

# GAMBIT: A Generalized Adaptive Multi-Armed Bandit for Intelligent Beam Tracking and Rate Adaptation

Sopan Sarkar, *Student Member, IEEE*, Marwan Krunz, *Fellow, IEEE*, Sherif Badran, *Student Member, IEEE*, Josep M Jornet, *Fellow, IEEE*, David Manzi, *Raytheon Corporation*, and Rajesh Kulkarni, *Raytheon Corporation*

**Abstract**—Beamforming provides transmission/reception directivity gains that compensate for the high propagation loss encountered at millimeter-wave (mmWave) and sub-THz bands. However, narrow beams introduce significant beam misalignment challenges. Providing fast and efficient beam tracking is vital for maintaining communications and minimizing service disruptions. This paper introduces *GAMBIT*, a restless multi-armed bandit (MAB) scheme for beam tracking and rate adaptation in mobile directional systems. *GAMBIT* aims to select the optimal beam and modulation and coding scheme (MCS) for upcoming transmissions through an online reinforcement learning technique called *Top-K Adaptive Thompson Sampling (Top-K-ATS)*. According to this technique,  $K$  beams ( $K \geq 1$ ) are chosen and ranked based on previously estimated beam quality information. This information is initially gathered during cell discovery and is updated periodically based on explicit or implicit feedback from user equipment (UE). The best of the  $K$  selected beams, called the “*leader*,” is used for communications. The remaining beams, referred to as “*scouts*,” provide contextual information about the RF environment. To prevent beam quality information from becoming stale due to channel/mobility dynamics, we use beam coherence time analysis to derive an upper bound on the time between consecutive beam selection instances. We evaluate the performance of *GAMBIT* through simulations at 28 GHz and over-the-air (OTA) measurements at 28 GHz and 130 GHz. We compare our scheme with  $\epsilon$ -greedy, upper confidence bound (UCB), and Thompson sampling (TS) beam tracking algorithms. Results indicate that *GAMBIT* outperforms its contenders in both achievable data rate and outage probability.

**Index Terms**—Millimeter-wave, sub-THz, beam tracking, rate adaptation, reinforcement learning, multi-armed bandit, beam coherence time.

## I. INTRODUCTION

AS the demand for higher data rates keeps rising, wireless systems are increasingly shifting to new spectrum in the millimeter-wave (mmWave) and sub-THz bands. The so-called “high bands” are key aspects of 5G and NextG cellular systems [1]. One of their main drawbacks is that the signal suffers from very high attenuation [2]. At the same time, the small wavelengths allows portable devices to integrate

large antenna arrays while maintaining small form-factors. By utilizing high-dimensional phased arrays for beamforming, narrow beams can be steered along any desired direction for transmission/reception. The high directivity (and gain) of these antennas makes it possible to achieve high data rates even in the harsh mmWave channel environment [3].

While beamforming provides high gains, establishing and maintaining a directional link can be challenging for several reasons [4]. During the initial access (IA) process, i.e., cell discovery, the base station (BS) establishes a communication link with new user equipment (UE) and updates the links of already connected UEs. IA overhead can be significant because of the large number of directions to be scanned [5], [6]. Furthermore, due to UE mobility and environmental obstacles (e.g., trees), significant beam misalignment can occur, resulting in delayed and incorrect channel state information reference signal (CSI-RS) measurements. This leads to improper selection of the modulation and coding scheme (MCS), reduced link throughput, and outages [7]. Therefore, it is critical to track mobile UEs, assign them to appropriate beams, and accurately select the best possible MCS index that maintains the quality of service of the served UE.

Efficient beam tracking in mmWave systems has been an active area of research. Several methods have been introduced to tackle this issue, as highlighted in a recent survey [8]. These methods include Bayesian filters [9]–[12], which utilize Kalman filters [9], [10] or particle filters [11], [12] to track the Angle of Arrival (AoA), Angle of Departure (AoD), and channel gain. These methods heavily depend on accurate signal propagation models. In practice, propagation models are inherently imperfect, which can cause the Bayesian filter-based algorithms to be ineffective.

Machine learning (ML) based beam tracking algorithms have also been proposed. These algorithms can be broadly categorized into two types: supervised learning (SL) and reinforcement learning (RL). SL is more commonly used due to its simplicity. Representative examples of SL algorithms for beam tracking are given in [13]–[15]. In particular, the authors in [13] introduced DeepBeam, a scheme that leverages convolutional neural network (CNN)’s feature extraction capabilities to passively eavesdrop on ongoing data transmissions, infer the AoA and the beam ID, and perform beam tracking. In [14] the authors proposed DeepIA, a deep neural network (DNN)-based beam tracking algorithm. Unlike multi-codebook-based approaches that use wide and narrow beams, DeepIA employs

An abridged version of this paper appeared in the *Proceedings of the Globecom’21 Conference*, 7–11 December 2021, Madrid, Spain.

S. Sarkar and M. Krunz are with Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721, USA (email: {sopansarkar, krunz}@arizona.edu). S. Badran and JM. Jornet are with Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, 02115, USA (email: {badran.s, j.jornet}@northeastern.edu). D. Manzi and R. Kulkarni are with Raytheon Corporation, USA (email: {dgmanzi, rajesh\_kulkarni}@raytheon.com)

a single-beam codebook. However, it optimizes the beam training process by selectively sweeping only a small subset of beams. The authors in [15] integrated an auto-encoder into a long short-term memory (LSTM) model for multi-cell multi-beam prediction using channel quality information (i.e., SNR).

SL techniques, in general, rely on labeled data, and hence may not be well-suited for handling the dynamic radio environment of mmWave communications. RL techniques can overcome such issues. Two RL techniques are commonly discussed in the literature: Q-learning and multi-armed bandits (MABs). Q-learning is a model-free algorithm that aims to identify the best action in a given state by evaluating the Q-value (beam quality) associated with each action. Its ultimate goal is to discover the optimal policy that maximizes the cumulative reward received over time. In [16] the authors introduced a beam management method that employs Q-learning. The objective was to determine the beam that yields the highest received power. During the exploration phase, the agent intentionally selects a specific set of serving beams to evaluate the rewards associated with this set. The agent may intentionally select suboptimal beams to explore alternative possibilities. Conversely, during the exploitation phase, the agent consistently opts for the best beam that maximizes the received power. Another related approach combines Q-learning with an auxiliary beam pair to further reduce the beam search space [17]. A significant limitation of Q-learning is its slow convergence due to the need to explore all possible state-action pairs. To reduce the training time of Q-learning, deep Q-networks (DQN) can be used, where the Q-values are predicted using a neural network. In [18] the authors proposed a DQN-based approach for beam tracking. The approach dynamically adjusts the beam probing range to adapt to environmental changes, making it suitable for highly mobile UEs. Considering both slow and fast mobile UEs, performance results in [18] demonstrate that the DQN-based approach outperforms traditional Q-learning in terms of the learning speed. However, a fundamental limitation of DQNs is that they can only be used with discrete and low-dimensional action spaces. Furthermore, for the beam tracking problem, selecting a new action (beam) does not change the system's state space. In other words, choosing a beam at time  $t$  does not affect the set of beams that can be chosen at time  $t + 1$ , so the problem can be adequately modeled as a single-state Markov decision process (MDP). In such cases, introducing Q-learning or DQN adds complexity without providing substantial benefits.

Due to the sequential nature of beam tracking, an inherent exploration vs. exploitation tradeoff exists, and the problem can be formulated as a MAB problem. In [19] a MAB-based beam tracking solution was proposed, which incorporates contextual information. In [20] beam tracking algorithms based on  $\epsilon$ -greedy and upper confidence bound (UCB) algorithm were proposed. The issue with such MAB technique is that it cannot tackle the time-varying nature of the channel and high-mobility scenarios. In [21] a MAB-based approach based on a modified Thompson sampling (TS) algorithm was proposed to address non-stationarity and utilize historical data. A similar technique was used in [22] to verify the effectiveness of mmWave beam tracking over the open-source COSMOS testbed.

In this paper, we propose GAMBIT, a restless multi-armed bandit framework for beam tracking and MCS index selection in directional systems. In GAMBIT, the BS acts as the agent and interacts with each beam to learn the changes in beam quality over time. Beam quality is reflected in the best MCS that the beam can support. We develop an RL algorithm, called the Top-K adaptive Thompson sampling (Top-K-ATS), to be used with GAMBIT. Top-K-ATS aims to maximize the expected transmission rate by determining the best beam and MCS index for the upcoming transmission/reception based on past channel observations.

The main contributions of this paper are as follows:

- We propose a restless multi-armed bandit framework called GAMBIT for joint beam and MCS index selection in directional wireless systems. GAMBIT integrates an RL-based algorithm called Top-K-ATS for selecting  $K$  beams and ranking them based on the probability of maximizing the average data rate of the underlying transmission. Among these  $K$  beams, the best beam, called the “*leader*,” is used for communication. The remaining  $K - 1$  beams, called “*scouts*,” are used for gathering information related to the dynamics of the wireless channel. Top-K-ATS also selects the optimal MCS index to be used by considering the fluctuations in the signal-to-noise ratio (SNR) during previous transmissions.
- Top-K-ATS relies on two parameters to capture the dynamics of the wireless channel: a “*forgetting factor*” ( $\gamma_1$ ), which discounts the relevance of past information, and a “*boost factor*” ( $\gamma_2$ ), which increases the weight of recent observations. We show through simulations that the performance of GAMBIT is particularly sensitive to the value of  $\gamma_1$ . Furthermore, we show how the throughput-optimal value of  $\gamma_1$  ( $\gamma_1^*$ ) depends on the change in SNR over time. Accordingly, we develop a multi-layer perceptron (MLP) neural network to predict the value of  $\gamma_1^*$ . The MLP is trained offline using a dataset collected via extensive simulations, and predictions are performed online based on SNR estimates.
- We obtain an upper bound on the optimal time between two execution instances of GAMBIT so as to avoid using obsolete past observations. This time is derived from the beam coherence time, which reflects the impact of beam misalignment due to UE mobility and dictates how frequently the channel should be observed and when to select a new beam.
- We conduct simulations for outdoor scenarios at 28 GHz using phased planner arrays and also utilize experimental datasets collected at 28 GHz and 130 GHz to verify the efficiency of GAMBIT in terms of average data rate, instantaneous data rate, and outage probability. Our results show that GAMBIT improves the amount of delivered traffic by up to 71.2% relative to the default beam management scheme used in 5G NR.

The rest of the paper is organized as follows. In Section II we describe the system model. Section III briefly overviews the main components of GAMBIT. The derived timing requirements for channel scanning, the description of the MLP

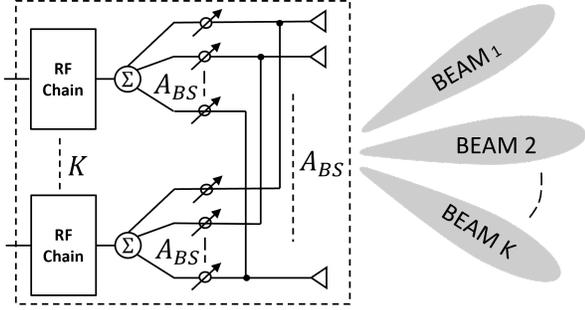


Fig. 1: System model.

framework, and the simulation setup are presented in Sections IV, V, and VI, respectively. We evaluate the performance of GAMBIT in Section VII. Finally, Section VIII concludes the paper.

## II. SYSTEM MODEL

### A. BS and UE Models

Without loss of generality, we consider an outdoor scenario in which a BS tracks a mobile UE. Beamforming is applied at both the BS and the UE, with wider beams used at the UE (reflecting its fewer antenna elements). The BS uses multiple receive beams to receive different copies of the same transmitted signal. We consider beam tracking from the BS side only.

Without loss of generality, we focus on a single UE and ignore any inter-user interference. In 5G systems, intra-cell interference between users is mitigated via orthogonal frequency division multiple access (OFDMA), which allocates different subcarriers to different users in the cell. Inter-cell interference can be addressed by assigning different channels to neighboring cells using graph coloring techniques. Furthermore, receive diversity techniques, such as maximum ratio combining (MRC) and interference rejection combining (IRC), can be applied to significantly enhance system performance. MRC maximizes the SNR by combining signals across multiple antennas, strengthening the desired signal while suppressing interference. In contrast, IRC specifically filters out interference from adjacent subcarriers.

### B. Multi-beam Antenna Model

We assume that the BS and UE are equipped with uniform planar arrays (UPAs). In practice, 5G systems employ hybrid beamforming, which combines analog and digital beamforming. For simplicity, we consider analog beamforming due its lower hardware complexity and reduced feedback overhead, particularly in multi-user mmWave systems. Our proposed GAMBIT scheme is in fact agnostic to the underlying beamforming strategy, and only requires the capability to generate multiple beams, which can be achieved through either hybrid beamforming or multi-beam analog beamforming.

At the BS, the multi-beam antenna model of [23] is assumed. According to this model, there are  $A_{BS}$  antenna elements and  $K$  RF chains that can transmit  $K$  simultaneous

beams (see Fig. 1). Let  $A_{UE}$  be the number of antenna elements at the UE. The transmitter (Tx) and receiver (Rx) beamforming vectors (a.k.a beamers) depend on the  $A_{UE} \times A_{BS}$  complex channel matrix  $\mathbf{H}$  between the BS and UE. We assume that following IA, both the BS and the UE have agreed on a particular beam. Suppose that the BS uses a Tx beamforming vector  $\mathbf{m}_q \in \mathbb{C}^{A_{BS} \times 1}$  and the UE uses an Rx beamforming vector  $\mathbf{n}_l \in \mathbb{C}^{A_{UE} \times 1}$  ( $q$  and  $l$  are the indices of the Tx/Rx beamforming vectors in their respective codebooks  $\mathbb{C}$ ). Let  $s(t)$  be the transmitted signal at time  $t$ . The received signal  $y_{ql}(t)$  can be expressed as:

$$y_{ql}(t) = \mathbf{n}_l^H \mathbf{H} \mathbf{m}_q s + \mathbf{n}_l^H \mathbf{z}(t) \quad (1)$$

where  $\mathbf{z} \in \mathbb{C}^{A_{UE} \times 1}$  is a vector representing complex circularly symmetric white Gaussian noise. Each  $(\mathbf{m}_q, \mathbf{n}_l)$  pair achieves a certain Rx power  $P_{ql}(t)$  at time  $t$ , where  $P_{ql}(t) = |y_{ql}|^2$ . Therefore, the received signal-to-noise ratio (SNR) is given by

$$\text{SNR}_{ql}(t) = \frac{|\mathbf{n}_l^H \mathbf{H} \mathbf{m}_q s|^2}{\mathbf{n}_l^H \mathbf{z}(t)}. \quad (2)$$

Note that the SNR is time-varying because  $\mathbf{H}$  is also time-varying.

## III. GAMBIT DESIGN

In the current 5G standards [24], the BS executes IA to admit a new UE to the network and then determines the best beam pair  $(\mathbf{m}_k, \mathbf{n}_l)$ . Simultaneously, the BS updates the beams of currently served UEs. A simple tracking approach maintains (exploits) the current best beam pair for a relatively long time until a new IA cycle is initiated. However, IA takes considerable time (approximately 5 ms), and performing more frequent IA due to beam misalignment incurs significant overhead. In addition, 5G mmWave systems rely on CSI-RS control messages to maintain reliable data rates. However, due to channel dynamics, the coherence time at mmWave frequencies is too short, which makes it challenging to perform frequent CSI-RS measurements. GAMBIT aims to extend the period between two IA cycles as much as possible, and yet maintain the required quality of service (QoS).

### A. Formulation of the Beam Tracking Problem

In GAMBIT, beam tracking is modeled as a single-state Markov decision process, where the BS acts as an agent that interacts with the UE through directional beams (modeled as arms of the MAB) to learn the changes in beam qualities over time. Beam quality impacts the best possible MCS index that can be supported at a given instance. GAMBIT utilizes an RL algorithm called Top-K-ATS to select the best beam and MCS index. According to Top-K-ATS,  $K$  out of  $M_{BS}$  beams at the BS are selected and ranked based on their average data rate. The beam with the highest probability, called the “*leader*,” is used to communicate with the UE. The UE sends a positive acknowledgment (ACK) or a negative ACK (NACK) to the BS. The remaining  $K - 1$  beams, “*scouts*,” are used to evaluate the channel quality along their respective directions by monitoring the ACK/NACK messages. These scouts provide additional

information regarding the channel and increase the rate of exploration without reducing the exploitation rate.

Top-K-ATS is specified by the tuple  $\langle \mathcal{A}, \mathcal{R} \rangle$ , where  $\mathcal{A} \triangleq \{\mathcal{J}_t \subseteq \mathcal{M} : |\mathcal{J}_t| = K\}$  is the set of actions, i.e., the selected  $K$  beams, and  $\mathcal{R}$  is the set of rewards, i.e., achievable rates associated with the selected beams (actions). Here,  $\mathcal{M}$  is the set of all beams at the BS. For each beam  $i \in \mathcal{J}_t \subseteq \mathcal{M}$ , a random reward  $\mathbf{r}_{i,t} \in \mathcal{R}$  is observed at time  $t$ , which follows a reward distribution  $\Theta_{i,t}$ . Let  $\mathbb{E}[\Theta_{i,t}] \stackrel{\text{def}}{=} \boldsymbol{\theta}_{i,t}$  (unknown). After IA, GAMBIT designs a beam tracking policy, which is defined as a  $T$ -element vector that specifies the actions to be taken at subsequent times  $t = 1, 2, \dots, T$ , so as to maximize the cumulative reward.

Through IA, the BS has prior belief about the reward distribution associated with each beam. Using Bayesian inference, the BS continuously updates the posterior of mean reward ( $\boldsymbol{\theta}$ ), given the observed data ( $\mathbf{u}$ ), as follows:

$$\Pr(\boldsymbol{\theta}|\mathbf{u}) = \Pr(\mathbf{u}|\boldsymbol{\theta})\Pr(\boldsymbol{\theta})/\Pr(\mathbf{u}). \quad (3)$$

The reward distribution is represented by the categorical random variable  $\mathbf{r}_{i,t} = [r_{i,t}^{(0)}, \dots, r_{i,t}^{(C-1)}]$ , where  $C$  is the number of MCS indices supported by the system plus one. Here,  $\mathbf{r}_{i,t}$  is a vector of zeros except for one element whose value is 1, representing the highest attainable MCS index  $c$ ,  $c \in \{1, \dots, C-1\}$ . The observed data rate at time  $t$  is given by  $\mathbf{r}_{i,t}\mathbf{w}^T$ , where  $\mathbf{w} \stackrel{\text{def}}{=} [w_0, w_1, \dots, w_{C-1}]$  is a vector whose entries correspond to the transmission rates associated with different MCS indices. The BS can communicate with the UE using one of the  $C-1$  MCS indices. If the BS fails to establish communication with the UE,  $w_0 \stackrel{\text{def}}{=} 0$  is selected, i.e.,  $r_{i,t}^{(0)} = 1$ . The BS performs its computation based on the feedback (ACK/NACK) received from the UE. Specifically, the BS calculates the SNR of the ACK/NACK packet and determines the highest possible MCS index for downlink communication. In fact, any metric of the wireless signal, such as received signal strength (RSS), reference signal receive power (RSRP), reference signal received quality (RSRQ), or the SNR, can be used with little to no modifications to the GAMBIT algorithm. In this paper, we choose SNR specifically due to the availability of channel quality indicator (CQI)-to-MCS-to-SNR mapping tables for 5G mmWave communications.

Given the above, the objective of the BS is to select an optimal policy  $\boldsymbol{\mu} = [\mathcal{J}_1, \dots, \mathcal{J}_T]$ , i.e., a sequence of  $K$  transmit beams over time slots  $t = 1, \dots, T$ , that maximizes the expected throughput. If the expected reward vectors  $\boldsymbol{\theta}_{i,t} = [\theta_{i,t}^{(0)}, \dots, \theta_{i,t}^{(C-1)}]$ , where  $\theta_{i,t}^{(c)}$  represents the probability of the  $c$ th element of  $\boldsymbol{\theta}_{i,t}$  such that  $\theta_{i,t}^{(c)} > 0 \forall c$  and  $\sum_{c=0}^{C-1} \theta_{i,t}^{(c)} = 1$ , are known for each beam  $i$  at every time  $t$ , the problem reduces to solving the following optimization formulation:

$$\underset{\boldsymbol{\mu}}{\text{maximize}} \quad \frac{1}{T} \frac{1}{K} \sum_{t=1}^T \sum_{i \in \mathcal{J}_t} \boldsymbol{\theta}_{i,t} \mathbf{w}^T \quad (4)$$

$$\text{s.t.} \quad \sum_{c=0}^{C-1} \theta_{i,t}^{(c)} = 1, \quad \forall i, t, \quad (5)$$

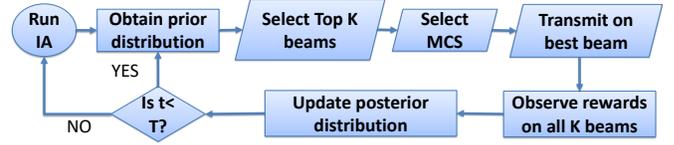


Fig. 2: Flowchart of the GAMBIT scheme at a BS.

$$\theta_{i,t}^{(c)} \geq 0, \quad \forall i, t, c. \quad (6)$$

In (4), the product  $\boldsymbol{\theta}_{i,t}\mathbf{w}^T$  quantifies the contribution of beam  $i$  to the throughput at time  $t$ . Note that the rewards distribution is unknown and nonstationary. As a result, (4) cannot be solved directly. Instead, an ATS-based RL algorithm is used to learn the expected rewards and select the optimal policy. A flowchart of the GAMBIT framework at a BS is shown in Fig. 2.

### B. Top-K Adaptive Thompson Sampling (Top-K-ATS)

We model the prior of the expected rewards as a Dirichlet random variable with parameter  $\boldsymbol{\alpha}_{i,t}$ ,  $\text{Dir}(\boldsymbol{\alpha}_{i,t})$ . The rationale for this choice is that the Dirichlet distribution is the conjugate prior of the categorical distribution, used to model the average reward per beam. According to (3), the posterior computed at each round will also follow a Dirichlet distribution. The update rule is simpler when the prior is the conjugate distribution of the likelihood.

Consider first a case where only a single beam is selected, i.e.,  $K = 1$ . At time  $t$ , suppose that action  $a_t$  results in selecting beam  $i$ , and reward  $\mathbf{r}_t$  is observed. For categorical rewards and Dirichlet priors, the posterior for each beam  $i$  is updated as:

$$\boldsymbol{\alpha}_{i,t+1} = \begin{cases} \boldsymbol{\alpha}_{i,t} + \mathbf{r}_t, & \text{if } a_t = i \\ \boldsymbol{\alpha}_{i,t}, & \text{if } a_t \neq i. \end{cases}$$

After such an update, the BS selects a beam for the next round by randomly sampling from the current posterior distributions. Specifically, at each time  $t$ , the BS samples  $\mathbf{s}_{i,t}$  from  $\text{Dir}(\boldsymbol{\alpha}_{i,t})$ ,  $\forall i \in \mathcal{A}$  (where  $|\mathbf{s}_{i,t}| = 1, \forall i \in \mathcal{A}$  and  $\forall t \in \{1, \dots, T\}$ ), and selects the action according to Thompson sampling as follows:

$$a_t = \arg \max_{i \in \mathcal{A}} \mathbf{s}_{i,t} \mathbf{w}^T. \quad (7)$$

This sampling implies that even though beams with currently high estimated means are more likely to be selected, other beams get a chance to be picked and updated, i.e., exploration versus exploitation.

The above scheme is suitable when the reward distribution is stationary. To capture the dynamic nature of directional mmWave channels under mobility, GAMBIT adopts ATS [21], where a “forgetting factor”  $\gamma_1$  is used to reduce the effect of past observations, and a “boost factor”  $\gamma_2$  is used to

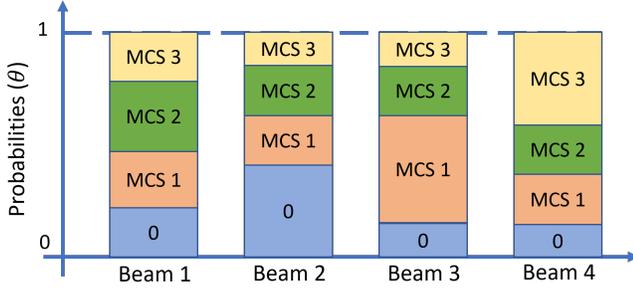


Fig. 3: Illustration of qualities of different beams in terms of achievable MCS indices ( $C = 4$ ). Probabilities refer to  $\theta_{i,t}^{(c)}$ ,  $\forall i \in \mathcal{A}$  and  $\forall m \in 1, \dots, C$  at a given  $t$ .

emphasize the most recent observation. For beam  $i \in \mathcal{A}$ , the updated posterior under these factors is as follows:

$$\alpha_{i,t+1} = \begin{cases} \gamma_1 \alpha_{i,t} + \gamma_2 \mathbf{r}_t, & \text{if } a_t = i \\ \gamma_1 \alpha_{i,t}, & \text{if } a_t \neq i \text{ and } \gamma_1 \|\alpha_{i,t}\|_1 > 1 \\ \mathbf{1}, & \text{otherwise.} \end{cases}$$

Here,  $\gamma_1 \in (0, 1]$ ,  $\gamma_2 \geq 1$ , and  $\|\cdot\|_1$  is the 1-norm of a vector. Multiplying  $\alpha_{i,t}$  by  $\gamma_1$  increases the variance of the posterior distribution for each beam, which increases the rate of exploration and, as a result, the BS tends to switch to new beams rather than exploiting the same one for a long time. Thus,  $\gamma_1$  dictates the rate of exploration. As for  $\gamma_2$ , this parameter emphasizes the currently selected beam by boosting its observed reward, thus dictating the exploitation rate. The final condition in the update rule ensures that if a beam is not selected for a long time,  $\alpha_{i,t}$  is updated in such a way that the belief on the beam's distribution converges to the multi-dimensional uniform distribution,  $\text{Dir}(\mathbf{1})$ . From the above discussion, it is evident that  $\gamma_1$  and  $\gamma_2$  significantly impact the performance of GAMBIT and how well it adapts to the dynamics of the wireless channel under mobility.

In the mmWave bands, as beams become narrower, more beams need to be explored. As a result, it becomes difficult for ATS to adapt to user mobility. To obtain more context about channel dynamics, we propose to use multiple beams to receive feedback from the UE and characterize the qualities of the specific beams. To do so, at each time  $t$ , the BS samples from each beam's updated distribution to obtain  $s_{i,t} \sim \text{Dir}(\alpha_{i,t})$ , (where  $|s_{i,t}| = 1$ ,  $\forall i \in \mathcal{M}$ ,  $\forall t \in \{1, \dots, T\}$ ), and calculates the probability of maximizing the average data rate as  $s_{i,t} \mathbf{w}^T$ . Next, the beams corresponding to the  $K$  highest values of  $s_{i,t} \mathbf{w}^T$  are selected and denoted by  $\mathcal{J}_t$ . For each beam  $i$ , the updated posterior under these factors is as follows:

$$\alpha_{i,t+1} = \begin{cases} \gamma_1 \alpha_{i,t} + \gamma_2 \mathbf{r}_t, & \text{if } i \in \mathcal{J}_t \\ \gamma_1 \alpha_{i,t}, & \text{if } i \notin \mathcal{J}_t \text{ and } \gamma_1 \|\alpha_{i,t}\|_1 > 1 \\ \mathbf{1}, & \text{otherwise.} \end{cases}$$

### C. MCS Selection

After selecting a beam using Top-K-ATS, the BS determines the optimal MCS index that will be used during the next transmission over that beam. This selection is important because

if the UE cannot support the MCS selected by the BS, the effective data rate will be zero. In this paper, we consider a conservative selection scheme where the MCS index that is *most likely* to be attained and can achieve a non-zero rate on the selected beam is used for transmission. After the BS collects  $s_{i,t}$ ,  $\forall i \in \mathcal{M}$ , and selects the “leader,” it will first select a transmission rate based on the probabilities of attaining different MCS indices on the selected beam. Specifically, given that action  $a_t$  has been selected, the initial rate selection problem by the leader can be written as:

$$m^* = \arg \max_{m \in \{0, \dots, M-1\}} s_{i,t}^{(m)}. \quad (8)$$

For example, Fig. 3 depicts the probabilities of achieving different MCS's for four beams ( $C = 4$ ). If we select the Top-2 beams ( $K = 2$ ), beams 4 and 1 will be selected by the framework where beam 4 will be selected as the “leader” and beam 1 as a “scout”. This is because  $s_{4,t} \mathbf{w}^T > s_{1,t} \mathbf{w}^T > s_{3,t} \mathbf{w}^T > s_{2,t} \mathbf{w}^T$ .

After the initial rate has been selected, the BS then calculates the SNR corresponding to  $m^*$  and adjusts the SNR based on the average of recently observed changes of the SNR within a given time window. Let  $\Delta_{\text{SNR}}$  be such an average. Then, the conservative rate selected is given by

$$m_c = \text{MCS}(f_{\text{mcs} \rightarrow \text{snr}}(m^*) - \Delta_{\text{SNR}}) \quad (9)$$

where  $f_{\text{mcs} \rightarrow \text{snr}}$  is a mapping between MCS and SNR values. If the UE is static for some time,  $\Delta_{\text{SNR}} \rightarrow 0$ . On the other hand, if the UE is mobile,  $\Delta_{\text{SNR}}$  will vary.

## IV. OPTIMIZING BEAM SELECTION TIMES

If mobility and beam misalignment are too rapid relative to the time between two beam selection instances, GAMBIT will be too slow to learn the dynamics of the environment. To address this issue, we establish an upper bound on the interval between two beam selection instances. The channel must be measured within this interval and a decision must be made whether or not to rerun GAMBIT.

Consider a scenario in which the UE moves within the coverage area of the BS, as shown in Fig. 4. Suppose that at time  $t$ , the BS runs IA to discover and connect with the UE at location  $A$ . Let  $d_t$  be the distance between the BS and UE at time  $t$ . We assume the BS and UE beams are perfectly aligned right after IA. Suppose that the UE starts moving at a fixed speed  $v$  towards location  $B$  at an angle  $\beta_t$ , and it reaches  $B$  after a short duration  $t'$ . At time  $t + t'$ , the new distance between the BS and the UE is  $d_{t+t'}$ . Due to UE mobility from  $A$  to  $B$ , beam misalignment may occur, and the SNR of the received signal could drop. The angular change in the AoA of the line-of-sight (LOS) signal at the UE between times  $t$  and  $t + t'$  is defined as the beam pointing error and is represented in Fig. 4 as  $\Delta\phi_{t+t'}$ . Note that  $\Delta\phi_t = 0^\circ$ .

To determine an appropriate beam selection time, we rely on the concept of *beam coherence time*,  $T_B$ , defined as the first time when the received SNR drops below a certain threshold,  $\xi$ , from its peak at time  $t$  [25]:

$$T_B = \inf_{t' > 0} \{t' \mid \frac{\text{SNR}(t+t')}{\text{SNR}(t)} < \xi\}. \quad (10)$$

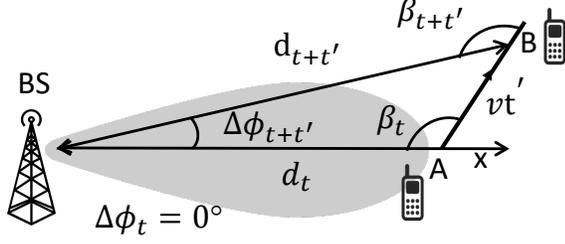


Fig. 4: Misalignment caused by UE mobility.

Suppose that the transmit power and antenna gain of the UE are fixed, and the noise power does not vary drastically over the duration  $T_B$ . Assume for the time being that no sudden blockages due to large obstacles occur in the interval  $[t, t+t']$ . Then, the received SNR at any given time will depend only on the receiver antenna gain and the path loss. Accordingly, we can express  $\xi$  as

$$\begin{aligned} \xi &= \frac{\text{SNR}(t+t')}{\text{SNR}(t)} \\ &= \frac{G(\Delta\phi_{t+t'}) \times L(d_{t+t'})}{G(\Delta\phi_t) \times L(d_t)}. \end{aligned} \quad (11)$$

Here,  $G(\phi)$  is the Rx antenna gain at an angle  $\phi$  from the boresight. For UPAs, we can express  $G(\phi)$  as [26]:

$$G(\phi) = G_{\max} 10^{-\frac{3}{10} \left(\frac{2\phi}{\phi_m}\right)^2}, \text{ for } |\phi| \leq \frac{\phi_m}{2} \quad (12)$$

where  $G_{\max} = G(0^\circ)$  is the maximum antenna gain:

$$G_{\max} = \frac{2\pi 10^{\frac{3}{10} \left(\frac{\phi_m}{\phi_m}\right)^2}}{V(\phi_m, \phi_m) + 2\pi - \phi_m} \quad (13)$$

$$V(\phi_m, \phi_m) = \int_0^{\phi_m} 10^{\frac{3}{10} \left(\frac{\phi_m^2 - \phi^2}{\phi_m^2}\right)} d\phi. \quad (14)$$

$\phi_\omega$  is the half-power beamwidth (HPBW) and  $\phi_m$  is the beamwidth of the main lobe. When  $\Delta\phi > \frac{\phi_m}{2}$ , we consider this situation as beam-alignment failure rather than beam misalignment.  $L(d)$  is the path-loss at a distance  $d$  from the BS, and is given by

$$L(d) = \left(\frac{\lambda}{4\pi}\right)^2 \times d^{-\alpha} \quad (15)$$

where  $\lambda$  is the wavelength and  $\alpha$  is the path-loss exponent. From (11), (12) and (15), we can calculate the value of  $\xi$ :

$$\begin{aligned} \xi &= \frac{G_{\max} 10^{-\frac{3}{10} \left(\frac{2\phi_{t+t'}}{\phi_\omega}\right)^2} \times \left(\frac{\lambda}{4\pi}\right)^2 \times d_{t+t'}^{-\alpha}}{G_{\max} 10^{-\frac{3}{10} \left(\frac{2\phi_t}{\phi_\omega}\right)^2} \times \left(\frac{\lambda}{4\pi}\right)^2 \times d_t^{-\alpha}} \\ &= 10^{-\frac{3}{10} \left(\frac{2\Delta\phi_{t+t'}}{\phi_\omega}\right)^2} \times \left(\frac{d_{t+t'}}{d_t}\right)^{-\alpha}. \end{aligned} \quad (16)$$

The first term of the RHS of (16) represents the impact of antenna gain on the SNR, whereas the second term represents

the effect of path loss. From Fig. 4 and utilizing the cosine laws of the triangle, we can express  $d_{t+t'}$  as

$$d_{t+t'} = \sqrt{d_t^2 + v^2(t')^2 - 2d_tv t' \cos \beta_t}. \quad (17)$$

Now, using (16) and (17), and replacing  $t'$  with  $T_B$ , we obtain:

$$d_{t+T_B} = d_t \left( 10^{-\frac{3}{10} \left(\frac{2\Delta\phi_{t+T_B}}{\phi_\omega}\right)^2} \times \xi^{-1} \right)^{\frac{1}{\alpha}}. \quad (18)$$

Accordingly,

$$T_B = \frac{d_t^2 \left( 10^{-\frac{3}{10} \left(\frac{2\Delta\phi_{t+T_B}}{\phi_\omega}\right)^2} \times \xi^{-1} \right)^{\frac{2}{\alpha}} - 1}{v^2 T_B - 2d_tv \cos \beta_t} \quad (19)$$

where

$$\Delta\phi_{t+T_B} = \arctan \left( \frac{vT_B \sin \beta_t}{d_t - vT_B \cos \beta_t} \right).$$

Equation (19) does not have a closed-form solution, so the value of  $T_B$  must be determined numerically. Note that  $T_B$  is defined for a given BS-UE distance, UE speed, and direction of travel. In practice, the motion of the UE can vary over time, so  $T_B$  will also change. If  $T_B$  is used as the basis for determining the beam selection instances and is initially set to a large value due to low UE speed, but shortly after that the UE increases its speed, beam misalignment will likely occur before the next beam selection time. To overcome this situation, we set  $T_s$  to the minimum value of  $T_B$ , *i.e.*,  $T_{B_{\min}}$ . This minimum value occurs when the UE is moving at the maximum speed  $v_{\max}$  and at a minimum distance  $d_{\min}$  from the BS. Moreover,  $T_{B_{\min}}$  also depends on  $\beta_t$ , which in turn depends on the HPBW of the BS antenna. While Fig. 4 illustrates a 2D scenario for clarity, it is important to note that the underlying assumptions and derivation of  $T_B$  remain applicable in a 3D environment, where the BS is situated at an elevated position relative to the UE.

To determine the value of  $\beta_t$  that results in  $T_{B_{\min}}$ , we first analyze the relationship between  $\beta_t$  and the HPBW. Fig. 5(a) depicts the rate of SNR reduction for a fixed distance ( $d = d_{\min}$ ) and speed ( $v = v_{\max}$ ) but for different  $\beta_t$  and HPBWs using (10). We set  $\xi = 0.5$  and  $\alpha = 3$ . We consider  $2 \times 2$ ,  $4 \times 4$ ,  $6 \times 6$ ,  $8 \times 8$ ,  $12 \times 12$ , and  $16 \times 16$  UPAs, with corresponding HPBWs of  $59.86^\circ$ ,  $26.28^\circ$ ,  $17.16^\circ$ ,  $12.78^\circ$ ,  $8.48^\circ$ , and  $6.36^\circ$ , respectively. A  $\beta_t$  of  $90^\circ$  means that the UE moves parallel to the planner array, so the SNR is mostly impacted by the antenna gain. On the other hand, a  $\beta_t$  of  $180^\circ$  means that the UE is moving perpendicular to the BS antenna. In this case, SNR changes are solely due to path loss. From Fig. 5(a), we observe that for a BS with large beamwidth, path-loss has more impact on the changes in SNR fluctuations than the impact of antenna gain. On the other hand, the impact of antenna gain is more pronounced for a narrower beam. Moreover, the value of  $\beta_t$  for which  $T_{B_{\min}}$  is obtained depends on the HPBW. For narrower beams (HPBW  $< 12.78^\circ$ ),  $\beta_t = 90^\circ$  results in the fastest rate of SNR reduction and thus  $T_{B_{\min}}$ . Fig. 5(b) shows how the SNR drops for different directions of travel when the BS is equipped with  $A_{\text{BS}} = 16 \times 16$  antenna (HPBW =  $6.36^\circ$ ).

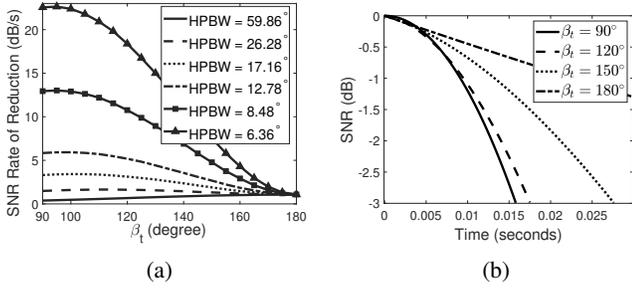


Fig. 5: (a) Rate of change in the SNR vs. direction of travel for different beamwidths; (b) time until SNR drops by 3 dB ( $\xi = 0.5$ ) for different values of  $\beta_t$  when HPBW is  $6.36^\circ$  and  $\alpha = 3$ .

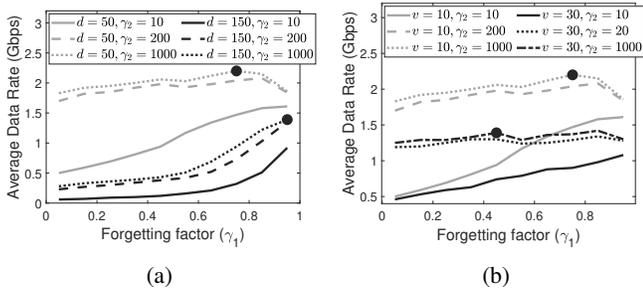


Fig. 6: Impact of  $\gamma_1$  and  $\gamma_2$  on GAMBIT's performance: (a) average data rate vs.  $\gamma_1$  for various distances, and (b) average data rate vs.  $\gamma_1$  for various speeds. The circles in (a) and (b) represents  $\gamma_1^*$ .

When  $\beta_t = 90^\circ$ , the SNR drops the fastest, by 3 dB within 15 ms.

So far, we assumed no sudden changes in the SNR (e.g., due to blockage). If a sudden blockage occurs, then according to (10) the beam coherence time will be defined by the time until the occurrence of the blockage, triggering an immediate rerun of GAMBIT. However, such recovery is not always desirable in real-world deployments due to hardware processing constraints and signaling overhead. To enhance its robustness under blockage conditions, GAMBIT uses the Top- $K$  beam selection strategy, which ensures that if the primary LOS beam fails due to blockage, the BS can seamlessly switch to an alternative NLOS beam from the  $K$ -best options. In cases of prolonged blockage (indicated by continuous NACKs), we allow GAMBIT to trigger an IA phase and rerun the algorithm, maintaining system reliability. In Section VII(B), we show the efficacy of such a strategy.

## V. PREDICTING THE OPTIMAL $\gamma_1$

In this sections, we first show how the performance of GAMBIT depends significantly on  $\gamma_1$ . We then study the impact of distance and speed on  $\gamma_1$ , and its relationship with SNR. Finally, we develop an MLP-based model to dynamically update the optimal value of  $\gamma_1$ .

### A. Dependence of GAMBIT on $\gamma_1$ and $\gamma_2$

As discussed in Section III,  $\gamma_1$  determines the *rate* at which previous information is forgotten, whereas  $\gamma_2$  specifies how

much the new information should be emphasized. Therefore, when executing GAMBIT, the BS must carefully select the values of both parameters.

Fig. 6(a) illustrates the impact of  $\gamma_1$  on the average data rate for three values of  $\gamma_2$  and two BS-UE distances. The UE speed is fixed at  $v = 10$  m/s. Fig. 6(b) demonstrates how the average data rate changes with  $\gamma_1$  at two UE speeds, 10 m/s and 30 m/s, when the BS-UE distance is fixed at 50 m. The two figures show that the average data rate varies significantly with  $\gamma_1$  and less with  $\gamma_2$ . After a certain value of  $\gamma_2$ , the effect of  $\gamma_2$  is more or less the same. From the figures, we see that for  $\gamma_2 > 200$ , the average data rate remains approximately the same for a given  $\gamma_1$ . Moreover, at the point where the UE achieves the highest data rate (the circles in the figures), the effect of a large  $\gamma_2$  on the data rate is insignificant.

### B. Correlation Between Optimal $\gamma_1$ and Rate of Change in the SNR

Because  $\gamma_1$  determines the significance of prior-beam quality information, the throughput-optimal value of  $\gamma_1$ , denoted as  $\gamma_1^*$ , depends on BS-UE distance and UE speed. If the UE is close to the BS, a small displacement will cause a large angular deviation. As a result, the BS is more likely to switch between beams. On the other hand, if the UE is far from the BS, it can be served by the same beam for a longer duration. The same intuition applies when the UE is at a fixed distance from the BS but is moving at a varying speed. Based on this intuition, we perform simulations to observe the effect of UE mobility on  $\gamma_1^*$ .

Fig. 7(a) and Fig. 7(b) depict  $\gamma_1^*$  and the corresponding data rate achieved for different UE distances and speeds, considering a  $16 \times 16$  UPA. From Fig. 7(a), we observe that at a fixed UE speed,  $\gamma_1^*$  increases with distance. A higher value of  $\gamma_1^*$  indicates that the variance of the posterior distribution is not changing much, and as a result, the rate of exploration is low. This justifies our previous assumption. The same argument can be used to justify the behavior of  $\gamma_1^*$  in Fig. 7(b), where the BS-UE distance is fixed but the speed is varied.

From the plots, it is evident that fixing the value of  $\gamma_1^*$  will negatively impact the performance of GAMBIT. Therefore, we need a way to update  $\gamma_1^*$  based on UE mobility. One way to do that is to design a predictor for  $\gamma_1^*$  based on the BS-UE distance and UE speed. However, obtaining distance and velocity directly from the received signal involves complicated signal processing. Indeed, we rely on the change in SNR of the received signal over time to adapt  $\gamma_1^*$ , given that SNR changes are also influenced by UE mobility. The change in the SNR at time  $t$  is defined as  $\Delta \text{SNR}_t = \text{SNR}_t - \text{SNR}_{t-1}$ , where  $\text{SNR}_t$  and  $\text{SNR}_{t-1}$  are the SNRs of packets received by the BS at times  $t$  and  $t-1$ , respectively. Fig. 7(c) and (d) depict how  $\text{SNR}_t$  varies with both distance and speed. Comparing Fig. 7(a) with Fig. 7(c), we observe that a higher value of  $\text{SNR}_t$  results in a lower value of  $\gamma_1^*$ , and vice versa. The same can be observed from Fig. 7(b) and Fig. 7(d).

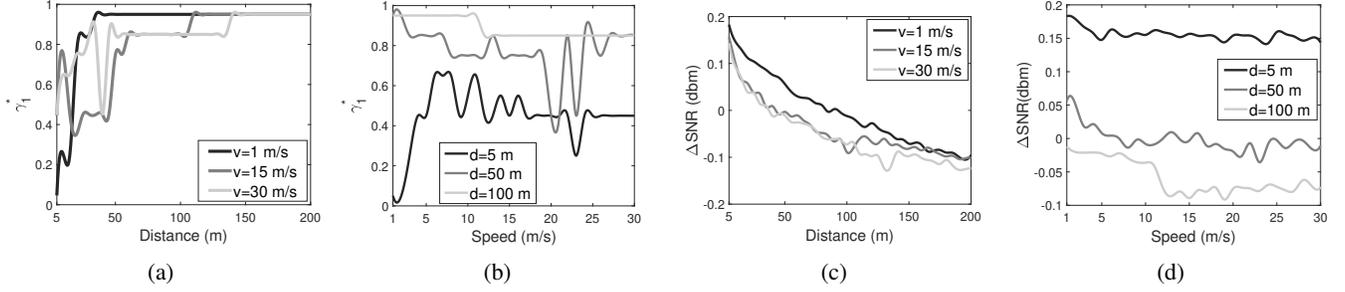


Fig. 7: (a)  $\gamma_1^*$  vs. BS-UE distance; (b)  $\gamma_1^*$  vs. UE speed; (c) change of SNR vs. BS-UE distance; and (d) change of SNR vs. UE speed.

### C. MLP-based Prediction of $\gamma_1^*$

Determining the relationship between  $\Delta \text{SNR}_t$  and  $\gamma_1^*$  is not trivial. For this reason, we employ an MLP model<sup>1</sup> to determine  $\gamma_1^*$  based on the changes in SNR values.

Let  $t_1, t_2, \dots$ , be the time instances at which GAMBIT is executed. At any of these instances, say  $t_n$ , let the input vector be denoted by  $X^{(t_n)} = X(\Delta \text{SNR}_{t-2}, \Delta \text{SNR}_{t-1}, \Delta \text{SNR}_t)$ , corresponding to the changes in the SNR of the past three execution instances. Let  $\gamma_1^*(t_n)$  denote the value of  $\gamma_1^*$  at any arbitrary time  $t_n$ . The prediction of  $\gamma_1^*(t_n)$  can be formulated as:

$$\gamma_1^*(t_n) = \mathcal{F}(X^{(t_n)}) \quad (20)$$

where  $\mathcal{F}(\cdot)$  defines the mapping from the input  $X^{(t_n)}$  to the output  $\gamma_1^*(t_n)$ . This mapping needs to be learned.

Our MLP predictor consists of three layers: an input layer, a hidden layer, and an output layer. The input and the hidden layers each consist of 32 hidden units, and the activation function is set to tanh. The output layer is constructed with ten hidden units corresponding to 10 different values of  $\gamma_1^*$  (from 0.05 to 0.95, with a step size of 0.1) and a softmax activation function. We use an Adam optimizer with a learning rate of 0.01 and select categorical cross entropy as the loss function.

To train the model, we use a dataset that consists of 48,000 data points, generated for 30 values of  $v$ , 40 values of  $d$ , 10 values of  $\gamma_1$ , and four values of  $\gamma_2$  (the simulation setup is detailed in Section VI). Because it takes a considerable time to simulate a MIMO mmWave channel and generate one data point, we obtain the dataset containing  $\gamma_1^*$  by considering only the value of  $\gamma_1$  that generated the highest data rate for  $\gamma_2 = 500$ . This resulted in a dataset of 1200 data points. We divide the dataset into 70% for training, 15% for validation, and 15% for testing. We train the MLP network offline by data collected under the random circular mobility model (RCM), while the predictions are performed online using random waypoint model (RWM). The two mobility models are explained in Section VI(A). The training batch size is set to 10. Training is completed when validation accuracy no longer increases with training epochs. The trained MLP network achieves an accuracy of 89.4%, precision of 0.855, recall of 0.894, and F1 score of 0.871.

<sup>1</sup>In [27], we used BS-UE distance and UE speed as inputs to an LSTM network to update  $\gamma_1^*$ .

---

### Algorithm 1 GAMBIT

---

```

1: for time instances  $t = t_1, t_2, \dots, T$  do
2:   Take samples:
3:   for  $i \in \mathcal{M}$  do
4:     Sample  $\mathbf{s}_{i,t} \sim \text{Dir}(\boldsymbol{\alpha}_{i,t})$ 
5:   end for
6:   Choose and apply action:
7:   Set of ranked arms,  $\mathcal{J}_t = \{\}$ 
8:   for  $i \in \mathcal{M}$  and  $i \notin \mathcal{J}_t$  do
9:      $\mathcal{J}_t \cup \arg \max_{i \in \mathcal{M}} \mathbf{s}_{i,t} \mathbf{w}^T$ 
10:  end for
11:  Select action  $\mathcal{J}_t$  and observe reward  $\mathbf{r}_{i,t}$ 
12:  Predict value of  $\gamma_1^*$ :
13:   $\gamma_1^* = \mathcal{F}(X(\Delta \text{SNR}_{t-2}, \Delta \text{SNR}_{t-1}, \Delta \text{SNR}_t))$ 
14:  Update distributions:
15:  for  $i \in \mathcal{M}$  do
16:    if  $i \in \mathcal{J}_t$  then
17:       $\boldsymbol{\alpha}_{i,t+1} \leftarrow \gamma_1^* \boldsymbol{\alpha}_{i,t} + \gamma_2 \mathbf{r}_{i,t}$ 
18:    else if  $i \notin \mathcal{J}_t$  and  $\gamma_1^* \|\boldsymbol{\alpha}_{i,t}\|_1 > 1$  then
19:       $\boldsymbol{\alpha}_{i,t+1} \leftarrow \gamma_1^* \boldsymbol{\alpha}_{i,t}$ 
20:    else
21:       $\boldsymbol{\alpha}_{i,t+1} \leftarrow \mathbf{1}$ 
22:    end if
23:  end for
24: end for

```

---

Algorithm 1 provides a procedure for implementing GAMBIT.

## VI. SIMULATION SETUP AND DATASET GENERATION

### A. UE Mobility Models

We consider two UE mobility models: the random waypoint model (RWM) and the random circular motion model (RCM).

Depicted in Fig. 8(a), RWM is a commonly used model that simulates the movement of UEs in an environment with random and unpredictable behavior. In this model, the UE is initially placed at a point  $p_1$  within the simulated region. The UE then selects at random a destination point  $p_2$  and a random speed  $v_1$ . It travels toward  $p_2$  at the specified speed in a straight line. Upon reaching  $p_2$ , the UE pauses for a fixed duration, referred to as the “pause time,” before selecting a new destination  $p_3$  and speed  $v_2$ . This process continues

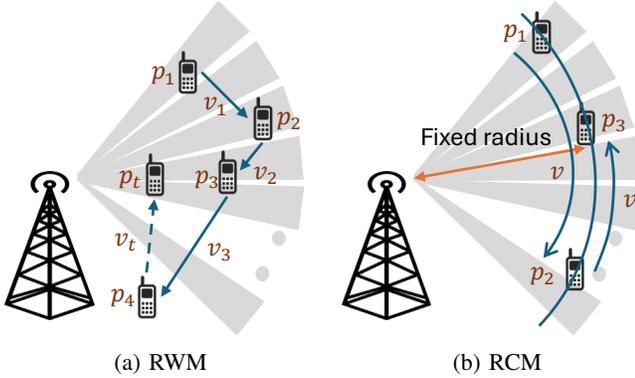


Fig. 8: Mobility models used in the simulations.

throughout the simulation, resulting in realistic yet variable movement patterns. RWM is used in our simulation to verify the performance of GAMBIT.

RCM (see Fig. 8 (b)) is a constrained version of the RWM, where the UE's movement is limited to a circular trajectory around the BS. The UE selects random points  $p_i$  along the circular path and moves towards them at a constant speed  $v$ . Upon reaching a point, the UE selects another random point along the circle and continues its movement randomly in either a clockwise or counterclockwise direction. RCM is used to study the relationship between the values of  $\gamma_1$  and  $\gamma_2$ , and UE mobility (i.e., distance and speed). Moreover, it is used to generate data for training the MLP predictor of  $\gamma_1^*$ .

### B. Channel Model

The channel model used in our simulation incorporates both LOS and NLOS components. The model captures path loss, shadowing, and small-scale effects. To generate a channel instance, we first determine the existence of an LOS path based on the probabilities defined in Equations (8a), (8b), and (8c) of [5], which were derived based on extensive measurements. Next, we account for large-scale effects, including path loss, by applying the channel parameters in Table I of [5].

To simulate multi-path components (MPCs), we position three point scatterers at random locations on the ellipsoid between the Tx and Rx, introducing both small-scale effects and three NLOS clusters in addition to the LOS cluster (see Fig. 9). At the Rx, each of the four clusters contains 32 rays (16 transmitted rays per cluster, with two Rx antenna elements). When arriving at a scatterer, each ray is characterized by unique Angle of Arrival (AoA) and Angle of Departure (AoD), which depend on the positions of the Tx element, scatterer, and Rx element.

The MIMO channel coefficients are computed based on the positions of the Tx, Rx, and scatterers. The small-scale channel gain for each path (ray) between a Tx antenna element and an Rx antenna element is sampled from a complex normal distribution with zero mean and unit variance. Beamforming vectors for both the Tx and Rx are derived from their respective azimuth angles, where only azimuthal tracking is performed. The Tx azimuth angle is determined by the given

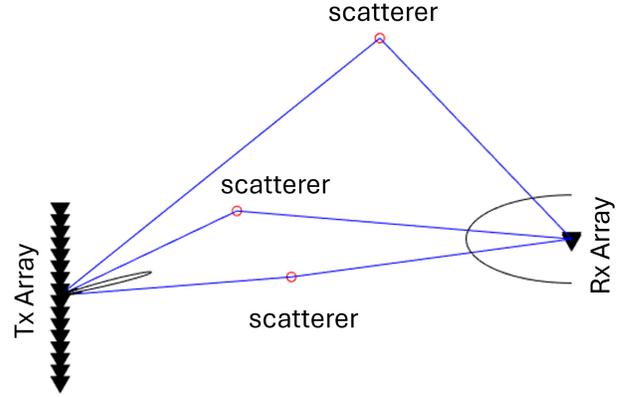


Fig. 9: Channel model with Tx and Rx arrays, and 3 point scatterers, illustrating the multi-path propagation and beamforming setup.

beam tracking algorithm. Using the beamforming vectors and the MIMO channel matrix, the beamforming gain is calculated. Together with the large-scale path loss, this gain is used to determine the RSS and subsequently the SNR.

The strength of the GAMBIT framework lies in its ability to adapt to varying channel conditions by updating posterior beliefs and modifying its policy based on observed data. While our study utilizes the above channel model, GAMBIT's flexibility makes it capable of accommodating more complex channel models, such as doubly-selective channels [28], which are characterized by time and frequency selectivity due to rapid mobility and wide bandwidth. Although the overall throughput may depend on the complexity and accuracy of the underlying channel model, GAMBIT's tracking capabilities are expected to remain consistent.

### C. Simulation Setup

A MATLAB program was developed to generate a dataset for analyzing the impact of  $\gamma_1$  and  $\gamma_2$  on GAMBIT's performance, training the MLP-based model, and evaluate the overall performance of GAMBIT. In the simulation setup, the UE moves according to mobility models described in Section VI(A). We set  $A_{BS}$  to  $16 \times 16$ ,  $A_{UE}$  to  $2 \times 2$ , and  $P_{Tx} = 30$  dBm. We run our simulations in the 28 GHz band. The Rx beam at the UE is kept the same. The BS performs beam tracking in the azimuthal plane with a scanning resolution of  $5^\circ$  and a beam scanning range  $\pm 30^\circ$  from the antenna's broadside. The BS-UE distance varies between 10 m to 200 m, and the UE speed varies between 1 m/s and 30 m/s. We rely on the CQI-to-MCS-to-SNR mapping table provided in [29]. We simulate different values of  $\gamma_1$  and  $\gamma_2$ , and collect UE mobility data (i.e., distance, speed, and SNR) at  $T_s$  instances. Other parameters used in the simulations are provided in Table I.

## VII. PERFORMANCE EVALUATION

To evaluate the performance of GAMBIT, we perform extensive simulations and we also run the algorithm on a publicly available experimental dataset. We compare GAMBIT with a

TABLE I: List of simulation parameters

Parameters	Value
Scenario	UMi
Center frequency	28 GHz
BS antennas ( $A_{BS}$ )	$16 \times 16$
UE antennas ( $A_{UE}$ )	$2 \times 2$
Sampling time ( $T_s$ )	1 ms, 15 ms, variable
UE speed ( $v$ )	1-30 m/s, in step of 1m/s
BS-to-UE distance ( $d$ )	10-200 m, in step of 5 m
Tx power ( $P_{Tx}$ )	30 dBm
No. of random scatters	3
$\gamma_1$	0.05-0.95, in step of 0.1
$\gamma_2$	10, 200, 500, 1000

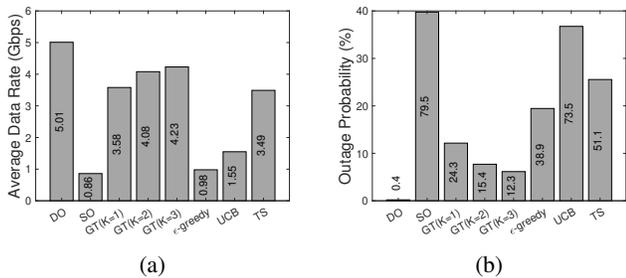


Fig. 10: Performance of various beam tracking algorithms (simulation dataset): (a) Average data rate, and (b) link outage probability for  $T_s = 15$  ms.

‘dynamic oracle’ (DO), a ‘static oracle’ (SO), and three state-of-the-art beam tracking approaches:  $\epsilon$ -greedy, UCB [20], and the TS algorithm in [22]. DO optimizes the expected reward at each step  $t$  by always selecting the best beam/MCS index. It has perfect knowledge that is updated instantaneously. On the other hand, SO represents the de facto beam management for 5G NR. In this case, the BS keeps exploiting the same beam until communication outage takes place. At that point, the algorithm reruns IA again [24].

#### A. Simulation Results

For testing and validation, we use the simulation parameters described in Section V except for the mobility model, where we test the various beam tracking approaches using RWM rather than RCM. Each simulation run is set to 90 seconds. For TS, we fix  $\gamma_1$  and  $\gamma_2$  at 0.3 and 12, respectively [22]. For GAMBIT, we let  $\gamma_1$  vary based on  $\Delta\text{SNR}_t$ , and set  $\gamma_2$  to 500. To study the effect of beam selection times, we consider three ways for selecting  $T_s$ : (i)  $T_s$  is varied based on the beam coherence time, as calculated using (19), (ii)  $T_s = T_{B_{\min}}$  (= 15 ms according to simulation parameters), and (iii)  $T_s = 1$  ms  $< T_{B_{\min}}$ .

Finally, to show the impact of the number of simultaneously selected beams in GAMBIT, we run GAMBIT for three different values of  $K$ , i.e.,  $K = 1, 2$ , and  $3$ . When  $K = 1$ , we only select the best beam, i.e., the “leader,” and there are no “scout” beams. For  $K = 2$  and  $3$ , there are 1 and 2 “scout” beams, respectively.

1) *Comparison With State-of-the-art Algorithms:* We compare GAMBIT with DO, SO,  $\epsilon$ -greedy, UCB, and TS algo-

gorithms. Here, we set  $T_s$  to 15 ms. Moreover, we run the  $\epsilon$ -greedy algorithm for different values of  $\epsilon$  and report the results for  $\epsilon = 0.1$  (for which we obtained the best results). We use the notation GT for GAMBIT for ease of representation in the figures.

Fig. 10(a) and (b) depict, respectively, the average data rate and link outage probability obtained under various algorithms. Outage occurs when the selected beam and/or MCS index is insufficient to meet the link budget. This could happen due to gradual or sudden changes (blockage) in the SNR. We define the outage probability as the ratio of the total outage instances to the total number of beam selection instances. As shown in the results, GAMBIT performs better than  $\epsilon$ -greedy, UCB, and TS in both metrics, even when ( $K = 1$ ). For the average data rate, GAMBIT achieves 3.58 Gbps, 4.08 Gbps, and 4.23 Gbps for  $K = 1$ ,  $K = 2$ , and  $K = 3$ , respectively, significantly higher than  $\epsilon$ -greedy (0.98 Gbps), UCB (1.55 Gbps), and TS (3.49 Gbps). Similarly, for link outage probability, GAMBIT demonstrates superior reliability with 24.3%, 15.4%, and 12.3% for  $K = 1$ ,  $K = 2$ , and  $K = 3$ , respectively, which are notably lower than  $\epsilon$ -greedy (38.9%), UCB (73.5%), and TS (51.1%).

2) *Impact of Beam Selection Times:* In Section III, we derived an upper bound on  $T_s$  and set it to  $T_{B_{\min}}$  (= 15ms in our case). We argued that GAMBIT must be run within this time, i.e.,  $T_s \leq T_{B_{\min}}$ , to ensure that previously observed data do not become obsolete. Fig. 11(a) and 11(b) depict the average data rate and outage probability for GAMBIT under different values of  $T_s$ . The figures also compare GAMBIT with TS, DO, and SO, respectively. We intentionally omit  $\epsilon$ -greedy and UCB algorithms in the following comparison as they performed poorly compared to other algorithms.

For a given value of  $K$ , we observe that the performance of GAMBIT depends on the beam selection time  $T_s$ . When  $T_s$  is set to vary with the beam coherence time, the performance of GAMBIT in terms of both link throughput and outage probability degrades severely compared to when  $T_s \leq T_{B_{\min}}$ . When  $T_s < T_{B_{\min}}$ , the performance improves compared to when  $T_s = T_{B_{\min}}$ . This is because running the beam tracking schemes frequently will result in better contextual information regarding the wireless environment. Moreover, we observe that the performance gain is more obvious in terms of reducing the number of outages than increasing the average data rate. However, setting  $T_s < T_{B_{\min}}$  comes with additional computational and communication overhead.

3) *Robustness to Outages:* We also obtain how long it takes a beam tracking algorithm to realign the BS-UE beams following an outage. Fig. 11(c) and 11(d) depict the CDF of the outage duration (up to 20 slots) for various algorithms when  $T_s = 15$  ms and  $T_s = 1$  ms, respectively. From Fig. 11(c), we observe that with a probability  $> 0.9$ , the outage will approximately last for 8, 9, and 10 slots for GT( $K = 1$ ), GT( $K = 2$ ), and GT( $K = 3$ ), respectively, whereas for TS, it lasts for seven slots, on average. Similarly, from Fig. 11(d), we observe that with a probability  $> 0.9$ , the outage will be approximately three slots for GT( $K = 1$ ), GT( $K = 2$ ), and GT( $K = 3$ ), respectively, and that for TS is 10 slots. This indicates that GAMBIT can quickly realign the Tx-Rx

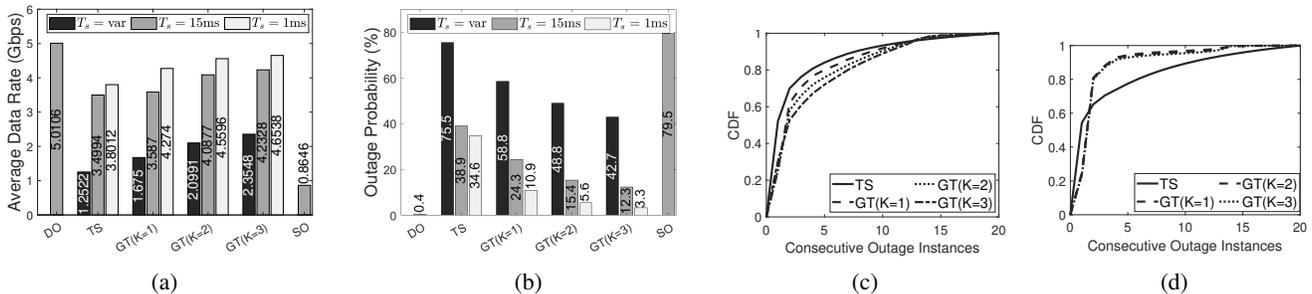


Fig. 11: Performance of various beam tracking algorithms (simulation dataset): (a) Average data rate, (b) link outage probability, (c) CDF of outage duration for  $T_s = 15\text{ms}$ , and (d) CDF of outage duration for  $T_s = 1\text{ms}$ .

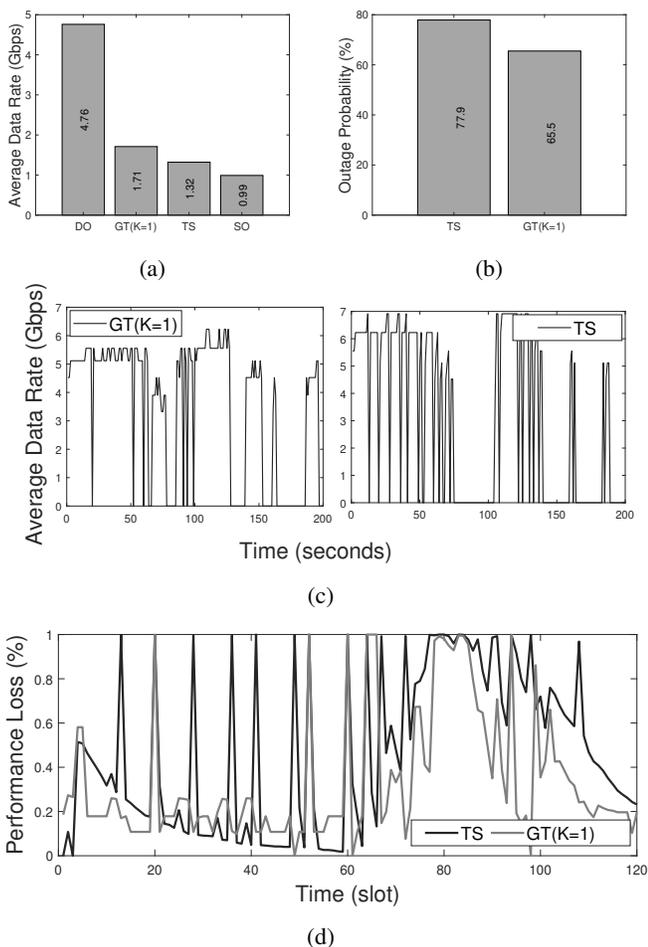


Fig. 12: Performance of various beam tracking algorithms (Lumos5G dataset): (a) average data rate, (b) link outage probability, (c) instantaneous data rate, and (d) instantaneous performance loss.

beams in case of a link outage, demonstrating its robustness. Moreover, there will be fewer IA cycles in GAMBIT compared to TS, thus extending the time between consecutive IA cycles and reducing overall system overhead.

## B. Experimental Results

We also study the effectiveness of GAMBIT using two experimental datasets, obtained at 28 GHz and 130 GHz.

### C. 28 GHz Experimental Dataset

The first dataset, called Lumos5G [30], is a publicly available dataset from the University of Minnesota (UMN), USA. It consists of almost 68,000 traces of 5G data collected at 28 GHz around the U.S. Bank Stadium in Minneapolis's downtown. The measurements were obtained along a 1300-meter loop that covers roads, railroad crossings, restaurants, coffee shops, and recreational outdoor parks. The collected data includes the UE's geolocation, its speed, 5G BS tower ID, 5G synchronization signal (SS) measurements, e.g., SS-RSRP, SS-RSRQ, and SS-SINR, throughput, etc., all sampled and logged every second. The UE speed varied from 0 m/s to 14 m/s. The dataset includes diverse conditions with LOS, NLOS, and blockages, allowing us to evaluate GAMBIT under realistic scenarios.

We select the BS with 'tower\_id' 16 from the dataset and calculate the BS-UE distance from the provided UE and BS locations. Since no information regarding the characteristics of the antenna array or beamwidth is available, we divided the provided UE locations so that at any given time, the UE is within the coverage of at least two beams. This represents an extreme scenario where the beam is either very narrow or the UE is moving close to the highest speed supported by the system. We run various beam tracking algorithms using this dataset. Here,  $T_s$  equals the sampling time, i.e., 1 second, and  $K = 1$ .

Fig. 12(a) depicts the average throughput under different beam tracking algorithms. GAMBIT achieves 29.104% and 71.2% higher throughput than TS and static oracle, respectively.

In Fig. 12(b), we compare GAMBIT and TS in terms of the outage probability. GAMBIT has an outage probability of 65.5%, which is 12.38% less than that of TS (77.88%).

The instantaneous data rates achieved by GAMBIT and TS are depicted in Fig. 12(c). GAMBIT demonstrates the ability to recover from sudden outages caused by blockages. For instance, extended blockages occur between 120–140 seconds and 160–180 seconds, during which communications temporarily halt. GAMBIT successfully restores communication



(a)

Fig. 13: TeraNova experimental setup used for performance evaluation at 130 GHz.

after these interruptions, as evidenced by the gradual recovery in the average data rate. This highlights the framework's adaptability in handling challenging real-world scenarios. In Fig. 12(d) we depict the *instantaneous performance loss*, defined as the percentage decrease in the data rate when the optimal beam or MCS index is not selected. From Fig. 12(c) and Fig. 12(d), we observe that the TS algorithm undergoes a lot of continuous outages and performance loss compared to GAMBIT, which explains the results in Fig. 12(a) and Fig. 12(b), respectively.

#### D. 130 GHz Experimental Dataset

We conducted measurements at 130 GHz using the TeraNova [31] platform at Northeastern University. In our experimental setup, illustrated in Fig. 13, we employ a 21 dBi horn antenna with a  $13^\circ$  3-dB beamwidth for both the Tx and Rx. The transmit power was set to 13 dBm, and the channel bandwidth was set to 20 GHz.

This experiment took place indoor in an atrium with various objects, e.g, pillars, chairs, tables, metallic objects, and more. It was challenging to precisely quantify the number of metallic objects within the setup and, more importantly, how many of them acted as reflectors. Notably, there were no blockages between the Tx and Rx, ensuring a consistent LOS path throughout the experiment. The Rx was positioned within an arc 3 meters away from the Tx. This position was changed in discrete steps, ranging from  $-16^\circ$  to  $+16^\circ$  with a  $2^\circ$  resolution. Meanwhile, the Tx beam was scanned from  $-90^\circ$  to  $+88^\circ$  with a  $2^\circ$  resolution. We gathered data on the SNR for different Tx beam directions and Rx locations.

Fig. 14(a) depicts the average link throughput for DO, SO,  $GT(K=1)$ ,  $GT(K=2)$ ,  $GT(K=3)$ , and TS algorithms. The figure clearly shows that GAMBIT outperforms TS. Fig. 14(b) depicts the outage probabilities for both GAMBIT and TS. GAMBIT consistently exhibits a lower outage probability than TS.

#### E. Impact of Beam Squint

Comment 5: In the sub-THz frequency range, beam squint poses a potential challenge due to the wide bandwidths

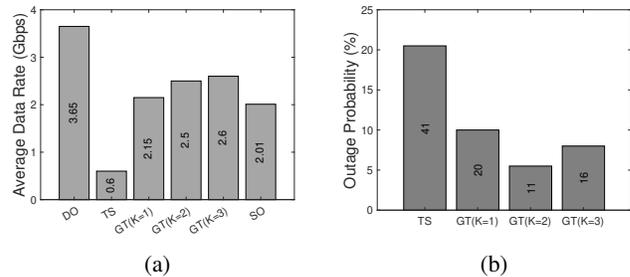


Fig. 14: Performance of various beam tracking algorithms (TeraNova dataset at 130 GHz): (a) average link throughput, and (b) outage probability.

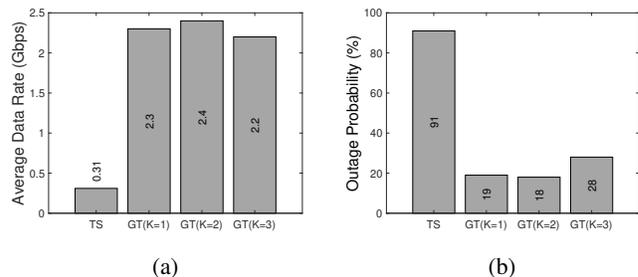


Fig. 15: Performance of various beam tracking algorithms under the effect of beam squint (TeraNova dataset at 130 GHz): (a) average link throughput, and (b) outage probability.

involved. Because horn antennas do not exhibit frequency-dependent beam steering, beam squint does not show in our experimentation. However, to evaluate its potential impact on GAMBIT, we conducted simulations incorporating beam squint. The effective beam direction due to beam squint  $\phi_{\text{squint}}$  was computed as [32]:

$$\phi_{\text{squint}}(f) = \sin^{-1} \left( \frac{f_c}{f} \sin \phi \right) \quad (21)$$

where,  $f_c = 130$  GHz is the center frequency,  $f$  is the operating frequency, and  $\phi$  is the intended beam angle at  $f_c$ . Given the 20 GHz bandwidth, we selected  $f = 120$  GHz to simulate the worst-case squint effect.

Figure 15(a) illustrates the average link throughput for the TS,  $GT(K=1)$ ,  $GT(K=2)$ , and  $GT(K=3)$  algorithms when beam squint is introduced. Figure 15(b) presents the corresponding outage probabilities for TS and GAMBIT.

As observed, with beam squint, the average data rate decreases from 0.6 to 0.31 Gbps for TS, from 2.5 to 2.4 Gbps for  $GT(K=2)$ , and from 2.6 to 2.2 Gbps for  $GT(K=3)$ . On the other hand, the average data rate increases from 2.15 to 2.3 Gbps for  $GT(K=1)$ . Similarly, the outage probability increases from 41% to 91% for TS, from 11% to 18% for  $GT(K=2)$ , and from 16% to 26% for  $GT(K=3)$ , but slightly decreases from 20% to 19% for  $GT(K=1)$ . These results indicate that beam squint generally degrades system performance by lowering throughput and increasing outages, but GAMBIT still consistently outperforms TS under such impairments.

To mitigate beam squint in practical deployments, true-time-delay (TTD) controllers have been proposed as an alternative to conventional phase-based beamforming [33]. Unlike phase shifters, which introduce frequency-dependent phase shifts, TTD-based approaches compensate for these variations across wide bandwidths, effectively mitigating beam squint. Although such arrays are not available in our current setup, they represent a promising direction for future high-bandwidth sub-THz beamforming systems.

## VIII. CONCLUSIONS

This paper introduced GAMBIT, a novel framework for intelligent beam tracking and rate adaptation in mmWave and sub-THz systems. GAMBIT leverages a restless multi-armed bandit approach with Top-K Adaptive Thompson Sampling to jointly optimize beam and MCS selection. This approach increases the exploration rate to adapt to dynamic channel conditions while maintaining a fixed exploitation rate. We also introduced a novel approach to determine the upper bound for beam selection time based on beam coherence time, ensuring that beam information remains up-to-date. Furthermore, an MLP was developed to dynamically predict the controlling parameters of GAMBIT (e.g.,  $\gamma_1$ ) based on changes in SNR, enhancing its adaptability to varying channel conditions.

The performance of GAMBIT was validated through extensive simulations and experimental evaluations at 28 GHz and 130 GHz. Results demonstrated that GAMBIT significantly outperforms state-of-the-art beam tracking algorithms, including  $\epsilon$ -greedy, UCB, and TS. Specifically, GAMBIT achieved up to 71.2% higher throughput compared to default 5G beam management schemes, reduced outage probabilities by over 12%, and delivered an average throughput of 4.23 Gbps with  $K = 3$ , surpassing TS (3.49 Gbps) and other baseline methods. These findings highlight GAMBIT's potential as a robust and efficient solution for beam tracking and rate adaptation in next-generation wireless systems. Future work could extend GAMBIT to support enhanced mobile broadband by incorporating multi-beam transmission, where multiple beams constructively combine at the UE to further increase overall throughput.

## REFERENCES

- [1] Q. Wu, J. Xu, Y. Zeng, D. W. K. Ng, N. Al-Dhahir, R. Schober, and A. L. Swindlehurst, "A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 2912–2945, 2021.
- [2] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, 2014.
- [3] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 173–196, 2018.
- [4] M. Krunz, I. Aykin, S. Sarkar, and B. Akgun, "Online reinforcement learning for beam tracking and rate adaptation in millimeter-wave systems," *IEEE Transactions on Mobile Computing*, pp. 1–16, 2023.
- [5] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, 2014.
- [6] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, 2014.
- [7] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.
- [8] M. Qurratulain Khan, A. Gaber, P. Schulz, and G. Fettweis, "Machine learning for millimeter wave and terahertz beam management: A survey and open challenges," *IEEE Access*, vol. 11, pp. 11 880–11 902, 2023.
- [9] S. G. Larew and D. J. Love, "Adaptive beam tracking with the unscented kalman filter for millimeter wave communication," *IEEE Signal Processing Letters*, vol. 26, no. 11, pp. 1658–1662, 2019.
- [10] S.-H. Hyun, J. Song, K. Kim, J.-H. Lee, and S.-C. Kim, "Adaptive beam design for V2I communications using vehicle tracking with extended kalman filter," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 489–502, 2022.
- [11] J. Lim, H.-M. Park, and D. Hong, "Beam tracking under highly non-linear mobile millimeter-wave channel," *IEEE Communications Letters*, vol. 23, no. 3, pp. 450–453, 2019.
- [12] J. Kang, I. Orikumhi, Y.-M. Park, and S. Kim, "A millimeter wave beam tracking in vehicular scenario via particle filter," in *Proc. of the International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, 2018, pp. 234–238.
- [13] M. Polese, F. Restuccia, and T. Melodia, "DeepBeam: Deep waveform learning for coordination-free beam management in mmwave networks," in *Proc. of the Twenty-Second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc)*, 2021, p. 61–70.
- [14] T. S. Cousik, V. K. Shah, T. Erpek, Y. E. Sagduyu, and J. H. Reed, "Deep learning for fast and reliable initial access in AI-Driven 6G mm Wave networks," *IEEE Transactions on Network Science and Engineering*, pp. 1–12, 2022.
- [15] S. H. A. Shah and S. Rangan, "Multi-cell multi-beam prediction using auto-encoder LSTM for mmWave systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10 366–10 380, 2022.
- [16] R. Wang, O. Onireti, L. Zhang, M. A. Imran, G. Ren, J. Qiu, and T. Tian, "Reinforcement learning method for beam management in millimeter-wave networks," in *UK/ China Emerging Technologies (UCET)*, 2019, pp. 1–4.
- [17] S. Kim, G. Kwon, and H. Park, "Q-learning-based low complexity beam tracking for mmWave beamforming system," in *Proc. of the International Conference on Information and Communication Technology Convergence (ICTC)*, 2020, pp. 1451–1455.
- [18] J. Zhang, Y. Huang, J. Wang, X. You, and C. Masouros, "Intelligent interactive beam training for millimeter wave communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 2034–2048, 2021.
- [19] M. Hashemi, A. Sabharwal, C. Emre Koksall, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," in *Proc. of the IEEE INFOCOM Conference*, 2018, pp. 2393–2401.
- [20] J. Zhang, Y. Huang, Y. Zhou, and X. You, "Beam alignment and tracking for millimeter wave communications via bandit learning," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5519–5533, 2020.
- [21] I. Aykin, B. Akgun, M. Feng, and M. Krunz, "MAMBA: A multi-armed bandit framework for beam tracking in millimeter-wave systems," in *Proc. of the IEEE INFOCOM Conference*, Toronto, Canada, Jul. 2020.
- [22] I. Nasim, P. Skrimponis, A. S. Ibrahim, S. Rangan, and I. Seskar, "Reinforcement learning of millimeter wave beamforming tracking over COSMOS platform," in *Proc. of the 16th ACM Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization*, 2022, p. 40–44.
- [23] L. Jiang and H. Jafarkhani, "Multi-user analog beamforming in millimeter wave mimo systems based on path angle information," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 608–619, 2018.
- [24] Y.-N. R. Li, B. Gao, X. Zhang, and K. Huang, "Beam management in millimeter-wave communications for 5G and beyond," *IEEE Access*, vol. 8, pp. 13 282–13 293, 2020.
- [25] V. Va, J. Choi, and R. W. Heath, "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5014–5029, 2016.
- [26] G. Yang, J. Du, and M. Xiao, "Analysis on 60 GHz wireless communications with beamwidth-dependent misalignment," *arXiv preprint arXiv:1611.07867*, 2016.
- [27] S. Sarkar, M. Krunz, I. Aykin, and D. Manzi, "Machine learning for robust beam tracking in mobile millimeter-wave systems," in *IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

- [28] S. Gao, X. Cheng, and L. Yang, "Estimating doubly-selective channels for hybrid mmWave massive MIMO systems: A doubly-sparse approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5703–5715, 2020.
- [29] A. K. Thyagarajan, P. Balasubramanian, D. Vydeki, and M. Karthik, "SNR-CQI mapping for 5G downlink network," in *Proc. of the IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, 2021, pp. 173–177.
- [30] A. Narayanan, E. Ramadan, R. Mehta, X. Hu, Q. Liu, R. A. K. Fezeu, U. K. Dayalan, S. Verma, P. Ji, T. Li, F. Qian, and Z.-L. Zhang, "Lumos5G: Mapping and predicting commercial mmwave 5G throughput," in *Proc. of the ACM Internet Measurement Conference*, 2020, pp. 176–193.
- [31] "A versatile experimental testbed for ultrabroadband communication networks above 100 GHz," *Computer Networks*, vol. 193, p. 108092, 2021.
- [32] H. Luo, F. Gao, W. Yuan, and S. Zhang, "Beam squint assisted user localization in near-field integrated sensing and communications systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 5, pp. 4504–4517, 2024.
- [33] T. Mewes, W. Rave, and G. Fettweis, "Beam squint reduction by combining phased subarrays with true time delays in baseband," *IEEE Wireless Communications Letters*, pp. 1–1, 2025.