Near-Instantaneously Adaptive Learning-Assisted and Compressed Sensing-Aided Joint Multi-Dimensional Index Modulation

Xinyu Feng, *Student Member, IEEE,* Mohammed EL-Hajjar, *Senior Member, IEEE,* Chao Xu, *Senior Member, IEEE,* and Lajos Hanzo, *Life Fellow, IEEE*

***Abstract*—Index Modulation (IM) is capable of striking an attractive performance, throughput and complexity trade-off. The concept of Multi-dimensional IM (MIM) combines the benefits of IM in multiple dimensions, including the space and frequency dimensions. On the other hand, IM has also been combined with compressed sensing (CS) for attaining an improved throughput. In this paper, we propose Joint MIM (JMIM) that can utilize the time-, space- and frequency-dimensions in order to increase the IM mapping design flexibility. Explicitly, this is the first paper developing a jointly designed MIM architecture combined with CS. Three different JMIM mapping methods are proposed for a space- and frequency-domain aided JMIM system, which can attain different throughput and diversity gains. Then, we extend the proposed JMIM design to three dimensions by combining it with the time domain. Additionally, to circumvent the high detection complexity of the proposed CS-aided JMIM design, we propose Deep Learning (DL) based detection. Both Hard-Decision (HD) as well as Soft-Decision (SD) detection are conceived. Additionally, we investigate the adaptive design of the proposed CS-aided JMIM system, where a learning-based adaptive modulation configuration method is applied. Our simulation results demonstrate that the proposed CS-aided JMIM (CS-JMIM) is capable of outperforming its CS-aided separate-domain MIM counterpart. Furthermore, the learning-aided adaptive scheme is capable of increasing the throughput while maintaining the required error probability target.**

***Index Terms*—Index Modulation (IM), Compressed Sensing-aided Multi-Dimensional Index Modulation (CS-MIM), Soft-Decision Detection, Machine Learning, Neural Network.**

NOMENCLATURE

Acronym

Approx-Log-MAP Approximate Logarithm MAP
AWGN Additive White Gaussian noise
BER Bit-Error Ratio
CE Channel Estimation
CoE Cost Efficiency
CP Cyclic Prefix
CS Compressed Sensing
CS-GFIM-SM CS-aided Generalized SIM with SM
CS-JMIM Compressed Sensing-aided JMIM
CS-MIM CS-aided MIM
CS-STFIM CS-Aided Space-Time-Frequency Index Modulation
CSI Channel State Information
DL Deep Learning
DM Dispersion Matrix
DNN Deep Neural Network
EE Energy Efficiency
FC Fully-Connected
FD Frequency Domain
FFT Fast Fourier Transform
GSFIM Generalized Space-and-Frequency IM
GSIM Generalized SIM
HD Hard-Decision
IFFT Inverse Fast Fourier Transform
IM Index Modulation
JIM Joint Index Mapping
JMIM Joint Multi-dimensional Index Modulation
KNN K-nearest Neighbour
LLR Log-Likelihood Ratio
MAP Maximum A Posteriori
MIM Multi-dimensional Index Modulation
MIMO Multiple-In and Multiple-Out
ML Maximum Likelihood
mmWave Millimeter Wave
MSE Mean Square Error
OFDM Orthogonal Frequency Division Multiplexing
OFDM-IM OFDM with Index Modulation
PSK Phase Shift Keying
QAM Quadrature Amplitude Modulation
QSF-IM Quadrature Space-Frequency IM
QSM Quadrature SM
RA Receive Antenna
RD Real Domain
Relu Rectified Linear Unit
RF Radio Frequency
RSC Recursive Systematic Convolutional
SD Soft-Decision
SE Spectral Efficiency
SF Space-Frequency
SFSK Space-Frequency Shift Keying
SIM Subcarrier-Index Modulation
SM Spatial Modulahttps://www.zhihu.com/tion
SNR Signal-to-Noise Ratio
SpD Spatial Domain
ST Space-Time

The authors are with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: xf2u18,meh,cx1g08,lh@ecs.soton.ac.uk).

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/X01228X/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

STSK	Space-Time Shift Keying
TA	Transit Antenna
TD	Time Domain
TS	Time Slot
TSF	Time-Space-Frequency
VD	Virtual Domain

Symbols

N_r	number of RAs
N_t	number of TAs
N_c	number of subcarrier for each frame
N_f	number of subcarrier per group in FD
N_{vt}	number of TAs in VD
N_v	number of subcarriers per group in VD
G	number of group
b	total bits number of each frame
b_g	bits number per group
K	number of activated index in JMIM matrix
Q	number of total realizations of JMIM mapping DMs
$\boldsymbol{S}$	ST block
$\boldsymbol{D}_q$	The q-th DM of Q total JIM DMs
$\boldsymbol{D}_i$	Selected DM of Q total JIM DMs
$\boldsymbol{\chi}$	M-ary PSK or QAM constellation
x	Selected constellation symbol from $\boldsymbol{\chi}$
$\boldsymbol{S}_g$	combined signal after JMIM and PSK/QAM
gs	number of sub-groups in grouped JIM methods of JMIM
gsx	number of split groups in virtual FD
gsy	number of split groups in virtual spatial domain
N_q	column length of coded JMIM DM
n_q	number of active subcarriers in each column of coded JMIM
T	number of TS for transmitting in TD
T_v	number of TS for transmitting in VD
gsz	number of split groups in virtual TD
$\boldsymbol{s}$	JMIM vector after block creator in VD
$\boldsymbol{A}$	CS measurement matrix
$\boldsymbol{s}^{(RD)}$	JMIM vector after CS
$\boldsymbol{h}_\alpha$	channel model in TD
$\boldsymbol{H}_\alpha$	channel model in FD
$\boldsymbol{Y}$	Received signal
$\boldsymbol{W}$	Additive White Gaussian noise
N_0	noise power of AWGN
$\bar{\mathcal{D}}(\beta)$	all the possible realizations of the JMIM DM
$\boldsymbol{W_n}$	weights of n-th neuron
$\bar{\mathcal{X}}$	all the possible realizations of the selected PSK/QAM symbol
$\boldsymbol{\theta_n}$	bias of n-th neuron
B	sample size of current iteration of DNN training phase
L_a	de-interleaved LLR sequence of L_e
L_e	output extrinsic LLR after soft demodulation
$\mathbf{c}$	coded bit sequence of information bits
$\mathbf{u}$	information bit sequence stream
$\mathcal{X}_{\gamma,\beta}$	the PSK/QAM symbol at the β-th CS-JMIM DM
$\mathcal{X}_1^l$	subset of the legitimate equivalent signal $\mathcal{X}$ corresponding to bit $u_l = 1$
$\mathcal{X}_0^l$	subset of the legitimate equivalent signal $\mathcal{X}$ corresponding to bit $u_l = 0$
ξ	the SNR value of a symbol with a BER lower than the target BER value for adaptive modulation
$\mathcal{T}$	training sets of SNR with different the adaptive mode
N_p	total number of instantaneous SNR values with BER under the target

I. Introduction

Index Modulation (IM) [1] can be considered as an energy-efficient candidate for next-generation wireless systems as a benefit of its flexible resource activation [2]. The concept of IM has been derived from that of Spatial Modulation (SM), which is a low-complexity Multiple-In and Multiple-Out (MIMO) scheme capable of striking a flexible performance *vs.* complexity trade-off using a single Radio Frequency (RF) chain [2]–[4]. Then, the concept of SM has also been extended to the frequency and time dimensions, where the philosophy of IM has been proposed [5], [6]. In the Frequency Domain (FD), the IM combined with Orthogonal Frequency Division Multiplexing (OFDM) is referred to as Subcarrier-IM (SIM), where only a fraction of the subcarriers is activated for signal transmission and the index of active subcarriers conveys extra information bits [7]. The effective signal power of the subcarriers activated in the FD is amplified, without increasing the time domain signal power after Inverse Fast Fourier Transform (IFFT). This results in a higher Signal-to-Noise Ratio (SNR) for the modulated symbols without requiring extra radiated power. Then, Tsonev *et al.* [8] proposed an enhanced SIM and Basar *et al.* [9] conceived a novel IM-aided OFDM (OFDM-IM) scheme for increasing the spectral efficiency. However, subcarrier-index modulated OFDM suffers from significant throughput reduction compared to the classic OFDM due to the deactivation of a number of subcarrers. Hence, Zhang *et al.* [10] proposed an improved SIM concept relying on Compressed Sensing (CS) [11], which benefits from the sparsity of symbols in the FD by compressing the sparse transmit vector [12].

To further increase the overall performance, Datta *et al.* proposed the concept of Generalized SIM (GSIM) and proved that Generalized Space-and-Frequency IM (GSFIM) achieves better performance than MIMO-OFDM. Their solution conveyed extra information in the SM part compared to GSIM [13]. However, the detection complexity of GSFIM escalates. Hence, Chakrapani *et al.* [14] proposed a message passing based low-complexity detection method for reducing the complexity of GSFIM detection. Furthermore, inspired by the SM and Quadrature SM (QSM) concepts [15], Quadrature Space-Frequency IM (QSF-IM) was proposed in [16], which applies a twin-antenna constellation for the in-phase and quadrature-phase transmission, in order to increase the throughput without extra energy consumption. Hence this solution struck a compelling Spectral Efficiency (SE), Energy Efficiency (EE) and Cost Efficiency (CE) trade-off.

Furthermore, several researchers considered the design of Multi-Dimensional Index Modulation (MIM) relying on both the Spatial Domain (SpD) and FD. For example, Space-Frequency Shift Keying (SFSK) [17] relies on an SFSK Dispersion Matrix (DM), which achieves beneficial transmit diversity in rapidly time-varying channels. Space-Time Shift

TABLE I: Contrasting our contributions to the literature

Contribution	proposed*	[10]	[24]	[22]	[25]	[26]	[27]	[28]	[29]	[30]
Index modulation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CS at transmitter	✓	✓		✓						
Learning aided detector	✓		✓		✓			✓	✓	✓
Soft decision detector	✓		✓				✓		✓	
Adaptive design for index modulation	✓									✓
Multi-dimensional index modulation	✓			✓					✓	
Joint index mapping design	✓			✓					✓	
3-Dimensional joint index modulation	✓									

Keying (STSK) constitutes another multi-functional MIMO technique in the family of MIM. It combines the Time Domain (TD) and the SpD and it is capable of striking a beneficial diversity versus multiplexing trade-off [18]. More specifically, in STSK, Q DMs are designed for spreading the signal over T Time Slots (TSs) and M Transmit Antennas (TA) in the TD and the SpD, respectively. Furthermore, the IM design activates one out of the Q DMs for transmission, hence $\log_2 Q$ extra IM bits may be conveyed. By appropriately adjusting these parameters, improved Bit Error Ratio (BER), throughput and complexity trade-offs may be struck [19].

Additionally, the concept of MIM was proposed in [20], which is capable of improving the degrees of freedom, hence achieving all the benefits of the IM concept in several domains without introducing extra deployment costs, such as extra RF chains or transmission power. Furthermore, Lu *et al.* [21] proposed Compressed-Sensing-Aided Space-Time Frequency Index Modulation (CS-STFIM) to combine CS techniques with STSK and OFDM-IM, which is an MIM system concept that inherits the benefits of both STSK and OFDM-IM. As a further advance, SM was also integrated into this MIM scheme for TA selection in [22]. In [6], the concept of multi-functional layered SM was proposed, which offers flexible trade-offs in terms of performance, hardware cost and power dissipation.

However, in previous MIM schemes, the index selection was performed separately in each dimension. By contrast, in this paper, we extend this concept to a Joint MIM system, where we jointly designs the IM in several dimensions. More specifically, the degrees of freedom of the IM mapping design is increased by harnessing multiple dimensions, which leads to a more flexible trade-off between the throughput, power efficiency, and cost. In this case, both SFSK and STSK can be considered as special cases of the proposed joint MIM (JMIM) family. JMIM may also be combined with CS techniques for increasing the spectral efficiency.

However, the joint detection of multiple dimensions leads to massive computational complexity at the receiver side. More specifically, conventional Maximum Likelihood (ML) detection, suffers from a rapidly escalating complexity upon increasing in the number of dimensions [31]. Coherent detection also requires the accurate knowledge of Channel State Information (CSI) at the receiver side, which leads to a substantial pilot overhead [32] as well as to a high Channel Estimation (CE) complexity [33], [34]. In [22], CS-aided MIM (CS-MIM) was presented, where multiple detection stages were required for recovering the data from the constituent CS, STSK, OFDM-IM and SM schemes. As a result, near-capacity operation can only be achieved, when Soft-Decision (SD) detection is used [35], but again, the complexity of MIM detection escalates with the number of IM dimensions.

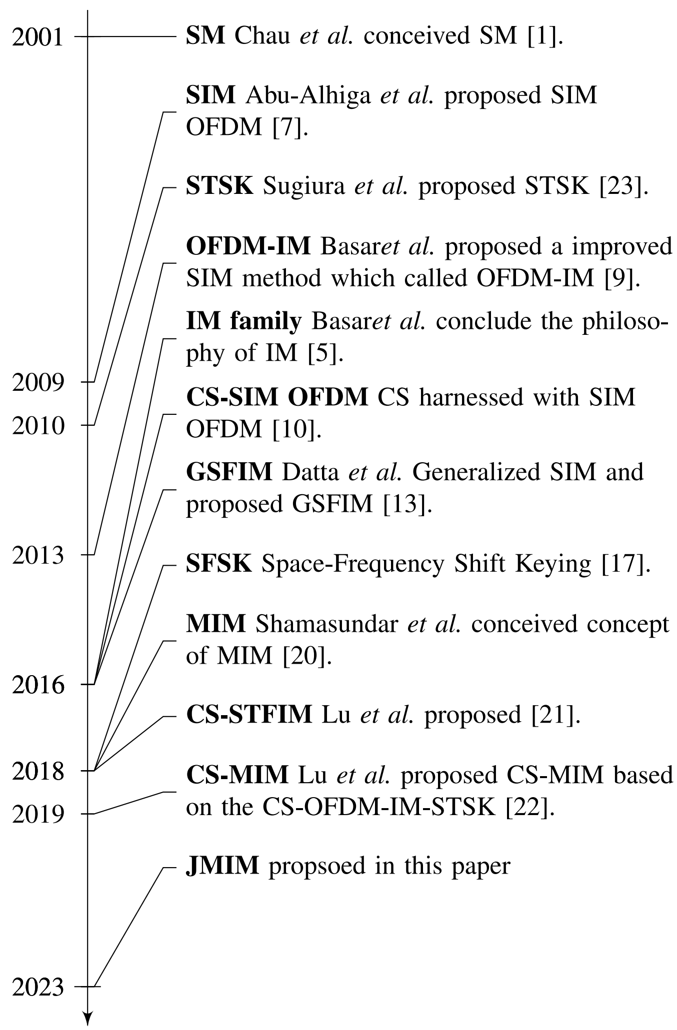


Fig. 1: Milestones of the index modulation family from single dimensional index modulation to MIM.

Recently, learning-based detection has been used as an efficient tool for reducing the complexity of detection, while dispensing with the requirement of explicit CSI estimation [36]. In [37], a Deep Neural Network (DNN) based model is proposed for detecting the OFDM-IM signal, while the authors of [38] and [39] harnessed convolutional neural networks for IM detection, when the CSI is available at the input of the detector. By contrast, blind learning based detection was designed for Millimeter Wave (mmWave) IM in [28] and for multi-set STSK in [29]. However, the authors of [29] only investigated the combination of basic SD and Deep Learning (DL). In [36], both DNN-based Hard-Decision (HD) and iterative SD assisted blind detection have been proposed for CS-MIM.

Additionally, given the flexibility of our CS-aided JMIM (CS-JMIM) design, we can adapt the JMIM mapping to hostile time-varying channel environments to improve the attainable performance. Hence, the concept of adaptive modulation can be intrinsically amalgamated with CS-JMIM to improve the attainable throughput, while maintaining a specific target BER. Yang *et al.* proposed machine learning aided adaptive SM [40], while Liu *et al.* [41] conceived learning-assisted IM for mmWave communications. In their follow-on research, they further developed the work by considering CE employing sparse Bayesian leaning for accurate CSI estimation [42].

Table I boldly contrasts the novelty of this paper to the literature. More explicitly, the contributions of this paper can be further detailed as follows:

1) We propose the CS-JMIM system concept and present several JMIM mapping matrix designs. Then, we demonstrate that the proposed JMIM mapping design is capable of striking an attractive trade-off between diversity and throughput.
2) We propose a DL-based HD detection aided CS-JMIM system that can achieve near-ML performance, while imposing significantly reduced complexity. Furthermore, we propose a DNN-aided SD detector for the proposed CS-JMIM that is capable of achieving near-capacity performance.
3) Both a K-nearest neighbour (KNN) algorithm based and a DL-assisted adaptive modulation scheme is proposed for CS-JMIM. We demonstrate that the learning-assisted adaptive CS-JMIM scheme is capable of selecting more appropriate CS-JMIM mapping design for transmission than its conventional threshold-based adaptive counterparts. Hence it can obtain a significant throughput gain over the conventional threshold-based adaptive method.
4) Our simulation results demonstrate that the proposed learning-based detector is capable of approaching the performance of the conventional coherent detection techniques at a reduced detection complexity. We also provide the associated capacity and throughput analysis, for characterising the trade-off between each mapping matrix and the benefits of the learning-assisted adaptive method.

The rest of the paper is organized as follows. In Section II, the system model of CS-JMIM is presented. In Section III, we characterize both HD and SD based learning-aided detectors. Then, in Section IV we present our proposed adaptive system design. In Section V, we present our simulation results, while our conclusions are offered in Section VI.

II. SYSTEM MODEL

In this section, we introduce the transceiver model of the proposed CS-JMIM system employing N_t TAs and N_r Receive Antennas (RAs). Fig. 2 shows the block diagram of the CS-JMIM system considered, where b bits are equally divided into G groups. We consider OFDM having Nc subcarriers, which are then split into G groups and each group has $N_f = N_c/G$ subcarriers in the FD[1], while N_{vt} TAs and N_v subcarriers of each group are applied for the CS-JMIM system in the Virtual Domain (VD) [2]. To be more specific, in each subcarrier group, there are N_v available subcarrier indices within the VD, where the dimension N_v of the VD is larger than the dimension N_f of the FD. Similarly, N_{vt} antennas in the VD are larger than the N_t antennas of the SpD. For each group of b bits as $b_g(g = 1, 2, \cdots, G)$, b_g^1 bits are used for generating K Phase Shift Keying/Quadrature Amplitude Modulation PSK/QAM symbols, while the remaining b_g^2 bits are mapped to the JMIM mapping matrix selector, which chooses a specific mapping matrix out of Q JMIM matrices. Then, these K PSK/QAM codewords and the selected JMIM mapping DM are combined to generate a Space-Time (ST) block $\boldsymbol{S}$. Afterwards, the block creator of Fig. 2 collects all codewords from the G groups for forming a frame, which is mapped to multiple index domains by the carrier index mapper, followed by the CS method and OFDM modulation, as shown in Fig. 2. Then, after transmission over the wireless channel, the receiver estimates the channel and detects the signal. At the receiver side, the signal is transformed back to the subcarrier symbols and each JMIM group signal is detected separately.

In the following, we present the details of the processing stages at the transmitter and the receiver. In this case, we only focus our attention on a single group instead of G groups, since the same procedure is applied to all groups, as shown in Fig.2. The transmitter model is introduced in Section II-A, followed by the receiver model in Section II-B.

A. Transmitter

As shown in Fig. 2, b bits are split into G groups, where the b_g bits, $(g = 1, 2, 3...G)$ of each group are split into two parts by the block splitter: b_g^1 bits are used for JMIM mapping matrix selection and b_g^2 bits for the classic PSK/QAM. In the following we explain in detail the Joint Index Mapping (JIM) part of the CS-JMIM transmitter of Fig. 2.

1) Joint Index Mapping: As shown in Fig.2, the N_c subcarriers of the OFDM symbol are divided into G groups of size N_f, with $N_f = N_c/G$. For each b_g group of bits, the first part b_g^1 is used for selecting the active DM from the Q candidates $\boldsymbol{D}_1, \boldsymbol{D}_2, \cdots, \boldsymbol{D}_q, \cdots, \boldsymbol{D}_Q$ with $\boldsymbol{D}_q \in \mathbb{C}^{N_v \times N_{vt}}, q = 1, 2, \cdots, Q$. The second part is used for determining the constellation symbol, which is employed for modulating the active DM. The classic constellation symbol is then selected from a M-ary PSK or QAM constellation $\boldsymbol{\chi}$.

Let us denote the selected DM and the selected constellation symbol, respectively, by $\boldsymbol{D}_i, i \in \{1, \cdots, Q\}$ and $x, x \in \chi$. Then the combined signal in group g can be expressed by

$$\boldsymbol{S}_g = x\boldsymbol{D}_i, g = 1, \cdots, G. \quad (1)$$

In the following, we introduce three designs of the DMs. Firstly, to leverage the multi-dimensionality of MIM systems,

[1]FD is the OFDM symbol domain after CS processing, as shown in Fig. 2.

[2]VD is the actual domain. This concept was firstly introduced in [10] to illustrate the CS techniques in IM systems to improve the spectral efficiency.

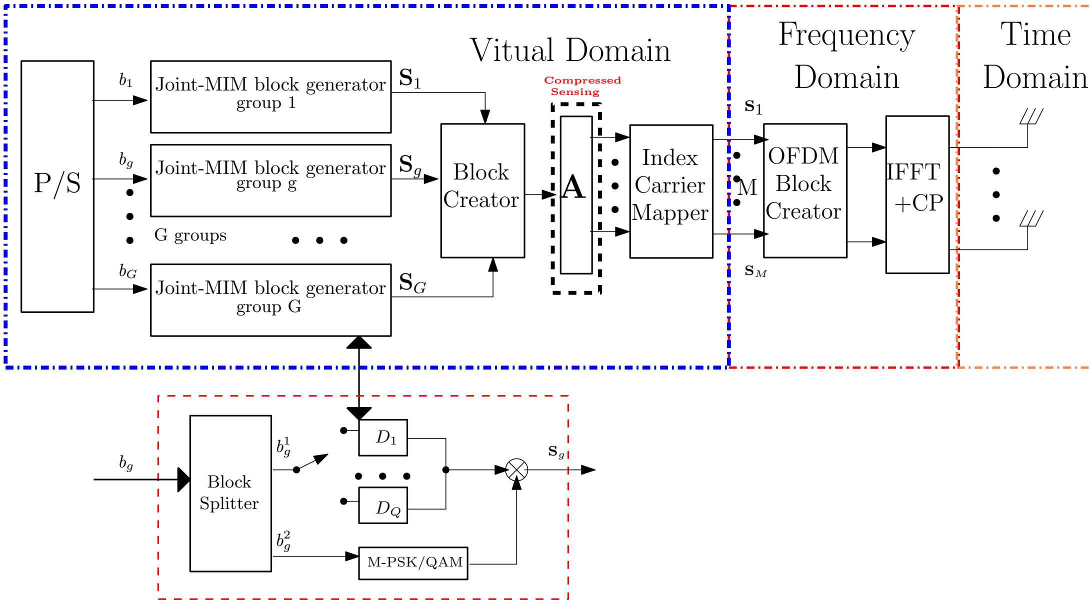


Fig. 2: CS-JMIM system transmitter block diagram.

the design of IM encompasses all dimensions. Then, the activation of the corresponding indices is guided by the coordinates of these joint dimensions, which is detailed in the following section in the context of a DM design referred to as General JIM. Secondly, to strike a design trade-off between the throughput and diversity gain attained, we can further split the joint multi-dimensional matrix into sub-group matrices, where different general JIM DMs can be selected for each sub-group matrix. We refer to this mapping design as Grouped JIM, which is further detailed in the following sections. Additionally, we introduce a coded DM design for achieving a high diversity gain, which is detailed in the following sections. Furthermore, we start a discussion considering the Space-Frequency (SF) dimensions and then we present a 3-dimensional mapping design for the Time-Space-Frequency (TSF) dimensions of JMIM.

a) General Joint Index Mapping: As JIM, first we consider joint SF DM design. The index is selected based on both dimensions' coordinates. We assign N_{vt} TAs and N_v subcarriers to a specific group, which results in $N_{vt}N_v$ possible active positions and to a total of $C(N_tN_{vt}, K)$ legitimate realizations. As an example, let us consider having $K = 2$ active subcarriers and $N_{vt} = 2, N_v = 2$ for each group. Then, we have $b_g^1 = \lfloor \log_2 C(N_{vt}N_v, K) \rfloor = \lfloor \log_2 C(4, 2) \rfloor = 2$ bits for selecting $K = 2$ active subcarriers out of 4 available subcarriers in each group, since we have $2^2 = 4$ legitimate combinations which equivalent to $Q = 4$ DMs, as shown in Table II. Fig.3 shows a block diagram of the general JIM example presented in Table II, where the activated index is then combined with the QAM symbol by the multiplier to form the combined symbol S. *Furthermore, when compared to the CS-aided separate MIM system, CS-JMIM can attain comparable throughput as CS-MIM with significant sparsity.*

b) Grouped Joint Index Mapping: Given a substantial number of TAs, subcarriers, and a limited quantity of active index elements K in each group, most elements in the DM remain inactive, leading to diminished SE. To address this, we propose grouped JIM, which divides the DM matrix into smaller sub-group matrices, each adopting a general JIM. Furthermore, striking a trade-off between throughput and diversity involves choosing either the same or different DMs across groups. To elaborate further, applying the same DM across all groups results in multiple copies of the information bits, which produces a diversity gain. On the other hand, employing different DMs for each group improve the throughput.

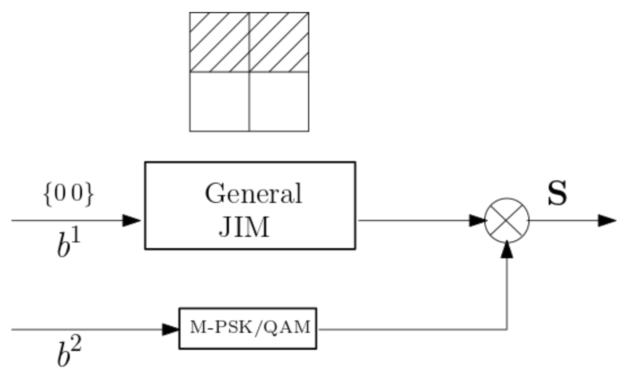


Fig. 3: Block diagram of the general JIM example in Table II with b_1=[0 0].

TABLE II: An example selection procedure of joint SF index selection in a CS-JMIM system having $K = 2, N_v = N_{vt} = 2$

b_2	matrix No.	Indices	Allocation
[00]	$\boldsymbol{D}_1$	$(1, 2)$	$\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}$
[01]	$\boldsymbol{D}_2$	$(1, 3)$	$\begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}$
[10]	$\boldsymbol{D}_3$	$(1, 4)$	$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
[11]	$\boldsymbol{D}_4$	$(2, 3)$	$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$

For example upon assuming $N_{vt}=4, N_v=4$ and $K=2$ for each groups DM results in $\boldsymbol{D}_q \in \mathbb{C}^{N_{vt}\times N_v}$. Then, we further split $\boldsymbol{D}_q$ into four equal sub-matrices expressed as

$$\boldsymbol{D}_q = \begin{bmatrix} \boldsymbol{D}_q^{1,1} & \boldsymbol{D}_q^{2,1} \\ \boldsymbol{D}_q^{1,2} & \boldsymbol{D}_q^{2,2} \end{bmatrix}, \tag{2}$$

where we have $\boldsymbol{D}_q^i \in \mathbb{C}^{N_{vt}/2\times N_v/2}, i=1,2,3,4$. For each sub-matrix $\boldsymbol{D}_q^{i,j}$, $(i=1,2,3...gsx), (j=1,2,3...gsy)$ general JIM can be applied. Here, gsx and gsy represent the number of sub-group's in the FD and SpD, respectively. In the above example, we can have a total of $gs = gsx \times gsy = 4$ sub-groups and $b_g^1 = [\log_2 C(4,2)] = 2$ bits for each sub-groups matrix. To maximize the throughput, four different sub-matrices can be aggregated to one DM $\boldsymbol{D}_q$ to obtain 8 bits in total. Fig.4 shows the block diagram of the grouped JIM, where we have four sub-groups of smaller general JMIM matrix. For a small general JMIM matrix we can apply $Q=4$ DMs in total, where we can assign 4×2 bits for all sub-groups. On the other hand, if four repeated sub-matrices are used, we can achieve similar structure of coded JMIM which will be discussed below.

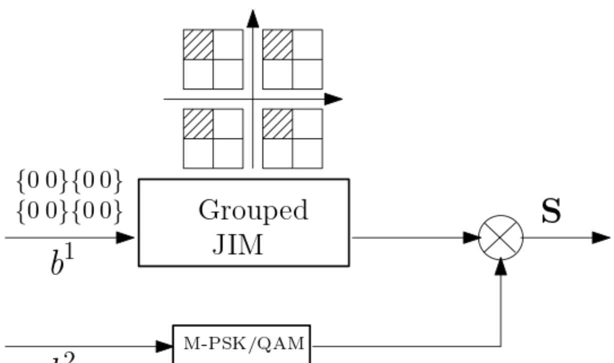


Fig. 4: Block diagram of a grouped JIM example with b_1=[0 0 0 0 0 0 0 0]

Subsequently, the grouped JIM optimally utilizes the available space of the VD matrix, albeit at the expense of sparsity. By adjusting the index mapping of each sub-group, it offers significant throughput or diversity gains. However, this leads to a substantial increase in detection complexity for conventional methods, such as the ML detector.

c) Coded Joint Index Mapping: Another way of further increasing the transmit diversity is to employ coded index mapping, where we use a circular shift based design of the DMs, which was proposed for SFSK in [17]. In this method, the number of active subcarriers in each column is n_q, with $N_q - n_q$ inactive subcarriers, where N_q is the column length of $\boldsymbol{D}_q$. Then, the second column is the circular down shift of the first column by one position. Similarly, other columns can be obtained based on the previous column distribution.

To elaborate a little further, using a 'toy' example, for $N_q = N_{vt} = 4, n_q = 2$, we can have $Q = C(N_q, n_q) = 6$ possible combinations, yielding $b_g^1 = \lfloor \log_2 C(N_q, n_q) \rfloor = 2$ bits. The following is an example of a circular shifting based DM:

$$D_1 = \begin{bmatrix} 1&0&0&1\\1&1&0&0\\0&1&1&0\\0&0&1&1 \end{bmatrix}, D_2 = \begin{bmatrix} 1&0&1&0\\0&1&0&1\\1&0&1&0\\0&1&0&1 \end{bmatrix},$$
$$D_3 = \begin{bmatrix} 1&1&0&0\\0&1&1&0\\0&0&1&1\\1&0&0&1 \end{bmatrix}, D_4 = \begin{bmatrix} 0&0&1&1\\1&0&0&1\\1&1&0&0\\0&1&1&0 \end{bmatrix},$$
$$D_5 = \begin{bmatrix} 0&1&0&1\\1&0&1&0\\0&1&0&1\\1&0&1&0 \end{bmatrix}, D_6 = \begin{bmatrix} 0&1&1&0\\0&0&1&1\\1&0&0&1\\1&1&0&0 \end{bmatrix}.$$

Given $b_g^1 = 2$ bits, then $2^2 = 4$ DMs are selected for the CS-JMIM system.

Fig.5 shows a block diagram of the coded JIM, where we can apply the first JIM DM for b1=[0 0] based on the code book used. *In this scenario, coded JIM offers the maximum diversity in the design of coded DMs, enabling reliable detection even in highly noisy environments.*

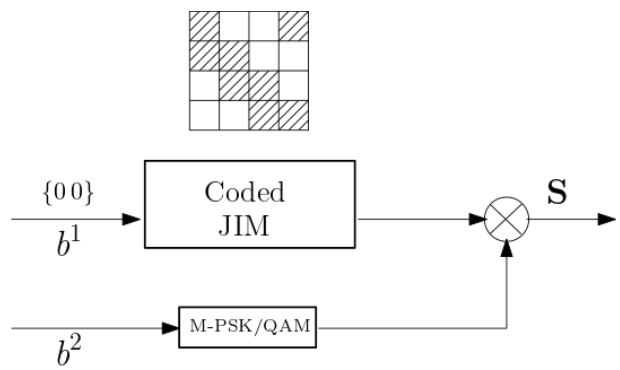
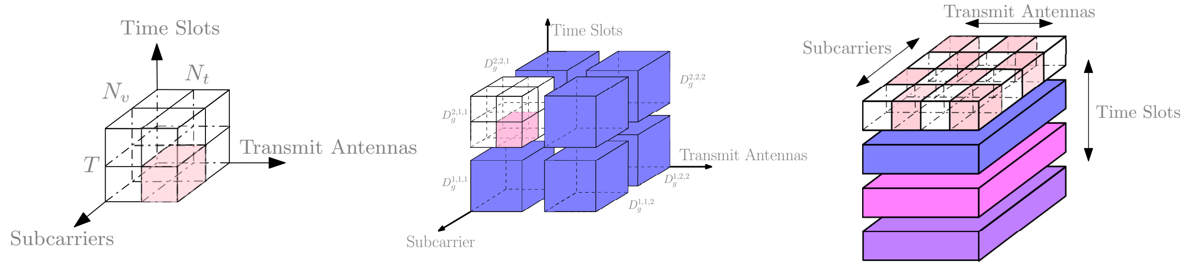


Fig. 5: Block diagram of a coded JIM example with b_1=[0 0]

d) 3-Dimensional Dispersion Matrix Design: In this design, the TD is introduced as an extra dimension for the JIM. We assume that T_v TSs are applied in the VD and T TSs are used in the TD, while we have $T_v > T$. Then, we can assign three-dimensional DMs $\boldsymbol{D}_q \in \mathbb{C}^{N_v\times N_{vt}\times T_v}$. In this case, the above-mentioned three mapping techniques can be applied.

Specifically, for the general JIM we may consider the following example for further illustration. Let $K=1$ and $N_{vt} = N_v = T_v = 2$ as shown in Fig. 6(a) and $b_g^1 = [001]$. More specifically, the three-dimensional matrix can be expressed in the coordinate form of (n_v, n_{vt}, t_v). In this case, given the IM bits $b_g^1 = [010]$, we activate the fourth element in a set of 8 elements in this three-dimensional matrix with the coordinates $(2,2,1)$ as shown in Fig. 6(a). Then, the number of bits of this JMIM applied for the DM selection becomes $b_g^1 = \lfloor \log_2 C(N_{vt}N_vT_v, K)\rfloor = \lfloor \log_2 C(8,1)\rfloor = 3$ bits.

Fig. 6(b) shows the structure of the grouped JIM applied in three dimensions. Similar to the SF matrix, the TSF matrix can be split into several equal sub-groups. As shown in Fig. 6(b), we assume $N_{vt} = N_v = T_v = 4$ and $K=1$ for each group's DM, which results in $\boldsymbol{D}_q \in \mathbb{C}^{N_v\times N_{vt}\times T_v}$. Then, we further split $\boldsymbol{D}_q$ into 8 equal sub-matrices. Each sub-group DM can



(a) Structure of general JIM DM, while having $K = 1$,$N_{vt} = 2$ TAs, $T_v = 2$ TSs, $N_v = 2$ subcarriers.

(b) Structure of group JIM DM, while having $K = 1$,$N_{vt} = 4$ TAs, $T_v = 4$ TSs, $N_v = 4$ subcarriers with 4 sub-group.

(c) Structure of coded JIM DM, while having $Q = 6$,$N_{vt} = 4$ TAs, $T_v = 4$ TSs, $N_v = 4$ subcarriers.

Fig. 6: Illustration of the structure for JIM DM in time-space-frequency domain

be expressed in the form of $\boldsymbol{D}_g^{gsx,gsy,gsz}$, where gsx, gsy, gsz represents the split size in the FD, the SpD and the TD, respectively. For each sub-matrix $D_g^{gsx,gsy,gsz}$, general JIM can be applied within a set of $gs = gsx \times gsy \times gsz = 8$ sub-group matrices. Then, we can have $\lfloor \log_2 C(8,1) \rfloor = 3$ bits for each sub-matrix. To maximize the throughput, we can also assign different information to each sub-group and then the 8 sub-matrices can be aggregated to form a single DM $\boldsymbol{D}_q$ to obtain $b_g^1 = gs\lfloor \log_2 C((N_v/gsx)(N_t/gsy)(T/gsz), K) \rfloor = 8\lfloor \log_2 C(8,1) \rfloor = 24$ bits for the JMIM design. Compared to the same DM size used in the general JIM, which has $b_g^1 = \lfloor \log_2 C(64,1) \rfloor = 6$ bits, the grouped JIM can provide a significant gain in the spectral efficiency. On the other hand, in order to attain a diversity gain, the sub-matrices can achieve maximum diversity gain, when all 8 sub-groups have the same active index.

Furthermore, for the coded JIM matrix design in three dimensions, the same method is applied for the first TS of the space-frequency matrix. Then, circular shifting is applied to the entire SF matrix to generate the next TS matrix with shifting by one position. As shown in Fig. 6(c), upon assuming $N_{vt} = N_v = T_v = 4$ for the DM size, as well as $N_q = n_q = 2$ for the activated subcarriers and $b^1 = [01]$, then the corresponding circular shifting based DM D_2 presented in the previous section is applied to the first TS of the 3D matrix. Then, we can generate each TS index mapping with the aid of a single position shifting, which can be represented as:

$$D_{t1} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, D_{t2} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix},$$
$$D_{t3} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, D_{t4} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

2) Compressed Sensing and Block Assembly: In order to exploit the sparsity of the JIM DM, CS is applied to all the dimensions of the joint multi-dimensional matrix symbol created by the block assembled to increase the throughput. As shown in Fig. 7, a matrix $\boldsymbol{S}_g$ associated with $N_{vt} = N_v = 4$ will be transformed from the matrix $\boldsymbol{S}, \boldsymbol{S} \in \mathbb{C}^{N_{vt} \times N_v}$ into the vector $\boldsymbol{s}, \boldsymbol{s} \in \mathbb{C}^{N_{vt} N_v \times 1}$.

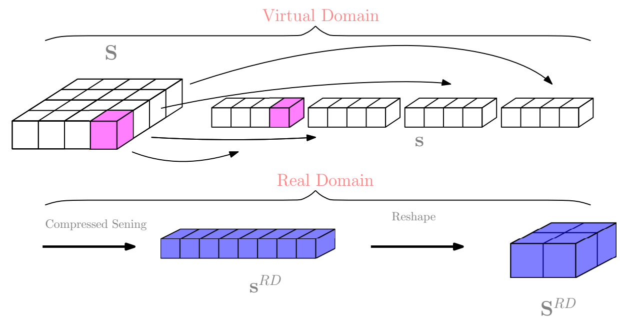


Fig. 7: Illustration of the process for compressing the JMIM DM in the SF domain with $N_{vt} = N_v = 4, K = 1$. Note that this example applies the general JIM with $b_g^1 = [0100]$.

The symbol vector $\boldsymbol{s}$ is then compressed by a CS measurement matrix $\boldsymbol{A} \in \mathbb{C}^{N_f N_t \times N_v N_{vt}}$ from the $N_v N_{vt}$-dimensional $\boldsymbol{s}$ in the VD into the $N_f N_t$-dimensional form in the Real Domain (RD)[3] denoted as $\boldsymbol{s}^{(RD)}$, which can be written as: $\boldsymbol{s}^{RD} = \boldsymbol{A}\boldsymbol{s}$. The RD vector $\boldsymbol{s}^{RD}$ after CS is then transferred into a compressed joint multidimensional symbol matrix $\boldsymbol{S}^{(RD)}$, where $\boldsymbol{S}^{(RD)} \in \mathbb{C}^{N_t \times N_f}$. Then, the index carrier mapper maps the corresponding joint multidimensional symbol elements to the OFDM subcarriers and the TAs to form the SF symbols. Afterwards, G groups of SF symbols $\boldsymbol{S}$ are assembled by the OFDM creator to a long SF symbol frame, as shown in Fig. 2. The RD SF symbol can be separated into N_t FD symbols, which means that N_t FD symbols are transmitted by N_t TAs. Similar to conventional OFDM, the FD symbol will be transformed into TD symbols to be transmitted by their corresponding TAs and then a Cyclic Prefix (CP) will be added. The G groups of SF symbols $\boldsymbol{S}$ are assembled by the block creator of Fig. 2 to form a long ST frame, which is processed by the ST mapper to output a symbol for transmission over multiple TAs and TSs, Equivalently, the ST symbols $\boldsymbol{S}$ of each subcarrier group are mapped to N_t TAs during T TSs, which have N_t symbol sequences $\{\boldsymbol{s}_1, ..., \boldsymbol{s}_{N_t}\}$ for transmission from the N_t TAs during each TS.

[3]RD is the joint dimension of DM after the CS process. For instance, the SF-based JMIM signal conveys more bits in the VD than in the RD.

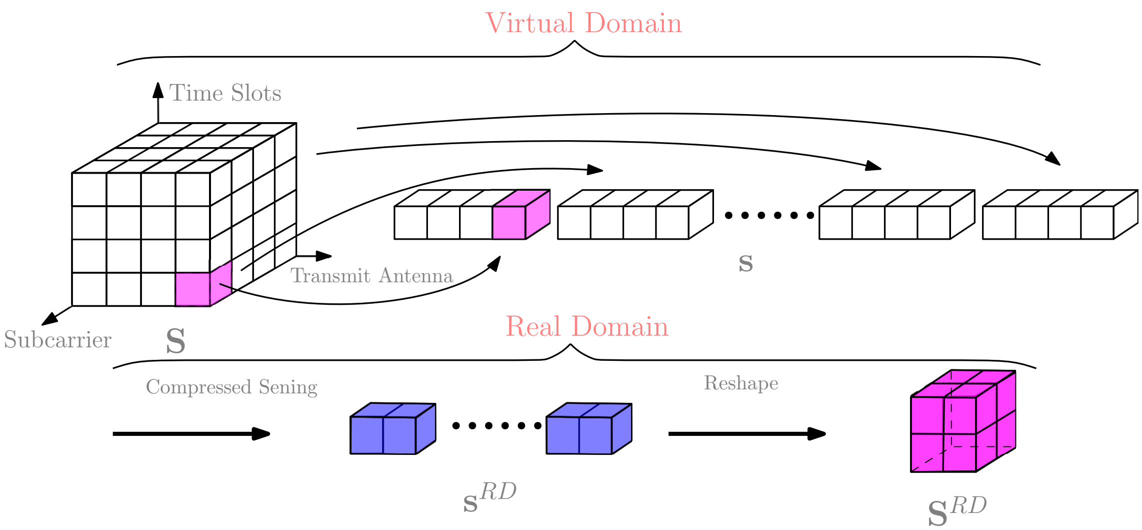


Fig. 8: Illustration of the process for compressing the JMIM DM in the TSF domain with $T = N_t = N_v = 4 K = 1$. Additionally, the example presented applies the general JIM for $b_g^1 = [000100]$.

For the three-dimensional JMIM, utilizing the TSF dimensions, the TD is also compressed by CS for improving the throughput, where T_v TSs are introduced in the VD for IM, complemented by T TSs in the TD. Specifically, for the general JMIM scheme, the TD is introduced for increasing the sparsity and for incorporating extra embedded information bits. As shown in Fig. 8, we apply CS to the TSF JMIM, where all the three dimensions are compressed for increasing the throughput. Specifically, a $(4 \times 4 \times 4)$-sized DM in the VD will be compressed to a $(2 \times 2 \times 2)$-sized DM of the RD. For example, when we have $T_v = N_{vt} = N_v = 4$, $b_g^1 = [000100]$ and $K = 1$, the element at the fourth subcarrier, fourth TA and first TS is activated, corresponding to the coordinate of $(4, 4, 1)$.

As for the coded JMIM scheme, additionally the TD is harnessed for further increasing the diversity gain, where CS is not considered for the TD. We assign either the same or different symbols in a sub-group matrix of the grouped JMIM scheme, which leads to a different CS approach. Given the different sub-group matrix symbols, the TD is exclusively harnessed for carrying extra copies of the symbol without CS. The design objective of this scheme is to increase the diversity gain.

B. Receiver Processing

As shown in Fig. 9, a receiver having N_r antennas is employed, where we assume that the transmitted signals are conveyed over a frequency-selective Rayleigh fading channel and the CSI is perfectly acquired at the receiver side. The G groups of signal are received by the receiver over N_r antennas and then the CP part of the received signals is removed. Finally, the processed signal is transformed into the FD by using the Fast Fourier Transform (FFT), as shown in Fig. 9.

The channel model can be expressed as $\boldsymbol{h}_\alpha \in \mathbb{C}^{N_r \times N_t}$, which represents the TD CSI between the N_t TAs and the N_r RAs. Then, the FD channel matrix can be expressed as $\boldsymbol{H}_\alpha \in \mathbb{C}^{N_r \times N_t}$ for $\alpha = 1, \ldots, M$, which are then

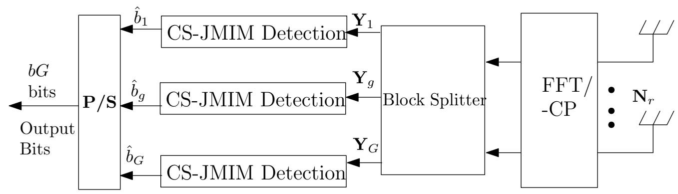


Fig. 9: CS-JMIM system receiver block diagram

split into G groups by the Block Splitter of Fig. 9. The symbols received by each subcarrier group are represented as $\boldsymbol{Y} = \{\boldsymbol{Y}[1], \ldots, \boldsymbol{Y}[\alpha], \ldots, \boldsymbol{Y}[N_f]\}$, with $\boldsymbol{Y} \in \mathbb{C}^{N_r \times N_f}$ and $\alpha = 1, 2, \cdots, N_f$.

As for the three-dimensional signal, the transmitted signal is mapped ST symbols, which are also collected by the receiver and split into G groups by the Block Splitter of Fig. 9. Afterwards, the symbols received in the three dimensions by each subcarrier group $\boldsymbol{Y} \in \mathbb{C}^{N_r \times M \times T}$ may be expressed as

$$\boldsymbol{Y} = \left\{ \begin{bmatrix} \boldsymbol{Y}_{1,1}^1 & \cdots & \boldsymbol{Y}_{1,N_f}^1 \\ \cdots & \cdots & \cdots \\ \boldsymbol{Y}_{1,1}^{N_r} & \cdots & \boldsymbol{Y}_{1,N_f}^{N_r} \end{bmatrix}_1, \ldots, \begin{bmatrix} \boldsymbol{Y}_{T,1}^1 & \cdots & \boldsymbol{Y}_{T,N_f}^1 \\ \cdots & \cdots & \cdots \\ \boldsymbol{Y}_{T,1}^{N_r} & \cdots & \boldsymbol{Y}_{T,N_f}^{N_r} \end{bmatrix}_T \right\}. \tag{3}$$

The received symbol of the t-th TS can be represented as $\boldsymbol{Y_t} = \{\boldsymbol{Y_t}[1], \ldots, \boldsymbol{Y_t}[\alpha], \ldots, \boldsymbol{Y_t}[N_f]\}$, with $\boldsymbol{Y_t} \in \mathbb{C}^{N_r \times T}$ and $\alpha = 1, 2, \ldots, N_f, t = 1, 2, \ldots, T$ characterizing the ST structure per group and the ST symbol received at the α-th subcarrier of each subcarrier group, respectively. Since the index is jointly decided in the multi-dimensional space, we can transform the ST symbol into a vectorial form $\boldsymbol{y}$ associated with $\boldsymbol{y} \in \mathbb{C}^{N_r N_f T \times 1}$.

Let the FD channel be $\mathbf{H}_\alpha \in \mathbb{C}^{N_r \times T}$ for $\alpha = 1, \ldots, N_f$. Then the signal $\boldsymbol{Y}_t[\alpha] \in \mathbb{C}^{N_r \times T} (\alpha = 1, \ldots, N_f)$ received during the T TSs for each subcarrier group can be expressed as [22]

$$\boldsymbol{Y}[\alpha] = \boldsymbol{H}_\alpha \boldsymbol{S}^{(\boldsymbol{RD})}[\alpha] + \boldsymbol{W}[\alpha], \tag{4}$$

where $\boldsymbol{S}^{RD}[\alpha] \in \mathbb{C}^{N_r \times T}$ denotes the ST symbols at the subcarrier α transmitted from the N_t TAs in the RD. Furthermore, $\boldsymbol{W}[\alpha] \in \mathbb{C}^{N_r \times T}$ represents the Additive White Gaussian noise (AWGN) obeying the distribution of $\mathcal{CN}(0, \sigma_N^2)$, and σ_N^2 is the noise variance.

III. CS-JMIM Signal Detection

Given the received signal model $\boldsymbol{Y}$ in (4), the receiver detects the information bits of the JMIM mapping matrix, which jointly conveys the index of the active subcarrier, the active TA and TS in the VD. Firstly, we reshape the received signal into a vectorial form $\boldsymbol{y}$ associated with $\boldsymbol{y} \in \mathbb{C}^{N_r N_f T \times 1}$.

The received signal $\boldsymbol{y}$ contains N_f ST symbols at N_f subcarriers in the FD of each subcarrier group. Then, we can rewrite $\boldsymbol{y}$ with the aid of (4) in the following form:

$$\boldsymbol{y} = \boldsymbol{H}\bar{\boldsymbol{A}}\bar{\boldsymbol{s}} + \boldsymbol{w}, \tag{5}$$

where $\bar{\boldsymbol{A}}$ is the equivalent measurement matrix $\boldsymbol{A}$ used for compressing the $\boldsymbol{s}$ VD vectors. In our three-dimensional CS-JMIM system, $\bar{\boldsymbol{A}}$ also compresses the TD, where we have $\bar{\boldsymbol{A}} \in \mathbb{C}^{N_v t N_v \times N_t N_f}$. Furthermore, $\boldsymbol{s} \in \mathbb{C}^{N_v N_v t T_v \times 1}$ denotes the vector of DM combined with the PSK/QAM symbol. In this case, we could rewrite $\bar{\boldsymbol{s}}$ in a matrix $\bar{\boldsymbol{S}}$ associated with $\bar{\boldsymbol{S}} = x\bar{\boldsymbol{D}}$, where $\bar{\boldsymbol{D}} \in \mathbb{C}^{N_v \times N_v t \times T_v}$ denotes the realization of the JMIM DM in each subcarrier group.

Conventional exhaustive search based maximum likelihood (ML) detection can be applied at the receiver, albeit this may lead to excessive complexity [5]. Furthermore, in the soft detection scenario, the received signal is converted into probability values, which are referred to as Log Likelihood Ratios (LLR) that are fed into the channel decoder for obtaining a near-capacity performance [43].

In the following section we present the conventional ML-based HD detector, followed by our proposed DNN aided HD detector, where the neural network replaces the exhaustive search by a learning-based classification model in order to significantly reduce the complexity. Afterwards, we discuss the SD detector, where we first present the conventional SD detectors followed by our learning-aided SD receiver.

A. Hard Decision Decoding

Again, we commence with the conventional ML-based detection of the CS-JMIM system, followed by the DNN-based detector.

1) Maximum Likelihood Detection: As shown in Fig. 9, we detect each group;s signal separately. In the CS-JMIM detector, according to the receiver model of (5), we have the modified joint JMIM and PSK/QAM symbol, which can be expressed as $\bar{\boldsymbol{S}} = \mathbf{x}\bar{\boldsymbol{D}}$. Here $\bar{\boldsymbol{D}}$ represents a specific realization of the selected JMIM DM and $\mathbf{x}$ represents K STSK PSK/QAM symbols. To detect the specific realization, we use $\bar{\mathcal{D}}(\beta)$ $(\beta = 1, 2, ..., N_{JMIM})$ to denote all the possible realizations of the JMIM DM. Furthermore, as there are $N_x = (X)^K$ realizations of $\mathbf{x}$, $\bar{\mathcal{X}}(\gamma)$ $(\gamma = 1, 2, ..., N_x)$ denotes all the possible realizations of the selected PSK/QAM symbol.

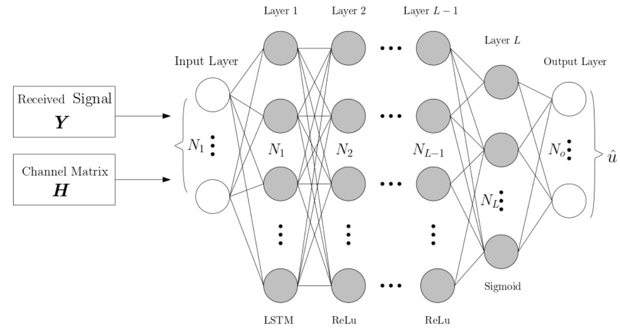


Fig. 10: Fully-connected DNN model for CS-JMIM HD detection.

The ML detector makes a joint decision concerning the JMIM DM and PSK/QAM with the aid of exhaustive search, which can be modelled as

$$\langle \hat{\gamma}, \hat{\beta} \rangle = \arg\min_{\gamma,\beta} \|\boldsymbol{Y} - \boldsymbol{H}\bar{\boldsymbol{A}}\mathcal{X}_{(}\gamma)\bar{\mathcal{I}}_D(\beta)\|^2, \tag{6}$$

where $\hat{\gamma}$ and $\hat{\beta}$ represent the estimates of the selected DM and the corresponding PSK/QAM constellation in each subcarrier group, respectively.

The excessively high search complexity of considering all possible candidates by the ML detector is given by $\mathcal{O}[N_{JMIM}(\mathcal{X})^K]$ per subcarrier group.

2) DNN-based Detection: To reduce the complexity of the ML detector, learning based detection is considered in this section, where a DNN based model is proposed for detecting the received CS-JMIM signal.

Detection may also be considered as a classification problem, where the corresponding bits of the harnessed CS-JMIM DM and PSK/QAM symbol constitute the DNN output. Under the assumption of perfect CSI at the receiver side, we use the received signal and the CSI as the input of the DNN model. The proposed DNN structure is shown in Fig. 10, where both the CSI $\boldsymbol{H}$ at the receiver and the received symbols $\boldsymbol{Y}$ constitute the inputs of the L-layer Fully-Connected (FC) network. Then, the output bits $\hat{\boldsymbol{u}}$ can be modelled as

$$\begin{aligned}\hat{u} =& f_{sigmoid}(\boldsymbol{W_n} ... f_{Relu}\{\boldsymbol{W_2}(f_{Relu_1}[\boldsymbol{W_1} f_{LSTM}(\boldsymbol{Y}) + \boldsymbol{\theta_1}]) \\ &+ \boldsymbol{\theta_2}\} + ... + \boldsymbol{\theta}_n),\end{aligned} \tag{7}$$

where $\boldsymbol{W_n}$ and $\boldsymbol{\theta_n}$, $n = 1, \cdots, L$ represent the weights and biases, respectively. In (7), the Rectified linear unit (Relu) function of $f_{Relu}(s) = max(0, s)$ is employed for activating the DNN during the training phase, while the sigmoid function of $f_{sigmoid}(s) = \frac{1}{1+e^{-s}}$ is used to obtain the detected bits $\hat{u}$. The raw input data represented in the complex-valued matrix form obtained from the received signal $\boldsymbol{Y}$ is vectorized first and then we rearrange the complex values by separately extracting the real as well as the imaginary parts and then merging them into a real-valued vector.

In the training phase, we employ randomly generated received signals, which are transmitted over a frequency selective Rayleigh fading channel for CS-JMIM. Afterwards, both the CSI and the received symbols are employed as the input data of the DNN. The number of training samples required is selected based on experimentation by gradually increasing the training size until acceptable mean square error (MSE) values are achieved. In this case, the MSE loss function of the DNN used for the training is

$$\mathcal{L}(\boldsymbol{u}, \hat{\boldsymbol{u}}; \boldsymbol{W_n}, \boldsymbol{\theta_n}) = \frac{1}{B}\sum_{i=1}^{B} \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|^2, \tag{8}$$

where B is the sample size of the current iteration. A stopping criterion can be defined either by the number of iterations or by an MSE threshold. Then, the parameter sets $\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\}$ can be updated in each training iteration based on our learning algorithm using gradient descent, which is formulated as

$$\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\} \leftarrow \{\boldsymbol{W_n}, \boldsymbol{\theta_n}\} - \alpha \nabla \boldsymbol{L}(\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\}),$$

where $\alpha > 0$ is the learning rate and $\nabla \boldsymbol{L}(\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\})$ represents the gradient of $\boldsymbol{L}(\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\})$. In our proposed network aided detection, we use $\alpha = 0.001$.

By the end of the training phase, the DNN has learnt the mapping from the received signal and stores both the weight as well as the bias information, which will be used for producing the desired outputs based on the input data in the testing phase. The statistical properties of the input/output data have to remain the same as those used during training.

The detection complexity of the learning algorithm is dominated by the calculation of the layer weights and biases, which may be considered to be of the order of $\mathcal{O}(n_i n_h) + \mathcal{O}(n_h^2) + \mathcal{O}(n_h n_o)$ [29], with n representing the number of neurons in each layer. Hence, the DNN complexity order is significantly lower than that of the ML detector.

B. *Soft Decision Decoding*

SD detection is employed for attaining near-capacity performance, when combined with channel coding. As the computational complexity of the maximum *a posteriori* probability in SD detector rapidly increases upon increasing the modulation order and the number of dimensions [44], the complexity of CS-JMIM rapidly becomes prohibitive, owing to the joint detection of JMIM signal in multiple dimensions. In the following, we present the conventional SD detector of CS-JMIM, followed by the correspond learning aided SD detector.

1) Conventional Soft Decision Detection: A channel coded CS-JMIM scheme is shown in Fig. 11, which was derived from the CS-MIM model of [22], [36] for achieving near-capacity performance. A Recursive Systematic Convolutional (RSC) encoder encodes the information bit sequence $\mathbf{b}$ followed by an interleaver, where the coded bit sequence $\mathbf{c}$ is interleaved to generate the stream $\mathbf{u}$ of Fig. 11. Then, the stream $\mathbf{u}$ is modulated in the CS-JMIM modulator of Fig. 2.

At the receiver side of Fig. 11, the received signal $\mathbf{Y}$ and CSI $\bar{\boldsymbol{H}}$ are input to the soft CS-JMIM that outputs LLRs. The LLRs output from the demodulator are then passed to the de-inteleaver and the RSC decoder performs soft decoding. In Fig. 11, $L(\cdot)$ represents the LLRs of the bit sequences, where $L_e(u)$ is the output extrinsic LLR after soft demodulation and $L_a(c)$ is the de-interleaved LLR sequence of $L_e(u)$.

The LLR of a bit is defined as the ratio of probabilities associated with the logical bits '1' and '0', which can be

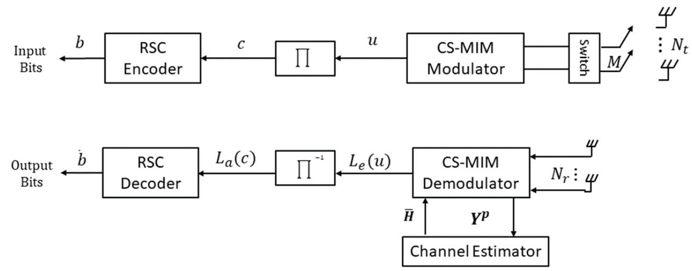


Fig. 11: The transceiver architecture of channel-coded CS-MIM.

written as $L(b) = log\frac{p(b=1)}{p(b=0)}$. The conditional probability $p(\boldsymbol{Y}|\mathcal{X}_{\beta,\gamma})$ of receiving the group signal $\boldsymbol{Y}$ is given by [45]

$$\begin{aligned} &p(\boldsymbol{Y}|\mathcal{X}_{\gamma,\beta}) \\ &= \frac{1}{(\pi \boldsymbol{N}_0)^{NT}} \exp(-\frac{||\boldsymbol{Y} - \boldsymbol{H}\bar{\boldsymbol{A}}x(\gamma)\bar{\boldsymbol{D}}(\beta)||^2}{N_0}), \end{aligned} \tag{9}$$

where $\mathcal{X}_{\gamma,\beta}$ represents the PSK/QAM symbol at the β-th CS-JMIM DM. Furthermore, N_0 is the noise power, where we have $\sigma_n^2 = N_0/2$ with $N_0/2$ representing the double-sided noise power spectral density.

Hence, we can formulate the LLR of bit u_i as

$$\begin{aligned} L_e(u_i) &= \ln\frac{p(\boldsymbol{y}|u_i=1)}{p(\boldsymbol{y}|u_i=0)} \\ &= \ln\frac{\sum_{\mathcal{X}_{\gamma,\beta \in \mathcal{X}_1^l}} p(\boldsymbol{Y}|\mathcal{X}_{\gamma,\beta})}{\sum_{\mathcal{X}_{\gamma,\beta \in \mathcal{X}_0^l}} p(\boldsymbol{Y}|\mathcal{X}_{\gamma,\beta})}, \end{aligned} \tag{10}$$

where $\mathcal{X}_1^l$ and $\mathcal{X}_0^l$ represent a subset of the legitimate equivalent signal $\mathcal{X}$ corresponding to bit u_l, when $u_l = 1$ and $u_l = 0$, respectively, yielding $X_1^l \equiv \{\mathcal{X}_{\gamma,\beta} \in \mathcal{X} : u_i = 1\}$ and $X_0^l \equiv \{\mathcal{X}_{\gamma,\beta} \in \mathcal{X} : u_i = 0\}$.

Upon using (9) and (10) we obtain the LLR $L(b_i)$ of the bit sequence conveyed by the received signal $\mathbf{Y}$. To simplify the calculation, the Approximate Log-MAP (Approx-Log-MAP) algorithm based on the Jacobian Maximum operation can be used, which is given by [46], [47]

$$\boldsymbol{L}_e(u_l) = \text{jac}_{\mathcal{X}_{\gamma,\beta} \in \mathcal{X}_1^l}(\lambda_{\gamma,\beta}) - \text{jac}_{\mathcal{X}_{\gamma,\beta} \in \mathcal{X}_0^l}(\lambda_{\gamma,\beta}), \tag{11}$$

where jac(.) denotes the Jacobian maximum operation and the intrinsic metric of $\lambda_{\gamma,\beta}$ is

$$\lambda_{\gamma,\beta} = -||\boldsymbol{Y} - \boldsymbol{H}\bar{\boldsymbol{A}}x(\gamma)\bar{\boldsymbol{D}}(\beta)||^2/N_0. \tag{12}$$

At the receiver, the soft demodulator evaluates the probability of each bit being logical '1' and '0'. Then it applies the approx-log-MAP algorithm for obtaining the extrinsic LLR of the coded bits, which has a complexity order $\mathcal{O}[2^{(c_g)}(N_{JMIM}(\mathcal{X})^K)]$, where c_g represents the number of coded bits after the RSC encoder and interleaver, and N_{JMIM} represents the number of possible realizations of JMIM.

2) DNN-based SD Detection: In this section, we propose a reduced-complexity SD detector using DNN, which considers a similar DNN architecture to that of [29]. Since the conventional SD detector obtains the LLRs of the received signal after the CS-MIM soft demodulator, we replace the detected

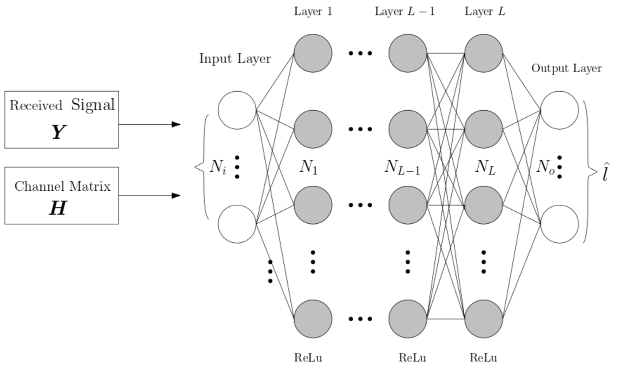


Fig. 12: Fully-connected DNN model for CS-JMIM SD detection.

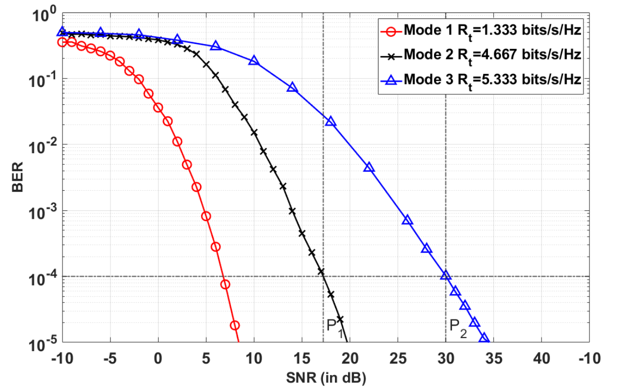


Fig. 13: BER *vs.* SNR performance of the CS-JMIM system for different mapping modes shown in Table III.

TABLE III: Configuration of the modes presented in Fig. 13

Mode	Mapping Type	Q	N_t	$N_v t$	N_f	N_v	K	R_t
1	Coded	4	4	8	4	8	2	1.333
2	General	4	4	8	4	8	3	4.666
3	Grouped	4	4	8	4	8	1	5.333

bits $\hat{\boldsymbol{u}}$ output by the DNN in Fig. 10 with the extrinsic LLR L_e at the output of the DNN, as shown in Fig. 12. Then, the output of the SD DNN model can be expressed as

$$\begin{aligned}\hat{\boldsymbol{L}}_e =&\boldsymbol{W}_{N_2}...f_{Relu}\{\boldsymbol{W}_2(f_{Relu_1}[\boldsymbol{W}_1(\boldsymbol{Y}_\tau)+\boldsymbol{b}_1]) \\ &+\boldsymbol{b}_2+...+\boldsymbol{b}_{N_2}),\end{aligned} \tag{13}$$

and the corresponding loss function is

$$\mathcal{L}(\boldsymbol{\theta},)=\frac{1}{BT}\sum_{i=1}^{B}\sum_{t=1}^{T}\|\hat{\boldsymbol{L}}_e(\tau)-\boldsymbol{L}_e(\tau)\|_2^2. \tag{14}$$

We can also define a stopping criterion, which can be either the number of iterations or meeting a maximum MSE threshold. Then, the parameter sets $\{\boldsymbol{W_n},\boldsymbol{\theta_n}\}$ can be updated in each training iteration based on the learning algorithm using gradient descent, which is formulated as

$$\{\boldsymbol{W_n},\boldsymbol{\theta_n}\}\leftarrow\{\boldsymbol{W_n},\boldsymbol{\theta_n}\}-\alpha\nabla\boldsymbol{L}(\{\boldsymbol{W_n},\boldsymbol{\theta_n}\}),$$

where $\alpha>0$ is the learning rate and $\nabla\boldsymbol{L}(\{\boldsymbol{W_n},\boldsymbol{\theta_n}\})$ represents the gradient of $\boldsymbol{L}(\{\boldsymbol{W_n},\boldsymbol{\theta_n}\})$.

In our proposed neural network aided detection, we use $\alpha=0.001$. Similar to the HD DNN detector described above, the model learns the parameters in the training phase and then outputs the LLR information.

The detection complexity of the learning algorithm is dominated by the calculation of the layer weights and biases, which may be considered to be of the order $\mathcal{O}(n_i n_h)+\mathcal{O}(n_h^2)+\mathcal{O}(n_h n_o)$ [29], with n representing the number of neurons in each layer.

IV. Adaptive design

Since the proposed CS-JMIM design provides flexibility in the design of the JMIM DM, we can design appropriate JMIM DMs for different channel conditions that can provide either an improved BER performance or an increased throughput. Furthermore, in our system, the transmitter can adapt both the JMIM DM $\boldsymbol{D}$ and the modulation order Q of PSK/QAM. Then, the system throughput may be adapted by appropriately adjusting the above parameters, while maintaining a target BER performance.

In the following two subsections, we highlight the classic threshold-based adaptive modulation, followed by its learning-aided counterpart. More specifically, both the KNN and DNN based adaptive model are applied for the proposed system.

1) Conventional Threshold-based Adaptive Design: In our adaptive scheme, we can adapt both the configuration of JMIM DM and of the PSK/QAM mode. We can define the different configurations as $Mode1$, $Mode2$, $Mode3$, $\ldots$, which can attain different BER performance and throughput. Based on the different modes, the parameters N_v, N_t, T and A of JMIM DM can be selected according to the SNR calculated at the receiver, where the SNR threshold values are selected for the different modes to satisfy a specific target BER [41], [42]. In the following, we present the scenario, where the different adaptive modes P refer to different configurations of the JMIM DM for characterising its design flexibility[4].

As an example, Fig.13 shows the BER performance of three different CS-JMIM mapping modes. The corresponding parameters and data rates provided by these modes are shown in Table III. For a target BER of 10^{-3}, as shown in Fig. 13 the SNR values of mode transition points P_1 and P_2 can be selected as the thresholds for operating the appropriate modes. Specifically, $Mode1$ is applied at low SNR values until the specific SNR reaches P_1. Then, the mode is changed to $Mode2$ to provide higher throughput, when the SNR range spans from P_1 to P_2. Finally, $Mode3$ is selected at SNRs higher than P_2, which has the highest throughput among the three modes.

For adaptive modulation, the receiver has to confidently infer the choice of the most appropriate transmission mode by comparing the instantaneous SNR of the received symbol against the Mode-switching threshold values. Then, the decision is fed back to the transmitter and applied for the next frame to be transmitted. Generally, with more available operation modes as well as faster and more accurate SNR feed-

[4]Note that the modulation scheme such as PSK/QAM can also be adapted, but in this design example, we aim to show the flexibility of the proposed CS-JMIM design.

back to the transmitter, we can obtain an increased throughput compared to non-adaptive designs. However, threshold-based adaptive modulation design ignores many of the hardware imperfections when deciding upon the threshold values, which results in sub-optimal performance of the adaptive system [41], [42]. Hence, in the next subsection, we propose the learning-based adaptive modulation scheme for our CS-JMIM system to further improve the adaptive system's performance.

2) Learning aided adaptive modulation: The adaptive modulation can be modelled as a classification problem, which can be solved using learning-based methods. The SNR of the received signal, which is evaluated at the receiver side, can be fed back to the transmitter and then given the SNR information, which also corresponds to the current channel state information, the transmitter can select a specific mode from a range of candidates to achieve the highest throughput, which still maintain the target BER. Therefore, for a given channel condition, adaptive modulation selects the most suitable mode to achieve the highest throughput, under the constraint of achieving the target BER. In this paper, both the KNN and DNN techniques are investigated in the context of adaptive modulation.

Before the training phase, the input data should be pre-processed to improve the learning efficiency. First, we randomly generate the training data of each mode under different instantaneous SNR values at the receiver. Then, the corresponding switching SNRs that can maintain a BER lower than the target BER are stored. Given these training data, we can use learning models to find the mode switching thresholds in the training phase. After training, the trained model becomes capable of predicting the next mode, given the knowledge of the SNR. In the following, we first employ KNN for our adaptive modulation scheme and then we propose a DNN-based adaptive model for further improving the performance.

a) KNN-based Adaptive Design: KNN is a popular classification techniques relying on low-complexity implementation and yet providing a good performance [48]. Yang *et al.* [40] developed KNN-assisted adaptive modulation schemes for SM, while Liu *et al.* [41] further developed DNN aided adaptive modulation to millimeter wave communication. To elaborate briefly on the KNN process, we define the training sets as

$$\mathcal{T}^{(i)} = [\xi_1^{(i)}, \cdots, \xi_n^{(i)}, \cdots, \xi_{N_p}^{(i)}]^T, \tag{15}$$

where ξ represents the SNR value of a symbol with a BER lower than the target BER value, with $i = 1, 2, \cdots, \mathcal{I}$ representing the adaptive mode index and N_p is the total number of instantaneous SNR values with BER under the target. Then, the total training set of each mode can be formulated as

$$\mathcal{T} = [\mathcal{T}^{(1)}, \cdots, \mathcal{T}^{(i)}, \cdots, \mathcal{T}^{(I)}]^T. \tag{16}$$

During runtime, for a given new data point, which corresponds to the instantaneous SNR ξ, the KNN model finds k nearest neighbours in the training set $\mathcal{T}$, using a distance metric $d(.)$, which can be expressed as

$$d(\xi_n^{(i)}), \xi_{new}) = ||\xi_n^{(i)}) - \xi_{new}||^2. \tag{17}$$

Then, the mode is decided by the majority mode of the k nearest neighbours to the input test point. With the possibility of several modes having the same number in the k nearest neighbours, the mode with the highest throughput will be selected.

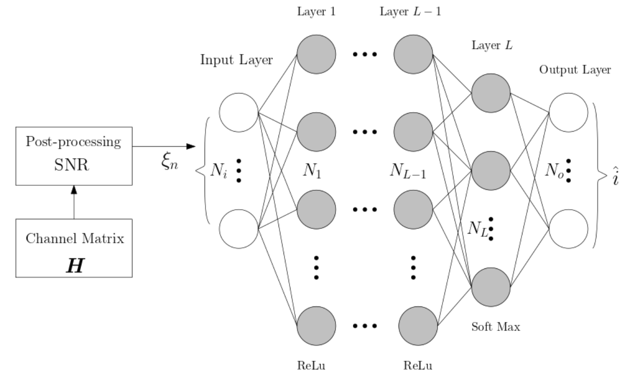


Fig. 14: Fully-connected DNN model for CS-MIM adaptive modulation selection.

The performance of KNN significantly depends on its parameters and on the value of k, where the best value of k can be selected empirically. In this adaptive system, the best value of k is determined by considering the trade-off between the BER and throughput. Furthermore, KNN results in a high computational complexity for the nearest neighbour search in addition to requiring a large memory for storing the training. Hence, in the following we present a DNN based design alternative.

b) DNN-aided Adaptive Design: In this section, we present the DNN-based adaptive modulation regime of Fig. 14. Similarly to KNN, we randomly generate the training data and then store the mode index and SNR value pairs, which have BERs lower than the target value. Then, the training set $\mathcal{T}$ constitutes the estimated SNR ξ of a symbol associated with a BER lower than the target BER. We use the DNN-based classification model, where the input corresponds to the instantaneous SNR and the output corresponds to the mode index of adaptive modulation.

The output mode index $\hat{i}$ of the DNN can be expressed as

$$\begin{aligned}\hat{i} =& f_{softmax}(\boldsymbol{W_n}...f_{Relu}\{\boldsymbol{W_2}(f_{Relu_1}[\boldsymbol{W_1}\xi + \boldsymbol{\theta_1}]) \\ &+ \boldsymbol{\theta_2}\} + ... + \boldsymbol{\theta}_n),\end{aligned} \tag{18}$$

where $\boldsymbol{W_n}$ and $\boldsymbol{\theta_n}$, $n = 1, \cdots, L$ represent the weights and biases, respectively. Relu is also employed for activating the DNN during the training phase, and the softmax function is used to obtain the mode index $\hat{i}$, which is

$$f_{softmax}(s) = \frac{e^s}{\sum_{c=1}^{C} e^{s_c}}. \tag{19}$$

The number of training samples required is selected based on experimentation by gradually increasing the training size until acceptable MSE values are achieved. In this case, the MSE loss function of the DNN used for the training is

$$\mathcal{L}(\xi, \hat{\xi}; \boldsymbol{W_n}, \boldsymbol{\theta_n}) = \frac{1}{B}\sum_{i=1}^{B} \|\xi - \hat{\xi}\|^2, \tag{20}$$

where B is the sample size of the current iteration.

TABLE IV: CS-MIM system simulation parameters.

<table>
<tr><td>Parameters</td><td>Scheme 1</td><td>Scheme 2</td><td>Scheme 3</td><td>Scheme 4</td><td>Scheme 5</td><td>Scheme 6</td><td>Scheme 7</td><td>Scheme 8</td><td>Scheme 9</td></tr>
<tr><td>Scheme type</td><td>CS-GFIM-SM</td><td>CS-JMIM</td><td>CS-GFIM-SM</td><td>CS-JMIM</td><td>CS-MIM</td><td colspan="4">CS-JMIM</td></tr>
<tr><td>Detection type</td><td colspan="7">HD</td><td colspan="2">SD</td></tr>
<tr><td>Multi-carrier System</td><td colspan="9">OFDM</td></tr>
<tr><td>Number of subcarriers, N_c</td><td colspan="9">128</td></tr>
<tr><td>Cyclic prefix</td><td colspan="9">16</td></tr>
<tr><td>Num of subcarrier group, G</td><td colspan="2">64</td><td colspan="2">32</td><td colspan="5">64</td></tr>
<tr><td>Num of active indices/gp, K</td><td colspan="2">1,2</td><td colspan="2">1,2,3</td><td colspan="2">1,2</td><td colspan="3">2</td></tr>
<tr><td>Receiver antennas, N_r</td><td colspan="2">2</td><td colspan="2">4</td><td>8</td><td colspan="4">2</td></tr>
<tr><td>RSC code, (n, k, K)</td><td colspan="7">-</td><td colspan="2">(2,1,3)</td></tr>
<tr><td colspan="10">Real Domain</td></tr>
<tr><td>Num of subcarrier/group, N_f</td><td colspan="2">2</td><td colspan="2">4</td><td colspan="5">2</td></tr>
<tr><td>Transmit antennas, N_t</td><td colspan="2">2</td><td colspan="2">4</td><td>8</td><td colspan="4">2</td></tr>
<tr><td>Activated antennas, N_{at}</td><td>1</td><td>-</td><td>1</td><td>-</td><td>2</td><td colspan="4"></td></tr>
<tr><td>Time Slots,T</td><td colspan="4"></td><td colspan="5">2</td></tr>
<tr><td colspan="10">Virtual Domain</td></tr>
<tr><td>Num of available subcarrier/group, N_v</td><td>8</td><td>4</td><td>16</td><td>8</td><td>8</td><td colspan="4">4</td></tr>
<tr><td>Transmit antennas, N_{vt}</td><td>-</td><td>4</td><td>-</td><td>8</td><td>-</td><td colspan="4">2</td></tr>
<tr><td>Time Slots,T_v</td><td colspan="5"></td><td colspan="4">4</td></tr>
<tr><td>STSK codeword, (m, n, t, q, l)</td><td colspan="4">-</td><td>(2,2,2,2,4)</td><td colspan="4">-</td></tr>
</table>

A stopping criterion can be defined either by the number of iterations or by the maximum tolerable MSE threshold. Then, the parameter sets $\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\}$ can be updated in each training iteration based on our learning algorithm using gradient descent, which is formulated as

$$\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\} \leftarrow \{\boldsymbol{W_n}, \boldsymbol{\theta_n}\} - \alpha \nabla \boldsymbol{L}(\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\}),$$

where $\alpha > 0$ is the learning rate and $\nabla \boldsymbol{L}(\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\})$ represents the gradient of $\boldsymbol{L}(\{\boldsymbol{W_n}, \boldsymbol{\theta_n}\})$. In our proposed DNN-aided detection, we use $\alpha = 0.001$.

V. Simulation Results and Analysis

In this section, we characterize the performance of the proposed CS-JMIM system, where conventional detection will be used for benchmarking the proposed learning aided detection methods. Furthermore, we consider the system employing SF CS-JMIM and TSF CS-JMIM. The BER performance is evaluated by Monte-Carlo simulations, where we use the simulation parameters summarized in Table IVThe parameters used by the learning models are presented in Table VI. In our simulations, we assume that the receiver has perfect channel knowledge, while in practice this is estimated using channel estimation techniques.

In the following, we present the different schemes considered in our simulations for comparison purposes. Firstly, we compared CS-aided separate multi-dimensional IM with CS-JMIM. More specifically, for our SF domain system, we compared CS-aided Generalized Subcarrier Index Modulation with SM (CS-GFIM-SM). These are termed as **Scheme 1, 3**, with CS-JMIM as **Scheme 2, 4**. Then, for the TSF domain, the CS-JMIM of **Scheme 5** is compared to **Scheme 6**, which represents the CS-MIM [22] [36]. Secondly, we compared the performance of different parameters in the context of **Schemes 2, 4, 6**. Thirdly, we characterized the performance of DNN-aided CS-JMIM both in HD and SD in **Schemes 6-9**. We also quantified the complexity and compared it to conventional ML detection. Finally, we also exploited the adaptation of CS-JMIM between different JMIM methods in **Scheme 10**. To elaborate:

1) **Scheme 1**: applies ML HD detection for the CS-GFIM-SM, which activated one of 2 TAs, 2 RAs, and 2 subcarriers per group, while considering 8 subcarriers per group in the VD and $K = 1, 2$ activated subcarriers.
2) **Scheme 2**: applies maximum likelihood hard decision detection for the CS-JMIM system in the SF domain along with 2 TAs, 2 RAs, and 2 subcarriers per group in the RD, while considering 4 antennas and 4 subcarriers per group in the VD. In this scheme, we consider the following mappings:
   a) General JMIM with $K = 1, 2$.
   b) Grouped JMIM with $gs = 4$ subgroups, and each subgroup applies general JMIM in conjunction with $K = 1$ (In this case, we can consider that both the FD and SpD is split into two sub groups, which have $gsx = gsy = 2$.).
   c) Coded JMIM with $n_q = 2$.
3) **Scheme 3**: applies ML HD detection for the CS-GFIM-SM, which activated one antenna out of 4 TAs, 4 RAs, and 4 subcarriers per group, while considering 16 subcarriers per group in the VD and $K = 1, 2, 3$ activated subcarriers.
4) **Scheme 4**: applies maximum likelihood hard decision detection for the CS-JMIM system in the SF domain along with 4 TAs, 4 RAs, and 4 subcarriers per group in this RD, with 8 antennas and 8 subcarriers per group in the VD. In this scheme, we consider the following mappings:
   a) General JMIM with $K = 1, 2, 3$.
   b) Grouped JMIM with $gs = 4, gsx = gsy = 2$ subgroups, with each subgroup applying the general JMIM along with $K = 1$.
   c) Coded JMIM with $n_q = 4$
5) **Scheme 5**: applies ML HD detection for the CS-MIM system in the TSF domain with 8 TAs, 8 RAs, 2 subcarriers per group and 2 TSs, while having 8 subcarriers per group in the VD. For the Space-Time-Shift-Keying (STSK) codeword $STSK(M, N, T, Q, L)$ used in CS-MIM [22], STSK(2,2,2,2,4) is applied. In this case, we have 2 activated antennas out of 8 and $K = 1, 2$ activated subcarrier out of 8 subcarrier in the VD.
6) **Scheme 6**: applies maximum likelihood hard decision

TABLE V: Simulation results and complexity analysis of each Scheme.

Scheme index			SNR at BER of 10^{-5}	Throughput(bits/s/Hz)	Complexity
HD Detection					
Scheme 1		K=1	20.8	2.667	1.4×10^5
		K=2	26.5	4	5.6×10^5
Scheme 2	a)	K=1	30.3	2.667	9.5×10^5
		K=2	30.2	4.444	3.8×10^6
	b)		34.9	7.111	5.1×10^8
	c)		22.4	1.778	1.8×10^5
Scheme 3		K=1	16.6	1.778	8.6×10^6
		K=2	23.4	2.667	3.4×10^7
		K=3	28.1	3.778	1.4×10^8
Scheme 4	a)	K=1	8.2	1.778	5.3×10^7
		K=2	13.4	3.111	2.1×10^8
		K=3	19.4	4.667	8.3×10^8
	b)		34.6	5.333	2.2×10^{10}
	c)		8.3	1.333	7.2×10^6
Scheme 5		K=1	9.6	3.556	1.2×10^7
		K=2	13.3	5.333	4.9×10^7
Scheme 6	a)	K=1	-0.4	3.556	4.1×10^7
		K=2	4.9	6.222	6.5×10^9
	b)		15.7	17.778	5.4×10^{11}
	c)		1.5	1.778	1.1×10^7
Scheme 7	a)		5.6	3.556	2.2×10^5
	b)		18.7	17.778	1.7×10^6
	c)		1.8	1.778	6.6×10^4
SD Detection					
Scheme 8	a)		1.1	1.778	2.2×10^{13}
	b)		6.2	8.889	3.2×10^{14}
	c)		0.1	0.889	3.4×10^{12}
Scheme 9	a)		4.3	1.778	1.3×10^6
	b)		8.9	8.889	8.3×10^6
	c)		4.1	0.889	1.2×10^5
Adaptive Modulation					
Scheme 10	a)		-	-	-
	b)		-	-	5.2×10^6
	c)		-	-	1.22×10^5

TABLE VI: Training configuration for learning-aided detection method of Scheme 7,9

Setting	Hard-decision	Soft-decision
Maximum training epoch	400	1000
Initial learning rate	0.001	
Target SNR for training	0dB-20dB	-10dB to 5dB
Mini batch size	1000	200 to 500
Optimizer	Adam	
Training data size	50000	
Validation data ratio	0.1	

TABLE VII: Training configuration for adaptive modulation of Scheme 10

Setting	value
Number of Channel realizations for training	100000
Number of Channel realizations for testing	20000
Target SNR for training	0dB-30dB
Number of neighbors in KNN searchin k	15
Number of FC layers in DNN	3
Number of neurons in each FC layer	(128,256,128)
Number of output layer size	3
Activation function for output layer	Soft Max

detection for the CS-JMIM system in the TSF domain with 2 TAs, 2 RAs, 2 subcarriers per group and 2 TSs in the RD, while using 4 antennas, 4 subcarriers per group and 4 TSs in the VD. In this scheme, we consider the following mappings:

a) General JMIM with $K = 1, 2$.
b) Grouped JMIM with $gs = 8, gsx = gsy = gsz = 2$ subgroups, where each subgroup applies general JMIM along with $K = 1$.(In this case, we further split the TD into two parts, which have $gsz = 2$.).
c) Coded JMIM $n_q = 2$.

7) **Scheme 7**: applies DNN based HD detection for the CS-JMIM system. Here, we consider 2 TAs, 2 RAs, 2 subcarriers per group, and 2 TSs in the RD, while using 4 antennas, 4 subcarriers per group and 4 TSs in the VD. In this scheme we consider the following mappings:
   a) General JMIM with $K = 2$.
   b) Grouped JMIM with $gs = 8, gsx = gsy = gsz = 2$ subgroups, where each subgroup applies general JMIM with $K = 1$.
   c) Coded JMIM with $n_q = 2$.
8) **Scheme 8**: applies conventional SD detection for the CS-JMIM system in the TSF domain, while using RSC channel coding RSC(2,1,3). Here, we consider 2 TAs, 2 RAs, 2 subcarriers per group, and 2 TSs in the RD, while using 4 antennas, 4 subcarriers per group and 4 TSs in the VD. In this scheme, we consider the following mappings:
   a) General JMIM with $K = 2$.
   b) Grouped JMIM with $gs = 8, gsx = gsy = gsz =$

2 subgroups, each subgroup applied general JMIM with $K = 1$.

c) Coded JMIM with $n_q = 2$.

9) **Scheme 9**: applies DNN-based SD detection for the CS-JMIM system in the TSF domain, while using RSC channel coding RSC(2,1,3). Here, we consider 2 TAs, 2 RAs, 2 subcarriers per group, and 2 TSs in the RD, while using 4 antennas, 4 subcarriers per group and 4 TSs in the VD. In this scheme, we consider the following mappings:

   a) General JMIM with $K = 2$.

   b) Grouped JMIM with $gs = 8, gsx = gsy = gsz = 2$ subgroups, each subgroup applied general JMIM with $K = 1$.

   c) Coded JMIM with $n_q = 2$.

10) **Scheme 10**: Adaptive HD-CS-JMIM system based on the TSF domain with 2 TAs, 2 RAs, 2 subcarriers per group, 2 TSs in RD and 4 antennas, 4 subcarriers per group and 4 TSs in VD. The details of the DNN based adaptive system design are shown in Table VII. In this system, we consider the following adaptation schemes:

   a) Conventional adaptation.

   b) KNN-based adaptation.

   c) DNN-based adaptation.

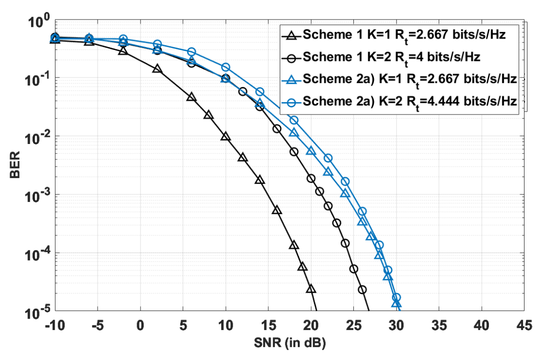


Fig. 15: BER performance comparison of Scheme 1 and Scheme 2a). Our simulation parameters are shown in Table IV.

As shown in Fig. 15, we compared the CS-aided separate MIM - namely the CS-GFIM-IM in this case - to CS-JMIM, which applied the general JMIM method of Section II-A1a). In this case, based on the transmission rate calculation formula $\frac{bG}{N_c+L_{CP}}$, we have the transmission rate of the CS-GFIM-IM associated with $K = 1$ in **Scheme 1** as $R_t^{k=1} = 2.667$ bits/s/Hz. This is the same as the CS-JMIM associated with $K = 1$ in **Scheme 2a)** under identical hardware configuration. However, the performance of **Scheme 2a)** is almost 10 dB worse than that of **Scheme 1** at a BER of 10^{-5}. Hence CS-JMIM is unattractive in this situation. For more activated index entities of both CS-JMIM and CS-GFIM-IM, the throughput of **Scheme 1** is increased to $R_t^{1,k=2} = 4$ bits/s/Hz and **Scheme 2a** has $R_t^{2,k=2} = 4.444$ bits/s/Hz. In this case, **Scheme 2a** of $K = 2$ has a 3.6 dB better performance than **Scheme 1** of $K = 2$ at a BER of 10^{-5}.

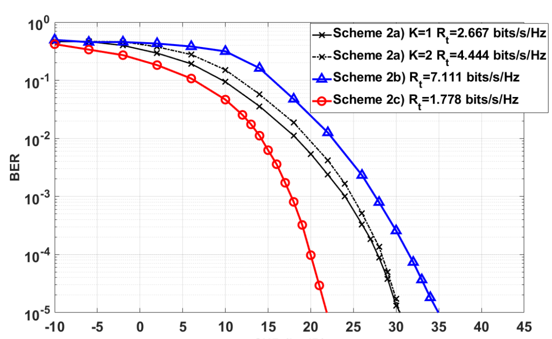


Fig. 16: BER performance comparison of CS-JMIM Scheme 2. Our simulation parameters are shown in Table IV.

Fig. 16 shows the performance of the proposed CS-JMIM **Scheme 2** for different JMIM methods. Observe that for a small index space of $N_t = N_f = 2$, the detector cannot beneficially exploit the sparsity. The transmission rate of **Scheme 2** is either $R_t^{k=1} = 2.667$ bits/s/Hz, or $R_t^{k=2} = 4.444$ bits/s/Hz and we have $R_t^b = 7.111$ bits/s/Hz, $R_t^c = 1.778$ bits/s/Hz. As shown in Fig. 16, **Scheme 2a)** associated with $K = 1, 2$ has a similar BER performance, while **Scheme 2a)** of $K = 2$ has a higher throughput. Additionally, **Scheme 2b)** has almost 4 times the transmission rate compared to **Scheme 2c)**, but the latter has an increased diversity gain. Hence the BER performance of **Scheme 2c)** is 12dB better than that of **Scheme 2c)**.

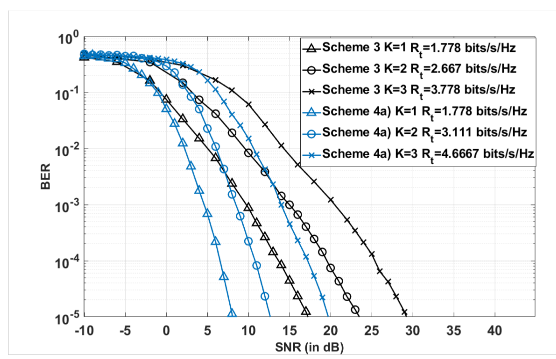


Fig. 17: BER performance comparison of Scheme 3 and Scheme 4a). Our simulation parameters are shown in Table IV.

To further exploit the sparsity of CS-JMIM, we also consider larger SF dimensions applied to the JMIM method, as shown in Fig. 17. We assume that both schemes have the same number of TAs and subcarriers per group along with an adjustable number of VD subcarriers. For $N_t = 4, N_f = 4$, the CS-JMIM of **Scheme 4a)** achieves better performance than the separate MIM in **Scheme 3** with the same K value. Specifically, both schemes have $R_t^{k=1} = 1.777$ bits/s/Hz and **Scheme 3** associated with $K = 1$ obtains 5 dB SNR gain over **Scheme 4a)** with $K = 1$ at BER of 10^{-5}. When relying on a higher K, CS-JMIM is capable of providing higher throughput as well as improved detection performance. With $K = 2, 3$, the throughput of **Scheme 3** is $R_t^{k=2} = 2.667$ bits/s/Hz and $R_t^{k=3} = 3.333$ bits/s/Hz, respectively, while **Scheme 4a)** could achieve $R_t^{k=2} = 3.111$ bits/s/Hz and $R_t^{k=3} = 4.667$ bits/s/Hz.

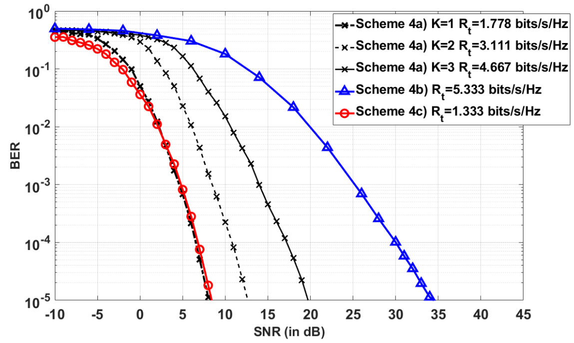


Fig. 18: BER performance comparison of CS-JMIM Scheme 4. Our simulation parameters are shown in Table IV.

Fig. 18 shows the BER performance of **Scheme 4**. A higher VD index mapping DM size allows for more flexible K value selection in **Scheme 4a)**. Observe that **Scheme 4a)** with $K = 1$ achieves a similar performance to **Scheme 4c)**, where **Scheme 4a)** with $K = 1$ has $R_t = 1.778$ bits/s/Hz and **Scheme 4c)** has $R_t = 1.333$ bits/s/Hz.

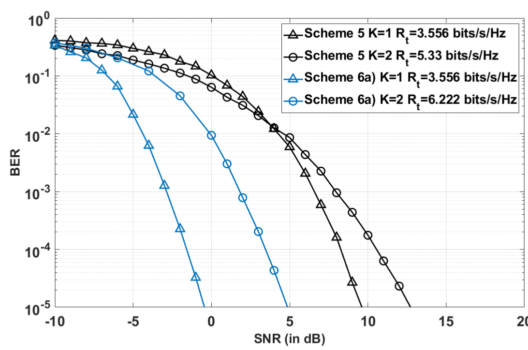


Fig. 19: BER performance comparison of Scheme 5 and Scheme 6a). Our simulation parameters are shown in Table IV.

For the TSF domain system of Fig.6(a), we consider a separate model termed as CS-MIM [22]. This model applied SIM and STSK in the FD and CS is applied for the FD. Then the symbol after IFFT is modulated using SM and transmitted by the activated antennas. The CS-MIM scheme is simulated using the parameters of Table IV for **Scheme 5**. In this case, to achieve the same throughput as **Scheme 5** and **Scheme 6a)** at $K = 1$, for **Scheme 5**, we deliver the signals over 8 TAs with the aid of 2 RF chains. Then both **Scheme 5** and **Scheme 6a)** can have a throughput of $R_t^{K=1} = 3.556$ bits/s/Hz with $K = 1$. Then, we can observe in Fig.19 that **Scheme 6a)** achieves a BER of 10^{-5} at -0.1 dB while **Scheme 5** requires about 9.8 dB at the same BER. For $K = 2$, **Scheme 5** requires 13.5 dB SNR at 10^{-5} BER for $R_t^{K=2} = 5.333$ bits/s/Hz and **Scheme 6a)** requires 7.5 dB lower SNR than **Scheme 5** for $R_t^{K=2} = 6.222$ bits/s/Hz.

In Fig. 20, the TSF domains are considered for the CS-JMIM using **Scheme 6**. As shown in Fig. 20, **Scheme 6a)** with $K = 1$ attains the best performance among all types in **Scheme 6**. Quantitatively, at a BER of 10^{-5}, it requires an SNR of -0.3 dB and has a throughput of $R_t = 3.555 bits/s/Hz$. **Scheme 6c)** achieves a BER of 10^{-5} at an SNR of 1.1 dB. When higher dimensions are introduced, both the general JMIM and grouped JMIM can provide a high throughput as well as a good BER performance, albeit at the cost of a huge detection complexity. In Fig. 20, **Scheme 6b)** represents the grouped JMIM associated with 8 sub-groups. When $K = 1$ and the general JMIM DM is applied, we have $R_t = 17.778 bits/s/Hz$. This scheme attains a BER of 10^{-5} at an SNR of 15.1 dB. **Scheme 6a)** with $K = 3$ has $R_t = 9.333 bits/s/Hz$ and achieves a BER of 10^{-5} at an SNR of 11 dB. Hence, for higher dimensions, the grouped JMIM outperforms the other two JMIM methods. However, the complexity of grouped JMIM is exponentially increasing. Specifically, the detection complexity order of the grouped JMIM can be expressed as $\mathcal{O}[(N_{JMIM}(\mathcal{X}^K)^{N_{sub}}]$ for the TSF domain CS-JMIM system. This can be simplified to $\mathcal{O}[((N_v N_{vt} T_v/(g_s))(M^K)^{N_{sub}}]$, where N_{sub} represents the number of sub-groups. On the other hand, the detection complexity order of the general JMIM is $\mathcal{O}[(N_v N_{vt} K T_v M^K)]$. Furthermore, the coded JMIM complexity order can be $\mathcal{O}[(N_q - n_q)n_q M]$. Then we can formulate the computational complexity order of ML for **Scheme 7a)** as $\mathcal{O}_{ML}[N_r N_f N_t T(N_r N_f N_t T N_{vt}^2 N_v^2 T_v^2 + N_{vt} N_v T_v M^2 N_f N_t T + MK)(N_{JMIM}(\mathcal{X}^K))]$. For **Scheme 7b)**, the sub-groups must be considered in each rounds, which have a complexity of $\mathcal{O}_{ML}[(N_{sub} N_r N_f N_t T/gs)(N_r N_f N_t T N_{vt}^2 N_v^2 T_v^2/(gs^2) + N_{vt} N_v T_v M N_f N_t T M/gs + MK)(N_{JMIM}(\mathcal{X}^K))^{N_{sub}}]$. For **Scheme 7c)**, we have a reduced complexity order of $\mathcal{O}_{ML}[N_r N_f N_t T N_{vt} N_v T_v M N_f N_t T M(N_q - n_q)n_q M]$ due to having multiple bit copies. Then we can calculate the computational complexity based on Table IV, as shown in Table V.

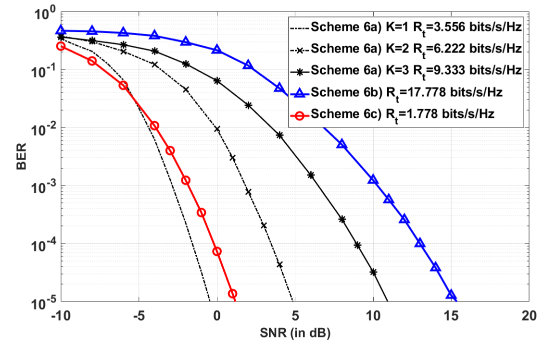


Fig. 20: BER performance comparison of CS-JMIM Scheme 6. Our simulation parameters are shown in Table IV.

Upon increasing the throughput excessive detection complexity is imposed by conventional ML detection. To reduce the detection complexity, we have to accept a performance *vs.* complexity trade-off. In this context, we compare our DNN-based detector of the TSF based CS-JMIM system to conventional maximum likelihood detection by comparing **Scheme 6** and **Scheme 7** in Fig 21. Observe that the DNN-assisted HD detector achieves a similar performance to the ML detector. Furthermore, the complexity of the NN is determined by that

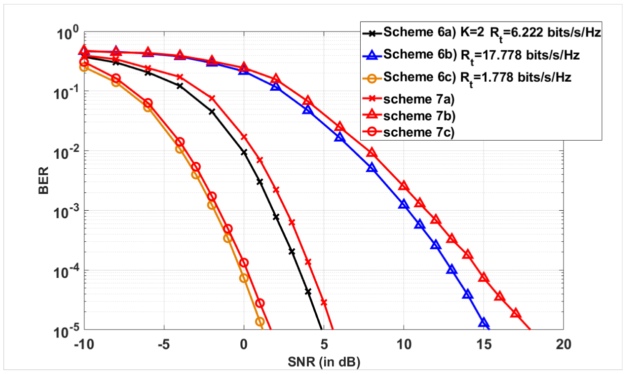


Fig. 21: BER performance comparison of CS-JMIM Scheme 6 and Scheme 7. Our simulation parameters are shown in Table IV.

of the forward and backward propagation, where we have the general DNN complexity order of $\mathcal{O}[n_i n_l n_{l+1} n_{h_L} n_o]$. Here n_i and n_o denote the number of neurons in the input and output layers, $n_l (l = 1, 2, \cdots)$ is the number of neurons in the hidden layer between the input and output. Then we can analyse each DNN model in **Scheme 8**. For a classification neural network, we have the LSTM layer as the activation layer of the input layer, which has the complexity of $\mathcal{O}_{LSTM}[4n_l(n_d+2+n_l)]$, where n_d is the number of neurons in the input layer and the popular sigmoid function is used as the activation layer of the output layer. The associated complexity is $\mathcal{O}[2n_L n_{L-1} - n_{L-1} + 2n_{L-1}]$. The complexity of the FC layer with the ReLu function is given by $\mathcal{O}[2n_l n_{l-1} - n_{l-1} +$. Then we have the computational complexity order of $\mathcal{O}[4n_l(n_1+2+n_l) + \sum^{L} -1_l(2n_{l+1}n_l - n_l) + 2n_{L-1}]$. Now we can also summarize the computational complexity of the DNN methods in Table V.

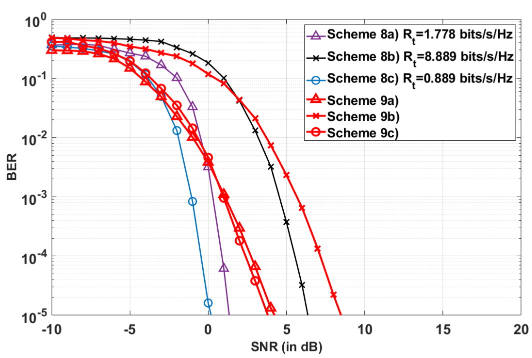


Fig. 22: BER performance comparison of CS-JMIM Scheme 8 and Scheme 9. Our simulation parameters are shown in Table IV.

Furthermore, we extend the DNN-assisted detector to the SD of the TSF domain CS-JMIM system in **Scheme 8** and **Scheme 9**, while using the half-rate RSC encoder RSC(2,1,3), having a memory of 3. As shown in Fig. 22, with the aid of channel coding, the performance of CS-JMIM can be further increased, as seen for **Scheme 8**. By comparing **Scheme 8** of Fig. 22 and **Scheme 6** of Fig. 21, the detection performance is 1dB better for **Scheme 8c)** than for **Scheme 6c)** at the BER of 10^{-5}. Furthermore, **Scheme 8a)** requires an SNR of 6.2 dBs at BER=10^{-5}, while **Scheme 6a)** necessitates SNR=1.6 dB. **Scheme 8b)** has the best performance, outperforming **Scheme 6b)** by about 8 dB at a BER of 10^{-5}. Fig.22 also shows the performance of DNN based detection for TSF CS-JMIM, where **Scheme 9a)** and **Scheme 9c)** exhibit similar performance. Quantitatively, they require about 4 and 3.2 dB at a BER of 10^{-5}. **Scheme 9b)** requires 3 dB higher SNR than the conventional SD detector, but it is still about 6 dB better than **Scheme 7b)**. The proposed learning method has a complexity order of $O[\mathcal{O}(n_i n_l) + \mathcal{O}(n_l^2) + \mathcal{O}(n_l n_o)]$ compared to $O[2^{c_g}(T_v N_t N_{vt}(Q\mathcal{X})^K]$ for the conventional scheme, where c_g denotes the RSC-coded number of bits in a transmitted symbol.

Finally, we present the performance of **Scheme 10** in

TABLE VIII: Configuration of mode used in conventional adaption with TSF domain CS-JMIM

No	Type	Scheme	R_t
1	Coded	Scheme 7a)	1.778
2	General	Scheme 7b)	6.222
3	Grouped	Scheme 7c)	17.778

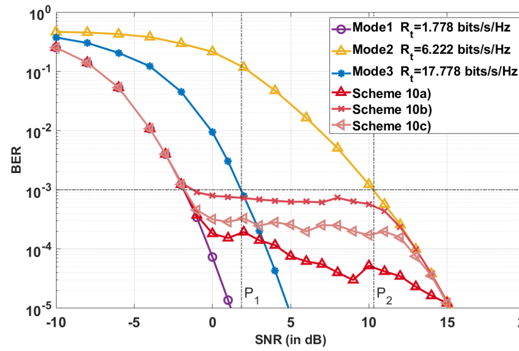


Fig. 23: Adaptive modulation performance comparison of CS-JMIM Scheme 7. Our simulation parameters are shown in Table IV.

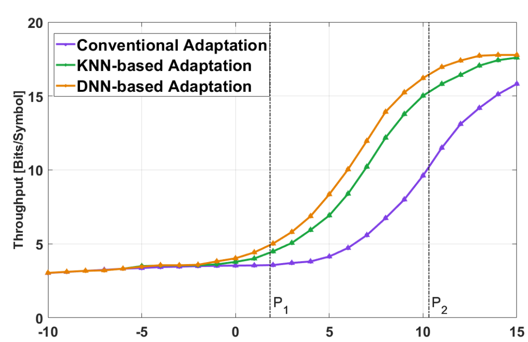


Fig. 24: Adaptive modulation performance comparison of CS-JMIM Scheme 7. Our simulation parameters are shown in Table IV.

Fig. 23. For the sake of a fair comparison, we use the data sets of the same size for both training and testing the KNN and DNN-based systems. Table VIII presents the configurations of the three modes of operation used in the adaptive system simulated. The switching thresholds for the conventional adaptive modulation are set as $P_1 = 1.85$ dB and $P_2 = 10.3$ dB, as shown in Fig. 23. Specifically, the conventional adaptive modulation scheme characterized in Fig. 23 uses $Mode_1$ when $SNR < P_1$, and $Mode_2$ for $P_2 > SNR$. After the instantaneous SNR becomes higher than P_2, $Mode_3$ is selected. Again, our KNN-based and DNN-based mode-selection algorithms are used in Fig. 23. Observe that the DNN based adaptive system attains a BER closer to the target of 10^{-3} than the KNN based adaptive system. Then we can further analyse the throughput of each mode selection scheme in Fig. 24. Observe that the DNN-based adaptive modulation scheme achieves a higher throughput than the KNN-based one, because more accurate decisions can be made by the DNN classifier than by the KNN classifier. Clearly, the learning assisted adaptive schemes are capable of selecting the best possible mode, while the conventional adaptive modulation uses the predefined average SNR-based thresholds for mode selection.

VI. Conclusions

A CS-JMIM system was proposed and DL-aided detection using both HD and SD was conceived for reducing the detection complexity. We demonstrated that the proposed JMIM system is capable of outperforming its individual domain based counterpart, striking more flexible trade-offs between the BER performance and throughput. The learning method constructed is capable of approaching the performance of the maximum likelihood detector at a significantly reduced complexity. Furthermore, we showed that adaptive modulation can be applied for the selection of the JMIM DM design. We demonstrated that the CS-JMIM can flexibly adjust the transmission mode for accommodating time-variant channel conditions. We presented both KNN and DNN based adaptive schemes. Our simulation results showed that both the KNN and DNN-based approaches outperform the conventional threshold-based adaptive modulation. We also demonstrated that the DNN based adaptive design has a lower computational complexity and higher throughput than the KNN based approach.

References

[1] Y. Chau and S.-H. Yu, "Space modulation on wireless fading channels," in *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No.01CH37211)*, vol. 3, 2001, pp. 1668–1671 vol.3.

[2] N. Ishikawa, S. Sugiura, and L. Hanzo, "50 years of permutation, spatial and index modulation: from classic RF to visible Llight communications and data storage," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 1905–1938, 2018.

[3] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[4] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2228–2241, 2008.

[5] E. Basar, "Index modulation techniques for 5G wireless networks," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 168–175, 7 2016.

[6] I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Hierarchical Multi-Functional Layered Spatial Modulation," *IEEE Access*, vol. 6, pp. 9492–9533, 2018.

[7] R. Abu-alhiga and H. Haas, "Subcarrier-index modulation OFDM," *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 177–181, 2009.

[8] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, pp. 728–732, 2011.

[9] E. Başar, Aygölü, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5536–5549, 2013.

[10] H. Zhang, L.-L. Yang, and L. Hanzo, "Compressed sensing improves the performance of subcarrier index-modulation-assisted OFDM," *IEEE Access*, vol. 4, pp. 7859–7873, 2016.

[11] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, apr 2006.

[12] L. Xiao, P. Yang, Y. Xiao, S. Fan, M. Di Renzo, W. Xiang, and S. Li, "Efficient compressive sensing detectors for generalized spatial modulation systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1284–1298, feb 2017.

[13] T. Datta, H. S. Eshwaraiah, A. Chockalingam, and S. Member, "Generalized space-and-frequency index modulation," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 4911–4924, 2016.

[14] B. Chakrapani, T. L. Narasimhan, and A. Chockalingam, "Generalized space-frequency index modulation: Low-complexity encoding and detection," in *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1–6.

[15] R. Mesleh, S. S. Ikki, and H. M. Aggoune, "Quadrature spatial modulation," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2738–2742, 2015.

[16] P. Patcharamaneepakorn, C.-X. Wang, Y. Fu, E.-H. M. Aggoune, M. M. Alwakeel, X. Tao, and X. Ge, "Quadrature space-frequency index modulation for energy-efficient 5G wireless communication systems," *IEEE Transactions on Communications*, vol. 66, no. 7, pp. 3050–3064, 2018.

[17] Z. Li and J. Zheng, "Space-frequency shift keying in rapidly time-varying MIMO OFDM channels," in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, 2018, pp. 1–5.

[18] C. Xu, P. Zhang, R. Rajashekar, N. Ishikawa, S. Sugiura, Z. Wang, and L. Hanzo, "“Near-perfect” ginite-cardinality generalized space-time shift keying," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 9, pp. 2146–2164, 2019.

[19] S. Sugiura, S. Chen, and L. Hanzo, "Generalized space-time shift keying designed for flexible diversity-, multiplexing- and complexity-tradeoffs," *IEEE Transactions on Wireless Communications*, vol. 10, no. 4, pp. 1144–1153, 2011.

[20] B. Shamasundar, S. Bhat, S. Jacob, and A. Chockalingam, "Multidimensional index modulation in wireless communications," *IEEE Access*, vol. 6, pp. 589–604, 2018.

[21] S. Lu, I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Compressed-sensing-aided space-time frequency index modulation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6259–6271, 2018.

[22] S. Lu, I. A. Hemadeh, M. El-hajjar, and L. Hanzo, "Compressed sensing-aided multi-dimensional index modulation," *IEEE Transactions on Communications*, vol. 67, no. 6, pp. 4074–4087, 2019.

[23] S. Sugiura, S. Chen, and L. Hanzo, "Space-time shift keying: A unified mimo architecture," in *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, 2010, pp. 1–5.

[24] J. Zheng and Q. Liu, "Low-complexity soft-decision detection of coded OFDM with index modulation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7759–7763, aug 2018.

[25] J. Liu and H. Lu, "Imnet: A learning based detector for index modulation aided mimo-ofdm systems," *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2020.

[26] S. Sugiura and L. Hanzo, "Effects of channel estimation on spatial modulation," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 805–808, 2012.

[27] Y. Acar, M. M. Leblebici, H. Doğan, and E. Panayirci, "Data detection based iterative channel estimation for coded SM-OFDM systems," in *2016 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2016, pp. 1–4.

[28] S. Katla, L. Xiang, Y. Zhang, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Deep learning assisted detection for index modulation aided mmwave systems," *IEEE Access*, vol. 8, pp. 202 738–202 754, 2020.

[29] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, P. Pietraski, and L. Hanzo, "Soft-decoding for multi-set space-time shift-keying mmWave systems: A deep learning approach," *IEEE Access*, vol. 8, pp. 49 584–49 595, 2020.

[30] 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. 191 910–191 919, 2020.

[31] X. Cheng, M. Zhang, M. Wen, and L. Yang, "Index modulation for 5G: Striving to do more with less," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 126–132, 2018.

[32] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223–229, 2002.

[33] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.

[34] 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. 178 246–178 299, 2019.

[35] L. Hanzo, J. P. Woodard, and P. Robertson, "Turbo decoding and detection for wireless applications," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1178–1200, 2007.

[36] X. Feng, M. EL-Hajjar, C. Xu, and L. Hanzo, "Deep learning-based soft iterative-detection of channel-coded compressed sensing-aided multi-dimensional index modulation," *IEEE Transactions on Vehicular Technology*, pp. 1–16, 2023.

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

[38] T. Wang, F. Yang, J. Song, and Z. Han, "Deep convolutional neural network-based detector for index modulation," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1705–1709, oct 2020.

[39] M. Turhan, E. Öztürk, and H. A. Çırpan, "Deep convolutional learning-aided detector for generalized frequency division multiplexing with index modulation," in *2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2019, pp. 1–6.

[40] P. Yang, Y. Xiao, M. Xiao, Y. L. Guan, S. Li, and W. Xiang, "Adaptive spatial modulation mimo based on machine learning," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 9, pp. 2117–2131, 2019.

[41] H. Liu, S. Lu, M. El-Hajjar, and L.-L. Yang, "Machine learning assisted adaptive index modulation for mmWave communications," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1425–1441, 2020.

[42] H. Liu, Y. Zhang, X. Zhang, M. El-Hajjar, and L.-L. Yang, "Deep learning assisted adaptive index modulation for mmWave communications With channel estimation," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 9, pp. 9186–9201, 2022.

[43] W. Koch and A. Baier, "Optimum and sub-optimum detection of coded data disturbed by time-varying intersymbol interference (applicable to digital mobile radio receivers)," in *[Proceedings] GLOBECOM '90: IEEE Global Telecommunications Conference and Exhibition*, 1990, pp. 1679–1684 vol.3.

[44] H. V. Nguyen, C. Xu, S. X. Ng, and L. Hanzo, "Near-capacity wireless system design principles," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1806–1833, 2015.

[45] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.

[46] M. El-Hajjar and L. Hanzo, "EXIT charts for system design and analysis," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 1, pp. 127–153, mar 2014.

[47] S. Benedetto and G. Montorsi, "Serial concatenation of block and convolutional codes," *Electronics Letters*, vol. 32, no. 10, pp. 887–888, 1996.

[48] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

Xinyu Feng received a BSc(Eng) degree with First Class Honours from University of Nottingham Ningbo, china and a BSc(Eng) in 2014. He obtained a MSc degree with distinction in wireless communication from the University of Southampton, UK in 2019. He is currently a PhD student at Next Generation Wireless Research Group under the supervision of Prof. L. Hanzo and Dr. M. El-Hajjar, University of Southampton, UK. His research interests include index modulation, Channel estimation and Detection and machine learning.

Mohammed El-Hajjar Mohammed El-Hajjar is an Associate Professor in the School of Electronics and Computer Science in the University of Southampton. He received his PhD in Wireless Communications from the University of Southampton, UK in 2008. Following the PhD, he joined Imagination Technologies as a design engineer, where he worked on designing and developing Imagination's multi-standard communications platform, which resulted in three patents. He is the recipient of several academic awards and has published a Wiley-IEEE book and in excess of 140 journal and conference papers. His research is funded by the UK government and many industrial collaborators. Mohammed's research interests include the development of intelligent communications systems, energy-efficient transceiver design, MIMO, millimeter wave communications and Radio over fiber network design.

Chao Xu (S'09-M'14-SM'19) received a B.Eng. degree from Beijing University of Posts and Telecommunications, China, and a BSc(Eng) with First Class Honours from Queen Mary, University of London, UK, through a Sino-UK joint degree program in 2008, both in Telecommunications. He obtained a MSc degree with distinction in Radio Frequency Communication Systems and a Ph.D. degree in Wireless Communications from the University of Southampton, UK in 2009 and 2015, respectively. He is currently a senior research fellow working at Next Generation Wireless Research Group, University of Southampton, UK. His research interests include index modulation, reconfigurable intelligent surfaces, noncoherent detection and turbo detection. He was awarded the Best M.Sc. Student in Broadband and Mobile Communication Networks by the IEEE Communications Society (United Kingdom and Republic of Ireland Chapter) in 2009. He also received 2012 Chinese Government Award for Outstanding Self-Financed Student Abroad and 2017 Dean's Award, Faculty of Physical Sciences and Engineering, the University of Southampton.

Lajos Hanzo (http://www-mobile.ecs.soton.ac.uk, https://en.wikipedia.org/wiki/Lajos_Hanzo) (FIEEE'04) received Honorary Doctorates from the Technical University of Budapest and Edinburgh University. He is a Foreign Member of the Hungarian Science-Academy, Fellow of the Royal Academy of Engineering (FREng), of the IET, of EURASIP and holds the IEEE Eric Sumner Technical Field Award.