On-Grid Angle-of-Arrival Estimation in Large-Scale MIMO Systems Using Channel Codes

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

***Abstract*—This paper presents a novel technique to design receive beamformers for on-grid angle-of-arrival (AoA) estimation in large-scale multiple-input multiple-output systems using channel codes. Specifically, the receive beamformers are designed so that the measurement model is effectively transformed to a Gaussian channel whose inputs are codewords in a channel code, with each codeword corresponding to a different AoA on the grid. Assuming that the number of antennas is larger than the desired angle resolution in the grid, the AoAs can be recovered by leveraging a suitable decoder on the resulting equivalent channel. The performance of the proposed method is derived in terms of the performance of the underlying channel code. Simulations results demonstrate the advantage of the proposed approach compared to existing beamforming strategies.**

I. Introduction

The angle-of-arrival (AoA) estimation problem entails estimating the direction upon which signals impinging on an antenna array are received. This is a classical problem in signal processing with many practical applications, including, but not limited to, target tracking [1], beam alignment [2], localization [3]–[5], and downlink beamforming in time-division duplex (TDD) systems [6]. In order to ensure accurate AoA estimation through highly directional beams, sixth-generation (6G) wireless networks are expected to employ extremely large-scale multiple-input multiple-output (XL-MIMO) systems with a huge number of antennas and an enlarged array aperture [7], but typically with limited number of radio-frequency (RF) chains. In such systems, the number of antennas is in the order of the desired angle resolution (but the number of RF chains is much lower).

The straightforward way of designing directional beams for AoA estimation is through exhaustive beam sweeping (e.g. [8]), which scans each possible direction through a dedicated beamformer, and outputs the direction which results in the highest received power. This is inefficient as the number of required measurements needs to grow linearly with the desired angle resolution. Later works on beamforming-based AoA estimation strategies show that the number of measurements can, in fact, grow only logarithmically with the angle resolution. For example, the work [9] proposes a non-adaptive beamforming strategy based on random hash functions and shows that a logarithmic number of measurements (in the angle resolution) is sufficient in the high signal-to-noise-ratio (SNR) regime. The same logarithmic scaling is obtained by exploiting the sparse nature of the problem in the angular domain using compressed sensing techniques, as done in [10], [11]. Adaptive beamforming strategies have also been applied to AoA estimation and show good performance with fewer number of measurements, e.g. in [12], [13], where a hierarchical codebook of beamformers is used, in [14], where the Bayesian Cramér-Rao bound (CRB) is optimized to design the beamformers, and in [15], [16], where a learning-based design is proposed. However, a common issue with all these methods is that resolving multiple AoAs is very challenging.

This paper takes a different approach to AoA estimation by leveraging a connection between designing beamforming vectors for AoA estimation and channel coding for reliable communication, under the simplifying assumption that the AoA's belong to a finite-size grid. The connection between these two problems has been previously pointed out in the literature through their common relation to the noisy searching problem with feedback [13], [17], where the goal is to estimate the location of a target by sequentially querying possible regions based on responses to previous queries. The application of channel coding to AoA estimation is also explored in [18], [19], where the beamforming vectors are designed to have particular beampatterns which are chosen based on a given channel code. However, in these works, it is not clear how the performance of the underlying coding scheme translates to a performance guarantee in the AoA estimation problem.

In this paper, we provide an explicit method to design receive beamformers for on-grid AoA estimation using point-to-point channel codes designed for Gaussian channels. Through successive decoding, the proposed method is able to resolve multiple AoAs. For estimating a single angle, the proposed method effectively transforms the measurement model to a Gaussian channel whose inputs are codewords in the code and whose noise statistics depend on the chosen beamformers. Through a careful analysis of the structure of the array response, the performance of the proposed strategy is bounded in terms of the performance of the channel code when simulated over the Gaussian channel. The performance guarantee holds as long as the number of antennas is larger than the desired angle resolution. This makes the proposed approach suitable for XL-MIMO systems in next-generation wireless networks.

II. Angle-of-Arrival Estimation Problem

Consider a mmWave communication channel in which a base station (BS) equipped with N antennas and Q RF chains

($Q < N$) aims to communicate with a single-antenna user. In order to estimate the channel between the user and the BS, an uplink pilot training phase consisting of T time frames is used, in which the user sends T pilot symbols $\{x_t\}_{t=1}^T$ with power P, i.e., $x_t = \sqrt{P}$, $\forall t$. Due to the limited number of RF chains, the BS can observe the pilot symbols only through a lower-dimensional analog beamformer $\mathbf{W}_t \in \mathbb{C}^{N\times Q}$. Therefore, the received symbol at time t, $1 \leq t \leq T$, can be expressed as

$$\mathbf{y}_t = \sqrt{P}\mathbf{W}_t^{\mathsf{H}}\mathbf{h} + \mathbf{W}_t^{\mathsf{H}}\mathbf{z}_t, \tag{1}$$

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 the Gaussian noise.

We assume that the channel between the BS and the user can be modeled geometrically by K paths [20], i.e.,

$$\mathbf{h} = \frac{1}{\sqrt{K}}\sum_{k=1}^{K}\alpha_k\mathbf{a}(\phi_k), \tag{2}$$

where ϕ_k is the AoA of the kth path, $\alpha_k \sim \mathcal{CN}(0, 1)$ is the small-scale fading coefficient affecting the kth path, and $\mathbf{a}(\cdot)$ is the array response vector. We consider a uniform linear array configuration with N antenna elements and half-wavelength antenna spacing for which the array response vector is

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

For many applications, including target tracking [1], beam alignment [2], localization [3]–[5], and downlink beamforming in TDD systems [6], it is crucial that the BS computes precise estimates of the AoAs $\{\phi_1, \ldots, \phi_K\}$ from the received symbols. Towards this end, we consider the simplifying assumption that each AoA ϕ_k belongs to a grid of M points in the angular space, i.e., we assume that $\phi \in \Theta$, where

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

which is a set of M angles spanning the region $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Furthermore, this paper considers the setting of XL-MIMO systems, in which case we assume that the number of antennas N is larger than the number of possible angles M.

This paper focuses on *non-adaptive* beamforming strategies, i.e., the beamformers $(\mathbf{W}_1, \ldots, \mathbf{W}_T)$ are chosen a priori before any symbols are received, and hence, can only depend on the array response $\mathbf{a}(\cdot)$. In this setting, a system model with a BS equipped with Q RF chains making T pilot measurements is mathematically equivalent to that of a BS equipped with a single RF chain making TQ pilot measurements. Hence, without loss of generality, this paper considers the special case of $Q = 1$, i.e., the BS is equipped with a single RF chain and uses T beamforming vectors $\mathbf{w}_1, \ldots, \mathbf{w}_T$.

The goal of this paper is to design the beamformers $\mathbf{w}_1, \ldots, \mathbf{w}_T$ to obtain the best estimates $\{\hat{\phi}_1, \ldots, \hat{\phi}_K\}$ of the AoAs from the received symbol vector $\mathbf{y} = (y_1, \ldots, y_T)$. The performance of such an estimation is measured by the mean-squared error (MSE) defined by

$$\text{MSE} = \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}\left[\left(\phi_k - \hat{\phi}_k\right)^2\right]. \tag{5}$$

To find good estimates, the pre-designed beamformers $(\mathbf{w}_1, \ldots, \mathbf{w}_T)$ should account for the known array response structure, while mitigating the effect of the additive noise. The main contribution of this paper is that, when $N \geq M$, such beamformers can be constructed by leveraging the error correcting capabilities of codes designed for Gaussian channels.

III. A Coding Perspective to AoA Estimation

This paper proposes to design beamforming strategies for AoA estimation using codes that are designed for the additive white Gaussian noise (AWGN) channel. In this section, we establish the connection between the two problems, which sets up the stage for the main result in the next section. For illustrative purposes, we make two simplifying assumptions in this and the subsequent sections. First, we consider a system model with a single dominant path (i.e., $K = 1$) in which case only one AoA ϕ is to be estimated. Second, we assume that the channel fading coefficient α is perfectly known at the BS. These assumptions are removed in Section V, where we consider the more practical setting in which multiple AoAs are to be estimated, and only the statistical information of the fading coefficients is available at the BS.

A. Code for AoA Estimation

To aid in the exposition, it is useful to define the notion of a *code* for AoA estimation as follows. An (M, N, T) code for AoA estimation consists of

- the set of angles Θ defined in (4) where $|\Theta| = M$,
- a sequence of beamformers $(\mathbf{w}_1, \ldots, \mathbf{w}_T)$ where $\mathbf{w}_j \in \mathbb{C}^N$,
- an estimator $f : \mathbb{C}^T \times \mathbb{C}^{N\times T} \to \Theta$ that assigns angle estimate $\hat{\phi}$ to each received sequence $\mathbf{y}$ and beamformer sequence $(\mathbf{w}_1, \ldots, \mathbf{w}_T)$.

The average probability of error of an AoA estimation code is $P_e^{(\text{AoA})} = \mathbb{P}(\hat{\phi} \neq \phi)$.

As we shall see later, the probability of error metric allows us to translate the performance guarantees from the channel coding problem to the AoA estimation problem. Nonetheless, in the simulation experiments of Section VI, the performance of an AoA estimation code is reported through the MSE defined in (5), which better matches the nature of the AoA estimation problem.

B. Coding for Binary-Input AWGN Channel

We begin by reviewing the channel coding problem over binary-input AWGN channels. Suppose a transmitter wishes to communicate a message S to a receiver over an AWGN channel with binary inputs, BI-AWGN(P), defined by

$$Y_t = \alpha\sqrt{P}X_t + N_t, \tag{6}$$

for $t = 1, \ldots, T$, where $X_t \in \{-1, +1\}$ denotes the channel input symbol at time t that encodes the message S, $\alpha \in \mathbb{C}$ is a channel gain that is known to the receiver, and $N_t \sim \mathcal{CN}(0, 1)$ is a sequence of i.i.d. circularly-symmetric complex-valued Gaussian random variables with unit variance. Upon receiving $Y_1, \ldots, Y_T$, the receiver wishes to find an estimate $\hat{S}$ of the transmitted message. The effective signal-to-noise ratio (SNR) in this model is $\gamma = |\alpha|^2 P$.

An (R, T) code for the BI-AWGN(P) channel consists of

- a message set $\mathcal{S} = [1 : 2^{RT}]$,
- an encoding function $f : \mathcal{S} \to \{-1, +1\}^T$ that assigns a codeword $\mathbf{x} = (x_1, \ldots, x_T)$ to each message $s \in \mathcal{S}$,
- a decoding function $g : \mathbb{C}^T \to \mathcal{S}$ that assigns a message estimate $\hat{s}$ to each received sequence $\mathbf{y} = (y_1, \ldots, y_T)$.

The message S is assumed to be uniformly distributed over the message set. The performance of the code is measured by the average probability of error $P_e^{(\mathrm{AWGN})} = \mathbb{P}(\hat{S} \neq S)$.

We define the channel codebook matrix $\mathbf{C} \in \{-1, +1\}^{T \times 2^{RT}}$ as the matrix whose columns are the codewords of the code, i.e.,

$$\mathbf{C} = \begin{bmatrix} | & & | \\ f(1) & \cdots & f(2^{RT}) \\ | & & | \end{bmatrix}, \tag{7}$$

where $f(\cdot)$ is the encoding function.

C. Connections Between AoA Estimation and Channel Coding

Now, we are ready to establish a connection between the AoA estimation problem and the coding problem over Gaussian channels. First, notice that, in the single path case ($K = 1$), the system model (1) for AoA estimation can be alternatively written as

$$y_t = \alpha\sqrt{P}\mathbf{w}_t^{\mathsf{H}}\mathbf{D}\mathbf{u} + \mathbf{w}_t^{\mathsf{H}}\mathbf{z}_t, \tag{8}$$

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

$$\mathbf{D} = \begin{bmatrix} | & & | \\ \mathbf{a}(\theta_1) & \ldots & \mathbf{a}(\theta_M) \\ | & & | \end{bmatrix}, \tag{9}$$

where $\theta_i = \arcsin\left(-1 + \frac{2i-1}{M}\right)$ is the ith angle in the set Θ defined in (4). Since $\phi = \theta_{k^*}$ for some k^*, the AoA estimation problem can now be cast as the problem of finding the support of $\mathbf{u}$ from the sequence of measurements $(y_1, \ldots, y_T)$ [15].

By contrasting (8) with the channel model in (6), one can make the following key connections between the two problems[1]:

1) The AoA ϕ can be represented by the message S in the channel coding problem.
2) The fading coefficient α, when perfectly known at the BS, plays the role of a channel gain in the coding problem.
3) The choice of the beamforming vector $\mathbf{w}_t$ can be made so that the baseband source signal $\mathbf{w}_t^{\mathsf{H}}\mathbf{D}\mathbf{u}$ is represented by the channel input symbol X_t in the coding problem.
4) The average probability of error is a common performance metric between the two problems.

However, there is a key distinction between the two problems. The distribution of the effective noise term $\tilde{n}_t \triangleq \mathbf{w}_t^{\mathsf{H}}\mathbf{z}_t$ in the AoA estimation problem is different than that of the noise term in the channel coding problem. In the following, we show that this distinction can be properly accounted for through a careful choice of the beamforming vectors $(\mathbf{w}_1, \ldots, \mathbf{w}_T)$ designed specifically with the structure of the array response matrix $\mathbf{D}$ in mind. This established connection between the two problems makes available a variety of coding techniques for AoA estimation in XL-MIMO systems. We present the main result of this paper in the next section.

IV. Main Result

The main result of this paper is the following theorem that shows how a code for the binary-input AWGN channel can be used to design a code for AoA estimation.

Theorem 1. *Let (M, N, T) be such that $N \geq M$, and let $R = \frac{\log_2 M}{T}$. Let $\epsilon > 0$. If there exists an (R, T) code for the BI-AWGN$\left(\lfloor \frac{N}{M} \rfloor P\right)$ channel whose average probability of error is $P_e^{(\mathrm{AWGN})} = \epsilon$, then we can construct an (M, N, T) code for the AoA estimation problem with average probability of error $P_e^{(\mathrm{AoA})} \leq \epsilon$.*

Proof. We give an explicit construction of the AoA estimation code using the code for the BI-AWGN$\left(\lfloor \frac{N}{M} \rfloor P\right)$ channel. Let $\mathbf{C} \in \{-1, +1\}^{T \times M}$ be the codebook matrix of the channel code, as defined in (7). Let $\mathbf{c}_1^{\mathsf{H}}, \ldots, \mathbf{c}_T^{\mathsf{H}}$ denote the rows of $\mathbf{C}$.

To construct a code for AoA estimation, consider the following optimization problem:

$$\underset{\mathbf{w}_t}{\text{minimize}} \quad \|\mathbf{w}_t\|_2 \tag{10a}$$

$$\text{subject to} \quad \mathbf{D}^H\mathbf{w}_t = \mathbf{c}_t. \tag{10b}$$

Since $N \geq M$, $\mathbf{D}^H$ has full row rank, so the solution to (10) is

$$\mathbf{w}_t^* = \left(\mathbf{D}^H\right)^{\dagger}\mathbf{c}_t, \tag{11}$$

where $(\cdot)^{\dagger}$ denotes the pseudoinverse of a matrix. For this choice of $\mathbf{w}_t^*$, the system model (8) can be written as

$$y_t = \alpha\sqrt{P}c_{t,k^*} + \tilde{n}_t, \tag{12}$$

where k^* denotes the position of the non-zero entry in $\mathbf{u}$, c_{t,k^*} denotes the k^*-th entry of $\mathbf{c}_t$, and $\tilde{n}_t = (\mathbf{w}_t^*)^H\mathbf{z}_t$ is a zero-mean Gaussian random variable with variance $\|\mathbf{w}_t^*\|_2^2$.

Notice the similarity of (12) with the BI-AWGN channel model given in (6). The only difference is that the effective SNR in (12) is

$$\gamma^{(\mathrm{AoA})} = \frac{|\alpha|^2 P}{\|\mathbf{w}_t^*\|_2^2}, \tag{13}$$

whereas for the channel code designed for a BI-AWGN$\left(\lfloor \frac{N}{M} \rfloor P\right)$ channel, the effective SNR is

$$\gamma^{(\mathrm{AWGN})} = |\alpha|^2 \left\lfloor \frac{N}{M} \right\rfloor P. \tag{14}$$

If we can show that $\gamma^{(\mathrm{AoA})} \geq \gamma^{(\mathrm{AWGN})}$, then one can leverage the decoder of the channel code to find an estimate of the codeword indexed by k^*. Hence, by (13) and (14), it is sufficient to show that

$$\|\mathbf{w}_t^*\|_2^2 \leq \frac{1}{\lfloor \frac{N}{M} \rfloor}. \tag{15}$$

[1] These connections are analogous to the connections that relate channel coding to the support recovery problem in the compressive sensing literature [21], with a key distinction as highlighted in this section.

To this end, notice that

$$\|\mathbf{w}_t^*\|_2^2 = \|(\mathbf{D}^H)^\dagger \mathbf{c}_t\|_2^2 \le \sigma_{\max}^2\left((\mathbf{D}^H)^\dagger\right)\|\mathbf{c}_t\|_2^2 = \frac{M}{\sigma_{\min}^2(\mathbf{D})}, \tag{16}$$

where $\sigma_{\max}(\cdot)$ and $\sigma_{\min}(\cdot)$ denote the largest and smallest singular value, respectively, and $\|\mathbf{c}_t\|_2^2 = M$ since each entry in $\mathbf{c}_t$ is a symbol in the matrix $\mathbf{C}$. Next, we show that

$$\sigma_{\min}^2(\mathbf{D}) = M\left\lfloor \frac{N}{M} \right\rfloor, \tag{17}$$

which implies the desired inequality (15).

One can show (17) by observing the resemblance of the matrix $\mathbf{D}$ to the discrete Fourier transform (DFT) matrix. In particular, notice that

$$\mathbf{D} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{N-1} & z_2^{N-1} & \cdots & z_M^{N-1} \end{bmatrix}, \tag{18}$$

where $z_i = e^{j\pi\left(-1+\frac{2i-1}{M}\right)}$ for $i = 1, \ldots, M$. Recall that the $M \times M$ DFT matrix is a matrix $\mathbf{F}$ with $[\mathbf{F}]_{i,j} = \frac{\omega^{(i-1)(j-1)}}{\sqrt{M}}$ for $i, j = 1, \ldots, M$, where $\omega = e^{-j\frac{2\pi}{M}}$. It can be checked that $\mathbf{D} = \mathbf{\Phi}\tilde{\mathbf{D}}$, where $\mathbf{\Phi}$ is an $N \times N$ unitary matrix, and

$$\tilde{\mathbf{D}}^{\mathsf{H}} = \sqrt{M}\,[\underbrace{\mathbf{F} \quad \mathbf{F} \quad \cdots \quad \mathbf{F}}_{\left\lfloor \frac{N}{M} \right\rfloor \text{ times}} \quad \mathbf{G}], \tag{19}$$

where $\mathbf{G}$ is the submatrix of $\mathbf{F}$ consisting of the first $N - M\left\lfloor \frac{N}{M} \right\rfloor$ columns. Note that, since $\mathbf{D} = \mathbf{\Phi}\tilde{\mathbf{D}}$ with unitary $\mathbf{\Phi}$, $\mathbf{D}$ and $\tilde{\mathbf{D}}$ have the same singular values. Furthermore,

$$\tilde{\mathbf{D}}^{\mathsf{H}}\tilde{\mathbf{D}} = M\left\lfloor \frac{N}{M} \right\rfloor \mathbf{F}\mathbf{F}^{\mathsf{H}} + M\mathbf{G}\mathbf{G}^{\mathsf{H}} = M\left\lfloor \frac{N}{M} \right\rfloor \mathbf{I}_M + M\mathbf{G}\mathbf{G}^{\mathsf{H}}, \tag{20}$$

where the last equality holds since $\mathbf{F}$ is a unitary matrix. It follows that the minimum eigenvalue of $\tilde{\mathbf{D}}^{\mathsf{H}}\tilde{\mathbf{D}}$ (or, equivalently, $\sigma_{\min}^2(\mathbf{D})$) is $M\left\lfloor \frac{N}{M} \right\rfloor$. In fact, the only other nonzero eigenvalue is $M\left\lceil \frac{N}{M} \right\rceil$. This implies the equality in (17), and thus, the inequality in (15), which in turn implies the fact that $\gamma^{(\mathrm{AoA})} \ge \gamma^{(\mathrm{AWGN})}$. Thus, the decoder of the channel code can be utilized over the channel model in (12) to recover an estimate $\hat{k}$ corresponding to the AoA. The output of the AoA estimator is $\hat{\phi} = \theta_{\hat{k}}$, and its average probability of error is at most that of the channel code, i.e., $P_e^{(\mathrm{AoA})} \le \epsilon$.

□

The proof of Theorem 1 gives an explicit construction of the beamformers and the AoA estimator based on a code for a BI-AWGN channel. The analysis suggests that when $N \ge M$, designing the beamformers to be in the range space of $(\mathbf{D}^H)^\dagger$ according to (11) boosts the SNR by at least a factor of $\left\lfloor \frac{N}{M} \right\rfloor$.

V. Practical Considerations

A. Unknown Channel Fading Coefficient

So far, we restricted attention to the case when the fading coefficient α is known to the BS. Now, we consider the practical scenario when α is unknown. One way to approach this setting is by using some of the received symbols to get an estimate $\hat{\alpha}$ of the fading coefficient, prior to AoA estimation. Then, the estimate $\hat{\alpha}$ is used in the AoA estimation phase.

In particular, we consider a fading coefficient estimation phase consisting of $\tilde{T}$ time frames that precede the measurements made for AoA estimation. Inspired by (11), the beamforming vectors in this phase are set to

$$\tilde{\mathbf{w}} = \left(\mathbf{D}^H\right)^\dagger \mathbf{1}_M, \tag{21}$$

for each $t = 1, \ldots, \tilde{T}$, where $\mathbf{1}_M$ is the all-ones vector of length M. For this choice of beamforming vector, the received symbol can be expressed as

$$\tilde{y}_t = \alpha\sqrt{P} + \tilde{n}_t, \tag{22}$$

which is independent of the AoA ϕ. Hence, the received symbols in this phase can be used to compute an estimate of the fading coefficient. For example, the minimum mean-squared error (MMSE) estimate of α given the received symbols $\tilde{y}_1, \ldots, \tilde{y}_{\tilde{T}}$ can be written as $\hat{\alpha} = \frac{\sqrt{P}}{\tilde{T}P + \|\tilde{\mathbf{w}}\|^2}\sum_{t=1}^{\tilde{T}} \tilde{y}_t$. This estimate can then be used for AoA estimation by regarding it as the true value of the fading coefficient.

B. Estimation of Multiple Angles

The proposed framework can be extended to the case where multiple angles are to be estimated. In particular, consider the setting in (1) in which the channel gain vector $\mathbf{h}$ is modeled with $K \ge 2$ paths, and there are K angles $\{\phi_1, \ldots, \phi_K\}$ to be estimated. In this setting, one can make the connection to the K-user multiple access channel (MAC), in the special case where the users of the MAC are restricted to use the same codebook. This is reminiscent of the connection between coding for MACs and the support recovery problem in compressive sensing [21]. However, practical codes and decoding strategies for such a MAC are difficult to design. Instead, one can use a single-user code and leverage a successive decoder for the MAC [22], in which case the single-path code described in previous sections can be applied. We remark that, when the fading coefficients $\alpha_1, \ldots, \alpha_K$ are unknown, it becomes more challenging to incorporate a training procedure to estimate the fading coefficients prior to AoA estimation.

VI. Numerical Experiments

In this section, we evaluate the performance of the proposed AoA estimation strategy over an XL-MIMO system equipped with $N = 1024$ antennas. We assume that the AoA is uniformly drawn over the grid set of size $M = 256$. The number of measurements made is $T = 16$. For the channel code, we use a Reed-Muller code [23], [24] with a rate $R = \frac{\log_2 M}{T} = \frac{1}{2}$. Since there are only $M = 256$ codewords in this code, we can efficiently implement the maximum likelihood (ML) decoder.

We compare the proposed strategy with two common beamforming strategies for AoA estimation: 1) Random Gaussian beamforming, and 2) Adaptive Bayesian CRB method. In random Gaussian beamforming, the beamformers are randomly and independently chosen according to a complex Gaussian distribution with unit variance, and an ML decoder is employed to recover the AoA. For the adaptive Bayesian

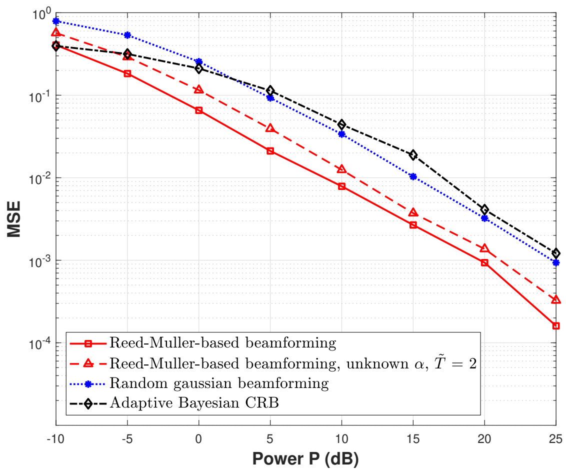

Fig. 1. Average MSE versus SNR for a single-path model with $N = 1024$ antennas, $M = 256$ possible AoAs, and $T = 16$ time frames.

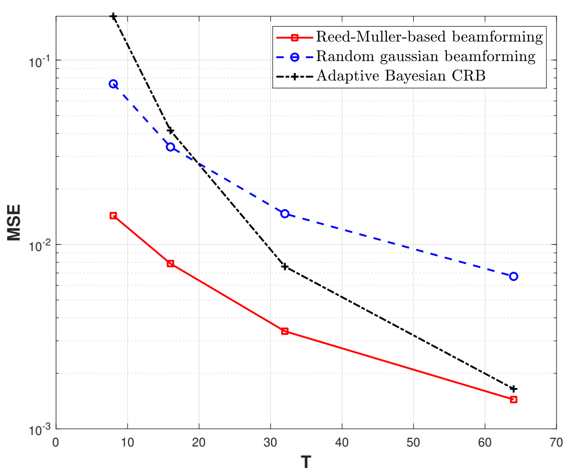

Fig. 2. Average MSE versus T for a single-path model with $N = 1024$ antennas and $M = 256$ possible AoAs at a power level $P = 10$ dB.

CRB method, the beamformers are designed sequentially to minimize the Bayesian CRB metric [14], which is a lower bound on the MSE of any unbiased estimator. In this case, we use an MMSE estimator to estimate the AoA.

Fig. 1 shows the plot of the average MSE for the different beamforming strategies versus the SNR in the case of a single dominant path (i.e., a single AoA to estimate) when the fading coefficient α is known. The performance of the proposed strategy when the fading coefficient α is unknown (Section V-A) is also shown, where $\tilde{T} = 2$ symbols are used for the estimation of α. The proposed Reed-Muller-based beamforming strategy achieves better AoA estimation performance compared to the other baseline strategies, even when the fading coefficient is unknown, and only two pilot symbols are used to estimate it. This highlights the potential of the proposed coding-based framework in AoA estimation.

Fig. 2 shows the average MSE performance of the beamforming strategies as a function of the number of measurements T for a single-path model and a fixed power level

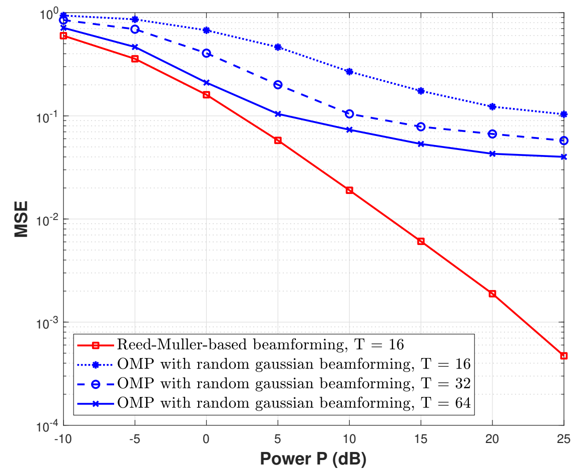

Fig. 3. Average MSE versus SNR for a system model with $N = 1024$ antennas, $M = 256$ possible AoAs, and $K = 2$ paths.

$P = 10$ dB. The results show that the baseline strategies require much more measurements to achieve a given MSE performance compared to Reed-Muller-based beamforming. For example, at an MSE of 8×10^{-3}, random Gaussian beamforming requires almost 4 times more measurements while the adaptive Bayesian CRB method requires around 2 times more measurements. This demonstrates that the proposed approach can significantly reduce the pilot overhead for AoA estimation.

Fig. 3 considers the average MSE performance of the beamforming strategies when there are two AoAs to be estimated (i.e., $K = 2$ paths). We assume that the two fading coefficients α_1 and α_2 are known at the BS. As highlighted in Section V-B, the proposed strategy in this case boils down to successive decoding of single-user codes, where the angle with the largest fading coefficient magnitude is decoded first. We again use a Reed-Muller code of rate $R = \frac{1}{2}$. For random Gaussian beamforming, we employ the orthogonal matching pursuit (OMP) [25] algorithm, a widely-used method in compressive sensing. As for the adaptive Bayesian CRB method, it has a high computational complexity due to the need to track the two-dimensional posterior distribution of the AoAs, which limits its practicality in the multipath scenario, particularly when the number of antennas is large ($N = 1024$ in our setup). The simulation results of Fig. 3 show a significant performance advantage of Reed-Muller-based beamforming compared to random Gaussian beamforming. This highlights the potential of the proposed beamforming approach in multipath settings.

VII. Concluding Remarks

In this paper, we develop a channel-coding-based technique to design receive beamformers for on-grid AoA estimation. The designed beamformers combine the received symbols at the antennas according to a prescribed channel codebook, while accounting for the structure of the array response. The proposed method shows superior performance compared to conventional methods, especially when multiple angles are to be estimated.

REFERENCES

[1] 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.

[2] T. Jiang and W. Yu, "Active sensing for reciprocal MIMO channels," *IEEE Trans. Signal Process.*, vol. 72, pp. 2905–2920, 2024.

[3] 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.

[4] 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.

[5] Z. Zhang and W. Yu, "Localization with reconfigurable intelligent surface: An active sensing approach," accepted in *IEEE Trans. Wireless Commun.*, 2025.

[6] L. Zhao, D. W. K. Ng, and J. Yuan, "Multi-user precoding and channel estimation for hybrid millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1576–1590, 2017.

[7] Z. Wang, J. Zhang, H. Du, D. Niyato, S. Cui, B. Ai, M. Debbah, K. B. Letaief, and H. V. Poor, "A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 3, pp. 1560–1605, 2024.

[8] 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 [invited paper]," *IEEE Commun. Mag.*, vol. 52, pp. 132–141, 12 2014.

[9] 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.*, ser. HotNets '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 169–175. [Online]. Available: https://doi.org/10.1145/3005745.3005766

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] S.-E. Chiu and T. Javidi, "Low complexity sequential search with size-dependent measurement noise," *IEEE Trans. Inf. Theory*, vol. 67, no. 9, pp. 5731–5748, 2021.

[18] Y. Shabara, C. E. Koksal, and E. Ekici, "Linear block coding for efficient beam discovery in millimeter wave communication networks," in *Proc. IEEE INFOCOM*, 2018, pp. 2285–2293.

[19] T. Zheng, J. Zhu, Q. Yu, Y. Yan, and L. Dai, "Coded beam training," 2024. [Online]. Available: https://arxiv.org/abs/2401.01673

[20] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, 2016.

[21] Y. Jin, Y.-H. Kim, and B. D. Rao, "Limits on support recovery of sparse signals via multiple-access communication techniques," *IEEE Trans. Inf. Theory*, vol. 57, no. 12, pp. 7877–7892, 2011.

[22] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.

[23] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *Trans. IRE Prof. Group Inf. Theory*, vol. 4, no. 4, pp. 38–49, 1954.

[24] D. E. Muller, "Application of boolean algebra to switching circuit design and to error detection," *Trans. IRE Prof. Group Electron. Comput.*, vol. EC-3, no. 3, pp. 6–12, 1954.

[25] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.