Novel Multilayer Extreme Learning Machine as a Massive MIMO Receiver for Millimeter Wave Communications

DIEGO FERNANDO CARRERA[1], (Member, IEEE),
CESAR VARGAS-ROSALES[2], (Senior Member, IEEE),
DAVID ZABALA-BLANCO[3], (Member, IEEE), NOE M. YUNGAICELA-NAULA[2],
CESAR A. AZURDIA-MEZA[4], (Member, IEEE), MOHAMED MAREY[5], (Senior Member, IEEE),
AND ALI DEHGHAN FIROOZABADI[6]

[1]School of Technologies, Universidad Tecnológica Empresarial de Guayaquil, Guayaquil 090511, Ecuador
[2]School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey 64849, Mexico
[3]Department of Computer Science and Industry, Universidad Catolica del Maule, Talca 3466706, Chile
[4]Department of Electrical Engineering, Universidad de Chile, Santiago 8370451, Chile
[5]Smart Systems Engineering Laboratory, College of Engineering, Prince Sultan University, Riyadh 11586, Saudi Arabia
[6]Department of Electricity, Universidad Tecnologica Metropolitana, Santiago 7800002, Chile

Corresponding author: Diego Fernando Carrera (dfcarrera@ieee.org)

ABSTRACT Wireless communication systems working in millimeter-wave (mmWave) frequency bands offer higher bandwiths than traditional radio frequency schemes. This technology allows multibeam steering and data multiplexing with the help of massive multiple-input multiple-output (MIMO) systems. However, supporting large bandwiths at mmWave frequencies is challenging due to the use of large antenna arrays with beamforming, sampling signals with large bandwiths, and baseband signal processing operations at gigabit data rates. Due to the wider bandwidth and higher signal processing requirements of mmWave systems, low-complexity receiver algorithms become important. Previously reported investigations assumed the use of hybrid beamforming structures that reduce power consumption and signal processing tasks. Therefore, the use of artificial neural networks (ANNs) becomes relevant for the processing of mmWave signals as reported in earlier works. In this article, to carry out MIMO combining processing for mmWave communications, we propose a fully complex multilayer extreme learning machine (M-ELM) neural network. We investigate the tuning of the number of neurons in each hidden layer for the proposed method to maximize the system performance and decrease the complexity of the receiver. We compare the results of the introduced M-ELM algorithm with a fully complex extreme learning machine (ELM), fully real ELM, and M-ELM defined in the real plane in terms of spectral efficiency, bit error rate, computational complexity, and processing time. Furthermore, we compare the novel M-ELM strategy with traditional linear MIMO receivers, such as Maximum Ratio and Minimum Mean Square Error, as well as to a multilayer perceptron (MLP) neural network trained offline. The numerical results show that with a good balance between the overall performance and computational cost of the ANN, the fully complex M-ELM MIMO receiver outperforms the other evaluated schemes.

INDEX TERMS 5G NR, beamforming, machine learning, massive MIMO, millimeter wave, multilayer ELM.

I. INTRODUCTION

Millimeter (mmWave) communication systems have received wide public attention as part of the development of the fifth generation (5G) New Radio (NR) technology, which provides low latency, more bandwidth, and higher data rates [1]. The mmWave massive multiple input, multiple output (MIMO) system can enhance the achievable Spectral Efficiency (SE) by providing a large array gain. However,

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Liu.

the high-resolution Analog-to-Digital (ADC) and Digital-to-Analog (DAC) converters and the fully digital precoding scheme result in high power consumption and unaffordable hardware costs for these systems [2]. Furthermore, due to the baseband signal processing at gigabit data rates, support for large bandwidths at mmWave frequencies is challenging. However, due to short wavelengths, mmWave communications suffer from fundamental technical challenges, such as severe path loss, blockage sensitivity, directivity, energy efficiency, and narrow beamwidth [3]. Essentially, Uplink Transmission (UL) in the mmWave MIMO system presents these challenges because of the low transmission power of mobile devices. Another related problem in UL communication is interference between users, which reduces the SE of the transmission. Therefore, it is relevant to obtain a solution that maximizes the SE for mmWave massive MIMO systems with low computational complexity.

Typical signal processing methods, such as the maximum ratio (MR) and minimum mean square error (MMSE) algorithms used in MIMO receivers, are not optimized to reduce computational complexity and communications latency. Nevertheless, Machine Learning (ML) and Deep Learning (DL) techniques can effectively address these challenges [4]–[8]. ML and DL methods leverage prior observations about channel estimation and precoding designs to solve the MIMO combining task for mmWave systems, significantly reducing the processing overhead. However, and as stated in [9], not all DL techniques can outperform known optimal solutions for wireless communications. As reviewed in [10], a DL technique like a trained Multilayer Perceptron (MLP) neural network cannot outperform algorithms like the MMSE receiver, from which the neural network learns the MIMO combining process, since the bit error rate (BER) is the same for both techniques. Thus, [9], [10], ML and DL may present improvements compared to known algorithms in communications, particularly in reducing the computational complexity, such as the extreme learning strategy described in [11]. Additionally, neural networks can effectively reduce the effects of non-linear hardware, interference, and frequency-selective channels [9].

Since channel estimation is relevant in conventional receiver designs, ML and DL algorithms are used to improve channel estimation efficiency and communication quality [12]–[14]. In addition, many studies have investigated the issue of joint channel estimation and signal detection [15]. Indeed, channel estimation is not an isolated process since channel equalization requires precise Channel State Information (CSI) to effectively remove signal distortions [16]. As a result, channel estimation and channel equalization can be performed together, and even symbol detection and channel decoding can be done together in one step. However, symbol detection using DL methods showed lower performance than with regular methods [17]. Furthermore, processing time should be considered, as several methods can achieve good wireless performance at the expense of more

computational complexity. Even DL techniques can take a lot of computational resources to learn specific tasks [15], [18].

Another important aspect of the ML and DL methods is that good data sets are required. In this study, the data set becomes the massive MIMO mmWave channels. Channel modeling is a hard task since there is the need to predict large- and small-scale propagation channel parameters with a limited amount of measurement data, especially for mmWave systems [10], [19], [20]. Several studies validate that ELM-based algorithms outperform deep learning neural networks in terms of computational cost for classification or regression issues [21]–[23].

In this paper, we propose a fully complex multilayer extreme learning machine (M-ELM) algorithm in order to learn the MIMO combining process to improve the SE of mmWave communications and minimize the use of computational resources.

In contrast to conventional MIMO receivers such as MR and MMSE, the M-ELM strategy does not require the CSI through channel estimation, which simplifies processing time, particularly useful for mmWave communications. This investigation introduces an M-ELM that learns feature representations by employing singular values based on autoencoders (AEs), strictly defined in the complex domain.

A. RELATED WORK

The Extreme Learning Machine (ELM), introduced by Huang in 2004 [24], comes to be a novel algorithm for optimizing the Single-Hidden-Layer Feedforward Networks (SLFN) and covers the disadvantages of the gradient descent method. In the ELM algorithm, the input weights and bias of the hidden layer of the SLFN are arbitrarily derived without several iterations to obtain the optimal value and remain unchanged during the training process [25]. In ELM approaches, the hidden layer does not need to be adjusted and is highlighted to solve the problem of generalization, since the global optimum is theoretically displayed. Moreover, setting hyperparameters takes a lot of time. In addition, the ELM is widely used in many technological areas, especially in telecommunications signal reception, due to the following: (i) Easy implementation, (ii) extremely fast training speed, and (iii) good generalization performance [11], [16], [26]–[28].

The work in [24] is the first to introduce the ELM concept as an SLFN network with a fast-learning speed thanks to the random adoption of input weights and biases and the minimum norm least-squares (LS) solution of the SLFN. The authors conclude that the standard ELM can be used generally in many cases, and this is the case for the work in [29] which investigates the problem of jointly solving equalization and symbol detection in orthogonal frequency division multiplexing (OFDM) systems with quadrature amplitude modulation (QAM). This study introduces the performance of the standard ELM with fast training speeds compared to other based methods. However, research is limited to single-input single-output (SISO) systems, as in works [10], [30], [31].

Nevertheless, the standard ELM requires setting a minimum number of neurons in the single hidden layer to achieve good performance. Therefore, for SISO systems, an ELM strategy can achieve good results, but processing time should also be considered [10].

Massive MIMO systems are a fundamental technology used for beamforming in mmWave systems. The work in [27] introduces an ELM receiver for MIMO light-emitting diode (LED) communications with nonlinearities and cross-LED interference. The work in [28] shows the performance of a standard ELM used as a MIMO-OFDM receiver, considering low-resolution ADCs. The investigation shows a MIMO combining process learned online with the help of reference signals. However, this study is limited to single-user scenarios. The work in [26] considers multiuser scenarios where the inter-user interference is the limiting factor. Finally, the studies in [26], [28] are limited to low-band frequencies.

With traditional methods such as MR and MMSE, it is necessary to estimate the MIMO channel and then find the MIMO combiner that performs channel equalization [16], [32]–[34]. On the contrary, with an ELM strategy, there is no need to perform channel estimation, as MIMO combining processing can be performed directly with this neural network as shown in [26] for low-band frequencies and in [11] for mmWave frequencies. However, these studies are limited to the single-layer ELM method. The work in [10] compares the performance of the ELM method with linear MIMO receivers and a DL approach called multilayer perceptron (MLP) considering mmWave band frequencies. However, the study is limited to SISO systems and a single-layer ELM strategy.

Since deep neural networks (DNNs) only work with real numbers, there is a need to separate the complex input for a DNN into real and imaginary parts. This process represents a limitation for DNNs, but not for an ELM strategy. The advantage of the latter is the fast-learning speed and simplicity since this neural network can work with complex input numbers such as the received OFDM symbols in a 5G mobile communication. Since an OFDM communication system requires processing complex numbers, an ELM strategy is applied directly to learn the MIMO processing task without requiring a real domain input. However, the investigation in [35] takes into account the standard ELM in the real domain.

An M-ELM network not only inherits the characteristics of the hyperparameters of a single layer ELM [36] but also achieves better performance [37]. However, previous studies outline the performance of real domain multi-layer ELM networks such as the works in [25], [38]–[40] for tasks that are not related to MIMO processing. In this article, we consider a complex M-ELM neural network to solve the MIMO combining problem. We compare the performance of this complex domain network to the real domain M-ELM, especially in terms of processing time.

B. RESEARCH CONTRIBUTIONS
The contributions of this work are summarized as follows:

1) We have designed an M-ELM network to perform MIMO processing of received signals with reduced computational complexity compared to a single-layer ELM. We have considered a single-cell, multiuser mmWave system, and no perfect CSI at the receiver. Specifically, there is no channel estimation process performed with the M-ELM strategy.

2) We have provided an analysis of the introduced M-ELM method by tuning the hyperparameters of the network, such as the number of hidden layers and the number of neurons per layer, to perform MIMO combining processing in one step, in order to reduce the processing time compared to a single-layer strategy.

3) We have tested the performance of the novel M-ELM receiver in terms of the achievable SE, bit error rate (BER), and average processing time. The performance of the M-ELM strategy was compared to that of the single-layer ELM method for the complex and real domains, as well as the MR and MMSE MIMO receivers.

The remainder of the paper is organized as follows. Section II presents the signal and channel model, as well as the channel estimation and equalization processes. The proposed M-ELM algorithm is introduced in Section III. The simulation results are provided in Section IV. Section V presents an analysis and discussion of the results. Finally, Section VI outlines the conclusions of this investigation.

Notation: Scalars are denoted in the lower case, whereas matrices and vectors are represented in bold upper and lower case, respectively. For any general matrix or vector, $\mathbf{x}^{\mathrm{T}}$ represents the transpose and $\mathbf{x}^*$ the conjugate transpose. $\mathbf{I}$ is an identity matrix of proper dimension. $\|.\|_F$ represents the Frobenius norm operator. $\mathbb{E}[.]$, $\mathbb{C}$, and $\mathbb{N}$ denotes the expected value, and the set of complex and natural numbers, respectively. Finally, a circular symmetric complex Gaussian stochastic vector is written as $\mathbf{x} \sim \mathcal{CN}\left(\mu_\mathbf{x}, \sigma_\mathbf{x}^2\right)$ with mean $\mu_\mathbf{x}$ and variance $\sigma_\mathbf{x}^2$.

II. SYSTEM DESCRIPTION
In this section, we present the multiuser mmWave signal and channel model, as well as the channel estimation and equalization processes considered in this work.

A. SIGNAL MODEL
We considered a single cell multiuser hybrid beamforming mmWave system based on the 5G NR standard. The simulation of radio links, based on cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) communication, employs s subcarriers to transmit N_s data symbols per link. The radio link consists of a base station equipped with a massive array of antennas, which communicates with multi-antenna user equipment (UE), as is illustrated in Fig. 1 [10].

Fig. 1 shows that random bits (data bitstream) are mapped to data symbols modulated with the quadrature phase shift
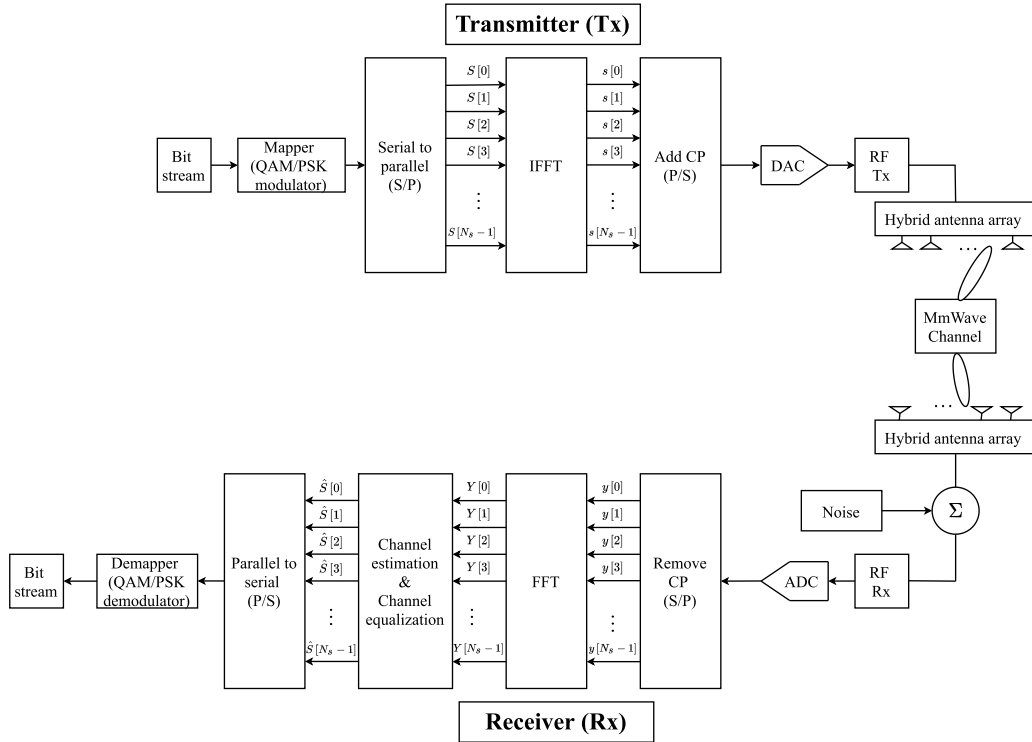
FIGURE 1. Simplified hardware block diagram of the CP-OFDM mmWave system.

keying (QPSK) or quadrature amplitude modulation (QAM) schemes. The data symbols must be multiplexed with CP-OFDM to perform a wideband transmission through the mmWave channel. The multiantenna receiver performs symbol demodulation to recover the original data. Therefore, for this study, we simulate OFDM symbols which are complex numbers that are configured following the 5G NR specifications [1], [41].

OFDM modulation offers high data rate transmission over a multipath fading channel, high SE, and low-complexity implementation due to the Fast Fourier Transform (FFT) algorithm [42]. This transmission technique can also scale the number of subcarriers, so the FFT size scales such that processing complexity does not increase unnecessarily for larger bandwidths. Furthermore, OFDM can work with massive MIMO systems to achieve high antenna diversity and spatial multiplexing [11], [43].

The hardware architecture shown in Fig. 2 represents a fully connected hybrid beamforming BS at the receiver side. At BS, there are L_r radio-frequency (RF) chains, whereas there are K multiantenna users, each equipped with a single RF chain [34], [44], [45].

As shown in Fig. 2, K users can transmit simultaneously to the BS by applying baseband precoding, $\mathbf{f}_{BB_k}$ for user $k = 1, 2, \ldots, K$, i.e., $\mathbf{F}_{BB_k} = [\mathbf{f}_{BB_1}, \mathbf{f}_{BB_2}, \ldots, \mathbf{f}_{BB_K}]$ to transmit their respective signals followed by an RF precoder $\mathbf{F}_{RF_k}$ using analog circuitry (phase shifters). The CP-OFDM multi-carrier scheme suppresses intersymbol interference by

mapping modulated symbols into N_s subcarriers, in this way the transmitted signal k is precoded as $\mathbf{F}_{RF_k}\mathbf{f}_{BB_k}[s]x_k[s]$ $\forall$ s $=$ $0, 1, 2, \ldots N_s - 1$, where the transmitted symbol of the kth user in subcarrier s is represented as $x_k[s]$ [34], [44].

Assuming that all users transmit simultaneously in the UL to the BS, the sth subcarrier reference signal $\mathbf{y}_k^\phi[s] \in \mathbb{C}^{N_r}$ received in the BS and transmitted by the kth user is given by

$$\mathbf{y}_k^\phi[s] = \mathbf{W}_{RF_k}^* \mathbf{H}_k[s] \sum_{k=1}^{K} \mathbf{F}_{RF_k}\mathbf{f}_{BB_k}[s]\phi_k[s] + \mathbf{v}_k[s],$$

(1)

where $\mathbf{W}_{RF_k}^*$ is the analog combining matrix for the kth user at the mmWave BS, $\mathbf{H}_k[s]$ is the channel matrix between the BS, equipped with $L_r N_r$ antennas and the kth user, each equipped with N_t antennas. $\phi_k[s]$ denotes the kth user transmitted reference symbol mapped at pilot subcarrier s. It is worth noting that not all subcarriers are used for pilot transmission. This is a characteristic of the pilot structure defined in the 5G NR standard, which allows multiuser multiplexing [41]. Finally, $\mathbf{v}_k[s]$ represents the circularly symmetric white Gaussian noise vector characterized as i.i.d $\mathcal{CN}(0, \sigma_k^2\mathbf{I})$ [34], [45].

On the other hand, the kth received data signal at the BS, $\mathbf{y}_k^x[s] \in \mathbb{C}^{N_r}$, for a frequency domain block fading channel is
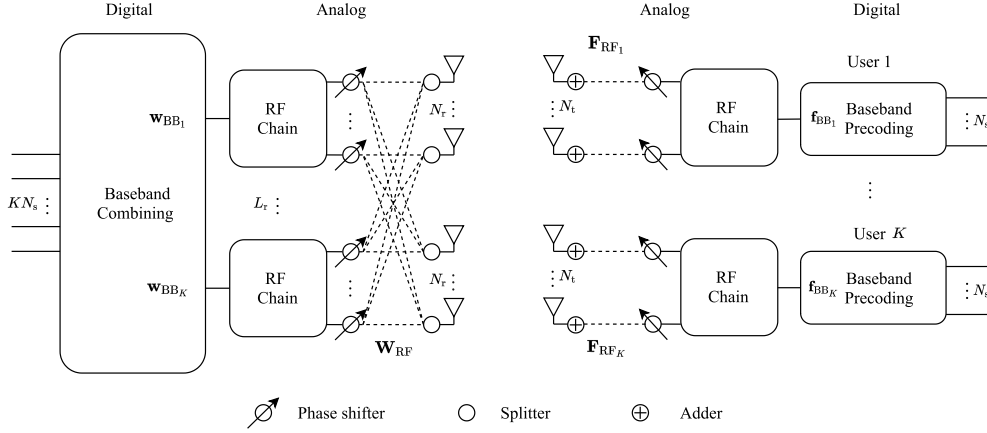
FIGURE 2. Simplified system blocks of an UL multiuser mmWave system with a fully-connected hybrid array of antennas.

written as

$$\mathbf{y}_k^x[s] = \mathbf{W}_{\mathrm{RF}_k}^* \mathbf{H}_k[s] \sum_{k=1}^{K} \mathbf{F}_{\mathrm{RF}_k} \mathbf{f}_{\mathrm{BB}_k}[s] x_k[s] + \mathbf{v}_k[s] \in \mathbb{C}^{N_{\mathrm{r}}}.$$

$$(2)$$

B. CHANNEL MODEL

Since we considered a massive MIMO mmWave system, the high path loss limits spatial selectivity, and the hybrid structure at the mmWave wavelength assumes the use of a tightly packet antenna array, which leads to high antenna correlation [44]. Therefore, the mmWave massive MIMO channel is assumed to be frequency selective and is represented by the clustered channel model. The dth delay tap of the kth user discrete-time $N_{\mathrm{r}} \times N_{\mathrm{t}}$ channel matrix $\mathbf{H}_{d,k}$, for $d = 1, 2, \ldots, N_{\mathrm{c}}$, assumed as the summation of N_{cl} scattering clusters, each contributing N_{ray} propagation paths [45], is written as

$$\mathbf{H}_{d,k} = \lambda \sum_{\eta=1}^{N_{\mathrm{cl}}} \sum_{\iota=1}^{N_{\mathrm{ray}}} \alpha_{\eta\iota,k} \mathbf{a}_{\mathrm{r}} \left(\theta_{\eta\iota,k}^{\mathrm{r}}, \phi_{\eta\iota,k}^{\mathrm{r}} \right) \mathbf{a}_{\mathrm{t}}^* \left(\theta_{\eta\iota,k}^{\mathrm{t}}, \phi_{\eta\iota,k}^{\mathrm{t}} \right),$$

$$(3)$$

where $\lambda = \sqrt{\frac{N_{\mathrm{r}} N_{\mathrm{t}}}{N_{\mathrm{cl}} N_{\mathrm{ray}}}}$ is the scalar normalization factor, and $\alpha_{\eta\iota,k}$ is the complex small-scale fading gain of the ιth ray in the ηth scattering cluster, characterized as i.i.d. $\mathcal{CN}\left(0, \sigma_\eta^2\right)$, where σ_η^2 denotes the average power of the ηth cluster.

The vectors $\mathbf{a}_{\mathrm{r}} \left(\theta_{\eta\iota,k}^{\mathrm{r}}, \phi_{\eta\iota,k}^{\mathrm{r}} \right)$ and $\mathbf{a}_{\mathrm{t}}^* \left(\theta_{\eta\iota,k}^{\mathrm{t}}, \phi_{\eta\iota,k}^{\mathrm{t}} \right)$ denote the array response functions, for the receive and transmit antenna arrays of the kth user, with respect to the angles of arrival $\phi_{\eta\iota}^{\mathrm{r}} \left(\theta_{\eta\iota,k}^{\mathrm{r}} \right)$ and departure $\phi_{\eta\iota,k}^{\mathrm{t}} \left(\theta_{\eta\iota,k}^{\mathrm{t}} \right)$, respectively, at an azimuth θ and an elevation angle ϕ [46]. Finally, based on [45], [47], the channel frequency response of the user k in subcarrier s, from N_c delay taps in the discrete time domain

is given by

$$\mathbf{H}_k[s] = \sum_{d=1}^{N_{\mathrm{c}}} \mathbf{H}_{d,k} e^{-j\frac{2\pi n}{N_{\mathrm{s}}}d} \quad \in \mathbb{C}^{N_r \times N_t}$$

$$(4)$$

C. CHANNEL ESTIMATION

To estimate the mmWave channel, we first define an effective channel that is equivalent to the convolution of the hybrid beamformer and the channel response in the frequency domain, written as

$$\mathbf{h}_{\mathrm{eff}_k}[s] = \mathbf{W}_{\mathrm{RF}_k}^* \mathbf{H}_k[s] \mathbf{F}_{\mathrm{RF}_k} \mathbf{f}_{\mathrm{BB}_k}[s] \in \mathbb{C}^{N_r}.$$

$$(5)$$

By transmitting reference symbols on variously known pilot subcarriers, CSI estimation can be performed with different channel estimation methods [48]. A simple estimation technique is the least squares (LS) algorithm. Based on the evidence presented in [49], the LS estimation can result in an imprecise measurement of the CSI when noise and interference are high, given that this technique does not consider the correlation properties of the channel [16]. The LS estimator is given by

$$\hat{\mathbf{h}}_{\mathrm{eff}_k}[s] = \hat{\mathbf{h}}_{\mathrm{eff}_k}^{\mathrm{LS}}[s] = \mathbf{y}_k^\phi[s] \, \phi[s]^* \quad \in \mathbb{C}^{N_r}.$$

$$(6)$$

With the LS estimation, we can obtain first- and second-order statistics of channel parameters such as the variance of the effective estimated channel vector $\sigma_{\mathbf{h}}^2$. These parameters of the mmWave channel must be known a priori to perform MMSE estimation [16], [50], [51], which is written as

$$\hat{\mathbf{h}}_{\mathrm{eff}_k}[s] = \hat{\mathbf{h}}_{\mathrm{eff}_k}^{\mathrm{MMSE}}[s]$$
$$= \sigma_{\mathbf{h}}^2 \phi^*[s] \left(\sigma_{\mathbf{h}}^2 \phi[s] \phi^*[s] + \sigma_k^2 \right)^{-1} \in \mathbb{C}^{N_r},$$

$$(7)$$

where σ_k^2 denotes the kth user noise variance at the receiver.

Consider that the proposed M-ELM strategy does not require performing channel estimation. However, this process is required by linear MIMO receivers, such as the MMSE and

MR receivers, in order to perform channel equalization [52], as described in the next subsection.

D. CHANNEL EQUALIZATION

In this section, we describe two common linear MIMO receivers that are frequently used to perform the channel equalization task at the receiver [32], [33].

MR processing computes the combining vector $\mathbf{w}_{\text{BB}_k}^{\text{MR}}[s]$ with the estimated effective channel at the sth subcarrier of the received OFDM signal [41], $\mathbf{h}_{\text{eff}_k}[s]$, as follows

$$\mathbf{w}_{\text{BB}_k}[s] = \mathbf{w}_{\text{BB}_k}^{\text{MR}}[s] = \frac{1}{\sqrt{\rho_k}} \frac{\hat{\mathbf{h}}_{\text{eff}_k}[s]}{\left\| \hat{\mathbf{h}}_{\text{eff}_k}[s] \right\|_{\text{F}}^2} \in \mathbb{C}^{N_r}. \quad (8)$$

Another MIMO receiver most studied in the literature is the MMSE combiner [16], [26], [32]–[34], [41], [44], [45], [53]. For MMSE processing, we can group the CSI of the K users in the cell of interest in a single matrix as follows

$$\hat{\mathbf{H}} = [\hat{\mathbf{h}}_{\text{eff}_1}[s], \ldots, \hat{\mathbf{h}}_{\text{eff}_k}[s], \ldots, \hat{\mathbf{h}}_{\text{eff}_K}[s]]^{\text{T}} \in \mathbb{C}^{K \times N_r}. \quad (9)$$

According to [41], the MMSE combining vector, $\mathbf{w}_{\text{BB}_k}^{\text{MMSE}}[s]$, can be written as

$$\mathbf{w}_{\text{BB}_k}[s] = \mathbf{w}_{\text{BB}_k}^{\text{MMSE}}[s]$$
$$= \frac{1}{\sqrt{\rho_k}} \left(\hat{\mathbf{H}}^{\text{H}} \hat{\mathbf{H}} + \sigma_k^2 \mathbf{I} \right)^{-1} \hat{\mathbf{h}}_{\text{eff}_k}[s] \in \mathbb{C}^{N_r}. \quad (10)$$

Therefore, the MIMO receiver problem reduces to finding a combining vector that recovers the transmitted data symbols written as

$$\hat{x}_k[s] = \mathbf{w}_{\text{BB}_k}^*[s] \mathbf{y}_k^x[s]$$
$$= \mathbf{w}_{\text{BB}_k}^*[s] \mathbf{h}_{\text{eff}_k}[s] x_k[s] + \sum_{\substack{j=1 \\ j \neq k}}^{K} \sqrt{\mathbf{w}_{\text{BB}_k}^*}[s] \mathbf{h}_{\text{eff}_j} x_j[s]$$
$$+ \mathbf{w}_{\text{BB}_k}^*[s] \mathbf{W}_{\text{RF}_k}^* \mathbf{v}_k[s], \quad (11)$$

where $\mathbf{w}_{\text{BB}_k}^*[s]$ is the kth user baseband (BB) combiner that maximizes the data symbol power, forcing to zero interference signals and nonlinear distortions, whereas $\mathbf{h}_{\text{eff}_j}$ and $x_j[s]$ represent the effective channel and the transmitted symbol of the jth interfering user, respectively.

E. SPECTRAL EFFICIENCY

In this work, we considered the UL bound capacity to get the SE expression. Therefore, as in [52], $\mathbf{C}_k[s]$ is defined as the corresponding estimation error covariance matrix of the kth user effective channel. This way, the kth user SE per subcarrier in the UL is given by

$$\text{SE}_k = \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k \right) \right\}, \quad (12)$$

where $\text{SINR}_k[s]$ represents the signal-to-interference plus noise ratio (SINR) of the kth user, which is written in (13), as shown at the bottom of the next page.

However, SE cannot be calculated in the same way with an ELM method, since there is no estimated channel with this

strategy [10], [26]. Therefore, the SINR can also be computed as

$$\text{SINR}_k[s] = -[\text{PAPR} + 20\log_{10}(\text{EVM}_k[s]/100\%)], \quad (14)$$

where PAPR is given per modulation scheme, i.e., the PAPR value corresponds to 0 dB, 2.6 dB, 3.7 dB, and 4.2 dB for the quadrature phase shift keying (QPSK) modulation, 16-quadrature amplitude modulation (16-QAM), 64-QAM, and 256-QAM, respectively, [26]. The error vector magnitude (EVM) of the sth subcarrier of the user k, ($\text{EVM}_k[s]$), is calculated with the combined symbols in the receiver based on the Euclidean distance between the combined symbols and their perfect constellation points.

III. EXTREME LEARNING MACHINE STRATEGIES

In this section, we introduce the design of a single-layer ELM receiver and the proposed M-ELM (both methods are strictly defined in the complex domain). It is important to mention that the former was initially presented in [26] for multiuser massive MIMO systems, as well as in [10], [11] for mmWave channels simulated at 28 GHz. In this study, the M-ELM receiver will be introduced for mmWave massive MIMO systems.

A. MIMO COMBINING BASED ON AN EXTREME LEARNING MACHINE

The input of the artificial neural network (ANN) is the signal received from the user k, and the output is the combined MIMO symbols. Traditional methods like MR and MMSE require first estimating the MIMO channel and then finding the MIMO combiner that equalizes the channel.

ELM learning can be applied in the complex domain taking as input the received pilot signal matrix $\mathbf{Y}_k^\phi = \{\mathbf{y}_k^\phi[1], \ldots, \mathbf{y}_k^\phi[N_s]\}^{\text{T}}$ and as output the pilot vector $\boldsymbol{\phi}_k = \{\phi_k[1], \ldots, \phi_k[N_s]\}^{\text{T}}$, known a priori in the BS, as illustrated in Fig. 3. The goal of the ELM strategy is to learn the MIMO combining process from the training data (in this case, the pilot symbols) that gives the best prediction when the unknown data symbols share the same modulation as the pilots [11], [26], [54]. By doing so, only the desired signal is maximized, while interference signals and non-linear distortions are attenuated.

A significant advantage of the ELM neural network is that this method can perform MIMO combining online because of the training phase. which consists of the random origination of the hidden neurons, so that the output weights can be found through a regularized least squares process [55]. This means that the ELM network does not require offline training as with ANNs based on the stochastic gradient descent algorithm [26], and DNN strategies [56].

The kth ELM receiver, where the pilot signal matrix $\mathbf{Y}_k^\phi$ is used at the input of the neural network, which consists of L hidden nodes between the input and output layers, is represented by $\boldsymbol{\phi}_k = \mathbf{O}_k^\phi \boldsymbol{\beta}_k$, where $\boldsymbol{\beta}_k \in \mathbb{C}^L$ is the output weight vector between the hidden layer of L nodes to the
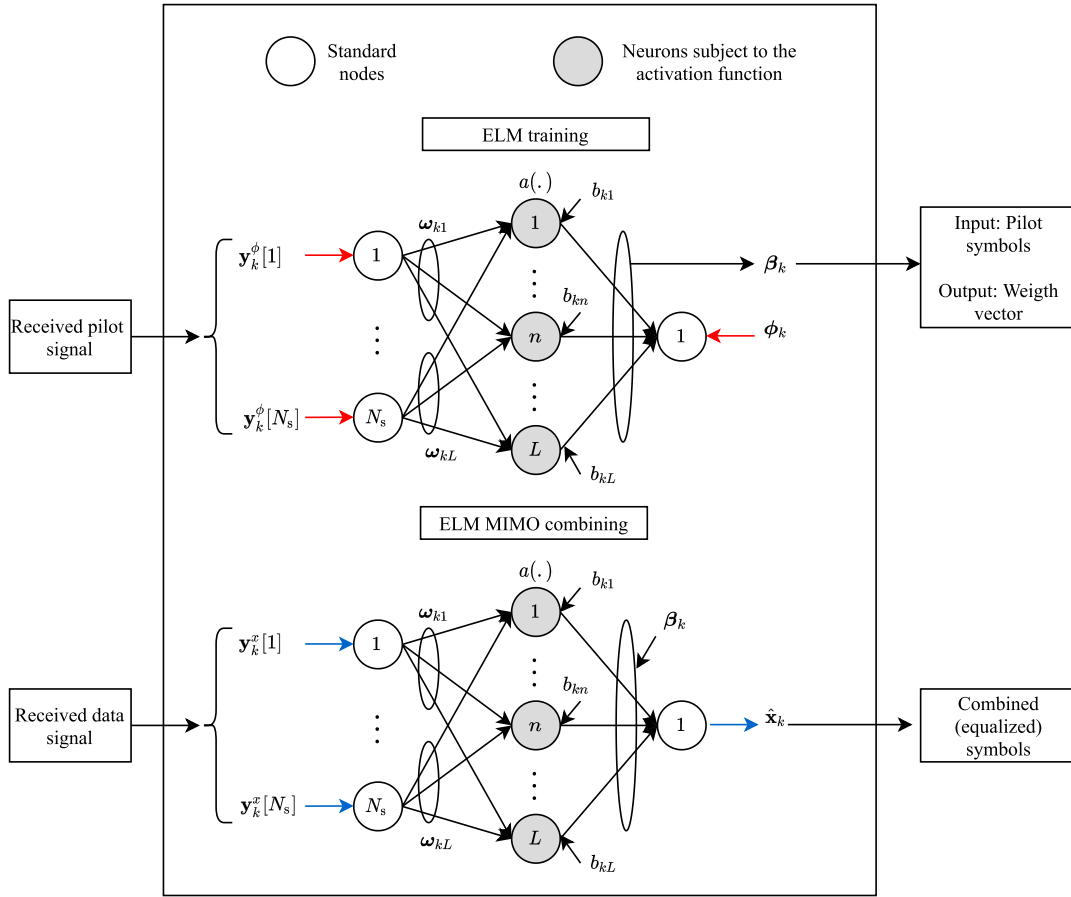
FIGURE 3. Structure of the complex ELM receiver. The received pilot signal $\mathbf{y}_k^\phi[s]$ and the pilots vector ϕ_k are used for online training of the ELM neural network, whereas the received data signals $\mathbf{y}_k^x[s]$ are directly combined into the detected data symbols $\hat{\mathbf{x}}_k$.

single output node, and $\mathbf{O}_k^\phi \in \mathbb{C}^{N_s \times L}$ represents the hidden layer output matrix, and is given by

$$\mathbf{O}_k^\phi = \begin{bmatrix} a(\boldsymbol{\omega}_{k1}^T \mathbf{y}_k^\phi[1] + b_{k1}) & \cdots & a(\boldsymbol{\omega}_{kL}^T \mathbf{y}_k^\phi[1] + b_{kL}) \\ \vdots & \ddots & \vdots \\ a(\boldsymbol{\omega}_{k1}^T \mathbf{y}_k^\phi[N_s] + b_{k1}) & \cdots & a(\boldsymbol{\omega}_{kL}^T \mathbf{y}_k^\phi[N_s] + b_{kL}) \end{bmatrix},$$

(15)

where $a(\cdot)$ is the complex activation function of the hidden layer, which includes circular, inverse circular, hyperbolic, and inverse hyperbolic functions, as is described in [57]. The kth input weights, $\boldsymbol{\omega}_{kn} = [\omega_{kn1}, \ldots, \omega_{knN_r}]^T \in \mathbb{C}^{N_r}$ and the biases b_{kn} of the hidden node n, for $n = 1, \ldots, L$, are randomly initialized and fixed without tuning for the ELM training step. Specifically, the output weight vector when

$N_s > L$ is written as

$$\boldsymbol{\beta}_k = \left(\mathbf{O}_k^{\phi H} \mathbf{O}_k^\phi\right)^{-1} \mathbf{O}_k^{\phi H} \boldsymbol{\phi}_k.$$

(16)

After the ELM receiver is trained, we can use the data signal $\mathbf{Y}_k = \{\mathbf{y}_k[1], \ldots, \mathbf{y}_k[N_s]\}^T$ at the input of the neural network, as is illustrated in Fig. 3. Therefore, a new data output weight vector $\mathbf{O}_k \in \mathbb{C}^{N_s \times L}$ is processed as

$$\mathbf{O}_k = \begin{bmatrix} a(\boldsymbol{\omega}_{k1}^T \mathbf{y}_k^x[1] + b_{k1}) & \cdots & a(\boldsymbol{\omega}_{kL}^T \mathbf{y}_k^x[1] + b_{kL}) \\ \vdots & \ddots & \vdots \\ a(\boldsymbol{\omega}_{k1}^T \mathbf{y}_k^x[N_s] + b_{k1}) & \cdots & a(\boldsymbol{\omega}_{kL}^T \mathbf{y}_k^x[N_s] + b_{kL}) \end{bmatrix}.$$

(17)

$$\text{SINR}_k[s] = \frac{\left|\mathbf{w}_{\text{BB}k}^H[s]\hat{\mathbf{h}}_k[s]\right|^2}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{K} \left|\mathbf{w}_{\text{BB}k}^H[s]\hat{\mathbf{h}}_i[s]\right|^2 + \sum\limits_{k=1}^{K} \mathbf{w}_{\text{BB}k}^H[s]\mathbf{C}_k[s]\mathbf{w}_{\text{BB}k}[s] + \left\|\mathbf{w}_{\text{BB}k}[s]\right\|^2}.$$

(13)

Finally, the last step is to perform MIMO processing with the trained ELM output weight vector as follows:

$$\hat{\mathbf{x}}_k = \mathbf{O}_k \boldsymbol{\beta}_k, \qquad (18)$$

where $\hat{\mathbf{x}}_k = \left\{ \hat{x}_k[1], \ldots, \hat{x}_k[N_s] \right\}^{\mathrm{T}}$ denotes the detected data symbols in the output layer of the ELM network [26], [58]. Note that $\boldsymbol{\omega}_{kn}$ and b_{kn} in (15) are fixed after training with ELM and reused in (17).

The steps for performing MIMO processing with a complex ELM algorithm are summarized below.

1) Set the complex activation function $a(.)$.
2) Randomly choose the complex input weight $\boldsymbol{\omega}_{kn}$ and the complex bias b_{kn}.
3) Calculate the complex hidden layer output matrix $\mathbf{O}_k^{\phi}$ with the received pilot signal matrix $\mathbf{Y}_k^{\phi}$ using (15).
4) Calculate the output weight vector $\boldsymbol{\beta}_k$ using (16).
5) Calculate the complex hidden layer output matrix $\mathbf{O}_k^{\phi}$ with the received data signal matrix $\mathbf{Y}_k$ in (16) using the same input weight and bias as in Step two.
6) Perform MIMO processing with the trained ELM output using (17).

B. MIMO COMBINING BASED ON A MULTILAYER EXTREME LEARNING MACHINE

The characteristics of the data set determine the generalizability of any machine learning algorithm [55]. Therefore, the design of the features that a prominent data structure can represent is relevant. However, this task requires subject matter expertise to identify appropriate characteristics [59]. AEs can perform feature engineering and may be used to train an ANN with several hidden layers [60]. In this sense, an AE represents an unsupervised ANN where the outputs and inputs are the same for the AE by reproducing the input signal as much as possible.

Fully complex AEs are adopted as the basic building block of M-ELMs, whose training architecture is structurally separated into unsupervised hierarchical feature representation and supervised feature classification [25]. While ELM-AE is being developed to obtain multilayer sparse features of the input data for the former stage; for the latter phase, the standard ELM is performed for final decision making. Based on the supervised learning performed by the standard ELM (previous subsection), the training procedure for the unsupervised building blocks (AEs) in the M-ELM architecture is highlighted here.

A single ELM-AE can be modeled by an input layer, a hidden layer, and an output layer, as shown in Fig. 4 [60]. In this subsection, the k index for the ELM explanation is removed for simplicity.

An ELM-AE is characterized by j input layer neurons, n hidden layer neurons, j output layer neurons, and the activation function of hidden neurons $a(\cdot)$ [40]. For N_s different samples $\mathbf{y}^{\phi}[i] \in \mathbb{C}_{N_s} \times \mathbb{C}_j$, $i = 1, \cdots, N_s$, the results of the hidden layer of ELM-AE and the relationship between
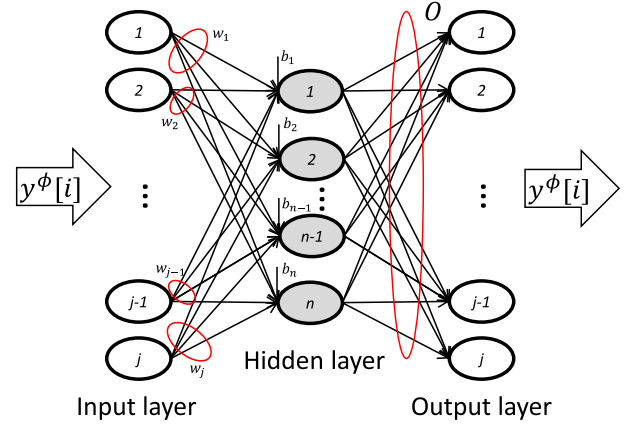

FIGURE 4. The network structure of the ELM-AE.

the outputs of the hidden layer can be written as

$$\mathbf{o} = a(\boldsymbol{\omega} \mathbf{y}^{\phi} + b), \qquad (19)$$

whereas the results of the hidden ELM-AE layer and the results of the output layer can be written as

$$\mathbf{o}(\mathbf{y}^{\phi}[i])\beta = (\mathbf{y}^{\phi}[i])^{\mathrm{H}}, \qquad (20)$$

where w_i denotes the input weight vector by joining the input layer with the hidden layer i, and b_j represents the bias weight for the hidden layer j.

The standard ELM presents the next modification to perform unsupervised learning: Input data is used as output data. An ELM-AE is oriented to properly enact the input features in three diverse forms: (1) compressed, which represents features from a higher-dimensional input data space to a lower-dimensional feature space, (2) sparse, which is the opposite of compressed, and (3) equal dimension where features maintain their dimensionality. The ELM-AE output weights make it possible to switch from feature space to input data. For all ELM-AE representations [25], the output weight acquires the form of

$$\beta = \begin{cases} (\mathbf{O}^{\mathrm{H}} \mathbf{O})^{-1} \mathbf{O}^{\mathrm{H}} \mathbf{Y}^{\phi} & N_s > n \\ \mathbf{O}^{\mathrm{H}} (\mathbf{O} \mathbf{O}^{\mathrm{H}})^{-1} \mathbf{Y}^{\phi} & N_s \leq n \end{cases}, \qquad (21)$$

with $\mathbf{O} = [\mathbf{o}_1, \cdots, \mathbf{o}_{N_s}]$ being the hidden layer outputs from ELM-AE and Y^{ϕ} the input and output data from ELM-AE. Since $\mathbf{O}$ becomes the projected feature space of Y^{ϕ} subject to the activation function, the output weight of ELM-AE can appropriately extract the features of the input data through singular values.

The ELM-AE is stacked layer by layer according to a hierarchical structure. Before the supervised least mean square (LMS) optimization, each weight of the M-ELM hidden layer is initialized utilizing ELM-AE, which performs unsupervised layer learning by excluding random feature mapping. Mathematically, as in [60], the output of every hidden layer is written as follows

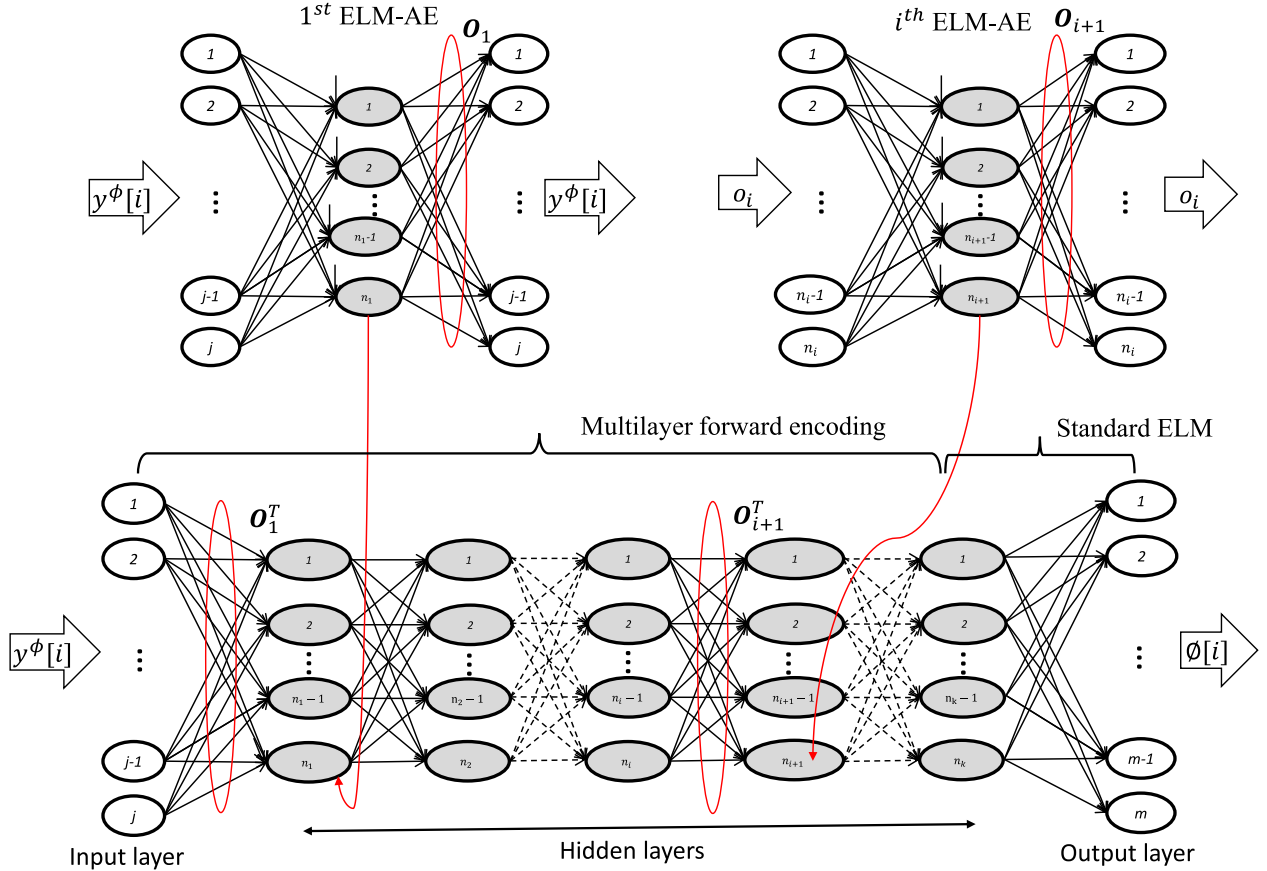$$\mathbf{O}^i = a[(\beta^i)^{\mathrm{H}} \mathbf{O}^{i-1}], \qquad (22)$$

FIGURE 5. Model structure of the M-ELM receiver.

where $\mathbf{O}^i$ represents the ith hidden layer output matrix (the input layer occurs for $i - 1 = 0$; hence, the inputs of M-ELM are given by $\mathbf{O}^{i-1}$). Once the feature of the previously hidden layer is calculated, the weights and hyperparameters (the activation function and the number of hidden neurons) of the current hidden layer are fixed. AE employs the encoded results to address the original inputs by reducing the reconstruction errors.

The output of the connections between the last hidden layer and the output node $\phi[i] \in \mathbb{C}_{N_s} \times \mathbb{C}_m, i = 1, \cdots, N_s$, can be analytically determined with the resolution of a linear system such as for the standard ELM [55]. The M-ELM model is shown in Fig. 5, where $\mathbf{y}^\phi[i]$ denotes the output weights of ELM-AE. The input of ELM-AE is $\mathbf{O}^{i-1}$, and the number of hidden layer nodes of ELM-AE matches the number of hidden nodes i^{th} of M-ELM in the case of the parameters between the hidden layer M-ELM i^{th} and the hidden layer $i - 1$ of M-ELM that ELM-AE learns. Finally, we represent the number of neurons in each of the hidden layers of i as n_i.

IV. RESULTS
In this section, we present the parameters of the propagation environment and the ELM strategies considered in this study,

as well as the simulation results to demonstrate the overall performance of the new M-ELM receiver introduced in Section III-B.

A. PARAMETERS OF THE EVALUATED SCENARIO
The simulation parameters are given in Table 1, based on the NR specifications developed by the Third Generation Partnership Project (3GPP), Release 15 [1], and parameters of the phased array antenna set for 5G radios in the 28 GHz band, as described in [41], [61], [62]. We simulated the channels between the K users and the BS with the quasi-deterministic radio channel generator (QuaDRiGa), which is software coded in Matlab and developed by the Fraunhofer Institute for Telecommunications [63]. Specifically, in QuaDRiGa software, we used the mmMAGIC urban microcellular (UMi) cluster channel model [64], a frequency-selective mmWave channel as the general model described in Section II-B.

In this work, we selected 16-QAM and 64-QAM as pilot and data modulation schemes for simulation. In addition, we did not have to do much data processing before the training module with the M-ELM neural network as this strategy can handle complex data entries. Finally, we used 10 dBm of equivalent isotropically radiated power (EIRP),

TABLE 1. Simulation parameters [1], [63].

| Parameter	Value
BS height	10 m
BS position	0, 0 (x,y coordinates in [m])
Users height	1.5 m for the four users
Users position	random position to the east-side of the BS
Cell radius	250 m
Channel model	mmMagic UMi
Operating band	n257
Carrier frequency	28 GHz
Carrier type	CP-OFDM (120 kHz of subcarrier spacing)
Channel coherence time	20 ms
Carrier bandwidth	100 MHz
EIRP	10 dBm
Occupied subcarriers	792 subcarriers per OFDM symbol
FFT size	1024
NR data slot	12 subcarriers per PRB, 14 OFDM symbols per slot / subframe, 66 PRBs
Channel estimation	LS (for the MR and MMSE receivers only)
MIMO combining	MR, MMSE, ELM, real ELM, M-ELM, and real M-ELM
Total noise power	-86.521 dBm
Fully-connected BS antenna properties [62], [63]	
Number of antenna panels	1
Number of antenna elements per panel	16×16 array with dual polarization ($M = 512$)
Element separation distance	0.55λ
Number of RF chains / beams	8
Peak beam gain (dBi)	27

the gain from the product of the transmitter power, and the antenna array gain.

Fig. 6 shows the fully connected hybrid simulated array composed of a single panel of 16×16 dual-polarized antennas. Each polarized array is connected to four RF chains. This hybrid structure consists of 512 antennas, where 256 antennas have vertical polarization, whereas the other 256 antennas have horizontal polarization [61], [62].

B. PARAMETERS OF THE ELM STRATEGIES
For ELM strategies, pilots must follow the same modulation scheme as data symbols to ensure a successful learning stage and thus maximize wireless performance [26], [35], [65]. The OFDM signals received with S subcarriers per symbol introduced in (1) are the complex numbers that we used as input to train the proposed M-ELM network. That is, with this strategy, we did not have to do extensive data processing prior to the training module. Although, we have used thousands of channel samples to simulate the stochastic behavior of a mobile scenario.

In this study, to explore the advantages and disadvantages of our proposal, we present various versions of the ELM neural network such as ELM, fully real ELM, M-ELM, and fully real M-ELM (M-ELM is strictly defined in the complex domain). In this sense, we present hyperparameters for the different ELM strategies to maximize their generalizability. We adjusted the hyperparameters of the proposed model by exhaustive search, such as the number of neurons per hidden layer, the number of hidden layers, and the activation
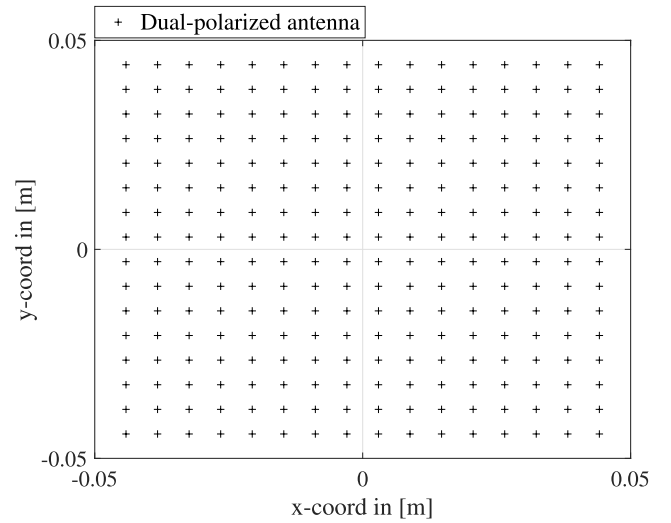
FIGURE 6. Antenna array at the mmWave BS equipped with a fully-connected hybrid structure.

function. We have also validated the proposed algorithm with respect to the achievable spectral efficiency parameter, as described in Section II-E.

For the approaches of the ELM and M-ELM neural networks (complex domain), we set the activation function *tanh* since a smaller BER can be achieved compared to other functions [26], [65]. The weights and biases between the input and hidden layers were generated following a uniform distribution in the interval [-0.005, 0.005], which consists of the activation function's region of convergence (ROC) [26]. For ELM and M-ELM in the real domain, we adopt the *sigmoid* activation function because it shows superiority in terms of computational cost and prediction accuracy in the context of ELM [66]. The hidden neuron parameters were randomly derived based on a uniform distribution defined from -1 to 1 [35]

The ELM hyperparameters that enhance the BER are briefly exposed as

1) **ELM:** According to [26], we set the number of neurons in the single hidden layer as the number of antennas dedicated per user in the BS, therefore $L = 64$.

2) **Real ELM:** The number of hidden neurons and the regularization parameter, which is useful to improve the stability of the ELM, are 120 and 228, respectively. This configuration was performed by looking at the performance of the system in terms of these hyperparameters for various levels of SNR. Note that this ELM must work with strictly real information (a single constellation symbol is decomposed in its in-phase and quadrature components). Consequently, the input layer of a real ELM strategy is composed of twice the number of antennas at the BS while it has two output neurons (one for each component of the constellation symbol). More details about the real ELM as an equalizer can be found in [35], where this strategy
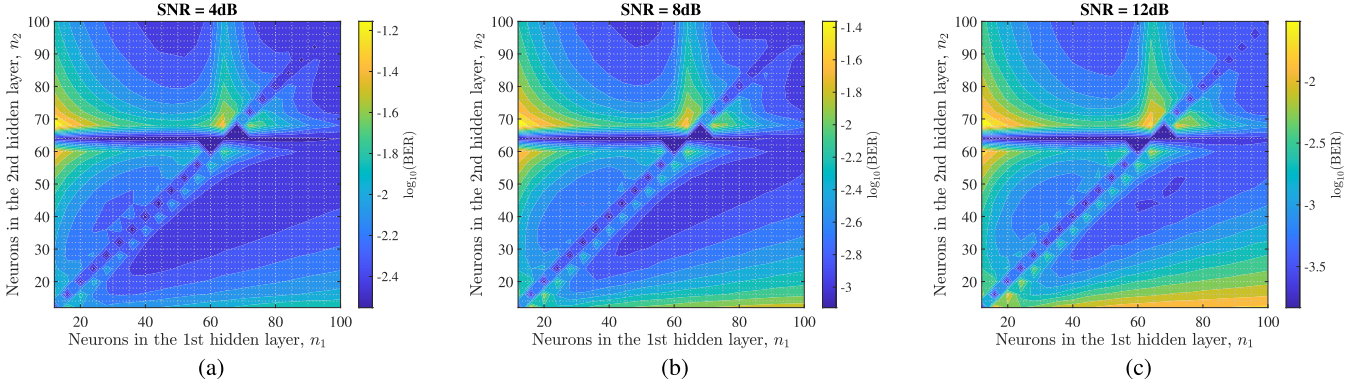
—

FIGURE 7. M-ELM hyperparameters optimization with fixed SNRs of (a) 4 dB, (b) 8 dB, (c) and 12 dB.

TABLE 2. Log_{10}(BER) numerical results for the proposed complex M-ELM receiver when the number of neurons in the first and second layer matches.

SNR	Neurons in the first (n_1) and second (n_2) hidden layers log_{10}(BER)								
	16	**24**	**32**	**40**	**48**	**56**	**64**	**72**	**80**
4 dB	-2.4864	-2.5219	-2.5401	-2.5448	-2.5472	-2.5471	-2.5519	-2.5219	-2.4962
8 dB	-3.0697	-3.1074	-3.1192	-3.1227	-3.1262	-3.1209	-3.1263	-3.0705	-3.0327
12 dB	-3.7928	-3.8076	-3.8268	-3.8242	-3.8073	-3.8209	-3.8257	-3.7268	-3.6509
Average	**-3.1163**	**-3.1456**	**-3.1620**	**-3.1639**	**-3.1602**	**-3.1629**	**-3.1679**	**-3.1064**	**-3.0599**

is introduced for diminishing the laser phase noise in coherent optical OFDM systems (SISO systems).

3) **M-ELM:** We set 16 neurons in each of the two hidden layers ($n_1 = n_2 = 16$). This adoption comes from the performance results presented in Section IV-C, as well as to reduce the computational cost of the proposed M-ELM receiver.

4) **Real M-ELM:** Based on an exhaustive optimization procedure, the number of neurons in the first and second hidden layers corresponds to 80 and 40, respectively. The three regularization parameters (one between each layer of the ELM) are set to 0 for the sake of simplicity of equalization. Since this ELM comes to be a modified/improved version of the real ELM [35], it also possesses 128 input neurons (a BS equipped with $Nr = 64$ receiving antennas) and two neurons defined in the real domain.

C. NUMERICAL RESULTS

In Fig. 7, we probed the performance of the proposed complex M-ELM receiver, varying the number of neurons in the first and second hidden layers. In this way, we tested the appropriate number of neurons to be configured in each of the hidden M-ELM layers that result in a lower BER.

As shown in Fig. 7, for the proposed M-ELM strategy, when the number of neurons in the first and second layers coincides, the achievable BER is smaller. These results reveal that it is not necessary to set a high number of neurons in each of the hidden layers of the M-ELM receiver since only a few neurons are needed to decrease the BER and enhance the SE. In general, the compressed representation (from a higher

dimensional input data space to a lower-dimensional feature space) is superior to the sparse representation in terms of the BER metric. Furthermore, in Fig. 7, we can see a horizontal line along the first layer of the M-ELM axis. This result reveals that with 64 neurons ($L = N_r$) in the first hidden layer, we can achieve a low BER. Therefore, we can achieve the same results with one hidden layer as with two hidden layers. However, with fewer neurons and more layers, the size of the matrices in each layer is smaller, so the processing task is easier.

Furthermore, in Table 2 we present the numerical results for BER achieved with different sets of neurons in the first and second layers of the proposed M-ELM strategy. These results correspond to Log_{10}BER with SNRs of 4dB, 8dB, and 12 dB, illustrated in Fig. 7.

Based on the data presented in Table 2, we can see that, on average, the combination of neurons in the hidden layers of the M-ELM receiver presents the same results. Furthermore, we present the SE and BER results in Fig. 8, with pairs of 16, 20, and 24 neurons configured in the first and second layers of the complex M-ELM strategy.

As shown in Fig. 8, there is no need to define a large number of neurons in each of the hidden layers of the proposed M-ELM receiver. With few neurons, we can achieve SE and BER results similar to those of a high number of neurons. Consequently, we set 16 neurons in each of the hidden layers of the M-ELM strategy to reduce the complexity of signal processing.

Second, to test the performance of the different MIMO combining techniques presented in this study, in Fig. 9, we show the average SE achieved by the kth user on the network with a 16-QAM modulation scheme.
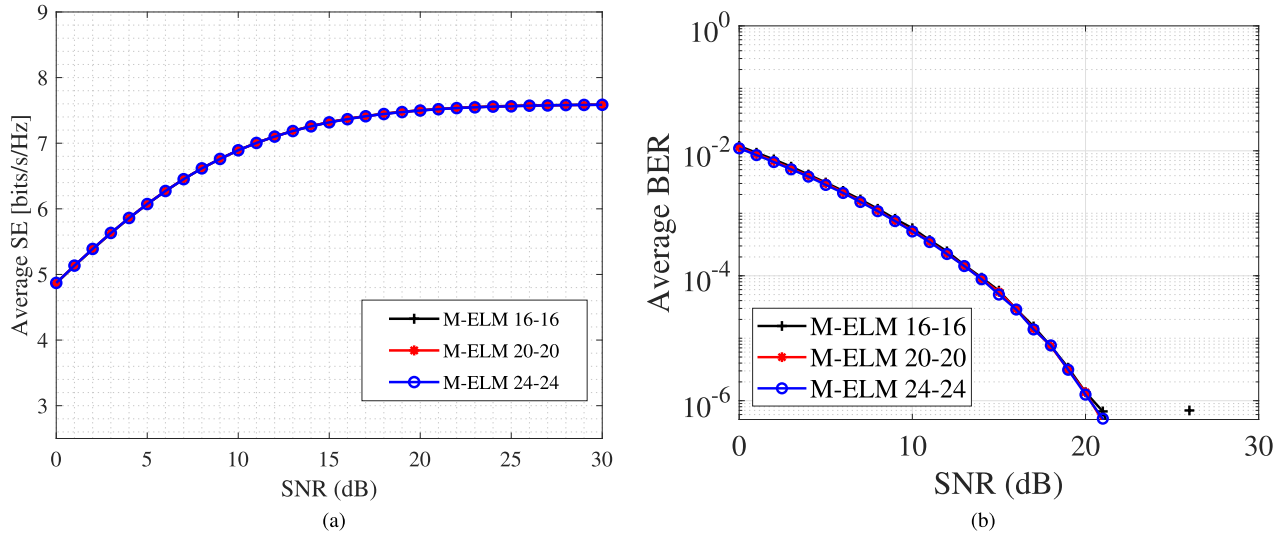
FIGURE 8. BER for 16-QAM demodulated symbols at fixed neurons of 16, 20, and 24 in the hidden layers of the M-ELM receiver. (a) Average SE, (b) average BER.
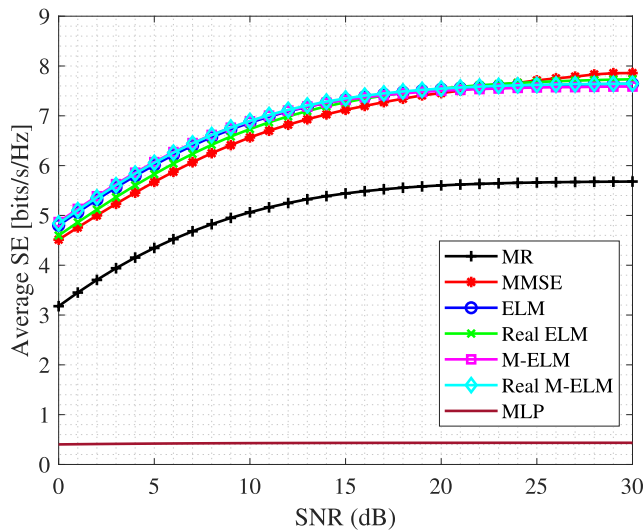

FIGURE 9. Achieved SE of the MR, MMSE, ELM, real ELM, M-ELM, and real M-ELM MIMO receivers for 16-QAM modulation.

FIGURE 10. Achieved SE of the MR, MMSE, ELM, real ELM, M-ELM, and real M-ELM MIMO receivers for 64-QAM modulation.

In Fig. 9, for the complex ELM and M-ELM strategies, the average SE is almost the same. For most of the SNR range, these ANNs outperform the real ELM and MMSE methods. For the MR receiver, the average SE is low due to the impact of non-linear distortions and the presence multiuser interference. To complement the results obtained in Fig. 9, we present the average SE for a 64-QAM modulation scheme in Fig. 10.

As expected, in Fig. 10 the SE results with 64-QAM are higher than 16-QAM for the techniques presented in this study. However, the SE results achieved for 64-QAM with MMSE are less than those achieved with 16-QAM.

In Fig. 11, we present the 16-QAM BER results achieved for the different receivers to account for transmission errors.
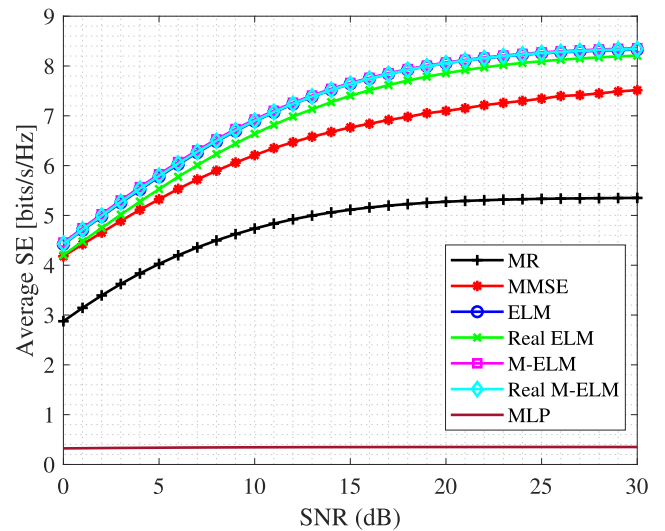
The results in Fig. 11 reveal that both the complex ELM and the real and complex M-ELM strategies have the lowest BER. An average BER of 10^{-6} can be reached at 10 dB of SNR with these ANN methods. The average BER achieved with real ELM is slightly higher than that achieved with the complex ELM and M-ELM (real and complex) techniques. However, both MMSE and MR receivers show high BER curves. These results show that ANN strategies can reduce the effects of multiuser interference with more precision than linear MIMO receivers (MMSE and MR). In Fig. 12 we present the average BER results achieved with a 64-QAM modulation scheme.

As expected, with a 64-QAM modulation scheme, the results in Fig. 12 present higher BER curves than the 16-QAM results in Fig. 11. We can see that the bit error rate
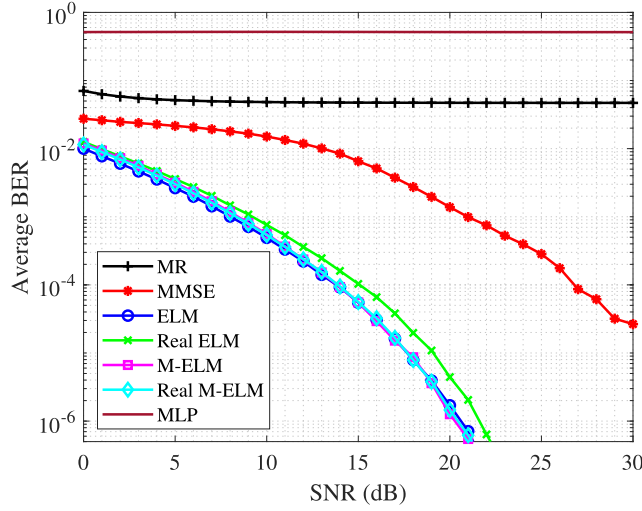
FIGURE 11. Achieved BER of the MR, MMSE, ELM, real ELM, M-ELM, and real M-ELM MIMO receivers for 16-QAM modulation.
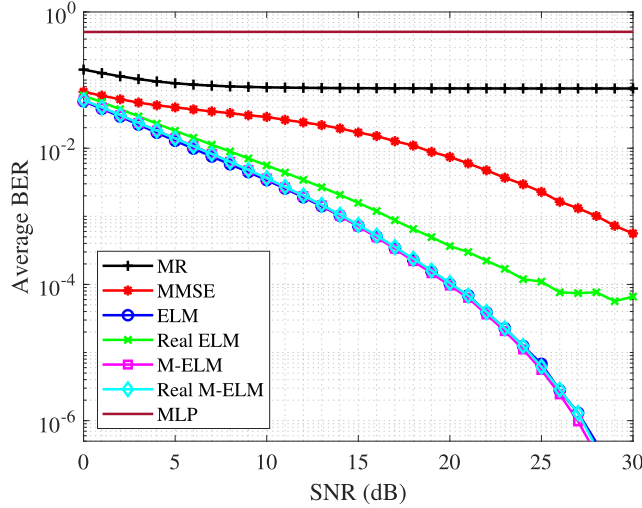

FIGURE 12. Achieved BER of the MR, MMSE, ELM, real ELM, M-ELM, and real M-ELM MIMO receivers for 64-QAM modulation.

with the Real ELM receiver increases, whereas the results of ELM, M-ELM, and Real-ELM are slightly the same.

Finally, in Table 3, we present the processing time, in terms of the sample mean and variance range over 2000 simulations, to perform the MIMO combination task with the proposed ANNs, as well as with the MR and MMSE receivers. However, only for MR and MMSE processing, we have also considered the processing time to perform channel estimation as is described in Section II-C, since this process is not performed with the proposed M-ELM strategy.

The results obtained in Table 3 depend on the hardware on which the simulations were performed. Taking this into account, the technical characteristics of the computer used to perform the simulations are as follows:

- Processor: 3.3 GHz 10 Core Intel Xeon W processor
- Memory: 128 GB 2933 MHz DDR4

TABLE 3. MIMO combining processing time.

Method	Time (ms)
MR	17.9 ± 1.8
MMSE	642.7 ± 40.6
ELM	20.0 ± 1.6
Real ELM	27.2 ± 2.2
M-ELM	15.5 ± 1.5
Real M-ELM	35.2 ± 2.3
MLP	340.5 ± 20.7

TABLE 4. Floating-point operations for the different receivers.

Method	Combining vector operations
MR	$N_s(4KN_r - 1)$
MMSE	$N_s \left(N_r^3 + N_r^2 \left(5K/2 + 2 \right) + N_r \left(K + 1/2 \right) \right)$
ELM	$K \left(L^3 + L^2 \left(3N_s + 1/2 \right) + L \left(2N_s - 1/2 \right) \right)$
M-ELM	$Ki \left(M^3 + M^2 \left(3N_s + 1/2 \right) + M \left(2N_s - 1/2 \right) \right)$

- Software: Matlab R2019b

As in [26], we complement the results presented in Table 3 with the number of floating-point operations in analytical form. The results obtained in Table 3 were obtained following the methodology used in [26]. Table 4 presents the number of floating-point operations for the receivers MR, MMSE, ELM (the best demodulator reported in the literature) and the novel M-ELM (proposed method), where M is the number of neurons in the hidden layers for the M-ELM strategy, and L is the number of neurons in the single layer of the ELM strategy. Furthermore, $M < L$, and i represents the number of hidden neurons for the proposed M-ELM strategy.

Taking into account the parameters of Table 1 and based on the data presented in Table 4, when $N_s = 792$, $K = 4$, $N_r = 64$, the number of floating-point operations required for MMSE is 246774528, whereas for ELM processing, the floating-point operations are 40390528, that is, the ELM algorithm requires 16.367 % of the operations required for MMSE. Furthermore, the number of floating-point operations for the M-ELM receiver is 5102528, which represents 2.067 % of the operations required for MMSE processing or 12.633 % of the operations required for the single layer ELM neural network.

D. COMPARISON WITH A DNN APPROACH

A multilayer perceptron (MLP) is implemented and tested to compare its performance with the ELM methods in terms of BER, time complexity, and implementation. The hyperparameters of this DNN were selected using a random search approach. In this sense, we first define the maximum and minimum values for the number of layers, the number of neurons in each layer, the minibatch size for model training, and the number of training epochs. The available activation functions are *tanh* and *ReLU*. Hyperparameters were randomly tested according to a uniform distribution. Finally, the hyperparameters with the best BER are selected.

The final structure of the MLP is the following: An input layer with $2N_s$ neurons, corresponding to the real and imaginary components of the N_s subcarriers in the mmWave signal. These are followed by 10 fully connected layers with 120, 100, 120, 100, 120, 100, 120, 100, 120, and 100 neurons, respectively. After each fully connected layer, there follows a batch normalization layer that helps to speed up the DNN training process and reduce sensitivity to network initialization. The activation function for each hidden layer is *tanh*. Finally, we set the output, a regression layer, with $2N_s$ neurons corresponding to the real and imaginary components of the N_s subcarriers. We used the stochastic gradient descent algorithm with a momentum of 0.9 and update the network parameters with a learning rate of 0.001 and a mini-batch size of 600 samples. Furthermore, we reduced the learning rate by a factor of 0.99 after each set of 10 epochs.

As we can see in Figs. 9-10, the SE results achieved with the MLP strategy for 16-QAM and 64-QAM modulations are significantly lower than the results obtained with the revised techniques in this work. Consequently, Figs. 11-12 present higher BER results for the MLP strategy with 16-QAM and 64-QAM. Although the MLP strategy produced efficient results for mmWave communications with single antenna systems, as presented in [10], the MLP neural network cannot perform MIMO processing as the M-ELM strategy and linear MIMO receivers can.

In terms of implementation, the ELMs methods present relevant advantages with respect to a DNN technique. First, a DNN requires much higher processing capabilities than ELM methods. In other words, the DNN requires adjusting the input to a vector of real numbers and a larger number of operations to predict the same output. Additionally, an ELM quickly adjusts (online training) to changes in the system. However, a DNN model requires offline training to adapt to new conditions in the system, such as a different location of the mobile user terminal.

V. ANALYSIS AND DISCUSSION

There are optimized signal processing techniques that can achieve better performance compared to MR and MMSE receivers. However, these techniques require more computational resources. This is a limitation because the signal processing task is higher for mmWave communications.

SE results with the ELM strategies are equal to or higher than those obtained with the MMSE receiver and higher than the SE results achieved with the MR combiner. Moreover, we can see that the SE results obtained with M-ELM can exceed 8 bits/s/Hz with a 64-QAM scheme, even though this modulation scheme can only transmit 6 bits/symbol. This result is due to the number of antennas in a communication system. Since mmWave systems use large arrays of antennas, SE increases with the number of antennas compared to SISO systems [52]. However, with linear MIMO receivers (MMSE, MR), the growth in spectral efficiency depends on how the multiple antenna channels

are estimated, since an error in the estimation process reduces the SINR and therefore reduces the SE, as described in (12).

We can also see that the SE results obtained with M-ELM can follow the same rule of SE for linear MIMO receivers, even though there is no direct channel estimation with this technique. However, the ELM strategies presented in this study show lower BER results than the linear MIMO receivers, that is, the MR and MMSE methods. However, the ELM strategies presented in this study show smaller BER results than the linear MIMO receivers, namely, the MR and MMSE methods. This is due to the advantage that ANNs present since the multiuser interference is attenuated more precisely than with linear MIMO combining methods; moreover, the computational complexity is reduced.

As presented in Section IV-C, in figures 9 and 10, the proposed complex M-ELM receiver can achieve the same SE and BER results as the ones achieved with the complex ELM method. However, the former is a more efficient receiver, as this strategy presents less computational complexity than the latter. The reason why the multilayer strategy presents less computational complexity than the single-layer ELM is that we can split the process into smaller matrices. The complex ELM is made up of a single hidden layer with $L = N_r$ neurons, that is, 64, the same number of BS antennas dedicated per user in the cell. This means that the internal processing of this method requires a large matrix operation. However, with the multilayer strategy, we can establish two hidden layers with at least 16 neurons in each layer. In this way, internal processing requires the operation of smaller matrices.

Splitting the signal processing task into smaller matrix operations is the primary function of a multilayer strategy. Therefore, the computational complexity relaxes with the proposed complex M-ELM receiver. This is the main contribution of our investigation because the mmWave signal processing is high, and effective methods are necessary to reduce complexity.

VI. CONCLUSION

There is a trade-off between the performance of the BER/SE metrics and the execution time of the MIMO combining processing with the proposed M-ELM. With only 16 neurons in each of the two configured hidden layers of the proposed ANN, we can achieve the same SE and BER results as the standard ELM receiver. Furthermore, the proposed M-ELM strategy requires less processing time than the standard ELM receiver to perform the same MIMO processing. Moreover, the lower processing time required by the proposed M-ELM is validated by the analytical number of floating-point operations required by the proposed scheme. Therefore, the M-ELM method requires a shorter processing time compared to other state-of-the-art schemes and even a shorter processing time than traditional linear MIMO receivers.

REFERENCES

[1] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. New York, NY, USA: Academic, 2018.

[2] L. Li, H. Ren, X. Li, W. Chen, and Z. Han, "Machine learning-based spectrum efficiency hybrid precoding with lens array and low-resolution ADCs," *IEEE Access*, vol. 7, pp. 117986–117996, 2019.

[3] A. N. Uwaechia and N. M. Mahyuddin, "A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges," *IEEE Access*, vol. 8, pp. 62367–62414, 2020.

[4] A. Mughees, M. Tahir, M. A. Sheikh, and A. Ahad, "Towards energy efficient 5G networks using machine learning: Taxonomy, research challenges, and future research directions," *IEEE Access*, vol. 8, pp. 187498–187522, 2020.

[5] S. Hashima, B. M. ElHalawany, K. Hatano, K. Wu, and E. M. Mohamed, "Leveraging machine-learning for D2D communications in 5G/Beyond 5G networks," *Electronics*, vol. 10, no. 2, p. 169, Jan. 2021.

[6] J. Chen, W. Feng, J. Xing, P. Yang, G. E. Sobelman, D. Lin, and S. Li, "Hybrid beamforming/combining for millimeter wave MIMO: A machine learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11353–11368, Dec. 2020.

[7] X. Li and A. Alkhateeb, "Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, Nov. 2019, pp. 800–805.

[8] H. Ren, L. Li, W. Xu, W. Chen, and Z. Han, "Machine learning-based hybrid precoding with robust error for UAV mmWave massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[9] E. Björnson and P. Giselsson, "Two applications of deep learning in the physical layer of communication systems [lecture notes]," *IEEE Signal Process. Mag.*, vol. 37, no. 5, pp. 134–140, Sep. 2020.

[10] D. F. Carrera, C. Vargas-Rosales, N. M. Yungaicela-Naula, and L. Azpilicueta, "Comparative study of artificial neural network based channel equalization methods for mmWave communications," *IEEE Access*, vol. 9, pp. 41678–41687, 2021.

[11] D. F. Carrera, C. Vargas-Rosales, C. A. Azurdia-Meza, and M. Morocho-Yaguana, "Extreme learning machine detector for millimeter-wave massive MIMO systems," *AEU Int. J. Electron. Commun.*, vol. 138, Aug. 2021, Art. no. 153875.

[12] S. Nie and I. F. Akyildiz, "Deep kernel learning-based channel estimation in ultra-massive MIMO communications at 0.06–10 THz," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2019, pp. 1–6.

[13] M. Y. Takeda, A. Klautau, A. Mezghani, and R. W. Heath, Jr., "MIMO channel estimation with non-ideal ADCS: Deep learning versus GAMP," in *Proc. IEEE 29th Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Oct. 2019, pp. 1–6.

[14] S. Kumar, S. K. Vankayala, B. S. Sahoo, and S. Yoon, "Continual learning-based channel estimation for 5G millimeter-wave systems," in *Proc. IEEE 18th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2021, pp. 1–6.

[15] L. Zhang and L.-L. Yang, "Machine learning for joint channel equalization and signal detection," in *Machine Learning for Future Wireless Communications*. Hoboken, NJ, USA: Wiley, 2020, pp. 213–241.

[16] D. F. Carrera, C. Vargas-Rosales, L. Azpilicueta, and J. A. Galaviz-Aguilar, "Comparative study of channel estimators for massive MIMO 5G NR systems," *IET Commun.*, vol. 14, no. 7, pp. 1175–1184, Apr. 2020.

[17] T. Van Chien, E. Bjornson, and E. G. Larsson, "Sum spectral efficiency maximization in massive MIMO systems: Benefits from deep learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[18] M.-W. Un, M. Shao, W.-K. Ma, and P. C. Ching, "Deep MIMO detection using ADMM unfolding," in *Proc. IEEE Data Sci. Workshop (DSW)*, Jun. 2019, pp. 333–337.

[19] L. Azpilicueta, P. Lopez-Iturri, J. Zuñiga-Mejia, M. Celaya-Echarri, F. A. Rodríguez-Corbo, C. Vargas-Rosales, E. Aguirre, D. G. Michelson, and F. Falcone, "Fifth-generation (5G) mmwave spatial channel characterization for urban environments' system analysis," *Sensors*, vol. 20, no. 18, p. 5360, 2020.

[20] F. A. Rodríguez-Corbo, L. Azpilicueta, M. Celaya-Echarri, P. Lopez-Iturri, I. Picallo, F. Falcone, and A. V. Alejos, "Deterministic 3D ray-launching millimeter wave channel characterization for vehicular communications in urban environments," *Sensors*, vol. 20, no. 18, p. 5284, Sep. 2020.

[21] D. Zabala-Blanco, M. Mora, R. J. Barrientos, R. Hernández-García, and J. Naranjo-Torres, "Fingerprint classification through standard and weighted extreme learning machines," *Appl. Sci.*, vol. 10, no. 12, p. 4125, Jun. 2020.

[22] A. Khellal, H. Ma, and Q. Fei, "Convolutional neural network features comparison between back-propagation and extreme learning machine," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 9629–9634.

[23] S. Pang and X. Yang, "Deep convolutional extreme learning machine and its application in handwritten digit classification," *Comput. Intell. Neurosci.*, vol. 2016, pp. 1–10, Jul. 2016.

[24] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, vol. 2, Jul. 2004, pp. 985–990.

[25] J. Tang, C. Deng, and G.-B. Huang, "Extreme learning machine for multilayer perceptron," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 4, pp. 809–821, Apr. 2015.

[26] D. F. Carrera, D. Zabala-Blanco, C. Vargas-Rosales, and C. A. Azurdia-Meza, "Extreme learning machine-based receiver for multi-user massive MIMO systems," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 484–488, Feb. 2021.

[27] D. Gao and Q. Guo, "Extreme learning machine-based receiver for MIMO LED communications," *Digit. Signal Process.*, vol. 95, Dec. 2019, Art. no. 102594.

[28] D. Gao, Q. Guo, and Y. C. Eldar, "Massive MIMO as an extreme learning machine," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 1046–1050, Jan. 2021.

[29] I. G. Muhammad, K. E. Tepe, and E. Abdel-Raheem, "QAM equalization and symbol detection in OFDM systems using extreme learning machine," *Neural Comput. Appl.*, vol. 22, nos. 3–4, pp. 491–500, Mar. 2013.

[30] D. Li, T. Kang, J. Zhou, and Q. Jin, "Dynamic extreme learning machine identification for nonlinear system with long time delay," in *Proc. 36th Chin. Control Conf. (CCC)*, Jul. 2017, pp. 2094–2098.

[31] M. B. Stojanovic, N. M. Sekulovic, and A. S. Panajotovic, "A comparative performance analysis of extreme learning machine and echo state network for wireless channel prediction," in *Proc. 14th Int. Conf. Adv. Technol., Syst. Services Telecommun. (TELSIKS)*, Oct. 2019, pp. 356–359.

[32] L. Sanguinetti, E. Bjornson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, Jan. 2019.

[33] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.

[34] D. F. Carrera, C. Vargas-Rosales, R. Villalpando-Hernandez, and J. A. Galaviz-Aguilar, "Performance improvement for multi-user millimeter-wave massive MIMO systems," *IEEE Access*, vol. 8, pp. 87735–87748, 2020.

[35] D. Zabala-Blanco, M. Mora, C. A. Azurdia-Meza, A. D. Firoozabadi, P. P. Játiva, and I. Soto, "Relaxation of the radio-frequency linewidth for coherent-optical orthogonal frequency-division multiplexing schemes by employing the improved extreme learning machine," *Symmetry*, vol. 12, no. 4, p. 632, Apr. 2020.

[36] D. Xiao, B. Li, and Y. Mao, "A multiple hidden layers extreme learning machine method and its application," *Math. Problems Eng.*, vol. 2017, pp. 1–10, Jan. 2017.

[37] L. Duan, M. Bao, J. Miao, Y. Xu, and J. Chen, "Classification based on multilayer extreme learning machine for motor imagery task from EEG signals," *Proc. Comput. Sci.*, vol. 88, pp. 176–184, Jan. 2016.

[38] L. L. C. Kasun, H. Zhou, G.-B. Huang, and C. M. Vong, "Representational learning with extreme learning machine for big data," *IEEE Intell. Syst.*, vol. 28, no. 6, pp. 31–34, Nov. 2013.

[39] L. L. C. Kasun, Y. Yang, G. B. Huang, and Z. Zhang, "Dimension reduction with extreme learning machine," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3906–3918, Aug. 2016.

[40] S. Ding, N. Zhang, X. Xu, L. Guo, and J. Zhang, "Deep extreme learning machine and its application in EEG classification," *Math. Problems Eng.*, vol. 2015, pp. 1–11, Nov. 2015.

[41] A. Zaidi, F. Athley, J. Medbo, U. Gustavsson, G. Durisi, and X. Chen, *5G Physical Layer: Principles, Models and Technology Components*. New York, NY, USA: Academic, 2018.

[42] H. Yang, "A road to future broadband wireless access: MIMO-OFDM-based air interface," *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 53–60, Jan. 2005.

[43] O. A. Saraereh, I. Khan, Q. Alsafasfeh, S. Alemaishat, and S. Kim, "Low-complexity channel estimation in 5G massive MIMO-OFDM systems," *Symmetry*, vol. 11, no. 5, p. 713, May 2019.

[44] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Jan. 2014.

[45] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, Jr., "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.

[46] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[47] C. A. Gutierrez and M. Cabrera, "Issues of the simulation of wireless channels with exponential-decay power-delay profl es," in *Proc. IEEE 16th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2005, pp. 507–511.

[48] S.-Y. Lien, S.-L. Shieh, Y. Huang, B. Su, Y.-L. Hsu, and H.-Y. Wei, "5G new radio: Waveform, frame structure, multiple access, and initial access," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 64–71, Jun. 2017.

[49] Y.-M. Liang, H.-W. Luo, and J.-G. Huang, "Extended Kalman filtering-based channel estimation for space-time coded MIMO-OFDM systems," *J. Shanghai Univ. English Ed.*, vol. 11, no. 5, pp. 469–473, Oct. 2007.

[50] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[51] A. K. Shrivas, "A comparative analysis of LS and MMSE channel estimation techniques for MIMO-OFDM system," *Int. J. Sci. Res. Develop.*, vol. 1, no. 8, pp. 44–48, Jan. 2015.

[52] T. L. Marzetta, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[53] L. Sanguinetti, E. Bjornson, M. Debbah, and A. L. Moustakas, "Optimal linear precoding in multi-user MIMO systems: A large system analysis," in *Proc. IEEE Global Commun. Conf.*, Dec. 2014, pp. 3922–3927.

[54] N. N. Pise and P. Kulkarni, "A survey of semi-supervised learning methods," in *Proc. Int. Conf. Comput. Intell. Secur.*, Dec. 2008, pp. 30–34.

[55] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in extreme learning machines: A review," *Neural Netw.*, vol. 61, pp. 32–48, Jun. 2015.

[56] E. Salazar, C. A. Azurdia-Meza, D. Zabala-Blanco, S. Bolufé, and I. Soto, "Semi-supervised extreme learning machine channel estimator and equalizer for vehicle to vehicle communications," *Electronics*, vol. 10, no. 8, p. 968, Apr. 2021.

[57] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "Fully complex extreme learning machine," *Neurocomputing*, vol. 68, nos. 1–4, pp. 306–314, Oct. 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231205000500

[58] L. Yang, Q. Zhao, and Y. Jing, "Channel equalization and detection with ELM-based regressors for OFDM systems," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 86–89, Jan. 2020.

[59] W. Wang, C.-M. Vong, Y. Yang, and P.-K. Wong, "Encrypted image classification based on multilayer extreme learning machine," *Multidimensional Syst. Signal Process.*, vol. 28, no. 3, pp. 851–865, Jul. 2017.

[60] G. Huang, L. Kasun, H. Zhou, and C. Vong, "Representational learning with extreme learning machine for big data," *IEEE Intell. Syst.*, vol. 28, no. 6, pp. 31–34, Dec. 2013.

[61] X. Gu, D. Liu, C. Baks, O. Tageman, B. Sadhu, J. Hallin, L. Rexberg, and A. Valdes-Garcia, "A multilayer organic package with 64 dual-polarized antennas for 28 GHz 5G communication," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2017, pp. 1899–1901.

[62] E. Degirmenci, "EMF test report: Ericsson AIR 5121," Ericsson AB, Stockholm, Sweden, Tech. Rep. GFTB 17:001589 Uen Rev B, Jan. 2018.

[63] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt, and E. Eberlein, "Quasi deterministic radio channel generator user manual and documentation," Fraunhofer Heinrich Hertz Inst., Ericsson AB, Stockholm, Sweden, Tech. Rep. GFTB-17:001589, 2016, pp. 1–4.

[64] M. Tercero, P. von Wrycza, A. Amah, J. Widmer, M. Fresia, V. Frascolla, J. Lorca, T. Svensson, M. Hamon, S. D. Roblot, A. Vijay, M. Peter, V. Sgardoni, M. Hunukumbure, J. Luo, and N. Vucic, "5G systems: The mmMAGIC project perspective on use cases and challenges between 6–100 GHz," in *Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW)*, Apr. 2016, pp. 200–205.

[65] D. Zabala-Blanco, M. Mora, C. A. Azurdia-Meza, and A. D. Firoozabadi, "Extreme learning machines to combat phase noise in RoF-OFDM schemes," *Electronics*, vol. 8, p. 921, Sep. 2019.

[66] S. Ding, X. Xu, and R. Nie, "Extreme learning machine and its applications," *Neural Comput. Appl.*, vol. 25, nos. 3–4, pp. 549–556, Sep. 2014.

DIEGO FERNANDO CARRERA (Member, IEEE) received the B.Sc. degree in electronics and telecommunications engineering from the Universidad Técnica Particular de Loja (UTPL), Ecuador, in 2012, the M.Sc. degree in information technology management, in 2015, and the Ph.D. degree in telecommunications engineering from the Tecnologico de Monterrey, Mexico, in 2020. His research interests include statistical signal processing and machine learning techniques for wireless communication systems, OFDM-based systems, channel estimation, symbol detection, 5G and beyond enabling technologies, massive MIMO systems, and mmWave communications. He has served as a Reviewer for journals, such as IEEE Access, IEEE Sensors Letters, IEEE Wireless Communications Letters, and *IET Communications*.

CESAR VARGAS-ROSALES (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering with major in communications and signal processing from Louisiana State University. He is the coauthor of the books *Position Location Techniques and Applications* (Academic Press/Elsevier) and *Radio Wave Propagation in Vehicular Environments* (IET). His research interests include personal communications, 5G/6G, cognitive radio, MIMO systems, stochastic modeling, intrusion/anomaly detection in networks, position location, interference, networks, and channel coding. He is a member of the Mexican National Researchers System (SNI), the Mexican Academy of Sciences (AMC), and the Academy of Engineering of Mexico. He has been the IEEE Communications Society Monterrey Chapter Chair and the Faculty Advisor of the IEEE-HKN Lambda-Rho Chapter at the Tecnologico de Monterrey. He was also the Technical Program Chair of the IEEE Wireless Communications and Networking Conference (IEEE WCNC). He is an Associate Editor of IEEE Access and *International Journal of Distributed Sensor Networks*.

DAVID ZABALA-BLANCO (Member, IEEE) received the B.S. degree in electronic systems engineering from the Escuela Militar de Ingeniería, La Paz, Bolivia, in 2011, and the M.Sc. and Ph.D. degrees in telecommunication engineering from the Tecnológico de Monterrey, Monterrey, México, in 2014 and 2018, respectively. From 2019 to 2021, he worked as a Postdoctoral Researcher at the Universidad Católica del Maule. He currently works as an Assistant Professor at the Department of Computer, Faculty of Engineering Science, Universidad Católica del Maule. His research interests include OFDM-based systems, optical communications, Nyquist-I pulses, and machine learning approaches. He has served as a Reviewer for several IEEE conferences, as well as in journals, such as IEEE Access, *Sensors*, IEEE Latin America Transactions, and *Wireless Personal Communications*.

NOE M. YUNGAICELA-NAULA received the B.Sc. degree in electronic and telecommunication engineering from the Universidad de Cuenca, Cuenca, Ecuador, in 2015, and the M.Sc. degree in intelligent systems from the Tecnológico de Monterrey, México, in 2018, where he is currently pursuing the Ph.D. degree. From November 2017 to March 2018, he was a Visiting Scholar at Concordia University, Montreal, QC, Canada. His current research interest includes the use of techniques of artificial intelligence to automate different tasks on new generation networks, such as traffic modeling, intrusion detection systems, and network resources optimization.

CESAR A. AZURDIA-MEZA (Member, IEEE) received the B.Sc. degree in electronics engineering from the Universidad del Valle de Guatemala, Guatemala, in 2005, the M.Sc. degree in electrical engineering from Linnaeus University, Sweden, in 2009, and the Ph.D. degree in electronics and radio engineering from Kyung Hee University, Republic of Korea, in 2013. He joined the Department of Electrical Engineering, University of Chile, as an Assistant Professor, in August 2013, and since August 2021, he has been an Associate Professor, where he is currently lecturing on wireless and mobile communication systems. His research interests include Nyquist's ISI criterion, OFDM-based systems, SC-FDMA, visible light communication systems, vehicular communications, 5G and beyond enabling technologies, and signal processing techniques for communication systems. He has served as a Technical Program Committee (TPC) Member for multiple conferences, as well as a Reviewer for journals, such as the IEEE COMMUNICATIONS LETTERS, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, *Wireless Personal Communications*, IEEE ACCESS, *IET Communications*, and *EURASIP Journal on Advances in Signal Processing*, among others. He is an IEEE Communications Society Member, as well as a member of the IEICE Communications Society. He was a co-recipient of the 2019 IEEE LATINCOM Best Paper Award, as well as the 2016 IEEE CONESCAPAN Best Paper Award.

and digital signal processing, with a particular focus on smart antennas, cooperative communications, signal classification for cognitive radio systems, synchronization and channel estimation, multiple-input multiple-output antenna systems, multicarrier systems, and error correcting codes. He was a recipient of the Young Scientist Award from the International Union of Radio Science in 1999. He serves as an Editor for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY.

MOHAMED MAREY (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from Menoufia University, Egypt, in 1999, and the Ph.D. degree in electrical engineering from Ghent University, Belgium, in 2008. From 2009 to 2014, he was a Research Associate and a Visiting Professor with the Faculty of Engineering and Applied Science, Memorial University, Canada. He is currently a Full Professor with the Faculty of Electronic Engineering, Menoufia University. He is on a sabbatical leave in order to join Prince Sultan University, Saudi Arabia, as a Research Laboratory Leader of the Smart Systems Engineering Laboratory. He authored the book *Multi-Carrier Receivers in the Presence of Interference: Overlay Systems* (VDM Publishing House Ltd., 2009) and around 100 scientific papers published in international journals and conferences. His main research interests include wireless communications

ALI DEHGHAN FIROOZABADI was born in Yazd, Iran, in 1985. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering-telecommunications from Yazd University, Iran, in 2007, 2009, and 2015, respectively. From 2015 to 2017, he worked as a Postdoctoral Researcher at the Universidad de Chile, Universidad de Santiago de Chile, and Pontificia Universidad Católica de Chile. He was appointed as an Associated Professor at the Departamento de Electricidad, Facultad de Ingeniería, Universidad Tecnológica Metropolitana, Chile, in 2017. His research interests include microphone array signal processing, speaker localization and tracking, speech enhancement, visible light communication, and light source localization.

● ● ●