2015

Multi-Resolution Codebook and Adaptive Beamforming Sequence Design for Millimeter Wave Beam Alignment

Song Noh
*Purdue University, songnoh@purdue.edu*

Michael D. Zoltowski
*Purdue University, michael.d.zoltowski.1@purdue.edu*

David J. Love
*Purdue University, djlove@purdue.edu*

Follow this and additional works at: http://docs.lib.purdue.edu/ecetr

Part of the [Electrical and Computer Engineering Commons](http://docs.lib.purdue.edu/ecetr)

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Multi-Resolution Codebook and Adaptive Beamforming Sequence Design for Millimeter Wave Beam Alignment

Song Noh, Michael D. Zoltowski, and David J. Love

Abstract—Millimeter wave (mmWave) communication is expected to be widely deployed in fifth generation (5G) wireless networks due to the substantial bandwidth available at mmWave frequencies. To overcome the higher path loss observed at mmWave bands, most prior work focused on the design of directional beamforming using analog and/or hybrid beamforming techniques in large-scale multiple-input multiple-output (MIMO) systems. Obtaining potential gains from highly directional beamforming in practical systems hinges on sufficient levels of channel estimation accuracy, where the problem of channel estimation becomes more challenging due to the substantial training overhead needed to sound all directions using a high-resolution narrow beam. In this work, we consider the design of multi-resolution beamforming sequences to enable the system to quickly search out the dominant channel direction for single-path channels. The resulting design generates a multilevel beamforming sequence that strikes a balance between minimizing the training overhead and maximizing beamforming gain, where a subset of multilevel beamforming vectors is chosen adaptively to provide an improved average data rate within a constrained time. We propose an efficient method to design a hierarchical multi-resolution codebook utilizing a Butler matrix, a generalized discrete Fourier transform (DFT) matrix implemented using analog RF circuitry, which extends to three-dimensional beamforming. Numerical results show the effectiveness of the proposed algorithm.

Index Terms—Millimeter wave, MIMO, multi-resolution beam alignment, 3D beamforming, Butler matrix

I. INTRODUCTION

The use of millimeter wave (mmWave) bands is considered a key technology for future wireless cellular communications due to the tens of GigaHertz of available licensed and unlicensed spectrum bands in the range of 20–100 GHz. The propagation properties of each of the mmWave bands are characterized by severe path loss and sparse scattering conditions [2]. Compared to the frequency bands used in most current wireless systems, directional beamforming design using large-scale MIMO systems is crucial to overcome the higher path loss at mmWave bands [3], [4].

Due to the radio frequency (RF) hardware limitations found in mmWave implementations, there is growing interest in the design of channel estimation and precoding algorithms [5]. To reduce hardware complexity, the works in [6]–[9] propose analog beam alignment techniques in which phase shifters or discrete lens arrays are employed. To provide increased multiplexing and beamforming gains while using a limited number of RF chains, [10]–[14] present the design of hybrid beam alignment using a combination of analog and digital beamforming.

Obtaining potential gains from highly directional beamforming using a large number of antennas hinges on sufficient levels of channel estimation accuracy. In large-scale MIMO, much of the prior work exploits a priori knowledge of the underlying channel statistics such as the channel covariance matrix and the spatial sparsity along with a specific antenna arrangement [15]–[20]. Without the prior information, which requires additional learning algorithms [16], [21]–[23], the problem of channel estimation in the mmWave case becomes more challenging because narrow pencil beams capable of supporting high beamforming gain should be aligned to dominant channel directions located randomly within a cell.

To tackle the problem of channel estimation in a mmWave system, there has been some work on channel estimation using a hierarchical multi-resolution beam codebook design [8], [10], [11], [24]. When the angles of departure or arrival (AoDs/AoAs) of the base station and the user are divided into multiple partitions, the beamforming vectors are sounded at increasing angular resolution to further divide either azimuth or elevation angular regions of interest. Ideally, the beamforming vectors should have beamwidths that grow narrow as the search level increases.

Since the entire area spanned by a beam pattern is bounded given a fixed transmit power, a wide beamwidth beamformer (i.e., low angular resolution) has relatively lower main-lobe directivity gain than a narrow beamwidth beamformer capable of supporting high angular resolution. Therefore, the probability of successfully estimating the AoA/AoD ranges of the previous level becomes lower and lower, and the initial estimation of angular partition can be the bottleneck to minimize the beam misalignment probability.

Most prior work on multilevel beamforming [8], [10], [11], [24] is premised on similar estimation performance at each training stage, which is possible at high signal-to-noise ratio (SNR). When the total signal power used for the training period is large, power allocation can
improve the estimation performance during early training stages [11], which requires an increased dynamic range for the power amplifiers. Since the performance achievable using a multilevel beamforming framework is closely related to both the training overhead and the beamforming gain, it is critical to understand how to design the multi-resolution beamforming sequences to improve the data rate performance.

In this work, we consider the design of a multilevel beamforming sequence that properly allocates a multi-resolution codebook to the training phase using a fixed signal power. Given a set of multi-resolution codebooks, a subset of the available codebooks is adaptively selected to provide an improved average data rate within a constrained time. We propose an efficient method to design a hierarchical multi-resolution codebook utilizing a Butler matrix implemented using analog RF circuitry. We also show that this codebook design framework can be extended to three-dimensional (3D) beamforming.

This manuscript is organized as follows: The system model and background are described in Section II. Section III describes the proposed multilevel beamforming sequence design method. A DFT-based codebook design is discussed in Section IV. Numerical results are provided in Section V, followed by conclusions in Section VI.

Notation
Vectors and matrices are written in boldface with matrices in capitals. All vectors are column vectors. For a matrix $A$, $A^H$ and $A^T$ indicate the transpose and Hermitian transpose of $A$, respectively. $[A]_{i,:}$ denotes the $i$-th row of $A$. $\text{diag}(d_1, \ldots, d_n)$ is the diagonal matrix composed of elements $d_1, \ldots, d_n$. For a vector $a$, $a(i)$ denotes the smallest integer greater than or equal to $a$ and the largest integer less than or equal to $a$, respectively. $E\{\cdot\}$ represents statistical expectation. $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$ denote the sets of natural, real, and complex numbers, respectively.

II. SYSTEM MODEL

A. System Setup

We consider a mmWave MIMO link composed of uniform linear arrays (ULAs) of $N_t$ transmit antennas and $N_r$ receive antennas. We assume block transmission with $K$ successive symbols for one block where the block length $K$ is smaller than the channel coherence time, and $K_p$ out of $K$ channel uses are set aside for channel sounding. For a unit norm transmit beamforming vector $f_k \in \mathbb{C}^{N_t}$, the received signal $y_k \in \mathbb{C}^{N_r}$ at a discrete symbol time $k$ is given by

$$ y_k = \sqrt{P}H_k s_k + z_k, \quad \text{for } k = 1, 2, \ldots, K $$

where $H \in \mathbb{C}^{N_r \times N_t}$ denotes the MIMO channel matrix, an $N_r \times 1$ Gaussian noise vector $z_k \sim \mathcal{CN}(0, \sigma_z^2 I_{N_r})$, $P$ denotes the transmit power, and $s_k \in \mathbb{C}$ represents either the pilot symbol during the training period (i.e., $1 \leq k \leq K_p$) or the data symbol during the data transmission period ($K_p < k \leq K$) with $\mathbb{E}\{|s_k|^2\} = 1$.

In (1), the transmitter sounds $K_p$ beamforming vectors and chooses a vector from them that achieves the largest received signal power, referred to as hard beam alignment [8]. The selected beamformer is then used for data transmission (i.e., channel uses satisfying $K_p < k \leq K$).

Transmission over mmWave frequencies using narrow transmit and receive beams can limit the multipath components of the effective mmWave channel [25]. Thus, we adopt a geometric-based channel model with $L$ scatterers [26] where the channel matrix $H$ for a ULA is given by

$$ H = \sqrt{N_t N_r} \sum_{l=1}^{L} g_l u_r(\theta_l^r) u_t^H(\theta_l^t), $$

where $u_r(\cdot)$ and $u_t(\cdot)$ denote the transmitter steering vector and the receiver response vector in the normalized steering phases, i.e.,

$$ u_r(\theta_l^r) = \frac{1}{\sqrt{N_s}} \begin{bmatrix} 1, e^{-j2\pi \frac{d}{\lambda} \sin(\theta_l^r)}, \ldots, e^{-j(N_r-1)2\pi \frac{d}{\lambda} \sin(\theta_l^r)} \end{bmatrix}^T, $$

where $j = \sqrt{-1}$, $s \in \{t, r\}$, $\lambda$ is the carrier wavelength, and $d$ is the antenna spacing given by half-wavelength spacing $d = \lambda/2$. For the possible physical angles $\phi_s^r \in (-\pi/2, \pi/2)$ that cover the entire (one-sided) spatial horizon, we define the spatial frequency $\theta_l^s \in (-1, 1)$ as $\theta_l^s = \sin(\phi_s^r)$ [27] where there is a one-to-one mapping between $\theta_l^s$ and $\phi_s^r$. Denote the complex small-scale fading gain of the $\ell$-th subpath by $g_{\ell}$ modeled as independent and identically distributed (i.i.d.) complex Gaussian, i.e., $g_{\ell} \sim \mathcal{CN}(0, \sigma_g^2)$ [25].

We first consider the case of quantized AoAs/AoDs assuming $\theta_l^s \in \{-1 + (2n - 1)/N_s : 1 \leq n \leq N_s\}$ for $s \in \{t, r\}$ as in [10], [11] and then extend to the case of continuous angles in Sections IV-A and V. In this scenario, the MIMO channel matrix $H$ in (2) can be rewritten as

$$ H = U_s H U_t^H $$

where $U_s = [u_s(\theta_{l_s}^s)]_{s \in \{t, r\}} \in \mathbb{C}^{N_r \times N_s}$ for $s \in \{t, r\}$. Here, $H$ has only $L$ non-zero elements, i.e., $[H]_{m,n} = \sqrt{N_t N_r} g_{\ell}$ where the indices $(m, n)$ are associated with the angles of $\theta_m^s$ and $\theta_n^s$ of the $\ell$-th path. Thus, we may require at most $N_t N_r$ channel measurements to find the non-zero entries of $H$.

To partition the set of AoAs into subintervals as in [24], [28], we assume that the receiver employs the filter bank $U_t^H$ to the received signal in (1) given by

$$ \tilde{y}_k = U_t^H y_k = \sqrt{P} H U_t^H f_k s_k + \tilde{z}_k, $$

where $\tilde{z}_k = U_t^H z_k$. For a ULA, the receive filter bank can be implemented using a DFT-based filter [15]. We convert a joint detection of the non-zero elements of $H$ in
the whole range of AoAs to separate detection problems in small angle intervals.\footnote{The receiver can use a combiner for the received signals as the transmitter uses the beamformer in the same way.}

B. Review of Prior Work on Multi-Resolution Beam Alignment

Much of the prior work considers a multi-resolution beam alignment framework \cite{8, 10, 11} where the hierarchical multi-resolution codebook consists of $M$ levels to achieve better angular resolution. For $1 \leq m \leq M$, the $m$-th level codebook $\mathcal{F}_m = \{f_1^{(m)}, \ldots, f_{\text{card}(\mathcal{F}_m)}^{(m)}\}$ is designed to uniformly cover all the spatial frequency range $(-1, 1)$ (i.e., one-sided physical angles $(-\pi/2, \pi/2)$) and to satisfy the relation $\text{card}(\mathcal{F}_m) < \cdots < \text{card}(\mathcal{F}_M)$, where $\text{card}(\mathcal{F})$ denotes the cardinality of $\mathcal{F}$. (An example of a multi-resolution codebook with $[\text{card}(\mathcal{F}_1), \text{card}(\mathcal{F}_3), \text{card}(\mathcal{F}_3)] = [8, 16, 64]$ is shown in Fig. \ref{fig:multilevel}. The beamwidth of each of the codewords in $\mathcal{F}_m$ is given by

$$\psi_m = 2/\text{card}(\mathcal{F}_m), \quad \text{for } 1 \leq m \leq M, \quad (6)$$

and satisfies $\psi_1 > \cdots > \psi_M$. Accordingly, the angular domains are partitioned into $\text{card}(\mathcal{F}_m)$ angular intervals for the $m$-th level codebook where we denote each angular interval covered by $f^{(m)}_v$ by $\text{AoD}_{m,v} \subset (-1, 1)$ such that $(-1, 1) = \cup_{v=1}^{\text{card}(\mathcal{F}_m)} \text{AoD}_{m,v}$ and $\text{AoD}_{m,v} \cap \text{AoD}_{m,w} = \emptyset$ for $v \neq w$. Thus, the length of the interval $\text{AoD}_{m,v}$ is given by $|\text{AoD}_{m,v}| = \psi_m$ for $1 \leq v \leq \text{card}(\mathcal{F}_m)$ where $|\text{AoD}_1|$ denotes the length of an angular interval $\text{AoD}_1$. The training period is partitioned into $M$ stages where the $m$-th training stage is allocated for sounding a subset of beamforming vectors in $\mathcal{F}_m$.\footnote{We use the term “level” to refer the codebook and the term “stage” to refer the training phase.}

The codebook $\mathcal{F}_m$ is designed to provide an approximately constant main-lobe beamforming gain over the angle interval, which is mathematically described as

$$\left| \mathbf{u}_1(\theta)^H f^{(m)}_v \right| = \begin{cases} \sqrt{C_m} & \text{if } \theta \in \text{AoD}_{m,v}, \\ 0 & \text{if } \theta \notin \text{AoD}_{m,v}, \end{cases} \quad (7)$$

where $C_m$ is a normalization constant that satisfies $\|f^{(m)}_v\|^2 = 1$. Therefore, the beam patterns of $\mathcal{F}_m$ are characterized by the main-lobe directivity gain $C_m$ and the main-lobe beamwidth $\psi_m$ with negligible sidelobe gain, in a manner similar to \cite{29, 30}, which is illustrated in Fig. \ref{fig:beam_pattern}.

Denote the angular interval to be searched at the $m$-th training stage by $\text{AoD}_m \subset (-1, 1)$ and the number of distinct beamforming vectors in $\mathcal{F}_m$ required for searching $\text{AoD}_m$ by $V_m \in \mathbb{N}$ given by $V_m = |\text{AoD}_m|/\psi_m$.

By abuse of notation, we write $\{f_1^{(m)}, \ldots, f_{V_m}^{(m)}\} \in \mathcal{F}_m$ to denote the selected $V_m$ codewords to cover the collection of disjoint and mutually exclusive $\text{AoD}_m$, i.e., $\text{AoD}_m = \cup_{v=1}^{V_m} \text{AoD}_{m,v}$. At the $m$-th training stage, the system model of (5) can be rewritten as

$$\begin{bmatrix} \mathbf{y}_1^{(m)} \\ \vdots \\ \mathbf{y}_{V_m}^{(m)} \end{bmatrix} = \sqrt{P} \mathbf{H} \mathbf{U}_1^H \begin{bmatrix} f_1^{(m)} \\ \vdots \\ f_{V_m}^{(m)} \end{bmatrix} + \begin{bmatrix} \tilde{z}_1^{(m)} \\ \vdots \\ \tilde{z}_{V_m}^{(m)} \end{bmatrix}, \quad (8)$$

where the pilot symbols are assumed normalized to unit power and omitted. From (8), the receiver selects a beam index $v'$ that achieves the largest received signal power given by

$$v' = \arg\max_{1 \leq v \leq V_m} \left| \mathbf{y}_v^{(m)}(n) \right|^2, \quad (9)$$

where $\mathbf{y}_v^{(m)}(n)$ denotes the $n$-th element of the vector $\mathbf{y}_v^{(m)}$ and the receiver selects the index $n$ that yields the largest average magnitude of $\mathbf{y}_1^{(m)}(n), \ldots, \mathbf{y}_{V_m}^{(m)}(n)$. For the sake of simplicity, we assume that the receiver knows the index $n$ and will omit the dependence on $(n)$. The beam index $v'$ in (9) which corresponds to $f_v^{(m)}$ and $\text{AoD}_{m,v'}$ is fed back to enable the transmitter to choose the next-level codewords in $\mathcal{F}_{m+1}$. (A hierarchical beam alignment system requires feedback of the codebook index across the training stages.)

At training stage $(m+1)$, the angle interval of interest is set to $\text{AoD}_{m+1} = \text{AoD}_{m,v'}$ and the $V_{m+1}$ codewords required for probing $\text{AoD}_{m+1}$ is selected from $\mathcal{F}_{m+1}$. Note from (A.1) that the value of $V_{m+1}$ is given by

$$V_{m+1} = \left\lfloor \frac{|\text{AoD}_{m+1}|}{\psi_{m+1}} \right\rfloor = \left\lfloor \frac{\psi_m}{\psi_{m+1}} \right\rfloor, \quad (10)$$

where $\psi_m = |\text{AoD}_{m,v'}|$ for $1 \leq m < M$ given the constrained training period length $K_p = \sum_{m=1}^M V_m$ within a block.

An example of a multilevel beamforming framework with $V_1 = V_2 = 3$ and $M = 2$ is shown in Fig. \ref{fig:multilevel}.
Given AoD$_1$, $f^{(1)}_2$ is chosen at stage 1 and then high-resolution beamformers $\{f^{(2)}_1, \ldots, f^{(2)}_m\}$ on AoD$_2 = \text{AoD}_{1,2}$ are sounded at stage 2.

III. PROPOSED MULTILEVEL BEAMFORMING FRAMEWORK

In this section, we first explain the design of a single-level beam alignment sequence that use a single codebook chosen from a set of multilevel codebooks during the training period. We then extend the single-level beam alignment framework to design a multilevel beam alignment sequence where a subset of the set of multilevel codebooks used for channel sounding is selected adaptively aimed at improving the data rate performance.

A. Adaptive Single Level Beam Alignment and Sequence Design

To estimate an AoD partition during the training period, the receiver chooses a codeword achieving the largest received signal power as in (9). However, if the desired signal power is less than the uncertainty in the noise power at moderate-to-low SNR, the receiver cannot robustly detect the desired signal. In contrast to the previous methods explained in Section II-B, we do not use the expression in (9). Instead we employ a threshold technique that is adapted to the SNR.

In this subsection, we consider a single-level beam alignment in which a single codebook selected from $\{\mathcal{F}_m : 1 \leq m \leq M\}$ is exclusively used during the training period. For the $m$th-level codebook $\mathcal{F}_m$, we replace the alignment decision in (9) with a hypothesis test given by

$$\begin{cases} H_{0}^{(m)} : \tilde{y}_{v}^{(m)} \sim \mathcal{CN}(0, \sigma_{z}^{2}) \\ H_{1}^{(m)} : \tilde{y}_{v}^{(m)} \sim \mathcal{CN}(G_m P [g]^{2}, \sigma_{z}^{2}) \end{cases},$$

(11)

where $G_m := N_l N_r C_m$ and the path gain $|g|^2$ may not be known a priori but follows a central chi-square distribution with $2\ell$ degrees of freedom, i.e., $|g|^2 \sim (\sigma_{z}^{2}/2) \chi^2_{2\ell}$. Here, $\ell \leq L$ represents the number of paths within the corresponding beamwidth. For tractability of the analysis, we consider the case of single dominant path channels (i.e., $L = 1$ in (2)), then we make some comments regarding the multi-path case.

Under a fixed signal power, the beamformer $f^{(m)}_v$ has a lower main-lobe directivity gain than $f^{(m+1)}_v$ due to its broader beamwidth. Thus, the probability of successfully estimating the angular interval becomes lower as the index $m$ decreases. However, using high-resolution beamforming vectors during the training period requires substantial training overhead due to its narrow beamwidth which can degrade the data rate performance. This tradeoff indicates that the index $m$ specifying the codebook level $\mathcal{F}_m$ used during the training period is a suitable design parameter.

Given the codebook $\mathcal{F}_m$, we introduce a new variable $W_m \in \mathbb{N}$ to specify the number of repetitions corresponding to each of the codewords in $\mathcal{F}_m$. Then, the codewords $\{f^{(m)}_1, \ldots, f^{(m)}_V\} \in \mathcal{F}_m$ selected to cover the sector of interest denoted by AoD$_1$ are sounded, where the received signal $\tilde{y}^{(m)}_{v} \in \mathbb{C}$ in (9) corresponding to $f^{(m)}_v$ is changed to $[\tilde{y}^{(m)}_{1,v}, \ldots, \tilde{y}^{(m)}_{W_m,v}]^T \in \mathbb{C}^{W_m}$ given by

$$\begin{bmatrix} \tilde{y}^{(m)}_{1,v} \\
\ldots \\
\tilde{y}^{(m)}_{W_m,v} \end{bmatrix} = \sqrt{P} \mathbf{H}_m : \mathbf{U}^H \begin{bmatrix} 1_{W_m} \otimes f^{(m)}_v \end{bmatrix} + \begin{bmatrix} z^{(m)}_{1,v} \\
\ldots \\
z^{(m)}_{W_m,v} \end{bmatrix},$$

(12)

where $1_N \in \mathbb{R}^N$ denotes the all-ones vector, $\otimes$ denotes the Kronecker product, and recall that we omit the index $(n)$ for a concise notation. Here, the variable $V_m$ is redefined as

$$V_m = \left\lceil \frac{|\text{AoD}_1|}{\psi_m} \right\rceil,$$

(13)

as to represent the size of the subset of $\mathcal{F}_m$ used to cover the angular interval AoD$_1$. The training overhead becomes

$$K_m = V_m W_m.$$ 

(14)

In Fig. 3, we illustrate the process of training and data transmission in a single level beam alignment framework. We will formulate an optimization problem over the variables $m$ and $W_m$ where the objective is to provide an improved average data rate. For the present model in (11) and (12), we adopt a Neyman-Pearson detector for $f^{(m)}_v$ given by [31]

$$\frac{1}{W_m} \sum_{n=1}^{W_m} \tilde{y}^{(m)}_{n,v} \begin{bmatrix} \tilde{r}^{(m)}_{n} \\
\tilde{r}^{(m)}_{m} \end{bmatrix}^T \begin{bmatrix} \tilde{r}^{(m)}_{m} \\
\tilde{r}^{(m)}_{n} \end{bmatrix} > \gamma_m,$$

(15)

where the threshold is set to $\gamma_m = -\sigma_{z}^{2} \ln(P_{FA})/W_m$. 

---

3For the multi-path case, the number of paths covered by a beamforming vector depends on the propagation geometry, which makes the closed-form expression of the data rate performance required for a sequence design intractable.
given a target false alarm probability $P_{FA}$, i.e., $P_{FA} = P\left\{ \frac{1}{W_m} \sum_{n=1}^{W_m} \gamma_{n,v} (m) > \gamma_m | \mathcal{A}_m \right\}$. If there are multiple codewords that satisfy (15), the receiver selects the single codeword index among them that achieves the largest received signal power (i.e., the left-hand side of (15)). The combined signal on the left-hand side of (15) can be rewritten as

$$1/W_m \sum_{n=1}^{W_m} \gamma_{n,v} (m) = \sqrt{G_m} P g + 1/W_m \sum_{n=1}^{W_m} z_{n,v} =: x_m + z_m,$$

(16)

where $x_m = \sqrt{G_m} P g$ and $z_m = (1/W_m) \sum_{n=1}^{W_m} z_{n,v}$ denote the desired signal and noise, respectively, i.e., $x_m \sim \mathcal{CN}(0, G_m \sigma_g^2)$ and $z_m \sim \mathcal{CN}(0, \sigma_z^2/W_m)$, and the variables are assumed to be mutually independent. That is, the magnitude of $x_m$ represents the instantaneous effective channel gain achievable after a decision $\mathcal{H}^m_1$, while we operate as if $|x_m| = 0$ for a decision $\mathcal{H}^m_0$ because there is no transmission. The beam misalignment analysis, similar to [8], [11], can be performed using the expressions of false alarm and miss detection probabilities of (15). However, this approach cannot be applied to the design of the optimal training length because reducing the probability of beam misalignment requires the maximum training length. Instead, we consider the average data rate as a performance metric and focus on designing a beam alignment sequence and the training length within a constrained time.

When a single propagation path is covered by a beamformer $\mathbf{f}_m^{(\text{opt})} \in \mathcal{F}_m$, an instantaneous data rate for $\mathcal{H}^m_1$ is obtained by

$$R(x_m) = \frac{K - K_m}{K} \cdot \log(1 + |x_m|^2),$$

(17)

where the pre-log factor in (17) represents the cost of channel sounding within $K$ channel uses. Note that since a signal is detected when the squared absolute value of the (combined) received signal is above the threshold, consideration of joint variations for the desired signal and noise is needed to provide an exact average data rate, i.e.,

$$R_m = \mathbb{E}\{R(x_m)|A_0\} p(A_0) + \mathbb{E}\{R(x_m)|A_1\} p(A_1) = \frac{K - K_m}{K} \int_{0}^{\infty} \int_{0}^{\infty} \log(1 + |x_m|^2) \cdot p(x_m, z_m | A_1) dx_m dz_m \cdot p(A_1),$$

(18)

where $A_0$ and $A_1$ denotes mutually exclusive events described in the terms of random variables $x_m$ and $z_m$ (i.e., $A_0 = \{ |x_m + z_m|^2 < \gamma_m \}$ and $A_1 = \{ |x_m + z_m|^2 \geq \gamma_m \}$) and $p(x_m, z_m | A_1)$ is a conditional joint probability distribution function (pdf) of $(x_m, z_m)$ given $A_1$. Here, $\mathbb{E}\{R(x_m)|\mathcal{H}^m_0\} = 0$ because we set $|x_m|^2 = 0$ for $(x_m, z_m) \in A_0$ from a consideration of no transmission. To the best of the authors’ knowledge, the integral in (18) does not admit a closed-form expression. Instead, we evaluate an average magnitude of the effective channel gain (i.e., $|x_m|^2$) to obtain a reasonable criterion for approximating the expression of (18). To gain a clear insight to the main factor, we now provide a simple expression for the desired signal power in the high SNR regime, while an exact expression without any approximation is given in Appendix A.

**Lemma 1:** The desired signal power in the high SNR regime (i.e., $G_m P \sigma_g^2 > \sigma_z^2/W_m$) is approximated as

$$\mathbb{E}\{|x_m|^2\} \approx \frac{1}{W_m} \frac{\gamma_m}{\sigma_z^2} (G_m P \sigma_g^2 + \gamma_m) + \frac{P_{FA}}{2} - f_{\sigma_z^2, P_{FA}, W_m},$$

(19)

where $f_{\sigma_z^2, P_{FA}, W_m} = \frac{\sigma_z^2 \ln(1/P_{FA})}{W_m} (\sqrt{P_{FA}} - P_{FA})$.

**Proof:** See Appendix A.

Note that the average signal power given $\mathcal{H}^m_1$ is characterized by the first term on the right-hand side of (19), which is the expected value of $|x_m|^2$ integrated over $(\gamma_m, \infty)$, because the second term is negligible due to a small target false alarm rate and we numerically find that the magnitude of $f_{\sigma_z^2, P_{FA}, W_m}$ is less than $0.53 \sigma_z^2$ for all $P_{FA} \in (0, 1)$ and $W_m \in \{1, \ldots, 10^2\}$.

From Lemma 1, we consider the case that an instantaneous data rate is given by (17) when $|x_m|^2 > \gamma_m$ (i.e., $G_m P |y|^2 > \gamma_m$), otherwise we have $R(x_m) = 0$. Then, $R_m$ in (18) is approximated as

$$R_m \approx \frac{K - K_m}{K} \int_{\gamma_m}^{\infty} \log(1 + |x_m|^2) \frac{e^{-|x_m|^2/\sigma_m^2}}{\sigma_m^2/2} d|x_m|$$

(20)

$$= \frac{K - K_m}{K} \left( \frac{1 - \beta_m}{\sigma_m^2} \log \beta_m + \frac{\beta_m}{\sigma_m^2} \ln 2 - \frac{\beta_m}{\sigma_m^2} \right),$$

(21)

where $\beta_m := G_m P \sigma_g^2$, $\beta_m = 1 + \gamma_m$, and $\Gamma(\cdot, \cdot)$ denotes the incomplete gamma function [32, Eq. 8.350.2]. Maximization of $R_m$ in (21) gives the design of the codebook level $m$ and the number of retransmissions of the corresponding codewords $W_m$.

**Problem 1:** Given the codebooks $\mathcal{F}_m : 1 \leq m \leq M$, the variables $m$ and $W_m$ must be chosen to maximize

$$\max_{\{m, W_m\}} R_m$$

subject to $1 \leq m \leq M$ and $K_m \leq K$,

where $K_m$ and $R_m$ are defined in (14) and (21), respectively.

Considering the integer constraint on the optimization variables, we tackle Problem 1 by exhaustive search. Due to space limitations, we omit the details of our algorithm.
for Problem 1. Instead, in the next subsection, we present an algorithm for a multilevel beam alignment sequence design that can be used to solve Problem 1.

Remark 1: By applying Alzer’s inequality to the incomplete gamma function \( \Gamma(\cdot, \cdot) \), we can use the lower bound of \( R_m \) in (23) for a beam alignment sequence design. Simulation results in Section V show that the proposed algorithm using the lower bound in (23) performs very closely to those results obtained using the expression in (21).

Note that the proposed beam alignment design can be directly used for the multi-path case where the selected single codeword is expected to cover at least a single dominant path. Alternatively, the codeword indices that achieve the received signal power above the threshold value in (15) can be fed back to the transmitter to enable multi-stream transmission using the combined codewords, which requires additional feedback. Joint processing of the multiple training snapshots for the multi-path case is beyond the scope of the current work.

B. Adaptive Multilevel Beam Alignment and Sequence Design

In this subsection, we focus on the design of a multilevel beam alignment sequence where a subset of codebooks (i.e., \( \mathcal{F}_m : m_F \leq m \leq m_L \)) can be used during the training period. Here, we define the indices specifying the codeword levels used at the first and last training stages as \( m_F \) and \( m_L \) for \( 1 \leq m_F \leq m_L \leq M \), respectively.

Since the codewords in \( \mathcal{F}_{m_F} \) sounded at the first training stage have relatively lower directivity gain due to its broader beamwidth than those in \( \mathcal{F}_m : m_F < m \leq m_L \), the received signal power at the first training stage can be easily swayed by noise impact. Thus, we assume retransmission of the codewords in \( \mathcal{F}_{m_F} \) sounded during the first training stage where \( W_{m_F} \) represents the number of retransmissions. The system model and the detection process at the first training stage can be described by replacing the index of the codebook in (12) and (15) with \( m_F \), respectively. During the intermediate training stages allocated for sounding the codewords in \( \mathcal{F}_m : m_F < m \leq m_L \), we assume that each codeword is transmitted once, i.e., we set \( W = 1 \) in (12) and (15) for \( m_F < m \leq m_L \).

![Fig. 4: Training and data transmissions where the codewords sounded at the first training stage are retransmitted over \( W_{m_F} \) channel uses, otherwise each codeword is transmitted once.](image)

Given this, the training overhead becomes

\[
K_{m_F}^{m_L} = V_{m_F} W_{m_F} + \sum_{m=m_F+1}^{m_L} V_m,
\]

where the variable \( V_m \), corresponding to the codebook \( \mathcal{F}_m \) in (A.1), is redefined as

\[
V_m = \lceil \text{AoD}_1 / \psi_m \rceil \quad \text{and} \quad V_m = \lfloor \psi_m - 1 / \psi_m \rceil,
\]

for \( m_F < m \leq m_L \). Here, \( \text{AoD}_1 \) denotes the sector angle to be searched at the first training stage. The overall process of training and data transmission is illustrated in Fig. 4. As a result, we will optimize the codeword levels \( (m_F, m_L) \) and \( W_{m_F} \) used during the consecutive training stages to provide an improved average data rate, which will be shown shortly.

We observe that the performance achievable by the multilevel beam alignment framework is characterized by the detection performance at the first stage because the codewords in \( \mathcal{F}_{m_F} \) have relatively low directivity gain than those in \( \mathcal{F}_m : m_F < m \leq m_L \). For the purpose of analysis, we assume that if a beamforming vector in \( \mathcal{F}_{m_F} \) is properly selected, then a beamforming vector in \( \mathcal{F}_m \) for \( m_F < m \leq m_L \) that covers the dominant channel path can be perfectly selected due to its relatively higher beamforming gain. We will refer to this assumption as (A.1). (As the comparison with numerical results without (A.1), this approximation does not significantly affect our results.)

From (A.1), an (approximate) instantaneous data rate

\[4\]
for $|x_{mF}|^2 > \gamma_{mF}$ can be written, similar to (17), as

$$R(x_{mF}, m_L) \approx \frac{K - K_{mF}^{mL}}{K} \cdot \log \left(1 + \frac{G_{mF}^{mL}}{G_{mF}^{mL}}|x_{mF}|^2\right),$$

(26)

where the operand in the log function denotes the beamforming gain achievable at the last training stage under (A.1). We then evaluate the (approximate) average throughput in the multilevel beamforming framework in semi-closed form given by

$$R_{mF}^{mL} \approx \mathbb{E}\{R(x_{mF}, m_L)\} \approx \frac{K - K_{mF}^{mL}}{K} \left(\frac{e^{1 - \beta_{mF}^{mL}}/\alpha_{mF}^{mL}}{e^{1/\alpha_{mF}^{mL}}/2} \cdot \log \beta_{mF}^{mL}ight),$$

(27)

where $\alpha_{mF}^{mL} = G_{mF}^L P \sigma^2_{\text{ph}}$, $\beta_{mF}^{mL} = 1 + (G_{mF}^L/G_{mF}^{mL})\gamma_{mF}$, and the expectation in (27) is performed over $|x_{mF}|^2 \in (\gamma_{mF}, \infty)$, similar to (20).

Therefore, using the objective function in (28), we design a multilevel beam alignment sequence optimized over the variables $(m_F, m_L, W_{mF})$ that strikes a balance between minimizing the training overhead and maximizing beamforming gain.

**Problem 2:** Given the codebooks $\{\mathcal{F}_m : 1 \leq m \leq M\}$ and (A.1), the level indices $(m_F, m_L)$ and $W_{mF}$ must be chosen to maximize

$$\max_{\{m_F, m_L, W_{mF}\}} R_{mF}^{mL}$$

subject to $1 \leq m_F \leq m_L \leq M$, $K_{mF}^{mL} \leq K$, and $1 \leq W_{mF} \leq |G_{mF+1}/G_{mF}|$, (29)

where $K_{mF}^{mL}$ and $R_{mF}^{mL}$ are defined in (24) and (28), respectively.

The upper bound on $W_{mF}$ in (30) guarantees that the effective SNR at level $m_F$ is smaller than that of the next level, i.e., $G_{mF}^L P |g|^2/\sigma^2_{\text{ph}} W_{mF} \leq G_{mF+1}^L P |g|^2/\sigma^2_{\text{ph}}$ because of $W_{mF} G_{mF} \leq G_{mF+1}$. The corresponding algorithm is summarized in Algorithm 1 with reasonable computational complexity of $\mathcal{O}(M^2 K)$ as only three to five levels are typically used in the multilevel beamforming framework.

**Remark 2:** Similar to Remark 1, the expression of (28) can be lower bounded by $R_{mF}^{mL} \geq R_{mF}^{mL}$ where

$$R_{mF}^{mL} = \frac{K - K_{mF}^{mL}}{K} \left(\frac{1}{e^{1 - \beta_{mF}^{mL}}/\alpha_{mF}^{mL}} \cdot \log \beta_{mF}^{mL} - \frac{1}{e^{1/\alpha_{mF}^{mL}}/2} \cdot \log \left(1 - e^{-\frac{1}{\alpha_{mF}^{mL}}}\right)\right).$$

(31)

**Algorithm 1 Multilevel Beam Sequence Design**

1. Set $R = 0$
2. for $m_L = M : -1 : 1$
3. for $m_F = m_L : -1 : 1$
4. $[\psi_{mF}; \ldots; \psi_{mL}] = \left[\frac{\psi_{mF}}{\psi_{mL}}, \ldots, \frac{\psi_{mL}}{\psi_{mL}}\right]$
5. $W_{max} = \left|\frac{1}{K - \sum_{mF=0}^{mF} W_{mF}}\right|$
6. if $m_F < m_L$
7. $W_{max} = \min\{W_{max}, |G_{mF+1}/G_{mF}|\}$
8. end if
9. for $W_{mF} = 1 : 1 : W_{max}$
10. Compute $R_{mF}^{mL}$
11. if $R_{mF}^{mL} > R$
12. $R = R_{mF}^{mL}$
13. $[\tilde{m}_F, \tilde{m}_L] \leftarrow [m_F, m_L]$
14. $W_{mF} \leftarrow W_{mF}$
15. end if
16. end for
17. end for
18. end for

(36)

(The notations follow those of (28) or (31). Here, the hat notation denotes the output of the algorithm.)

IV. PROPOSED CODEBOOK DESIGN TECHNIQUE

Throughout the manuscript, we assumed that the beam pattern is approximated by constant directivity gain $C_m$ within the main-lobe bandwidth $\psi_m$. To design accurate approximations to the ideal beam patterns, much of the prior work considers analog beamforming [6]–[8] and hybrid beamforming techniques [10]–[12] where the codebook is assumed to be designed off-line constructed using a limited number of RF chains.

However, when a user or scatterers are around the boundary of two adjacent (pre-designed) beamforming patterns, it might be detected in both angular partitions. Then, shifting the main-lobe beam direction to the target angle or redesigning the codebook is needed to provide an improved beamforming gain. Considering many prior works on DFT-based codebooks in [7], [34]–[36], we present an efficient method to design a multi-resolution codebook using the DFT vectors.

A. DFT-based Multilevel Codebook Design

Considering the quantized spatial angles $\{\theta_n = -1 + (2n - 1)/N_t : 1 \leq n \leq N_t\}$, the $m$-th level codebook $\mathcal{F}_m$ can be designed by combining the array steering vectors corresponding to the subset of quantized angles given by

$$\mathcal{F}_m = \left\{e^{(m)}_{1}, e^{(m)}_{2}, \ldots, e^{(m)}_{N_m/m}\right\}$$

(36)

$$= \left\{\frac{1}{\sqrt{N_m}} \sum_{p=1}^{N_m} u_t(\theta_p) e^{j\theta_m/p}, \frac{1}{\sqrt{N_m}} \sum_{p=N_m+1}^{2N_m} u_t(\theta_p) e^{j\theta_m/p}, \ldots, \frac{1}{\sqrt{N_m}} \sum_{p=N_t-N_m+1}^{N_t} u_t(\theta_p) e^{j\theta_m/p}\right\},$$

(37)
where we introduce a new variable $\omega_m$ to enable the beam patterns of $f_i^{(m)}$ to have a little to no variation in main-lobe directivity gain as best as possible. Here, we assume that $N_m$ is a divisor of $N_t$ for simplicity. The array steering vector on the quantized angles $\{\theta_m\}$ in (37) becomes a generalized DFT vector with offset $-(1+N_t)/2$, i.e.,

$$ u_i(\theta_m) = \frac{1}{\sqrt{N_t}} \left[ 1, e^{-j \frac{2\pi}{N_t} (n - \frac{N_t+1}{2})}, \ldots, e^{-j \frac{2\pi}{N_t} (n - \frac{N_t+1}{2})} \right]^T, \quad (38) $$

which is the $n$-th column of the Butler matrix [37] and can be easily implemented using millimeter microstrip technology [38]-[40]. We denote the number of combined vectors at level $m$ by $N_m$ and the number of active RF chains by $N_{RF}$, respectively. From (37), we need at least $N_{RF} = N_m$ RF chains for combining the vectors.

The codewords in (37) form ideal beam patterns over the quantized angles such that the codebook with arbitrary $\omega_m \in \mathbb{R}$ achieves the (constant) maximum beamforming gain $N_t/N_m$ on the quantized angles $\{\theta_m\}$. For $N_t = 64$, $M = 3$, and $\{N_1, N_2, N_3\} = \{8, 4, 1\}$, Fig. 5(a) shows the corresponding beam patterns over the quantized angles, which is useful for theoretical analysis and can be used as a benchmark showing what can be achieved under idealized assumptions. For realizing this scenario in practice, large-scale antenna arrays can be applied in the transmitter and/or the receiver required for finer spatial sampling of continuous angles.

Note that the codebook $\mathcal{F}_m$ in (37) is parameterized by the phase shift $\omega_m$, and we refer to this scheme as a phase-shifted DFT method. Optimizing the single parameter $\omega_m$ is important to achieve constant main-lobe directivity gain as possible. An example of the codebook for $\omega_m = 0$ (i.e., a simple addition of the array steering vectors as used in [24]) is shown in Fig. 5(b). Especially at lower level codewords, there are a large number of null points within the main beamwidth and picking points near the boundary of beam patterns due to the overlapped DFT beam patterns on the continuous angles.

For the $i$-th codeword $f_i^{(m)} \in \mathcal{F}_m$ in (36), we evaluate a closed-form expression for the array response over the continuous angle $\theta$ in (33) where $u_i(\theta)^H u_i(\theta_p)$ represents the discrete Fourier transform using the $p$-th vector in (38) evaluated at the spatial location parameter $\theta$. Given this, the array response magnitude can be de-

\[
\begin{align*}
    u_i(\theta)^H f_i^{(m)} &= u_i(\theta)^H \left( \frac{1}{\sqrt{N_m}} \sum_{p=(i-1)N_m+1}^{iN_m} u_i(\theta_p) \right) \\
    &= \frac{1}{N_t \sqrt{N_m}} \sum_{p=(i-1)N_m+1}^{iN_m} e^{j\omega_m p} e^{-j\frac{2\pi}{N_t} \left( p - N_t \frac{\theta + 1}{2} \right)} \sin \left( \frac{\pi (p - N_t \frac{\theta + 1}{2})}{N_t} \right) \\
    f_{\theta,i,\omega_m} &= \left| u_i(\theta)^H f_i^{(m)} \right|^2 \\
    &= \frac{1}{N_t^2 N_m} \left\{ \sum_{p=(i-1)N_m+1}^{iN_m} \sin^2 \left( \frac{\pi (p - \theta)}{N} \right) \right. \\
    &\quad+ \left. \sum_{p=(i-1)N_m+1}^{iN_m} 2 \cos \left( p' - p \right) \left( \frac{N - 1}{N} \pi - \omega_m \right) \right\} \times \frac{\sin \left( \frac{\pi (p - \theta)}{N_t} \right) \sin \left( \frac{\pi (p' - \theta)}{N_t} \right)}{\sin \left( \frac{\pi (\theta_p - \theta)}{N_t} \right) \sin \left( \frac{\pi (\theta_p' - \theta)}{N_t} \right)},
\end{align*}
\]

where $\hat{\theta} := \frac{N_t \theta + 1}{2}$. 

---

Fig. 5: Phase-shifted DFT beam patterns where $N_t = 64$, $M = 3$, $|\text{card}(\mathcal{F}_1), \text{card}(\mathcal{F}_3), \text{card}(\mathcal{F}_3)| = \{8, 16, 64\}$, and $\{N_1, N_2, N_3\} = \{8, 4, 1\}$: (a) ideal case on the quantized spatial domain with an arbitrary $\omega_m \in \mathbb{R}$ and (b) non-ideal case on the continuous spatial domain with $\omega_m = 0$ for $1 \leq m \leq M$. 

---

(32)
scribed by a function $f^{(m)}_{\theta,i,\omega_m}$ defined in (34). As a result, we focus on an exhaustive search for the proper phase shift $\omega_m$ that minimizes the variation of the main-lobe directivity gain evaluated with $f^{(m)}_{\theta,i,\omega_m}$ over the quantized version of main beam range given by

$$[-1 + \frac{2(i - 1)N_m + 1}{N_1} - 1, -1 + \frac{2(iN_m - 1)}{N_1}].$$  

(39)

In this case, it is important to reduce the computational complexity of the exhaustive search, and thus we derive a property of the objective function $f^{(m)}_{\theta,i,\omega_m}$.

**Theorem 1:** Given the level $m$ and $1 \leq i \leq N/N_m$, the function $f^{(m)}_{\theta,i,\omega_m}$ in (35) satisfies the following properties:

- (P.1) for $\omega_m$, $2\pi$-periodic and an even function at $\pi(N_1 - 1)/N_1$,
- (P.2) for $\theta$, an even function at $-1 + (2i - 1)N_m/N_1$,
- (P.3) independent of the index $i$ where $1 \leq i \leq N/N_m$.

**Proof:** See Appendix B.

**Remark 3:** Theorem 1 yields a much simpler implementation of the exhaustive search without loss in performance because: the search space for finding the phase shift $\omega_m$ is reduced to the interval $[-\pi/N_1, \pi(1 - 1/N_1)]$ of length $\pi$ by (P.1); the continuous angular interval for $\theta$ can be set to one-half of the main beam interval of (39) by (P.2), e.g. $[-1 + (2i - 1)N_m/N_1, -1 + (2iN_m - 1)/N_1]$; the optimized phase shift $\omega_m$ for the $i$-th codeword $f^{(m)}_i$ is also an optimal design for all other codewords in $\mathcal{F}_m$ by (P.3).

Therefore, the problem of the codebook design at level $m$ in (37) is formulated as

$$\omega_m \in \left[\frac{\pi}{N_1}, \pi(1 - \frac{1}{N_1})\right],$$  

(40)

for any fixed index $i \in \{1, \ldots, N/N_m\}$. Here, $\theta_b \in \mathbb{R}^b$ denotes the uniformly quantized points of the half of the main beam interval $[-1 + (2i - 1)N_m/N_1, -1 + (2iN_m - 1)/N_1]$ using $b$ bits where we set $b = 9$. Since the objective function of (35) can operate on the vector input $\theta_b$, the design parameter $\omega_m$ is swept over the possible range to minimize the sample variance in (40).

**Remark 4:** Since each codeword $f^{(m)}_i \in \mathcal{F}_m$ should provide the maximal constant level of concentration of energy over the spatial interval $[\theta_{(i - 1)N_m + 1}, \theta_{iN_m}]$, one can relate the codebook design problem in (40) to the design of bandpass filters for estimating the power spectral density (PSD) over $(-\pi, \pi)$. The finite impulse response (FIR) bandpass filter that concentrates its energy to the interval $[\pi\theta_{(i - 1)N_m + 1}, \pi\theta_{iN_m}]$ can be designed by using $N_m$ dominant eigenvectors of $\Gamma \in \mathbb{C}^{N_1 \times N_1}$, referred to as Slepian sequences [41], given by

$$\Gamma = \frac{1}{2\pi} \int_{-\pi\theta_{(i - 1)N_m + 1}}^{\pi\theta_{iN_m}} \mathbf{u}_i(\theta)\mathbf{u}_i(\theta)^H d\theta,$$

(41)

where $\mathbf{u}_i(\cdot)$ denotes the steering vector defined in (3) and $(p,q)$ element of $\Gamma$ in (41) is given by

$$[\Gamma]_{p,q} = e^{-j\pi(p-q)}(-1 + \frac{(2i - 1)\pi}{N_1}) \sin \left(\frac{(p-q)(N_m-1)}{N_1}\right) \pi(p-q).$$

Since the power of the filtered signal is approximately equal to the PSD value at $\theta$, the auto-PSD estimation over the interval is given by

$$\frac{1}{N_m} \sum_{p=1}^{N_m} |\mathbf{u}_i(\theta)^H \mathbf{v}_p|^2$$

where $\mathbf{v}_p \in \mathbb{C}^{N_1}$ is the $p$-th dominant eigenvector of $\Gamma$, referred to as the $p$-th dominant Slepian sequence. However, simulations of a codebook design using the Slepian sequences as the basis vectors and optimizing over a single phase-shift design parameter, $\omega_m$, similar to the codebook design for generalized DFT (Butler) beams, yielded distorted beamforming patterns. Thus, multidimensional optimization is needed to linearly combine Slepian sequences to achieve a maximal constant level of energy concentration over the sector of interest as best as possible. However, with respect to practical implementation, our proposed codebook using DFT beams according to (37) can be implemented via the well-known Butler Matrix Beamformer [37] in analog RF circuitry which, as discussed previously, can be implemented using millimeter microstrip technology [38]–[40]. In contrast, the values of the Slepian sequences are not constant magnitude and also their phase values are arbitrary, i.e. are not confined to a finite alphabet.

The proposed design obtained from (40) yields good performance over the continuous angular domain as shown in Fig. 6 for $\{N_1, N_2, N_3\} = \{8, 4, 1\}$, $M = 3$, and $N_1 = 64$. The proposed codebook has constant main-lobe directivity gain per beam similar to the ideal beam patterns in Fig. 5(a). For the codebook design in Fig. 6(a), we use at most $N_{RF} = 8$ RF chains at the first level codebook because of $N_1 = 8$. To accommodate the case that we cover the same angular interval with the small number of RF chains, we can reduce the number of active antenna elements by turning off a subset of successive antenna elements while having a coarse spatial
where and elevation patterns obtained with $N_{RF} = N_u = 4$ and $N_t = 32$. Although the estimation accuracy of the angular interval is restricted by the reduced number of antenna elements, the proposed beam patterns still yield reasonably good performance, which will be further demonstrated in Section V.

B. 3D Beamforming

When the array manifold of a two-dimensional uniform planar array (UPA) is modeled as the Kronecker product of vectors from the array manifolds of two (horizontal/vertical) ULAs as in [42]–[44], a 3D beamforming vector can be constructed from the Kronecker product of two (horizontal/vertical) codebooks. From the phase-shifted DFT method in (37), two (horizontal/vertical) codebooks, denoted as $\mathcal{F}_H$ and $\mathcal{F}_V$, are designed by optimizing the horizontal and vertical phase shifts $(\omega_H, \omega_V)$ separately. Mathematically, this means

$$\mathcal{F} = \mathcal{F}_H \otimes \mathcal{F}_V,$$  

(42)

where for $s \in \{\mathcal{H}, \mathcal{V}\}$

$$\mathcal{F}_s = \{f_{s,1}, f_{s,2}, \ldots, f_{s,N_s/t/N_s}\}$$

$$= \left\{ \frac{1}{\sqrt{N}} \sum_{p=(i-1)N_s+1}^{iN_s} u_{r,s}(\theta_{p,s}) e^{j\omega_x p} \middle| 1 \leq i \leq \frac{N_s t}{N_s} \right\}.$$  

(43)

Here, $u_{r,s}(\theta_{p,s})$ denotes the array steering vector on the quantized angles $\{\theta_{p,s} = -1 + (2p - 1)/N_{s,t} : 1 \leq p \leq N_{s,t}\}$, and $(N_{s,t}, N_s)$ denote the number of antenna elements and combined vectors in each dimension, respectively.

For a size $N_{V,t} \times N_{H,t}$ UPA with the azimuth $\phi_a$ (deg) and elevation $\phi_e$ (deg) AoDs, the channel response is given by [42]–[44]

$$u(\phi_a, \phi_e) = u_H(\phi_a, \phi_e) \otimes u_V(\phi_e),$$  

(44)

where

$$u_H(\phi_a, \phi_e) = \left[1, e^{-j\frac{\pi d}{\lambda} \sin(\phi_a) \cos(\phi_e)}, \ldots, e^{-j(\frac{N_{H,t} - 1}{\lambda} \pi d \sin(\phi_a) \cos(\phi_e))} \right]^T$$

(45)

$$u_V(\phi_e) = \left[1, e^{-j\frac{\pi d}{\lambda} \sin(\phi_e)}, \ldots, e^{-j(\frac{N_{V,t} - 1}{\lambda} \pi d \sin(\phi_e))} \right]^T.$$  

(46)

From (42) and (44), the array response for the codeword in $\mathcal{F}$ is evaluated as

$$|u(\phi_a, \phi_e)^H(f_{H,v} \otimes f_{V,w})|^2$$

$$= |u_H(\phi_a, \phi_e)^H f_{H,v}|^2 \times |u_V(\phi_e)^H f_{V,w}|^2,$$  

(47)

where $f_{H,v} \in \mathcal{F}_H$ and $f_{V,w} \in \mathcal{F}_V$ defined in (43).

In (47), if $f_{H,v}$ is a codeword designed for covering the azimuth angle $\phi_a \in (0^\circ, 30^\circ)$ (i.e., $\sin(\phi_a) \in (0, 0.5)$) for a ULA, the first term on the right hand side of (47) yields the maximum beamforming gain over $\sin(\phi_a) \cos(\phi_e) \in (0, 0.5)$ for a UPA due to a similar structure of the array steering vectors in (3) and (45). As a result, the range of the azimuth angle for a UPA that achieves the maximum beamforming gain becomes larger as much as the elevation angle $\phi_e$ changes from $0^\circ$ to $\pm 90^\circ$ (i.e., decreasing the value of $\cos(\phi_e)$). Similarly, the area spanned by the codeword that covers off-center in azimuth gets narrower by changing the elevation angle towards $\pm 90^\circ$. This is a manifestation of the well-known reduced-aperture effect off boresite.

For a $32 \times 128$ UPA, Fig. 7 shows the beam patterns of $\mathcal{F}$ obtained from the proposed method. Note that the beamwidth of the codeword designed for covering center azimuth grows larger while maintaining (nearly) the same beamforming gain as elevation angle $\phi_e$ goes to $\pm 90^\circ$. Each codeword yields reasonably constant beamforming gain over the azimuth and elevation angular partitions.

Therefore, the proposed method provides an efficient way to design the multi-resolution codebook on a need-based basis. Due to the separable 3D beamforming structure in (42), Algorithm 1 can be used to generate the multilevel beam alignment sequence in each of the azimuth and elevation directions by averaging the beamforming gains within a sector angle to characterize the main-lobe directivity gain of the codebooks.

V. NUMERICAL RESULTS

In this section, we provide some numerical results to evaluate the proposed beam sequence design and the corresponding algorithm. Throughout the simulation, we fixed the noise power $\sigma_n^2 = 1$ and $K = 40$. The sector angle at the transmitter is

$^5$We observe that $\omega_s = 2$ can be used as a rule of thumb.
shown in Fig. 6(a) on the quantized angles. It is seen [45]. In Fig. 8, we examine the data rate performance

**R** mean angle

$L$ Laplacian distributed angles of channel paths with the direction is set to

$P$ loss attenuation.

Thus, in examining the results, it is important to look at

$M$ gain, which is substantial in a Massive MIMO setting.

Note that the SNR values in the abcissa of each plot does not include the beamforming gain, which is substantial in a Massive MIMO setting. Thus, in examining the results, it is important to look at SNRs significantly below 0 dB in consideration of path loss attenuation.

We consider a MIMO system with $N_t = 64$, $N_r = 4$, $M = 3$, and AoD$_1 = (-30^\circ, 30^\circ)$ where we adopt Laplacian distributed angles of channel paths with the mean angle $\mu_\theta = 0$ and the standard deviation $\sigma_\theta = 7$ [45]. In Fig. 8, we examine the data rate performance based on the proposed method along with the codebook shown in Fig. 6(a) on the quantized angles. It is seen that the proposed algorithms using $R_{m_F}^L$ in (28) and the lower bound $\bar{R}_{m_F}^L$ in (31) yield almost the same performance. Thus, the simpler algorithm with (31) can be used without much performance loss. For $L = 1$, simulation result without (A.1) matches well the analytic result, i.e., the result in (28) gives expressions for the average data rate for a single path channel $L = 1$. For $L > 1$, simulation results are obtained by choosing a single codeword that achieves the largest received signal power.

For purpose of comparison, we consider the case that the transmitter fixes the codebook levels $\{F_m : m_F \leq m \leq m_L\}$ for $1 \leq m_F \leq m_L \leq M$ with $W_{m_F} = 1$. With the same setup used in Fig 8, we compared the data rate performance of the proposed method to those of others using the analytic result of (28) in Fig. 9. It appears that the proposed method using adaptively selected codebooks outperforms other methods. The optimized beamforming sequence obtained by Algorithm 1 is summarized in Table I. At low SNR, the proposed method retransmits the codewords sounded at the first training stage by two times aimed at improving the effective SNR. As the operating SNR increases, the proposed method uses all codebook levels with a single codeword transmission to decrease the training overhead. We observed that the use of a multilevel codebook (i.e., $m_F < m_L$) yields good performance compared to the use of a single level codebook (i.e., $m_F = m_L$) for most of the SNR range.

**TABLE I: Optimized beam alignment sequence**

<table>
<thead>
<tr>
<th>SNR(dB)</th>
<th>$M$</th>
<th>${m_F, \ldots, m_L}$</th>
<th>${W_{m_F}, \ldots, W_{m_L}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-16 ~ -12</td>
<td>2</td>
<td>${2, 3}$</td>
<td>$(2, 1)$</td>
</tr>
<tr>
<td>-10 ~ -2</td>
<td>2</td>
<td>${2, 3}$</td>
<td>$(1, 1)$</td>
</tr>
<tr>
<td>0 ~ 2</td>
<td>3</td>
<td>${1, 2, 3}$</td>
<td>$(1, 1, 1)$</td>
</tr>
</tbody>
</table>

Next, we compare the performance of the proposed methods to those of several codebook design techniques [8], [11]. For $M = 3$ and $N_t = 64$, we considered the channel model with a number of path $L = 1$ and the AoDs are assumed to take continuous values, i.e., not quantized and uniformly distributed in AoD$_1$ given by AoD$_1 = (-50^\circ, 50^\circ)$. Fig. 10 shows the beam patterns obtained by analog beamforming [8] (i.e., $N_{RF} = 1$) and hybrid beamforming [11] where we set $N_{RF} = 4$ and adopt enough phase quantization bits so that the phase quantization error on the performance is negligible. These should be compared to that obtained with our proposed method in Fig. 6. Given $\{N_1, N_2, N_3\} = \{8, 4, 1\}$, the proposed codebook design (i.e., the phase-shifted DFT method shown in Fig. 6(a)) needs at most $N_{RF} = 8$ RF chains because of $N_1 = 8$. For a fair comparison in terms of $N_{RF}$, we also considered a modified method where we replace the codebook in Fig. 6(a: Level 1) by the codebook in Fig. 6(a') in order to maintain $N_{RF} = 4$.

To evaluate the performance of the phase-shifted DFT
method without beam sequence optimization, we first fix the number of codewords levels as $M = 3$ where $m_F = 1, m_L = 3$, and $W_{m_F} = 1$, marked (a) in Fig. 11. Hybrid beamforming design achieves better performance than analog beamforming design due to the well-designed codebooks especially at level 1. For the fixed levels, the phase-shifted DFT method is superior to the other methods over the whole SNR range relative to data rate performance, even when we keep the same number of RF chains $N_{RF} = 4$. Using more RF chains $N_{RF} = 8$ in the phase-shifted DFT method yields slightly better performance than using $N_{RF} = 4$ due to the different codebooks at level 1 shown in Fig. 6(a, a').

By averaging the main-lobe directivity gain of the multilevel codebook over the interval AoD$_1$, we investigate the performance of all the methods including analog, hybrid, and the proposed methods in conjunction with Algorithm 1. The results marked (b) in Fig. 11 show that the data rate performance is further enhanced especially at moderate-to-low SNR due to beam sequence optimization. The methods for analog and hybrid beamforming yield similar performance up to the SNR around -4dB since only the codebooks at levels (2, 3) are used at low SNR and its beam patterns are comparable to each other shown in Fig. 10. Similarly, the phase-shifted DFT methods that use $N_{RF} \in \{4, 8\}$ also show the same performance up to mid-range SNR.

VI. CONCLUSIONS

We considered multi-resolution beam alignment sequence design to improve the data rate performance achievable in mmWave systems. The resulting multi-level beamformer alignment design generates a beamformer sequence that counterbalances minimizing training overhead and maximizing beamforming gain. We then provided a phase-shifted DFT method that efficiently designs the multi-resolution codebook and easily extends to three-dimensional beamforming codebook. The proposed codebook design and beam alignment sequence design provide an efficient means to improve the data rate performance in large-scale mmWave MIMO systems.

APPENDIX

A. Proof of Lemma 1

Define random variables $X$ and $Y$ by $X \sim \mathcal{CN}(0, \sigma_X^2)$ and $Z \sim \mathcal{CN}(0, \sigma_Z^2)$ where $\sigma_X^2 := G_m P_0 \sigma_0^2$ and $\sigma_Z^2 := \sigma^2/W_m$. We will omit the codebook index ($m$) for notational simplicity. The average power of $X$ is given by

\[
\mathbb{E}\{|X|^2\} = \mathbb{E}\{|X|^2|A_0\}p(A_0) + \mathbb{E}\{|X|^2|A_1\}p(A_1)
\]

\[
= p(A_1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x|^2 \cdot p_{XZ}(x, z|A_1) \, dx \, dz,
\]

where $A_0$ and $A_1$ denotes mutually exclusive events described in terms of the random variables $X$ and $Z$, i.e., $A_0 = \{|X + Z|^2 < \gamma\}$ and $A_1 = \{|X + Z|^2 \geq \gamma\}$. $p_{XZ}(x, z|A_1)$ denotes a conditional joint probability distribution function (pdf) given by [46]

\[
p_{XZ}(x, z|A_1) = \frac{p_{XZ}(x, z)}{p(A_1)},
\]

for $(X, Z) \in A_1$, otherwise $p_{XZ}(x, z|A_1) = 0$, where $p(A_1) = \int_{(X,Z) \in A_1} p_{XZ}(x, z) dx \, dz$. (When there is no

\[
\mathbb{E}\{|X|^2\} = p(A_1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x|^2 \cdot p_{XZ}(x, z|A_1) \, dx \, dz,
\]

the resulting multilevel codebook over the interval AoD$_1$, we investigate the performance of all the methods including analog, hybrid, and the proposed methods in conjunction with Algorithm 1. The results marked (b) in Fig. 11 show that the data rate performance is further enhanced especially at moderate-to-low SNR due to beam sequence optimization. The methods for analog and hybrid beamforming yield similar performance up to the SNR around -4dB since only the codebooks at levels (2, 3) are used at low SNR and its beam patterns are comparable to each other shown in Fig. 10. Similarly, the phase-shifted DFT methods that use $N_{RF} \in \{4, 8\}$ also show the same performance up to mid-range SNR.

VI. CONCLUSIONS

We considered multi-resolution beam alignment sequence design to improve the data rate performance achievable in mmWave systems. The resulting multi-level beamformer alignment design generates a beamformer sequence that counterbalances minimizing training overhead and maximizing beamforming gain. We then provided a phase-shifted DFT method that efficiently designs the multi-resolution codebook and easily extends to three-dimensional beamforming codebook. The proposed codebook design and beam alignment sequence design provide an efficient means to improve the data rate performance in large-scale mmWave MIMO systems.

APPENDIX

A. Proof of Lemma 1

Define random variables $X$ and $Y$ by $X \sim \mathcal{CN}(0, \sigma_X^2)$ and $Z \sim \mathcal{CN}(0, \sigma_Z^2)$ where $\sigma_X^2 := G_m P_0 \sigma_0^2$ and $\sigma_Z^2 := \sigma^2/W_m$. We will omit the codebook index ($m$) for notational simplicity. The average power of $X$ is given by

\[
\mathbb{E}\{ |X|^2 \} = \mathbb{E}\{ |X|^2 |A_0 \} p(A_0) + \mathbb{E}\{ |X|^2 |A_1 \} p(A_1)
\]

\[
= p(A_1) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x|^2 \cdot p_{XZ}(x, z|A_1) \, dx \, dz,
\]

where $A_0$ and $A_1$ denotes mutually exclusive events described in terms of the random variables $X$ and $Z$, i.e., $A_0 = \{|X + Z|^2 < \gamma\}$ and $A_1 = \{|X + Z|^2 \geq \gamma\}$. $p_{XZ}(x, z|A_1)$ denotes a conditional joint probability distribution function (pdf) given by [46]

\[
p_{XZ}(x, z|A_1) = \frac{p_{XZ}(x, z)}{p(A_1)},
\]

for $(X, Z) \in A_1$, otherwise $p_{XZ}(x, z|A_1) = 0$, where $p(A_1) = \int_{(X,Z) \in A_1} p_{XZ}(x, z) dx \, dz$. (When there is no
ambiguity, we omit the subscript \( X \) from \( p_X(x) \) and write \( p(x) \). Note that \( \mathbb{E}\{|X|^2|A_0\} = 0 \) in (48) because we set \( |X|^2 = 0 \) for \((X, Z) \in A_0\) from a consideration of no transmission.

For computational purposes, we separate the complex random variables \( X \) and \( Z \) into real part and imaginary part as
\[
X = X_R + jX_I \quad \text{and} \quad Z = Z_R + jZ_I,
\]
and change the random variable from \( X \) to \( R = \sqrt{X_R^2 + X_I^2} \in (0, \infty) \) and \( \Theta = \tan^{-1}(X_I/X_R) \) where \( p_R(r) = \frac{1}{\pi r} e^{-r^2/\sigma^2} \) and \( p_\Theta(\theta) = 1/(2\pi) \). Then, the event of \( A_1 \) is rewritten as
\[
\{ |X + Z|^2 \geq \gamma \} \Leftrightarrow \{ R^2 + 2(rZ_R \cos \Theta + Z_I \sin \Theta) + (Z_R^2 + Z_I^2) \geq \gamma \} \Leftrightarrow \{(R + Z_R)^2 \geq \gamma - Z_I^2\} =: \tilde{A}_1, \quad (51)
\]
where \( N_R := Z_R \cos \Theta + Z_I \sin \Theta \) and \( N_I := -Z_R \sin \Theta + Z_I \cos \Theta \). When \((X_R, X_I)\) and \((Z_R, Z_I)\) are initially aligned with the real and imaginary axes, \( Z_R \) and \( Z_I \) denote the noise components in the direction of \( R e^{j\Theta} \) and perpendicular to those of \( R e^{j\Theta} \), respectively. Note that \( Z_R \) and \( Z_I \) are two independent Gaussian variables, i.e., \( Z_R \sim \mathcal{N}(0, \sigma_R^2/2) \) and \( Z_I \sim \mathcal{N}(0, \sigma_I^2/2) \) due to circular symmetry assumed for noise modeling. Therefore, (49) can be rewritten in terms of \( R, Z_R, \) and \( Z_I \) given by
\[
\mathbb{E}\{|X|^2\} = p(\tilde{A}_1) \int_{\infty}^{\infty} \int_{\infty}^{\infty} \int_{0}^{\infty} r^2 \times p(r, \tilde{z}_R, \tilde{z}_I|\tilde{A}_1) d\tilde{z}_R d\tilde{z}_I
\]
\[
= \mathbb{E}\text{Re}\{X\} = \mathbb{E}\text{Re}\{X^2\} = \mathbb{E}\{X^2\} = \mathbb{E}\{R^2\}, \quad (52)
\]
where \( p(r, \tilde{z}_R, \tilde{z}_I|\tilde{A}_1) \) is defined similar to (50) and the nuisance variable \( \Theta \) is integrated out in the analysis. Since the variables \( R, Z_R, \) and \( Z_I \) are mutually independent, the joint pdf for \( R, Z_R, \) and \( Z_I \) can be factored as \( p(r, \tilde{z}_R, \tilde{z}_I) = p(r)p(\tilde{z}_R)p(\tilde{z}_I) \). Then, (53) can be rewritten as
\[
\mathbb{E}\{X^2\} = \int_{\infty}^{\infty} p(\tilde{z}_I) \int_{\infty}^{\infty} p(\tilde{z}_R) \times \int_{R+\tilde{z}_R^2 > \gamma - \tilde{z}_I^2} r^2 \cdot p(r)d\tilde{z}_R d\tilde{z}_I
\]
\[
= \int_{\tilde{z}_R^2 > \gamma} p(\tilde{z}_R) \int_{\infty}^{\infty} p(\tilde{z}_R) \times \int_{0}^{\tilde{z}_R^2} r^2 \cdot p(r)d\tilde{z}_R d\tilde{z}_I
\]
\[
+ \int_{\tilde{z}_R^2 < \gamma} p(\tilde{z}_R) \int_{\infty}^{\tilde{z}_R^2} p(\tilde{z}_R) \times \left( \int_{0}^{\tilde{z}_R^2} r^2 p(r)dr + \int_{\tilde{z}_R^2}^{\infty} r^2 p(r)dr \right) d\tilde{z}_R d\tilde{z}_I, \quad (54)
\]
where \( I_1 := \max\{0, -\tilde{z}_R - \sqrt{\gamma - \tilde{z}_I^2}\} \) and \( I_2 := \max\{0, -\tilde{z}_R + \sqrt{\gamma - \tilde{z}_I^2}\} \).

In (54), we consider two separated cases for \( \tilde{z}_I \): if \( \tilde{z}_I^2 > \gamma \), all values of \( \tilde{z}_R \) and \( r \) satisfy \( A_1 \) (i.e., the first term on the right-hand side of (54)), and if \( \tilde{z}_I^2 < \gamma \), there is a range of \( r \) for each \( \tilde{z}_R \) as to satisfy \( A_1 \) (i.e., the second term on the right-hand side of (54)). The first term of (54) can be evaluated as
\[
2\sigma^2 \left(1 - \Phi \left(\sqrt{\gamma/(2\sigma^2)}\right)\right). \quad \text{If} \ \tau \text{ denotes the second term of (54), we provide a complete form of} \ \tau \text{ in (55).}
\]

After some tedious integration, we obtain the analytical formula for the third term of (55), denoted as \( \tau' \), shown in (56).

Note that equation (54) depends on \( \sigma^2, \sigma_R^2 \), and \( \gamma \) as seen from equations (55) and (56). The result may look complicated, so we apply an approximation in high SNR regime (i.e., \( \sigma^2 \gg \sigma_R^2 \)) to derive a simple formula of \( \sigma^2, \sigma_R^2 \), and \( \gamma \), where we use \( \sigma^2 = G_m P \sigma_g^2 \) and \( \sigma_R^2 = \sigma_z/W_m \). In this case, (54) is given by
\[
\mathbb{E}\{X^2\} = e^{-\frac{\sigma}{\sigma_R^2}(\sigma^2 + \gamma)} + \frac{1}{2} e^{-\frac{\sigma}{\sigma_R^2}} - \gamma \left(e^{-\frac{\sigma}{\sigma_R^2}} - e^{-\frac{\sigma}{\sigma_R^2}}\right) = e^{-\frac{\sigma}{\sigma_R^2}}\left(\frac{G_m \sigma_g^2 + \gamma}{P_{FA}}\right) + \frac{\sigma^2 \ln(1/P_{FA})}{W_m} \left(\sqrt{P_{FA} - P_{FA}}\right) - \frac{\sigma^2 \ln(1/P_{FA})}{W_m} \left(\sqrt{P_{FA} - P_{FA}}\right), \quad (57)
\]
where \( \gamma = \sigma^2 \ln(1/P_{FA})/W_m \) and \( \int_{\gamma}^{\infty} r^2 \cdot p(r)dr = e^{-\frac{\sigma}{\sigma_R^2}}\left(\frac{G_m \sigma_g^2 + \gamma}{P_{FA}}\right) \).

B. Proof of Theorem 1

(P1) Since only the cosine function in the second term in (35) depends on the phase shift \( \omega_m, g_{\theta_i, \omega_m} \) becomes periodic with the period \( 2\pi \). Define \( c^+ = \pi(N-1)/N + \epsilon \) and \( c^- = \pi(N-1)/N - \epsilon \) for \( \epsilon > 0 \). Then, we show that the values of cosine function evaluated at \( c^+ \) and \( c^- \) are the same, i.e.,
\[
\cos ((p'-p)(\pi(N-1)/N - c^-)) = \cos ((p'-p)c) = \cos ((p'-p)(\pi(N-1)/N - c^+)).
\]

(P2) For the \( i \)-th codeword \( f_{i,m}^{(m)} \in \mathcal{F}_m \), recall that the summation index \( p \) in (35) is given by \( p = (i-1)N_m + q \) for \( 1 \leq q \leq N_m \). From the center of the main beam range in (39) (i.e., \( -1 + (2i-1)N_m/N_i \)), we set the continuous angle as
\[
\theta^+ = -1 + (2i-1)N_m/N_i + \epsilon \quad (58)
\]
for \( \epsilon > 0 \). In (33), we consider the argument of the sin and exponential functions that correspond to \( \theta \) given by
\[
p = \frac{N_i(\theta^+ + 1) + 1}{2} = q - \frac{1 + N_m + \epsilon N_i}{2}, \quad (59)
\]
where \( 1 \leq q \leq N_m \). Then, we define \( p' := N_m(2i-1) + 1 - p = -q + iN_m + 1 \in \{i-1,N_m + 1, \ldots, iN_m\} \) that gives a one-to-one correspondence between \( p \) and
\[ \tau = 2\sigma^2 \int_{-\sqrt{\gamma}}^{\sqrt{\gamma}} p(\tilde{z}_I) \cdot \left( 1 - \Phi \left( \frac{\gamma - \tilde{z}_I^2}{\sigma^2} \right) \right) d\tilde{z}_I \]

\[ + \int_{-\sqrt{\gamma}}^{\sqrt{\gamma}} p(\tilde{z}_I) \left[ - \int_{-\infty}^{-\sqrt{\gamma - \tilde{z}_I^2}} p(R). e^{-\left( -z_R - \sqrt{\gamma - \tilde{z}_I^2} \right)^2} \left( \sigma^2 + \left( -\tilde{z}_I^2 + \sqrt{\gamma - \tilde{z}_I^2} \right)^2 \right) d\tilde{z}_R \right] d\tilde{z}_I \]

\[ + \int_{-\infty}^{-\sqrt{\gamma - \tilde{z}_I^2}} p(\tilde{z}_R) \cdot e^{-\left( -z_R + \sqrt{\gamma - \tilde{z}_I^2} \right)^2} \left( \sigma^2 + \left( -\tilde{z}_I^2 - \sqrt{\gamma - \tilde{z}_I^2} \right)^2 \right) d\tilde{z}_R \]

\[ = 2\sigma^2 \left( -1 + \Phi \left( \frac{\gamma - \tilde{z}_I^2}{\sigma^2} \right) + \frac{1}{2} e^{-\frac{\gamma - \tilde{z}_I^2}{\sigma^2}} \right) e^{-\frac{\gamma - \tilde{z}_I^2}{\sigma^2}} \left[ \frac{1}{\sigma^2} \right] \]

\[ + \int_{-\sqrt{\gamma}}^{\sqrt{\gamma}} P(\tilde{z}_I) \cdot e^{-\frac{\gamma - \tilde{z}_I^2}{\sigma^2}} \left( \frac{1}{\sigma^2} \right) \]

\[ \text{erf}(\cdot) \text{ denotes the error function} \ [32, \text{Eq. 8.250.1}]. \]

\[ \tau' = \frac{1}{\sigma^2} \left[ \left( \frac{1}{\sigma^2} \right) - \left( \frac{1}{\sigma^2} \right) \right] \]

\[ - \left( \frac{\sigma^2}{2} \right) \left( \frac{\sigma^2}{2} \right) \frac{1}{\sigma^2} \left( \frac{1}{\sigma^2} \right) \]

\[ \text{Note from (60) that for each term in (59) derived at } \theta^+ \text{, there is the same term with a negative sign derived at } \theta^- \text{. Then, we have} \]

\[ u_i(\theta^-)^H f_i^{(m)} = e^{j\omega_i(N_m(2i-1)+1)} \left( u_i(\theta^+)^H f_i^{(m)} \right)^H, \]

\[ \text{because for the summand in (33), two minus signs on the sin function cancel each other and the phase of the exponential function changes from negative to positive. Since the function } f_i^{(m)} \text{ is obtained from the square of (33), } f_i^{(m)} = f_i^{(m)} \text{ holds.} \]

\[ \text{(P.3) Given any } i \in \{1, \ldots, N/N_m\}, \text{ the function } f_i^{(m)} \text{ is evaluated on the spatial angle in one-half of the main beam interval in (39) by (P.2), e.g., set } \theta = -1 + (2i - 1)N_m/N_t + \epsilon \in [-1 + (2i - 1)N_m/N_t, -1 + (2iN_m - 1)/N_t] \text{ as in (58). For the } i \text{-th codeword, we consider the indices } p = (i - 1)N_m + q \text{ and } p' = (i - 1)N_m + q' \text{ for } q, q' \in \{1, \ldots, N_m\}. \text{ Then, } f_{\theta, i, \omega_m}^{(m)} \text{ can be rewritten as} \]

\[ f_{\theta, i, \omega_m}^{(m)} = \frac{1}{N_t N_m} \left( \sum_{q=1}^{N_m} \sin^2 \left( \frac{\pi q}{N_t} \right) \right) \]

\[ + \sum_{q=1}^{N_m} \sum_{q'=1}^{N_m} 2 \cos \left( \frac{N - 1}{N} \right) \left( \frac{N - 1}{N} \right) \]

\[ \times \sin \left( \frac{\pi q}{N_t} \right) \sin \left( \frac{\pi q'}{N_t} \right) \]

\[ \text{where } \tilde{q} = q - (N_m + \epsilon N_t - 1)/2 \text{ and } \tilde{q}' = q' - (N_m + \epsilon N_t - 1)/2. \text{ Since the function } f_{\theta, i, \omega_m}^{(m)} \text{ does not depend on the index } i, \text{ we have the claim.} \]

\[ \text{REFERENCES} \]
