Computers, Materials & Continua
DOI:10.32604/cmc.2022.022034
Article

Tech Science Press

An Efficient Machine Learning Based Precoding Algorithm for Millimeter-Wave Massive MIMO

Waleed Shahjehan[1], Abid Ullah[1], Syed Waqar Shah[1], Ayman A. Aly[2], Bassem F. Felemban[2] and Wonjong Noh[3,*]

[1]Department of Electrical Engineering, University of Engineering and Technology Peshawar, Pakistan
[2]Department of Mechanical Engineering, College of Engineering, Taif University, Taif, 21944, Saudi Arabia
[3]School of Software, Hallym University, Chuncheon, 24252, Korea
*Corresponding Author: Wonjong Noh. Email: wonjong.noh@hallym.ac.kr
Received: 25 July 2021; Accepted: 16 November 2021

Abstract: Millimeter wave communication works in the 30–300 GHz frequency range, and can obtain a very high bandwidth, which greatly improves the transmission rate of the communication system and becomes one of the key technologies of fifth-generation (5G). The smaller wavelength of the millimeter wave makes it possible to assemble a large number of antennas in a small aperture. The resulting array gain can compensate for the path loss of the millimeter wave. Utilizing this feature, the millimeter wave massive multiple-input multiple-output (MIMO) system uses a large antenna array at the base station. It enables the transmission of multiple data streams, making the system have a higher data transmission rate. In the millimeter wave massive MIMO system, the precoding technology uses the state information of the channel to adjust the transmission strategy at the transmitting end, and the receiving end performs equalization, so that users can better obtain the antenna multiplexing gain and improve the system capacity. This paper proposes an efficient algorithm based on machine learning (ML) for effective system performance in mmwave massive MIMO systems. The main idea is to optimize the adaptive connection structure to maximize the received signal power of each user and correlate the RF chain and base station antenna. Simulation results show that, the proposed algorithm effectively improved the system performance in terms of spectral efficiency and complexity as compared with existing algorithms.

Keywords: MIMO; phased array; precoding scheme; machine learning optimization

1 Introduction

In the past ten years, the rapid development of various business systems such as the Internet of Things (IoT) and the Internet of Vehicles (IoV), as well as the advancement of wireless equipment

manufacturing processes, have promoted the development and deployment of 5G mobile communication systems with high-speed, large connections and low latency. In general, the improvement of spectrum efficiency is achieved through network densification of micro-cell millimeter wave and massive MIMO technology [1]. The current low-frequency spectrum resources can no longer meet the people,s needs for high-rate data transmission in people,s lives. The development of wireless communication technology has forced researchers to include the millimeter wave frequency band into the scope of research. The millimeter wave frequency is distributed in the range of 30 to 300 GHz, and its ample bandwidth has become a hot spot in the academic and industrial circles. However, millimeter waves are severely absorbed by the atmosphere and rain in the process of space transmission, and the limited number of propagation paths, resulting in their short effective propagation distance, which is very suitable for micro-cell communication with small coverage and high data transmission rate. It provides sufficient array gain to form a needle beam to reduce interference, which is regarded as one of the key technologies of 5G [2]. When a large number of antennas are deployed at the millimeter wave transmitting end, all-digital precoding needs to be equipped with a dedicated radio frequency chain for each transmitting antenna, and a radio frequency chain with the number of antennas (composed of digital-to-analog converters, mixers, etc.) which has high cost and unacceptable energy consumption. Therefore, the research and design of low-dimensional baseband digital precoders and high-dimensional radio frequency analog precoders instead of all-digital precoders has aroused great interest. In a single-user communication system, there are ways to improve the spectrum efficiency by minimizing the Euclidean distance between hyrbid analog and digital precoding and fully-digital precoding [3-6], and also through a joint transmitter analog precoder and receiver analog research on the maximized spectral efficiency of the combiner [7–9]. In multi-user communication systems, the hybrid precoding with a fully connected structure has also been studied [10–14]. For example, reference [15] directly uses the phase of the channel conjugate transpose as an analog precoder, and then uses the zero forcing (ZF) technique to design the baseband digital precoder. Reference [16] proposed a two-stage hybrid precoding and designed analog precoding in which the phases of all digital precoding obtained by maximum ratio transmission (MRT) and ZF precoding were extracted respectively, and then combined with channel moments. The equivalent matrix obtained by the matrix eliminates inter-user interference through ZF to obtain the baseband digital precoding. However, these hybrid precoding based on the fully connected structure require the use of more RF chains and high-precision phase devices, resulting in expensive hardware costs and power consumption proportional to accuracy, hindering the hybrid precoding structure in the base station and mobile end millimeter wave cellular network systems with strict size and power restrictions in the deployment.

In order to reduce the number and precision of the hardware used, further research on hybrid precoding of partial connection structures has been carried out [17,18]. Reference [19] first proposed the hybrid precoding of switch and inverter based on machine learning adaptive cross entropy. The authors in [20–22] applied machine learning adaptive cross entropy to the hybrid precoding of lens array switch structure, and further analyzed the impact of important parameters based on the sum rate and energy efficiency of the hybrid precoding of the switch and inverter structure. Reference [23] proposes a precoding with an adaptive connection structure, which can better achieve beam gain, but still requires a higher-precision phase shifter (at least 6-bit accuracy) to achieve close to the optimal fully-digital precoding and lower precision such as 1-bit quantized adaptive connection structure is achievable and rate performance is severely reduced. In order to solve the problem of poor accessibility a rate performance of the 1-bit quantized phase shifter of the adaptive connection structure, this paper proposes a 1-bit quantized phase shift based on machine learning adaptive cross-entropy hybrid precoding. The adaptive connection structure that obtains the matching relationship

between the RF chain and the base station antenna by maximizing the user's received signal power is more flexible than the fixed sub-connection. According to the probability distribution, the analog precoder is randomly generated, and the classic ZF precoding is used to obtain the corresponding digital precoder. The analog precoder is adaptively weighted according to the reach and rate. Then, the probability distribution of the simulated precoding is updated by reducing the cross entropy and adding a constant smoothing parameter, and repeating this way, a hybrid precoder with almost optimal performance and rate is finally obtained. Numerical simulations are performed to evaluate the effectiveness of the proposed scheme.

2 System Model

As shown in Fig. 1, consider the massive MIMO system of multi-user downlink , the base station deploys N antennas, N_{RF} radio frequency chains, and simultaneously serves K non-cooperative users with a single antenna. Generally, the massive MIMO system satisfies $K \leq N_{\mathrm{RF}} \leq N$, and $N_{\mathrm{RF}} = K$ is set in this article. The analog precoder $\boldsymbol{F}_{\mathrm{RF}}$ is composed of a small number of radio frequency (RF) chains, an adaptive connection network and a large number of antennas [24]. It is assumed that the symbols sent to K users $\boldsymbol{s} = [s_1, , s_2 \ldots, s_K]^{\mathrm{T}}$ are independently and identically distributed, and they are all symbols that obey the zero-mean complex Gaussian distribution, satisfying $E(\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}) = P/K\boldsymbol{I}_K$ where P is the total transmitted signal power of the base station [25].

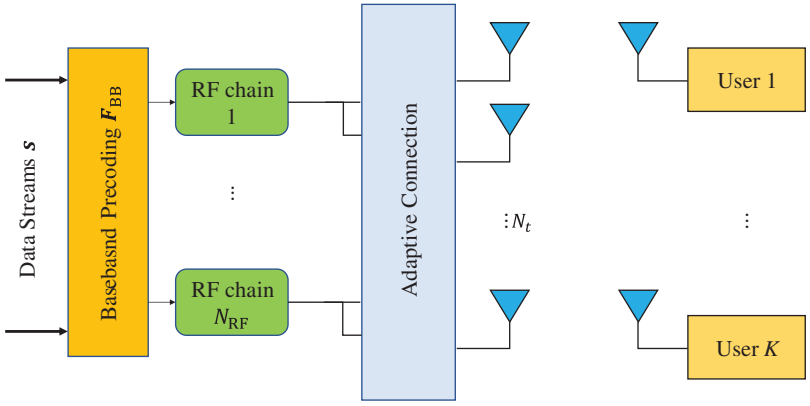

Figure 1: Proposed system model

The signal received by K users can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}^{\mathrm{T}}\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}\boldsymbol{s} + \boldsymbol{n} \tag{1}$$

Among them, $\boldsymbol{H}^{\mathrm{T}} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_{\mathrm{K}}]^{\mathrm{T}} \in \mathbb{C}^{K\times N}$ represents the downlink channel from the base station to all users, $\boldsymbol{h}_k \in \mathbb{C}^{N\times 1}$ represents the channel gain from the base station to the k th user. $\boldsymbol{F}_{\mathrm{RF}} = [\boldsymbol{f}_1^{\mathrm{RF}}, \boldsymbol{f}_2^{\mathrm{RF}}, \ldots, \boldsymbol{f}_{N_{\mathrm{RF}}}^{\mathrm{RF}}] \in \mathbb{C}^{K\times N_{\mathrm{RF}}}$ represents a high-dimensional constant modulus analog precoder that can only adjust the signal phase, $\boldsymbol{F}_{\mathrm{BB}} = [\boldsymbol{f}_1^{\mathrm{BB}}, \boldsymbol{f}_2^{\mathrm{BB}}, \ldots, \boldsymbol{f}_K^{\mathrm{BB}}] \in \mathbb{C}^{N_{\mathrm{RF}}\times K}$ stands for a low-dimensional digital precoder that can adjust the signal amplitude and phase. $\boldsymbol{n}$ represents $K \times 1$ dimension vector additive white Gaussian noise AWGN, satisfying $E(\boldsymbol{n}\boldsymbol{n}^{\mathrm{H}}) = \sigma^2\boldsymbol{I}_K$, σ^2 represents the power of additive white Gaussian noise [26]. In order to ensure the total transmit power constraint, after the analog precoding matrix $\boldsymbol{F}_{\mathrm{RF}}$ is designed, the corresponding digital precoding matrix $\boldsymbol{F}_{\mathrm{BB}}$ should meet the power constraint $\|F_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}\|_{\mathrm{F}}^2 = K$. For the channel gain vector, in order to better describe the characteristics of limited space selection and limited scattering caused by the extremely high

propagation path loss of the millimeter wave channel, and the large number of antenna arrays of the millimeter wave transmitter leads to a large degree of antenna correlation, using the geometric Saleh-Valenzuela channel model, assuming that the transmitting end is a uniform linear array (ULA) antenna, the channel gain of the k-th user can be expressed as:

$$\boldsymbol{h}_k = \sqrt{\frac{N}{L_k}} \sum\nolimits_{l=1}^{L_k} \alpha_k^{(l)} \boldsymbol{a}(\Phi_k^{(l)}) \tag{2}$$

Among them, N is the number of antennas of the base station ULA antenna. L_k represents the number of propagation paths of the k-th user, $\alpha_k^{(l)}$ and $\Phi_k^{(l)}$ respectively represent the complex gain and departure azimuth angle of the l-th path of the k-th user, and $\boldsymbol{a}(\Phi)$ represents the response vector of the transmission array with a dimension of $N \times 1$, which can be expressed for

$$\boldsymbol{a}(\Phi) = \frac{1}{\sqrt{N}} \left[1, e^{j\frac{2\pi d}{\lambda}\sin(\Phi)}, e^{j\frac{4\pi d}{\lambda}\sin(\Phi),\ldots,e^{j\frac{2\pi d\,(N-1)}{\lambda}\sin(\Phi)}} \right]^{\mathrm{T}} \tag{3}$$

Here, λ is the wavelength of the millimeter wave, and d is the distance between the ULA antenna elements, usually set to half the wavelength $d = \lambda/2$ [27–28].

3 Proposed Algorithm

The analog precoding of the traditional low-precision phase shifter partial connection structure often fails to achieve the array gain of the millimeter wave large-dimensional antenna. Therefore, this article simulates the precoder adaptive connection and deploy adaptive connection network instead of fixed sub-connection switch and reverse vectorizer (equivalent to a phase shifter with 1-bit quantization). The same as the fixed sub-connection structure, the adaptive connection only requires N number of 1-bit quantized phase shifters, N_{RF} number of radio frequency chains, and no adders. Compared with fully connected $N \times N_{\mathrm{RF}}$ high-precision phase shifters, N_{RF} radio frequency chains, and N adders, it greatly reduces the hardware complexity, cost and energy consumption. The adaptive connection network can better match the downlink to improve the user's reach and rate. In order to make better use of adaptive cross-entropy optimization to apply to the adaptive connection structure, it is necessary to match the RF chain and the antenna under specific channel state information, that is, to find the position of the non-zero element in the corresponding analog precoding matrix $\boldsymbol{F}_{\mathrm{RF}}$.

$$\boldsymbol{F}_{\mathrm{RF}} = \begin{bmatrix} f_{1,1}^{\mathrm{RF}} & f_{2,1}^{\mathrm{RF}} & \cdots & f_{K,1}^{\mathrm{RF}} \\ f_{1,2}^{\mathrm{RF}} & f_{2,2}^{\mathrm{RF}} & \cdots & f_{K,2}^{\mathrm{RF}} \\ \vdots & \vdots & \vdots & \vdots \\ f_{1,N}^{\mathrm{RF}} & f_{2,N}^{\mathrm{RF}} & \cdots & f_{K,N}^{\mathrm{RF}} \end{bmatrix}_{N\times K} \overset{\text{sum by line}}{\rightarrow} \sum_{k=1}^{K} f_{k,j}^{\mathrm{RF}}$$

$$\overrightarrow{f_x^{\mathrm{RF}}} = \begin{bmatrix} f_{x_1,1}^{\mathrm{RF}} \\ f_{x_2,2}^{\mathrm{RF}} \\ \vdots \\ f_{x_N,N}^{\mathrm{RF}} \end{bmatrix}_{N\times 1} \tag{4}$$

Due to the special structure of the adaptive connection and the normal mode constraint of the elements in the analog precoder, the corresponding analog precoding $\boldsymbol{F}_{\mathrm{RF}}$ is a sparse matrix, as shown in Eq. (4), $f_{k,j}^{\mathrm{RF}}$ represents the element in the j th row and k th column, and the k th radio frequency

chain is connected to the j th base station antenna, and its element $f_{k,j}^{\mathrm{RF}} \in \frac{1}{\sqrt{N}}\{-1,0,1\}$, $1 \le k \le K$; $1 \le j \le N$. Sum each row of $\boldsymbol{F}_{\mathrm{RF}}$ to get vector $\overrightarrow{f_x^{\mathrm{RF}}} = [f_{x_1,1}^{\mathrm{RF}}, f_{x_2,2}^{\mathrm{RF}}, \ldots, f_{x_N,N}^{\mathrm{RF}}]^{\mathrm{T}}$, and its j th element $f_{x_j,j}^{\mathrm{RF}} \in \frac{1}{\sqrt{N}}\{-1,1\}$. Therefore, the adaptive connection structure simulation precoding has two constraints:

$$\begin{cases} \sum_{j=1}^{N} |f_{k,j}^{\mathrm{RF}}| = \dfrac{M}{\sqrt{N}} \\ \sum_{k=1}^{K} |f_{k,j}^{\mathrm{RF}}| = \dfrac{1}{\sqrt{N}} \end{cases} \tag{5}$$

Indicates the connection relationship between all radio frequency chains and all antennas of the base station. Assuming that $N/K = M$ is an integer, divide the N base station antennas into K independent sets, use S_k to represent the set of k-th radio frequency chain connected to base station antennas, and $S_k(j)$ to denote the k-th radio frequency chain connected to j th base station antenna, and $|S_k| = M$. So Eq. (5) is also equivalent to

$$S = \{\{S_k\}_{k=1}^{K} | \bigcup_{k=1}^{K} S_k = \{1,2,\ldots,N\}\},\ S_k \cap S_q = \psi,\ \forall\, k \ne q \tag{6}$$

The designed analog precoding $\boldsymbol{F}_{\mathrm{RF}}$ and digital precoding $\boldsymbol{F}_{\mathrm{BB}}$ should maximize the downlink reachability and rate of the K users served

$$\begin{cases} \boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{BB}} = \arg\max_{\boldsymbol{F}_{\mathrm{RF}}, \boldsymbol{F}_{\mathrm{BB}}} \sum_{k=1}^{K} \log(1 + \mathrm{SINR}_k) \\ \text{s. t.} \quad \boldsymbol{F}_{\mathrm{RF}} \in \quad \mathcal{F}_{\mathrm{RF}} \\ \|\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{F}_{\mathrm{BB}}\|_{\mathrm{F}}^2 = K \end{cases} \tag{7}$$

where SINR_k represents the signal-to-interference and noise ratio of the k-th user, which is expressed as

$$\mathrm{SINR}_k = \frac{P/K\left|\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_k^{\mathrm{BB}}\right|^2}{P/K\sum_{k\ne i}\left|\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{F}_{\mathrm{RF}}\boldsymbol{f}_k^{\mathrm{BB}}\right|^2 + \sigma^2} \tag{8}$$

It can be seen that Eq. (7) is a non-convex optimization problem under the constraints of power constraints and adaptive connection structure. To maximize the total reachability and rate, that is, to maximize the SINR of the symbols received by each user, the k-th column $\boldsymbol{f}_k^{\mathrm{RF}}$ of the $\boldsymbol{F}_{\mathrm{RF}}$ must satisfy

$$\boldsymbol{f}_k^{\mathrm{RF}} = \arg\max_{\boldsymbol{f}_k^{\mathrm{RF}}} \left\{\left|\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{f}_k^{\mathrm{RF}}\right| : \text{costraints (5)}\right\} \tag{9}$$

There are only M non-zero elements in $\boldsymbol{f}_k^{\mathrm{RF}}$, and the first M maximum values of the channel $\boldsymbol{h}_k$ element modulus need to be matched, one element is matched each time and the antenna position is returned, and stored in the set S_k

$$S_k = \arg\max_{1\le j\le N} \left\{\left|\boldsymbol{h}_{k,j}\right| : \sum\nolimits_{k'=1}^{K} \left|\boldsymbol{f}_{k',j}^{\mathrm{RF}}\right| = 0\right\} \tag{10}$$

The constraint $\sum_{k'=1}^{K} \left|\boldsymbol{f}_{k',j}^{\mathrm{RF}}\right| = 0$ means that the j th line of the $\boldsymbol{F}_{\mathrm{RF}}$ has not been evaluated or the j th antenna has not been matched.

Performing Eq. (10) once can match the k-th radio frequency chain with a base station antenna, that is, get a non-zero element position in the $\boldsymbol{F}_{\mathrm{RF}}$. To ensure the fairness of the radio frequency chain, K radio frequency chains are matched in turn, and M turns are performed. The position of the N non-zero elements of the $\boldsymbol{F}_{\mathrm{RF}}$ that meets the constraint (5) can be obtained, or the set of matching relationships between the RF chain and the antenna that meets the constraint (6) $\{S_k\}_{k=1}^{K}$.

In the ACN-MLACE algorithm, since the phase shifter is quantized by 1 bit, the non-convex optimization problem of Eq. (7) has been transformed into a combinatorial optimization problem. How to obtain the $\boldsymbol{F}_{\mathrm{RF}}$ and $\boldsymbol{F}_{\mathrm{BB}}$ with the largest reach and rate? It is necessary to search exhaustively to find N determinations. The $2N$ combinations of non-zero element values in the position involve extremely high computational complexity. For millimeter-wave massive MIMO systems, N is often large, such as $N = 56$, and an exhaustive search is required for $2^{56} \approx 7.2 \times 10^{16}$. Take another path, deploy ML adaptive algorithm based on cross-entropy optimization to intelligently look for the best adaptive connection precoding. In each iteration, $Z \cdot N \times 1$ dimensional vectors are generated according to a probability distribution. These vectors can be set according to the matching relationship between each radio frequency chain and the base station antenna $\{S_k\}_{k=1}^{K}$ to obtain Z candidates analog precoding matrices $\boldsymbol{F}_{\mathrm{RF}}$, each candidate analog precoding matrix $\boldsymbol{F}_{\mathrm{RF}}$ has a corresponding reach and rate. Select the $Z_{\mathrm{elite}} \cdot \boldsymbol{F}_{\mathrm{RF}}$ with the best performance and rate, and update the probability distribution by minimizing cross entropy and adding smoothing parameters. In this way, it will be generated with probability 1 and close to the optimal candidate simulates the probability distribution of non-zero elements in the precoding matrix. The vector $\overrightarrow{f_x^{\mathrm{RF}}} = [f_{x_1,1}^{\mathrm{RF}}, f_{x_2,2}^{\mathrm{RF}}, \ldots, f_{x_N,N}^{\mathrm{RF}}]^{\mathrm{T}}$ composed of non-zero elements in each row of the analog precoder is reconstructed according to the found matching relationship, and $\boldsymbol{p} = [p_1, p_2, \ldots, p_N]^{\mathrm{T}}$ is used to represent the probability of the corresponding element in $\overrightarrow{f_x^{\mathrm{RF}}}$, and the j-th element $f_{x_j,j}^{\mathrm{RF}} = 1/\sqrt{N}$ is Bernoulli random variables, that is, the probability of $f_{x_j,j}^{\mathrm{RF}} = 1/\sqrt{N}$ is p_j, and the probability of $f_{x_j,j}^{\mathrm{RF}} = -1/\sqrt{N}$ is $1 - p_j$.

By initializing the probability distribution parameter $p^{(0)}$ to $p^{(0)} = 1/2 \times \boldsymbol{I}_{N\times 1}$, ($\boldsymbol{I}$ is a vector of all 1 s), according to the probability distribution $\xi(\overrightarrow{\boldsymbol{f}_x^{\mathrm{RF}}}, p^{(i)})$, generate Z candidate vectors $\{\overrightarrow{\boldsymbol{f}_x^{\mathrm{RF},z}}\}_{z=1}^{Z}$ samples, the matching relationship is reconstructed to generate Z candidate analog precoding $\{\boldsymbol{f}_{\mathrm{RF}}^{z}\}_{z=1}^{Z}$ samples, and then according to the equivalent channel $\boldsymbol{H}_{\mathrm{eq}}^{z} = \boldsymbol{H}^{\mathrm{H}}\boldsymbol{F}_{\mathrm{RF}}^{z}$, $1 \le z \le Z$, the corresponding Z digital precoding $\{\boldsymbol{F}_{\mathrm{BB}}^{z}\}_{z=1}^{Z}$. In this article, the classic digital ZF precoding is used to eliminate the inter-user interference to obtain the corresponding z-th digital precoding matrix

$$\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{temp},\,z} = (\boldsymbol{H}_{\mathrm{eq}}^{z})^{\mathrm{H}}((\boldsymbol{H}_{\mathrm{eq}}^{z})(\boldsymbol{H}_{\mathrm{eq}}^{z}))^{-1} \tag{11}$$

Put power constraints on it: $\boldsymbol{F}_{\mathrm{BB}}^{z} = \sqrt{K}/\|\boldsymbol{F}_{\mathrm{RF}}^{z}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{temp},z}\|_{\mathrm{F}}\boldsymbol{F}_{\mathrm{BB}}^{\mathrm{temp},z}$.

After that, the achievable sum rate $R(\boldsymbol{F}_{\mathrm{BB}}^{z})$ is obtained by substituting $\boldsymbol{F}_{\mathrm{RF}}^{z}$ and $\boldsymbol{F}_{\mathrm{BB}}^{z}$ into Eqs. (7) and (8), and sort Z and $\{R(\boldsymbol{f}_{\mathrm{RF}}^{z})\}_{z=1}^{Z}$ in descending order. In order to adaptively update the next probability $p^{(i+1)}$, it is necessary to obtain the first Z_{elite} and $\{R(\boldsymbol{f}_{\mathrm{RF}}^{z})\}_{z=1}^{Z_{\mathrm{elite}}}$ with the best achievable rate, and define the best analog precoding achievable rate average value is T which is expressed as

$$\begin{cases} T = \dfrac{1}{Z_{\mathrm{elite}}} \sum_{z=1}^{Z_{\mathrm{elite}}} R(\boldsymbol{F}_{\mathrm{RF}}^{z}) \\ T = \dfrac{1}{Z_{\mathrm{elite}}} \sum_{z=1}^{Z_{\mathrm{elite}}} R(\overrightarrow{\boldsymbol{f}_x^{\mathrm{RF},z}}) \end{cases} \tag{12}$$

Here, the z-th analog precoding matrix $\boldsymbol{F}_{\mathrm{RF}}^{z}$ is the first vector $\overrightarrow{\boldsymbol{f}_{x}^{\mathrm{RF},z}}$ obtained according to the matching relationship between the radio frequency chain and the base station antenna, so $R(\boldsymbol{F}_{\mathrm{RF}}^{z}) = R(\overrightarrow{\boldsymbol{f}_{x}^{\mathrm{RF},z}})$. The weight of the achievable rate corresponding to the z-th analog precoding matrix is $w_z = R(\boldsymbol{F}_{\mathrm{RF}}^{z})/T$. Update the probability of the next time adaptively according to the current probability distribution and weight

$$p^{(i+1)} = \underset{p^{(i)}}{max}\ \frac{1}{Z}\sum_{z=1}^{Z_{\mathrm{elite}}} w_z \ln \xi(\overrightarrow{\boldsymbol{f}_{x}^{\mathrm{RF},z}}, p^{(i)}) \tag{13}$$

There is also $\xi(\overrightarrow{\boldsymbol{f}_{x}^{\mathrm{RF},z}}, p^{(i)}) = \xi(\boldsymbol{F}_{\mathrm{RF}}^{z}, p^{(i)})$, where

$$\overrightarrow{\boldsymbol{f}_{x}^{\mathrm{RF},z}} = \prod_{j=1}^{N} (p_j^{(i)})^{\frac{1}{2}\left(1+\sqrt{N}\boldsymbol{f}_{x_j,j}^{\mathrm{RF},z}\right)} (1-p_j^{(i)})^{\frac{1}{2}\left(1-\sqrt{N}\boldsymbol{f}_{x_j,j}^{\mathrm{RF},z}\right)} \tag{14}$$

Substituting Eq. (14) into Eq. (13), and then Eq. (13) finds the first derivative of its j-th probability element $p_j^{(i)}$ to obtain

$$\frac{1}{Z}\sum_{z=1}^{Z_{\mathrm{elite}}} w_z \left(\frac{1+\sqrt{N}\boldsymbol{f}_{x_j,j}^{\mathrm{RF},z}}{2p_j^{(i)}} - \frac{1}{2}\frac{1-\sqrt{N}\boldsymbol{f}_{x_j,j}^{\mathrm{RF},z}}{2(1-p_j^{(i)})}\right) \tag{15}$$

Setting Eq. (15) equal to zero, the j-th element of the next probability distribution can be obtained

$$p_j^{(i+1)} = \frac{\sum_{z=1}^{Z_{\mathrm{elite}}} w_z \left(1+\sqrt{N}\boldsymbol{f}_{x_j,j}^{\mathrm{RF},z}\right)}{2\sum_{z=1}^{Z_{\mathrm{elite}}} w_z} \tag{16}$$

In order to ensure that the adaptive cross entropy optimization converges to the optimal solution to avoid local convergence, a constant smoothing parameter $\ddot{\Theta}$ can be further added between the current probability distribution and the next probability distribution.

$$p^{(i+!)} = \ddot{\theta} p^{(i+!)} + (1-\ddot{\theta}) p^{(i)} \tag{17}$$

Here $0 < \ddot{\theta} \leq 1$, until the end of the l iteration is reached, the probability distribution $p^{(l)}$ for generating the optimal analog precoding will be obtained, and the optimal analog precoding $\boldsymbol{F}_{\mathrm{RF}}^{l}$ and the optimal number in the l generation sample will be selected precoding $\boldsymbol{F}_{\mathrm{BB}}^{l}$, which is the almost optimal adaptive connection hybrid precoding under the corresponding channel state information. The specific algorithm is as follows in Algorithm 1.

Algorithm 1: Proposed algorithm

Input: Channel matrix $\boldsymbol{H}^{\mathrm{T}}$, number of iterations I, number of candidates Z, Optimal number of candidates Z_{elite}, smoothing parameter $\ddot{\Theta}$,

Initialization: BS antenna set $S = \{1, 2, \ldots, N\}$, matching relation $S_k = \psi,\ 1 \leq k \leq K,\ i = 0,\ \boldsymbol{p}^{(0)} = \frac{1}{2} \times 1_{N\times 1}$.

1: for $m = 1$ to M

2: for $k = 1$ to K

(Continued)

Algorithm 1: Continued

3: $j_{\text{opt}} = \arg\max_{j \in S} \{|h_{k,j}|\}$
4: $S_k = j_{\text{opt}}$
5: $S = S - j_{\text{opt}}$
6: end for
7: end for
8: for $i = 1$ to I
9: Randomly generate Z candidate vectors $\{\overrightarrow{\boldsymbol{f}_x^{\text{RF},z}}\}_{z=1}^{Z}$ according to $\xi(\overrightarrow{\boldsymbol{f}_x^{\text{RF},z}}, p^{(i)})$
10: According to the matching relationship set $\{S_k\}_{k=1}^{K}$ and $\{\overrightarrow{\boldsymbol{f}_x^{\text{RF},z}}\}_{z=1}^{Z}$, reconstruct $Z\{\boldsymbol{F}_{\text{RF}}^{z}\}_{z=1}^{Z}$, namely: $\boldsymbol{F}_{\text{RF}}^{z}(S_1, S_2, \ldots, S_K) = \overrightarrow{\boldsymbol{f}_x^{\text{RF},z}}$
11: Calculate digital precoding $\{\boldsymbol{F}_{\text{BB}}^{z}\}_{z=1}^{Z}$ according to Eq. (11)
12: Calculate the achievable rate $\{R(F_{\text{RF}}^{z})\}_{z=1}^{Z}$ using Eqs. (7) and (8)
13: Arrange achievable in descending order $R(F_{\text{RF}}^{1}) \geq R(F_{\text{RF}}^{2}), \ldots, \geq R(F_{\text{RF}}^{Z})$
14: Choose the first $Z_{\text{elite}} R(\boldsymbol{F}_{\text{RF}}^{1}), R(\boldsymbol{F}_{\text{RF}}^{2}), \ldots, R(\boldsymbol{F}_{\text{RF}}^{Z_{\text{elite}}})$ to get $\{\boldsymbol{F}_{\text{RF}}^{z}\}_{z=1}^{Z_{\text{elite}}}$.
15: Determine the Z_{elite} weighting coefficients w_z
16: Update $p^{(i+1)}$ using Eqs. (16) and (17)
17: End for
18: **Output:** $\boldsymbol{F}_{\text{RF}}^{1}$, $\boldsymbol{F}_{\text{BB}}^{1}$

4 Simulation Results

This section provides the simulation results and analysis. The proposed machine learning based precoding algorithm is compared with fully digital precoding, hybrid precoding of adaptive connection structure, and the conventional OMP precoding of structure. The combined precoding has the same lower hardware complexity and eliminates the NN_{RF} phase shifters and N adders required by the fully connected hybrid structure. Therefore, the sum rate and complexity are used here as a comparison of the performance of different precoding schemes. The simulation parameters are set as follows in Tab. 1.

Table 1: Simulation parameters

Parameter	Value
Distance between ULA d	0.5
Number of beam paths L_k	3
Number of transmitter antennas N_t	1024
Number of receiver antennas N_r	64
SNR	25 dB
Number of RF chains	16
Number of data streams N_8	8
Number of phase shifter N_c	40

4.1 Achievable Sum Rate Comparison with Different Number of RF Chains and Data Streams with Fixed Antennas

Fig. 2 compares the achievable sum rate of the proposed algorithm, fully digital, and other algorithms *vs.* SNR for system configuration when the number of RF chains $\boldsymbol{N}_{\boldsymbol{RF}}^{t} = \boldsymbol{N}_{\boldsymbol{RF}}^{r} = 4$ and number of data streams $\boldsymbol{N}_{s} = 4$. As can be seen from Fig. 2 that, the achievable sum rates of all algorithms increase with increasing SNR. Moreover, the proposed algorithm gives close performance with optimal fully digital scheme which indicates its effectiveness over the existing algorithms.

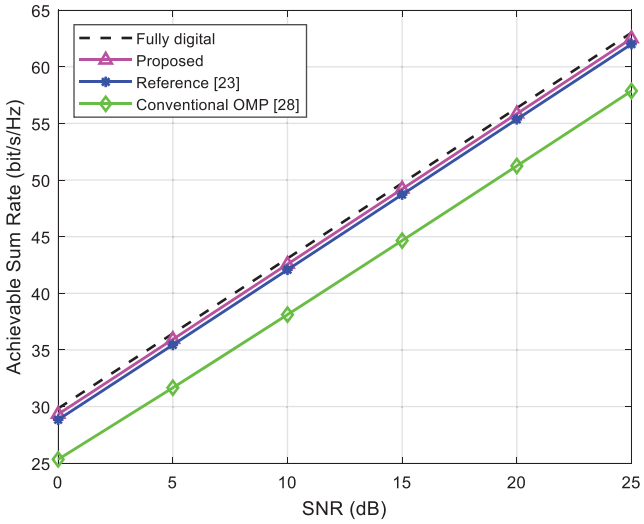

Figure 2: Comparison of achievable sum rate of algorithms *vs.* SNR when $N_{\mathrm{RF}}^{t} = N_{\mathrm{RF}}^{r} = N_{s} = 4$

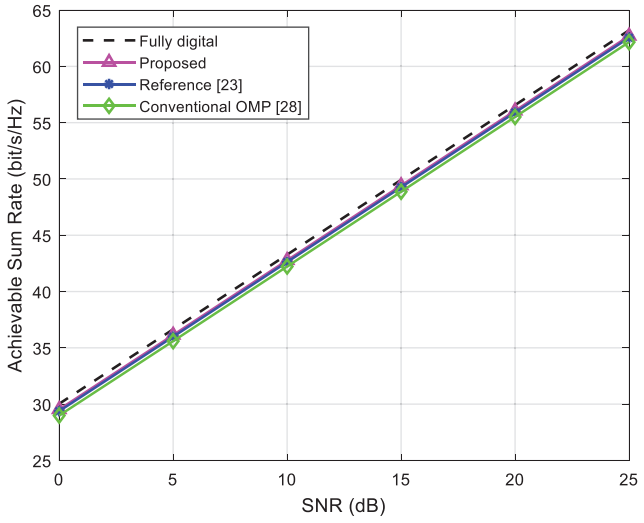

Figure 3: Comparison of achievable sum rate of algorithms *vs.* SNR when $N_{\mathrm{RF}}^{t} = N_{\mathrm{RF}}^{r} = 8$, $N_{s} = 4$

Fig. 3 illustrates the achievable sum rate of the proposed algorithm, fully digital, and other algorithms *vs.* SNR for system configuration when the number of RF chains $\boldsymbol{N}_{\mathbf{RF}}^{t} = \boldsymbol{N}_{\mathbf{RF}}^{r} = 8$ and number of data streams $\boldsymbol{N}_{s} = 4$. As can be seen from Fig. 3, the achievable sum rates of all

algorithms increases with increasing SNR. Moreover, the proposed algorithm gives close performance with optimal fully-digital scheme which indicates its effectiveness over the existing algorithms. Here, the results are closed for all algorithms because the number of RF chains are increased. But the energy consumption drastically increases in the existing algorithms in contrast, which makes them unsuitable for deployment. Also, increasing the number of RF chains increases the computational complexity and hardware structure.

4.2 Achievable Sum Rate Comparison with Different Number of Antennas with Fixed Number of RF Chains and Data Streams

Fig. 4 compares the achievable sum rate of the algorithms versus SNR when the number of transmitter antennas $N_t = 256$, the number of receiver antennas $N_r = 16$ and number of RF chains and data streams is $N_{\mathrm{RF}}^t = N_{\mathrm{RF}}^r = N_s = 4$. As can be seen from Fig. 4, the achievable rate of all algorithms increases with SNR. Moreover, the proposed algorithm gives better performance and shows close sum rate with optimal fully digital precoding. It is also clear from Fig. 4 that, due to increasing number of antennas, the sum rate is about 72 bps/Hz for SNR = 25 dB, whereas the sum rate is 63 bps/Hz for SNR = 25 dB in Figs. 3 and 4, respectively. This proves that the sum rate increases with increasing the number of antennas, which is one of the main features of massive MIMO. Fig. 5 compares the achievable sum rate of the algorithms versus SNR when the number of transmitter antennas $N_t = 1024$, the number of receiver antennas $N_r = 64$ and number of RF chains and data streams is $N_{\mathrm{RF}}^t = N_{\mathrm{RF}}^r = N_s = 4$. As can be seen from Fig. 5, the achievable rate of all algorithms increases with SNR. Moreover, the proposed algorithm gives better performance and shows close sum rate with optimal fully digital precoding. It is also clear from Fig. 5 that, due to increasing number of antennas, the sum rate is about 87 bps/Hz for SNR = 25 dB, whereas the sum rate is 63 bps/Hz for SNR = 25 dB in Figs. 3 and 4, respectively. This further proves that the sum rate increases with increasing the number of antennas, which is one of the main features of massive MIMO.

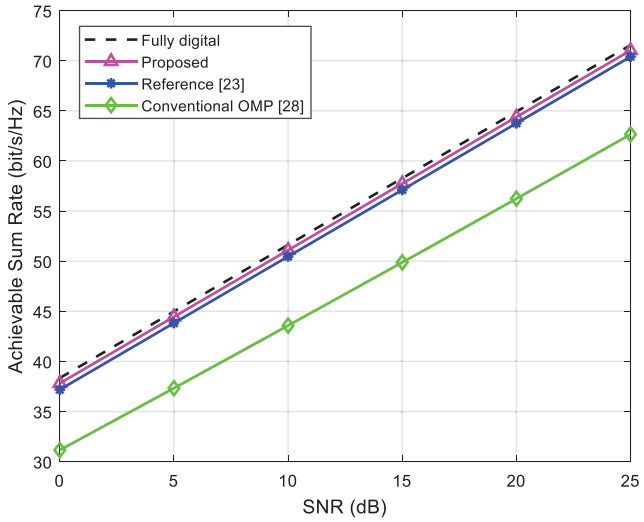

Figure 4: Comparison of achievable sum rate of algorithms *vs.* SNR when $N_{\mathrm{RF}}^t = N_{\mathrm{RF}}^r = N_s = 4$, $N_t = 256$, $N_r = 16$

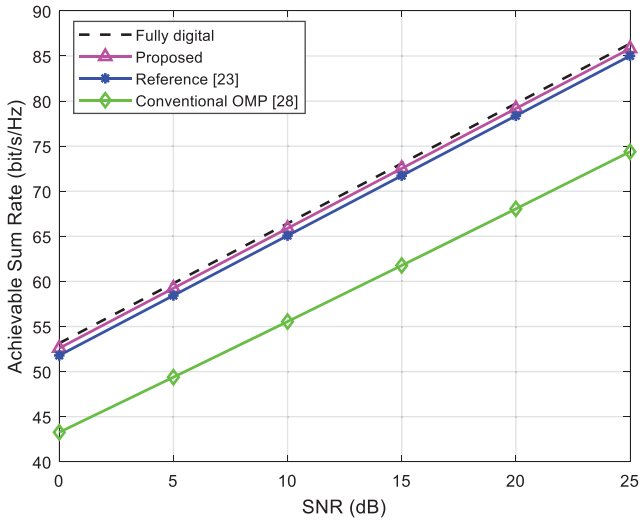

Figure 5: Comparison of achievable sum rate of algorithms *vs.* SNR when $N_{\mathrm{RF}}^{t} = N_{\mathrm{RF}}^{r} = N_s = 4$, $N_t = 1024$, $N_r = 64$

4.3 Complexity Analysis

Fig. 6 compares complexity of the algorithms with increasing number of antennas at the transmitters and $N_{\mathrm{RF}}^{t} = N_{\mathrm{RF}}^{r} = N_s = 4$, and $N_r = 64$. As can be seen from Fig. 6, the complexities of all algorithms increase with increasing the number of antennas at the BS. Moreover, the complexity of the proposed algorithm is lower than the complexities of existing algorithms and also closed to the optimal fully digital precoding. This means that the proposed algorithm requires less number of iterations to achieve the same performance as compared with existing algorithms.

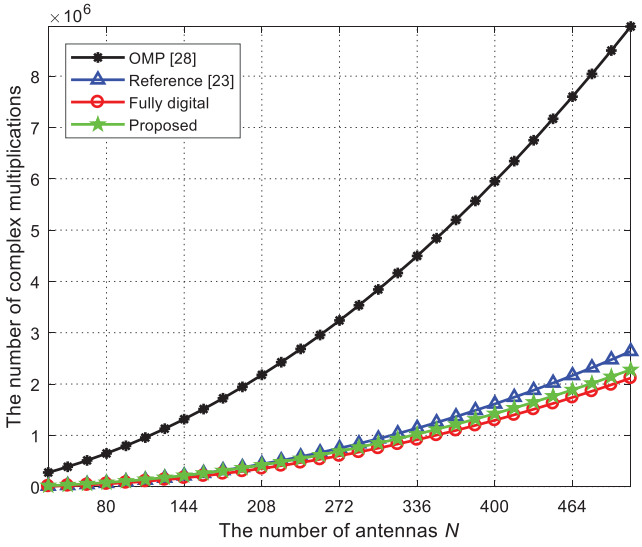

Figure 6: Complexity comparison of the algorithms *vs.* number of antennas

5 Conclusions

This paper proposes an adaptive connection network hybrid precoding with 1-bit quantization, and applies the adaptive algorithm based on machine learning to the adaptive connection structure hybrid precoding, which improves the 1-bit quantization phase shift of the adaptive connection structure. Under the same low hardware complexity, the proposed solution has a higher computational complexity than the switch and inverter hybrid precoding based on the fixed sub-connection of machine learning and the hybrid precoding based on the adaptive connection structure and achievable rate performance. Recently, highly efficient deep learning methods have been applied to hybrid precoding, and precoding with lower computational complexity and better spectral efficiency is worthy of further research.

Acknowledgement: Taif University Researchers Supporting Project Number (TURSP-2020/260), Taif University, Taif, Saudi Arabia.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai and J. Rodriguez, "Millimeter-wave massive MIMO communication for future wireless systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 836–869, 2018.

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

[3] A. Silva, S. Teodoro, R. Dinis and A. Gameiro, "Iterative frequency-domain detection for IA-precoded MC-cDMA systems," *IEEE Transactions on Communications*, vol. 62, no. 4, pp. 1240–1248, 2014.

[4] A. Silva, J. Assuncai, R. Dinis and A. Gameiro, "Performance evaluation of IB-dFE based strategies for SC-fDMA systems," *EURASIP Journal of Wireless Communications and Networking*, vol. 13, pp. 1–10, 2013.

[5] D. Castanheira, A. Silva, R. Dinis and A. Gameiro, "Efficient transmitter and receiver designs for SC-fDMA based heterogeneous networks," *IEEE Transactions on Communications*, vol. 63, no. 7, pp. 2500–2510, 2015.

[6] D. Castanheira, A. Silva and A. Gameiro, "Set optimization for efficient interference alignment in heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5648–5660, 2014.

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

[8] D. Castanheira, S. Teodoro, R. Simoes, A. Silva and A. Gameiro, "Multi-user linear equalizer and precoder scheme for hybrid sub-connected wideband systems," *Electronics*, vol. 8, no. 4, pp. 1–16, 2019.

[9] Y. Wang and W. Zou, "Hybrid digital and analog precoding algorithm for millimeter wave MIMO systems," in *IEEE Wireless Communications and Networking Conf. (WCNC)*, Morocco, pp. 1–6, 2019.

[10] W. Ni, X. Dong and W. Lu, "Near-optimal hybrid processing for massive MIMO systems via matrix decomposition," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3922–3933, 2017.

[11] Z. Wang, M. Li, Q. Liu and A. L. Swindlehurst, "Hybrid precoder and combiner design with low-resolution phase shifters in mmwave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 2, pp. 256–269, 2018.

[12] R. Zhang, W. Zhou, Y. Wang and M. Cui, "Hysbrid precoder and combiner design with finite resolution PSs for mmWave MIMO systems," *China Communications*, vol. 16, no. 2, pp. 95–104, 2019.

[13] F. Dong, W. Wang and Z. Wei, "Low-complexity hybrid precoding for multi-user mmwave systems with low-resolution phase shifters," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9774–9784, 2019.
[14] L. Liang, W. Xu and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 653–656, 2014.
[15] W. Zhang, Z. Wei and J. Wu, "UAV beam alignment for highly mobile millimeter wave communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8577–8585, 2020.
[16] Y. Ren, Y. Wang and G. Xu, "Two-stage hybrid precoding for massive MIMO systems," in *IEEE 22nd Int. Conf. on Telecommunications (ICT)*, Australia, pp. 294–297, 2015.
[17] R. M. Rial, C. Rusu, A. Alkhateeb, N. G. Prelcic and R. W. Heath, "Channel estimation and hybrid combining for mmwave: Phase shifters or switches?," in *Information Theory and Applications Workshop*, USA, pp. 90–97, 2015.
[18] T. Mir, M. Z. Siddiqi, U. Mir, R. Mackenzie and M. Hao, "Machine learning inspired hybrid precoding for wideband millimeter-wave massive MIMO systems," *IEEE Access*, vol. 7, pp. 62852–62864, 2019.
[19] X. Gao, L. Dai, Y. Sun, S. Han and I. C. Lin, "Machine learning inspired energy-efficient hybrid precoding for mmwave massive MIMO systems," in *IEEE Int. Conf. on Communications (ICC)*, Paris, France, pp. 1–6, 2017.
[20] H. Huang, Y. Song, J. Yang, G. Gui and F. Adachi, "Deep-learning based millimeter-wave massive MIMO for hybrid precoding," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 3027–3032, 2019.
[21] T. Ding, Y. Zhao, L. Li, D. Hu and L. Zhang, "Hybrid precoding for beamspace MIMO systems with Sub-connected switches: A machine learning approach," *IEEE Access*, vol. 7, pp. 143273–143281, 2019.
[22] M. Tian, J. Zhang, Y. Zhou, L. Yuan, J. Yang *et al.*, "Switch and inverter based hybrid precoding algorithm for mmwave massive MIMO systems: Analysis on sum-rate and energy-efficiency," *IEEE Access*, vol. 7, pp. 49448–49455, 2019.
[23] X. Zhu, Z. Wang and L. Dai, "Adaptive hybrid precoding for multiuser massive MIMO," *IEEE Communications Letters*, vol. 20, no. 4, pp. 776–779, 2016.
[24] L. Y. Deng, "The cross-entropy method: A unified approach to combinational optimization, monte-carlo simulation, and machine learning," *Technometrics*, vol. 48, no. 1, pp. 147–148, 2012.
[25] A. Costa, O. D. Jones and D. Kroess, "Convergence properties of the cross-entropy method for discrete optimization," *Operation Research Letters*, vol. 35, no. 5, pp. 573–580, 2007.
[26] Z. Wu and M. Kolonko, "Asymptotic properties of a generalized cross-entropy optimization algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 5, pp. 658–673, 2014.
[27] A. M. Elbir and A. K. Papazafeiropoulos, "Hybrid precoding for multiuser millimeter wave massive MIMO systems: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 552–563, 2020.
[28] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar *et al.*, "Millimeter wave mobile communication for 5G cellular: It will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.