# Leveraging Machine Learning for Millimeter Wave Beamforming in Beyond 5G Networks

Basem M. ElHalawany, Sherief Hashima, Kohei Hatano, Kaishun Wu, and Ehab Mahmoud Mohamed

***Abstract*—Millimeter wave (mmWave) communication has attracted considerable attention as a key technology for the next-generation wireless communications thanks to its exceptional advantages. MmWave leads the way to achieve a high transmission quality with directed narrow beams from source to multiple destinations by adopting different antenna beamforming (BF) techniques, which have a pivotal role in establishing and maintaining robust links. However, realizing such BF gains in practice requires overcoming several challenges, such as severe signal deterioration, hardware constraints, and design complexity. The elevated complexity of configuring mmWave BF vectors encourages researchers to leverage relevant machine learning (ML) techniques for better BF configurations deployment in 5G and beyond. In this article, we summarize mmWave BF strategies employed for future wireless networks. Then, we provide a comprehensive overview of ML techniques plus its applications and promising contributions toward efficient mmWave BF deployment. Furthermore, we discuss mmWave BF's future research directions and challenges. Finally, we discuss a single and concurrent mmWave BF case study by applying multiarmed bandit to confirm the superiority of ML-based methods over conventional ones.**

***Index Terms*—Beamforming training (BT), deep learning, machine learning (ML), millimeter wave (mmWave), multiarmed bandit (MAB).**

## I. Introduction

THE urgent need to deliver higher throughput, lower latency, massive connectivity, and reliable services has encouraged academia and industry to investigate different technologies and spectrum regions for the next-generation wireless networks. Although the fifth-generation (5G) of mobile networks is still in planning stage, many research efforts have been explored to propose hybrid models, where different technologies coexist in a heterogeneous architecture at which billions of devices are communicating via the same network. Several technologies have recently received profound attention and can be considered pillars for the 5G and beyond 5G (B5G) networks, including massive multiple-input–multiple-output (massive MIMO) antenna arrays, beamforming (BF), and nonorthogonal multiple access (NOMA), etc. On the other hand, to support such massive connectivity, the capacity of the cellular network needs to be extensively boosted (1000 fold) by efficiently using any available spectrum regions while optimizing the system energy efficiency, which constitutes a substantial part of the operating cost [1].

Due to its large swath of available spectrum, millimeter wave (mmWave) band, 30–300 GHz, is considered one of the key components of 5G and B5G. Yet, mmWave channel owns bad propagation characteristics due to its high operating frequencies, vulnerability to path blocking, and oxygen absorption [2]. Consequently, massive MIMO arrays are needed to mitigate such severe channel impairments by leveraging the high BF gain to obtain sufficient signal-to-noise ratio (SNR) [1]. However, traditional sub-6 GHz MIMO digital precoding techniques are not feasible for mmWave, since it requires an incredible number of radio-frequency (RF) chains, digital-to-analog (D/A) converters, and analog-to-digital (A/D) converters with excessive power consumption. Instead, mmWave BF can be done either using analog beamformers, where a single RF chain is shared by all antenna elements or using hybrid analog/digital beamformers. In this context, hybrid beamformers provide an efficient tradeoff between the low-complexity/limited-performance analog BF and high-complexity/good-performance fully digital BF using lower RF chains than the number of antenna elements, as shown in Fig. 1 [1]. However, both analog-only and hybrid BF are employed to realize multiuser and multistream transmissions in indoor and outdoor mmWave communications. Consequently, a variety of BF training (BT) techniques were suggested in literature to determine the best mmWave transmit-receive beam pair for different considerations, such as improving signal-to-interference-plus-noise ratio (SINR) or security [3].

On the other hand, the recent progress in machine learning (ML) has provided momentum to employ it in a large number of applications in wireless networks, including cells association, spectrum management, intelligent resources allocation [4], routing [5], recovery of channel state information (CSI), modulation, detection, and device-to-device communication [6], beside its

Manuscript received August 29, 2020; revised March 3, 2021, April 18, 2021, and May 31, 2021; accepted June 5, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant U2001207 and Grant 61872248, in part by Guangdong NSF 2017A030312008, in part by Shenzhen Science and Technology Foundation under Grant ZDSYS20190902092853047, in part by the Project of DEGP under Grant 2019KCXTD005, and in part by the Guangdong "Pearl River Talent Recruitment Program" under Grant 2019ZT08X603. *(Corresponding author: Kaishun Wu.)*

Basem M. ElHalawany is with the School of Computer Science, Shenzhen University, Shenzhen 518060, China, and also with the Benha University, Cairo 11241, Egypt (e-mail: basem.mamdoh@feng.bu.edu.eg).

Sherief Hashima is with the RIKEN AIP, Kyushu, Saitama, Japan, and also with the Egyptian Atomic Energy Authority, Cairo 13759, Egypt (e-mail: Sherief.hashima@riken.jp).

Kohei Hatano is with the RIKEN AIP, Kyushu, Saitama, Japan, and also with the Kyushu University, Fukuoka 819-0395, Japan (e-mail: Kohei.hatano@riken.jp).

Kaishun Wu is with the School of Computer Science, Shenzhen University, Shenzhen 518060, China, and also with the Guangzhou HKUST Fok Ying Tung Research Institute, Guangzhou 511458, China (e-mail: wu@szu.edu.cn).

Ehab Mahmoud Mohamed is with the College of Engineering, Prince Sattam Bin Abdulaziz University, Wadi Al Dwaser 11991, Saudi Arabia, and also with the Faculty of Engineering, Aswan University, Aswan 81542, Egypt (e-mail: ehab_mahmoud@aswu.edu.eg).

Digital Object Identifier 10.1109/JSYST.2021.3089536

TABLE I
NOMENCLATURE

| | | | | | |
|---|---|---|---|---|---|
| mmWave | Millimeter-wave | BF | Beamforming | D2D | Device to device |
| 5G | Fifth generation | B5G | Beyond fifth generation | 6G | Sixth generation |
| BS | Base station | ML | Machine-learning | DL | Deep Learning |
| WiGig | Wireless Gigabit | Li-Fi | Light-Fidelity | UAV | Unmanned aerial vehicles |
| ANN | Artificial Neural network | DNN | Deep NN | DCNN | Deep convolution NN |
| DRL | Deep RL | FNN | Forward neural network | MIMO | Multiple-input multiple-output |
| RNN | Recurrent NN | NOMA | Non-orthogonal multiple access | RL | Reinforcement-Learning |
| V2X | Vehicle to everything | KNN | K-nearest neighbor | COMP | Coordinated multi point |
| AWV | Antenna weight vectors | GMM | Gaussian mixture model | AP | Access point |
| SINR | Signal to interference plus noise ratio | MAB | Multi armed bandit | MP-MAB | Multiplayer MAB |
| UCB | Upper confidence bound | A/D | Analog-to-digital | MP-UCB | Multiplayer UCB |
| LOS | Line of sight | NLOS | Non line of sight | BT | Beamforming training |
| SVM | Support vector machine | FL | Federated learning | MDP | Markov decision process |
| EX | Exhaustive search | AoA | Angle of Arrival | AoD | Angle of departure |
| SLS | Sector level sweep | MIDC | Multiple sector ID capture | AWGN | Additive white Gaussian noise |

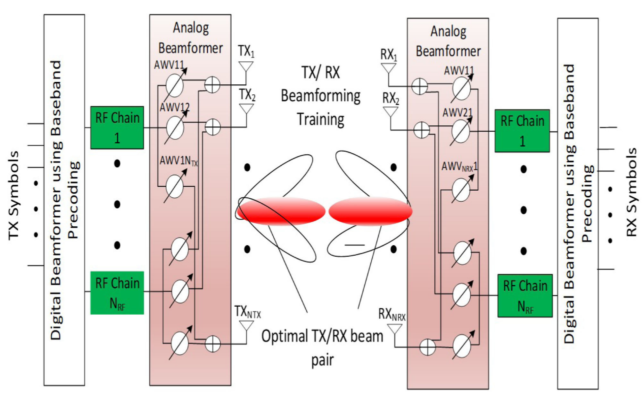


Fig. 1. Structure of hybrid analog/digital BF for mmWave communications.

potential in other paradigms, such as load-balancing in software-defined networks [7], underwater communication [8], and robot vision and image processing [9]. Additionally, leveraging ML for developing efficient mmWave BF has become a hot research area, where researchers have investigated several ML algorithms for solving BF-related technical problems [10]–[12]. These include beam pairs selection [10], [13]–[17], location prediction, angle-of-arrival (AoA) estimation [10], [13], [18], [19], BF scheme selection [11], and mmWave beam alignment [19]–[21], etc. The expected massive data traffic and number of users increment over the next few years encourages researchers to propose several solutions to these challenges. In highly dynamic mmWave scenarios, conventional mmWave BF techniques will be inappropriate as they require huge training overhead associated with adjusting large array BF vectors. One promising solution is to learn the surrounding environment via utilizing a suitable ML technique. Thus, just based on some selected environmental signatures, fast ML-based BF and beam tracking algorithms can be developed. In this article, motivated by providing intelligent self-decision-making communication networks, we investigate the ML-based research efforts for improved BF techniques. Consequently, we highlighted a brief discussion of the BT optimization problem in mmWave channels in Section II, followed by a summary of mmWave BT strategies in Section III. A short overview of the classification of ML techniques is presented in Section IV. Section V summarizes several ML applications with promising contributions toward efficient mmWave deployment. Additionally, we shed light on some open issues and challenges for future research in Section VI. In Section VII, we elaborated in explaining a case study to show the superiority of ML-based methods in single and concurrent mmWave BF using single-player and multiplayer multiarmed bandit (MAB) schemes. Finally, Section VIII concludes this article. Table I shows a list of the abbreviations used throughout this article.

## II. MMWAVE CHANNEL MODEL AND BT OPTIMIZATION PROBLEM

As mmWave channel has low number of scatters, it can be represented by adopting $L$ geometric scatters as [22]

$$\mathbf{H} = \sqrt{\frac{N_{\mathrm{TX}} N_{\mathrm{RX}}}{\alpha}} \sum_{\ell=1}^{L} \tau_\ell A_{\mathrm{RX}}(\phi_\ell) A_{\mathrm{TX}}^{\mathcal{H}}(\theta_\ell) \tag{1}$$

where $1 \leq \ell \leq L$ is the path index and $L$ is the total number of paths, and $N_{\mathrm{TX}}$ and $N_{\mathrm{RX}}$ are the number of TX and RX antenna elements shown in Fig. 1. $\alpha$ is the distance-dependent average path loss, and $\tau_\ell$ is the $\ell$th path complex gain. $A_{\mathrm{RX}}(\phi_\ell)$ and $A_{\mathrm{TX}}^{\mathcal{H}}(\theta_\ell)$ denote the antenna array response vectors of the RX and TX, respectively, and $\mathcal{H}$ denotes the Hermitian transpose. $\phi_\ell$ and $\theta_\ell$ denote the AoA and angle of departure (AoD) of $\ell$th channel path, respectively. $A_{\mathrm{TX}}(\theta_\ell)$ can be represented as

$$\mathbf{A}_{\mathrm{TX}}(\theta_\ell) = [1, e^{j\frac{2\pi}{\lambda} d\sin(\theta_\ell)}, .., e^{j(N_{\mathrm{TX}}-1)\frac{2\pi}{\lambda} d\sin(\theta_\ell)}]^T \tag{2}$$

where $T$ means transpose, $d$ is antenna elements spacing, and $\lambda$ is the signal wavelength. The same equation can be applied for evaluating $A_{\mathrm{RX}}(\phi_\ell)$ except that TX is replaced by RX and $\theta_\ell$ is replaced by $\phi_\ell$. By only considering one TX/RX RF chain in (1) to emphasize the operation of the analog beamformer responsible for the BT process, the signal at the RX side can be expressed as

$$y = \mathbf{F}_{\mathrm{RX}}^{\mathcal{H}}(:, b_{\mathrm{rx}})\, \mathbf{H} \mathbf{F}_{\mathrm{TX}}(:, b_{\mathrm{tx}})\, x + \mathbf{F}_{\mathrm{RX}}^{\mathcal{H}}(:, b_{\mathrm{rx}})\, \mathbf{n} \tag{3}$$

where $x$ is the TX signal, and $\mathbf{n}$ is the additive white Gaussian noise of length $N_{\mathrm{RX}} \times 1$. $\mathbf{F}_{\mathrm{TX}}(:, b_{\mathrm{tx}})$ and $\mathbf{F}_{\mathrm{RX}}(:, b_{\mathrm{rx}})$ are the antenna weight vectors (AWVs) of the TX and RX of lengths $N_{\mathrm{TX}} \times 1$ and $N_{\mathrm{RX}} \times 1$ in the directions of $b_{tx}$ and $b_{rx}$ beams, i.e., corresponding to columns $b_{\mathrm{tx}}$ and $b_{\mathrm{rx}}$ in $\mathbf{F}_{\mathrm{TX}}$ and $\mathbf{F}_{\mathrm{RX}}$, respectively. The target of the BT optimization problem is to find out the best beam pair $(b_{\mathrm{tx}}^*, b_{\mathrm{rx}}^*)$ maximizing the achievable

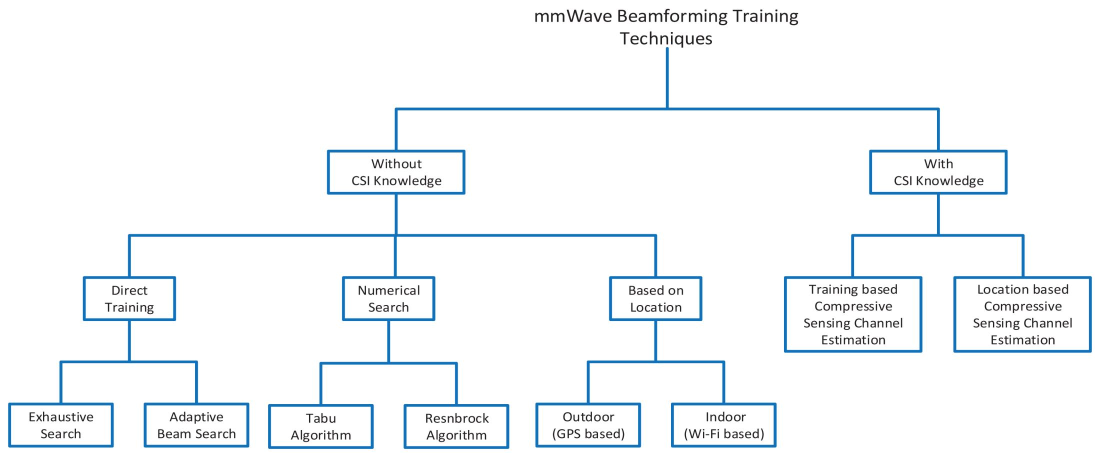


Fig. 2. Taxonomy of mmWave BT techniques.

throughput in bit per second. This can be represented as

$$(b_{\text{tx}}^*, b_{\text{rx}}^*) = W\left(\frac{T_D}{KT_{BT}+T_D}\right)$$
$$\arg\max_{\forall(b_{\text{tx}},b_{\text{rx}})} \log_2\left(1+\frac{|\mathbf{F}_{\text{RX}}^{\mathcal{H}}(:,b_{\text{rx}})\,\mathbf{H}\mathbf{F}_{\text{TX}}(:,b_{\text{tx}})|^2}{\sigma_0}\right)$$
$$\text{s.t. } b_{\text{tx}} \in \phi_{B_{\text{TX}}},\ b_{\text{rx}} \in \phi_{B_{\text{RX}}}, \mathbf{F}_{\text{TX}} \in \phi_{\mathbf{F}_{\text{TX}}}, \mathbf{F}_{\text{RX}} \in \phi_{\mathbf{F}_{\text{RX}}} \quad (4)$$

where $W$ is the available bandwidth, and $\sigma_0$ is the noise power. $\phi_{B_{\text{TX}}}$, $\phi_{B_{\text{RX}}}$, $\phi_{\mathbf{F}_{\text{TX}}}$, and $\phi_{\mathbf{F}_{\text{RX}}}$ are the spaces of the TX beam directions, RX beam directions, TX AWVs, and RX AWVs, respectively. The ratio $(T_D/(KT_{BT}+T_D))$ represents the BT overhead, where $T_D$ and $T_{BT}$ are the times required for data transmissions and BT using one $(b_{\text{tx}}, b_{\text{rx}})$ pair, and $K$ is the total number of $(b_{\text{tx}}, b_{\text{rx}})$ beam pairs involved in the BT process. Optimally, the $(b_{\text{tx}}^*, b_{\text{rx}}^*)$ should be obtained just using one BT duration, i.e., $K = 1$.

## III. Literature Review in Recent mmWave BT Techniques

In this section, we classify various BT techniques suggested in the literature and summarize their pros/cons. These techniques can be classified into two main categories, as shown in Fig. 2, based on the availability of the mmWave CSI [23] as follows.

### A. BT Without CSI Knowledge

In this category, BT is done by testing all available or a subset of transmit/receive (TX/RX) beam pairs to determine the optimal one for constructing the mmWave link without estimating the mmWave CSI. This category includes several BT algorithms, such as the exhaustive, numerical, and location-based search algorithms, etc.

*1) Exhaustive Search (EX) BT:* EX is standardized by IEEE 802.11ad wireless gigabit (WiGig) standard [24]. EX BT is based on analog beamformers using structured antenna codebooks. Three phases are conducted to complete the BT process, as shown in Fig. 3, namely, the sector level sweep (SLS), multiple sector ID capture (MIDC), and beam combining (BC). In the SLS phase shown in Fig. 3, the mmWave WiGig transmitter (initiator) scans its available transmit (TX) antenna sectors, i.e., AWVs, whereas the WiGig receiver is kept in quasi-omni antenna mode [24]. This is done by sending sector sweep frames in all available TX beam directions, as shown in Fig. 3. Then, the WiGig receiver (responder) makes the same process to train its TX beams, as shown in Fig. 3. In the MIDC phase, both initiator and responder trains their RX beams using the same process used in the SLS phase by sweeping their RX beams instead. After SLS and MIDC phases, tables of best candidate TX /RX AWVs are constructed to be used in the BC phase. In this phase, best TX/RX beams collected from the SLS and MIDC phases are combined and tested to find out the best candidates TX/RX beam pairs for constructing the mmWave link and performing fast switching if blocking happens. EX takes a considerable amount of time for completing the BT process, which opens the door for more sophisticated BT techniques. For example, it is stated in [25] that 1.8 ms is needed to accomplish the EX search using 32 TX/RX antenna sectors in the SLS phase and 49 TX/RX beam pairs in the BC phase.

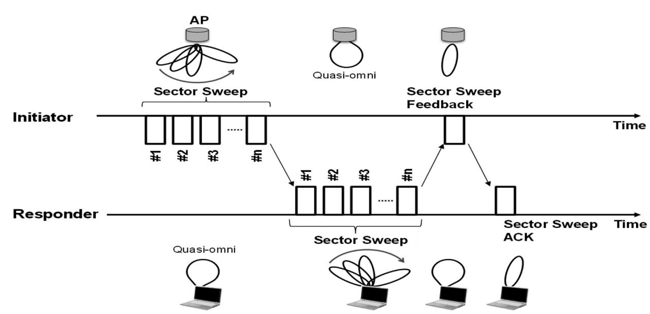


Fig. 3. Example of SLS phase.

*2) Adaptive Beam Search BT:* Multilevel beamwidths can be used to reduce the complexity of the EX BT by adopting wider beams in the earlier BT stages. According to their attained SNRs, narrower beams are adjusted. However, this BT scheme suffers from low BF gain due to the little coverage of the wide beams [26].

*3) Numerical Search BT:* Here, the BT starts with a random TX/RX beam pair, and then the beam search is refined using numerical algorithms, such as the Rosenbrock or Tabu algorithm to reach at a suboptimal pair. However, these algorithms require large training steps, and it is highly based on the initially selected beam pair [27].

*4) Location-Based BT:* Location-based BT was proposed to reduce the number of beams used by the BT process to those expected to cover the mmWave device at estimated location [28]–[31]. The BT complexity of these techniques is based on the localization error of the used localization methodology, including global positioning system for outdoor and Wi-Fi/Li-Fi for indoor localization.

### B. Based on mmWave CSI Knowledge

In this category, the parameters of the mmWave channel, e.g., channel gains, AoDs, and AoAs, are estimated for both line-of-sight (LOS) and non-LOS paths. Then, based on these estimated parameters, the TX/RX antenna beams are constructed.

*1) Training-Based Compressive Sensing (CS) BT:* Utilizing sparsity inherent in mmWave channel, CS was used to estimate mmWave channel parameters at the transmitter side for constructing either analog or hybrid precoders. In this regard, adaptive multiresolution BT was proposed for constructing the sensing matrices [22].

*2) Location-Based CS BT:* To further reduce the complexity of the sensing matrices and increase the BF gain, adaptive multilevel sensing matrices can be constructed based on user localization. Users' positioning can be utilized to assist the construction of the sensing matrices to overcome the poor accuracy of the mmWave channel. In this scheme, the adaptive BT levels are adjusted based on the expected localization error relative to the expected angular spread of the mmWave channel [30], [31].

## IV. Overview of ML Techniques and Terminologies

The classification of ML techniques has been explained in detail in several literature, including [13], [32]–[34]. For the sake of clarity and completeness of the discussions, we give a brief overview of ML techniques' training methodologies and terminologies. Generally speaking, the ML's training methodology can be categorized into one of the following three broad categories.

1) *Supervised Learning:* The ML model is trained to learn a mapping function, $y = f(x)$, using a historical dataset that gives samples of the input–output $(x - y)$ relationship, whereas the objective of the model is to predict future output $(y_o)$ for a given test input $(x_o)$. The learning is done by estimating the probability $p(y|x)$ of the samples in the dataset or specific properties of that distribution [32], [34]. Two main predictive models can be employed, namely the regression and the classification models. The regression models use statistical techniques to model the relationship between explanatory variables and real-valued outcomes to predict the output by using either linear or sigmoid function approximations. On the other hand, classification as the most widely used ML techniques classifies data samples into one out of several classes. In other words, it learns how to map an input to one of the possible outputs. Several classical classification models can be used for mmWave BF applications, including K-nearest neighbor (KNN), support vector machines (SVMs), and decision tree [19]. Additionally, the recent breakthrough in graphical processing unit designs allows more sophisticated and deep artificial neural networks (ANNs) to be used for large-size datasets. Such deep neural network (DNN) architectures, including the convolutional neural network (CNN), recurrent neural network (RNN), Hopfield networks, and Boltzmann machine, have been used in many novel areas in wireless and optical communication networks [13], [33]. We briefly highlight the main ideas of some of the most widely used types as follows.

a) *K-nearest neighbor:* KNN is a classical supervised ML algorithm, which can be applied for classification as well as for regression tasks. When KNN is used for classification, we search for the $K$th nearest neighbors of the test sample $x_o$ in the training dataset and then classify it following the majority of the samples among the $K$th nearest neighbors [32], [35]. Consequently, a particular case is when $K = 1$, where $x_o$ is assigned to the class of the nearest sample using any particular selected distance function, such as the Minkowski norm $L_p$ distance, etc. The performance of KNN depends on the selection of the value of $K$, where less values lead to more accurate classification results but with more sensitivity to noise. On the other hand, larger values reduce noise sensitivity but may lead to less distinct classes' boundaries.

b) *Support vector machine:* SVM is a binary-classification model that finds the hyperplane that maximizes the classification margin (i.e., the distance away from the hyperplane) to separate two classes of the training samples [32].

c) *Deep neural network (also known as deep multilayer neural network):* ANN can be organized either in shallow or deep structures consisting of several layers, which are known as shallow NN or deep NN. One of the main usage of DNN is the function approximation by a weighted combinations of simple units (neurons) in a sequence of layers (input, hidden, and output layers). The neurons in those layers resemble the perception process in a brain, where a specific neurons are activated based on the excitation. Different structures have been investigated in the literature to approximate several types of functions, such as a mapping between input signals (images, sounds, and videos) and their class labels (i.e., classification), and computing future output based on historical values of the inputs and/or outputs (regression) [13], [33].

d) *Convolutional neural network:* CNN is a DNN architecture capable of automatically extracting high-level features from raw input features better than the manual or human-designed features. CNN assumes a locally connected filters instead of fully connected architectures between layers to capture the spatial correlations [36]. CNN exploits two operations, namely convolution and pooling. The convolution uses multiple filters to extract features from the dataset in addition to preserving their corresponding spatial information. On the other hand, pooling (also known as subsampling) is used to reduce the dimensionality of the feature map via either max-pooling or average-pooling [33]. CNN has provided improvements in many fields, including image processing and recognition. Recently, different CNN architectures have been used in many areas through transfer learning technique, which will be explored in more details in Section VI-B.

e) *Recurrent neural network:* RNN is a class of neural networks designed for modeling sequential data. In RNN, it is assumed that the current output of the network is a function of the current input and the previous output via introducing memory cells. However, with long sequences, RNN shows gradient vanishing and exploding problems frequently, which leads to designing other advanced types of RNN, such as the long short-term memory (LSTM) and gated recurrent units (GRU) by introducing a set of gates in the design [37]. Investigating different RNN architectures are promising for analyzing time series data in mobile networks as well as their success stories in speech recognition and natural language processing [33].

2) *Unsupervised Learning:* This methodology searches for hidden patterns and structures of the input data in the absence of data labels. The tasks of unsupervised learning can be divided into clustering, density estimation, and dimension reduction [32]. First, the main goal of clustering is to divide samples into groups or clusters. This means that initially we do not know the class that each sample belongs to in the dataset. Clustering has many applications in data analysis, image/audio/video processing, and recently in wireless communication. *K*-means is one of the most widely used clustering techniques. Second, the aim of density estimation is to estimate the density of the data distribution in the feature space, which may reveal several important characteristics in the high-density regions. The Gaussian mixture model (GMM) is a famous technique of this type. Finally, the dimension reduction aims to transform the data from a high-dimensional space into a low-dimensional space while reserving the principal structures of the data. Principal component analysis and autoencoder (AE) are examples of such type. We briefly highlight the main ideas of some of the most widely used unsupervised learning types as follows.

a) *K-means:* It is a simple clustering algorithm that finds *k* representative optimal points in the feature space for the *k* clusters. Each sample in the dataset is assigned to one cluster according to the distance between the point and each representative. However, selecting the optimal *k*-points is NP-hard problem that can be approximated using a less complex iterative algorithm by randomly selecting initial *k*-points, followed by assigning all samples to the initial *k*-points. Then, we get the mean of each cluster and repeat the process again and again until convergence [38].

b) *Gaussian mixture model:* GMM is an efficient density estimation technique with the objective of fitting the data into a mixture (weighted linear combination) of *k* Gaussian probability distributions. This allows GMM to handle complicated cluster forms. The parameter *k* controls the complexity of GMM, where increasing *k* allows GMM to approximate any continuous distribution to some degree of accuracy. However, the larger *k*, the higher the probability of overfitting and time cost to estimate the mixture parameters using the log-likelihood [39].

c) *Autoencoder:* AE is used for dimension reduction of images, audio, and video signals besides its application in communication systems [4], [40]. An AE is an NN that can be used to learn an effective representation for a dataset in unsupervised learning way (i.e., encoded mode), where the transformed code has lower dimensions compared with the original data. AE consists of two NN: an encoder $f(-|\eta)$ and a decoder $g(-|\theta)$, where $\eta$ and $\theta$ denote the parameters of the encoder and the decoder, respectively [32]. If we have an input $x \in R^d$, the encoder is responsible of finding a latent distribution (or a code) $z \in R^t$, $f(x|\eta) = z$, where $d$ and $t$ are the lengths of $x$ and $z$, respectively, and $t < d$. On the other hand, the decoder tries to recover the original feature $x$ from the code $z$ such that $g(z|\theta) = \bar{x}$ and $x \approx \bar{x}$. Given the dataset $\{x_i\}_{i=1}^{n}$, the object of the AE training process is to learn the parameter set that minimizes the sum of squared error $\sum_{i=1}^{n} ||x_i - \bar{x}_i||^2$. It is noteworthy that when the AE limits the code's length, the training will force the code to capture critical structure of the input features and ignore trivial ones, such as sparse noises, which makes it suitable for denoising.

3) *Reinforcement Learning (RL):* RL adopts different learning methodologies based on trial-and-error similar to humans. An RL's agent is rewarded or penalized for its action in order to maximize the long-term rewards. A recursive environmental feedback is provided to the agent to help in selecting the proper actions in each step by following certain policy that maps agent behavior from state to action. With uncertainty in the environment, the system's dynamics can be modeled using a Markov decision process to optimize the objectives [34]. In the following, we briefly highlight the main idea of *Q*-learning as an example of RL techniques, whereas MAB is explained in Section VI in more details [14], [20].

a) *Q-learning:* This is a model-free RL technique at which the agent does not need to know or have a model of the environment. The agent calculates a *Q*-value corresponding to each state-action pair through experience, which is stored in *Q*-table. The *Q*-value can be considered as a long-term reward. However, it is not suitable for large-scale problems because the tables become too large with hard complicated complexity [41].

## V. ML Applications in mmWave BF

Herein, we will list several interesting applications of ML algorithms for enabling efficient mmWave BF for 5G and B5G networks, which are shown in Fig. 4. Table II summarizes these different ML applications, emphasizing its importance as discussed below.

### A. Beam Selection/Alignment/Tracking

*1) Beam Selection:* ML techniques have been widely used for improving the beam searching, selection, and alignment especially in highly mobile systems, such vehicular, high-speed

TABLE II
ML-BASED MMWAVE BF

| Reference and Scenario | Objective | Machine learning methods | Main conclusion |
|---|---|---|---|
| [35] Downlink hybrid BF | Efficient hybrid BF design | KNN | Achieve the target BER at higher data rates. |
| [13] Multiuser SIMO system | Beam selection using ML/DL based AoA | KNN,SVC,feed forward DNN, and MLP | Near sum rate and classification accuracy results to ES. |
| [50] High speed trains | Fast beam searching with low latency at high speeds | RL, upper confidence bound (UCB) | Proposed scheme can achieve theoretical limit. |
| [20] MISO system | Beam alignment and tracking | RL, Bayesian MAB | Faster learning rate. |
| [14] Vehicular scenario plus LIDAR data | Beam alignment and tracking | DL, DCNN | Efficient configuration of mmWave v2I links. |
| [15] V2I situational awareness | Beam selection with low latency | DL and classification methods | 99% of throughput can be achieved. |
| [37] Vehicular scenario | Intelligent beam/ channel tracking | LSTM model | Accurate prediction of CSI with less overhead |
| [45] One user moving and served by N BSs | mmWave blockage prediction/ handoff requirement | Deep learning, RNN | Successful blockage/ handoff by 95% |
| [4] mmWave coordinated BF | High mobility with reducing latency overhead | DL model that predicts BF vectors | High data rates with adaptation to changing environments. |
| [10] Uplink mmWave | Beam selection using AoA | Multiclass classification KNN, SVM,MLP | MLP outperforms both KNN and SVM. |
| [18] mmWave antenna array | Estimating AoA for radio waves | DNN | Computational complexity reduction compared to MUSIC. |
| [46] MIMO channel transmitter | Predict CSI | LSTM and GRU | Reduce channel estimation and pilot overhead. |
| [11] mmWave BF | Select the best BF technique | DNN | Improve BF performance. |
| [12] Multiuser beam allocation | Beam allocation | DNN multiclass classification | Achieve 91.6 to 97.7% accuracy compared to EX. |
| [49] mmWave user allocation and handover | Maximize user-BS connection time after handover | MAB | Enhance handover decision and network performance. |
| [36] mmWave blockage prediction | Predict blockage using RGB image of the surrounding | DCNN | Enhance handover decision. |
| [51] SVM based BF | Reduce complexity of calculated SVD | DNN | Reduce complexity and enhance performance. |
| [16] Beam selection | Select primary and redundant beams | Iterative random search | Improve handover. |
| [38] Clustering | User clustering and PA in mmWAve-NOMA settings | K-means | Improve user clustering with low complexity. |
| [52] Pilot contamination | Pilot contamination attack detection in mmWave-NOMA | SVM, ELM, LOG | 100% to 95% static and dyunamic detection rates. |
| [53] RRM (user association and PA) | Maximize the system's EE under different constraints | Semi-Supervised learning, DNN | Higher EE with low complexity. |
| [41] Inter-Beam Inter-Cell Interference Mitigation | User association, inter-beam PA in mmWAve-NOMA | Q-learning | 30% sum rate enhancement. |
| [54] Wireless power transfer (WPT) | Enhance the power transfer efficiency | Random forest, DNN | Improved performance and pblackiction accuracy. |
| [55] Hybrid BF/ combining | High transmission capacity with low complexity | DNN, self supervised learning | Higher BER performance. |
| [21] Blind beam alignment | Blind BF with multiple BS and users | DRL | Improved considerable data rates. |
| [56] Hybrid precoding | Effective hybrid precoding with low complexity | DNN | Higher performance. |
| [57] BF in mMIMO HetNets | Efficient BF technique | DNN,RNN,MLP | Higher weighted sum rates. |
| [17] Beam selection in mmWave Hetnets | Develop efficient beam selection design | DNN | 95% accuracy. |
| [19] Initial Beam Alignment | AoA Estimation | DT | Significant improvement of the rate. |
| [40] Hybrid BF in mmWave. | Find the BF that maximize the achieved rates | Autoencoder | 60 - 70% gains in rates |
| [39] Modeling Blockage Loss | Robust BF using statistical Blockage modeling | GMM | Better understanding of blockage effects on users |

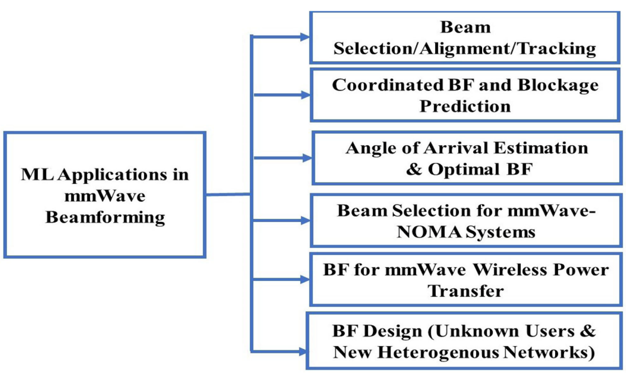


Fig. 4. Sample applications of ML in mmWave BF.

trains (HSTs), and unmanned aerial vehicles (UAVs) communications [15], [20], [37], respectively. A deep conventional neural network (DCNN) model has been used for reducing mmWave beam selection overhead and configuring mmWave vehicle to infrastructure links (V2I) [15]. The mmWave beam selection problem has been formulated as a multiclass classification problem to learn the optimal beam pair taking into consideration the nearby vehicles locations [37]. A bandit inspired beam searching methods for mmWave HSTs was proposed in [20] that not only speed up the highly variable channel estimation calculations but also to provide sufficient time for data transmission. The number of beam directions or propagation paths, which represent the arms, are reduced by exploiting the previous propagation information. Those algorithms performed extensive evaluations using realistic traffic patterns derived from Google maps, which not only allows mmWave BSs to reach near-optimal performance but also remains within 5% of the optimal performance by swift adaptation to system changes (i.e., blockage and traffic).

*2) Beam Alignment:* Experimental results cleared out that in a 7° beam width system, a misalignment of 18° diminishes the link budget by 17 dB, which might cut the link entirely [42]. Hence, precise mmWave BA is required to guarantee high data rate communications. Several works have formulated environmental-aware beam alignment learning algorithms as contextual MAB problems by exploiting the historic information, which have been used for vehicular and HST communications [14], [43]. The distributed BA search problem in a point to point mmWave MIMO system was formulated as adversarial MAB problem in [44]. An exponential weight algorithm algorithm was applied independently at both TX and RX to explore different beams and identify the best one for data transmission using a single bit feedback information.

*3) Beam Tracking:* mmWave beam tracking is a challenging issue, especially in high mobile scenarios, such as UAVs and V2I, due to the fast mobility and narrow transmission beams. It is imperative to employ artificial intelligence and ML techniques that predicts the surrounding environment in order to enhance the performance [45], [46]. A promising RNN-based beam tracking model that tracks the AoA of a mobile user is proposed in [47]. ML-based beam tracking will enhance the data transmission quality and speeds up the beam switching process.

## *B. Coordinated BF and Blockage Prediction*

The sensitivity of mmWave signals to blockage greatly impact the coverage and reliability of mobile users. One interesting paradigm to enhance the mmWave network coverage is integrating ML and coordinated BF techniques [48], where a number of coordinated BSs simultaneously serve mobile users to avoid blockage. Each user transmits only one uplink training pilot sequence, which is jointly received at the BSs and draw a defining signature for the user location and its surrounding environment. The authors in [48] have developed a deep learning model that learns these signatures to predict the BF vectors at the BSs. The proposed system adopts an online training technique by learning and adapting to the environment.

Blockage naturally leads to disconnecting the communication session between the user and the BS while reconnecting the user to another LOS BS incurs high BT overhead and latency. ML-based BF techniques can solve such a problem by switching to another unblocked beams by learning and predicting that certain link will experience blockage in the next few time frames [13]. This capability allows the serving BS to proactively handover the user to another BS with highly probable LOS link. The work in [49] has exploited empirical distribution of users, posthandover trajectory, and LOS blockage, which is learned online via an MAB framework in order to maximize the expectation of the user-BS connection time after each handover. This approach is capable of enhancing handover decision and greatly improve the network performance. On the other hand, the authors in [36] have exploited a framework that incorporate computer vision and deep learning tools to predict mmWave beams and blockages using RGB images captured by a camera-enabled BS. The author in [39] have used GMM for modeling blockage loss data for BF.

### C. AoA Estimation and Optimal BF Technique Selection

*1) AoA Estimation:* The work in [10] investigates how AoA information can be exploited by both DL and ML approaches to perform beam selection in the uplink of a mmWave communication system. The uplink mmWave beam selection task was formulated as a multiclass classification problem and solved by two supervised ML algorithms (KNN and SVM classifiers) and one feed forward DNN technique [multilayer perceptron (MLP)]. The main target of these algorithms is to choose the optimal formation for the analog BF network based on the estimated AoAs of different users devices. The results show that MLP outperforms both KNN and SVM methods in terms of classification accuracy. Additionally, a multiple signal classification (MUSIC) algorithm is used for estimating both AoA and received powers more precisely (close to 80%). Additionally, the work in [18] proposed another DNN approach for detecting and estimating the AoAs of radio waves. Their network architecture learns a mapping technique that relates the received antenna array signals with its associated AoAs of the impinging wave. In [19], the authors have considered a hierarchical posterior tree-based matching algorithm for active learning in mmWave initial alignment and AoA estimation.

*2) Optimal BF Technique Selection:* Due to the existence of several BF techniques in the literature, it is imperative to select the proper and the most efficient one under different constraints and environments. The authors in [11] have investigated a DNN-based BF selection scheme for a two-user multi-input single-output (MISO) interference channels. In their design, each user chooses between two popular BF schemes, which are the maximum ratio transmission BF and the zero-forcing BF, where the transmit power and the channel vectors are the inputs, and the output is the recommended BF scheme.

### D. Beam Selection for mmWave NOMA Systems

The author in [58] has investigated a mmWave NOMA system, where a localization-based joint user selection and BF are proposed. Developing practical algorithms that efficiently handle different problems in mmWave NOMA scenarios can also be handled using the promising ML techniques. In [59], a set of ML-based algorithms have been proposed for beam selection and power allocation in mmWave NOMA systems. A low complex and promising $k$-means user clustering algorithm implemented in NOMA scenario is developed in [38]. An ML detection scheme was proposed to achieve better attack detection policy in mmWave NOMA systems in [52]. Zhang *et al.* [53] discuss an efficient DL framework to handle the user association, subchannel, and power allocation problems in NOMA mmWave heterogeneous networks to maximize EE under QoS constraints. The authors proposed a semisupervised DL-based subchannel allocation scheme and DNN- based power optimization algorithms. The work in [41] discusses an improved online $Q$-learning based algorithm that handles joint user-cell association and interbeam power allocation for sum rate maximization of a downlink mmWave NOMA system.

### E. BF for mmWave Wireless Power Transfer (WPT)

ML algorithms can be used to accelerate efficient realization of mmWave WPT techniques. The authors in [54] developed an online technique for receiver positioning using random forest and DNN to efficiently charge the power receivers. A stochastic geometry based approach for directional power transfer using BF for WPT has been discussed in [60]. In [61], an ML-based method for controlling power with fast control at the transmitter was proposed, where Bluetooth low energy technology is used for communication.

### F. BF Design

*1) Unknown Users Scenario:* Flexible ML-based design techniques that are compatible with different BF/combing architectures have been studied and investigated by researchers to accelerate the BT especially in mmWave systems. In [55], the authors proposed a neural hybrid BF/combining MIMO system with a significant BER performance improvements. A blind BF on a multiple BS cellular environment with multiple mobile users using DRL has been investigated in [21]. A DL-based mmWave massive MIMO design for effective hybrid precoding with minimum BER, in which each selection of the precoders for obtaining the optimized decoder is regarded as a mapping relation in the DNN, was proposed in [56]. In [35], a hybrid BF design structure for the DL of multiuser mmWave scenario using KNN for grouping RF chains at the BS was given.

*2) New Heterogeneous Networks Scenario:* Probing optimal beamformers in massive MIMO systems is a critical nonconvex optimization problem, solved by high computational cost optimization techniques. The problem is further complicated in mmWave-based HetNet scenarios, where the BSs are merged with large numbers of transmit antennas and have different inter-BS distances. Trials of solving such problems by using ML techniques have been conducted [17], [57]. A 85% precise DNN design for beam selection in mmWave HetNets that utilizes the CSI of sub-6 GHz network as input features and its output is the mmWave BS and beam selection is discussed in [17]. An RNN learning based BF for MISO interference channel in HetNet is investigated in [57].

## VI. Challenges and Future Research Directions

In spite of the progress in ML algorithms and how they can reduce complexity in wireless network and BF, it is vital to recognize the existence of different issues and challenges that must be addressed if those algorithms are to be adopted in practical systems. In this section, we present how ML-based algorithms affect and be affected by various challenges in mmWave BF. We present the following interesting and challenging future research topics that are worth for further investigations.

### A. Datasets Availability

The implementation of ML-based algorithms requires large amounts of real data samples for training purpose. Additionally, those huge samples need to be labeled for supervised learning algorithms. However, it is not always feasible to acquire enough amounts of real labeled data in practice, which forms a bottleneck in communication networks. Regarding datasets in the area of mmWave BF, few datasets have been published for public use [62]–[64].

### B. Dealing With Small Labeled Data

The problem of small labeled data can be solved by semisupervised learning techniques that can efficiently deal with small labeled data and large set of unlabeled data. Another interesting method to solve the lack of massive data samples is to use generative models, such as generative adversarial networks and AEs to expand the available sample set by generating a synthetic data following the same distribution of the original real data [4]. This approach is beneficial for all ML-based BF algorithms while the performance of such method of synthetic data-driven ML algorithms needs to be investigated in more details. Another interesting concept has been explored in [18] to deal with small datasets. The author investigated a DNN-based network for AoA estimation for antenna array under the assumption of low number of samples. The proposed DNN is divided into two subnetworks, where the first one which is known as the detection network is used to divide the search area of antenna array into subsectors in order to reduce the training angles combinations. Then, another subnetwork is used for estimating the AoA within the specified subsector. However, no deep investigations have been performed to confirm the performance of such approach. Finally, one of the most promising techniques to explore is to improve the learning process through the transfer of knowledge from a pretrained related task in other fields. It is noteworthy that pretrained CNN networks that have been trained on very large datasets, such as ImageNet [65], AlexNet [66], VGG Net [67], or GoogLeNet [68], can be used to fine-tune their weights on a different dataset with limited number of samples, and a different vector of features [34].

### C. Frequent Hand-Off and Mobility

In order to support high mobility in mmWave systems, we have to overcome the sensitivity of mmWave signals to blockage and to dynamically maintain mmWave beams aligned. Since the optimal BF in mmWave systems with large antenna array requires considerable training overhead and complex processing, it is interesting to explore other ML algorithms beside the coordinated BF to solve such problems. One interesting idea to exploit ML techniques to select primary and several redundant beams for each node [16], where the node could perform self-handover if the primary beam is interrupted. Additionally, it is imperative to jointly consider channel prediction techniques to support hand-off decision. While LSTM has been used in recent the literature, other time series techniques can be investigated for this task, such as the echo state network or long-short term echo state network [69].

### D. Fast ML Inference Versus Conventional Beam Selection Algorithms

Due to the inherent complexity of the beam training algorithms in mmWave BF, proposing an optimal/suboptimal ML-based algorithm with fast inference time is a challenging task specially if we take into consideration the latency, frequent hand-off, and users' mobility in vehicular-to-everything scenarios. The work in [12] has considered a supervised ML algorithm for beam selection and switching (BSS) to provide the best performance at the receiver. The BSS is interpreted using DNN as a multiclass classification algorithm. However, it is possible to exploit other ML techniques to support efficient BSS especially under uncertainty in the network, such as MAB and actor critic RL techniques.

### E. Adaptive ML for Easy/Adversarial Environment

Current existing ML algorithms are implemented directly to mmWave BF. However, nontrivial robust ML algorithms specialized for mmWave environments are required. ML-based easy data methodology tries to develop adaptive simultaneous algorithms for both best and worst cases [70]. For example, an application of the learning with easy data framework (e.g., [70]) to beam learning might work not only in the worst situations but also take advantage of easy situations to obtain better performance.

## VII. Case Study: MAB for mmWave BT

In this section, we elaborate in explaining single-player MAB technique and its extension to multiplayer case. Also, we provide a case study with two different scenarios exploiting MAB for single and concurrent mmWave BT to show the superiority of ML-based BT over conventional techniques.

### A. MAB Dilemma

MAB is an entirely online ML, where the player attempts to obtain the maximum reward from several arms of slot machines, where the rewards of the arms follow random distributions. It has been adopted for solving many practical optimization issues in wireless communications (network routing and resource allocation, etc.) [71], [72]. It is divided into single-player and multiplayer MAB as follow.

*1) Single-Player MAB:* In single-player MAB, a player tries to identify the arm having the maximum long-term reward within finite trials [73]. The player gathers information on every slot machine (exploration) by examining several arms as feasible, later determining the arm with the most considerable reward. Hence, the player tries to compromise between playing with the arm having the maximum achievable reward so far, i.e., exploitation, or investigating new arms, i.e., exploration. For long horizon time (investigation period), the player can accurately anticipate the due reward of each arm. According to the distribution of the rewards, the MAB problem is divided into stochastic or adversarial. In stochastic MAB, the arms' rewards are pulled from independent and identical distributions, which are undetermined for the players. Meanwhile, in adversarial MAB, the rewards are decided according to the hostile environment.

*2) Multiplayer MAB:* In multiplayer MAB, each player acts in sequential trials to obtain an unidentified reward too. If more than one player chooses the same arm, collisions occur [71]. Then, based on the collision model, players might share the rewards or no player obtains the reward. According to the exchanged information between the players, multiplayer MAB algorithms are classified as centralized and decentralized. For decentralized MAB, no data exchanges among the players, and each player plays his future actions only based on his collected reward remarks. On the other hand, in the centralized MAB, the game is performed cooperatively among the players by sharing complete observations making the game seems like a single-player one. Collisions in the decentralized setup are unavoidable compared to the centralized counterparts. Hence, each player plays selfishly to study collisions and attempts to be far from them while interfacing with the environment toward increasing his profit.

### *B. Case Study*

In the following case study, we model the mmWave BF problem as a stochastic MAB, such as the upper confidence bound (UCB) algorithm [73].

UCB is one of the famous MAB algorithms that efficiently addresses the exploitation–exploration dilemma. The authors in [74] proposed to utilize UCB for mmWave BT, where they used locations indexed offline database containing mmWave beam identifications and their related channel strengths. Then, a digest from this database is utilized by the UCB-based BT algorithm for online learning based on the current user's location. However, no concurrent mmWave BT was proposed in [74]. In this article and for the purpose of surveying, we will propose a simple form of the UCB-based BT suitable for both single and concurrent mmWave BT, where the player(s) learns the best beam pair time by time selfishly without any prior knowledge provided by any preconstructed databases or subside information. During the MAB game, UCB provides a compromise between selecting the arm having the maximum average reward so far or exploring new ones. At the beginning of the UCB algorithm, the player checks all available arms and obtains their corresponding rewards. Afterward, the arm satisfying the following maximization equation is selected by the player:

$$k_t^\star = \arg\max_{1\le k\le K}\left(\bar{\Upsilon}_{k,t-1} + \sqrt{\frac{2\ln(t)}{s_{k,t-1}}}\right), K+1 \le t \le T \quad (5)$$

where $k_t^\star$ indicates the selected arm $k^\star$ at time $t$, where $K$ is the total number of arms, and $T$ is the time horizon. $\bar{\Upsilon}_{k,t-1}$ and $s_{k,t-1}$ are the average reward of arm $k$ and the number of times it was selected up to $t-1$. After selecting arm $k^\star{}_t$, its number of selections and average reward are updated for the next round, as follows:

$$s_{k^\star,t} = s_{k^\star,t-1} + 1 \quad (6)$$

$$\bar{\Upsilon}_{k^\star,t} = \frac{1}{s_{k^\star,t}} \sum_{j=1}^{s_{k^\star,t}} \Upsilon_{k^\star,j}. \quad (7)$$

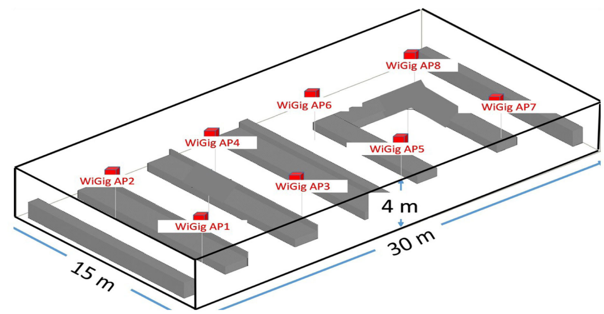


Fig. 5. mmWave WLAN area under study.

**Algorithm 1:** UCB Algorithm.

1: **Initialize**: each arm $k$, $1 \le k \le K$ will be selected once, and its corresponding $\Upsilon_{k,t}$ is obtained.
2: **For** $t = K+1 : T$
3: Draw a beam pair and obtain the reward $\Upsilon_{m,t}$
$k_t^\star = \arg\max_{1\le k\le K}(\bar{\Upsilon}_{k,t} + \sqrt{\frac{2\ln(t)}{s_{k,t}}})$
4: $s_{k^\star,t} = s_{k^\star,t-1} + 1$
5: $\bar{\Upsilon}_{k^\star,t} = \frac{1}{s_{k^\star,t}} \sum_{j=1}^{s_{k^\star,t}} \Upsilon_{k^\star,j}$
6: **End For**

In the conducted simulations, realistic mmWave channels are generated using ray tracing through commercial wireless Insite software. Fig. 5 shows the ray tracing simulation area where an indoor mmWave WLAN system is considered with eight access points (APs) that are fixed on the ceiling. The room size is $30 \times 15 \times 4\text{m}^3$, and each AP uses 10-dBm TX power and 2-D BF. In the concurrent mmWave BT scenario, all APs are operating, whereas in the single mmWave BT scenario, we consider a single AP fixed at the center of the room's ceiling. A dataset of 7200 channels is generated at different locations in the room area, where the locations are separated by 0.25 m in both horizontal and vertical directions. This dataset in addition to the source code of the proposed UCB-based BT is set publicly available in [75] and [76]. For BF, the azimuth coverage angle of the mmWave AP, $\vartheta_{\text{azm}}$ is divided into a number of beam tiers, which is equal to $N_{\text{tier}} = \lfloor \frac{\vartheta_{\text{azm}}}{\vartheta_{-3\text{dB}}} \rfloor$, where $\vartheta_{-3\text{dB}}$ is the beamwidth. Then, the total number of beams is equal to $K = 1 + \frac{6N_{\text{tier}}(N_{\text{tier}}+1)}{2}$, as given in [77]. For example, using $\vartheta_{-3\text{dB}} = 60^\circ$ and $\vartheta_{\text{azm}} = 85^\circ$, then $N_{\text{tier}} \approx 2$ and $K = 19$ beams.

*1) Single mmWave BT:* In this scenario, we assume a single mmWave AP located at the center of the room with quasi-omni antenna pattern, whereas the mmWave users are uniformly distributed inside the room area. Both cases of no blockage and

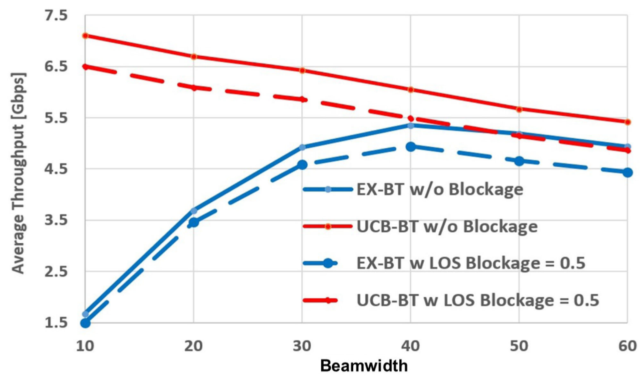


Fig. 6. Average throughput comparisons against beamwidth.

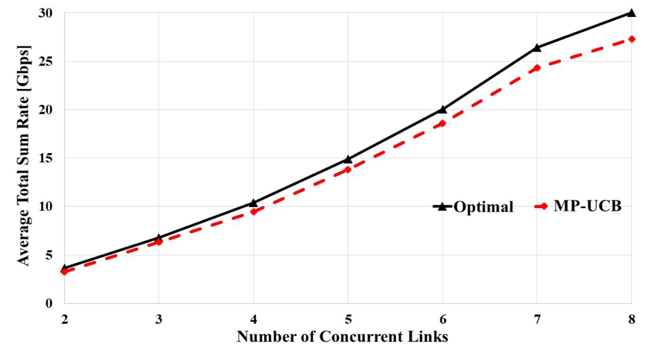


Fig. 7. Average total system rate against the number of concurrent links using LoS blockage probability of 0 and $-3$ dB beamwidth of 50°.

LOS blockage are modeled as Bernoulli random variables [78] with a probability of 0.5.

The problem is formulated as a single-player MAB. The objective is to allow the AP to interact with the environment via MAB-based algorithm in order to improve the beam selection based on its observations. The AP is considered as a player or agent trying to maximize its long-term reward/throughput, through playing over the available beam pairs $(b_{\mathrm{tx},k}, b_{\mathrm{rx},k})$, where $K$ indicates the total number of beam pairs, i.e., the bandit's arms. In this case, the reward for the UCB algorithm is expressed as

$$\Upsilon_{k,t} = W\left(\frac{T_D}{T_{BT}+T_D}\right) \log_2\left(1+\frac{\left|F_{\mathrm{RX}}^H\left(:, b_{\mathrm{rx},k,t}\right) H F_{\mathrm{TX}}\left(:, b_{\mathrm{tx},k,t}\right)\right|^2}{\sigma_0}\right) \quad (8)$$

where $W, T_D$, $T_{BT}$, and $\sigma_0$ in dBm are equal to 2.16 GHz, 1 ms, $14\mu$s, and $-174+10\log_{10}(W)+10$, respectively, whereas the adopted carrier frequency is 60 GHz. Then, the algorithm is running as previously explained and shown in Algorithm 1 steps.

Fig. 4 compares the average throughput of both the conventional EX-BT and the UCB-BT versus $-3$ dB beamwidth. The simulation results in Fig. 4 shows that for sharp beams, i.e., $-3$ dB beamwidth of 10°, a five times improvement in average throughput is obtained using UCB-BT over using the EX-BT. This comes from the online learning capability of the UCB algorithm, which can successively learn the environment and enhance the throughput performance without the need of frequently EX all available beams. As the beamwidth increases, the performance of the UCB-BT decreases due to the decrease in the BF gain as UCB-BT uses one beam switching only at a time to learn the environment. However, the performance of the EX-BT increases influenced by the high decrease in the BT overhead, i.e., $K\,T_{BT}$, as the exhaustively searched beam space $K$ is decreased. For example, at $\vartheta_{-3\mathrm{dB}} = 10°$, $K$ is equal to 271, whereas it is equal to 36 when $\vartheta_{-3\mathrm{dB}} = 30°$ as previously explained. This happens till reaching a point where the effect of the low BF gain becomes dominant, at which the average throughput of the EX-BT tends to decrease. Additionally, Fig. 6 shows that ML-based UCB-BT approach effectively outperforms the EX-BT at all tested beamwidths for both cases of with and without LOS blockage.

*2) Concurrent mmWave BT:* In this scenario, multiple mmWave APs are operating simultaneously without central coordination or permitting information exchange among them. Thus, selfish multiplayer MAB will be employed to achieve the suboptimal set of concurrent beams. Specifically, an MAB algorithm will be implemented in each mmWave AP to interact with the environment independently, and timely enhance its concurrent beam selection based on its successive observations. In this scenario, a mmWave AP is acting as the player trying to maximize its own long- term profit (i.e., spectral efficiency) at each time via playing over its available beam space (i.e., the arms of the bandit). This is done through utilizing its own observations irrespective of the other APs selections/observations.

In this case, the reward for the UCB algorithm is expressed as

$$\Upsilon_{k,t} = W\left(\frac{T_D}{T_{BT}+T_D}\right) \times \log_2\left(1+\frac{P_{rk}\left(b_{\mathrm{tx},k,t}, b_{\mathrm{rx},k,t}\right)}{\sum_{m=1,\ m\neq k}^{K} P_{rm}\left(b_{\mathrm{tx},m,t}, b_{\mathrm{rx},m,t}\right)+\sigma_0}\right) \quad (9)$$

where the term $P_{rk}(b_{\mathrm{tx},k,t}, b_{\mathrm{rx},k,t})$ indicates the received power of mmWave link $k$ using beam pair $(b_{\mathrm{tx},k,t}, b_{\mathrm{rx},k,t})$ at time $t$, where $K$ indicates the total number of concurrent links. $\sum_{m=1,\ m\neq k}^{K} P_{rm}(b_{\mathrm{tx},m,t}, b_{\mathrm{rx},m,t})$ represents the sum of the interference powers from other concurrent links at time $t$.

Fig. 7 shows the average total sum rate of the MP-UCB MAB algorithm in addition to the optimal performance against the number of concurrent links, which are uniformly distributed inside the WLAN area, where LoS blockage probability of 0 and $-3$ dB beamwidths of 50°, i.e., 19 beam IDs, are used. As shown in this figure, as the number of concurrent links increases, the average total sum rate increases as well. However, the curves tend to saturate after using seven concurrent links due to the increase in mutual interference. It is interesting to note that the proposed MP-UCB MAB based concurrent BT algorithm shows comparable performance to the baseline optimal one. At eight concurrent mmWave links, about 93% of the optimal performance is obtained using MP-UCB MAB algorithm by just testing one TX/RX beam pair at a time. Compared to the existing mmWave BT techniques, in [29], to obtain 93% of the optimal performance using the highly accurate Li-Fi and Wi-Fi localization techniques, about 8 and 26 beam pairs should be used in the BT process. Moreover, the schemes given in [22], [29], and [30] used 64, 961, and 94 beam pairs to obtain 93% of optimal performance, respectively. Also, a high number of beam pairs are required by the numerical search BT to obtain 93% of the optimal performance, as given in [27]. These compared values come even without considering mutual interference, such as our case [29]–[31]. Hence, these performance comparisons assure the high potency of the proposed MAB-based BT over highly accurate non-ML localization-based BT techniques, e.g., Li-Fi localization, and even over complicated non-ML BT schemes based on mmWave channel estimation or numerical search.

## VIII. Conclusion

This article has provided a comprehensive overview of the applicability of ML algorithms in the area of mmWave BF. The aforementioned discussion has identified challenges and hurdles that need to be addressed by the community to establish viable ML-based protocols for supporting BF in B5G networks. The scope of future research when ML meets BF is broad; therefore, we presented a few interesting and challenging research topics we believe are worth further investigations. Moreover, we presented a case study with two scenarios to show the efficiency of MAB method over EX method in mmWave single and concurrent BT.

## References

[1] I. Ahmed *et al.*, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surv. Tuts.*, vol. 20, no. 4, pp. 3060–3097, Oct.–Dec. 2018.

[2] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.

[3] Z. Lin, M. Lin, B. Champagne, W. Zhu, and N. Al-Dhahir, "Secure beamforming for cognitive satellite terrestrial networks with unknown eavesdroppers," *IEEE Syst. J.*, vol. 15, no. 2, pp. 2186–2189, Jun. 2021.

[4] A. B. Zaky, J. Z. Huang, K. Wu, and B. M. ElHalawany, "Generative neural network based spectrum sharing using linear sum assignment problems," *China Commun.*, vol. 17, no. 2, pp. 14–29, Feb. 2020.

[5] A. A. A. El-Banna, B. M. ElHalawany, A. B. Zaky, J. Z. Huang, and K. Wu, "Machine learning-based multi-layer multi-hop transmission scheme for dense networks," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2238–2242, Dec. 2019.

[6] S. Hashima, B. M. ElHalawany, K. Hatano, K. Wu, and E. M. Mohamed, "Leveraging machine-learning for D2D communications in 5G/beyond 5G networks," *Electronics*, vol. 10, no. 2, 2021, Art. no. 169.

[7] M. I. Hamed, B. M. ElHalawany, M. M. Fouda, and A. S. T. Eldien, "A novel approach for resource utilization and management in SDN," in *Proc. 13th Int. Comput. Eng. Conf.*, 2017, pp. 337–342.

[8] A. A. A. El-Banna, K. Wu, and B. M. ElHalawany, "Application of neural networks for dynamic modeling of an environmental-aware underwater acoustic positioning system using seawater physical properties," *IEEE Geosci. Remote Sens. Lett.*, to be published, doi: 10.1109/LGRS.2020.3039283.

[9] B. M. ElHalawany, H. M. Abdel-Kader, A. TagEldeen, A. E. S. Ahmed, and Z. B. Nossair, "Vision-based obstacles detection for a mobile robot," in *Proc. 8th Int. Conf. Informat. Syst.*, 2012, pp. MM- 93–MM-99.

[10] C. Antón-Haro and X. Mestre, "Learning and data-driven beam selection for mmWave communications: An angle of arrival-based approach," *IEEE Access*, vol. 7, pp. 20404–20415, 2019.

[11] H. J. Kwon, J. H. Lee, and W. Choi, "Machine learning-based beamforming in two-user MISO interference channels," in *Proc. Int. Conf. Artif. Intell. Inf. Commun.*, 2019, pp. 496–499.

[12] H. P. Tauqir and A. Habib, "Deep learning based beam allocation in switched-beam multiuser massive MIMO systems," in *Proc. 2nd Int. Conf. Latest Trends Elect. Eng. Comput. Technol.*, 2019, pp. 1–5.

[13] A. Alkhateeb, I. Beltagy, and S. Alex, "Machine learning for reliable mmWave systems: Blockage prediction and proactive handoff," in *Proc. IEEE Global Conf. Signal Inf. Process.*, 2018, pp. 1055–1059.

[14] M. B. Booth, V. Suresh, N. Michelusi, and D. J. Love, "Multi-armed bandit beam alignment and tracking for mobile millimeter wave communications," *IEEE Commun. Lett.*, vol. 23, no. 7, pp. 1244–1248, Jul. 2019.

[15] A. Klautau, N. González-Prelcic, and R. W. Heath, "LIDAR data for deep learning-based mmWave beam-selection," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 909–912, Jun. 2019.

[16] M. A. Jasim, N. Siasi, A. Aldalbahi, and N. Ghani, "Soft self-handover scheme for mmWave communications," in *Proc. SoutheastCon*, Huntsville, AL, USA, 2019, pp. 1–6.

[17] D. Jagyasi and M. Coupechoux, "DNN based beam selection in mmW heterogeneous networks," *Int. Conf. Network Games, Control Optimisation (Netgcoop)*, 2020, Cargèse, France.

[18] M. Chen, Y. Gong, and X. Mao, "Deep neural network for estimation of direction of arrival with antenna array," *IEEE Access*, vol. 8, pp. 140688–140698, 2020.

[19] S. Chiu, N. Ronquillo, and T. Javidi, "Active learning and CSI acquisition for mmWave initial alignment," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2474–2489, Nov. 2019.

[20] M. Cheng, J. Wang, J. Wang, M. Lin, Y. Wu, and H. Zhu, "A fast beam searching scheme in mmWave communications for high-speed trains," in *Proc. IEEE Int. Conf. Commun.*, May 2019, pp. 1–6.

[21] V. Raj, N. Nayak, and S. Kalyani, "Deep reinforcement learning based blind mmWave MIMO beam alignment,"2021. [Online]. Available: https://arxiv.org/abs/2001.09251

[22] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[23] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE Commun. Surv. Tuts.*, vol. 18, no. 2, pp. 949–973, Apr.–Jun. 2016.

[24] *IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band*, IEEE Standard 802.*11ad-2012*, pp. 1–628, 2012.

[25] K. Hosoya *et al.*, "Multiple sector ID capture (MIDC): A novel beamforming technique for 60-GHz band multi-Gbps WLAN/PAN systems," *IEEE Trans. Antennas Propag.*, vol. 63, no. 1, pp. 81–96, Jan. 2015.

[26] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Multilevel millimeter wave beamforming for wireless backhaul," in *Proc. IEEE GLOBECOM Workshops*, 2011, pp. 253–257.

[27] B. Li, Z. Zhou, H. Zhang, and A. Nallanathan, "Efficient beamforming training for 60-GHz millimeter-wave communications: A novel numerical optimization framework," *IEEE Trans. Veh. Technol.*, vol. 63, no. 2, pp. 703–717, Feb. 2014.

[28] A. S. Mubarak, H. Esmaiel, and E. M. Mohamed, "LTE/Wi-Fi/mmWave RAN-level interworking using 2 C/U plane splitting for future 5G networks," *IEEE Access*, vol. 6, pp. 53473–53488, 2018.

[29] A. M. Nor and E. M. Mohamed, "Millimeter wave beamforming training based on Li-Fi localization in indoor environment," in *Proc. IEEE Global Commun. Conf.*, 2017, pp. 1–6.

[30] A. Abdelreheem, E. M. Mohamed, and H. Esmaiel, "Location-based millimeter wave multi-level beamforming using compressive sensing," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 185–188, Jan. 2018.

[31] A. Abdelreheem, E. M. Mohamed, and H. Esmaiel, "Adaptive location-based millimetre wave beamforming using compressive sensing based channel estimation," *IET Commun.*, vol. 13, no. 9, pp. 1287–1296, 2019.

[32] R. He and Z. Ding, *Applications of Machine Learning in Wireless Communications*. London, U.K., 2019.

[33] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surv. Tuts.*, vol. 21, no. 3, pp. 2224–2287, Jul.–Sep. 2019.

[34] M. E. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine learning for 5G/B5G mobile and wireless communications: Potential, limitations, and future directions," *IEEE Access*, vol. 7, pp. 137184–137206, 2019.

[35] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Multi-user hybrid beamforming relying on learning-aided link-adaptation for mmWave systems," *IEEE Access*, vol. 7, pp. 23197–23209, 2019.

[36] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter wave base stations with cameras: Vision-aided beam and blockage prediction," in *Proc. IEEE 91st Veh. Technol. Conf.*, Antwerp, Belgium, 2020, pp. 1–5.

[37] Y. Wang, A. Klautau, M. Ribero, A. C. K. Soong, and R. W. Heath, "MmWave vehicular beam selection with situational awareness using machine learning," *IEEE Access*, vol. 7, pp. 87479–87493, 2019.

[38] J. Cui, Z. Ding, and P. Fan, "The application of machine learning in mmWave-NOMA systems," in *Proc. IEEE 87th Veh. Technol. Conf.*, 2018, pp. 1–6.

[39] V. Raghavan *et al.*, "Statistical blockage modeling and robustness of beamforming in millimeter-wave systems," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3010–3024, Jul. 2019.

[40] T. Peken, R. Tandon, and T. Bose, "Unsupervised mmWave beamforming via autoencoders," in *Proc. IEEE Int. Conf. Commun.*, 2020, pp. 1–6.

[41] M. Elsayed, K. Shimotakahara, and M. Erol-Kantarci, "Machine learning-based inter-beam inter-cell interference mitigation in mmWave," in *Proc. IEEE Int. Conf. Commun.*, 2020, pp. 1–6.

[42] T. Nitsche, A. B. Flores, E. Knightly, and J. Widmer, "Steering with eyes closed: Mm-Wave beam steering without in-band measurement," in *Proc. IEEE Conf. Comput. Commun.*, 2015, pp. 2416–2424.

[43] M. Hashemi, A. Sabharwal, C. Emre Koksal, and N. B. Shroff, "Efficient beam alignment in millimeter wave systems using contextual bandits," in *Proc. IEEE Conf. Comput. Commun.*, 2018, pp. 2393–2401.

[44] I. Chafaa, E. V. Belmega, and M. Debbah, "Adversarial multi-armed bandit for mmWave beam alignment with one-bit feedback," in *Proc. EAI Int. Conf. Perform. Eval. Methodologies Tools*, 2019, pp. 23–30.

[45] Y. Guo, Z. Wang, M. Li, and Q. Liu, "Machine learning based mmWave channel tracking in vehicular scenario," in *Proc. IEEE Int. Conf. Commun. Workshops*, May 2019, pp. 1–6.

[46] W. Jiang and H. D. Schotten, "Deep learning for fading channel prediction," *IEEE Open J. Commun. Soc.*, vol. 1, no. 1, pp. 320–332, 2020.

[47] D. Burghal, N. A. Abbasi, and A. F. Molisch, "A machine learning solution for beam tracking in mmWave systems," in *Proc. 53rd Asilomar Conf. Signals, Syst., Comput.*, 2019, pp. 173–177.

[48] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, "Deep learning coordinated beamforming for highly-mobile millimeter wave systems," *IEEE Access*, vol. 6, pp. 37328–37348, 2018.

[49] L. Sun, J. Hou, and T. Shu, "Optimal handover policy for mmWave cellular networks: A multi-armed bandit approach," in *Proc. IEEE Global Commun. Conf.*, Waikoloa, HI, USA, 2019, pp. 1–6.

[50] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, pp. 7599–7605, 2019.

[51] T. Peken, S. Adiga, R. Tandon, and T. Bose, "Deep learning for SVD and hybrid beamforming," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6621–6642, Oct. 2020.

[52] N. Wang, L. Jiao, A. Alipour-Fanid, M. Dabaghchian, and K. Zeng, "Pilot contamination attack detection for NOMA in 5G mm-Wave massive MIMO networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 15, no. 1, pp. 1363–1378, 2020.

[53] H. Zhang, H. Zhang, K. Long, and G. K. Karagiannidis, "Deep learning based radio resource management in NOMA networks: User association, subchannel and power allocation," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 4, pp. 2406–2415, Oct.–Dec. 2020.

[54] T. Bai, B. Mei, L. Zhao, and X. Wang, "Machine learning-assisted wireless power transfer based on magnetic resonance," *IEEE Access*, vol. 7, pp. 109454–109459, 2019.

[55] J. Chen *et al.*, "Hybrid beamforming/combining for millimeter wave MIMO: A machine learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11353–11368, Oct. 2020.

[56] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.

[57] M. Zhu, T.-H. Chang, and M. Hong, "Learning to beamform in heterogeneous massive MIMO networks," 2020, *arXiv: 2011.03971*. [Online]. Available: https://arxiv.org/abs/2011.0397

[58] E. M. Mohamed, "Joint users selection and beamforming in downlink millimetre-wave NOMA based on users positioning," *IET Commun.*, vol. 14, no. 8, pp. 1234–1240, 2020.

[59] Y. Chiu and K. Liu, "Beam selection and power allocation for massive connectivity in millimeter wave NOMA systems," *IEEE Access*, vol. 8, pp. 53868–53882, 2020.

[60] J. Guo, X. Zhou, and S. Durrani, "Wireless power transfer via mmWave power beacons with directional beamforming," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 17–20, Feb. 2019.

[61] K. Na, J. Kim, and Y. J. Park, "Machine learning-based power control technique for wireless power transfer," in *Proc. IEEE Wireless Power Transfer Conf.*, 2020, pp. 443–446.

[62] M. Alrabeiah, A. Hredzak, Z. Liu, and A. Alkhateeb, "ViWi: A deep learning dataset framework for vision-aided wireless communications," in *Proc. IEEE 91st Veh. Technol. Conf.*, 2020, pp. 1–5.

[63] Y. G. Lim, Y. J. Cho, M. S. Sim, Y. Kim, C. B. Chae, and R. A. Valenzuela, "Map-based millimeter-wave channel models: An overview, data for B5G evaluation and machine learning," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 54–62, Aug. 2020.

[64] A. Klautau, P. Batista, N. González-Prelcic, Y. Wang, and R. W. Heath, "5G MIMO data for machine learning: Application to beam-selection using deep learning," in *Proc. Inf. Theory Appl. Workshop Workshop*, 2018, pp. 1–9.

[65] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2009, pp. 248–255.

[66] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[67] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Representations*, 2015, pp. 1–14.

[68] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2015, pp. 1–9.

[69] K. Zheng, B. Qian, S. Li, Y. Xiao, W. Zhuang, and Q. Ma, "Long-short term echo state network for time series prediction," *IEEE Access*, vol. 8, pp. 91961–91974, 2020.

[70] W. M. Koolen and T. V. Erven, "Second-order quantile methods for experts and combinatorial games," in *Proc. 28th Conf. Learn. Theory*, Jul. 2015, vol. 40, pp. 1155–1175.

[71] F. W. *et al.*, "Collaborative spatial reuse in wireless networks via selfish multi-armed bandits," *Ad Hoc Netw.*, vol. 88, pp. 129–141, 2019.

[72] S. Hashima, K. Hatano, E. Takimoto, and E. Mahmoud Mohamed, "Neighbor discovery and selection in millimeter wave D2D networks using stochastic MAB," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1840–1844, Aug. 2020.

[73] N. C. P. Auer and P. Fischee, "Finite time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, no. 6, pp. 235–256, May 2002.

[74] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Online learning for position-aided millimeter wave beam training," *IEEE Access*, vol. 7, pp. 30507–30526, 2019.

[75] E. M. Mohamed, K. Sakaguchi, and S. Sampei, "Millimeter wave indoor channel dataset," 2021. [Online]. Available: https://dx.doi.org/10.21227/ehb1-9j69

[76] E. M. Mohamed, S. Hashima, B. M. ElHalawany, K. Hatano, and K. Wu, "MAB algorithms for single and concurrent millimeter wave indoor beamforming training," 2021. [Online]. Available: https://dx.doi.org/10.21227/7vkq-9164.

[77] E. M. Mohamed, K. Sakaguchi, and S. Sampei, "Wi-Fi coordinated WiGig concurrent transmissions in random access scenarios," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10357–10371, Nov. 2017.

[78] E. M. Mohamed, B. M. Elhalawany, H. S. Khallaf, M. Zareei, A. Zeb, and M. A. Abdelghany, "Relay probing for millimeter wave multi-hop D2D networks," *IEEE Access*, vol. 8, pp. 30560–30574, 2020.