Beamforming Codebook Optimization for Angle-of-Arrival Estimation

Nadim Ghaddar*, Lele Wang*, and Wei Yu†
*University of British Columbia, Vancouver, Canada, {nadim.ghaddar, lelewang}@ece.ubc.ca
†University of Toronto, Toronto, Canada, weiyu@ece.utoronto.ca

***Abstract*—This paper presents a novel optimization framework for designing beamforming codebooks for angle-of-arrival (AoA) estimation, by utilizing a close relation between the AoA estimation problem and channel coding. Specifically, we optimize the beamforming codebooks under two different metrics depending on whether the fading coefficient is known: the minimum Euclidean (ℓ_2) distance between the beamforming codewords, or the minimum chordal distance between the subspaces spanned by the codewords. In both cases, it is shown that the optimization problems can be formulated as a rank-constrained semidefinite program, which can be approached via a convex iteration method. The proposed optimization framework works for any number of antennas, number of channel measurements or desired angle resolution. Simulation results demonstrate the advantage of the proposed codebook designs as compared to existing beamforming strategies.**

I. Introduction

With the advent of millimeter-wave (mmWave) communications, and the rise of integrated sensing and communications (ISAC) [1], [2], the need for highly-directional signaling becomes crucial for mitigating the effect of severe path loss and blockage. Realizing such directional beams requires precise and reliable estimates of the channel during the initial access phase. This need is further compelled by practical hardware constraints that often limit the number of radio-frequency (RF) chains at the base station (BS), thus restricting channel measurements to be made through low-dimensional linear beamformers. This paper treats the simplest scenario, in which the BS is equipped with a single RF chain and aims to estimate a channel dominated by a single line-of-sight path. Specifically, we focus on the channel estimation task of learning the angle of arrival (AoA) of the dominant path.

Several designs of the sequence of beamformers (i.e., the beamforming codebook) have been proposed in the literature. Exhaustive beam sweeping (e.g., [3]), which scans over all possible beam directions before selecting the best one, requires a large pilot training overhead that grows linearly with the desired angle resolution. To reduce the training overhead, [4] proposes a design of the beamforming codebook based on random hash functions and shows that for such a random design, a logarithmic number of channel measurements (in the angle resolution) is sufficient in the high signal-to-noise ratio (SNR) regime. Designs based on compressive sensing that exploit the sparse nature of the channel in the angular domain have also shown a similar logarithmic scaling when used in conjunction with randomly constructed beamforming codebooks [5], [6]. The pilot training overhead can be further reduced by adaptively designing the beamformers based on previous measurements [7]–[12], but it incurs additional implementation complexity.

The logarithmic scaling of the number of measurements can be also inferred from a recently established connection between AoA estimation and channel coding [13], [14]. In particular, [13] shows that, when the fading coefficient of the single-path channel is known at the BS, the AoA estimation problem can be posed as a coding problem over a Gaussian channel. In light of this observation, a beamforming codebook based on Reed-Muller codes is constructed for the AoA estimation problem. However, the construction in [13] requires that the number of antennas is at least as large as the desired angle resolution, which is not necessarily the case in practical applications [15]–[17]. On the other hand, when the fading coefficient is unknown (and not estimated), the work in [14] draws a connection between the antenna selection problem (which is one way of approaching AoA estimation) and the coding problem over non-coherent channels. Based on this observation, [14] proposes to use a subspace distance metric for beamforming codebook design and presents an antenna selection method based on Golomb rulers [18]. However, the construction in [14] works only when the desired angle resolution is roughly equal to the square of the number of channel measurements.

Motivated by the aforementioned connection between AoA estimation and channel coding, this paper proposes a methodology for designing beamforming codebooks based on two different distance metrics. In particular, when the fading coefficient is known (or can be estimated), we formulate an optimization problem that maximizes the minimum Euclidean distance between the beamforming codewords. Indeed, in the coherent setting, a large minimum Euclidean distance guarantees optimal coding performance in the high SNR regime, and thus high AoA estimation accuracy. On the other hand, when the fading coefficient is unknown (and not estimated), we design the beamforming codebook to maximize the chordal distance between the *subspaces* spanned by the beamforming codewords. A large chordal distance between codeword subspaces guarantees optimal coding performance in the high SNR regime for non-coherent communications [19], which translates to a performance guarantee for AoA estimation [14]. We show in this paper that for both cases, the optimization problems can be formulated as a rank-constrained semidefinite program (SDP), which can be approached via a convex iteration method. The solutions to the optimization problems give rise to beamforming codebooks with guaranteed minimum distance properties. The optimization framework is general, in the sense that no restrictions are made on the angle resolution, the number of antennas or the number of measurements.

II. Angle-of-Arrival Estimation Problem

A. System Model

Consider a mmWave communication system in which a base station (BS) equipped with N antennas and a single RF chain serves a single-antenna user. To establish a reliable link between the BS and the user, an uplink training phase is used, in which the user sends T pilot symbols within the coherence interval in order to estimate the channel. Due to the single RF chain limitation, the BS can only observe the received pilot signals through a low-dimensional beamformer (i.e., combiner) $\mathbf{w}_t \in \mathbb{C}^N$ in each symbol period. The received symbols at the BS over the T symbol periods can be expressed as

$$y_t = \sqrt{P}\mathbf{w}_t^{\mathsf{H}}\mathbf{h} + \mathbf{w}_t^{\mathsf{H}}\mathbf{z}_t, \qquad t = 1, \dots, T, \tag{1}$$

where $\mathbf{h} \in \mathbb{C}^N$ is a vector of channel gains between the user and the BS, and $\mathbf{z}_t \sim \mathcal{CN}(0, \mathbf{I})$ is independent Gaussian noise.

The channel $\mathbf{h}$ is modeled by a single line-of-sight path and takes the form [20, Chapter 7]

$$\mathbf{h} = \alpha\mathbf{a}(\phi), \tag{2}$$

where ϕ is the AoA, $\alpha \sim \mathcal{CN}(0, 1)$ is the small-scale fading coefficient, and $\mathbf{a}(\cdot)$ is the array response vector. In this paper, we consider two separate settings, either α is known, or α is unknown. We consider a uniform linear array configuration with half-wavelength antenna spacing for which the array response vector is

$$\mathbf{a}(\phi) = \begin{bmatrix} 1 & e^{j\pi \sin\phi} & \cdots & e^{j(N-1)\pi \sin\phi} \end{bmatrix}^{\mathsf{T}}. \tag{3}$$

This paper considers the simplifying assumption that the AoA belongs to a grid of M points in the angular space, i.e., we assume that ϕ is uniformly distributed over the set

$$\Theta = \left\{ \arcsin\left(-1 + \frac{2i-1}{M} \right) : i = 1, \dots, M \right\}, \tag{4}$$

which is a set of M angles spanning the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Under this assumption, the collection of received signals $\mathbf{y} = (y_1, \dots, y_T)$ can be expressed as

$$\mathbf{y} = \alpha\sqrt{P}\mathbf{W}^{\mathsf{H}}\mathbf{D}\mathbf{u} + \mathrm{diag}(\mathbf{W}^{\mathsf{H}}\mathbf{Z}), \tag{5}$$

where $\mathbf{u} \in \{0, 1\}^M$ is a 1-sparse vector, and

$$\mathbf{D} = \begin{bmatrix} | & & | \\ \mathbf{d}_1 & \dots & \mathbf{d}_M \\ | & & | \end{bmatrix}, \tag{6}$$

is the collection of all array response vectors $\mathbf{d}_i \triangleq \mathbf{a}(\theta_i)$ with $\theta_i = \arcsin\left(-1 + \frac{2i-1}{M}\right)$ being the ith angle in the grid Θ. We also write $\mathbf{W} = \begin{bmatrix} \mathbf{w}_1 & \cdots & \mathbf{w}_T \end{bmatrix} \in \mathbb{C}^{N \times T}$ and $\mathbf{Z} = \begin{bmatrix} \mathbf{z}_1 & \cdots & \mathbf{z}_T \end{bmatrix}$. The matrix $\mathbf{W}^{\mathsf{H}}\mathbf{D}$ is referred to as the *beamforming codebook*, and its columns as the associated *codewords*. Note that, for a given beamforming matrix $\mathbf{W}$, the effective noise term $\mathbf{n} \triangleq \mathrm{diag}(\mathbf{W}^{\mathsf{H}}\mathbf{Z})$ is a zero-mean, circularly symmetric complex Gaussian vector with covariance matrix $\mathbf{\Gamma_W} = \mathrm{diag}(\|\mathbf{w}_1\|^2, \dots, \|\mathbf{w}_T\|^2)$.

This paper focuses on *non-adaptive* beamforming strategies, i.e., the beamforming matrix $\mathbf{W}$ is designed a priori before any symbols are received, and hence, can only depend on the array response matrix $\mathbf{D}$. As such, the goal of this paper is to design the beamforming matrix $\mathbf{W}$ to obtain the best estimate $\hat{\phi}$ of the AoA from the received symbol vector $\mathbf{y}$. In this paper, we formulate the AoA estimation problem as a classification task over the grid Θ. In this case, the performance of a particular beamforming codebook design is measured via the *detection error probability* defined as $P_e = \mathbb{P}(\hat{\phi} \neq \phi)$.

B. Connection to Channel Coding

The connection between the AoA estimation problem and the channel coding problem has been previously pointed out in [13], [14]. In the channel coding setting, a transmitter aims to communicate a message S to a receiver over a complex-valued additive white Gaussian noise (AWGN) channel, which can be modeled by

$$\tilde{\mathbf{y}} = \alpha\sqrt{P}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}, \tag{7}$$

where $\tilde{\mathbf{x}} \in \mathbb{C}^T$ is a codeword representing the message S, $\alpha \in \mathbb{C}$ is the fading coefficient and $\tilde{\mathbf{n}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is Gaussian noise. The receiver produces an $\hat{S}$ based on the received symbol vector $\tilde{\mathbf{y}}$. The performance is measured by the decoding error probability $\tilde{P}_e = \mathbb{P}(\hat{S} \neq S)$.

By contrasting (7) with the channel model in (5), one can observe that the codeword $\tilde{\mathbf{x}}$ in the channel coding problem corresponds to $\mathbf{W}^{\mathsf{H}}\mathbf{D}\mathbf{u}$ in the AoA estimation problem. In this context, the message S to be decoded in the channel coding problem represents the AoA ϕ to be estimated. Despite this correspondence, there are two key distinctions between the two problems. Firstly, the AoA estimation problem imposes a structure on the codewords; namely, the codewords are restricted to take the form $\{\mathbf{W}^{\mathsf{H}}\mathbf{d}_i\}_{i=1}^M$ (where $\mathbf{W}$ is a fixed beamforming matrix to be designed and $\mathbf{d}_i$ is given by the antenna array geometry) whereas no such structure is naturally imposed in the channel coding problem. Secondly, the choice of the beamforming matrix $\mathbf{W}$ in the channel model (5) affects the noise term $\mathbf{n}$, whose covariance matrix is given by $\mathbf{\Gamma_W}$, whereas the noise term $\tilde{\mathbf{n}}$ in the channel coding problem (7) is isotropic and independent of the transmitted codeword.

C. Distance Metrics

Motivated by the connection between AoA estimation and channel coding, this paper proposes two constructions of the beamforming matrix that optimize two different distance metrics pertaining to the beamforming codebook: 1) the minimum *Euclidean* (ℓ_2) distance between the codewords, and 2) the minimum *chordal* distance between the subspaces spanned by the codewords, depending on whether the fading coefficient α is known. The proposed constructions utilize the known array response structure, while taking into account the effect on the additive noise term.

These two metrics are inspired by channel coding for coherent and non-coherent communications, respectively. In the coherent setting, the fading coefficient α is assumed to be known (i.e., the receiver has access to an accurate estimate of it). It is well known that the minimum Euclidean distance largely dictates the performance of a code over AWGN channels in the coherent setting (especially in the high SNR regime). In contrast, in the non-coherent setting, the fading coefficient α is unknown (i.e., the receiver operates without estimating it), in which case its effect is to scale and rotate the transmitted codeword, preserving only the *subspace* that the signal vector spans rather than its absolute position. In the high SNR regime, it is well understood that maximizing the minimum chordal distance between the codeword subspaces minimizes the error probability of detecting an incorrect subspace under a maximum likelihood detector [19]. In the following sections, we formulate two different codebook optimization problems for the two metrics.

III. Known α: Maximizing Euclidean Distance

When the fading coefficient α is known, the AoA estimation problem boils down to a channel coding problem where the encoded sequences have the form $\{\mathbf{W}^{\mathsf{H}}\mathbf{d}_i\}_{i=1}^{M}$, and the noise realizations have a covariance matrix given by $\mathbf{\Gamma_W} = \mathrm{diag}(\|\mathbf{w}_1\|^2, \ldots, \|\mathbf{w}_T\|^2)$. We design the beamforming matrix $\mathbf{W}$ in this coherent setting to maximize the minimum Euclidean distance between the codewords. Since the choice of $\mathbf{W}$ affects the noise term only through the covariance matrix $\mathbf{\Gamma_W}$, there is no loss of generality in designing the beamformers to have unit norm. In particular, we consider the following optimization problem in the coherent setting:

$$(\mathsf{P}_1): \quad \underset{\mathbf{W}}{\text{maximize}} \quad \min_{i\neq j} \left\|\mathbf{\Gamma}_\mathbf{W}^{-1/2}\mathbf{W}^{\mathsf{H}}\big(\mathbf{d}_i - \mathbf{d}_j\big)\right\|^2, \tag{8}$$

where the multiplication by $\mathbf{\Gamma}_\mathbf{W}^{-1/2}$ induces unit-norm beamformers. Using the epigraph formulation, we can write the optimization problem (P_1) as

$$\underset{\mathbf{W},\gamma}{\text{maximize}} \quad \gamma \tag{9a}$$

$$\text{subject to} \quad \big(\mathbf{d}_i - \mathbf{d}_j\big)^{\mathsf{H}}\mathbf{W}\mathbf{\Gamma}_\mathbf{W}^{-1}\mathbf{W}^{\mathsf{H}}\big(\mathbf{d}_i - \mathbf{d}_j\big) \geq \gamma, \quad \forall\, i \neq j, \tag{9b}$$

where the maximization in (9) is over $(\mathbf{W}, \gamma)$. To solve the optimization problem (9), we consider two cases.

1) Case 1 ($T \geq N$): When $T \geq N$, the optimization problem (9) can be equivalently written as a semidefinite program (SDP). In particular, by denoting $\mathbf{R} = \mathbf{W}\mathbf{\Gamma}_\mathbf{W}^{-1}\mathbf{W}^{\mathsf{H}}$, we can write (9) as

$$\underset{\mathbf{R},\gamma}{\text{maximize}} \quad \gamma \tag{10a}$$

$$\text{subject to} \quad \big(\mathbf{d}_i - \mathbf{d}_j\big)^{\mathsf{H}}\mathbf{R}\big(\mathbf{d}_i - \mathbf{d}_j\big) \geq \gamma, \quad \forall\, i \neq j, \tag{10b}$$

$$\mathrm{tr}(\mathbf{R}) = T, \quad \mathbf{R} \succeq 0, \tag{10c}$$

where the trace constraint holds since $\mathrm{tr}(\mathbf{W}\mathbf{\Gamma}_\mathbf{W}^{-1}\mathbf{W}^{\mathsf{H}}) = T$ for any matrix $\mathbf{W}$. Problem (10) is an SDP, whose solution $(\mathbf{R}^*, \gamma^*)$ can be found via numerical solvers.

To find the optimal beamforming matrix, we need to find a matrix $\mathbf{W}^*$ such that $\mathbf{R}^* = \mathbf{W}^*\mathbf{\Gamma}_{\mathbf{W}^*}^{-1}(\mathbf{W}^*)^{\mathsf{H}}$. Note that spectral decomposition of $\mathbf{R}^*$ does not work on its own, because if $\mathbf{R}^* = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathsf{H}}$ is the spectral decomposition, and $\mathbf{V} = \mathbf{U}\mathbf{\Lambda}^{1/2}$, then $\mathbf{V}\mathbf{\Gamma}_\mathbf{V}^{-1}\mathbf{V}^{\mathsf{H}} = \mathbf{U}\mathbf{U}^{\mathsf{H}} \neq \mathbf{R}^*$. Instead, we propose the following technique to achieve the desired matrix factorization, which is motivated by the Schur–Horn Theorem [21], [22]. Let $\mathbf{F}$ denote the $T \times T$ discrete Fourier transform (DFT) matrix, where $[\mathbf{F}]_{k,\ell} = \frac{1}{\sqrt{T}}\omega^{(k-1)(\ell-1)}$ for $k, \ell = 1, \ldots, T$ and $\omega = e^{-j\frac{2\pi}{T}}$, and let

$$\mathbf{W}^* = \mathbf{U}\begin{bmatrix}\mathbf{\Lambda}^{1/2} & \mathbf{0}_{N\times(T-N)}\end{bmatrix}\mathbf{F}, \tag{11}$$

where $\mathbf{0}_{N\times(T-N)}$ denotes the all-zero matrix. It can be checked that $\mathbf{W}^*\mathbf{\Gamma}_{\mathbf{W}^*}^{-1}(\mathbf{W}^*)^{\mathsf{H}} = \mathbf{R}^*$, and all columns of $\mathbf{W}^*$ have unit norm. This follows because multiplying by the DFT matrix spreads the total "energy" in the spectrum of $\mathbf{R}^*$ (which is equal to $\mathrm{tr}(\mathbf{R}^*) = T$) equally across all columns. The choice of $\mathbf{F}$ may not be unique, but the DFT matrix has the advantage of being a universal choice that works for all $\mathbf{R}^*$. The proposed factorization gives an optimal solution $\mathbf{W}^*$ of the optimization problem (9) when $T \geq N$.

2) Case 2 ($T < N$): When $T < N$, the optimization problem (9) can be equivalently written as a *rank-constrained* SDP. As before, we denote $\mathbf{R} = \mathbf{W}\mathbf{\Gamma_W}\mathbf{W}^{\mathsf{H}}$. In addition to previous constraints, since $T < N$, the matrix $\mathbf{R}$ is constrained to have a rank at most T. Hence, in this case, the optimization problem (9) can be equivalently written as

$$\underset{\mathbf{R},\gamma}{\text{maximize}} \quad \gamma \tag{12a}$$

$$\text{subject to} \quad \big(\mathbf{d}_i - \mathbf{d}_j\big)^{\mathsf{H}}\mathbf{R}\big(\mathbf{d}_i - \mathbf{d}_j\big) \geq \gamma, \quad \forall\, i \neq j, \tag{12b}$$

$$\mathrm{rank}(\mathbf{R}) \leq T, \tag{12c}$$

$$\mathrm{tr}(\mathbf{R}) = T, \quad \mathbf{R} \succeq 0. \tag{12d}$$

To solve the optimization problem (12), we first express it as a feasibility problem for a given parameter γ. That is, for a fixed value of γ, we search the space of $N \times N$ matrices to find a matrix $\mathbf{R}$ that satisfies constraints (12b)–(12d). In other words, for fixed γ, we aim to solve the following rank-constrained feasibility problem:

$$\text{find} \quad \mathbf{R} \tag{13a}$$

$$\text{subject to} \quad \big(\mathbf{d}_i - \mathbf{d}_j\big)^{\mathsf{H}}\mathbf{R}\big(\mathbf{d}_i - \mathbf{d}_j\big) \geq \gamma, \quad \forall\, i \neq j, \tag{13b}$$

$$\mathrm{rank}(\mathbf{R}) \leq T, \tag{13c}$$

$$\mathrm{tr}(\mathbf{R}) = T, \quad \mathbf{R} \succeq 0. \tag{13d}$$

Despite the non-convex rank constraint, the feasibility problem (13) can be approached via a convex iteration method [23, Chapter 4.5] that expresses the problem as an iteration of the following two (convex) SDP's:

$$\underset{\mathbf{R}}{\text{minimize}} \quad \mathrm{tr}(\mathbf{R}\mathbf{Q}^*) \tag{14a}$$

$$\text{subject to} \quad \big(\mathbf{d}_i - \mathbf{d}_j\big)^{\mathsf{H}}\mathbf{R}\big(\mathbf{d}_i - \mathbf{d}_j\big) \geq \gamma, \quad \forall\, i \neq j, \tag{14b}$$

$$\mathrm{tr}(\mathbf{R}) = T, \quad \mathbf{R} \succeq 0, \tag{14c}$$

where $\mathbf{Q}^*$ is the solution to the SDP given by

$$\underset{\mathbf{Q}}{\text{minimize}} \quad \mathrm{tr}(\mathbf{R}^*\mathbf{Q}) \tag{15a}$$

$$\text{subject to} \quad \mathbf{0} \preceq \mathbf{Q} \preceq \mathbf{I}, \tag{15b}$$

$$\mathrm{tr}(\mathbf{Q}) = N - T, \tag{15c}$$

where $\mathbf{R}^*$ is the solution of (14). For a given γ, the feasibility problem (13) can be solved by iterating (14) and (15) for some given number of iterations $I_{\max}$. Note that problem (15) can be solved in closed form, and the solution is $\mathbf{Q}^* = \mathbf{V}^*(\mathbf{V}^*)^{\mathsf{H}}$ where $\mathbf{V}^* \in \mathbb{C}^{N\times(N-T)}$ is a matrix whose columns are the eigenvectors corresponding to the smallest $N-T$ eigenvalues of $\mathbf{R}^*$. Hence, the matrix $\mathbf{Q}^*$ points in the direction of rank-T (or less) matrices whose null space contains that of $\mathbf{R}^*$.

The inspiration behind this convex iteration method is the fact that, for a semidefinite feasibility problem, a solution of least rank must be an extreme point of the feasible set. Thus, there must exist a hyperlane that supports the feasible set at that extreme point, and hence, there must exist a linear objective function that is minimized at this least-rank solution. The goal of the iteration between (14) and (15) is to find the matrix $\mathbf{Q}^*$ that describes this linear function.

Note that if the value of the objective function after the i-th iteration of (14) and (15) is denoted by $\beta_i \triangleq \operatorname{tr}(\mathbf{R}^*\mathbf{Q}^*)$, then the sequence $\{\beta_i\}_{i=1}^{I_{\max}}$ is guaranteed to be non-increasing and convergent. Moreover, if this sequence converges to 0, then a rank-T solution is guaranteed to exist for the optimization problem (14), and hence for the feasibility problem (13) [23]. This implies the existence of a beamforming codebook with a squared minimum distance equal to the value of γ that we started with. Hence, in this case, γ can be increased, and the same convex iteration method can be used to search for a beamforming codebook with a squared minimum distance equal to the new value of γ. Otherwise, if no solution of rank T is found, γ is reduced. This procedure continues until the value of γ can be narrowed down to a sufficiently small interval. Finally, let $\mathbf{R}_{\text{opt}}$ be the rank-T matrix found through this process. The corresponding optimal beamforming matrix can be obtained by setting

$$\mathbf{W}^* = \mathbf{V}\mathbf{\Lambda}^{1/2}\mathbf{F}, \tag{16}$$

where $\mathbf{R}_{\text{opt}} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathsf{H}}$ is the (thin) spectral decomposition of $\mathbf{R}_{\text{opt}}$, and $\mathbf{F}$ is the $T\times T$ DFT matrix.

Remark 1. The complexity of the proposed design is largely dictated by the $\binom{M}{2}$ constraints in (14b) and the number of antennas N. This complexity may be high for large M and N, but is acceptable for practical values such as $N = 16$ and $M = 128$, because the codebook design is done offline.

Remark 2. The beamformer design proposed in this section assumes that an accurate estimate of α is available at the receiver. To estimate α, we can use a fading coefficient estimation phase consisting of T_{pilot} symbols that precede the measurements made for AoA estimation. During this phase, the beamforming vector can be set to $\tilde{\mathbf{w}} = \begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix} \in \mathbb{C}^N$, for which the received symbol can be expressed as

$$\tilde{y}_t = \alpha\sqrt{P} + \tilde{n}_t, \tag{17}$$

for $t = 1,\ldots,T_{\text{pilot}}$. Since (17) is independent of the AoA ϕ, the received symbols during this phase can be used to compute an estimate of α. In particular, the minimum mean-squared error (MMSE) estimate of α given the received symbols $\tilde{y}_1,\ldots,\tilde{y}_{T_{\text{pilot}}}$ is given by $\hat{\alpha} = \frac{\sqrt{P}}{T_{\text{pilot}}P+1}\sum_{t=1}^{T_{\text{pilot}}}\tilde{y}_t$.

IV. Unknown α: Maximizing Chordal Distance

In contrast to coherent communications, coding for non-coherent settings (often referred to as Grassmannian signaling [19], [24]–[26]) aims to ensure that the codewords remain distinguishable despite the scaling/rotation caused by the unknown fading coefficient α. The use of the squared chordal distance between codeword subspaces as a metric for AoA estimation is first proposed in [14]. For two codewords $\mathbf{c}_i$ and $\mathbf{c}_j$, the chordal distance between their subspaces is defined as

$$d_c(\mathbf{c}_i,\mathbf{c}_j) = \sqrt{1 - \frac{|\langle \mathbf{c}_i,\mathbf{c}_j\rangle|^2}{\|\mathbf{c}_i\|^2\|\mathbf{c}_j\|^2}}, \tag{18}$$

where $\langle\cdot,\cdot\rangle$ denotes complex inner product. Note that the chordal distance corresponds to the sine of the angle between the two vectors.

It can be seen from (18) that maximizing the chordal distance corresponds to minimizing the square of the cosine similarity between the two codewords. Hence, in the non-coherent setting, maximizing the minimum chordal distance between any two subspaces is equivalent to minimizing the maximum cosine similarity between any two codewords:

$$(\mathsf{P}_2): \underset{\mathbf{W}}{\text{minimize}} \quad \max_{i\neq j} \frac{\left|\langle \mathbf{\Gamma}_{\mathbf{W}}^{-1/2}\mathbf{W}^{\mathsf{H}}\mathbf{d}_i, \mathbf{\Gamma}_{\mathbf{W}}^{-1/2}\mathbf{W}^{\mathsf{H}}\mathbf{d}_j\rangle\right|^2}{\|\mathbf{\Gamma}_{\mathbf{W}}^{-1/2}\mathbf{W}^{\mathsf{H}}\mathbf{d}_i\|^2\|\mathbf{\Gamma}_{\mathbf{W}}^{-1/2}\mathbf{W}^{\mathsf{H}}\mathbf{d}_j\|^2}, \tag{19}$$

where the multiplication by $\mathbf{\Gamma}_{\mathbf{W}}^{-1/2}$ induces unit-norm beamformers. By denoting $\mathbf{R} = \mathbf{W}\mathbf{\Gamma}_{\mathbf{W}}^{-1}\mathbf{W}^{\mathsf{H}}$ and using the epigraph formulation, we can write the optimization problem (P_2) for $T < N$ as

$$\underset{\mathbf{R},\gamma}{\text{minimize}} \quad \gamma \tag{20a}$$

$$\text{subject to} \quad \frac{\left|\mathbf{d}_i^{\mathsf{H}}\mathbf{R}\mathbf{d}_j\right|^2}{(\mathbf{d}_i^{\mathsf{H}}\mathbf{R}\mathbf{d}_i)(\mathbf{d}_j^{\mathsf{H}}\mathbf{R}\mathbf{d}_j)} \le \gamma, \quad \forall i \neq j, \tag{20b}$$

$$\operatorname{rank}(\mathbf{R}) \le T, \tag{20c}$$

$$\operatorname{tr}(\mathbf{R}) = T, \quad \mathbf{R} \succeq 0. \tag{20d}$$

Note that condition (20b) can be equivalently expressed using the Schur complement relation as

$$\begin{bmatrix} \sqrt{\gamma}\mathbf{d}_i^{\mathsf{H}}\mathbf{R}\mathbf{d}_i & \mathbf{d}_i^{\mathsf{H}}\mathbf{R}\mathbf{d}_j \\ \mathbf{d}_j^{\mathsf{H}}\mathbf{R}\mathbf{d}_i & \sqrt{\gamma}\mathbf{d}_j^{\mathsf{H}}\mathbf{R}\mathbf{d}_j \end{bmatrix} \succeq 0, \quad \forall i \neq j. \tag{21}$$

Therefore, optimization problem (20) is a rank-constrained SDP, which can be solved using a similar approach as in Section III. Namely, the convex iteration method is used to search for a matrix $\mathbf{R}$ with a given squared cosine similarity γ, and the parameter γ is updated based on whether a rank-T solution is found through convex iteration or not. In this process, the constraint (14b) in the convex iteration is replaced with the constraint (21).

Remark 3. Note that, if some $\mathbf{d}_i$ lies in the null space of $\mathbf{W}^{\mathsf{H}}$, the cosine similarity in (19) is not well-defined. In our simulations, we observe that this is typically not the case at the optimal $\mathbf{W}^*$. (This is true either with constraint (20b) or equivalently with constraint (21).)

V. Numerical Experiments

The main objective of this paper is to show that it is possible to use the proposed optimization methodology to design beamforming codebooks for the AoA estimation problem. To this end, we explicitly construct practical codebooks under the two different metrics for a scenario in which the BS is equipped with $N = 16$ antennas, the AoA estimation takes place over a coherence interval of $T = 8$ symbol periods, and the AoA belongs to a grid of size $M = 128$. We consider the maximum a posteriori (MAP) estimator of ϕ which, when α is known (or estimated), can be expressed as[1]

$$\hat{\phi}^{(1)}_{\mathrm{MAP}}(\mathbf{y}) = \underset{\phi\in\Theta}{\arg\min} \ \left\|\mathbf{y} - \alpha\sqrt{P}\mathbf{W}^{\mathsf{H}}\mathbf{a}(\phi)\right\|^2, \tag{22}$$

and when α is unknown, can be expressed as

$$\hat{\phi}^{(2)}_{\mathrm{MAP}}(\mathbf{y}) = \underset{\phi\in\Theta}{\arg\max} \ \frac{\exp\left\{\frac{P\left|\mathbf{y}^{\mathsf{H}}\mathbf{W}^{\mathsf{H}}\mathbf{a}(\phi)\right|^2}{1+P\|\mathbf{W}^{\mathsf{H}}\mathbf{a}(\phi)\|^2}\right\}}{1+P\|\mathbf{W}^{\mathsf{H}}\mathbf{a}(\phi)\|^2} \tag{23}$$

The proposed beamforming codebooks are compared with two conventional beamforming strategies for AoA estimation: 1) Random Gaussian beamforming, and 2) Bayesian Cramér-Rao bound (CRB) beamforming. In random Gaussian beamforming, the beamformers are randomly and independently chosen according to a complex Gaussian distribution with unit variance. For the Bayesian CRB method, the beamformers are designed to minimize the Bayesian CRB metric [12], which is a lower bound on the MSE of any unbiased estimator. A uniform prior distribution of the AoA over the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$ is used to derive the Bayesian CRB metric. The beamforming strategies are compared for the same received $\mathrm{SNR} = |\alpha|^2 P$. Note that, since $T < N$, the convex iteration method is used for the proposed constructions, for which we set $I_{\max} = 5$ iterations.

Fig. 1 shows the performance of the different beamforming strategies in the case when the fading coefficient α is perfectly known at the BS. Clearly, the proposed design that maximizes the minimum Euclidean distance outperforms the other beamforming strategies. This is expected since, in the coherent setting, the performance of a codebook is largely dictated by its minimum Euclidean distance.

Fig. 2 shows the performance of the beamforming strategies when α is unknown to the BS. To estimate α, the BS employs the fading coefficient estimation scheme described in Remark 2. Note that the beamforming strategies are compared for the same total number of measurements $T = 8$. That is, if $T_{\mathrm{pilot}} = 2$ symbols are used for estimating α, then the remaining $T - T_{\mathrm{pilot}} = 6$ symbols are used for AoA estimation. Fig. 2 shows that the proposed design that maximizes the minimum chordal distance between the codeword subspaces has a comparable performance as the design that maximizes the minimum Euclidean distance while using $T_{\mathrm{pilot}} = 4$ symbols to estimate α prior to AoA estimation, and both significantly outperform the baseline schemes. These results show that the proposed codebook constructions are viable in practice, and outperform the state-of-the-art methods in the literature.

[1]Equations (22) and (23) are derived for the case when the beamformers have unit norm. If that is not the case, proper normalization of the beamformers and the received symbols should be done before using these equations.

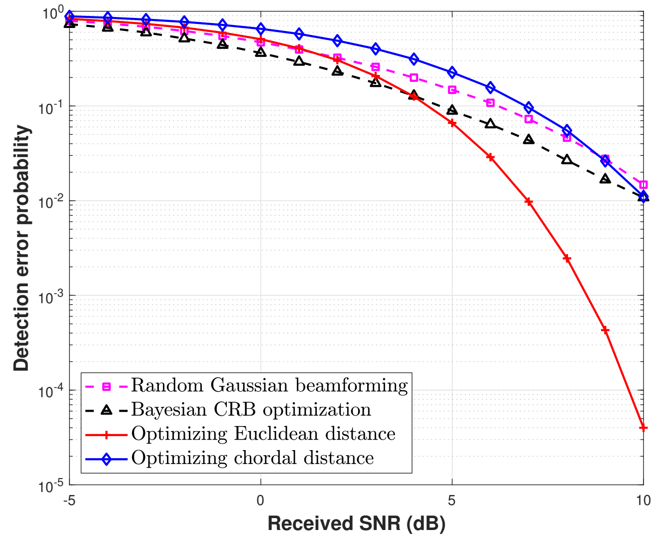

Fig. 1. Performance of the proposed beamforming codebooks for a system with $N = 16$ antennas, $M = 128$ possible AoAs, and $T = 8$ measurements, for the case when the fading coefficient α is perfectly known at the BS.

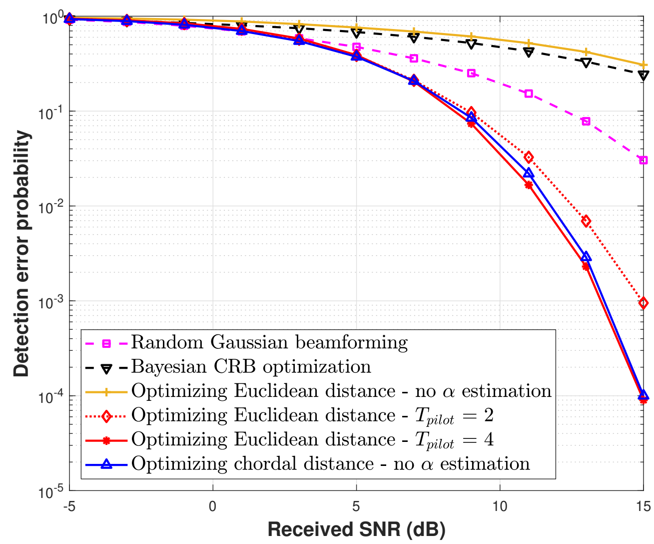

Fig. 2. Performance of the proposed beamforming codebooks for a system with $N = 16$ antennas, $M = 128$ possible AoAs, and $T = 8$ measurements, for the case when the fading coefficient α is unknown at the BS.

VI. Conclusion

In this paper, we develop a framework for optimizing beamforming codebooks for AoA estimation under two metrics: the minimum Euclidean distance between codewords, and the minimum chordal distance between subspaces spanned by the codewords. The proposed codebook designs work for all underlying system parameters and have superior performance compared to conventional beamforming strategies.

References

[1] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, 2020.

[2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6g and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] T. Nitsche, C. Cordeiro, A. Flores, E. Knightly, E. Perahia, and J. Widmer, "IEEE 802.11ad: Directional 60 GHz communication for multi-Gigabit-per-second Wi-Fi," *IEEE Commun. Mag.*, vol. 52, pp. 132–141, 12 2014.

[4] O. Abari, H. Hassanieh, M. Rodriguez, and D. Katabi, "Millimeter wave communications: From point-to-point links to agile network connections," in *Proc. ACM Workshop Topics Netw.*, 2016, p. 169–175.

[5] Y. Ding, S.-E. Chiu, and B. D. Rao, "Bayesian channel estimation algorithms for massive MIMO systems with hybrid analog-digital processing and low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 499–513, 2018.

[6] X. Song, S. Haghighatshoar, and G. Caire, "Efficient beam alignment for millimeter wave single-carrier systems with hybrid MIMO transceivers," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1518–1533, 2019.

[7] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[8] S.-E. Chiu, N. Ronquillo, and T. Javidi, "Active learning and CSI acquisition for mmWave initial alignment," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2474–2489, 2019.

[9] W. Huleihel, J. Tabrikian, and R. Shavit, "Optimal adaptive waveform design for cognitive MIMO radar," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 5075–5089, 2013.

[10] F. Sohrabi, Z. Chen, and W. Yu, "Deep active learning approach to adaptive beamforming for mmWave initial alignment," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2347–2360, 2021.

[11] F. Sohrabi, T. Jiang, W. Cui, and W. Yu, "Active sensing for communications by learning," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1780–1794, 2022.

[12] N. Ghaddar and W. Yu, "Active uplink sensing beamformer design via Bayesian Cramér-Rao Bound dual optimization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2025, pp. 5736–5741.

[13] N. Ghaddar, L. Wang, and W. Yu, "On-grid angle-of-arrival estimation in large-scale MIMO systems using channel codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2025, pp. 1–6.

[14] H. Mahdavifar, R. Rajamäki, and P. Pal, "Subspace coding for spatial sensing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2024, pp. 2394–2399.

[15] M. Li and Y. Lu, "Angle-of-arrival estimation for localization and communication in wireless networks," in *Eur. Signal Process. Conf. (EUSIPCO)*, 2008, pp. 1–5.

[16] Z. Zhang, T. Jiang, and W. Yu, "Localization with reconfigurable intelligent surface: An active sensing approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7698–7711, 2024.

[17] H. Han, T. Jiang, and W. Yu, "Active sensing for multiuser beam tracking with reconfigurable intelligent surface," *IEEE Trans. Wireless Commun.*, vol. 24, no. 1, pp. 540–554, 2025.

[18] S. Sidon, "Ein satz über trigonometrische polynome und seine anwendung in der theorie der fourier-reihen," *Mathematische Annalen*, vol. 106, no. 1, pp. 536–539, 1932.

[19] D. Agrawal, T. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2618–2626, 2001.

[20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[21] I. Schur, "Über eine klasse von mittelbildungen mit anwendungen auf die determinantentheorie," *Sitzungsber. Berl. Math. Ges.*, vol. 22, pp. 9–20, 1923.

[22] A. Horn, "Doubly stochastic matrices and the diagonal of a rotation matrix," *Amer. J. Math.*, vol. 76, pp. 620–630, 1954.

[23] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*. Meboo Publishing, 2005.

[24] D. Love, R. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.

[25] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, 1999.

[26] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, 2000.