

# Multi-beam Transmissions for Blockage Resilience and Reliability in Millimeter-wave Systems

Irmak Aykin, *Student Member, IEEE*, Berk Akgun, *Student Member, IEEE*, and Marwan Krunz, *Fellow, IEEE*

**Abstract**—Directionality in millimeter-wave (mmW) systems make link establishment and maintenance challenging, due to the search-time overhead of beam scanning and the vulnerability of directional links to blockages. In this paper, we propose a communication protocol called SmartLink, which exploits the clustering phenomenon at mmW frequencies to establish a multi-beam link between a base station and a user. By utilizing multiple clusters, SmartLink enables efficient link maintenance and sustained throughput. We develop a logarithmic-time search algorithm called multi-lobe beam search (MLBS), which is used in SmartLink to discover the clusters. MLBS probes several directions simultaneously, using *multi-lobe* beam patterns. The number of simultaneous lobes is selected to minimize the search time of the clusters. We provide detailed analysis of the false alarm and misdetection probabilities for the designed beam patterns. Following cluster discovery, SmartLink divides antennas into sub-arrays to generate the optimal multi-lobe beam pattern that maximizes the average data rate under blockage. Extensive simulations using actual channel traces obtained by utilizing phased-array antennas at 29 GHz are used to verify the efficiency of SmartLink. MLBS decreases the discovery time by up to 88% compared to common existing search schemes, and exploiting multiple clusters improves the average data rate by 10%.

**Index Terms**—Millimeter-wave; initial access; analog beamforming; blockage; multi-lobe beams; beam scanning.

## I. INTRODUCTION

The rapid growth in mobile data is forcing next-generation wireless systems like 5G NR [1] and WiGig [2] to explore RF communications at millimeter-wave (mmW) bands. Unlike the heavily congested and fragmented sub-6 GHz spectrum, mmW bands provide abundant spectrum (about 1.3 GHz in the 28 GHz band, 2.1 GHz in the 39 GHz band, and 14 GHz in the 60 GHz band) [1]. This allows mmW systems to achieve very-high-throughput and ultra-low-latency communications, and to support a diverse set of applications [3]. However, mmW signals experience much higher propagation and penetration losses compared to their sub-6 GHz counterparts (e.g., up to 40 dB through glass and brick walls), making them very vulnerable to blockage [4]–[6]. Fortunately, the smaller wavelengths of mmW signals allow large antenna arrays to be implemented into small form-factor radios. With proper analog and/or digital processing of the signals fed to/received from these antenna elements, transmissions/receptions can be

beamed towards desired directions. The high beamforming gain can then compensate for the severe signal degradation and provide adequate link margin [7].

There are several ways to apply beamforming at mmW frequencies. Analog beamforming can be used with only one RF chain and multiple phase shifters that feed an antenna array and steer the beam. In contrast, digital beamforming relies on several RF chains and multiplies a particular precoding matrix by the modulated baseband signals from these chains. Despite its superior performance, pure digital beamforming is less favored due to its higher complexity and energy consumption. As a good tradeoff, hybrid beamforming, in which the signal processing is divided between the analog and digital domains, has been proposed [8]. In this paper, we consider analog beamforming, whose low power consumption and simple hardware make it a desirable architecture, especially for user equipment (UE) radios.

While beamforming allows for high gains, establishing and maintaining network connectivity based on directional links can be quite challenging [9], [10]. Due to the limited scattering at mmW frequencies, the channel matrix between the base station (BS) and the UE is typically sparse [7], [11]. The transmitted signal reaches the receiver along a few (typically less than 5) distinct angular *clusters*. Identifying the directions of these clusters takes a considerable amount of time.

In this paper, we propose an efficient and reliable communication protocol for mmW systems. Our protocol, called SmartLink, utilizes multiple clusters between the BS and the UE to provide an effective mechanism for maintaining communications under random blockage. It uses a unique beam scanning technique called multi-lobe beam search (MLBS). Implementing shortest-depth decision trees, MLBS utilizes beam patterns with multiple lobes to simultaneously discover multiple channel clusters. The decision tree indicates the next beam pattern that should be scanned, based on previous measurement results. As the worst-case beam scanning time depends on the depth of the decision tree, the number of simultaneous lobes to be used by MLBS is optimized to minimize the depth. Through rigorous analysis, we show that MLBS reduces the search time from linear to logarithmic with respect to the total number of beam directions. The creation of multiple *main* lobes naturally reduces the beamforming gain, leading to higher false alarm and misdetection probabilities. Here, misdetection is defined as the inability to discover the UE due to low received power. False alarm refers to the event when a channel cluster is aligned with an undesired side lobe of the transmitter (Tx) and/or the receiver (Rx) antenna array (resulting in sub-optimal communication). We provide detailed

An abridged version of this paper appeared in the Proceedings of the INFOCOM '19 Conference, Apr. 29 - May. 2, 2019, Paris, France.

I. Aykin, B. Akgun, and M. Krunz are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721 USA (email: {aykin,berkagun,krunz}@email.arizona.edu)

M. Krunz is also with School of Electrical and Data Engineering, University of Technology Sydney.

Manuscript received May XX, 2019; revised August XX, 2019.

analyses of both probabilities. Depending on the discovered channel clusters and their relative gains, we virtually divide the Tx and the Rx antenna arrays into several sub-arrays. Each sub-array generates a beam towards one of the inferred channel clusters, so as to maximize the average data rate in a blockage-prone environment.

The main contributions of this paper are as follows:

- We develop a logarithmic-time search scheme called MLBS for the discovery of multiple channel clusters in a directional mmW system. MLBS utilizes decision trees whose nodes indicate the sequence of beam directions that need to be scanned. As the worst-case beam scanning time depends on the depth of the decision tree, the number of simultaneous lobes to be used by MLBS is selected in a way to minimize the depth. Upper and lower bounds on the depth of the resulting tree are derived.
- We provide detailed analysis of the misdetection and false alarm probabilities under MLBS.
- We present a technique for splitting the antenna array into sub-arrays to generate an optimal multi-lobe pattern. This splitting depends on the discovered channel clusters, their relative strengths, their blockage probabilities, and the data transmission duration.
- We introduce the SmartLink protocol, which defines the required message exchange between the BS and the UE to establish the multi-directional link.
- Through experimental channel measurements at 29 GHz frequency using  $4 \times 4$  uniform planar arrays (UPAs), we verify the efficiency of SmartLink in terms of the reduction in search time and the increase in data rate. With a proper number of main lobes, SmartLink reduces the search latency by up to 88% compared to 5G NR and 65% compared to 802.11ad-like beam scanning approaches. Our results also show that utilizing multiple clusters provides an efficient mechanism against blockages, and improves the average data rate by 10%.

## II. RELATED WORK

Several previous works on mmW communications focused on determining the best channel cluster and communicating through a single beam [2], [12]–[15]. Maintaining an active link over a relatively long period of time may not be feasible in this case, due to the dynamic nature of mmW channels and their susceptibility to frequent blockage events [6], [16].

To identify one cluster, three main approaches have been discussed in the literature: Exhaustive search [12], two-stage hierarchical search [2], and context-information-based (CI-based) search [17]. Exhaustive search is a sequential brute-force approach and it is considered for 5G NR [17]. It comes at a significant cost in terms of discovery time, as each pair of transmit/receive directions must be probed sequentially. On the other hand, the two-stage beam search used in the 802.11ad standard for WiGig devices employs a hierarchical multi-resolution beamforming codebook to reduce the overhead of exhaustive search. In the first stage, the access point (AP) sequentially transmits synchronization signals over relatively wide (quasi-omnidirectional) sectors and tries to determine the

best coarse direction. In the second stage (beam refinement), the AP refines its search within the selected coarse sector by switching to narrow beams [2]. Although this approach reduces the initial access (IA) delay, the search time still scales linearly with the total number of narrow beams. Note that when multiple clusters are to be found, the performance of hierarchical beam search approaches that of the exhaustive search. This is due to the fact that multiple quasi-omni sectors that received relatively strong signals in the first stage needs to be scanned with narrow beams in the second stage. Finally, the CI-based search relies on location information, where the UE simply selects the closest BS based on location and steers its beam towards that BS [17].

In addition to the above works, others in the literature considered transmissions through multiple antenna lobes. The authors in [18] used a unique approach based on hashing functions to identify the best beam. Although the multi-beam hash functions reduce the search time, they were not used to identify multiple channel clusters. [19] improves on [18] and reduces beam alignment delay by orders of magnitude. Similar to [19], the authors in [20] and [21] also aimed at scanning the environment with multi-lobe beam patterns and establishing communications through multiple lobes. However, they did not optimize their search scheme and simply rotated a pre-designed multi-lobe pattern to decide the best Rx pattern. As a result, their scheme does not guarantee discovering multiple clusters and combating blockage. On a different perspective, [22] introduced a heterogeneous multi-beam cloud radio access network architecture, with the goal of providing seamless mobility and coverage, rather than exploiting multiple channel clusters. However, none of the above works considered the tradeoff between simultaneously searching multiple directions and the reduced beamforming gain for a system that utilizes analog beamforming. Such tradeoff was studied in [23], but the evaluation was done through antenna patterns with a fixed number of lobes. As a result, the search time is not minimized.

Our proposed MLBS approach computes the optimum number of simultaneous beams to be used, given the number of channel clusters and beam directions. As a result, we guarantee the minimum search time. We also provide a detailed analysis of the misdetection and false alarm probabilities under MLBS. After the clusters are discovered, they are used to simultaneously receive copies of the same signal arriving from multiple directions. The phases of these signals are adjusted for coherent combination. Because the probability that all channel clusters are blocked at the same time is small, our proposed scheme provides an effective mechanism to combat blockage and maintain communications at high link rates.

Our contributions in this paper show the potential benefits of discovering multiple clusters/links between two communicating nodes in mmW systems and encourage further research on developing multi-beam transmission methods. Even though we focus on dealing with dynamic blockage at the mmW bands through the usage of multiple clusters and analog beamforming (i.e., higher diversity gain), our method can be applied in a hybrid beamforming architecture for multiplexing gain and can also be integrated into coordinated multipoint (CoMP) techniques for the 5G networks.

### III. SYSTEM MODEL

In this paper, we consider the IA process at a BS that wishes to discover the UEs in its range. Electronically steerable phased-array antennas are used by the BS and the UEs. In addition, we assume analog beamforming, as it is currently the most energy-efficient beamforming solution available. In particular, a beamforming architecture with a much lower number of transceivers than the total antenna number is more practical and cost effective to deploy, especially in the UE side [24]. For that reason, UEs operating on mmW bands are typically envisioned to exploit analog beamforming. The use of digital and hybrid beamforming architectures can reduce the IA time to meet the low-latency 5G requirements, but at the expense of much higher energy consumption. As described in [9], low-resolution digital architectures can also be a viable solution in digital/hybrid architectures. Still, even low resolution digital beamforming is suitable for the BS side, due to stringent power requirements at the UE [25].

Following the IA process, the BS and any discovered UE can communicate over multiple beams. Without loss of generality, we let the BS be the Tx and one arbitrary UE be the Rx, i.e., the downlink.

#### A. Preliminaries - Antenna Arrays

In this section, we demonstrate how multi-lobe patterns are created. To do so, we first explain the calculation of the array factor<sup>1</sup> (AF) at a UE antenna array. Extension to the BS is straightforward. Consider a UPA that consists of  $N$  antenna elements with a horizontal inter-element distance  $d_x$  and a vertical inter-element distance  $d_y$ . Suppose that the incident wave of the received signal arrives at zenith angle  $\theta$  and azimuthal angle  $\alpha$ , and the antennas are placed on an  $N_x \times N_y$  2D grid (i.e.,  $N = N_x N_y$ ). The phase of the received signal at element  $(n_x, n_y)$ ,  $n_x \in \{1, \dots, N_x\}$  and  $n_y \in \{1, \dots, N_y\}$ , leads the phase at element  $(n_x - 1, n_y - 1)$  by  $2\pi(d_x \cos \alpha \sin \theta + d_y \sin \alpha \sin \theta)/\lambda$ , where  $\lambda$  is the wavelength of the signal. Let  $\kappa \triangleq d_x n_x \cos \alpha \sin \theta + d_y n_y \sin \alpha \sin \theta$ . The received signal at antenna  $(n_x, n_y)$  can then be written as:

$$s_{n_x, n_y}(\theta, \alpha) = R I_{n_x, n_y} e^{j \frac{2\pi}{\lambda} \kappa} \quad (1)$$

where  $I_{n_x, n_y}$  is the amplitude excitation of the element  $(n_x, n_y)$  and  $R$  is the individual gain of each antenna element. Since  $I_{n_x, n_y}$  does not affect the analog beamforming weights, for simplicity, we let  $I_{n_x, n_y} = 1$ . Let  $w_{n_x, n_y}$  be a complex phase shifter weight associated with  $s_{n_x, n_y}$ . The total received signal is given by:

$$s(\theta, \alpha) = R \sum_{n_x=1}^{N_x} \sum_{n_y=1}^{N_y} w_{n_x, n_y} e^{j \frac{2\pi}{\lambda} \kappa} = R F_{\text{UPA}} \quad (2)$$

where  $F_{\text{UPA}}$  is the AF of the UPA. The signal power can then be maximized by maximizing  $|R F_{\text{UPA}}|$ . Assuming the same signal magnitude at each antenna,  $|F_{\text{UPA}}|$  is maximum when  $w_{n_x, n_y}$  is selected in a way to ensure that the received signals

<sup>1</sup>The AF is the factor by which the element factor of an individual antenna is multiplied to get the total firing pattern of the entire array.

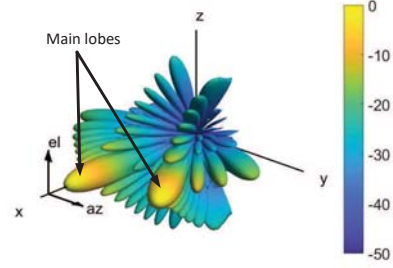


Fig. 1. Normalized directivity pattern of a  $16 \times 16$  UPA divided into two sub-arrays, beaming along  $(\theta_1, \alpha_1) = (0^\circ, 0^\circ)$  and  $(\theta_2, \alpha_2) = (30^\circ, 0^\circ)$  (antennas placed on the Y-Z plane).

are in phase, i.e., by setting  $w_{n_x, n_y} = e^{-j \frac{2\pi}{\lambda} \kappa}$ . This way, the UE beam can be steered along the direction  $(\theta, \alpha)$ .

To create a pattern with multiple lobes, the antenna array is divided into several sub-arrays. The elements in each sub-array are then assigned weights for different steering angles. An example of a two-lobe beam pattern with uniform sub-array division is illustrated in Fig. 1, where the steering angles are  $(\theta_1, \alpha_1) = (0^\circ, 0^\circ)$  and  $(\theta_2, \alpha_2) = (30^\circ, 0^\circ)$ .

#### B. Channel and Beamforming

To develop a multi-lobe beamforming design, we must express the channel between a BS and a UE, when both are equipped with UPAs. We use the notation  $x \times y$  to denote a matrix of  $x$  rows and  $y$  columns. Let the total number of antennas at the BS and UE be  $N_{\text{BS}} = N_{\text{BS},x} \times N_{\text{BS},y}$  and  $N_{\text{UE}} = N_{\text{UE},x} \times N_{\text{UE},y}$ , respectively. Denote the  $N_{\text{UE}} \times N_{\text{BS}}$  channel matrix between the BS and the UE by  $\mathbf{H}$ .

To express the received signal, Tx and Rx beamforming should be applied to channel  $\mathbf{H}$ . In practice, the beamforming vectors are computed offline for a set of directions and stored in the codebooks at the Tx and the Rx. During directional communications, if the BS uses the transmit beamforming vector  $\mathbf{f}_i \in \mathbb{C}^{N_{\text{BS}} \times 1}$  and the UE uses the receive beamforming vector  $\mathbf{q}_j \in \mathbb{C}^{N_{\text{UE}} \times 1}$  ( $i$  and  $j$  are the indices of the beamforming vectors in their respective codebooks), then the received signal  $y_{ij}$  can be written as

$$y_{ij} = \mathbf{q}_j^* \mathbf{H} \mathbf{f}_i x + \mathbf{q}_j^* \mathbf{n} \quad (3)$$

where  $(\cdot)^*$  denotes conjugate transpose operation,  $x$  is the transmitted signal and the entries of  $\mathbf{n} \in \mathbb{C}^{N_{\text{UE}} \times 1}$  are complex circularly symmetric white Gaussian noise. Here,  $\mathbf{f}_i$  and  $\mathbf{q}_j$  consist of the complex antenna weights  $w_{n_x, n_y}$  (computed as discussed in Section III-A) to construct Tx beam  $i$  and Rx beam  $j$ , respectively. Note that patterns generated by  $\mathbf{f}_i$  and  $\mathbf{q}_j$  can have single or multiple main lobes. Let us define the codebook at the Rx side as  $\mathcal{Q}$ , where  $\mathcal{Q}$  consists of  $L$  single-lobe beams,  $\binom{L}{2}$  two-lobe beams, and so on. Here,  $L$  is the maximum number of single-lobe beam patterns that can be generated at the Rx, which is a function of the resolution of the phase shifters.

Our aim is to find the optimum Rx beamformer that maximizes the average data rate within a given time period, for a given Tx beamformer. Because mmW channel clusters experience significant blockages, utilizing several channel

clusters simultaneously prevents the Tx and the Rx from losing communication, and hence improves the average data rate.

#### IV. CLUSTER DISCOVERY AND MULTI-LOBE COMMUNICATIONS

In this section, we explain the main aspects of the proposed multi-directional communications scheme. Specifically, the proposed scheme consists of three main parts: identifying cluster directions, aligning the phases of the signals coming from different paths, and optimally allocating antennas/power to different clusters while taking blockages into account. We then explain the protocol design that specifies the steps taken at the BS and the UE to execute these parts.

##### A. MLBS Algorithm

For a given Tx direction, the MLBS algorithm aims at identifying the directions of the dominant channel clusters at the Rx. Recall that  $L$  is the total number of available narrow beam directions. We assume that the number of clusters  $P$  is unknown. To infer the  $P$  cluster directions, we propose to take multiple measurements using beam patterns with  $B$  lobes,  $B \geq 1$ , generated using  $B$  sub-arrays. This problem is similar to the game Mastermind, with  $L$  colors (directions) and  $B$  pegs (simultaneous beams). The outcome of taking a measurement with a pattern at each round will either be a 1 if the power of the received signal is above a threshold  $T$  or a 0 if the power of the received signal is below  $T$ . In general, the solution to Mastermind can be found in a relatively small number of rounds using the *minimax* method, i.e., choosing a test pattern that minimizes the maximum number of remaining possibilities at each round. However, finding the optimal solution for the overall problem is NP-hard [26]. As a result, online computation of the optimal sequence of test patterns is not feasible. Thankfully, the optimal strategy can be computed offline, considering all possible responses, and stored in a binary decision tree. Then, at each round, the receiver measures the channel with a given antenna pattern. Based on the received power and the decision tree, it can select the next test pattern optimally.

To build the decision tree, we first initialize an  $\binom{L}{P} \times \binom{L}{B}$  matrix  $\mathbf{D}$ , whose rows correspond to all possible cluster combinations in the environment and its columns correspond to all possible  $B$ -lobe beam patterns. For example, consider the row labeled as  $(m, n)$  and the column labeled as  $(i, j)$ . Their intersection represents the scenario where the Rx uses a beam pattern with two main lobes pointing along directions  $i$  and  $j$  to measure the channel that has two clusters along the directions  $m$  and  $n$ . If the Rx pattern captures at least one cluster, the corresponding matrix entry is set to 1; otherwise it is set to 0. The selection of the threshold  $T$  for this binary decision will be discussed in the next section.

Recall that  $P$  is not known before any channel measurements are taken. However, statistical channel models can be utilized to determine an appropriate value for  $P$  at the frequency of interest. MLBS can then be executed with that value. In case that the channel exhibits fewer clusters than initially assumed, the algorithm will still return  $P$  output

---

#### Algorithm 1 MLBS Algorithm

---

```

1: procedure MAIN( $L, P, B$ )
2:   rows  $\leftarrow \binom{L}{P}$  direction combinations
3:   global columns  $\leftarrow \binom{L}{B}$  direction combinations
4:   global  $\mathbf{D} \leftarrow \text{zeros}(\binom{L}{P}, \binom{L}{B})$ 
5:   for  $i$  in rows do
6:     for  $j$  in columns do
7:       if  $i$  and  $j$  have a common direction then
8:          $\mathbf{D}[i][j] \leftarrow 1$ 
9:   root  $\leftarrow (1, 2, \dots, B)$ 
10:  MLBS(rows, root)
11: procedure MLBS(REMAININGROWS, PARENT)
12:  if size(remainingRows) > 2 then
13:    remainingLeft  $\leftarrow$  remainingRows - {rows whose
14:    intersection with parent is 0 in  $\mathbf{D}$ }
15:    remainingRight  $\leftarrow$  remainingRows - {rows whose
16:    intersection with parent is 1 in  $\mathbf{D}$ }
17:    ll  $\leftarrow$  column that divides remainingLeft as evenly
18:    as possible, in terms of 1s and 0s
19:    rr  $\leftarrow$  column that divides remainingRight as evenly
20:    as possible, in terms of 1s and 0s
21:    parent.leftchild  $\leftarrow$  ll
22:    parent.rightchild  $\leftarrow$  rr
23:    MLBS(remainingLeft, ll)
24:    MLBS(remainingRight, rr)
25:  else
26:    remainingLeft  $\leftarrow$  remainingRows - {rows whose
27:    intersection with parent is 0 in  $\mathbf{D}$ }
28:    remainingRight  $\leftarrow$  remainingRows - {rows whose
29:    intersection with parent is 1 in  $\mathbf{D}$ }
30:    parent.leftchild  $\leftarrow$  remainingLeft
31:    parent.rightchild  $\leftarrow$  remainingRight

```

---

directions, but no antenna/power will be allocated to weaker clusters. Antenna allocation will be explained in detail in Section IV-D.

The MLBS algorithm is summarized in Algorithm 1. The output of the algorithm is the resulting decision tree, which is defined by its root and the left and right children of each node. The nodes of the tree correspond to multi-lobe measurement patterns and the leaves correspond to identified cluster directions. The root of the decision tree does not affect the performance of the MLBS. Specifically, an optimal tree can be constructed with any arbitrary pattern selected as the root. Without loss of generality, we select the root as the pattern  $(1, 2, \dots, B)$ , i.e., the pattern with  $B$  main lobes in the directions  $1, \dots, B$ . Once a measurement result for this initial pattern is obtained, the next pattern is selected in a way to divide the remaining cluster directions as evenly as possible, so as to minimize the maximum remaining possibilities. In other words, the next test pattern should have even (or close to even) number of 1s and 0s in the corresponding column of  $\mathbf{D}$ , for the remaining possible cluster directions. This way, the remaining candidate directions are halved at each round, until the unique cluster directions are identified. As a result, although the complexity of constructing the decision tree from

TABLE I  
DECISION MATRIX FOR  $L = 5$  AND  $P = B = 2$ .

	(1,2)	(1,3)	(1,4)	(1,5)	(2,3)	(2,4)	(2,5)	(3,4)	(3,5)	(4,5)
(1,2)	1	1	1	1	1	1	1	0	0	0
(1,3)	1	1	1	1	1	0	0	1	1	0
(1,4)	1	1	1	1	0	1	0	1	0	1
(1,5)	1	1	1	1	0	0	1	0	1	1
(2,3)	1	1	0	0	1	1	1	1	1	0
(2,4)	1	0	1	0	1	1	1	1	0	1
(2,5)	1	0	0	1	1	1	1	0	1	1
(3,4)	0	1	1	0	1	1	0	1	1	1
(3,5)	0	1	0	1	1	0	1	1	1	1
(4,5)	0	0	1	1	0	1	1	1	1	1

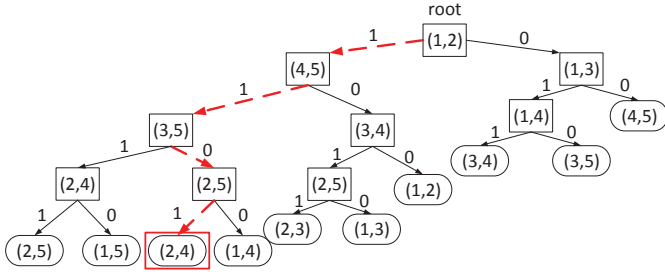


Fig. 2. Decision tree for matrix  $\mathbf{D}$ . Path traversed in the example is shown with dashed lines. Identified cluster pair is shown in the red box.

matrix  $\mathbf{D}$  is exponential, it could be done once and offline. Once the tree is constructed, traversing it and identifying the cluster directions take logarithmic time, as the remaining possibilities are reduced approximately by half after testing an antenna pattern. On the other hand, exhaustive 5G beam search scales linearly with  $L$ , regardless of how many clusters the BS and the UE aim at finding, as they probe each beam direction sequentially. Finally, 802.11ad search also scales linearly with  $L$ . Even though its search time may be relatively low when the aim is to find a single cluster, the search time approaches that of exhaustive search when the algorithm aims at finding multiple clusters.

*Example:* We illustrate the execution of MLBS algorithm with  $L = 5$  and  $P = B = 2$ . For this case, the matrix  $\mathbf{D}$  can be constructed as in Table I. Notice that when the row and column tuples overlap by at least one entry, i.e., one or more clusters can be captured with the given pattern, the corresponding entry in  $\mathbf{D}$  is set to 1.

Using  $\mathbf{D}$ , the decision tree can be constructed as in Fig. 2. Specifically, after taking a measurement with the root (first column), either the first seven or the last three rows will be eliminated, depending on the measurement outcome. Afterwards, if the first seven rows remain, the next pattern can be chosen as (4, 5), i.e., the last row, as it divides the remaining rows as evenly as possible in terms of 1s and 0s. Similarly, if the last three rows remain, the next pattern can be chosen as (1, 3), i.e., second column. The rest of the tree can be constructed in the same manner. Note that the decision tree is not unique. Fig. 2 depicts one of the optimal (minimum depth) decision trees for the given  $L$ ,  $P$ , and  $B$ . In the decision tree, rectangles correspond to nodes (i.e., antenna patterns to be tested), ellipses correspond to leaves (i.e., inferred directions), and left and right branches correspond to strong or weak signal relative to  $T$  (1 or 0, respectively).

Suppose that the strongest clusters that the algorithm aims at discovering are along the directions 2 and 4 (i.e., the row labeled as (2, 4)). The algorithm starts at the root by applying the initial pattern (1, 2). Since pattern (1, 2) will capture cluster 2, the outcome of this measurement will be 1. As a result, the algorithm proceeds with the left branch that originates from the root, eliminating the cluster pairs (3, 4), (3, 5) and (4, 5). The next measurement pattern to be selected according to the decision tree is (4, 5). Since this pattern also captures a cluster (direction 4), the result will again be 1. After this step, only four possible cluster pairs remain: (1, 4), (1, 5), (2, 4) and (2, 5). Next, pattern (3, 5) is chosen according to the tree, and the outcome of this measurement is 0. This leaves only two possible cluster pairs: (1, 4) and (2, 4). The algorithm then proceeds with the final measurement using pattern (2, 5), and using its outcome, cluster directions can be uniquely identified as (2, 4), in just four steps. 5G beam search, on the other hand, requires measuring all five directions to find these clusters. Finally, 802.11ad search will initially take two measurements using quasi-omni beams (say, (1, 2, 3) and (4, 5)), and then will do beam refinement in both of these quasi-omni beams. As a result, it requires seven test trials. Clearly, the reduction in discovery time that MLBS provides is more significant when  $L \gg 5$ , which is typical for a mmW system.

*Complexity Analysis:* In this section, we compute the time complexity of MLBS. We also derive upper and lower bounds on the depth of the decision tree. Let  $\zeta$  be the probability of capturing a cluster using a  $B$ -lobe pattern. This probability can be computed as:

$$\zeta = \frac{P}{L} \left( 1 + \sum_{i=0}^{B-2} \prod_{j=0}^i \frac{L-P-j}{L-1-j} \right) \quad (4)$$

where the first term,  $\frac{P}{L}$ , corresponds to capturing a cluster with the first main lobe, the second term,  $\frac{L-P}{L} \frac{P}{L-1}$ , corresponds to not capturing a cluster with the first lobe but capturing it with the second lobe, and so on. Capturing at least one cluster with a pattern results in a measurement outcome of 1. As explained before, the aim of the algorithm is to find a pattern at each round such that it divides the remaining rows as evenly as possible, in terms of measurement outcomes of 1s and 0s.

In each round, the algorithm finds a pattern that reduces the number of remaining rows by at least a fraction  $f$  of the rows from the previous round. Note that since we want to analyze the worst-case complexity, the number of rows can never be reduced by more than half. Specifically,

$$f = \begin{cases} \zeta & \text{if } \zeta < 0.5 \\ 1 - \zeta & \text{otherwise.} \end{cases} \quad (5)$$

In the best-case scenario, the remaining rows are halved at each round, i.e.,  $f = 0.5$ , which results in the following complexity:

$$\mathcal{O} \left( \log_2 \binom{L}{P} \right) = \mathcal{O} \left( \log_2 L^P \right) = \mathcal{O} \left( P \log_2 L \right) = \mathcal{O} \left( \log_2 L \right) \quad (6)$$

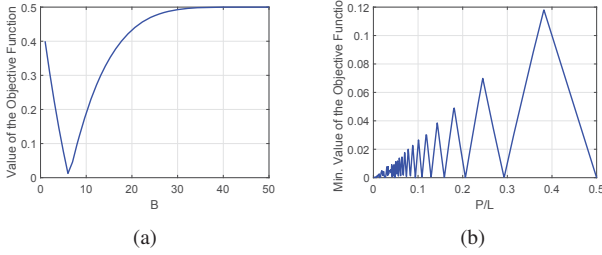


Fig. 3. Numerical analysis of Equation (7). (a) The value of the objective function in (7) vs.  $B$  ( $P = 5$  and  $L = 50$ ), (b) minimum achievable value of the objective function at the optimum  $B$  vs.  $P/L$ .

where  $\binom{L}{P}$  is the number of rows in the first round and the last equality comes from the fact that  $P \ll L$ . For the best performance,  $B$  should be selected properly, ensuring that  $f$  is as close to 0.5 as possible. Thus, the optimum  $B$  can be found by solving

$$\begin{aligned} & \underset{B}{\text{minimize}} \quad |\zeta - 0.5| \\ & \text{s.t.} \quad 1 \leq B < L, \quad P, L, B \in \mathbb{Z}. \end{aligned} \quad (7)$$

For the algorithm to achieve logarithmic complexity, we must ensure that a suitable pattern can be found in each round.

*Theorem 1:* At each round of the algorithm, a pattern can be found such that it reduces the number of remaining rows by at least a fraction  $f$  of the rows from the previous round.

*Proof:* Let us use proof by induction. In step 1 of the algorithm, the optimal  $B$  can be found by solving (7). The resulting  $B$  reduces the number of rows at least by a fraction  $f$  and at most by 0.5. Although (7) does not have a closed-form solution, it can be solved via gradient descent since the objective function is unimodal, i.e., guaranteed to have a single global minimum. Note that if  $P/L \geq 0.5$ , then the objective function is monotonically increasing, as all the terms in the sum are positive. In that case, the optimum  $B = 1$ . On the other hand, if  $P/L < 0.5$ , then the objective function first decreases until the sum exceeds 0.5; afterwards, it monotonically increases, which results in an optimum  $B > 1$ . An example graph of the objective function versus  $B$  is shown in Fig. 3(a) for  $P = 5$  and  $L = 50$ .

Suppose now that at step  $s$ , we select  $B = B_s$  and the theorem holds. Then, we can write

$$f \leq \frac{P}{L} \left( 1 + \sum_{i=0}^{B_s-2} \prod_{j=0}^i \frac{L-P-j}{L-1-j} \right) \leq 0.5. \quad (8)$$

To complete the proof, we must show that at step  $s+1$ , we can find a  $B_{s+1}$  such that

$$f \leq \frac{P}{L} \left( 1 + \sum_{i=0}^{B_{s+1}-2} \prod_{j=0}^i \frac{L-P-j}{L-1-j} \right) \leq 0.5. \quad (9)$$

To satisfy (9), we can simply select  $B_{s+1} = B_s$ , as that would make (8) and (9) equivalent. Thus, by selecting the same  $B$  at each round, the remaining number of rows can be reduced by at least the fraction  $f$  at each round, leaving  $1-f$  fraction of rows. Denote the depth of the decision tree by  $\eta$ . Then,

$\eta \leq \log_{\frac{1}{1-f}} \binom{L}{P}$  when the same  $B$  is selected at each round. Accordingly, the worst-case complexity of the algorithm is  $\mathcal{O} \left( \log_{\frac{1}{1-f}} L \right) = \mathcal{O}(\log_2 L)$ . ■

Note that this is an upper bound on  $\eta$ , and  $f$  approaches 0.5 at each round ( $f = 0.5$  at the last round). Thus,  $\log_2 \binom{L}{P} \leq \eta \leq \log_{\frac{1}{1-f}} \binom{L}{P}$ .

If  $P \ll L$ , one can conclude that  $P/L < 0.5$ . In this case, we can numerically show that the value of the objective function in (7) gets arbitrarily close to 0 with the optimal selection of  $B$ . Fig. 3(b) depicts the minimum achievable value of the objective function vs.  $P/L$ . The figure shows that this value is at most 0.12 (at  $P/L = 0.4$ ). A typical  $P/L$  in a real system is likely to be much smaller, given that the environment typically exhibits  $P < 5$  clusters and the number of directions  $L > 50$ . When  $P/L < 0.1$ , the minimum value of the objective function is less than 0.03, meaning that  $\eta = \log_{1.9} \binom{L}{P} \approx \log_2 \binom{L}{P}$ . Note that the sawtooth shape in Fig. 3(b) is due to  $B$  being an integer.

So far, we discuss the time complexity of MLBS, which is  $\mathcal{O}(\log_2 L)$ . However, this logarithmic-time complexity comes at the expense of higher memory requirement. The Rx needs to store the decision tree (nodes and leaves) to execute MLBS. As the depth of the decision tree scales logarithmically with  $L$ , the number of nodes and leaves of the decision tree also increases with  $L$ . In particular, a binary tree with a depth of  $\mathcal{O}(\log_2 L)$  has at most  $2L - 1$  nodes and  $2L$  leaves, leading to  $\mathcal{O}(L)$  memory complexity.

*Discussion:* The first-null beamwidth (FNBW) of an antenna pattern and the beam scanning resolution are typically not the same. Therefore, the neighboring beams usually overlap (see the lined region in Fig. 4(a)). When a channel cluster is captured by the overlapping region as in Fig. 4(a), it may cause the measurement outcomes of both patterns to be interpreted as 1, as shown in 4(b). We eliminate these duplicate channel clusters during the phase alignment stage, as explained in Section IV-C. Because this will potentially eliminate some duplicates during phase alignment, it is a good idea to aim at identifying more than  $P$  clusters with MLBS, if beam overlaps are significant. Even if MLBS identifies some weak clusters that remain after phase alignment, these clusters will not be utilized during data communication, thanks to our optimal antenna allocation scheme, explained in Section IV-D.

## B. Threshold Selection

MLBS utilizes the received signal power,  $P_{\text{Rx}}$ , to determine whether a cluster is captured or not by a given antenna pattern. In particular, if  $P_{\text{Rx}} \geq T$  ( $P_{\text{Rx}} < T$ ), the binary outcome of the measurement is set to 1 (0). However, due to multipath effects and the side-lobes of Tx and Rx beams, wrong decisions may be made, depending on the value of  $T$ . For example, if  $T$  is too small, an outcome of 1 may be observed even when Tx and Rx beams are not perfectly aligned. On the other hand, if  $T$  is too large, the measurement outcome may be 0 even though it should be 1, due to the incoherent combining of signals arriving from different clusters or insufficient beamforming gain. Therefore,  $T$  must

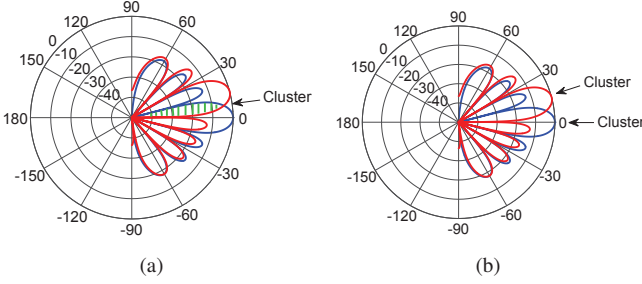


Fig. 4. Visualization of the overlapping beams in an 8-element uniform linear array (FNBW = 28.6°). Blue curve corresponds to the beam pattern when the antenna steers towards 0°, the red curve corresponds to the beam pattern when the antenna steers towards 15°, and the green lines show the overlapping region. Cluster in (a) can be interpreted as in (b), due to the measurement outcome of both beams being 1.

be selected in a way that accounts for false positives (false alarm) and false negatives (misdetction).

To decide whether the current measurement outcome should be set to 1 or 0, we apply a likelihood-ratio test based on  $P_{\text{Rx}}$ . In particular, our hypothesis testing includes the following two events:

- $H_1$ : At least one cluster is captured by the main lobes of a given pattern
- $H_0$ : No cluster is captured by the main lobes of a given pattern

To select an appropriate value for  $T$ , we employ the *Neyman-Pearson* hypothesis test and evaluate the misdetection probability,  $\Pr\{\text{MD}\} = \Pr\{P_{\text{Rx}} < T | H_1\}$ , and the false alarm probability,  $\Pr\{\text{FA}\} = \Pr\{P_{\text{Rx}} \geq T | H_0\}$ . This test can be expressed by the following optimization problem:

$$\begin{aligned} & \underset{T}{\text{minimize}} \quad \Pr\{\text{MD}\} \\ & \text{s.t.} \quad \Pr\{\text{FA}\} \leq \gamma_{\text{FA}} \end{aligned} \quad (10)$$

where  $\gamma_{\text{FA}}$  is a given maximum tolerable false alarm probability. From the definition, it is easy to see that  $\Pr\{\text{MD}\}$  ( $\Pr\{\text{FA}\}$ ) is monotonically increasing (decreasing) in  $T$ . Therefore, to minimize  $\Pr\{\text{MD}\}$ , the smallest  $T$  should be chosen while satisfying the constraint in (10). Hence, the optimal  $T$  value is obtained when (10) holds with equality. Let  $T_{\text{opt}}$  denote the optimal threshold value, i.e.,  $\Pr\{\text{FA}\} = \Pr\{P_{\text{Rx}} \geq T_{\text{opt}} | H_0\} = \gamma_{\text{FA}}$ . To find  $T_{\text{opt}}$ , we need to evaluate  $P_{\text{Rx}}$ .

For analytical tractability, we approximate the actual antenna patterns by a sectored antenna model, as commonly done in the literature [27]–[29]. Specifically, let  $G(\theta, \alpha)$  be the antenna gain of the Rx antenna array, where  $\{\theta, \alpha\} \in [0, 2\pi]$  are the zenith and azimuthal angles, respectively. Let  $\theta_b \in [0, 2\pi]$  and  $\alpha_b \in [0, 2\pi]$ ,  $\forall b \in \{1, \dots, B\}$ , denote the zenith and azimuthal directions of the  $b$ th main lobe, respectively. Then,

$$G(\theta, \alpha) = \begin{cases} G_{\text{max}}, & \text{if } |\theta - \theta_b| \leq \omega_y/2, \\ & |\alpha - \alpha_b| \leq \omega_x/2, \quad b \in \{1, \dots, B\} \\ G_{\text{min}}, & \text{otherwise} \end{cases} \quad (11)$$

where  $\omega_y$  and  $\omega_x$  are the beamwidths of a main lobe in the zenith and azimuthal coordinates, respectively. Furthermore,  $G_{\text{max}}$  and  $G_{\text{min}}$  are the antenna gains of the main and side lobes, respectively.

Let  $r_p e^{j\varphi_p}$  be the received signal from the  $p$ th cluster, where  $r_p$  and  $\varphi_p$  represent its magnitude and phase, respectively. Let  $\mathcal{P}$  denote the set of clusters. Then,  $P_{\text{Rx}}$  is then given by:

$$P_{\text{Rx}} = \left( \sum_{p \in \mathcal{P}} r_p e^{j\varphi_p} \right) \left( \sum_{p \in \mathcal{P}} r_p e^{j\varphi_p} \right)^* + P_N \quad (12)$$

where  $P_N = nn^*$  and  $n \sim \mathcal{CN}(0, \sigma_N^2)$  is the additive white Gaussian noise. In our analysis, we rely on the 3GPP channel model [30]. Based on this model, the phase of a received signal from a particular cluster is a uniformly distributed random variable between 0 to  $2\pi$ , i.e.,  $U(0, 2\pi)$ . Let us evaluate the received signal power from the  $p$ th cluster, i.e.,  $r_p^2$ . Let  $P_p$ ,  $P_{\text{Tx}}$ , and  $P_L$  denote the  $p$ th cluster power, the transmit power of BS (including the Tx antenna array gain), and the path loss of the channel between BS and UE, respectively. Based on the assumed sectored antenna model,  $H_0$  is the event in which all the dominant channel clusters between the BS and the UE are captured by the side lobe. Therefore, given that  $H_0$  occurs,  $r_p^2$  is given by:

$$r_p^2 = G_{\text{min}} P_p P_{\text{Tx}} P_L. \quad (13)$$

In dB scale, the path loss can be written as  $-10 \log_{10}(P_L) = \alpha + 10\beta \log_{10}(d) + \xi$ , where  $d$  is the distance between the BS and the UE,  $\alpha$  and  $\beta$  are frequency- and environment-dependent constants, and  $\xi \sim \mathcal{N}(0, \sigma_{\text{SF}}^2)$  is the shadow fading. IA is performed when a new UE enters the range of a BS, or a UE attempts to change its status from idle to connected. Therefore, instead of relying on specific distributions for UE locations, the maximum possible distance between a BS and a UE,  $D_{\text{max}}$ , can be considered in the path-loss. In that case,  $d$  can be replaced by  $D_{\text{max}}$ . Alternatively, one can assume that UEs are randomly and uniformly located on a ring around the BS. Hence, the CDF of the distance between a BS and a UE is given by  $\Pr\{D < d\} = \frac{d^2 - D_{\text{min}}^2}{D_{\text{max}}^2 - D_{\text{min}}^2}$ , where  $D_{\text{min}}$  is the minimum distance between a BS and the UE, and  $d \in [D_{\text{min}}, D_{\text{max}}]$ .

According to [30], the cluster powers are calculated as follows. Let  $X_p \sim U(0, 1)$ ,  $\forall p \in \mathcal{P}$ . Let  $P'_p \triangleq X_p^{\tau_p - 1} 10^{-0.1Z_p}$ , where  $Z_p \sim \mathcal{N}(0, \zeta^2)$ ,  $\forall p \in \mathcal{P}$ .  $\tau_p$  and  $\zeta$  are given environment-dependent constants. Finally, the cluster power  $P_p$  is given by:

$$P_p = \frac{P'_p}{\sum_{n \in \mathcal{P}} P'_n}. \quad (14)$$

The CDF of  $P_{\text{Rx}}$  under  $H_0$ , i.e.,  $\Pr\{P_{\text{Rx}} \leq T | H_0\}$ , can then be calculated numerically. In particular,  $\Pr\{P_{\text{Rx}} \leq T | H_0\} = \Pr\{P_N \leq T - (\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})(\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})^* | H_0\}$ , where  $P_N$  is a random variable with chi-square distribution. Let  $f_X(x)$  and  $F_X(x)$  denote the PDF and CDF of an arbitrary random variable  $x$ , respectively. Hence, for a given  $T$ ,  $\Pr\{P_N \leq T - (\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})(\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})^* | H_0\}$  can be computed by utilizing the PDF of random variables  $\varphi_p$ ,

$d, \xi, X_p, Z_p, \forall p \in \mathcal{P}$ , and the CDF of  $P_N$ . Note that from the above discussion, the PDF and CDF of these random variables are known. More formally,  $\Pr\{P_{\text{Rx}} \leq T|H_0\}$  is given in (15). The integral in (15) is taken over the set  $\mathcal{A}$ , where  $\mathcal{A} = \{\varphi_p, d, \xi, X_p, Z_p \forall p \in \mathcal{P} | T \geq (\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})(\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})^* | H_0\}$ . In other words, the values of  $\varphi_p, d, \xi, X_p, Z_p \forall p \in \mathcal{P}$  that satisfy the condition  $T \geq (\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})(\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})^*$  are considered in (15). In Fig. 5, we numerically evaluate this integral by exploiting the midpoint rule. We use the channel model parameters as given in [30], where  $P_{\text{Tx}} = 46$  dBm, center frequency is 28 GHz, signal bandwidth is 57.6 MHz, and the cell radius is 200 m.  $P$  and  $G_{\text{min}}$  are set to 2 and 1, respectively. Note that determination process of the threshold value  $T_{\text{opt}}$  is performed offline, i.e., no online calculation is required.

### C. Phase Alignment

After cluster directions have been discovered, we need to determine the phases and magnitudes of the signals arriving at the Rx through each cluster, in order to add the multipath components coherently. This process is similar to how Rake receivers work [31]. Rake receivers use several sub-receivers called fingers to cope with multipath fading. In particular, each finger is assigned to a different multipath component, and independently performs decoding. Then, the contribution of all fingers are coherently combined. However, this method requires multiple fingers and results in a significantly more complex system compared to a single-RF-chain receiver. Here, as the cluster directions are already determined previously, instead of employing a separate finger per multipath component, each cluster is probed with the optimal single-lobe beam for that direction, in a different time slot. The received signals are then decoded, and the phases and magnitudes of the signals coming from each cluster are computed.

To coherently add the signals, consider the  $P$  clusters identified with MLBS. The phase alignment problem with  $P$  clusters can be formally written as:

$$\begin{aligned} & \underset{\{c_1, \dots, c_P\}}{\text{maximize}} && |c_1 r_1 e^{j\varphi_1} + \dots + c_P r_P e^{j\varphi_P}| \\ & \text{s.t.} && |c_i| = 1, \forall i \in \{1, \dots, P\} \end{aligned} \quad (16)$$

where  $r_i$  and  $\varphi_i$  represent the magnitude and phase of the signal arriving from cluster  $i$ , respectively,  $\forall i \in \{1, \dots, P\}$ . Here, the objective function is the magnitude of the superimposed signals received from all clusters, decision variable is the vector of unit-modulus complex weights, and the constraint is a result of analog beamforming (having only phase shifters and no amplifiers). Note that  $r_i e^{j\varphi_i}$  already accounts for transmit and receive beamformers, i.e.,  $c_i$  are selected to coherently add the signals, not to steer the beams. By setting the first cluster as reference (e.g.,  $c_1 = 1$ ), we can easily compute  $c_i, \forall i \in \{2, \dots, P\}$ . In particular,  $c_i = e^{j(\varphi_1 - \varphi_i)}$ , so that the phases of two signals coming from the first and  $i$ th clusters are aligned. Specifically, this corresponds to adding a phase offset  $c_i$  to all antennas in the sub-array that beams towards the cluster  $i$ . As all antennas in the sub-array are exposed to the same phase offset,  $c_i$  does not affect the beam directions.

Note that the phase alignment is only required after the signal directions are found. Thus, (16) needs to be solved only once, after the clusters are determined. The alignment of phases is not possible with 802.11ad search as all antennas are used to create a single wide beam, which can capture more than one signal adding destructively [32].

As explained in Section IV-A, MLBS can misidentify some clusters due to the partial overlap of neighboring beams. These duplicate clusters can be eliminated as follows: If the phases of the signals captured by neighboring beams are exactly the same, we can determine that they are in fact the same cluster. Then, the cluster with less power can be eliminated, since less received power indicates that the cluster direction is further away from the boresight of the beam. This way, even if the MLBS is set to discover more than  $P$  clusters, the excessive clusters can be eliminated at this stage.

### D. Optimal Antenna Allocation

Without loss of generality, the optimal antenna allocation at the Rx is studied in this section. Note that, MLBS identifies  $P$  cluster directions using  $B$ -lobe beams. After MLBS is completed and the cluster directions are found, here the antenna array will be divided up to  $P$  sub-arrays to beam towards the discovered clusters. If the environment exhibits fewer clusters than we expect and MLBS returns some eminently weak clusters, they will not be assigned any antennas thanks to our optimization problem presented below.

A UPA steers a beam towards a direction with a gain that is directly proportional to its number of antennas under half-wavelength spacing (which is the recommended value by the 3GPP standards [30]), thanks to ‘‘array gain’’ [33]. Let  $N_{\text{UE},x}$  and  $N_{\text{UE},y}$  denote the number of columns and number of rows, respectively, in the Rx antenna array, where  $N_{\text{UE}} = N_{\text{UE},x} \times N_{\text{UE},y}$  is the total number of Rx antenna elements. During phase alignment (see Section IV-C), the Rx acquires the SNR from each inferred cluster by allocating all its antenna elements to a single beam and measuring the corresponding received power. Let  $K_p$  denote the measured SNR during this process when the Rx beam is directed towards  $p$ th inferred cluster,  $\forall p \in \mathcal{P} \triangleq \{1, \dots, P\}$ . Note that  $K_p$  includes the effects of Tx and Rx antenna gains, Tx transmission power, cluster powers, and the path-loss.

Here, we propose to dynamically divide the Rx antenna array into sub-arrays during the data transmission phase such that each sub-array forms a beam towards the direction of a different cluster. In particular, let the array be divided into  $P$  sub-arrays, and let  $N_p$  denote the number of antennas of the  $p$ th sub-array,  $\forall p \in \mathcal{P}$ . Further suppose that the  $p$ th sub-array is beamforming towards the direction of the  $p$ th cluster. By coherently combining signals from multiple directions as explained in Section IV-C, the SNR at the Rx during the data transmission is given by  $\text{SNR} = \left( \sum_{p \in \mathcal{P}} N_p \sqrt{K_p} / N_{\text{UE}} \right)^2$ . Note that existing approaches find the strongest cluster and allocate all antennas to its direction, i.e.,  $\text{SNR} = K_p^*$  if  $\max(K_1, \dots, K_P) = K_p^*$ .

By exploiting different beam directions, we aim at achieving higher average transmission rate between the Tx and the Rx



$$\begin{aligned} \Pr\{P_{\text{Rx}} \leq T | H_0\} &= \int \cdots \int f_{\varphi_1}(\varphi_1) \cdots f_{\varphi_P}(\varphi_P) f_D(d) f_{\Xi}(\xi) f_{X_1}(X_1) \cdots f_{X_P}(X_P) \\ &\quad f_{Z_1}(Z_1) \cdots f_{Z_P}(Z_P) F_{P_N}(T - (\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p}) (\sum_{p \in \mathcal{P}} r_p e^{j\varphi_p})^* | H_0) \\ &\quad d\varphi_1 \cdots d\varphi_P dd d\xi dX_1 \cdots dX_P dZ_1 \cdots dZ_P \end{aligned} \quad (15)$$

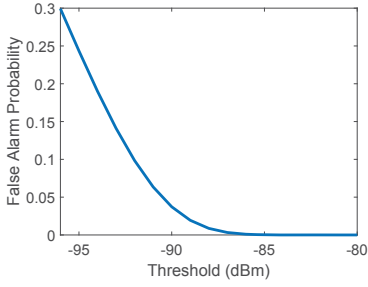


Fig. 5. False alarm probability under various threshold values.

under dynamic blockage. This way, even when one or more beams are blocked, the Tx-Rx link will remain operational. Dynamic blockage is commonly modeled in time by a Poisson process [5], [6]. Thus, the arrival time of a blocker is exponentially distributed. Here, we assume that the blockers are impenetrable, and the blockage processes of various directions are independent and identically distributed. Let  $\rho$  denote the blockage rate, i.e., the rate parameter of the exponential distribution. The exact formulation of the associated optimization problem depends on  $P$ , i.e., the size of  $\mathcal{P}$ . Therefore, for simplicity, here we focus on the special case  $P = 2$ . The problem formulation for other  $P$  values is straightforward. To generate 2 lobes, the antenna array is horizontally divided into two, meaning that the decision parameter of the optimization problem is the number of rows allocated to a sub-array. Let  $N_1 = N_{\text{UE},x} \times N_{\text{UE},y}^{(1)}$  and  $N_2 = N_{\text{UE},x} \times N_{\text{UE},y}^{(2)}$  denote the number of antennas of the first and second sub-arrays, respectively, where  $N_{\text{UE},y}^{(1)}$  and  $N_{\text{UE},y}^{(2)}$  represent the numbers of rows allocated to the two directions, respectively. Let  $t_1$  denote the time that passes until the one of the beams is blocked since the start of data transmission. Similarly, let  $t_2$  denote the time that passes until the blockage of the second beam, starting from when the first direction was blocked. Note that due to the merging and memoryless properties of Poisson process,  $\mathbb{E}[t_1] = (2\rho)^{-1}$  and  $\mathbb{E}[t_2] = \rho^{-1}$ . Let  $T_d$  denote the duration of data transmission phase. Accordingly, the optimization problem that maximizes the total transmitted data over the period of time that starts with the data transmission phase and lasts until all the beams are blocked or the data transmission duration expires (whichever occurs first) is formulated as:

$$\begin{aligned} &\underset{\{N_{\text{UE},y}^{(1)}, N_{\text{UE},y}^{(2)}\}}{\text{maximize}} && \mathbb{E} [\min\{T_d, t_1\} \log_2(1 + U_B) + \\ &\text{s.t.} && \max\{\min\{T_d - t_1, t_2\}, 0\} \log_2(1 + U_l)] \quad (17) \\ &&& N_{\text{UE},y}^{(1)} + N_{\text{UE},y}^{(2)} = N_{\text{UE},y}, \quad N_{\text{UE},y}^{(1)}, N_{\text{UE},y}^{(2)} \in \{0\} \cup \mathbb{Z}^+ \end{aligned}$$

Here,  $U_B \triangleq (N_{\text{UE},y}^{(1)} \sqrt{K_1} / N_{\text{UE},y} + N_{\text{UE},y}^{(2)} \sqrt{K_2} / N_{\text{UE},y})^2$  and  $U_l \triangleq (N_{\text{UE},y}^{(l)} \sqrt{K_l} / N_{\text{UE},y})^2$  where the  $l$ th beam represents the last blocked beam,  $l \in \{1, 2\}$ . The expectation in (17) is taken with respect to  $t_1$  and  $t_2$ , as they represent the random blockage.  $U_B$  is the SNR at the Rx when both directions are used for data transmission. Therefore, the term  $\min\{T_d, t_1\} \log_2(1 + U_B)$  is the amount of transmitted data until the one of the beams is blocked since the start of data transmission or the data transmission duration expires (whichever occurs first). Similarly,  $U_l$  is the SNR at the Rx when only the  $l$ th direction is used for data transmission, so the second term in the objective function is the amount of transmitted data until the blockage of the  $l$ th beam ( $t_2$ ) or the data transmission duration expires ( $T_d - t_1$ ), i.e., whichever occurs first after the first blockage event. Finally, the objective function in (17) can also be written as  $\mathbb{E}[\min\{T_d, t_1\} \log_2(1 + U_B) + \mathbb{E}[\max\{\min\{T_d - t_1, t_2\}, 0\} (\log_2(1 + U_1) + \log_2(1 + U_2)) / 2]$ . Note that the probability of the first cluster being blocked before the other one is 0.5 (and vice versa) due to the equiprobable events. The closed-form expressions for the expected values in this expression can be given by:

$$\mathbb{E} [\min\{T_d, t_1\}] = \frac{1 - e^{-2\rho T_d}}{2\rho} \quad (18)$$

$$\mathbb{E} [\max\{\min\{T_d - t_1, t_2\}, 0\}] = \frac{1 - e^{-2\rho T_d}}{\rho} - \frac{2e^{-\rho T_d} (1 - e^{-\rho T_d})}{\rho} \quad (19)$$

Hence, expectation in the objective function can be eliminated. The final optimization problem is still a nonlinear integer programming problem, which is NP-hard. However, it can be converted to a convex optimization problem, if the integer restrictions of (17) are relaxed such that  $N_{\text{UE},y}^{(p)}$ 's are allowed to take any nonnegative real number, i.e.,  $N_{\text{UE},y}^{(p)} \geq 0, \forall p \in \mathcal{P}$ . Particularly, the objective and constraint functions of this relaxed problem are concave and convex with respect to  $N_{\text{UE},y}^{(p)}$   $\forall p \in \mathcal{P}$ , respectively. Hence, this problem can be numerically and efficiently solved using existing methods, such as the gradient descend. Let  $N_{\text{UE},y}^{(p)*} \forall p \in \mathcal{P}$  denote the optimal solutions of the relaxed problem. That solution can be used to determine  $N_{\text{UE},y}^{(p)}$ 's of the original problem. Even though it does not guarantee the optimal solution,  $N_{\text{UE},y}^{(p)}$ 's can be set to  $\lceil N_{\text{UE},y}^{(p)*} \rceil$  or  $\lfloor N_{\text{UE},y}^{(p)*} \rfloor \forall p \in \mathcal{P}$ , such that  $\sum_{p \in \mathcal{P}} N_{\text{UE},y}^{(p)} = N_{\text{UE},y}$  ( $\lceil \cdot \rceil$  and  $\lfloor \cdot \rfloor$  are ceiling and floor functions, respectively). This way, the integer restrictions can be satisfied. Note that the solution of the relaxed problem provides an upper-bound for the original problem. We can compare the performance gap be-

tween this upper-bound and the one obtained by using ceiling and floor functions. If the gap is relatively small, this efficient solution can be considered good enough. Alternatively, we can apply *branch and bound* algorithm, which is one of the most common methods to solve integer programming problems. This method exploits the solution of the relaxed problem to add certain bounding constraints to the original one. Although these extra constraints lead to the optimal solution, which is purely integer, the complexity of the algorithm is high. Thankfully, given that  $P$  is typically a small number, the size of our problem in (17) is not large, meaning that branch and bound method can also be efficiently applied here.

The problem formulation in the case where  $P > 2$  is straightforward and omitted here for brevity. For example, in the case of  $P = 3$ , one more decision variable, which is  $N_{UE,y}^{(3)}$ , is added to the problem. Furthermore,  $t_3$  that denotes the time that passes until the blockage of the third beam (since the blockage of the second beam) needs to be incorporated.

Note that in a given problem instance, the utilization of more than  $P$  clusters may result in a higher average data rate than utilizing only  $P$  of them. However, it is not possible to obtain this optimal value before running our beam-search algorithm MLBS to infer the cluster directions and their associated powers. As discussed before, due to fewer reflections in mmW spectrum, only a few channel clusters are present in mmW channels. In fact, a typical number of clusters is less than five [7]. Therefore, using the mmW channel statistics, an appropriate number for  $P$  can be selected, i.e., five. Furthermore, the initial selection of  $P$  may depend on some system requirements. In particular, as  $P$  increases, the discovery time of the clusters increases as well, see Fig. 9(c). For example, if the maximum allowable discovery time for a specific system is 20 slots, at most four clusters can be discovered within that duration (according to Fig. 9(c)), leading to a condition of  $P \leq 4$ .

### E. Protocol Design

In this section, we present the SmartLink protocol for the IA process in mmW systems. SmartLink utilizes the aforementioned steps and defines the required message exchange between the BS and the UE to establish the multi-directional link. We first briefly explain the IA process in LTE and current 5G standard, and then discuss our proposed protocol.

In LTE systems, the IA procedure utilizes an omnidirectional signal called the Cell Reference Signal (CRS), which is regularly monitored by each UE to create a wideband channel estimate that can be used both for demodulating downlink transmissions and for estimating the channel quality [17]. However, IA needs to be done directionally in 5G mmW systems, to achieve full coverage potential. As a result, when the Tx and the Rx beams are not aligned, the directional link cannot be established.

To find a suitable directional link, recent 5G specifications require that the BS periodically broadcasts synchronization signal (SS) blocks towards pre-defined number of beam directions in a sequential manner [12]. These SS blocks carry primary synchronization signals (PSS), secondary synchronization signals (SSS), and physical broadcast channel

(PBCH) information [34]. PSS is mainly used for initial symbol boundary synchronization to the NR cell and the SSS is used for detection of cell and beam IDs. When the UE enters the coverage area of a BS, it listens to an SS burst (consisting of multiple SS blocks) and measures the signal quality of different beams. It then determines the beam for which the received power is maximum (and above a predefined threshold). This beam will be chosen for subsequent transmissions/receptions. After determining the best BS beam, the UE has to wait for the BS to schedule the random access channel (RACH) opportunity for the beam direction that the UE has selected [17]. During a RACH opportunity, UE performs random access, implicitly informing the BS of its selected beam direction. Note that current 5G specifications do not standardize how beam sweeping will be performed at the UE.

Our MLBS algorithm can be directly applied at the UE side, without changing the default 5G IA process. This reduces the search time at the UE side and combats blockage to a certain extent, if the BS beam is relatively large. With some small changes in the 5G IA structure, MLBS can be employed at both the BS and the UE, as shown in Fig. 6. In SmartLink, the BS first selects a quasi-omnidirectional beam and transmits multiple copies of the same SS block over that beam. A relatively large quasi-omnidirectional beam spans multiple channel clusters. During this transmission, the UE constantly measures the received power, while steering its receive beams according to the MLBS algorithm. This way, the UE identifies the best Rx beam directions corresponding to the current quasi-omnidirectional beam of the BS. This process is repeated for all quasi-omnidirectional beams at the BS. The UE compares each BS quasi-omnidirectional beam in terms of maximizing (17), i.e., average data rate. Similar to 5G, after the BS finishes transmitting the SS blocks, the UE waits for the BS to switch to Rx mode, i.e., waits for RACH opportunities. There is a separate RACH resource in time and frequency dedicated to each BS quasi-omni beam. The information of these resources is obtained at the UE via PBCH and system information block (SIB) transmitted by the BS. The UE then sends its RACH preamble on the RACH resources corresponding to the selected quasi-omnidirectional BS beam. Note that the UE transmits the preamble along the best beam directions it found via MLBS, while the BS receives it using a single-lobe beam, which is the selected quasi-omnidirectional beam. After receiving the RACH preamble, the BS runs MLBS algorithm to find the best narrow beam directions for the subsequent transmissions. To do that, the UE transmits consecutive uplink reference signals, i.e., sounding reference signals (SRS), on scheduled resources specific to the UE (see Fig. 6(b)).

Because the BS and the UE operate on the same frequency, the number of clusters they expect to observe will be the same. Thus, they will aim at identifying the same number of clusters when running MLBS. Therefore, the BS can compute the depth of UEs decision tree  $\eta_{UE}$ , without prior communication with the UE. This allows the BS to know how many each SS block should be transmitted for (UE to run MLBS). Specifically, BS sends  $\tau_{UE}$  consecutive SS blocks through

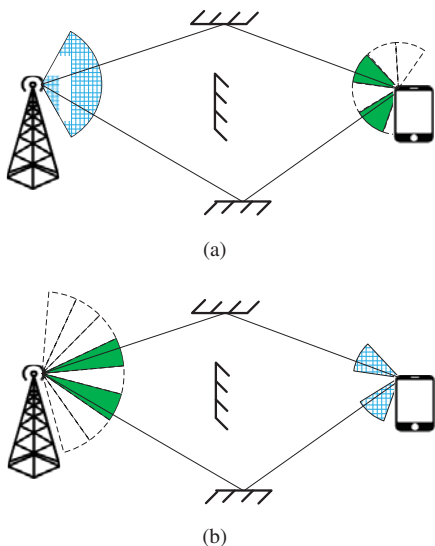


Fig. 6. SmartLink protocol, where squared blue shows the transmitting beams and solid green shows the receiving beams. (a) Step 1, where the BS sends consecutive SS blocks over the same quasi-omnidirectional beam for the UE to run MLBS. (b) Step 2, where the UE sends the uplink reference signals to the BS over the best beam directions found in Step 1 and the BS runs MLBS.

a selected quasi-omnidirectional beam, where  $\tau_{UE} = \eta_{UE}$  is the maximum required number of SS blocks for the UE to run the MLBS algorithm. This is repeated  $L_{\text{omni}}$  times where  $L_{\text{omni}}$  is the number of quasi-omnidirectional beams at the BS ( $L_{\text{omni}} \ll L$ ). Following that, the UE determines the best receive beams for the best quasi-omnidirectional BS transmit beam and sends  $\tau_{BS} = \eta_{BS}$  SRS messages to the BS for BS beam training, where  $\eta_{BS}$  is the depth of the BSs decision tree. After collecting SRS messages from the UE, the BS finally selects the best transmit beams corresponding to the best UE receive beams. Note that as the UE has already selected its best beam directions previously, the BS training is performed only one time. This way, the multi-directional link can be established in  $\tau_{BS} + L_{\text{omni}}\tau_{UE}$  signal transmissions.

After the narrow beam directions are identified, either the BS or the UE proceeds with the phase alignment to ensure that the signals are in phase. Note that the party that does the phase alignment will continue to do so for the subsequent uplink and downlink transmissions. Finally, the BS solves the optimal antenna allocation problem in (17) for the identified cluster directions, while the UE is already transmitting optimally. Solving the joint optimization problem is left for future work.

## V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed scheme through extensive trace-driven emulations. We compare our scheme with 802.11ad-like and 5G beam search approaches, in which the search time scales linearly with  $L$ .

### A. Experimental Results

We first conduct extensive experiments to obtain the mmW channel characteristics to be used in our simulations.  $4 \times 4$  UPAs are used in our experiments with  $d_x = 0.5\lambda$  and



Fig. 7. Test setup using  $4 \times 4$  UPAs both at the Tx and the Rx side, signal generator, and vector signal analyzer.

$d_y = 0.6\lambda$ . The antenna gain at the broadside of the array is 12 dB. For the sake of measuring the received power, a continuous wave with 5 dBm amplitude is transmitted over the 29 GHz band, which is a candidate band for 5G communications. Keysight E8257D-ATO-8384 PSG signal generator is used to generate the waveform. At the Rx side, the array is connected to Keysight PXA-550-MY55002004 vector signal analyzer (VSA). To steer the transmit/receive beams to desired directions, antenna arrays are connected to microcontrollers, which are interfaced with the PC through serial port. The whole setup with the Tx, Rx, PSG, and the VSA can be seen in Fig. 7.

We test several LOS and NLOS scenarios with a Tx-Rx separation of 3 m, where the NLOS path is created by a  $1.2 \text{ m} \times 1.2 \text{ m}$  metal reflector. The effective beam scanning range of the UPAs in our experiments are  $\pm 60^\circ$  from broadside, in both azimuth and elevation (effective FOV =  $\sim 120^\circ$  in each direction). Beyond that, the antenna gain drops significantly as a result of the non-ideal behavior of the antenna elements. To experimentally obtain the AoA profile, we exhaustively scanned the 3D space within the effective beam scanning range of the antenna arrays under different scenarios and collected RSS measurements from the 3D space.

In Fig. 8, we evaluate the effect of the detection threshold  $T$ , the improvement in data-rate via using a multi-directional link, and the performance of the MLBS algorithm using trace-driven simulations. We take the average values obtained from all LOS and NLOS scenarios. In Fig. 8(a), when  $T$  increases,  $\Pr\{\text{MD}\}$  increases and  $\Pr\{\text{FA}\}$  decreases, as expected. Thus, for this experimental environment, when  $T$  is approximately equal to  $-68 \text{ dBm}$ , both  $\Pr\{\text{MD}\}$  and  $\Pr\{\text{FA}\}$  stay below 0.2. We then compare the throughput performances of a benchmark scheme (single-beam) that allocates all antennas to form a beam towards the strongest cluster and the proposed multi-beam scheme, under certain blockage probabilities of the underlying links. In Fig. 8(b), we set the data transmission duration  $T_d$  to 10 ms. The results indicate that when the blockage rate is low, benchmark scheme performs as good as the proposed scheme. In that case, there is no benefit of using multiple beams. However, when blockage rate is above 5 bl/s, the multi-directional scheme outperforms the single-beam scheme up to 10%. Note that when we allocate the antennas to form multiple beams, the instantaneous transmission rate decreases. However, under the dynamic blockage, the proposed scheme

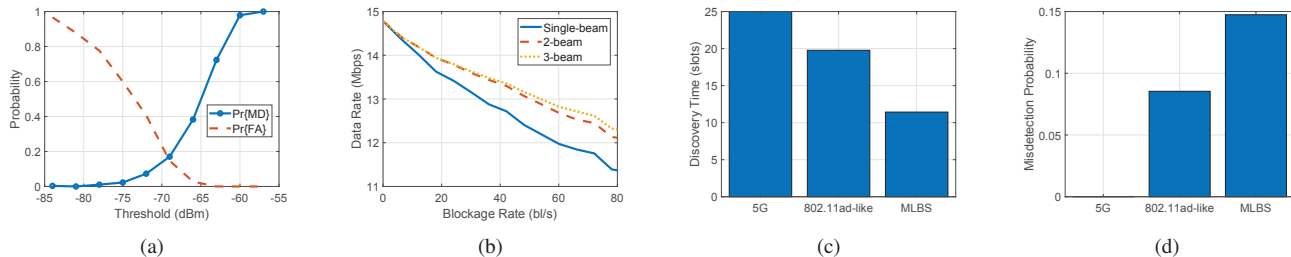


Fig. 8. Trace-driven simulation results: (a) comparison of misdetection and false alarm probabilities for a given threshold, (b) data rate vs. blockage rate (comparison of single-beam and multi-beam schemes when  $T_d = 10$  ms), (c) Discovery time comparison of MLBS with other search schemes, (d) Misdetection probability comparison of MLBS with other search schemes.

allows an uninterrupted transmission due to the utilization of multiple beams, leading to higher data rates on average. Fig. 8(c) shows the discovery time comparison of MLBS with 5G and 802.11ad-like beam searching schemes, when all algorithms aim at identifying three channel clusters. 5G beam search corresponds to the traditional exhaustive beam scan and 802.11ad-like search represents the hierarchical beam search with the goal of identifying multiple channel clusters. For computational efficiency, the comparison is done over a subsampled version of the original channel, where the new channel matrix is  $5 \times 5$  with 25 total directions. The results indicate that MLBS offers significant reduction in search time (54% compared to 5G and 43% compared to 802.11ad-like beam searches). On the other hand, it experiences a slightly higher misdetection probability compared to 802.11ad-like search, as observed in 8(d). This 0.065 increase in the misdetection probability is due to a higher detection threshold ( $T$ ) selection for limiting the false alarm probability. Specifically, MLBS uses a detection threshold of  $T = -68$  dBm (as found in Fig. 8(a)), whereas the 802.11ad-like scheme uses  $T = -80$  dBm. Finally, note that the misdetection probability in Fig. 8(a) is the misdetection of a single measurement taken with a multi-lobe beam, while the misdetection probability in Fig. 8(d) is the overall misdetection of MLBS. Deviating from  $T = -68$  dBm found in Fig. 8(a) results in a higher overall misdetection.

## B. Simulation Results

In our simulation, we use 3GPP channel models [30] where  $P_{Tx} = 46$  dBm, the operating frequency is 29 GHz, the bandwidth is 57.6 MHz, and the cell radius is 200 m. A UPA of half-wavelength antenna spacing is implemented with various numbers of antennas.

A numerical comparison of MLBS with 5G, 802.11ad-like, and Rapid-Link [19] beam searching schemes are shown in Fig. 9. For the first stage of the 802.11ad-like scheme, we used 6 quasi-omnidirectional beams with  $60^\circ$  beamwidth each. In addition, we used 4 bins within Rapid-Link, as in [19]. Fig. 9(a) depicts the performances of the algorithms, when they aim at identifying 2 channel clusters. Clearly, discovery times of all schemes increase with increasing  $L$ . However, the increase is linear for 5G and 802.11ad-like scheme, whereas it is logarithmic for MLBS and Rapid-Link. MLBS decreases the discovery time by 88% compared to 5G beam scan, 65% compared to 802.11ad-like scheme, and 56% compared to

Rapid-Link. In Fig. 9(b), the algorithms aim at identifying 6 channel clusters. Note that compared to Fig. 9(a), curves in Fig. 9(b) are closer to each other. The search time of 5G and Rapid-Link remains the same, whereas the search time of the 802.11ad-like scheme is higher, as more clusters means more quasi-omnidirectional regions to search in the second stage. Similarly, the search time of MLBS is also higher, since the number of rows in the decision matrix,  $\binom{L}{P}$ , increases with increasing  $P$  (when  $P < L/2$ ). As a result, the gap between the curves shrink when  $P$  is larger. Notice that when  $P = 6$ , the Rapid-Link and MLBS perform almost identically. As the clusters to be discovered increases, the performance of both MLBS and 802.11ad-like gets closer to that of the exhaustive beam scan, as we can observe in Fig. 9(c). For discovering more than 6 clusters, Rapid-Link outperforms MLBS. However, due to fewer reflections in mmW spectrum,  $P$  is typically less than 5 [7].

In Fig. 10(a), the comparison of  $\Pr\{\text{MD}\}$  and  $\Pr\{\text{FA}\}$  for various number of Rx beams is provided when  $N_{UE} = 256$  and  $L = 60$ . The results are obtained for various  $\gamma_{FA}$  values that is used to determine the threshold  $T$  as explained in Section IV-B. When  $B = 2$  and  $\gamma_{FA} = 0.04$ , both  $\Pr\{\text{MD}\}$  and  $\Pr\{\text{FA}\}$  are roughly equal to 0.04. Fig. 10(a) also shows that when the antenna gain per cluster increases,  $\Pr\{\text{MD}\}$  decreases. (The case of  $B = 15$  is shown here to present the misdetection and false alarm probabilities of Rapid-Link per measurement when 4 bins are used.) Note that  $\Pr\{\text{FA}\} = \gamma_{FA}$  for all cases, which shows that our analysis on  $T$  is valid. Fig. 10(b) demonstrates the effect of blockage rate on the data rate when  $T_d = 10$  ms and  $N_{UE} = 256$ . The results here verify the ones obtained via experiments, which are shown in Fig. 8(b). We note that  $T_d$  also has the same impact as  $\rho$  on the data rate. In Fig. 10(c), we investigate the outage performance of single-beam and multi-beam schemes when  $N_{UE} = 256$  and  $T_d = 10$  ms. Especially for the wireless systems that require very-high reliability, outage duration is an important performance metric. Even when  $\rho = 80$  bl/s, the average outage duration of 3-beam scheme is almost zero. On the other hand, single-beam scheme leads to 3 ms outage duration, meaning that on average 30% of the transmission time is wasted.

## VI. CONCLUSIONS

In this paper, we proposed an efficient communication protocol for mmW systems called SmartLink. SmartLink

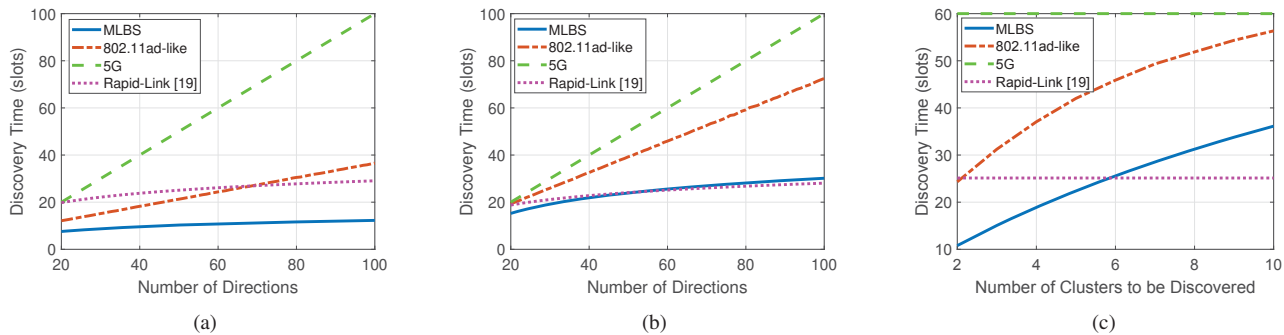


Fig. 9. Numerical results: comparison of cluster discovery times between MLBS, 5G, 802.11ad-like, and Rapid-Link [19] beam search approaches.  $B$  used in MLBS is selected optimally by solving (7), for the given  $P$  and  $L$ . Rapid-Link uses 4 bins. (a) Discovering 2 clusters, (b) discovering 6 clusters, (c) number of clusters is varied and  $L = 60$ .

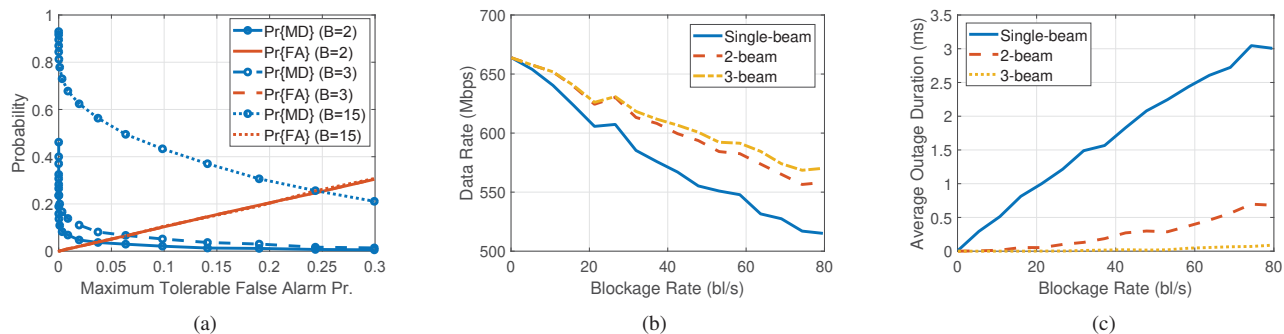


Fig. 10. Simulation results: (a) comparison of misdetection and false alarm probabilities for various number of simultaneous beams when  $N_{UE} = 256$  and  $L = 60$ , (b) data rate vs. blockage rate and (c) average outage duration vs. blockage rate (comparison of single-beam and multi-beam schemes when  $N_{UE} = 256$  and  $T_d = 10$  ms).

utilizes multiple channel clusters in the mmW channel between the BS and the UE to combat blockage. It uses a novel search scheme called MLBS, which discovers multiple clusters in logarithmic time with respect to the number of beam directions. Discovered clusters are then simultaneously used for transmission/reception, and the signals coming from different directions are coherently combined. As the probability of all channel clusters being blocked at the same time is low, SmartLink provides an effective mechanism to maintain communications and improve the data rate by up to 10%.

In our experiments, we were limited by the capabilities of our setup, such as analog-only beamforming and short Tx-Rx separation. Our future work will focus on multi-lobe beam search methods under hybrid beamforming. With hybrid beamforming, inferred clusters can be exploited for multiplexing gain, to transmit a separate independent data stream from each one of the inferred clusters. We will also investigate applications of SmartLink to other network architectures such as coordinated multipoint (CoMP).

#### ACKNOWLEDGMENTS

Authors would like to thank Keysight Technologies for providing the experimental setup used in this paper. We would also like to thank Dean Gienger and Andrew Smail for providing technical expertise regarding RF measurements and help building the setup. This research was supported in part

by NSF (grants CNS-1409172, IIP-1265960, CNS-1513649, CNS-1563655, CNS-1731164, and IIP-1822071) and by the Broadband Wireless Access & Applications Center (BWAC). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF.

#### REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [2] IEEE Computer Society, "IEEE Standard-part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 3: Enhancements for very high throughput in the 60 GHz band (adoption of IEEE std 802.11ad-2012)," 2014. [Online]. Available: <https://standards.ieee.org/findstds/standard/802.11ad-2012.html>.
- [3] R. Ford, M. Zhang, M. Mezzavilla, S. Dutta, S. Rangan, and M. Zorzi, "Achieving ultra-low latency in 5G millimeter wave cellular networks," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 196–203, 2017.
- [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] B. Han, L. Wang, and H. D. Schotten, "A 3D human body blockage model for outdoor millimeter-wave cellular communication," *Physical Communication*, vol. 25, pp. 502–510, Dec. 2017.
- [6] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, M. R. Akdeniz, E. Aryafar, N. Himayat, S. Andreev, and Y. Koucheryavy, "On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10 124–10 138, Nov. 2017.

- [7] 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 J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.
- [8] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.
- [9] C. N. Barati, S. A. Hosseini, S. Rangan, P. Liu, T. Korakis, S. S. Panwar, and T. S. Rappaport, "Directional cell discovery in millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6664–6678, 2015.
- [10] I. Aykin and M. Krunz, "Fastlink: An efficient initial access protocol for millimeter wave systems," in *Proc. of the 21st ACM MSWiM Conference*, Montreal, CA, Oct. 2018, pp. 109–117.
- [11] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 70–77, 2014.
- [12] 3GPP TR 38.802 v14.2.0, "Study on new radio access technology-physical layer aspects (release 14)," Sep. 2017. [Online]. Available: <http://www.3gpp.org/ftp/Specs/archive/38series/38.802/38802e20.zip>.
- [13] A. Zhou, L. Wu, S. Xu, H. Ma, T. Wei, and X. Zhang, "Following the shadow: Agile 3-D beam-steering for 60 GHz wireless networks," in *Proc. of the IEEE INFOCOM*, Honolulu, HI, Apr. 2018.
- [14] D. Steinmetzer, D. Wegemer, M. Schulz, J. Widmer, and M. Hollick, "Compressive millimeter-wave sector selection in off-the-shelf IEEE 802.11ad devices," in *Proc. of the 13th ACM International Conference on Emerging Networking Experiments and Technologies*, Seoul, South Korea, Dec. 2017, pp. 414–425.
- [15] H. S. Ghadikolaei, Y. Yang, M. Petrova, K. W. Sung, and C. Fischione, "Fast and reliable initial cell-search for mmWave networks," in *Proc. of the 2nd ACM Workshop on Millimeter Wave Networks and Sensing Systems*. New Delhi, India: ACM, Oct. 2018, pp. 57–62.
- [16] M. Hashemi, C. E. Koksall, and N. B. Shroff, "Out-of-band millimeter wave beamforming and communications to achieve low latency and high energy efficiency in 5G systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 875–888, 2018.
- [17] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, Firstquarter 2019.
- [18] O. Abari, H. Hassanieh, M. Rodriguez, and D. Katabi, "Millimeter wave communications: From point-to-point links to agile network connections," in *Proc. of the 15th ACM Workshop on Hot Topics in Networks*, Atlanta, GA, Nov. 2016, pp. 169–175.
- [19] H. Hassanieh, O. Abari, M. Rodriguez, M. Abdelghany, D. Katabi, and P. Indyk, "Fast millimeter wave beam alignment," in *Proc. of the 2018 Conference of the ACM Special Interest Group on Data Communication*, Budapest, Hungary, Aug. 2018, pp. 432–445.
- [20] Q. Xue, X. Fang, and C.-X. Wang, "Beamspace SU-MIMO for future millimeter wave wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1564–1575, 2017.
- [21] Q. Xue, X. Fang, M. Xiao, S. Mumtaz, and J. Rodriguez, "Beam management for millimeter-wave beamspace MU-MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 205–217, Jan. 2019.
- [22] Y. Liu, X. Fang, M. Xiao, and S. Mumtaz, "Decentralized beam pair selection in multi-beam millimeter-wave networks," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2722–2737, 2018.
- [23] I. Aykin, B. Akgun, and M. Krunz, "Smartlink: Exploiting channel clustering effects for reliable millimeter wave communications," in *Proc. of the IEEE INFOCOM*, Paris, France, Apr. 2019.
- [24] 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.
- [25] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmWave cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 40–47, 2016.
- [26] J. Stuckman and G.-Q. Zhang, "Mastermind is NP-complete," *INFO-COMP Journal of Computer Science*, vol. 5, no. 2, pp. 25–28, June 2006.
- [27] M. Di Renzo, "Stochastic geometry modeling and analysis of multi-tier millimeter wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 5038–5057, Sep. 2015.
- [28] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.
- [29] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Tech.*, vol. 64, no. 7, pp. 2207–2225, 2016.
- [30] 3GPP, "Study on channel model for frequency spectrum above 6 GHz," 3GPP, Tech. Rep. TR 38.900 v14.3.1, July 2017.
- [31] S. Hara and R. Prasad, "Overview of multicarrier CDMA," *IEEE Commun. Mag.*, vol. 35, no. 12, pp. 126–133, 1997.
- [32] M. E. Rasekh, Z. Marzi, Y. Zhu, U. Madhow, and H. Zheng, "Noncoherent mmWave path tracking," in *Proc. of the 18th ACM International Workshop on Mobile Computing Systems and Applications*, Sonoma, CA, Feb. 2017, pp. 13–18.
- [33] R. S. Elliot, *Antenna theory and design*. John Wiley & Sons, 2006.
- [34] 3GPP, "Study on new radio access technology physical layer aspects," 3GPP, Tech. Rep. TR 38.802 v14.2.0, Sep. 2017.



She is the recipient of ACM MSWiM 2018 Best Paper Award.



characterization, robust mmWave system design, wireless communications and networking, with emphasis on designing secure multiuser MIMO systems.



Krunz was a Visiting Chair of Excellence at the University of Carlos III de Madrid. He held visiting research positions at UTS, INRIA-Sophia Antipolis, HP Labs, University of Paris VI, University of Paris V, University of Jordan, and US West Advanced Technologies. Dr. Krunz's research interests are in wireless communications and networking, with emphasis on resource management, adaptive protocols, and security issues. He has published more than 280 journal articles and peer-reviewed conference papers, and is a co-inventor on several US patents. He is an IEEE Fellow, an Arizona Engineering Faculty Fellow (2011-2014), and an IEEE Communications Society Distinguished Lecturer (2013 and 2014). He was the recipient of the 2012 IEEE TCCC Outstanding Service Award. He received the NSF CAREER award in 1998. He currently serves as the Editor-in-Chief for the IEEE Transactions on Mobile Computing. He previously served on the editorial boards for the IEEE Transactions on Cognitive Communications and Networks, IEEE/ACM Transactions on Networking, IEEE TMC, IEEE Transactions on Network and Service Management, Computer Communications Journal, and IEEE Communications Interactive Magazine. He was the general vice-chair for WiOpt 2016 and general co-chair for WiSec12. He was the TPC chair for WCNC 2016 (Networking Track), INFOCOM'04, SECON'05, WoWMoM'06, and Hot Interconnects 9. He has served on the steering and advisory committees of numerous conferences and on the panels of several funding agencies. He was a keynote speaker, an invited panelist, and a tutorial presenter at numerous international conferences. See <http://www2.engr.arizona.edu/~krunz/> for more details.

**Irmak Aykin** received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey in 2012. In 2014, she earned the M.S. degree from Bilkent University. She is currently pursuing her Ph.D. degree in the Department of Electrical and Computer Engineering at the University of Arizona, working in Wireless Communications and Networking Laboratory directed by Dr. Marwan Krunz. Her current research interests include millimeter-wave systems, wireless networking, algorithm design and machine learning.

**Berk Akgun** received the B.S. and M.S. degrees in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2012 and in 2014, respectively. From 2012 to 2014, he was a software design engineer at the Communication and Information Technologies Division of Aselsan, Ankara, Turkey. He is currently working towards the Ph.D. degree at the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ, USA. His current research

**Marwan Krunz** (S'93-M'95-SM'04-F'10) is the Kenneth VonBehren Endowed Professor in the ECE Department at the University of Arizona. He is also an affiliated faculty member of the University of Technology Sydney. He directs the Broadband Wireless Access and Applications Center, a multi-university industry-focused NSF center that includes affiliates from industry and government labs. Previously, he served as the UA site director for Connection One, an NSF IUCRC that focuses on wireless communication circuits and systems. In 2010, Dr.