

Received October 15, 2020, accepted October 17, 2020, date of publication October 20, 2020, date of current version November 2, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032627

Deep-Learning-Aided Joint Channel Estimation and Data Detection for Spatial Modulation

LUPING XIANG^{ID}, (Member, IEEE), YUSHA LIU^{ID}, THIEN VAN LUONG^{ID}, (Member, IEEE), ROBERT G. MAUNDER^{ID}, (Senior Member, IEEE), LIE-LIANG YANG^{ID}, (Fellow, IEEE), AND LAJOS HANZO^{ID}, (Fellow, IEEE)

NGW Group, School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K.

Corresponding author: Yusha Liu (yl6g15@ecs.soton.ac.uk)

This work was supported by the Engineering and Physical Sciences Research Council Project under Grant EP/P034284/1. The work of Lajos Hanzo was supported by the Engineering and Physical Sciences Research Council Projects under Grant EP/N004558/1, Grant EP/PO34284/1, and Grant COALESCE; and in part by the Royal Society's Global Challenges Research Fund Grant, as well as of the European Research Council's Advanced Fellow Grant QuantCom.

ABSTRACT Deep neural network (DNN)-aided spatial modulation (SM) is conceived. In particular, a pair of DNN structures are designed for replacing the conventional model-based channel estimators and detectors. As our first prototype, the conventional DNN estimates the channel relying on the pilot symbols and then carries out data detection in a data-driven manner. By contrast, our new DeepSM scheme is proposed for operation in more realistic time-varying channels, which updates the channel state information (CSI) at each time-slot (TS) before detecting the data. Hence, our novel DeepSM scheme is capable of performing well even in highly dynamic communication environments. Finally, our simulations show that the proposed DeepSM outperforms the conventional model-based channel estimation and data detection for transmission over time-varying channels.

INDEX TERMS Deep neural network (DNN), spatial modulation (SM), channel estimation, data detection.

NOMENCLATURE

Adam	Adaptive Moment Estimation
ADC	Analogue-to-Digital Convertor
APM	Amplitude-Phase Modulation
AR	Autoregressive
AWGN	Additive White Gaussian Noise
BER	Bit Error Ratio
BPSK	Binary Phase Shift Keying
CSI	Channel State Information
DNN	Deep Neural Network
FEC	Forward Error Correction
GSM	Generalized Spatial Modulation
i.i.d.	independently and identically distributed
IoT	Internet of Things
IRS	Intelligent Reconfigurable Surfaces
LDPC	Low-Density Parity-Check
MFNN	Multilayer Feedforward Neural Network
MI	Mutual Information
MIMO	Multiple-Input Multiple-Output
mMTC	Massive Machine Type Communications

MSE	Mean Squared Error
NN	Neural Network
QPSK	Quadrature Phase Shift Keying
RA	Receiver Antenna
Relu	Rectified Linear Unit
RF	Radio Frequency
SGD	Stochastic Gradient Descent
SM	Spatial Modulation
SNR	Signal-to-Noise Ratio
SSK	Space-Shift Keying
STBC	Space-Time Block Coded
TA	Transmit Antenna
TD	Time-Domain
TS	Time-Slot
V-BLAST	Vertical Bell Laboratories Layered Space-Time

I. INTRODUCTION

For coping with the dramatically increasing data traffic, novel transmission schemes are in urgent demand for improving the throughput while minimizing the deployment complexity. Among numerous cutting-edge techniques, spatial modulation (SM) [1]–[7] has constitute a promising next generation

The associate editor coordinating the review of this manuscript and approving it for publication was Zeeshan Kaleem^{ID}.

TABLE 1. Contrasting our contributions to the state-of-the-art.

Contributions	This work	[25]	[26]	[27]	[28]	[29]	[30]	[31]	[32]
SM	✓			✓	✓	✓	✓	✓	✓
DNN	✓	✓	✓			✓	✓	✓	✓
Data-driven pattern	✓			✓					
Data detection	✓	✓	✓				✓	✓	✓
Imperfect CSI	✓	✓							
Dynamic channel	✓								
Low-resolution ADCs	✓								

scheme for massive machine type communications (mMTC) [8]–[10], for TeraHertz communications [11] and for intelligent reconfigurable surfaces (IRS) [12]–[14].

The roots of SM can be traced back to 2001, when Chau and Yu [15] proposed the space-shift keying (SSK) concept, which only conveys information by the transmit antenna (TA) indices. By contrast, the SM [1], [2] activates a single TA, which transmits a single amplitude-phase modulation (APM) symbol. This unique TA activation scheme allows the transmitter to implicitly convey additional information bits ‘hidden’ in the active TA index patterns, hence achieving energy-efficient communications. A subsequent development of SM activating a small fraction of TAs is known as generalised SM (GSM) [16], which simultaneously activates a group of TAs for conveying multiple APM symbols.

Since only a low number of radio frequency (RF)-chains are active during each symbol instant, the optimal maximum-likelihood SM detector is capable of jointly detecting both the active TA indices and the classic APM signal at a low complexity. Since the RF-chains are the priciest and highest-power transceiver components, SM has compelling benefits over the family of conventional multiple-input multiple-output (MIMO) systems, such as the vertical Bell Laboratories layered space-time (V-BLAST) coded or space-time block coded (STBC) systems [17].

At the time of writing, machine learning is attracting increasing attention in wireless communications [18]–[21]. In particular, by employing neural networks (NNs), near-optimal low-complexity channel estimation and/or data detection can be achieved for different physical-layer communication schemes, relying either on a data-driven approach [22], [23], where no mathematical model is required, or on a model-driven basis [24]–[26], which exploits the benefits of both well-established mathematical physical-layer communication models and of NNs.

Machine learning techniques have also demonstrated good performance in SM systems [27]–[32]. First, adaptive SM can be designed with the aid of different deep learning techniques. For instance, Yang *et al.* [27] integrated machine learning

with adaptive SM-MIMO systems, while Tato *et al.* [28] adopted multilayer feedforward NNs (MFNNs) for accurately evaluating the mutual information (MI) in support of adapting the SM modes. Furthermore, deep learning techniques may also contribute to the TA selection of SM. In particular, Zhang *et al.* [29] formulated a data-driven NN architecture for TA selection by exploiting the channel state information (CSI), while a deep neural network (DNN)-aided TA selection scheme was designed for secure SM in [30]. Recently, Albinsaid *et al.* [31] investigated the block DNN-aided detection of GSM, whereas Shamasundar and Chockalingam [32] proposed a modularized DNN architecture for GSM detection.

However, to the best of our knowledge, there is no research published on DNN-aided joint channel estimation and data detection conceived for SM. To fill this knowledge-gap, a pair of DNN architectures are designed for joint channel estimation and data detection in SM systems. The main contributions of this paper are contrasted to the literature in Table 1 and are summarized as follows:

- Firstly, we apply the fully-connected multi-layer DNNs to SM and conceive a conventional DNN architecture for joint channel estimation and data detection of SM, operating in a data-driven manner, which replaces the conventional model-based channel estimator and detector by DNNs. Our simulations demonstrate that the conventional DNN-aided channel estimation and data detection of SM is capable of approaching the bit error ratio (BER) performance to that of the conventional model-based SM channel estimation and data detection approach for transmission over time-invariant channels, despite its reduced detection complexity.
- We then propose a novel DNN architecture, referred to as DeepSM, which detects the transmitted bits and in the meantime, updates the CSI of each time slot (TS) for transmission over time-varying fading channels in a data-driven manner. We demonstrate that our proposed DeepSM is capable of improving the BER performance of the conventional model-based channel estimation and data detection, despite its reduced detection complexity

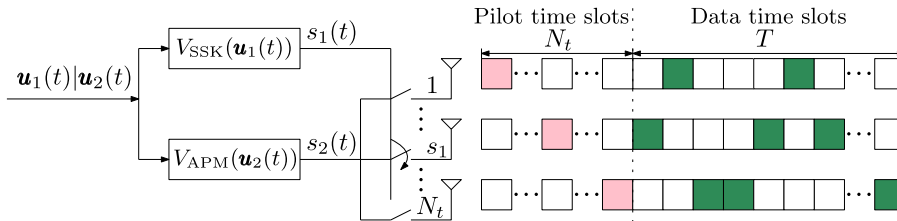


FIGURE 1. The transmitter structure of the SM system.

in terms of runtime, when communicating over more dynamic channels.

- Finally, we investigate the effects of the quantization errors imposed by finite-resolution Analogue-to-Digital Convertors (ADCs) on the BER performance attained. When our DeepSM is employed, better BER performance can be observed in the presence of 4- or 6-bit ADCs than that of conventional model-based approaches.

The rest of this paper is structured as follows. Section II reviews the SM system model, while Section III discusses the conventional model-based SM channel estimation and data detection. Next, Section IV details a pair of DNN architectures designed for joint channel estimation and data detection of SM. Then, the performance of SM evaluated for transmission over different wireless channels is quantified in Section V. Finally, our conclusions and future research ideas will be provided in Section VI.

II. SYSTEM MODEL

In this section, we detail the transceiver model of a SM system employing N_t TAs and N_r receive antennas (RAs), as shown in Fig. 1, transmitting over wireless channels, along with the main assumptions of the paper.

As shown in Fig. 1, first, the pilots are transmitted during the N_t pilot TSs, followed by T TSs for transmitting a single frame of data symbols. Within the t -th data transmission TS, the u -bit information $\mathbf{u}(t)$ is transmitted only by one of the N_t TAs. More specifically, $\mathbf{u}(t)$ is first divided into two sub-groups expressed as $\mathbf{u}(t) = \mathbf{u}_1(t)|\mathbf{u}_2(t)$. The first $u_1 = \log_2 N_t$ bits of $\mathbf{u}_1(t)$ are SSK modulated [15], [33], resulting in a SSK symbol $s_1(t) = V_{SSK}[\mathbf{u}_1(t)] \in \mathcal{M}_1$, where $V_{SSK}[\cdot]$ refers to the SSK mapping and \mathcal{M}_1 is the SSK symbol set, which can be expressed as $\mathcal{M}_1 = \{1, \dots, N_t\}$. The remaining $u_2 = \log_2 M_2$ bits $\mathbf{u}_2(t)$ are mapped to an APM symbol $s_2(t) = V_{APM}[\mathbf{u}_2(t)] \in \mathcal{M}_2$, where $V_{APM}[\cdot]$ refers to the APM mapping and \mathcal{M}_2 is defined as the symbol set of M_2 -ary APM. Then the APM symbol $s_2(t)$ together with the active TA index $s_1(t)$ mapped to the SSK symbol are transmitted by a single active TA. The transmitted signal can be expressed as

$$\mathbf{x}(t) = \mathbf{e}_{s_1(t)}(t)s_2(t) = [0, \dots, 0, s_2(t), 0, \dots, 0]^T, \quad (1)$$

where $\mathbf{e}_{s_1(t)}(t)$ is a $(N_t \times 1)$ antenna selection vector comprising only a single non-zero element ‘1’ in the $s_1(t)$ -th position, which indicates the active TA’s index.

Given the channel vector $\mathbf{h}_{s_1(t)}$ between the $s_1(t)$ -th TA and N_r RAs in the t -th TS, the signal received by the BS can be expressed as

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{e}_{s_1(t)}s_2(t) + \mathbf{n}(t), \quad (2)$$

where $\mathbf{H}(t) = [\mathbf{h}_1(t), \mathbf{h}_2(t), \dots, \mathbf{h}_{N_t}(t)]$ and $\mathbf{n}(t)$ is the additive white Gaussian noise (AWGN) obeying a zero-mean complex Gaussian distribution with a covariance matrix of $\sigma^2\mathbf{I}_{N_r}$. Hence, $1/2\sigma^2$ is the signal-to-noise ratio (SNR) per SM symbol. For time-varying channels, $\mathbf{h}_{s_1(t)}$ is expressed as [34]

$$\mathbf{h}_{s_1(t)} = \sqrt{\alpha}\mathbf{h}_{s_1(t-1)} + \sqrt{1-\alpha}\mathbf{v}(t), \quad t = 1, 2, \dots, T, \quad (3)$$

where $\mathbf{h}_{s_1(0)}$ obeys the independently and identically distributed (i.i.d.) complex Gaussian distribution and $\mathbf{v}(t)$ is the unpredictable difference from the channel vector, which also obeys the i.i.d. complex Gaussian distribution with a mean of 0 and a covariance of \mathbf{I}_{N_r} , while α is the autoregressive (AR) coefficient, which can be defined as $\alpha^{\tau/2} = \frac{E_t[h_{s_1(t)}^*h_{s_1(t+\tau)}]}{E_t[h_{s_1(t)}^*h_{s_1(t)}]}$, where $\alpha^{\tau/2}$ is the time-domain (TD) correlation and τ is the discrete time-lag. The detailed discussion of α can be found in [34].

III. CONVENTIONAL CHANNEL ESTIMATION AND DATA DETECTION FOR SM

In this section, we briefly review the conventional pilot-aided channel estimation and maximum-likelihood data detection regime of SM in Sections III-A and III-B, respectively.

A. CONVENTIONAL CHANNEL ESTIMATION

As shown in Fig. 1, we assume that prior to the data transmission, N_t TSs are employed to transmit the pilot symbols over the N_t TAs, where the proportion η of the N_t pilot symbols to the T -length data frame is defined as

$$\eta = \frac{N_t}{N_t + T}. \quad (4)$$

More specifically, in the p -th ($1 \leq p \leq N_t$) pilot TS, the p -th TA transmits its signal $x(p) = 1$ over the wireless channel. The pilot signal $\bar{\mathbf{y}}(p)$ received at the BS in the p -th ($1 \leq p \leq N_t$) TS can be expressed as

$$\bar{\mathbf{y}}(p) = \mathbf{h}_p x(p) + \mathbf{n}(p) = \mathbf{h}_p + \mathbf{n}(p), \quad (5)$$

where \mathbf{h}_p is an $N_r \times 1$ vector representing the CSI between the p -th TA and the N_r RAs.

Then the received signal is considered as the estimated CSI $\hat{\mathbf{h}}_p$, which is expressed as

$$\hat{\mathbf{h}}_p = \bar{\mathbf{y}}(p) = \mathbf{h}_p + \mathbf{n}(p), \quad 1 \leq p \leq N_t. \quad (6)$$

After the N_t pilot TSs, we can obtain the $N_r \times N_t$ estimated channel matrix, which can be expressed as

$$\hat{\mathbf{H}}(0) = [\hat{\mathbf{h}}_{1(0)}, \hat{\mathbf{h}}_{2(0)}, \dots, \hat{\mathbf{h}}_{N_t(0)}], \quad (7)$$

where $\hat{\mathbf{h}}_{p(0)} = \hat{\mathbf{h}}_p$, $p = 1, 2, \dots, N_t$, which will be employed for the maximum-likelihood data detection in the next subsection.

B. MAXIMUM-LIKELIHOOD DATA DETECTION

Following the pilot transmission, data symbols are transmitted from $t = 1$ -st to $t = T$ -th data TSs, which are detected with the aid of the estimated channel matrix $\hat{\mathbf{H}}(0)$ obtained from (6). Specifically, given the received signal $\mathbf{y}(t)$ and the estimated channel $\hat{\mathbf{H}}(0)$, the optimal maximum-likelihood detection finds the estimate $\hat{s}_1(t)|\hat{s}_2(t) \in \mathcal{M} = \mathcal{M}_1 \otimes \mathcal{M}_2$ by solving the following optimization problem:

$$\hat{\mathbf{x}}(t) = \min_{\tilde{s}_1(t) \in \mathcal{M}_1, \tilde{s}_2(t) \in \mathcal{M}_2} \|\mathbf{y}(t) - \hat{\mathbf{H}}(0)\mathbf{e}_{\tilde{s}_1(t)\tilde{s}_2(t)}\|, \quad (8)$$

where $\tilde{s}_1(t)$ represents the legitimate active TA index at the t -th TS and $\tilde{s}_2(t)$ is the potential candidate of the APM symbol in the symbol set \mathcal{M}_2 at the t -th TS. We can see that the maximum-likelihood detector has to visit all $M = N_t \times M_2$ possible combinations of $\tilde{s}_1(t) \in \mathcal{M}_1$ and $\tilde{s}_2(t) \in \mathcal{M}_2$, yielding a complexity order of $\mathcal{O}(N_t M_2)$.

IV. DATA-DRIVEN DNN-AIDED CHANNEL ESTIMATION AND DATA DETECTION

In this section, we first introduce DNN-aided channel estimation and data detection assisted SM in Section IV-A, followed by the proposed data-driven DNN architecture in Section IV-B.

A. CONVENTIONAL DNN ARCHITECTURE

The multi-layer fully-connected DNN architecture of Fig. 2 can be employed for replacing the conventional channel estimator and data detector discussed in Section III. More specifically, in time-invariant stationary scenarios, the channels $\hat{\mathbf{H}}(0)$ estimated from the pilot symbols can be expressed as

$$\hat{\mathbf{H}}(0) = \hat{\mathbf{H}}(1) = \dots = \hat{\mathbf{H}}(T) = \mathbf{Y}(0), \quad (9)$$

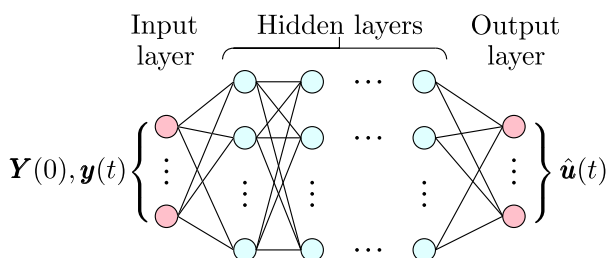


FIGURE 2. A fully-connected DNN architecture for channel estimation and data detection in SM system.

where $\mathbf{Y}(0) = [\bar{\mathbf{y}}(1), \bar{\mathbf{y}}(2), \dots, \bar{\mathbf{y}}(N_t)]$ represents the pilot symbols of N_t pilot TSs. Hence, as shown in Fig. 2, the received pilots $\mathbf{Y}(0)$ and the signal $\mathbf{y}(t)$ received during the t -th data TS constitute the inputs of the L -layer fully-connected DNNs, where each layer is comprised of $Z = 64$ nodes, yielding the output of

$$\hat{\mathbf{u}}(t) = f_{\text{sigmoid}}\{\mathbf{W}_L \dots f_{\text{Relu}}\{\mathbf{W}_2[f_{\text{Relu}}(\mathbf{W}_1 \mathbf{y}(t) + \mathbf{b}_2)] + \mathbf{b}_2\} + \dots + \mathbf{b}_L\}, \quad (10)$$

where \mathbf{W}_l and \mathbf{b}_l , $l = 1, \dots, L$, represent the weights and biases, respectively, of the l -th hidden layer of the DNNs of Fig. 2. Here, the rectified linear unit (Relu) function of $f_{\text{Relu}}(a) = \max(0, a)$ is employed for activating the DNN training, and the sigmoid function of $f_{\text{sigmoid}}(a) = \frac{1}{1+e^{-a}}$ is used at the output layer to obtain the detected bits $\hat{\mathbf{u}}(t)$. The network weights, which comprise the parameters that are being optimized in the DNN, can be expressed as

$$\boldsymbol{\theta} = \{\mathbf{W}_1, \mathbf{b}_1, \mathbf{W}_2, \mathbf{b}_2, \dots, \mathbf{W}_L, \mathbf{b}_L\}. \quad (11)$$

The training process optimizes the network weights $\boldsymbol{\theta}$ by minimizing the loss function. In this paper, the mean squared error (MSE) between the output bits $\hat{\mathbf{u}}(t)$ of the DNN and the transmitted bits $\mathbf{u}(t)$ in the t -th TS is adopted as the loss function, which can be expressed as

$$\varphi_1(\boldsymbol{\theta}) = \frac{1}{BT} \sum_{t=1}^T \sum_{i=1}^B \|\mathbf{u}_i(t) - \hat{\mathbf{u}}_i(t)\|^2, \quad (12)$$

where B is the size of a mini-batch.

The network weights $\boldsymbol{\theta}$ are updated for the batches and randomly picked up from the data samples, using the classic stochastic gradient descent (SGD) algorithm [35] expressed as

$$\boldsymbol{\theta}^+ := \boldsymbol{\theta} - \varepsilon \nabla \varphi_1(\boldsymbol{\theta}), \quad (13)$$

where ε is the learning rate of the SGD and $\nabla \varphi_1(\boldsymbol{\theta})$ represents the gradient of $\varphi_1(\boldsymbol{\theta})$. In this paper, $\varepsilon = 10^{-3}$ is selected for the performance characterization of both the DNN and of our proposed DeepSM architectures.

Note that during the training process, the popular adaptive moment estimation (Adam) optimizer is employed for the off-line learning. Our experiments show that 1.5×10^5 training samples are sufficient for the training set.

B. DEEPSM

Both the maximum likelihood detection as well as the conventional DNN architecture assume the CSI $\hat{\mathbf{H}}(0)$ to remain near constant during the data transmission period. In order to dispense with this idealized simplifying assumption, we now propose a novel DNN architecture, referred to as the DeepSM, which updates the estimated CSI $\hat{\mathbf{h}}_{\tilde{s}_1(t)}$ of the active TA at the t -th TS for all $t \in [1, T]$ and operates in a data-driven fashion, as shown in Fig. 3. Similar to the DNN of Fig. 2, the proposed DeepSM technique replaces the conventional channel estimation and data detection operations relying on channel-models by the DNN of Fig. 3. A joint

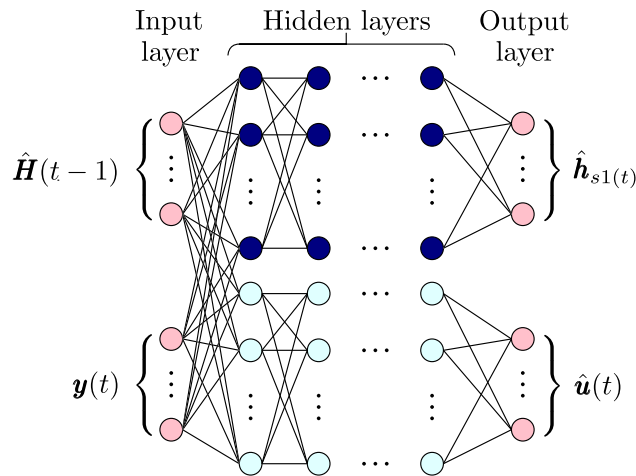


FIGURE 3. The proposed DeepSM architecture for joint channel estimation and data detection of SM.

data-driven channel estimation and data detection technique can be designed with the aid of the multiple-layer fully-connected DNNs, which updates the CSI at each TS.

More specifically, both the CSI matrix $\hat{\mathbf{H}}(t - 1)$, which is obtained from the DeepSM in the $(t - 1)$ -st TS, and the received data $\mathbf{y}(t)$ at the t -th TS are input to the DeepSM, resulting in a $[2N_t N_r + 2N_r]$ -node input layer. As shown in Fig. 3, the hidden layers of our DeepSM architecture are comprised of two subgroups. The upper subgroup comprising L_1 hidden layers is employed for updating $\hat{\mathbf{h}}_{\hat{s}_1(t)}$ at the t -th TS and $\hat{\mathbf{h}}_{\hat{m}(t)} = \hat{\mathbf{h}}_{\hat{m}(t-1)}$ for $\hat{m}(t) \neq \hat{s}_1(t) \in \mathcal{M}_1$, while the lower subgroup comprising L_2 hidden layers learns to detect the transmitted bits $\hat{\mathbf{u}}(t)$ at the t -th TS.

For DeepSM, the number of nodes employed in the hidden layers for both subgroups is fixed to $Z_1 = Z_2 = 64$, which is sufficiently high for attaining a superior BER performance for SM systems, while maintaining a moderate detection complexity. Additionally, $L_1 = L_2 = 3$ hidden layers are selected after the trial experiments.

The operations of DeepSM are summarized in Algorithm 1. More specifically, the DeepSM relies on the following steps. To start with, in the $t = 1$ -st TS, the real and imaginary parts of each element in $\hat{\mathbf{H}}(0)$ are input to the DNN via the $2N_t N_r$ -node input layer of the upper subgroup in DeepSM. In the mean time, the real and imaginary parts of the elements in the signal vector $\mathbf{y}(1)$ received during the $t = 1$ -st TS are input to the lower subgroup of the DeepSM via the $2N_r$ -node input layer of the lower subgroup. Following this, given the inputs of $\hat{\mathbf{H}}(0)$ and $\mathbf{y}(1)$, the lower subgroup of the proposed DeepSM is activated to obtain the output $\hat{\mathbf{u}}(1)$. In the meantime, the upper subgroup updates the CSI $\hat{\mathbf{h}}_{\hat{s}_1(1)}$ at the $t = 1$ -st TS. Note that $\hat{s}_1(1) = V_{SSK}[\hat{\mathbf{u}}(1)]$ is determined by the output $\hat{\mathbf{u}}(1)$ of the lower subgroup, which indicates the active TA index in the $t = 1$ -st TS. The output of the proposed DNN architecture can be expressed as

$$\hat{\mathbf{h}}_{\hat{s}_1(t)} = \mathbf{W}_{L_1}^{(1)} \cdots f_{\text{Relu}} \left\{ \mathbf{W}_2^{(1)} \left[f_{\text{Relu}} \left(\mathbf{W}_1^{(1)} \mathbf{H}(t - 1) \right. \right. \right.$$

$$\left. \left. + \mathbf{b}_1^{(1)} \right) \right] + \mathbf{b}_2^{(1)} \left. \right\} + \cdots + \mathbf{b}_{L_1}^{(1)}, \quad (14)$$

$$\hat{\mathbf{u}}(t) = f_{\text{sigmoid}} \left\{ \mathbf{W}_{L_2}^{(2)} \cdots f_{\text{Relu}} \left\{ \mathbf{W}_2^{(2)} \left[f_{\text{Relu}} \left(\mathbf{W}_1^{(2)} \mathbf{y}(t) \right. \right. \right. \right. \left. \left. \left. + \mathbf{b}_1^{(2)} \right) \right] + \mathbf{b}_2^{(2)} \right. \right. \left. \left. + \cdots + \mathbf{b}_{L_2}^{(2)} \right\} \right\}, \quad (15)$$

where $\mathbf{W}_{l_1}^{(1)}$ and $\mathbf{b}_{l_1}^{(1)}$, $l_1 = 1, \dots, L_1$, represent the weights and biases of the l_1 -th hidden layer of the upper subgroup, respectively, whereas $\mathbf{W}_{l_2}^{(2)}$ and $\mathbf{b}_{l_2}^{(2)}$, $l_2 = 1, \dots, L_2$ are the weights and biases of the l_2 -th hidden layer of the lower subgroup, respectively.

The network weights of DeepSM may be expressed as

$$\theta^{(1)} = \left\{ \mathbf{W}_1^{(1)}, \mathbf{b}_1^{(1)}, \mathbf{W}_2^{(1)}, \mathbf{b}_2^{(1)}, \dots, \mathbf{W}_{L_1}^{(1)}, \mathbf{b}_{L_1}^{(1)} \right\}, \quad (16a)$$

$$\theta^{(2)} = \left\{ \mathbf{W}_1^{(2)}, \mathbf{b}_1^{(2)}, \mathbf{W}_2^{(2)}, \mathbf{b}_2^{(2)}, \dots, \mathbf{W}_{L_2}^{(2)}, \mathbf{b}_{L_2}^{(2)} \right\}. \quad (16b)$$

After the output $\hat{\mathbf{h}}_{s_1(1)}$ of the $t = 1$ -st TS is obtained, $\hat{\mathbf{h}}_{s_1(1)}$ will be employed for updating the $s_1(1)$ -th column of $\hat{\mathbf{H}}(0)$, obtaining the updated CSI in the $t = 1$ -st TS, which can be expressed as

$$\hat{\mathbf{H}}(1) = [\hat{\mathbf{h}}_{1(0)}, \dots, \hat{\mathbf{h}}_{s_1(1)}, \dots, \hat{\mathbf{h}}_{N_T(0)}]. \quad (17)$$

Later in the t -th ($t > 1$) TS, the same operations as those in the $t = 1$ -st TS are performed, obtaining $\hat{\mathbf{h}}_{s_1(t)}$ and $\hat{\mathbf{u}}(t)$.

Algorithm 1 DeepSM for Joint Channel Estimation and Data Detection of SM

Input:

CSI matrix $\hat{\mathbf{H}}(0)$, received data symbols $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(T)$.

Output:

Detected bits $\hat{\mathbf{u}}(1), \hat{\mathbf{u}}(2), \dots, \hat{\mathbf{u}}(T)$; Estimated CSI $\hat{\mathbf{h}}_{s_1(1)}, \hat{\mathbf{h}}_{s_1(2)}, \dots, \hat{\mathbf{h}}_{s_1(T)}$.

- 1: **for** $t = 1, \dots, T$ **do**
- 2: Activate the proposed DeepSM of Fig. 3 with the inputs, $\mathbf{y}(t)$ and $\hat{\mathbf{H}}(t - 1)$;
- 3: Use the DeepSM with the optimized parameters of (16b) to obtain detected bits $\hat{\mathbf{u}}(t)$;
- 4: Use the DeepSM with the optimized parameters of (16a) to obtain $\hat{\mathbf{h}}_{\hat{s}_1(t)}$;
- 5: Update the $\hat{s}_1(t)$ -th column in $\hat{\mathbf{H}}(t - 1)$ with $\hat{\mathbf{h}}_{\hat{s}_1(t)}$;
- 6: **end for**

During the training process summarized in Algorithm 2, the MSE is employed as the loss function, which can be expressed as

$$\varphi_2(\theta^{(1)}, \theta^{(2)}) = \frac{1}{BT} \sum_{t=1}^T \sum_{i=1}^B \left\| \mathbf{u}_i(t) - \hat{\mathbf{u}}_i(t) \right\|^2 + \frac{1}{BT} \sum_{t=1}^T \sum_{i=1}^B \left\| \mathbf{h}_{i,s_1(t)} - \hat{\mathbf{h}}_{i,s_1(t)} \right\|^2. \quad (18)$$

Algorithm 2 Training of the DeepSM**Input:**

CSI matrix $\hat{\mathbf{H}}(0)$ and received data samples $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(T)$;
 Perfect CSI $\mathbf{h}_{s_1(1)}, \mathbf{h}_{s_1(2)}, \dots, \mathbf{h}_{s_1(T)}$ and transmitted bits $\mathbf{u}(1), \mathbf{u}(2), \dots, \mathbf{u}(T)$.

Output:

Trained network weights parameters $\theta^{(1)}$ and $\theta^{(2)}$.

Initialization:

Randomize initial weights $\theta^{(1)}$ and $\theta^{(2)}$.

- 1: **repeat**
- 2: **for** $t = 1, \dots, T$ **do**
- 3: Use samples of $\hat{\mathbf{H}}(t-1)$ and $\mathbf{y}(t)$ to update $\theta^{(1)}$ and $\theta^{(2)}$ by minimizing $\varphi_2(\theta^{(1)}, \theta^{(2)})$ of (18);
- 4: Replace the $s_1(t)$ -th column in $\hat{\mathbf{H}}(t-1)$ with $\mathbf{h}_{s_1(t)}$;
- 5: **end for**
- 6: **until** Convergence of $\theta^{(1)}$ and $\theta^{(2)}$

V. PERFORMANCE RESULTS

In this section, we characterize the performance of the conventional DNN and of the proposed DeepSM architectures for SM channel estimation and data detection in terms of the associated loss function, BER performance and complexity, in Sections V-A to V-C, respectively.

A. LOSS FUNCTION

First, Fig. 4 shows the loss function vs. the number of epochs of the proposed DeepSM architecture employing $L_1 = L_2 = 1, 2, 3$ or 4 hidden layers, with each layer adopting $Z_1 = Z_2 = 64$ nodes, where a single epoch refers to a complete training session relying on a training data set. The SM system considered in Fig. 4 employs $N_t = 4$ TAs and $N_r = 4$ RAs using binary phase shift keying (BPSK) modulation and communicates over time-varying Rayleigh fading channels associated with $\alpha = 0.98$. It can be observed from Fig. 4 that a single or a pair of hidden layers, i.e. $L_1 = L_2 = 1$ or 2 are insufficient for the proposed DNN architecture, whereas the employment of $L_1 = L_2 = 4$ layers may result in over-fitting, hence requiring slightly more epochs to achieve convergence of the loss function. Hence, $L_1 = L_2 = 3$ hidden layers are adopted for the remaining simulations.

Furthermore, we investigate the influence of the modulation order on the loss function φ_2 of the proposed DNN architecture for the channel estimation and data detection of the SM system for transmission over time-varying Rayleigh fading channels at SNR= 25 dB, where $N_t = 4, N_r = 4, \alpha = 0.99, 0.98$ or 0.97 , and BPSK or quadrature phase shift keying (QPSK) are employed. Here we adopt $L_1 = L_2 = 3$ hidden layers with each layer comprising $Z_1 = Z_2 = 64$ nodes. As shown in Fig. 5, a lower number of epochs are required for the convergence of the loss function, when BPSK is employed.

Additionally, observe in Fig. 5 that there is a rapid initial convergence leading to a plateau, followed by a second rapid

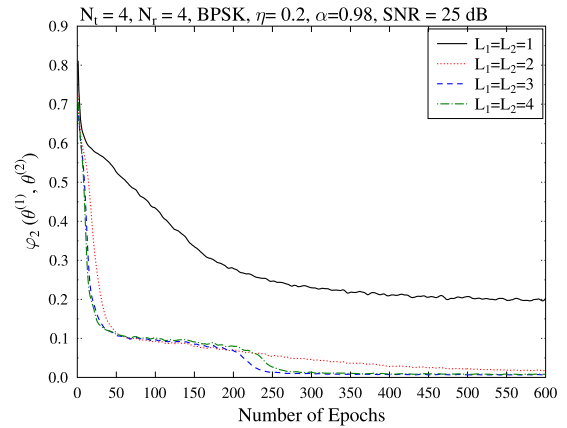


FIGURE 4. Loss function of the proposed DeepSM for the channel estimation and data detection of the SM system employing $N_t = 4$ TAs and $N_r = 4$ RAs over time-varying Rayleigh fading channels with different number of hidden layers under SNR= 25 dB, where $\eta = 0.2, \alpha = 0.98$ and BPSK are employed.

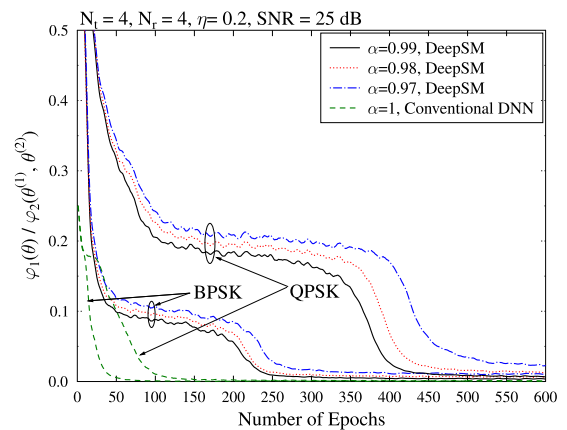


FIGURE 5. Loss function of the conventional DNN architecture for the channel estimation and data detection of the SM system employing $N_t = 4$ TAs and $N_r = 4$ RAs over time-variant and that of the proposed DeepSM over time-varying Rayleigh fading channels with $\alpha = 0.99, 0.98$ and 0.97 under SNR= 25 dB, where $\eta = 0.2$ and BPSK or QPSK are employed.

convergence phase. The reason lies in that the loss function of (18) is comprised of the sum of two MSEs, which corresponds to the output $\hat{\mathbf{H}}(t)$ and $\hat{\mathbf{u}}(t)$, respectively.

B. BER PERFORMANCE

In this section, the BER performance of the proposed DeepSM channel estimation and data detection is investigated in Fig. 6 over time-invariant Rayleigh fading channels, while in Figs. 7 and 8 for transmission over time-varying Rayleigh fading channels¹ and in Fig. 9 over quantized time-varying Rayleigh fading channels, where low-resolution ADCs are employed.

To elaborate, Fig. 6 characterizes the BER performance of the conventional DNNs for the joint channel estimation and

¹The block Rayleigh fading is assumed, where the fading coefficient is fixed for each block and SNR is taken as an average over all blocks.

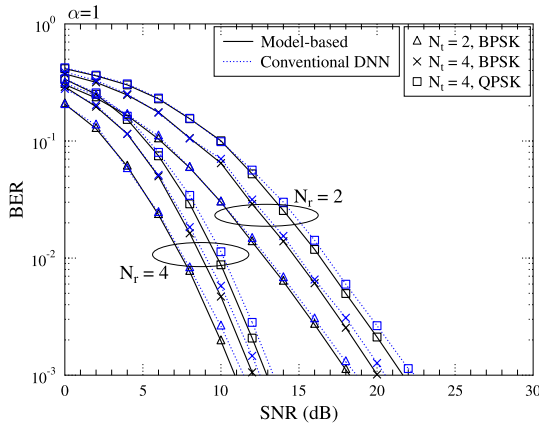


FIGURE 6. Uncoded BER performance of the conventional DNNs for the data detection of the SM system over time-invariant Rayleigh fading channels employing $N_r = 2$ or 4 and $N_t = 2$ or 4.

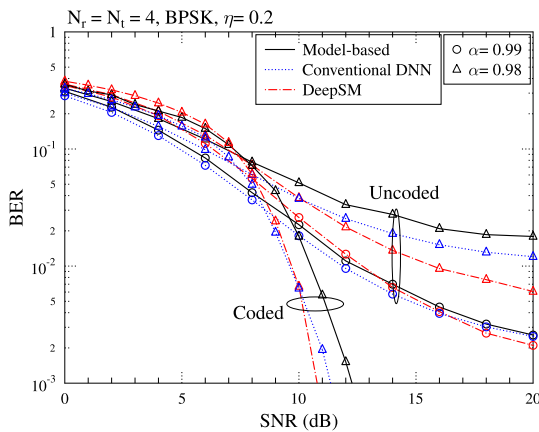


FIGURE 7. Uncoded and LDPC-coded BER performance of the model-based, conventional DNN and DeepSM architecture for the joint channel estimation and data detection of the SM systems employing $N_t = 4$ TAs and $N_r = 4$ RAs over time-varying Rayleigh fading channels under imperfect CSI with $\eta = 0.2$.

data detection scheme of our SM system for transmission over time-invariant channels employing $N_r = 2$ or 4 and $N_t = 2$ or 4, which shows a similar performance to that of the conventional approach discussed in Section III. Hence, a direct application of the DNNs is capable of achieving near model-based channel estimation and data detection performance for the SM channel estimation and data detection over time-invariant channels.

However, since the conventional DNN architecture fails to update the CSI at each TS in realistic time-varying channels, it fails to facilitate reliable detection, hence resulting in a poor BER performance. By contrast, the proposed DeepSM outperforms the conventional DNN architecture in the high-SNR region, which is an explicit benefit of exploiting the CSI at each TS. This is shown in Figs. 7 and 8, where the uncoded SM system employs $N_t = 4$ TAs and $N_r = 4$ RAs for communicating over time-varying Rayleigh fading channels under imperfect CSI in conjunction with $\eta = 0.2$ and 0.1. Furthermore, by comparing Figs. 7 and 8, we can see the

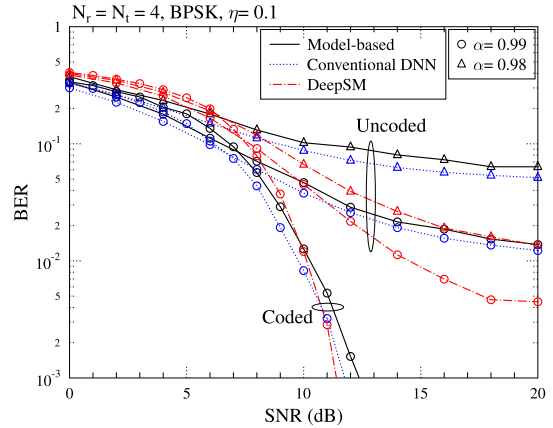


FIGURE 8. Uncoded and LDPC-coded BER performance of the model-based, conventional DNN and DeepSM architecture for the joint channel estimation and data detection of the SM systems employing $N_t = 4$ TAs and $N_r = 4$ RAs over time-varying Rayleigh fading channels under imperfect CSI with $\eta = 0.1$.

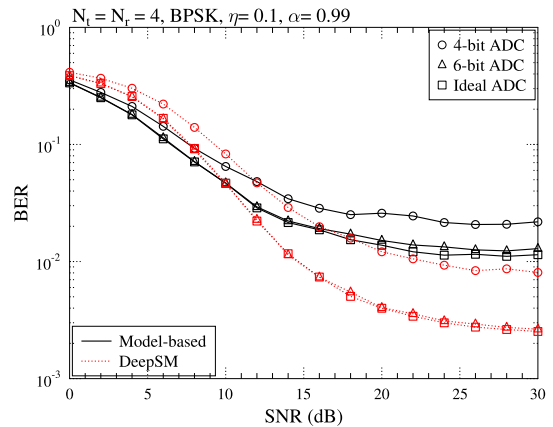


FIGURE 9. Uncoded BER performance of the conventional model-based and the DeepSM for the joint channel estimation and data detection of the SM systems employing $N_t = 4$ TAs and $N_r = 4$ RAs over the time-invariant and time-varying quantized Rayleigh fading channels with 4-bit or 6-bit ADCs where $\eta = 0.1$ and $\alpha = 0.99$.

influence of η . Naturally, having a higher percentage of pilot symbols in a frame results in an improved BER performance. Hence, for an uncoded system, where the target BER lies between 10^{-2} to 10^{-3} , the DeepSM advocated shows the best performance among the three approaches.

Furthermore, the residual BER can be mitigated by forward error correction (FEC) codes, as demonstrated in Figs. 7 and 8, where a half-rate low-density parity-check (LDPC) code is employed. We can see from both figures that for the LDPC-coded SM systems, DeepSM still achieves the best BER performance at high SNRs.

Fig. 9 further characterizes the BER performance of both the conventional model-based and of the DeepSM aided joint channel estimation and data detection in the context of SM systems for transmission over fading channels, where 4- or 6-bit ADCs are employed by the receiver for quantizing the received signal. The SM system of Fig. 9 adopts the same

TABLE 2. MATLAB Runtime (ms) comparison of conventional SM channel estimation and data detection, conventional DNN as well as DeepSM, where $N_r = 4$.

Parameters	Maximum-likelihood $\mathcal{O}(N_t M_2)$	Conventional DNN $\mathcal{O}(Z^2)$	DeepSM $\mathcal{O}(Z_1^2 + Z_2^2)$
$N_t = M_2 = 2$	4.79	3.62	5.85
$N_t = 4, M_2 = 2$	8.42	3.96	7.52
$N_t = M_2 = 4$	15.88	4.09	7.76

parameters as those of Fig. 7. We can see that our DeepSM exhibits a lower error floor than the conventional channel estimation and data detection approach in the case of low-cost 4-bit or 6-bit ADCs. Additionally, Fig. 9 shows that the BER performance of the DeepSM scheme employing a 4-bit ADC approaches that of an ideal ADC.

C. COMPLEXITY

Now we characterize the complexity when employing the different channel estimation and data detection approaches conceived in this paper, as shown in Table 2. In particular, the complexity order of the conventional maximum-likelihood approach, conventional DNN-aided approach as well as DeepSM are $\mathcal{O}(N_t M_2)$, $\mathcal{O}(Z^2)$ and $\mathcal{O}(Z_1^2 + Z_2^2)$, respectively. Specifically, the complexity order of the maximum-likelihood approach is determined by the modulation order, whereas the complexity order of the DNN-aided approaches are determined by the size of the hidden layer in the DNNs. This indicates that when the DNN architecture is fixed, the modulation order will not influence the detection complexity of either conventional DNN or DeepSM much.

Additionally, we can see from Table 2 that the MATLAB runtime [36], [37] of the maximum-likelihood approach increases significantly when either N_t or M_2 increases. By contrast, when DNN-aided approaches are employed, the MATLAB runtime can be greatly reduced and remains similar as the increase of modulation order N_t or M_2 , since the DNN structure allows the parallel operations in MATLAB, significantly reducing the communication latency. Additionally, by comparing the two DNN-aided approaches, we can see that the DeepSM requires a higher running time due to the higher size of hidden layers, but achieves significantly better BER performance over the time-varying channels, as demonstrated in Section V-B.

VI. CONCLUSION AND FUTURE RESEARCH

We have first used the conventional DNN for joint channel estimation and signal detection in SM systems communicating over time-invariant channels. Then, a DeepSM structure relying on a pair of DNN subgroups has been proposed for channel estimation and data detection in SM systems communicating over time-varying fading channels. Our studies and simulation results have shown that the

conventional DNN based detector is capable of achieving a similar BER performance to that of the model-based channel estimator and detector over idealized time-invariant channels. However, the proposed DeepSM outperforms both the model-based approach and the conventional DNN structure, even when communicating over time-varying and non-linear channels, since the proposed DeepSM estimates the CSI at each TS.

Our future work will investigate data-driven DNN-aided detection in multiuser communications in the scenarios of non-coherent as well as coherent mMTC and Internet of Things. Furthermore, reducing the gap to the capacity by iterative detection and decoding approaches is worth investigating.

REFERENCES

- [1] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.
- [2] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.
- [3] P. Yang, M. Di Renzo, Y. Xiao, S. Li, and L. Hanzo, "Design guidelines for spatial modulation," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 6–26, 1st Quart., 2015.
- [4] E. Basar, M. Wen, R. Mesleh, M. Di Renzo, Y. Xiao, and H. Haas, "Index modulation techniques for next-generation wireless networks," *IEEE Access*, vol. 5, pp. 16693–16746, 2017.
- [5] H. S. Hussein, M. Elsayed, U. S. Mohamed, H. Esmail, and E. M. Mohamed, "Spectral efficient spatial modulation techniques," *IEEE Access*, vol. 7, pp. 1454–1469, 2019.
- [6] M. Wen, B. Zheng, K. J. Kim, M. Di Renzo, T. A. Tsiftsis, K.-C. Chen, and N. Al-Dhahir, "A survey on spatial modulation in emerging wireless systems: Research progresses and applications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 1949–1972, Sep. 2019.
- [7] C. Xu, N. Ishikawa, R. Rajashekar, S. Sugiura, R. G. Maunder, Z. Wang, L.-L. Yang, and L. Hanzo, "Sixty years of coherent versus non-coherent tradeoffs and the road from 5G to wireless futures," *IEEE Access*, vol. 7, pp. 178246–178299, 2019.
- [8] Z. Wang and J. Cao, "NOMA-based spatial modulation," *IEEE Access*, vol. 5, pp. 3790–3800, 2017.
- [9] Y. Liu, L.-L. Yang, and L. Hanzo, "Spatial modulation aided sparse code-division multiple access," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1474–1487, Mar. 2018.
- [10] Y. Liu, L.-L. Yang, and L. Hanzo, "Joint user-activity and data detection for grant-free spatial-modulated multi-carrier non-orthogonal multiple access," *IEEE Trans. Veh. Technol.*, early access, Aug. 14, 2020, doi: 10.1109/TVT.2020.3016898.
- [11] 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.

- [12] D.-T. Phan-Huy, Y. Kokar, K. Rachedi, P. Pajusco, A. Mokh, T. Magounaki, R. Masood, C. Buey, P. Ratajczak, N. Malhouroux-Gaffet, J.-M. Conrat, J.-C. Prévotet, A. Ourir, J. De Rosny, M. Crussière, M. Héliard, A. Gati, T. Sarrebourg, and M. Di Renzo, "Single-carrier spatial modulation for the Internet of Things: Design and performance evaluation by using real compact and reconfigurable antennas," *IEEE Access*, vol. 7, pp. 18978–18993, 2019.
- [13] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3187–3196, May 2020.
- [14] A. E. Canbilen, E. Basar, and S. S. Ikki, "Reconfigurable intelligent surface-assisted space shift keying," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1495–1499, Sep. 2020.
- [15] Y. A. Chau and S.-H. Yu, "Space modulation on wireless fading channels," in *Proc. IEEE 54th Veh. Technol. Conf. (VTC Fall)*, vol. 3, Oct. 2001, pp. 1668–1671.
- [16] J. Wang, S. Jia, and J. Song, "Generalised spatial modulation system with multiple active transmit antennas and low complexity detection scheme," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1605–1615, Apr. 2012.
- [17] S. Yang and L. Hanzo, "Fifty years of MIMO detection: The road to large-scale MIMOs," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1941–1988, 4th Quart., 2015.
- [18] Z. Qin, H. Ye, G. Y. Li, and B.-H.-F. Juang, "Deep learning in physical layer communications," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 93–99, Apr. 2019.
- [19] M. Mehrabi, M. Mohammadkarimi, M. Ardakani, and Y. Jing, "Decision directed channel estimation based on deep neural network k -step predictor for MIMO communications in 5G," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2443–2456, Nov. 2019.
- [20] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: The road to Pareto-optimal wireless networks," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1472–1514, 3rd Quart., 2020.
- [21] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, "Deep learning for wireless communications: An emerging interdisciplinary paradigm," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 133–139, Aug. 2020.
- [22] Y. Yang, F. Gao, X. Ma, and S. Zhang, "Deep learning-based channel estimation for doubly selective fading channels," *IEEE Access*, vol. 7, pp. 36579–36589, 2019.
- [23] X. Ma and Z. Gao, "Data-driven deep learning to design pilot and channel estimator for massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5677–5682, May 2020.
- [24] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M.-S. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9223–9228, Aug. 2020.
- [25] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.
- [26] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Model-driven deep learning for MIMO detection," *IEEE Trans. Signal Process.*, vol. 68, pp. 1702–1715, 2020.
- [27] 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.
- [28] A. Tato, C. Mosquera, P. Henarejos, and A. Pérez-Neira, "Neural network aided computation of mutual information for adaptation of spatial modulation," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2809–2822, May 2020.
- [29] 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.
- [30] F. Shu, L. Liu, Y. Zhang, G. Xia, X. Liu, J. Li, S. Jin, and J. Wang, "A Deep-learning-based joint inference for secure spatial modulation receiver," 2019, *arXiv:1907.02215*. [Online]. Available: <http://arxiv.org/abs/1907.02215>
- [31] H. Albinsaid, K. Singh, S. Biswas, C.-P. Li, and M.-S. Alouini, "Block deep neural network-based signal detector for generalized spatial modulation," *IEEE Commun. Lett.*, early access, Aug. 11, 2020, doi: [10.1109/LCOMM.2020.3015810](https://doi.org/10.1109/LCOMM.2020.3015810).
- [32] B. Shamasundar and A. Chockalingam, "A DNN architecture for the detection of generalized spatial modulation signals," *IEEE Commun. Lett.*, early access, Aug. 20, 2020, doi: [10.1109/LCOMM.2020.3018260](https://doi.org/10.1109/LCOMM.2020.3018260).
- [33] J. Jegathanan, A. Ghraryeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3692–3703, Jul. 2009.
- [34] N. Ishikawa, R. Rajashekar, C. Xu, M. El-Hajjar, S. Sugiura, L.-L. Yang, and L. Hanzo, "Differential-detection aided large-scale generalized spatial modulation is capable of operating in high-mobility millimeter-wave channels," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 6, pp. 1360–1374, Oct. 2019.
- [35] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. COMPSTAT*. Cham, Switzerland: Springer, 2010, pp. 177–186.
- [36] T. V. Luong, Y. Ko, N. A. Vien, D. H. N. Nguyen, and M. Matthaiou, "Deep learning-based detector for OFDM-IM," *IEEE Wireless Commun. Lett.*, vol. 8, no. 4, pp. 1159–1162, Aug. 2019.
- [37] T. Van Luong, Y. Ko, N. A. Vien, M. Matthaiou, and H. Q. Ngo, "Deep energy autoencoder for noncoherent multicarrier MU-SIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3952–3962, Jun. 2020.



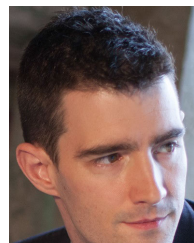
LUPING XIANG (Member, IEEE) received the B.Eng. degree (Hons.) from Xiamen University, China, in 2015, and the Ph.D. degree from the University of Southampton, in 2020. He is currently a Research Fellow with the Next Generation Wireless Group, University of Southampton. His research interests include mMTC, URLLC, machine learning, and channel coding.



YUSHA LIU received the B.Eng. degree in communication engineering from Xiamen University, China, and the M.Sc. degree in wireless communications from the University of Southampton, U.K., where she is currently pursuing the Ph.D. degree with the Next Generation Wireless Group. Her research interests include spatial modulation, non-orthogonal multiple access, and physical-layer security for wireless communications.



THIEN VAN LUONG (Member, IEEE) received the B.S. degree from the Hanoi University of Science and Technology, Vietnam, in 2015, and the Ph.D. degree from Queen's University Belfast, U.K., in 2019. In 2016, he was a Research Assistant with the Singapore University of Technology and Design. He is currently a Research Fellow with the University of Southampton, U.K. His research interests include quantum communications and applied machine learning in wireless signal processing.



ROBERT G. MAUNDER (Senior Member, IEEE) received the degree from the School of Electronics and Computer Science, University of Southampton, U.K., in October 2000, the B.Eng. degree (Hons.) in electronic engineering, in July 2003, and the Ph.D. degree in telecommunications, in December 2007. He began a lectureship in November 2007 and was promoted to an Associate Professor in March 2013 and to a Professor in August 2017. In addition to his role in the School

of Electronics and Computer Science at the University of Southampton, he is the Founder and CTO of AccelerComm Ltd., which is commercializing his research as soft-IP. His research interests include joint source/channel coding and the holistic design of algorithms and hardware implementations for wireless communications. He has published a number of IEEE articles in these areas. He was awarded a Chartered Engineer status of the IET, in November 2013, and a Fellow status of the IET, in January 2017.



LIE-LIANG YANG (Fellow, IEEE) received the B.Eng. degree in communications engineering from Shanghai TieDao University, Shanghai, China in 1988, and the M.Eng. and Ph.D. degrees in communications and electronics from Northern Jiaotong University, Beijing, China, in 1991 and 1997, respectively. From June 1997 to December 1997, he was a Visiting Scientist with the Institute of Radio Engineering and Electronics, Academy of Sciences, Czech Republic. Since

December 1997, he has been with the University of Southampton, U.K., where he is currently a Professor of Wireless Communications with the School of Electronics and Computer Science. His research interests include wireless communications, wireless networks, and signal processing for wireless communications, as well as molecular communications and nano-networks. He has published more than 390 research articles in journals and conference proceedings, authored or coauthored three books and also published several book chapters. He is a Fellow of the IET. He served as an Associate Editor to the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and *Journal of Communications and Networks (JCN)*. He is also an Associate Editor to IEEE ACCESS and a Subject Editor to the *Electronics Letters*. He was a Distinguished Lecturer of the IEEE VTS.



LAJOS HANZO (Fellow, IEEE) received the D.Sc. degree from the Technical University of Budapest, in 2009, and the Honorary Doctorate degree from the University of Edinburgh, in 2015. He is the Head of the Next Generation Wireless Group, University of Southampton. He is currently a Foreign Member of the Hungarian Academy of Sciences and the former Editor-in-Chief of the IEEE Press. He has served as Governor of both IEEE ComSoc and of VTS. He has published more than 1900 contributions at IEEE Xplore and 19 Wiley-IEEE Press books. He has helped the fast-track career of 119 Ph.D. degree students. More than 40 of them are Professors at various stages of their careers in academia and many of them are leading scientists in the wireless industry. He is a Fellow of REng, IET, and EURASIP.

...