Yet, due to the limits of orthogonal resources, it is hard for these orthogonal schemes to support

Manuscript received May 9, 2019; revised November 11, 2019, March 17, 2020, July 7, 2020, and September 30, 2020; accepted November 13, 2020. This work was supported by the EPSRC, U.K., under Grant EP/N509541/1. The work of M. Matthaiou was supported by a research grant from the Department for the Economy Northern Ireland under the US-Ireland R&D Partnership Programme and by the EPSRC, U.K., under Grant EP/P000673/1. The associate editor coordinating the review of this article and approving it for publication was J. Hoydis. *(Corresponding author: Youngwook Ko.)*

T. V. Luong is with the Faculty of Computer Science, Phenikaa University, Hanoi 12116, Vietnam, and also with the Phenikaa Research and Technology Institute (PRATI), A&A Green Phoenix Group JSC, Hanoi 11313, Vietnam, (e-mail: thien.luongvan@phenikaa-uni.edu.vn).

Y. Ko is with the University of York, Heslington, York, YO10 5DD, UK, (email: youngwook.ko@york.ac.uk).

M. Mathhaiou is with the Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast, Belfast, BT3 9DT, UK, (e-mail: m.matthaiou@qub.ac.uk).

N. A. Vien is with the Bosch Center for Artificial Intelligence, 71272 Renningen, Germany, (e-mail: anhvien.ngo@bosch.com).

M.-T. Le is with the MobiFone R&D Center, MobiFone Corporation, Hanoi 11312, Vietnam, (e-mail: tuan.minh@mobifone.vn).

V.-D. Ngo is with the School of Electronics and Telecommunications, Hanoi University of Science and Technology (HUST), Hanoi 11657, Vietnam, (e-mail: duc.ngovu@hust.edu.vn).

massive connectivity in future wireless networks. For these scenarios, various non-orthogonal multiple access (NOMA) techniques have been explored, in which sparse code multiple access (SCMA) [11] appears to be the most promising, particularly in overloaded transmissions. In SCMA, data bits of multiple users are mapped to sparse multi-dimensional codewords, enabling an iterative multiuser detection (MUD) based on message passing algorithm (MPA) to achieve a near-optimal performance. Note that these schemes require sophisticated receivers, which involve either the maximum likelihood (ML) or iterative MUD to ensure a good performance, otherwise their low-complexity detectors would severely degrade performance. This issue will also be taken into account in our work.

DL [12] has recently been applied to numerous aspects in the field of communications, in particular the physical layer issues. For example, deep neural networks (DNNs) were employed for efficient signal detection of OFDM [13] and OFDM-IM [14], especially under channel impairments. In [15], a deep autoencoder (AE) architecture was adopted to reduce the peak-to-average power ratio (PAPR) of OFDM. Specifically, a novel concept of an end-to-end AE-based system over an additive white Gaussian noise (AWGN) channel was proposed in [16], where both the transmitter and receiver are represented by DNNs based on an AE architecture. For fading channels, this work proposed a radio transformer network (RTN), based on domain knowledge, to suppress the effects of fading on the received signal. However, such a model-driven method that acts as an equalizer does not exploit the inherent multipath diversity gain of fading channels. In [17], the AE concept was applied to each independent sub-carrier of OFDM, hence, the obtained scheme is also unable to provide any diversity advantage. In [18], the SCMA codewords were optimized over the AWGN channel, using a DNN-based AE. To the best of our knowledge, none of existing works has explored the potential of DL in optimizing the diversity and coding gains of MC systems over fading channels.

This paper presents the first attempt of applying DL in MC systems, aiming to address all the issues raised above. Our main contributions are summarized as follows:

- We propose a single-user MC AE-based (MC-AE) system operating on fading channels, whose modulation and demodulation blocks are performed by DNNs based on an AE architecture. Unlike current AE-based systems [16]–[18][1], MC-AE directly feeds the decoder with the CSI and received signal, without any domain knowledge of a channel equalizer. Such a novel fully data-driven system can effectively learn the encoder and decoder to maximize the diversity and coding gains in fading channels.
- The block error rate (BLER) of MC-AE is then analyzed, which reveals that our DL-based scheme achieves better performance gains than the existing hand-crafted baselines, such as various IM-based MC schemes.
- We then extend MC-AE to multiuser scenarios, coined as MU-MC-AE, where two novel DNN structures of MU-MC-AE are proposed for both uplink and downlink transmissions. To efficiently train MU-MC-AE, we design a new loss function which ensures not only a fairness among users but also fast training convergence.
- Finally, extensive simulations are provided to show that the proposed DL-based schemes outperform various hand-crafted baselines, at even lower receiver complexity.

[1]These schemes neither rely on the CSI at the receiver, nor provide frequency diversity gain in the presence of independent fading channels across different sub-carriers. A detailed discussion on the novelty of the proposed MC-AE against the state-of-the-art is provided in Section II.A.

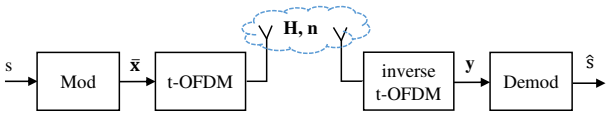

Fig. 1. Block diagram of the classical MC system.

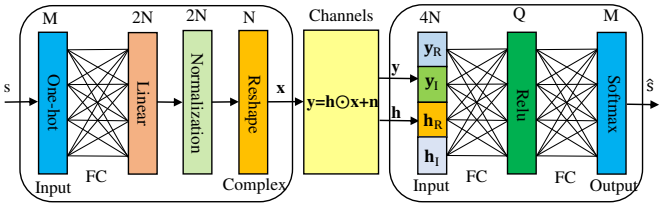

Fig. 2. Network structure of MC-AE.

The rest of the paper is organized as follows. Section II presents the MC-AE system, while its BLER performance is analyzed in Section III. Uplink and downlink MU-MC-AE are presented in Section IV. Simulation results are provided in Section V. Finally, the paper is concluded in Section VI.

Notation: Upper-case bold and lower-case bold letters present matrices and vectors, respectively; $(.)^T$ and $\|.\|$ stand for the transpose operation and the Frobenius norm, respectively. $\mathcal{CN}\left(0,\sigma^2\right)$ denotes the complex Gaussian distribution with zero mean and variance σ^2. The element-wise product is presented by $\odot$, while the Little-O and Big-O notations are denoted by $o(.)$ and $\mathcal{O}(.)$, respectively.

II. Proposed Single-User MC-AE System

A. MC-AE Structure

The classical MC system is briefly illustrated in Fig. 1. In particular, an incoming message s is passed through a modulation block to obtain a transmit data vector $\bar{\mathbf{x}} = [x_1, x_2, ..., x_{N_c}]^T$, where N_c is the number of sub-carriers. This vector is then fed to a time-domain OFDM (t-OFDM) operation, where an inverse fast Fourier transform (IFFT) and a cyclic prefix are performed and added respectively prior to being sent to the receiver, through a fading channel and additive noise. At the receiver, an inverse t-OFDM operation is first applied to recover the received signal in the frequency domain $\mathbf{y}$, which is then fed into a demodulation block to output the estimate $\hat{s}$. The modulation schemes can be the M-ary QAM/PSK modulation or a number of recently emerged IM techniques [19]. Note that modulation and demodulation blocks of existing MCM schemes are separately designed and optimized, and, thus are unable to provide an optimal performance with a reasonably low complexity. For this reason, we intend to exploit DL in order to break the limits of current MC systems in terms of both reliability and receiver complexity.

In particular, the traditional modulation and demodulation blocks are replaced with DNNs, while the t-OFDM block remains unchanged. The resulting scheme can be considered as a multicarrier autoencoder-based (MC-AE) system, where the transmitter and receiver act as the encoder and decoder of an AE structure, respectively. The MC-AE structure is detailed in Fig. 2, where the t-OFDM block is removed for simplicity as the signal processing performed by DNNs is in the frequency domain only. Note that similar to the IM-based schemes [3], we divide N_c sub-carriers into G blocks of N sub-carriers, i.e., $N_c = NG$, and then independently apply the AE structure to these blocks. Fig. 2 demonstrates the network structure of only one MC-AE block for simplicity.

As for the encoder of MC-AE, one out of $M = 2^m$ possible incoming messages $s \in \mathcal{S} = \{s_1, ..., s_M\}$, which can be a bitstream of m bits, is mapped to an one-hot vector $\mathbf{s}$ of the size $M \times 1$, whose entries are all zeros except for a single entry being one. Next, a linear fully-connected (FC) layer with a weight matrix $\mathbf{W}$ and bias vector $\mathbf{b}$ is applied to obtain an $2N$-dimensional vector $\mathbf{z} = \mathbf{W}\mathbf{s}+\mathbf{b}$. A normalization layer is then used to constrain the average transmit power per sub-carrier to be a given constant E_s, as follows: $\mathbf{v} = \sqrt{NE_s}\mathbf{z}/\left\|\mathbf{z}\right\|$, where $\mathbf{v}$ is the output of this layer. Finally, $\mathbf{v}$ is reshaped into an $N \times 1$ complex-valued vector $\mathbf{x}$. The proposed encoder can be represented through the function $\mathbf{x} = f_{\theta_{\text{enc}}}(\mathbf{s})$, where $\theta_{\text{enc}} = \{\mathbf{W}, \mathbf{b}\}$ is its parameters.

After being processed by the t-OFDM block, $\mathbf{x}$ is sent to the receiver, first passing through the fading channel $\mathbf{h}$ and then is being impaired by additive noise $\mathbf{n}$. The received signal in frequency domain is given by

$$\mathbf{y} = \mathbf{h} \odot \mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{h} = [h_1, ..., h_N]^T$ represents the flat Rayleigh fading channel vector across N sub-carriers with $h_i \sim \mathcal{CN}(0, 1)$ and $\mathbf{n}$ denotes the AWGN vector with its entries $n_i \sim \mathcal{CN}\left(0, \sigma^2\right)$, $i = 1, ..., N$.[2] Thus, the average received signal-to-noise ratio (SNR) is expressed by $\bar{\gamma} = E_s/\sigma^2$.

For the signal decoding, we assume that the CSI $\mathbf{h}$ is perfectly known at the receiver and considered as the input of the decoder, along with the received signal $\mathbf{y}$. Particularly, the complex vectors $\mathbf{y}$ and $\mathbf{h}$ are transformed into an $4N$-dimensional real vector $\mathbf{u} = [\mathbf{y}_R, \mathbf{y}_I, \mathbf{h}_R, \mathbf{h}_I]^T$ as shown in Fig. 2, where $\mathbf{y}_R$, $\mathbf{y}_I$ and $\mathbf{h}_R, \mathbf{h}_I$ are real and imaginary parts of $\mathbf{y}$ and $\mathbf{h}$, respectively. As seen in Fig. 2, the proposed decoder has two nonlinear FC layers, in which a hidden layer with Q nodes uses the rectifier linear unit (Relu) as the activation function, whilst an output layer of M nodes employs the softmax activation function [12]. By using the softmax layer as the output layer, the decoder outputs a probability vector $\hat{\mathbf{s}} = [\hat{s}_1, ..., \hat{s}_M]^T$ whose i-th entry is the probability that the message s_i was be transmitted. Finally, the estimated message $\hat{s}$ is determined according to the largest entry of $\hat{\mathbf{s}}$. More specifically, denote by $\theta_{\text{dec}} = \{\mathbf{W}_i, \mathbf{b}_i\}_{i=1,2}$ the parameters of the decoder whose output is expressed by

$$\begin{aligned}\hat{\mathbf{s}} &= f_{\theta_{\text{dec}}}(\mathbf{y}, \mathbf{h}) = f_{\theta_{\text{dec}}}(\mathbf{h} \odot f_{\theta_{\text{enc}}}(\mathbf{s}) + \mathbf{n}, \mathbf{h}) \\ &= \sigma_{\text{Softmax}}\left(\mathbf{W}_2 \sigma_{\text{Relu}}\left(\mathbf{W}_1 \mathbf{u} + \mathbf{b}_1\right) + \mathbf{b}_2\right), \end{aligned} \tag{2}$$

where σ_{Softmax} and σ_{Relu} denote the element-wise softmax and Relu functions, respectively. Then, the transmitted data is recovered as follows $\hat{s} = s_{\hat{i}}$, $\hat{i} = \arg\max \hat{s}_i$ for $i = 1, ..., M$.[3]

The proposed MC-AE is very different from the traditional AE-based schemes, which are based on the RTN [16], [17].[4] In particular, RTN is designed under block fading channels [17], where the sub-carrier channels remain constant over several channel uses. Such channel condition ensures RTN to be constructed based on domain knowledge of a channel equalizer (e.g., zero forcing), without any knowledge of CSI. Hence, RTN can be considered as a model-driven and noncoherent approach. By contrast, the proposed MC-AE is designed for a time-varying channel condition, where the channel coefficients change randomly in every channel use. This means that our approach is not limited to block fading channels as in RTN. In this context, the proposed scheme processes CSI together with the received signal in a fully data-driven manner, where perfect CSI at the receiver is used to harness the frequency diversity across different sub-carriers, instead of relying on domain knowledge of a channel equalizer as in RTN.[5]

When M is very large, i.e., higher data rates, the one-hot encoding makes the MC-AE model too complicated, hence unstable and time-consuming in training. To address this issue, we can use an embedding layer as the input of the encoder, where the resulting system is termed as Emb-MC-AE. Particularly, the incoming message s is mapped to a real-valued embedding vector $\mathbf{s}$ of length L, whose entries are trainable parameters, which are updated during the training. The remaining layers of the encoder, as well as, the entire decoder are retained as those in the previous MC-AE.[6]

B. Training procedure of MC-AE

The MC-AE model is trained offline by a set of random incoming messages s or its corresponding one-hot vectors $\mathbf{s}$, while the channel $\mathbf{h}$ and noise $\mathbf{n}$ are randomly generated and added to the output of the encoder while training, based on their statistical models, as described in (1). We adopt the mean-squared error (MSE) loss function for training MC-AE as

$$\mathcal{L}(\theta) = \frac{1}{T}\sum_{i=1}^{T}\left\|\mathbf{s}_i - \hat{\mathbf{s}}_i\right\|^2, \tag{3}$$

[2]In practice, since OFDM often experiences frequency selective fading with correlated sub-carriers, we can interleave sub-carriers over G MC-AE blocks to make the sub-carriers within each block nearly independent in order to enhance transmit diversity [1].

[3]The numbers of hidden layers needed for the encoder and decoder of MC-AE have been minimized to reduce the complexity as we observed via experiments that further increasing this number does not improve performance.

[4]RTN was proposed for time-domain systems in [16]. It was then extended to frequency-domain systems in [17], which is more relevant to our work.

[5]The use of CSI in MC-AE is reasonable as the recent multicarrier schemes [2]–[8] also require knowledge of CSI at the receiver for reliable signal detection. Note that this is a very common consideration in the context of coherent communications. Nevertheless, an explicit CSI estimation is deemed as a penalty for diversity enhancement of MC-AE compared to the existing AE schemes with RTN [17].

[6]For example, in Fig. 7(b) in Section V, where $M = 1024$, $N = 4$, and $Q = 256$, using one-hot encoding, the encoder of MC-AE has $MQ + Q = 262400$ trainable parameters, while that of Emb-MC-AE with $L = 16$ in the embedding layer is much lower with $LQ + Q = 4352$ trainable parameters.

TABLE I
BLER COMPARISON BETWEEN MSE AND CE LOSSES

SNR	$(N, M) = (4, 32)$		$(N, M) = (4, 64)$	
	MSE	CE	MSE	CE
0 dB	0.36104	0.36344	0.43242	0.43241
5 dB	0.07756	0.07615	0.10052	0.10076
10 dB	0.00569	0.00568	0.0083	0.0083
15 dB	0.00024	0.00032	0.0005	0.00052
20 dB	2e-05	3e-05	8.75e-05	0.0001

where $\theta = \{\theta_{\text{dec}}, \theta_{\text{enc}}\}$ denotes the overall MC-AE model parameters, $\hat{\mathbf{s}}_i$ is the prediction of $\mathbf{s}_i$ and T is the training batch size. Based on (3), θ is updated for each batch of training data, using the stochastic gradient descent (SGD) method as follows:

$$\theta := \theta - \eta \nabla \mathcal{L}(\theta), \tag{4}$$

where η stands for the learning rate. For improved training, the adaptive moment estimation (Adam) optimizer [20] (an advanced SGD-based method) and the Xavier initialization [21] are employed as they are are widely available in a number of off-the-shelf DL libraries such as Tensorflow [22].[7]

It is worth noting that we also considered the cross-entropy (CE) loss in our experiments and found that it yields slightly worse performance than MSE at high SNRs, while at low and medium SNRs, both losses have similar performance. This is shown in Table I, where we compare the block error rate (BLER) of MC-AE achieved by the two losses. The MSE loss is also a reasonable choice for minimizing the BLER of our scheme as will be analyzed in Section III. Interestingly, if an outer channel code and bit-metric decoder are used, we can use the total binary cross-entropy as the loss function for the bit-level optimization to further improve the bit error rate [23].

As for Emb-MC-AE, the MSE loss is not applicable since $\mathbf{s}$ and $\hat{\mathbf{s}}$ have different lengths. Instead, the sparse categorical cross-entropy will be used [22], which takes the true label as a single integer that is the index of the incoming message as follows: $\mathcal{L}(\theta) = -\frac{1}{T}\sum_{i=1}^{T} \log(\hat{s}_i)$, where i is the true label.

Since the system may operate at different noise levels that are represented by the average SNR $\bar{\gamma}$, it is not efficient to train the model multiple times with different SNRs. Instead, we deliberately choose an appropriate training SNR $\bar{\gamma}_{\text{tr}}$ such that the model trained by this SNR still works well at any other SNR levels of interest. Note that $\bar{\gamma}_{\text{tr}}$ has a huge impact on the training performance and is selected based on experiments for certain system parameters (N, M) and channel model. For this, details of selecting $\bar{\gamma}_{\text{tr}}$ as well as other training parameters, such as batch size, learning rate, epoch and training/testing data size will be provided for each experiment in Section V.

III. PERFORMANCE ANALYSIS OF MC-AE

The block error rate of MC-AE is analyzed by investigating the transmit diversity and coding gains and comparing them with that of baseline schemes. Herein, a block error occurs when a message s transmitted by a block of N sub-carriers is incorrectly decoded at the receiver. Accordingly, the BLER can be approximated, in terms of the diversity gain G_d and coding gain G_c, at high SNRs, as follows [24]:

$$P_e = (G_c \bar{\gamma})^{-G_d} + o\left(\bar{\gamma}^{-G_d}\right), \tag{5}$$

where P_e denotes the BLER and $\bar{\gamma}$ is the average SNR.

We first estimate the pairwise error probability (PEP) that the transmitted message s is erroneously detected into another message $\hat{s} \neq s$ for given channel $\mathbf{h}$ as follows:

$$P(s \to \hat{s} | \mathbf{h}) = \mathcal{Q}\left(\sqrt{\frac{\|\mathbf{h} \odot (\mathbf{x} - \hat{\mathbf{x}})\|^2}{2\sigma^2}}\right), \tag{6}$$

where $\mathbf{x} = [x_1, ..., x_N]^T$ and $\hat{\mathbf{x}} = [\hat{x}_1, ..., \hat{x}_N]^T$ are the outputs of the encoder corresponding to inputs s and $\hat{s}$, and $\mathcal{Q}(.)$ is the Gaussian Q-function [24]. Using $\beta_i = |x_i - \hat{x}_i|^2$ for $i = 1, ..., N$, the PEP in (6) can be rewritten as

$$P(s \to \hat{s} | \mathbf{h}) = \mathcal{Q}\left(\sqrt{\frac{\sum_{i=1}^{N} \beta_i |h_i|^2}{2\sigma^2}}\right). \tag{7}$$

Following some derivation steps borrowed from [6], the unconditional PEP of MC-AE is obtained from (7) as

$$P(s \to \hat{s}) = \frac{1}{\pi}\int_0^{\pi/2} \prod_{i=1}^{N} \left(\frac{\sin^2\phi}{\sin^2\phi + \frac{\beta_i \bar{\gamma}}{4}}\right) d\phi, \tag{8}$$

and its approximation is given by

$$P(s \to \hat{s}) \approx \frac{(\bar{\gamma}/4)^{-|\mathcal{K}_{s,\hat{s}}|}}{2\prod_{i \in \mathcal{K}_{s,\hat{s}}} \beta_i}, \tag{9}$$

where $\mathcal{K}_{s,\hat{s}} = \{i | \beta_i \neq 0\}$ and $|\mathcal{K}_{s,\hat{s}}|$ is its cardinality. Herein, without loss of generality, we assume that $E_s = 1$, i.e., unit average transmit power per sub-carrier to attain $\bar{\gamma} = 1/\sigma^2$.

Using (9), the BLER of MC-AE is approximated based on the total probability theory as follows [6]:

$$P_e \approx \frac{1}{2M} \sum_{s \in \mathcal{S}} \sum_{\hat{s} \neq s} \frac{(\bar{\gamma}/4)^{-|\mathcal{K}_{s,\hat{s}}|}}{\prod_{i \in \mathcal{K}_{s,\hat{s}}} \beta_i}, \tag{10}$$

where $\mathcal{S} = \{s_1, ..., s_M\}$ is the set of all M possible messages. As a result, the diversity gain achieved by MC-AE is given by

$$G_d = \min_{\hat{s} \neq s \in \mathcal{S}} |\mathcal{K}_{s,\hat{s}}|. \tag{11}$$

Hence, we can represent P_e in (10) according to (5) as follows:

$$P_e \approx \frac{1}{2^{2G_d+1}M} \sum_{s \in \mathcal{S}} \sum_{\hat{s} \neq s,\, |\mathcal{K}_{s,\hat{s}}| = G_d} \frac{\bar{\gamma}^{-G_d}}{\prod_{i \in \mathcal{K}_{s,\hat{s}}} \beta_i} + o\left(\bar{\gamma}^{-G_d}\right), \tag{12}$$

in order to obtain the coding gain given by

$$G_c = \left(\frac{1}{2^{2G_d+1}M} \sum_{s \in \mathcal{S}} \sum_{\hat{s} \neq s, |\mathcal{K}_{s,\hat{s}}| = G_d} \frac{1}{\prod_{i \in \mathcal{K}_{s,\hat{s}}} \beta_i}\right)^{-\frac{1}{G_d}}. \tag{13}$$

Note that the formula of G_c in (13) differs from that in [6, Eq. (12)] in the sense that we take into account all PEPs with $|\mathcal{K}_{s,\hat{s}}| = G_d$, while the former is based only on a maximum

[7]The Xavier initialization [21] is used to achieve better local optimum, while the Adam optimizer [20] is used for faster convergence than SGD [20]. Here, except for the learning rate, the Adam optimizer with default parameters on Tensorflow is used, such as $\beta_1 = 0.9$ and $\beta_2 = 0.999$.

TABLE II
DIVERSITY AND CODING GAINS OF MC-AE AND BASELINES

(N, M)	$(2,4)$		$(4,16)$		$(4,64)$	
	G_d	G_c	G_d	G_c	G_d	G_c
OFDM [1]	1	16	1	8	-	-
OFDM-IM [2]	1	64	1	25.6	1	6.4
DM-OFDM [8]	-	-	-	-	1	8
CI-OFDM-IM [4]	-	-	2	7.07	2	2.36
S-OFDM [7]	2	7.54	4	3.09	-	-
S-OFDM-IM [6]	2	7.54	4	2.50	4	1.19
Proposed MC-AE	2	8.14	4	3.34	4	1.48

PEP. Hence, our formula is more accurate than that in [6], which is very important for a fair comparison in the following.

Table II compares the diversity and coding gains achieved by MC-AE and the state-of-the-art baselines.[8] Herein, several gain values are left empty since the relative schemes do not work at corresponding (N, M). It is shown that MC-AE can attain a maximum diversity gain which is N similar to two spreading MC schemes, i.e., S-OFDM and S-OFDM-IM, and much higher than that of classical OFDM and IM-based schemes. In addition, our learning-based scheme achieves better coding gains than both spreading schemes. These observations confirm the powerful capability of DL in learning and optimizing the performance gains of MC systems over fading channels. By contrast, current hand-crafted schemes are unable to produce such high performance gains.

In order to gain a deeper insight into where these gains come from, we demonstrate the learned constellation of MC-AE with $(N, M) = (2, 8)$ compared with the constellations of OFDM-IM and S-OFDM-IM in Fig. 3. It is shown that the constellation points of the benchmarks are overlapped and this limits their diversity gain, while that of MC-AE are well-separated to achieve a higher gain. The reason is that the diversity gain of a PEP in (9) is the number of non-zeros elements of $\mathbf{x} - \hat{\mathbf{x}}$, which intuitively has more zeros when the constellation is more overlapped. In particular, in Fig. 3, MC-AE and S-OFDM-IM achieve the same diversity gain of two, while that of OFDM-IM is only one. More importantly, the coding gain of MC-AE is 3.68, while that of S-OFDM-IM is lower with 3.53. More learned constellations of MC-AE and their gain calculation can be found at https://github.com/ThienVanLuong/multicarrier_autoencoder.

IV. PROPOSED MULTIUSER MC-AE SYSTEM

We extend the proposed MC-AE to multiuser systems for both uplink and downlink transmissions, termed as MU-MC-AE. Particularly, two novel DNN structures of uplink and downlink MU-MC-AE are designed in the presence of fading channels and a new loss function is proposed, aiming at fairness among users and fast convergence speed of training.

[8]Denote by $\mathcal{X} = \{\mathbf{x}_1, ..., \mathbf{x}_M\}$ the set of all M possible transmitted vectors of MC-AE with $\mathbf{x}_k = f_{\theta_{\text{enc}}}(\mathbf{s}_k)$, where $\mathbf{s}_k$ is the one-hot vector corresponding to $s_k \in \mathcal{S}$ for $k = 1, ..., M$. Note that $\mathcal{X}$ can be referred to as a codebook, where different schemes (including baselines) have different codebooks which produce different coding and diversity gains. Hence, as long as the codebook of each scheme is known, it is straightforward to calculate its performance gains according to (11) and (13). The codebooks of the baselines can be found in the references listed in Table II, while that of MC-AE with (M, N) as in Table II are obtained after trained with the training parameters similar to those given at column "Fig. 8" of Table IV in Section V.

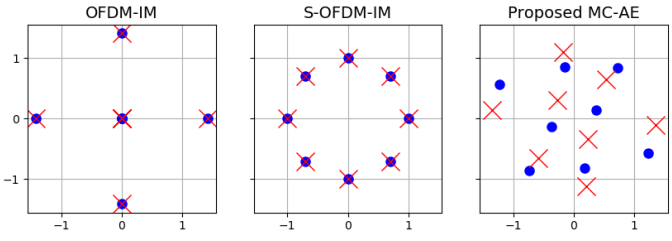

Fig. 3. The constellations of MC-AE with $(N, M) = (2, 8)$ learned over the Rayleigh channel and the benchmarks, where each marker corresponds to the complex constellation points of a sub-carrier.

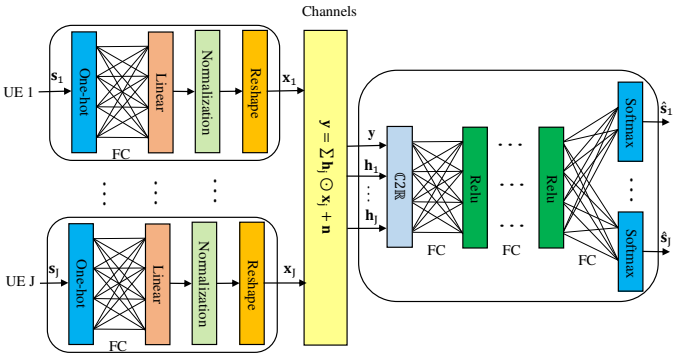

Fig. 4. Network structure of uplink MU-MC-AE.

A. Uplink MU-MC-AE

The uplink MU-MC-AE structure is illustrated in Fig. 4, where J users simultaneously transmit their data to the base station (BS), using the same N sub-carriers. The encoder structures of all users are the same as designed in Section II. Accordingly, the transmitted signal vector in the frequency domain of user j can be represented by $\mathbf{x}_j = f_{\theta^j_{\text{enc}}}(\mathbf{s}_j)$, where θ^j_{enc} is its encoder parameters for $j = 1, ..., J$. It is assumed that the transmit powers across N sub-carriers are the same among users. At the BS, the received signal is given by

$$\mathbf{y} = \sum_{j=1}^{J} \mathbf{h}_j \odot \mathbf{x}_j + \mathbf{n}, \tag{14}$$

where $\mathbf{h}_j$ is the channel vector from user j to the BS and $\mathbf{n}$ is the noise vector, whose statistical models are assumed to be the same as in Section II. We also assume that $\mathbf{h}_j$ are perfectly known at the BS, and used as an input to the BS decoder.

Specifically, the real and imaginary parts of complex vectors $\mathbf{y}$ and $\mathbf{h}_j$ are transformed into a real $2(J+1)N$-dimensional vector (denoted by $\mathbf{u}$), which is then fed to the DNN decoder. This transformation can be represented by the $\mathbb{C}2\mathbb{R}$ layer as in Fig. 4. The decoder structure of uplink MU-MC-AE consists of several FC hidden layers which use the Relu as the activation function, while the last layer is split into J independent FC sub-layers, each employing the softmax function to output the decoded signal of the corresponding user. More precisely, let $\mathbf{W}_u$ and $\mathbf{b}_u$ denote the weight and bias, respectively, of the u-th hidden layer of the decoder and let Q_u denote the

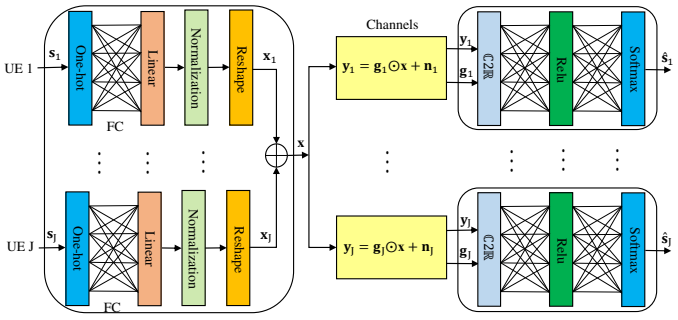

Fig. 5. Network structure of downlink MU-MC-AE.

corresponding number of hidden nodes, for $u = 1, ..., U$, where U is the number of hidden layers. The output vector at the u-th hidden layer of the proposed decoder is given by

$$\mathbf{u}_u = \sigma_{\text{Relu}} \left(\mathbf{W}_u \mathbf{u}_{u-1} + \mathbf{b}_u \right), \tag{15}$$

where $\mathbf{u}_0 = \mathbf{u}$ and $u = 1, ..., U$. Denote by $\mathbf{W}_{U+1}^{(j)}$ and $\mathbf{b}_{U+1}^{(j)}$ the weight and bias of the last sub-layer of user j, then we can obtain the output at each sub-layer as follows:

$$\hat{\mathbf{s}}_j = \sigma_{\text{Softmax}} \left(\mathbf{W}_{U+1}^{(j)} \mathbf{u}_U + \mathbf{b}_{U+1}^{(j)} \right),$$

where the dimension of $\hat{\mathbf{s}}_j$ is M as the previous MC-AE. Finally, we can easily estimate the transmitted message of user j based on the largest entry of $\hat{\mathbf{s}}_j$.

It is worth noting that the decoder structure parameters, such as U and Q_u are selected based on the MU-MC-AE system parameters, such as J, N and M in the sense that when these system parameters get larger, U and Q_u also need to be larger to provide a sufficient model capacity for training.

B. Downlink MU-MC-AE

The proposed downlink MU-MC-AE structure is depicted in Fig. 5, where the BS simultaneously transmits J different signals to J users, using the same N sub-carriers. In particular, the encoder structure at the BS consists of J independent sub-networks, each of which is similar to the encoder network of MC-AE in Fig. 2. The encoder at the BS would produce the frequency-domain vector, which is the summation of the outputs of J sub-networks, i.e., $\mathbf{x} = \sum_{j=1}^{J} \mathbf{x}_j$, where $\mathbf{x}_j = f_{\theta_{\text{enc}}^j}(\mathbf{s}_j)$ with θ_{enc}^j being the model parameters of sub-network j and $\mathbf{s}_j$ being the corresponding one-hot encoded vector of user j. It can be seen from the design of the downlink encoder that the BS allocates evenly the transmit power for every user. The received signal of user j is given by

$$\mathbf{y}_j = \mathbf{g}_j \odot \mathbf{x} + \mathbf{n}_j, \tag{16}$$

where $\mathbf{g}_j$ and $\mathbf{n}_j$ are the downlink channel and noise vectors of user j, respectively, which are assumed to have the same statistical models as the previous section.

At the receiver side, the decoder structure of each user is the same as that of MC-AE. Particularly, the complex channel and received signal vectors $\mathbf{g}_j$, $\mathbf{y}_j$ are first converted into a real $4N$-dimensional input vector of the decoder of user j, which is denoted by $\mathbf{u}_j$. Denote by $\mathbf{W}_i^{(j)}$ and $\mathbf{b}_i^{(j)}$ the weight and bias of the i-th layer of user j's decoder for $i = 1, 2$. Based on (2), the output of the decoder of user j is given by

$$\hat{\mathbf{s}}_j = \sigma_{\text{Softmax}} \left(\mathbf{W}_2^{(j)} \sigma_{\text{Relu}} \left(\mathbf{W}_1^{(j)} \mathbf{u}_j + \mathbf{b}_1^{(j)} \right) + \mathbf{b}_2^{(j)} \right), \tag{17}$$

and the transmitted message of user j will be obtained according to the largest element of $\hat{\mathbf{s}}_j$.

It is worth noting from the proposed downlink MU-MC-AE scheme that although the decoder of each user employs only one hidden layer as the single-user MC-AE, it still effectively decodes the desired signal in the presence of the interference caused by other users. This is achieved by properly training the model as shown in the following.

C. Training procedure of MU-MC-AE

In order to efficiently deploy MU-MC-AE at both the uplink and downlink, we need to train the models offline with data samples including $\{\mathbf{s}_j, \mathbf{h}_j, \mathbf{n}\}$ and $\{\mathbf{s}_j, \mathbf{g}_j, \mathbf{n}_j\}$ for uplink and downlink schemes, respectively. Particularly, the one-hot vectors $\mathbf{s}_j$ are generated first to create J sets of one-hot labels for J users, while the channel and noise vectors are randomly generated and added to the channel layer while training.

Regarding the design of the loss function of MU-MC-AE, we can simply adopt a generalized MSE function for multiple users, which is expressed for a single data point as follows:

$$\mathcal{L}_1 = \sum_{j=1}^{J} \left\| \mathbf{s}_j - \hat{\mathbf{s}}_j \right\|^2. \tag{18}$$

This loss is intuitively able to minimize the difference between the transmitted $\mathbf{s}_j$ and its prediction $\hat{\mathbf{s}}_j$ for every $j = 1, ..., J$. Yet, we found that the traditional approach makes the MSE of each user (denoted by $\mathcal{E}_j = \left\| \mathbf{s}_j - \hat{\mathbf{s}}_j \right\|^2$) vary substantially across them. This means that some users may exhibit small MSE, whilst others exhibit much larger even after training for a long time. As a result, the objective in (18) is not stable and, more importantly, often gets stuck to a bad local optimum, leading to not only a poor overall performance but also a seriously unfair performance among users. Based on these observations, we now introduce a novel loss function for MU-MC-AE, which can address the drawbacks of the traditional approach.

Let us denote $\overline{\mathcal{E}} = \frac{1}{J} \sum_{j=1}^{J} \mathcal{E}_j$. The proposed loss function for MU-MC-AE is expressed, per single data point, by

$$\mathcal{L}_2 = \mathcal{L}_1 + \lambda \sum_{j=1}^{J} \left(\mathcal{E}_j - \overline{\mathcal{E}} \right)^2, \tag{19}$$

where λ is a loss scaling factor. Note that the second term in (19) is a constraint which aims to make the MU-MC-AE model converge faster to a better local optimum than the MSE loss in (18).[9] Furthermore, by minimizing this term, our proposed loss can ensure fairness among users better than the MSE. To illustrate these benefits, Fig. 6 compares the convergence behavior between the two losses. It is interesting that although

[9]In fact, the additional term in (19) has been found experimentally by trial and error, and it happens to have a positive effect on the convergence rate and fairness. Note that the reason for this is not yet clear to us, and therefore, it is an open question that requires further investigation.

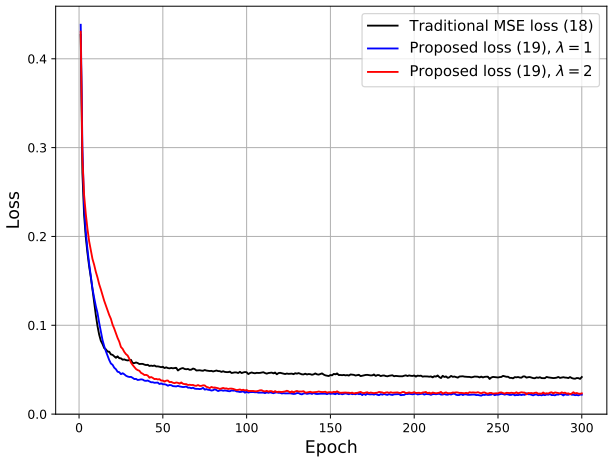

Fig. 6. The proposed loss (19) versus the traditional MSE loss (18) in training uplink MU-MC-AE with $(J, N, M) = (3, 2, 4)$ under the AWGN channel.

TABLE III
BLER COMPARISON BETWEEN THE PROPOSED AND TRADITIONAL LOSSES

Loss	User 1	User 2	User 3	Overall
Traditional loss	0.0111	0.0197	0.0084	0.013
Proposed loss	0.0088	0.0079	0.0043	0.007

$\mathcal{L}_1$ is bounded by $\mathcal{L}_2$, the proposed loss is still much lower than the MSE loss after a few epochs. As a result, our loss can achieve better error performance than the MSE loss, as shown via Table III, where MU-MC-AE has the same setup as in Fig. 6 and the testing SNR is 8 dB. Moreover, the proposed loss achieves better fairness since the BLER difference among users trained with the traditional loss is bigger than that trained with our proposed loss.

Finally, similar to training MC-AE, we employ the Adam optimizer and Xavier initialization method [21] for updating the model parameters of MU-MC-AE. The training SNR $\bar{\gamma}_{\text{tr}}$ also needs to be properly chosen for certain system parameters, such as J, N and M as well as the channel model, in order to result in the best performance. Further details of selecting the training parameters of uplink and downlink MU-MC-AE are provided for each specific experiment in the next section.

V. SIMULATION RESULTS

We conduct various simulations to compare the proposed schemes with the state-of-the-art baselines in terms of the BLER performance and computational complexity. In particular, the BLER of MC-AE is first presented, followed by the MU-MC-AE performance and then the complexity comparison. Apart from the Rayleigh fading channel, we also consider the AWGN channel for a more comprehensive comparison. Note that under the AWGN channel, we feed the decoder of our schemes with the received signals only. The BLER versus the SNR per bit E_b/σ^2 is evaluated for all schemes, where $E_b = mE_s/N$ is the average energy per bit. We also note that all figures in this section are with perfect CSI at the receiver except for Fig. 10 with imperfect CSI.

TABLE IV
(EMB-)MC-AE TRAINING PARAMETERS

Parameter	Fig. 7(a)	Fig. 7(b)	Fig. 8	Fig. 9
Epoch	10^3	2×10^3	2×10^3	5×10^3
Batch size	512	512	512	512
Train size	5×10^4	5×10^4	10^5	10^5
η	0.0001	0.0002	0.0002	0.0002
Q	128	256	256	256
$\bar{\gamma}_{\text{tr}}$	5 dB	10 dB	7 dB	10 dB
L	-	16	-	16

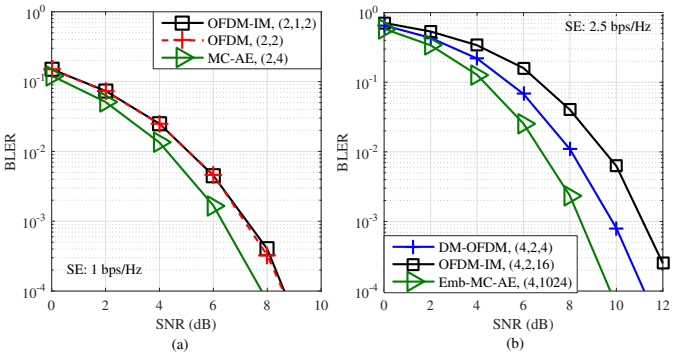

Fig. 7. BLER comparison between (a) MC-AE, (b) Emb-MC-AE and baselines, such as OFDM, OFDM-IM and DM-OFDM in a AWGN channel.

A. BLER Performance of MC-AE

We consider a range of the state-of-the-art MC schemes presented in Table II in Section III as baseline schemes of MC-AE. Particularly, the configuration of OFDM [1] and S-OFDM [7] are denoted by (N, M), while that of IM-based schemes, such as OFDM-IM [2], DM-OFDM [8], CI-OFDM-IM [4] and S-OFDM-IM [6] are denoted by (N, K, M), where N is the number of sub-carriers per block, in which K out of them are active, and M is the size of the conventional M-ary modulation. Meanwhile, the configuration of MC-AE and Emb-MC-AE are represented by (N, M), wherein note that M in our scheme differs from that in the baseline schemes as it refers to the size of the transmitted message per block of N sub-carriers, i.e., $M = 2^m$, where m is the number of data bits per message. The detailed training parameters of our proposed schemes associated with Figs. 7-9 are illustrated in Table IV, where we use 10^6 testing samples for all cases.[10]

Fig. 7 demonstrates the BLER performance of (a) MC-AE of $(2, 4)$, (b) Emb-MC-AE of $(4, 1024)$, in comparison with OFDM, OFDM-IM and DM-OFDM under AWGN channels. Herein, we do not include CI-OFDM-IM, S-OFDM and S-OFDM-IM as the benchmarks in this figure since they are designed to provide transmit diversity under fading channels and are, therefore, not suitable for AWGN channels. Also, DM-OFDM and OFDM do not work at the SEs of 1 and 2.5 bps/Hz, thus, are not included in Fig 7(a) and (b), respectively. It is clearly shown in Fig. 7 that our schemes considerably outperform all baselines at every SNR. For instance, in Fig. 7(b), our scheme achieves an SNR gain of 1 dB and 2 dB over

[10]As shown in Table IV, a large batch size of 512 is selected for all cases since it was found via our experiments that smaller batch sizes, such as 64, 128 and 256, require longer training time, but do not perform better.

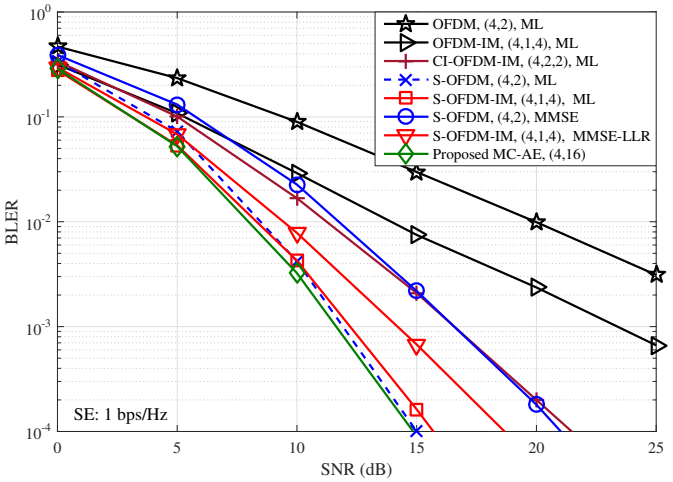

Fig. 8. BLER comparison between MC-AE and baseline schemes, such as OFDM, OFDM-IM, CI-OFDM-IM, S-OFDM and S-OFDM-IM, under Rayleigh fading channels, at the SE of 1 bps/Hz.

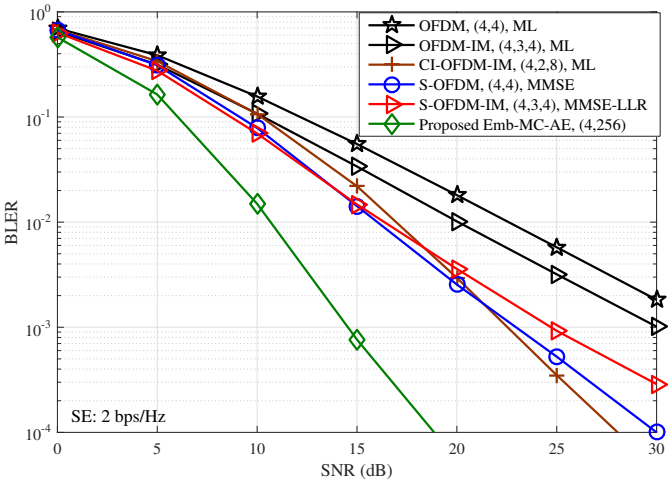

Fig. 9. BLER comparison between Emb-MC-AE and baseline schemes, such as OFDM, OFDM-IM, CI-OFDM-IM, S-OFDM and S-OFDM-IM, under Rayleigh fading channels, at the SE of 2 bps/Hz.

DM-OFDM and OFDM-IM, respectively. By investigating the minimum Euclidean distance (MED) of the learned multi-dimensional codewords $\mathbf{x}$ in comparison with the baselines, we found that MC-AE can provide larger MED and, thus, better BLER than the baselines as shown via Fig. 7.[11]

Fig. 8 compares the proposed MC-AE of $(4, 16)$ with various baselines using the ML detector under Rayleigh fading, at a SE of 1 bps/Hz. We also include the BLER of S-OFDM and S-OFDM-IM with the MMSE-based detector as the ML detector is not practically appropriate for these two spreading schemes due to its high complexity [6]. As observed from Fig. 8, in fading channels, MC-AE performs better than all baselines since it can learn to optimize the diversity and coding gains as analyzed in Section III. Interestingly, this superior performance is achieved when the decoder of MC-AE is very simple with only one hidden layer of $Q = 256$ nodes.[12]

In Fig. 9, we present the BLER of Emb-MC-AE of $(4, 256)$ under Rayleigh fading, in comparison with the baseline schemes at a higher SE, i.e., 2 bps/Hz. At higher SEs, the ML detector of S-OFDM and S-OFDM-IM which has an extremely high complexity is not realistic in practice and, thus, not included in Fig. 9. Instead, we present the BLERs of their MMSE-based low complexity detectors for comparison. It is observed from Fig. 9 that the BLER of Emb-MC-AE is much better than that of baselines as it harnesses the joint optimization of both the transmitter and receiver, which results in higher diversity and coding gains as presented in Section III. Specifically, at a BLER of 10^{-2}, there are significant SNR gains of about 5 dB, 6dB and 12 dB achieved by our scheme over S-OFDM-IM, CI-OFDM-IM and OFDM, respectively.

[11]In AWGN channels, the proposed scheme becomes the conventional AE-based scheme in [16], [17], so we have not included them in Fig. 7 for comparison. Particularly, note that unlike [16], [17], we intend to show via Fig. 7 that the AE-based scheme outperforms the state-of-the-art IM schemes, which has been overlooked in the literature.

[12]Note that the conventional AE scheme [17] for OFDM performs similar to classical OFDM as it applies AE with RTN per single sub-carrier, and is, thus, unable to achieve any diversity gain in fading channels (see Section II.A). Therefore, we do not include [17] in our comparative evaluation.

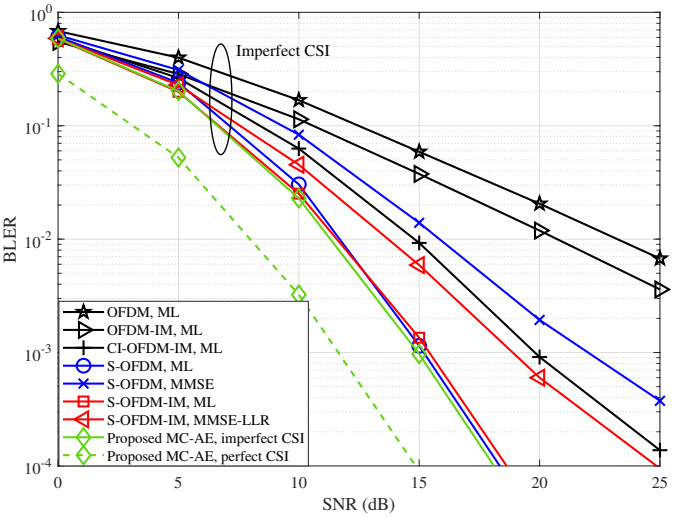

Fig. 10. The BLER comparison between the proposed MC-AE with $(N, M) = (4, 16)$ and the benchmarks under imperfect CSI. The configurations of the benchmark schemes are the same as in Fig. 8.

Fig. 10 illustrates the BLER comparison between MC-AE with $(N, M) = (4, 16)$ and the benchmarks under imperfect CSI. In particular, considering the MMSE channel estimation (see Appendix B in [3]), the actual channel $\mathbf{h} \sim \mathcal{CN}(0, 1)$ is represented by $\mathbf{h} = \hat{\mathbf{h}} + \mathbf{e}$, where $\hat{\mathbf{h}} \sim \mathcal{CN}\left(0, 1 - \epsilon^2\right)$ is the estimated channel, $\mathbf{e} \sim \mathcal{CN}\left(0, \epsilon^2\right)$ is the channel estimation error independent of $\hat{\mathbf{h}}$, in which $\epsilon^2 \in (0, 1)$ varies as a function of the average SNR per sub-carrier, i.e., $\epsilon^2 = (1 + \bar{\gamma})^{-1}$. Note that the proposed scheme is trained with perfect CSI and then tested with MMSE imperfect CSI.[13] It is shown via Fig. 10 that similar to the perfect CSI scenario, the proposed MC-AE still performs well under imperfect CSI and

[13]In practice, the channel dataset should be collected with sufficiently high power of pilot transmission, and the obtained channel can be regarded as nearly perfect. In fact, we found via experiments that when the pilot has a power larger than 30 dB (i.e., $\bar{\gamma} > 30$ dB), the performance of MC-AE trained with imperfect CSI is almost identical to that trained with perfect CSI.

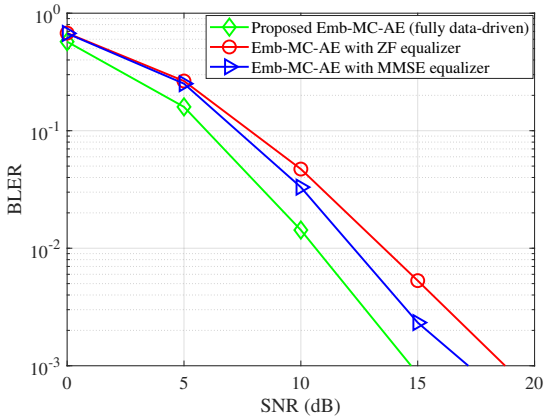

Fig. 11. The BLER comparison between the proposed Emb-MC-AE with $(N, M) = (4, 256)$ and Emb-MC-AE using ZF and MMSE equalizers. Here, all schemes employ perfect CSI at the receiver.

TABLE V
UPLINK AND DOWNLINK MU-MC-AE TRAINING PARAMETERS

Figure	Epoch	η	$\bar{\gamma}_{\text{tr}}$	Q/Q_u	λ
Fig. 12(a)	10^3	0.001	5 dB	128	1
Fig. 12(b)	10^3	0.002	7 dB	256	0.01
Fig. 13(a)	5×10^3	0.0002	17 dB	256	1
Fig. 13(b)	5×10^3	0.0002	20 dB	256-128	5
Fig. 14(a)	2×10^3	0.0005	7 dB	128	1
Fig. 14(b)	2×10^3	0.0005	13 dB	256	2

outperforms the benchmarks. In addition, compared with the perfect CSI case, MC-AE under imperfect CSI suffers from about a 3 dB performance loss.

Fig. 11 shows the BLER of the proposed Emb-MC-AE in comparison with the Emb-MC-AE schemes using ZF and MMSE equalizers to process the received signal before feeding the DNN decoder. It is shown via Fig. 11 that our proposed fully data-driven scheme outperforms the baselines, which confirms the benefit of directly feeding the decoder with the CSI instead of relying on channel equalization. Moreover, unlike MC-AE with the MMSE equalizer, the proposed method does not require an estimation of the noise variance.

B. BLER Performance of MU-MC-AE

A number of current MU-MC schemes, such as MC-CDMA [9], IM-MC-CDMA [10] and SCMA [11] are considered as baselines of MU-MC-AE. Since SCMA outperforms other code domain NOMA schemes [25], we select it as the only NOMA baseline for simplicity. The configuration of all of these schemes is represented by (J, N, M), where J, N are the numbers of users and sub-carriers, respectively, and M is the M-ary modulation size in the baselines, while it is the size of the message sent in MC-MC-AE. The loading factor of a MU-MC system is defined as $\xi = J/N$. Herein, MC-CDMA and IM-MC-CDMA employ MMSE and MRC detectors, respectively, while SCMA uses the MPA with 4 iterations. A range of key training parameters of MU-MC-AE that were fine-tuned in different experiment settings are detailed for Fig. 12-14 in Table V. Besides, the proposed

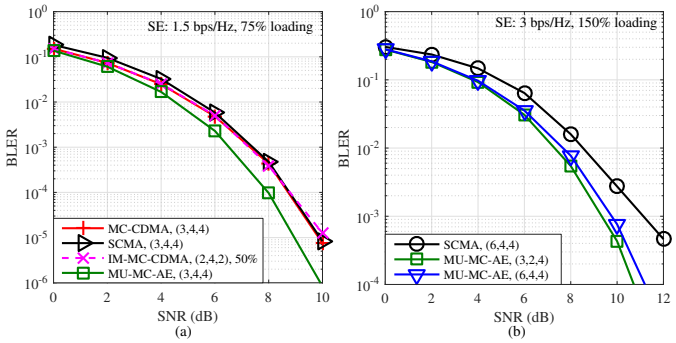

Fig. 12. BLER comparison between uplink MU-MC-AE and baselines including MC-CDMA, IM-MC-CDMA and SCMA, under AWGN channels, at (a) 1.5 bps/Hz - 75% loading, and (b) 3 bps/Hz - 150% loading.

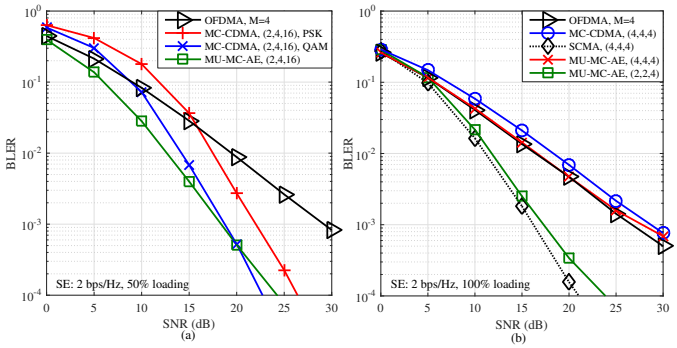

Fig. 13. BLER comparison between uplink MU-MC-AE and baselines including OFDMA, MC-CDMA, IM-MC-CDMA, SCMA under Rayleigh fading, at (a) 2 bps/Hz - 50% loading, and (b) 2 bps/Hz - 100% loading.

schemes are trained and tested by 10^5 and 10^6 data samples, respectively, with a batch size of 512, for all relative figures.

Fig. 12 depicts the BLER comparison between the proposed uplink MU-MC-AE and baselines under the AWGN channel at two different SEs and loading factors. In particular, Fig. 12(a) shows the performance of systems with 75% loading and the SE of 1.5 bps/Hz when $(J, N, M) = (3, 4, 4)$ for all schemes, except for IM-MC-CDMA of $(2, 4, 2)$ with only 50% loading as it is unable to support over 50% loading transmission [10]. Meanwhile, the performance of higher SE and loading systems, i.e., 3 bps/Hz - 150%, is illustrated in Fig 12(b), where both MC-CDMA and IM-MC-CDMA do not work at such overloading with $J > N$, while our scheme can work even at the two overloading configurations of $(6, 4, 4)$ and $(3, 2, 4)$. It is shown in both figures that the proposed scheme exhibits higher reliability than baselines. For example, at a BLER of 10^{-3} in Fig. 12(b), our scheme with $(3, 2, 4)$ provides an SNR gain of 2 dB over SCMA. Interestingly, also in Fig. 12(b), the BLER of the proposed scheme with $(3, 2, 4)$ is slightly better than that with $(6, 4, 4)$, while SCMA is not capable of working at such small N, i.e., $N = 2$. It is noteworthy that such the superior performance of our scheme is attained with a very simple decoder structure which has one hidden layer of $Q = 128$ or 256 nodes as shown in Table V.

Fig. 13 shows the BLER of the uplink MU-MC-AE in comparison with that of baselines under Rayleigh fading at the SE of 2 bps/Hz and (a) $\xi = 50\%$ and (b) $\xi = 100\%$. Note

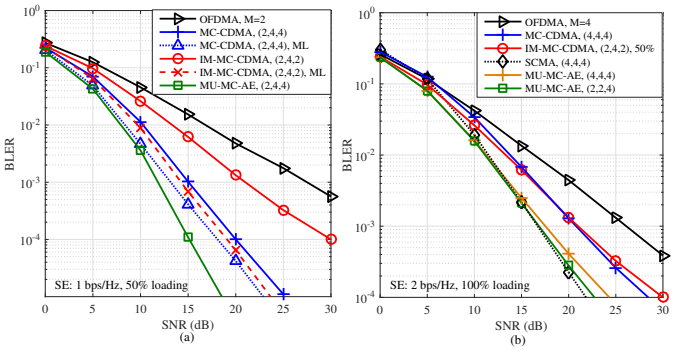

Fig. 14. BLER comparison between downlink MU-MC-AE and baselines including OFDMA, MC-CDMA, IM-MC-CDMA, under Rayleigh fading, at (a) 1 bps/Hz - 50% loading, and (b) 2 bps/Hz - 100% loading.

that we ignore the performance of uplink IM-MC-CDMA in this figure since it is very poor when the MRC detector is used [10]. Instead, the BLER of OFDMA with the single-carrier detection is included for comparison. Since SCMA is designed mainly to support transmissions with high loading factors of over 100%, we do not include this scheme in Fig. 13(a), while we illustrate the BLER of MC-CDMA with both 16-QAM and 16-PSK modulations. It is shown from Fig. 13(a) that the proposed scheme which employs only one hidden layer of 256 nodes at the receiver still outperforms baselines, especially at low-to-medium SNRs. Meanwhile, in Fig. 13(b) with 100% loading, the proposed scheme needs two hidden layers at the decoder with $Q_1 = 256$ and $Q_2 = 128$, as shown in Table V, to efficiently decode the desired signals under higher inter-user interference than the case in Fig. 13(a). As seen from Fig. 13(b), the BLER of MU-MC-AE of $(4, 4, 4)$ is similar to that of OFDMA and better than that of MC-CDMA. However, the proposed scheme of $(2, 2, 4)$ achieves a promising performance which is relatively close to SCMA and much better than OFDMA and MC-OFDMA with over 10 dB SNR gains at the BLER of 10^{-3}. This is because our scheme of $(2, 2, 4)$ with fewer users experiences less inter-user interference than the case of $(4, 4, 4)$.

Fig. 14 illustrates the BLER comparison between the proposed downlink MU-MC-AE and baselines under Rayleigh fading, at two different SEs and loading factors. Particularly, we include the BLER of MC-CDMA and IM-MC-CDMA with the ML detector in Fig. 14(a) in addition to the performance of their low-complexity detectors. It is worth pointing out from Fig. 14(a) that MU-MC-AE significantly outperforms all baseline schemes, even when they use the ML detector, while our scheme only requires a simple decoder structure similar to the single-user MC-AE. This indicates that by training with lots of data samples, the proposed scheme is able to achieve a better mapping at the transmitter, i.e., $\mathbf{x}_j = f_{\theta^j_{\text{enc}}}(\mathbf{s}_j)$, than current hand-crafted schemes that are based on orthogonal codes. Moreover, for a higher SE and loading factor in Fig. 14(b), MU-MC-AE has the BLER relatively close to SCMA, while still performing much better all remaining baselines. Finally, it should be noted that MU-MC-AE tends to perform worse at high SNRs as it is trained with only $\bar{\gamma}_{\text{tr}}$, thus, does not always guarantee to perform best at every SNR.

TABLE VI
RUNTIMES OF (EMB-)MC-AE AND BASELINES IN MILLISECOND

Scheme	Fig. 7(a)	Fig. 7(b)	Fig. 8	Fig. 9
OFDM [1]	0.019	-	0.032	0.039
OFDM-IM [2]	0.046	13.7	0.107	1.7
DM-OFDM [8]	-	20.7	-	-
CI-OFDM-IM [4]	-	-	0.185	2.04
S-OFDM [7]	-	-	0.036/0.121	0.037
S-OFDM-IM [6]	-	-	0.055/0.145	0.075
(Emb-)MC-AE	0.027	0.413	0.028	0.14

TABLE VII
RUNTIMES OF MU-MC-AE AND BASELINES IN MILLISECOND

Scheme	Fig. 12(b)	Fig. 13(b)	Fig. 14(a)	Fig. 14(a)
OFDMA	-	0.045	0.034	0.045
MC-CDMA	-	0.11	0.055/0.253	0.058
IM-MC-CDMA	-	-	0.071/0.123	0.071
SCMA	6.3	1.9	-	1.8
MU-MC-AE	0.103	0.311	0.067	0.071

C. Complexity Comparison

One major issue of current advanced MC systems is the heavy computational burden of decoding data, which is considered as the penalty for their performance improvement. For this, we investigate the decoding complexity of the proposed DL-based MC schemes in terms of the decoding runtime per sample and compare with baselines. We convert the trained model of the proposed schemes from Tensorflow to MATLAB in order to compute its runtime. The runtimes of baselines are also evaluated on MATLAB of the same computer to ensure a fair comparison. Note that the transmitter complexity of our schemes is negligible compared with that of the receiver since the proposed encoders require only one linear FC layer.

Table VI compares the decoding complexity between the proposed (Emb-)MC-AE and baselines corresponding to four figures in Section V.A, where the runtime is measured in millisecond (ms). We recall that OFDM [1], OFDM-IM [2], DM-OFDM [8] employ the ML detector in all considered figures, while S-OFDM [7] and S-OFDM-IM [6] use both the ML and MMSE-based detectors in Fig. 8 and MMSE-based detector only in Fig. 9. Accordingly, the first and second values of the cells that have two values correspond to the runtimes of low-complexity and ML detectors, respectively. As observed from Table VI, our scheme requires much less runtime than almost all baselines. For example, in Fig. 8, the runtime of MC-AE is 0.029 ms which is even lower than that of S-OFDM and S-OFDM-IM with MMSE-based detectors with 0.036 and 0.055 ms, respectively. Moreover, in contrast to MMSE-based counterparts, the decoder of MC-AE does not require the noise level in the detection process. Such observations clearly show that our proposed scheme benefits from not only higher reliability but also lower complexity than current schemes.

We also demonstrate the receiver complexity of MU-MC-AE and baseline schemes in Table VII, which is associated with the four figures in Section V.B. As seen from Table VII, the proposed scheme has much lower runtime than SCMA with the MPA detector [11] or MC-CDMA [9] and IM-MC-CDMA [10] with the ML detector. For instance, in Fig. 12(b),

MU-MC-AE has a runtime of 0.103 ms which is 60 times less than SCMA with 6.3 ms. Interestingly, despite requiring less runtime, MU-MC-AE still achieves a better BLER than SCMA as shown in Fig. 12(b). Besides, the runtime of MU-MC-AE is comparable to that of baselines with their low-complexity detectors, such as MC-CDMA and IM-MC-CDMA.

Finally, to get a better insight into the decoding complexity of the proposed schemes compared with that of the baselines, we analyze the number of floating point operations (flops) for each decoder. In particular, the number of flops required by the MC-AE decoder in Fig. 2 is approximated by $\mathcal{O}\left(4NQ + QM\right) \sim \mathcal{O}\left(QM\right)$, since M and Q are generally much larger than N, while that of the ML and MMSE decoders as the baselines are $\mathcal{O}\left(11NM\right)$ and $\mathcal{O}\left(8N^2 + 5M\right)$, respectively [6]. Hence, it can be seen that when Q is large, the decoder in MC-AE may require more flops than the ML and MMSE-based decoders. However, the runtime of our scheme is still less than that of the baselines, as shown via Table IV. This is because the DNN decoder does not involve any iterations such as "for" or "while" loops as in the baseline decoders. The same observation can be seen for multiuser scenarios.

VI. Conclusion

We proposed a novel DL-based MC-AE system, which is capable of learning both the encoder and decoder to optimize the diversity and coding gains over fading channels in a fully data-driven manner. The BLER analysis clearly showed higher performance gains achieved by MC-AE over hand-crafted baselines including various recent IM-based schemes. We then proposed two novel DNN structures for uplink and downlink MU-MC-AE, along with a new loss function that ensures both the user fairness and the fast training convergence. Simulation results showed that our schemes outperform a wide range of current hand-crafted schemes under both AWGN and Rayleigh fading channels, while still yielding comparable or lower decoding complexity. However, under fading channels, the performance of MU-MC-AE degrades when transmissions are highly overloaded, especially during the uplink transmission. This issue will be left for our future research.

References

[1] T. Hwang, C. Yang, G. Wu, S. Li, and G. Y. Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.

[2] E. Basar, U. Aygolu, E. Panayirci, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.

[3] T. V. Luong and Y. Ko, "Impact of CSI uncertainty on MCIK-OFDM: Tight, closed-form symbol error probability analysis," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1272–1279, Feb. 2018.

[4] E. Basar, "OFDM with index modulation using coordinate interleaving," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 381–384, Aug. 2015.

[5] T. V. Luong, Y. Ko, and J. Choi, "Repeated MCIK-OFDM with enhanced transmit diversity under CSI uncertainty," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4079–4088, June 2018.

[6] T. V. Luong and Y. Ko, "Spread OFDM-IM with precoding matrix and low-complexity detection designs," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 619–11 626, Dec. 2018.

[7] A. Bury, J. Egle, and J. Lindner, "Diversity comparison of spreading transforms for multicarrier spread spectrum transmission," *IEEE Trans. Commun.*, vol. 51, no. 5, pp. 774–781, May 2003.

[8] T. Mao, Z. Wang, Q. Wang, S. Chen, and L. Hanzo, "Dual-mode index modulation aided OFDM," *IEEE Access*, vol. 5, pp. 50–60, 2017.

[9] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, vol. 35, no. 12, pp. 126–133, Dec. 1997.

[10] Q. Li, M. Wen, E. Basar, and F. Chen, "Index modulated OFDM spread spectrum," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2360–2374, Apr. 2018.

[11] H. Nikopour and H. Baligh, "Sparse code multiple access," in *Proc. IEEE PIMRC*, Sep. 2013, pp. 332–336.

[12] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.

[13] H. Ye, G. Y. Li, and B. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[14] T. V. Luong, Y. Ko, N. A. Vien, D. H. N. Nguyen, and M. Matthaiou, "Deep learning-based detector for OFDM-IM," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1159–1162, Aug. 2019.

[15] M. Kim, W. Lee, and D. Cho, "A novel PAPR reduction scheme for OFDM system based on deep learning," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 510–513, Mar. 2018.

[16] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[17] A. Felix, S. Cammerer, S. Dorner, J. Hoydis, and S. T. Brink, "OFDM-autoencoder for end-to-end learning of communications systems," in *Proc. IEEE SPAWC*, June 2018, pp. 1–5.

[18] M. Kim, N. Kim, W. Lee, and D. Cho, "Deep learning-aided SCMA," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 720–723, Apr. 2018.

[19] E. Basar, M. Wen, R. Mesleh, M. D. Renzo, Y. Xiao, and H. Haas, "Index modulation techniques for next-generation wireless networks," *IEEE Access*, vol. 5, pp. 16 693–16 746, 2017.

[20] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *ICLR*, 2014.

[21] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. AISTATS*, 2010, pp. 249–256.

[22] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. [Online]. Available: https://www.tensorflow.org/

[23] S. Cammerer, F. A. Aoudia, S. Dorner, M. Stark, J. Hoydis, and S. ten Brink, "Trainable communication systems: Concepts and prototype," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5489–5503, Sep. 2020.

[24] M. K. Simon and M. S. Alouini, *Digital Communication over Fading Channels*. 2nd edition. John & Wiley, 2005.

[25] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, "Comparison study of non-orthogonal multiple access schemes for 5G," in *Proc. IEEE Int.Symp. BMSB*, June 2015, pp. 1–5.