A Low-Complexity DNN-Based DoA Estimation Method for EHF and THF Cell-Free Massive MIMO

Seyyed Saleh Hosseini
Dept. of Electrical & Computer Eng.
McGill University
Montreal, Canada
seyyed.hosseini@mail.mcgill.ca

Benoit Champagne
Dept. of Electrical & Computer Eng.
McGill University
Montreal, Canada
benoit.champagne@mcgill.ca

Xiao-Wen Chang
School of Computer Science
McGill University
Montreal, Canada
chang@cs.mcgill.ca

***Abstract*—We study the problem of direction of arrival (DoA) estimation for cell-free massive MIMO (m-MIMO) systems operating over extremely high frequency (EHF) and terahertz (THF) bands, where the wireless channel can effectively be modeled by a line-of-sight path. For this model, a low-complexity deep neural network (DNN)-based method is proposed to estimate the DoA of a radio wave impinging on an access point (AP) equipped with an antenna array. To train the DNN, a special feature set is proposed obtained from the first superdiagonal entries of the spatial correlation matrix. This selection of features makes it possible to employ a DNN with only a few low-dimensional layers, which considerably speeds up training and processing. More importantly, it is shown that the trained DNN is robust against quantization noise in the array snapshot data. This property makes the centralized implementation of the proposed DNN-based method feasible, which is particularly well-suited for cell-free m-MIMO. Through extensive simulations, the new method is shown to achieve an estimation performance that nearly matches or exceeds that of classical bechmark methods, but with considerably reduced complexity.**

***Index Terms*—DoA estimation, cell-free massive MIMO, neural network, LoS communications.**

I. Introduction

Network densification is a highly effective approach to increase capacity, as needed to support high-data-rate applications envisaged for 5G and beyond. An emerging framework for achieving network densification is provided by cell-free massive (m-MIMO), wherein a large number of access points (APs) are scattered over a wide area without partitioning the network into bounded cells [1]–[3]. The APs are connected via high-speed fronthaul links to a central processing unit (CPU), which coordinates the various signal processing tasks. Such a dense network offers several advantages, e.g.: a high spectral efficiency, dispensing with the need for frequent handover requests, and taking advantage of large-scale fading diversity.

Besides network densification, data rates can be dramatically enhanced by transmitting over the extremely high frequency (EHF) or terahertz frequency (THF) bands, where tremendous bandwidth is accessible. However, the high-level of frequency-dependent path loss over the EHF and THF bands, due to spreading, scattering and molecular absorption [4], often leads to very high attenuation in the received signal power, especially for non-line-of-sight (NLoS) paths. For example, the observed attenuation for the *main* NLoS paths, compared to the LoS, can be as high as 20dB in THF communications [4], [5]. This phenomenon results in the dominance of the LoS over the NLoS paths [5]–[10].

The use of the EHF and THF bands in cell-free m-MIMO systems is highly auspicious [11], [12], especially in large public environments where the densification of APs can be realized while maintaining strong LoS radio links. To achieve a high spectral efficiency in cell-free m-MIMO, demanding signal processing tasks such as distributed beamforming, pilot assignment, and power allocation, must be performed by the APs and the CPU. These tasks require the system to accurately estimate physical parameters of the transmission channels between the system users and the APs, in particular the directions of arrival (DoAs) of the radio signals at the APs [3]. Furthermore, it may be desirable to perform the DoA estimation at the CPU to reduce the computational burden on the APs or avoid quantization errors in the DOAs caused by transmission over fronthaul links.

There is a growing trend to endow communication systems with artificial intelligence (AI) by means of deep neural network (DNN)-based approaches. The chief reasons behind this tendency are twofold: i) DNN architectures have offered a remarkable level of accuracy in numerous applications such as image classification, data analyses, speech recognition, etc. [13], ii) a low-complexity processing is generally attainable in the prediction (or estimation) phase once the DNN has been properly trained.

There exist several works in the literature which address the problem of DoA estimation via DNN-based approaches for a general multi-source system [14]–[17]. In [14], the authors design a framework

Funding for this work was provided in part by an Alliance Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada, with industrial sponsor InterDigital Canada.

in which multiple parallel trained DNN classifiers are first used to determine subregions of angular spectrum over which DoAs of received signals lie. Once angular subregions are found, received signals are preprocessed and fed into DNNs, each trained exclusively for a particular subregion, to estimate received DoAs. In [15], the angular spectrum is estimated by means of a convolutional DNN-based approach. Both approaches require an extra post-processing stage, particularly an interpolation module, to procure accurate results, especially for off-grid DoAs. An DNN-based approach is adopted in [16] where multiple parallel DNNs are trained to estimate DoAs in a massive MIMO system. This approach requires parallel maximum likelihood modules to refine the coarse estimates obtained in the prediction phase. In [17], a large complex DNN is developed to estimate DoAs for a massive MIMO system, but this approach requires a long training time.

The problem of DoA estimation in a cell-free m-MIMO system is addressed in [3], where a discrete Fourier transform (DFT)-based approach, implemented in a distributed manner, is adopted. The method requires a double search whose complexity is proportional to the product of the number of array snapshots and the size of the grid search. Moreover, in context of cell-free m-MIMO, the classical subspace-based DoA estimation methods such as multiple signal classification (MUSIC) [18] and Root-MUSIC [19] impose high computational complexity and generally suffer from a limited array operational range [20], [21].

The limitations of the above approaches motivate us to investigate the problem of DoA estimation for a cell-free m-MIMO system operating over the EHF and THF bands. Our aim is to conceive a DNN-based DoA estimation method yielding high accuracy while keeping the processing complexity low. Moreover, the method must be robust against noises caused by low-resolution quantizers to make it implementable at the CPU. To this end, we herein take advantage of a special feature set which is obtained from the superdiagonal entries of the spatial correlation matrix. This choice of features makes it possible to employ a DNN with only a few low-dimensional layers, which considerably speeds up training while ensuring high level of accuracy in the DoA estimates.

Simulation results shown that the proposed DNN-based method can achieve a DoA estimation performance that exceeds or nearly matches that of conventional benchmark approaches (i.e., MUSIC, root-MUSIC and DFT-based method [3]), but with considerably reduced processing complexity. The trained DNN is also robust against quantization noise in the array snapshot data, which allows the centralized implementation of the proposed method at the CPU.

The rest of the paper is organized as follows: The system model is introduced in Section II. The proposed DNN-based method for DoA estimation in cell-free m-MIMO, is presented in Section III. The simulation methodology and results are demonstrated in Section IV. A conclusion is drawn in Section V. music

II. System Model

We consider a cell-free m-MIMO system comprised of APs connected to a common CPU via fronthaul links. The APs, equipped with multiple antennas, serve single-antenna users distributed across a geographical area. The system operates over EHF or THF bands such that the propagation channel is accurately modeled by an LoS path [5]–[10]. As in [3], [22], we assume that the DoAs between the APs and the users remain constant within a coherence time, which is long enough to allow the use of orthogonal pilots for the different users. This assumption is not restrictive since DoAs typically change over a time interval whose scale is similar to that of large-scale fading [3].

Due to the orthogonality of pilots and highly attenuated NLoS paths, we focus our investigation on the DoA estimation of the LoS between a particular pair of user and AP during the uplink training phase. Letting M denote the number of antennas at the AP, the tth sample of the preprocessed[1] signal vector at the antenna array output can be expressed as [3], [5]

$$\mathbf{y}[t] = \sqrt{\rho}\,\alpha[t]\mathbf{a}(\omega) + \mathbf{v}[t], \tag{1}$$

where t is the snapshot index, ρ is the average signal-to-noise ratio (SNR) at each individual antenna, $\alpha[t] \in \mathbb{C}$ is the small-scale fading modeled as a circular complex Gaussian random variable with zero-mean and unit variance, $\mathbf{a}(\omega) \in \mathbb{C}^M$ is the *spatial signature vector* induced on the antenna array, and $\mathbf{v}[t] \in \mathbb{C}^M$ is a spatially white circular complex Gaussian noise with zero-mean vector and an identity covariance matrix.

In this work, we consider for simplicity a uniform linear array (ULA); however, the extension of our approach to the case of a uniform rectangular array is possible. The spatial signature vector of the ULA can be expressed as

$$\mathbf{a}(\omega) = [1, e^{j\omega}, \ldots, e^{j(M-1)\omega}]^T, \tag{2}$$

where ω is the spatial frequency defined as

$$\omega = 2\pi\eta\sin(\theta), \tag{3}$$

η is the array element separation (AES) normalized to wavelength unit, and θ is the unknown DoA (expressed in degrees). The latter is assumed to lie in the interval $[-\theta^{\text{AOR}}, \theta^{\text{AOR}}]$ which is referred to as the array operational range (AOR).

[1]Preprocessing refers to baseband demodulation followed by projection on user's pilot at each antenna element.

In this paper, we focus on the problem of DoA estimation using a finite sequence of observed snapshots $\mathbf{y}[t]$ for $t = 1, 2, \dots T$, where T is the number of snapshots. Specifically, we investigate the design and training of a DNN engine that can accurately estimate the unknown DoA for the above cell-free m-MIMO system model. The trained DNN must exhibit low processing complexity and a robust performance against quantization noise to facilitate centralized implementation.

III. Proposed Neural Network-Based Method for DoA Estimation

In this section, we first propose a new feature set for DNN-based DoA estimation in cell-free m-MIMO. We then introduce the architecture of the employed DNN. Finally, we describe the training and online estimation phases of the DNN.

A. Feature Extraction for DoA Estimation

Here, we propose a new way of feature extraction, which stipulates what type of data must be fed to the DNN used for DOA estimation. To obtain the proposed features, we first estimate the spatial correlation matrix of the array as

$$\widehat{\mathbf{R}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{y}[t]\mathbf{y}[t]^H \in \mathbb{C}^{M\times M}. \tag{4}$$

Then, for $1 \le m \le M-1$, let $\mathcal{S}_m$ denote the set of entries on the mth superdiagonal of $\widehat{\mathbf{R}}$, that is,

$$\mathcal{S}_m = \{\hat{r}_{i,i+m} | 1 \le i \le M-m\}. \tag{5}$$

Next, to each element of $\mathcal{S}_m$, we can associate an angle $\hat{\omega}_{i,i+m} = \arctan2(\hat{r}_{i,i+m})$, interpreted as spatial frequency, and form a corresponding set

$$\mathcal{F}_m = \{\hat{\omega}_{i,i+m} | 1 \le i \le M-m\}. \tag{6}$$

Different from other works in the literature (e.g., [14], [15]) which use $\widehat{\mathbf{R}}$ or $\{\mathcal{S}_m\}_{m=1}^{M}$, we propose to select $\mathcal{F}_1$ as the feature set. The rationale behind this selection will be explained in the sequel.

For the LoS signal model in (1)-(2), the spatial correlation matrix at the array output is given by

$$\mathbf{R} = \mathbb{E}\{\mathbf{y}[t]\mathbf{y}[t]^H\} = \rho\, \mathbf{a}(\omega)\mathbf{a}(\omega)^H + \mathbf{I}_M, \tag{7}$$

where the entries on the first superdiagonal takes the special form

$$r_{i,i+1} = \rho \exp(j\omega), \;\; 1 \le i \le M-1, \tag{8}$$

and therefore, $\arctan2(r_{i,i+1}) = \omega$ for al i in the above range. Hence, it is clear that the set $\mathcal{F}_1$ contains sufficient information for the estimation of the spatial frequency ω when the noise is averaged out, which in light of (4) amounts to using a sufficiently large value of T. Furthermore, as will be demonstrated in Section IV, the feature set $\mathcal{F}_1$ makes it possible to employ a DNN with only a few low-dimensional layers, which considerably speeds up the training and processing processes. The selection of $\mathcal{F}_1$ also allows to decrease the computational complexity of the feature extraction from an order of $\mathcal{O}(TM^2)$ to $\mathcal{O}(TM)$ since not all entries of $\widehat{\mathbf{R}}$ are required.

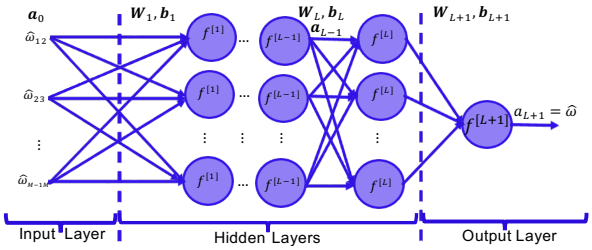


Fig. 1: General Architecture of a DNN

B. Neural Network Architecture

Fig. 1 depicts the general architecture of a DNN used in this work for scalar parameter estimation. The DNN is composed of one input layer, L hidden layers and one output layer, respectively indexed by $\ell = 0$, $1 \le \ell \le L$ and $\ell = L+1$. Layer 0 consists of input data vector $\mathbf{a}_0 \in \mathbb{R}^{q_0}$, while for $1 \le \ell \le L+1$, layer ℓ is comprised of q_ℓ neurons, with $q_{L+1} = 1$. Specifically, for $\ell \ge 1$, layer ℓ takes $\mathbf{a}_{\ell-1} \in \mathbb{R}^{q_{\ell-1}}$ as its input and outputs the vector $\mathbf{a}_\ell \in \mathbb{R}^{q_\ell}$ as

$$\mathbf{a}_\ell = f_\ell(\, \mathbf{W}_\ell \mathbf{a}_{\ell-1} + \mathbf{b}_\ell\,), \tag{9}$$

where $\mathbf{W}_\ell \in \mathbb{R}^{q_\ell \times q_{\ell-1}}$ is the weighting matrix, $\mathbf{b}_\ell \in \mathbb{R}^{q_\ell}$ is the bias vector, and $f_\ell(\cdot)$, which acts in a component-wise manner, is the non-linear activation function. In the sequel, we let $a_{L+1} = \mathbf{a}_{L+1}$ denote the scalar output of the DNN.

In our proposed method, we use the angles of the superdiagonal of $\widehat{\mathbf{R}}$ as input to the DNN. That is, the elements of $\mathcal{F}_1$ are vectorized into $\mathbf{a}_0 = [\hat{\omega}_{1,2}, \dots, \hat{\omega}_{M-1,M}]^T \in \mathbb{R}^{M-1}$. Since the variable $\hat{\omega}_{i,i+1}$ are mutually independent for large T, the vectorization can be done either in-order or out-of-order. For each such input vector, the trained DNN produces a spatial frequency estimate at its output, i.e. $a_{L+1} = \hat{\omega}$.

C. Training Phase

We assume a dataset consisting of N_s pairs of ground-truth and training samples, indexed by superscript (j) where $1 \le j \le N_s$. The ground-truth samples are spatial frequencies corresponding to selected DoAs while the training samples are the superdiagonal entries in the set $\mathcal{F}_1$. Specifically, for a fixed value of θ^{AOR} and the jth pair of samples, a DoA is chosen from a discretized AOR, i.e. $\theta^{(j)} \in [-\theta^{\mathrm{AOR}}, \theta^{\mathrm{AOR}}]$, and the corresponding spatial frequency $\omega^{(j)}$ calculated from (3) is used as the ground truth. For the corresponding

training sample, we generate T received signal vector $\mathbf{y}^{(j)}[t]$'s by using $\omega^{(j)}$ (see (1)). Then, we compute the spatial correlation matrix $\widehat{\mathbf{R}}^{(j)}$ from the sequence $\mathbf{y}^{(j)}[t]$, $t = 1, 2, \dots T$ (see (4)). Next, we form the set $\mathcal{F}_1^{(j)}$ and use it as the jth training sample (see (6)). Finally, we collect all ground-truth and training samples into a dataset

$$\mathcal{D} = \left\{ \left(\mathcal{F}_1^{(j)}; \omega^{(j)} \right) \right\}_{j=1}^{N_\mathrm{s}}. \tag{10}$$

For training, the elements of dataset (10) are first divided evenly into B subsets or *batches* of size N_s/B, assumed to be integer. Then, each training sample of a batch is fed as input to the DNN in Fig. 1 (forward propagation) to produce a corresponding output a_{L+1}. After one batch is passed through the DNN, the values of $\{\mathbf{W}_\ell, \mathbf{b}_\ell\}_{\ell=1}^{L+1}$ are refined (back propagation) by applying steepest descent to the mean square error (MSE) loss function

$$\mathrm{MSE}_\omega = \frac{1}{N_\mathrm{s}/B} \sum_j (\omega^{(j)} - a_{L+1}^{(j)})^2, \tag{11}$$

where the summation is taken over the index range of the batch. Once all the B batches have been processed and the DNN sees the whole training dataset, referred to as one *epoch*, the process is repeated.

D. Online Estimation Phase

Once the DNN is trained, it is ready to be used as a DoA estimator for received signals in a cell-free m-MIMO system. Specifically, the correlation matrix $\widehat{\mathbf{R}}$ is first estimated by using T snapshots of a received signal $\mathbf{y}[t]$ and $\mathcal{F}_1$ is then calculated according to (6). Next, the DNN is fed with entries of $\mathcal{F}_1$ to output the estimation of the spatial frequency as $\hat{\omega}$. Finally, the estimation of DoA is obtained as

$$\hat{\theta} = \arcsin\left(\hat{\omega}/2\pi\eta\right). \tag{12}$$

IV. Simulation Study

In this section, we first present our methodology for training the proposed DNN and evaluating its performance. We then provide justification for choosing the feature set $\mathcal{F}_1$ and investigate the impact of AES and AOR upon the training behavior based on numerical analysis. Finally, we present simulation results to evaluate the performance of the proposed DNN-based DoA estimation method.

A. Methodology

We consider the signal model in (1)-(3) along with the feature extraction procedure described in Subsection III-A to generate data for training and performance evaluation of the proposed DNN-based and other benchmark DoA estimation methods. The default values of the system model parameters are summarized in Table I. For the proposed DNN in Subsection III-B, we choose 2 hidden layers, where the first and second hidden layers have 8 and 4 neurons, respectively. The rectified linear unit (ReLU) and the identity functions are employed as activation functions for hidden and output layers, respectively. The DNN is built by using the libraries `TensorFlow` and `Keras` in a Functional API fashion [23].

TABLE I: System parameters

Parameter	Symbol	Range / Value(s)
Number of antennas	M	$\{4, 8, 16\}$
Number of snapshots	T	$[10, 50]$
Endpoint of the AOR	θ^{AOR}	$\{60^\circ, 90^\circ\}$
AES	η	$\{0.33, 0.50, 0.67\}$
SNR	ρ	$[0, 20]$dB

TABLE II: Training parameters and specifications

Parameter	Value / Type
Number of batches	64
Learning rate	10^{-6}
Number of epochs	1000
Training–validation split	80%–20%
Weights initializer	He/Glorot uniform
Optimizer	Adam

To train the DNN, we generate data by following the general procedure described in Subsection III-C. Specifically, the AOR $[-\theta^{\mathrm{AOR}}, \theta^{\mathrm{AOR}}]$ is discretized in steps of 1° and each DoA is selected R times. The value of R is chosen as 100 unless otherwise indicated. Hence, the number of samples in the dataset is equal to $N_\mathrm{s} = R(2\theta^{\mathrm{AOR}} + 1)$, while $B = 64$ is set as the number of batches[2]. For each $\theta^{(j)}$, $(j = 1, \dots, N_s)$, $T = 50$ independent array snapshot vectors are generated based on (1) with SNR set to $\rho = 10$ dB. These vectors are used to calculate a sample covariance matrix $\hat{\mathbf{R}}^{(j)}$ from which $\mathcal{F}_1^{(j)}$ is obtained. For purpose of comparison (as explained below), we also generate the larger dataset

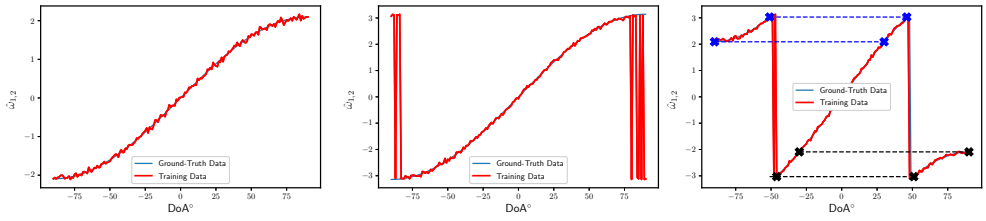
$$\mathcal{D}' = \left\{ (\mathcal{S}_1^{(j)}, \mathcal{S}_2^{(j)}, \dots, \mathcal{S}_{M-1}^{(j)}; \omega^{(j)}) \right\}_{j=1}^{N_\mathrm{s}}, \tag{13}$$

whose training samples are obtained by considering all superdiagonal entries of $\widehat{\mathbf{R}}^{(j)}$. Unless otherwise specified, the DNN is trained by using the parameters and specifications given in Table II. In particular, the Adam algorithm is used as optimizer with He-Glorot uniform weight initialization.

We evaluate the performance of the proposed DNN-based DOA estimation method against that of the following methods:

- Subspace-based MUSIC algorithm [18];
- Subspace-based Root-MUSIC algorithm [19];
- DFT-based approach from [3].

[2]The size of each batch is $\lfloor N_s/64 \rfloor$, except for the last batch with size $N_\mathrm{s} \bmod 64$.



(a) $\eta = 0.33$ (b) $\eta = 0.5$ (c) $\eta = 0.67$

Fig. 2: $\hat{\omega}_{1,2}$ versus DoA for different values of η ($\theta^{\text{AOR}} = 90°$, $T = 50$, $M = 16$, and $\rho = 10$dB).

TABLE III: Training and validation RMSE$_\omega$ of the DNN trained with datasets $\mathcal{D}$ and $\mathcal{D}'$

	Training RMSE$_\omega$		Validation RMSE$_\omega$		Training Time (min.)	
M	$\mathcal{D}$	$\mathcal{D}'$	$\mathcal{D}$	$\mathcal{D}'$	$\mathcal{D}$	$\mathcal{D}'$
4	0.164	0.525	0.169	1.590	14.23	14.32
8	0.112	0.747	0.166	2.154	15.03	16.10
16	0.091	0.717	0.129	2.243	14.75	17.60

As it is well known, MUSIC (with fine grid search) and root-MUSIC achieve a RMSE performance comparable to the Cramer-Rao lower bound [18], [19]. Hence they are used here as fundamental benchmarks against which the performance of the proposed DNN-based method can be assessed. The performance metrics for the trained DNN include the root mean square error on the DOA estimates, denoted as RMSE$_\theta$ and the processing time. The number of runs for the evaluation of RMSE$_\theta$ is set to 5000. Simulations are conducted on a desktop computer with an Intel®core™-i7 CPU (4×3.6-GHz cores).

B. Justification of Feature Set

We first provide justification for our choice of $\mathcal{F}_1$ as input feature set to the DNN. To this end, we conduct an experiment in which the considered DNN is trained separately with the proposed set $\mathcal{D}$ (10) and with the extended data set $\mathcal{D}'$ (13). In the latter case, the DNN is fed with the real and imaginary parts of the training samples in $\mathcal{D}'$. The dataset $\mathcal{D}'$ has M times more (real) elements per training sample than $\mathcal{D}$ and it is commonly used in the literature [14], [15]. The training and validation RMSE$_\omega$ of the DNN for the two different data sets, along with the training time, are listed in Table III for different values of M.[3] It can be seen that the DNN trained with the proposed data set $\mathcal{D}$, as opposed to $\mathcal{D}'$, always yields the lowest average training and validation RMSE$_\omega$. Besides, as M grows to 16, the average training time with the extended data set $\mathcal{D}'$ is increased by almost 20%.

[3]The values of RMSE$_\omega$ are obtained by taking an average over 100 different initial values of the weighting matrices $\mathbf{W}_\ell$ and bias vectors $\mathbf{b}_\ell$. In each case, the epoch whose validation RMSE$_\omega$ is minimum is selected for the averaging.

C. Impact of AOR and AES on Training

Next, we investigate the impact of the AOR and AES upon the training behavior. Indeed, we have observed that the training performance degrades as the DoA θ and/or sensor separation η increase beyond a certain limit. To get more insight, we plot a single realization of the spatial frequency estimate $\hat{\omega}_{1,2}$ in (6) versus the true DoA $\theta \in [-90°, 90°]$ for different values of η in Fig. 2. For small values of η (Fig. 2a), the training feature $\hat{\omega}_{1,2}$ generally increases with the DoA, except for the small fluctuations due to the measurement noise in matrix $\widehat{\mathbf{R}}$. For larger values of η however (Fig 2b and 2c), we note the presence of phase jumps in the training feature as $|\theta|$ approaches the limiting value $\theta_{\text{lim}} = \arcsin(\frac{1}{2\eta})$. This effect is due to the measurement noise which changes the sign of $\Im\{\hat{r}_{1,2}\}$ when $\Im\{\hat{r}_{1,2}\} \approx 0$ and $\Re\{\hat{r}_{1,2}\} < 0$;[4] in turn, this introduces a factor of $\pm\pi$ when computing $\hat{\omega}_{1,2} = \arctan\left(\frac{\Im\{\hat{r}_{1,2}\}}{\Re\{\hat{r}_{1,2}\}}\right)$. These phase jumps negatively impact the DNN training, and hence, its online DoA estimation performance.

To cope with this issue, one might attempt to unwrap the spatial frequencies and feed the DNN with the new resulting dataset. This approach, which is possible here because of the particular scheme used to generate the training data, would however introduce uncertainties in the feature extraction at the prediction phase. Indeed, since some values of $\hat{\omega}_{1,2}$ are repeated over the different DoA intervals (as shown by crosses in Fig. 2c), it is not feasible in practice to perform phase unwrapping before passing spatial frequencies to the trained DNN.

As explained above, the useful range for the AOR and AES, i.e. over which no phase jumps are likely to occur, are intimately related. In particular, the value of θ^{AOR} should be less than θ_{lim}, as can be seen from Figs. 2b and 2c. The existence of such phase jumps further explains why only the first superdiagonal is chosen and other superdiagonals are ignored for the feature set. Specifically, the entries of $\mathcal{F}_m$ are rough estimations for the mth multiple of ω. Hence, when $m \geq 2$, the range of $|\hat{\omega}_{i,i+m}|$ is m times larger than that of

[4]$\Re\{z\}$ and $\Im\{z\}$ denote the real and imaginary parts of the complex random variable z, respectively.

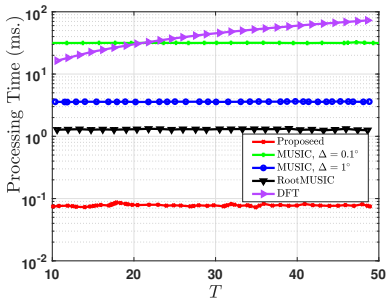


Fig. 3: Average processing complexity of different methods versus number of snapshots T ($\theta = 60°$, $M = 16$, $\eta = 0.5$, and $\rho = 10$dB).

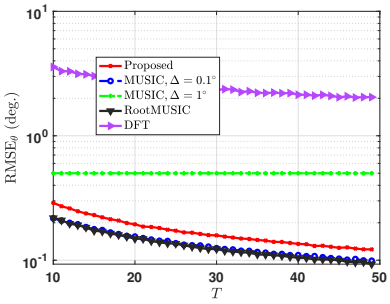


Fig. 4: Average RMSE$_\theta$ of different methods versus number of snapshots T ($M = 8$, $\theta = 15.5°$, $\eta = 0.5$, and $\rho = 10$dB).

$|\hat{\omega}_{i,i+1}|$, which in turn leads to the onset of phase jumps at a smaller value of $|\theta_{\text{lim}}|$, i.e., $\arcsin(\frac{1}{2m\eta})$. Based on the above discussion, we choose $\theta^{\text{AOR}} = 60°$ and $\eta = 0.5$ for DNN training in the rest of the paper; these values are similar to those used in related works, e.g., [14], [15].

D. DoA Estimation Results

Herein, we compare the performance of the proposed DNN-based DoA estimation method with that of the DFT-based method recently proposed in [3] for cell-free m-MIMO systems, and the subspace-based MUSIC and root-MUSIC algorithms. Moreover, the performance of the proposed method in the presence of quantization noise caused by low-resolution quantizers at the APs, will be compared for the centralized and decentralized implementations.

First, we compare the processing complexity of different methods in Fig. 3, where two step sizes, i.e., $\Delta = 1°$ and $0.1°$ are considered for the grid search of MUSIC. We first note that the processing complexity of the proposed DNN-based method is remarkably less than that of all other methods. Specifically, the processing time is about 0.07ms for the former, 4ms and 30ms for MUSIC with $\Delta = 1°, 0.1°$, respectively, 1.3ms for Root-MUSIC, and 20 to 70ms for the DFT-based method. This observation confirms that the low-complexity processing is the dominant feature of the proposed method. In Table IV, we also compare the total number of model parameters of the proposed method to that of other recent DNN-based methods for DoA estimation. It can be seen that compared to the methods in [14] and [15], the proposed method requires much fewer parameters.

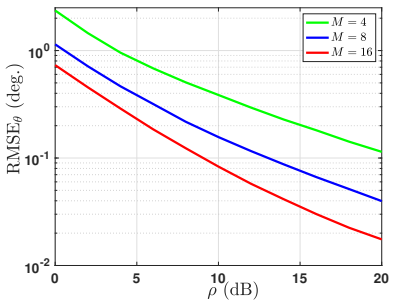


Fig. 5: Average RMSE$_\theta$ of proposed DNN-based method versus SNR ($\theta = 15.5°$, $\eta = 0.5$, and $T = 30$).

TABLE IV: Total number of DNN model parameters for different methods

Method	Total no. of hidden layers	Total no. of parameters
Proposed	2	105
[15]	4	1801
[14]	13	25482

Second, we compare the RMSE$_\theta$ performance of the DoA estimation methods. To this end, Fig. 4 depicts the RMSE$_\theta$ of different approaches versus the number of snapshots T for $\theta = 15.5°$, which is not part of the training dataset. From the figure, the DFT-based method does not perform well compared to other methods while MUSIC with $\Delta = 1°$ always estimates an angle next to the true DoA value. The proposed DNN-based method outperforms the MUSIC algorithm with $\Delta = 1°$ for different number of snapshots. MUSIC with $\Delta = 0.1°$ and Root-MUSIC slightly outperform the proposed DNN-based method (up to $0.06°$ in terms of RMSE$_\theta$), but this is achieved by incurring significant processing costs (see Fig. 3).

Third, we present the RMSE$_\theta$ of the proposed DoA estimation method versus the SNR ρ in Fig. 5. Note that the proposed DNN has only seen the value of $\rho = 10$dB during the training phase. As observed from the figure, the RMSE$_\theta$ decreases in nearly linear fashion as ρ increases, suggesting that the performance of the proposed DNN is robust against unseen values of SNR.

Finally, we compare the performance of the proposed method when it is implemented in centralized and decentralized manners for DoA estimation in a cell-free m-MIMO system. Here, following [24], we assume that a low-resolution uniform quantizer with Q_L levels is applied to the received signal $\mathbf{y}[t]$ in (1) and the estimated DoA $\hat{\theta}$ in (12) for the central-

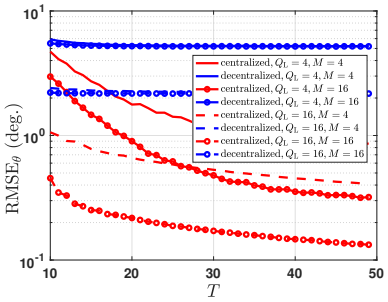


Fig. 6: Average RMSE_θ versus number of snapshots T for centralized and decentralized implementations ($\eta = 0.5$ and $\rho = 10$dB).

ized and decentralized implementations, respectively. Fig. 6 displays the RMSE_θ performance of the proposed method for the centralized and decentralized implementations, where the RMSE is averaged over the range $[-60°, 60°]$. As seen from the figure, the centralized implementation outperforms considerably the decentralized one for different number of antennas and $Q_\mathrm{L} = 4$ and 16. This observation suggests that the centralized implementation of the proposed method is more efficient when low-resolution quantizers are available.

V. Conclusion

We studied the problem of DoA estimation for cell-free m-MIMO systems operating over EHF and THF bands and proposed a low-complexity DNN-based method. To train the DNN, a special feature set was prescribed which makes it possible to employ a DNN with only a few low-dimensional layers. It was shown that the trained DNN is robust against quantization noise which makes the centralized implementation of the proposed method feasible. Through extensive simulations, the new method was shown to achieve an estimation performance that nearly matches or exceeds that of bechmark methods, but with considerably reduced complexity.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[2] S. S. Hosseini, B. Champagne, and X.-W. Chang, "A green downlink power allocation scheme for cell-free massive MIMO systems," *IEEE Access*, vol. 9, pp. 6498–6512, Dec. 2020.

[3] A. Abdallah and M. M. Mansour, "Efficient angle-domain processing for FDD-based cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2188–2203, 2020.

[4] C. Han, A. O. Bicen, and I. F. Akyildiz, "Multi-ray channel modeling and wideband characterization for wireless communications in the terahertz band," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2402–2412, May 2015.

[5] X. Gao, L. Dai, Y. Zhang, T. Xie, X. Dai, and Z. Wang, "Fast channel tracking for terahertz beamspace massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5689–5696, July 2017.

[6] P. Liu, M. D. Renzo, and A. Springer, "Line-of-sight spatial modulation for indoor mmwave communication at 60 GHz," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7373–7389, Nov. 2016.

[7] N. S. Perovic, P. Liu, M. D. Renzo, and A. Springer, "Receive spatial modulation for LOS mmwave communications based on TX beamforming," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 921–924, Dec. 2017.

[8] L. Zhu and J. Zhu, "Optimal design of uniform circular antenna array in mmwave LOS MIMO channel," *IEEE Access*, vol. 6, pp. 61 022–61 029, Sept. 2018.

[9] D.-W. Yue and H. H. Nguyen, "On the multiplexing capability of mmwave doubly-massive MIMO systems in LOS environments," *IEEE Access*, vol. 7, pp. 126 973–126 984, Sept. 2019.

[10] N. Kaur, S. S. Hosseini, and B. Champagne, "Enhanced channel tracking in THz beamspace massive MIMO: A deep CNN approach," in *Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*. Auckland, New Zealand, 2020, pp. 76–81.

[11] Z. Shen, K. Xu, and X. Xia, "2D fingerprinting-based localization for mmwave cell-free massive MIMO systems," *IEEE Commun. Lett.*, vol. 25, no. 11, pp. 3556–3560, Nov. 2021.

[12] D. Umut and A. Alkhateeb, "Enabling cell-free massive MIMO systems with wireless millimeter wave fronthaul," *arXiv preprint arXiv:2110.01798*, pp. 1–30, 2021.

[13] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*. Cambridge Univ. Press, Cambridge, U.K., 2014.

[14] Z.-M. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7315–7327, Dec. 2018.

[15] L. Wu, Z.-M. Liu, and Z.-T. Huang, "Deep convolution network for direction of arrival estimation with sparse prior," *IEEE Signal Process. Lett.*, vol. 26, no. 11, pp. 1688–1692, Nov. 2019.

[16] D. Hu, Y. Zhang, L. He, and J. Wu, "Low-complexity deep-learning-based DOA estimation for hybrid massive MIMO systems with uniform circular arrays," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 83–86, Jan. 2020.

[17] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sept. 2018.

[18] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Dept. of Elect. Eng., Stanford Univ., Stanford, CA, USA, 1982.

[19] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. 83th IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, vol. 8. Boston, MA, USA, 1983, pp. 336–339.

[20] O. Bialer, N. Garnett, and T. Tirer, "Performance advantages of deep neural networks for angle of arrival estimation," in *Proc. 83th IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*. Brighton, UK, 2019, pp. 3907–3911.

[21] Z. Dai, Y. He, V. Tran, N. Trigoni, and A. Markham, "DeepAoANet: Learning angle of arrival from software defined radios with deep neural networks," *IEEE Access*, vol. 10, pp. 3164–3176, 2022.

[22] T. M. Hoang, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and A. Marshall, "Cell-free massive MIMO networks: Optimal power control against active eavesdropping," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4724–4737, Oct. 2018.

[23] The Functional API. May 2022. [Online]. Available: https://www.tensorflow.org/guide/keras/functional

[24] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, "Cell-free massive MIMO with limited backhaul," in *IEEE Int. Conf. Commun. (ICC)*. Kansas City, MO, USA, May 2018, pp. 1–7.