

# Entropy-based Probing Beam Selection and Beam Prediction via Deep Learning

Fan Meng, Cheng Zhang, *Member, IEEE*,  
Yongming Huang, *Senior Member, IEEE*, Zhaohua Lu

**Abstract**—Hierarchical beam search in mmWave communications incurs substantial training overhead, necessitating deep learning-enabled beam predictions to effectively leverage channel priors and mitigate this overhead. In this study, we introduce a comprehensive probabilistic model of power distribution in beamspace, and formulate the joint optimization problem of probing beam selection and probabilistic beam prediction as an entropy minimization problem. Then, we propose a greedy scheme to iteratively and alternately solve this problem, where a transformer-based beam predictor is trained to estimate the conditional power distribution based on the probing beams and user location within each iteration, and the trained predictor selects an unmeasured beam that minimizes the entropy of remaining beams. To further reduce the number of interactions and the computational complexity of the iterative scheme, we propose a two-stage probing beam selection scheme. Firstly, probing beams are selected from a location-specific codebook designed by an entropy-based criterion, and predictions are made with corresponding feedback. Secondly, the optimal beam is identified using additional probing beams with the highest predicted power values. Simulation results demonstrate the superiority of the proposed schemes compared to hierarchical beam search and beam prediction with uniform probing beams.

**Index Terms**—mmWave communication, beam prediction, probing beam selection, deep learning, entropy minimization.

## I. INTRODUCTION

In B5G/6G wireless communications, millimeter wave (mmWave) communication is emerging as an appealing solution to provide abundant available spectrum to meet the critical demands of exploding data traffic [1,2]. However, the high path loss of mmWave signals poses a significant challenge to data transmission, resulting in limited coverage area. The small carrier wavelength enables packing a large number of antenna elements into small form factors. Leveraging the large antenna arrays employed at the transmitter and receiver, mmWave systems perform directional beamforming [3,4] to overcome the high free-space path loss of mmWave signals and also reduce spatial interference. Nevertheless, the massive

antennas bring significant challenges for channel estimation, beam alignment/tracking (BA/T), especially in highly mobile and/or complex environments such as high-speed railway, unmanned aerial vehicle (UAV), urban macro (UMa).

Pencil-like beamforming for mmWave scenarios such as UMa is challenging. Environmental factors, i.e., wind flow and precipitation, moving vehicles and pedestrians, can cause drastic variations in received power fluctuation. As the transceiver units of mmWave base stations (BS) are mounted to facilities such as poles, pillars or street lamps, vibration and movement can cause unacceptable outage probability if BA/T is not frequently performed [3]. Meanwhile, link stability is improved with frequent BA/T, but at the cost of high beam training overheads. Therefore, low-overhead mmWave BA/T while maintaining link stability is essential.

Traditional model-driven BA/T including exhaustive [5] and hierarchical beam search [6]–[8], fail to adequately exploit the (partial) channel state information (CSI) prior, resulting in large overheads. Furthermore, the CSI prior is difficult to analytically characterize especially in complex scenarios, thus the model-driven methods inherently are inappropriate for the implicit CSI prior. Deep learning (DL) [9] has been identified as an enabling technology for future wireless mobile networks [10,11], and it has received extensive attention for precoding [12]–[14], positioning [15,16], CSI compression and reconstruction [17,18], beam management [19]–[23], network optimization [24]. Data-driven or -aided BA/T is a promising technique that automatically learns and exploits the underlying correlations of CSI across different times, frequencies, spaces or other out-of-band information [25,26], to reduce CSI acquisition overhead and improve system spectral efficiency and robustness.

### A. Related Work

Conventional BA/T is measurement-based, the transmit/receive beam is an element within the measured beam set, e.g., the beam with maximal received power is selected in a single-user link-level system. Predictive BA/T, on the other hand, derives the maximum received power or RSRP in the entire beamspace with few or no measurements. The main difference is that the selected beam is not necessarily measured in the beam prediction, but can be inferred from measured beams at different times, frequencies and spaces. In some special scenarios, even instantaneous measurement is not required and the beam prediction is performed only with out-of-band information such as location, motion and orientation of the mobile user (MU) [21,22,27].

This work was supported in part by the National Key R&D Program of China under Grant 2020YFB1806600 and the National Natural Science Foundation of China under Grant No. 62225107, 62001103 and 62201394, and the Fundamental Research Funds for the Central Universities under Grant 2242022k60002. (Corresponding authors: Y. Huang)

F. Meng, C. Zhang and Y. Huang are with the Purple Mountain Laboratories, Nanjing 211111, China (e-mail: mengfan@pmlabs.com.cn; zhangcheng\_seu@seu.edu.cn; huangym@seu.edu.cn). C. Zhang and Y. Huang are also with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, China. Z. Lu is with the ZTE Corporation and State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518057, China (e-mail: lu.zhaohua@zte.com.cn).

Considering the time correlation, some studies establish the dynamics of the mmWave channel to realize beam prediction [28]–[30]. The studies in [28] propose a two-level probabilistic beam prediction. On a long-time scale, a variational auto-encoder uses noisy beam-training feedback to learn a probabilistic model of beam dynamics and enable predictive beam-tracking; on a short-time scale, an adaptive beam-training procedure is formulated as a partially observable Markov decision process and solved by reinforcement learning. In [29], the neural ordinary differential equation is exploited to predict the arbitrary-instant optimal beam between the current and next beam training instants.

Due to the same physical environment, channels in low and high carrier frequencies are correlated. To reap this benefit, many studies learn to predict the mmWave channel with full sub-6 GHz channel and limited low-overhead measurement of the mmWave channel [31]–[33]. In [32], a dual-input neural network (NN) architecture is designed to merge the sub-6 GHz channel and the mmWave channel of a few active antennas. In addition, an antenna selection method is introduced to better match the mmWave channel instead of a uniform pilot design. Moreover, the work in [31] considers blockage prediction which is important to proactively improve the link reliability.

Location-awareness is becoming a fundamental feature to support various mobile applications, and the radio access network is a promising infrastructure to efficiently and intelligently utilize wireless resources by integrating sensing and communication [34]–[36]. Therefore, recent studies estimate or directly use the MU location or geometric environment between transceivers as side information for beam prediction [27,37]–[40]. The Gaussian process (GP) is a probabilistic machine learning model that performs inference with uncertainties, and it has been well-applied in small-sample low-dimensional beam prediction [27,39]–[41]. In mmWave fixed wireless access, researchers develop an explicit mapping between transmit/receive beams and MU physical coordinates via a GP [39]. Similarly in the mmWave UAV network, GP-enabled beam management scheme utilizing angular domain information is proposed [27,40] to rapidly establish and reliably maintain the communication links. Small-sample learning is promising for real-time online adaption, but it inherently lacks the ability to exact complex priors from plenty of data. Moreover, GP training involves matrix inversion, making it difficult for large-sample high-dimensional problems where DL tools are better options [37,38,42,43]. As a large amount of training data is presented, offline learning and online inference with DL is well-investigated. In [37], a mapping from the user location to the beam pairs (fingerprints) is realized by a deep neural network (NN), with labeled data collected in different locations. Meanwhile, single spatial information is insufficient to accurately infer the reference signal receive power (RSRP) of narrow beams, and low-overhead probing is necessary [44].

### B. Motivation and Contribution

Considering a mmWave system with massive antenna arrays, we investigate the location-aware probing beam selection and the probabilistic beam prediction problems. In general,

we use the multivariate Gaussian model to approximate the distribution of RSRP in beamspace, and a beam predictor is to estimate the RSRP distribution with MU location. Given the learned beam predictor, the probing beam selection is to minimize the conditional entropy of the unmeasured beams, by selecting the training beam combination which is a subset of the discrete Fourier transform (DFT) codebook. The involved technical difficulties are as follows.

- **Prediction Model.** The beamspace of RSRP with massive antennas is high-dimensional and the underlying channel prior is implicit and complex. Thus, model-driven or shallow data-driven schemes such as GP cannot work well. On the other hand, DL has the merit of efficiently extracting the channel prior, but the existing literature rarely considers probabilistic inference.
- **Uncertainty Evaluation.** Global uncertainty can be evaluated by entropy, but this is incompatible with beamforming which usually concerns only with the optimal beam with maximum RSRP. Meanwhile, the optimal beam can be in any direction, and local uncertainty cannot address this issue. Apparently, there is a trade-off between global and local uncertainty in the probing beam selection.
- **Computational Complexity.** The probing beam selection is a combinatorial optimization problem requiring exhaustive search, which has extremely high computational complexity in high-dimensional beamspace and is infeasible for practical use. Furthermore, each search consumes one beam prediction involving matrix inversion, which exacerbates this issue.

To address these difficulties, first to realize probabilistic inference, we propose a DL-based beam predictor which is trained by the maximum likelihood (ML) criterion, to estimate the conditional RSRP distribution with the MU location and probing beams. Second, to achieve a trade-off between global and local uncertainty, we mask the entropy with a weight matrix, which is designed by the predicted RSRP values. Third, to reduce the computational complexity, we design a greedy solution to iteratively solve the probing beam selection problem. Furthermore, to reduce the number of interactions and the computational complexity of the iterative solution, we propose a two-stage probing beam selection scheme. The two-level scheme is feasible for practical implementation, which has only two interactions and one beam prediction operation. Our technical contributions are summarized as follows.

- We establish a generic probabilistic model of the RSRP in beamspace, and formulate the joint probing beam selection and probabilistic beam prediction as an entropy minimization problem. To obtain a tradeoff between the overall entropy and the local entropy w.r.t. large-power beams, we extend the problem as a weighted entropy minimization with a tunable diagonal matrix.
- We propose a greedy scheme, i.e., Iter-BP&PBS, to iteratively solve the weighted entropy minimization problem, to reduce the computational complexity. During an iteration, a beam predictor is trained to estimate the conditional RSRP distribution with the probing beams and the MU location. The learned predictor then selects

an unmeasured beam to minimize the weighted entropy of the remaining unmeasured beams.

- We propose a two-stage probing beam selection scheme, i.e., 2S-BP&PBS, to further reduce the number of interactions and the computational complexity of Iter-BP&PBS. Firstly, the BS selects probing beams in a location-specific codebook by the MU location, and roughly locates the optimal beam by the prediction with the RSRP feedbacks. Secondly, the BS probes several beams with top-predicted RSRP values to find the optimal beam.
- We design a scalable beam predictor composed of a mean network and a variance network, using the transformer as the backbone. We design an ML-based cost function to realize probabilistic inference, and simplify the variance network output as a diagonal covariance matrix, to further reduce computational complexity and achieve numerical stability.
- Simulation results for an urban scenario demonstrate the superior performance of the proposed schemes compared to the existing hierarchical beam search and beam prediction with uniform probing beams. In addition, the two-stage scheme has low computational, storage and interaction requirements, which is important for real-time deployment.

The rest of this paper is organized as follows. The system model and problem formulation are described in Section II. The iterative algorithm for probing beam selection and beam prediction is clarified in Section III, and the two-stage probing beam selection is introduced in Section IV. Furthermore, the design of the transformer-based beam predictor is given in Section V. The numerical results are shown in Section VI, and the conclusions are drawn in Section VII.

*Notations:* We use lowercase (uppercase) boldface  $\mathbf{A}(\mathbf{a})$  to denote the vector (matrix), and  $a$  is a scalar. Calligraphy letter  $\mathcal{A}$  represents the set. Superscripts  $(\cdot)^T$  represents the transpose.  $\det$ ,  $\text{diag}$ ,  $|\cdot|$ ,  $\|\cdot\|_2$ ,  $\otimes$  respectively denote determinant, diagonal, absolute,  $\ell_2$  norm, and Kronecker product operators.  $\mathbb{E}\{\cdot\}$ ,  $\mathbb{R}$  and  $\mathbb{C}$  respectively represent the expectation, the real and complex fields.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

### A. System Model

We consider a link-level mmWave massive multiple-input single-output (MISO) communication system composed of one BS and one MU. The BS is equipped with a massive planer antenna array with  $N$  antennas connected to one radio frequency (RF) chain, and the MU is equipped with one isotropic antenna. The scenarios can be easily extended to cellular or cell-free systems with multiple BSs and MUs.

1) *Channel Model:* Without loss of generality, the wireless mmWave propagation is characterized by multi-path propagation due to interactions (reflections, diffractions, penetrations, scattering) at stationary obstacles (hills, buildings, towers) and mobile objects (cars, pedestrians). According to the 3GPP channel modeling [45], the downlink mmWave channel  $\mathbf{h} \in$

$\mathbb{C}^{N \times 1}$  is modeled as a combination of line-of-sight (LoS) and non-LoS (NLoS) channels, i.e.,

$$\mathbf{h} = \mathbf{h}_{\text{LoS}} + \mathbf{h}_{\text{NLoS}}. \quad (1)$$

The LoS channel  $\mathbf{h}_{\text{LoS}}$  has only a dominant path, the NLoS  $\mathbf{h}_{\text{NLoS}}$  is consisting of  $N_{\text{cl}}$  dominant clusters, and each cluster is composed of  $N_{\text{ray}}$  rays. Thus, the narrow-band channel vectors in the antenna domain respectively are described as

$$\mathbf{h}_{\text{LoS}} = \alpha_{\text{LoS}} \psi(\phi_{\text{LoS}}, \theta_{\text{LoS}}), \quad (2)$$

$$\mathbf{h}_{\text{NLoS}} = \sum_{m=1}^{N_{\text{cl}}} \sum_{n=1}^{N_{\text{ray}}} \alpha_{m,n} \psi(\phi_{m,n}, \theta_{m,n}), \quad (3)$$

where  $\alpha$  is a complex channel gain,  $\phi$  and  $\theta$  respectively are the angles of departure in horizontal and vertical directions,  $\psi$  is the planer array response at the BS which is given as follows

$$\psi(\phi, \theta) = \mathbf{a}_{\text{xy}}(\phi, \theta) \otimes \mathbf{a}_z(\theta), \quad (4)$$

where

$$\mathbf{a}_{\text{xy}}(\phi, \theta) = \frac{1}{\sqrt{N_\phi}} [1, e^{j\pi \sin \phi \sin \theta}, \dots, e^{j\pi (N_\phi - 1) \sin \phi \sin \theta}]^T, \quad (5)$$

$$\mathbf{a}_z(\theta) = \frac{1}{\sqrt{N_\theta}} [1, e^{j\pi \cos \theta}, \dots, e^{j\pi (N_\theta - 1) \cos \theta}]^T, \quad (6)$$

where  $N_\phi$  and  $N_\theta$  respectively are the antenna numbers in the horizontal and vertical dimensions, and  $N = N_\phi N_\theta$ . The rays in cluster  $m$  are closely distributed around the center of this cluster in angles  $\{\phi_m, \theta_m\}$ . Particularly,  $\alpha_{\text{LoS}} = 0$  indicates the LoS path is blocked.

2) *Data Model:* The BS exhaustively sweeps the beams in the DFT codebook  $\mathbf{A} \in \mathbb{C}^{N \times N}$ , and the observed RSRP at the MU side is represented as

$$\mathbf{x} = |\mathbf{A}\mathbf{h} + \mathbf{n}_x|^2, \quad (7)$$

where  $\mathbf{n}_x \in \mathbb{C}^{N \times 1}$  is the measurement noise following  $\mathcal{N}(\mathbf{0}, \sigma_x^2 \mathbf{I}_N)$  where  $\sigma_x^2$  is the noise variance. The relative position of the MU w.r.t. the BS is

$$\mathbf{s} = \mathbf{s}_r - \mathbf{s}_t + \mathbf{n}_s, \quad (8)$$

where  $\mathbf{s}_t \in \mathbb{R}^{2 \times 1}$  and  $\mathbf{s}_r \in \mathbb{R}^{2 \times 1}$  respectively are 2D locations of the BS and the MU,  $\mathbf{n}_s$  denotes the positioning error following  $\mathcal{N}(\mathbf{0}, \sigma_s^2 \mathbf{I}_2)$ .

Measurement-based BA/T typically consumes large beam training overhead to align the optimal beam. Using the potential mapping from the side information to the RSRP, can significantly reduce the overhead. In this work, we propose to realize BA/T with a small number of probing beams and the UE location, thus the design of probing beams is crucial. Concretely, given the relative MU location  $\mathbf{s}$  and the maximal probing beam number  $L$ , the BS sends  $L$  probing beams in  $\mathbf{A}$  and receives the counterpart MU feedbacks, then predicts the RSRPs of other beams in  $\mathbf{A}$ .

### B. Problem Formulation

The RSRP  $\mathbf{x}$  follows an unknown distribution  $\Pr(\mathbf{x})$ . For the convenience of analysis and implementation, we assume that  $\mathbf{x}$

approximately follows the multivariate Gaussian distribution, i.e.,

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{\det(2\pi\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})^\top\right), \quad (9)$$

where  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$  respectively are the mean vector and covariance matrix of  $\mathbf{x}$ , and they are regarded as a function w.r.t. the side information  $\mathbf{s}$ .

Given  $\mathbf{s}$ , the beam prediction problem is to estimate the distribution of  $\mathbf{x}$ , which is modeled as a ML estimation, i.e.,

$$\max_{\boldsymbol{\Theta}_f, \boldsymbol{\Theta}_g} \mathbb{E}_{\mathbf{x}, \mathbf{s}} \left\{ \ln \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \right\}, \quad (10)$$

where

$$\boldsymbol{\mu} = f(\mathbf{s}; \boldsymbol{\Theta}_f), \quad (11)$$

$$\boldsymbol{\Sigma} = g(\mathbf{s}; \boldsymbol{\Theta}_g), \quad (12)$$

where  $f$  and  $g$  respectively are the functions of the mean network and the variance network,  $\boldsymbol{\Theta}_f$  and  $\boldsymbol{\Theta}_g$  are the corresponding learnable parameters. Evidently, Mean square error (MSE) minimization-based beam prediction is a special case of (10) with the simpler setting  $\boldsymbol{\Sigma} = \mathbf{I}_N$ .

We define a set of measured beam indices  $\mathcal{Q} \subseteq \{1, \dots, N\}$ , and a set of un-measured beam indices  $\mathcal{P} \subseteq \{1, \dots, N\}$ . Apparently, the measured beam number  $|\mathcal{Q}| \leq L$ ,  $\mathcal{P} \cup \mathcal{Q} = \{1, \dots, N\}$  and  $\mathcal{P} \cap \mathcal{Q} = \emptyset$ . Equivalently,  $\{x_i | i \in \mathcal{Q}\}$  and  $\{x_i | i \in \mathcal{P}\}$  respectively can be vectorized as  $\mathbf{x}_{\mathcal{Q}}$  and  $\mathbf{x}_{\mathcal{P}}$ .

For convenience of express, we equivalently re-arrange the order of  $\mathbf{x}$  to get  $\mathbf{x} = [\mathbf{x}_{\mathcal{P}}^\top, \mathbf{x}_{\mathcal{Q}}^\top]^\top$ , thus the counterpart mean and covariance respectively are  $\boldsymbol{\mu} = [\boldsymbol{\mu}_{\mathcal{P}}^\top, \boldsymbol{\mu}_{\mathcal{Q}}^\top]^\top$  and

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathcal{P}\mathcal{P}} & \boldsymbol{\Sigma}_{\mathcal{P}\mathcal{Q}} \\ \boldsymbol{\Sigma}_{\mathcal{Q}\mathcal{P}} & \boldsymbol{\Sigma}_{\mathcal{Q}\mathcal{Q}} \end{bmatrix}.$$

Apparently,  $\mathbf{x}_{\mathcal{P}} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathcal{P}}, \boldsymbol{\Sigma}_{\mathcal{P}\mathcal{P}})$  and  $\mathbf{x}_{\mathcal{Q}} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathcal{Q}}, \boldsymbol{\Sigma}_{\mathcal{Q}\mathcal{Q}})$ . Given measured beams  $\mathbf{x}_{\mathcal{Q}}$ , the distribution of the un-measured beams  $\mathbf{x}_{\mathcal{P}}$  is a conditional multivariate Gaussian variable [46] following  $\mathcal{N}(\mathbf{x}_{\mathcal{P}}; \boldsymbol{\mu}_{\mathcal{P}|\mathcal{Q}}, \boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}})$  where

$$\boldsymbol{\mu}_{\mathcal{P}|\mathcal{Q}} = \boldsymbol{\mu}_{\mathcal{P}} + \boldsymbol{\Sigma}_{\mathcal{P}\mathcal{Q}}\boldsymbol{\Sigma}_{\mathcal{Q}\mathcal{Q}}^{-1}(\mathbf{x}_{\mathcal{Q}} - \boldsymbol{\mu}_{\mathcal{Q}}), \quad (13)$$

$$\boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}} = \boldsymbol{\Sigma}_{\mathcal{P}\mathcal{P}} - \boldsymbol{\Sigma}_{\mathcal{P}\mathcal{Q}}\boldsymbol{\Sigma}_{\mathcal{Q}\mathcal{Q}}^{-1}\boldsymbol{\Sigma}_{\mathcal{Q}\mathcal{P}}. \quad (14)$$

Based on the learned statistics  $\boldsymbol{\mu}, \boldsymbol{\Sigma}$  in (11) and (12), the conditional distribution  $\mathcal{N}(\mathbf{x}_{\mathcal{P}}; \boldsymbol{\mu}_{\mathcal{P}|\mathcal{Q}}, \boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}})$  has a closed-form expression based on (13) and (14).

The target of probing beam selection is to minimize the uncertainty of the unmeasured beams  $\mathbf{x}_{\mathcal{P}}$ , by finding a combination of measured beams  $\mathbf{x}_{\mathcal{Q}}$ . The conditional entropy of  $\mathbf{x}_{\mathcal{P}}$ , i.e.,

$$\mathbf{H}(\mathbf{x}_{\mathcal{P}}) = \frac{1}{2} \ln(2\pi e \det \boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}}), \quad (15)$$

can perform as an uncertainty measure of  $\mathbf{x}_{\mathcal{P}}$ . Obviously, the entropy  $\mathbf{H}(\mathbf{x}_{\mathcal{P}})$  is monotonous w.r.t.  $\det \boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}}$ . Given the maximal probing beam number  $L$ , the estimate statistics  $\boldsymbol{\mu}$  and  $\boldsymbol{\Sigma}$ , the subsequent probing beam selection problem considers the minimization of conditional entropy  $\mathbf{H}(\mathbf{x}_{\mathcal{P}})$  w.r.t. the set of measured beam indices  $\mathcal{Q}$ . This problem can be equivalently

written as a combinatorial optimization problem as follows

$$\begin{aligned} & \min_{\mathcal{Q}} \det \boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}} \\ & \text{s.t. } |\mathcal{Q}| \leq L. \end{aligned} \quad (16)$$

### III. ITERATIVE BEAM PREDICTION AND PROBING BEAM SELECTION

In Section II-B we have generally formulated the beam prediction and probing beam selection problems, and we will discuss the counterpart solutions below.

#### A. Weighted Entropy Minimization

In practice, most BA/T only considers the beams with high or maximum RSRP values, so we propose to rewrite problem (16) as

$$\begin{aligned} & \min_{\mathcal{Q}} \det(\boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}}^{\frac{1}{2}} \boldsymbol{\Sigma}_{\mathcal{P}|\mathcal{Q}} \boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}}^{\frac{1}{2}}) \\ & \text{s.t. } |\mathcal{Q}| = L, \end{aligned} \quad (17)$$

where the diagonal matrix  $\boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}}$  performs as a mask with its diagonal element indicating the weight. Obviously, the minimum of (16) is only obtained with  $|\mathcal{Q}| = L$ , and (16) is a special case of (17) with  $\boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}} = \mathbf{I}_{N-|\mathcal{Q}|}$ . We propose a mask function  $h$  w.r.t. the corresponding conditional mean, i.e.,

$$\boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}} = h(\boldsymbol{\mu}_{\mathcal{P}|\mathcal{Q}}). \quad (18)$$

When the mask is designed as  $[\boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}}]_{i^*i^*} = 1$  and  $[\boldsymbol{\Delta}_{\mathcal{P}|\mathcal{Q}}]_{jj} = 0, \forall j \neq i^*$  for  $i^* = \arg \max \boldsymbol{\mu}_{\mathcal{P}|\mathcal{Q}}$ , it indicates that only the beam with the maximum mean of RSRP is considered. In this work, we propose a heuristic mask design as

$$h(\mathbf{a}) = \text{sigmoid}(\beta(\mathbf{a} - \max \mathbf{a} + \alpha)), \quad (19)$$

where  $\mathbf{a}$  is an input vector with element scalar  $a$ ,  $\text{sigmoid}(a) = \frac{1}{1+\exp(-a)}$ ,  $\alpha$  and  $\beta$  respectively are defined as the threshold and fairness coefficients.

#### B. Iterative Beam Prediction and Probing Beam Selection

The primitive beam prediction problem (10) and the combinatorial optimization problem (16) are intractable, especially when the beam space  $N$  is large. First, the estimation of the covariance matrix  $\boldsymbol{\Sigma}$  is difficult, requiring a large amount of offline data and having a high computational complexity  $\mathcal{O}(N^3)$ . Besides, learning to generate a covariance matrix involves the matrix inversion operation in (9), which easily causes numerical stability issues in engineering. Second, to obtain the optimal solution of the combinatorial optimization problem (16), the computational complexity of the exhaustive search is  $\mathcal{O}(C_N^{|\mathcal{Q}|}(N - |\mathcal{Q}|)^3)$ , which causes a prohibitive overhead.

To address these issues and obtain a low-complexity feasible alternation, one way is to reduce the covariance matrix  $\boldsymbol{\Sigma}$  to be a diagonal matrix  $\boldsymbol{\Lambda}$ , and the subsequent problem (16) is simplified to a trivial question where the optimal combination is a subset of the first  $L$  indices with top diagonal values in



$\Lambda$ . Apparently, this reduction completely ignores the correlations across beams, resulting in non-negligible performance degradation.

To achieve a better tradeoff between performance and computational complexity, we propose to alternatively and iteratively address the problems (10) and (16) in a greedy manner. The iteration times is equal to  $L$ . The covariance matrix  $\Sigma$  is simplified to a diagonal matrix  $\Lambda$ . During each iteration, we learn and estimate  $\Lambda_{\mathcal{P}|\mathcal{Q}}$  using an iteration-specific variance network, instead of the closed-form solutions in (13) and (14). Subsequently, we select only one probing beam by  $\Lambda_{\mathcal{P}|\mathcal{Q}}$ . Although the conditional variance  $\Lambda_{\mathcal{P}|\mathcal{Q}}$  can be deducted with  $\Lambda$ , the estimation of  $\Lambda$  is coarse and thus the estimation error will propagate with the deduction. The proposed retraining of the variance network can reduce the error propagation issue. Specifically, in the  $l$ -th iteration, the beam prediction problem (10) is rewritten as

$$\max_{\Theta_f^l, \Theta_g^l} \mathbb{E}_{\mathbf{x}, \mathbf{s}, q^l} \left\{ \ln \mathcal{N}(\mathbf{x}_{\mathcal{P}^l}; \boldsymbol{\mu}_{\mathcal{P}^l|\mathcal{Q}^l}, \Lambda_{\mathcal{P}^l|\mathcal{Q}^l}) \right\}, \quad (20)$$

where

$$\boldsymbol{\mu}_{\mathcal{P}^l|\mathcal{Q}^l} = f(\mathbf{x}_{\mathcal{P}^l|\mathcal{Q}^l}, \mathcal{Q}^l, \mathbf{s}; \Theta_f^l), \quad (21)$$

$$\Lambda_{\mathcal{P}^l|\mathcal{Q}^l} = g(\mathcal{Q}^l, \mathbf{s}; \Theta_g^l), \quad (22)$$

and  $\mathcal{P}^l = \mathcal{P}^{l-1,*} \setminus \{q^l\}$ ,  $\mathcal{Q}^l = \mathcal{Q}^{l-1,*} \cup \{q^l\}$ ,  $\forall q \in \mathcal{P}^{l-1,*}$  with  $q^l$  being the candidate probing beam index for selection in this round,  $\mathcal{P}^{l-1,*}$  and  $\mathcal{Q}^{l-1,*}$  respectively are the set of candidate probing beam indices and the set of probing beam indices in the previous round. In the initial round,  $\mathcal{P}^{0,*} = \{1, \dots, N\}$ ,  $\mathcal{Q}^{0,*} = \emptyset$ . Then, the combinatorial optimization problem (16) is simplified as a one-dimensional search problem as

$$\min_{q^l \in \mathcal{P}^{l-1,*}} \det \Lambda_{\mathcal{P}^l|\mathcal{Q}^l}. \quad (23)$$

We denote the computational complexity of one inference of the variance network as  $\mathcal{O}(\omega)$ . Then, the computational complexity of (23) is  $\mathcal{O}(\omega(N-l+1))$ , and the total computational complexity of the greedy algorithm is  $\mathcal{O}(\omega \frac{L(2N-L+1)}{2})$ . At iteration  $l$ , the selected probing beam index is represented as  $q^{l,*}$ ,  $\mathcal{Q}^{l,*} = \mathcal{Q}^{l-1,*} \cup \{q^{l,*}\}$ , and  $\mathcal{P}^{l,*} = \{1, \dots, N\} \setminus \mathcal{Q}^{l,*}$ .

Considering the iterative solution with masking, the problem (23) is reformulated as

$$\min_{q^l \in \mathcal{P}^{l-1,*}} \det(\Delta_{\mathcal{P}^l|\mathcal{Q}^l} \Lambda_{\mathcal{P}^l|\mathcal{Q}^l}), \quad (24)$$

where  $\Delta_{\mathcal{P}^l|\mathcal{Q}^l}$  is the weight derived by (18) with  $\boldsymbol{\mu}_{\mathcal{P}^l|\mathcal{Q}^l}$ .

The offline collected training data is defined as  $\mathcal{D} = \{\mathbf{x}_j, \mathbf{s}_j\}_{j=1}^{N_s}$  where  $N_s$  is the sample number. The parameter sets  $\Theta_f$  and  $\Theta_g$  are iteratively updated by mini-batch gradient descent (MBGD) until convergence. Fig. 2(a) shows an illustrative ML prediction of RSRP, where the x and y axes respectively are the antenna numbers in the horizontal and vertical dimensions. The ground-truth and mean estimation of RSRP are given, with the optimal beams marked with red circles. The counterpart variance estimate and the probing beams (in the black square) are also depicted. In summary, at the offline training stage, the iterative beam prediction and

**Algorithm 1:** Iterative beam prediction and probing beam selection (offline training).

**Input:** Dataset  $\mathcal{D}$ , maximal probing beam number  $L$ .

**Output:** Learned networks  $\{f^{l,*}, g^{l,*}\}_{l=1}^L$ .

- 1 Initialize the set of candidate probing beam indices  $\mathcal{P}^{0,*} = \{1, \dots, N\}$ , the set of probing beam indices  $\mathcal{Q}^{0,*} = \emptyset$ .
- 2 **for**  $l = 1$  to  $L$  **do**
- 3     Choose  $\forall q^l \in \mathcal{P}^{l-1,*}$  randomly, and obtain  $\mathcal{P}^l = \mathcal{P}^{l-1,*} \setminus \{q^l\}$ ,  $\mathcal{Q}^l = \mathcal{Q}^{l-1,*} \cup \{q^l\}$ .
- 4     Train the mean network  $f$  and the variance network  $g$  by (20) with  $\mathcal{P}^l$ ,  $\mathcal{Q}^l$ .
- 5     Select the probing beam index  $q^{l,*}$  by (24) with learned  $f^{l,*}$  and  $g^{l,*}$ .
- 6     Update  $\mathcal{P}^{l,*} \leftarrow \mathcal{P}^{l-1,*} \setminus \{q^{l,*}\}$  and  $\mathcal{Q}^{l,*} \leftarrow \mathcal{Q}^{l-1,*} \cup \{q^{l,*}\}$ .
- 7 **end**

**Algorithm 2:** Iter-BP&PBS: beam selection for data transmission (online inference).

**Input:** Relative MU location  $\mathbf{s}$ , maximal probing beam number  $L$ .

**Output:** Beam for data transmission  $\hat{i}^*$ .

- 1 Initialize  $\mathcal{P}^{0,*} = \{1, \dots, N\}$ ,  $\mathcal{Q}^{0,*} = \emptyset$ , reload the learned mean and variance networks  $\{f^{l,*}, g^{l,*}\}_{l=1}^L$ .
- 2 **for**  $l = 1$  to  $L$  **do**
- 3     BS searches the optimal probing beam index  $q^{l,*}$  by (24) with learned  $f^{l,*}$  and  $g^{l,*}$ , and transmits the selected beam.
- 4     MU reports the counterpart RSRP  $x_{q^{l,*}}$  to the BS.
- 5     BS updates  $\mathcal{P}^{l,*} \leftarrow \mathcal{P}^{l-1,*} \setminus \{q^{l,*}\}$  and  $\mathcal{Q}^{l,*} \leftarrow \mathcal{Q}^{l-1,*} \cup \{q^{l,*}\}$ .
- 6 **end**
- 7 BS estimates the RSRP with the mean network  $f^{L,*}$ , and select the one with maximal predicted RSRP, i.e.,  $\hat{i}^*$ .

probing beam selection are given in Algorithm 1, and the counterpart online inference is named as Iter-BP&PBS and described in Algorithm 2.

#### IV. TWO-STAGE PROBING BEAM SELECTION

In Section III, we have proposed an iterative probing beam selection and beam prediction algorithm where the training beams are sequentially determined with RSRP feedback. As shown in Fig. 1(a), the proposed algorithm, i.e., Iter-BP&PBS, is similar to the binary search with  $\log_2(N)$  interactions. The difference mainly is in the selection of the probing beams, where our proposed scheme decides by an entropy-based criterion, and the latter decides by binary comparison of RSRP values. Iter-BP&PBS is still difficult to implement in practice for the following reasons.

- **Interaction latency.** In each iteration, the mask in (24) is updated by the RSRP feedback. Hence, the number of

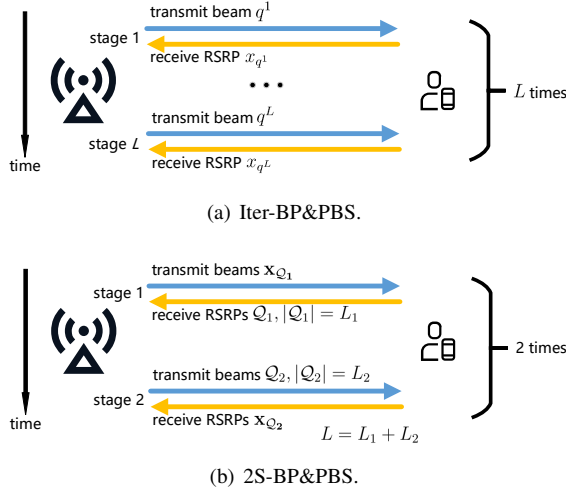


Fig. 1. Interactions between the BS and the MU.

interactions between the BS and the MU is  $L$ , and the latency linearly grows with the number of interactions.

- **Computational complexity.** In the  $l$ -th iteration, the computational complexity is  $\mathcal{O}(\omega(N - l + 1))$ , so the total computational complexity of  $L$  iterations is  $\mathcal{O}(\omega(\frac{L(2N-L+1)}{2}))$ .

Therefore, the significant interaction latency and computational complexity of Iter-BP&PBS maybe be unacceptable for a real-time system.

To address these issues, it is beneficial to design a location-aware probing codebook which is offline designed and online executed, and does not rely on the instantaneous feedback. As shown in Fig. 1(b), we propose a two-stage beam prediction and probing beam selection, i.e., 2S-BP&PBS. Compared to Iter-BP&PBS, the main revision is:

- In Iter-BP&PBS, the mask (18) is both location- and measurement-specific; but in 2S-BP&PBS, the mask is only location-specific by approximating the input  $\mu_{\mathcal{P}|\mathcal{Q}} \approx f(\mathbf{s})$ .

This means that all probing beams can be fully determined by a location-aware codebook. However, without instantaneous measurement, the location-aware mask is not precise enough to guide the probing beam selection. Thus, we propose a two-stage probing method, where  $L_1 < L$  beams are measured at the first stage, the BS receives the feedbacks and subsequently decides on  $L_2 (L = L_1 + L_2)$  probing beams in the second stage. Moreover, the feedbacks of the  $L_1$  beams experimentally are able to coarsely locate the strongest channel cluster. To further reduce the search complexity of the training beams in the second stage, we propose to select the top- $L_2$  beams w.r.t. the predicted RSRP for probing instead of the entropy-based beam selection.

Fig. 2(b) shows an illustrative ML prediction of RSRP. The optimal beams are marked with red circles in the ground-truth and mean estimation. The probing beams at the first and second stages respectively are labeled in black and grey squares. In the following, we will respectively introduce the two-stage probing beam selection, i.e., codebook- and prediction-based beam probings.

### A. Codebook-based Beam Probing

The area covered by the BS is assumed to be a rectangle, and is evenly divided into  $N_x \times N_y$  square grids. As shown in Fig. 3, the 2D location-specific codebook  $\mathcal{C}$  for probing is composed of  $N_x \times N_y$  grids, and grid  $i, j$  stores a codeword including a probing beam set  $\mathcal{Q}_{i,j}$  and a counterpart location  $\mathbf{s}_{i,j}$ , where  $i, j$  respectively are the indices of x-y coordinate. The length of a probing beam set is  $L_1$ , i.e.,  $|\mathcal{Q}_{i,j}| = L_1, \forall i, j$ . Regarding the  $l$ -th element in  $\mathcal{Q}_{i,j}$ , i.e.,  $q_{i,j}^l$ , it is obtained by (24), where  $\mu_{\mathcal{P}|\mathcal{Q}}$  to compute  $\Delta_{\mathcal{P}|\mathcal{Q}^l}$  in (18) is approximated by  $f(\mathbf{s}_{i,j})$ <sup>1</sup>. Thus, the probing codebook design is still entropy-based, and the corresponding procedure is summarized in Algorithm 3.

During online inference, we propose to select the probing beams in  $\mathcal{C}$  with the closest corresponding distance  $d(\mathbf{s}, \mathbf{s}_{i,j}), \forall i, j$  where  $d$  is a distance function. Formally, considering the Euclidean distance as  $d$ , the selected probing beams are given by

$$\begin{aligned} \mathcal{Q}^{L_1,*} &= \mathcal{Q}_{i,j}|_{i=i^*, j=j^*} \text{ where} \\ i^*, j^* &= \arg \min_j \|\mathbf{s} - \mathbf{s}_{i,j}\|_2^2, 1 \leq i \leq N_x, 1 \leq j \leq N_y. \end{aligned} \quad (25)$$

The codebook  $\mathcal{C}$  is represented in the form of a binary tree, and the computational complexity of (25) is  $\mathcal{O}(\log_2 N_x N_y)$  with binary search.

---

#### Algorithm 3: Probing codebook design.

---

**Input:** Dataset  $\mathcal{D}$ , first probing beam number  $L_1$ .

**Output:** Probing codebook  $\mathcal{C} = \{\{\mathcal{Q}_{i,j}, \mathbf{s}_{i,j}\}_{i=1}^{N_x}\}_{j=1}^{N_y}$ .

---

- 1 Initialize  $\mathcal{P}^{0,*} = \{1, \dots, N\}$ ,  $\mathcal{Q}^{0,*} = \emptyset$ , iteration times  $L$ , location-specific codebook  $\mathcal{C} = \{\{\mathcal{Q}_{i,j}, \mathbf{s}_{i,j}\}_{i=1}^{N_x}\}_{j=1}^{N_y}$  where  $\mathcal{Q}_{i,j} = \emptyset, \forall i, j$ .
  - 2 Train the mean network  $f$  only with MU location.
  - 3 **for**  $l = 1$  to  $L_1$  **do**
  - 4   Choose  $\forall q^l \in \mathcal{P}^{l-1,*}$  randomly, and obtain  $\mathcal{P}^l = \mathcal{P}^{l-1,*} \setminus \{q^l\}$ ,  $\mathcal{Q}^l = \mathcal{Q}^{l-1,*} \cup \{q^l\}$ .
  - 5   Train the mean network  $f^l$  and the variance network  $g^l$  by (20) with  $\mathcal{P}^l, \mathcal{Q}^l$ .
  - 6   **Output:** Learned networks  $f^{l,*}$  and  $g^{l,*}$ .
  - 7   Select the probing beam index  $q^{l,*}$  by (24) with learned  $f^{l,*}$  and  $g^{l,*}$ .
  - 8   Update  $\mathcal{P}^{l,*} \leftarrow \mathcal{P}^{l-1,*} \setminus \{q^{l,*}\}$  and  $\mathcal{Q}^{l,*} \leftarrow \mathcal{Q}^{l-1,*} \cup \{q^{l,*}\}$ .
  - 9   **for**  $i = 1$  to  $N_x$  **do**
  - 10    **for**  $j = 1$  to  $N_y$  **do**
  - 11      Select the probing beam index  $q_{i,j}^{l,*}$  by (24) with learned  $f^{l,*}$  and  $g^{l,*}$ , where  $\mu_{\mathcal{P}|\mathcal{Q}}$  in (18) is approximated by  $f(\mathbf{s}_{i,j})$ .
  - 12      Update  $\mathcal{Q}_{i,j} \leftarrow \mathcal{Q}_{i,j} \cup \{q_{i,j}^{l,*}\}$ .
  - 13    **end**
  - 14 **end**
- 

<sup>1</sup>Noting that  $\mathbf{A}_{\mathcal{P}^l|\mathcal{Q}^l}$  in (24) remains unchanged.

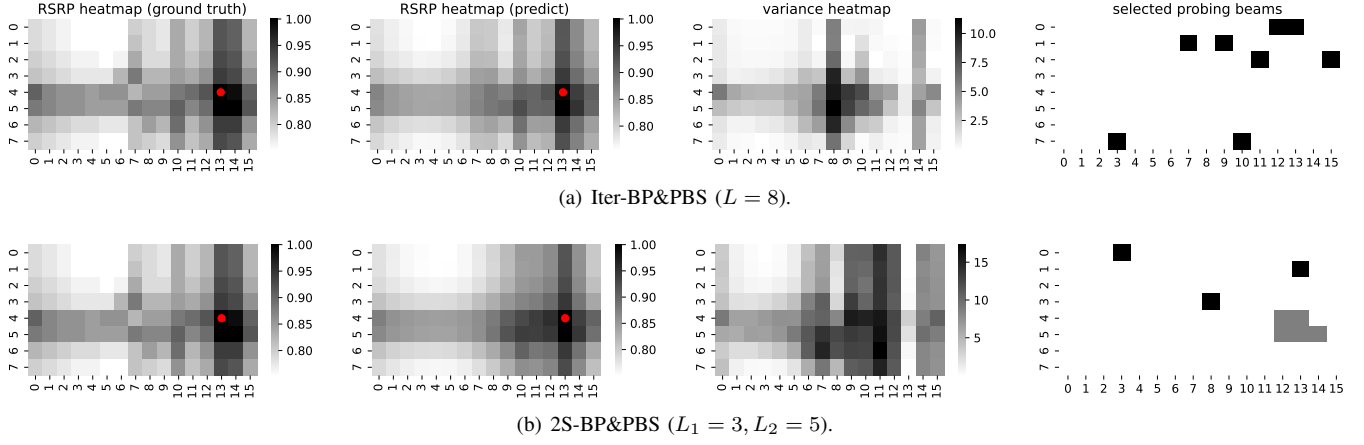


Fig. 2. Illustrations of selected probing beams and prediction results.

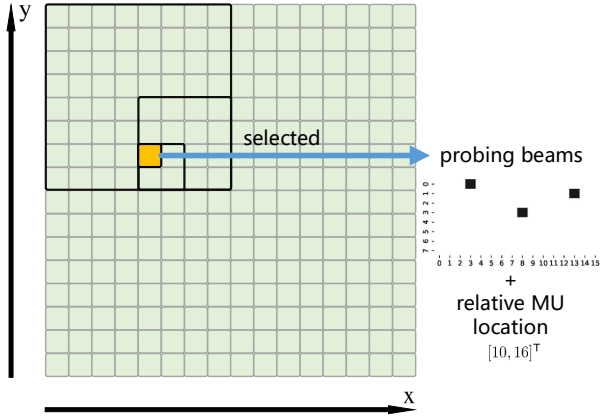


Fig. 3. Binary searching in the 2D location-specific probing codebook.

### B. Prediction-based Beam Probing

In the second stage, the BS transmits the probing beams with indices  $Q^{L_1,*}$  and receives the corresponding RSRPs  $\mathbf{x}_{Q^{L_1,*}}$ , predicts the RSRP  $\mu_{\mathcal{P}^{L_1,*}|Q^{L_1,*}}$  with the mean network  $f^{L_1,*}$ . Using the RSRPs  $\mathbf{x}_{Q^{L_1,*}}$ , accurate estimation of the optimal beam poses a challenge for the predictor. However, the predictor is still capable to roughly locate the optimal beam. Hence, we directly select the top- $L_2$  beams from the predicted RSRP  $\mu_{\mathcal{P}^{L_1,*}|Q^{L_1,*}}$  as the second probing beams, i.e.,  $Q^{L_2,*}$ . The MU then reports the counterpart RSRPs  $\mathbf{x}_{Q^{L_2,*}}$  to the BS. The BS replaces the corresponding prediction result with the measurement  $\mathbf{x}_{Q^{L_2,*}}$ , and selects the beam with the maximum RSRP for data transmission, i.e.,  $\hat{i}^*$ .

In summary, the procedure of the 2S-BP&PBS is clarified as follows. In the first interaction, the BS is aware of the MU location, selects  $L_1$  beams from the codebook  $\mathcal{C}$  for probing, and receives the RSRP feedbacks from the MU. In the second interaction, the BS predicts the RSRP in beamspace with the MU location and the RSRP feedbacks, selects the top- $L_2$  beams as the probing beams to re-measure the RSRP, and decides the beam for data transmission with the measurement. The proposed 2S-BP&PBS only has twice interactions and a computational complexity  $\mathcal{O}(\omega + \log_2 N_x N_y) \approx \mathcal{O}(\omega)$ , and

---

### Algorithm 4: 2S-BP&PBS: two-level probing beam selection (online inference).

---

**Input:** Relative MU location  $\hat{\mathbf{s}}$ , first and second probing beam numbers  $L_1, L_2$ .

**Output:** Beam for data transmission  $\hat{i}^*$ .

---

- 1 Reload the mean network  $f^{L_1,*}$ , probing codebook  $\mathcal{C}$ .
  - 2 **Stage 1**
  - 3 BS searches the nearest probing beams  $Q^{L_1,*}$  in  $\mathcal{C}$  with  $\mathbf{s}$  by (25), and transmits the beams.
  - 4 MU reports the counterpart RSRPs  $\mathbf{x}_{Q^{L_1,*}}$  to the BS.
  - 5 BS estimates the RSRP  $\mu_{\mathcal{P}^{L_1,*}|Q^{L_1,*}}$  with the mean network  $f^{L_1,*}$ .
  - 6 **Stage 2**
  - 7 BS selects the top- $L_2$  beams by the predicted RSRP  $\mu_{\mathcal{P}^{L_1,*}|Q^{L_1,*}}$  as the second probing beams, i.e.,  $Q^{L_2,*}$ .
  - 8 MU reports the counterpart RSRPs  $\mathbf{x}_{Q^{L_2,*}}$  to the BS.
  - 9 BS replaces the corresponding prediction result with the measurement  $\mathbf{x}_{Q^{L_2,*}}$ , and selects the beam with maximal RSRP as the beam for data transmission, i.e.,  $\hat{i}^*$ .
- 

the corresponding online inference is given in Algorithm 4.

## V. DEEP LEARNING-ENABLED BEAM PREDICTOR

In Section III-B, we have generally proposed the DL-enabled mean and variance networks, i.e.,  $f$  and  $g$ . To achieve formidable learning capabilities, we design  $f$  and  $g$  with the transformer [47] which is a transduction model that relies on self-attention to compute representations of its input and output. As shown in Fig. 4, both the networks are composed of three sequential blocks, i.e., an embedding layer, a transformer, and an output layer. In this work, we focus on the designs of the embedding and the output layers, while the transformer is quite mature in the literature, so we directly use it as the backbone.

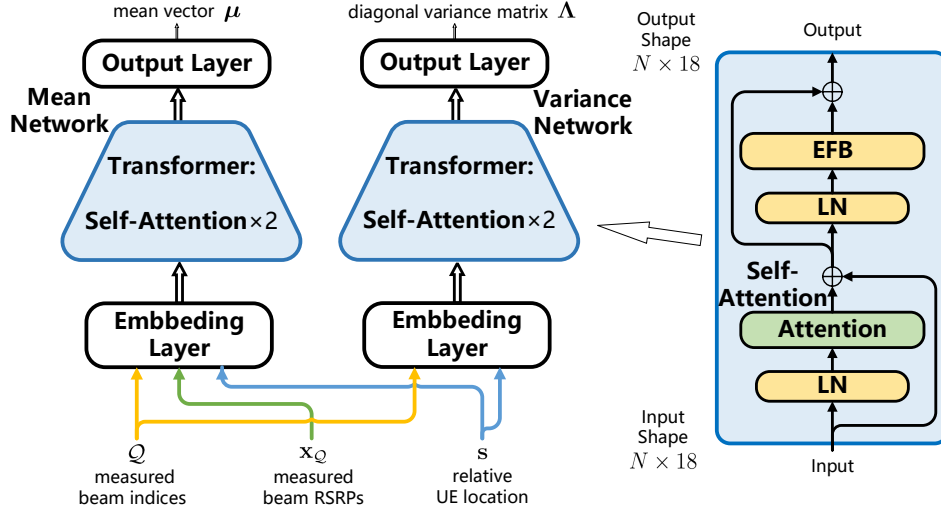


Fig. 4. An illustration of the mean and variance networks (left) and a transformer (right).

#### A. Embedding and Output Layers

In Section III-B, the dimension of the input  $\mathbf{x}_Q \in \mathbb{R}^{|Q| \times 1}$  is a variable w.r.t. the measured beam number  $|Q|$ . To design a network scalable to  $|Q|$ , we propose to equivalently transform  $\mathbf{x}_Q$  into  $\mathbf{x}'_Q \in \mathbb{R}^{N \times 2}$  where the  $i$ -th row vector  $\mathbf{x}'_i$  is

$$\mathbf{x}'_i = \begin{cases} [x_q, 1], & i \in Q, \\ [0, 0], & \text{otherwise,} \end{cases} \quad (26)$$

where the first element is the RSRP value, and the second element indicates whether the corresponding beam is selected or not, to distinguish the measured and unmeasured beams both having zero RSRP values. Then,  $\mathbf{x}'_Q$  is projected as  $\mathbf{x}''_Q \in \mathbb{R}^{N \times 16}$ , with a one-dimensional convolution operator<sup>2</sup> and the RELU function, i.e.,  $\tilde{a} = \max(0, a)$ . The relative MU location  $s$  is linearly projected as  $s' \in \mathbb{R}^{N \times 1}$ . Additionally, a bias  $\text{cls} \in \mathbb{R}^{N \times 1}$ , i.e., the class token in [47], is introduced.  $\mathbf{x}''_Q$ ,  $s'$  and  $\text{cls}$  are concatenated as an input in  $\mathbb{R}^{N \times 18}$  for the cascaded transformer. The above embedding for the mean network  $f$  is also applicable for the variance network  $g$ . The only difference is that the map  $g$  in (22) is conducted without measured RSRPs, so the first column of  $\mathbf{x}'_Q$  are all zero.

The output of the transformer, i.e., the input for the output layer, is in the space  $\mathbb{R}^{N \times 18}$ , and the counterpart output is in the space  $\mathbb{R}^{N \times 1}$ , with a one-dimensional convolution and a RELU.

#### B. Transformer

As shown in Fig. 4, the proposed transformer  $g$  is a stack of different operations, layers and modules. The layer normalization (LN) operation is used to speed up training by normalizing the data into a standard normal distribution. In an expansion forward block (EFB), the input is linearly projected into an expanded space, after an activation layer, the expanded vector is projected back into the primary space. Residual connection

is used to solve the training loss degradation problem in very deep networks by introducing an identity map. In Fig. 4, the residual connection operations are denoted by  $\oplus$ .

The core module in a transformer is self-attention, which is an attention mechanism that relates different positions of a single sequence to compute a representation of the sequence. An attention function can be described as mapping a query and a set of key-value pairs to an output, where the query, keys, values, and output are all vectors. The output is computed as a weighted sum of the values, where the weight assigned to each value is computed by a compatibility function of the query with the corresponding key. Given the input matrix  $\mathbf{A} \in \mathbb{R}^{M \times \tilde{M}}$ , the key, query and value vectors are

$$\mathbf{Q} = \mathbf{W}_q \mathbf{A}, \quad (27a)$$

$$\mathbf{K} = \mathbf{W}_k \mathbf{A}, \quad (27b)$$

$$\mathbf{V} = \mathbf{W}_v \mathbf{A}, \quad (27c)$$

where superscript  $(\tilde{\cdot})$  denotes the output marker,  $\mathbf{W}_q$ ,  $\mathbf{W}_k$ ,  $\mathbf{W}_v$  respectively are the corresponding trainable linear transformation matrices,  $\mathbf{W}_q$  and  $\mathbf{W}_k$  both are in  $\mathbb{R}^{N_d \times M}$  where  $N_d$  is the feature dimension of the key matrix, and  $\mathbf{W}_v \in \mathbb{R}^{M \times M}$ . The result of the scaled dot-product attention is

$$\tilde{\mathbf{A}} = \mathbf{V} \text{softmax} \left( \frac{\mathbf{K}^T \mathbf{Q}}{\sqrt{N_d}} \right), \quad (28)$$

where  $\text{softmax}(\tilde{\mathbf{a}}) = \frac{\exp(\tilde{\mathbf{a}})}{\sum_j \exp(\tilde{\mathbf{a}}_j)}$ . Since the variance of the inner product of  $\mathbf{K}$  and  $\mathbf{Q}$  increases with increasing embedding size, the result of the product is scaled by  $\sqrt{N_d}$ .

## VI. SIMULATIONS

#### A. Configuration

To simulate spatial consistency and location-aware beam probing and prediction, the mmWave channel is established as a map-based hybrid model according to 3GPP 38.901 clause 8 [45], consisting of deterministic and stochastic components. The deterministic ray tracing model is established by Feko

<sup>2</sup>In this work, we only use a kernel size of 1, to make the input and output dimensions be equal, and also increase or decrease the data channel.





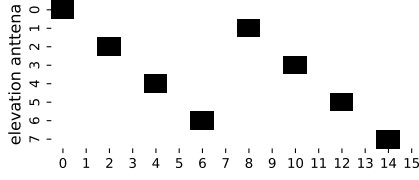
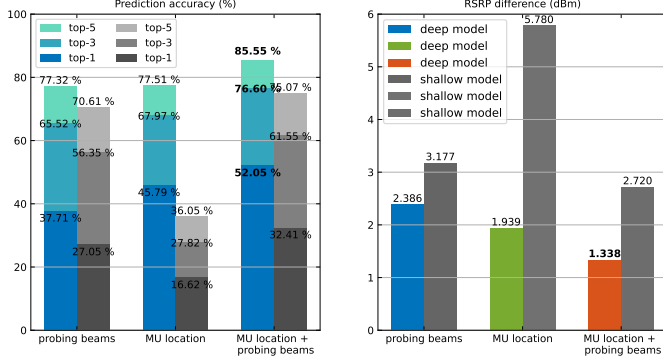
Fig. 7. Uniform probing beams plotted in black squares  $L = 8$ .

Fig. 8. A comparison of prediction performance with different input combinations. The results with shallow and the proposed deep networks respectively are shown in grey and color.

the shallow network is obtained by replacing the transformer backbone with 2 fully connected layers and each layer contains 128 neurons. As shown in Fig. 8, the prediction performance of the proposed deep network (in color) significantly outperforms that of the shallow network (in grey), indicating that the potential map between the probing beams and/or MU location to the RSRP distribution is complex, and thus sufficient network depth can improve the prediction.

On the other hand, considering different input combinations, the prior of MU spatial information significantly improves the RSRP prediction performance, and the top-1 accuracy of MU location is up to 45.79% which is better than 37.71% of prob-

ing beams. This implies that location-aware probing beam-free BA/T is feasible if the requirement of RSRP difference is not critical. In the following, the scheme of the deep model with 8 probing beams & MU location performs as a benchmark of uniform probing.

#### D. Location-aware Beam Probing

The above subsections consider uniform probing beams, here we further discuss the influence of non-uniform probing beams. The location-specific probing beams can be designed using Iter-BP&PBS or 2S-BP&PBS, weighted or uniform entropy minimizations. In 2S-BP&PBS, the coverage area is uniformly divided into  $2 \times 2$  m<sup>2</sup> square grids.

As plotted in Fig. 9,  $L_1 = 3$  indicates that the 2S-BP&PBS transmits probing beams twice, and the first and second probing beam numbers respectively are  $L_1 = 3$  and  $L_2 = 5$ . Compared to Iter-BP&PBS, 2S-BP&PBS significantly improves the top-1 prediction accuracy and reduces the RSRP difference, by directly measuring the top- $L_2$  beams in the second interaction. Hence, the top- $K$  ( $K > 1$ ) accuracies of 2S-BP&PBS are worse than those of Iter-BP&PBS. Meanwhile, the proposed Iter-BP&PBS and 2S-BP&PBS significantly outperform the baseline with about 5 dB gain in RSRP difference.

In terms of prediction MSE with mask in (19), the proposed Iter-BP&PBS without mask has achieved the best result. This is because the use of weighted entropy minimization sacrifices the overall MSE performance to achieve a local entropy minimization on the areas with high predicted RSRP values. The mask design is consistent with the link-level BA/T whereas only the maximum RSRP is concerned. Moreover, our proposed schemes are also potential for the system-level BA/T optimizations, since we have predicted the RSRP in the whole beamspace and thus the estimation of the channel in the inference direction is feasible.

Considering Iter-BP&PBS, the RSRP difference versus probing beam number  $L$  is depicted in Fig. 10(a). When  $L \leq 3$ , the performance of Iter-BP&PBS with the mask rapidly improves, then the improvement slows down as  $L$  continues

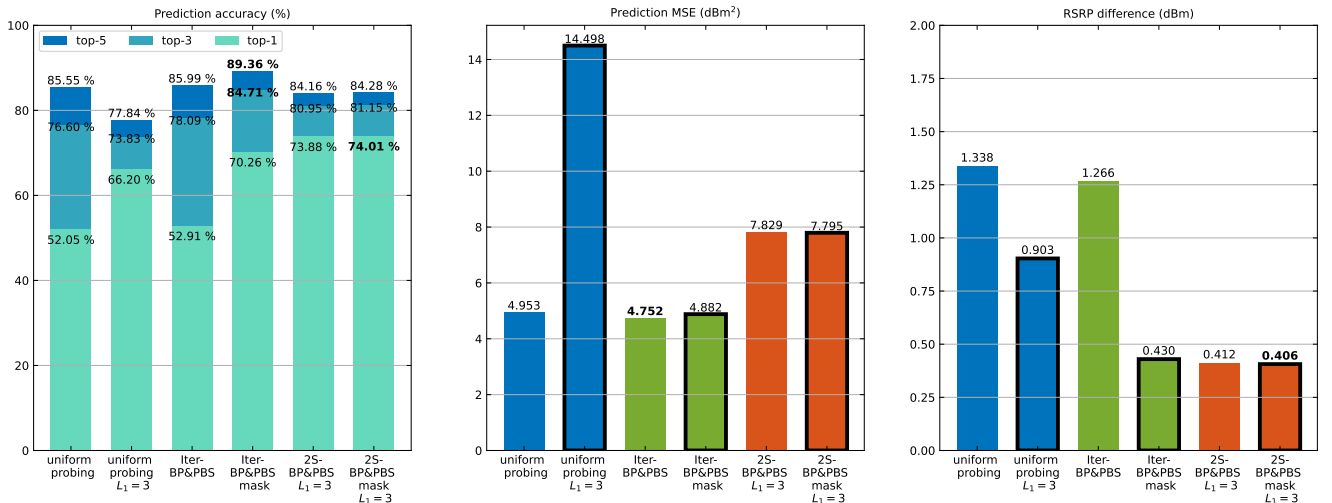


Fig. 9. Prediction results of the Iter-BP&amp;PBS and 2S-BP&amp;PBS with/without mask.

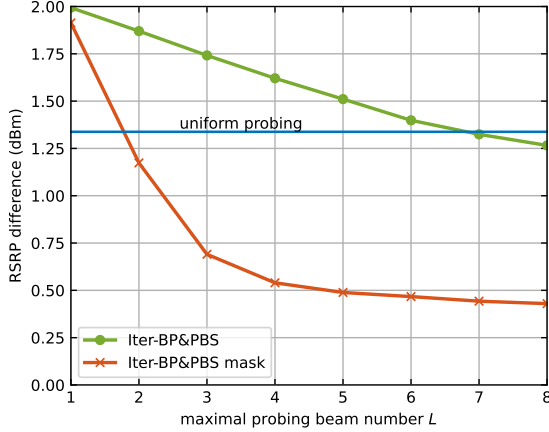
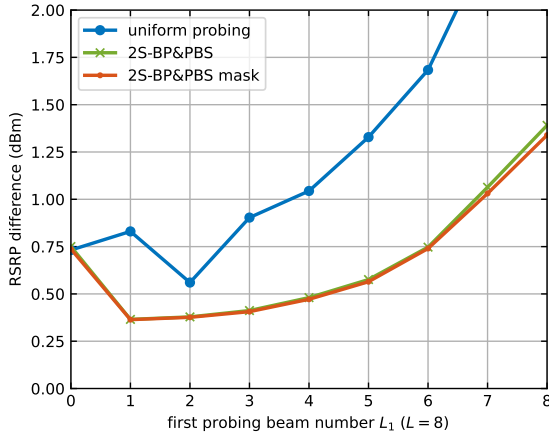
(a) Maximal probing beam number  $L$ .(b) First probing beam number  $L_1$  ( $L = 8$ ).

Fig. 10. RSRP difference versus first probing beam number.

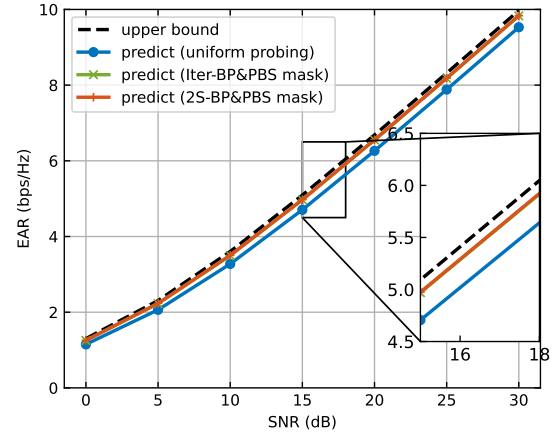
to increase. Meanwhile, Iter-BP&PBS without the mask has a bad performance, and it is comparable to the baseline when  $L = 7$ . This indicates that the proposed mask can help the beam predictor roughly locate the strongest cluster with a few probing beams, but an accurate prediction of the optimal beam is difficult. Thus, when the strongest cluster is roughly located, it would be better to re-design the mask in (19) or just probe the beams with the top- $L_2$  predicted RSRPs.

Considering 2S-BP&PBS, the RSRP difference versus first probing beam number  $L_1$  is shown in Fig. 10(b). The uniform probing beams with  $L < 8$  are designed so that the beams are uniformly located in beamspace. In particular,  $L_1 = 0$  is the scheme that probes the top- $L$  beams predicted by the networks with only MU location, and the beam probing at the first stage is removed, thus requiring an interaction. The scheme with  $L_1 = 8$  is the one that probes the top- $L$  beams predicted by the networks with only MU location, the beam probing at the second stage is removed and an interaction is also required. The scheme with  $L_1 = 0$  is better than the one with  $L_1 = 8$ , indicating the importance of obtaining the optimal beam

by measurement in the second stage. On the other hand, the proposed 2S-BP&PBS significantly outperforms the baseline, and schemes with or without mask have similar performance, implying that the gain is mainly achieved by the entropy-based probing beam selection.

### E. Data Transmission

The performance degradation of single-user BA/T is mainly determined by the beam alignment ratio, i.e., top-1 accuracy. Regarding the influence of mis-alignment, there exists an EAR performance gap between the prediction-based schemes and the upper bound where the transmission beam is assumed to be correctly aligned without any training overhead. As shown in Fig. 9, the proposed schemes can align the transmission beam with a top-1 accuracy about 70%, which is far from 100% in a numerical sense. However, the mis-alignment does not mean that the transmission will fail to work, since the sub-optimal beams also have near-optimal RSRP. Thus, as shown in Fig. 11, the EAR performance versus signal-to-noise-ratio (SNR) of the proposed Iter-BP&PBS and 2S-BP&PBS is very close to the upper bound.

Fig. 11. EAR versus SNR ( $U = 1$ ).

In terms of hierarchical measurement-based BA/T, the single-user overhead of two-level search is 24 including 16 wide beams and 8 narrow beams, and the overhead of binary search is  $2\log_2 N = 14$  with 7 interactions. As we consider multiple MUs, the curves of EAR versus user number  $U$  are demonstrated in Fig. 12. When  $U$  grows up to 100, all the time resources are consumed in the binary search, more than 60% time resources are consumed in the two-level search. Meanwhile, the prediction-based scheme has about 35% drop due to overhead cost.

Moreover, we list the average computational time and storage costs of the proposed schemes in Table III. Considering the storage cost, 3.144 MB and 0.321 MB respectively are the storage space of a mean and variance network and a location-aware codebook. At the online inference stage, the mean network performs only one inference in 2S-BP&PBS and the computational time cost is 2.10 ms, which roughly satisfies

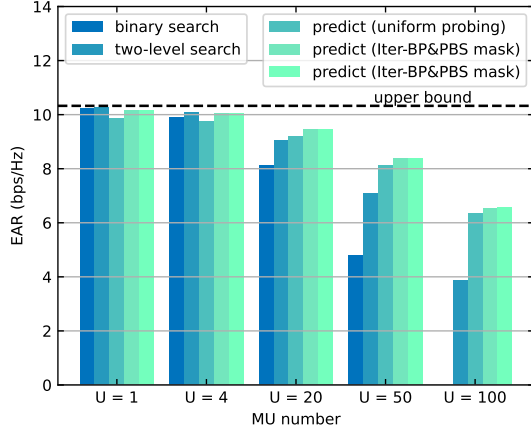
Fig. 12. EAR versus user number  $U$ .

TABLE III  
AVERAGE COMPUTATIONAL TIME AND STORAGE COSTS

	two-level search	binary search	Iter-BP&PBS	2S-BP&PBS
storage cost (MB)	\	\	25.152	3.144 + 0.321
computational time cost (ms)	$\approx 0$	$\approx 0$	34.08	2.10
number of interactions	2	7	8	2

real-time deployment. Meanwhile, the computational time cost in Iter-BP&PBS is 34.08 ms, which is greatly accelerated by the parallel computation in the GPU.

## VII. CONCLUSIONS

In this work, we investigated the joint probing beam selection and probabilistic beam prediction, and formulated it as an entropy minimization problem. To solve this problem, we proposed an iterative scheme (Iter-BP&PBS) with a simplified diagonal covariance matrix. To further reduce the number of interactions and the computational complexity of the Iter-BP&PBS, we proposed a two-stage probing beam selection scheme, i.e., 2S-BP&PBS. Simulation results demonstrated the superiority of the proposed schemes compared to the existing hierarchical beam search and beam prediction with uniform probing beams. In our future study, we will extend the entropy-based method and utilize the channel prior in frequency and time domains, for probing beam selection and beam prediction.

## REFERENCES

- [1] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, and T. Thomas, "LTE-advanced: Next-generation wireless broadband technology [Invited Paper]," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 10–22, 2010.
- [2] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Björnson, K. Yang, C.-L. I, and A. Ghosh, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, 2017.
- [3] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, 2013.
- [4] S. Han, C.-I. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5g," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, 2015.
- [5] C. Jeong, J. Park, and H. Yu, "Random access in millimeter-wave beamforming cellular networks: issues and approaches," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 180–185, 2015.
- [6] J. Wang, Z. Lan, C. woo Pyo, T. Baykas, C. sean Sum, M. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-gbps millimeter-wave wpan systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, 2009.
- [7] J. Wang, Z. Lan, C.-S. Sum, C.-W. Pyo, J. Gao, T. Baykas, A. Rahman, R. Funada, F. Kojima, I. Lakkis, H. Harada, and S. Kato, "Beamforming codebook design and performance evaluation for 60ghz wideband wpans," in *2009 IEEE 70th Veh. Technol. Conf. Fall*, 2009, pp. 1–6.
- [8] 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, 2016.
- [9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.
- [10] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, AI-based, or both?" *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7331–7376, 2019.
- [11] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-driven deep learning for physical layer communications," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 77–83, 2019.
- [12] J. Zhang, G. Zheng, I. Krikidis, and R. Zhang, "Fast specific absorption rate aware beamforming for downlink SWIPT via deep learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16 178–16 182, 2020.
- [13] C. Xu, S. Liu, Z. Yang, Y. Huang, and K.-K. Wong, "Learning rate optimization for federated learning exploiting over-the-air computation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 12, pp. 3742–3756, 2021.
- [14] J. Su, F. Meng, S. Liu, Y. Huang, and Z. Lu, "Learning to predict and optimize imperfect MIMO system performance: Framework and application," in *Proc. 41-th IEEE Global Commun. Conf. (GLOBECOM'22)*, 2022, pp. 335–340.
- [15] J. Yang, S. Jin, C.-K. Wen, J. Guo, M. Matthaiou, and B. Gao, "Model-based learning network for 3-D localization in mmWave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5449–5466, 2021.
- [16] S. Fan, Y. Wu, C. Han, and X. Wang, "SIABR: A structured intra-attention bidirectional recurrent deep learning method for ultra-accurate terahertz indoor localization," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 2226–2240, 2021.
- [17] Z. Liu, M. del Rosario, and Z. Ding, "A markovian model-driven deep learning framework for massive mimo csi feedback," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1214–1228, 2022.
- [18] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Overview of deep learning-based csi feedback in massive mimo systems," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8017–8045, 2022.
- [19] J. Zhang, Y. Huang, Y. Zhou, and X. You, "Beam alignment and tracking for millimeter wave communications via bandit learning," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5519–5533, 2020.
- [20] J. Zhang, Y. Huang, J. Wang, X. You, and C. Masouros, "Intelligent interactive beam training for millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 2034–2048, 2021.
- [21] F. Meng, S. Liu, Y. Huang, and Z. Lu, "Learning-aided beam prediction in mmWave MU-MIMO systems for high-speed railway," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 693–706, 2022.
- [22] K. Ma, Z. Wang, W. Tian, S. Chen, and L. Hanzo, "Deep learning for mmwave beam-management: State-of-the-art, opportunities and challenges," *IEEE Wireless Commun.*, pp. 1–8, 2022.
- [23] R. Yang, Z. Zhang, X. Zhang, C. Li, Y. Huang, and L. Yang, "Meta-learning for beam prediction in a dual-band communication system," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 145–157, 2023.
- [24] W. He, C. Zhang, Y. Huang, and X. You, "Intelligent optimization of base station array orientations via scenario-specific modeling," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 2117–2130, 2022.
- [25] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter wave beam-selection using out-of-band spatial information," *IEEE Trans. Wireless Commun.*, vol. 17, no. 2, pp. 1038–1052, 2018.
- [26] W. Xu, F. Gao, X. Tao, J. Zhang, and A. Alkhateeb, "Computer vision aided mmwave beam alignment in v2x communications," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2022.
- [27] H.-L. Song and Y.-C. Ko, "Beam alignment for high-speed uav via angle prediction and adaptive beam coverage," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 185–10 192, 2021.



- [28] M. Hussain and N. Michelusi, "Learning and adaptation for millimeter-wave beam tracking and training: A dual timescale variational framework," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 37–53, 2022.
- [29] K. Ma, F. Zhang, W. Tian, and Z. Wang, "Continuous-time mmwave beam prediction with ode-lstm learning architecture," *IEEE Wireless Commun. Letters*, vol. 12, no. 1, pp. 187–191, 2023.
- [30] S. H. A. Shah and S. Rangan, "Multi-cell multi-beam prediction using auto-encoder lstm for mmwave systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10366–10380, 2022.
- [31] 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, 2020.
- [32] F. Gao, B. Lin, C. Bian, T. Zhou, J. Qian, and H. Wang, "Fusionnet: Enhanced beam prediction for mmwave communications using sub-6 ghz channel and a few pilots," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8488–8500, 2021.
- [33] K. Ma, S. Du, H. Zou, W. Tian, Z. Wang, and S. Chen, "Deep learning assisted mmwave beam prediction for heterogeneous networks: A dual-band fusion approach," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 115–130, 2023.
- [34] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1G to 5G," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 1124–1148, 2018.
- [35] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmwave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, 2017.
- [36] Y. Zeng and X. Xu, "Toward environment-aware 6g communications via channel knowledge map," *IEEE Wireless Commun.*, vol. 28, no. 3, pp. 84–91, 2021.
- [37] K. Satyanarayana, M. El-Hajjar, A. A. M. Mourad, and L. Hanzo, "Deep learning aided fingerprint-based beam alignment for mmWave vehicular communication," *IEEE Trans. Veh. Technol.*, vol. 68, no. 11, pp. 10858–10871, 2019.
- [38] T.-H. Chou, N. Michelusi, D. J. Love, and J. V. Krogmeier, "Fast position-aided mimo beam training via noisy tensor completion," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 3, pp. 774–788, 2021.
- [39] J. Zhang and C. Masouros, "Learning-based predictive transmitter-receiver beam alignment in millimeter wave fixed wireless access links," *IEEE Trans. Signal Process.*, vol. 69, pp. 3268–3282, 2021.
- [40] W. Xu, Y. Ke, C.-H. Lee, H. Gao, Z. Feng, and P. Zhang, "Data-driven beam management with angular domain information for mmwave uav networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7040–7056, 2021.
- [41] C. Zhang, M. Wang, W. He, L. Zhang, and Y. Huang, "Channel beam pattern extension for massive MIMO via deep gaussian process regression," in *2021 IEEE/CIC Int. Conf. Commun. China (ICCC)*, 2021, pp. 172–177.
- [42] Y. Heng and J. G. Andrews, "Machine learning-assisted beam alignment for mmWave systems," in *2019 IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, 2019.
- [43] 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, 2020.
- [44] Y. Heng, J. Mo, and J. G. Andrews, "Learning site-specific probing beams for fast mmwave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 5785–5800, 2022.
- [45] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3GPP TR 38.901, Jan. 2020, version 16.1.0.
- [46] M. L. Eaton, "Multivariate statistics: a vector space approach," *JOHN WILEY & SONS, INC., 605 THIRD AVE., NEW YORK, NY 10158, USA*, 1983.
- [47] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, "Attention is all you need," in *NIPS*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30, 2017.
- [48] WinProp, "Wave propagation and radio network planning software (part of altair hyperworks)." [Online]. Available: <https://www.altair.com/>