Linear Block Coding for Efficient Beam Discovery in Millimeter Wave Communication Networks

Yahia Shabara, C. Emre Koksal and Eylem Ekici
Department of Electrical and Computer Engineering
The Ohio State University, Columbus, Ohio 43210
Email: {shabara.1, koksal.2, ekici.2}@osu.edu

Abstract—The surge in mobile broadband data demands is expected to surpass the available spectrum capacity below 6 GHz. This expectation has prompted the exploration of millimeter wave (mm-wave) frequency bands as a candidate technology for next generation wireless networks. However, numerous challenges to deploying mm-wave communication systems, including channel estimation, need to be met before practical deployments are possible. This work addresses the mm-wave channel estimation problem and treats it as a beam discovery problem in which locating beams with strong path reflectors is analogous to locating errors in linear block codes. We show that a significantly small number of measurements (compared to the original dimensions of the channel matrix) is sufficient to reliably estimate the channel. We also show that this can be achieved using a simple and energy-efficient transceiver architecture.

I. Introduction

We investigate the problem of channel estimation in millimeter wave (mm-wave) wireless communication networks. Mm-wave refers to the wavelength of electromagnetic signals at 30-300 GHz frequency bands. At these high frequencies, channel measurement campaigns revealed that wireless communication channels exhibit very limited number of scattering clusters in the angular domain [1]–[3]. A *cluster* refers to a propagation path or continuum of paths that span a small interval of transmit Angles of Departure (AoD) and receive Angles of Arrival (AoA). Moreover, signal attenuation is very significant at mm-wave frequencies. This motivates the use of large antenna arrays at the transmitter (TX) and receiver (RX) to provide high antenna gains that compensate for high path losses [4]. Nevertheless, due to the high power consumption of mixed signal components, e.g., Analog to Digital Converters (ADCs) [5], conventional digital transceiver architectures that employ a complete RF chain per antenna is not practical. Hence, alternate architectures have been proposed for mm-wave radios with the objective of maintaining a close performance to channel capacity. Among the proposed solutions are the use of i) hybrid analog/digital beamforming [6]–[8] and ii) fully digital beamforming with low resolution ADCs [9]–[11].

For all proposed solutions, *channel estimation* remains one of the most critical determinants of performance in communication. Due to the large number of antennas at TX and RX,

estimation of the full channel gain matrix may require a large number of measurements, proportional to the product of the number of transmit and receive antennas. This imposes a great burden on the estimation process. To address this issue, various methods have been used, the most prevalent among them, is compressed sensing [6], [11]–[14], which leverages channel sparsity. Performance of compressed sensing based approaches is heavily dependent on the design of system (sensing) matrices. For instance, while random sensing matrices are known to perform well, in practice, sensing matrices involve the design of transmit and receive beamforming vectors and the choice of dictionary matrices, which are used to express channels in a sparse form. Hence, purely random matrices have not been used in practice [15]. On the other hand, no design that involves deterministic sensing matrices has been considered for sparse channel estimation.

Despite the efforts, we do not have a full understanding of the dependence of channel estimation performance on the channel parameters and number of measurements. In an effort to understand this relationship, the study in [16] proposed a multi-user mm-wave downlink framework based on compressed sensing in which the authors evaluate the achievable rate performance against the number of measurements.

In this work, we follow a different approach. We propose a systematic method in which we use sequences of error correction codes chosen in a way to control the channel estimation performance. To demonstrate our approach, consider the following simple example. Let a point to point communication channel be such that, there exists 3 possible receive AoA directions, only one of which may have a strong path to TX. We need to obtain the correct AoA at RX, if it exists. Instead of exhaustively searching all 3 possible AoA directions, we alternatively measure signals from combined directions. For instance, by combining directions 1&2 in one measurement and 2&3 in the next measurement, we can find the AoA in just two measurements. Specifically, four different scenarios might occur, namely, i) only the 1st, or ii) only the 2nd measurement contains a strong path, iii) both 1st and 2nd measurements contain a strong path, and finally, iv) neither measurement reveals a strong path. Interpretation of those cases is: AoA is in i) direction 1, ii) direction 3, iii) direction 2, and iv) none exists. Therefore, only 2 measurements are sufficient for beam detection instead of 3 that are needed for exhaustive search.

We will generalize this idea to develop a systematic method

for beam detection, inspired by linear block coding. Specifically, we show that linear block error correcting codes (LBC) possess favorable properties that fit in with the desirable behavior of sparse channel estimation. As a result, we are able to **i) provide rigorous criteria for solving the channel estimation problem**, **ii) significantly decrease the number of required measurements**, and **iii) utilize a fairly simple and energy-efficient transceiver architecture.** We design the system using LBCs that leverage the fact that transmission errors are typically sparse in transmitted data streams, and hence, only a few number of erroneous bits need to be corrected per transmitted codeword. Similarly, mm-wave channels are also sparse, i.e., only a small number of AoAs/AoDs carry strong signals. LBCs can correct sparse transmission errors by identifying their location in a transmitted sequence (followed by flipping them). We are inspired by LBC's ability to locate erroneous bits and exploit it to identify the AoAs/AoDs that carry strong signals (and their path gains) among all possible AoA/AoD values. To this end, we exploit hard decision decoding of LBCs, in which the receiver obtains an *error syndrome* that maps to one of the correctable error patterns. An obtained error pattern determines the positions where errors have occurred. Likewise, for channel estimation, the receiver will be designed to do a sequence of measurements that would result in a *channel syndrome*. The resultant channel syndrome shall identify the positions (and values) of non-zero angular channel components.

Contributions of this work can be summarized as follows:

- We set an analogy between beam discovery and channel coding to utilize low-complexity decoding techniques for efficient beam discovery.
- We provide rigorous criteria for setting the number of channel measurements based on the size of the channel and its sparsity level.
- We show that the number of measurements required for beam discovery is linked to the rate of a used linear block code. Hence, maximizing the rate of the underlying code is equivalent to minimizing the number of measurements.
- We develop a simple receiver architecture that enables us to measure signals arriving from multiple directions.

Related Work: The main objective of mm-wave channel estimation is to find a mechanism that can reliably estimate the channel using as few measurements as possible. For instance, in [6], a compressed sensing based algorithm to estimate single-path channels is proposed and an upper bound on its estimation error is derived. Further, the authors propose a multipath channel estimation algorithm based on that of single-path channels. The proposed algorithms in [6] use an adaptive approach with a hierarchical codebook (i.e., the set of all possible beamforming vectors) of increasing resolution. Similarly, the work in [13] proposes an adaptive compressive sensing channel estimation algorithm that accounts for off-the-grid AoAs and AoDs by using continuous basis pursuit [17] dictionaries. Such adaptive algorithms divide the estimation process into stages and demand frequent feed back to the

TX after each stage. Hence, while the number of required measurements are shown to decrease, these methods may add a considerable overhead.

Other works like [18] and [19] have proposed channel estimation algorithms using overlapped beam patterns. For instance, the algorithm in [19] can estimate multipath channel components by sequentially estimating each path gain using an algorithm designed to estimate single-path channels followed by recursively removing the estimated paths' effect from subsequent measurements. Similar to [6], [13] adaptive beams with increasing resolution that require feedback to TX are used to refine the AoA/AoD estimates.

On the other hand, in [18], a beam alignment algorithm is proposed assuming a multipath mm-wave channel. The presented algorithm, with a high probability, can find the best beam in a logarithmic number of measurements (with respect to the total number of available AoA directions). Nonetheless, despite the possible existence of multiple paths, the algorithm in [18] is designed to find one path to TX.

Notations: A vector and a matrix are denoted by $\boldsymbol{x}$ and $\boldsymbol{X}$, respectively, while x denotes a scalar or a complex number depending on the context. The transpose, conjugate transpose and frobenius norm of $\boldsymbol{X}$ are given by $\boldsymbol{X}^T$, $\boldsymbol{X}^H$ and $\|\boldsymbol{X}\|_F$, respectively. The sets of real and complex numbers are $\mathbb{R}$ and $\mathbb{C}$. The $k \times k$ identity matrix is $\boldsymbol{I_k}$. A set is denoted by $\mathcal{X}$, while $|\mathcal{X}|$ is its cardinality. Finally, $\mathbb{1}()$ is the indicator function.

II. MOTIVATING EXAMPLE

To elaborate, we present the following example: consider a point to point communication link between a TX with single antenna ($n_t = 1$) and RX with $n_r = 15$ antennas. Therefore, the vector of channel gains $\boldsymbol{q}$, is a 15×1 vector (let all the channels have one single significant tap), and its corresponding angular (virtual) channel, $\boldsymbol{q}^a$, is a vector of the same size and can be derived using the DFT matrix $\boldsymbol{U_r}$ as $\boldsymbol{q}^a = \boldsymbol{U_r^H} \boldsymbol{q}$ [20] (this is merely a linear transformation that maps the sequence of channel gains into a sequence of gains from different AoAs. This mapping will be presented in more detail in Section III). Assume a single-path channel, i.e., the channel has only one cluster with a single path in it. Let the path gain be denoted by α. For simplicity assume $\alpha = 1$. Further, let us assume perfect sparsity such that the AoA is along one of the directions defined in the DFT matrix $\boldsymbol{U_r}$, i.e., the channel path will only contribute to one angular bin. Finally, let us also neglect the channel noise.

Based on the channel description above, we get an angular channel vector of the form

$$\boldsymbol{q}^a = \begin{pmatrix} q_0^a & q_1^a & \cdots & q_{14}^a \end{pmatrix}^T, \tag{1}$$

such that $q_i^a \in \{0, 1\}$ and the number of non-zero elements in $\boldsymbol{q}^a$ is 1. Any component of $\boldsymbol{q}^a$ can be measured using one of the beam patterns shown in Fig. 1.

Objective: Suppose the transmitter sends pilot symbols of the form $x = 1$. Thus, the received vector $\boldsymbol{y}$ of size 15×1 can be obtained as

$$\boldsymbol{y} = \boldsymbol{q}x = \boldsymbol{q} \iff \boldsymbol{y}^a = \boldsymbol{q}^a \tag{2}$$
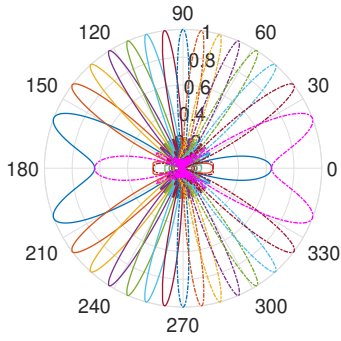
Fig. 1: Beam patterns of all possible angular directions

where $\boldsymbol{y}^a$ is the received vector in the angular domain. So, with change of basis, we can think of $\boldsymbol{q}^a$ as a received sequence with just one non-zero component. To identify the position of this non-zero component, the receiver performs a sequence of channel measurements. Let y_{s_i} denote the ith measurement such that

$$y_{s_i} = \boldsymbol{w}_i^H \boldsymbol{y} = \boldsymbol{w}_i^H \boldsymbol{q}, \quad (3)$$

where $\boldsymbol{w}_i$ denotes the ith receive (rx-)combining vector.

Our aim is to design channel measurements (i.e., $\boldsymbol{w_i}$'s) such that the correct AoA is identified using the minimum number of measurements.

Proposed Solution: We consider this non-zero component to be an anomaly to a normally all-zero 15-bin angular channel. Hence, the goal of identifying its position is analogous to finding the most likely 1-bit error pattern of a 15-bit codeword in a linear block code. Now, we need to identify an error correction code with codewords of length 15 and with 1-bit error correction capability [21]. Hence, we can use the binary $(15, 11, 3)$ Hamming code with parity check matrix $\boldsymbol{H}$ of size 4×15 and given by

$$\boldsymbol{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (4)$$

where $h_{i,j}$ represents the component at the intersection of row i and column j of $\boldsymbol{H}$. Using hard decision decoding of LBCs, error syndrome vectors of length 4 are obtained. Every possible syndrome vector maps to only one correctable error pattern[1]. Similarly, for channel estimation, several measurements should be performed at RX where each measurement mimics the behavior of a corresponding element in the error syndrome vector. Each measurement boils down to adding signals from a subset of the available 15 directions. Since each measurement can either include the direction of the incoming strong path of gain $\alpha = 1$ or no strong paths at all, then the elements of the channel syndrome vector are in $\{0, 1\}$.

For every measurement y_{s_i}, we design $\boldsymbol{w}_i$ based on the entries of the ith row of $\boldsymbol{H}$ such that: if $h_{i,j} = 1$, then we

---

[1]A correctable error pattern of a $(15, 11, 3)$ Hamming code is any 15×1 binary vector that contains only one '1' (at the error's position).

---

include the beam pattern that points to direction j in $\boldsymbol{w}_i$. For example, the 0th row of $\boldsymbol{H}$ is given by $[100010011010111]$. Hence, $\boldsymbol{w}_0$ should include beam patterns pointing to the set of directions $\{0, 4, 7, 8, 10, 12, 13, 14\}$.

Fig. 2 illustrates this operation for $\boldsymbol{w}_0$. We can see that the resultant beam pattern of $\boldsymbol{w}_i$ combines signals coming from a set of selected directions dictated by the ith row of $\boldsymbol{H}$. We call the obtained measurement vector, $\boldsymbol{y_s}$, the *channel syndrome* which is analogous to error syndromes in hard decision decoding of LBCs. Then, a table that maps every possible channel syndrome to a unique corresponding channel can be constructed.

In this example, we are able to estimate the channel based on only 4 measurements as opposed to 15, which is the number of measurements with exhaustive search. Important aspects of our proposed method include the choice of codes, the design of precoding and rx-combining measurement vectors, the effect of variable gains and phases of different paths and the occurrence of measurement errors.

Remark (Receiver Architecture). Note that, to achieve beam patterns similar to the one shown in Fig. 2, the receiver architecture needs to be a bit different from those of classical analog/hybrid beamforming architectures. Specifically, in addition to low-noise amplifiers (LNA) typically placed at the output of each antenna, we will need to add controllable low-power amplifiers, as well. The resultant architecture is still quite simple (see Fig. 3). That is, besides the low-power amplifiers, the proposed architecture is similar to those of simple analog beamforming. Moreover, low-resolution ADCs can be used which mitigates the high power consumption problem associated with high-resolution ADCs.

Motivation for LBC-inspired approach: LBCs are designed to discover and correct a certain maximum number of errors in a codeword of a specified length. This objective is achieved by adding redundant *parity check* bits to the original information sequence. What makes our devised approach attractive is that the number of measurements needed for channel estimation can be shown to be equal to the number of parity bits of some corresponding code. Hence, we can control the estimation performance via appropriate code selection. In this work, we will propose a method to specify the number of necessary channel measurements as a function of the rate of the underlying code.

III. System Model

Consider a point-to-point millimeter-wave wireless communication system with a transmitter (TX) equipped with n_t antennas and a receiver (RX) with n_r antennas placed at fixed locations. Uniform Linear Arrays (ULA) are assumed at both TX and RX where each antenna element is connected to a phase shifter and a variable gain amplifier. A single RF chain at the receiver, with in-phase (I) and quadrature (Q) channels, is fed through a linear combiner (see Fig. 3). Only two midtread ADCs, with $2^b + 1$ quantization levels, are utilized, i.e., quantization levels are $\{-2^{b-1}, \ldots, -1, 0, 1, \ldots, 2^{b-1}\}$.
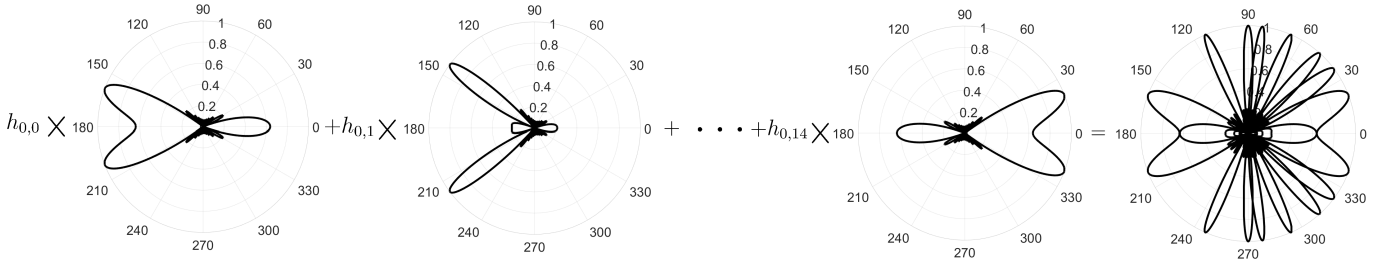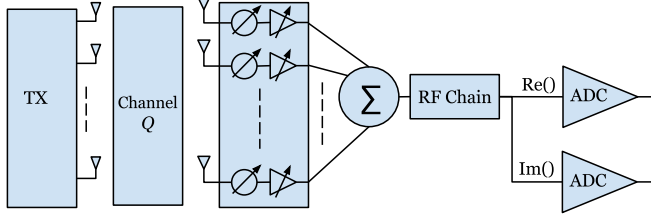
Fig. 2: Beam pattern of receive combining vector $\boldsymbol{w_0}$

Fig. 3: Hardware Block Diagram: Every antenna is connected to a phase shifter and low-power variable gain amplifier. Then, all outputs are combined using an adder and passed to an RF chain with in-phase and quadrature channels.

We adopt a single-tap channel model where $\boldsymbol{Q} \in \mathbb{C}^{n_r \times n_t}$ denotes the channel matrix between TX and RX. Assume that the channel has L clusters, where each cluster contains a single path with attenuation α_l, AoD θ_l, and AoA ϕ_l. The channel is assumed to be sparse such that $L \ll n_t, n_r$. Let $\alpha_l^b \in \mathbb{C}$ denote the baseband channel gain and is defined as

$$\alpha_l^b = \alpha_l \sqrt{n_t n_r} \, e^{-j \frac{2\pi \rho_l}{\lambda_c}} \qquad (5)$$

where ρ_l is the length of path l and λ_c is the carrier frequency. The angular cosines of AoD and AoA associated with path l are denoted by Ω_{tl} and Ω_{rl}, respectively. The transmit and receive spatial signatures along the direction Ω is given by $\boldsymbol{e_t}(\Omega)$ and $\boldsymbol{e_r}(\Omega)$ such that

$$\boldsymbol{e_t}(\Omega) = \frac{1}{\sqrt{n_t}} \begin{pmatrix} 1 \\ e^{-j2\pi\Delta_t\Omega} \\ e^{-j2\pi 2\Delta_t\Omega} \\ \vdots \\ e^{-j2\pi(n_t-1)\Delta_t\Omega} \end{pmatrix}, \qquad (6)$$

where $\boldsymbol{e_r}(\Omega)$ has a similar definition to $\boldsymbol{e_t}(\Omega)$, and Δ_t and Δ_r are the antenna separations at TX and RX normalized by the wavelength λ_c. Thus, $\boldsymbol{Q}$ can be written as

$$\boldsymbol{Q} = \sum_{l=1}^{L} \alpha_l^b \, \boldsymbol{e_r}(\Omega_{rl}) \, \boldsymbol{e_t}^H(\Omega_{tl}). \qquad (7)$$

We define $\boldsymbol{U_t}$ and $\boldsymbol{U_r}$ as the unitary Discrete Fourier Transform (DFT) matrices whose columns constitute an orthonor-mal basis for the transmit and receive signal spaces $\mathbb{C}^{n_t}$ and $\mathbb{C}^{n_r}$, respectively. $\boldsymbol{U_t}$ (and similarly $\boldsymbol{U_r}$) is given by

$$\boldsymbol{U_t} = \begin{pmatrix} \boldsymbol{e_t}(0) & \boldsymbol{e_t}(\frac{1}{L_t}) & \cdots & \boldsymbol{e_t}(\frac{n_t-1}{L_t}) \end{pmatrix}, \qquad (8)$$

where L_t and L_r are the normalized lengths of the transmit and receive antenna arrays such that $L_t = n_t \Delta_t$ and $L_r = n_r \Delta_r$. Let $\boldsymbol{Q^a}$ be the channel matrix in the angular domain [20], where

$$\boldsymbol{Q^a} = \boldsymbol{U_r^H} \boldsymbol{Q} \boldsymbol{U_t}. \qquad (9)$$

The rows and columns of $\boldsymbol{Q^a}$ divide the channel into re-solvable RX and TX bins, respectively. Further, we assume a perfect sparsity model in which AoDs θ_l, and AoA ϕ_l, are along the directions defined in $\boldsymbol{U_t}$ and $\boldsymbol{U_r}$ [6], [14], [19]. Hence, each channel cluster will only contribute to a single pair of TX and RX bins. Therefore, $\boldsymbol{Q^a}$ has a maximum of L non-zero TX and RX bins.

The baseband channel model is given by

$$\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{n}, \qquad (10)$$

where $\boldsymbol{x} \in \mathbb{C}^{n_t}$ is the transmitted signal, $\boldsymbol{y} \in \mathbb{C}^{n_r}$ is the received signal and $\boldsymbol{n} \sim \mathcal{CN}(\boldsymbol{0}, N_0 \boldsymbol{I}_{n_r})$ is an i.i.d. complex Gaussian noise vector.

Let $\boldsymbol{f} \in \mathbb{C}^{n_t}$ and $\boldsymbol{w} \in \mathbb{C}^{n_r}$ be the precoding and rx-combining vectors, respectively. The transmit signal $\boldsymbol{x}$ is given by $\boldsymbol{x} = \boldsymbol{f}s$ where s is the transmitted symbol with average power $\mathbb{E}(ss^H) = P$. After the receiver applies the rx-combining vector $\boldsymbol{w}$, the resultant symbol y_s can be written as

$$y_s = \boldsymbol{w}^H \boldsymbol{Q}\boldsymbol{f}s + \boldsymbol{w}^H \boldsymbol{n} \qquad (11)$$

where y_s constitutes a single unit measurement obtained using specific $\boldsymbol{f}$ and $\boldsymbol{w}$ vectors. We assume that $\boldsymbol{Q}$ remains fixed throughput the entire estimation process. The noise component $\boldsymbol{w}^H \boldsymbol{n}$ normalized by $\|\boldsymbol{w}\|$ is also a complex gaussian random variable such that $\frac{\boldsymbol{w}^H}{\|\boldsymbol{w}\|} \boldsymbol{n} \sim \mathcal{CN}(0, N_0)$. We define the signal to noise ratio (SNR) on a per path basis such that SNR of path l is given by

$$SNR_l = \frac{P}{N_0} |\alpha_l^b|^2. \qquad (12)$$

Note that the actual received SNR depends on all path gains included in a measurement. Finally, we define a *detectable path* as a path l for which $|\alpha_l^b|$ is such that the received signal power on that path is above a minimum detectable signal power (MDS).

IV. PROBLEM STATEMENT

Suppose a maximum number of L clusters need to be discovered in the channel where $L \ll n_t, n_r$. Under the prefect sparsity assumption, $\boldsymbol{Q}^a$ has a maximum of L non-zero RX and TX angular bins. Our objective is to identify the angular positions at which channel clusters exist and identify their path gain values using the **least possible number of measurements**. Let the number of measurements be m such that each measurement, $y_{s_{i,j}}$, is obtained using the precoder $\boldsymbol{f_j}$ and rx-combiner $\boldsymbol{w_i}$. Measurements take the form $y_{s_{i,j}} = \boldsymbol{w_i}^H \boldsymbol{Q} \boldsymbol{f_j} s + \boldsymbol{w_i}^H \boldsymbol{n}$. Let $\xi()$ be a mapping function that takes in the measurements $\{y_{s_{i,j}}\}_{\forall i,j}$ as inputs and returns the estimated channel $\widehat{\boldsymbol{Q}}^a$. We stack the measurements $\{y_{s_{i,j}}\}_{\forall j}$ in a single syndrome vector such that $\boldsymbol{y_{s_j}} = [y_{s_{0,j}} \; y_{s_{1,j}} \ldots y_{s_{m_1-1,j}}]^T$. Our design variables are the precoding vectors $\boldsymbol{f_j}$, rx-combining vectors $\boldsymbol{w_i}$, the number of measurements m, the mapping function $\xi()$, and the transmitted symbol power P.

In its essence, solving this problem boils down to finding the optimal set of measurements $\{y_{s_{i,j}}\}_{\forall i,j}$ and the mapping function $\xi()$ such that $\boldsymbol{Q}^a$ can be estimated using the minimum number of measurements. For ease of explanation, we first consider a channel with a single transmit antenna and n_r receive antennas. Therefore, no precoding is needed and the design of measurements is reduced to designing the rx-combining vectors $\boldsymbol{w_i}$. Recall that in the motivating example in Section II, we dealt with a special case of $n_r \times 1$ channels where we sought to find the direction of arrival of a channel with a single path of known gain, $\alpha = 1$. In the general case, we should consider arbitrary path gains $\alpha \in \mathbb{C}$ and channels with multiple paths.

V. BEAM DISCOVERY

In this section, we present our proposed solution. As an initial step, we solve a simplified version of the problem where communication channels have a single transmit antenna and multiple receive antennas. Afterwards, we will build on it to provide the solution for general channels with multiple transmit and receive antennas.

A. Beam Detection using LBC-inspired approach

To identify the exact number of measurements and their corresponding design, we follow a decoding-like approach of LBC[2]. First, we need to find an LBC, C, that has an error correction capability e_n such that i) the maximum number of clusters in the channel, L, is equal to e_n and ii) the length of

---

[2]In channel coding, the convention is to use row vectors. Thus, let $\boldsymbol{x}$ and $\boldsymbol{c}$ be $1 \times k$ and $1 \times n$ binary row vectors that represent an information sequence and its corresponding codeword of an LBC, respectively. Also let $\boldsymbol{r} = \boldsymbol{c} + \boldsymbol{e}$ be a received sequence corrupted by $1 \times n$ error pattern $\boldsymbol{e}$. To decode $\boldsymbol{r}$, we calculate an error syndrome vector $\boldsymbol{s}$, of size $1 \times n - k$, such that $\boldsymbol{s} = \boldsymbol{r} \boldsymbol{H}^T$, where $\boldsymbol{H}$ is the parity check matrix of the used LBC. Then, a most likely error pattern $\hat{\boldsymbol{e}}$ can be uniquely identified by $\boldsymbol{s}$ using a look-up table called the *standard array*. Finally, the decoded codeword is obtained using $\hat{\boldsymbol{c}} = \boldsymbol{r} - \hat{\boldsymbol{e}}$. A decoding error occurs if the number of errors, identified using 1's in $\boldsymbol{e}$, is beyond the error correction capability of the used code, denoted by e_n. Note that in this context, all vectors, matrices and math operations are over GF(2).

its codewords n is equal to the number of antennas n_r (n_r is also the number of resolvable directions). The code C has a parity check matrix $\boldsymbol{H}$ which represents the link between channel decoding and beam detection problems. Binary codes deal with data and error sequences defined over the finite field $GF(2)$, i.e., addition and multiplication operations are defined over $GF(2)$ with binary inputs and outputs, i.e., 1's and 0's. However, mm-wave channel parameters are defined over the complex numbers field $\mathbb{C}$. Therefore, to account for arbitrary path gains, we should be able to extend this concept to $\mathbb{C}$.

Although $\boldsymbol{H}$ is defined over $GF(2)$, we interpret its $'1'$ and $'0'$ entries as real numbers. Then, similar to channel decoding, we seek to obtain a channel syndrome, $\boldsymbol{y_s}$, such that $(\boldsymbol{y_s})^T = (\boldsymbol{q}^a)^T \boldsymbol{H}^T \implies \boldsymbol{y_s} = \boldsymbol{H} \boldsymbol{q}^a$. This matrix multiplication can be realized using channel measurements such that each measurement gives one component in $\boldsymbol{y_s}$. Measurements $\{y_{s_i}\}_{\forall i}$ make up the components of the channel syndrome vector $\boldsymbol{y_s}$. Then, we need to find a mapping function $\xi()$ that takes in the channel syndrome vector $\{\boldsymbol{y_s}\}$ as an input and returns the estimated channel $\widehat{\boldsymbol{q}}^a$. The position of each non-zero component in $\widehat{\boldsymbol{q}}^a$ identifies a path's AoA, and its value identifies its baseband path gain. Finally, for this to work, we need to show that such channel measurements provide one-to-one mapping to the channel. In other words, $\boldsymbol{y_s}$ must be a *sufficient statistic* for estimating the channel. In Section V-C, we will show that our design results in the sufficient statistic we seek to achieve.

Remark (Number of Measurements). The solution we obtain is dependent on channel parameters, namely, the number of antennas and the sparsity level of the channel. That is, at a fixed sparsity level, i.e., fixed number of clusters L, a larger number of antennas necessitates more channel measurements. In other other words, the high resolution realized by large n_r comes at a price of an increased number of measurements. Similarly, at fixed n_r, more channel clusters involve more measurements for correct channel estimation.

B. Measurements Design

Recall that each component in $\boldsymbol{q}^a$ represents a resolvable angular direction at the receiver. Let each resolvable direction be given an identification number ($dir_{rx}\#i$). Also let $beam_{rx}\#i$ denote the beam pattern pointing to $dir_{rx}\#i$, i.e., a signal coming from $dir_{rx}\#i$ can be individually measured using $beam_{rx}\#i$ (similar to beam patterns in Fig. 1).

Now, we seek to obtain $\boldsymbol{y_s} = \boldsymbol{H} \boldsymbol{q}^a$ using careful design of $\boldsymbol{w_i}$'s (note that we drop the noise terms for clarity), i.e.,

$$\boldsymbol{y_s} = \begin{pmatrix} y_{s_0} \\ y_{s_1} \\ \vdots \\ y_{s_{m-1}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{w_0}^H \boldsymbol{q} \\ \boldsymbol{w_1}^H \boldsymbol{q} \\ \vdots \\ \boldsymbol{w_{m-1}}^H \boldsymbol{q} \end{pmatrix} \equiv \boldsymbol{H} \boldsymbol{q}^a. \qquad (13)$$

To achieve this, each rx-combining vector $\boldsymbol{w_i}$ is designed as a multi-armed beam, i.e., composed of several sub-beams similar to the beam pattern in Fig. 2. The sub-beams included in each $\boldsymbol{w_i}$ are identified by the ith row of the matrix $\boldsymbol{H}$. That is,

only if $h_{i,j}$, the intersection of the ith row and j^{th} column, is $= 1$, do we include $beam_{rx}\#j$ as a sub-beam in w_i. (also refer to our discussion in Section II).

The design of rx-combining vectors is a crucial aspect of this work. As an initial step towards obtaining proper rx-combining vectors, we consider designing w_i's using linear summation of all analog beamformers that correspond to $beam_{rx}\#j$'s $\forall j$: $h_{i,j}=1$. Let $\Omega_j = \cos(\phi_j) = \frac{j}{L_r}$ $\forall j \in \{0, \ldots, n_r-1\}$, such that $e_r(\Omega_j)$ is the spatial signature of $beam_{rx}\#j$. Then, w_i can be designed as

$$w_i = \sum_{j=0}^{n_r-1} \mathbb{1}(h_{i,j}=1) e_r(\Omega_j) \qquad (14)$$

C. Sufficient Statistic

We will show in this section that each channel syndrome can only be mapped to a single *measurable channel*. A measurable channel in this context refers to $n_r \times 1$ channels with L non-zero components such that $L \leq e_n$, where e_n is the error correction capability of the underlying code C and $n_r = n$ is its CWs length.

Since each measurement combines signals coming from multiple directions, each element in the channel syndrome vector is a linear combination of a subset of the available paths. In other words, each measurement has the possibility that one or more paths are included in it. This setting is rather challenging. To understand why, consider a channel that has two paths with gains $\alpha_1, \alpha_2 \in \mathbb{C}$. Suppose that α_1 and α_2 are of equal magnitudes but are out-of-phase (i.e., phase shift = 180°). Hence, if signals coming from both paths are combined in a single measurement, the resultant value is 0 which is similar to the result we get if no paths exist in the measured directions. Also each channel measurement can be a result of endless possibilities for the combined path gain values. So, a natural question to ask is: does this ambiguity cause measurement errors? The direct answer to this question is: **No**. In the sequel we will show that the resulting channel syndrome, i.e., the combination of all channel measurements, is sufficient to correctly estimate the channel.

First, recall our discussion in Footnote 2. Then, consider all *single-bit error patterns* $e^{(i)}$ of a code C, with maximum number of correctable errors $= e_n$, such that

$$e_k^{(i)} = \begin{cases} 1, & k = i \\ 0, & k \neq i \end{cases}$$

where $e_k^{(i)}$ is the kth component of $e^{(i)}$. Also let $s^{(i)}$ be the corresponding error syndrome of $e^{(i)}$. Recall that $s^{(i)} = e^{(i)} H^T$. Hence, we can see that $s^{(i)}$ is exactly the ith row of H^T, i.e., i^{th} column of H. Now, write any correctable error pattern e as a linear combination of all single-bit error patterns over the finite field $GF(2)$ such that

$$e = \omega_1 e^{(1)} + \omega_2 e^{(2)} + \cdots + \omega_n e^{(n)} \qquad (15)$$

and its corresponding error syndrome is

$$s = \omega_1 s^{(1)} + \omega_2 s^{(2)} + \cdots + \omega_n s^{(n)} \qquad (16)$$

where $\omega_i \in \{0, 1\}$ and $|\omega_i : \omega_i = 1| \leq e_n$.

Lemma 1 . *For an error pattern e_t with number of bit errors identical to e_n, its syndrome s_t is a linear combination of e_n linearly independent vectors $s^{(i)}$.*

Proof. We are going to prove this lemma by contradiction. First, assume that s_t is a linear combination of e_n linearly **dependent** vectors $s^{(i)}$ over $GF(2)$. Therefore, there exists another error syndrome s_t^* composed of only linear combination of independent vectors $s^{(i)}$ such that $s_t = s_t^*$. Therefore, there exists another error patter e_t^* with number of errors strictly less than e_n such that its syndrome $s_t^* = s_t$. Since e_t^* has a number of errors less than e_n, then it is a correctable error pattern, and since all error syndromes of correctable error patterns are different, then s_t^* should be $\neq s_t$. Hence, we arrive at a contradiction. □

It is also easy to see that if e_{t_1} and e_{t_2} are two different correctable error patterns, then their error syndromes s_{t_1} and s_{t_2} are composed of a linear combination of different sets of single-bit error syndromes $s^{(i)}$.

Lemma 2 . *Any $n-$dimensional linearly independent vectors over $GF(2)$, are also linearly independent over $\mathbb{C}^n$.*

Proof. Let $v_1, \ldots, v_m$ be a set of $n-$dimensional vectors defined over $GF(2)$. The vectors v_i can be made the columns of an $n \times m$ matrix Ψ. Since all v_i's are linearly independent over $GF(2)$, then Ψ is an invertible matrix. Therefore, there exists a non-zero (modulo 2) $m \times m$ minor of Ψ. Now, suppose the entries in Ψ are interpreted as real numbers. Therefore, Ψ, now taken over $\mathbb{R}$, has an $m \times m$ sub-matrix whose determinant is non-zero, which proves that it is invertible. Therefore, the vectors v_i's, i.e., columns of Ψ, are linearly independent over $\mathbb{R}$ which, using the same argument, can also be shown to be linearly independent over $\mathbb{C}$. □

Suppose that entries of H and $e^{(i)}$ are interpreted as real numbers, then we can write the channel q^a as

$$(q^a)^T = \alpha_1 e^{(1)} + \alpha_2 e^{(2)} + \cdots + \alpha_n e^{(n)} \qquad (17)$$

where $\alpha_i \in \mathbb{C}$ and $\sum_{i=1}^n \mathbb{1}(\alpha_i \neq 0) \leq e_n$. Therefore, each channel syndrome $(y_s)^T = (q^a)^T H^T \implies y_s = H q^a$ is a linear combination of independent vectors in $\mathbb{C}^{n-k}$ (columns of H). Therefore, all possible measurable channels yield unique channel syndromes which implies that they are sufficient for the channel estimation problem.

D. Mapping Function $\xi()$

Now that we have shown that each measurable channel can be mapped to a unique channel syndrome, we need to find this mapping function, i.e., $\xi : y_s \rightarrow \hat{q}^a$, where $\hat{q}^a$ denotes the estimated channel. Again, we resolve to a technique used in hard decision decoding where a look-up table is constructed that maps every error syndrome to a corresponding error pattern. Likewise, we construct a look-up table that indicates which channel corresponds to an obtained channel syndrome.

Since we employ ADCs with finite resolution, only a finite number of realizable syndromes, y_s, exist (and a finite number of corresponding channels). Therefore, a look-up table method is feasible. We construct the table by, first, generating all possible sparse angular channels. Then, we find the corresponding channel syndromes using $y_s = Hq^a$, where $q^a \in \mathcal{Q}^a$ such that $\mathcal{Q}^a$ is the set of all measurable channels. Let the actual, arbitrary, noise-corrupted, received channel syndrome be given by y_s^r. Therefore, y_s^r might not match exactly one of the channel syndrome vectors in the look-up table. Hence, we instead search for the y_s table entry that has the closest *distance* δ to y_s^r, and pick its corresponding channel as the estimated channel $\hat{q}^a$. We define the distance between the two complex $m-$dimensional vectors y_s, y_s^r, as follows:

$$\delta(y_s, y_s^r) = \sqrt{\sum_{i=0}^{m-1} |y_{s_i} - y_{s_i}^r|^2}. \qquad (18)$$

By obtaining $\hat{q}^a$, we not only identify the AoA at the Rx, but we also obtain the magnitude and phase information associated with every strong path to the TX.

E. Multiple Transmit and Receive Antennas

So far, we have considered channels with single transmit antennas and shown how to perform beam discovery at RX. To extend our approach to a general setting, we consider channels with n_t antennas at TX, and n_r antennas at RX. Thus, instead of the TX just sending signals omnidirectionally, now it can perform highly directional transmission. Recall that the RX is able to perform channel measurements using multi-armed beams. Similarly, the TX can send signals using multi-armed beams to simultaneously focus on multiple directions using precoding vectors f_j.

The design of precoding vectors can also be obtained using an LBC approach. Similar to the method of designing rx-combining vectors w_i, we look for an LBC, C_2, that has CWs of length $n_2 = n_t$ and can correct for $e_n = L$ errors. Let the parity check matrix of C_2 be H_2, using which, we will design the precoding vectors f_j. Let $beam_{tx}\#i$ denote the ith TX beam which points to TX direction dir_{tx} #i. Then, just as before, we envisage H_2 as an array whose columns are associated with resolvable TX directions such that: i) its jth column corresponds to $dir_{tx}\#j$, and ii) its ith row corresponds to the ith measurement. We note that no actual measurements are performed at TX; we use the word *measurement* to refer to precoding, consistent with the case of RX. That is, the ith TX measurement is actually the ith precoder f_i. Thereby, we design the ith precoder as a multi-armed TX beam such that, only if $h_{i,j}$, the intersection of the ith row and j^{th} columns of H_2, is $= 1$, do we include sub-beam $beam_{tx}\#j$ in f_i. Each TX measurement provides a component in a TX channel syndrome vector y_s^{TX}. The total number of TX measurements (i.e., precoding vectors), denoted by m_2, is equal to the number of parity check bits of the code C_2. That is, $m_2 = n_2 - k_2$, where k_2 is the length of C_2's information sequences. To obtain AoDs of strong paths at TX, we define the function $\xi_2()$ as the

---

Algorithm 1: Beam discovery of multiple TX/RX antennas.

input : $\{w_i\}_{\forall i \in \{1,\dots,m_1\}}$, $\{f_j\}_{\forall j \in \{1,\dots,m_2\}}$,
$\qquad\quad \xi_1() : y_s \to \hat{q}^a$, $\xi_2() : y_s^{TX} \to \hat{q}^{a^{TX}}$
output: $\{y_{s_i}\}_{\forall i \in \{1,\dots,m_2\}}$

1 **begin**
2 $\quad$ $j = 0$;
3 $\quad$ **while** $j < m_2$ **do**
4 $\qquad$ $i = 0$;
5 $\qquad$ **while** $i < m_1$ **do**
6 $\qquad\quad$ $y_{s_{i,j}} = w_i^H Q f_j s + w_i^H n$; $\quad$ // channel measurement
7 $\qquad\quad$ $i \leftarrow i + 1$
8 $\qquad$ **end**
9 $\qquad$ $y_{s_j} \leftarrow \{y_{s_{i,j}}\}_{\forall i \in \{1,\dots,m_1\}}$; $\quad$ // construct channel syndrome y_{s_j}
10 $\qquad$ /* find corresponding channel
$\qquad\quad q^{a(j)} = [q_1^{a}(j), q_2^{a}(j), \dots, q_{n_r}^{a}(j)]^T \qquad$ */
11 $\qquad$ $q^{a(j)} \leftarrow \xi_1(y_{s_j})$;
12 $\qquad$ **for** $p \leftarrow 1$ **to** n_r **do**
13 $\qquad\quad$ /* construct TX channel
$\qquad\qquad$ syndromes $y_s^{TX(p)}$, where
$\qquad\qquad y_s^{TX(p)} = $
$\qquad\qquad [y_{s_1}^{TX(p)}, y_{s_2}^{TX(p)}, \dots, y_{s_{m_2}}^{TX(p)}]^T \qquad$ */
14 $\qquad\quad$ $y_{s_j}^{TX(p)} \leftarrow q_p^{a(j)}$
15 $\qquad$ **end**
16 $\qquad$ $j \leftarrow j + 1$;
17 $\quad$ **end**
18 $\quad$ **for** $p \leftarrow 1$ **to** n_r **do**
19 $\qquad$ $q^{aTX(p)} \leftarrow \xi_2(y_s^{TX(p)})$
20 $\quad$ **end**
21 $\quad$ $\hat{Q}^a = \begin{pmatrix} q^{aTX(1)} & q^{aTX(2)} & \dots & q^{aTX(n_r)} \end{pmatrix}^T$
22 **end**

---

mapping function between all possible TX channel syndromes and their corresponding angular channels denoted by q^{aTX}. Note that, for every $dir_{rx}\#i$, there exists a corresponding $q^{aTX(i)}$ which represents the ith row of Q^a. Also, since the maximum number of clusters is L, then, the number of non-zero vectors $q^{aTX(i)}$ is $\leq L$.

To see the whole picture, assume that a code C_1, with CWs of length $n_1 = n_r$, is an LBC code associated with beam discovery at RX side. Let the number of RX measurements, i.e., the number of rx-combining vectors, be m_1 such that $m_1 = n_1 - k_1$, where k_1 is the length of information sequences of C_1. Also let $\xi_1()$ be the mapping function between RX channel syndromes and its corresponding angular channel. Under this setting, the beam discovery problem is performed as follows: i) The TX starts starts sending its training sequence using the precoder $f_j, \forall j \in \{0, \dots, m_2 - 1\}$. ii) The RX performs m_1 channel measurements while f_i is being used at TX and obtains a channel syndrome y_{s_j}. iii) Based on y_{s_j}, the RX obtains a corresponding channel, $q^{a(j)}$ with path components $\{q_p^{a(j)}\}_{\forall p \in \{1,\dots,n_r\}}$. Notice that $q^{a(j)}$'s do not necessarily represent individual path gains, but rather, combinations of paths accumulating at a single $dir_{rx}\#$. Therefore, there exists a resemblance to channel syndromes which we exploit. iv) We construct a set of n_r

TX channel syndromes, $\boldsymbol{y_s^{TX(p)}}$ where their j^{th} component $y_{s_j}^{TX(p)}=q_p^{a(j)}$, i.e., $[\boldsymbol{y_s^{TX(1)}}, \boldsymbol{y_s^{TX(2)}}, \ldots, \boldsymbol{y_s^{TX(n_r)}}] = [\boldsymbol{q^{a(1)}}, \boldsymbol{q^{a(2)}}, \ldots, \boldsymbol{q^{a(m_2)}}]^T$. v) Finally, we find the angular TX channel for every dir_{rx} #p, i.e., p^{th} row of $\boldsymbol{Q^a}$, using the mapping function $\boldsymbol{q^{aTX(p)}}=\xi_2(\boldsymbol{y_s^{TX(p)}})$. Notice that, since no more than $L \ll n_r$ clusters exist, and since $\boldsymbol{0}$ channels correspond to $\boldsymbol{0}$ channel syndromes, we only need to apply $\xi_2()$ a maximum of L times -unless measurement error occurs. This whole process is highlighted in Algorithm 1.

Remark. The estimated channel $\widehat{\boldsymbol{Q}}^{\boldsymbol{a}}$ may contain more than L non-zero components. The reason is that the receiver obtains a channel $\boldsymbol{q^{a(j)}}$ for every precoder $\boldsymbol{f_j}$ which may contain erroneous component estimates. Incorrect estimates occur as a result of measurement errors which happen due to i) channel noise, ii) quantization error. Now every $\boldsymbol{q^{a(j)}}$ may contain a maximum of L non-zero components, however, some of which may be due to measurement errors. Afterwards, potentially noise-corrupted $\{\boldsymbol{q^{a(j)}}\}_{\forall j \in \{1, \ldots, n_r\}}$ are used to obtain TX channel syndromes as shown in Algorithm 1. That is, for every dir_{rx}#i we obtain a TX channel syndrome to identify the corresponding dir_{tx}#j's that have strong components. Thus, we may obtain a maximum of L non-zero components per dir_{rx}#i.

VI. PERFORMANCE EVALUATION

A. Performance Metrics

Our performance metrics focus on three basic criteria; accuracy of beam detection, number of measurements, and accuracy of path gain value estimates. To that end, we use the following performance parameters:

i) **Number of measurements.** An important performance objective is to achieve a similar performance as the exhaustive search, yet with a much smaller number of measurements.

ii) **Perfect beam detection probability:** probability that **only** correct AoAs/AoDs of all available detectable paths are identified (i.e, no incorrect paths are detected along with true ones).

iii) **All-beam detection probability:** probability of correctly identifying all true detectable paths (with possibility of other incorrect paths).

iv) **Partial beam detection probability:** probability that at least one correct path is detected (with possibly other incorrect paths).

v) **Number of incorrect beams:** the number of incorrect paths identified between TX and RX.

vi) **Normalized mean squared error (MSE):** $\frac{\|\boldsymbol{Q^a} - \widehat{\boldsymbol{Q}}^{\boldsymbol{a}}\|_F^2}{\|\boldsymbol{Q^a}\|_F^2}$.
We assess these performance parameters against different *SNR* levels which capture the effect of noise power on performance.

B. Simulation Parameters

Unless otherwise specified, we use the following simulation parameters. We consider a mm-wave channel with L paths between TX and RX[3]. Path gains α_l^b are generated such that

---

[3]Note that we always generate L channel paths, however, according to the preset *SNR*, some path gains are too weak to be detected at RX (recall that in Eq. 12, *SNR* is defined per path). Thus, only paths with enough signal power to cross the threshold of the 1st level (±1) of the utilized ADCs are considered as the ground truth to which we evaluate our performance parameters.

SNR_l lies in a range between SNR_{min} and $(SNR_{min}+20)$ dB for paths that contain strong reflectors. The noise power is $N_0=-95$dBm. The signal power P is adjusted such that for the weakest reflector, an SNR_{min} value is achieved. For example, at $SNR =10$dB, the required transmitted power for a 15×15 channel is $P \approx 43$dBm and for an 8×8 channel is ≈ 46dBm. Further, we use ADCs with 9 quantization levels, i.e., $\{-4, \ldots, 4\}$, where the maximum output is adjusted to the maximum expected received signal power according to the preset *SNR*. Note that the maximum received power depends on the available number of paths. Also negative quantization levels, e.g. -4, represent signals with the same power level as 4 but with phase shift $= 180°$ where the phase reference can be set arbitrarily at RX. For every simulation scenario, we average across 10^5 runs.

C. SinglePath Channels

Consider a 15×15 mm-wave channel with $L=1$ path between TX and RX. Hence, the parity check matrix of $(15, 11, 3)$ Hamming code can be used for the design of both precoders, $\boldsymbol{f_j}$, and rx-combiners, $\boldsymbol{w_i}$, i.e., $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$, respectively. Hence, we need a number of TX measurements m_1, which is identical to the number of RX measurements $m_2=15-11=4$. Hence, the total number of measurements is $m=16$. On the other hand, the exhaustive search method needs 255 measurements to check every possible TX and RX beam combination. Thus, our approach results in $\sim 92.8\%$ reduction in the number of measurements needed for beam discovery. We investigate the probability of perfect, all and partial beam detection at different levels of *SNR*. Fig. 4a depicts these probabilities. Note that since there only exists 1 path, both all and partial beam discoveries are the same.

Further, we investigate the performance in terms of the probability of obtaining a specific number of incorrect paths/beams. We notice that the number of incorrect beams decreases with increasing *SNR*. For instance, at $SNR =10$dB, the probability of obtaining no incorrect paths is ~ 0.92. Hence, the probability of obtaining incorrect paths is ~ 0.08 which is distributed in a decreasing fashion with the number of incorrect paths. Moreover, the number of incorrect beams does not exceed 4 at any *SNR* level which is still good information to obtain. That is, out of the 255 possible transmit and receive beam pairs, a maximum of 4 incorrect $+1$ true beam pairs are discovered which may need further search to obtain the true path. Figures are omitted due to space limitations. Finally, Fig. 4c depicts the normalized mean squared error of the channel estimate $\widehat{\boldsymbol{Q}}^{\boldsymbol{a}}$. The very high MSE value at -5dB indicates very poor channel estimates which indicates that $\widehat{\boldsymbol{Q}}^{\boldsymbol{a}}$ has large components at truly zero components in $\boldsymbol{Q^a}$. Nevertheless, MSE drops fast at higher values of *SNR*. We can see that starting at $SNR =10$dB, the performance becomes highly reliable.

D. MultiPath Channels

Consider an 8×8 channel with $L=2$ paths. Thus, an $(8, 2, 5)$ code is a proper choice for both $\boldsymbol{H_1}$ and $\boldsymbol{H_2}$. A total number, 36, of measurements is needed for beam discovery. As opposed

(a) Beam detection probability
(15×15 channel with L=1).

(b) Beam detection probability
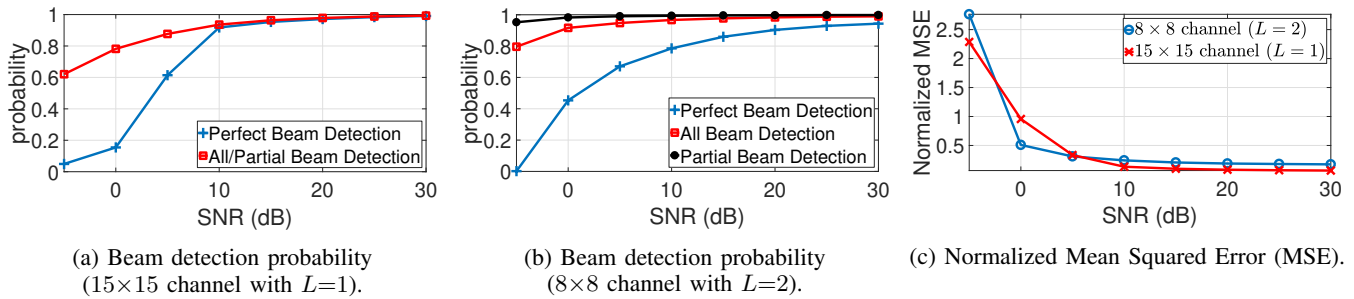(8×8 channel with L=2).

(c) Normalized Mean Squared Error (MSE).

Fig. 4: Performance Evaluation.

to 64 measurements needed for exhaustive search, we achieve ≈43.7% reduction in the number of measurements under this scenario. Fig. 4b shows the probability of perfect, all, and partial beam detection. We notice a large gap between these probability measures at low *SNR*. However, this gap becomes quickly narrower as *SNR* increases. For the number of incorrect paths, we obtain similar performance trends to the single-path channel case. However, the maximum number of incorrect paths experienced is slightly larger. Again, we observe that the reliability of beam discovery increases significantly with increasing *SNR* which is depicted in terms of MSE in Fig. 4c.

VII. CONCLUSION

This work provides a solution for the mm-wave channel estimation problem by exploiting its sparse nature in the angular domain. The proposed solution is a beam discovery technique that is similar to error discovery in channel coding. We show that our proposed technique can significantly reduce the number of measurements required for reliable channel estimation. Our solution takes into account the size of the channel and its sparsity level when determining the number of measurements. As future work, we will investigate more efficient ways of executing the mapping function $\xi()$, possible improvements for designing precoders and rx-combiners, and mitigating the effect of noise-corrupted channel measurements.

REFERENCES

[1] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE journal on selected areas in communications*, vol. 32, no. 6, pp. 1164–1179, 2014.

[2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.

[3] C. R. Anderson and T. S. Rappaport, "In-building wideband partition loss measurements at 2.5 and 60 ghz," *IEEE transactions on wireless communications*, vol. 3, no. 3, pp. 922–928, 2004.

[4] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE access*, vol. 1, pp. 335–349, 2013.

[5] R. W. Heath, N. Gonzlez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave mimo systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, April 2016.

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

[7] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, "Channel estimation and hybrid combining for mmwave: Phase shifters or switches?" in *Information Theory and Applications Workshop (ITA), 2015*. IEEE, 2015, pp. 90–97.

[8] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5g," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 186–194, 2015.

[9] O. Orhan, E. Erkip, and S. Rangan, "Low power analog-to-digital conversion in millimeter wave systems: Impact of resolution and bandwidth on performance," in *2015 Information Theory and Applications Workshop (ITA)*, Feb 2015, pp. 191–198.

[10] J. Mo and R. W. Heath, "High snr capacity of millimeter wave mimo systems with one-bit quantization," in *2014 Information Theory and Applications Workshop (ITA)*, Feb 2014, pp. 1–5.

[11] C. Rusu, R. Mendez-Rial, N. Gonzalez-Prelcic, and R. W. Heath, "Adaptive one-bit compressive sensing with application to low-precision receivers at mmwave," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–6.

[12] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, June 2010.

[13] S. Sun and T. S. Rappaport, "Millimeter wave mimo channel estimation based on adaptive compressed sensing," in *2017 IEEE International Conference on Communications Workshops*, May 2017, pp. 47–53.

[14] J. Mo, P. Schniter, N. G. Prelcic, and R. W. Heath, "Channel estimation in millimeter wave mimo systems with one-bit quantization," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, Nov 2014, pp. 957–961.

[15] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications : Useful tips and tricks," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2017.

[16] A. Alkhateeby, G. Leusz, and R. W. Heath, "Compressed sensing based multi-user millimeter wave systems: How many measurements are needed?" in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015, pp. 2909–2913.

[17] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, Oct 2011.

[18] H. Hassanieh, O. Abari, M. Rodreguez, M. Abdelghany, D. Katabi, and P. Indyk, "Agile Millimeter Wave Networks with Provable Guarantees," *ArXiv e-prints*, Jun. 2017.

[19] M. Kokshoorn, H. Chen, P. Wang, Y. Li, and B. Vucetic, "Millimeter wave mimo channel estimation using overlapped beam patterns and rate adaptation," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 601–616, Feb 2017.

[20] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2563–2579, 2002.

[21] J. H. Van Lint, *Introduction to coding theory*. Springer Science & Business Media, 2012, vol. 86.