Fast Beam Training for FDD Multi-User Massive MIMO Systems with Finite Phase Shifter Resolution

Ke Xu, Fu-Chun Zheng, *Senior Member, IEEE*, Pan Cao, *Member, IEEE*,
Hongguang Xu, and Xu Zhu, *Senior Member, IEEE*

***Abstract*—A fast beam training scheme for the massive antenna arrays is a key to frequency division duplex (FDD) mmWave systems, as channel reciprocity between the down-link (DL) and up-link (UL) channels does not hold in general, requiring feedback mechanism for DL beam selection. Due to the large number of antennas in a massive array, it is not very practical to conduct an exhaustive search, especially when considering the small angular coverage of one directional narrow beam (hence the number of candidate beams). To address such a challenge, we consider the 3D beam matching problem for a multi-user massive multiple-input multiple-output (MIMO) system, and propose a two-stage hierachical codebook along with the corresponding fast beam training scheme. The proposed codebook contains a primary codebook for hierarchical beams and an auxiliary codebook for narrow beams following the limited resolution of the phase shifters (PSs). The fast beam training scheme based on the codebook not only reduces beam training overhead, but also is applicable to the scenario where there exist multiple propagation paths for one mobile station (MS). Numerical results show that the proposed scheme not only enjoys a lower beam matching complexity (i.e. the training overhead), but also achieves a comparable performance to the exhaustive search scheme.**

I. Introduction

TREATED as a promising technology for next-generation wireless communications, massive MIMO systems working at the mmWave frequency bands have attracted much attention in recent years. The advantage of using the mmWave bands is two-fold: firstly it provides abundant frequency spectrum resource that enables higher system capacity, and secondly the massive array, comprising a large number of antennas, is able to generate multiple high-gain directional beams, compensating the heavy path loss incurred due to mmWave signal propagation [1]–[3].

As the number of antennas is very large for a mmWave massive MIMO system, the conventional fully-digital beamforming architecture becomes costly, considering the hardware cost and power consumption. Instead, beamforming is usually performed by a single analog beamformer [4], or a hybrid architecture where a small digital precoder is connected to a large analog precoder via RF chains [5] [6]. Both can achieve great reduction in hardware cost and power consumption at the expense of some performance loss, and extensive works have been reported in recent years. Nevertheless, instantaneous channel state information (CSI) is required for both above beamforming schemes. For a practical system, however, an accurate entry-wise estimation for the channel matrix cannot be expected for such a massive array [7]. This has limited the practical use of the adaptive beamforming schemes. As a result, another type of beamforming schemes, which works on a predefined codebook [8], has been proposed and widely studied by the mobile communications industry. In such codebook based schemes, the beam searching space is represented by a predefined codebook that contains a set of codewords, each of which corresponds to a directional beam. The beamforming is then achieved by picking the optimal codeword (hence the optimal beam) through some appropriate metric. By carefully designing the codewords and search algorithm, a codebook based scheme can obtain a favorable performance while maintaining low complexity and robustness against imperfectly estimated CSI.

Recently, extensive works on beam selection schemes have turned the attention to the sparsity nature of mmWave channels, since it has been widely acknowledged that the channel matrix involves only a very limited number of angle-of-arrival (AoA) and angle-of-departure (AoD) pairs [9]. This has inspired the compressed sensing (CS) approaches where the limited number of propagation paths are sparsely represented to obtain a faster beam search [10] [11]. For these CS based approaches, assumptions that instantaneous channel parameters can be obtained accurately have to be made. Unfortunately, such assumptions may be difficult to meet as considerable time-variation can occur on mmWave channel parameters, even for MSs with low mobility [12]. Such time-variations lead to a very short coherence time for mmWave channels, making a fast beam selection and tracking procedure very critical.

Manuscript received March 8, 2020; revised July 19, 2020 and October 15 2020; accepted December 9, 2020. This work was supported in part by the National Major Research and Development Program of China under Grant 2020YFB1805005, in part by the Shenzhen Science and Technology Program (Grants Nos. KQTD20190929172545139 and JCYJ20180306171815699), and in part by the project "The Verification Platform of Multi-tier Coverage Communication Network for Oceans" under Grant PCL2018KP002. The associate editor coordinating the review of this article and approving it for publication was Prof. L. Zhao. (*Corresponding author: F.-C. Zheng*.)

K. Xu and H. Xu are with the School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, and also with the Peng Cheng Laboratory, Shenzhen 518055, China (e-mail: 18B952055@stu.hit.edu.cn; xhg@hit.edu.cn).

F.-C. Zheng are with the School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China (e-mail: fzheng@ieee.org; xuzhu@ieee.org)

P. Cao is with the School of Engineering and Technology, University of Hertfordshire, Hertfordshire AL10 9AB, U.K. (e-mail: p.cao@herts.ac.uk).

X. Zhu is with the Department of Electrical and Electronics Engineering, University of Liverpool, Liverpool L69 3GJ, U.K. (e-mail: xuzhu@liverpool.ac.uk)

The large number of antennas for a massive array produces highly directional but very narrow beams. This brings another challenge for the beam training procedure, as a large number of beams ought to be contained in the codebook in order to cover the whole angular space with sufficient granularity. For the time division duplex (TDD) systems, the base station (BS) can directly obtain CSI from uplink pilot signals through channel reciprocity (but still subject to downlink channel calibration). Yet for the frequency division duplex (FDD) systems, the BS cannot acquire CSI via uplink pilots due to the frequency gap between uplink and downlink channels. Therefore, the beam training process, where the BS scans different training beams and the optimal one is selected through the feedback (e.g. the precoding matrix indicator (PMI) in 5G NR) of the MSs, is a reasonable approach for a non-massive MIMO system. However, in a massive MIMO cellular system, this can severely slow down the initial access (IA) phase, where an MS tries to establish a connection when entering a new cell or when a handover between cells is taking place [13]. The large size of the codebook may incur a prohibitively large beam training overhead, especially if exhaustive search is directly employed. This, given the beam switching time (i.e. 20-40 microseconds currently [14]) of mmWave arrays, will have a direct impact on all the key performance indicators (KPIs, including the end-to-end latency) of 5G and B5G systems. For the current Long Term Evolution (LTE) systems, omnidirectional beamforming is usually used at the beginning of the IA phase for a quicker physical link establishment. However, for mmWave systems, due to the heavy path loss, such an omnidirectional start-up leads to a mismatch: the MS can only be detected when it is very close to the BS, but in fact data streams should be transmitted between the MS and BS at a much longer range [15]. This mismatch further motivates the need to obtain a fast beam selection scheme, in order to perform beamforming in both IA and data transmission phases.

To this end, beamforming schemes based on hierarchical codebooks, rather than exhaustive search, have been proposed [16]–[18]. In such schemes, the beams on the upper layers contain only several narrow beams, and a hierarchical search can be performed among them [19]. Yet training overhead can still be a key issue when multiple MSs need to be served at the same time. Furthermore, merely using a hierarchical codebook with analog beams cannot provide a flat gain for all spatial directions since the hierarchical codebook can generate irregular beam patterns, which amplify paths close to the center of beams but weaken those close to the edges [20]. This may cause a mis-selection of the codewords when some propagation paths happen to point at the boundary of two adjacent beams, hence is not ideal for mmWave massive array beam training [21]. Recent work [18] has considered this scenario and included a further detection step in the transition gap between adjacent beams. However, this approach would result in extra beam training overhead due to the extra search in the gap between the adjacent beams.

To address the aforementioned challenges, in this paper we focus on the beam training for a multi-user mmWave massive MIMO system and propose a 3D codebook design, as well as the corresponding fast beam training scheme. We also investigate the scenarios where multiple propagation paths exist, and propose an improved beam training scheme to account for them. The main contributions of this paper are as follows:

- We develop a multi-layer two-stage 3D codebook that can be used in hierarchical beam training for a multi-user system. The codebook consists of a hierarchical primary codebook for a fast multi-user beam search, along with an auxiliary codebook with high resolution for further beam refinement but meets the finite resolution constraint of the PSs. The hierarchical multi-user beam search saves beam training overhead by ignoring the beams that point to no MS on each layer.
- We develop a novel codeword configuration for the proposed two-stage codebooks, where the spatial coverage of each beam on the bottom layer of the primary codebook is set to be slightly smaller than the total coverage of the auxiliary codebook. Such beam coverage configuration not only assures that the beams meet the constant amplitude constraint of the PSs, but also avoids potential false beam detection when the paths point at the boundary between two adjacent beams.
- Based on the hierarchical beam training, the case where multiple propagation paths exist is taken into consideration, in order to obtain a better performance when multiple codewords per MS can be selected for some MSs. The scenarios where paths are close to or far away from each other are both investigated, and the corresponding improved beam training approaches are explored accordingly.

Notations: In the rest of this paper, we use the following notations: boldface uppercase $\boldsymbol{A}$ denotes a matrix, boldface lowercase $\boldsymbol{a}$ a vector, $\mathcal{A}$ a set, α a scalar, $\boldsymbol{A}^H$, $\boldsymbol{A}^*$ and $\boldsymbol{A}^T$ the conjugate-transpose (Hermitian), conjugate and transpose of matrix $\boldsymbol{A}$ respectively.

II. System Model

A BS equipped with a half-wavelength uniform planar array (UPA) of $M \times N$ antennas, serving K single-antenna MSs simultaneously is considered, as shown in Fig. 1, where K data streams are precoded firstly with a small baseband precoder $\boldsymbol{F}_{BB}$, then a large analog beamformer (RF precoder) $\boldsymbol{F}_{RF}$ using phase shifters (PSs) via N_{RF} RF chains.

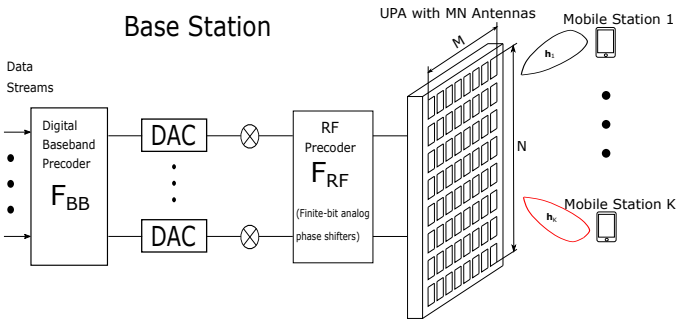


Fig. 1: The block diagram of a multi-user system.

The steering vector for the UPA, $\boldsymbol{a}(\theta,\phi)$, is expressed by stacking the phases of all the antennas as

$$\boldsymbol{a}(\theta,\phi)=\Big[1,\cdots,e^{-j\pi\sin\theta[(m-1)\cos\phi+(n-1)\sin\phi]},\\ \cdots,e^{-j\pi\sin\theta[(M-1)\cos\phi+(N-1)\sin\phi]}\Big]^T, \tag{1}$$

where θ and ϕ are the elevation and azimuth angles for a propagation path, as shown in Fig. 2.

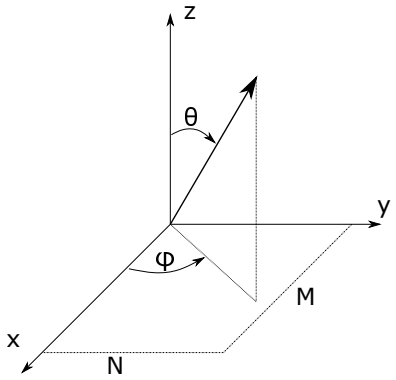


Fig. 2: Geometric structure of the UPA.

It is commonly acknowledged that mmWave channels consist of only a few propagation paths. Assuming that there are L_k propagation paths for the k-th MS , with channel gain $\alpha_{k,l}$ for the lth path, the mmWave channel $\boldsymbol{h}_k$ can be written as

$$\boldsymbol{h}_k=\sum_{l=1}^{L_k}\alpha_{k,l}\boldsymbol{a}\left(\theta_{k,l},\phi_{k,l}\right), \tag{2}$$

where $(\theta_{k,l},\phi_{k,l})$ denotes the AoD of the lth path of the kth MS. From the expression of the steering vector $\boldsymbol{a}(\theta,\phi)$ in (1), it can be observed that the elevation and azimuth angles $(\theta_{k,l},\phi_{k,l})$ are coupled for the phases of the antennas. For the ease of illustration, we introduce a pair of spatial variables $(\psi_{k,l},\zeta_{k,l})$ to decouple $(\theta_{k,l},\phi_{k,l})$, as

$$\psi_{k,l}\triangleq\frac{\sin\theta_{k,l}\cos\phi_{k,l}}{2}, \tag{3}$$

and

$$\zeta_{k,l}\triangleq\frac{\sin\theta_{k,l}\sin\phi_{k,l}}{2}. \tag{4}$$

Therefore, the channel vector $\boldsymbol{h}_k$ becomes

$$\boldsymbol{h}_k=\sum_{l=1}^{L_k}\alpha_{k,l}\boldsymbol{a}'\left(\psi_{k,l},\zeta_{k,l}\right), \tag{5}$$

with reformed steering vector

$$\boldsymbol{a}'(\psi_{k,l},\zeta_{k,l})=\Big[1,\cdots,e^{-2j\pi[(m-1)\psi_{k,l}+(n-1)\zeta_{k,l}]},\\ \cdots,e^{-j\pi[(M-1)\psi_{k,l}+(N-1)\zeta_{k,l}]}\Big]^T. \tag{6}$$

In this paper, we assume that the RF precoder $\boldsymbol{F}_{RF}$ uses PSs with limited resolution, which leads to the constraint that has to be followed when designing the codebook $\mathcal{F}$. Codewords in the codebook $\boldsymbol{f}\in\mathcal{F}$ contain the quantized phase values applied to the antennas in the array, namely for the p-th antenna, we have

$$[\boldsymbol{f}]_p=\frac{1}{\sqrt{MN}}e^{j\gamma_p},\forall p=1,\cdots,MN \tag{7}$$

where γ_p is the quantized phase value. For the BS with $N_{BS,RF}$ RF chains, the RF precoder $\boldsymbol{F}_{RF}$ can be written as $\boldsymbol{F}_{RF}=[\boldsymbol{f}_1,\cdots,\boldsymbol{f}_{N_{RF}}]$. The signal to be transmitted at the BS side is denoted as $\boldsymbol{s}=[r_1,\cdots,r_K]^T=\sum_{k=1}^K\boldsymbol{s}_k$, where $\boldsymbol{s}_k=[0,\cdots,r_k,\cdots,0]^T$ is the symbol for the k-th MS, and $\mathbb{E}\left[\boldsymbol{s}^H\boldsymbol{s}\right]=(P_t/K)\boldsymbol{I}$ with the total transmitting power P_t evenly allocated to all MSs. The received signal y_k of the k-th MS can be written as

$$y_k=\boldsymbol{h}_k^T\boldsymbol{F}_{RF}\boldsymbol{F}_{BB}\boldsymbol{s}_k+\boldsymbol{h}_k^T\sum_{u\neq k}^{K}\boldsymbol{F}_{RF}\boldsymbol{F}_{BB}\boldsymbol{s}_u+z, \tag{8}$$

where $z\sim\mathcal{N}\left(0,\sigma_n^2\right)$ is the additive Gaussian noise. The first term is the desired signal while the second term represents the interference caused by data symbols of other MSs.

Adopting the zero-forcing(ZF) precoder, the digital precoder $\boldsymbol{F}_{BB}$ is given by $\boldsymbol{F}_{BB}=\boldsymbol{H}_e^H\left(\boldsymbol{H}_e\boldsymbol{H}_e^H\right)^{-1}$, where $\boldsymbol{H}_e^H=\left[\boldsymbol{h}_1^T\boldsymbol{F}_{RF},\cdots,\boldsymbol{h}_K^T\boldsymbol{F}_{RF}\right]$ assuming that $K=N_{RF}$. Then with the selected beamforming vector $\boldsymbol{f}_k$ and channel $\boldsymbol{h}_k$, signal-to-interference-noise-ratio (SINR) can be expressed as

$$\mathrm{SINR}_k=\frac{\left|\boldsymbol{h}_k^T\boldsymbol{F}_{RF}\left[\boldsymbol{F}_{BB}\right]_{:,k}\right|^2}{\sigma^2+\sum_{u\neq k}^{K}\left|\boldsymbol{h}_k^T\boldsymbol{F}_{RF}\left[\boldsymbol{F}_{BB}\right]_{:,u}\right|^2}, \tag{9}$$

and the overall spectrum efficiency R (bit/Hz/s) is given by

$$R=\sum_{k=1}^{K}\log_2\left(1+\mathrm{SINR}_k\right). \tag{10}$$

III. Two-Stage Codebook Design

A. The Auxiliary Codebook Design

To reduce the beam training overhead, we adopt a two-stage codebook design, where a primary codebook $\mathcal{F}_P$ and an auxiliary codebook $\mathcal{F}_A$ are jointly used. Specifically, the auxiliary codebook $\mathcal{F}_A$ is a set of codewords corresponding to the beams with high angular resolution but very limited spatial coverage, produced by the UPA steering vector as

$$\boldsymbol{f}_A(\theta,\phi)=\boldsymbol{f}(\theta,\phi),\boldsymbol{f}_A\in\mathcal{F}_A,\theta\in\mathcal{X},\phi\in\mathcal{Y} \tag{11}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are the angular range of θ and ϕ respectively.

For a practical system using limited resolution PSs, phases of $\boldsymbol{f}_A$ ought to meet the hardware resolution constraint. Assuming that every PS has K_p possible states, for any antenna element (m,n), we would like to assure that the phase resolution does not exceed the resolution of the PS, namely

$$\angle\left\{\frac{[\boldsymbol{f}_A(\theta_a,\phi_a)]_{m,n}}{[\boldsymbol{f}_A(\theta_a^*,\phi_a^*)]_{m,n}}\right\}\geq\frac{2\pi}{K_p},\\ \forall\theta_a,\theta_a^*\in\mathcal{X},\phi_a,\phi_a^*\in\mathcal{Y},\theta_a\neq\theta_a^*,\phi_a\neq\phi_a^*, \tag{12}$$

where $\angle\{\alpha\}$ denotes the phase of a complex value α. Inequality (12) assures that different beams in the auxiliary codebook have different PS states. For a practical implementation of a codeword $\boldsymbol{f}$, the phase of each element in $\boldsymbol{f}$ should be

an integral multiple of the phase resolution unit $2\pi/K_p$. An approximated solution can be obtained by taking the closest integral multiple of $2\pi/K_p$, and the approximation is expected to be accurate enough as the error is at most π/K_p for any single antenna.

B. The Primary Codebook Design

The primary codebook $\mathcal{F}_P$ consists of wide beams pointing at different spatial directions, with lower resolution but larger spatial coverage, which enables a hierarchical search for fast beam training. Specifically, the codewords in $\mathcal{F}_P$ are hierarchically organized: the top layer has smallest number of beams but largest beam width, and each beam on the top layer is divided into several narrower beams on the following layer. As the codewords in the auxiliary codebook $\mathcal{F}_A$ are used after the codewords on the bottom layer of primary codebook $\mathcal{F}_P$ have been properly selected for a hierarchical search, concordance between the codewords on the bottom layer of $\mathcal{F}_P$, and the spatial angle or coverage of $\mathcal{F}_A$, should be established.

Therefore, we design the codewords in the primary codebook $\mathcal{F}_P$ by starting with the codewords on the bottom layer. As the angular direction of one beam is determined jointly by θ and ϕ, to decouple the spatial coverage analysis, we turn to the variables (ψ, ζ) as shown in (3) and (4) respectively. To quantify the beamforming gain and beam coverage, array factor $A(\boldsymbol{f}, \psi, \zeta)$ is introduced as follows:

$$A(\boldsymbol{f}, \psi, \zeta) = \sqrt{MN}\left|\boldsymbol{a}(\psi, \zeta)\, \boldsymbol{f}^H\right|, \tag{13}$$

where $\boldsymbol{f}$ is a codeword and $\boldsymbol{a}(\psi, \zeta)$ the array steering vector corresponding to spatial direction ψ, ζ according to the definition in (1). Beam coverage $\mathcal{C}(\boldsymbol{f})$ for the 3D case (ψ, ζ) can be defined with (13), as

$$\mathcal{C}(\boldsymbol{f}) = \left\{\psi, \zeta | A(\boldsymbol{f}, \psi, \zeta) > \rho \max_{\psi^*, \zeta^*} A(\boldsymbol{f}, \psi^*, \zeta^*)\right\}, \tag{14}$$

where $\rho \in (0, 1)$ is a factor that can be flexibly assigned to follow the beamforming gain requirement. When $\rho = 1/\sqrt{2}$, the case corresponds to the widely used 3dB beamwidth. Without loss of generality, in the rest of this paper we adopt this beamwidth.

Beams in the primary codebook $\mathcal{F}_P$ are produced by beams in the auxiliary codebook $\mathcal{F}_A$. Denoting the auxiliary codebook's total beam coverage, determined by $\mathcal{X}$ and $\mathcal{Y}$, as Δ_ψ and Δ_ζ with respect to ψ and ζ respectively, we let every beam on the bottom layer of the primary codebook $\mathcal{F}_P$ have a spatial coverage slightly smaller than Δ_ψ and Δ_ζ, where the margin is represented by $\{\Theta_\psi \times \Theta_\zeta\}$. Namely, one primary beam covers an area of $\left\{\Delta_\psi^P \times \Delta_\zeta^P\right\}$, with $\Delta_\psi^P = \Delta_\psi - \Theta_\psi$ and $\Delta_\zeta^P = \Delta_\zeta - \Theta_\zeta$. With these notations, Q_A, the number of auxiliary beams required in a auxiliary codebook is given by

$$Q_A = \lceil \mathrm{Max}\left(\frac{\Delta_\psi}{\psi_a}, \frac{\Delta_\zeta}{\zeta_a}\right)\rceil, \tag{15}$$

where $\lceil \circ \rceil$ denotes the ceiling function, $\mathrm{Max}(\alpha, \beta)$ returns the larger value between α and β, $\{\psi_a, \zeta_a\}$ is the beam coverage of one auxiliary beam $\mathcal{C}(\boldsymbol{f}_A)$ as defined in (14), following the PS resolution constraint (12). For each primary beam on the bottom layer of the primary codebook, the number of auxiliary beams needed for one beam on the bottom layer of the primary codebook $\mathcal{F}_P$ is given by

$$S_A = \lceil \mathrm{Max}\left(\frac{\Delta_\psi^P}{\psi_a}, \frac{\Delta_\zeta^P}{\zeta_a}\right)\rceil. \tag{16}$$

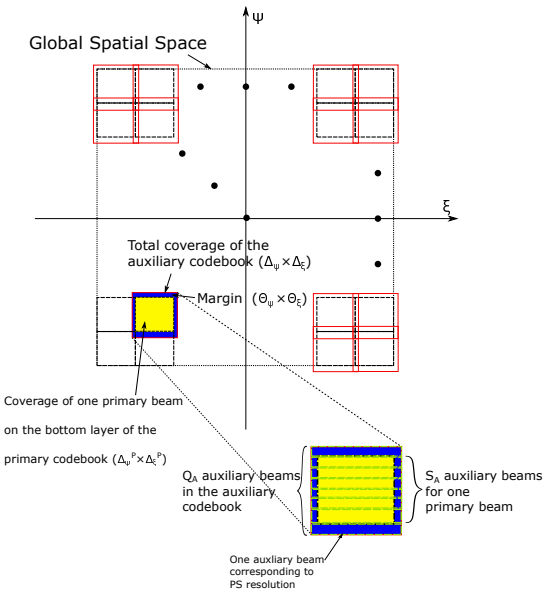


Fig. 3: Beam coverages of the auxiliary codebook and bottom layer of the primary codebook.

An illustration is shown in Fig. 3, where margin $\{\Theta_\psi \times \Theta_\zeta\}$ is painted blue, and coverage of one beam on the bottom layer of $\mathcal{F}_P$, $\left\{\Delta_\psi^P \times \Delta_\zeta^P\right\}$ yellow. The auxiliary codebook's total beam coverage $\{\Delta_\psi \times \Delta_\zeta\}$ is bounded by red, and each auxiliary beam, which follows the PS resolution constraint (12) by green. This requirement assures that once one optimal beam on the bottom layer of the primary codebook has been determined, an exhaustive search over the codewords can be executed within the auxiliary codebook according to the selected primary beam.

Another major benefit of such codebook design is that it improves the beamforming gain for an incorrect beam selection. When the actual spatial direction of one path is located at the boundary of two adjacent beams, the beamforming gain will be quite similar for both beams. For the proposed approach, as the auxiliary codebook covers a larger angle than an individual beam of the primary codebook, an optimal beam can still be selected thereafter even though the beam in the primary codebook may not be the optimal one. As the 2D example has shown in Fig. 4, though the optimal primary beam may not be selected correctly if the path points at the boundary of two adjacent beams, the larger coverage of the auxiliary beams can still allow an optimal beam to be reached with high probability, near the edge of the mis-selected primary beam. In all, the procedure of designing the primary and auxiliary codebook is summarized graphically in Fig. 5.

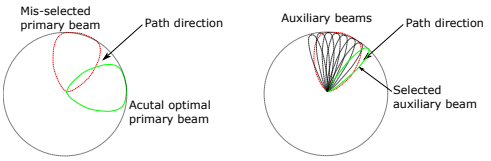


Fig. 4: Auxiliary beams for a mis-selected primary beam.

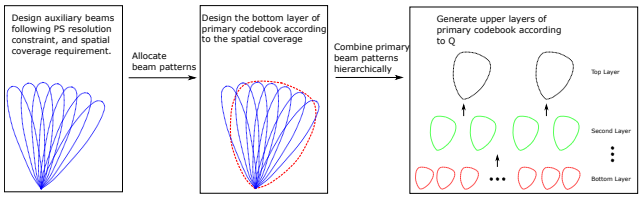


Fig. 5: Procedure of designing $\mathcal{F}_A$ and $\mathcal{F}_P$.

IV. Multi-user Hierarchical Beam Training

The proposed multi-user hierarchical beam training contains a hierarchical search in the primary codebook $\mathcal{F}_P$ and then an exhaustive search in the auxiliary codebook $\mathcal{F}_A$, as shown in Fig. 6. The spatial coverage of $\mathcal{F}_A$ and $\mathcal{F}_P$ has been discussed previously.

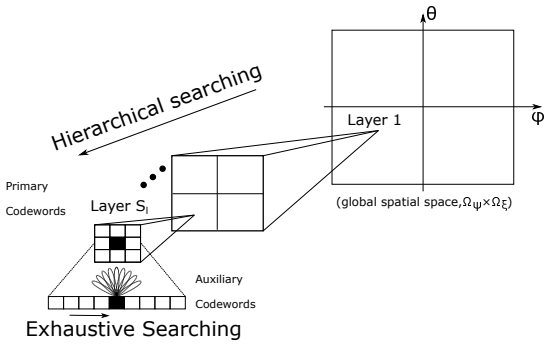


Fig. 6: Beam searching process in the codebooks.

A. Wide beams in the Primary Codebook

To achieve hierarchical beam training, a multi-layer primary codebook is required. The codewords on the bottom layer, which can be used to derive the codewords on other top layers, have been determined in the previous section. If coverage of one primary beam $\left\{\Delta_\psi^P, \Delta_\zeta^P\right\}$ have been determined, we have a total of

$$S_l = \lceil \mathrm{Max}\left(\frac{\Omega_\psi}{\Delta_\psi^P}, \frac{\Omega_\zeta}{\Delta_\zeta^P}\right)\rceil, \tag{17}$$

codewords on the bottom layer of the primary codebook, where $\Omega_\psi, \Omega_\zeta$ denotes the global spatial space. Specially for a beam training over the entire spatial area (i.e. $\theta \in (0, \pi/2), \phi \in (0, 2\pi)$), $\Omega_\psi = \Omega_\zeta = 2$. Then, in order to design a multi-layer codebook with Q codewords on each layer, $S = \lceil \log_Q S_l \rceil$ layers are required, where Q is called the hierarchical factor in previous works such as [21] [22], and $\lceil A \rceil$ denotes the ceiling function. For simplicity it is assumed that $\log_Q S_l$ is an integer, namely $S = \log_Q S_l$, and each of the wide beams on one layer will be divided into Q parts to perform a Q-divided hierarchical beam training. In this way, compared to the conventional exhaustive search scheme that takes $\mathcal{O}(MN)$ tests to obtain an optimal beam, the hierarchical codebook reduces the number to $\mathcal{O}(QS)$ for a single MS scenario.

While the on the bottom layer, beams are produced by the combination of auxiliary beams, in the multi-layer primary codebook, summations of beams on the lower layers are adopted to create wide beams for upper layers. For a Q-divided codebook, denoting the primary codebook for the sth layer from the top as $\mathcal{F}_P^s$, the n_sth codeword can be written as

$$\mathcal{F}_P^s(:, n_s) = \frac{Q^s}{MN} \sum_{i=1}^{S_l/Q^s} \mathcal{F}_P\left(:, n_s + (i-1)Q^s\right) \\ s = 1, \cdots, S, n = 1, \cdots, Q^s, \tag{18}$$

where $\mathcal{F}_P(:, n)$ denotes the nth codeword on the bottom layer, and can be determined with the method proposed in Section III-B. We denote the phase of the m_p-th element for the n_s-th codeword on the bottom layer as

$$\mathcal{F}_P(m_p, n_s) = \exp\left[2j\pi\left(\alpha_{m_p}\psi_{ns} + \beta_{m_p}\zeta_{ns}\right)\right], \tag{19}$$

where α_{m_p} and β_{m_p} are some coefficients depending on antenna index m_p, according the steering vector (1), and ψ_{n_s}, ζ_{n_s} are related to the direction represented by the n_s-th codeword. Then by assuming a beam coverage on the bottom layer of $\left(\bar{\Delta_\psi}, \bar{\Delta_\zeta}\right)$, which are fixed values and properly larger than $(\Delta_\psi, \Delta_\zeta)$ as shown in Fig. 3, the phase of the $n_s + i$th element is given by

$$\mathcal{F}_P(m_p, n_s + i) = \\ \exp\left[2j\pi\left(\alpha_{m_p}\left(\psi_{n_s} + i\bar{\Delta_\psi}\right) + \beta_{m_p}\left(\zeta_{n_s} + i\bar{\Delta_\zeta}\right)\right)\right]. \tag{20}$$

Expression (20) reveals that, for a determined antenna m_p in the array, the phase delay, between one codeword and the ith one after it, has a fixed value of $\Delta_i = i\left(\alpha_{m_p}\bar{\Delta_\psi} + \beta_{m_p}\bar{\Delta_\zeta}\right)$. Then the summation on the right hand of expression (18) for antenna index m_p, can be treated as a geometric series with common ration $e^{j2\pi\Delta_i}$, namely

$$\begin{aligned}\mathcal{F}_P^s(m_p, n_s) &= \sum_{i=1}^{S_l/Q^s} \mathcal{F}_P\left(m_p, n_s + (i-1)Q^s\right) \\ &= \frac{e^A\left(1 - e^{j2\pi\Delta_i Q^s S_l/Q^s}\right)}{1 - e^{j2\pi\Delta_i Q^s}} \\ &= \frac{e^A\left(1 - e^{j2\pi\Delta_i S_l}\right)}{1 - e^{j2\pi\Delta_i Q^s}},\end{aligned} \tag{21}$$

where $A = 2j\pi\left(\alpha_{m_p}\psi_{ns} + \beta_{m_p}\zeta_{ns}\right)$ is the phase of $\mathcal{F}_P(m_p, n_s)$ as shown in (17). However, for a practical system using PSs with limited resolution, directly using the proposed summation beams are not practical, as the amplitude of (21)

is given by

$$\begin{aligned}\left|\mathcal{F}_P^s\left(m_p,n_s\right)\right| &= \left|\frac{e^A\left(1-e^{j2\pi\Delta_i S_l}\right)}{1-e^{j2\pi\Delta_i Q^s}}\right| \\ &= \sqrt{\frac{\left(1-\cos 2\pi\Delta_i S_l\right)^2+\sin^2 2\pi\Delta_i S_l}{\left(1-\cos 2\pi\Delta_i Q^s\right)^2+\sin^2 2\pi\Delta_i Q^s}} \\ &= \sqrt{\frac{1-\cos 2\pi\Delta_i S_l}{1-\cos 2\pi\Delta_i Q^s}} \\ &= \left|\frac{\sin \pi\Delta_i S_l}{\sin \pi\Delta_i Q^s}\right|,\end{aligned} \tag{22}$$

which does not necessarily equal 1. This means that direct summation does follow the constant amplitude constraint of analog beamforming, therefore cannot be directly implemented by PSs. To account for this, a further refinement based on appropriate approximation of the amplitude is applied as follows.

When we inspect the values of $S_l\Delta_\psi$ and $S_l\Delta_\zeta$, it can be observed from Fig. 3 that the S_l codewords almost occupy the whole global spatial space, except for the relatively small margin $\{\Theta_\psi \times \Theta_\zeta\}$ between the coverage of auxiliary codebook and one codeword on the bottom layer. Therefore we have $S_l\Delta_\psi \approx 1$ and $T\Delta_\zeta \approx 1$, which means that $T\Delta_i$ is very close to an integer, as α_{m_p} and β_{m_p} both denote the index of the m_pth antenna and are integers, thus $\sin \pi\Delta_i S_l \approx 0$. With this approximation, for the amplitude (22), we have

$$\left|\mathcal{F}_P^s\left(m_p,n_s\right)\right| \approx \begin{cases} S_l/Q^s, & \Delta_i \approx Q \\ 0, & \text{otherwise} \end{cases}, \tag{23}$$

where Q is an integer. Since $S_l = Q^S$, S_l/Q^s can be further written as Q^{S-s}. Physically, expression (23) means that, for an analog beamformer using PSs only, the constant amplitude constraint is almost met for any given layer s. This reveals that approximation (23) can be used to meet such amplitude constraint.

An illustrative example of a three-layer primary codebook for 256 antennas, along with an auxiliary codebook which contains 6 beams is shown in Fig 7. Only the beams belonging to one beam on the immediate higher layer is shown on each layer for a clear illustration.

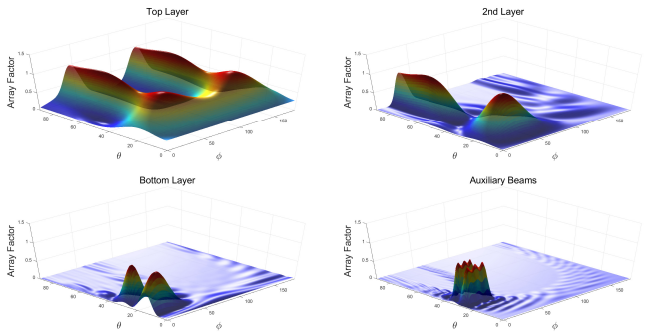


Fig. 7: An example of the codebooks with $Q = 2$.

B. Hierarchical beam search in the primary codebook

After the wide beams on each layer of the primary codebook have been created, the hierarchical beam search in the primary codebook can be conducted. We consider a scenario where K MSs communicate simultaneously with the BS, as shown in Fig. 1. The BS first broadcasts the training symbols with Q wide beams on the top layer of the primary codebook. The MSs then detect the power distribution of the beams and select the one with the highest beamforming gain. The indices of selected beams are then fed back to the BS, marked as *effective beams*. Each of the effective beams is then divided into Q beams and the training process moves on to the next layer, until each MS picks up its optimal beam on the bottom layer of the primary codebook.

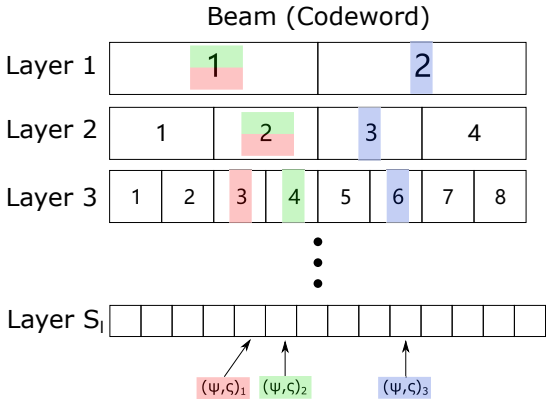


Fig. 8: Multi-user hierarchical beam searching in the primary codebook.

An illustrative example of a $K = 3$ MSs hierarchical beam training is shown in Fig. 8 with each color representing one MS, where a $Q = 2$ divided beam training is performed. On the first layer, BS trains the beams with indices $1, 2$ and the 3 MSs simultaneously receive the training symbols and feed back the optimal beams. On the first layer, Beam $[1]$ is fed back by both MS 1 and 2, and Beam $[2]$ is fed back by MS 3, hence the both are effective beams. Then the two are divided into Beam $[1, 2]$ and $[3, 4]$ respectively. The returned effective beams are Beam $[2]$ for MS 1 and 2, and and Beam $[3]$ for MS 3. Therefore, on Layer 2, only Beam $[2, 3]$ are effective beams. This process will be chronologically executed until the bottom layer has been reached.

For the broadcast beam training rather than the unicast one, the BS does not have information about the characteristics of the MSs. Therefore, the beam feedback will be kept and updated at the BS. The beam training process moves on to the next layer only after all the K MSs have accomplished their beam training on this layer, and the BS has received their feedback, in order to prevent ambiguity in beam indices. For layer $s = 1, 2, \cdots, S$, the BS transmits the training sequences to all K MSs using each of the beamforming vectors on this layer of codebook $\mathcal{F}_P^s$, and performs the proposed hierarchical beam training. The received signal of the uth MS on the sth layer from the top can be written as

$$\boldsymbol{y}_u = \sqrt{\frac{P_T}{MN}}\boldsymbol{h}_u \boldsymbol{f}\left(s,n_s\right)\boldsymbol{s} + \boldsymbol{z}, \tag{24}$$

where P_T denotes the transmitting power, MN the total number of antennas at BS, $\boldsymbol{z}$ the Gaussian white noise, $\boldsymbol{f}(s, n_s) \in \mathcal{F}_p^s$ is the n_sth codeword from the sth layer of the primary codebook, and $\boldsymbol{s}$ the normalized training sequence with unitary amplitude. By detecting the amplitude of $\boldsymbol{y}$, each MS sends back the indices of optimal beams, and the searching goes on until the bottom layer has been reached. A summary of the proposed hierarchical beam training scheme is shown in Algorithm 1, where a vector $\boldsymbol{m}$ is used to keep a record of the effective beams on each layer.

Algorithm 1 Multi-user Beam Training in $\mathcal{F}_p$

Input: Primary codebook $\mathcal{F}_p$, S_l beams on the bottom layer, MS number K, division coefficient Q, Layers $S = \log_Q S_l$
Output: Indices of optimal beams $\boldsymbol{m} \in \mathbb{C}^{1\times K}$

```
for s = 1 : S do
    m → Effective Beams in F_P^s : m = 1, ··· , Q^s
    for u = 1 : K do
        for b_1 = 1 : length(F_P^s) do
            y_u(b_1 * Q − 1 : b_1 * Q) =
            h_u f(s, QF_P^s[b_1] + 1 : QF_P^s[b_1] + Q) + z
        end for
    end for
    [m]_u = arg max_p |y_u(p)|
    update m
end for
```

C. Exhaustive Beam Search for the Auxiliary Codebook

After the previously proposed hierarchical beam search, the optimal beams within the primary beam can be determined, which indicates that the possible optimal beams are located within a very small spatial area $(\Delta_\psi, \Delta_\zeta)$. As this area can be fully covered by the auxiliary codebook as shown in Fig. 3, an exhaustive search will be executed for a more precise beam training result. Assuming that the auxiliary codebook contains Q_A codewords, following the PS resolution constraint (12), then each beam will test the Q_A beams exhaustively and the associated MSs return the optimal beams. Namely, after the optimal beam $\boldsymbol{f}_P^{opt}$ on the bottom layer has been determined, each MS picks up its optimal beam

$$\boldsymbol{f}_A^* = \underset{\boldsymbol{f}_A^* \in \mathcal{F}_A}{\arg} \max \left| y\left(\boldsymbol{f}_P^{opt}, \boldsymbol{f}_A^*\right)\right|, \tag{25}$$

where $y\left(\boldsymbol{f}_P^{opt}, \boldsymbol{f}_A^*\right)$ denotes the received signal using beam $\boldsymbol{f}_P^{opt}$ and $\boldsymbol{f}_A^*$ at the same time. Due to the relative small size of $\mathcal{F}_A$, an exhaustive search will work with affordable overhead.

This exhaustive search will cost an extra of $Q_A K$ time slots, yet thanks to the relative small number of Q_A, it will not bring too much overhead. Further, the actual number of the extra cost should be smaller in most cases, since the MSs sharing the same effective beams can be trained within the same time slot for their common effective beams.

D. Beam Training Overhead Analysis

After the hierarchical search in the primary codebook $\mathcal{F}_P$, and the exhaustive search in the auxiliary codebook $\mathcal{F}_A$, the proposed scheme gives out the optimal beam that ought to be adopted for data transmission. The beam training overhead, which is related to the configuration of the two codebooks $\mathcal{F}_P$ and $\mathcal{F}_A$, is defined as the number of time slots required in order to select the optimal beams for all MSs. Here numerical analysis on the training overhead will be conducted for a typical codebook setup, in order to make an illustrative comparison between the proposed and other similar hierarchical beam training schemes.

For the primary codebook $\mathcal{F}_P$, the number of beams on the bottom layer S_l, as defined in (18), along with the hierarchical factor Q, determines the complexity of the hierarchical search. In the analysis, we take an example of $S_l = 72$ which represents a configuration of 6×12 beams in $\mathcal{F}_P$. In this case, each beam on the bottom layer covers an area of $\Delta_\theta = 90^\circ/6 = 15^\circ$ over the elevation angle domain, and $\Delta_\phi = 360^\circ/12 = 30^\circ$ over the azimuth angle domain. The hierarchical factor Q is set as 2 or 3 here, giving a binary or 3-divided hierarchical search respectively. For the auxiliary codebook $\mathcal{F}_A$, its size Q_A, which is subject to the resolution of the PS as shown in (12), is also directly related to the overhead. In the analysis we take the examples of $Q_A = 3$ and $Q_A = 6$, representing the 4-bit PS and 5-bit PS scenarios respectively.

The proposed two-stage scheme includes a hierarchical search in $\mathcal{F}_P$ and an exhaustive search in $\mathcal{F}_A$, while in a number of previous works, such as [23]–[25], a hierarchical search is often conducted directly over all possible beams. Comparison of beam training overhead between such 'all hierarchical' schemes and the proposed scheme, is shown in Table I. The 'all hierarchical' search here is performed over the all $S_l \times Q_A$ possible beams, whose overhead determined by the primary and auxiliary codebooks.

Since in the proposed scheme, the training overhead can be reduced when multiple MSs share the same effective beams, it will be of interest to analyze the best and the worse case of the proposed scheme. The best case is where all MSs are within the same effective beam, while the worst one is where they all belong to different effective beams. From Table I, it can be observed that, the proposed scheme, which employs a hierarchical search in the primary codebook and an exhaustive search in the auxiliary codebook, has a training overhead that is almost the same as that of all hierarchical search schemes in the worst case. With S_l primary beams on the bottom layer, and an auxiliary codebook that contains Q_A auxiliary beams, the overall beam training overhead T_d for all K MSs is given by

$$T_d^{max} = KQ\lceil \log_Q S_l \rceil + KQ_A, \tag{26}$$

for the worst case, and

$$T_d^{min} = K\left(\lceil \log_Q S_l \rceil + Q - 1\right) + KQ_A, \tag{27}$$

for the best case, where $\lceil \circ \rceil$ denotes the ceiling function. Most practical cases will somewhere in between, and the corresponding training overhead will therefore be between T_d^{min} and T_d^{max}, representing a significant reduction in training overhead. For instance, with the codebook configurations in Table. I, the maximum overhead reduction varies from 47.2%

TABLE I: Overhead comparison between all hierarchical schemes and the proposed two-stage scheme.

Schemes/Overhead (time slots)		All hierarchical		Proposed (worst case)		Proposed (best case)	
MSs	Hierarchical factor $/\mathcal{F}_A$ size	$Q_A = 4$	$Q_A = 8$	$Q_A = 4$	$Q_A = 8$	$Q_A = 4$	$Q_A = 8$
$k = 4$	$Q = 2$	72	80	72	88	30	46
	$Q = 3$	72	72	64	80	31	47
$k = 8$	$Q = 2$	144	160	144	176	54	86
	$Q = 3$	144	144	128	160	59	91

(38 versus 72) to 72.2% (40 versus 144).

V. Beam Training for Multi-path Channels

In Algorithm 1, the MSs select the optimal beams by simply picking up the one with the highest beamforming gain, which makes the beam point at the spatial direction of the dominant path of one MS. For the multi-user system, this works for the assumption $K = N_{RF}$ for (9) where one beam is selected for each MS for fairness concern. When $K > N_{RF}$, it is clearly not possible to allocate at least one beam to each MS, and other means will have to be employed to distinguish some of the MSs, e.g. by allocating them different time-frequency resource blocks [26].

In this section, we extend the beam training problem to the case of $K < N_{RF}$, which allows multiple beams to be selected for one MS to maximize overall sum-rate (10). As shown in Fig. 9, such multi-beam selection can be performed corresponding to each path $P_1, \cdots, P_{L_k}$, rather than one beam aiming at the dominant path P_{dom} to improve beamforming gain, providing that all paths have been identified. Given the synchronization requirements of some recent works (e.g. [27]), and based on the previously proposed two-stage codebook, we then propose a multi-path beam training scheme that identifies paths in an angular manner, with low complexity but high robustness.

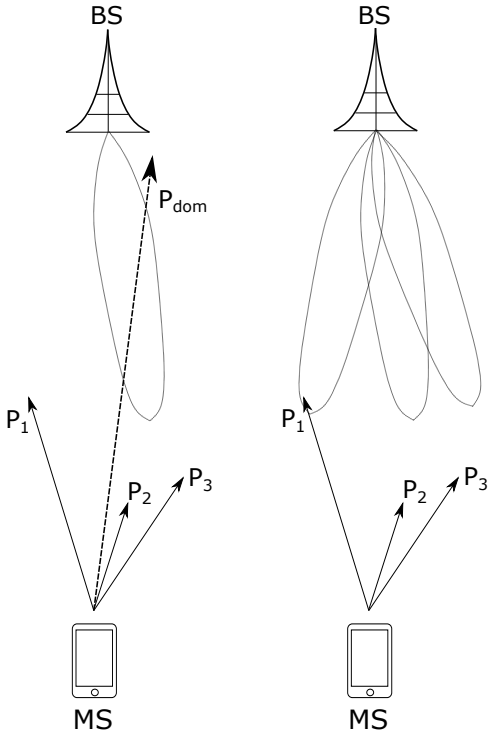


Fig. 9: Multi-path channel beamforming (left: one beam points at the dominant path; right: multiple beams point at all paths).

A. Multi-path beam training for LOS scenario

As measurements have suggested that for the line-of-sight (LOS) scenario, angular spreadings for the propagation paths are very small [28], they can be detected within the adjacent spatial area when a dominant path of one MS has been determined. Specifically, we assume that the possible angular spreading for the propagation paths of one MS is within the coverage of the auxiliary codebook $(\Delta_\psi \times \Delta_\zeta)$. Therefore, once the hierarchical search in the primary codebook has been executed, the possible spatial area that covers most propagation paths is also obtained. Then, instead of conducting an exhaustive search in the auxiliary codebook for only one optimal beam, an iterative search can be performed to extract all the beams for all possible propagation paths.

The proposed iterative approach for the LOS multi-path case is shown in the flow-chart of Fig. 10. The main idea of the iterative approach is to first determine the dominant path, then use several other beams to replace the original beam. After one substitution proposal that improves the beamforming gain has been determined, the process can move on to deal with the new beams sequentially with the same approach. For a LOS case, as the paths are likely to come within a very small spatial area, spatial directions of the propagation paths are assumed to belong to the same codeword in the primary codebook. Therefore, once the an optimal beam $\boldsymbol{f}_P^{dom} \in \mathcal{F}_P$ has been selected according to the dominant path, beams for the actual L_k paths can be found in the spatial coverage of $\boldsymbol{f}_P^{dom}$. Then within the corresponding auxiliary codebook, each path looks for its optimal codeword sequentially with an independently exhaustive search.

Algorithm 2 provides a summary of the proposed iterative algorithm, accounting for the LOS multi-path channel, where $y(\mathcal{C})$ denotes the received signal with beams from the set $\mathcal{C}$, and $\mathcal{C} \pm \boldsymbol{f}$ means add/remove beam $\boldsymbol{f}$ from the set $\mathcal{C}$. In Algorithm 2, the same hierarchical-exhaustive approach is used to find the optimal beam $\boldsymbol{f}_A^{dom}$ for the dominant path, which is added to the set of optimal beams $\mathcal{C}$. Then, for any beam in the set, we check if the beamforming gain improves when the beam is replaced by two other beams. If so, then the beam is replaced by the two beams of the

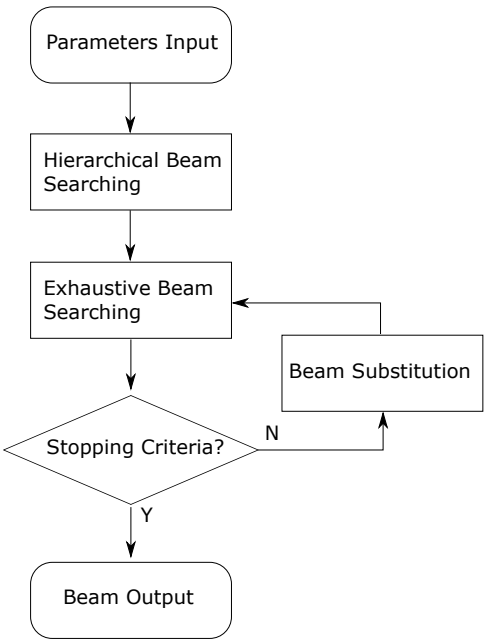


Fig. 10: Iterative beam training for LOS multi-path channel.

highest beamforming gain. An illustrative example is shown in Fig. 11, where a 3-path scenario is shown. Each time one beam is substituted by two other best-performing beams, if any performance improvement can be achieved through this. When the number of beams reaches the actual number of propagation paths, performance cannot be improved by simply adding beams, so the algorithm stops to output the result. Providing L_k paths exist, in each iteration the set is expected to add one beam corresponding to one path. Therefore, the maximum number of iterations I can be set as the expected largest L_k possible, which is often a small integer according to measurements [28].

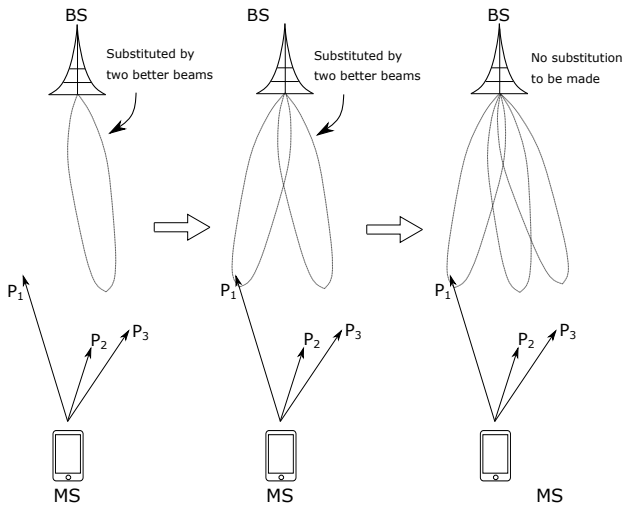


Fig. 11: Multi-path channel beamforming with Algorithm 2.

B. Multi-path beam training for NLOS scenario

For a non-line-of-sight (NLOS) scenario when the LOS path is blocked, the same approach for the LOS scenario may be applied at first glance, if a rough result is enough for the system. However, since for a NLOS case, these propagation paths do not necessarily come within a relatively small area covered by one auxiliary codeword, as the paths can be produced by reflection of different scatterers. Thus the proposed iterative search can potentially result in some errors when directly applied for the NLOS case, since the paths beyond the coverage of auxiliary codebook $\mathcal{F}_A$ cannot be detected. If a more precise result is required for the NLOS case, the searching range ought to be extended to the whole spatial area, which may result in prohibitively large training overhead. Therefore, an alternative approach should be taken.

Algorithm 2 LOS Multi-path Beam Training in $\mathcal{F}_p$

Input: Dominant beam $\boldsymbol{f}_P^{dom}$ by Algorithm 1, limit I
Output: Auxiliary beams for paths $\boldsymbol{f}_A^1, \cdots, \boldsymbol{f}_A^l$

Exhaustive search for auxiliary beam $\boldsymbol{f}_A^{dom} \in \mathcal{F}_A$
Set of beams $\mathcal{C}^1 = \left\{\boldsymbol{f}_A^{dom}\right\}$, index $i = 1$
while $i \leq I$ **do**
  for $\boldsymbol{f} \in \mathcal{C}^i$ **do**
    $\boldsymbol{f}_1, \boldsymbol{f}_2 = \underset{\boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathcal{F}_A}{\arg\max} \left| y\left(\mathcal{C}^i - \boldsymbol{f} + \boldsymbol{f}_1 + \boldsymbol{f}_2\right)\right|$
    if $\left| y\left(\mathcal{C}^i - \boldsymbol{f} + \boldsymbol{f}_1 + \boldsymbol{f}_2\right)\right| > \left| y\left(\mathcal{C}^i\right)\right|$ **then**
      $\mathcal{C}^{i+1} = \mathcal{C}^i - \boldsymbol{f} + \boldsymbol{f}_1 + \boldsymbol{f}_2$
    else
      $\mathcal{C}^{i+1} = \mathcal{C}^i$
    end if
    $i = i + 1$
  end for
end while
Output $\mathcal{C}^i$

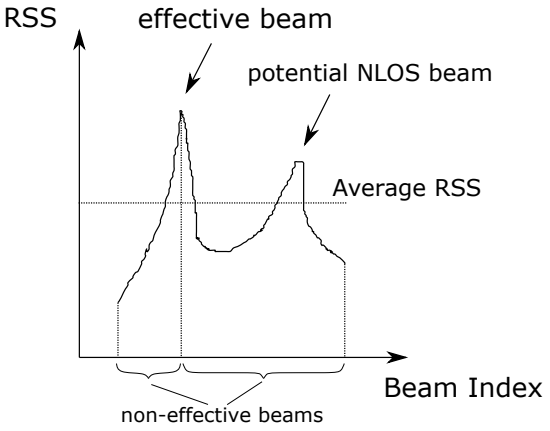


Fig. 12: A potential beam for NLOS multi-path scenario.

Since beamforming gain will be significantly greater when an optimal beam is applied, for the NLOS multi-path channel, some non-effective beams with large enough beamforming gain may also be applied, to account for the paths other than the dominant path. As the example in Fig. 12 shows, on one layer of the hierarchical search, by taking the optimal beam and moving to the next beam, if there exists one beam in the non-effective beam set whose received signal strength (RSS) is significantly larger than the average RSS (quantified by a

threshold ρ_R), then it is kept as a potential beam for the NLOS paths. Once the optimal beams, along with the potential NLOS multi-path ones, on the bottom layer have been determined, beamforming gains are compared for all the combination of these potential beams. Those which enjoy the highest beamforming gains, when used with the optimal beams, are selected, and the same exhaustive search in the auxiliary codebook is performed as in the LOS case. A summarized procedure of the beam training for the NLOS case is shown in Algorithm 3. Since for a practical scenario, both NLOS and LOS multi-paths can exist, the beam substitution proposed for the LOS case is also employed in Algorithm 3 (Steps 16-26) to obtain a more complete algorithm. Namely, any NLoS paths are detected outside the auxiliary codebook (i.e. during the hierarchical search in the primary codebook) while the LoS paths are detected within the auxiliary codebook.

Algorithm 3 Multi-path Beam Training in $\mathcal{F}_p$

Input: Input as in Algorithm 1
Threshold for NLOS multi-path RSS ρ_R
Output: Output as in Algorithm 2
1: **for** $s = 1 : S$ **do**
2: $\boldsymbol{m} \rightarrow$ Effective Beams in $\mathcal{F}_P^s : m = 1, \cdots, Q^s$
3: **for** $u = 1 : K$ **do**
4: **for** $b_1 = 1 : \text{length}\left(\mathcal{F}_P^s\right)$ **do**
5: $\boldsymbol{y}_u\left(b_1 * Q - 1 : b_1 * Q\right) = \sqrt{\frac{P_T}{MN}}\boldsymbol{h}_u \boldsymbol{f}\left(s, Q\mathcal{F}_P^s[b_1] + 1 : Q\mathcal{F}_P^s[b_1] + Q\right) + \boldsymbol{n}$
6: **end for**
7: **end for**
8: $[\boldsymbol{m}]_u = \arg\max_p |y_u(p)|$
9: **if** $|y_u(q)| > \rho_R \times$ Average RSS **then**
10: Add q to the effective beams
11: **end if**
12: update $\boldsymbol{m}$
13: **end for**
14: Dominant beam $\boldsymbol{f}_P^{dom}$, NLOS beams $\boldsymbol{f}_P^{nlos1}, \cdots$
15: Exhaustive search for $\boldsymbol{f}_A^{dom}, \boldsymbol{f}_A^{nlos} \in \mathcal{F}_A$
16: Set of beams $\mathcal{C}^1 = \left\{\boldsymbol{f}_A^{dom}, \boldsymbol{f}_A^{nlos1}, \cdots\right\}$, index $i = 1$
17: **while** $i \leq I$ **do**
18: **for** $\boldsymbol{f} \in \mathcal{C}^i$ **do**
19: $\boldsymbol{f}_1, \boldsymbol{f}_2 = \underset{\boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathcal{F}_A}{\arg\max} \left|y\left(\mathcal{C}^i - \boldsymbol{f} + \boldsymbol{f}_1 + \boldsymbol{f}_2\right)\right|$
20: **if** $\left|y\left(\mathcal{C}^i - \boldsymbol{f} + \boldsymbol{f}_1 + \boldsymbol{f}_2\right)\right| > \left|y\left(\mathcal{C}^i\right)\right|$ **then**
21: $\mathcal{C}^{i+1} = \mathcal{C}^i - \boldsymbol{f} + \boldsymbol{f}_1 + \boldsymbol{f}_2$
22: **else**
23: $\mathcal{C}^{i+1} = \mathcal{C}^i$
24: **end if**
25: $i = i + 1$
26: **end for**
27: **end while**
28: Output $\mathcal{C}^i$

C. Complexity analysis of the multi-path scenario

Complexity analysis when the problem is extended to the multi-path scenario is presented in Table. II, where the complexity of proposed multi-path extraction scheme under the best and worst cases, is compared to that of a brutal force one, i.e all the possible combinations of paths are tested based on the proposed two-stage codebooks directly assuming that L_k is known in advance. A recent work [18] that also considers potential incorrect beam selection when the paths point at the boundary of two adjacent beams is included as well. Different from our work, [18] employs an extra search within the interval of adjacent beams.

For the LoS scenario, it is assume that all the paths are within the spatial coverage of the auxiliary codebook. The worst case takes place when three beam splitting operations are performed, resulting in an extra overhead of $4 + 3 + 2 = 9$ ($Q_A = 4$) or $8 + 7 + 6 = 21$ ($Q_A = 8$) time slots. In contrast, the best case happens when beams cannot be split as no beamforming gain improvement is observed, and the algorithm stops after only one time slots, which means that the 3 LoS paths share the same optimal codeword. For the brutal force scheme, selecting 3 beams from $Q_A = 4$ codewords is equivalent to selecting one beam from 4 codewords for the single-path scenario, hence the complexity is not increasing. However, it is worth noting that the number of paths L_k is usually not known in advance, hence the brutal force approach may not be practical even though it is of less complexity under some particular circumstances.

For the NLoS scenario, the worst case takes place when 2 potential optimal beams that corresponds to the 2 NLoS paths are identified on the top layer of the primary codebook, resulting in an iterative search of 2 extra times though the entire codebooks. Thus the complexity is 3 times of that of the single-path scenario, showing no reduction. The best case happens when the potential optimal beams locates at the bottom layer with the same optimal codewords on the bottom layer of the primary codebook. In this case, no extra complexity is involved and the complexity remains the same as that of LoS scenario.

For the extra interval scheme [18], since it only uses one hierarchical codebook, we suppose that it takes the all hierarchical configuration in Table. I. For a fair comparison, it is assumed that it includes an extra search within the intervals of every two adjacent beams only on the bottom layer (LoS) or the top layer (NLoS), in order to keep accordance with the assumption made in this work. Apart from the overhead shown in Table. II, the extra interval scheme requires extra bits to keep a record of the RSS of adjacent beams, in order to determine whether to search in the intervals. In general, a remarkable overhead reduction can be expected with the proposed scheme, compared to that of the brutal force scheme or extra interval scheme [18], even in the worst case.

VI. Simulation Results

The proposed beam training scheme is evaluated trough numerical simulations conducted under the following configurations, unless otherwise mentioned:

- Carrier central frequency $f_c = 60$ GHz
- UPA with 256 antennas: $M = N = 16$
- Number of MSs $K = 8$
- Path gain $\alpha_l \sim \mathcal{N}(1, 0.25)$

TABLE II: Overhead comparison for the multi-path scenario with $L_k = 3$ paths and $K = 4$ MSs.

Schemes/Overhead (time slots)		Proposed (worst)		Proposed (best)		Brutal Force		Extra Interval [18]	
Paths	Hierarchical factor $/\mathcal{F}_A$ size	$Q_A = 4$	$Q_A = 8$	$Q_A = 4$	$Q_A = 8$	$Q_A = 4$	$Q_A = 8$	$Q_A = 4$	$Q_A = 8$
LoS	$Q = 2$	92	140	31	47	72	280	84	172
	$Q = 3$	84	164	32	48	64	272	84	186
NLoS	$Q = 2$	216	420	31	47	216	420	252	456
	$Q = 3$	192	396	32	48	192	396	228	432

- Spatial direction angle $\theta \in (0, \pi/2)$, $\phi \in (0, 2\pi)$
- Threshold for a potential NLOS path $\rho_R = 0.75$
- Size of auxiliary codebook $Q_A = 6$ (representing 5-bit PSs with $K_p = 2^5 = 32$ possible states)
- Hierarchical factor $Q = 2$ and $Q = 3$ to perform a binary or 3-divided hierarchical search

Simulation results are averaged over 3000 test cases.

A. Beamforming performances of the codebooks

In Section IV-D, beam training overhead has been compared between the all hierarchical schemes and the proposed two-stage scheme. Since both schemes work on the same proposed two-stage codebook, they will obtain almost the same beamforming gains, but the proposed scheme can offer a much lower overhead, as shown in Table I. Therefore, here we would like to investigate the beamforming results of the proposed two-stage codebook, and compare them with those of other existing codebooks adopted in industrial standards.

The physical layer standard of IEEE 802.11ad [29] at 60 GHz will be used as a benchmark, which is an exhaustive and sector-based beamforming protocol. Specifically in IEEE 802.11ad, beamforming is completed in two phases: the sector sweep level (SLS) and the beam refinement protocol (BRP) phase. A sector is a set that contains several beams. BS transmits unique frames from each sector and collect feedback from the MSs to determine the optimal sector. Once the optimal sector has been determined, the optimal beam is found with an exhaustive linear search within the sector during the BRP phase. Note that the sector in IEEE 802.11ad can be regarded as the the counterpart of the primary codebook in this paper, while the beams in one sector correspond to the auxiliary codebook. The major difference between the proposed codebooks in this paper and IEEE 802.11ad standard or other similar protocols for mmWave systems such as IEEE 802.15.3c [30], is that the latter usually construct a fixed number of sectors without considering the spatial coverage and the relationship between sectors and beams. For a more complete comparison, the "extra interval" scheme from [18] that also deals with incorrect beam selection, along with the case of analog beamforming using ideal continuous PSs, is presented.

In the simulations, auxiliary codebook $\mathcal{F}_A$ consists of $Q_A = 2 \times 3 = 6$ beams (2 elevation angles and 3 azimuth angles) with inter-beam spacing $\theta = 6^\circ$ and $\phi = 12^\circ$ to meet the 4-bit PS resolution constraint. Spatial coverages of the auxiliary codebook under this configuration are $\Delta\theta = 6^\circ \times (2-1) = 6^\circ$ and $\Delta\phi = 12^\circ \times (3-1) = 24^\circ$. A $Q = 2$ or $Q = 3$ primary codebook for hierarchical search is then constructed, whose bottom layer uses a beam coverage of $\Delta\theta = 5^\circ$ and $\Delta\phi = 20^\circ$ to account for any mis-selection of the optimal beams, giving a margin of $(6^\circ - 5^\circ) \times (24^\circ - 20^\circ) = 1^\circ \times 4^\circ$. The upper layers are generated using the summation of adjacent beams of the lower layer. The benchmark, IEEE 802.11ad beamforming approach, is performed using 8 sectors, which are evenly located over the whole spatial area, with 2 beams in each sector. This configuration indicates $\Delta\theta = 90^\circ/(8 \times 2) = 5.625^\circ, \Delta\phi = 360^\circ/(8 \times 2) = 22.5^\circ$, which is very close to that of the proposed codebooks ($\Delta\theta = 6^\circ$ and $\Delta\phi = 24^\circ$) for a fair comparison. It is assumed that each MSs are evenly distributed over the whole spatial area and power allocation is identical for all MSs.

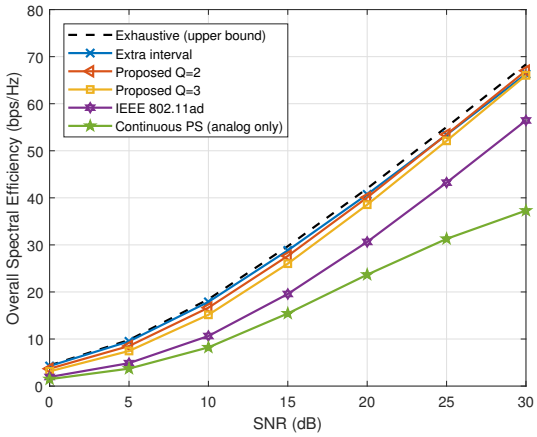


Fig. 13: Beamforming gains for one dominant path $L_k = 1$.

Simulation results of the beamforming gains for different scenarios and structures of the codebooks are shown in Fig. 13-15. The three scenarios considered are: only one dominant path (Fig. 13), 3 LOS paths with a maximum spreading 6° for each MS (Fig. 14), and 3 NLOS paths with spreadings from 6° to 30° (Fig. 15). These scenarios are tested along with two different codebook structures: $Q = 2$ or $Q = 3$ in the simulation. Also, we use an exhaustive search to extract the theoretically optimal beams, which plays the role of the performance upper bound. It can be observed that, the proposed beam training scheme enjoys a spectral efficiency very close to the upper bound obtained by the fully exhaustive search, as well as that by the extra interval scheme, indicating that the proposed scheme selects the actual optimal beams almost for sure.

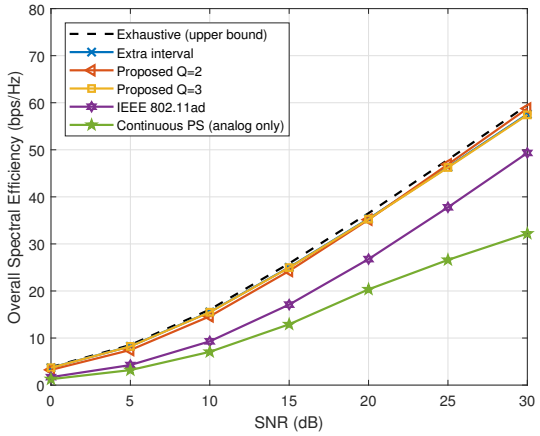


Fig. 14: Beamforming gains for 3 LOS paths $L_k = 3$.

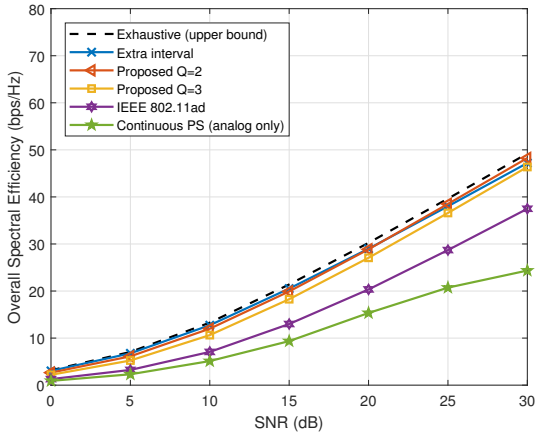


Fig. 15: Beamforming gains for 3 NLOS paths $L_k = 3$ ($\rho_R = 0.75$).

In addition, the proposed codebooks outperform the IEEE 802.11ad codebook remarkably. The reason of this superiority is that the proposed primary codebooks are designed hierarchically with respect to the auxiliary beams and therefore have a better spatial coverage than the IEEE 802.11ad codebook (whose sectors are constructed with fixed number of beams), leading to a more favorable beamforming performance. In particular, for a primary codebook with different numbers of beams on one layer, say $Q = 2$ or $Q = 3$, smaller Q is of more interest. The explanation is that a smaller Q provides a better resolution but requires larger search range on each layer during the hierarchical search. This effect can be expressed in a two-fold way: a better resolution leads to more accurate results but more layers ought to be used, bringing extra overhead, as even less common effective beams can potentially be used for some MSs.

B. Beam Training Efficiency

Next, the efficiency of the proposed beam training scheme is analyzed with numerical examples. Though theoretical upper bound and lower bound of training overhead for the single path scenario have been shown in Table I, for a real scenario training efficiency depends on the number of common effective beams on each layer, as proposed scheme reduces the beam training overhead by eliminating the noneffective beams that no MS selects. For the employed simulation configuration, $S_l = 6 \times 12 = 72$ beams are on the bottom layer of the primary codebook $\mathcal{F}_P$, and the auxiliary codebook $\mathcal{F}_A$ contains $Q_A = 6$ beams. Then for a certain hierarchical factor Q, number of layers is defined by $S = \lceil \log_Q 72 \rceil$. Again, different T_d for the single-path, LOS and NLOS scenarios are tested. For comparison, the IEEE 802.11ad sequential scheme with 8 sectors and 2 beams in each sector, independently for every MS, along with the extra internal scheme [18], is employed as the benchmark. For both LOS and NLOS cases, the IEEE 802.11ad and extra interval schemes assume the number of paths ($L_k = 3$) in advance. Specifically, for the LOS case where the 3 paths are relatively close to each other, it first uses an exhaustive search to select one optimal beam for the dominant path, then executes exhaustive search within the optimal sector and its two adjacent sectors for the other 2 paths. For the NLOS scenario, arbitrary path location is assumed, so both the IEEE 802.11ad and extra interval schemes are executed for 3 times independently.

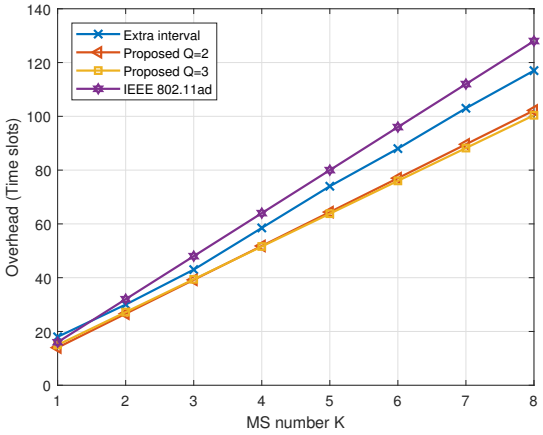


Fig. 16: Training overhead for one dominant path $L_k = 1$.

Numerical training overhead for the single-path, 3 LOS path and 3 NLOS path scenarios are shown in Fig. 16-18. Generally speaking, by eliminating the beams marked as non-effective in the proposed scheme, training overhead can be significantly reduced. For the singe-path scenario, it can be reduced by around 20% and 25% by the proposed approach with $Q = 2$ and $Q = 3$ respectively. The overhead can be further reduced for the multi-path cases, as traditional hierarchical search does not take this into consideration, and has to execute extra full search operations to detect all the paths, within or across sectors (in IEEE 802.11ad) or adjacent beams (in extra interval), depending on the distribution of the paths. For the proposed scheme, extra paths can be detected within the auxiliary codebook for the LOS case, or kept in the hierarchical searching in the primary codebook for the NLOS case. Adaptiveness to the multi-path channel can be obtained for both cases. Note that, for the scenarios where

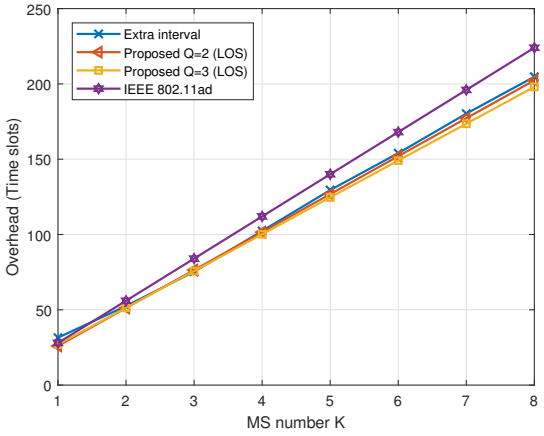


Fig. 17: Training overhead for 3 LOS paths $L_k = 3$.

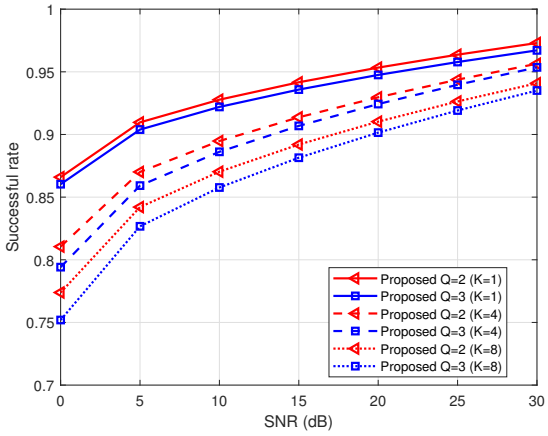


Fig. 19: Beam Training successful rate for one dominant path with different numbers of MSs.

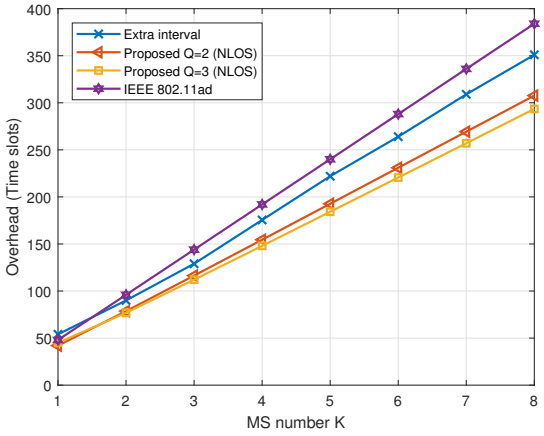


Fig. 18: Beam training overhead for 3 NLOS paths $L_k = 3$ ($\rho_R = 0.75$).

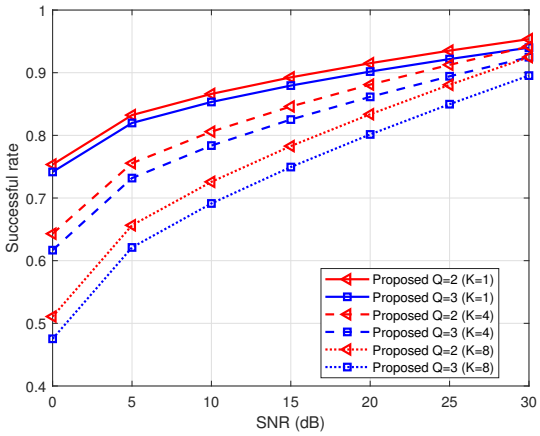


Fig. 20: Beam training successful rate for 3 random paths (LOS or NLOS) with different numbers of MSs ($\rho_R = 0.75$).

MSs are densely located within a small area, (e.g. on the crowded underground or a stadium), more training overhead can be saved as most MSs share a limited number of effective beams.

C. Beam Training Accuracy

The angular correctiveness of the selected beams by the proposed scheme is evaluated by the beam training successful rate. A *successful* beam training is defined as the selected beam falls within the 3dB range of the actual optimal beam. The successful rate is also evaluated for the single-path and multi-path scenarios, with simulation results shown in Fig. 19 and Fig. 20 respectively. The multi-path scenarios generate a channel consisting of 3 arbitrary paths, either LOS or NLOS. Simulation results for the single-path case indicates a high probability that the correct beams are selected. As expected, smaller number of Q is favorable, but the gap can be marginal when SNR is higher, since in this case both $Q = 2$ and $Q = 3$ can give the correct beam training result. When more MSs join in the beam training, it gets more difficult to find out the exact direction of all the MSs at the same time, but this is still very likely to be achieved when SNR is large enough. However for the multi-path, multi-user scenarios, the successful rate deteriorates heavily for both $Q = 2$ and $Q = 3$. In this case, substitution of beams for the LOS paths, as well as the recording of a potential NLOS paths, ought to be performed, resulting the deterioration, especially when SNR is lower and incorrect beam substitution can easily occur.

VII. Conclusion

In this paper we have considered the 3D beam training problem for a multi-user massive FDD MIMO system, and proposed a primary-auxiliary joint codebook, as well as a hierarchical-exhaustive search beam training scheme. The proposed scheme reduces the training overhead for a multi-user system by combining the MSs that share the same optimal beams on the same layer of the codebook. The beam training problem for a multi-path channel with both LOS and NLOS is also considered and solved by beam substitution and trans-layer beam recording. Simulation results have shown a comparable performance to that of independent, exhaustive search for all the MSs, but with a much lower training overhead.

For future work, it would be worthwhile to investigate how to perform a new beam search, with the prior knowledge of previously trained beams.

REFERENCES

[1] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Pearson Education, 2015.

[2] S. Rajagopal, S. Abu-Surra, and M. Malmirchegini, "Channel feasibility for outdoor non-line-of-sight mmwave mobile communication," in *2012 IEEE vehicular technology conference (VTC Fall)*, 2012, pp. 1–6.

[3] H. Zhao, R. Mayzus, S. Sun, M. Samimi, J. K. Schulz, Y. Azar, K. Wang, G. N. Wong, F. Gutierrez, and T. S. Rappaport, "28 ghz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in new york city," in *2013 IEEE International Conference on Communications (ICC)*, 2013, pp. 5163–5167.

[4] 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, 2016.

[5] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave mimo systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, 2016.

[6] L. Zhang, L. Gui, K. Ying, and Q. Qin, "Clustering based hybrid precoding design for multi-user massive mimo systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 12 164–12 178, 2019.

[7] J. Du, J. Li, J. He, Y. Guan, and H. Lin, "Low-complexity joint channel estimation for multi-user mmwave massive mimo systems," *Electronics*, vol. 9, no. 2, p. 301, 2020.

[8] J. Wang, Z. Lan, C.-W. Pyo, T. Baykas, C.-S. Sum, M. A. Rahman, R. Funada, F. Kojima, I. Lakkis, H. Harada *et al.*, "Beam codebook based beamforming protocol for multi-gbps millimeter-wave wpan systems," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, 2009, pp. 1–6.

[9] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.

[10] A. Alkhateeb, G. Leus, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 2909–2913.

[11] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, 2017.

[12] R. J. Weiler, M. Peter, W. Keusgen, and M. Wisotzki, "Measuring the busy urban 60 ghz outdoor access radio channel," in *2014 IEEE International Conference on Ultra-WideBand (ICUWB)*, 2014, pp. 166–170.

[13] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, "Initial access in millimeter wave cellular systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 7926–7940, 2016.

[14] C. U. Bas, R. Wang, S. Sangodoyin, D. Psychoudakis, T. Henige, R. Monroe, J. Park, C. J. Zhang, and A. F. Molisch, "Real-time millimeter-wave mimo channel sounder for dynamic directional measurements," *IEEE Trans. Veh. Technol*, vol. 68, no. 9, pp. 8775–8789, 2019.

[15] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3gpp nr at mmwave frequencies," *IEEE Commun.Surveys Tut.*, vol. 21, no. 1, pp. 173–196, 2018.

[16] W. Wu, D. Liu, Z. Li, X. Hou, and M. Liu, "Two-stage 3d codebook design and beam training for millimeter-wave massive mimo systems," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, 2017, pp. 1–7.

[17] C. Chang, F.-c. Zheng, and S. Jin, "Fast beam training in mmwave multiuser mimo systems with finite-bit phase shifters," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–5.

[18] H. Yu, P. Guan, W. Qu, and Y. Zhao, "An improved beam training scheme under hierarchical codebook," *IEEE Access*, vol. 8, pp. 53 627–53 635, 2020.

[19] M. Li, C. Liu, S. V. Hanly, I. B. Collings, and P. Whiting, "Explore and eliminate: Optimized two-stage search for millimeter-wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4379–4393, 2019.

[20] P. Cao, J. S. Thompson, and H. Haas, "Constant modulus shaped beam synthesis via convex relaxation," *IEEE Antennas Wirel. Propag. Lett.*, vol. 16, pp. 617–620, 2017.

[21] K. Chen, C. Qi, and G. Y. Li, "Two-step codeword design for millimeter wave massive mimo systems with quantized phase shifters," *IEEE Trans. Signal Process.*, vol. 68, pp. 170–180, 2019.

[22] Z. Xiao, P. Xia, and X.-G. Xia, "Channel estimation and hybrid precoding for millimeter-wave mimo systems: A low-complexity overall solution," *IEEE Access*, vol. 5, pp. 16 100–16 110, 2017.

[23] S. Gao, Y. Dong, C. Chen, and Y. Jin, "Hierarchical beam selection in mmwave multiuser mimo systems with one-bit analog phase shifters," in *2016 8th International Conference on Wireless Communications & Signal Processing (WCSP)*, 2016, pp. 1–5.

[24] Z. Xiao, H. Dong, L. Bai, P. Xia, and X.-G. Xia, "Enhanced channel estimation and codebook design for millimeter-wave communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9393–9405, 2018.

[25] K. Chen, C. Qi, O. A. Dobre, and G. Li, "Simultaneous multiuser beam training using adaptive hierarchical codebook for mmwave massive mimo," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[26] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5g networks: Research challenges and future trends," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2181–2195, 2017.

[27] Y. Ghasempour, M. K. Haider, C. Cordeiro, and E. W. Knightly, "Multi-user multi-stream mmwave wlans with efficient path discovery and beam steering," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 12, pp. 2744–2758, 2019.

[28] 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, 2015.

[29] B. Satchidanandan, S. Yau, P. Kumar, A. Aziz, A. Ekbal, and N. Kundargi, "Trackmac: an ieee 802.11 ad-compatible beam tracking-based mac protocol for 5g millimeter-wave local area networks," in *2018 10th International Conference on Communication Systems & Networks (COMSNETS)*, 2018, pp. 185–182.

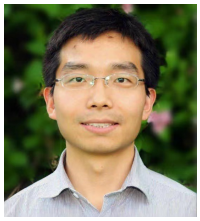
[30] H. Singh, S.-K. Yong, J. Oh, and C. Ngo, "Principles of ieee 802.15. 3c: Multi-gigabit millimeter-wave wireless pan," in *2009 Proceedings of 18th International Conference on Computer Communications and Networks*, 2009, pp. 1–6.

Ke Xu received the B.Eng. degree in communication engineering from the Harbin Institute of Technology, Harbin, China in 2014, and the M.Eng. degree in electrical engineering from the University of California at Los Angeles, CA, USA, in 2016. He is currently pursuing Ph.D. degree with the School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen, China. His research interests include massive MIMO systems and millimeter-wave communications.

Fu-Chun Zheng obtained the B.Eng. (1985) and M.Eng. (1988) degrees in radio engineering from Harbin Institute of Technology, China, and the PhD degree in Electrical Engineering from the University of Edinburgh, UK, in 1992.

From 1992 to 1995, he was a post-doctoral research associate with the University of Bradford, UK, Between May 1995 and August 2007, he was with Victoria University, Melbourne, Australia, first as a lecturer and then as an associate professor in mobile communications. He was with the University of Reading, UK, from September 2007 to July 2016 as a Professor (Chair) of Signal Processing. He has also been a distinguished adjunct professor with Southeast University, China, since 2010. Since August 2016, he has been with Harbin Institute of Technology (Shenzhen), China, as a distinguished professor. He has been awarded two UK EPSRC Visiting Fellowships - both hosted by the University of York (UK): first in August 2002 and then again in August 2006. Over the past two decades, Dr. Zheng has also carried out many government and industry sponsored research projects - in Australia, the UK, and China. He has been both a short term visiting fellow and a long term visiting research fellow with British Telecom, UK. Dr. Zheng's current research interests include ultra-dense networks (UDN), ultra-reliable low latency communications (URLLC), multiple antenna systems, green communications, and machine learning based communications.

He has been an active IEEE member since 1995. He was an editor (2001 – 2004) of IEEE Transactions on Wireless Communications. In 2006, Dr Zheng served as the general chair of IEEE VTC 2006-S, Melbourne, Australia (www.ieeevtc.org/vtc2006spring) - the first ever VTC held in the southern hemisphere in VTC's history of six decades. More recently he was the executive TPC Chair for VTC 2016-S, Nanjing, China (the first ever VTC held in mainland China: www.ieeevtc.org/vtc2016spring).



Pan Cao received the B.Eng. and M.Eng. degrees from Xidian University, China, in 2008 and 2011, respectively, and the Dr. Ing (Ph.D.) degree in electrical engineering from TU Dresden, Germany, in 2015. He worked as a Post-Doctoral Research Associate at the University of Edinburgh from 2015 to 2017 and a visiting researcher at Princeton University in 2017. He has been a Senior Lecturer in electronics and communications at the University of Hertfordshire, U.K., since 2017. His current research interests mainly include millimeter-wave communication, antenna array signal processing. He received the Best Paper Awards of IEEE SPAWC 2012 and International Conference on Machine Learning and Networking 2018, and the Qualcomm Innovation Fellowship Award (QInF) in 2013. He serves as an Associate Editor for the EURASIP Journal on Wireless Communications and Networking since 2017.

Hongguang Xu received the B.Eng. degree in radio technology, the M.Eng. degree in signal and information processing, from the Harbin Institute of Technology, Harbin, China, in 1985 and 1988. He is currently an associate professor of electronic and information at the Harbin Institute of Technology, Harbin, China. His research focuses on the array signal processing, the deep-space communication, the UWB communication, the metal-tag-identification applied in IoT, the heart sound signal collecting and processing, and etc. He has authored or coauthored more that 30 technical papers and patents.

Xu Zhu (S'02-M'03-SM'12) received the B.Eng. degree (Hons.) in Electronics and Information Engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1999, and the Ph.D. degree in Electrical and Electronic Engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2003. She joined the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, U.K., in 2003, as an Academic Member, where she is currently a Reader. She has more than 200 peer-reviewed publications on communications and signal processing. Her research interests include MIMO, channel estimation and equalization, resource allocation, cooperative communications, green communications etc.. She has served as an Editor for the IEEE Transactions on Wireless Communications and a Guest Editor for several international journals such as Electronics. She has acted as a Chair for various international conferences, such as the Vice-Chair of the 2006 and 2008 ICARN International Workshops, the Program Chair of ICSAI 2012, the Symposium Co-Chair of IEEE ICC 2016, ICC 2019 and Globecom 2021, and the Publicity Chair of IEEE IUCC 2016.