

A Codebook Design for FD-MIMO Systems With Multi-Panel Array

Zhilin Fu¹, Sangwon Hwang², Jihwan Moon³, *Member, IEEE*,
Haibao Ren⁴, and Inkyu Lee⁵, *Fellow, IEEE*

Abstract—In this work, we study codebook designs for full-dimension multiple-input multiple-output (FD-MIMO) systems with a multi-panel array (MPA). We propose novel codebooks which allow precise beam structures for MPA FD-MIMO systems by investigating the physical properties and alignments of the panels. We specifically exploit the characteristic that a group of antennas in a vertical direction exhibit more correlation than those in a horizontal direction. This enables an economical use of feedback bits while constructing finer beams compared to conventional codebooks. The codebook is further improved by dynamically allocating the feedback bits on multiple parts such as beam amplitude and co-phasing coefficients using reinforcement learning. The numerical results confirm the effectiveness of the proposed approach in terms of both performance and computational complexity.

Index Terms—Full-dimension MIMO, limited feedback, codebook design, reinforcement learning.

I. INTRODUCTION

Recently, demands for real-time connectivity and insufficient bandwidth usage have become critical issues in the telecommunications field. In response to these needs, a multiple input multiple output (MIMO) technology have been actively studied as a key feature for the fifth generation (5G) system [1], since they provide increased throughput with extra degrees of freedom and high-speed connections [2]. However, its large linear antenna array at a base station (BS) within a limited space generally makes the operational work cumbersome [3].

Long-Term Evolution (LTE) by Third Generation Partnership Project (3GPP) has included a full-dimension MIMO (FD-MIMO) architecture where antenna elements are positioned in a uniform planar array (UPA) [4] to reduce the required space. Furthermore, a multi-panel based FD-MIMO has been proposed in [5], which assembles antenna elements on different panels. Such a multi-panel array (MPA) has raised a lot of interest for millimeter-wave (mmWave) MIMO systems due to the cost reduction and power savings [3].

The MPA systems can also provide high-resolution channel state information (CSI) and allow an efficient beam management with a higher degree of spatial freedom [6], which result in beamforming gains and broader coverage. One may consider adopting conventional

mmWave MIMO codebooks [7]–[9] to the MPA in a similar manner. The authors in [9] presented a codebook that minimizes the average distortion with a fully digital precoding architecture. For hybrid precoding which composes both analog and digital precoders, hierarchical codeword search schemes were proposed in [7], [8] and [10]. In these works, based on discrete Fourier transform (DFT) codebooks, the mmWave subarray characteristics were utilized to obtain the optimal beam direction. Besides, some research has been done which exploits the correlation between codebooks and channel spatial characteristics [11][12]. However, in the MPA systems, each panel carries its own radio frequency (RF) circuits, and thus incurs a phase offset problem during calibrating RF chains in different panels, even if the panels are closely spaced [13]. As a result, existing single-panel codebook designs based on Kronecker product and discrete Fourier transform (DFT) [14] cannot be directly applied to the MPA systems.

Recently, a multi-panel codebook, which is called the Type-I multi-panel (MP) codebook [5], was introduced which quantifies channel direction information (CDI) and generates a DFT codebook of the entire MPA. The integral DFT codebook is then calibrated by co-phasing factors in the panel. Also, the authors of [15] proposed the independent panel codebook (IPC), which is a modified Type-I MP codebook by independently quantifying each panel's CDI to address the phase ambiguity (PA) among panels.

Beside Type-I codebook, another codebook called Type-II was presented in [16], which has a higher precision of a beamforming structure by simultaneously adjusting the amplitudes and power of the over-sampled DFT beams. This Type-II scheme can potentially enhance the performance of the MPA systems, but to the best of the authors' knowledge, the Type-II based MPA codebooks have not been fully optimized yet. One of the main reasons is that an uncaredful adoption of Type-II onto the MPA may result in an overwhelmingly large number of feedback bits.

In this work, we propose novel Type-II codebooks for MPA systems by investigating the physical properties and alignments of the panels to reduce the feedback burden. Noting that quantizing each panel may require too many feedback bits, we group some panels as a single one in a way that an accurate directivity property of Type-II can be preserved as much as possible. We particularly utilize the fact that multiple antennas stacked in the vertical direction create a radiation pattern that is thinner and wider [17]. Hence, in our scheme, panels in the vertical direction are treated as a single panel and only the PA among these groups is adjusted.

The proposed codebook is composed of different components such as beam amplitude, beam combining coefficients, DFT size, and co-phasing coefficients. To further improve the codebook performance, we also introduce a deep Q-learning (DQL) based reinforcement learning (RL) algorithm to identify the best bit allocation for each component. The numerical results verify that our proposed schemes exhibit an outstanding performance-complexity trade-off compared with conventional schemes.

II. SYSTEM MODEL

As illustrated in Fig. 1, we consider an MPA system where a BS is equipped M_v vertical panels and M_h horizontal panels, each of which contains a uniformly spaced rectangular antenna array (URA) with N_v and N_h antennas in vertical and horizontal direction, respectively. At the receiving end, K single-antenna users are served with the same frequency/time resources from the BS.

Manuscript received 7 February 2022; revised 26 July 2022; accepted 26 July 2022. Date of publication 2 August 2022; date of current version 19 December 2022. This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) under Grant 2022R1A5A1027646, and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF), the Ministry of Education under Grant 2021R111A3050126. The review of this article was coordinated by Dr. Maged El Kashlan. (*Corresponding author: Inkyu Lee.*)

Zhilin Fu, Sangwon Hwang, and Inkyu Lee are with the School of Electrical Engineering, Korea University, Seoul 02841, South Korea (e-mail: celynfu@korea.ac.kr; tkddnjs3510@korea.ac.kr; inkyu@korea.ac.kr).

Jihwan Moon is with the Department of Mobile Convergence Engineering, Hanbat National University, Daejeon 34158, South Korea (e-mail: an-schino@staff.hanbat.ac.kr).

Haibao Ren is with the Wireless Network Research Department, Huawei Technologies, Shanghai 201206, China (e-mail: renhaibao@huawei.com).

Digital Object Identifier 10.1109/TVT.2022.3195529

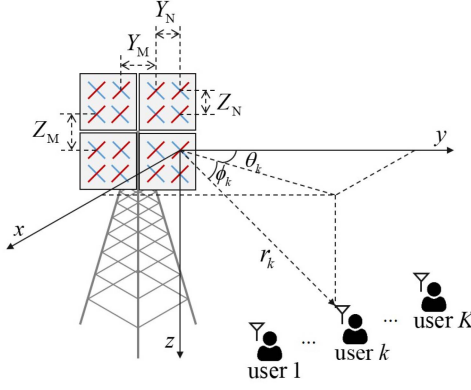


Fig. 1. Multi-user FD-MIMO systems with MPA.

A. Spatial Channel Model for Multi-Panel MIMO Systems

We adopt the spatial channel model (SCM) proposed by the 3GPP standards [4][18] that characterizes three dimensional (3D) correlation among antenna elements. Particularly, the authors in [15] have studied the 3D correlation model with different spacing in horizontal and vertical directions. Assuming a multi-path channel with single polarized antenna elements, the aggregate channel matrix $\mathbf{H}_k \in \mathbb{C}^{M_v N_v \times M_h N_h}$ between the BS and the k -th user is given by

$$\mathbf{H}_k = \frac{1}{\sqrt{S}} \sum_{i=1}^S \mathbf{H}_k^i, \quad (1)$$

where S represents the number of paths, and \mathbf{H}_k^i stands for the channel matrix along the i -th path.

Let us define $\rho_k^i = z_i 10^{-\gamma(r_k)/20}$ as the large-scale fading coefficient for the i -th path where z_i is the random complex coefficient with zero mean and unit variance, and $\gamma(r_k)$ indicates the path-loss with the distance r_k between the BS and the user. In the SCM, the phase of channel elements varies by the panel arrays, while the magnitude is fixed regardless of the panel arrays. For the i -th path and the k -th user, denote ϕ_k^i and θ_k^i as the azimuth angle of departure (AOD) and the zenith angle of departure (ZOD), respectively. Here, we designate $[\mathbf{A}]_{m,n}$ as the (m,n) -th element of a matrix \mathbf{A} . For the MPA systems, the (m,n) -th element of the channel matrix ($m = 1, 2, \dots, M_v N_v$, $n = 1, 2, \dots, M_h N_h$) along the i -th path \mathbf{H}_k^i is expressed as

$$[\mathbf{H}_k^i]_{m,n} = \rho_k^i \exp \left(-j \frac{2\pi}{\lambda} (\Delta_v^m \sin \phi_k^i + \Delta_h^n \cos \theta_k^i \cos \phi_k^i) \right), \quad (2)$$

where the λ means the wavelength, and Δ_v^m and Δ_h^n specify the spacing of antenna elements for the vertical and horizontal domain, respectively as

$$\Delta_v^m = Z_M \left(\left\lceil \frac{m}{N_v} \right\rceil - 1 \right) + Z_N \left(m - \left\lceil \frac{m}{N_v} \right\rceil \right),$$

$$\Delta_h^n = Y_M \left(\left\lceil \frac{n}{N_h} \right\rceil - 1 \right) + Y_N \left(n - \left\lceil \frac{n}{N_h} \right\rceil \right).$$

Here, Z and Y correspond to the spacing values in the vertical and horizontal domain, respectively, Z_M and Y_M indicate the panel spacing distance, Z_N and Y_N equal the distance between antenna elements on the same panel, and $\lceil \cdot \rceil$ denotes the ceiling operation.

With these in hand, the vectorized form of the wireless channel \mathbf{H}_k for the k -th user can be introduced by utilizing the vectorization operation $\text{vec}(\cdot)$ as

$$\mathbf{h}_k = \frac{1}{\sqrt{M_v N_v M_h N_h}} \text{vec}(\mathbf{H}_k), \quad (3)$$

in which \mathbf{h}_k becomes a complex column vector of length $M_v N_v M_h N_h$.

B. 1-to-M Mapping

Ideally, each antenna on FD-MIMO can be equipped with a dedicated RF chain [19]. This approach, however, requires high cost in terms of hardware and energy. To address this issue, separate analog and digital beamforming called 1-to- M mapping [4] can be utilized in which M physical antenna elements are combined into one RF chain to reduce the complexity of digital beamforming, while additional analog beam adjustments are carried out by phase shifters.

For each RF chain, R_v vertical antenna elements are assembled through the 1-to- M mapping approach with $M = R_v$. The m -th element of the weight vector $\mathbf{u} \in \mathbb{C}^{R_v \times 1}$ with these R_v antenna elements is obtained as

$$[\mathbf{u}]_m = \frac{1}{\sqrt{R_v}} \exp \left(-j \frac{2\pi}{\lambda} (m-1) \Delta_v^m \cos \theta_{\text{tilt}} \right), \quad (4)$$

where θ_{tilt} indicates the steering angle in the vertical direction.

The number of RF chains on each panel equals $N_v^R \times N_h$, where $N_v^R = N_v / R_v$. Then, by applying the analog beamforming weight component to the 3D channel coefficients for each antenna element, the (m,n) -th elements ($m = 1, 2, \dots, M_v N_v^R$, and $n = 1, 2, \dots, M_h N_h$) can be represented as

$$[\tilde{\mathbf{H}}_k]_{m,n} = \mathbf{u}^H \tilde{\mathbf{h}}_k, \quad (5)$$

in which $\tilde{\mathbf{h}}_k \in \mathbb{C}^{R_v \times 1}$ is the channel coefficient corresponding to the antenna elements with the same RF chain. By operating the vectorization on $\tilde{\mathbf{H}}_k$, the corresponding channel vector for the k -th user can be denoted as $\tilde{\mathbf{h}}_k \in \mathbb{C}^{M_v N_v^R M_h N_h \times 1}$ similar to (3).

III. TYPE-II SINGLE PANEL CODEBOOK

In this section, we first review Type-II single panel (SP) codebook systems [5], where the CSI feedback framework consists of two matrices selected from separate codebooks. To apply this SP codebook design to our system model, we assume the MPA as a whole panel. The total number of vertical antenna ports and horizontal antenna ports are then expressed as $N_v^t = M_v N_v^R$ and $N_h^t = M_h N_h$, respectively. For further consideration with the cross-polarization antenna model, the total number of antenna ports should be doubled as $2N_v^t N_h^t$. Defining L and R as the number of over-sampled 2D DFT beams and the rank of the codebook, respectively, the basic form of the precoding matrix index (PMI) codebook is given by

$$\mathbf{W} = \mathbf{W}_1 \mathbf{W}_2, \quad (6)$$

where $\mathbf{W}_1 \in \mathbb{C}^{2N_v^t N_h^t \times 2L}$ represents the wide-band feedback which provides a group of beams based on long-term channel statistic, and $\mathbf{W}_2 \in \mathbb{C}^{2L \times R}$ denotes the relatively short-term and wide-band/sub-band (WB/SB) feedback [20].

According to this mechanism, the Type-II SP codebook based CSI feedback can report both the WB and the SB amplitude information of the selected beams. To support a cross-polarized antenna array as adopted in 3GPP standard, the beam selection matrix \mathbf{W}_1 is composed of unconstrained beam selection from an orthogonal basis [16] as

$$\mathbf{W}_1 = \text{diag}(\mathbf{B}, \mathbf{B}), \quad (7)$$

where $\text{diag}(\cdot)$ defines a block diagonal matrix containing each matrix on the diagonal and $\mathbf{B} \in \mathbb{C}^{N_v^t N_h^t \times L}$ accounts for the antenna subgroup with the same polarization.

Since the correlation is high within the same antenna subgroup, it makes sense to use a grid of beam codebooks implemented from the

DFT based precoder vectors [4], and thus \mathbf{B} is constructed by adjacent DFT vectors. To be specific, in the Type-II SP codebook design, the DFT codebook \mathbf{B} consists of $L \in \{2, 3, 4\}$ over-sampled 2D DFT beams with length $N_v^t N_h^t$. Each over-sampled beam can be denoted as $\mathbf{b}_{\theta_v^{(i)}, \theta_h^{(i)}}$ for $i = 1, \dots, L$, where

$$\theta_v^{(i)} = 2^{B_v} n_v^{(i)} + q_v, \text{ and } \theta_h^{(i)} = 2^{B_h} n_h^{(i)} + q_h$$

refer to the indices for the i -th over-sampled beam vector in vertical and horizontal domain, respectively. Here, B_v and B_h are the number of feedback bits corresponding to the selected orthogonal beams, and $q_v \in \{0, \dots, 2^{B_v} - 1\}$ and $q_h \in \{0, \dots, 2^{B_h} - 1\}$ represent the beam rotation factors. In order to make the over-sampled beams orthogonal, the beam indices $n_v^{(i)} \in \{0, \dots, N_v^t - 1\}$ and $n_h^{(i)} \in \{0, \dots, N_h^t - 1\}$ should be $(n_v^{(i)}, n_h^{(i)}) \neq (n_v^{(j)}, n_h^{(j)})$ for $i \neq j$. Consequently, the required number of bits for feedback information on the beam selection is

$$B_{\text{DFT}} = B_v + B_h + \lceil \log_2 \binom{N}{L} \rceil. \quad (8)$$

where the $\lceil \log_2 \binom{N}{L} \rceil$ bits determine which L beams are over-sampled among $N = N_v^t N_h^t$ possible beams.

We now consider the short-term feedback matrix \mathbf{W}_2 which performs a weighted combination of L beams and co-phase adjustment between polarization. We denote $c_{r,l,i}$ as the beam combining coefficient and $p_{r,l,i}^{(\text{WB/SB})}$ as the WB/SB beam amplitude scaling factor for beam i for polarization r and layer l . Without loss of generality, we adopt the rank one scheme ($R = 1$) with the WB-only amplitude scaling mode of the Type-II SP codebook [16] for simplicity. This means that \mathbf{W}_2 has only one layer ($l = 1$) and $p_{r,i}^{(\text{SB})} = 1$.

Assuming that the feedback information of the beam combining coefficient $c_{r,i}$ and the beam amplitude scaling factor $p_{r,i}^{(\text{WB})}$ are respectively expressed with B_c and B_p feedback bits, \mathbf{W}_2 can be formed as

$$\mathbf{W}_2 = \begin{bmatrix} 1, & c_{1,2} p_{1,2}^{(\text{WB})}, & \dots, & c_{r,i} p_{r,i}^{(\text{WB})}, & \dots, & c_{2,L} p_{2,L}^{(\text{WB})} \end{bmatrix}^T, \quad (9)$$

where $c_{r,i} \in \{\exp(j \frac{n\pi}{2}), n = 0, 1, \dots, 2^{B_c} - 1\}$ and $p_{r,i}^{(\text{WB})} \in \{\sqrt{2^0}, \sqrt{2^{-1}}, \dots, \sqrt{2^{-(B_p-2)}}\}$. For the over-sampled beams with different polarization, $c_{r,i}$ and $p_{r,i}^{(\text{WB})}$ are independent.

Note that one of the over-sampled beams is determined with initial values ($c_{1,1} p_{1,1}^{(\text{WB})} = 1$), so that $(2L - 1)(B_p + B_c)$ bits are required for the short-term feedback information. Hence, the total required number of feedback bits B_{II} in cross-polarization antenna systems is computed as

$$B_{\text{II}} = B_{\text{DFT}} + \lceil \log_2(2L) \rceil + (2L - 1)(B_p + B_c), \quad (10)$$

where $\lceil \log_2(2L) \rceil$ bits are used to label the over-sampled $2L$ beams. The codebook length which accounts for the search complexity of the channel quantization for the Type-II codebook design is then calculated as

$$\Omega_{\text{II}} = \Omega_{\mathbf{W}_1} \Omega_{\mathbf{W}_2} = 2L N_v^t N_h^t 2^{B_v+B_h} \times (2L - 1) 2^{B_p+B_c}. \quad (11)$$

This indicates that the search complexity grows with the number of antenna ports for typical MPA systems. To address this issue, we propose a new codebook scheme for MPA designs in the next section.

IV. PROPOSED TYPE-II MULTI-PANEL CODEBOOK

In this section, we propose a new line-panel (LP) codebook which extends the Type-II SP codebook to the multi-panel case with cross-polarization antennas. The LP codebook groups panels in the same

vertical direction as $M_{\text{LP}} = M_h$ line-panels with $N_{\text{LP}} = M_v N_v^R \times N_h$ antenna ports on each line-panel. For the cross-polarization antenna model, the number of antenna ports on each line-panel becomes $2N_{\text{LP}}$.

Let us denote $\mathcal{C}_{\text{SLP}} = \{\mathbf{c}_{\text{SLP}}^{(1)}, \mathbf{c}_{\text{SLP}}^{(2)}, \dots, \mathbf{c}_{\text{SLP}}^{(\Omega_S)}\}$ as the single line-panel (SLP) codebook consisting of Ω_S candidate precoding vectors $\mathbf{c}_{\text{SLP}}^{(i)} \in \mathbb{C}^{2N_{\text{LP}} \times 1}$ for $i = 1, \dots, \Omega_S$. In \mathcal{C}_{SLP} , each SLP codeword candidate $\mathbf{c}_{\text{SLP}}^{(i)}$ is generated following the Type-II SP codebook scheme as a unique combination of different L DFT beams with diverse co-phasing coefficients, amplitude scaling factors, and cross-polarization. Referring to (11), the length of \mathcal{C}_{SLP} can be calculated as

$$\Omega_S = 2L(2L - 1)N_{\text{LP}} 2^{B_v+B_h+B_p+B_c}. \quad (12)$$

Then, each user quantizes the channel vector corresponding to each line-panel, independently, for $m = 1, 2, \dots, M_{\text{LP}}$ as

$$\tilde{\mathbf{h}}_{m,k} = \mathbf{c}_{\text{SLP}}^{(i_{m,k}^*)}, \text{ where } i_{m,k}^* = \arg \max_{1 \leq i \leq \Omega_S} |\tilde{\mathbf{h}}_{m,k}^H \mathbf{c}_{\text{SLP}}^{(i)}|, \quad (13)$$

where $\tilde{\mathbf{h}}_{m,k}$ is formed by a part of elements in $\tilde{\mathbf{h}}_k$ which corresponds to the same line-panel. As a result, the DFT beam candidate $\mathbf{c}_{\text{SLP}}^{(i_{m,k}^*)}$ is chosen which generates the maximum inner product with the corresponding channel vector $\tilde{\mathbf{h}}_{m,k}^H$.

Based on the SLP codebook, we now design the LP codebook for multiple panels. The LP codebook adds the panel co-phasing factors, which capture panel-wise channel characteristics. With B_{LP} feedback bits for the panel co-phasing coefficients, the uniform PA codebook \mathcal{C}_{PA} is defined as a $M_{\text{LP}} \times (M_{\text{LP}} - 1) 2^{B_{\text{LP}}}$ matrix whose columns are given by $[1, e^{j\theta_2}, \dots, e^{j\theta_{M_{\text{LP}}}}]^T$ with $\theta_m \in \{0, \frac{2\pi}{2^{B_{\text{LP}}}}, \dots, \frac{2\pi(2^{B_{\text{LP}}}-1)}{2^{B_{\text{LP}}}}\}$ for $m = 2, 3, \dots, M_{\text{LP}}$. Here, the original phase of the first line-panel is adopted as a reference for the phase adjustment of other panels, and the length of the whole panel co-phasing codebook \mathcal{C}_{PA} can be expressed as $\Omega_P = (M_{\text{LP}} - 1) 2^{B_{\text{LP}}}$.

The LP codebook $\mathcal{C}_{\text{LP}} = \{\mathbf{c}_{\text{LP}}^{(1)}, \mathbf{c}_{\text{LP}}^{(2)}, \dots, \mathbf{c}_{\text{LP}}^{(\Omega_P)}\}$ can be obtained by performing dot products between the aggregation of quantized SLP codewords $[\tilde{\mathbf{h}}_{1,k}^T, \tilde{\mathbf{h}}_{2,k}^T, \dots, \tilde{\mathbf{h}}_{M_{\text{LP}},k}^T]^T$ and each column of \mathcal{C}_{PA} , where $\mathbf{c}_{\text{LP}}^{(i)} \in \mathbb{C}^{M_{\text{LP}} N_{\text{LP}} \times 1}$ for $i = 1, \dots, \Omega_P$ is given by

$$\mathbf{c}_{\text{LP}}^{(i)} = [\tilde{\mathbf{h}}_{1,k}^T, e^{j\theta_2} \tilde{\mathbf{h}}_{2,k}^T, \dots, e^{j\theta_{M_{\text{LP}}}} \tilde{\mathbf{h}}_{M_{\text{LP}},k}^T]^T. \quad (14)$$

Then, each user separately quantizes its aggregate channel vector $\tilde{\mathbf{h}}_k \in \mathbb{C}^{M_{\text{LP}} N_{\text{LP}} \times 1}$ and obtains the suitable panel co-phasing coefficients from \mathcal{C}_{LP} as

$$\hat{\mathbf{h}}_k = \mathbf{c}_{\text{LP}}^{(i_k^*)}, \text{ where } i_k^* = \arg \max_{1 \leq i \leq \Omega_P} |\tilde{\mathbf{h}}_k^H \mathbf{c}_{\text{LP}}^{(i)}|. \quad (15)$$

The total required number of feedback bits B_L in this proposed scheme is then calculated as

$$B_L = M_{\text{LP}} (B_{\text{DFT}} + \lceil \log_2(2L) \rceil + (2L - 1)(B_p + B_c)) + (M_{\text{LP}} - 1) B_{\text{LP}}, \quad (16)$$

where B_{DFT} can be calculated by (8) with $N = N_{\text{LP}}$. Note that if the Type-II SP codebook in (10) is directly applied to MPA by merely multiplying the required number of bits by panels, the total number of bits becomes multiples of B_L in (16). It is clear that our proposed LP codebook design requires much less bits.

Compared with the Type-II SP scheme that neglects panel interspace, the LP codebook scheme can mitigate the PA issue with only a small amount of extra feedback bits. Furthermore, in the LP codebook scheme, the channel vectors corresponding to each line-panel are quantized individually. In contrast, the Type-II SP codebook scheme

quantizes the channel vectors corresponding to the whole MPA at once. As a result, for the LP codebook, the size of the codeword candidate matrices is much smaller, and the resulting search complexity is substantially lower as

$$\begin{aligned}\Omega_L &= M_{LP}\Omega_S + \Omega_P \\ &= 2L(2L-1)M_{LP}N_{LP}2^{B_v+B_h+B_p+B_c} + (M_{LP}-1)2^{B_{LP}}.\end{aligned}\quad (17)$$

In the meantime, assuming that the total number of feedback bits is B , the search complexity of the conventional 2D DFT codebook is

$$\Omega_{DFT} = 2^B. \quad (18)$$

The advantage of the LP codebook is demonstrated in Section VI which compares Ω_{DFT} , Ω_{II} , and Ω_L .

V. REINFORCEMENT LEARNING BASED DYNAMIC BIT ALLOCATION

The 3GPP has given guidelines for the payload calculation for the Type-II SP codebook with common long-term evolution (LTE) MIMO configurations, with $B_p = 3$ and $B_c = 2$ [16]. One feasible approach for the bit allocation issue is to directly adopt the 3GPP guideline to the LP codebook with $B_{LP} = 2$ while allocating the rest of the feedback bits on B_{DFT} . However, it is apparent that such an approach does not guarantee the optimal performance. The complexity of exhaustive search for finding the best bit allocation can also become prohibitive even with a practical number of limited feedback bits. Therefore, in this section, we propose the RL based dynamic bit allocation algorithm to further optimize the codebook.

From the expression of (16), one can see that the bit allocations for B_{LP} , B_v , B_h , B_p , and B_c need to be determined. In order to identify these bits, we adopt an DQL based RL algorithm. To this end, the Markov decision process (MDP) can be designed as follows. First, at the n -th time step, the state s_n can be defined as an aggregation of the bit components in (16) as

$$s_n = \{B_{LP}, B_v, B_h, B_p, B_c\}. \quad (19)$$

Each component in s_n is dynamically selected according to the action a_n from the action space \mathcal{A} as

$$\mathcal{A} = \{B_{LP}^+, B_{LP}^-, \dots, B_p^+, B_p^-, B^0\}, \quad (20)$$

where $(\cdot)^\pm$ means plus/minus one bit on one component, and B^0 indicates retaining the same state. After acting a_n , the remaining bits will be allocated to B_c from (16). Then, the state s_n is updated to s_{n+1} while the reward r_n is received. Notice that the transition to the state s_{n+1} is only dependent on s_n and a_n , satisfying the memoryless characteristic of the Markov property.

Next, we represent G_{n+1} as the average sum-rate value corresponding to the codebook generated through s_{n+1} , and denote \bar{G} as the average sum-rate value simulated from conventional settings. We also define G_{\max} as the maximum average sum-rate value. Then, we utilize the reward r_n as

$$r_n = \begin{cases} \eta b_{n+1} \left(1 + 2^{G_{n+1} - \bar{G}}\right), & \text{if } G_{n+1} = G_{\max}, \\ \eta b_{n+1} 2^{G_{n+1} - \bar{G}}, & \text{if } \bar{G} \leq G_{n+1} < G_{\max}, \\ \eta \log_2 \left(b_{n+1} \frac{G_{n+1}}{\bar{G}}\right), & \text{otherwise,} \end{cases} \quad (21)$$

where η indicates a constant to emphasize the performance difference among bit allocations, $b_{n+1} = B_{n+1}/B_n$ and B_n means the total number of feedback bits at the n -th step. Through the reward r_n ,

Algorithm 1: Proposed Bit Allocation Algorithm.

Initialize the parameter θ randomly, and set the target Q-network as $\theta^- = \theta$

repeat

Select action a_n and calculate s_{n+1} by (16)

Calculate the G_{n+1} and observe r_n by (21)

Store (s_n, a_n, r_n, s_{n+1}) into \mathcal{D}

Replace the oldest tuple if $|\mathcal{D}| \geq l_{\mathcal{D}}$

Sample a random mini-batch $(s, a, r, s') \in \mathcal{B}$

Update θ with the gradient descent method in (22)

if $G_{n+1} > G_{\max}$ **then**

Update $G_{\max} \leftarrow G_{n+1}$

Set $s_{n+1} \leftarrow s_0$

else if $G_{n+1} < \bar{G}/2$

Set $s_{n+1} \leftarrow s_0$

end if

Update $n \leftarrow n + 1$ and $\theta^- \leftarrow \theta$

until convergence

s_{n+1} can be evaluated whether it is an optimal bit allocation scheme or not.

Based on the MDP design above, we propose the DQL based dynamic bit allocation method. In the proposed DQL based method, an agent efficiently determines a bit allocation strategy. Specifically, at each time step, the deep Q-network (DQN) with a trainable parameter θ approximates the Q-value $Q(s_n, a_n; \theta)$ which represents the expected cumulative reward [21]. The action a_n is obtained and conducted based on the ϵ -greedy algorithm. Then, the tuple (s_n, a_n, r_n, s_{n+1}) is stored into the experience-replay memory \mathcal{D} with length $l_{\mathcal{D}}$.

With mini-batch samples $(s, a, r, s') \in \mathcal{B}$, the DQN's parameter θ is updated by the gradient descent method $\theta \leftarrow \theta - \alpha \nabla_{\theta} \mathcal{L}$, where α and $\nabla_{\theta} \mathcal{L}$ represent the learning rate and the gradient of the loss value \mathcal{L} , respectively. The loss function \mathcal{L} is formulated as

$$\mathcal{L} = \frac{1}{|\mathcal{B}|} \sum_{\mathcal{B}} \left[r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right]^2, \quad (22)$$

where γ is the discount factor and $Q(s', a'; \theta^-)$ stands for the estimated Q-value of the target Q-network with parameter θ^- [22]. The detailed procedure of the proposed dynamic bit allocation scheme is summarized in Algorithm 1. One can note that our proposed bit allocation algorithm is flexible enough to find the best bit allocation method according to different panel configurations and various B . Moreover, our proposed LP codebook is applicable even when B is smaller than the minimum requirement of the 3GPP-based bit allocation method.

VI. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we compare the performance of the proposed codebook with conventional schemes. Table I lists the detailed simulation configurations. To conduct practical simulations, all the experiments are based on cross-polarized antenna systems. We name the proposed LP codebook with the 3GPP based bit allocation introduced in Section V as “3GPP based LP”. Also, the LP codebook scheme with the optimal bit allocation scheme given by Algorithm 1 in Section V is named as “RL based LP”.

Fig. 2 provides the search complexity of the conventional 2D DFT codebook, the Type-II SP codebook and the proposed 3GPP based LP codebook with respect to the number of feedback bits B . The antenna ports N_v^t and N_h^t in (11) correspond to $M_v N_v / 8 = 1$ and $M_h N_h = 4$, respectively. For the LP scheme, (17) leads to $M_{LP} = M_h = 2$ and

TABLE I
SIMULATION SETUP

Beamforming	zero-forcing (ZF)
1-to- M mapping R_v	8
Number of over-sampled DFT beams L	2
Number of users K	3
Ground distance between a BS and a user d_k	100 m
Carrier frequency	900 MHz
Path-loss $\gamma(r_k)$	$8 + 37.6 \log r_k$
Antenna distance Y_N, Z_N	$0.7\lambda, 0.5\lambda$
Height of a BS and a UE	30 m, 2 m
Angle-of-arrival in azimuth domain θ_k^s	$[0, \pi]$
Angle-of-arrival in elevation domain ϕ_k^s	$[0, \frac{\pi}{36}]$
Number of paths S	20
Transmit power at a BS, P_t	10 dBm
Bandwidth	4 MHz
Noise figure	3 dB
Amplification factor for reward calculation η	1000
Mini-batch size $ \mathcal{B} $	128
Experience-replay memory size $l_{\mathcal{D}}$	2000
Reward discount factor γ	0.99
Learning rate α	0.001

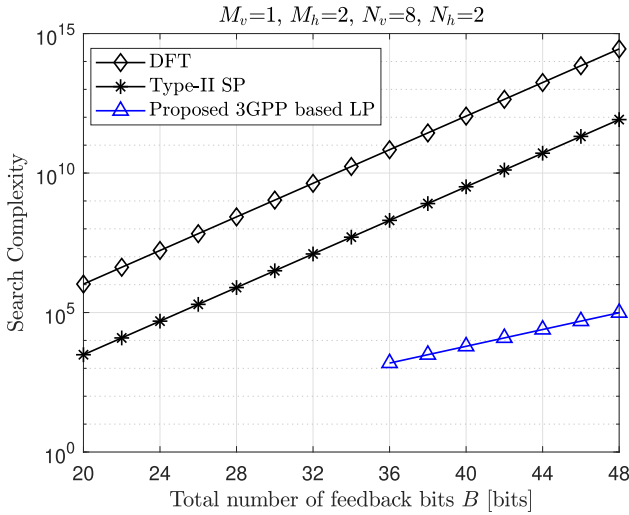


Fig. 2. Search Complexity with respect to the total number of feedback bits B .

$N_{LP} = M_v N_v N_h / 8 = 2$. The required minimum bits on each component are $B_{LP} = 2$, $B_v = 0$, $B_h = 0$, $B_p = 3$, $B_c = 2$, and the minimum B_L is 36 in (16). We notice here that although the required minimum bit B for the 3GPP based LP codebook is as large as 36 bits, the advantage of the proposed LP codebook in terms of search complexity is significant. For instance, the search complexity of the 3GPP based LP codebook when $B = 48$ is close to that of the Type-II SP codebook with $B = 24$.

Fig. 3 exhibits the average sum-rate in terms of the number of feedback bits B . In the figure, the DFT codebook shows no performance change over B . A theoretical proof of the inadequacy of the conventional DFT was given in [23] for MU-MIMO systems. According to [23], when the angle spread between antennas is large, or the size of the DFT codebook is large, the sum rate of the MU-MIMO system does not improve with the feedback bits. This explains why the DFT scheme in our numerical results does not exhibit noticeable gains. The plot also shows that the proposed RL based bit allocation outperforms other schemes for all B . Especially when $B = 40$, a performance gain of the

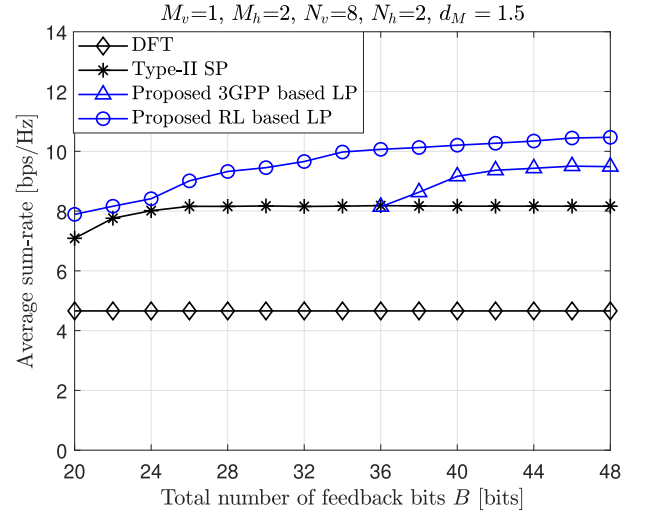


Fig. 3. Average sum-rate with respect to the number of feedback bits B .

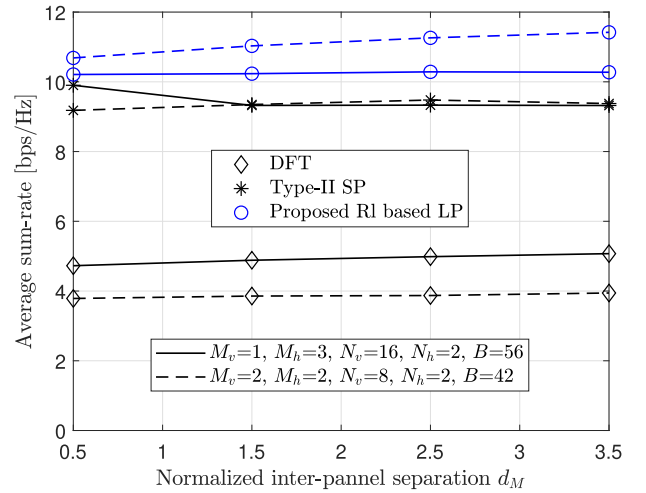


Fig. 4. Average sum-rate with respect to the normalized panel distance d_M .

proposed LP based on RL bit allocation over the Type-II SP scheme reaches 24.8%.

Fig. 4 presents the average sum-rate performance with respect to the normalized panel distance d_M . Here, the bit allocation obtained from the RL approach is adopted. First, it can be observed that the RL based LP codebook outperforms the Type-II SP codebook, since the LP codebook exploits panel co-phasing information. The average sum-rate performance gap between these two codebook designs is narrow when d_M equals the antenna distance. Since the PA issue becomes distinct as more panels are combined in the system, the performance of the RL based LP codebook improves with d_M .

VII. CONCLUSION

In summary, our proposed LP codebook scheme compensates the PA issue by the panel co-phasing factors and individually quantizes the CDI of each line-panel. Moreover, the LP codebook inherits a high degree of freedom from the Type-II SP codebook. These features make the proposed LP codebook achieve the highest gain among all codebook schemes. For the Type-II SP codebook, neglecting the discontinuity between different panels results in a performance loss. Also, we confirm that the LP codebook with RL-based bit allocation designs is an efficient

and low-complexity scheme. Furthermore, MPA can be considered as a promising technique for massive MIMO and mmWave systems, and our future work may explore hybrid precoding designs for MPA systems. Reducing feedback bits and feedback delay can also be an interesting future work.

REFERENCES

- [1] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [2] E. Hong et al., "6G R&D vision: Requirements and candidate technologies," *J. Commun. Netw.*, vol. 24, pp. 232–245, Apr. 2022.
- [3] Y. Huang, Y. Li, H. Ren, J. Lu, and W. Zhang, "Multi-panel MIMO in 5G," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 56–61, Mar. 2018.
- [4] Q.-U.-A. Nadeem, A. Kammoun, and M.-S. Alouini, "Elevation beamforming with full dimension MIMO architectures in 5G systems: A tutorial," *IEEE Commun. Surv. Tut.*, vol. 21, no. 4, pp. 3238–3273, Fourthquarter 2019.
- [5] 3GPP, "NR; Physical layer procedures for data," 3GPP Sophia Antipolis Techn., France, Tech. Specification TS 38.214, V16.7.0, 2021.
- [6] S. Ahmadi, *5G NR: Architecture, Technology, Implementation, and Operation of 3GPP New Radio Standards*. New York, NY, USA: Academic Press, 2019.
- [7] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.
- [8] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook design for millimeter-wave channel estimation with hybrid precoding structure," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 141–153, Jan. 2017.
- [9] A. W. Shaban, O. Damen, Y. Xin, and E. Au, "Statistically-aided codebook-based hybrid precoding for millimeter wave channels," *IEEE Access*, vol. 8, pp. 101500–101513, 2020.
- [10] C. Qi, K. Chen, O. A. Dobre, and G. Y. Li, "Hierarchical codebook-based multiuser beam training for millimeter wave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8142–8152, Dec. 2020.
- [11] D. Yang, L.-L. Yang, and L. Hanzo, "DFT-Based beamforming weight-vector codebook design for spatially correlated channels in the unitary precoding aided multiuser downlink," in *Proc. IEEE Int. Conf. Commun.*, 2010, pp. 1–5.
- [12] J. Li et al., "Codebook design for uniform rectangular arrays of massive antennas," in *Proc. IEEE 77th Veh. Technol. Conf.*, 2013, pp. 1–5.
- [13] 3GPP, "Multi-panel/multi-TRP transmission," 3GPP TSG-RAN WG1, Sophia Antipolis Techn., France, Tech. Rep. TR R1-1702071, Feb. 2017.
- [14] J. Song, J. Choi, and D. J. Love, "Common codebook millimeter wave beam design: Designing beams for both sounding and communication with uniform planar arrays," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1859–1872, Apr. 2017.
- [15] H. M. Shin et al., "Performance evaluation of codebook designs for FD-MIMO with multiple panel array systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2018, pp. 1–5.
- [16] 3GPP, "WF on Type I and II CSI codebooks," 3GPP TSG-RAN WG1-89, Sophia Antipolis Techn., France, Tech. Rep. TR R1-1709232, May 2017.
- [17] C. A. Balanis, *Antenna Theory: Analysis and Design*, 4th ed. Hoboken, NJ, USA: Wiley, 2016.
- [18] W. Liu, Z. Wang, C. Sun, S. Chen, and L. Hanzo, "Structured non-uniformly spaced rectangular antenna array design for FD-MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3252–3266, May 2017.
- [19] J. Lee, T. Oh, J. Moon, C. Song, B. Lee, and I. Lee, "Hybrid beamforming with variable RF attenuator for multi-user mmWave systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9131–9134, Aug. 2020.
- [20] H. Asplund et al., *Advanced Antenna Systems for 5G Network Deployments: Bridging the Gap Between Theory and Practice*. Cambridge, MA, USA: Academic Press, Jun. 2020.
- [21] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, Feb. 2015.
- [22] S. Hwang, H. Kim, H. Lee, and I. Lee, "Multi-agent deep reinforcement learning for distributed resource management in wirelessly powered communication networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 14055–14060, Nov. 2020.
- [23] F. Yuan, S. Han, C. Yang, Y. Zhang, G. Wang, and M. Lei, "Weighted DFT codebook for multiuser MIMO in spatially correlated channels," in *Proc. IEEE 73rd Veh. Technol. Conf.*, 2011, pp. 1–5.