

# A Novel Extreme-Learning-Machine Aided Receiver Design for THz-SM with Hardware Imperfections

Ke Jiang, Ping Yang, *Senior Member, IEEE*, Bo Zhang, *Member, IEEE*, Zilong Liu, *Senior Member, IEEE*, Jialiang Fu, Shaoqian Li, *Fellow, IEEE*

**Abstract**—Terahertz (THz) communication is promising as it can enable ultra-wide-band and ultra-high-rate for various emerging communication services. In this letter, we propose to exploit the extreme learning machine (ELM) network based regressor for simple and low-complexity joint channel estimation (CE) and signal detection (SD) for THz-band spatial modulation (THz-SM) communications impaired by hardware imperfections. Computer simulations show the performance superiority of the proposed joint CE/SD scheme when compared with the state-of-the-art schemes, and other machine learning-based ones, including the support vector machine (SVM), deep neural network (DNN) and some variants of ELM. Specifically, we show that its bit error rate (BER) performance approaches to that of the recently derived maximal likelihood (ML) SD. In addition, the robustness of the proposed scheme is validated by considering two types of background impulsive noises.

**Index Terms**—Terahertz (THz) communication, spatial modulation (SM), hardware imperfections, extreme learning machine (ELM), channel estimation (CE), signal detection (SD).

## I. INTRODUCTION

Terahertz communication band (from 0.1 THz  $\sim$  10 THz) has drawn increasing research attention from the fields of industry and academia, due to its capabilities of supporting ultra-high speed transmissions [1]–[3]. In recent years, spatial modulation (SM) [4] has been studied in THz-band and/or sub-THz-band [5]–[7]. Specifically, the authors of [5] pioneered the study of THz-band SM with non-negligible hardware imperfections, by showing that conventional channel estimation/signal detection (CE/SD) schemes in low-speed SM may not work. In [6], SM was introduced to THz-band ultra-massive MIMO systems as a feasible paradigm to increase capacity and spectral efficiency. More recently, generalized SM (GSM) is exploited to work over dual-polarized antennas to achieve higher spectral efficiency with lower spatial correlation and space occupancy in sub-THz-band MIMO systems [7].

As pointed out in [5], classical CE/SD schemes for low-frequency-band SM are generally designed by assuming simple (Gaussian) noise distributions [3], [8], which may not be applicable in THz-SM systems consisting of various types of interferences caused by hardware imperfections. Also

acquiring the probability distribution functions of the non-negligible interferences/noises imposes great challenges to enhanced designs of THz-SM receivers. In addition, to the best of our knowledge, no existing works have carefully considered the unique structural features of THz-band transmissions and SM transmissions for efficient joint designs of CE/SD. In the past few years, technologies in the field of artificial intelligence and machine learning have made major breakthroughs and some have been applied to SM and its variants [9]–[16]. For example, [9] proposed a novel framework to exploit machine learning technologies to help implement low-complexity link-adaptive SM systems; the conventional optimization-based problems, such as transmit antennas selection (TAS) and power allocation, are converted to data-based prediction problems. [10]–[12] all focused on the problem of antenna selection of SM-based systems. Authors of [10] developed a dynamic and flexible framework for large-scale GSM systems and adopted the decision tree and multi-layer perceptron to realize TAS to improve the system reliability. In [11], the TAS pipeline was formulated in both neural networks and gradient boosting decision trees, where the latter was demonstrated to be capable of achieving better efficiency in TAS. Different from the first two works, [12] considered the TAS problem of full-duplex SM and proposed two novel TAS methods based on the support vector machine (SVM) and deep neural network (DNN) to reduce the residual self-interference effects. By contrast, machine learning based signal detection of SM is the main research topic in [13]–[16]. In particular, [13] proposed a modularized DNN architecture by detecting the active antenna indices and complex symbols with small sub-DNNs. [14] proposed a block DNN based architecture and achieved better bit error ratio (BER) performance than the block zero-forcing/minimum mean-squared error schemes. Besides the signal detection, [15] also considered the channel estimation and proposed a joint scheme based on DNN for SM. Besides DNN, efficient signal detection of GSM-based visible light communication was developed in [16] with the aid of SVM by casting it as a multiple classification problem.

Without stacking convolution/pooling layers like the convolutional neural network, the extreme learning machine (ELM) [17], as a fast learning structure for the feedforward neural network, was applied to the wireless physical-layer designs due to its excellent learning efficiency [18]–[23]. In [18] [19], the ELM-based schemes were proposed for channel estimation/equalization, and detection of OFDM systems. In [20]–[23], ELM was applied to (massive) MIMO transmissions for very appealing receiver designs. However, few works have considered ELM for THz or SM transmissions.

Against above backgrounds, we propose an ELM-based

This work is supported by the National Key R&D Program of China under Grant 2020YFB1807203, the National Science Foundation of China under Grant 61876033 and the National Science Foundation of China under Grant U19B2014. (*Corresponding author: Ping Yang.*)

Ke Jiang, Ping Yang, Jialiang Fu, and Shaoqian Li are with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 610056, China (e-mail: yang.ping@uestc.edu.cn).

Bo Zhang is with the Artificial Intelligence Research Center, National Innovation Institute of Defense Technology, Beijing 100010, China.

Zilong Liu is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K.

network to implement a low-complexity joint CE/SD receiver for THz-SM. Main contributions of this paper are as follows.

- 1) To deal with very complex types of interference and noise caused by hardware imperfections in ultra-high-rate THz-band communication systems, and also by considering the unique transmit principle of SM, we propose to exploit a three-layer ELM to perform the joint CE and SD with low-complexity.
- 2) We compare the proposed ELM-based scheme with the classical CE/SD schemes, as well as that based on other machine learning algorithms, including SVM, DNN, and some variants of ELM, such as complex ELM (CELM), non-regularized ELM (NR-ELM). The BER performance superiority of the proposed scheme is validated by computer simulations, showing that it approaches to that of the recently derived maximal likelihood (ML) scheme with full consideration of hardware imperfections of THz-SM systems in [5] given perfect CSI. In addition, we consider two more types of background impulsive noises, i.e. the Gaussian Mixture Models (GMM) and Middleton (Mid) noises, to further demonstrate the robustness of the proposed scheme.

## II. SYSTEM MODEL OF THZ-SM

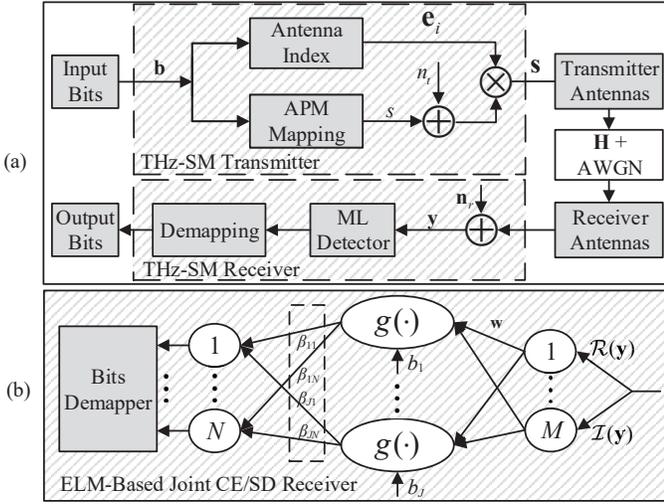


Fig. 1. (a). System diagram of THz-SM systems with hardware imperfections; (b). Diagram of ELM-based receiver for THz-SM systems.

Fig. 1 (a) is the diagram of a THz-SM system with  $N_t \times N_r$  antenna configuration. Different from low-frequency band communication, THz communication signals suffer from inevitable severe path loss mainly caused by the interaction of the spreading and absorption effects. Thus in [5] the THz channel link coefficient  $h_{i,j}$  is modeled as a product,

$$h_{i,j} = h_{PL}^{i,j} h_{AM}^{i,j} h_{MF}^{i,j}, \quad (1)$$

where  $h_{PL}^{i,j}$ ,  $h_{AM}^{i,j}$  and  $h_{MF}^{i,j}$  denote the path loss coefficient, the antenna misalignment fading coefficient, the multipath fading coefficient of the  $(i,j)$ -th channel link, respectively. Except for these, another key factor that causes the performance degradation of THz communication is that the hardware imperfections are non-negligible in ultra-high-frequency bands, including the phase noise, the in-phase and quadrature

imbalance, and the non-linearities of amplifiers. In [3], distortions caused by hardware imperfections at THz transceiver are modeled as additive complex Gaussian scalar/vectors, i.e.  $n_t \sim \mathcal{CN}(0, k_t^2 P)$  and  $\mathbf{n}_r \sim \mathcal{CN}(0, k_r^2 P \mathbf{H}_i)$  of size  $N_r \times 1$ ,  $P$  is the average power of  $M_s$ -ary amplitude and phase modulation (APM) symbol  $s$ ,  $k_t$  and  $k_r$  characterize the hardware imperfections levels of transmitter and receiver. Thus the received signal vector can be expressed as [3]:

$$\mathbf{y} = \mathbf{h}_i(s + n_t) + \mathbf{n}_r + \mathbf{n}, \quad (1 \leq i \leq N_t), \quad (2)$$

where  $\mathbf{h}_i = [h_{1,i}, h_{2,i}, \dots, h_{N_r,i}]^T$  is the  $i$ -th column THz channel matrix  $\mathbf{H}$ ,  $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_r})$  denotes the additive white gaussian noise (AWGN),  $\mathbf{I}_{N_r}$  is an  $N_r \times N_r$  identity matrix,  $\mathbf{H}_i$  is a diagonal matrix related to channel vector  $\mathbf{h}_i$ ,

$$\mathbf{H}_i = \text{diag}(|h_{1,i}|^2, |h_{2,i}|^2, \dots, |h_{N_r,i}|^2), \quad (3)$$

whose diagonal elements are the modulus of elements of  $\mathbf{h}_i$ .

At the receiver side, the THz channel is firstly estimated and then the activated antenna index and APM symbols are detected according to the ML criterion. Specifically, due to the fact that classical CE/SD schemes are designed assuming over ideal hardware environments, while this is not applicable for ultra-high-frequency transmissions, where the hardware imperfections are non-neglectable. A low-complexity mean-least-squares (MLS) CE scheme was proposed in [5] for THz-SM systems. Assume a single pilot symbol  $s_p$  is transmitted from the  $i$ -th activated TA, whose average transmit power is  $\mathbb{E}\|s_p\| = P$ . Then the received pilot signal at the  $j$ -th RA is

$$y_{j,i} = h_{j,i}(s_p + n_t) + n_{r,j} + n = h_{j,i}(s_p + \tilde{n}_t) + n, \quad (4)$$

where  $\tilde{n}_t \sim \mathcal{CN}(0, k^2 P)$ ,  $k^2 = k_t^2 + k_r^2$ . Thus the MLS estimation of  $h_{j,i}$  is [5]

$$\hat{h}_{j,i,MLS} = \frac{s_p^* y_{j,i}}{(k^2 + 1)P}. \quad (5)$$

And the ML SD of THz-SM considering hardware imperfections is also proposed in [5]. Rewrite Eq.(2) as

$$\mathbf{y} = \mathbf{h}_i s + (\mathbf{h}_i n_t + \mathbf{n}_r) + \mathbf{n} = \mathbf{h}_i s + \tilde{\mathbf{n}} + \mathbf{n}, \quad (1 \leq i \leq N_t), \quad (6)$$

where  $\tilde{\mathbf{n}} \sim \mathcal{CN}(0, (\mathbf{h}_i \mathbf{h}_i^H k_t^2 + \mathbf{H}_i k_r^2) P)$ ,  $\mathbf{y} \sim \mathcal{CN}(\mathbf{h}_i s, \mathbf{C}_i)$ ,  $\mathbf{C}_i = (\mathbf{h}_i \mathbf{h}_i^H k_t^2 + \mathbf{H}_i k_r^2) P + \sigma^2 \mathbf{I}_{N_r}$ . The likelihood function is

$$f(\mathbf{y} | \mathbf{h}_i, s) = \frac{\exp\left\{-\frac{(\mathbf{y} - \mathbf{h}_i s)^H \mathbf{C}_i^{-1} (\mathbf{y} - \mathbf{h}_i s)}{\pi^{N_r} \det(\mathbf{C}_i)}\right\}}{\pi^{N_r} \det(\mathbf{C}_i)}, \quad (7)$$

and the proposed SD criterion of [5] is

$$\begin{aligned} (\hat{i}, \hat{s}) &= \arg \min_{i,s} \{-\log(f(\mathbf{y} | i, s))\} \\ &= \arg \min_{i,s} \left\{ \log(\det(\mathbf{C}_i)) + (\mathbf{y} - \mathbf{h}_i s)^H \mathbf{C}_i^{-1} (\mathbf{y} - \mathbf{h}_i s) \right\}. \end{aligned} \quad (8)$$

## III. ELM BASED JOINT CE AND SD THZ-SM RECEIVER

### A. Extreme Learning Machine

Assume a training set  $\mathbf{G}$  with  $I$  groups  $[\mathbf{x}_i, \mathbf{t}_i]$ ,  $i=1, \dots, I$ ,

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{iM}]^T, \quad \mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{iN}]^T, \quad (10)$$

and all the elements of  $\mathbf{x}_i$  and  $\mathbf{t}_i$  are real numbers.  $M$  and  $N$  indicate that there are  $M$  input layer nodes and  $N$  output

$$\mathbf{U} = \begin{bmatrix} g\left(\left(\mathbf{w}_1\right)^T \mathbf{x}_1 + b_1\right) & g\left(\left(\mathbf{w}_2\right)^T \mathbf{x}_1 + b_2\right) & \cdots & g\left(\left(\mathbf{w}_J\right)^T \mathbf{x}_1 + b_J\right) \\ g\left(\left(\mathbf{w}_1\right)^T \mathbf{x}_2 + b_1\right) & g\left(\left(\mathbf{w}_2\right)^T \mathbf{x}_2 + b_2\right) & \cdots & g\left(\left(\mathbf{w}_J\right)^T \mathbf{x}_2 + b_J\right) \\ \vdots & \vdots & \ddots & \vdots \\ g\left(\left(\mathbf{w}_1\right)^T \mathbf{x}_I + b_1\right) & g\left(\left(\mathbf{w}_2\right)^T \mathbf{x}_I + b_2\right) & \cdots & g\left(\left(\mathbf{w}_J\right)^T \mathbf{x}_I + b_J\right) \end{bmatrix}_{I \times J}. \quad (9)$$

layer nodes. Thus the  $i$ -th output vector  $\mathbf{s}_i$  of an ELM network corresponding to input  $\mathbf{x}_i$  is

$$\mathbf{s}_i = \begin{bmatrix} \sum_{j=1}^J \beta_{j,1g} \left( (\mathbf{w}_j)^T \mathbf{x}_i + b_j \right) \\ \sum_{j=1}^J \beta_{j,2g} \left( (\mathbf{w}_j)^T \mathbf{x}_i + b_j \right) \\ \vdots \\ \sum_{j=1}^J \beta_{j,Ng} \left( (\mathbf{w}_j)^T \mathbf{x}_i + b_j \right) \end{bmatrix}^T, i = 1, \dots, I, \quad (11)$$

where  $g(\cdot)$  denotes the activation function,  $\beta_{j,n}$  ( $n = 1, \dots, N$ ) denotes the weight of the connect between the  $j$ -th hidden layer and the  $n$ -th output layer node.  $\mathbf{w}_j = [w_{1j}, w_{2j}, \dots, w_{Mj}]^T$  denotes the weights of connects from all the input layer nodes to the  $j$ -th hidden layer node.  $b_j$  denotes the bias value of the  $j$ -th hidden layer node. Let  $\mathbf{U}$  denote the hidden layer output and expressed as Eq. (9).

Let  $\mathbf{B}$  denote the matrix of weights of connections of all hidden layer nodes and output layer nodes

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1N} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{J1} & \beta_{J2} & \cdots & \beta_{JN} \end{bmatrix}_{J \times N}. \quad (12)$$

Thus the input-output of the whole ELM network can be expressed in the form of a matrix multiplication as  $\mathbf{UB} = \mathbf{S}$ ,  $\mathbf{S} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_I^T]^T$ .

Similar to [19], the cost function of the ELM network is:

$$\mathcal{F}_{ELM} = \frac{1}{2} \|\mathbf{B}\|^2 + \frac{C}{2} \|\mathbf{T} - \mathbf{UB}\|^2, \quad (13)$$

where  $C$  denotes the balancing factor for empirical and structural risks,  $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_I]^T$ . By minimizing the Eq.(13) and the regularized least squares optimization solution is [24]

$$\mathbf{B}_{ls} = \begin{cases} \left( \mathbf{U}^T \mathbf{U} + \frac{\mathbf{I}_J}{C} \right)^{-1} \mathbf{U}^T \mathbf{T}, I < J, \\ \mathbf{U}^T \left( \mathbf{U} \mathbf{U}^T + \frac{\mathbf{I}_I}{C} \right)^{-1} \mathbf{T}, I \geq J. \end{cases} \quad (14)$$

### B. ELM Based THz-SM Receiver

As in Fig. 1 (b), for a THz-SM system which exploits an ELM network for joint CE/SD, the received signal vector  $\mathbf{y}$  is a complex vector and is the input of the ELM-based receiver. If all the processing units of the ELM network are modified for the complex input, it will greatly increase the computational complexity. Thus as shown in Fig. 1 (b), the real and imaginary parts of the input complex vectors are firstly obtained and then input into the real-value ELM network. Main parts of the proposed ELM-based detector are described as follows:

1) Input layer: Considering a  $4 \times 4$  THz-SM system exploiting QPSK. In receiver side, the original input signal  $\mathbf{y}$  of

TABLE I  
MAPPING OF ONE-HOT ENCODING  $k \in \mathbb{K}$  WITH THE ANTENNA INDEX AND APM CONSTELLATION POINTS COMBINATIONS  $\mathbf{a}_n \in \mathbb{A}$

| $k = n, \mathbf{a}_n$ |
|-----------------------|-----------------------|-----------------------|-----------------------|
| 1, [1,1]              | 2, [1,2]              | 3, [1,3]              | 4, [1,4]              |
| 5, [2,1]              | 6, [6,2]              | 7, [2,3]              | 8, [2,4]              |
| 9, [3,1]              | 10, [3,2]             | 11, [3,3]             | 12, [3,4]             |
| 13, [4,1]             | 14, [4,2]             | 15, [4,3]             | 16, [4,4]             |

$4 \times 1$  is complex vector. Since we only consider a real-value ELM network for low computational complexity, thus as shown in Fig. 1 (b), the real and imaginary parts of the input  $\mathbf{y}$  are firstly obtained and form a real-value vector  $\hat{\mathbf{y}}$  of  $8 \times 1$  as the input of the ELM network,

$$\mathbf{y} = [y_1, y_2, y_3, y_4]^T \rightarrow \hat{\mathbf{y}} = [\mathcal{R}(y_1), \dots, \mathcal{R}(y_4), \mathcal{I}(y_1), \dots, \mathcal{I}(y_4)]^T. \quad (15)$$

So the input layer consists of  $M = 8$  nodes corresponding to 4 real parts and 4 imaginary parts of  $\mathbf{y}$ .

2) Hidden layer: The hidden layer contains  $J$  hidden nodes the output of each hidden layer is

$$h_j = g\left(\left(\mathbf{w}_j\right)^T \mathbf{x}_i + b_j\right) = g\left(\left(\mathbf{w}_j\right)^T \hat{\mathbf{y}} + b_j\right), j = 1, \dots, J \quad (16)$$

3) Output layer: For the THz-SM systems, there are 16 possible transmit vectors and they are encoded as one-hot codes  $k \in \mathbb{K}$  according to Table I, which indicates combinations  $\mathbf{a}_n \in \mathbb{A}$  of the activated antenna index and APM symbol index. Thus the output layer contains  $N = 16$  nodes and the output of each node is

$$s_{i,n} = \sum_{j=1}^J \beta_{j,ng} \left( (\mathbf{w}_j)^T \mathbf{x}_i + b_j \right), n = 1, \dots, N \quad (17)$$

4) Generate the training set. Specifically, randomly generate  $I$  THz-SM transmit vectors  $\mathbf{X}_{\text{training}} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I]$  and the corresponding received complex vectors with distortion noises  $\mathbf{Y}_{\text{training}} = [y_1, y_2, \dots, y_I]$  according to Eq.(2). Then the  $\mathbf{Y}_{\text{training}}$  is transformed to its real-value form  $\hat{\mathbf{Y}}_{\text{training}}$ . According to TABLE I, each SM modulated vector  $\mathbf{x}_i$  is mapped to a one-hot code  $\mathbf{r}_i$  with the length of  $N = 16$ , and forms  $\mathbf{R}_{\text{training}} = [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_I]$ . Thus the whole training set consisting of two parts is denoted as  $\mathbf{G}_{\text{training}} = (\hat{\mathbf{Y}}_{\text{training}}, \mathbf{R}_{\text{training}})$  beforehand according to the first  $\log_2(M_{APM} N_i)$  bits.

5) Generate the test set. Similar to the generation of training set and obtain the  $\mathbf{G}_{\text{test}} = (\hat{\mathbf{Y}}_{\text{test}}, \mathbf{R}_{\text{test}})$  of size  $I_{\text{test}} \times 24$ ;  
6) Joint CE/SD and BER calculation. Input  $\hat{\mathbf{Y}}_{\text{test}}$  to the trained ELM and the output is  $\hat{\mathbf{R}}$ . The transmitted signals are recovered from  $\hat{\mathbf{R}}$  according to the one-hot decoding. BER is calculated by comparing  $\hat{\mathbf{R}}$  and  $\mathbf{R}_{\text{test}}$ .

### C. Complexity Analysis and Comparisons

In this part, we analyze the complexity of the compared CE/SD schemes by calculating the number of real-value multiplications and additions needed for the transmission of a THz-SM vector. In addition, for machine-learning based schemes, training phases are off-line, and  $M=2N_t$ ,  $N=M_s N_t$ . For the proposed approach, according to Eq. (17), the complexity is  $JM_s N_t(4N_t + 1)$ . Similarly, the complexity of the CELM-based, NR-ELM-based and DNN-based schemes are  $8JM_s N_t(N_t + 6)$ ,  $JM_s N_t(4N_t + 1)$ ,  $(4JN_t + 4J^2 + 2JM_s N_t)/3 - M_s N_t$ , respectively. In addition, for the ML scheme in [5], the major complexity is incurred due to the calculation of  $\det(\mathbf{C}_i)$  and  $\mathbf{C}_i^{-1}$  in Eq. (8), both of which are the order of  $O(N_r^3)$ , and there are  $M_s N_t$  pairs of  $(i, s)$ . Thus the overall complexity of the ML in [5] is approximated to be  $M_s N_t(12N_r^2 + 36N_r - 4 + O(N_r^3))$ . It can be seen that the proposed ELM-based has the lowest complexity.

#### Algorithm 1 ELM Based Joint CE/SD for THz-SM

- 1: Initialize the ELM network as  $M$  input nodes,  $J$  hidden nodes and  $N$  output nodes;
- 2: Randomly assign the real-value input weights  $\mathbf{w}_j$  and bias factors  $b_j$ ,  $j = 1, \dots, J$ ;
- 3: Input training set  $\mathbf{G}_{\text{training}}$  to the ELM network, then calculate Eq. (14) to acquire the weight matrix  $\mathbf{B}$ ;
- 4: Input  $\hat{\mathbf{Y}}_{\text{test}}$  of test set  $\mathbf{G}_{\text{test}}$  to the trained ELM and obtain the corresponding one-hot codes  $\hat{\mathbf{R}}$ ;
- 5: Apply the mapping rule in TABLE I and demodulate the data bits from the obtained  $\hat{\mathbf{R}}$  in Step 4.

### IV. SIMULATIONS AND DISCUSSIONS

In this section, we evaluate the design feasibility on the CE and SD of THz-SM systems are evaluated by simulations at given perfect/imperfect CSI. The numbers of transmit and receive antennas are  $N_t = N_r = 4$ . In addition, without loss of generality and for more intuitive observations, i.i.d Rayleigh fading channel is assumed, meanwhile, the path loss coefficient  $h_{PL}$  and the antenna misalignment fading coefficient  $h_{AM}$  are set to 1 [5]. The signal-to-noise power ratio (SNR) is defined as  $\frac{P}{\sigma^2}$  at the receiver side. Two hardware imperfections factors are assumed to be equal, i.e.  $k_t = k_r$ . The number of hidden nodes for ELM and its variants is  $J = 200$ , activation functions are *ReLU* for ELM, NR-ELM and *tanh* for the CELM. The DNN has 3 hidden layers with  $\frac{J}{3}$  nodes and *ReLU* for each. The SVM is the C-support vector classification with radial basis function,  $C = 1$ ,  $\gamma = 5$ . Lengths of training/test sets are  $I = I_{\text{test}} = 10000$ . The validation is omitted since we regard the BER as the performance indicator of considered networks.

Fig. 2 (a) compares the BER performance of THz-SM with hardware imperfections exploiting conventional ML (Con-ML) and the ML in [5]. MLS-CSI means the CSI is obtained by MLS criterion. For cases of MLS-CSI, two types of ML SD schemes are compared with different hardware imperfections of levels  $k_t^2 = k_r^2 = -16$  dB and  $-10$  dB. Fig. 2 (b) evaluates the effects of different levels of hardware imperfections on the performance of THz-SM over perfect/imperfect CSI. In

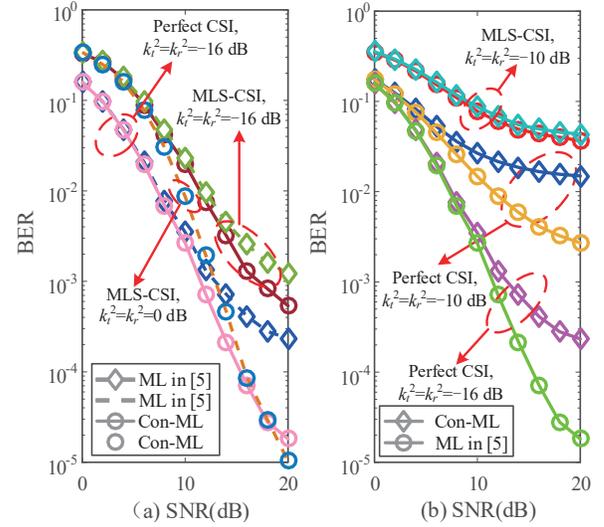


Fig. 2. BER of conventional ML and the ML in [5] for THz-SM-QPSK with perfect/LMS CSI. In (a), the circle markers or solid lines represent the results of ML in [5], the diamond markers or dash lines represent the conventional ML of low-frequency band SM.

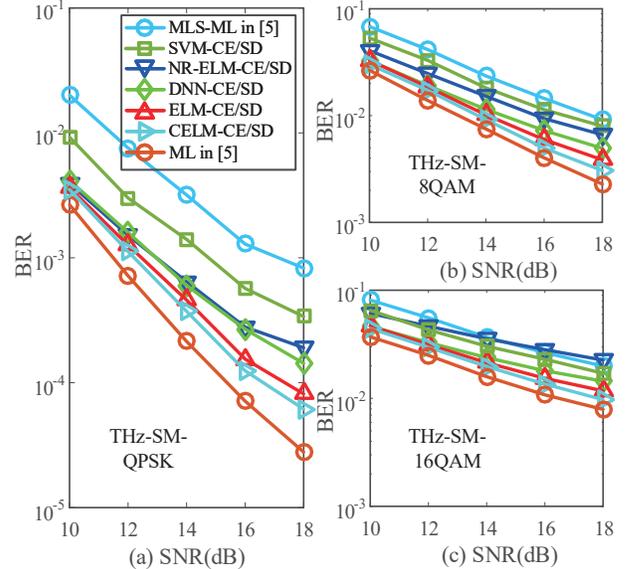


Fig. 3. BER of the proposed ELM-based, CELM-based, NR-ELM-based, DNN-based, SVM-based schemes, and the ML in [5] with MLS-based/perfect CSI for THz-SM exploiting QPSK, 8QAM, 16QAM,  $k_t^2 = k_r^2 = -16$ .

general, at the two considered hardware imperfections levels,  $k_t^2 = k_r^2 = -16$  dB and  $-10$  dB, even for the same SD scheme, i.e. the ML in [5] and the Con-ML, they all achieve considerably different BER performance with perfect CSI. This validates that the hardware imperfections are indeed non-negligible and have great effect on the systems performance. In addition, with a higher level of hardware imperfections and imperfect CSI ( $-10$  dB), the two considered MLS-ML SD schemes achieve almost the same poor BER performance. This reveals that the MLS-based CE cannot acquire satisfactory estimated channels, no matter if the hardware imperfections are considered in the stage of SD.

In Fig. 3, we present the BER performances of ELM-based joint CE/SD schemes for THz-SM with hardware imperfections and compare with that of schemes based on

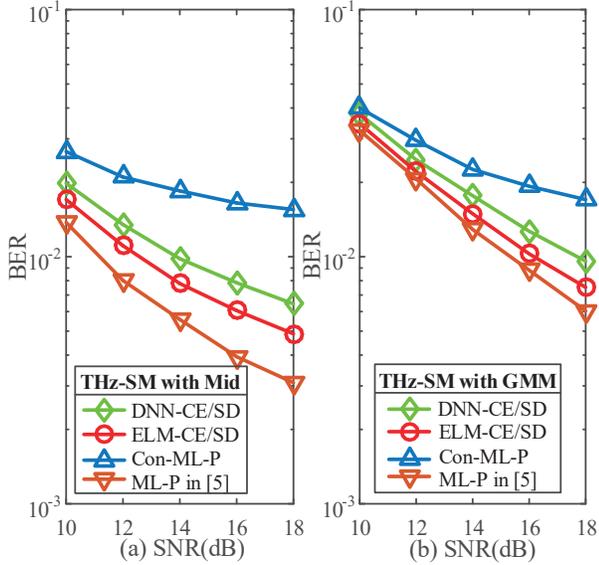


Fig. 4. BER of the proposed ELM-based scheme, DNN-based scheme, conventional ML-based scheme and the ML in [5]-based scheme with perfect CSI, QPSK and  $k_t^2 = k_r^2 = -10$  under Mid noise and GMM noise.

other machine learning methods, including DNN-based, SVM-based, CELM-based and NR-ELM-based schemes. Here QPSK, 8QAM and 16QAM are considered. And the channels are assumed to be quasi-static during offline training. From the curves in Fig. 3, we can see that the BER of the ELM-based scheme is only inferior to those of the CELM-based and ML schemes in [5], but considering the paid complexity, our proposed ELM-based scheme outperforms in terms of efficiency. While the performance of MLS-based ML scheme is far worse than that of the proposed scheme and the performance gap is about 6 dB at the BER of  $10^{-3}$ . This indicates that the ELM-based structure is feasible for CE/SD and it outperforms the MLS-based scheme in channel estimation. In addition, compared with other machine learning methods based schemes, the ELM-based scheme achieves the best performance with respect to the efficiency of achieved BER and paid computational complexity. The above results demonstrated that when exploiting the ELM network in THz-SM receivers, it is capable of achieving considerable performance with simple structure and also shows its superiority in performing fitting.

In addition, in order to further demonstrate the robustness of the proposed ELM-based joint CE/SD approach, we consider two types of impulsive background noise in the THz-SM system, i.e. the Mid noise [25] and GMM noise [26]. Limited by the paper length, only consider a higher level of hardware imperfections of  $k_t^2 = k_r^2 = -10$  dB in this part. In Fig. (4), ML-P means that the ML SD in [5] is performed with perfect CSI. From the two sets of comparisons, the proposed scheme still achieves appealing BER performance, especially in the case of GMM noise, the proposed joint CE/SD achieves very close BER performance to that of the ideal ML in [5], even at a higher level of hardware imperfections.

## V. CONCLUSION

In this letter, we have proposed to exploit the ELM-based network with only three layers for joint CE/SD in the THz-SM systems to cope with the very complex hardware imperfections. Simulation results have shown the appealing performance advantages of the ELM-based joint CE/SD design over the MLS-based ML SD scheme, and other machine learning based schemes (such as CELM, NR-ELM, DNN and SVM). It is important to point out that the BER performances approach to the recently derived ML scheme which have fully considered the hardware imperfections of THz-SM systems. In addition, the robustness of the proposed scheme under different impulsive background noises has also been validated.

## VI. ACKNOWLEDGEMENT

Many thanks to Miss Yongyang Li for her preliminary work.

## REFERENCES

- [1] K. Huang and Z. Wang, "Terahertz terabit wireless communication," *IEEE Microwave Mag.*, vol. 12, no. 4, pp. 108-116, Jun. 2011.
- [2] H. Elayan *et al.*, "Terahertz band: The last piece of RF spectrum puzzle for communication systems," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1-32, Nov. 2019.
- [3] A.-A. A. Boulogeorgos *et al.*, "Analytical performance assessment of THz wireless systems," *IEEE Access*, vol. 7, pp. 11436-11453, Jan. 2019.
- [4] R. Mesleh *et al.*, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228-2241, Jul. 2008.
- [5] T. Mao, Q. Wang, and Z. Wang, "Spatial modulation for Terahertz communication systems with hardware impairments," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4553-4557, Apr. 2020.
- [6] H. Saeeddeen "Terahertz-band ultra-massive spatial modulation MIMO," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2040-2052, Sep. 2019.
- [7] N. Bouhlel, M. Saad, and F. Bader, "Sub-Terahertz wireless system using dual-polarized generalized spatial modulation with RF impairments," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1636-1650, Jun. 2021.
- [8] X. Wu *et al.*, "Channel estimation for spatial modulation," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4362-4372, Dec. 2014.
- [9] P. Yang, Y. Xiao, M. Xiao, Y. L. Guan, S. Li, and W. Xiang, "Adaptive spatial modulation MIMO based on machine learning," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2117-2131, Sep. 2019.
- [10] S. Gecgel, C. Goztepe, and G. Karabulut Kurt, "Transmit antenna selection for large-scale MIMO GSM with machine learning," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 113-116, Jan. 2020.
- [11] Y. Zhang, J. Wang, X. Wang, Y. Xue, and J. Song, "Efficient selection on spatial modulation antennas: Learning or boosting," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1249-1252, Aug. 2020.
- [12] H. Liu, Y. Xiao, P. Yang, J. Fu, S. Li, and W. Xiang, "Transmit antenna selection for full-duplex spatial modulation based on machine learning," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10695-10708, Oct. 2021.
- [13] B. Shamasundar and A. Chockalingam, "A DNN architecture for the detection of generalized spatial modulation signals," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2770-2774, Dec. 2020.
- [14] H. Sun *et al.*, "SVM aided signal detection in generalized spatial modulation VLC system," *IEEE Access*, vol. 9, pp. 80360-80372, 2021.
- [15] L. Xiang, Y. Liu, T. Van Luong, R. G. Maunder, L. -L. Yang, and L. Hanzo, "Deep-learning-aided joint channel estimation and data detection for spatial modulation," *IEEE Access*, vol. 8, pp. 191910-191919, 2020.
- [16] H. Albinsaid *et al.*, "Block deep neural network-based signal detector for generalized spatial modulation," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2775-2779, Dec. 2020.
- [17] G.-B. Huang *et al.*, "Extreme learning machine: Theory and applications", *Neurocomputing*, vol. 70, no. 13, pp. 489-501, Dec. 2006.
- [18] J. Liu, K. Mei, X. Zhang, D. Ma, and J. Wei, "Online extreme learning machine-based channel estimation and equalization for OFDM systems," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1276-1279, Jul. 2019.
- [19] L. Yang *et al.*, "Channel equalization and detection with ELM-based regressors for OFDM systems," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 86-89, Jan. 2020.

- [20] D. F. Carrera *et al.*, "Extreme learning machine-based receiver for multi-user massive MIMO systems," *IEEE Commun. Lett.*, vol. 25, no. 2, pp. 484-488, Feb. 2021.
- [21] D. F. Carrera *et al.*, "Novel multilayer extreme learning machine as a massive MIMO receiver for millimeter wave communications," *IEEE Access*, vol. 10, pp. 58965-58981, May. 2022.
- [22] D. Gao *et al.*, "Massive MIMO as an extreme learning machine," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 1046-1050, Jan. 2021.
- [23] D. Gao, Q. Guo, "Extreme learning machine-based receiver for MIMO LED communications," *Digital Signal Processing*, vol. 95, Dec. 2019.
- [24] E. Salazar *et al.*, "Semi-supervised extreme learning machine channel estimator and equalizer for vehicle to vehicle communications," *Electronics*, vol. 10, no. 8, pp. 968, Apr. 2021.
- [25] R. Savoia and F. Verde, "Performance analysis of distributed spacetime block coding schemes in Middleton class-A noise," *IEEE Trans. Vehi. Technol.*, vol. 62, no. 6, pp. 2579-2595, Jul. 2013.
- [26] E. Axell, K. C. Wiklundh, and P. F. Stenumgaard, "Optimal power allocation for parallel two-state Gaussian mixture impulse noise channels," *IEEE Wireless Commun. Lett.*, vol. 4, no. 2, pp. 177-180, Apr. 2015.