

A Deep Learning Approach to Location- and Orientation-Aided 3D Beam Selection for mmWave Communications

Sajad Rezaie^{ID}, *Graduate Student Member, IEEE*, Elisabeth de Carvalho^{ID}, *Senior Member, IEEE*,
and Carles Navarro Manchón^{ID}

Abstract—Position-aided beam selection methods have been shown to be an effective approach to achieve high beamforming gain while limiting the overhead and latency of initial access in millimeter wave (mmWave) communications. Most research in the area, however, has focused on vehicular applications, where the orientation of the user terminal (UT) is mostly fixed at each position of the environment. This paper proposes a location- and orientation-based beam selection method to enable context information (CI)-based beam alignment in applications where the UT can take arbitrary orientation at each location. We propose three different network structures, with different amounts of trainable parameters that can be used with different training dataset sizes. A professional 3-dimensional ray tracing tool is used to generate datasets for an IEEE standard indoor scenario. Numerical results show the proposed networks outperform a CI-aided benchmark such as the generalized inverse fingerprinting (GIFP) method as well as hierarchical beam search as a non-CI-based approach. Moreover, compared to the GIFP method, the proposed deep learning-based beam selection shows higher robustness to different line-of-sight blockage probability in the training and test datasets and lower sensitivity to inaccuracies in the position and orientation information.

Index Terms—Millimeter wave, initial access, beam alignment, location-aided, orientation-aided, deep learning.

I. INTRODUCTION

EMERGENT services such as virtual and augmented reality or high definition multimedia applications are gaining increasing popularity and will be commonplace in beyond 5G systems. Their large data-rate demands can only be satisfied using large portions of available bandwidth, which millimeter wave (mmWave) bands currently offer [1]. Although higher propagation and penetration losses in mmWave bands make mmWave communication a good choice for establishing a network in small cells, those properties, besides the lack of diffraction in mmWave frequencies, make it difficult to

establish reliable mmWave links [2]. Multiple-input multiple-output beamforming is an inseparable part of mmWave communication, as it allows for compensating for the higher path loss compared to sub-6 GHz communications. To establish a high-quality directional link, transceivers need to transmit and receive over the direction of the line-of-sight (LOS) path or a powerful non-line-of-sight (NLOS) path in case the LOS is blocked. Thus the selection of appropriate beamforming and combining solutions at mmWave transceivers' antenna arrays is a crucial and challenging task [3].

Codebook-based beamforming is an attractive solution to reduce the complexity of the beam selection problem as well as the radio-frequency (RF) implementation of beamformers. It considers predefined directional beam configuration sets for precoder and combiner at access point (AP) and user terminal (UT), respectively [4]. Analog or hybrid beamforming is usually employed to construct the beams in the codebook. Although the optimal precoder and combiner configurations can be determined by exhaustively searching over all possible combinations of beamformers at the AP and UT, this results in undesirably high overhead and latency. Hierarchical beam search (HBS) is an alternative approach that reduces the search space of the beam alignment (BA) procedure by progressively sensing the environment with narrower beam widths based on the counterpart's feedback [4], [5]. However, HBS methods suffer a degradation in accuracy due to low antenna gain and insufficient spatial resolution when using large beam widths, which lead to wrong decisions in the search procedure. To address these downsides, machine learning (ML)-based methods can be used to obtain solutions that are optimized using measurement data obtained at specific deployments without the need for a model of the channel or the propagation scenario. The authors of [6] propose an ML-based framework that designs adaptive compressed sensing beamformers, which helps estimate the posterior distribution of the angle of arrival (AoA) of the dominant path. Estimates of the AoA and power of the dominant path of users are fed to support vector classifiers and neural networks to perform beam selection in a mmWave communication link in [7]. However, these methods are designed to perform initial BA in a mmWave environment assuming a single-path channel and are therefore susceptible to performance degradation in multipath conditions. As a new approach to estimate the AoA, a deep-learning

Manuscript received 11 October 2021; revised 27 April 2022; accepted 5 July 2022. Date of publication 18 July 2022; date of current version 12 December 2022. This work was supported by the Danish Council for Independent Research under Grant DFF 8022-00371B. The associate editor coordinating the review of this article and approving it for publication was A. El Gamal. (Corresponding author: Sajad Rezaie.)

The authors are with the Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark (e-mail: sre@es.aau.dk; edc@es.aau.dk; cnm@es.aau.dk).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2022.3189788>.

Digital Object Identifier 10.1109/TWC.2022.3189788

1536-1276 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

based framework is proposed in [8] that uses the unique patterns in the in-phase-quadrature representation of each beam.

Another avenue to achieve high beamforming accuracy while limiting the overhead and latency of BA is to exploit context information (CI) about the users and/or environment. In CI-based methods, different types of sensors on transceivers or out-band channel measurements are used to acquire CI, at the cost of power and device complexity [9]. As an example of such methods, the inverse fingerprinting (IFP) method in [10] recommends a candidate beam list based on the location of the user and the history of optimal beam pairs in the training examples. Furthermore, ML-based methods have been widely used to exploit context information and achieve higher accuracy owing to their unparalleled capability in dealing with non-linear problems. Reinforcement learning (RL) and deep neural networks (DNNs) are the two areas of machine learning most frequently employed in mmWave BA [11]–[13]. For instance, an RL-based method that uses historical beam management data in a specific environment to reduce the beam training overhead is presented in [14]. However, a large number of interactions with the environment are needed before a RL agent is able to learn the optimal policy [15]. A neural network (NN)-based method is proposed in [16] that uses sub-6 GHz channel measurements as input and exploits the spatial correlation between the sub-6 GHz and mmWave links to predict the best beam for data transmission in the mmWave band. For vehicular-to-everything (V2X) application, RADAR signals from joint sensing and communication functionalities at road side units are used for the beamforming of vehicular links in [17]. Also, the authors of [18] propose an ML-based structure that fuses LIDAR data and the position of vehicles to propose a more accurate beam selection method by exploiting knowledge about the obstacles in the environment. As another way of detecting static and mobile obstacles and predicting future link blockage, images from cameras placed in the environment are processed using computer vision and deep learning techniques in [13], [19], [20].

Although NN-based BA methods achieve the best performance in beam initialization, beam tracking, and blockage prediction, they require large amounts of training data at each deployment site and for each array configuration [16]. As a solution, the transfer learning technique is used in [21] to reuse the knowledge learned by the NN from a previous environment and array configuration in a new setup. Moreover, one of the key challenges for beam management in beyond 5G systems is device rotation [22]. It becomes more critical as beams in directional codebooks usually have unequal beamwidths and, therefore, the effects of rotation on the received signal strength (RSS) at each beam cannot be fully predicted by the beams' RSS before rotation. In addition, the mentioned machine learning approaches are mainly focused on the V2X application, where the orientation of the UT array can typically be inferred from the vehicle position. The authors of [23] proposed a beam management method using particle filters that exploits the knowledge of changes in the orientation of the user device to track the best beam at the user side. However, this method is limited to beam tracking only at the UT, and it

is assumed that the BS has genie-aided knowledge of the best transmit beam.

A. Contributions

To address the above-mentioned gaps, we consider the initial BA problem for mmWave communication where the UT can take an arbitrary orientation and position. This study focuses on indoor environments and handset terminals, contrary to many of the works focusing on outdoor V2X communications. We propose a deep learning-based initial BA framework that considers the handset's position and rotation. The proposed solution yields a list of recommended beam pairs that should, later, be sensed by AP and UT. This approach results in a drastic reduction of the overhead compared to exhaustive search and a significantly improved accuracy compared to HBS. The beam recommendation solution operates using only the position and orientation information of the UT handset as an input. A first version of the method was presented in [24] where an ML-based beam selection method can handle the handset's position and rotation in a 2-dimensional environment. Compared to it, this paper has the following contributions:

- 1) We generated datasets using Altair Feko-Winprop software [25] as a professional ray-tracing tool. Since UTs can take any arbitrary rotation in a 3-dimensional (3D) environment, it significantly increases the problem's non-linearity. In addition, instead of a uniform linear array (ULA), we use a uniform planar array (UPA) at both sides of the communication links to investigate the effects of rotations in 3D space on the BA of horizontal and vertical beamforming.
- 2) We present three different neural networks: single task (ST), multi-task (MT), and extended multi-task (EMT) structures to make accurate beam recommendations with different training dataset sizes. In the MT structure, we solve the problem of the AP and UT beamforming as separate tasks. We extend this idea to also separate vertical and horizontal beamforming in the EMT structure. The MT and EMT structures have significantly fewer trainable parameters than the ST one, resulting in faster performance in the training and running mode at the expense of a slight accuracy loss.
- 3) To enhance the performance of the deep learning-based method, we use the multi-labeling technique to train the NNs. This technique helps the network learn more about the environment and the possible NLOS paths at each location. The effects of the multi-labeling technique on all the three structures DNN-ST, DNN-MT, and DNN-EMT are investigated.
- 4) We evaluate the sensitivity of the data-driven-based methods to estimation errors in the location and orientation information by training and evaluating the proposed algorithms with inaccurate CI. In another experiment, we assess the robustness of the proposed deep learning-based methods against mismatch between the channel conditions present in the training and test datasets.

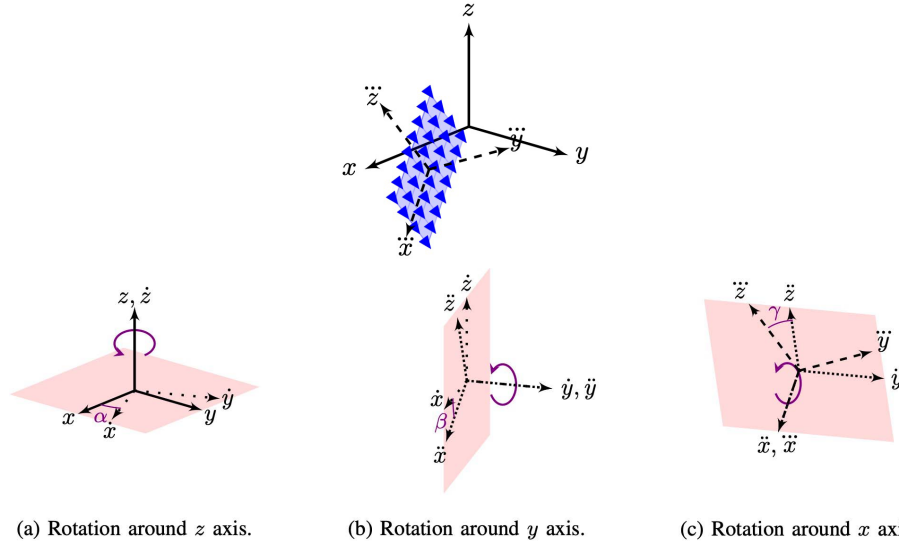


Fig. 1. By rotating around the x , y , and z axes, a transition from the global coordinate system (xyz) to any arbitrary local coordinate system $(\ddot{x}\ddot{y}\ddot{z})$ is possible. The red rectangles show the rotation planes at three rotation steps. The coordinate systems $(\dot{x}\dot{y}\dot{z})$ and $(\ddot{x}\ddot{y}\ddot{z})$ show the rotated versions of (xyz) after the first and second rotation around z and \dot{y} axes, respectively. (a) Rotation around z axis. (b) Rotation around y axis. (c) Rotation around x axis.

5) Our proposed method's performance is compared against several benchmarks. Besides ideal BA, we generalize the inverse fingerprinting method in [10] to deal also with arbitrary orientation of user handset. In order to compare with a non-CI based method, we extend the deactivation (DEACT) method in [4] to perform 3D HBS using UPAs. The proposed deep learning-based beam selection methods significantly outperform the generalized inverse fingerprinting (GIFP) and DEACT methods in terms of accuracy and latency of initial access BA.

B. Organization and Notations

First, the system model, channel model, and codebook definition are described in Section II. The general description of the beam alignment problem and two possible approaches for dealing with the problem are given in Section III. In Section IV, we present the three proposed structures for the DNN. The simulation setup and numerical results are presented in Section V. Conclusions are drawn in Section VI.

In this paper, \mathbb{R} and \mathbb{C} respectively denote the fields of real and complex numbers. \mathcal{A} denotes a finite set, with $|\mathcal{A}|$ being its cardinality. Also, a , \mathbf{a} , and \mathbf{A} denote a scalar, a vector, and a matrix, respectively, with a_i being the i th entry of the column vector \mathbf{a} , and $A_{i,j}$ is the entry in the i th row and j th column of the matrix \mathbf{A} . Transposition and complex transposition of vectors and matrices are respectively represented as $(\cdot)^T$ and $(\cdot)^H$, and \otimes is the Kronecker product. In addition, $\arg \max_{i,j} A_{i,j}$ is the index of the maximum entry of matrix \mathbf{A} as a tuple, and $\arg \text{sort}_{i,j} A_{i,j}$ denote the list of indices (as tuples) of entries of matrix \mathbf{A} sorted in descending order.

II. SYSTEM MODEL

We consider a downlink communication system consisting of a fixed AP and a mobile UT operating in the mmWave

frequency band. Both AP and UT are equipped with UPAs and are placed in a 3D propagation scenario. The received signal at the UT may be written as

$$y = \sqrt{P_{\text{AP}}} \mathbf{v}^H \mathbf{H} \mathbf{u} s + \mathbf{v}^H \mathbf{n} \quad (1)$$

where \mathbf{u} and \mathbf{v} denote the precoder and combiner at AP and UT, respectively. Also, P_{AP} , $s \in \mathbb{C}$, and $\mathbf{n} \in \mathbb{C}^{N_{\text{UT}}}$, respectively, are the transmission power, the transmitted symbol with unit power, and a complex Gaussian noise vector with zero mean and variance σ_n^2 .

We assume that AP and UT UPAs are made respectively of $\{N_{\text{APH}}, N_{\text{APV}}\}$ and $\{N_{\text{UTH}}, N_{\text{UTV}}\}$ elements in horizontal and vertical dimensions, with the elements spaced half a wavelength apart in both dimensions. The total number of antenna elements at AP and UT is $N_{\text{AP}} = N_{\text{APH}} N_{\text{APV}}$ and $N_{\text{UT}} = N_{\text{UTH}} N_{\text{UTV}}$, respectively. We consider a global coordinate system (GCS) for the environment, and the positions and orientations of AP and UT are defined in the GCS. In addition, both AP and UT have their own local coordinate systems (LCS), shown in Fig. 1. The LCS is selected such that the UPA is oriented parallel to the xz plane. We assume the AP is placed at fixed position $\mathbf{p}_{\text{AP}} = (x_{\text{AP}}, y_{\text{AP}}, z_{\text{AP}}) \in \mathbb{R}^3$ with fixed angles of antenna arrays $\psi_{\text{AP}} = (\alpha_{\text{AP}}, \beta_{\text{AP}}, \gamma_{\text{AP}})$ rotated around z , y , and x axes, respectively. Also, the UT can be placed randomly in the environment at position $\mathbf{p}_{\text{UT}} = (x_{\text{UT}}, y_{\text{UT}}, z_{\text{UT}}) \in \mathbb{R}^3$ with orientation $\psi_{\text{UT}} = (\alpha_{\text{UT}}, \beta_{\text{UT}}, \gamma_{\text{UT}})$ where the orientation angles are uniformly random in the ranges $\alpha_{\text{UT}} \in [-\pi, \pi)$, $\beta_{\text{UT}} \in [-\pi/4, \pi/4)$, and $\gamma_{\text{UT}} \in [-\pi/4, \pi/4)$. The relation between the GCS, LCS and the orientation angles is illustrated in Fig. 1. This study assumes that all the devices share the same reference system for estimating the device orientation.

A. Channel Model

We use Altair WinPropTM software package [25] to generate channel responses at each UT point in the environment based

on the ray-tracing technique. The outputs of the ray-tracing tool such as the angle of departure (AoD), AoA, path gains, etc. for all paths between the AP and each user point are reported as results of the ray-tracing tool. We collect channel responses of UT points at each scene of the environment. The dataset generation steps and assumptions are explained in detail in Section IV-A.

By applying a narrow band channel model on the contributions of one LOS and L NLOS paths, the channel matrix $\mathbf{H} \in \mathbb{C}^{N_{\text{UT}} \times N_{\text{AP}}}$ between the AP and UT is modeled as

$$\mathbf{H} = \sum_{l=0}^L \sqrt{\rho_l} e^{j\vartheta_l} \mathbf{a}_{\text{UT}}(\phi_l^A, \theta_l^A) \mathbf{a}_{\text{AP}}^H(\phi_l^D, \theta_l^D) \quad (2)$$

where ρ_l and ϑ_l are the receive power and the phase of the l th path, respectively. Also, ϕ_l^A and θ_l^A denote the azimuth and elevation AoAs with respect to the UT array plane. Likewise, ϕ_l^D and θ_l^D are the azimuth and elevation AoDs with reference to the AP array plane. The antenna array response of the UT and AP, respectively $\mathbf{a}_{\text{UT}} \in \mathbb{C}^{N_{\text{UT}} \times 1}$ and $\mathbf{a}_{\text{AP}} \in \mathbb{C}^{N_{\text{AP}} \times 1}$, can be written as

$$\mathbf{a}_{\text{UT}}(\phi_l^A, \theta_l^A) = \frac{1}{\sqrt{N_{\text{UT}}}} \mathbf{a}_{\text{E}}(N_{\text{UTV}}, \theta_l^A) \otimes \mathbf{a}_{\text{A}}(N_{\text{UTH}}, \phi_l^A, \theta_l^A), \quad (3)$$

$$\mathbf{a}_{\text{AP}}(\phi_l^D, \theta_l^D) = \frac{1}{\sqrt{N_{\text{AP}}}} \mathbf{a}_{\text{E}}(N_{\text{APV}}, \theta_l^D) \otimes \mathbf{a}_{\text{A}}(N_{\text{APH}}, \phi_l^D, \theta_l^D) \quad (4)$$

where $\mathbf{a}_{\text{A}} \in \mathbb{C}^{N \times 1}$ and $\mathbf{a}_{\text{E}} \in \mathbb{C}^{N \times 1}$ are

$$\mathbf{a}_{\text{A}}(N, \phi, \theta) = [1, e^{j\pi \sin(\theta) \cos(\phi)}, \dots, e^{j\pi(N-1) \sin(\theta) \cos(\phi)}]^T, \quad (5)$$

$$\mathbf{a}_{\text{E}}(N, \theta) = [1, e^{j\pi \cos(\theta)}, \dots, e^{j\pi(N-1) \cos(\theta)}]^T. \quad (6)$$

B. Beam Codebook

We consider analog phased antenna array at both AP and UT. Analog electrically controlled phase shifters are passive devices that can control the signal phase at each array element. We use discrete Fourier transform (DFT)-based codebook to simplify the beamforming procedure where there is one RF chain at each transceiver side. The analog precoder at the AP and the analog combiner at the UT are defined as

$$\mathbf{u}_{p,q} = \mathbf{a}_{\text{AP}}(\phi_p^D, \theta_q^D), \quad p \in \{1, \dots, N_{\text{APH}}\}, \quad q \in \{1, \dots, N_{\text{APV}}\}, \quad (7)$$

$$\mathbf{v}_{m,n} = \mathbf{a}_{\text{UT}}(\phi_m^A, \theta_n^A), \quad m \in \{1, \dots, N_{\text{UTH}}\}, \quad n \in \{1, \dots, N_{\text{UTV}}\} \quad (8)$$

where ϕ_p^D , θ_q^D , ϕ_m^A and θ_n^A are

$$\phi_p^D = \arccos \frac{2p-1-N_{\text{APH}}}{N_{\text{APH}}}, \quad \theta_q^D = \arccos \frac{2q-1-N_{\text{APV}}}{N_{\text{APV}}}, \quad (9)$$

$$\phi_m^A = \arccos \frac{2m-1-N_{\text{UTH}}}{N_{\text{UTH}}}, \quad \theta_n^A = \arccos \frac{2n-1-N_{\text{UTV}}}{N_{\text{UTV}}}. \quad (10)$$

By mapping all tuples (p, q) and (m, n) into the sets of integers $\{1, 2, \dots, N_{\text{AP}}\}$ and $\{1, 2, \dots, N_{\text{UT}}\}$, the sets

$\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_{N_{\text{AP}}}\}$ and $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_{N_{\text{UT}}}\}$ indicate all available analog precoders and combiners at the AP and the UT, respectively. By employing the precoder \mathbf{u}_i and combiner \mathbf{v}_j at the transceivers, the (i, j) th entry of the RSS matrix, $\mathbf{R} \in \mathbb{R}^{N_{\text{AP}} \times N_{\text{UT}}}$, can be set as

$$R_{i,j} = \left| \sqrt{P_{\text{AP}}} \mathbf{v}_j^H \mathbf{H} \mathbf{u}_i s + \mathbf{v}_j^H \mathbf{n} \right|^2. \quad (11)$$

In the BA phase, $s \in \mathbb{C}$ is a pilot symbol with unit power.

III. BEAM ALIGNMENT PROCEDURES

A. Context-Aware Beam Selection

The position and orientation of the UT can be used to predict the LOS direction. Although knowledge of the LOS direction is valuable, it is insufficient due to the frequent blockage occurrence in mmWave communications. Besides, the direction and strength of NLOS paths are environment-dependent. Information about the properties of the propagation environment can be obtained from extensive measurements obtained in the specific environment. It is important to note that obtaining the location and orientation information of the device comes at a price, and it may imply additional overhead or implementation cost to the system. However, we expect that the introduction of joint communication, sensing, and localization in upcoming beyond 5G systems will facilitate position- and orientation-aided BA [26]. In Section IV, we discuss two data-driven methods that use position and orientation information to recommend a list of beam pair candidates.

Consider \mathcal{B} as a set of all possible combinations of precoders and combiners in the transceivers. In the exhaustive search approach, the environment is sensed with all members of \mathcal{B} , but this yields unacceptably high overhead [3]. Context information from the environment and the transceivers can boost the sensing phase of BA. In particular, context information such as geometrical properties of environment, position and orientation information of transceivers, etc. can be used to reduce the space of sensed precoders and combiners. Thus, in this approach, a candidate list \mathcal{S} including a few beam pair indices is proposed based on the context information, where $|\mathcal{S}| \ll |\mathcal{B}|$. The probability of including the optimal beam pair increases with the cardinality of \mathcal{S} , at the expense of increased latency in the beam-alignment process.

Fig. 2(a) shows the beam selection process including the beam-training request and feedback in sub-6 GHz communication links. First, a beam-training request is sent from the UT, including the location and orientation information of the UT array. The AP uses the provided information to recommend a candidate beam list \mathcal{S} . Subsequently, the AP shares the beam list \mathcal{S} with the UT. Then, as the AP and UT know their beamformer at each time interval, they start to sense all the beam pairs in \mathcal{S} . Then, the UT feeds back the index of the beam pair which yielded the largest RSS. Note that in this work we assume that all control information exchanged over the sub-6 GHz link is conveyed error-free.

B. Hierarchical Beam Search

In contrast to the context-aware beam selection approach, the hierarchical beam search method does not need extra

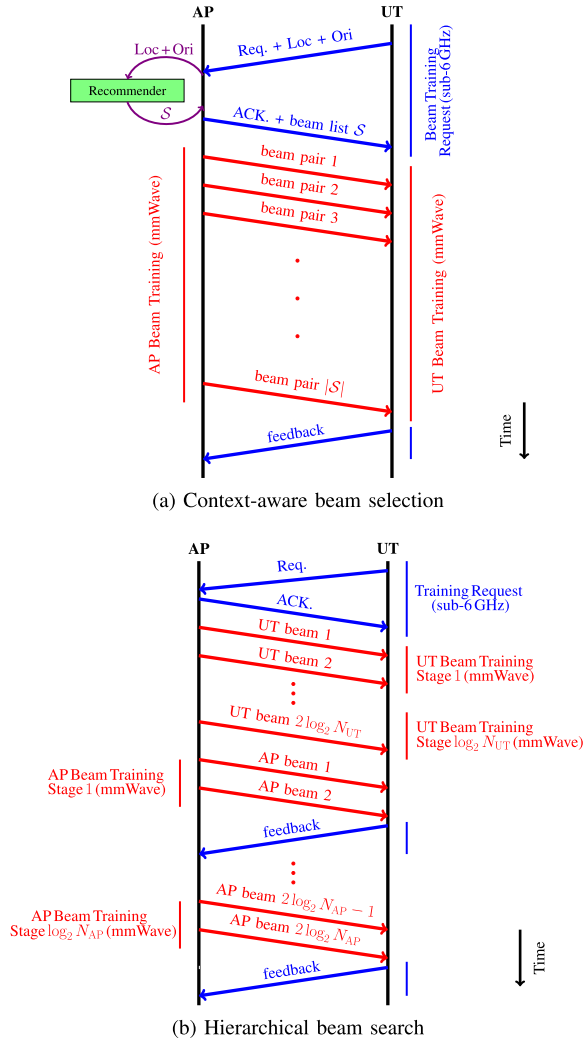


Fig. 2. The mmWave beam training procedure for the context-aware beam selection and hierarchical beam search approaches. In the former, the beam list is proposed based on the UT position and orientation.

information about the environment or transceivers. In HBS, the AP and UT sense the environment with wide beams and gradually narrow down the beam-width to find the best beam pairs for transmission [3], [4]. As it is shown in Fig. 2(b), at the first stages AP transmits omnidirectionally, and UT finds the best combiner. Subsequently, AP tries to find its best precoder. At each step of the hierarchical search, it needs feedback from UT to find out the right direction for further search. Thus, $\log_2 N_{AP}$ feedbacks are needed to be transmitted in HBS [4]. Again, we assume sub-6 GHz feedback to be error-free.

IV. DATA-DRIVEN BEAM SELECTION

In this section, we discuss the data collection phase as an important step for data-driven based methods and how the labeling can be done for different purposes. Then, we propose the deep-learning based beam selection methods with three novel structures: single-task, multi-task, and extended multi-task. In addition, the motivations for proposing each structure are explained.

A. Dataset Structure and Construction

As mentioned in the previous section, the properties of the propagation environment can be extracted from measurements in the specific site. The measurements captured in the particular environment can be collected in a dataset consisting of a set of inputs and outputs. Here, the position and orientation of measurements are the inputs, and the labels indicating the preferred beams for each input are the outputs.

In the data collection phase, the UT takes an arbitrary location and orientation at each scene of the environment. At each UT position and orientation, the measurements are captured by sensing the environment with all the possible combinations of precoders \mathcal{U} and combiners \mathcal{V} at the transceivers. Thus, the entries of \mathbf{R} in (11) are set using the RSS measured using all beam pairs in the codebook. The entries of \mathbf{R} are used to set the labels used in the collected datasets. In our work, we consider two types of label encoding options: the one-hot encoding scheme, and the multi-labeling scheme. In the one-hot encoding scheme, the beam pair with the highest RSS is marked in \mathbf{L} as the label matrix, where

$$L_{i,j}^{(1)} = \begin{cases} 1, & \text{if } (i,j) = \arg \max_{m,n} R_{m,n}, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

In this case, in blockage-free situations the beam pair which covers the LOS path is marked.¹ When blockage occurs, we mark the beam pair covering the strongest NLOS path.

The one-hot encoding scheme labels only the strongest beam pair for each position and orientation, while all other beam pairs are treated equally. Since the purpose of the recommender in data-driven methods is to propose a list of several candidate beams, it may be useful to, instead, label multiple strong beam pairs as relevant, such that the recommender has information about the M -best beam pairs, rather than just the strongest one. As an example, in a LOS situation such labeling would allow marking both the beam pair capturing the LOS path and a few beam pairs aligned in the direction of strong NLOS paths. This would allow the recommender to, after training, provide a list of alternative candidate beams that can be used in the event that the LOS is blocked. Hence, in the M -hot encoding scheme ($M \in \{1, 2, \dots, N_{AP}N_{UT}\}$), we mark the M best beam pairs with the M highest RSS in the label matrix $\mathbf{L}^{(M)}$ as

$$\mathbf{r} = \arg \text{sort}_{i,j} (R_{i,j}), \quad (13a)$$

$$\mathcal{T} = \{r_k | k = 1, \dots, M\}, \quad (13b)$$

$$L_{i,j}^{(M)} = \begin{cases} 1, & \text{if } (i,j) \in \mathcal{T}, \\ 0, & \text{otherwise.} \end{cases} \quad (13c)$$

Therefore, each sample of the dataset contains the receivers' coordinates (\mathbf{p}_{UT}) and orientations (ψ_{UT}), the RSS matrix (\mathbf{R}), the type of encoding scheme (M), and the label matrix ($\mathbf{L}^{(M)}$). We represent the dataset as $\mathbb{D}^\Xi = \{(\mathbf{x}_m, \mathbf{L}_m^{(M)})\}, m = 1, \dots, N^\Xi$, where $\mathbf{x} = [x_{UT}, y_{UT}, z_{UT}, \alpha_{UT}, \beta_{UT}, \gamma_{UT}]$ and N^Ξ shows the dataset

¹The presence of measurement noise may impair labeling the LOS direction.

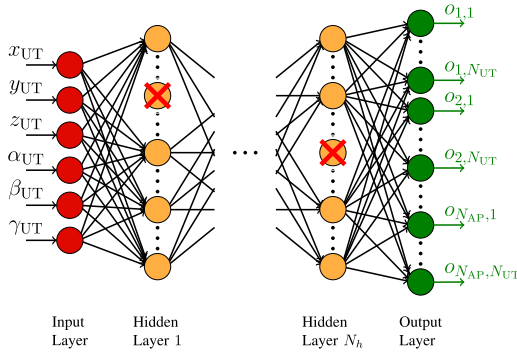


Fig. 3. The proposed DNN architecture of the beam selection method using position and orientation information. We refer to this network with the term “single-task structure” as $N_{AP}N_{UT}$ neurons at its output layer correspond to all possible beam pairs.

size. Indeed, $\mathbf{L}^{(M)}$ implicitly holds information about the propagation environment, as it is the post-processed form of \mathbf{R} . Thus, for given \mathbf{x} the dataset contains information about the M strongest beam pairs for transmission, which can be exploited by probabilistic or machine learning based methods [24].

B. Proposed Deep-Learning Based Methods

To extract the knowledge about the environment and beam patterns at the transceivers from the training data, probabilistic and machine learning based approaches are the two most popular strategies. In Section IV-C1, we discuss a probabilistic beam selection method as a baseline to the proposed methods. Here, we discuss the motivations to use machine learning in beam selection and introduce our machine learning based beam selection methods.

As static and mobile scatterers in the site may change the propagation properties of the environment by blocking or reflecting paths with high power, beam selection leveraging both location and orientation information becomes a highly nonlinear classification problem. DNNs have shown remarkable achievements in learning complex nonlinear input-output relationships using training data in different applications [27]. Therefore, we choose DNNs as classifiers to predict the best beam pairs for transmission.

For training and evaluation of all the three proposed network structures, we use the labeling procedure described in (12) and (13). Categorical cross entropy is used as the loss function, \mathcal{L} , for training of all the three network structures, i.e.,

$$\mathcal{L}(\mathbf{o}, \mathbf{L}) = - \sum_{i=1}^{N_{AP}} \sum_{j=1}^{N_{UT}} L_{i,j} \log(o_{i,j}) \quad (14)$$

where $o_{i,j}$ and $L_{i,j}$ denote, respectively, the neural network's output and the label corresponding to the (i, j) th beam pair.

1) *Single-Task Structure*: As shown in Fig. 3, the coordinates and orientations of the UT, $\mathbf{x} = [x_{UT}, y_{UT}, z_{UT}, \alpha_{UT}, \beta_{UT}, \gamma_{UT}]$, are fed as inputs to a feed-forward, fully connected, DNN. As the AP has fixed position and orientation, there is no need to explicitly feed this information to the NN. There are $N_{AP} \times N_{UT}$ neurons at the output layer, which correspond to all possible combinations of the N_{AP} and N_{UT} beams at the AP and UT, respectively. The network has N_h

hidden layers with n_h neurons at each hidden layer. Neurons at the i th layer of the network, \mathbf{y}^i , are calculated as a non-linear function of all the neurons at the previous layer, \mathbf{y}^{i-1} , and the trainable parameters of the i th layer:

$$\mathbf{y}^i = \sigma(\mathbf{W}_i^c \mathbf{y}^{i-1} + \mathbf{w}_i^b), \quad i \in \{1, \dots, N_h + 1\} \quad (15)$$

where σ , \mathbf{W}_i^c and \mathbf{w}_i^b are the activation function, weights and biases of the i th layer, respectively. As it is illustrated in Fig. 3, \mathbf{y}^0 and \mathbf{y}^{N_h+1} are equal to the input and output vectors, \mathbf{x} and \mathbf{o} , respectively. We define $\mathbf{W}_i = \{\mathbf{W}_i^c, \mathbf{w}_i^b\}$ which includes all the trainable parameters in the i th layer. We use *tanh* function as the non-linear activation function of the hidden layers. By using *softmax* function as the activation function at the output layer, each network's output represents the predicted probability of each beam pair being the best, meaning providing the largest RSS. The outputs of the network, \mathbf{o} , can be expressed as a non-linear function of the input and the weights of different layers:

$$\mathbf{o} = f_{\mathbf{W}_o}^{(o)} \left(f_{\mathbf{W}_{N_h}}^{(N_h)} (\dots f_{\mathbf{W}_1}^{(1)}(\mathbf{x})) \right) \quad (16)$$

where $f_{\mathbf{W}_o}^{(o)}$ and $f_{\mathbf{W}_i}^{(i)}$, $i = 1, \dots, N_h$ are the non-linear functions of the output layer and the i th hidden layer, respectively. With the datasets described in Section IV-A, the network can be trained using a stochastic gradient descent algorithm to learn the most favorable beam pairs for any position in the environment

To prevent overfitting, the dropout technique is used as a regularization mechanism at hidden layers [28]. In this technique, some neurons of the hidden layers are randomly dropped in each training data batch. Dropout helps the network to avoid memorizing training data by effectively using different network structures in the learning process. In Fig. 3, we show the dropped out neurons in the hidden layers with red cross marks.

Once the model has been trained, it can be used to recommend a set of beam pairs by feeding the UT coordinates and orientation to the network as inputs. A list of recommended beams \mathcal{S} is obtained from the output values \mathbf{o} , which represent the probability of each beam pair in the codebooks being the best for the input position and orientation. If the sensing phase of the environment is limited by N_b measurements, the N_b beam pairs with highest probabilities are chosen as candidates as follows:

$$\mathbf{g} = \underset{i,j}{\text{arg sort}} (o_{i,j}), \quad (17a)$$

$$\mathcal{S} = \{g_k | k = 1, \dots, N_b\}, \quad (17b)$$

where $|\mathcal{S}| = N_b$.

2) *Multi-Task Structure*: Although the single-task structure of the NN can propose good beam candidates for a given UT position and orientation, it needs large training datasets to be trained well due to the large number of trainable weights it contains. By inspecting the number of trainable parameters at each layer, we see that there are $7n_h$, $(n_h + 1)n_h$, and $(n_h + 1)N_{AP}N_{UT}$ trainable weights, respectively, at the first hidden layer, other hidden layers, and the output layer. For large number of antenna elements, it is likely that $N_{AP}N_{UT} \gg n_h$.

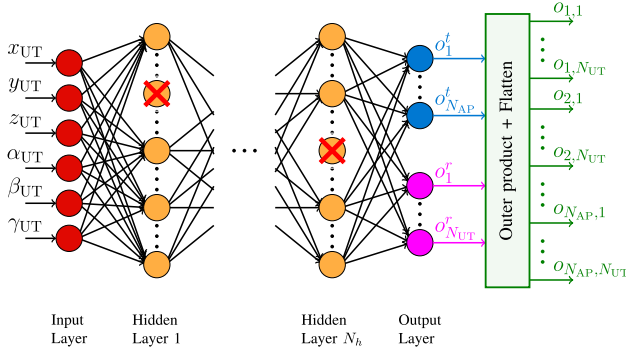


Fig. 4. Structure of the proposed multi-task DNN. We coin this network a multi-task DNN since it includes two separate sets of neurons at the output layer: each set represents the selected AP and UT beams, respectively. This results in $N_{AP} + N_{UT}$ neurons at the output layer, leading to fewer trainable parameters than the single-task structure.

In this case, the output layer contains most of the trainable parameters. By changing the structure of the output layer, we can decrease the number of trainable weights significantly.

We propose a network structure that attempts to solve the AP and UT beamformer selection as separate problems or tasks. As shown in Fig. 4, we consider two separate sets of neurons at the output layer, with different tasks for each branch: beam selection at the AP and beam selection at the UT. The two sets of neurons, $\mathbf{o}^t = [o_1^t, \dots, o_{N_{AP}}^t]^T$ and $\mathbf{o}^r = [o_1^r, \dots, o_{N_{UT}}^r]^T$, show the probability of each beam being the best at AP and UT for a given UT position and orientation. As the network aims to propose good beam pairs to be sensed, we synthetically combine the probabilities of beams at the AP and the UT to construct probabilities of all possible beam pairs. Following this simplifying assumption, we use the outer product to calculate the probability of each beam pair being the best, $o_{i,j}$, i.e.

$$o_{i,j} = o_i^t o_j^r, \quad i \in \{1, \dots, N_{AP}\}, \quad j \in \{1, \dots, N_{UT}\}. \quad (18)$$

As the outer product of two vectors \mathbf{o}^t and \mathbf{o}^r is a matrix by size $N_{AP} \times N_{UT}$, we flatten the result of the outer product to obtain a vector with $N_{AP}N_{UT}$ entries, \mathbf{o}^2 . As it is clear in (18), there is no trainable parameter in the construction of probabilities of beam pairs. As the output layer has $N_{AP} + N_{UT}$ neurons, the number of trainable weights at the output layer is $(n_h + 1)(N_{AP} + N_{UT})$. In practical scenarios where the training dataset is limited, the multi-task structured network can be trained better than the single-task structured one, as the former has far fewer trainable parameters. In the Appendix, we discuss the relation between the training of the multi-task network with outer-product and that of classical multi-task classifiers.

3) *Extended Multi-Task Structure*: To further decrease the number of trainable weights, we consider the assumption that the horizontal and vertical beamforming are independent at both AP and UT. Based on this simplifying assumption, we propose an extended version of the multi-task structured network, which is shown in Fig. 5. There are 4 separate sets of neurons for horizontal and vertical beamforming at AP and

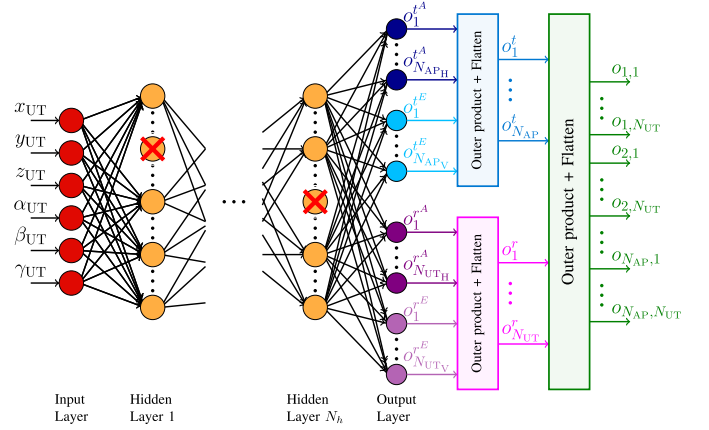


Fig. 5. The extended multi-task structured DNN with $N_{AP_H} + N_{AP_V} + N_{UT_H} + N_{UT_V}$ neurons at its output layer.

UT. Similar to the previous structure, we use the outer product to obtain the probabilities of being the best for all possible beams at transceivers, i.e.,

$$o_i^t = o_m^t o_n^t, \quad i \in \{1, \dots, N_{AP}\}, \quad m = \left\lfloor \frac{i}{N_{AP_H}} \right\rfloor, \quad n = i \bmod N_{AP_H}, \quad (19)$$

$$o_j^r = o_p^r o_q^r, \quad j \in \{1, \dots, N_{UT}\}, \quad p = \left\lfloor \frac{j}{N_{UT_H}} \right\rfloor, \quad q = j \bmod N_{UT_H}, \quad (20)$$

where o_m^t , o_n^t , o_p^r , and o_q^r denote the probability of m th, n th, p th, and q th beams being the best at horizontal and vertical beamforming at AP and UT, respectively. As an example, neuron o_m^t denotes the output probability for the AP beams in the m th azimuth angle of the codebook, while o_n^t represents the probability for the AP beams pointing at the n th elevation angle. As in the new structure the output layer includes $N_{AP_H} + N_{AP_V} + N_{UT_H} + N_{UT_V}$ neurons, the number of trainable parameters in the output layer decreases to $(n_h + 1)(N_{AP_H} + N_{AP_V} + N_{UT_H} + N_{UT_V})$. Such a reduction is advantageous when the amount of training data is limited.

C. Baselines

1) *Generalized Inverse Fingerprinting Method*: The IFP is a data-driven method that uses training samples to make a lookup table of best beam pairs for different environment locations. However, the IFP method in [10] uses only the position information of the UT, not considering the arbitrary orientation of the UT. The bin definition of the IFP method is extended in a way that considers the rotation of the UT array about x, y , and z axes. This is a straightforward extension of the GIFP algorithm we proposed in [24]. By extending discretization to both the spatial and angular domains, position and orientation of the UT, \mathbf{p}_{UT} and ψ_{UT} respectively, determine the corresponding bin for each observation. The k th bin may be defined as

$$\begin{aligned} B_k = & [x_k, x_k + \Delta_s) \times [y_k, y_k + \Delta_s) \times [z_k, z_k + \Delta_s) \\ & \times [\alpha_k, \alpha_k + \Delta_a) \times [\beta_k, \beta_k + \Delta_a) \times [\gamma_k, \gamma_k + \Delta_a) \end{aligned} \quad (21)$$

²The operation “outer product + flatten” is equivalent to the Kronecker product, i.e. $\mathbf{o} = \mathbf{o}^t \otimes \mathbf{o}^r$.

where Δ_s and Δ_a , as hyper parameters of the generalized inverse fingerprinting (GIFP) method, are spatial bin size (SBS) and angular bin size (ABS), respectively. The values $x_k, y_k, z_k, \alpha_k, \beta_k$, and γ_k are selected to obtain a set of bins that have no overlap and cover all the spatial and angular coordinates of interest. Note that the bin definition of the IFP method in [10] can be recovered by setting $\Delta_a = 2\pi$.

Following the IFP method approach in proposing a candidate list for the k th bin, \mathcal{S}_k , the beam pairs of this list are chosen to minimize the misalignment probability, i.e., the probability of not including the beam pair with the highest RSS in the list. The misalignment probability may be expressed as

$$P_m(\mathcal{S}_k) = \mathbb{P} \left[\max_{(i,j) \in \mathcal{S}_k} R_{i,j} < \max_{(p,q) \in \mathcal{B}} R_{p,q} \right], \quad (22)$$

where \mathcal{B} is a set including all possible beam pairs. Using the training samples assigned to the k th bin, the members of \mathcal{S}_k are the top N_b ranked beam pairs which have the highest frequencies of being best [10]. In test mode, for a given new UT coordinate and orientation, the candidate list of the associated bin is used, and the beam selection follows the procedure in Fig. 2(a). Because of the discretization in the angular domain added to the spatial domain, there are significantly more bins in GIFP compared to classical IFP. Thus, the GIFP method becomes more data-demanding, and it needs a much larger training dataset to have the same density of training samples over bins as IFP.

2) *Hierarchical Beam Search*: As mentioned before, in hierarchical beam search, the space is covered and sensed in several stages with different beam width. For a UPA with $N = N_h \times N_v$ antennas, there are $\log_2(N_h) + 1$ and $\log_2(N_v) + 1$ levels in horizontal and vertical directions. At the k_h th and k_v th horizontal and vertical level, there are $N_{k_h} \times N_{k_v}$ codewords, where $N_{k_h} = 2^{k_h}$ and $N_{k_v} = 2^{k_v}$. By extending the DEACT approach proposed in [4], the n_h th and n_v th codeword in the k_h th and k_v th horizontal and vertical level of the codebook can be written as

$$\begin{aligned} \mathbf{w}(k_h, k_v, n_h, n_v) &= [\mathbf{a}_E^T(N_{k_v}, \theta_v), \mathbf{0}_{(N_v - N_{k_v}) \times 1}^T]^T \\ &\quad \otimes [\mathbf{a}_A^T(N_{k_h}, \phi_h, \theta_v), \mathbf{0}_{(N_h - N_{k_h}) \times 1}^T]^T, \\ \phi_h &= \arccos \frac{2n_h - 1 - N_{k_h}}{N_{k_h}}, \\ \theta_v &= \arccos \frac{2n_v - 1 - N_{k_v}}{N_{k_v}}, \end{aligned} \quad (23)$$

where \mathbf{a}_A and \mathbf{a}_E are defined in (5) and (6). Algorithm 1 illustrates the procedure of horizontal and vertical beamforming using a 3D codebook. The device first searches the vertical codewords while considering omnidirectional horizontal beamforming during this part. Then, the device tries to find the best horizontal codeword with the knowledge of the best vertical one. At each stage of the HBS, the beam decision region of a codeword is the subspace in the angular domain where the specified codeword has the highest gain among all the possible codewords. As an example, Fig. 6 shows the beam decision regions for different levels of the DEACT method for a UPA with $N_h = 4$ and $N_v = 4$.

Algorithm 1 Vertical and Horizontal Analog Beamforming Based on Codebook

Input: N_h, N_v, \mathbf{w}

Result: n_h, n_v

Initialize: Omnidirectional horizontal beamforming ($k_h = 0$), $n_h = 1$, $n_v = 1$

for $k_v = 1$ *to* $\log_2(N_v) + 1$ **do**

1. $n_v^{(1)} = 2n_v - 1$, $n_v^{(2)} = 2n_v$

2. Sensing the environment with the two selected codewords: $\mathbf{w}(k_h, k_v, n_h, n_v^{(1)})$ and $\mathbf{w}(k_h, k_v, n_h, n_v^{(2)})$

if $\mathbf{w}(k_h, k_v, n_h, n_v^{(1)})$ provides higher RSS **then**
 $n_v = n_v^{(1)}$;

else $n_v = n_v^{(2)}$;

end

for $k_h = 1$ *to* $\log_2(N_h) + 1$ **do**

1. $n_h^{(1)} = 2n_h - 1$, $n_h^{(2)} = 2n_h$

2. Sensing with the two selected codewords: $\mathbf{w}(k_h, k_v, n_h^{(1)}, n_v)$ and $\mathbf{w}(k_h, k_v, n_h^{(2)}, n_v)$

if $\mathbf{w}(k_h, k_v, n_h^{(1)}, n_v)$ provides higher RSS **then**
 $n_h = n_h^{(1)}$;

else $n_h = n_h^{(2)}$;

end

V. PERFORMANCE AND COMPLEXITY EVALUATION

We start this section by introducing the ray-tracing setup. Following we compare the performance of the methods in terms of misalignment probability and effective spectral efficiency (ESE). In order to quantify the training data requirements of each of the analyzed data-driven methods, we investigate their performance when they are trained using datasets of different sizes. Finally, we evaluate the robustness of the deep learning based and GIFP methods with inaccurate position and orientation information.

A. Simulation Setup and Performance Measures

We consider the living room (LR) described and defined precisely in the IEEE 802.11ad task group [29]. Fig. 7 illustrates the LR as the considered indoor environment. According to the standard, the LR dimensions are $7\text{m} \times 7\text{m} \times 3\text{m}$ ($W \times L \times H$). Two sofas, a table, and an armchair are placed in the LR, and a cabinet is placed between two windows of one of the outer walls. The AP is placed in the middle of one of the LR walls. The UT can take a position in a sector with the same height as the AP, where the sector has dimensions of $4\text{m} \times 7\text{m}$ ($W \times L$). In the ray tracing tool, we defined 70,000 UT positions in the user sector. In the standard, a cluster blockage model is considered where a NLOS path can be blocked with a probability that is defined in the standard. This probability depends on the order of reflection and the reflective surfaces. In addition, the probability of LOS blockage is considered a model parameter that can be set to 0 or 1. To construct datasets with arbitrary LOS blockage probability, we can combine the samples proportionally from datasets with 0 and 1 LOS

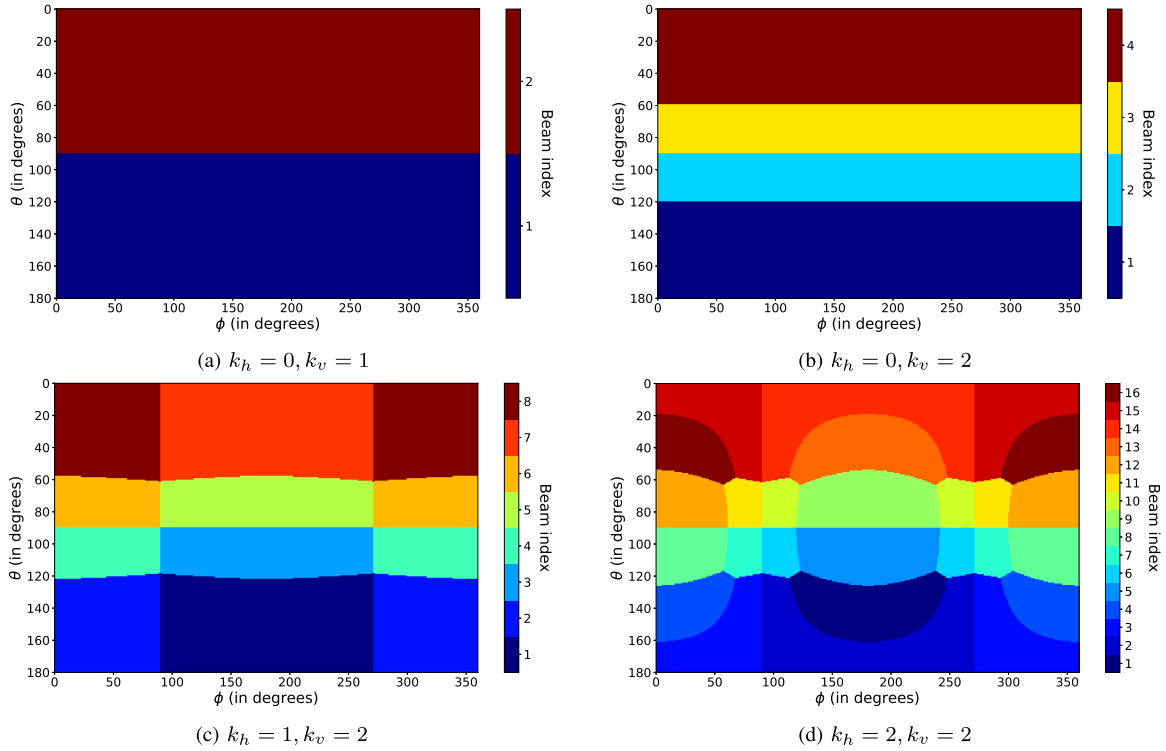


Fig. 6. Beam decision regions in different levels of the DEACT method as a hierarchical beam search. The antenna array size is $\{N_h = 4, N_v = 4\}$. At each level of the search, different colors are representing different codewords in the codebook.

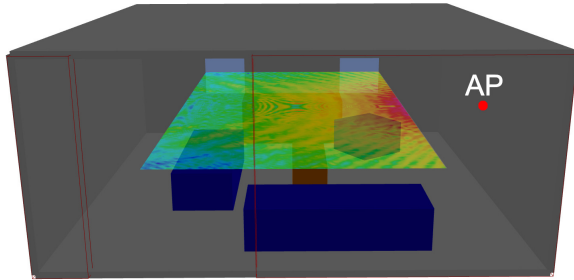


Fig. 7. The ray-tracing simulation of the living room as the simulation scenario. The AP position is shown with the red circle. Also, the received LOS power is depicted in the user grid of the scenario.

TABLE I
GEOMETRICAL INFORMATION OF THE LIVING ROOM

	Dimension (m)	AP's pos. (m)	User grid (m)	# UT positions
LR	$7 \times 7 \times 3$	(7, 3.5, 1.5)	$[1.5 - 5.5] \times [0 - 7] \times [1.5]$	70,000

blockage probability. Table I gives a summary of the most important information about the LR.

To generate channel responses, we use Altair Feko-Winprop software [25] as a professional ray-tracing software, which considers empirical losses for transmission, reflection and diffraction in the simulations. Using the outputs of the ray-tracing tool including the AoD, AoA, and path gains of all the paths, the channel response can be constructed using (2). The 25 strongest multipath components are used at any UT position to generate the channel response.

For all samples of the dataset, we consider $P_{AP} = 0$ dBm and $\sigma_n^2 = -84$ dBm as the transmit power and noise variance,

respectively. We consider $\{8, 8\}$ and $\{4, 4\}$ UPA at the AP and UT, respectively. After constructing the channel response for each UT location, we calculate the RSS for each beam pair using (11). According to the type of encoding scheme, M , the labels are calculated and stored in \mathbb{D}^Ξ . We consider 80% of each dataset for the training scheme and the remaining 20% as the test samples. To reduce the dependencies between the training samples and the performance of the data-driven methods, we use the 5-fold technique. In this technique, the dataset is divided into 5 parts. Then, the training and testing experiments are repeated 5 times, each of them with a different part of the dataset selected as test data, with the end result obtained by averaging over all 5 experiments.

There are 70,000 UT positions in the user grid of the LR, which are stored in \mathbb{D}^Ξ . To evaluate the performance of the data-driven methods with limited training datasets, we make two sub-datasets $\mathbb{D}_{10\%}^\Xi$ and $\mathbb{D}_{1\%}^\Xi$, respectively, consisting of 7,000 and 700 samples randomly selected from \mathbb{D}^Ξ . In the performance evaluations, we consider 5 beam alignment approaches:

- 1) **DNN-ST**: The DNN method with single-task structure described in Section IV-B1,
- 2) **DNN-MT**: The DNN method with multi-task structure explained in Section IV-B2,
- 3) **DNN-EMT**: The DNN method with extended multi-task structure described in Section IV-B3,
- 4) **GIFP**: The generalized inverse fingerprinting method explained in Section IV-C1,
- 5) **HBS**: The hierarchical beam search method with the deactivation technique represented in Section IV-C2.

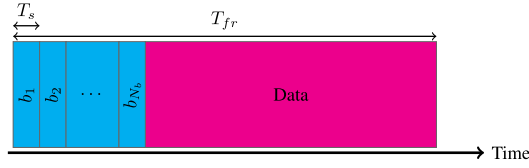


Fig. 8. Frame structure of the BA procedures. In the beam selection method, the transceivers sense the environment with N_b beam pairs, and then use the rest of the frame to transmit the actual data.

All the NNs structures have $N_h = 5$ hidden layers with $n_h = 128$ neurons at each hidden layer. Also, to prevent overfitting in the NNs, 10% dropout for all the hidden layers is used. We use Adam optimizer [30] in the training phase with 50 epochs, while the minibatch size is progressively increased from 32 to 8192 samples [24]. To reduce the effects of initial weights of NNs on the BA performance, we averaged over 3 results with 3 random weight initializations for each experiment. To reproduce the numerical results, the code and datasets are released at <https://github.com/SajadRezaie/3DOriLocBeamSelection>.

For a channel with a given coherence time, the channel resources used for the BA phase should be deducted from the achievable data rate of the system [31]. Fig. 8 shows the frame structure of the initial BA procedures. T_{fr} , T_s , and N_b are the frame duration, beam scanning time of one beam pair, and the number of sensed beam pairs in the BA phase, respectively. We assume the frame duration is less than the coherence time of the channel, and the channel response is not varying during the whole frame. In contrast to the beam selection methods, in the HBS approach the number of sensed beam pairs is fixed and entirely determined by the antenna array configurations as $N_b = 2\log_2(N_{UT}) + 2\log_2(N_{AP})$. The transceivers use the selected beam pair in the initial BA phase to communicate the actual data.

In the beam selection methods, the transceivers sense the environment with all the beam pairs in the candidate list \mathcal{S} ($|\mathcal{S}| = N_b$) and select the one that provides the highest RSS, i.e.

$$i^*, j^* = \arg \max_{i,j} (R_{i,j}). \quad (24)$$

As a performance measure, we define the misalignment probability by reformulating (22) as:

$$P_m(\mathcal{S}_k) = \mathbb{P} \left[R_{i^*,j^*} < \max_{(p,q) \in \mathcal{B}} R_{p,q} \right] \quad (25)$$

where \mathcal{B} ($|\mathcal{B}| = N_{AP}N_{UT}$) includes all the possible beam pair combinations. For the HBS approach, i^*, j^* are found by sensing the environment with different beam width, and then we use (25) to calculate the misalignment probability. In addition, the ESE as another performance measure may be written as

$$\text{SE}_{\text{eff}} = \frac{T_{fr} - N_b T_s}{T_{fr}} \log_2(1 + \text{SNR}_{i^*,j^*}), \quad N_b T_s \leq T_{fr} \quad (26)$$

where $\text{SNR}_{i,j}$ is the SNR of (i,j) th beam pair, i.e.

$$\text{SNR}_{i,j} = \frac{\left\| \sqrt{P_{AP}} \mathbf{v}_j^H \mathbf{H} \mathbf{u}_i \right\|^2}{\sigma_n^2}. \quad (27)$$

We use $T_{fr} = 20\text{ms}$ and $T_s = 0.1\text{ms}$ in the numerical evaluations [31]. In addition to the described baselines in Section IV-C, we use perfect alignment as a genie-aided method that always selects the beam pair with highest SNR out of all the possible beam combinations. The perfect alignment provides an upper bound for ESE.

B. Numerical Evaluation

In Fig. 9, the performance of the different investigated beam selection methods is depicted. The plots show the misalignment probability and spectral efficiency of the methods as a function of the number of beam pairs scanned. For data-driven methods, this corresponds to the size N_b of the candidate beam list \mathcal{S} . For HBS, on the other hand, the number of scanned beams is fixed and determined by the size of the hierarchical codebooks defined at AP and UT; in the selected configurations, this corresponds to $2\log_2(16) + 2\log_2(64) = 20$ beam pairs. As it can be seen from Fig. 9(a), the data-driven beam selection methods have much lower misalignment probability and up to 60% higher ESE than the HBS method. In addition, in the data-driven beam selection methods, the candidate list may include a different number of beam pairs for scanning, N_b . For example, the DNN-ST structure with a candidate list size of $N_b = 3$ can reach more than 50% higher ESE and 85% less latency than the HBS method, which scans 20 beam pairs. Also, all the deep-learning based methods perform significantly better than GIFF methods with different spatial and angular bin sizes. For example, the GIFF with $\text{SBS} = 1\text{m}$ and $\text{ABS} = 22.5^\circ$ has $4 \times 7 \times 1 \times 16 \times 4 \times 4 = 7168$ bins, which end up with around 8 samples on average at each bin. The excellent performance of DNN-ST is a consequence of training it with a large dataset of sufficient size. It clearly outperforms the DNN-MT and DNN-EMT methods, which suffer from the decoupling of the beam selection process at AP and UT for both cases, and between horizontal and vertical beam directions in the latter case. On the other hand, the DNN-MT and DNN-EMT networks have much less computational cost than the DNN-ST structure in the training and test. The computational complexity is proportional to the number of trainable parameters, and there are significantly fewer trainable weights in the DNN-MT and DNN-EMT.

As the dataset size is decreased, the performance advantages of DNN-ST progressively fade away. This is illustrated in Fig. 9(b) where, with a dataset size of 7000 samples, the DNN-MT method with fewer trainable parameters has the best performance. In spite of the 90% reduction of the training dataset size, the deep learning-based methods work significantly better than the HBS method. The performance of GIFF is further deteriorated with the decrease in dataset-size, and the DNN-MT method reaches up to 30% higher spectral efficiency than the GIFF methods with different spatial and angular bin sizes. Also, Fig. 9(c) shows the performance of different BA approaches with the minimal dataset, $\mathbb{D}_{1\%}^\Xi$, including only 700 samples. As it can be seen, the DNN-EMT method has the best performance in terms of misalignment probability and ESE, regardless of the size of the candidate list. Following our expectation, by limiting the dataset size significantly, the performance gap between data-driven BA methods and the

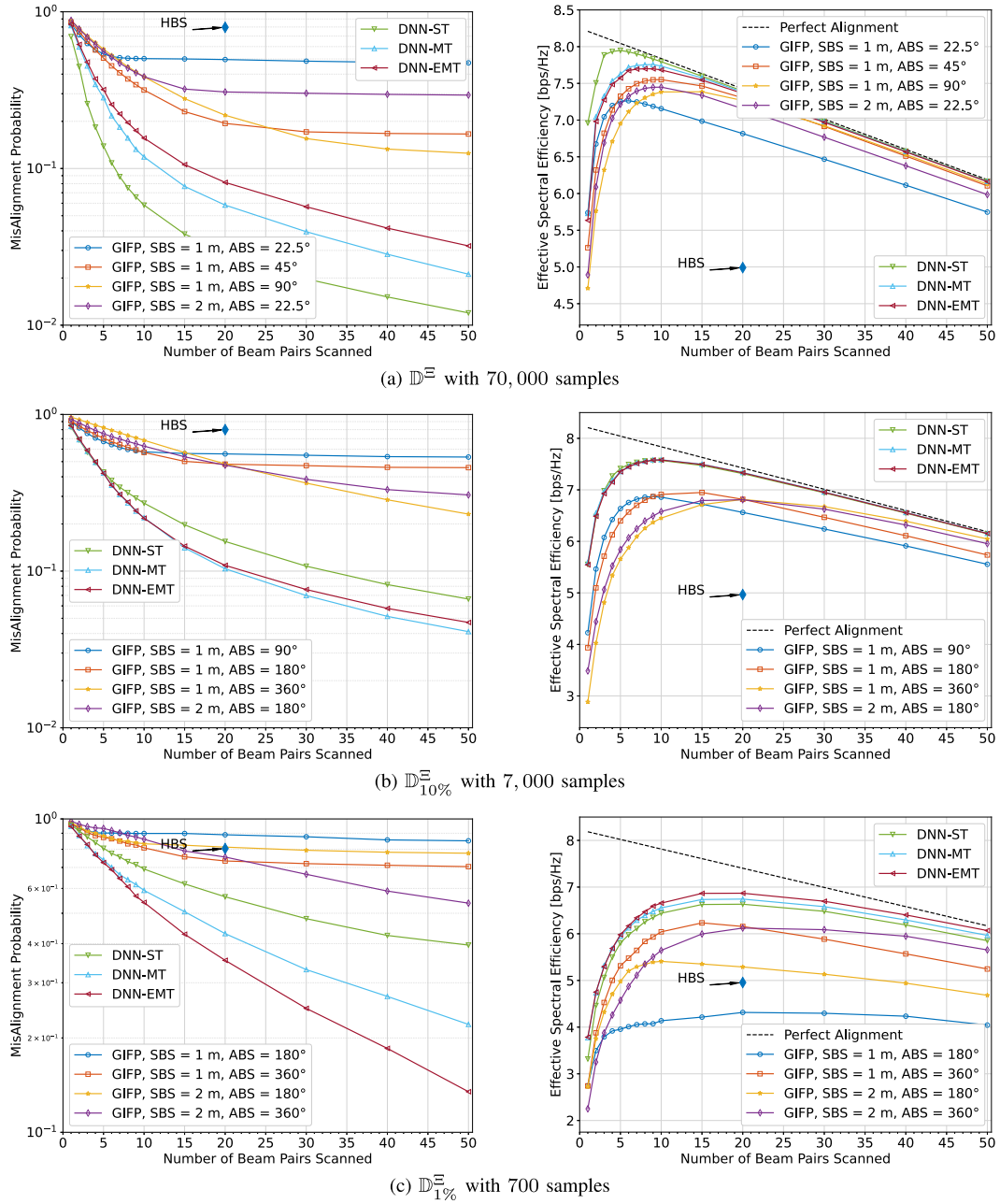


Fig. 9. Misalignment probability and effective spectral efficiency for different dataset sizes.

HBS method is reduced. These results illustrate the trade off between accuracy and dataset size in DNN based methods: with large datasets, the DNN-ST model with a large number of trainable parameters leads to more accurate results than DNN-MT and DNN-EMT; when training data is scarce, however, the DNN-MT and DNN-EMT achieve better performance due to their much smaller number of trainable parameters.

Fig. 10 shows the misalignment probability of the deep learning-based beam selection methods using the labels with M -hot encoding scheme. We use both \mathbb{D}^Ξ and $\mathbb{D}_{1\%}^\Xi$ datasets to see the effects of the multi-labeling technique on the performance of the methods with different structures. As there are many trainable parameters in the DNN-ST structure, multi-labeling works like data augmentation to fill the performance

gap because of insufficient training samples. However, the DNN-MT and DNN-EMT structures have significantly fewer trainable parameters. Thus, training these methods using multi-labeling with the large dataset \mathbb{D}^Ξ does not show any advantage over using a single label. When the training examples are limited in $\mathbb{D}_{1\%}^\Xi$, however, the multi-labeling technique is helpful to have a lower misalignment probability.

The robustness of the data-driven based BA methods against different propagation properties in the training and test samples is evaluated in an experiment. We consider to have two datasets with 70,000 samples, \mathbb{D}^Ξ and \mathbb{D}^Υ , with LOS blockage probability of 0.5 and 0.2, respectively. For the rest of simulations, we use $SBS = 1\text{m}$ and $ABS = 45^\circ$ parameters for the GIFF method. Fig. 11 shows the performance of the GIFF and deep

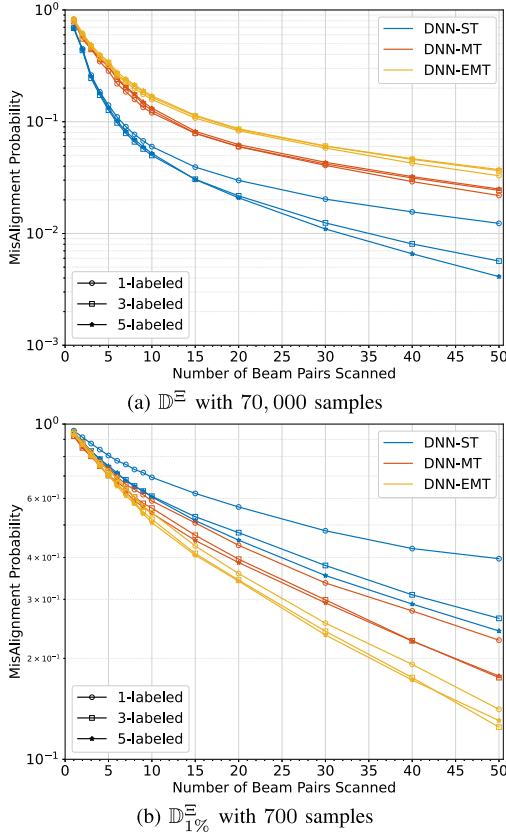


Fig. 10. Effects of the multi-labeling scheme on the performance of the proposed DNNs with large and small training datasets.

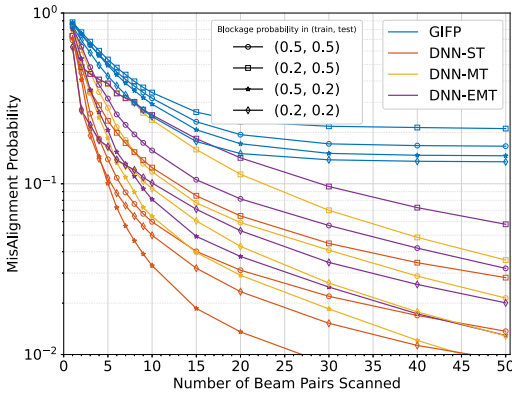


Fig. 11. The misalignment probability of different data-driven BA methods with various LOS blockage probability in the training and test datasets.

learning-based methods when there is a match/mismatch of LOS blockage rate between the training and test samples. As it can be seen in Fig. 11, in the case where the LOS blockage rate is lower in the training set than in the test set, the network learns better the LOS direction but has fewer opportunities to learn NLOS directions. Thus, in comparison to the matched case when there is high LOS blockage probability in the training and test sets, the performance is degraded. Contrary, by training a NN at a higher LOS blockage probability compared to the test samples, the network learns better the NLOS paths at the expense of a worse learning of the LOS one. Due to this, the network trained with matched LOS

blockage probability performs better when few beam pairs are scanned. On the other hand, when more beam pairs are scanned there is a good chance to scan candidates for both LOS and NLOS paths. Thus, the network trained in a higher LOS blockage rate has a better performance compared to the matched case with large N_b . To summarize, the NNs are robust to mismatch in the LOS blockage probability, and we see more robustness in the networks if trained in a high LOS blockage probability.

Although context information can reduce the overhead of initial BA, in practical scenarios the position and orientation of the UT are subject to estimation errors. To account for this in our numerical assessment, we use a simple model of such inaccuracies. Consider \mathbf{p}_{UT} and ψ_{UT} as the true position and orientation vector of the UT. By adding random perturbation to these vectors, we can model the measured position and orientation, $\hat{\mathbf{p}}_{UT}$ and $\hat{\psi}_{UT}$, as

$$\hat{\mathbf{p}}_{UT} = \mathbf{p}_{UT} + \epsilon_p, \quad (28)$$

$$\hat{\psi}_{UT} = \psi_{UT} + \epsilon_\psi, \quad (29)$$

where ϵ_p and ϵ_ψ , respectively, denote the random perturbation of the position and orientation sensors on the UT. In line with the known statistics of the measuring error, the entries of the random perturbation vectors ϵ_p and ϵ_ψ are generated from independent Gaussian distributions with zero mean and variance σ_p^2 and σ_ψ^2 , respectively.

In Fig. 12, the performance of the GIFF, DNN-ST, and DNN-MT methods including inaccuracies in the position and/or orientation information of the UT device are shown. The dataset with 70,000 samples are used, and the estimation error is included in both the train and test sets. When applying perturbation only on the position information, we see a slight degradation with $\sigma_p = 0.1m$. This level of accuracy can be obtained with, e.g., ultra wide-band (UWB) positioning [32]. As shown in Fig. 12(a), by increasing σ_p to $0.5m$, both the GIFF and deep learning-based methods have performance loss, respectively, around 10% and 5% in the maximum of achievable ESE. The sensitivity of data-driven based BA methods to position inaccuracy depends on the dimensions of the environment. Thus, we expect to see less performance degradation due to inaccurate positions in big rooms or outdoor scenarios. Fig. 12(b) shows the performance of different BA methods by feeding inaccurate orientation information. All the data-driven approaches offer good robustness to the different amounts of orientation perturbations. When we apply the perturbation model on both position and orientation information, as we expected, more degradation in their performance can be seen in Fig. 12(c). At all the considered σ_p and σ_o cases, the data-driven based methods show better performance than the HBS method. Thus, the data-driven methods still leverage inaccurate context information to reduce the latency of the BA procedure and reach higher spectral efficiency. By looking at the ESE curves, it can be seen that the DNN-ST method has the least sensitivity to inaccurate location and orientation information. This shows that, during the training process, the DNN is able to learn how to deal with different degrees of CI uncertainty, leading to a robust performance against these inaccuracies.

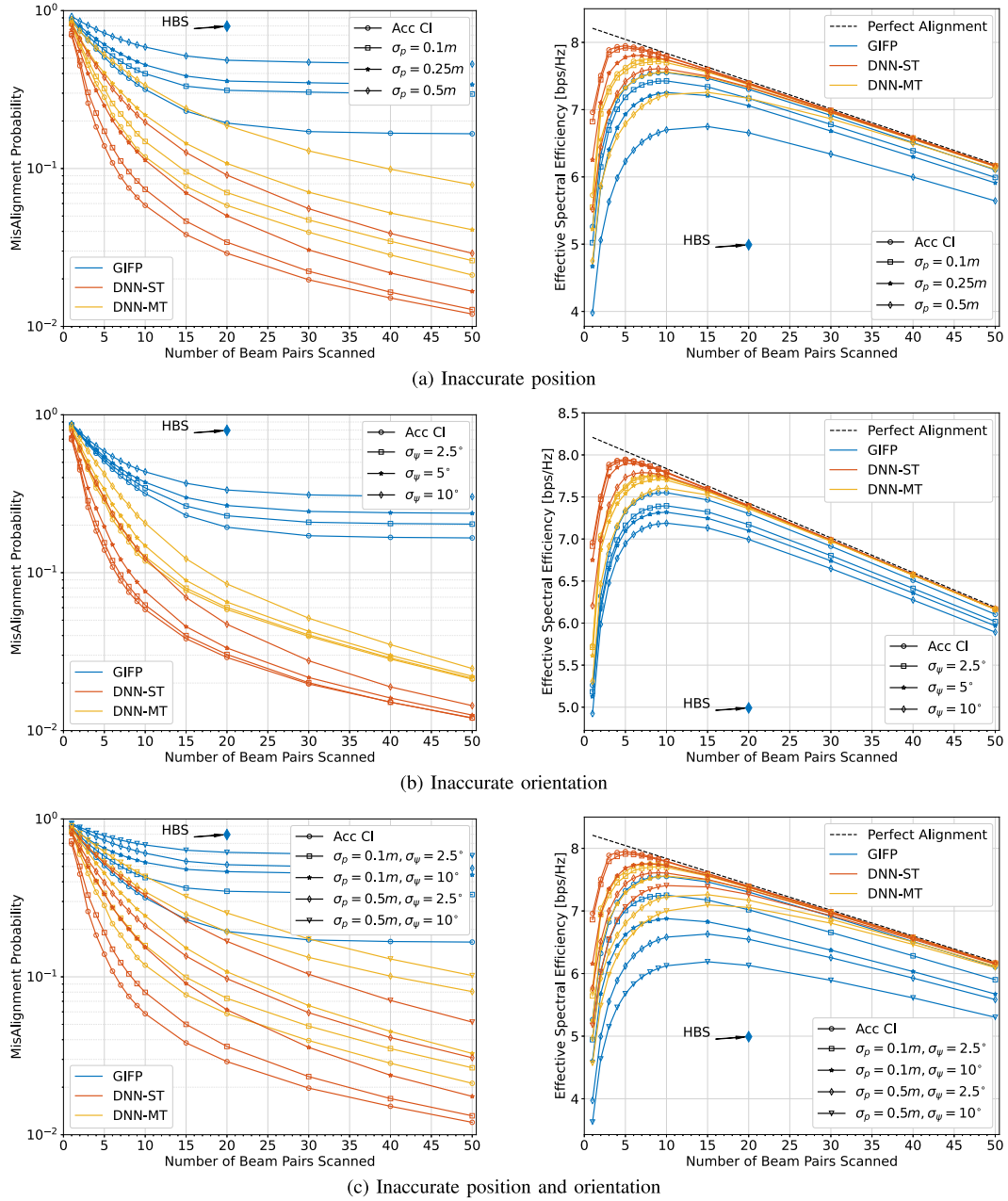


Fig. 12. Misalignment probability and effective spectral efficiency with inaccurate position and orientation information.

C. Computational Complexity Analysis

The DNN-ST, DNN-MT, and DNN-EMT methods respectively need $6n_h + (N_h - 1)n_h^2 + n_h(N_{AP}N_{UT}) = 197376$, $6n_h + (N_h - 1)n_h^2 + n_h(N_{AP} + N_{UT}) + N_{AP}N_{UT} = 77568$, and $6n_h + (N_h - 1)n_h^2 + n_h(N_{AP_H} + N_{AP_V} + N_{UT_H} + N_{UT_V}) + N_{AP_H}N_{AP_V} + N_{UT_H}N_{UT_V} + N_{AP}N_{UT} = 70480$ real multiply-accumulate operations in the evaluation phase for each test sample [33]. Using the MT and EMT structures reduces the computational complexity by more than 50% with respect to the ST structure. However, the GIFF has no significant computational load, as the best beam pairs for the corresponding cell of the UT location and orientation just need to be loaded from memory. Also, the computational complexity of the HBS method is negligible as the beamforming

vectors at each level of the hierarchical search are predefined.

VI. CONCLUSION

In this work, we have shown that position and orientation of UT as a kind of context information can play an important role in the initial beam alignment procedure of mmWave links in pedestrians applications. We have presented a deep-learning based approach for beam recommendation based solely on the position and orientation of the UT. Three neural network structures are designed for the task: a standard feed forward network coined DNN-ST, and two alternative and simplified designs. The simplified structures, coined DNN-MT and DNN-EMT, are inspired by multitask networks and separate

the beamforming tasks at AP and UT, as well as the horizontal and vertical beam directions in the latter design. This results in a significant reduction of trainable parameters and size of the DNN model.

Our extensive performance assessment based on ray-traced channel responses reveals that the DNN-ST design shows excellent performance when enough training samples are in the dataset. However, the DNN-MT and DNN-EMT designs have less computational cost in the training and evaluation phase, and these networks perform better with small training datasets. The proposed beam selection methods outperform the GIFFP method as a CI-based lookup table approach and the DEACT method as a CI-agnostic hierarchical beam search solution in terms of latency and spectral efficiency. Our proposed methods show low sensitivity to changes in the propagation properties in the evaluation compared to the data collection time as well as displaying high robustness against measurement inaccuracies in the position and orientation of the UT.

APPENDIX

The lemma shows the relation between the loss function of the proposed multi-task structure calculated before and after the outer product. The lemma highlights that even though the outer product operation is not typical in NNs structures, the training of the multi-task structured network is similar to the standard training procedure in multi-task applications.

Lemma 1: For the multi-task structured network, the loss function of the form $\mathcal{L}(\mathbf{o}, \mathbf{L}^{(M)})$ is equivalent to $\mathcal{L}(\mathbf{o}^t, \mathbf{l}^t) + \mathcal{L}(\mathbf{o}^r, \mathbf{l}^r)$, where $\mathbf{l}_i^t = \sum_{j=1}^{N_{UT}} L_{i,j}^{(M)}$ and $\mathbf{l}_j^r = \sum_{i=1}^{N_{AP}} L_{i,j}^{(M)}$.

Proof:

$$\mathcal{L}(\mathbf{o}, \mathbf{L}^{(M)}) = - \sum_{i=1}^{N_{AP}} \sum_{j=1}^{N_{UT}} L_{i,j}^{(M)} \log(o_{i,j}) \quad (30a)$$

$$= - \sum_{i=1}^{N_{AP}} \sum_{j=1}^{N_{UT}} L_{i,j}^{(M)} \left(\log(o_i^t) + \log(o_j^r) \right) \quad (30b)$$

$$= - \sum_{i=1}^{N_{AP}} l_i^t \log(o_i^t) - \sum_{j=1}^{N_{UT}} l_j^r \log(o_j^r) \quad (30c)$$

$$= \mathcal{L}(\mathbf{o}^t, \mathbf{l}^t) + \mathcal{L}(\mathbf{o}^r, \mathbf{l}^r) \quad (30d)$$

In the single-label case $M = 1$, training the network with label $L_{i,j} = 1$ after the outer product is equal to training with labels $\mathbf{l}_i^t = 1$ and $\mathbf{l}_j^r = 1$ at the two output branches before the outer product. In the case $M > 1$, however, AP and UT beams that appear in multiple non-zero entries in the beam-pair label matrix $\mathbf{L}^{(M)}$ are given larger weight in the equivalent AP and UT beam label vectors \mathbf{l}^t and \mathbf{l}^r .

REFERENCES

- [1] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.
- [2] C. Liu, M. Li, S. V. Hanly, P. Whiting, and I. B. Collings, "Millimeter-wave small cells: Base station discovery, beam alignment, and system design challenges," *IEEE Wireless Commun.*, vol. 25, no. 4, pp. 40–46, Aug. 2018.
- [3] M. Giordani, M. Mezzavilla, and M. Zorzi, "Initial access in 5G mmWave cellular networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 40–47, Nov. 2016.
- [4] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.
- [5] C. N. Manchón, E. De Carvalho, and J. B. Andersen, "Ping-pong beam training with hybrid digital-analog antenna arrays," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–7.
- [6] F. Sohrabi, Z. Chen, and W. Yu, "Deep active learning approach to adaptive beamforming for mmWave initial alignment," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2347–2360, Aug. 2021.
- [7] 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.
- [8] M. Polese, F. Restuccia, and T. Melodia, "DeepBeam: Deep waveform learning for coordination-free beam management in mmWave networks," in *Proc. 21st Int. Symp. Theory, Algorithmic Found., Protocol Des. Mobile Netw. Mobile Comput.*, Jul. 2021, pp. 61–70.
- [9] M. Kok, J. D. Hol, and T. B. Schön, "Using inertial sensors for position and orientation estimation," *Found. Trends Signal Process.*, vol. 11, nos. 1–2, pp. 1–153, 2017.
- [10] V. Va, J. Choi, T. Shimizu, G. Bansal, and R. W. Heath, Jr., "Inverse multipath fingerprinting for millimeter wave V2I beam alignment," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 4042–4058, May 2018.
- [11] M. S. Mollel *et al.*, "A survey of machine learning applications to handover management in 5G and beyond," *IEEE Access*, vol. 9, pp. 45770–45802, 2021.
- [12] Y. Zhang, M. Alrabeiah, and A. Alkhateeb, "Reinforcement learning of beam codebooks in millimeter wave and terahertz MIMO systems," 2021, *arXiv:2102.11392*.
- [13] W. Xu, F. Gao, S. Jin, and A. Alkhateeb, "3D scene-based beam selection for mmWave communications," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 1850–1854, Nov. 2020.
- [14] L.-H. Shen, T.-W. Chang, K.-T. Feng, and P.-T. Huang, "Design and implementation for deep learning based adjustable beamforming training for millimeter wave communication systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2413–2427, Mar. 2021.
- [15] Z. Zhu, K. Lin, A. K. Jain, and J. Zhou, "Transfer learning in deep reinforcement learning: A survey," 2020, *arXiv:2009.07888*.
- [16] M. Alrabeiah and A. Alkhateeb, "Deep learning for mmWave beam and blockage prediction using sub-6 GHz channels," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5504–5518, Sep. 2020.
- [17] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7704–7719, Nov. 2020.
- [18] M. Dias, A. Klautau, N. Gonzalez-Prelcic, and R. W. Heath, Jr., "Position and LIDAR-aided mmWave beam selection using deep learning," in *Proc. IEEE 20th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [19] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-aided dynamic blockage prediction for 6G wireless communication networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.
- [20] 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. (VTC-Spring)*, May 2020, pp. 1–5.
- [21] S. Rezaie, A. Amiri, E. de Carvalho, and C. N. Manchón, "Deep transfer learning for location-aware millimeter wave beam selection," *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 2963–2967, Sep. 2021.
- [22] Y. Heng *et al.*, "Six key challenges for beam management in 5G and 6G systems," *IEEE Commun. Mag.*, vol. 59, no. 7, pp. 74–79, Jul. 2021.
- [23] A. Ali, J. Mo, B. L. Ng, V. Va, and J. C. Zhang, "Orientation-assisted beam management for beyond 5G systems," *IEEE Access*, vol. 9, pp. 51832–51846, 2021.
- [24] S. Rezaie, C. N. Manchón, and E. De Carvalho, "Location- and orientation-aided millimeter wave beam selection using deep learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.
- [25] Altair Feko WinProp. Accessed: Aug. 10, 2021. [Online]. Available: <https://www.altair.com/feko>
- [26] C. De Lima *et al.*, "Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges," *IEEE Access*, vol. 9, pp. 26902–26925, 2021.
- [27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, Nov. 2016. [Online]. Available: <https://www.deeplearningbook.org/>

- [28] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [29] A. Maltsev, *Channel Models for 60 GHz WLAN Systems*, IEEE Standard 802.11.09/0334r8, May 2010.
- [30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–15.
- [31] M. Hussain and N. Michelusi, "Second-best beam-alignment via Bayesian multi-armed bandits," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [32] C. Zhang, M. J. Kuhn, B. C. Merkl, A. E. Fathy, and M. R. Mahfouz, "Real-time noncoherent UWB positioning radar with millimeter range accuracy: Theory and experiment," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 1, pp. 9–20, Jan. 2010.
- [33] P. J. Freire *et al.*, "Performance versus complexity study of neural network equalizers in coherent optical systems," *J. Lightw. Technol.*, vol. 39, no. 19, pp. 6085–6096, Oct. 1, 2021.



Sajad Rezaie (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 2013 and 2016, respectively. Since 2019, he has been a Ph.D. fellow with the Wireless Communication Section, Aalborg University, Aalborg, Denmark. His research interests include MIMO communications, mmWave communications, signal processing, and machine learning applications in wireless communications. Some sample projects of his work are available at <https://github.com/SajadRezaie/>



Elisabeth de Carvalho (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Télécom ParisTech, France. She was a Post-Doctoral Fellow with Stanford University, Stanford, CA, USA; and then with industry in the field of DSL and wireless LAN. Since 2005, she has been an Associate Professor with Aalborg University, where she has led several research projects in wireless communications. She has coauthored the text book *Practical Guide to MIMO Radio Channel*. Her main expertise are in signal processing for MIMO communications, with recent focus on massive MIMO, including channel measurements, channel modeling, beamforming, and protocol aspects.



Carles Navarro Manchón received the degree in telecommunication engineering from Miguel Hernández University of Elche, Spain, in 2006, and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2011. Since 2006, he has been with Aalborg University, where he is currently an Associate Professor with the Department of Electronic Systems, Wireless Communication Networks Section. His research interests lie in the area of statistical signal processing for wireless communications, including joint channel estimation and decoding algorithms, estimation and reconstruction of sparse signals, signal processing for multiple antenna systems, and machine learning methods for wireless networks.