Received 25 July 2020; revised 24 August 2020; accepted 14 September 2020. Date of publication 18 September 2020; date of current version 8 October 2020.

Digital Object Identifier 10.1109/OJCOMS.2020.3024724

Machine Learning Assisted Adaptive Index Modulation for mmWave Communications

HAOCHEN LIU, SIYAO LU (Member, IEEE), MOHAMMED EL-HAJJAR (Senior Member, IEEE), AND LIE-LIANG YANG (Fellow, IEEE)

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

CORRESPONDING AUTHOR: M. EL-HAJJAR (e-mail: meh@ecs.soton.ac.uk)

This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) under Project EP/P034282/1, and in part by the Royal Academy of Engineering.

ABSTRACT In this article, we propose an orthogonal frequency-division multiplexing system supported by the compressed sensing assisted index modulation, termed as (OFDM-CSIM), applied to millimeter-wave (mmWave) communications. In the OFDM-CSIM mmWave system, information is conveyed not only by the classic constellation symbols but also by the on/off status of subcarriers, where the size of constellation symbols and the number of active subcarriers can be beneficially configured for maximizing the system's throughput. We conceive a machine learning (ML) assisted adaptive OFDM-CSIM mmWave system, which simultaneously benefits from the OFDM with index modulation (IM), compressed sensing (CS) and the hybrid beamforming techniques. Specifically, a ML-assisted link adaptation scheme is designed based on the k-nearest neighbors (k-NN) algorithm with the objective to maximize the system's throughput. Our studies show that the proposed ML-assisted link adaptation is capable of providing higher throughput than the conventional threshold-based link adaptation when different antenna structures are considered. Furthermore, the achievable data rates of four types of antenna arrays, including uniform linear array (ULA), uniform rectangular planar array (URPA), uniform circle planar array (UCPA) and uniform cylindrical array (UCYA), are investigated and compared over mmWave channels. The simulation results show that the UCYA achieves the highest data rate among these antenna arrays.

INDEX TERMS OFDM, mmwave, index modulation, compressed sensing, hybrid beamforming, link-adaptation, machine learning, k-nearest neighbour.

I. INTRODUCTION

With the development of wireless industry and the worldwide deployment of mobile networks, there has been a dramatic increase in the number of mobile devices, since powerful smart phones, laptops as well as wearable devices for entertainment have been becoming more popular and essential in our daily life. In recent years, mmWave frequency band has attracted the community's attention for development of the future generations of wireless systems, owing to its potential to meet the ever-growing wireless capacity demanded in the coming years [1]. However, mmWave signals suffer from high path-loss and shadowing losses due to its short wavelength, spanning from 1 mm to 10 mm [2]. Fortunately, the short wavelength of mmWave enables a large number of antennas distributed over a relatively compact area at both the transmitter and receiver sides, which allows the employment of advanced beamforming techniques for overcoming the pathloss [2].

According to practical measurements [3], beamforming can provide a significant gain to overcome the propagation pathloss and other absorption in mmWave frequency band. In conventional multiple-input multiple-output (MIMO) systems, the full-digital beamforming is usually employed, which simultaneously controls the phase and amplitude of the transmit/receive signals in digital domain. However, this kind of full-digital beamforming is infeasible for operation with mmWave systems, as each antenna element requires a radio frequency chain, leading to high cost and power consumption [4]–[6]. To mitigate this hardware limitation, full-analog beamforming can be employed, which however

results in significant performance loss in comparison with the full-digital beamforming. In order to attain low-complexity transceivers while approaching the performance of full-digital beamforming, the principle of hybrid beamforming was proposed and widely studied [4]–[9]. In principle, hybrid beamforming divides the beamforming into an analog part and a digital part. This beamforming arrangement allows to reduce the hardware complexity and power consumption, while maintaining a desirable system performance that may be close to that achievable by a full-digital beamformer [10], [11].

Beamforming enhances the gain of the signals in particular directions via controlling the signal outputs from the different antenna elements in an antenna array, so as to achieve directional transmission. The direction resolution of an array is increased with the increase of the number of antenna elements, which can be placed in various geometries [12], [13]. Generally, the array geometries can be divided into three categories based on their dimensions, namely, linear, planar, and volumetric arrays [14]. Linear array is the most common array considered in the research on hybrid beamforming [4]–[6]. However, linear array is a one-dimensional antenna array that has no resolution capability in the azimuth direction. On the other hand, two- and three-dimensional antenna arrays have recently drawn increasing attention in research [13]–[16].

mmWave channels are in general wideband channels, owing to the fact that high bandwidth is usually used for attaining high data rate [12], [17]. According to the 5G New Radio [18], OFDM has been considered as a typical signalling scheme for operation with mmWave communications [6], [8]. On the other hand, index modulation (IM) has been regarded as a promising modulation scheme, which is capable of providing additional degrees of freedom for data modulation, as well as the flexibility to strike a good trade-off between spectral efficiency (SE) and energy efficiency (EE). In the literature, OFDM with subcarrier IM (OFDM-IM) has been proposed in [19], which was extended and studied by following the principles of spatial modulation (SM) [20]–[25]. In OFDM-IM [21], information is conveyed by both the conventional amplitude phase modulation (APM) and the indices of subcarriers. In comparison with the conventional OFDM, OFDM-IM may achieve a better error performance [21] and can also provide an opportunity to attain a good design trade-off among SE, EE and complexity [19], [26], [27]. Moreover, to improve the flexibility of OFDM-IM, the compressed sensing (CS) concept can be introduced to form a virtual index space for IM operation [26]. This enables OFDM-IM to benefit from CS for further improving the trade-off between SE and EE [26]–[28]. Additionally, by introducing appropriate interleaving, the OFDM with CS-assisted index modulation (CSIM) is also capable of achieving a higher diversity gain and better error performance than the OFDM-IM [26].

In wireless communications system design, the time-varying wireless channel is among the biggest challenges, where link adaptation can be employed to optimize the SE. Conventionally, link adaptation is achieved via adaptive modulation and coding (AMC), which selects the appropriate modulation order and coding rate to maximize throughput at a given reliability [29]–[31]. In the conventional link adaptation schemes, transmission modes are determined according to the corresponding thresholds set based on the channel statistics [32]–[34]. However, the thresholds in the conventional link adaptation are usually hard to be set to near optimum, as the result of the deficiencies introduced at the various stages of a wireless system, including time-varying channel, non-linearity of amplifier, transmission frequency instability, *etc.* [35]. By contrast, machine learning (ML)-based approaches treat the physical layer as the transfer between system state and data observations [36], which have the potential to overcome the disadvantages of the conventional link adaptation techniques. Hence, the application of ML for link adaptation in wireless communications has drawn attention [31], [33], [37]. Specifically, a supervised learning algorithm was proposed in [31], where the observation data correlated with the conditions are directly used to control the transmission modes.

Bearing in mind the above-mentioned issues, in this contribution, the advantages of adaptive transmission and ML for information transmission over mmWave channels are combined. Explicitly, we first propose an OFDM-CSIM system working with the conventional adaptive concept, in order to take advantage of the individual components of the proposed system. Then, a ML-assisted adaptation scheme is introduced to the OFDM-CSIM for further enhancing its performance beyond its conventional counterpart. To summarize, our novel contributions can be listed as follows.

- A comprehensive survey of the different antenna structures considered in mmWave communications is provided. The effect of the antenna structures on the achievable SE performance is also investigated in the context of the hybrid beamforming applied in mmWave communications.
- We intrinsically amalgamate the concept of adaptive modulation with the OFDM-CSIM system communicating over mmWave channels. In our adaptive OFDM-CSIM systems, the modulation order and the number of active subcarriers are adaptively adjusted according to the communication conditions, with the objective of maximizing the system's throughput for a constrained target bit error ratio (BER). The main advantage of our proposed work, as compared to the state-of-the-art such as [29], [32], is the attainable flexibility of the design in terms of providing a flexible trade-off among SE, EE and complexity.
- Based on the k-nearest neighbour (k-NN) algorithm, an ML-assisted adaptive modulation scheme for operation with the OFDM-CSIM is proposed. Our studies show that the ML-assisted adaptive OFDM-CSIM scheme is capable of choosing better transmission modes than the conventional adaptive schemes, and hence providing a

significantly increased throughput than the conventional ones.

The rest of the paper is organized as follows. In Section II, a survey of the antenna arrays for wireless beamforming is presented. Section III introduces the model of the proposed system, including the design of hybrid beamforming, mmWave channel model and the joint maximum likelihood (JML) detector. In Section IV, both the conventional adaptive modulation and the ML-assisted adaptive modulation are addressed. In Section V, the simulation results are presented and analyzed. Finally, the conclusions and some suggestions for future research are provided in Section VI.

Notations: Lower-case boldface letter $\boldsymbol{a}$ and upper-case boldface letter $\boldsymbol{A}$ denote vectors and matrices, respectively; $[\cdot]_{\mathrm{i}}$ denotes the i-th element of a vector; $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ express transpose and conjugate transpose operations, respectively; Λ denotes complex signal symbol; $\mathbb{C}^{A\times B}$ is the set of $(A\times B)$-element in the complex field; $\{\cdot\}_a^b$ is a sequence with the indices from a to b; $\mathbb{E}[\cdot]$ is the expectation operator; $\mathrm{diag}(\boldsymbol{a})$ expresses a diagonal matrix formed from vector $\boldsymbol{a}$; $\mathrm{Tr}(\cdot)$ and $|\cdot|$ represent trace and absolute value, respectively; $\binom{n}{k}$ is the combination of the selection of k items from a collection of n items; $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm; $\langle\cdot\rangle$ express inner product operations.

II. SURVEY OF THE ANTENNA ARRAYS FOR BEAMFORMING

As introduced above, large-scale antenna arrays can provide beamforming gain to overcome the propagation losses in mmWave systems, while the directional transmission with narrow beams can also improve the throughput by reducing the interference from other transmissions [38], [39]. Furthermore, large-scale antenna arrays provide the opportunity for multiple data stream transmission, which improves the SE of the systems [40] by exploiting spatial reuse. Although the short wavelength of mmWave frequency bands allows to pack a large number of antenna elements within a finite area, the geometrical shapes of antenna arrays impose some challenges on mmWave transceivers and affect the system performance, which still require further investigations. The geometry of antenna arrays in wireless communications can be classified as linear arrays, planar arrays and volumetric arrays, which are one-dimensional, two-dimensional and three-dimensional arrays, respectively. In this section, a survey of these different antenna arrays is provided.

In general, the array manifold vector of an antenna array can be represented as

$$[\boldsymbol{v}(\boldsymbol{k})]_n = e^{-j\boldsymbol{k}^{\mathrm{T}}\boldsymbol{p}_n},\ \ n=0,1,\ldots,N-1, \tag{1}$$

where

$$\boldsymbol{k} = -\frac{2\pi}{\lambda}\begin{bmatrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\\ \cos\theta\end{bmatrix} \tag{2}$$

expresses the wavenumber, $\boldsymbol{p}_n$ denotes the location of the n-th element. θ and φ are the grazing angles with respect to the xz-plane and xy-plane, respectively. The array manifold vector includes all the spatial characteristics of the array and is also the main factor to distinguish different antenna structures [41].

Fig. 1 shows some examples of the typical antenna arrays. Specifically, Fig. 1(a) shows the geometry of a uniform linear array (ULA), where the elements are located on the z-axis, and d_z is the distance between two adjacent elements. It can be shown that the array manifold vector of ULA is

$$[\boldsymbol{v}_{\mathrm{ULA}}(\theta)]_n = e^{-jk_z p_{z_n}} = e^{j\left(n-\frac{N-1}{2}\right)d_z\frac{2\pi}{\lambda}\cos\theta},\quad n=0,\ldots,N-1, \tag{3}$$

where λ is the signal's wavelength, p_{z_n} is the location of an element on the z-axis, and $k_z = -\frac{2\pi}{\lambda}\cos\theta$. Usually for ULA, $d_z = \frac{\lambda}{2}$. Also, note that the linear arrays can only control the signals in the θ-direction.

Fig. 1(b) shows the geometry of a uniform rectangular planar array (URPA) with $M\times N$ elements, which are distributed as a square in the xy-plane. When the distance between two adjacent elements is set to $d_x = d_y = \lambda/2$, the array manifold vector can be found to be

$$[\boldsymbol{v}_{\mathrm{URA}}(\theta,\varphi)]_{mn} = e^{-j\boldsymbol{k}_{x,y}^{\mathrm{T}}\boldsymbol{p}_{x_n,y_m}} = e^{j(n\psi_x+m\psi_y)},\quad n=0,\ldots,N-1;\ \ m=0,\ldots,M-1, \tag{4}$$

where $\boldsymbol{k}_{x,y} = -\frac{2\pi}{\lambda}\begin{bmatrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\end{bmatrix}$ and $\boldsymbol{p}_{x_n,y_m} = \begin{bmatrix}nd_x\\ md_y\end{bmatrix}$ giving

$$\psi_x = \frac{2\pi}{\lambda}\sin\theta\cos\varphi d_x,$$
$$\psi_y = \frac{2\pi}{\lambda}\sin\theta\sin\varphi d_y.$$

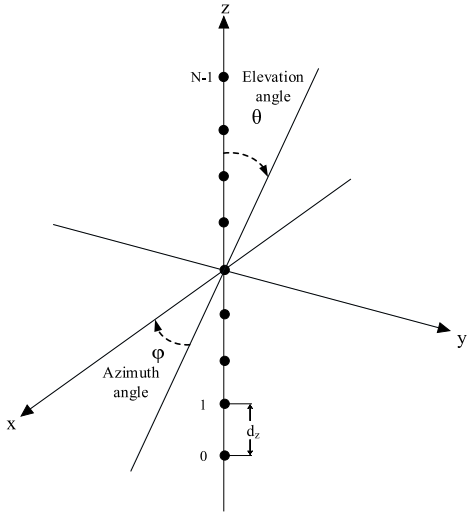
For the URPA array shown in Fig. 1(b), ψ_x, ψ_y account for the spatial characteristics of the horizontal elements.

By contrast, Fig. 1(c) illustrates the geometry of a uniform circle planar array (UCPA), which is also a two-dimensional array but with N elements located uniformly on a circle in the xy-plane. In Fig. 1(c), R is the radius of the array, φ_n is the angle of the n-th element, with $\varphi_n = \frac{2\pi}{N}(n-1)$. The array manifold vector of UCPA can be expressed as

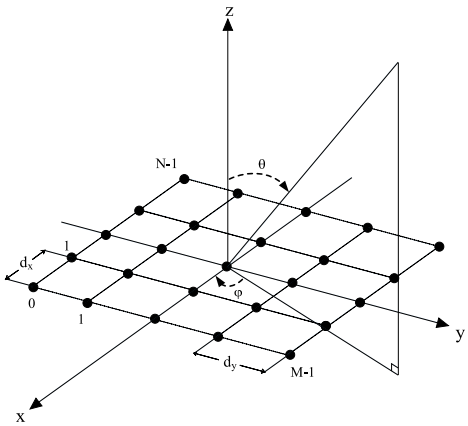
$$[\boldsymbol{v}_{\mathrm{UCPA}}(\theta,\varphi)]_n = e^{-j\boldsymbol{k}_{x,y}^{\mathrm{T}}\boldsymbol{p}_{x_n,y_n}} = e^{j\frac{2\pi}{\lambda}R\sin\theta\cos(\varphi-\varphi_n)},\quad n=0,\ldots,N-1. \tag{5}$$

obtained by setting $\boldsymbol{k}_{x,y} = -\frac{2\pi}{\lambda}\begin{bmatrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\end{bmatrix}$ and $\boldsymbol{p}_{x_n,y_n} = \begin{bmatrix}R\cos\varphi_n\\ R\sin\varphi_n\end{bmatrix}$. One of the advantages of UCPA over URPA is that it can be rotated with respect to the z-axis without changing the shape of the beams significantly [42].

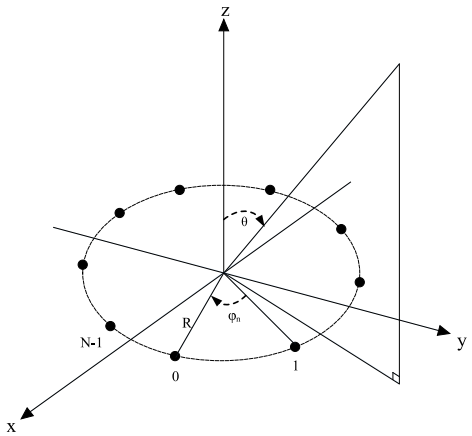
In contrast to the one-dimensional ULA, the two-dimensional antenna arrays are capable of forming the beams in both the θ-direction and φ-direction, and can also distribute more antenna elements in a given area [43]. However,



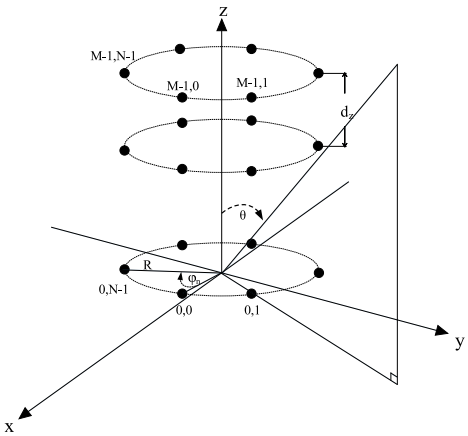
(a) One-dimensional ULA with N elements along z-axis.



(b) Two-dimensional URPA having $M \times N$ array elements arranged on a rectangular planar.



(c) Two-dimensional UCPA having N array elements uniformly arranged on a circle.



(d) Three-dimensional UCYA with $M \times N$ array elements arranged as a cylinder.

FIGURE 1. **Examples of antenna arrays.**

the studies in [13] show that for a given number of elements, UCPA experiences higher inter-user interference than ULA. Besides, ULA can always achieve the highest sum-rate, as it does not have the resolution capability in the azimuth direction, which increases the channel correlation. In [43], the authors compared the achievable SE of different planar antenna arrays in mmWave massive MIMO systems, showing that the two-dimensional arrays with the same number of elements as the UCPA and URPA have no significant benefit on the achievable SE.

In many practical applications of interest, the array elements may not be distributed on a flat surface. Instead, three-dimensional arrays may be necessary, which may be in different structures, such as spherical, cylindrical, *etc.* As an example, Fig. 1(d) shows a uniform cylindrical array (UCYA), which is the most commonly used one [16]. As shown in Fig. 1(d), the UCYA has M sub-arrays, each of which is a UCPA with radius R and contains N elements uniformly distributed on the circle, with the phases expressed as $\varphi_n = \frac{2\pi}{N}(n-1)$, for $n = 0, \ldots, N-1$. The distance between two adjacent UCPA sub-arrays is expressed as d_z. Hence, we have $\boldsymbol{k}_{x,y,z} = -\frac{2\pi}{\lambda}\begin{bmatrix}\sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta\end{bmatrix}$, $\boldsymbol{p}_{x_n,y_n,z_m} = \begin{bmatrix}R\cos\varphi_n \\ R\sin\varphi_n \\ md_z\end{bmatrix}$, and it can be shown that the array manifold vector is given by

$$[\boldsymbol{v}_{\mathrm{UCYA}}(\theta,\varphi)]_{m,n} = e^{j\frac{2\pi}{\lambda}[md_z\cos(\theta)+R\sin\theta\cos(\varphi-\varphi_n)]}, \quad m = 0,\ldots,M-1,\ n = 0,\ldots,N-1. \tag{6}$$

It can be shown that the array manifold vector of the UCYA is the product of the manifold vector of the ULA on the z-axis and that of the UCPA in the xy-plane. This can be explained by the fact that the UCYA can be geometrically seen as the ULA sub-arrays isometrically distributed in parallel along the

surface of a cylinder [44]. In contrast to ULA and UCPA, for a given number of elements, UCYA is capable of achieving the highest gain, narrowest beam and lowest sidelobes [45].

Three-dimensional antenna arrays are useful for some applications, owing to their high efficiency and robustness [15], [16], [46], [47]. One of the advantages of the three-dimensional antenna arrays is the increase of the coverage gain without changing the number of antenna elements. This is because the three-dimensional antenna arrays can provide wider coverage in both azimuth and elevation direction than the one- or two-dimensional arrays [16], [47]. Three-dimensional antenna arrays are also attractive for the implementation of adaptive antennas, owing to their high degrees of freedom [48]. Moreover, the studies in [14] show that when operated in mmWave MIMO propagation environments, the three-dimensional UCYA outperforms the two-dimensional URPA, when both the effective interference and the achievable rate are considered.

According to [6], [11], mmWave channel can be modeled as a clustered channel with N_{cl} clusters and N_{ray} rays in each cluster. Specifically, the frequency-domain channel matrix $\boldsymbol{H}(\beta)$ for the β-th subcarrier of OFDM systems can be expressed as

$$\boldsymbol{H}(\beta) = r \sum_{l=1}^{N_{\text{cl}}} \sum_{i=1}^{N_{\text{ray}}} \alpha_{li} \boldsymbol{v}_r\left(\theta_{li}^r, \varphi_{li}^r\right) \boldsymbol{v}_t\left(\theta_{li}^t, \varphi_{li}^t\right)^{\text{H}} e^{-j2\pi \psi_l \frac{\beta}{M}}, \quad (7)$$

where $r = \sqrt{\frac{N_t N_r}{N_{\text{cl}} N_{\text{ray}}}}$ is a normalization factor, ψ_l is proportional to the phase shift of the l-th cluster, α_{li} is the complex gain of the i-th ray in the l-th cluster, θ_{li}^r and φ_{li}^r are the elevation angle and azimuth angle of the arrival of the i-th ray in the l-th cluster, respectively, while θ_{li}^t and φ_{li}^t are the corresponding elevation angle and azimuth angle of departure, respectively. Finally, $\boldsymbol{v}_r(\cdot)$ and $\boldsymbol{v}_t(\cdot)$ express the array manifold vectors of the transmitter antenna array and the receiver antenna array, respectively. It can be shown that (7) can also be written as

$$\boldsymbol{H}(\beta) = \boldsymbol{V}_r \text{diag}(\boldsymbol{\alpha}(\beta)) \boldsymbol{V}_t^{\text{H}}, \quad (8)$$

where, by definition

$$\boldsymbol{V}_r = \left[\boldsymbol{v}_r\left(\theta_{11}^r, \varphi_{11}^r\right), \boldsymbol{v}_r\left(\theta_{12}^r, \varphi_{12}^r\right), \ldots, \boldsymbol{v}_r\left(\theta_{N_{\text{cl}}N_{\text{ray}}}^r, \varphi_{N_{\text{cl}}N_{\text{ray}}}^r\right)\right],$$

$$\boldsymbol{V}_t = \left[\boldsymbol{v}_t\left(\theta_{11}^t, \varphi_{11}^t\right), \boldsymbol{v}_t\left(\theta_{12}^t, \varphi_{12}^t\right), \ldots, \boldsymbol{v}_t\left(\theta_{N_{\text{cl}}N_{\text{ray}}}^t, \varphi_{N_{\text{cl}}N_{\text{ray}}}^t\right)\right],$$

$$\boldsymbol{\alpha}(\beta) = \left[\alpha_{11} e^{-j2\pi \psi_1 \frac{\beta}{M}}, \alpha_{12} e^{-j2\pi \psi_1 \frac{\beta}{M}}, \ldots, \alpha_{N_{\text{cl}}N_{\text{ray}}} e^{-j2\pi \psi_{N_{\text{cl}}} \frac{\beta}{M}}\right].$$

$\boldsymbol{V}_r \in \mathbb{C}^{N_r \times N_{\text{cl}} N_{\text{ray}}}$ and $\boldsymbol{V}_t \in \mathbb{C}^{N_t \times N_{\text{cl}} N_{\text{ray}}}$ include all the array manifold vectors of the receiver and transmitter, and $\text{diag}(\boldsymbol{\alpha}(\beta)) \in \mathbb{C}^{N_{\text{cl}} N_{\text{ray}} \times N_{\text{cl}} N_{\text{ray}}}$ is a diagonal matrix that contains the complex gains of the rays in all the clusters.

Fig. 2 shows the achievable data rate against average SNR for the different array configurations operated in the MIMO-OFDM systems communicating over mmWave channels. The propagation environment is assumed to have $N_{\text{cl}} = 8$ clusters with each cluster consisting of $N_{\text{ray}} = 10$ rays, while the azimuth and elevation angles of both arrival and departure are assumed to obey the Laplacian distribution with the angular spreads of 7.5 degrees [4]. The transmitter is equipped with the ULA, URPA, UCPA and UCYA having 64 elements, while the receiver employs a ULA of 32 elements. The singular value decomposition (SVD) beamformer is assumed to be employed at both transmitter and receiver [4]. As shown in Fig. 2, UCYA slightly outperforms UCPA, URPA and ULA. The one-dimensional array ULA achieves a higher data rate than the two-dimensional arrays of URPA and UCPA, due to the fact that ULA has no resolution in the azimuth direction, which reduces channel correlation. UCPA and URPA achieve almost identical data rate. UCYA shows no significant advantages over the other arrays in terms of the achievable data rates. However, UCYA is capable of providing a wider coverage range and also higher degrees of freedom than the other arrays. Moreover, the simulation results in [14] show that UCYA demonstrates advantages over the two-dimensional antenna arrays with respect to the radiation pattern, the channel eigenvalue distribution and the effective interference suppression.

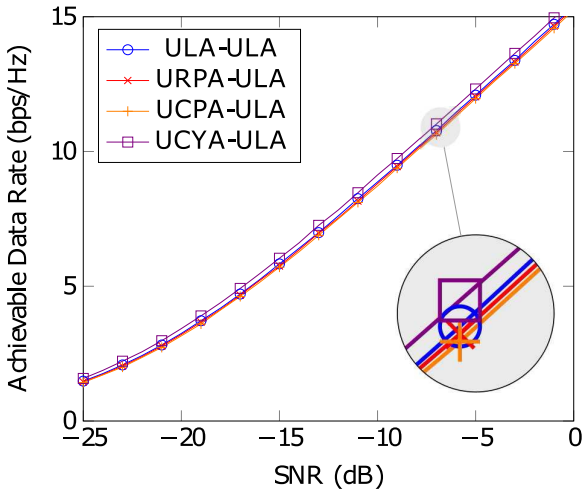


FIGURE 2. Achievable data rate versus the SNR for a system employing a transmitter with 64 elements and ULA, URPA, UCPA and UCYA and a receiver employing ULA with 32 elements.

III. PROPOSED SYSTEM MODEL

Fig. 3 shows the block diagram of the proposed system, which can be divided into four parts shown in different colors identified by Roman numerals in Fig. 3. The first part in blue color illustrates the CSIM, where the input bit sequence is split into G groups, and a part of the bits in each group are used for selecting active subcarriers, while the others are employed to modulate amplitude-phase modulation (APM) symbols. The outputs of both parts of the G groups are combined to form an OFDM block, which is then compressed using a measurement matrix from high (N) dimensions in the virtual domain to low (M) dimensions in the frequency domain. Then the compressed OFDM block is input to the red part of Fig. 3, where the main operation is hybrid beamforming. Specifically, the compressed

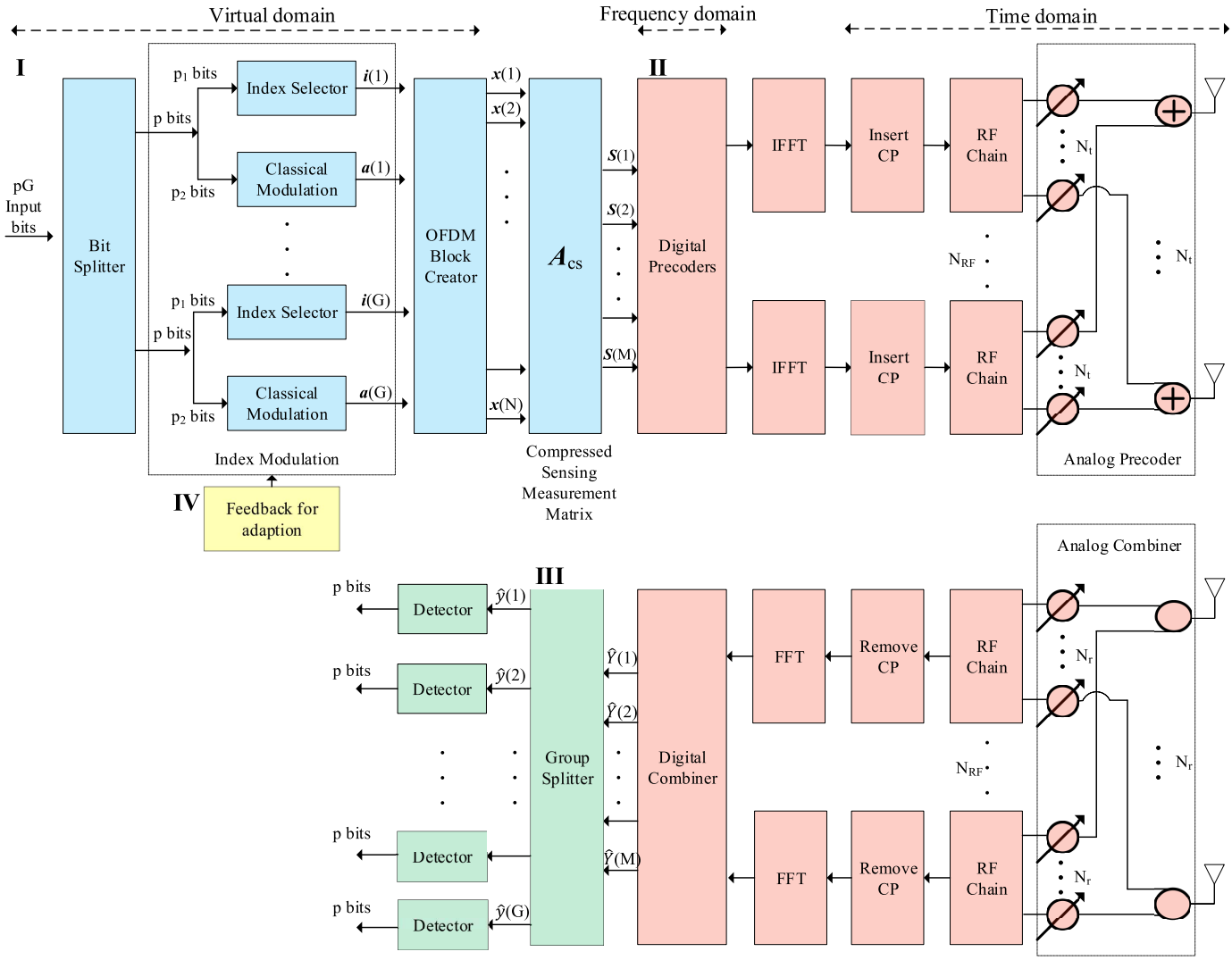


FIGURE 3. Block diagram of the adaptive modulation-aided OFDM-CSIM transceiver with hybrid beamforming.

OFDM block is first processed by a digital precoder, and then transformed from frequency domain to time domain by the inverse fast Fourier transforms (IFFT). Then, after adding the cyclic prefixes (CP), an analog precoder is applied to form the signals transmitted from the antennas.

At the receiver side, the received signal is first processed by an analog combiner, the outputs of which are then transformed back to frequency domain by the fast Fourier transforms (FFT) after removing the CP. Then, the received signal is combined using a digital combiner, which outputs the compressed OFDM block. Afterwards, the outputs of the digital combiner are divided into several sub-blocks for executing signal detection, as shown in the green part in Fig. 3, where JML detection can be employed. The details of the three parts mentioned above are provided in the following sub-sections. The fourth part of Fig. 3 colored in yellow is for the adaptive modulation, which is detailed in Section IV. Note that to implement adaptive modulation, the channel information is assumed to be fed back from the receiver to the transmitter, which is used to adjust the parameters in the index modulation part, with the aim of maximizing the throughput for a given target BER.

A. COMPRESSED SENSING ASSISTED INDEX MODULATION

As in [26], the notions of virtual domain and frequency domain are used to differentiate the signal before and after the compressed sensing. As shown in the blue part of Fig. 3, in the virtual domain, we assume that a total of pG input bits are partitioned into G groups, with each group including p bits. Let N and N_{v} denote the total number of subcarrier indices and the number of indices per group, respectively, in the virtual domain. Hence, we have $N = nG$. As shown in Fig. 3, the p bits of a group are separated into two parts of p_1 and p_2 bits, with $p = p_1 + p_2$. Then, p_1 bits are used to select F_{v} virtual indices out of the N_{v} virtual indices of a group. This gives the total number of bits carried by the

TABLE 1. A look-up table example for $F_v = 1$ and $N_v = 4$.

p_1	Index	Sub-Blocks
$[0\ \ 0]$	$\{1\}$	$[a_{g,1}\ \ 0\ \ 0\ \ 0]$
$[0\ \ 1]$	$\{2\}$	$[0\ \ a_{g,1}\ \ 0\ \ 0]$
$[1\ \ 0]$	$\{3\}$	$[0\ \ 0\ \ a_{g,1}\ \ 0]$
$[1\ \ 0]$	$\{4\}$	$[0\ \ 0\ \ 0\ \ a_{g,1}]$

positions of the active indices in a virtual OFDM block as

$$m_1 = p_1 G = \left\lfloor \log_2 \left[\binom{N_v}{F_v} \right] \right\rfloor G, \tag{9}$$

where $\lfloor a \rfloor$ returns the largest integer less than a. After the selections of indices, the active virtual indices of G groups can be represented as

$$\boldsymbol{I} = \{\boldsymbol{i}(1), \ldots, \boldsymbol{i}(g), \ldots, \boldsymbol{i}(G)\}, \tag{10}$$

where $\boldsymbol{i}(g) = \{i_{g,1}, i_{g,2}, \ldots, i_{g,f}\}$, $i_{g,\omega} \in \{1, 2, \ldots, n\}$, contains the indices activated in the g-th group, which has in total 2^{p_1} possible realizations. Following the virtual index selections, p_2 bits are used to generate the APM symbols (such as BPSK, *etc.*), to be transmitted on the active virtual indices of a group. The total number of bits conveyed by the APM symbols of G groups is

$$m_2 = p_2 G = F_v \big(\log_2(Q)\big) G, \tag{11}$$

where Q is the modulation order of APM. Let these APM symbols be collected to a matrix

$$\boldsymbol{A} = \big[\boldsymbol{a}(1), \ldots, \boldsymbol{a}(g), \ldots, \boldsymbol{a}(G)\big], \tag{12}$$

where $\boldsymbol{a}(g) = [a_{g,1}, a_{g,2}, \ldots, a_{g,F_v}]^{\mathrm{T}}$ contains the APM symbols to be transmitted on the F_v active virtual indices of group g, and the APM symbols are normalized as $\mathbb{E}[|a_{g,\omega}|^2] = 1$. Hence, based on the above discussion, the total number of information bits transmitted per symbol period is $pG = (p_1 + p_2)G = m_1 + m_2$.

Note that, the value of F_v for different virtual groups may be different. However, for simplicity, the same value of F_v for all groups is assumed in this article. For example, Table 1 shows the mappings for $p_1 = 2$, $F_v = 1$ and $N_v = 4$. Since one out of 4 indices is activated per symbol period, there are in total 4 possible realizations, which can carry 2 bits, in addition to the bits conveyed by the AMP symbol transmitted on the active virtual index.

After the index selections and symbol mappings, the OFDM block creator seen in Fig. 3 combines all the G groups of symbols in $\boldsymbol{A}$ to generate a virtual OFDM block of $(N = N_v G)$-length, while referring to $\boldsymbol{I}$ of (10) for the locations of the APM symbols. Specifically, the AMP symbol $a_{g,w}$ is located on the active virtual index $i_{g,w}$. Furthermore, owing to the employment of transmit antenna array, we assume that each active subcarrier conveys N_s APM symbols, as detailed in Section III-B. Then, the virtual OFDM

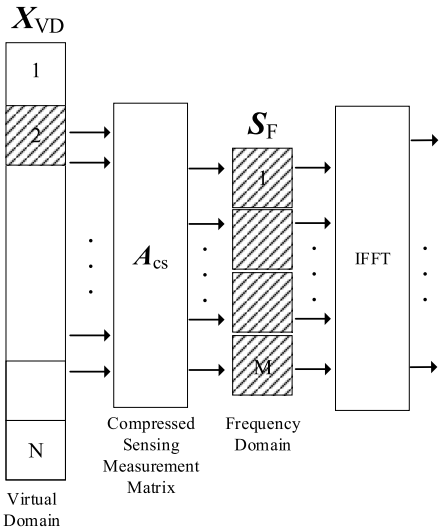


FIGURE 4. Illustration of subcarrier index modulation with compressed sensing. N-dimensional vector X_{VD} in the virtual domain is compressed to a M-dimensional vector S_F using measurement matrix A_{CS}, where $N > M$. Active subcarriers and virtual indices are expressed as shaded boxes.

block $\boldsymbol{X}_{\mathrm{VD}}$ containing the $N_s F_v G$ APM symbols transmitted in one OFDM symbol period can be expressed as

$$\boldsymbol{X}_{\mathrm{VD}} = \big[\boldsymbol{x}(1), \ldots, \boldsymbol{x}(\gamma), \ldots, \boldsymbol{x}(N)\big]^{\mathrm{T}}, \tag{13}$$

where $\boldsymbol{X}_{\mathrm{VD}} \in \mathbb{C}^{N \times N_s}$, $\boldsymbol{x}(\gamma) \in \mathbb{C}^{N_s \times 1}$ contains N_s data symbols associated with the γ-th virtual index, and $\boldsymbol{x}_{\gamma,j} \in \{0, \Lambda\}$ with '0' indicating an inactive index, where $1 < j \leq N_s$, and Λ expressing an APM symbol. To benefit from the compressed sensing, $\boldsymbol{X}_{\mathrm{VD}}$ should be a sparse matrix containing mostly zero elements. However, this virtual vector needs to be transmitted on M subcarriers. Hence, a measurement matrix $\boldsymbol{A}_{\mathrm{cs}} \in \mathbb{C}^{M \times N}$, as illustrated in Fig. 4, is employed to map the virtual matrix $\boldsymbol{X}_{\mathrm{VD}}$ of N-dimension to the M-dimensional frequency domain, $M < N$, which can be expressed as

$$\boldsymbol{S}_{\mathrm{F}} = \boldsymbol{A}_{\mathrm{cs}} \boldsymbol{X}_{\mathrm{VD}}, \tag{14}$$

where $\boldsymbol{S}_{\mathrm{F}} \in \mathbb{C}^{M \times N_s}$, $\boldsymbol{S}_{\mathrm{F}} = [\boldsymbol{s}(1), \ldots, \boldsymbol{s}(\beta) \ldots, \boldsymbol{s}(M)]^{\mathrm{T}}$, $\boldsymbol{s}(\beta) \in \mathbb{C}^{N_s \times 1}$ is the N_s data symbols transmitted by the β-th subcarrier with the aid of N_t transmit antenna elements. Correspondingly, M subcarriers are divided into G groups with each having $M_{\mathrm{f}} = \frac{M}{G}$ subcarriers. It can be shown that the OFDM-CSIM scheme can effectively improve the throughput of the system. For example, for a virtual dimension of $N_v = 256$ per group, $M_{\mathrm{f}} = 8$, $F_v = 2$ and $Q = 2$, where Q is the modulation order of APM, the OFDM-CSIM system can achieve a throughput of 2 bit/s/Hz. By contrast, the conventional OFDM-IM system with $N_v = M_{\mathrm{f}} = 8$, $F_v = 2$ and $Q = 2$ can only achieve a throughput of 0.75 bit/s/Hz.

However, in order for the receiver to have a good recovery performance, the virtual matrix $\mathbf{X}_{\mathrm{VD}}$ needs to have a low sparsity, i.e., $F_v \ll N_v$. In principle, to guarantee the efficiency of compressed sensing, the columns of the

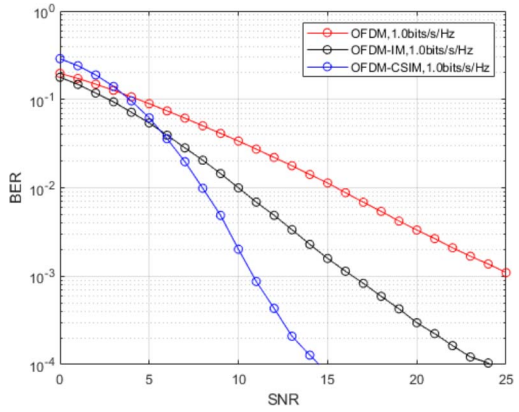


FIGURE 5. BER performance comparison of OFDM, OFDM-IM and OFDM-CSIM systems with the same throughput of 1.0 bits/s/Hz.

measurement matrix $\boldsymbol{A}_{\rm cs}$ as shown in (14) should be as uncorrelated as possible [26]. According to [26], [49], the simplest method to generate a measurement matrix is to make $\mathbf{A}_{\rm cs}$ a random matrix, such as random Gaussian matrix, which has a high probability to satisfy the requirements for attaining a good recovery performance. The rigorous approach to verify the quality of measurement matrix is computing its mutual coherence $\mu(\mathbf{A}_{\rm cs})$ [26], as

$$\mu(\mathbf{A}_{\rm cs}) = \max_{i\neq j} \frac{\left|\langle \mathbf{A}_{{\rm cs},i}, \mathbf{A}_{{\rm cs},j}\rangle\right|}{\left\|\mathbf{A}_{{\rm cs},i}\right\|_F \left\|\mathbf{A}_{{\rm cs},j}\right\|_F}, \tag{15}$$

where $\mathbf{A}_{{\rm cs},i}$ and $\mathbf{A}_{{\rm cs},j}$ are different columns in the measurement matrix $\mathbf{A}_{\rm cs}$. Then, as shown in [26], there exists one sparse vector $\mathbf{X_{VD}}$ satisfying $\mathbf{S}_{\rm F} = \mathbf{A}_{\rm cs}\mathbf{X}_{\rm VD}$, provided that $F_{\rm v} < \frac{1}{2}(1 + \frac{1}{\mu(\mathbf{A}_{\rm cs})})$, which means an ideal recovery.

Fig. 5 shows that OFDM-CSIM outperforms both OFDM-IM and OFDM, which becomes more significant in the high SNR region, when the same spectral-efficiency is considered. Note that, for the results shown in Fig. 5, we assumed 58 sub-carriers used in the conventional OFDM system with QPSK modulation. The parameters for OFDM-IM system are $M_{\rm f} = 256$, $F_{\rm v} = 2$, $Q = 4$, while for the OFDM-CSIM system these are $N_{\rm v} = 256$, $M_{\rm f} = 32$, $F_{\rm v} = 2$, $Q = 4$. Here, we should point out that the number of transmitted subcarriers will cause a certain reduction in power gain, rather than a change in the slope of the BER curve [26].

B. HYBRID BEAMFORMING

Consider a single-user OFDM-based mmWave system, which employs N_t transmit antennas, N_r receiver antennas and supports N_s data streams per subcarrier. As mentioned previously in Section I, beamforming is essential in mmWave communication, in order to obtain a desirable antenna gain. In this article, we employ the hybrid beamforming as shown by the red-colored part of Fig. 3, owing to its reduced complexity and near-full-digital beamforming performance. In our system, we assume that $N^t_{\rm RF}$ and $N^r_{\rm RF}$ RF chains are employed by the transmitter and receiver, respectively, where the number of RF chains is smaller than the corresponding number of antennas, i.e., $N^t_{\rm RF} \ll N_t$, $N^r_{\rm RF} \ll N_r$. Additionally, the fully connected hybrid beamforming architecture [4] is employed in this article.

As shown in Fig. 3, N_s transmitted symbols $\mathbf{s}(\beta) \in \mathbb{C}^{N_s\times 1}$ on the βth subcarrier, $\beta = 1, \ldots, M$, are precoded using a digital precoder $\boldsymbol{F}_D(\beta) \in \mathbb{C}^{N^t_{\rm RF}\times N_s}$, yielding $N^t_{\rm RF}$ outputs in the frequency domain. Then, the outputs of the M numbers of digital precoders are re-arranged to form $N^t_{\rm RF}$ OFDM blocks, where each digital precoder contributes one symbol to one OFDM block according to its assigned subcarrier. Then, the $N^t_{\rm RF}$ OFDM blocks are transformed to the time-domain using the M-point IFFTs. Finally, after adding the CP and processing by a common analog precoder $\boldsymbol{F}_{RF} \in \mathbb{C}^{N_t\times N^t_{\rm RF}}$, the final transmitted signal on subcarrier β can be represented in a baseband equivalent form as

$$\boldsymbol{u}(\beta) = \boldsymbol{F}_{RF}\boldsymbol{F}_D(\beta)\boldsymbol{s}(\beta), \tag{16}$$

where $\mathbb{E}[\boldsymbol{s}(\beta)\boldsymbol{s}^{\rm H}(\beta)] = \frac{1}{N_s}\boldsymbol{I}_{N_s}$.

Since the analog beamforming is implemented by phase shifters, the elements in $\boldsymbol{F}_{RF}$ are constrained with constant magnitude. The constraint on the transmitted power is given by $\|\boldsymbol{F}_{\rm RF}\boldsymbol{F}_{\rm D}(\beta)\|_2 = N_s$. It is worth noting that the design of analog beamformer for each subcarrier in a multi-carrier system (e.g., OFDM) is identical. This gives the main difference from the design in single-carrier system, and is also the challenging part of precoder design in multi-carrier systems [6]. When the signal of (16) is transmitted over mmWave channels, the received signal can be expressed as

$$\boldsymbol{y}(\beta) = \boldsymbol{H}(\beta)\boldsymbol{F}_{\rm RF}\boldsymbol{F}_{\rm D}(\beta)\boldsymbol{s}(\beta) + \boldsymbol{n}(\beta), \tag{17}$$

where $\boldsymbol{H}(\beta) \in \mathbb{C}^{N_r\times N_t}$ is the mmWave channel matrix, and $\boldsymbol{n}(\beta) \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{N_r})$ is the vector of additive white Gaussian noise (AWGN), with σ^2 representing the noise variance.

At the receiver side, as shown by the red-colored part of Fig. 3, the received signal is first processed by an analog combiner $\boldsymbol{W}_{\rm RF} \in \mathbb{C}^{N_r\times N^r_{\rm RF}}$, which is common to all subcarriers. Following this, the CP is removed and then the signal is transformed from time domain to frequency domain by the M-point FFTs. Then, the signal is processed by a digital combiners $\boldsymbol{W}_D(\beta) \in \mathbb{C}^{N^r_{\rm RF}\times N_s}$ to generate the final output. When considering all the above mentioned processing, the output can be represented as:

$$\begin{aligned}\hat{\boldsymbol{Y}}(\beta) = {} & \boldsymbol{W}^{\rm H}_{\rm D}(\beta)\boldsymbol{W}^{\rm H}_{\rm RF}\boldsymbol{H}(\beta)\boldsymbol{F}_{\rm RF}\boldsymbol{F}_{\rm D}(\beta)\boldsymbol{s}(\beta) \\ & + \boldsymbol{W}^{\rm H}_{\rm D}(\beta)\boldsymbol{W}^{\rm H}_{\rm RF}\boldsymbol{n}(\beta).\end{aligned} \tag{18}$$

Based on [4]–[7], hybrid beamforming is usually designed to optimize the system SE. However, the joint optimization of the precoder and combiner matrices is highly challenging, as explained in [6], [21]. To simplify the design of hybrid beamforming, the joint optimization is typically decoupled into the transmitter and receiver optimization. Specifically at the transmitter side, the design of the precoder ($\boldsymbol{F}_{\rm RF}$, $\boldsymbol{F}_{\rm D}$)

is carried out to maximize the SE. This can be modeled mathematically as [6]:

$$\left(\boldsymbol{F}_{\mathrm{RF}}^{\mathrm{opt}}, \boldsymbol{F}_{\mathrm{D}}^{\mathrm{opt}}(\beta)\right) = \underset{\boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{D}}(\beta)}{\arg\min} \sum_{\beta=1}^{M} \left\|\boldsymbol{F}_{\mathrm{opt}}(\beta) - \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{D}}(\beta)\right\|_{\mathrm{F}},$$
$$s.t. \quad \boldsymbol{F}_{\mathrm{RF}} \in \left\{\boldsymbol{v}_t\left(\theta_{li}^t, \varphi_{li}^t\right) \mid 1 \le l \le N_{\mathrm{cl}}, 1 \le i \le N_{\mathrm{ray}}\right\},$$
$$\|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{D}}(\beta)\|_{\mathrm{F}}^2 = N_s, \tag{19}$$

where $\boldsymbol{F}_{\mathrm{opt}}(\beta) \in \mathbb{C}^{N_t \times N_s}$ is the optimum precoder for subcarrier β, which can be obtained from the SVD of the channel matrix for the β-th subcarrier. However, the focus of this article is not on the design of hybrid beamforming. Instead, the approach in [4] is directly adopted. At the receiver side, similarly, the solutions for $\boldsymbol{W}_{\mathrm{RF}}$ and $\boldsymbol{W}_{\mathrm{D}}(\beta)$ can be attained as above by first deriving the joint combiner. When the minimum mean square error (MMSE) criterion is applied, the joint combiner is [4]

$$\boldsymbol{W}_{\mathrm{MMSE}}(\beta) = (\boldsymbol{H}(\beta)\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}(\beta)^{\mathrm{H}} + \sigma_n^2 N_s \boldsymbol{I})^{-1}\boldsymbol{H}(\beta)\boldsymbol{F}, \tag{20}$$

where $\boldsymbol{F} = \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{D}}(\beta)$.

C. JOINT MAXIMUM LIKELIHOOD DETECTION

Finally, after the received signal is processed by the digital combiner, detection is carried out to recover the transmitted information, which is shown in the green-colored area in Fig. 3. Specifically, corresponding to the transmitter, the received signals are divided into groups in the frequency domain. Then, for each group, if the joint maximum likelihood (JML) detection is employed, the information is detected as:

$$\hat{\boldsymbol{X}}_{\mathrm{VD},g} = \arg\min_{\tilde{\boldsymbol{X}}_{\mathrm{VD},g}} \left\|\hat{\boldsymbol{y}}_g - \hat{\boldsymbol{H}}_g \boldsymbol{A}_{\mathrm{cs}} \widetilde{\boldsymbol{X}}_{\mathrm{VD},g}\right\|_{\mathrm{F}}^2,$$
$$g = 1, 2, \ldots, G, \tag{21}$$

where $\hat{\boldsymbol{y}}_g$ is the received signal of the g-th group corresponding to the transmitted signal $\widetilde{\boldsymbol{X}}_{\mathrm{VD},g}$, while

$$\hat{\boldsymbol{H}}_g = \boldsymbol{W}_{\mathrm{D},g}^{\mathrm{H}} \boldsymbol{W}_{\mathrm{RF},g}^{\mathrm{H}} \boldsymbol{H}_g \boldsymbol{F}_{\mathrm{RF},g} \boldsymbol{F}_{\mathrm{D},g},$$

and $\boldsymbol{H}_g$ can be written as:

$$\boldsymbol{H}_g = \begin{bmatrix} \boldsymbol{H}_{g,1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{H}_{g,2} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \boldsymbol{H}_{g,i} \end{bmatrix}, \tag{22}$$

where $\boldsymbol{H}_{g,i}$ represents the channel gain of the i-th subcarrier of the g-th group. After the JML detection, the p bits conveyed in each of the groups are recovered.

mmWave communication environment is typically time varying and hence the mmWave communications systems' throughput can be improved with the adoption of link adaptation. However, in real-time adaptive modulation, it is often difficult to switch the modulation mode at the right point, as the switching point depends on the SNR, which is in turn related to different issues, as described in Section I. Therefore, it is highly challenging for the transmitter to make the switching at the near optimum point. In order to circumvent these problems, in the next section, our learning-assisted adaptive system design is introduced, which does not depend on thresholds but only on the observation of the communication environment.

IV. ADAPTIVE MODULATION

In wireless communications, adaptive modulation aims to maximize the throughput, while maintaining a target BER, via adjusting the modulation order. Adaptive modulation can also be applied to the mmWave systems, as mmWave channels are also time-varying channels, as shown in Fig. 6 and Fig. 7. To obtain the results in these figures, the cluster channel model of (7) operating at 28 GHz carrier frequency is considered, where we assume $N_{\mathrm{cl}} = 8$ clusters and each cluster has $N_{\mathrm{ray}} = 10$ rays, while the azimuth and elevation angles of arrival and departure are assumed to obey the Laplacian distribution with the uniformly distributed angles over $[0, 2\pi)$, and the angular spreads of 7.5 degrees within each cluster [4], [6]–[8]. Furthermore, we assume the ULA at both transmitter and receiver. Specifically, as shown in Fig. 6, the channel gains of subcarriers change over time, which may have a dynamic range up to 30 dB. Furthermore, as shown in Fig. 7, the mean value of a channel gain may also have a dynamic range of more than 8 dB. Therefore, adaptive transmission, including adaptive modulation, is required to attain a near-capacity throughput.

In our OFDM-CSIM system, the parameters that may be adapted include the modulation order Q, the number of active indices F_{v} per virtual sub-block, the length N_{v} of virtual sub-blocks, and the number of subcarriers M_{f} of a sub-block. As introduced in Section III, the transmitted information is not only carried by the APM symbols, but also by the virtual indices, which together are mapped to the symbols transmitted by a sub-block of subcarriers. Therefore, the system throughput can be increased by adjusting the number of active indices F_{v} in the virtual domain. For example, assume $Q = 2$, $N_{\mathrm{v}} = 16$ and $M_{\mathrm{f}} = 8$. When $F_{\mathrm{v}} = 1$, the system can achieve a data rate of 0.625 bit/s/Hz, while with $F_{\mathrm{v}} = 2$, the throughput increases to 1 bit/s/Hz. Due to the strict requirement of sparsity level introduced in Section III-A, in some cases, the length M_{f} of OFDM sub-blocks and the number of active indices F_{v} per virtual sub-block should be adjusted according to the communication environments.

In the following two subsections, we present the conventional threshold-based adaptive modulation, followed by the learning-aided adaptive modulation. In our adaptive modulation schemes, the post-processing SNR is used as the metric for selecting the parameters Q, F_{v}, N_{v}, and M_{f} in the OFDM-CSIM system. Based on [33], the post-processing SNR can be expressed as:

$$\xi = \mathbb{E}\left[\frac{\mathrm{Tr}\left(\left(\boldsymbol{W}^{\mathrm{H}}(\beta)\boldsymbol{H}(\beta)\boldsymbol{F}(\beta)\right)^{\mathrm{H}}\boldsymbol{W}^{\mathrm{H}}(\beta)\boldsymbol{H}(\beta)\boldsymbol{F}(\beta)\right)}{\mathrm{Tr}\left(\boldsymbol{W}(\beta)\boldsymbol{W}^{\mathrm{H}}(\beta)\sigma^2(\beta)\right)}\right], \tag{23}$$

where $\boldsymbol{W}(\beta) = \boldsymbol{W}_{\mathrm{RF}}\boldsymbol{W}_{\mathrm{D}}(\beta)$ and $\boldsymbol{F}(\beta) = \boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{D}}(\beta)$.

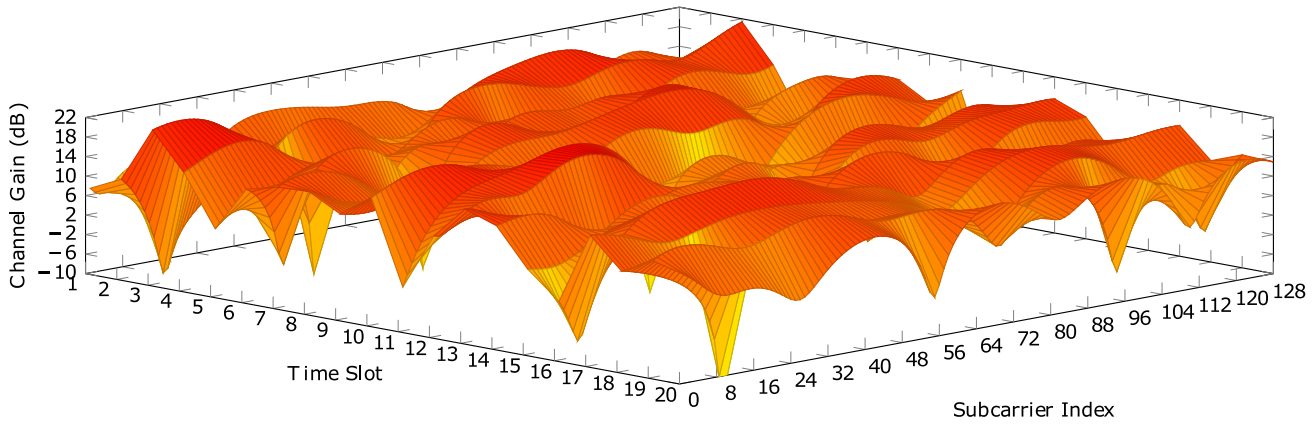


FIGURE 6. An illustration of the channel gains over a time-frequency plane.

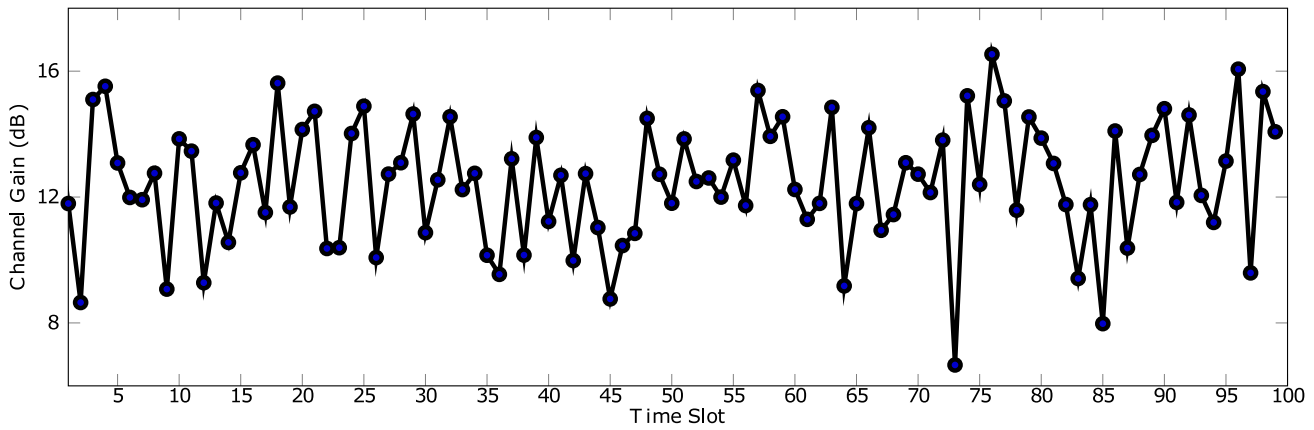
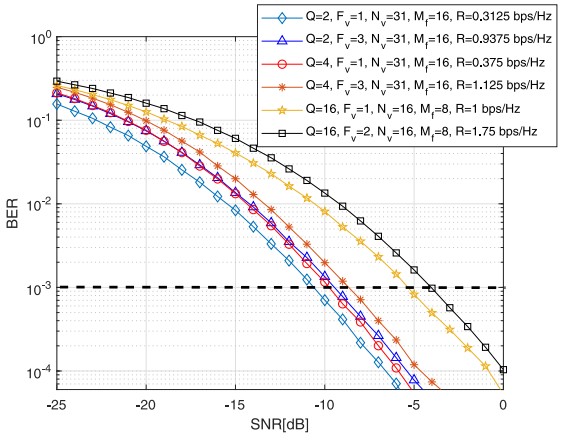


FIGURE 7. Time-varying channel magnitude of a subcarrier.

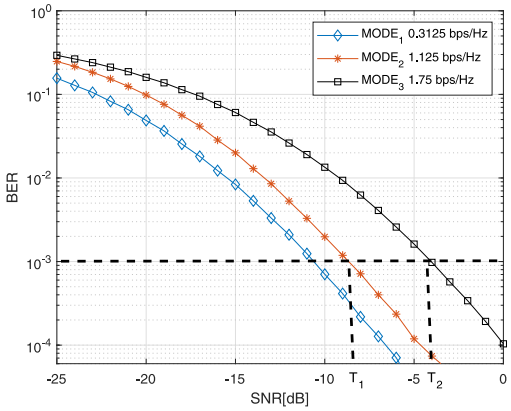
A. CONVENTIONAL ADAPTIVE MODULATION

In conventional adaptive modulation, the parameters Q, F_{v}, N_{v}, and M_{f} can be selected according to the post-processing SNR by referring to pre-defined thresholds, in order to maximize the throughput, while maintaining a target BER. Since the number of combinations of Q, F_{v}, N_{v}, and M_{f} are limited, different combinations can be defined as MODEs, expressed as MODE_i, for $i = 1, \ldots, I$, and each MODE corresponds to a given data rate. The number of selected modes I can take any value less than the total number of combinations dependent on the available parameters to adapt, which would result in variable throughputs. For example, Fig. 8(a) shows the BER performance of six different transmission modes, when the OFDM-CSIM system employs a UCYA transmit antenna array with 64 elements, and a ULA receive antenna array with 16 elements to support four RF chains at both the transmitter and receiver. As shown in Fig. 8(a), varying the system parameters Q, F_{v}, N_{v} and M_{f} results in variable throughputs and different BER performance. Then, given the total number of modes, we can decide the adaptive modes for the system via selecting them from the available ones as follows: a) More uniformly distributed modes in terms of SNR are desired; b) when two or more modes have similar BER performance, the mode yielding the highest throughput is desired, so that the highest possible throughput at a given SNR can be achieved. Based on these consideration, in Fig. 8(b) we show an example where 3 out of 6 modes are selected. Observe from Fig. 8(a) that of the other 3 modes that are not selected, 2 modes marked by blue triangles and red circles yield the throughputs of 0.9375 bps/Hz and 0.375 bps/Hz, respectively and have similar BER performance to one of the selected mode marked by orange stars, which gives the throughput of 1.125 bps/Hz. In this case, the mode generating the highest throughput of 1.125 bps/Hz is selected. Finally, 3 of the transmission modes may be selected, and the system parameters as well as the corresponding data rates are shown in Table 2.

In Fig. 8(b), assuming that the target BER is 10^{-3}, the specific SNR value T_1 and T_2 can be selected as the thresholds for the system to be operated in MODE_1, MODE_2 or MODE_3. After the thresholds are obtained, the procedure of the conventional adaptive modulation can be described as Fig. 9. Therefore, in order to implement the adaptive modulation, the receiver needs to decide the transmission mode by comparing the instantaneous post-processed SNR against the threshold values. A higher data rate can be achieved if more optional thresholds can be set, and the post-processing SNR can be more accurately measured. However, in the



(a) BER performance of OFDM-CSIM systems with different parameters.



(b) BER performance of OFDM-CSIM systems operated at MODE_1, MODE_2 and MODE_3 as shown in Table II. T_1 and T_2 are marked as the first and second thresholds based on the intersection of the target BER of 10^{-3} and BER performance curves of MODE_2 and MODE_3, respectively.

FIGURE 8. BER performance of OFDM-CSIM systems.

TABLE 2. System parameters used in simulations.

	Q	F_v	N_v	M_f	R (bps/Hz)
MODE_1	2	1	31	16	0.3125
MODE_2	4	3	31	16	1.125
MODE_3	16	2	16	8	1.75

OFDM-CSIM systems, these are hard to achieve, due to the many factors involved. Hence, in the next subsection, we propose the learning-based adaptive modulation for the OFDM-CSIM system to maximize its throughput.

B. SUPERVISED LEARNING AIDED ADAPTIVE MODULATION

The operation of adaptive modulation can be understood as solving a classification problem, which can hence be solved using ML techniques. Specifically in our system, the adaptive modulation is the process of selecting a transmission mode

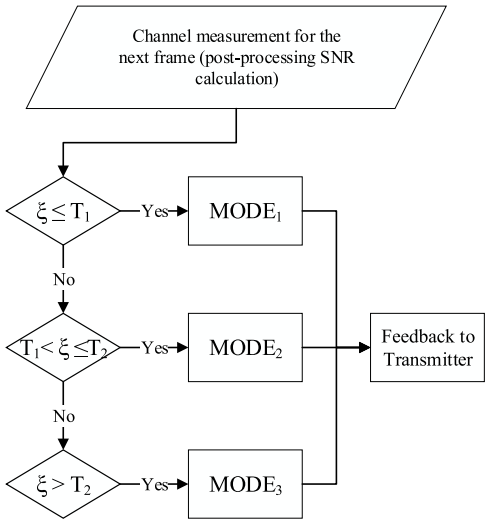


FIGURE 9. Flow chart to explain the conventional adaptive modulation.

for a given realization of the post-processing SNR, which reflects the channel state. Specifically, index i is referred to as a class, and each MODE_i corresponds to a given data rate. Therefore, given a channel realization, adaptive modulation is the process of selecting a class i that maps to a specific transmission mode MODE_i, so as to achieve the highest data rate under the constraint of the target BER. Normally, the choice of the classification algorithm depends on the structure of the feature set [31], [33]. In our adaptive modulation, the feature set includes the post-processing SNR constituting a one-dimensional link-quality matrix. The transmission mode MODE_i, which includes the combination of different orders of modulation, length of sub-block, number of active indices per virtual sub-block, and the size of virtual domain vector, has limited mapping relationships with the feature set. Therefore, we consider the k-NN supervised learning algorithm, which is capable of achieving the classification without requiring the information about the functional mapping between classifier and feature sets, and also has good performance for noisy training data [35], [50]–[52]. Another reason to employ the k-NN algorithm is that the data collected from some practical scenarios may not obey a specific distribution, while the k-NN algorithm does not require any knowledge or assumptions about the data distribution [35].

The k-NN is trained with the realizations of the feature set and its corresponding classes, which can be considered as a training set. Specifically, the training set for MODE_i can be expressed as a matrix with indices from 1 to J_i as

$$\boldsymbol{T}^{(i)} = \left[\xi_1^{(i)}, \ldots, \xi_j^{(i)}, \ldots, \xi_{J_i}^{(i)}\right]^\text{T}, \quad i = 1, 2, \ldots, I, \quad (24)$$

where J_i is the number of realizations of class i and the total number of realizations of the training set is expressed as $J = \sum_{i=1}^{I} J_i$. In (24), $\xi_j^{(i)}$ is a feature set in the training set $\boldsymbol{T}^{(\text{i})}$, which belongs to the transmission MODE_i.

To elaborate further, Fig. 10 is an example showing the training sets of MODE_1, MODE_2 and MODE_3 for 100

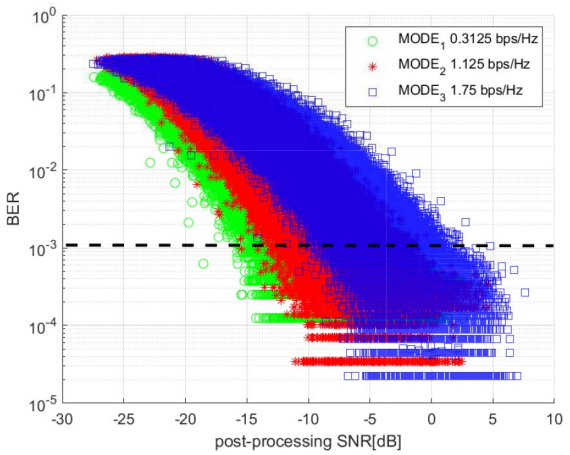


FIGURE 10. The training sets for MODE_1, MODE_2 and MODE_3.

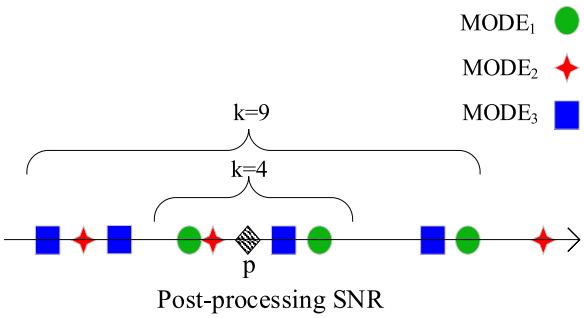


FIGURE 11. Illustration for explaining the principle of k-NN Algorithm, where $\mathcal{P}$ represents a testing data.

channel realizations, with each of the channel realizations used for transmitting 100 symbols at each of the 26 different noise levels (SNR $\in$ [−25dB 0dB]). Hence, there are in total $J_i = 260\ 000$ training symbols available for each transmission mode. In Fig. 10, a point is determined by its post-processing SNR computed from (23) and the corresponding BER.

Note that during training, we only record the post-processing SNR as a feature set in the training set, when the target BER below 10^{-3} is achieved. This results in that the actual size of the training set used in this example is $J_1 \approx 65\ 000$, $J_2 \approx 46\ 000$ and $J_3 \approx 54\ 000$. As shown in Fig. 10, due to the ambiguous boundaries of the modes resulted from the varying communication environment, it is difficult to accurately determine a transmission mode by using fixed threshold values as in the conventional adaptive modulation.

In the context of the ML-based adaptive modulation supported by the k-NN based classification, the principle can be explained with the aid of Fig. 11, where the location of a new observation point $\mathcal{P}$ is given by the value of the post-processing SNR. In Fig. 11, the elements of the training data corresponding to the three classes are distinguished by different colors and shapes. As shown in Fig. 11, given a $\mathcal{P}$, the classifier searches over $\boldsymbol{T}^{(\mathrm{i})}$ for $i = 1, 2, 3$ to find the k nearest points in the training sets, based on the distance metric $d(\cdot)$. When the Euclidian distance is employed, we have

$$d\left(\xi_j^{(i)}, \xi_{\mathcal{P}}\right) = \left\|\xi_j^{(i)} - \xi_{\mathcal{P}}\right\|_{\mathrm{F}}^2, \tag{25}$$

where $\xi_{\mathcal{P}}$ is the post-processing SNR of the new observation data. Based on (25) and for a given value of k, let the number of neighbors associated with MODE_1, MODE_2 and MODE_3 be expressed as k_1, k_2 and k_3. Then, the transmission mode is selected as the one providing the highest number of neighbors to the observation data, i.e., the one corresponding to the largest value in $\{k_1, k_2, k_3\}$. Specifically in this example, as shown in Fig. 11, if we set $k = 4$, we can see that there are two of the nearest neighbors labelled as MODE_1, while only one nearest neighbor is labelled as MODE_2 and MODE_3, respectively. Hence in this case, MODE_1 is selected. By contrast, when $k = 9$, MODE_3 provides the new observation data with the most number of neighbors, which is 4. Correspondingly, MODE_3 is selected as the transmission mode. Note that, there is possibility that two or more MODEs provide the same maximum number of neighbors. In this case, the mode yielding the highest throughput is selected.

There are two phases in the proposed learning aided adaptive modulation, namely the training phase and testing phase. The training data and testing data are generated independently so as to avoid any correlation between them [31], [33]. During the training phase, the post-processing SNR of each realization is calculated and stored as the training set for different MODE_i, for example $\boldsymbol{T}^{(\mathrm{i})}$. During the testing phase, the post-processing SNR of the new observation data, for example $\xi_{\mathcal{P}}$, is generated. Then, the k nearest neighbors are identified based on the minimum distance between the new observation data $\xi_{\mathcal{P}}$ and each of the training data in $\boldsymbol{T}^{(\mathrm{i})}$. Finally, the specific transmission mode is determined as the one giving the majority nearest neighbors.

The complexity of the k-NN algorithm can be analyzed from two perspectives, which are the memory required for saving the training data and the computational requirement for searching the nearest neighbors. The requirement of a large memory space is a major disadvantage of the k-NN algorithm, especially when single-chip devices are considered. However, the memory space required can be reduced with the aid of the pre-processing of training data. For example in this article, only the elements achieving the target BER are considered, which is usually substantially smaller than the number of elements in the training set. Moreover, the high numerical precision (floating point) is not required for storing each dimension in the feature set [31]. Additionally, the training data can be updated during run-time by exploiting the new observation data, in order to improve the performance over time-varying environments. On the other hand, a brute-force search for the nearest neighbors would demand a search complexity of $O(k\hat{d}J)$, where $\hat{d}$ represents the dimensions of the feature set. In this article, however, only two dimensions

TABLE 3. Parameters used in simulations.

Parameters	Value
Carrier frequency (f_c)	28 GHz
Number of subcarriers (N)	1024
Number of transmit antenna (N_t)	64
Number of receive antenna (N_r)	16
Number of RF chains in transmitter (N^t_{RF})	4
Number of RF chains in receiver (N^r_{RF})	4
k in k-NN	20
Number of data streams (N_s)	2
Target BER	10^{-3}

of feature set are considered, which are the post-processing SNR and data rate. Furthermore, there are many references that addressed the search complexity reduction for the k-NN algorithm [52], [53]. For example, in [52], the training set is restructured to form different clusters, where the distance calculation only considers the clusters of data relatively close to the new observation data, while ignoring the clusters far away from it. This method is capable of reducing the search complexity to $O(k\hat{d}\sqrt{J}\log J\hat{d} + kJ)$, which can be significantly lower than that of the brute-force search.

Note that in the conventional adaptive modulation, transmission modes are selected by referring to the pre-defined thresholds, with the objective to achieve the maximum data rate. Although the conventional adaptive modulation does not consider the reserved memory or search complexity, the achievable data rate is significantly reduced, when compared with the ML-based approaches.

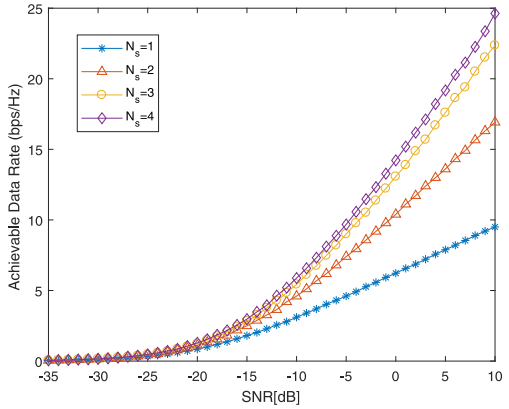


FIGURE 12. Data rate achieved by a 64 × 16 MIMO system with the ULA at both transmitter and receiver, when both transmitter and receiver employ 4 RF chains.

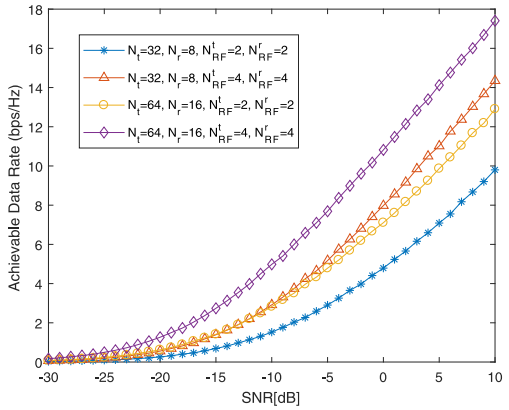


FIGURE 13. Achievable data rate by various number of antenna elements and RF chains at both transmitter and receiver in mmWave system with ULA and two data streams.

V. SIMULATION RESULTS

In this section, the performance results of the OFDM-CSIM systems with the proposed learning-assisted adaptive modulation and conventional adaptive modulation are demonstrated. For all simulations, the propagation environment is assumed to have $N_{\text{cl}} = 8$ clusters and each cluster has $N_{\text{ray}} = 10$ rays with equal power. The elevation angle and azimuth angle for both the arrival and departure rays are assumed to obey the Laplace distribution. The angle spreads in both elevation and azimuth directions and at both transmitter and receiver are assumed to be the same value of 7.5°. To train the learning-assisted adaptive system, $J \approx 160\,000$ channel realizations are employed in the training phase and 5 000 channel realizations are used in the testing phase. Note that, $J = J_1 + J_2 + J_3$. Finally, the distance metric $d(\cdot)$ is selected as the Euclidean distance. The other parameters have the values listed in Table 3.

Fig. 12 shows the achievable data rate of a 64×16 MIMO system with ULA at both transmitter and receiver, when 4 RF chains are used at both transmitter and receiver to support $N_s = 1, 2, 3$ or 4 data streams. Explicitly, the achievable data rate increases, as the number of data streams increases. Hence the number of data streams can be considered as one of the adjustable parameters for implementing adaptive rate transmission. Fig. 13 shows the effect of the number of transmit and receive antenna elements as well as the number of RF chains on the data rate of the mmWave systems. When given the numbers of transmit and receive antennas, the achievable data rate increases, as the numbers of RF chains increases. Similarly, for given numbers of RF chains, the achievable data rate increases as the numbers of transmit/receive antennas increases.

In Fig. 14, the achievable data rate of the proposed system as shown in Fig. 3 is depicted against the average SNR, when different array configurations are considered, which include the ULA, URPA, UCPA and UCYA employed at transmitter, while the receiver employs ULA. As shown in Fig. 14, for a given SNR, the UCYA achieves the highest date rate followed by the ULA and then the URPA, while the UCPA attains the lowest data rate. The one-dimensional ULA achieves a higher data rate than the two-dimensional URPA and UCPA arrays. However, ULA has no resolution capability in the azimuth direction and furthermore, it occupies more space than the other two- or three-dimensional arrays.

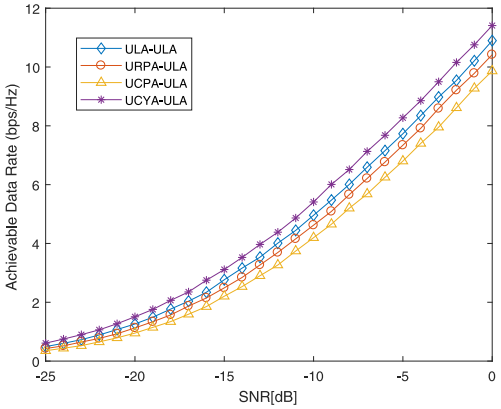
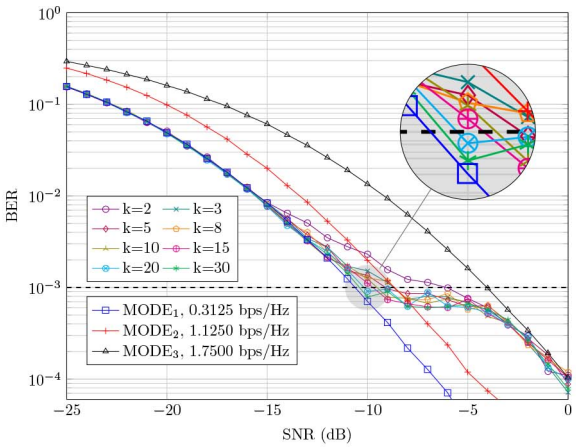
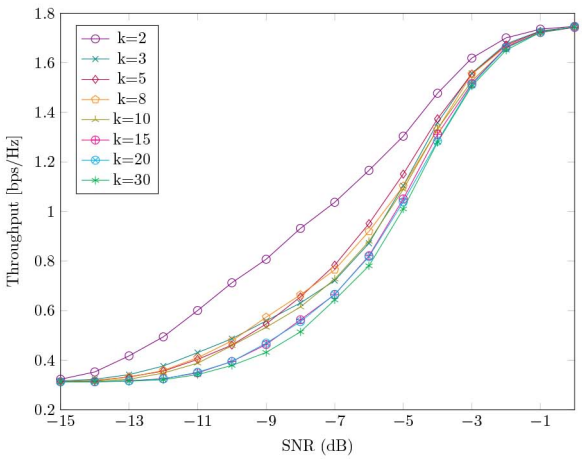


FIGURE 14. Achievable data rate versus SNR performance by adapting hybrid beamforming for ULA, URPA, UCPA and UCYA.



(a) BER versus SNR performance of the k-NN assisted adaptive modulations with different k values.



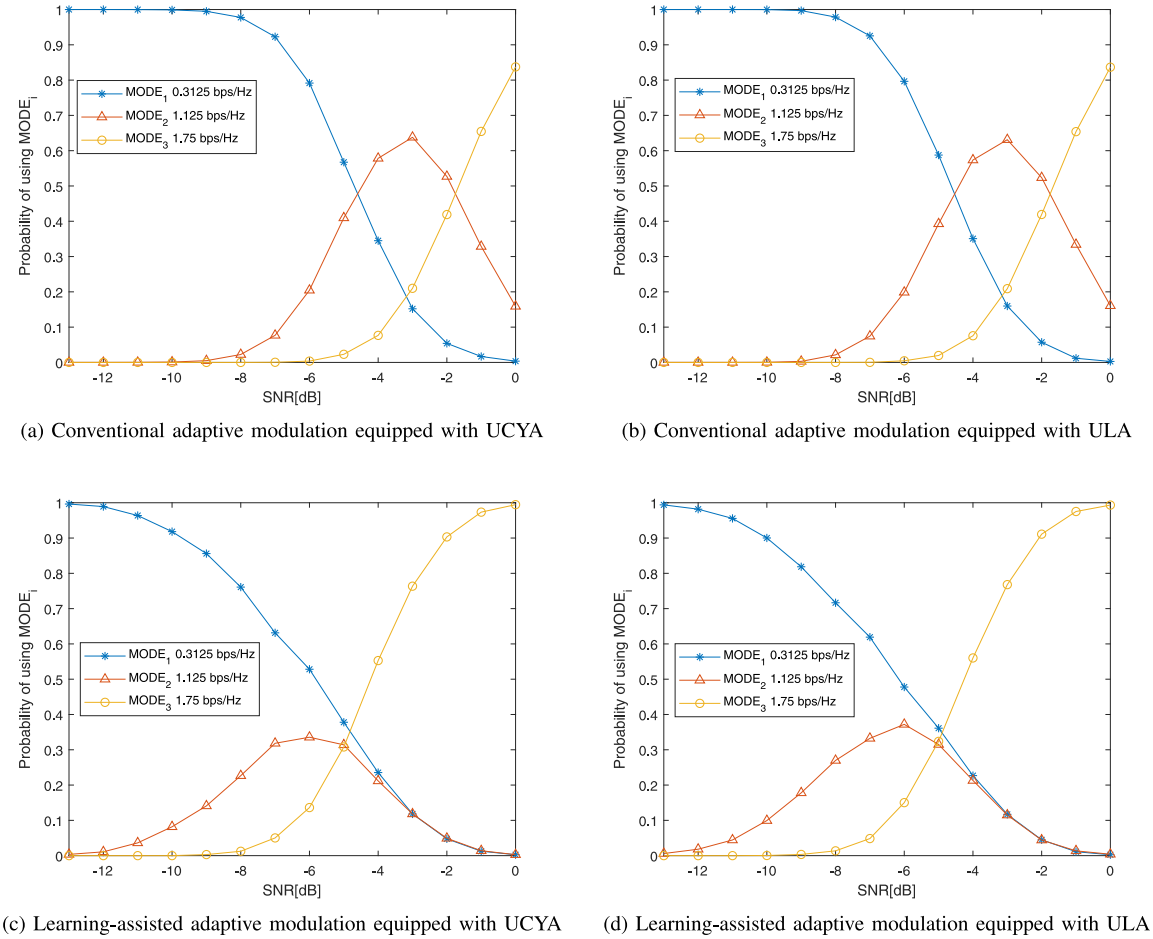
(b) Achievable throughput versus SNR performance of the k-NN assisted adaptive modulations with different k values.

FIGURE 15. Performance comparison of the *k*-NN assisted adaptive modulations with different *k* values.

Let us now demonstrate the trade-off between the throughput (Fig. 15(b)) and BER (Fig. 15(a)) performance of the adaptive modulation schemes implemented by the k-NN with different k values, in order to determine a suitable k value. As shown in Fig. 15(a), in general, at the turning point of about -10 dB to reach the target BER of 10^{-3}, the BER performance improves as k increases. By contrast, from Fig. 15(b) we observe that the throughput decreases with the increase of k, as the result of the higher reliability shown in Fig. 15(a). Since our target BER is 10^{-3} and a smaller value of k means a lower computational complexity of the k-NN, we choose $k = 20$ for the following studies. Note that we choose $k = 20$ instead of $k = 15$, because both cases provide similar throughput while the case of $k = 15$ reaches the target BER of 10^{-3} about 1 dB later than the case of $k = 20$.

The probabilities that the adaptive OFDM-CSIM system is operated in different MODEs, i.e., MODE_1, MODE_2 or MODE_3, versus SNR are shown in Fig. 16, when the conventional adaptive modulation and the proposed learning-assisted adaptive modulation equipped with different transmit antenna array structures are considered. In more detail, Fig. 16(a) and Fig. 16(b) show the mode probabilities for the conventional adaptive modulation, when a UCYA transmit antenna array with 64 elements and a ULA receive antenna array with 16 elements are respectively employed to support four RF chains at both the transmitter and receiver. According to Fig. 8(b), when the adaptive switching thresholds are set to $T_1 = -8$ dB and $T_2 = -4$ dB, as shown in Fig. 16(a), the transmitter activates MODE_1 for transmission with a probability of nearly 1, when the SNR is below $T_1 = -8$ dB. Then, when the SNR increases beyond T_1, the probability of activating MODE_2 increases, while simultaneously, the probability of using MODE_1 reduces. This process continues until reaching about the second threshold of $T_2 = -4$ dB, when the system begins to activate MODE_3. Correspondingly, we can see that the probability of employing MODE_3 grows gradually as the SNR increases. A similar trend can also be observed from Fig. 16(b), when the UCYA transmit antenna array is replaced by ULA with an identical number of elements.

Correspondingly, the mode probabilities for the learning-assisted adaptive modulation are depicted in Fig. 16(c) and Fig. 16(d), when the UCYA and ULA transmit antenna arrays are respectively employed. In comparison with Fig. 16(a) and Fig. 16(b) for the conventional adaptive modulation, explicitly, the starting points of using MODE_2 and MODE_3 occur at lower SNR values and the probability of using MODE_3 is significantly increased, for example, at around -2 dB. Specifically at SNR$= -4$ dB and considering the UCYA, as shown in Fig. 16(a) for the conventional adaptive modulation, more than half of transmissions are at MODE_2. By contrast, as shown in Fig. 16(c), over 50% of the transmissions are at MODE_3 in the learning-assisted adaptive modulation. Similar effects can also be observed when comparing Fig. 16(b) with Fig. 16(d). The reason behind the observations is that the learning-assisted adaptive modulation is capable of making use of the instantaneous post-processing SNR to select a best possible mode, while the conventional adaptive modulation



(a) Conventional adaptive modulation equipped with UCYA

(b) Conventional adaptive modulation equipped with ULA

(c) Learning-assisted adaptive modulation equipped with UCYA

(d) Learning-assisted adaptive modulation equipped with ULA

FIGURE 16. Probabilities for the system to be operated with MODE_1, MODE_2 and MODE_3. The parameters of propagation medium is summarized in Table 3, and the elevation angle and azimuth angle for both the arrival and departure rays are assumed to obey the Laplace distribution with the angular spread of 7.5°.

has to use the pre-defined average SNR-based thresholds for mode selection. Consequently, the learning-assisted design is able to make more accurate decision than the conventional adaptive design, resulting in a significantly improved throughput for a given SNR, as seen in Fig. 17. Furthermore, Fig. 17 shows that employing UCYA or ULA at the transmitter yields similar throughput in both the conventional modulation and the learning-aided adaptive modulation.

Finally, Fig. 18 shows the BER performance of the learning-assisted and conventional adaptive modulation systems, where the BER performance of the three individual modulation modes is also included. Explicitly, both adaptive modulation systems are capable of attaining the target BER of 10^{-3}, when the SNR is higher than -9 dB. As seen in Fig. 18, the learning-assisted adaptive modulation can make the BER closer to 10^{-3} than the conventional adaptive modulation, and hence attain a higher throughput, as shown in Fig. 17.

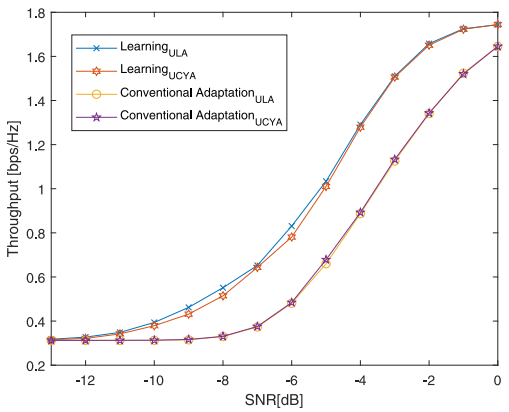


FIGURE 17. Throughput comparison of the conventional and learning-assisted adaptive modulation.

VI. CONCLUSION

In this article, we first provided a survey for the various antenna structures, and compared their SE performance in the context of mmWave communications, showing that the UCYA as a three-dimensional array has some advantages over the ULA, URPA and UCPA. Then, the principles of the OFDM-CSIM system with the hybrid beamforming operated at both transmitter and receiver is addressed. Our studies show that in OFDM-CSIM systems, adaptive rate transmission can be implemented via changing the values

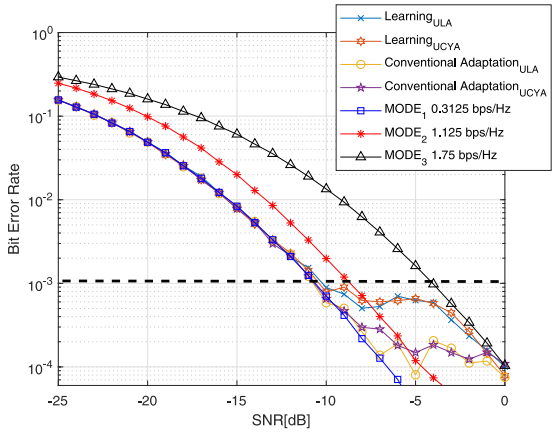


FIGURE 18. BER performance of the conventional and learning-assisted adaptive modulations, as well as of the individual MODEs.

of the parameters involved. Based on these observations, the link adaptation schemes for the OFDM-CSIM mmWave systems is proposed. Both the conventional link adaptation approach relying on thresholds and the ML-assisted link adaptation operated on k-NN algorithm were studied and compared. It is demonstrated that the conventional link adaptation has the challenge to set the near-optimum thresholds, hence resulting in the loss of achievable performance. By contrast, the ML-assisted approach is capable of efficiently exploiting the time-varying nature of wireless channels, and choosing the best possible transmit mode when considering different antenna structures. Consequently, the ML-assisted link adaptation can significantly outperform the conventional counterpart in terms of the throughput achieved.

One extension of this work is to consider the impact of the propagation environment on the ML-assisted link adaptive system, such as the generalization capability of the trained classifier when different channel models are considered. Another extension of the work is to use other classification algorithms, which are able to attain a similar performance as the k-NN, but come with lower computational cost and are less demanding on memory storage [35], [54].

ACKNOWLEDGMENT

The research data supporting the paper can be obtained from the University of Southampton institutional repository: http://dx.doi.org/10.5258/SOTON/D151910.5258/SOTON/D1519.

REFERENCES

[1] X. Wang *et al.*, "Millimeter wave communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1616–1653, 3rd Quart., 2018.

[2] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[3] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

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

[5] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2013, pp. 1–5.

[6] F. Sohrabi and W. Yu, "Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1432–1443, Jul. 2017.

[7] X. Yu, J. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[8] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.

[9] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Hybrid beamforming design for full-duplex millimeter wave communication," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1394–1404, Feb. 2019.

[10] K. Satyanarayana, M. El-Hajjar, P. Kuo, A. Mourad, and L. Hanzo, "Dual-function hybrid beamforming and transmit diversity aided millimeter wave architecture," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2798–2803, Mar. 2018.

[11] J. Lee and Y. H. Lee, "AF relaying for millimeter wave communication systems with hybrid RF/baseband MIMO processing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 5838–5842.

[12] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, 2nd Quart., 2018.

[13] W. Tan, S. Jin, J. Wang, and Y. Huang, "Achievable sum-rate analysis for massive MIMO systems with different array configurations," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, New Orleans, LA, USA, Mar. 2015, pp. 316–321.

[14] W. Tan, X. Li, D. Xie, W. Tan, L. Fan, and S. Jin, "On the performance of three-dimensional antenna arrays in millimetre wave propagation environments," *IET Commun.*, vol. 12, no. 14, pp. 1743–1750, Aug. 2018.

[15] V. Kallnichev, "Analysis of beam-steering and directive characteristics of adaptive antenna arrays for mobile communications," *IEEE Antennas Propag. Mag.*, vol. 43, no. 3, pp. 145–152, Jun. 2001.

[16] M. G. M. Hussain, "Theory and analysis of adaptive cylindrical array antenna for ultrawideband wireless communications," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 3075–3083, Nov. 2005.

[17] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmWave) for 5G: Opportunities and challenges," *Wireless Netw.*, vol. 21, no. 8, pp. 2657–2676, Apr. 2015.

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

[19] R. Abu-Alhiga and H. Haas, "Subcarrier-index modulation OFDM," in *Proc. IEEE 20th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Tokyo, Japan, Sep. 2009, pp. 177–181.

[20] D. Tsonev, S. Sinanovic, and H. Haas, "Enhanced subcarrier index modulation (SIM) OFDM," in *Proc. IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, TX, USA, Dec. 2011, pp. 728–732.

[21] E. Başar, Ü. Aygölü, E. Panayırcı, and H. V. Poor, "Orthogonal frequency division multiplexing with index modulation," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5536–5549, Nov. 2013.

[22] R. Y. Mesleh, H. Haas, S. Sinanovic, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228–2241, Jul. 2008.

[23] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proc. IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014.

[24] I. Al-Nahhal, E. Basar, O. A. Dobre, and S. Ikki, "Optimum low-complexity decoder for spatial modulation," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2001–2013, Jul. 2019.

[25] I. Al-Nahhal, O. A. Dobre, and S. Ikki, "Low complexity decoders for spatial and quadrature spatial modulations—Invited paper," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Porto, Portugal, Jun. 2018, pp. 1–5.

[26] H. Zhang, L. Yang, and L. Hanzo, "Compressed sensing improves the performance of subcarrier index-modulation-assisted OFDM," *IEEE Access*, vol. 4, pp. 7859–7873, 2016.
[27] I. A. Hemadeh, S. Lu, M. El-Hajjar, and L. Hanzo, "Compressed sensing-aided index modulation improves space-time shift keying assisted millimeter-wave communications," *IEEE Access*, vol. 6, pp. 64742–64756, 2018.
[28] S. Lu, M. El-Hajjar, and L. Hanzo, "Two-dimensional index modulation for the large-scale multi-user MIMO uplink," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7904–7918, Aug. 2019.
[29] S. Catreux, V. Erceg, D. Gesbert, and R. W. Heath, "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Commun. Mag.*, vol. 40, no. 6, pp. 108–115, Jun. 2002.
[30] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications," *Proc. IEEE*, vol. 88, no. 5, pp. 609–640, May 2000.
[31] R. C. Daniels, C. M. Caramanis, and R. W. Heath, "Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and SNR ordering," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 114–126, Jan. 2010.
[32] S. S. Das, E. D. Carvalho, and R. Prasad, "Performance analysis of OFDM systems with adaptive sub carrier bandwidth," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1117–1122, Apr. 2008.
[33] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Multi-user hybrid beamforming relying on learning-aided link-adaptation for mmWave systems," *IEEE Access*, vol. 7, pp. 23197–23209, 2019.
[34] L. Hanzo, C. H. Wong, and M.-S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalised and Space-Time Coded TDMA, CDMA and OFDM Systems*. Hoboken, NJ, USA: Wiley, 2002.
[35] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2012.
[36] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY, USA: Springer, 2009.
[37] R. Daniels and R. W. Heath, "Online adaptive modulation and coding with support vector machines," in *Proc. Eur. Wireless Conf. (EW)*, Lucca, Italy, Apr. 2010, pp. 718–724.
[38] L. Hanzo, L.-L. Yang, E. Kuan, and K. Yen, *Adaptive DSCDMA Networking Using Variable Spreading Factors and Beamforming*, IEEE, Piscataway, NJ, USA, 2004, pp. 929–978.
[39] L. Hanzo, J. Blogh, and S. Ni, *3G, HSPA and FDD Versus TDD Networking: Smart Antennas and Adaptive Modulation*, 2nd ed. New York, NY, USA: Wiley, 2008.
[40] E. Torkildson, C. Sheldon, U. Madhow, and M. Rodwell, "Millimeter-wave spatial multiplexing in an indoor environment," in *Proc. IEEE GLOBECOM Workshops*, Honolulu, HI, USA, Nov. 2009, pp. 1–6.
[41] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY, USA: Wiley, 2004.
[42] P. Ioannides and C. A. Balanis, "Uniform circular and rectangular arrays for adaptive beamforming applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 4, pp. 351–354, 2005.
[43] W. Tan, S. D. Assimonis, M. Matthaiou, Y. Han, X. Li, and S. Jin, "Analysis of different planar antenna arrays for mmWave massive MIMO systems," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Sydney, NSW, Australia, Jun. 2017, pp. 1–5.
[44] E. Yaacoub, M. Al Husseini, A. Chehab, A. E. Hajj, and K. Kabalan, "Hybrid linear and circular antenna arrays," *Iran. J. Elect. Comput. Eng.*, vol. 6, no. 1, pp. 48–54, 2007.
[45] E. Yaacoub, K. Y. Kabalan, A. El-Hajj, and A. Chehab, "Cylindrical antenna arrays for WCDMA downlink capacity enhancement," in *Proc. IEEE Intl. Conf. Commun.*, vol. 11. Istanbul, Turkey, Jun. 2006, pp. 4912–4917.
[46] C. Kumar, B. P. Kumar, V. S. Kumar, and V. V. Srinivasan, "Dual circularly polarized spherical phased-array antenna for spacecraft application," *IEEE Trans. Antennas Propag.*, vol. 61, no. 2, pp. 598–605, Feb. 2013.
[47] A. K. Chan, A. Ishimaru, and R. A. Sigelmann, "Equally spaced spherical arrays," *Radio Sci.*, vol. 3, no. 5, pp. 401–404, May 1968.
[48] T. N. Kaifas and J. N. Sahalos, "Design and performance aspects of an adaptive cylindrical beamforming array," *IEEE Antennas Propag. Mag.*, vol. 54, no. 1, pp. 51–65, Feb. 2012.
[49] S. Lu, I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Compressed-sensing-aided space-time frequency index modulation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 6259–6271, Jul. 2018.
[50] Y. Zhang and X. Wu, "Integrating induction and deduction for noisy data mining," *Inf. Sci.*, vol. 180, no. 14, pp. 2663–2673, 2010.
[51] H. Liu and S. Zhang, "Noisy data elimination using mutual k-nearest neighbor for classification mining," *J. Syst. Softw.*, vol. 85, no. 5, pp. 1067–1074, 2012.
[52] X. Wang, "A fast exact k-nearest neighbors algorithm for high dimensional search using k-means clustering and triangle inequality," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2011, pp. 1293–1299.
[53] N. Suguna and K. Thanushkodi, "An improved k-nearest neighbor classification using genetic algorithm," *Int. J. Comput. Sci. Issues*, vol. 7, no. 2, pp. 18–21, 2010.
[54] P. Yang, Y. Xiao, M. Xiao, Y. L. Guan, S. Li, and W. Xiang, "Adaptive spatial modulation MIMO based on machine learning," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 9, pp. 2117–2131, Sep. 2019.