Online Learning-based Beamwidth Optimization for Initial Access in Millimeter Wave Cellular Networks

Mingjie Feng, *Member, IEEE* and Marwan Krunz, *Fellow, IEEE*

***Abstract*—The use of highly directional antennas in millimeter wave (mmWave) cellular networks necessitates precise beam alignment between a base station (BS) and a user equipment (UE), which requires beam sweeping over a large number of directions and causes high initial access (IA) delay. Intuitively, wider beams could lower this delay by requiring fewer sweeping directions. However, this results in a weak received signal and a higher risk of misdetection, which potentially increases the expected IA delay by requiring more rounds of sweeping to discover a UE. In this paper, we propose a beamwidth optimization framework for both single-link and dual-link mmWave cellular networks, aiming to minimize the beam sweeping delay for a successful IA. We first analyze the impact of beamwidth on misdetection probability and formulate the beamwidth optimization problem accordingly. Then, we present the beam sweeping protocols that support beamwidth optimization. After that, we formulate the beamwidth optimization problem based on the multi-armed bandit framework and propose an online learning-based solution. Simulation results show that the proposed solutions can decrease the beam sweeping delay by more than 50% compared to the benchmark schemes.**

***Index Terms*—Millimeter wave communications; 5G NR; 6G cellular network; initial access; beam sweeping; beamwidth optimization.**

I. Introduction

Millimeter wave (mmWave) communication is one of the enabling technologies for Fifth Generation (5G) wireless systems [2]. By operating at mmWave bands, multi-Gbps data rates per user can be achieved. Due to their great potential, mmWave bands are utilized by both next-generation WLANs (e.g., 802.11ad and 802.11ay) as well as 5G New Radio (NR), and are expected to be more intensively utilized in Sixth Generation (6G) wireless systems [3].

Meanwhile, mmWave communications suffer from high propagation loss, limited scattering, and vulnerability to blockage. To compensate for channel losses, large electronically steerable antenna arrays need to be employed at both the transmitter (Tx) and the receiver (Rx) to achieve highly directional transmissions/receptions with high antenna gains. However, the use of narrow beams complicates the initial access (IA) process in mmWave cellular networks. IA is the process of network discovery and establishing a connection between a user equipment (UE) and a base station (BS). In mmWave cellular networks, IA requires beam alignment to find the best BS-UE beam pair through a *beam sweeping* process, which is performed before data transmission. Beam sweeping is implemented by the transmission and reception of control signals with directional beams, allowing the link qualities of different Tx/Rx directions to be measured and the best beam pair to be found. In 5G NR systems, synchronization signals (SS) are sent by the BS and received by the UE along different directions, allowing the BS and UE to identify the best beam directions. Such an approach is likely to be applied in B5G/6G mmWave cellular systems, which also rely on narrow directional beams to compensate for channel loss. As narrow beams are used, a large number of directions need to be sequentially scanned to cover the entire angular domain, resulting in a high IA delay.

An intuitive approach to lower the IA delay is using wider beams to decrease the number of directions used during beam sweeping. Such an approach, which has been considered for both WLANs [4] and cellular networks [5], is typically based on a two-stage hierarchical search. In the first stage, a small number of wide sectors (or cones in 3D search) are swept (known as P1 in 5G NR). In the second stage, the search is refined within the best found coarse sector (known as P2 at the BS side and P3 at the UE side in 5G NR). However, with wider beams, the received power decreases significantly due to lower antenna gains. In addition, when signals are transmitted and received across wide sectors, it is more likely that several signal clusters will be received within the same sector. Due to the phase difference, these clusters may add destructively, which potentially degrades the received signal. If no signal is detected during one round of beam sweeping, additional rounds will be required, which in turn increases the IA delay.

Considering the tradeoff between the increased likelihood of misdetection under wider beams and the higher beam sweeping overhead under narrower beams, the optimal beamwidths at the BS and the UE need to be studied so as to minimize the beam sweeping delay for a successful IA. From the perspective of a UE, the misdetection probability (i.e., the probability that no signal can be successfully detected during a complete round of beam sweeping) is low when the channel between the UE and the BS is strong. Then, it has higher confidence in using wide beams to reduce the beam sweeping overhead. Conversely, when the channel between the BS and the UE is weak, narrow beams should be used to improve the reliability of SS detection so as to prevent a large number of beam sweeping rounds. For a BS, the selection of its beamwidth during IA impacts the delays of all UEs that are trying to connect to the BS. Thus, the beamwidth at the BS needs to be adjusted based on the channel gains and beamwidths of all UEs. Beamwidth optimization in mmWave cellular systems is

M. Feng is with Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, 430074 China. M. Krunz is with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721 USA. Email: mingjiefeng@hust.edu.cn, krunz@email.arizona.edu.

An abridged version of this paper was presented at IEEE ICC 2021, Online, June 2021 [1].

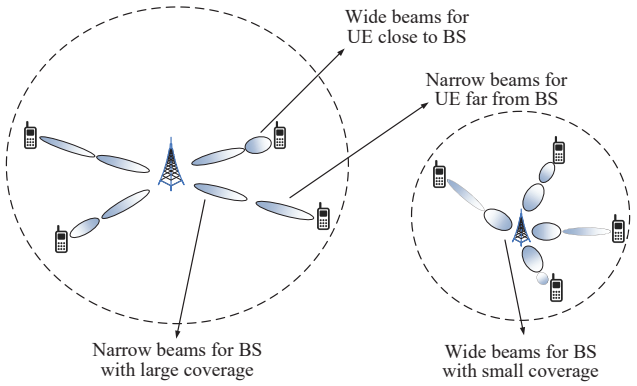

Fig. 1. Illustration of beamwidth optimization in mmWave cellular networks.

shown in Fig. 1.

Depending on the link availability between the BS and the UEs, mmWave systems can be classified into single-link (e.g., standalone 5G NR) and dual-link (e.g., non-standalone 5G NR with LTE links, future 6G ground-satellite links) systems. In single-link systems, the UE can only connect to the BS via mmWave link. In dual-link systems, the UE maintains dual connectivity with mmWave links and another link (e.g., sub-6 GHz LTE link). The sub-6 GHz link enables information exchange between the BS and the UE before beam sweeping, which contributes to reduced beam sweeping delay. Specifically, a UE can know the beamwidth at the BS it will be connected to and estimate the channel gain to that BS via the sub-6 GHz link, then optimize its beamwidth accordingly. In single-link systems, each UE does not know such information before IA. With such a limitation, each UE can only select a default beamwidth that is optimized based on networkwide BS-UE channel statistics, rather than local BS parameters, resulting in less performance gain compared to dual-link systems. Due to such differences, the solution algorithm to obtain the optimal beamwidths should be specifically designed for both types of systems. Moreover, the beam sweeping protocols (i.e., the way that various directions are scanned by the BS and the UE) that support beamwidth optimization in both types of systems need to be designed.

In this paper, we propose an online learning-based beamwidth optimization framework for IA in both single-link and dual-link mmWave cellular networks. We first derive an initial solution for beamwidth optimization based on the distribution of BS-UE distance and channel statistics. Then, we develop an online learning-based scheme for the BS to refine beamwidth selection. Specifically, we formulate beamwidth selection as a multiple-armed bandit (MAB), since selecting various beamwidths results in different expected beam sweeping delays, which is similar to playing various arms with corresponding average rewards in an MAB problem. The formulated MAB problem is solved with a Thompson sampling (TS)-based algorithm, which was shown to outperform other classical algorithms for a wide range of problems [6], and can be easily implemented with a simple calculation at each iteration, making it amenable to practical applications. The main contributions are as follows:

- We formulate the beamwidth optimization problem with the objective of minimizing the average beam sweeping delay, considering the effect of misdetection.
- We analyze how beamwidth impacts UE detection and derive the misdetection probability.
- Considering the mmWave cellular standards, we present beam sweeping protocols that support beamwidth optimization for both single-link and dual-link systems.
- We propose a two-stage framework for optimizing the BS and UE beamwidths. The first stage provides an initial solution for the BS beamwidth and the optimal solution for the UE beamwidth under a given BS beamwidth. In the second stage, the BS beamwidth is further optimized by formulating an MAB problem, which is solved by a TS-based algorithm.
- The performance of the proposed solutions is evaluated via simulations. Our results show that the proposed algorithms lower the beam sweeping delay by more than 50% compared to the benchmark schemes.

In the remainder of this paper, we first review related literature in Section II. Then, we present the system model in Section III, followed by the misdetection analysis and problem formulation in Section IV and V, respectively. Next, the beam sweeping protocols and solution algorithms are introduced in Sections VI and VII, respectively. Last, we depict the simulation results and conclude the paper in Sections VIII and IX, respectively.

II. Related Work

IA protocols for mmWave systems have been extensively studied. A few analytical frameworks for IA in mmWave cellular networks were presented in [20], [22]. The IA specified by 3GPP for mmWave cellular networks was summarized in [10]. Several methods were proposed to reduce the IA delay. Two-stage hierarchical beam sweeping was considered in [4], [5]. In [23] and [24], the sparsity of mmWave channels was utilized to reduce the search overhead. In [25], the unimodularity of the received power across the angular domain was used to reduce the beam alignment overhead. The search overhead can also be reduced with random beamforming, where instead of sequentially searching all directions, a random subset of directions are selected for beam sweeping [7], [26], or certain directions are searched based on the information of the environment and user distribution [27]. Although the effectiveness of IA delay reduction was demonstrated in these works, the solutions were not designed based on cellular mmWave standards. Thus, the performance of these solutions in a practical system is not guaranteed. The IA delay can also be reduced with the assistance of sub-6 GHz LTE link [28] or by using the location information [29], at the cost of additional signaling. However, the potential of using sub-6 GHz links for IA delay reduction with beamwidth optimization has not been harnessed. In this paper, we aim to minimize the beam sweeping delay via beamwidth optimization. Our approach can be combined with many approaches mentioned above.

Beamwidth optimization for mmWave systems was considered in existing works. Considering the impact of beamwidth on beam alignment time and antenna gain, beamwidth se-

lection was optimized to enhance data transmission performance [30], [31]. For beam optimization during IA, the tradeoff between the received SNR loss and UE discovery latency during beam sweeping has been demonstrated in [11]. However, the interplay among the beamwidth, the misdetection probability, and the beam sweeping overhead under the mmWave cellular standards has not been analyzed. Besides, the analysis and design are based on a single-path LOS channel model, which can not capture the effect of receiving signals from multiple clusters. In contrast, we analyze the impact of beamwidth on the misdetection probability and the expected beam sweeping delay based on current/future mmWave cellular standards, and design beamwidth optimization for both single-link and dual-link systems. Moreover, our analysis is based on a multipath propagation environment where the effect of receiving multiple clusters is considered.

Misdetection-aware beamwidth optimization was investigated recently (e.g., in [8], [26]). In these works, beamwidth optimization was only applied to the BS, hence the potential of adjusting UE beamwidth for delay reduction has not been harnessed. Optimizing beamwidths at both the BS and UE not only completes the problem formulation and solution, but also necessitates dedicated beam sweeping protocol designs. Besides, the analytical frameworks in these works were not based on cellular standards, thus the analysis and solution cannot be applied to our problem. In contrast, our analysis is specifically designed based on current/future standards, and we provide beamwidth optimization solutions for both BS and UE.

III. System Model

A. Network Model

We consider a cellular mmWave network with multiple BSs and UEs, whose locations are randomly distributed according to Poisson Point Processes (PPP) of densities $\rho_{\rm BS}$ and $\rho_{\rm UE}$, respectively. Let r be the distance between an arbitrary UE and its serving BS. Assuming that each UE is served by the nearest BS, the probability density function (PDF) of r is given by [9]:

$$f_r(r) = 2\pi\rho_{\rm BS} r e^{-\rho_{\rm BS}\pi r^2}, r > 0. \tag{1}$$

B. Antenna Model

We assume that the BSs and UEs are equipped with uniform linear arrays (ULAs) with $M_{\rm BS}$ and $M_{\rm UE}$ *active* antenna elements, respectively. The beamwidth at BS and UE can be changed by adjusting $M_{\rm BS}$ and $M_{\rm UE}$, respectively. Specifically, when the beam is directed toward the broadside of the antenna array, the half-power beamwidth (HPBW) can be estimated as [16]:

$$\theta_{\rm BS} \approx \frac{0.886\lambda}{d \cdot M_{\rm BS}},\ \theta_{\rm UE} \approx \frac{0.886\lambda}{d \cdot M_{\rm UE}} \tag{2}$$

where λ and d are the wavelength and antenna separation distance, respectively. For analytical tractability, we approximate the antenna patterns by a sectored antenna model, as often done in the literature [33]. Let $G_{\rm BS}(\eta')$ and $G_{\rm UE}(\eta)$ be the antenna gains of the BS and the UE when the angles off the broadside are η' and η, respectively, they are given by:

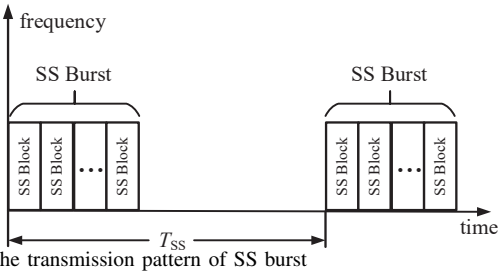

Fig. 2. The transmission pattern of SS burst

$$G_{\rm BS}(\eta') = \begin{cases} M_{\rm BS}, & \text{if } |\eta'| \le \frac{\theta_{\rm BS}}{2} \\ 0, & \text{otherwise,} \end{cases}, \eta' \in [0, 2\pi]$$
$$G_{\rm UE}(\eta) = \begin{cases} M_{\rm UE}, & \text{if } |\eta| \le \frac{\theta_{\rm UE}}{2} \\ 0, & \text{otherwise.} \end{cases}, \eta \in [0, 2\pi] \tag{3}$$

We assume that both the BS and the US adopt a two-stage beam sweeping, which has been applied in 5G NR. Let $\theta_{\rm BS}^{(1)}$ and $\theta_{\rm BS}^{(2)}$ be the BS beamwidths used in the first stage (coarse beam sweeping) and the second stage (refined beam sweeping) respectively; $\theta_{\rm BS}^{(1)}$ and $\theta_{\rm BS}^{(2)}$ be the BS beamwidths used in the first stage (coarse) and the second stage (refined) respectively. The minimum required number of beam directions to cover the entire 2-dimension angular space at the BS and the UE during the two stages are given by:

$$N_{\rm BS}^{(1)} = \left\lceil \frac{2\pi}{\theta_{\rm BS}^{(1)}} \right\rceil,\ N_{\rm UE}^{(1)} = \left\lceil \frac{2\pi}{\theta_{\rm UE}^{(1)}} \right\rceil;$$
$$N_{\rm BS}^{(2)} = \left\lceil \frac{\theta_{\rm BS}^{(1)}}{\theta_{\rm BS}^{(2)}} \right\rceil,\ N_{\rm UE}^{(2)} = \left\lceil \frac{\theta_{\rm UE}^{(1)}}{\theta_{\rm UE}^{(2)}} \right\rceil. \tag{4}$$

C. Control Signals Used in Beam Sweeping

In mmWave cellular systems, control signals are used to measure the link qualities of various Rx/Tx directions. In 5G NR, synchronization signal (SS) blocks are periodically sent by the BS and received by the UE along different directions during beam sweeping. As shown in Fig. 2, SS blocks are sent within SS bursts. An SS block consists of 4 consecutive OFDM symbols with 240 subcarriers [12]. The SS blocks carry the Primary Synchronization Signal (PSS), the Secondary Synchronization Signal (SSS), and Physical Broadcast Channel (PBCH). In particular, the PBCH is used to estimate the reference signal received power (RSRP) of the SS block. The PBCH also contains the serving beam index information, which is used by the UE when reporting the best beam to the BS. Due to the coherence of standards, such a pattern will likely continue to be applied in B5G/6G systems. In the remainder of this paper, we use SS blocks as an example of control signals to illustrate the beam sweeping process.

D. Communication Model

The beam directions at the BS and the UE are usually generated by corresponding codebooks [5], [27]. Let $i = 1, ..., N_{\rm BS}$ and $j = 1, ..., N_{\rm UE}$ be the indices of codebooks/beam directions at the BS and the UE, respectively. Then, when the BS

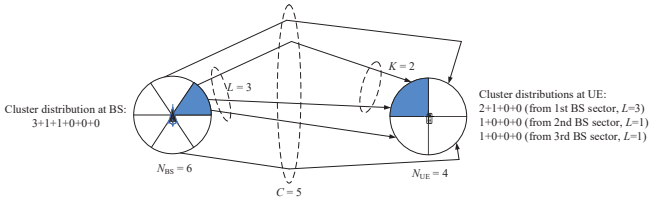

Fig. 3. Illustration of multipath transmission between the BS and the UE.

transmits with the ith beam direction and the UE receives with the jth beam direction, the received signal is given by:

$$y_{i,j} = \mathbf{q}_j^H \mathbf{H} \mathbf{f}_i s + \mathbf{q}_j^H \mathbf{z} \tag{5}$$

where s is the transmitted signal, $\mathbf{H}$ is the channel between the BS and the UE, $\mathbf{f}_i$ and $\mathbf{q}_j$ are the ith BS beamforming vector and the jth UE beamforming vector, respectively, and $\mathbf{z}$ is the noise vector. Each BS-UE beam pair achieves a certain received power $P_{i,j} = |y_{i,j}|^2$. For analytical tractability, we use the statistical channel model from NYU Tandon [18] to estimate the received power. The channel between the BS and the UE is composed of a set of distinctive clusters, each corresponding to a scattering path. The AoD/AoA of each cluster at the BS/UE is uniformly distributed in $(0, 2\pi)$. During beam sweeping, if the AoD/AoA of a cluster is not in the range of a BS/UE sector, the cluster cannot be utilized to transmit/receive SS blocks.

Let C be the total number of clusters. During a specific stage of beam sweeping, due to directional Tx/Rx, only a subset of the C clusters would be utilized by each BS sector for SS block transmission. Similarly, each UE sector would only utilize a subset of the clusters that are utilized by the BS to receive SS blocks. Let L_i^C and K_j^C be the numbers of clusters utilized by the ith BS sector and the jth UE sector when the total number of clusters is C, respectively. For notational simplicity, we denote L and K as the number of clusters at a certain pair of BS and UE sectors. Fig. 3 illustrates an example with $C = 5$, $L = 3$, and $K = 2$, where the beam sweeping is performed at the stage between the colored BS sector and the colored UE sector. The value of L determines the distribution of power fraction for each cluster, and the value of K determines Rx power gain when multiple clusters are combined. Then, the power gain between the BS and the UE is a function of both L and K. Considering the major components of channel power loss, including the TX/Rx antenna gains, path loss, and multipath interference under directional Tx/Rx, the received power of an SS block at UE is given by:

$$P_{\mathrm{UE}} = P_{\mathrm{BS}} \cdot G_{\mathrm{BS}} \cdot G_{\mathrm{UE}} \cdot \mathrm{PL}^{-1} \cdot \gamma_{\mathrm{L,K}} \tag{6}$$

where P_{BS} is the BS transmission power, PL is the distance-dependent path loss between BS and UE, which can be obtained with existing models (e.g., [18]). $\gamma_{L,K}$ is the power scaling factor when L clusters are transmitted from a BS sector and K of them are received by the UE in the same sector. Using the model from [18], the distribution of $\gamma_{L,K}$ can be numerically derived.

IV. Misdetection Probability Analysis

Recall that misdetection is defined as the incident where no SS block is successfully received by the UE during a complete round of beam sweeping. The UE fails to receive an SS block when the received power is below a threshold P_{th}. We first consider the reception of SS blocks at a fixed UE sector. Let $\mathcal{P}_{L,K}^{\mathrm{Sec}}$ be the probability that a UE fails to detect any SS block in that sector with given L and K. Following (6), $\mathcal{P}_{L,K}^{\mathrm{Sec}}$ is given by:

$$\mathcal{P}_{L,K}^{\mathrm{Sec}} = \Pr\left\{\gamma_{L,K} \leq \frac{P_{\mathrm{th}} \cdot \mathrm{PL}}{P_{\mathrm{BS}} \cdot G_{\mathrm{BS}} \cdot G_{\mathrm{UE}}}\right\}. \tag{7}$$

Using the numerically derived PDF of $\gamma_{L,K}$ and empirical path loss model (e.g.,from [18]), $\mathcal{P}_{L,K}^{\mathrm{Sec}}$ can be calculated as a function of M_{BS} and M_{UE}, and thus a function of N_{BS} and N_{UE}.

To calculate the misdetection probability of a complete round of beam sweeping, the misdetection probabilities of all pairs of BS and UE sectors need to be integrated. The misdetection probability of each pair under given L and K is calculated with (7), we then calculate the probabilities of various cluster distributions at the BS and UE (i.e., how the clusters are distributed among different BS and UE sectors), which are determined by the values of L_i^C and K_j^C for various BS and UE sectors. We index the possible distributions at BS and UE sectors by $u = 1, \ldots, U$ and $v = 1, \ldots, V$, respectively. Then, we define two sets of binary variables, δ_u and π_v, as indicators of these distributions. Specifically, $\delta_u = 1$ indicates that the uth BS cluster distribution occurs and $\delta_u = 0$ indicates otherwise; $\pi_v = 1$ indicates that the vth UE cluster distribution occurs and $\pi_v = 0$ indicates otherwise. With C clusters in total, we denote the uth BS sector distribution by $[L_1^C(u), ..., L_{N_{\mathrm{BS}}}^C(u)]$, where $\sum_{i=1}^{N_{\mathrm{BS}}} L_i^C(u) = C$ holds for $u = 1, ..., U$. With L clusters used by the BS, we denote the uth UE sector distribution by $[K_1^L(v), ..., K_{N_{\mathrm{UE}}}^L(v)]$, then $\sum_{j=1}^{N_{\mathrm{UE}}} K_j^L(v) = L$ holds for $v = 1, ..., V$. The probabilities of various distributions (i.e., $u = 1, ..., U$ and $v = 1, ..., V$) are calculated with a combination formula in probability theory. Due to the page limit, we only present the cluster distribution at the BS when $C = 4$ as an example in Table I.

Next, we consider the misdetection probability when SS blocks are sent via a fixed BS sector and are received by *all* UE sectors. This probability is calculated by multiplying the misdetection probabilities of all UE sectors. With multiple possible cluster distributions at the UE ($v = 1, ..., V$), the average misdetection probability is a weighted sum of misdetection probabilities under different distributions, given by:

$$\mathcal{P}_L^{\mathrm{UE}} = \sum_{v=1}^{V} \left\{ \Pr\{\pi_v = 1\} \cdot \prod_{j=1}^{N_{\mathrm{UE}}} \mathcal{P}_{L,K_j^L(v)}^{\mathrm{Sec}} \right\}. \tag{8}$$

Note that no SS block will be received for the UE sectors with no path/cluster in their ranges. Thus, the misdetection probabilities for these sectors are 1, which do not impact the product given in (8). This also applies to the BS sectors. As a result, we only need to consider the sectors with paths/clusters in their ranges in our calculation.

We then consider the scenario that SS blocks are sequentially sent from *all* BS sectors, which is a complete round of beam sweeping for all $N_{\mathrm{BS}} N_{\mathrm{UE}}$ sectors. The misdetection probability under this scenario is the product of misdetection

TABLE I
EXAMPLE OF DISTRIBUTION OF CLUSTERS AT BS SECTORS WITH DIFFERENT N_{BS} ($C = 4$)

$N_{\mathrm{BS}} \geq 4$		$N_{\mathrm{BS}} = 3$		$N_{\mathrm{BS}} = 2$	
Distribution	Probability	Distribution	Probability	Distribution	Probability
$[1,1,1,1,0,...]$	$\frac{\binom{N_{\mathrm{BS}}}{4}\cdot 4!}{N_{\mathrm{BS}}^4}$	$[2,1,1]$	$\frac{\binom{3}{1}\cdot\binom{4}{2}\cdot 2!}{3^4} = \frac{4}{9}$	$[2,2]$	$\frac{\binom{4}{2}}{2^4} = \frac{3}{8}$
$[2,1,1,0,0,...]$	$\frac{\binom{N_{\mathrm{BS}}}{4}\cdot\binom{4}{2}\cdot 3!}{N_{\mathrm{BS}}^4}$	$[2,2,0]$	$\frac{\binom{3}{2}\cdot\binom{4}{2}}{3^4} = \frac{2}{9}$	$[3,1]$	$\frac{\binom{4}{1}\cdot 2!}{2^4} = \frac{1}{2}$
$[2,2,0,0,0,...]$	$\frac{\binom{N_{\mathrm{BS}}}{2}\cdot\binom{4}{2}}{N_{\mathrm{BS}}^4}$	$[3,1,0]$	$\frac{\binom{4}{2}\cdot\binom{4}{1}}{3^4} = \frac{8}{27}$	$[4,0]$	$\frac{\binom{2}{1}}{2^4} = \frac{1}{8}$
$[3,1,0,0,0,...]$	$\frac{\binom{N_{\mathrm{BS}}}{2}\cdot\binom{4}{1}\cdot 4!}{N_{\mathrm{BS}}^4}$	$[4,0,0]$	$\frac{\binom{3}{1}}{3^4} = \frac{1}{27}$		
$[4,0,0,0,0,...]$	$\frac{\binom{N_{\mathrm{BS}}}{1}}{N_{\mathrm{BS}}^4}$				

probabilities when the SS blocks are sent from all BS sectors. With different cluster distributions at the BS ($u = 1, ..., U$), and for a given C, the average misdetection probability is a weighted sum of misdetection probabilities under different cluster distributions at BS, given by:

$$\mathcal{P}_{\mathrm{mis}}^{C} = \sum_{u=1}^{U} \left\{ \Pr\{\delta_u = 1\} \cdot \prod_{i=1}^{N_{\mathrm{BS}}} \mathcal{P}_{L_i^C(u)}^{\mathrm{UE}} \right\}. \tag{9}$$

Based on [18], C is a random variable that follows: $C \sim \max\{\mathrm{Poisson}(\kappa), 1\}$, where $\kappa = 1.8$ at 28 GHz and $\kappa = 1.9$ at 73 GHz.

Finally, the average misdetection probability for a complete round of beam sweeping is:

$$\mathcal{P}_{\mathrm{mis}} = \sum_{C'=1}^{\infty} \left\{ \Pr\{C = C'\} \cdot \mathcal{P}_{\mathrm{mis}}^{C'} \right\}. \tag{10}$$

V. PROBLEM FORMULATION

As mentioned in Section III-C, periodical transmission of control signals has been applied for beam sweeping in 5G NR, and will likely be applied in B5G/6G systems as well. Thus, we formulate and solve the beamwidth optimization problem under the 5G NR framework. Our solutions can be extended to future mmWave cellular systems.

In the 5G NR framework, each round of beam sweeping begins with the transmission of the first SS block in an SS burst. The number of SS blocks required to complete one round of beam sweeping is determined by the total number of beam pairs, given by $N_{\mathrm{BS}}^{(1)} N_{\mathrm{UE}}^{(1)}$ for the first stage (coarse sweeping) and $N_{\mathrm{BS}}^{(2)} N_{\mathrm{UE}}^{(2)}$ for the second stage (refined sweeping), respectively[1] During beam sweeping, misdetection can occur when none of the SS blocks are successfully received by a UE, due to insufficient received power. Thus, the misdetection probability is dictated by the channel and antenna gains at the BS and the UEs. Let $\mathcal{P}_{\mathrm{mis}}^{(1)}$ and $\mathcal{P}_{\mathrm{mis}}^{(2)}$ denote the misdetection probability for one round of beam sweeping during the first and second stage, respectively. When misdetection occurs, the UE needs to wait for the next round of beam sweeping, until the successful reception of an SS block. Then, the probability that n rounds of first stage beam sweeping are required follows a geometric distribution, given by $(1 - \mathcal{P}_{\mathrm{mis}}^{(1)}){\mathcal{P}_{\mathrm{mis}}^{(1)}}^{n-1}$. The same result applies to the second stage of beam sweeping.

[1]Note that, hybrid beamforming can be applied to improve to the beam sweeping performance by allowing the BS and/or the UE to scan multiple directions simultaneously. However, regardless of the beamforming scheme, the tradeoff in beamwidth selection always exists, hence optimizing beamwidth can still reduce beam sweeping delay when hybrid beamforming is used. Thus, applying hybrid beamforming is compatible with beamwidth optimization, and the proposed solution can be applied to hybrid beamforming systems with minor modifications. For simplicity, we assume an analog beamforming architecture in which the BS and/or the UE scan various directions sequentially.

Denote N_{SS} as the maximum number of SS blocks in each SS burst, which is set to 64 for mmWave systems [13]. Let n and m be the numbers of rounds in the first and second stages of beam sweeping, respectively. The total number of beam sweeping directions is given by:

$$N_{\mathrm{tot}} = n N_{\mathrm{BS}}^{(1)} N_{\mathrm{UE}}^{(1)} + m N_{\mathrm{BS}}^{(2)} N_{\mathrm{UE}}^{(2)}. \tag{11}$$

Among all SS bursts used for beam sweeping, $\left\lfloor \frac{N_{\mathrm{tot}}}{N_{\mathrm{SS}}} \right\rfloor$ of them are completely used (i.e., all SS blocks in the SS burst are used for beam sweeping) and the remaining one is partially used (i.e., only part of the SS blocks in the SS burst are used). The number of SS bursts that are completely used is $\left\lfloor \frac{N_{\mathrm{tot}}}{N_{\mathrm{SS}}} \right\rfloor$. In the last SS burst, the number of utilized SS blocks is $N_{\mathrm{last}} = N_{\mathrm{tot}} - N_{\mathrm{SS}} \left\lfloor \frac{N_{\mathrm{tot}}}{N_{\mathrm{SS}}} \right\rfloor$. In 5G NR, two SS blocks are transmitted in each slot. Hence, the number of slots needed in the last SS burst is $\frac{N_{\mathrm{last}}}{2}$. Given the period of SS bursts T_{SS} and the time slot duration T_{slot}, the expected beam sweeping delay of a UE is given by:

$$\overline{D} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (1 - \mathcal{P}_{\mathrm{mis}}^{(1)}){\mathcal{P}_{\mathrm{mis}}^{(1)}}^{n-1} (1 - \mathcal{P}_{\mathrm{mis}}^{(2)}){\mathcal{P}_{\mathrm{mis}}^{(2)}}^{m-1} \left[T_{\mathrm{SS}} \left\lfloor \frac{N_{\mathrm{tot}}}{N_{\mathrm{SS}}} \right\rfloor + \frac{T_{\mathrm{slot}}}{2} \left(N_{\mathrm{tot}} - N_{\mathrm{SS}} \left\lfloor \frac{N_{\mathrm{tot}}}{N_{\mathrm{SS}}} \right\rfloor \right) \right]. \tag{12}$$

Given the distribution of r, the expected delay of UEs in the coverage of a BS is given by:

$$\mathbb{E}[D] = \int_0^{\infty} \overline{D}(r) f_r(r) \mathrm{d}r \tag{13}$$

where $f_r(r)$ is the PDF of r given in (1). Then, the delay

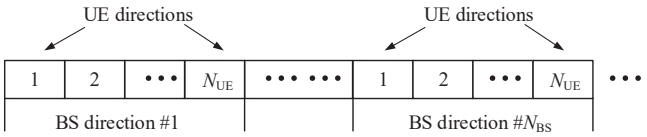

Fig. 4. Beam sweeping in single-link mmWave cellular.

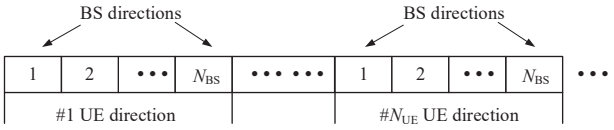

Fig. 5. Beam sweeping in dual-link mmWave cellular.

minimization problem is given as:

$$\min_{\left\{N_{\mathrm{BS}}^{(1)},N_{\mathrm{BS}}^{(2)},N_{\mathrm{UE}}^{(1)},N_{\mathrm{UE}}^{(2)}\right\}} \mathbb{E}[D]$$
$$\text{subject to: } N_{\mathrm{BS}}^{(1)}, N_{\mathrm{BS}}^{(2)} \in \Omega_{\mathrm{BS}};$$
$$N_{\mathrm{UE}}^{(1)}, N_{\mathrm{UE}}^{(2)} \in \Omega_{\mathrm{UE}} \tag{14}$$

where Ω_{BS} and Ω_{UE} are the sets of feasible values of BS and UE beamwidths, respectively.

VI. Beam Sweeping Protocols for Beamwidth Optimization

A. Single-link Systems

In a single-link mmWave cellular network, the UEs do not know N_{BS} before receiving a complete round of SS blocks for all BS directions. Thus, the beam sweeping protocol can only be designed as the one shown in Fig. 4, where each UE needs to sequentially scan each SS block along N_{UE} directional Rx beams. Since there is no coordination between the BS and the UE before IA, N_{UE} has to be set to a default value. This way, the BS knows how many SS blocks should be sent along each direction, in order for the UE to complete Rx beam sweeping.

During beam sweeping, the UE measures the link qualities of all possible BS-UE beam pairs. After a complete round of beam measurements, the UE determines the pair with the maximum received power. Then, the UE waits for the BS to schedule the RACH opportunity towards the best BS direction that it just determined, and performs random access by sending a RACH preamble to the BS with the RACH resources specified by the BS. This way, the UE implicitly notifies the BS about the best discovered BS beam, which will be used in the subsequent transmissions [15].

B. Dual-link Systems

In a dual-link mmWave cellular network, the UE can use a sub-6 GHz connection to learn the value of N_{BS} before beam sweeping takes place. As a result, the beam sweeping protocol in Fig. 4 can be replaced by the one shown in Fig. 5. Specifically, the BS sends the SS blocks in a round-robin fashion to cover N_{BS} directions in the angular domain, and the UE alters its Rx beam only after the BS completes one round of SS blocks transmission along all N_{BS} directions. This allows a more flexible system configuration compared to the one shown in Fig. 4, since each UE can dynamically select N_{UE} to optimize its performance (see Section VII-B).

VII. Solution Algorithms

As indicated in Section III-B, optimizing the beamwidth is equivalent to optimizing the number of active antennas. For notational simplicity, we denote $\mathbf{N}_{\mathrm{BS}} \triangleq [N_{\mathrm{BS}}^{(1)}, N_{\mathrm{BS}}^{(2)}]$ and $\mathbf{N}_{\mathrm{UE}} \triangleq [N_{\mathrm{UE}}^{(1)}, N_{\mathrm{UE}}^{(2)}]$. The solution framework has two stages. At the first stage, the optimal $[\mathbf{N}_{\mathrm{BS}}, \mathbf{N}_{\mathrm{UE}}]$, denoted by $[\mathbf{N}_{\mathrm{BS}}^*, \mathbf{N}_{\mathrm{UE}}^*]$, are obtained based on the BS density ρ_{BS}. In the second stage, each BS further optimizes $\mathbf{N}_{\mathrm{BS}}$ by solving an MAB problem with a TS-based algorithm.

A. Single-link Systems

1) First Stage: In a single-link system, each UE has no information about the BS that it will be connecting to until IA is completed. Thus, $\mathbf{N}_{\mathrm{UE}}$ should be set to a default value that is optimized in advance and is known by all BSs owned by the same operator. As a UE may roam in areas with varying BS densities, $\mathbf{N}_{\mathrm{UE}}$ is optimized based on the average BS density (which represents the average BS-UE distance), i.e., by setting $\rho_{\mathrm{BS}} = \overline{\rho}_{\mathrm{BS}}$. Once N_{UE}^* is obtained, all UEs set the beamwidth to N_{UE}^*. Then, $\mathbf{N}_{\mathrm{UE}}^*$ is obtained by:

$$\left[\mathbf{N}_{\mathrm{BS}}^*, \mathbf{N}_{\mathrm{UE}}^*\right] = \arg \min_{\{\mathbf{N}_{\mathrm{BS}}, \mathbf{N}_{\mathrm{UE}}\}} \mathbb{E}[D] \tag{15}$$

The search in (15) is performed offline by the operator.

Based on $\mathbf{N}_{\mathrm{UE}}^*$ and the knowledge of ρ_{BS} (which implicitly determines the distribution of BS-UE distance), each BS calculates $\mathbb{E}[D]$ with (13) and optimizes its beamwidth by:

$$\mathbf{N}_{\mathrm{BS}}^* = \arg \min_{\mathbf{N}_{\mathrm{BS}}} \mathbb{E}[D]. \tag{16}$$

The optimization of (16) is performed offline by each BS.

2) Second Stage: We formulate the beamwidth optimization of each BS as an MAB problem and solve it with a TS algorithm. In the formulated MAB problem, each BS acts as an agent who plays one arm at each time step and learns to find the arm with minimum average penalty.[2] The arms to be played are the possible selections of BS beamwidth, denoted by $\mathbf{N}_{\mathrm{BS}}^i$, $i = 1, \ldots, \Phi_{\mathrm{BS}}$, where Φ_{BS} is the number of possible $\mathbf{N}_{\mathrm{BS}}^i$. Let d_i be the mean penalty for playing arm i, it is defined as the *normalized* average delay when the BS beamwidth is set to $\mathbf{N}_{\mathrm{BS}}^i$, given by:

$$d_i = \min\left(\frac{\overline{D}(\mathbf{N}_{\mathrm{BS}}^i)}{D_{\mathrm{ref}}}, 1\right) \tag{17}$$

where $\overline{D}(\mathbf{N}_{\mathrm{BS}}^i)$ is the average delay of all UEs when $\mathbf{N}_{\mathrm{BS}} = \mathbf{N}_{\mathrm{BS}}^i$, D_{ref} is a sufficiently large reference delay predetermined by the system. The normalization given by (17) is necessary for applying the TS algorithm given in [36]. At each iteration of the TS algorithm, the prior distribution of the played arm is updated according to the outcome of a

[2]While the objective of most MAB problems is maximizing the sum reward, we consider an MAB problem that aims to minimize the sum penalty for simplicity, since we aim to find the beamwidth with the minimum expected delay.

Bernoulli trial; the success probability of the Bernoulli trial is positively related to the observed penalty for playing the selected arm. Thus, to generate a valid success probability (in [0,1]), the normalized observed penalty is used in the proposed TS algorithm.

As the channel varies over time, the BS collects the average delay of all UEs over a relatively long period of time (e.g., multiple rounds of SS bursts) for learning. This way, the beamwidth selection is performed on a larger time scale than that of beam sweeping, and the effect of channel variation is averaged out. To obtain the average delay, each UE has a timer that records its beam sweeping delay and reports it to the BS when IA is completed. At the end of each time step, the BS calculates the average delay during that period.

The main objective of the TS algorithm is to find the arm with the minimum d_i. The BS keeps a belief about the distribution of d_i, $i = 1, \ldots, \Phi_{\mathrm{BS}}$. At each step of TS, the BS samples all arms according to such belief. Then, it plays the arm with the minimum sampled value and observes the penalty for playing that arm. After that, it updates the belief of the played arm based on the observed penalty. Such an update is implemented with Bayesian inference, which calculates the posterior distribution based on the observed data and prior distribution as follows:

$$\Pr(\boldsymbol{\theta}|\mathbf{x}) = \frac{\Pr(\mathbf{x}|\boldsymbol{\theta}) \Pr(\boldsymbol{\theta})}{\Pr(\mathbf{x})} \tag{18}$$

where $\Pr(\mathbf{x}|\boldsymbol{\theta})$ is the distribution of the observed data, $\Pr(\boldsymbol{\theta})$ is the prior distribution before any observation, and $\Pr(\mathbf{x})$ is the marginal distribution of the evidence. Based on the TS algorithm described in [36] and given that $0 \leq d_i \leq 1$, Beta distribution is the conjugate prior for the distribution of d_i. This is because the posterior of a Beta distribution is also a Beta distribution, which makes the parameter update at each round of the TS algorithm easy to implement. Let $\hat{d}_i^{[t]}$ be the prior distribution of d_i at time t, we have $\hat{d}_i^{[t]} \sim \mathrm{Beta}(\alpha_i^{[t]}, \beta_i^{[t]})$, $i = 1, \ldots, \Phi_{\mathrm{BS}}$, where $\alpha_i^{[t]}$ and $\beta_i^{[t]}$ are the parameters of the Beta distribution. The PDF of a Beta distribution is $f(x; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, with mean value $\frac{\alpha}{\alpha+\beta}$. Beta distribution has been used to model the reward of MAB problems with Bernoulli trials, where α and β are the numbers of accumulated successes (reward $= 1$) and failures (reward $= 0$), respectively. Recently, Beta distribution has also been applied to MAB problems with reward in $[0, 1]$ by adjusting the TS algorithm [36].

For TS problems without prior information, the initial distribution of each arm is set to be $\mathrm{Beta}(1, 1)$, which corresponds to the uniform distribution on $[0, 1]$. In our problem, the result in the first stage is used to generate the initial distribution of each d_i, which can be used to accelerate the convergence of the TS algorithm. Without such initialization, the BS may spend a significant amount of time on the arms with a large delay. To obtain the initial distribution of d_i, we first calculate $\mathbb{E}[D(\mathbf{N}_{\mathrm{BS}}^i)]$ by (13) based on $\mathbf{N}_{\mathrm{BS}}^i$, ρ_{BS}, and $\mathbf{N}_{\mathrm{UE}}^*$. Then, we find a set of $(\alpha_i^{[1]}, \beta_i^{[1]})$ to approximate the normalized $\mathbb{E}[D(\mathbf{N}_{\mathrm{BS}}^i)]$ as follow:

$$\frac{\alpha_i^{[1]}}{\alpha_i^{[1]} + \beta_i^{[1]}} \approx \min\left(\frac{\mathbb{E}[D(\mathbf{N}_{\mathrm{BS}}^i)]}{D_{\mathrm{ref}}}, 1\right). \tag{19}$$

For example, if $\frac{\mathbb{E}[D(\mathbf{N}_{\mathrm{BS}}^i)]}{D_{\mathrm{ref}}} = 0.78$ (keep two numbers after decimal), we set $\alpha_i^{[1]} = 78, \beta_i^{[1]} = 22$.

With the initial distributions of d_i ($i = 1, \ldots, \Phi_{\mathrm{BS}}$), each BS performs the TS algorithm to find the arm with the minimum sampled value e_i. At each time t, the BS samples the arms according to their PDFs, plays the arm with minimum sampled value, observes the penalty, and updates the parameters of the selected arm. The observed penalty is the *normalized* average beam sweeping delay of all UEs, calculated by:

$$\Theta^{[t]} = \min\left(\frac{\overline{D}(\mathbf{N}_{\mathrm{BS}}^{i^*})}{D_{\mathrm{ref}}}, 1\right). \tag{20}$$

In (20), we set a timeout for learning at each iteration. Specifically, if a beam sweeping process is not successful within a period of D_{ref}, the observed penalty will be set to be 1 and the system will proceed to the next round of algorithm. This way, the system would not wait to observe a beam sweeping process with a delay larger than D_{ref}, as the corresponding beamwidth is highly unlikely to be optimal. The procedure of the TS-based algorithm is summarized in Algorithm 1. Note that, the duration of each time period in Algorithm 1 should be carefully selected. On the one hand, each period should be long enough to observe the delay under various beamwidths. On the other hand, faster convergence can be achieved within a shorter period.

Algorithm 1: TS-based BS Beamwidth Optimization

```
Initialize: Set (α_i^[1], β_i^[1]) (i = 1, ..., Φ_BS) according to (19) ;
d̂_i^[1] ~ Beta(α_i^[1], β_i^[1]), i = 1, ..., |Ω_BS| ;
for t = 1 : T do
    for i = 1 : |Ω_BS| do
        Sample d̂_i^[t] from Beta(α_i^[t], β_i^[t]) with outcome e_i^[t] ;
    end
    Play arm i*[t] = arg min_i e_i^[t] and observe penalty Θ^[t] ;
    Perform a Bernoulli trial with success probability Θ^[t]
      and record outcome Θ̃^[t] ;
    if Θ̃^[t] = 1 then
        α_{i*}^[t] = α_{i*}^[t] + 1 ;
    else
        β_{i*}^[t] = β_{i*}^[t] + 1 ;
    end
end
```

B. Dual-link Systems

1) First Stage: In a dual-link system, a UE can obtain the values of $\mathbf{N}_{\mathrm{BS}}$ of the serving BS and estimate the PL of the sub-6 GHz link before IA takes place. The PL of the sub-6 GHz link can be used to estimate the distance r, which is then used to estimate the PL of the mmWave link $\mathrm{PL}(r)$ [34], [35]. Then, each UE can select its optimal $\mathbf{N}_{\mathrm{UE}}$ by:

$$\mathbf{N}_{\mathrm{UE}}^* = \arg \min_{\{\mathbf{N}_{\mathrm{UE}}\}} \overline{D} \tag{21}$$

where $\overline{D}$ is calculated by (12) with given $\mathbf{N}_{\mathrm{BS}}$ and $\mathrm{PL}(r)$. The optimization is performed by each UE in an online manner.

For each BS, it optimizes $\mathbf{N}_{\mathrm{BS}}$ based on the knowledge that each UE selects its own optimal $\mathbf{N}_{\mathrm{UE}}$. With ρ_{BS} as input, the optimal BS beamwidth is determined by:

$$\mathbf{N}_{\mathrm{BS}}^{*} = \arg \min_{\{\mathbf{N}_{\mathrm{BS}}\}} \mathbb{E}[D] \tag{22}$$

where $\mathbb{E}[D]$ is calculated by (13). In particular, $\mathbf{N}_{\mathrm{UE}}$ is set to be $\mathbf{N}_{\mathrm{UE}}^{*}$ which is a function of r and $\mathbf{N}_{\mathrm{UE}}$. The optimization given in (21) is performed by each BS in an offline manner.

2) Second Stage: Same as the single-link systems, we formulate the beamwidth optimization in the second stage as an MAB problem and apply Algorithm 1 to solve it. In particular, the initial distributions of d_i $(i = 1, \ldots, \Phi_{\mathrm{BS}})$ are derived by (21) and (22) from the first stage.

C. Regret Analysis

To quantify the performance of the proposed TS-based solution, we derive an upper bound for the expected total regret in this part. Such regret is defined as the accumulated additional penalty for not playing the optimal arm over $t = 1, \ldots, T$, which is given by:

$$\mathbb{E}[\mathcal{R}(T)] = \mathbb{E}\left[\sum\nolimits_{t=1}^{T}\left(d_i^{[t]} - d_{i^*}\right)\right] = \sum_{i} \Delta_i \mathbb{E}\left[\kappa_i(T)\right] \tag{23}$$

where $i^{[t]}$ is the arm played at time t and i^* is the optimal arm (i.e., $d_{i^*} = \arg\min_i d_i$); $\Delta_i = d_{i^{[t]}} - d_{i^*}$, and $\kappa_i(T)$ is the number of times arm i is played.

Theorem 1. *The expected regret of the proposed TS-based solution is upper bounded by:*

$$\mathbb{E}[\mathcal{R}(T)] \leq \mathcal{O}\left(\left(\sum_{i \neq i^*} \frac{1}{\Delta_i^2}\right)^2 \ln T\right). \tag{24}$$

The proof of Theorem 1 is presented in the appendix.

VIII. Performance Evaluation

In this section, we first show the impact of beamwidth on beam sweeping delay via numerical results. Then, we evaluate the performance of the proposed schemes with simulations. The NYU channel model in [18] is implemented with the operating frequency set to 28 GHz. The beam sweeping parameters are based on 5G NR. Between the two bandwidth options for an SS block in 5G standards, we select numerology 3, which corresponds to a 28.8 MHz bandwidth. N_{SS} is set to 64 and T_{SS} is set to 20 ms. The BS transmit power is set to 30 dBm, and the minimum SNR required for signal detection is set to 0 dB[3] Unless otherwise specified, the number of UE antennas M_{UE} is set to 4 and the BS density ρ_{BS} is set to 10 $\mathrm{BS/km}^2$.

[3]Obviously, the minimum required SNR for detection impacts the beam sweeping. When the detection threshold is high, the optimal beamwidths would be small, since narrow beams have to be used to reach a high SNR, resulting in a high beam sweeping delay. When the detection threshold is low, the optimal beamwidths would be large, since wide beams can be used to allow a relatively weak received signal, resulting in low beam sweeping delay.

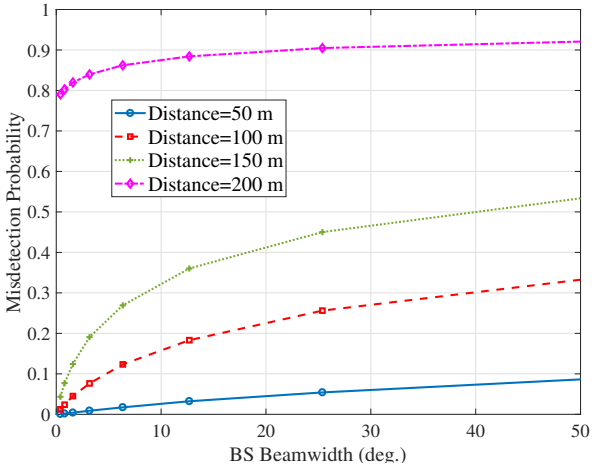

Fig. 6. Misdetection probability vs. BS beamwidth.

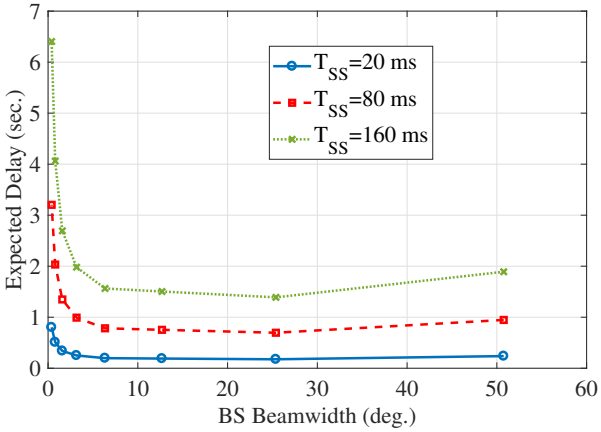

Fig. 7. Average delay vs. BS beamwidth.

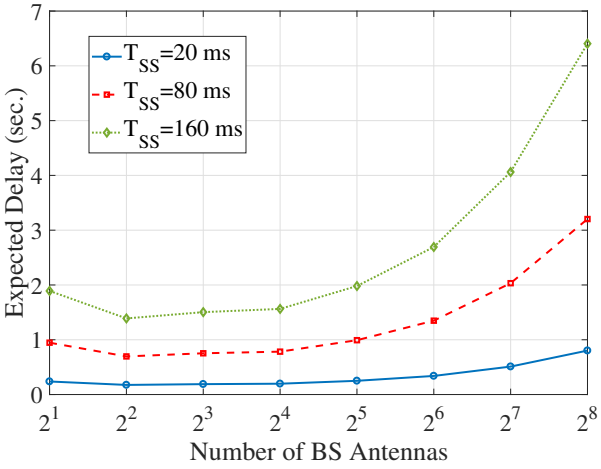

Fig. 8. Average delay vs. the number of BS antennas.

We first present the numerical results to show the impacts of beamwidth on misdetection probability and beam sweeping delay. Fig. 6 shows the misdetection probabilities under different BS beamwidths when a UE is located at different distances to the BS. The results indicate that when a UE is relatively far from the BS, narrow beams need to be employed by the BS to compensate for the propagation loss and avoid excessively high misdetection probability. On the other hand,

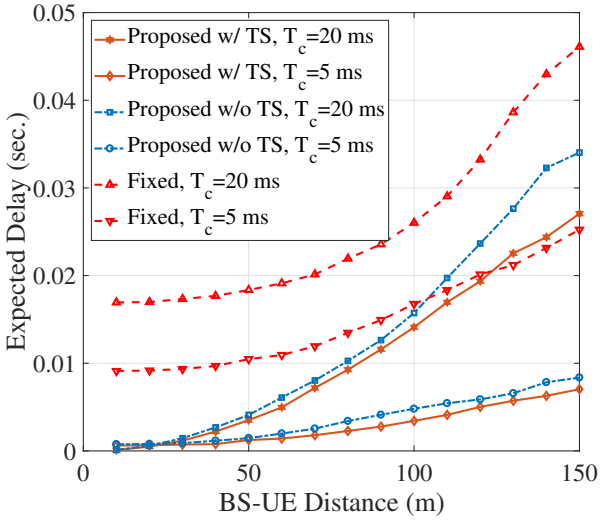

Fig. 9. Average delay vs. BS-UE distance in single-link systems.

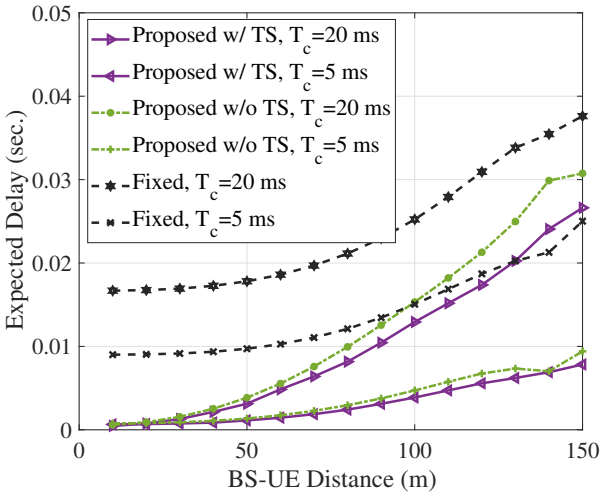

Fig. 10. Average delay vs. BS-UE distance in dual-link systems.

when the UE is close to the BS, wide beams can be used due to the low risk of misdetection. Fig. 7 depicts the average delay under different BS beamwidths. We observe that the delay first decreases as the beam sweeping overhead reduces, and then increases as the effect of misdetection becomes dominant, causing more rounds of beam sweeping. The impact of BS antenna number on the average delay is shown in Fig. 8 where a similar trend can be observed. Figs. 7 and 8 show that there exists a unique value for the BS beamwidth that minimizes the expected beam sweeping delay. Note that the BS beamwidth and UE beamwidth are equivalent in determining the delay performance. Therefore, the same results would be observed if we alter the UE beamwidth under a fixed BS beamwidth.

We then present the simulation results. To show the effectiveness of solutions in the two stages, we consider the scheme that only applies first-stage optimization (termed proposed w/o TS) and the scheme that applies optimization of both stages (termed proposed w/ TS). We compare the delay performance of the proposed schemes with the classical fixed beamwidth scheme (termed fixed) that sets N_{tot} to be equal to N_{SS}. In the simulations, we take the coherence time of a wireless channel into account and generate a new channel instance for every

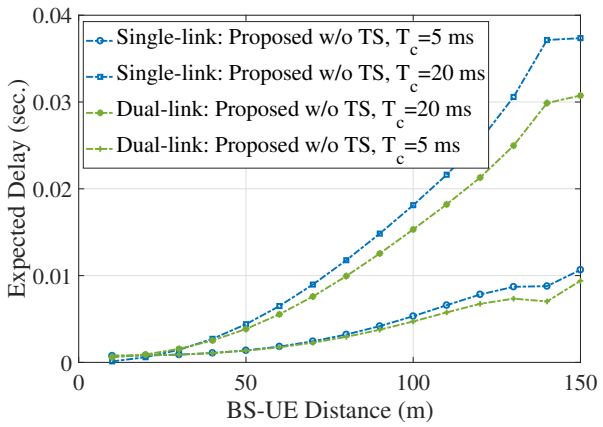

Fig. 11. Comparison between single-link and dual-link systems with the proposed algorithms.

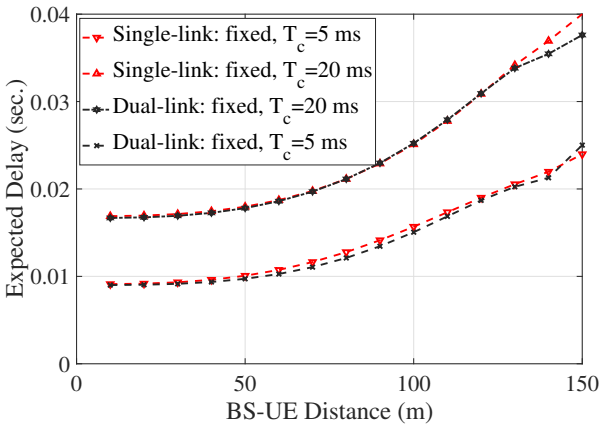

Fig. 12. Comparison between single-link and dual-link systems with fixed beamwidth setting.

T_c seconds. When a new channel is generated, it may be in three different states: LOS, NLOS, and outage, following the probabilities given in [18]. The delay performances of different schemes in single-link and dual-link systems are presented in Figs. 9 and 10, respectively. We can see that the proposed schemes outperform the fixed beamwidth scheme with much lower delays, showing the effectiveness of beamwidth optimization in the first stage. With the beamwidth adaptation in the second stage, the delay can be further lowered by more than 15% on average. In particular, the performance gain is larger when the distance is small, since the proposed schemes select the wider beams when the channel condition is good, while the fixed beamwidth scheme sticks to a relatively narrow beam. As the distance increases, the performance gap gets smaller, since the optimal beamwidths at the BS and UE are close to the ones in the fixed beamwidth scheme.

Figs. 11 and 12 show the comparisons between single-link and dual-link systems for the proposed schemes and the fixed beamwidth schemes, respectively. The delay performances of single-link and dual-link systems are close to each other when the fixed beamwidth scheme is applied. In contrast, under the proposed schemes, the dual-link systems achieve considerable delay reduction over the single-link systems, showing that the proposed schemes can exploit the sub-6 GHz link for delay reduction.

Figs. 13 and 14 present the average delay of single-link

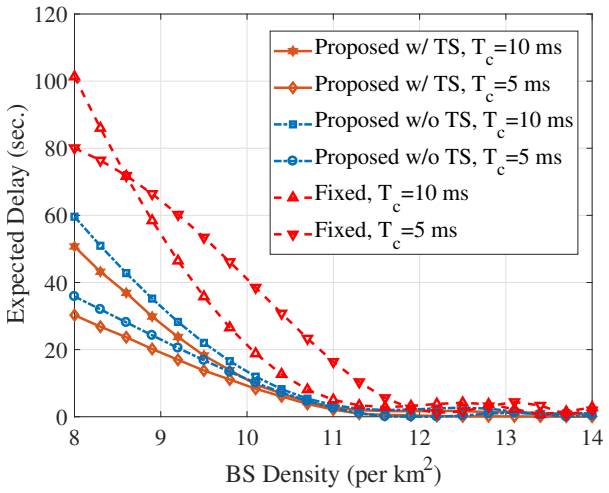

Fig. 13. Average delay versus BS density $\rho_{\rm BS}$ in single-link systems.

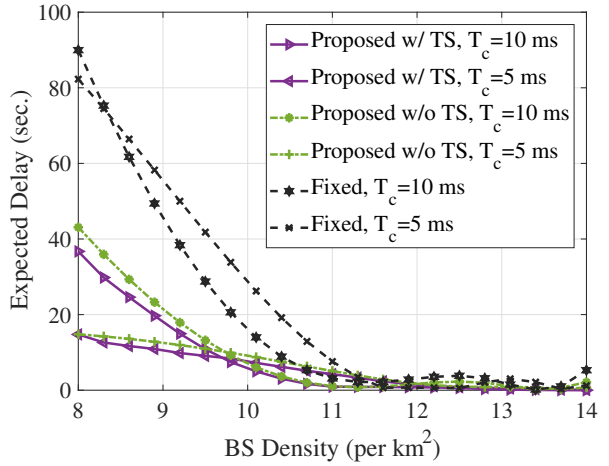

Fig. 14. Average delay versus BS density $\rho_{\rm BS}$ in dual-link systems.

and dual-link systems under varying BS densities $\rho_{\rm BS}$, respectively. The proposed schemes achieve much lower average delays, especially when $\rho_{\rm BS}$ is at a moderate range. This is because the BSs can select their optimal beamwidths according to $\rho_{\rm BS}$. When $\rho_{\rm BS}$ is high, the average BS-UE distance is small, then the proposed schemes would select wide beams at the BS, resulting in lower beam sweeping delays. When $\rho_{\rm BS}$ is low, the proposed schemes would select narrow beams to lower the misdetection probability.

IX. Conclusions

In this paper, we investigated the problem of beamwidth optimization in mmWave cellular networks, with the objective of minimizing the beam sweeping delay during IA. We first formulated the beamwidth optimization problem. Then, we proposed solution algorithms to obtain the optimal beamwidths at BS and UE. Simulation results showed that the proposed schemes achieve more than 50% lower delay on average compared to the benchmark schemes.

Appendix

In this part, we present the proof for Theorem 1.

Proof. At each t, we divide the suboptimal arms ($i \neq i^*$) into two sets, the *saturated* arms and the *unsaturated* arms, and calculate the regret bounds for playing both kinds of arms. An arm i is in the saturated set $\mathcal{C}^{[t]}$ at time t if it has been played for at least $\mathcal{L}_i = \frac{24 \ln T}{\Delta_i}$ times up to t.

We first analyze the regret due to playing the saturated arms. The key is to bound the number of times that saturated arms are played between two consecutive plays of the optimal arm. We define an *interval* as a set of continuous time steps. Let I_j be the interval between (and excluding) the j th and $(j+1)$ th play of the optimal arm, we define event $\mathcal{M}^{[t]}$ at time t as:

$$\mathcal{M}^{[t]} : e_{i^*}^{[t]} < d_i - \frac{\Delta_i}{2}, \forall i \in \mathcal{M}^{[t]}. \tag{25}$$

Let π_j be the number of times event $\mathcal{M}^{[t]}$ occur during interval I_j. Events $\mathcal{M}^{[t]}$ ($t \in I_j$) divide the interval I_j into multiple sub-intervals, given by $I_j(l)(l = 1, \ldots, \pi_j + 1)$. Specifically, $I_j(l)$ denotes the sub-interval between the $(l-1)$ th and l th occurrences of $\mathcal{M}^{[t]}$ in I_j. In particular, $I_j(1)$ denotes the sub-interval before the first occurrence of $\mathcal{M}^{[t]}$ in I_j and $I_j(\pi_j+1)$ denotes the sub-interval after the last occurrence of $\mathcal{M}^{[t]}$ in I_j. Next, we define event $\mathcal{Q}^{[t]}$ as:

$$\mathcal{Q}^{[t]} : d_i - \frac{\Delta_i}{2} \leq e_i^{[t]} \leq d_i + \frac{\Delta_i}{2}, \forall i \in \mathcal{M}^{[t]} \tag{26}$$

By definition, $\mathcal{Q}^{[t]}$ indicates the event that the sampled value of each saturated arm is close to its mean. As the saturated arms have been played a sufficient number of times, $\mathcal{Q}^{[t]}$ should occur with a high probability. According to [36], we have:

$$\Pr\left\{\mathcal{Q}^{[t]}\right\} \geq 1 - \frac{4(N-1)}{T^2}, \forall t \tag{27}$$

where N is the number of arms. In our problem, $N = |\Omega_{\rm BS}|$. During any time step t, a saturated arm i will be played only when the sampled value $e_i^{[t]} < e_{i^*}^{[t]}$. Suppose $\mathcal{M}^{[t]}$ holds at time t and arm i is played, it must be $e_i^{[t]} < d_i - \frac{\Delta_i}{2}$. Thus, as long as the high probability event $\mathcal{Q}^{[t]}$ holds, the occurrence of $\mathcal{M}^{[t]}$ corresponds to one play of an unsaturated arm at time t. The number of times that saturated arms are played in the interval I_j is upper bounded by:

$$\sum_{l=1}^{\pi_j+1} |I_j(l)| + \sum_{t \in I_j} I\left(\overline{\mathcal{Q}}^{[t]}\right) \tag{28}$$

where the first term is the number of times that $\mathcal{M}^{[t]}$ does not occur, which excludes the plays of unsaturated arms; the second term corresponds to the rare events that the sampled values of all saturated arms fall out of their normal ranges.

The total regret caused by playing saturated arms is determined by both the number of plays and the selection of the arm played at each time step. Let $Y_j^{l,a}$ be the number times that arm a is the best saturated arm in sub-interval $I_j(l)$, i.e., $Y_j^{l,a} = |t \in I_j(l) : d_a = \min_{i \in \mathcal{C}^{[t]}} d_i|$. Recall that sub-intervals $I_j(l)(l = 1, \ldots, \pi_j + 1)$ correspond to the plays of saturated arms in interval I_j. Hence, we have $|I_j(l)| = \sum_{a \neq i^*} Y_j^{l,a}$.

We then analyze the regret due to playing any saturated arm when arm a is the best-saturated arm in one of the $Y_j^{l,a}$ steps.

A saturated arm may be played under two cases: (i) event $\mathcal{Q}^{[t]}$ holds, arm a or an arm with a mean very close to d_a is played; (ii) event $\mathcal{Q}^{[t]}$ is violated. Under case (i), if arm i is played, the following inequalities must hold:

$$d_i - \frac{\Delta_i}{2} \leq e_i^{[t]} \leq e_a^{[t]} \leq d_a + \frac{\Delta_a}{2}. \tag{29}$$

Then, Δ_i should satisfy:

$$\Delta_i = d_i - d_{i^*} \leq d_a - d_{i^*} + \frac{\Delta_a}{2} + \frac{\Delta_i}{2}. \tag{30}$$

Thus, we have $\Delta_i \leq 3\Delta_a$. Considering both cases, the regret due to playing a saturated arm when arm a is the best-saturated arm is upper bounded by $3\Delta_a + I\left(\overline{\mathcal{Q}}^{[t]}\right)$. Then, the expected total regret caused by playing saturated arms in the interval I_j is upper bounded by:

$$\begin{aligned}\mathbb{E}[\mathcal{R}^s(I_j)] &\leq \mathbb{E}\left[\sum_{l=1}^{\pi_j+1}\sum_{a\neq i^*}\sum_{t\in\mathcal{Y}_j^{l,a}}\left(3\Delta_a + I\left(\overline{\mathcal{Q}^{[t]}}\right)\right)\right] \\ &\quad + \sum_{t\in I_j} I\left(\overline{\mathcal{Q}^{[t]}}\right) \\ &= \mathbb{E}\left[\sum_{l=1}^{\pi_j+1}\sum_{a\neq i^*} 3\Delta_a Y_j^{l,a}\right] + 2\mathbb{E}\left[\sum_{t\in I_j} I\left(\overline{\mathcal{Q}^{[t]}}\right)\right]\end{aligned} \tag{31}$$

where $\mathcal{Y}_j^{l,a}$ is the set of time steps in which arm a is the best saturated arm. Following (31), the total regret caused by playing saturated arms over $t = 1, \ldots, T$ is upper bounded by:

$$\begin{aligned}\mathbb{E}[\mathcal{R}^s(T)] &= \sum_j \mathbb{E}[\mathcal{R}^s(I_j)] \leq \sum_j \mathbb{E}\left[\sum_{l=1}^{\pi_j+1}\sum_{a\neq i^*} 3\Delta_a Y_j^{l,a}\right] \\ &\quad + 2T\sum_t \Pr\left(\overline{\mathcal{Q}^{[t]}}\right).\end{aligned} \tag{32}$$

The upper bound for the second term of (32) can be obtained from (15). According to [36], the upper bound for the first term is given by:[4]

$$\begin{aligned}\sum_j \mathbb{E}\left[\sum_{l=1}^{\pi_j+1}\sum_{a\neq i^*} 3\Delta_a Y_j^{l,a}\right] &\leq 1152\ln T\left(\sum_i \frac{1}{\Delta_i^2}\right)^2 \\ &\quad + 288\ln T\sum_i \frac{1}{\Delta_i^2} + 48\ln T\sum_{a\neq i^*}\Delta_a \\ &\quad + 192N\ln T\sum_{a\neq i^*}\frac{1}{\Delta_a^2} + 96(N-1).\end{aligned} \tag{33}$$

For an unsaturated arm u, it becomes saturated after $\mathcal{L}_u = \frac{24\ln T}{\Delta_u}$ plays. Thus, the total regret due to plays of unsaturated arms is upper bounded by:

$$\mathbb{E}[\mathcal{R}^u(T)] \leq \sum_{u\neq i^*}\mathcal{L}_u\Delta_u = 24\ln T\sum_{u\neq i^*}\frac{1}{\Delta_u}. \tag{34}$$

[4]We omit some steps of the proof due to the space limit. The detailed proof can be found in Appendix D of [36].

Combining the results for saturated and unsaturated arms, we conclude that the total regret follows the upper bound given in Theorem 1. □

REFERENCES

[1] M. Feng, B. Akgun, I. Aykin, and M. Krunz, "Beamwidth optimization for initial access in 5G NR millimeter wave cellular networks," in *IEEE Proc. ICC'21,* Online, June 2021.

[2] J. G. Andrews, *et al.,* "What will 5G be?" *IEEE J. Sel. Areas Commun.,* vol. 32, no. 6, pp. 1065–1082, June 2014.

[3] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE,* vol. 109, no. 7, pp. 1166–1199, July 2021.

[4] J. Wang *et al.,* "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WLAN systems," *IEEE J. Sel. Areas Commun.,* vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

[5] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook design for millimeter-wave channel estimation with hybrid precoding structure," *IEEE Trans. Wireless Commun.,* vol. 16, no. 1, Jan. 2017.

[6] O. Chapelle and L. Li, "An empirical evaluation of Thompson sampling," in *Proc. of the Advances in Neural Information Processing Systems,*, 2011, pp. 2249–2257.

[7] H. Shokri-Ghadikolaei, C. Fischione, G. Fodor, P. Popovski, and M. Zorzi, "Millimeter wave cellular networks: A MAC layer perspective," *IEEE Trans. Commun.,* vol. 63, no. 10, pp. 3437–3458, Oct. 2015.

[8] Y. Li, J. G. Andrews, F. Baccelli, T. D. Novlan, and C. J. Zhang, "Design and analysis of initial access in millimeter wave cellular networks," *IEEE Trans. Wireless Commun.,* vol. 16, no. 10, pp. 6409–6425, Oct. 2017.

[9] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.,* vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[10] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Survey and Tutorials,* DOI: 10.1109/COMST.2018.2869411.

[11] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems", *IEEE J. Sel. Topics Signal Process.,* vol. 10, no. 3, pp. 543-559, Apr. 2016.

[12] 3GPP, "NR - Physical channels and modulation - Release 15," TS 38.211, V15.0.0, 2018.

[13] 3GPP, "Discussion on remaining issues of SS block and SS burst set," Motorola Mobility, Lenovo - Tdoc R1-1714212, 2017.

[14] Ericsson, "5G New Radio: Designing for the future," Ericsson Technology Review, 2017.

[15] 3GPP, "NR PRACH preamble resource allocation," Ericsson - Tdoc R1-1611905, 2016.

[16] R. S. Elliot, Antenna theory and design. John Wiley & Sons, 2006.

[17] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.,* vol. 53, no. 1, pp. 186–194, Jan. 2015.

[18] 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, June 2014.

[19] 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, Dec. 2015.

[20] C. N. Barati, S. A. Hosseini, M. Mezzavilla, T. Korakis, S. S. Panwar, S. Rangan, and M. Zorzi, "Initial access in millimeter wave cellular systems," *IEEE Trans. Wireless Commun.,* vol. 15, no. 12, pp. 6664–6678, Dec. 2016.

[21] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmWave cellular networks," *IEEE Commun. Mag.,* vol. 54, no. 11, pp. 40–47, Nov. 2016.

[22] A. Alkhateeb, Y.-H. Nam, M. S Rahman, J. Zhang, and R. W. Heath, "Initial beam association in millimeter wave cellular systems: Analysis and design insights," *IEEE Trans. Wireless Commun.,* vol. 16, no. 5, pp. 2807–2821, May 2017.

[23] H. Hassanieh, O. Abari, M. Rodriguez, M. Abdelghany, D. Katabi, and P. Indyk, "Fast millimeter wave beam alignment," in *Proc. ACM SIGCOMM'18,* Budapest, Hungary, Aug. 2018, pp. 432–445.

[24] I. Aykin and M. Krunz, "FastLink: An efficient initial access protocol for millimeter wave systems," in *Proc. ACM MSWiM'18,* Montreal, Canada, Oct. 2018, pp. 109–117.

[25] Morteza Hashemi, A. Sabharwal, C. E. Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," in *Proc. IEEE INFOCOM'18,* Honolulu, HI, Apr. 2018, pp. 2393–2401.
[26] Y. Yang, H. S. Ghadikolaei, C. Fischione, M. Petrova, and K. W. Sung, "Reducing initial cell-search latency in mmwave networks," *IEEE INFOCOM'18 WKSHPS,* Honolulu, HI, Apr. 2018, pp. 686–691.
[27] M. Alrabeiah, Y. Zhang, and A. Alkhateeb, "Neural networks based beam codebooks: Learning mmWave massive MIMO beams that adapt to deployment and hardware," *IEEE Transactions on Communications,* vol. 70, no. 6, pp. 3818–3833, June 2022.
[28] A. Ali, N. G.-Prelcic, and R. W. Heath, "Estimating millimeter wave channels using out-of-band measurements," *IEEE ITA WKSHPS 2016,* La Jolla, CA, Jan. 2016, pp. 1–6.
[29] A. Hu and J. He, "Position-aided beam learning for initial access in mmWave MIMO cellular networks," *IEEE Systems J.,* vol. 16, no. 1, pp. 1103–1113, Mar. 2022.
[30] J. Gao, C. Zhong, X. Chen, H. Lin, and Z. Zhang, "Deep reinforcement learning for joint beamwidth and power optimization in mmWave systems," *IEEE Communications Letters,* vol. 24, no. 10, pp. 2201–2205, Oct. 2020.
[31] M. A. Almasi, L. Jiang, H. Jafarkhani, and H. Mehrpouyan, "Joint beamwidth and power optimization in mmWave hybrid beamforming-NOMA systems," *IEEE Transactions on Wireless Communications,* vol. 20, no. 4, pp. 2442–2456, Ap. 2021.
[32] I. Aykin, B. Akgun, and M. Krunz, "Smartlink: Exploiting channel clustering effects for reliable millimeter wave communications," in *Proc. IEEE INFOCOM'19,* Paris, France, Apr. 2019.
[33] 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.
[34] 3GPP, "Study on 3D channel model for LTE," Tech. Rep. 3GPP 36.873 (V12.2.0), July 2015.
[35] T. S. Rappaport, G. R. MacCartney, Jr., M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, Sept. 2015.
[36] S. Agrawal and N. Goyal, "Analysis of Thompson sampling for the multi-armed bandit problem," in *Proc. of the 25th Annual Conference On Learning (COLT)*, vol. 23, pp. 39.1–39.26, June 2012.