Max-min Hierarchical Codebooks for Millimeter-Wave Channel Sensing with Bandit Algorithms

Parthasarathi Khirwadkar, Sina Shahsavari, Mehmet Can Hücümenoğlu, Piya Pal
University of California San Diego
pkhirwad@ucsd.edu, sshahsav@ucsd.edu, mhucumen@ucsd.edu, pipal@ucsd.edu

2023 57th Asilomar Conference on Signals, Systems, and Computers | 979-8-3503-2574-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/IEEECONF59524.2023.10476938

***Abstract*—We consider the problem of millimeter-wave beam alignment using only one RF chain and beamformers implementable with just phase shifters and antenna deactivation. Recent works have proposed Multi-Armed Bandit (MAB) based approaches for fast and efficient beam alignment. In this work, we investigate the role of codebook design in the performance of a specific MAB beam alignment algorithm called HOSUB which exploits the so-called "Unimodality" of rewards from hierarchical codebooks. We study the relationship between unimodality and beampattern. Motivated by this analysis, we design a new hierarchical codebook for HOSUB that leverages our past design of robust analog beamformers. Extensive simulations are presented to showcase the efficacy of the proposed codebook in improving the performance of HOSUB in comparison to other hierarchical codebooks in the literature which obey the same hardware constraints.**

***Index Terms*—Millimeter wave communication, beam alignment, beamforming, codebook design, Multi-Armed Bandits, Hierarchical codebook**

I. Introduction

Millimeter-Wave (mmWave) systems promise to offer expanded signal bandwidth for applications such as 5G, and the evolving 6G technologies [1], [2]. However, mmWave signals suffer from significant path loss, which leads to low Signal to Noise Ratio (SNR) and increased link outage [3]. Despite these challenges, short wavelength allows for the deployment of an array featuring an extensive number of antennas within a relatively confined space [4], enabling massive Multiple-Input Multiple-Output (MIMO) arrays in compact devices.

A massive MIMO mmWave systems can potentially overcome path loss and low SNR challenges with highly directional beams. Achieving effective communication using directional beams requires spatial beam alignment. It is important to note that the problem of beam alignment is an online decision-making problem since the beamformer needs to be determined "on-the-fly" by adaptively sensing the environment (without prior knowledge of the direction of the user). In 5G standard's beam sweeping approach [5], an exhaustive search across all possible directions spatially aligns the beam. Such beam scanning might require the scan time to scale linearly with the number of directions, leading to high pilot overhead for beam training.

A recent line of work employs Bayesian decision-making over hierarchical beamforming codebooks to find the most aligned beam. In particular, in [6] a Hierarchical Posterior Matching (HPM) algorithm is proposed. This algorithm utilizes complete Channel State Information (CSI) to update the posterior probabilities of the incoming beam direction. However, precise and reliable CSI might not be available during beam alignment.

An effective approach for designing adaptive strategies for beam alignment without the knowledge of CSI and in the presence of noise utilizes ideas from Multi-Armed Bandits (MAB) [7]–[9]. Given a codebook $\mathcal{W}$, each beamformer $\mathbf{w} \in \mathcal{W}$ can be considered as an arm of a multi-armed bandit. One possible choice of reward associated with $\mathbf{w} \in \mathcal{W}$ can be the Received Signal Strength (RSS) obtained by sampling the arm $\mathbf{w}$. The problem of finding the beam with maximum RSS can then be cast as finding the arm with the maximum reward in a multi-armed bandit.

One such MAB algorithm for beam alignment is the Hierarchical Optimal Sampling of Unimodal Bandits (HOSUB), which was recently proposed in [10] to find a beamformer with maximum expected RSS in a sample efficient manner. The HOSUB algorithm exploits the so-called "unimodal structure" of rewards associated with a hierarchical codebook, meaning rewards monotonically increase as one moves in the direction of the highest reward. One important consideration is that the success of the HOSUB algorithm is dependent on the extent to which the given codebook satisfies the "unimodal structure".

A conventional approach for designing a hierarchical codebook is to consider a grid of spatial frequencies $\mathcal{G}$ over the region of interest and construct beamformers to probe grid locations $\mathcal{D} \subseteq \mathcal{G}$ using the pseudo-inverse of array response matrix [6], [11]. It is not clear if such codebooks satisfy the unimodal property. Moreover, these beamformers may not be realizable due to physical hardware constraints. To address this issue, the authors in [11], [12], suggested approximating such ideal beamformers using multiple RF chains and phase shifters in a hybrid architecture. However, in the case of large antenna arrays, the installation of an RF chain for each antenna becomes prohibitively expensive, rendering per-antenna digital processing impractical. A cost-effective solution for the practical implementation of large array systems in mmWave bands involves the deployment of a single RF chain.

In this work, our main goal is to design a hierarchical codebook suitable for the HOSUB algorithm using only a single RF chain, phase shifters, and antenna deactivation.

This work was supported in part by grants ONR N00014-19-1-2256, NSF 2124929, and DE-SC0022165.

To do so, we first take a closer look at the relationship between the unimodal structure on which the effectiveness of HOSUB depends, and the properties of beampatterns that need to be satisfied. With this insight, we tailor our max-min optimal beamformers, introduced in our previous work [13], and combine them with the idea of subarray splitting developed in [14].

II. Problem Formulation

Consider a base station (BS) equipped with $N-$antenna uniform linear array (ULA) with $\frac{\lambda}{2}$ spacing, where λ is the carrier wavelength. The $K-$path uplink channel between a single-antenna user and the BS can be written as [11]

$$\mathbf{h} = \sum_{k=1}^{K} \alpha_k \mathbf{a}_N(\sin(\theta_k))$$

where $\{\alpha_k\}_{k=1}^K$ are the complex path gains, $\{\theta_k\}_{k=1}^K \in [-\frac{\pi}{2}, \frac{\pi}{2})$ are the angles of arrival and $\mathbf{a}_N(f) = \begin{bmatrix}1 & e^{j\pi f} & e^{j2\pi f} & \dots & e^{j(N-1)\pi f}\end{bmatrix}^T$ represents the array response (or steering) vector corresponding to the spatial frequency $f = \sin(\theta) \in [-1, 1)$. Each user $u = 1, \cdots, U$, transmits orthonormal pilot symbol $x_u(t)$ i.e., $x_{u'}(t)^H x_u(t) = 1$ if $u = u'$ and zero otherwise. We consider a single Radio Frequency (RF) chain at the BS. Under the aforementioned assumptions and after matched filtering, the received signal from a particular user, at time t, at the output of the RF chain is expressed as [6], [10]

$$y_t = \sqrt{P}\mathbf{w}_t^H \mathbf{h} + \mathbf{w}_t^H \mathbf{n}_t, \quad t = 1, 2, \dots, T$$

where $\sqrt{P}$ is the combined effect of transmit power and large-scale fading, $\mathbf{w}_t$ is the beamforming vector used at time t, T is the length of pilot sequence, and $\mathbf{n}_t \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ is i.i.d complex Gaussian noise at the antennas. The *received signal strength* (RSS) at time t corresponding to the beamformer $\mathbf{w}_t$, is denoted by $r(\mathbf{w}_t) := |\mathbf{w}_t^H(\sqrt{P}\mathbf{h} + \mathbf{n}_t)|^2$.

Given a codebook $\mathcal{W}$, the objective of beam alignment is to find the beamformer $\mathbf{w}^*(\mathbf{h}) \in \mathcal{W}$ corresponding to the channel $\mathbf{h}$ which maximizes the expected RSS:

$$\mathbf{w}^*(\mathbf{h}) = \underset{\mathbf{w}\in\mathcal{W}}{\arg\max}\, \mathbb{E}[r(\mathbf{w})] \tag{1}$$

If the channel $\mathbf{h}$ is known, we can choose the beamformer $\mathbf{w}^*(\mathbf{h}) \in \mathcal{W}$ which is aligned with $\mathbf{h}$. However, the channel is usually unknown and there is no single $\mathbf{w}$ that maximizes RSS for all possible channels. One naive approach to beam alignment is an exhaustive search over the codebook which could be highly sample-inefficient. The design of sample efficient strategies for finding $\mathbf{w}^*(\mathbf{h})$ without channel information and in the presence of noise is a challenging problem. One line of work on designing sample efficient strategies for beam alignment is based on the design of adaptive strategies such as Bayesian posterior matching [6] and Multi-Armed Bandit (MAB) algorithms [7]–[10]. We consider MAB-based approaches over posterior matching since they do not require the knowledge of the channel path gain. In particular, we consider the HOSUB algorithm [10] which is a recently proposed MAB beam alignment strategy. HOSUB exploits a certain reward structure called "Unimodality" [15], [16] in order to find the beamformer with maximum RSS in a sample efficient manner. The effectiveness of HOSUB algorithm depends on the extent to which the hierarchical codebook satisfies this unimodality property.

It is to be noted that analog beamformers can only be implemented with phase shifters and antenna deactivation, i.e. $|[\mathbf{w}]_i| \in \{0, 1\}, i = 1, 2, \dots, N$. These hardware constraints further exacerbate the complexity of the beam alignment problem since it is not clear whether a suitable hierarchical codebook can be designed for HOSUB algorithm that obeys these constraints.

III. Understanding Unimodality

We first take a closer look at the "Unimodality" property required by HOSUB. A hierarchical codebook $\mathcal{W}$ can be associated with a binary tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ defined as follows

$$\mathcal{V} = \{(\ell, n_\ell) \text{ s.t. } \ell \in [0 : L_{\max}], n_\ell \in [1 : 2^\ell]\ \}$$

$$\mathcal{E} = \bigcup_{\ell=0}^{L_{\max}-1} \bigcup_{n_\ell=1}^{2^\ell} \Big\{ \big((\ell, n_\ell), (\ell+1, 2n_\ell - 1)\big), \big((\ell, n_\ell), (\ell+1, 2n_\ell)\big) \Big\}$$

where $\ell = L_{\max}$ is the last level in the binary tree and we use the notation $[a : b] := \{a, a+1, \dots, b\}$. Each vertex $(\ell, n_\ell) \in \mathcal{V}$ corresponds to a unique beamformer $\mathbf{w}_\ell^{n_\ell} \in \mathcal{W}$. In order to focus on the relationship between beamformers and unimodality, we will assume a *single-path channel* i.e., $K = 1$ with $\mathbf{h} = \alpha\mathbf{a}_N(f)$. Considering RSS as the reward function, the expected reward associated with vertex (ℓ, n_ℓ) corresponding to the (ground truth) spatial frequency f is given as

$$R_f(\mathbf{w}_\ell^{n_\ell}) := \mathbb{E}\left[\left|(\mathbf{w}_\ell^{n_\ell})^H(\sqrt{P}\alpha\mathbf{a}_N(f) + \mathbf{n})\right|^2\right]. \tag{2}$$

Here, the expectation is taken with respect to the noise distribution (Note that α and f are deterministic unknowns).

Considering this setup and for a fixed α and P, we first provide the definition for unimodality following [16].

Definition III.1 (Unimodality)**.** *Given a reward function $R_f(.)$, for every $f \in [-1, 1)$ and every node $(\ell_1, n_{\ell_1}) \in \mathcal{V}$, the node with the highest expected reward $(\ell_f^*, n_f^*) = \arg\max_{(\ell,n_\ell)\in\mathcal{V}} R_f(\mathbf{w}_\ell^{n_\ell})$ is unique and there exists a path $p = \big((\ell_1, n_{\ell_1}), (\ell_2, n_{\ell_2}), \dots, (\ell_m, n_{\ell_m})\big)$ where $(\ell_m, n_{\ell_m}) = (\ell_f^*, n_f^*)$ such that*

$$R_f(\mathbf{w}_{\ell_i}^{n_{\ell_i}}) < R_f(\mathbf{w}_{\ell_{i+1}}^{n_{\ell_{i+1}}}) \text{ for } (\ell_i, n_{\ell_i}) \in p,\ 1 \le i \le m-1 \tag{3}$$

Although this definition of unimodality has been used to design MAB algorithms [16], it is difficult to gain insights about the design of hierarchical beamforming codebooks $\mathcal{W}$ that satisfy unimodality. It can be observed that the unimodal property (3) is directly related to beam gain as follows

$$R_f(\mathbf{w}_\ell^{n_\ell}) = P|\alpha|^2|(\mathbf{w}_\ell^{n_\ell})^H\mathbf{a}_N(f)|^2 + \mathbb{E}[|(\mathbf{w}_\ell^{n_\ell})^H\mathbf{n}|^2] \tag{4}$$

Let $B_f(\mathbf{w}_\ell^{n_\ell}) = |(\mathbf{w}_\ell^{n_\ell})^H \mathbf{a}_N(f)|^2$ be the beam gain in the direction corresponding to spatial frequency f. For $\mathbf{n} = \mathbf{0}$ (or high SNR regime), whether (3) is satisfied or not will be mainly determined by the corresponding beam gains. Hence the reward unimodality structure is explicitly tied to the unimodality structure of the beam gains.

We will study guiding principles for a hierarchical codebook to satisfy the unimodality of beam gains in terms of main lobe and sidelobe of neighboring beamformers. Following [14], we define the main lobe (coverage) region of $\mathbf{w}_\ell^{n_\ell}$ as follows (for a fixed $\lambda \in (0,1)$)

$$\mathcal{M}(\mathbf{w}_\ell^{n_\ell}) = \{f' \in [-1,1) |\ B_{f'}(\mathbf{w}_\ell^{n_\ell}) \geq \lambda \max_f B_f(\mathbf{w}_\ell^{n_\ell})\}$$

and the sidelobe region is defined as the complement set of the main lobe, i.e. $\mathcal{S}(\mathbf{w}_\ell^{n_\ell}) = [-1,1) \setminus \mathcal{M}(\mathbf{w}_\ell^{n_\ell})$.

To motivate our insights on the beamformer design, we briefly describe the HOSUB algorithm. The HOSUB algorithm uses a truncated hierarchical codebook consisting of levels $\{\ell_0, \ldots, L_{\max}\}$ where $\ell_0 \geq 1$. The algorithm initially samples every node at level $\ell = \ell_0$ once and updates their sample mean reward. For the remaining steps $t = 2^{\ell_0} + 1, \ldots, T$, the algorithm iterates between the following two steps. Firstly, the algorithm chooses the node with the largest sample mean reward as the leader node $L(t)$. Next, it samples the neighbor of $L(t)$ with the largest Upper Confidence Bound (UCB) on the expected reward, followed by updating the sample mean reward and UCB of the sampled node. This procedure continues until either $L(t)$ is one of the leaf nodes or the sample limit T is reached. Please refer to [10] for a detailed description of the algorithm.

From this description, it is evident that HOSUB tends to select the neighbor of the current leader with the highest expected reward $R_f(\mathbf{w}_\ell^{n_\ell})$ as the next leader. Thus, following (4), the relationship between beam gains $B_f(\mathbf{w}_\ell^{n_\ell})$ of neighboring nodes is an important indicator of the performance of HOSUB.

We now provide guiding principles for beamformer design in terms of the main lobe and sidelobe that are needed to satisfy unimodality. To simplify notations, for vertex $v = (\ell, n_\ell) \in \mathcal{V}$, we use $\mathcal{M}(v)$ to denote $\mathcal{M}(\mathbf{w}_\ell^{n_\ell})$ and $B_f(v)$ to denote $B_f(\mathbf{w}_\ell^{n_\ell})$. For $v = (\ell, n_\ell) \in \mathcal{V}$, let $c_1 = (\ell+1, 2n_\ell - 1)$ and $c_2 = (\ell+1, 2n_\ell)$ be the two children of v. Then, it is desirable to design a beamformer with the following properties for all $v = (\ell, n_\ell), \ell \in [0 : L_{\max} - 1], n_\ell \in [1 : 2^\ell]$.

$$B_f(c_i) > B_f(v) \text{ for all } f \in \mathcal{M}(c_i), i \in \{1,2\} \tag{5}$$

$$B_f(v) > B_f(v') \text{ for all } v, v' \in \mathcal{V}, f \in \mathcal{M}(v) \cap \mathcal{S}(v') \tag{6}$$

These statements can be motivated as follows: The condition (5) helps HOSUB to *converge to the correct leaf node* starting from level ℓ_0 since HOSUB traverses the path of monotonically increasing rewards and hence we need the beam gain to increase monotonically as well. The condition (6) ensures that i) a wrong child does not become the leader over its correct sibling (i.e., the sibling with f in its main lobe) since the side lobe of the wrong sibling will have lower beamgain and hence lower expected reward compared to the main lobe of the correct child, and ii) the algorithm is able to correct itself in the presence of noise.

In the following section, we utilize the proposed conditions in (5) and (6) to evaluate a given hierarchical codebook and suggest suitable modifications.

IV. Towards Designing Unimodal Hierarchical Codebook

Following [14], the design of hierarchical codebooks is simplified by discretizing the region of interest $[-1,1)$ into a uniform grid of size $2^{L_{\max}}$ and spacing between adjacent gridpoints $\delta = \frac{2}{2^{L_{\max}}}$ given by

$$\mathcal{G} = \left\{-1 + \left(n - \frac{1}{2}\right)\delta,\ n = 1, 2, \ldots, 2^{L_{\max}}\right\} \tag{7}$$

The number of grid points is assumed to be equal to the number of antennas i.e., $N = 2^{L_{\max}}$.

Although the conditions in (5) and (6) are applicable to the entire spatial frequency domain, we can restrict the verification of these conditions to the grid $\mathcal{G}$ under the simplified assumption of user being located at one of the grid points. Moreover, we can focus on the design of only $\mathbf{w}_\ell^1$ at each level $\ell \in \{1, \ldots, L_{\max}\}$ since the beampattern of $\mathbf{w}_\ell^{n_\ell}, n_\ell \in \{2, 3, \ldots, 2^\ell\}$ can be obtained by shifting the beampattern of $\mathbf{w}_\ell^1$ by appropriate amounts as done in [14].

A. Max-Min Codebook Design

In our earlier work [13] we proposed a novel max-min criterion for beamformer design using phase shifters and antenna deactivation, which aims to maximize the minimum beam gain over a region of interest $\mathbb{T} := [f_c - \frac{\Delta}{2}, f_c + \frac{\Delta}{2}] \subseteq [-1,1)$ centered at frequency f_c and width Δ as follows

$$\mathbf{w}_{\mathbb{T}}^* := \arg\max_{\mathbf{w} \in \mathcal{F}} \min_{f \in \mathbb{T}} B_f(\mathbf{w}), \quad \text{s.t. } |w_n| \in \{0,1\}, n = 1, 2, \cdots, N \tag{8}$$

where $\mathcal{F}$ is the set of all DFT-type beamformers with varying center frequencies and beamwidths. The Max-min criterion can serve as a useful tool for designing a codebook that satisfies unimodality conditions (5) and (6). Thus, we utilise the beamformer design in [13] to design the Max-min hierarchical codebook by choosing the appropriate region $[f_c - \frac{\Delta}{2}, f_c + \frac{\Delta}{2}]$ for each $\widehat{\mathbf{w}}_\ell^{n_\ell} \in \mathcal{W}_{\text{max-min}}$. For each $(\ell, n_\ell) \in \mathcal{V}$, we choose center frequency $f_c(\ell, n_\ell)$ and width Δ_ℓ as

$$\Delta_\ell = (2^{L_{\max} - \ell} - 1)\delta \tag{9}$$

$$f_c(\ell, n_\ell) = -1 + (n_\ell - 1)(\Delta_\ell + \delta) + \frac{\Delta_\ell + \delta}{2} \tag{10}$$

Note that the width Δ_ℓ is independent of n_ℓ since $\widehat{\mathbf{w}}_\ell^{n_\ell}$ is a modulated version of $\widehat{\mathbf{w}}_\ell^1$ given as follows

$$\widehat{\mathbf{w}}_\ell^{n_\ell} = \widehat{\mathbf{w}}_\ell^1 \circ \mathbf{a}_N(f_c(\ell, n_\ell) - f_c(\ell, 1)) \tag{11}$$

where $\circ$ represents elementwise product of two vectors. Following [13], the beamformer $\widehat{\mathbf{w}}_\ell^{n_\ell}$ is given by

$$\widehat{\mathbf{w}}_\ell^{n_\ell} = \left[\mathbf{a}_{M_\ell}(f_c(\ell, n_\ell))^T \quad \mathbf{0}^T\right]^T \in \mathbb{C}^N \tag{12}$$

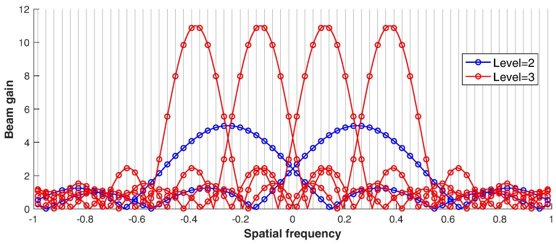


Fig. 1: Max-min codebook design for $N = 64$ antennas

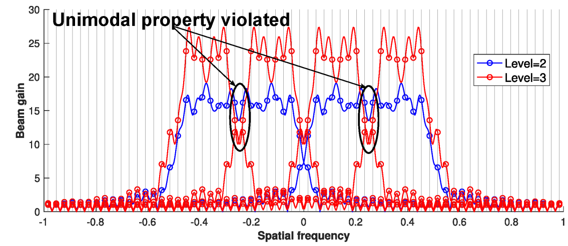


Fig. 2: Codebook design using [14] with $N = 64$ antennas. The locations where (5) is violated are highlighted.

where $M_\ell = \left\lfloor \frac{2}{\max(\gamma_\ell \Delta_\ell, \delta)} \right\rceil$ with $\lfloor x \rceil$ denoting closest integer to x and γ_ℓ is a tuning parameter that allows slight modification to the beamwidth which helps in improving performance of HOSUB.[1] Empirically, we noticed that $\gamma_\ell = 0.8,\ \ell \in [0 : L_{\text{max}} - 2]$ and $M_{L_{\text{max}}-1} = 3N/4$ led to better performance of HOSUB.

We observed that the Max-min codebook satisfies conditions (5) and (6) to a large extent for levels $\ell \geq \ell_0$ where ℓ_0 is sufficiently large, and in the on-grid setting, thereby improving the performance of HOSUB. [2] Figure 1 shows the beampatterns corresponding to levels $\ell = 2$ and 3.

In early levels of the hierarchy, the beampattern is required to cover a wide region in its main lobe. To achieve this, the above Max-min beamformer design deactivates a large number of antennas which leads to a significant reduction in the maximum mainlobe beam gain of the beamformers. This makes HOSUB prone to mistakes in initial timesteps, especially in low SNR regimes. Correcting these early mistakes can be costly in terms of sample complexity. Thus, in the low SNR regimes, the performance of HOSUB is highly dependent on the maximum main lobe beam gain of the early levels of the hierarchy.

To improve the maximum mainlobe beam gain while maintaining the same mainlobe coverage region, Xia et al [14] introduced the idea of splitting the array into multiple subarrays *(Subarray Splitting)* which allows the design of beamformers with wide coverage region without antenna deactivation. Although Xia et al's codebook provides a larger maximum beam gain compared to the Max-min codebook, it violates (5) at certain levels as shown in Figure 2, since the beam gain of both children beamformers falls below the beam gain of their parent in the highlighted regions.

B. Proposed Codebook

We wish to modify the Max-min codebook by ensuring that for each $(\ell, n_\ell) \in \mathcal{V}$, the mainlobe $\mathcal{M}(\mathbf{w}_\ell^{n_\ell})$ of beamformer $\mathbf{w}_\ell^{n_\ell}$ contains the following set of grid points $\mathcal{C}(\ell, n_\ell) \subseteq \mathcal{G}$

$$\mathcal{C}(\ell, n_\ell) = \left\{ f_c(\ell, n_\ell) - \frac{\Delta_\ell}{2} + j\delta,\ j \in [0 : 2^{L_{\text{max}} - \ell} - 1] \right\}$$

[1]At $\ell = L_{\text{max}}$, the width $\Delta_{L_{\text{max}}} = 0$ since the region of interest corresponding to $(L_{\text{max}}, n) \in \mathcal{V},\ n = 1, 2, \ldots, L_{\text{max}}$ contains only $f_c(L_{\text{max}}, n)$ which is the n^{th} gridpoint in $\mathcal{G}$.

[2]We do not satisfy unimodality completely for all levels due to the larger sidelobe of child beamformers compared to their parents and the shape of the beampattern sidelobe is hard to control.

We improve the minimum beam gain over the coverage region $\mathcal{C}(\ell, n_\ell)$ by combining the Max-min criteria [13] with the subarray splitting approach of [14]. This leads to a hierarchical codebook which rectifies the violation of (5) in Xia's codebook and has higher minimum beam gain over the coverage region compared to the Max-min codebook.[3]

Inspired by [14], a beamformer can be split into S subarrays ($2 \leq S \leq N$), each of size $N_S = \lfloor \frac{N}{S} \rfloor$ with parameters $(\mathbf{f}, \boldsymbol{\rho})$ where $\mathbf{f} = (f_1, f_2, \ldots, f_S) \in [-1, 1]^S$ and $\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots, \rho_S) \in \mathbb{R}^S$. Such a beamformer $\mathbf{w}_S(\mathbf{f}, \boldsymbol{\rho})$ is of the form

$$\mathbf{w}_S(\mathbf{f}, \boldsymbol{\rho}) = [e^{j\rho_1} \mathbf{a}_{N_S}(f_1)^T, \ldots, e^{j\rho_S} \mathbf{a}_{N_S}(f_S)^T, \mathbf{0}]^T \quad (13)$$

We design the hierarchical codebook using beamformers of the form shown in equation (13). Further, for each level $\ell = 0, 1, \ldots, L_{\text{max}}$, we describe the design of $\mathbf{w}_\ell^1$ since $\mathbf{w}_\ell^{n_\ell}$ can be obtained as shown in equation (11).

For level ℓ, we are interested in finding the number of subarrays S_ℓ^* and corresponding best parameters $(\mathbf{f}_\ell^*, \boldsymbol{\rho}_\ell^*)$ such that $\mathbf{w}_\ell^1 = \mathbf{w}_{S_\ell^*}(\mathbf{f}_\ell^*, \boldsymbol{\rho}_\ell^*)$ covers the grid points in $\mathcal{C}(\ell, 1)$ in its mainlobe. We propose to choose these parameters based on the Max-min criteria as follows

$$S_\ell^*, (\mathbf{f}_\ell^*, \boldsymbol{\rho}_\ell^*) = \underset{S, (\mathbf{f}, \boldsymbol{\rho})}{\arg\max} \min_{\omega \in \mathcal{C}(\ell, 1)} B_\omega(\mathbf{w}_S(\mathbf{f}, \boldsymbol{\rho})) \quad (14)$$
$$\text{s.t} \min_{\omega \in \mathcal{C}(\ell, 1)} B_\omega(\mathbf{w}_S(\mathbf{f}, \boldsymbol{\rho})) > \max_{\omega' \in \mathcal{G} \setminus \mathcal{C}(\ell, 1)} B_{\omega'}(\mathbf{w}_S(\mathbf{f}, \boldsymbol{\rho}))$$

However, the optimization problem shown in equation (14) is analytically intractable. In this paper, we choose to optimize (14) by discretizing $\mathbf{f}$ and computing the maximum and minimum beam gain only over a finite set of frequencies ω. Note that the optimization for our proposed codebook design is done offline and hence it does not affect the beam training overhead.

For levels $\ell = 0, 1, ..., L_{\text{max}} - 2$, the search of S is restricted to the set $\{2, 3, \ldots, 2^{L_{\text{max}} - \ell} - 1\}$. For levels $\ell = L_{\text{max}} - 1$ and L_{max}, we use the Max-min codebook design since $C(\ell, 1)$ only contains 2 and 1 point respectively. For a given $\ell \in [0 : L_{\text{max}} - 2]$ and $S \in [2 : 2^{L_{\text{max}} - \ell} - 1]$, we further restrict the frequencies $\mathbf{f}$ to the following form parameterized by a single variable m:

$$\mathbf{f} = \mathbf{f}_S(m) := (f_1, f_1 + m\delta, \ldots, f_1 + (S-1)m\delta) \quad (15)$$

[3]The proposed design still does not satisfy unimodality completely at every node. Provably achieving this at every node can be an interesting future direction of work.

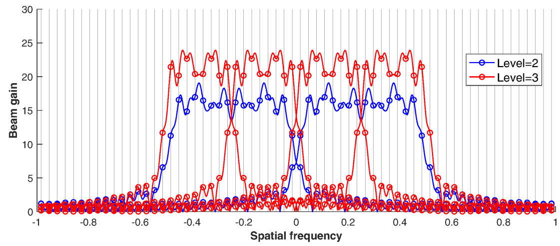
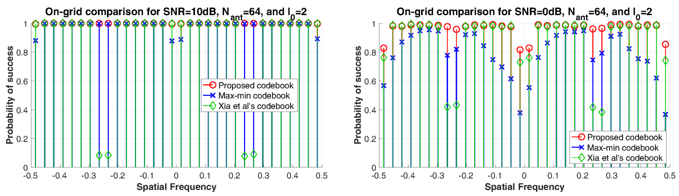


Fig. 3: Proposed codebook design for $N = 64$ antennas. The unimodality violations of Xia et al's codebook [14] have been rectified in our codebook design.



(a) SNR=10dB

(b) SNR=0dB

Fig. 4: Probability of success of HOSUB with fixed SNR and user location varying over the grid $\mathcal{G}$

where $f_1 = f_c(\ell, 1) - (\frac{S-1}{2})m\delta$, $m \in [1 : \zeta_S]$ and $\zeta_S = \left\lfloor \frac{2^{L_{\max}-\ell}-2}{S-1} \right\rfloor$ is chosen such that the frequencies lie within the region $[f_c(\ell, 1) - \frac{\Delta_\ell}{2}, f_c(\ell, 1) + \frac{\Delta_\ell}{2}]$.

Given the set of frequencies $\mathbf{f}$, the phases $\boldsymbol{\rho}$ are computed by setting $\rho_1 = 0$ and choosing

$$\rho_i = \rho_{i-1} + \pi \left(f_{i-1} N_S + \frac{f_i - f_{i-1}}{2} \right), i \in [2 : S] \quad (16)$$

where ρ_i expression is obtained by maximizing $|\mathbf{w}_S(\mathbf{f}, \boldsymbol{\rho})^H \mathbf{a}_N(\frac{f_{i-1}+f_i}{2})|$ using the approximation mentioned in [14].

Therefore, for a given $\ell \in \{0, 1, \ldots, L_{\max} - 2\}$, we consider all possible $S \in [2 : 2^{L_{\max}-\ell} - 1]$ and $m \in [1 : \zeta_S]$, and for each (S, m) pair we compute $B_{\min}(S, m) = \min_{\omega \in C(\ell,1)} B_\omega(\mathbf{w}_S(\mathbf{f}_S(m), \boldsymbol{\rho}))$ and $B_{\max}(S, m) = \max_{\omega \in \mathcal{G} \setminus C(\ell,1)} B_\omega(\mathbf{w}_S(\mathbf{f}_S(m), \boldsymbol{\rho}))$ [4] where $\boldsymbol{\rho}$ is obtained from (16). We obtain (S_ℓ^*, m_ℓ^*) by solving

$$\begin{aligned}(S_\ell^*, m_\ell^*) = &\arg\max_{(S,m)} B_{\min}(S, m) \\ &\text{s.t } B_{\min}(S, m) > B_{\max}(S, m)\end{aligned}$$

We choose $\mathbf{f}_\ell^* = \mathbf{f}_{S_\ell^*}(m_\ell^*)$ and $\boldsymbol{\rho}_\ell^*$ is obtained from $\mathbf{f}_\ell^*$ using equation (16). Finally, we obtain $\mathbf{w}_\ell^1 = \mathbf{w}_{S_\ell^*}(\mathbf{f}_\ell^*, \boldsymbol{\rho}_\ell^*)$.

V. Simulations

In order to validate the effectiveness of the proposed codebook, we compare the probability of success of HOSUB using the proposed codebook, the codebook from [14] (referred to as Xia's codebook), and Max-min codebook [13]. We consider $N = 64$ antennas at the BS, and a single-path channel ($K = 1$) with $\alpha = 1$ and noise power $\sigma^2 = 1$. The SNR is defined as $\text{SNR} = 10\log_{10}(\frac{|\alpha|^2 P}{\sigma^2})$. We simulate an on-grid setting in which the user is located at one of the grid points in $\mathcal{G}$ and HOSUB succeeds if it converges to the leaf node

[4] For $\ell = 0$, $C(0, 1) = \mathcal{G}$ and hence we set $B_{\max}(S, m) = 0$

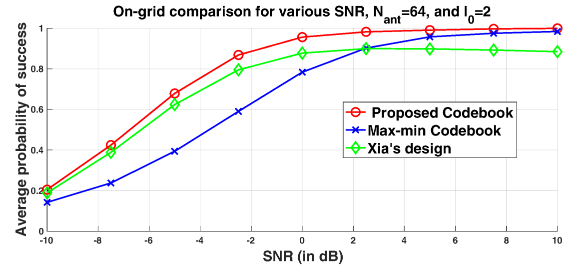


Fig. 5: Average probability of success of HOSUB across the grid $\mathcal{G}$ with varying SNR

corresponding to the user location. HOSUB is simulated with truncated hierarchical codebooks starting from $\ell_0 = 2$ for all codebooks.

Figures 4a and 4b illustrate the performance of HOSUB at fixed SNR levels with the user location varying across the grid. In both figures, we observe that the proposed codebook performs better than the other two codebooks across the grid. In Fig. 4a, we see that HOSUB with Xia's codebook fails at grid locations where the codebook does not satisfy (5). On the other hand, Max-min codebook and the proposed codebook perform better than Xia's codebook uniformly over all grid locations. At SNR=0 dB (Fig. 4b), HOSUB's performance with the Max-min codebook degrades more than the proposed codebook due to low beam gain.

Figure 5 shows the average success probability of HOSUB (over the entire grid) with varying SNR. Our proposed codebook consistently performs better on average compared to the other two codebook designs. The success probability of HOSUB with Xia's codebook saturates around 0.9 at high SNR since the codebook does not satisfy unimodality at certain grid points. On the other hand, the performance of HOSUB with Max-min codebook deteriorates quickly with decreasing SNR. The proposed codebook design combines the advantages of Max-min and Xia's codebook to achieve a higher success rate over the entire considered SNR regime.

VI. Conclusion

In this work, we took a closer look into the effectiveness of a Multi-Armed Bandit algorithm, namely HOSUB, for mmWave beam alignment. We provided certain physically meaningful conditions on the beampattern for achieving the unimodality of hierarchical codebooks associated with HOSUB, which serves as a guiding principle for codebook design. By utilizing these insights and combining them with subarray splitting and robust beamformer design, we propose a suitable hierarchical codebook for the HOSUB algorithm which is *restricted to only one RF chain and can be implemented using phase shifters*. In future work, we plan to further optimize and analyze the codebook design so that unimodality is provably satisfied at every node of the tree.

References

[1] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5g: opportunities and challenges," *Wireless networks*, vol. 21, pp. 2657–2676, 2015.

[2] W. Hong, Z. H. Jiang, C. Yu, D. Hou, H. Wang, C. Guo, Y. Hu, L. Kuai, Y. Yu, Z. Jiang, Z. Chen, J. Chen, Z. Yu, J. Zhai, N. Zhang, L. Tian, F. Wu, G. Yang, Z.-C. Hao, and J. Y. Zhou, "The role of millimeter-wave technologies in 5g/6g wireless communications," *IEEE Journal of Microwaves*, vol. 1, no. 1, pp. 101–122, 2021.

[3] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5g) wireless networks—with a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.

[4] X. Wang, L. Kong, F. Kong, F. Qiu, M. Xia, S. Arnon, and G. Chen, "Millimeter wave communication: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1616–1653, 2018.

[5] M. Giordani, M. Mezzavilla, C. N. Barati, S. Rangan, and M. Zorzi, "Comparative analysis of initial access techniques in 5g mmwave cellular networks," in *2016 Annual Conference on Information Science and Systems (CISS)*, 2016, pp. 268–273.

[6] S.-E. Chiu, N. Ronquillo, and T. Javidi, "Active learning and csi acquisition for mmwave initial alignment," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2474–2489, 2019.

[7] M. Hashemi, A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 2393–2401.

[8] W. Wu, N. Cheng, N. Zhang, P. Yang, W. Zhuang, and X. Shen, "Fast mmwave beam alignment via correlated bandit learning," *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5894–5908, 2019.

[9] Y. Wei, Z. Zhong, and V. Y. F. Tan, "Fast Beam Alignment via Pure Exploration in Multi-Armed Bandits," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3264–3279, 05 2023.

[10] N. Blinn, J. Boerger, and M. Bloch, "mmwave beam steering with hierarchical optimal sampling for unimodal bandits," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.

[11] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

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

[13] S. Shahsavari, P. Sarangi, and P. Pal, "Beamspace esprit for mmwave channel sensing: Performance analysis and beamformer design," *Frontiers in Signal Processing*, vol. 1, 2022.

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

[15] J. Y. Yu and S. Mannor, "Unimodal bandits," in *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ser. ICML'11, 2011, p. 41–48.

[16] R. Combes and A. Proutiere, "Unimodal bandits: Regret lower bounds and optimal algorithms," in *International Conference on Machine Learning*. PMLR, 2014, pp. 521–529.