Hybrid precoding codebook design in millimetre-wave massive MIMO systems with low-resolution phase shifters

Jingbo Tan | Shiqiang Suo | Haichao Qin

1 Beijing National Research Center for Information Science and Technology (BNRist), Department of Electronic Engineering, Tsinghua University, Beijing, China
2 DaTang Mobile Communications Equipment CO., LTD, Beijing, China

Abstract

Millimetre-wave (mmWave) massive multiple-input multiple-output (MIMO) is one of the promising techniques for 5G wireless communications and beyond. Low-resolution hybrid precoding using low-resolution phase shifters (PSs) is considered to be promising for mmWave massive MIMO, since it can realize an acceptable performance with significantly reduced energy consumption. However, to realize accurate channel state information acquisition, the traditional channel feedback codebooks that quantize the channel with high resolution are not suitable for low-resolution hybrid precoding. To solve this problem, angle-based codebook is proposed here. In the proposed codebook, the analog codebook is designed based on the channel angle-of-departures (AoDs) and the digital codebook is generated by the random vector quantization. Specifically, the analog codewords are optimized by a neighbour search algorithm under the constraint of low-resolution PSs. These analog codewords are designed to be aligned with channel AoDs. In this way, they can remain unchanged in a much larger time scale, since the angle-coherence time is much longer than the channel-coherence time. Therefore, the channel feedback overhead can be significantly reduced. Both theoretical analyses and simulation results illustrate that the proposed codebook can achieve the acceptable achievable rate performance with low channel feedback overhead.

1 | INTRODUCTION

Millimetre-wave (mmWave) massive multiple-input multiple-output (MIMO) has been widely considered as one of the promising techniques to provide a very high data rate for the 5G communications and beyond [1–6]. To realize the multiplexing gain of mmWave massive MIMO systems, channel-adaptive techniques such as precoding should be used. However, the traditional fully digital precoding requires an equal number of radio-frequency (RF) chains as the large antenna number at the base station (BS). This will cause unacceptable power consumption and hardware complexity in mmWave massive MIMO systems. To solve this challenging problem, hybrid precoding has been proposed to significantly reduce the number of RF chains [7–15]. The main idea of hybrid precoding is to decompose the large-size fully digital precoder into a large-size analog beamformer realized by phase shifters (PSs) and a small-size digital precoder realized by a small number of RF chains. Due to the sparsity of mmWave channels, it has been proved that hybrid precoding is able to achieve the near-optimal achievable rate performance [14].

1.1 | Prior works

Accurate channel state information at the transmitter (CSIT) is essential to realize the near-optimal achievable rate performance for massive MIMO systems [16–19]. To this end, various channel feedback schemes have been proposed to realize accurate CSIT for massive MIMO. Specifically, the authors in [20] proposed a distributed compressive sensing-based channel feedback scheme to exploit the common support of the channel to achieve better channel quantization accuracy. In [21], a codebook that linearly combines traditional discrete Fourier transform (DFT) codewords was designed to quantize the channel more accurately. In addition, an adaptive codebook was proposed by quantizing the low-dimensional channel subspace in [22]. These schemes work well for massive MIMO using...
the fully digital precoding. However, the inevitable channel quantization error in these schemes will cause an obvious performance loss for mmWave massive MIMO using hybrid precoding, where the number of RF chains is much smaller. One possible solution to this problem is to directly quantize the optimal precoder to avoid the error propagation caused by the channel quantization error. Specifically, a two-stage hybrid precoding scheme with limited feedback was proposed in [23]. The beamsteering codebook and random vector quantization (RVQ) codebook were used for analog beamformer and digital precoder, respectively [23]. In [24, 25], a hierarchical codebook was designed to form the analog beamformer for hybrid precoding via the deactivation antenna processing and the multi-RF-chain sub-array technique. However, these existing codebooks only consider high-resolution PSs (e.g., 8 bits or more [25]). When large numbers of high-resolution PSs are used for hybrid precoding, the system may suffer from high energy consumption (e.g., 30 – 60 mW for 8 bits PSs in mmWave frequencies [26–29]), and $8 \times 1024 = 8192$ PSs will be used at the BS with 1024 antennas for the fully connected hybrid precoding to serve 8 users [30].

Compared with the high-resolution hybrid precoding utilizing high-resolution PSs (e.g., 8 bits or more), the recently proposed low-resolution hybrid precoding using low-resolution PSs (e.g., 1–2 bits) is attractive due to its acceptable achievable rate performance and higher energy efficiency [31–33]. In [31], the near-optimal low-resolution hybrid precoder was designed by projecting the optimal precoder to the subspace determined by the constraint of low-resolution PSs. In [32], two types of low-resolution hybrid precoding methods which respectively quantize real parts and imaginary parts of the optimal precoder were proposed and analysed. A machine learning-based low-resolution hybrid precoding method was proposed for 1-bit PSs realized by switches and inverters in [33]. Moreover, an iterative phase matching algorithm was introduced to design optimal low-resolution precoder and combiner in [34]. However, all of these low-resolution hybrid precoding schemes [31–34] were designed under the ideal assumption of perfect CSIT at the base station. Unfortunately, the traditional channel feedback schemes [16, 35] which quantize the channel cannot realize the perfect CSIT, especially with the large antenna number in massive MIMO systems. Therefore, the channel quantization error will propagate to the design of low-resolution hybrid precoding, and result in an obvious achievable rate performance loss. In addition, the existing channel feedback schemes in [23, 25] are designed to quantize the optimal precoder with high resolution. They are not suitable for low-resolution hybrid precoding. Therefore, it is of great importance to design a codebook suitable for low-resolution hybrid precoding. This has not been well addressed in the literature to the best of our knowledge.

1.2 Contributions

To fill in this gap, here, we propose an angle-based codebook to realize low-overhead channel feedback for low-resolution hybrid precoding. The main contributions are summarized as follows:

- We propose an angle-based codebook for low-resolution hybrid precoding. Unlike existing codebooks which quantize the channel [16, 21], the proposed codebook directly quantizes the analog beamformer and digital precoder to avoid error propagation. Its basic idea is to design the high-dimensional analog codebook based on channel angle-of-departures (AoDs), while the traditional RVQ codebook is utilized for the low-dimensional digital precoder. To be suitable for low-resolution hybrid precoding, each element of the analog codeword in the proposed codebook satisfies the PSs resolution constraint.
- We formulate an optimization problem to design the high-dimensional analog codebook under the PSs resolution constraint. The target of the optimization problem is to make analog codewords have directional array gain aligned with channel AoDs. A neighbour search algorithm (NSA) is utilized to solve the optimization problem. Under the framework of angle-coherence time [36], the analog codewords designed by solving the optimization problem can remain unchanged for a much longer time than the channel-coherence time. Therefore, the feedback frequency of the analog codewords can be significantly reduced, which leads to a low channel feedback overhead.
- We analyse the required channel feedback overhead and the achievable rate of the proposed angle-based codebook. These analyses show that the proposed angle-based codebook can achieve an acceptable achievable rate performance with a much lower channel feedback overhead. Numerical simulation results illustrate that the proposed angle-based codebook can form narrow and directional beams. Also, it can achieve a better tradeoff between the achievable rate and power consumption with a low channel feedback overhead for low-resolution hybrid precoding in mmWave massive MIMO systems.

1.3 Organization and notations

The remainder of this paper is organized as follows. Section 2 introduces the channel model of mmWave massive MIMO and low-resolution hybrid precoding. The proposed angle-based codebook and the corresponding analysis are provided in Section 3. The numerical simulation results are shown in Section 4. Finally, in Section 5, the conclusions are drawn.

Notation: Lower-case and upper-case boldface letters represent vectors and matrices, respectively; $(\cdot)^T$, $(\cdot)^H$, $\| \cdot \|_F$, and $\| \cdot \|_k$ denote the transpose, conjugate transpose, Frobenius norm, and $k$-norm of a vector or matrix, respectively; $E(\cdot)$ denotes the expectation; $| \cdot |$ denotes the absolute operator; $I_N$ represents the identity matrix of size $N \times N$; $\binom{n}{k}$ denotes the number of $k$-combinations of a set with $n$ elements; $\mathcal{C} \mathcal{N}(\mu, \Sigma)$ and $\mathcal{U}(a, b)$ denote the Gaussian distribution with mean $\mu$ and covariance $\Sigma$, and the uniform distribution between $a$ and $b$. 


respective. Finally, \( A[i, j] \) denotes the element of matrix \( A \) at the \( i \)th row and the \( j \)th column, respectively.

## 2. SYSTEM MODEL

We consider the typical mmWave massive MIMO systems utilizing hybrid precoding in this paper. The BS and the user deploy the widely used uniform linear array (ULA). The antenna number is \( M \) for the BS and \( N \) for the user. Considering hybrid precoding, \( M_{RF} \) RF chains are applied at the BS, where \( M_{RF} < M \) [23]. Therefore, the received signal vector \( y \) of size \( N \times 1 \) at the user can be represented as

\[
y = \sqrt{\rho} H F_{RF} F_{BB} s + n.
\]

The \( N \times M \) matrix \( H = [h_1, h_2, ..., h_N]^T \) denotes the channel in which \( h_\ell \) of size \( M \times 1 \) denotes the channel vector between the \( \ell \)th antenna of the user and the \( M \)-antenna BS. The \( M \times M_{RF} \) matrix \( F_{RF} \) is the analog beamformer realized by \( P \)-bit PSs subject to the PSs resolution constraint, that is, the elements of the analog beamformer \( F_{RF}[i, j], \ i = 1, 2, ..., M, \ j = 1, 2, ..., M_{RF} \), belong to \( \mathcal{P} \) with

\[
F_{RF}[i, j] \in \mathcal{P} = \frac{1}{\sqrt{M}}\{1, e^{\frac{2\pi i}{M-1}}, ..., e^{\frac{2\pi i}{M-1}}\}. \tag{2}
\]

\( F_{BB} \) of size \( M_{RF} \times M \) denotes the baseband digital precoder restricted by the transmit power \( \|F_{RF} F_{BB}\|^2_F = M_s \). \( s \) denotes the transmitted signal vector of size \( M \times 1 \) which satisfies \( E(\langle s s^H \rangle) = \frac{1}{M}I_M \). \( \rho \) presents the average received power. \( n \) is the additive white Gaussian noise vector distributing as \( n \sim \mathcal{CN}(0, \sigma_n^2 I_N) \) with \( \sigma_n^2 \) representing the noise power.

Specifically, we exploit the classical ray-based channel model for the channel matrix \( H \) to describe the mmWave massive MIMO channel characteristics as [1, 2]

\[
H = \sqrt{\frac{MN}{L}} \sum_{\ell=1}^{L} a(\phi_\ell, N) a(\theta_\ell, M) H_\ell, \tag{3}
\]

where \( L \) represents the number of the resolvable paths of the channel, \( a_\ell \) denotes the complex gain of the \( \ell \)th path following the distribution \( \mathcal{CN}(0, 1) \), \( \phi_\ell \) and \( \theta_\ell \) represent the \( \ell \)th path’s spatial AoD and angle-of-arrival (AoA), respectively, with \( \phi_\ell, \theta_\ell \in (-1, 1) \). For instance, \( a(\theta_\ell, M) \) is the steering vector of size \( M \times 1 \) for the spatial AoD \( \theta_\ell \) with the following form

\[
a(\theta_\ell, M) = \frac{1}{\sqrt{M}} \left[1, e^{\frac{2\pi i}{2} \theta_\ell}, ..., e^{\frac{2\pi i (M-1) \theta_\ell}}\right]^T. \tag{4}
\]

where \( d \) is antenna spacing, \( \lambda \) denotes the signal wavelength, and \( \theta_\ell = \sin \theta_\ell \) where \( \theta_\ell \) is the physical AoD. Usually, at mmWave frequencies, the distance \( d \) is set as \( \lambda/2 \) [1].

![FIGURE 1 Hybrid precoding structure for mmWave massive MIMO, where the PSs are realized by: (a) high resolution PSs (e.g. 8 bits or more [25]); (b) low resolution PSs (e.g. 1-bit PSs realized by switches and inverters [33]).](image)

In the existing hybrid precoding structure for mmWave massive MIMO shown in Figure 1a, every element of the analog beamformer \( F_{RF} \) is realized by a PS. As shown in Figure 1a, high-resolution PSs (8 bits or more) are utilized to guarantee good achievable rate performance [14]. Nevertheless, the power consumption of high-resolution PSs is high (e.g. \( 30-60 \) mW for 8 bits PSs in mmWave frequencies [26–29]). This will result in the high power consumption of hybrid precoding, since a large number of PSs will be utilized. To alleviate the high power consumption problem, substituting low-resolution PSs (1–2 bits, e.g. \( P = 1, 2 \)) for high-resolution PSs (8 bits or more, e.g. \( P = 8 \)) is considered to be a feasible solution without an obvious loss in the achievable rate [31–34]. The low-resolution PSs can be realized by power-efficient components, for example, one possible implementation is to use switches and inverters as shown in Figure 1b. The power consumption of one inverter or one switch is merely 5 mW at mmWave frequencies [37, 38]. Therefore, the low-resolution hybrid precoding can realize a better tradeoff between the power consumption and the system achievable rate performance. However, the existing low-resolution hybrid precoding methods such as [33] assume perfect CSIT at the BS. In practical scenarios, due to the large antenna number in massive MIMO systems, the quantization error for channel feedback will cause a serious performance loss for these methods. On the other hand, in existing
The overall design of the angle-based codebook is to directly quantize the hybrid precoder. Considering that the hybrid precoder is composed of two parts, that is, analog beamformer and digital precoder, the proposed codebook also has two parts as analog codebook $\mathcal{A}$ and digital codebook $\mathcal{D}$.

Inspired by the beam steering codebook [23], the analog codebook $\mathcal{A}$ is designed to contain analog codewords that have array gains aligned with different AoDs under PSs resolution constraint. Each codeword quantizes one column of the analog beamformer. In this way, the array gain of each channel path can be sufficiently achieved. Specifically, the analog codebook $\mathcal{A}$ can be designed as

$$\mathcal{A} = \{f_{\delta_1}, f_{\delta_2}, ..., f_{\delta_2^{N_A}} \}.$$  

$N_A$ is the feedback bits of each analog codeword, and $f_{\delta}$ is of size $M \times 1$ is the codeword designed to have the directional array gain aligned with angle $\delta_i = -1 + (2i - 1) \frac{1}{2^{N_A}}$ for $i = 1, 2, ..., 2^{N_A}$. We denote these $\delta_i$ as a set $\Delta_{N_A}$ with $\Delta_{N_A} = \{\delta_1, \delta_2, ..., \delta_{2^{N_A}} \}$. Considering these angles $\delta_i$ distribute
in the angle domain uniformly, the analog codebook can cover the whole angle domain \((-1, 1)\) uniformly with near-optimal array gains.

On the other hand, the digital codebook is designed to contain codewords that directly quantize the digital precoder. In specific, the digital codebook \(D\) can be denoted as

\[
D = \{D_1, D_2, \ldots, D_{2^{N_D}}\}. \tag{10}
\]

\(N_D\) denotes the feedback bits of each digital codeword, and \(D_i\) of size \(M_{RF} \times M_i\) is the \(i\)th digital codebook. Next, we will explain how to design analog codebook \(A\) and digital codebook \(D\) in detail.

### 3.2.2.1 Analog codebook design

The major problem of analog codebook design is how to design codewords \(f_2\) and make \(f_2\) have directional array gain. Such analog codewords are hard to design due to the PSs resolution constraint in low-resolution hybrid precoding. To solve this problem, we formulate an optimization problem under the PSs resolution constraint to design analog codewords. The key idea is to make analog codewords have directional array gains similar to steering vectors. To simplify the expression, we denote \(f_2\) of size \(M \times 1\) as the analog codeword that has the directional array gain on AoD \(\theta\) under PSs resolution constraint.

First, we give a definition of the array gain. Specifically, denote \(\eta(\alpha, M)\) as the array gain of the steering vector \(a(\alpha, M)\) on the AoD \(\alpha \in (-1, 1)\). It can be denoted as

\[
\eta(\alpha, M) = |\bar{\eta}(\alpha, M)^H a(\alpha, M)|
\]

\[
= \frac{1}{M} \left| \sum_{m=1}^{M} e^{j2\pi \frac{m}{M} (\alpha - \alpha)} \right| = \frac{1}{M} |\Gamma(\alpha - \alpha)|, \tag{11}
\]

where \(\Gamma(\alpha) = \frac{\sin \frac{M\pi\alpha}{2}}{\sin \frac{\pi\alpha}{2}}\) is the Dirichlet function [40]. It should be noticed that due to the definition of steering vectors in (4), the array gain has a value range as \(\eta(\alpha, M) \in [0, 1]\). From (11), the following two observations on the array gain of \(a(\alpha, M)\) can be obtained [40] to provide instructions for analog codewords design:

- Due to the power-focusing property of Dirichlet function, the power of \(\eta(\alpha, M, \alpha)\) concentrates on the mainlobe of \(\Gamma(\alpha - \alpha)\) with the range of \(\alpha \in (\theta - \frac{2}{M}, \theta + \frac{2}{M})\).
- On other AoDs \(\alpha \in (-1, 1) - (\theta - \frac{2}{M}, \theta + \frac{2}{M})\), the power leakage of \(a(\alpha, M)\) will induce interference between different analog codewords when they form the analog beamformer \(\mathbf{F}_{RF}\), which leads to achievable rate loss.

Therefore, for a specific angle \(\theta\), the analog codeword which has similar directional array gain as the steering vector should have high array gain on \((\theta - \frac{2}{M}, \theta + \frac{2}{M})\) and low array gain on \((-1, 1) - (\theta - \frac{2}{M}, \theta + \frac{2}{M})\). To this end, we can formulate the design of the analog codeword by maximizing the ratio between the array gain on \((\theta - \frac{2}{M}, \theta + \frac{2}{M})\) and the array gain on \((-1, 1) - (\theta - \frac{2}{M}, \theta + \frac{2}{M})\). Following the above idea, we propose to formulate the design of analog codeword \(f_2\) as an optimization problem

\[
f_2 = \arg \max_{\mathbf{f} \in \mathcal{F}} \frac{||\eta(\Omega)||^2_2}{||\eta(\Omega)||^2_2}, \tag{12}
\]

In (12), the numerator \(||\eta(\Omega)||^2_2\) represents the array gain on the angle range \((\theta - \frac{2}{M}, \theta + \frac{2}{M})\), and the denominator \(||\eta(\Omega)||^2_2\) represents the array gain on the angle range \((-1, 1) - (\theta - \frac{2}{M}, \theta + \frac{2}{M})\). Specifically, \(\mathcal{F}\) denotes the set including all possible analog codewords that satisfies the PSs resolution constraint (2). \(\eta\) represents the array gain vector of \(f\) over the whole angle domain as

\[
\eta = [\eta(f, \alpha_1), \eta(f, \alpha_2), \ldots, \eta(f, \alpha_{\gamma})]^T = \mathbf{G}^H f, \tag{13}
\]

with \(\mathbf{G} = [a(\alpha_1, M), a(\alpha_2, M), \ldots, a(\alpha_{\gamma}, M)]\) and \(\alpha_i = -1 + \frac{2i-1}{T}, i = 1, 2, \ldots, T\). \(\Omega\) denotes the indexes of \(\alpha_i\) which locates in the range \((\theta - \frac{2}{M}, \theta + \frac{2}{M})\) as

\[
\Omega = \{ i | \alpha_i \in (\theta - \frac{2}{M}, \theta + \frac{2}{M}) \}. \tag{14}
\]

It should be pointed out that when \(P = 1\), the elements of array gain vector \(\eta\) is symmetric as \(\eta(f, \alpha_i) = \eta(f, -\alpha_i)\). This is because

\[
\eta(f, \alpha_i) = |a(\alpha_i, M)^H f| = |\text{conj}(a(\alpha_i, M)^H f)|
\]

\[
\overset{(a)}{=} |\text{conj}(a(\alpha_i, M)^H f)| = |a(-\alpha_i, M)^H f| = \eta(f, -\alpha_i), \tag{15}
\]

where (a) comes from the PSs resolution constraint with \(P = 1\) as \(f(\alpha) \in \mathbb{E}_1\), and (b) comes from the form of steering vectors as shown in (4). Therefore, for the case \(P = 1\), we slightly modify the definition of \(\Omega\) to take this symmetry property of the codeword \(f\) into consideration as

\[
\Omega = \{ i | \alpha_i \in (\theta - \frac{2}{M}, \theta + \frac{2}{M}) \cup (-\theta - \frac{2}{M}, -\theta + \frac{2}{M}) \}. \tag{16}
\]

By maximizing the array gain ratio, (12) can guarantee a near-optimal directional array gain of \(f_2\), which is aligned with the AoD \(\theta\). The next question is how to resolve the optimization problem (12). We can find that (12) is a non-convex problem
Inputs:
Antenna number $M$; Target angle $\theta$; PSs bits $P$
Search interval $D$; Iterative number $I$; Grid size $T$

Output:

1. Analog codebook $f_\delta$ with directional array gain aligned with $\theta$
2. Randomly generate $G_1$ with PSs constraint that each element belongs to $\Xi_p$
3. For $i = 0$ to $T$, $i < T$, $i + 1$
4. $\Phi_{D,i} = \{f | f \in F, \|f - x_i\|_0 = D\}$
5. $f_{\text{opt}}(i) = \arg \max_{f \in D, \Xi_p} \|f\|_2^2 / \|f\|_0^2$
6. If $f_{\text{opt}}(i)$ is not equal to $f_{\text{opt}}(i)$, then
7. $f_\delta = f_{\text{opt}}(i)$; break
8. Else
9. $f_{\text{opt}}(i) = f_{\text{opt}}(i)$; $f_\delta = f_{\text{opt}}(i)$
10. End if
11. End for
12. Return $f_\delta$.

Algorithm 1 The neighbour search algorithm (NSA)

Because of the PSs resolution constraint in (2), Note that for the codebook-based hybrid precoding, the codebook is designed in advance so that both the BS and the user know the codebook. Therefore, the exhausted search (ES) algorithm can be adopted to solve (12). However, although the PSs resolution $P$ is small, the number of candidates in $F$ is still quite large (e.g. 2$^{64}$ when $P = 1, M = 64$). Such a large number of candidates will cause unacceptable search complexity. To avoid the high search complexity in the ES algorithm, we utilize a NSA to solve (12) in a near-optimal manner. The main idea of the NSA algorithm is to search the neighbour set iteratively to find out the codebook with the maximum array gain ratio. In the NSA, to make the neighbour set suitable for the low-resolution hybrid precoding, we design the neighbour set elaborately to meet the constraint of low-resolution PSs in (2). Specifically, the neighbour set $\Phi_D$ of the codebook $f \in F$ is defined as

$$\Phi_D = \{f | f \in F, \|f - i\|_0 = D\}.$$  \hfill (17)

(17) means the neighbour set $\Phi_D$ consists of all the codewords which have $D$ elements different from $i$. For example, when $N = 4, P = 1, D = 1$, the neighbour set of the codeword $f = [1, 1, -1, 1]^T$ is

$$\Phi_D = \begin{bmatrix} -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$  \hfill (18)

The pseudo-code of the NSA is shown in Algorithm 1. Specifically, $G$ is first calculated in step 1, and in step 2, the initial codeword $f_{\text{opt}}(0)$ of size $M \times 1$ whose elements belong to $\Xi_p$ is randomly generated. Then, in step 4, a neighbour set $\Phi_{D,i}$ is constructed by selecting all the codewords which have $D$ elements different from $f_{\text{opt}}(i)$ for the $i$th iteration. After that, in step 5, the new near-optimal codeword $f_{\text{opt}}(i)$ is searched out from the neighbour set $\Phi_{D,i}$ by finding the best candidate that maximizes the array gain ratio as shown in (12). Finally, we can obtain the near-optimal codeword $f_\delta$ with directional array gain on $\theta$ when $f_{\text{opt}}(i)$ equals $f_{\text{opt}}(i)$ in step 6, or the iteration number $I$ is approached in step 9. Due to low-resolution PSs, the size of the neighbour set $\Phi_{D,i}$ is small. Under these circumstances, the search range in each iteration in step 5 is limited. Hence, the NSA possesses favourable convergence. The convergence of the NSA will be verified later by simulation results in Section 4.

By utilizing the NSA, the analog codebook $A$ in (9) can be easily obtained. For example, by setting the target angle $\theta = \delta$, we can obtain the $i$th codeword $f_i$ in the analog codebook $A$.

3.2.3 Digital codebook design

After the analog beamformer $F_{\text{RF}}$ has been determined, the optimal digital precoder $F_{\text{BB}}$ can be computed by the singular value decomposition (SVD) principle as discussed in [16]. Specifically, we define $H_{\text{eq}} = HF_{\text{RF}}$ as the equivalent channel and the ordered SVD of $H_{\text{eq}}$ can be defined as $H_{\text{eq}} = U_{\text{eq}} \Sigma_{\text{eq}} V_{\text{eq}}^H$. Then, we have the optimal digital precoder $F_{\text{BB}}$ as

$$V_{\text{eq}} = [V_{\text{eq},1} V_{\text{eq},2}], B = V_{\text{eq},1}, F_{\text{BB}} = \beta B.$$  \hfill (19)

$V_{\text{eq},1}$ is of size $M_{\text{RF}} \times N_s$ and $\beta = \sqrt{M_{\text{RF}}/\|F_{\text{RF}}B\|_F}$ is the power normalization factor. Therefore, the digital codebook $D$ should quantize the digital precoder $F_{\text{BB}}$. Regarding that in hybrid precoding the number of RF chains $M_{\text{RF}}$ is usually much smaller than the number of antennas $M$ (e.g. $M_{\text{RF}} = 64, M_{\text{RF}} = 4$), the digital precoder $F_{\text{BB}}$ is low dimensional. Therefore, to quantize the low-dimensional randomly distributed digital precoder, we utilize the traditional RVQ codebook which can guarantee an acceptable quantization accuracy [16]. Specifically, the digital codebook $D$ can be generated as $D = \{D_1, D_2, ..., D_{N_D}\}$. $D_i$ is an $M_{\text{RF}} \times M_{\text{RF}}$ matrix with all randomly distributed elements, and $N_D$ is the feedback bits of a digital codebook.

3.3 Complexity and performance analysis

Here, we analyse the computational complexity of the NSA algorithm and the ES. From Algorithm 1, we can notice that steps 4 and 5 occupies the major part of the complexity of the NSA. Specifically, we construct the neighbour set $\Phi_{D,i}$ in step 4. Since we select $f \in F$ that has $D$ elements different from $f_{\text{opt}}(i)$, the size of the neighbour set $\Phi_{D,i}$ can be easily obtained.
TABLE 1  Complexity comparison

|   | NSA      | ES          |
|---|----------|-------------|
| $M = 32, P = 1$ | $3.18 \times 10^8$ | $1.37 \times 10^{10}$ |
| $M = 32, P = 2$ | $6.34 \times 10^8$ | $5.90 \times 10^{10}$ |
| $M = 64, P = 1$ | $2.58 \times 10^9$ | $1.18 \times 10^{11}$ |
| $M = 64, P = 2$ | $5.16 \times 10^9$ | $2.18 \times 10^{10}$ |

as $2^P \binom{M}{I}$, $\binom{M}{I}$ denotes the number of $D$-combinations of a set with $M$ elements. In step 5, the codeword $f_{\text{opt}}(i)$ is searched through $2^P \binom{M}{I}$ times array gain calculation of $\eta_i$. For each time of calculating the array gain vector $\eta_i$, $\Theta(M^2T)$ is required. Therefore, the complexity of step 5 is $\Theta(2^PM^2T \binom{M}{I})$. In summary, after $I$ iterations of searching, the total complexity of the NSA $C_{\text{NSA}}$ is

$$C_{\text{NSA}} = \Theta \left( 2^P IM^2T \binom{M}{I} \right).$$

(20)

By contrast, the complexity $C_{\text{ES}}$ of the ES scheme can be denoted as

$$C_{\text{ES}} = \Theta \left( MP^M \right).$$

(21)

Comparing (20) and (21), we can notice that the computational complexity of the NSA increases with the square of the antenna number $M$, while the complexity of the ES is an exponential function of $M$. This indicates that the NSA enjoys much lower complexity when $M$ is large in mmWave massive MIMO systems. Specifically, Table 1 shows the complexity comparison between the NSA and the ES, where we set $T = 200$, $I = 50$, $D = 2$ in the NSA. From Table 1, we can observe that the NSA has much lower complexity than the ES especially for a large antenna number. For example, when $M = 64$, $P = 1$, the searching complexity of the NSA is only $2.58 \times 10^9$. However, the complexity of the ES is $1.18 \times 10^{21}$, which is unacceptable in practice.

It should be pointed out that the selection of the parameter $D$ may simultaneously influence the complexity and performance of the NSA algorithm. Specifically, on the one hand, as illustrated in (20), when $D$ increases, $\binom{M}{I}$ will increase and in contrast the iteration number $I$ will decrease because of the enlarged search range. Hence, the overall impact on the complexity of an increased $D$ may be a larger complexity or a smaller complexity. On the other hand, since the search range becomes larger when $D$ grows up, the NSA may fall into local optimum with a larger probability. This means a large $D$ may cause performance decline to some degree. Therefore, we can conclude that the selection of the parameter $D$ may depend on the system requirements and the system parameters such as antenna number $M$. Specifically, if the system requires a lower complexity, a large $D$ may be promising due to the fast convergence speed. While, if the performance, that is, array gain ratio, is more important to the system, a small $D$ such as $D = 1$ may be better.

An example on how to select the parameter $D$ will be shown in Figure 5 in Section 4.

### 3.4 Channel feedback overhead analysis

Here, we will analyse the channel feedback overhead of the proposed angle-based codebook. Under the assumption that efficient channel estimation is carried out, the channel matrix $H$ and the AoDs $\theta_\ell$ are known at the user [19, 20, 41]. Without loss of generality, we assume that the AoDs are sorted in a descending order according to the power of the path gain $g_\ell$. Then, the analog beamformer $F_{\text{RF}} = [f_1^{\text{RF}}, f_2^{\text{RF}}, \ldots, f_M^{\text{RF}}]$ and the digital precoder $\mathbf{F}_{\text{BB}}$ can be selected from analog codebook $A$ with size of $2^{N_A}$ and digital codebook $D$ with size of $2^{N_D}$, respectively. Specifically, the analog beamformer $F_{\text{RF}}$ is composed of $M_{\text{RF}}$ codewords which are selected out as

$$f_i^{\text{RF}} = \hat{\theta}_i, \hat{\theta}_i = \arg \min_{\delta \in \Delta_{\text{RF}}} |\theta_i - \delta|, i = 2, 3, \ldots, M_{\text{RF}}.$$  

(22)

(23)

In (23), we define $\Delta_{\text{RF}} = \{\hat{\theta}_2, \hat{\theta}_3, \ldots, \hat{\theta}_{M_{\text{RF}}-1}\}$ where $\Delta_{\text{RF}} \subset \Delta_{\text{RF}} \notin \{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{M_{\text{RF}}-1}\}$. In (22) and (23), the analog codewords aligned with the nearest angle to the AoDs $\theta_\ell$ are selected out one by one, to guarantee the near-optimal directional array gain. After the analog beamformer has been determined, the digital precoder can be chosen in digital codebook $D$ to maximize the achievable rate $R$ as

$$\mathbf{F}_{\text{BB}} = \arg \max_{\mathbf{D} \in D} R(F_{\text{RF}}, \mathbf{D}).$$

(24)

After the analog codewords and digital codeword are selected out at the user, their indexes are fed back to the BS through the uplink channel. Then, the hybrid precoding is carried out based on these indexes. Since $M_{\text{RF}}$ analog codewords and one digital codeword are selected, the channel feedback overhead of the proposed angle-based codebook is $M_{\text{RF}}N_A$ bits for the analog beamformer, and $N_D$ bits for the digital precoder. Therefore, the total channel feedback overhead is $M_{\text{RF}}N_A + N_D$ bits. We can observe that the channel feedback overhead consists of two parts. One part is $M_{\text{RF}}N_A$ which is for the selected analog codewords to indicate analog beamformer, and the other part is $N_D$ which is for the selected digital codeword to indicate digital precoder.

We first analyse the channel feedback overhead for analog beamformer $M_{\text{RF}}N_A$. Because the scatters surrounding the BS do not change their positions frequently, the path angles which are mainly determined by these scatters, vary much slower than the path gains [36]. Therefore, the angle-coherence time, during which the AoDs can be regarded as unchanged, is much
longer than the channel-coherence time [42, 43]. For example, [36] shows that the angle-coherence time is over 20 times longer than the channel-coherence time. Recalling that in angle-based codebook the analog codewords are designed to have directional array gains, which can be seen determined by the channel AoDs, the analog codewords selected in (22) and (23) will remain unchanged for a much longer time than the digital codeword. Thus, we can conclude that the average channel feedback overhead can be considered mainly decided by the part of the digital precoder. That is, the average channel feedback overhead of the proposed codebook is not $N_{D}M_{RF}+N_{D}$ bits but close to $N_{D}$ bits.

Fortunately, due to the low-dimensional property of the digital precoder, the channel feedback overhead $N_{D}$ could be very small. Specifically, to achieve common quantization accuracy, the required channel feedback overhead should linearly increase with the dimension of the target matrix of the feedback [17]. For instance, when $M = 64$, $M_{RF} = 2$, $M_{s} = 2$, the dimension of the digital precoder, for example, $M_{RF}, M_{s} = 2 \times 2$, is much smaller than the dimension of massive MIMO channel, for example, $MM_{s} = 64 \times 2$. Consequently, the channel feedback overhead $N_{D}$ can be significantly reduced by utilizing the proposed angle-based codebook. This is because a low-dimensional digital precoder instead of a large-dimensional massive MIMO channel is fed back.

### 3.5 Achievable rate performance analysis

Under the codewords selection principle in (22), (23), and (24), the achievable rate of our proposed angle-based codebook can be analysed. Specifically, the achievable rate $R$ can be rewritten by the SVD of $H$ as

$$R = \log_{2} \left( \left| I_{d} + \frac{\rho}{M_{s}\sigma^{2}} \Sigma_{i}^{2} V_{i}^{H} F_{BB} F_{RF}^{H} V_{i}^{H} \right| \right),$$

where $\Sigma_{i} = \text{diag}([\lambda_{i}, \lambda_{2}, ..., \lambda_{d}]) \in C^{d \times d}$ and $V_{i} \in C^{N \times d}$ with $V_{i}^{H} V_{i} = I_{d}$ comes from the ordered SVD of $H$ as $H = U_{i} \Sigma_{i} V_{i}^{H}$. $d$ denotes the rank of $H$. $\lambda_{i}, i = 1, 2, ..., d$ denotes the $i$th largest singular value of $H$. Without losing generality, we assume that the system parameters for hybrid precoding $(M, M_{RF}, M_{s})$ are dedicatedly designed so that the multiplexing gain from the multi-path channel $H$ can be fully utilized [7]. This assumption, specifically, can be easily satisfied in practical massive MIMO systems as: $M_{s} \leq d \leq M_{RF} \leq M$. Under this assumption, according to [7], (24) can be rewritten as

$$R = \log_{2} \left( \left| I_{M_{s}} + \frac{\rho}{M_{s}\sigma^{2}} \Sigma_{i}^{2} V_{i}^{H} F_{RF}^{H} F_{BB} F_{RF}^{H} V_{i}^{H} \right| \right),$$

where $V = V_{i} \upharpoonright_{1:M_{s}}$ and $\Sigma = \Sigma_{i} |_{1:M_{s}, 1:M_{s}}$. Notice that in (3), the steering vectors of AoDs $a(\theta_{i}, M)$ form an orthogonal basis of $H^{H}H$ due to the approximate orthogonality of steering vectors with large number of antennas [22]. Considering that the columns of $V_{i}$ are also an orthogonal basis of $H^{H}H$, we can approximate the columns of the matrix $V$ by the linear combination of $a(\theta_{i}, M)$ as

$$V \approx A D_{opt}^{H} \tag{27}$$

In (27), $A_{i} = [a(\theta_{i}, M), a(\theta_{2}, M), ..., a(\theta_{M_{RF}} M)]$ with $\theta_{i}$ sorted by $|\xi_{i}| > |\xi_{2}| > ... > |\xi_{M_{RF}}|$, and $D_{opt} \in C^{M_{RF} \times M_{s}}$. By substituting (27) into (26), we can obtain

$$R = \log_{2} \left( \left| I_{M_{s}} + \frac{\rho}{M_{s}\sigma^{2}} \Sigma_{i}^{2} D_{opt}^{H} A_{i} F_{RF} F_{BB} F_{RF}^{H} A_{i}^{H} D_{opt} \right| \right) \tag{28}$$

According to (22) and (23), the analog precoder $F_{RF}$ is composed of the codewords selected from the proposed angle-based codebook. Because the codewords in the analog codebook $A$ are designed to have directional array gains aligned with channel AoDs in (12), we can expect that $f_{i}$ in $A$ is near-optimally directional and approximately orthogonal to each other, just as the steering vectors. This property of the analog codewords in the proposed angle-based codebook will be verified in Section 4. Hence, we can obtain that the matrix $A_{i}^{H} F_{RF}$ is approximately diagonal and satisfies

$$A_{i}^{H} F_{RF} F_{RF}^{H} A_{i} \approx \begin{bmatrix} \eta(f_{i}, \theta_{1})^{2} & 0 & \cdots & 0 \\ 0 & \ddots & \vdots & \vdots \\ 0 & \cdots & \eta(f_{i}, \theta_{M_{RF}})^{2} & \end{bmatrix},$$

where $\eta(f_{i}, \theta_{j})^{2}, i = 1, 2, ..., M_{RF}$ denotes the array gain of $f_{i}$ on the angle $\theta_{j}$. For the analog codebook $A$, the size $N_{D}$ should be sufficiently large such that all the codewords can provide near-optimal array gain over the whole angle domain. This can be easily realized since the feedback of the analog codewords is infrequent based on the discussion in Section 3.4. Thus, we have $\eta(f_{i}, \theta_{1})^{2} \approx \eta(f_{i}, \theta_{2})^{2} \approx \cdots \approx \eta(f_{i}, \theta_{M_{RF}})^{2} \approx E(\eta(f_{i}, \theta)^{2})$. Then, from substituting (29) into (28), we can obtain

$$R \approx \log_{2} \left( \left| I_{M_{s}} + \frac{\rho}{M_{s}\sigma^{2}} E(\eta(f_{i}, \theta)^{2}) \Sigma_{i}^{2} D_{opt}^{H} F_{BB} F_{RF}^{H} D_{opt} \right| \right) \tag{30}$$

Notice that the high effective signal-to-noise rate (SNR) assumption can be easily realized in mmWave system due to the high array gain. (30) can be reformulated according to [7] under high SNR situation as

$$R \approx \log_{2} \left( \left| I_{M_{s}} + \frac{\rho}{M_{s}\sigma^{2}} E(\eta(f_{i}, \theta)^{2}) \Sigma_{i}^{2} \right| \right) - \left( M_{s} - \| D_{opt}^{H} F_{BB} \|_{F}^{2} \right).$$

\[ \tag{31} \]
In (31), we can observe that the achievable rate performance based on our proposed angle-based codebook is decided by two terms. The first one \( \log_2 \left( \frac{1}{M} + \frac{\rho}{M \sigma^2} \mathbb{E} (\eta(f, \theta)^2 \Sigma^2) \right) \) reveals the array gain loss caused by the low-resolution analog code-words, and the second term \( \max \{0, \epsilon \} \) reflects the quantization error of the digital precoder.

We first analyse the performance of the analog codebook, which is mainly determined by the expectation of the array gain achieved by the codewords in the analog codebook \( \mathbb{E} (\eta(f, \theta)^2) \). However, due to the low resolution of the PSs, this expectation cannot be easily computed. To this end, we utilize the array gain of the random generated codeword \( f \) when the antenna number tends to infinity as \( M \to \infty \) to replace \( \mathbb{E} (\eta(f, \theta)^2) \). We denote this array gain as \( \lim_{M \to \infty} \eta(f, \theta)^2 \). As shown in Appendix, Lemma 1 proves the property of \( \lim_{M \to \infty} \eta(f, \theta)^2 \) satisfies

\[
\lim_{M \to \infty} \eta(f, \theta)^2 = \frac{4^p}{\pi^2} \sin^2 \frac{\pi}{2^p},
\]

where \( P \) is the resolution of the PSs. Based on (32), we have

\[
\log_2 \left( \frac{1}{M} + \frac{\rho}{M \sigma^2} \mathbb{E} (\eta(f, \theta)^2 \Sigma^2) \right) \approx \log_2 \left( \frac{1}{M} + \frac{4^p \rho}{\pi^2 M \sigma^2} \sin^2 \frac{\pi}{2^p} \Sigma^2 \right).
\]

(33) reveals that the achievable rate is influenced by the resolution of PSs \( P \), where PSs with lower resolution induces achievable rate loss. However, we can notice from (33) that although the application of low-resolution PSs induces achievable rate loss, this loss is acceptable and maintains asymptotically constant as the number of BS antennas \( M \) increases to infinity.

For the digital precoder, we utilize a RVQ codebook to quantize and feedback. In (24), the digital precoder is chosen to maximize the achievable rate. Under these circumstances, the received signals of \( K \) users \( \tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_K]^T \) can be denoted as

\[
\tilde{y} = \tilde{H}_{RF} F_{BB} \tilde{s} + \tilde{n}.
\]

(35) reflects that the achievable rate achieved by the analog beamformer and digital precoder selected from our proposed angle-based codebook is mainly determined by the resolution of PSs \( P \) and the feedback overhead \( N_D \). On the one hand, the achievable rate loss caused by the low-resolution PSs is limited and acceptable. On the other hand, since the dimension of digital precoder \( M, M_{RF} \) is much smaller than the antenna number \( M \), a small \( N_D \) is able to guarantee a sufficiently small quantization error \( 2^\frac{N_D}{\min(M_{RF}, M)} \). In conclusion, our proposed angle-based codebook for low-resolution hybrid precoding can achieve a better tradeoff between the achievable rate performance and the power consumption. This conclusion will be confirmed by simulation results in Section 4.

3.6 Extension to multi-user scenario

Here, we will introduce how to apply the proposed angle-based codebook into multi-user scenarios. In this case, we consider \( K \) single-antenna users are served by the BS. Without loss of generality, we assume the number of users are equal to the number of RF chains as \( K = M_{RF} \) so that each user is served by one RF chain.

Under these circumstances, the received signals of \( K \) users \( \tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_K]^T \) can be denoted as

\[
\tilde{y} = \tilde{H}_{RF} F_{BB} \tilde{s} + \tilde{n}.
\]

(36) reflects that the achievable rate achieved by the analog beamformer and digital precoder selected from our proposed angle-based codebook is mainly determined by the resolution of PSs \( P \) and the feedback overhead \( N_D \). On the one hand, the achievable rate loss caused by the low-resolution PSs is limited and acceptable. On the other hand, since the dimension of digital precoder \( M, M_{RF} \) is much smaller than the antenna number \( M \), a small \( N_D \) is able to guarantee a sufficiently small quantization error \( 2^\frac{N_D}{\min(M_{RF}, M)} \). In conclusion, our proposed angle-based codebook for low-resolution hybrid precoding can achieve a better tradeoff between the achievable rate performance and the power consumption. This conclusion will be confirmed by simulation results in Section 4.

3.6.1 Extension to multi-user scenario

Here, we will introduce how to apply the proposed angle-based codebook into multi-user scenarios. In this case, we consider \( K \) single-antenna users are served by the BS. Without loss of generality, we assume the number of users are equal to the number of RF chains as \( K = M_{RF} \) so that each user is served by one RF chain.

Under these circumstances, the received signals of \( K \) users \( \tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_K]^T \) can be denoted as

\[
\tilde{y} = \tilde{H}_{RF} F_{BB} \tilde{s} + \tilde{n}.
\]

(35) reflects that the achievable rate achieved by the analog beamformer and digital precoder selected from our proposed angle-based codebook is mainly determined by the resolution of PSs \( P \) and the feedback overhead \( N_D \). On the one hand, the achievable rate loss caused by the low-resolution PSs is limited and acceptable. On the other hand, since the dimension of digital precoder \( M, M_{RF} \) is much smaller than the antenna number \( M \), a small \( N_D \) is able to guarantee a sufficiently small quantization error \( 2^\frac{N_D}{\min(M_{RF}, M)} \). In conclusion, our proposed angle-based codebook for low-resolution hybrid precoding can achieve a better tradeoff between the achievable rate performance and the power consumption. This conclusion will be confirmed by simulation results in Section 4.

3.6.1 Extension to multi-user scenario

Here, we will introduce how to apply the proposed angle-based codebook into multi-user scenarios. In this case, we consider \( K \) single-antenna users are served by the BS. Without loss of generality, we assume the number of users are equal to the number of RF chains as \( K = M_{RF} \) so that each user is served by one RF chain.

Under these circumstances, the received signals of \( K \) users \( \tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_K]^T \) can be denoted as

\[
\tilde{y} = \tilde{H}_{RF} F_{BB} \tilde{s} + \tilde{n}.
\]

(35) reflects that the achievable rate achieved by the analog beamformer and digital precoder selected from our proposed angle-based codebook is mainly determined by the resolution of PSs \( P \) and the feedback overhead \( N_D \). On the one hand, the achievable rate loss caused by the low-resolution PSs is limited and acceptable. On the other hand, since the dimension of digital precoder \( M, M_{RF} \) is much smaller than the antenna number \( M \), a small \( N_D \) is able to guarantee a sufficiently small quantization error \( 2^\frac{N_D}{\min(M_{RF}, M)} \). In conclusion, our proposed angle-based codebook for low-resolution hybrid precoding can achieve a better tradeoff between the achievable rate performance and the power consumption. This conclusion will be confirmed by simulation results in Section 4.

3.6.1 Extension to multi-user scenario

Here, we will introduce how to apply the proposed angle-based codebook into multi-user scenarios. In this case, we consider \( K \) single-antenna users are served by the BS. Without loss of generality, we assume the number of users are equal to the number of RF chains as \( K = M_{RF} \) so that each user is served by one RF chain.

Under these circumstances, the received signals of \( K \) users \( \tilde{y} = [\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_K]^T \) can be denoted as

\[
\tilde{y} = \tilde{H}_{RF} F_{BB} \tilde{s} + \tilde{n}.
\]
array gain. Then, the digital precoder is designed based on the equivalent channel. Specifically, the proposed multi-user hybrid precoding scheme has two stages as follows.

1. Each user select a codeword from the analog codebook $\mathcal{A}$ based on the channel $\mathbf{h}_k$. The label $i_k$ of the selected codeword for the $k$th user can be denoted as

$$i_k = \arg \max_{\alpha \in \{1, 2, \ldots, 2^{N_A}\}} \left\| \mathbf{h}_k \right\|_2$$

Then, each user feed the label $i_k$, $k = 1, 2, \ldots, K$ back to the BS. After that, the BS can determine the analog beamformer by $\mathbf{F}_{\text{RF}} = [\mathbf{f}_{i_1}, \mathbf{f}_{i_2}, \ldots, \mathbf{f}_{i_K}]$.

2. The BS acquires the equivalent channel $\mathbf{H}_{\text{eq}} = \mathbf{HF}_{\text{RF}}$ through traditional channel estimation methods [15]. Then, the digital precoder $\mathbf{F}_{\text{BB}}$ can be designed based on the equivalent channel $\mathbf{H}_{\text{eq}}$ by using existing method, for example, zero forcing [33].

By utilizing the above two-stage hybrid precoding scheme, the array gain can be sufficiently obtained by stage 1 and the inter-user interference can be mitigated by stage 2. In this way, an acceptable achievable rate performance can be expected, which will be verified by simulation results in Section 4.

4 | SIMULATION RESULTS

Numerical simulations are operated to verify the performance of our proposed angle-based codebook obtained by the NSA in this section. The main parameters of the simulations are set as: (1) $(M, N, M_{\text{RF}}, L) = (64, 2, 4, 3)$; (2) The AoDs and AoAs follow the uniform distribution $U(-1, 1)$.

Figure 2 shows the array gain of the codewords designed by the NSA where both 1-bit and 2-bit PSs are considered. The array gain is calculated according to (6). The array gain of the ideal steering vector with infinite-resolution PSs is also shown as the benchmark. For the NSA, the parameters are set as: target angle $\theta = \frac{\sqrt{3}}{2}$, grid size $T = 200$, iterative number $I = 10$, search interval $D = 1$. For comparison, Figure 2a illustrates the array gain of the steering vector $\mathbf{a}(\frac{\sqrt{3}}{2}, 64)$, which can be seen as the benchmark. The array gain of the proposed angle-based codeword with 1-bit and 2-bit PSs are shown in Figure 2b and 2c, respectively. We can observe from Figure 2 that the codewords designed by the NSA with 1-bit and 2-bit PSs can achieve equal beam width together with about 70% and 90% array gain compared to the ideal steering vector with infinite resolution. This reveals that the proposed angle-based codebook can generate narrow enough beams to distinguish different AoDs, and supports the analysis in Sections 3.4 and 3.5. On the other hand, we can notice that when $P = 1$, the codeword may induce interference in direction $\theta = -\frac{\sqrt{3}}{2}$. This interference may cause achievable rate performance loss when there are two users locate in a symmetric location. Fortunately, the probability of this case is quite small due to the narrow beam width, and performance loss caused by this interference can be avoided by efficient user scheduling scheme [44].

In Figure 3, we simulate the convergence of the NSA with different search interval $D = 1, D = 2$, and $D = 3$. Specifically, the array gain ratio of the codeword searched by ES is also shown for comparison. From Figure 3, we can observe that the NSA can converge with a small number of iterations, for example, $I = 10$. The NSA can achieve close array gain ratio performance to ES with $D = 1, D = 2$, and $D = 3$. On the other hand, we can see that with the search interval $D$ growing up, the array gain ratio performance decreases, while the required iteration number $I$ becomes smaller. Hence, to determine how to select $D$, the specific complexity and performance of the NSA with parameters $M = 16, P = 1$ should be considered. Specifically, considering

---

2 Notice that since the equivalent channel $\mathbf{H}_{\text{eq}}$ is low dimensional, the channel estimation overhead is small.
FIGURE 3 The convergence of the NSA with different $D$.

FIGURE 4 Achievable rate performance comparison for the low-resolution hybrid precoding with 1-bit PSs.

DFT-combined codebook [21] and the RVQ codebook [16], the channel feedback overheads to quantize the channel are 7 bits, and low-resolution hybrid precoding method in [33] is carried out. In addition, the high-resolution hybrid precoding with 8-bit PSs [7] and fully digital hybrid precoding with perfect CSIT [1] are also considered as benchmarks for comparison. From Figures 4 and 5, We can observe that the proposed angle-based codebook outperforms both of the DFT-combined codebook and the RVQ codebook with low-resolution hybrid precoding [33], under equal channel feedback overhead. Moreover, the performance gap between the proposed angle-based codebook and the traditional high-resolution hybrid precoding is small, for example, the 1-bit and 2-bit angle-based codebook can approach $85\%$ and $90\%$ rate performance of high-resolution hybrid precoding when SNR is $10$ dB. On the other hand, the performance gap between fully digital hybrid precoding and the angle-based codebook keeps constant when the SNR increases, which is consistent with Lemma 1.

We provide the achievable rate performance of the proposed angle-based codebook versus different antenna number $M$ in Figure 6. The parameters are set the same as those in Figure 4. The performance of fully digital precoding and high-resolution hybrid precoding with perfect CSIT are also illustrated. We can observe from Figure 6 that the achievable rate performance achieved by the proposed angle-based codebook increases as the antenna number grows. Moreover, the performance ratio between the proposed angle-based codebook and fully digital precoding with perfect CSIT maintains constant when the antenna number $M$ increases. This reflects that the proposed angle-based codebook can realize acceptable achievable rate performance with different antenna number $M$.

To further compare the effect of ES and NSA on the codebook design, we provide an achievable rate comparison between the angle-based codebook designed by the ES and the NSA. The parameters are set as $M = 16$ and $P = 1$. The other parameters
are the same as those in Figure 4. We generate two angle-based codebooks by solving the optimization problem in (12) with ES and NSA. From Figure 7, we can see that the ES can obtain a very small performance gain compared with the NSA. Considering the huge complexity of the ES, this small performance gain can be ignored. Therefore, we can conclude that the NSA can achieve nearly the same performance as the ES with a much smaller complexity. The NSA is effective to be utilized to design the proposed angle-based codebook.

Figure 8 illustrates the achievable sum-rate performance of the proposed angle-based codebook in multi-user scenarios. We set $M = 64$, $P = 1$, and $K = 4$. The size of the analog codebook is $N_A = 7$. The overhead for estimating the equivalent channel is set as 12 bits. For the RVQ codebook [16], the channel feedback overhead is also set as 12 bits to quantize the channel, and the low-resolution hybrid precoding method in [33] is carried out at the BS. Just as described in Section 3.4, the channel feedback overhead of the analog codewords can be ignored due to the long angle-coherence time. Hence, the total channel feedback overheads of the proposed angle-based codebook and the RVQ codebook are the same. In addition, fully digital precoding with perfect CSIT is also provided for comparison. We can observe from Figure 8 that the proposed angle-based codebook can achieve much better performance than the RVQ codebook. This is because the RVQ codebook suffers from the propagation error between channel estimation and precoder design. In contrast, the proposed angle-based codebook avoids this performance decline since it directly quantizes and feeds back the precoder. Besides, we also find that the proposed angle-based codebook can approach 70% sum-rate performance of the fully digital precoding with perfect CSIT. The performance gap keeps constant as the SNR increases. This indicates that the proposed angle-based codebook can also realize acceptable achievable sum-rate performance in multi-user scenarios.

Figure 9 illustrates the energy efficiency comparison of low-resolution hybrid precoding with 1-bit PSs using the proposed angle-based codebook, the existing high-resolution hybrid precoding with 8-bit PSs [7] and perfect CSIT, and the fully digital precoding [1] with perfect CSIT. We set $M = 64$, $SNR = 10$ dB, and $M_{RF} = N$ varies from 1 to 16. The parameters of the angle-based codebook and high-resolution hybrid precoding are set the same as those in Figure 4. According to [30], the energy efficiency can be defined as the ratio between the achievable rate and the energy consumption. Specifically, the energy efficiency of fully digital precoding $R_{FD}$, high-resolution hybrid precoding $R_{HH}$, and low-resolution hybrid precoding $R_{LH}$, can be respectively represented as

$$R_{FD} = R_1 + MP_{RF} + P_{BB},$$

(39)
An NSA was utilized to optimize the analog codewords for having directional array gain. Based on the observation that the channel vectors are composed of slow-varying path AoDs and fast-varying path gains, the analog beamformer selected from the proposed angle-based codebook can maintain unchanged for a long time. Thus, the channel feedback overhead can be reduced significantly. Theoretical analysis and numerical simulation results verified that our proposed angle-based codebook can achieve the acceptable achievable rate performance and higher energy efficiency for low-resolution hybrid precoding in mmWave massive MIMO systems, with low channel feedback overhead. The proposed codebook also has the potential to be applied in THz massive MIMO systems [46] and reflectable intelligent surface-assisted systems [47, 48], which may be investigated in the future work.

**ACKNOWLEDGEMENTS**

This work was supported in part by National Key R&D Program of China (Grant No. 2020YFB1805000), in part by the National Natural Science Foundation of China (Grant No. 62031019), and in part by the European Commission through the H2020-MSCA-ITN META WIRELESS Research Project under Grant 956256.

**ORCID**

Jingbo Tan https://orcid.org/0000-0003-4365-7435

**REFERENCES**

1. Mumtaz, S., et al.: Introduction to mmWave massive MIMO. MmWave Massive MIMO Opp. 1–18, Elsevier, Amsterdam (2016)
2. Rusek, F., et al.: Scaling up MIMO: Opportunities and challenges with very large arrays. IEEE Signal Process. Mag. 30(1), 40–60 (2013)
3. Bai, T., et al.: Coverage and capacity of millimeter-wave cellular networks. IEEE Commun. Mag. 52(9), 70–77 (2014)
4. Ge, X., et al.: Joint optimization of computation and communication power in multi-user massive MIMO systems. IEEE Trans. Wireless Commun. 17(6), 4051–4063 (2018)
5. Xiao, Z., et al.: Channel estimation and hybrid precoding for millimeter-wave MIMO systems: A low-complexity overall solution. IEEE Access 5, 16100–16110 (2017)
6. Dai, L., et al.: Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer. IEEE J. Sel. Areas Commun. 37(1), 131–141 (2019)
7. Ayach, O.E., et al.: Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans. Wireless Commun. 13(3), 1499–1513 (2014)
8. Gao, X., et al.: Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays. IEEE J. Sel. Areas Commun. 34(4), 998–1009 (2016)
9. Yu, X., et al.: Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. IEEE J. Sel. Topics Signal Process. 10(3), 485–500 (2016)
10. Wang, B., et al.: Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array. IEEE J. Sel. Areas Commun. 35(10), 2370–2382 (2017)
11. Ramadhan, Y.R., et al.: Hybrid analog-digital precoding design for secrecy mmwave MISO-OFDM systems. IEEE Trans. Commun. 65(11), 5009–5026 (2017)
12. Lee, C.S., Chung, W.H.: Hybrid RF-baseband precoding for cooperative multiuser massive MIMO systems with limited RF chains. IEEE Trans. Commun. 65(4), 1575–1589 (2017)
13. Du, J., et al.: Hybrid precoding architecture for massive multiuser MIMO with dissipation: Sub-connected or fully connected structures? IEEE Trans. Wireless Commun. 17(8), 5465–5479 (2018)
14. Heath, R.W., et al.: An overview of signal processing techniques for millimeter-wave MIMO systems. IEEE J. Sel. Top. Signal Process. 10(3), 436–453 (2016)
15. Xie, T., et al.: Geometric mean decomposition based hybrid precoding for millimeter-wave massive MIMO. China Commun. 15(5), 229–238 (2018)
16. Jindal, N.: MIMO broadcast channels with finite-rate feedback. IEEE Trans. Inf. Theory 52(11), 5045–5060 (2006)
17. Love, D.J., et al.: Grassmannian beamforming for multiple-input multiple-output wireless systems. IEEE Trans. Inf. Theory 49(10), 2735–2747 (2003)
18. Choi, J., et al.: Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory. IEEE J. Sel. Topics Signal Process. 8(5), 802–814 (2014)
19. Gao, Z., et al.: Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels. IEEE Commun. Lett. 20(6), 1299–1302 (2016)
20. Yao, X., et al.: Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems. IEEE Trans. Signal Process. 62(12), 3261–3271 (2014)
21. Choi, J., et al.: Advanced limited feedback designs for FD-MIMO using uniform planar arrays. In: Proceedings of the IEEE Global Communications Conference, San Diego, pp. 1–6, Dec. 2015
22. Shen, W., et al.: Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO systems. IEEE Trans. Commun. 66(11), 5235–5248 (2018)
23. Alkhateeb, A., et al.: Limited feedback hybrid precoding for multi-user millimeter wave systems. IEEE Trans. Wireless Commun. 14(11), 6481–6494 (2015)
24. Xiao, Z., et al.: Hierarchical codebook design for beamforming training in millimeter-wave communication. IEEE Trans. Wireless Commun. 15(5), 3380–3392 (2016)
25. Xiao, Z., et al.: Codebook design for millimeter-wave channel estimation with hybrid precoding structure. IEEE Trans. Wireless Commun. 16(1), 141–153 (2017)
26. Yu, Y., et al.: A 60 GHz phase shifter integrated with LNA and PA in 65 nm CMOS for phased array systems. IEEE J. Solid-State Circuits 45(9), 1697–1709 (2010)
27. Kuo, J., et al.: 60-GHz four-element phased-array transmit/receive system-in-package using phase compensation techniques in 65-nm flip-chip CMOS process. IEEE Trans. Microw. Theory Techn. 60(3), 743–756 (2010)
28. Uzunkol, M., Rebeiz, G.M.: A 65 GHz LNA/phase shifter with 4.3 dB NF using 45 nm CMOS SOI. IEEE Micro. Wireless Compon. Lett. 22(10), 530–532 (2012)
29. Belov, L.A., et al.: Handbook of RF, Microwave, and Millimeter-Wave Components. Artech House, Norwood (2012)
30. Mendez-Rial, R., et al.: Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches? IEEE Access 4, 247–267 (2016)
31. Arai, D.C., et al.: Hybrid beamforming design with finite-resolution phase-shifters for frequency selective massive MIMO channels. In: Proceedings of the IEEE International Conf. Acoustics, Speech, and Signal Processin. New Orleans, pp. 1–6, Mar 2017
32. Lin, Y.P.: On the quantization of phase shifters for hybrid precoding systems. IEEE Trans. Signal Process. 65(9), 2237–2246 (2017)
33. Gao, X., et al.: Machine learning inspired energy-efficient hybrid precoding for mmwave massive MIMO systems. In: Proceedings of the IEEE International Conference on Communications, Paris, pp. 1–6, Mar 2017
34. Wang, Z., et al.: Hybrid precoder and combiner design with low-resolution phase shifters in mmwave MIMO systems. IEEE J. Sel. Topics Signal Process. 12(2), 256–269 (2018)
35. Love, D.J., Heath, R.W.: Limited feedback diversity techniques for correlated channels. IEEE Trans. Veh. Technol. 55(2), 718–722 (2006)
36. Va, V., et al.: The impact of beamwidth on temporal channel variation in vehicular channels and its implications. IEEE Trans. Veh. Technol. 66(4), 5014–5029 (2017)
37. Schmid, R.L., et al.: On the analysis and design of low-loss single-pole double-throw Wband switches utilizing saturated SiGe HBTs. IEEE Trans. Microw. Theory Techn. 62(11), 2755–2767 (2014)
38. Ziegler, V., et al.: Metamorphic pin diodes for high isolation and low power consumption mmwave-channel switching MMICs. Electron. Lett. 36(18), 1556–1557 (2000)
39. He, S., et al.: Codebook-based hybrid precoding for millimeter wave multi-user systems. IEEE Trans. Signal Process. 65(20), 5289–5304 (2017)
40. Gao, X., et al.: Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays. IEEE Trans. Signal Process. 67(18), 4809–4824 (2019)
41. Liao, A., Gao, Z.: Super-resolution channel estimation for mmwave massive MIMO. In: Proceedings of the IEEE International Conference on Communications, Kansas, pp. 1–5, Mar 2018
42. Zhao, J., et al.: Time-varying channel tracking with spatial and temporal BEM for massive MIMO systems. IEEE Trans. Wireless Commun. 17(8), 5653–5666 (2018)
43. Tan, J., et al.: Angle-based codebook for low-resolution hybrid precoding in millimeter-wave massive MIMO systems. In: Proceedings of the Internation Conference on Communications, Qingdao, pp. 1–5, Oct 2017
44. Zhang, J., et al.: Joint user scheduling and beam selection optimization for beam-based massive MIMO downlinks. IEEE Trans. Wireless Commun. 17(4), 2190–2204 (2018)
45. Alkhateeb, A., et al.: Massive MIMO combining with switches. IEEE Wireless Commun. Lett. 5, (3), 232–235 (2016)
46. Tan, J., Dai, L.: Delay-phase precoding for THz massive MIMO with beam split. In Proc. IEEE Global Communications Conference (IEEE GLOBECOM’19), Hawaii, Dec 2019
47. Wei, X., et al.: Channel estimation for RIS assisted wireless communications—part I: Fundamentals, solutions, and future opportunities. IEEE Commun. Lett. 25(5), 1398–1402 (2021)
48. Wei, X., et al.: Channel estimation for RIS assisted wireless communications—part II: An improved method based on double-structured sparsity. IEEE Commun. Lett. 25(5), 1403–1407 (2021)

How to cite this article: Tan J, Suo S, Qin H. Hybrid precoding codebook design in millimetre-wave massive MIMO systems with low-resolution phase shifters. IET Commun. 2021;1–15. https://doi.org/10.1049/cmu2.12229

APPENDIX

Lemma A1. Without losing generality, let $\mathbf{f}$ denote the codeword, and the elements of $\mathbf{f}$ are generated randomly from $\mathbb{S}_P$, that satisfies the $P$-bit PSIs constraint. When $M \to \infty$, array gain of $\mathbf{f}$ has

$$
\lim_{M \to \infty} \eta(\mathbf{f}, \Theta)^2 = \frac{4^P}{\pi^2} \sin^2 \frac{\pi P}{2}. \quad (A1)
$$

Proof. The array gain of $\mathbf{f}$ towards the angle $\Theta$ can be denoted as

$$
\eta(\mathbf{f}, \Theta) = a(\Theta, M)^H \mathbf{f} = \frac{1}{M} \sum_{n=1}^{M} e^{i\Theta_n}, \quad (A2)
$$

where $a(\Theta, M)$ is the array gain of $\mathbf{f}$.
where elements in \( f \) belong to \( \Xi_P \) due to the constraint of P-bit PSs as shown in (2), and \( \beta_m \) represents the phase quantization error. Then, we obtain

\[
\eta(f, \theta)^2 = \frac{1}{M^2} \left| \sum_{m=1}^{M} e^{j \beta_m} \right|^2 \tag{A3}
\]

\[
= \frac{1}{M^2} \left( \left| \sum_{m=1}^{M} \cos(\beta_m) \right|^2 + \left| \sum_{m=1}^{M} \sin(\beta_m) \right|^2 \right),
\]

where (a) is obtained by using Euler's formula. Since the elements in \( f \) belong to \( \Xi_P \), the quantization error \( \beta_m \) can be assumed to follow the uniform distribution \( U(-\pi/2^b, \pi/2^b) \) for \( m = 1, 2, \ldots, M \) \([45]\). Therefore, we have

\[
\lim_{M \to \infty} \eta(f, \theta)^2 = \frac{1}{M^2} \left( \sum_{m=1}^{M} \cos(\beta_m) \right)^2 + \left( \sum_{m=1}^{M} \sin(\beta_m) \right)^2 \tag{A4}
\]

\[
= \mathbb{E} \left( \cos(\beta_m) \right)^2 + \mathbb{E} \left( \sin(\beta_m) \right)^2.
\]

Considering the uniform distribution of \( \beta_m \), we have

\[
\lim_{M \to \infty} \eta(f, \theta)^2 = \frac{4^b}{\pi^2} \sin^2 \frac{\pi}{2^b}, \tag{A5}
\]

which accomplishes the proof of Lemma 1.