Deep Neural Networks for Direction of Arrival Estimation of Multiple Targets with Sparse Prior for Line-of-Sight Scenarios

Saiqin Xu, *Student Member, IEEE,* Alessandro Brighente, *Member, IEEE*
Baixiao Chen, Mauro Conti *Fellow, IEEE,* Xiancheng Cheng and Dongchen Zhu

***Abstract*—Received signal Direction of Arrival (DOA) estimation represents a significant problem with multiple applications, ranging from wireless communications to radars. This problem presents significant challenges, mainly given by a large number of closely located transmitters being difficultly separable. Currently available state of the art approaches fail in providing sufficient resolution to separate and recognize the DOA of closely located transmitters, unless using a large number of antennas and hence increasing the deployment and operation costs.**

In this paper, we present a deep learning framework for DOA estimation under Line-of-Sight scenarios, which able to distinguish a number of closely located sources higher than the number of receivers' antennas. We first propose a formulation that maps the received signal to a higher dimensional space that allows for better identification of signal sources. Secondly, we introduce a Deep Neural Network that learns the mapping from the receiver antenna space to the extended space to avoid relying on specific receiver antenna array structures. Thanks to our approach, we reduce the hardware complexity compared to state of the art solutions and allow reconfigurability of the receiver channels. Via extensive numerical simulations, we demonstrate the superiority of our proposed method compared to state-of-the-art deep learning-based DOA estimation methods, especially in demanding scenarios with low Signal-to-Noise Ratio and limited number of snapshots.

***Index Terms*—DOA estimation, deep neural network, sparse representation, multiple targets.**

I. Introduction

DIRECTION of Arrival (DOA) estimation has been advocated as an attractive field of research with many applications including wireless communications, astronomical observation, radar, and sonar [1]–[4]. In general, DOA estimation addresses the problem of locating transmitting sources by looking at the signal received by an array of antennas whose spatial positions are known [5]. The main challenge in DOA estimation is the difficulty of jointly developing approaches with a minimal hardware complexity [6] in terms of receiver costs and power consumption, while providing a desired level of estimation precision and robustness in the presence of multiple sources or multiple paths [7].

Saiqin Xu, Baixiao Chen and Xiancheng Cheng are with National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China (e-mail: xusaiqin@stu.xidian.edu.cn;bxchen@xidian.edu.cn; xcchengxd@gmail.com;dongchen_zhu@163.com;)

Alessandro Brighente and Mauro Conti are with the Department of Mathematics and HIT Research Center, University of Padova, 35131 Padova, Italy (e-mail: alessandro.brighente@unipd.it;conti@math.unipd.it;).

Manuscript received X X, 2022; revised X X, 2022. This work was supported by the National Natural Science Foundation of China (No. 61571344).

Traditionally, DOA estimation algorithms focus on ideal antenna arrays whose models often significantly vary from realistic scenarios [8]. These approaches, referred to as model-driven, first formulate forward parametric models that define the transformation from signal directions to array outputs, and then estimate the directions by exploiting the properties of the employed array structures [9]–[11]. Recently, research on estimating solid angle using super resolution processing has largely focused on uniform array for which many efficient parameter estimators have been applied. The Uniform Linear Array (ULA) provides a simple regular array structure and well-analyzed signal processing techniques [12], [13]. However, under the assumption of stationary sources, it generally allows for the resolution of a number of sources smaller than the number of antennas [14]. Some eminent model-driven algorithms are applied to estimate DOA, such as beamformer methods [15], maximum likelihood methods [16]–[18], subspace-based methods [19], [20], and sparsity-inducing methods [21]–[23]. However, these methods assume that there are no antenna calibration errors not array imperfections, and that the number of available snapshots is large. Hence, some robust methods for DOA estimation in the presence of array errors have been proposed [24]–[28]. These methods require the knowledge of the statistics of the array model errors [24], [25] and calibration signals with known directions [26]–[28], but these may not be available in practice. Zhang et al. [29], proposed an alternative iterative method, which can simultaneously provide estimates of the DOA of signals and calibration of the gain-phase error of each sensor. However, it requires that at least two signals are spatially far separated from each other. Furthermore, the performance of model-driven methods heavily depend on the array's aperture due to the strict physical constraints. In addition, the distance between adjacent elements of the antenna array should be equal to or less than half the wavelength of the impinging planar wavefronts. Otherwise it leads to grating lobes in the spectrum which correspond to ambiguities in the array manifold [6]. At the same time, to achieve DOA estimation with a high resolution, the receiving arrays should have a relatively large aperture [30], [31]. This implies that, to obtain a high resolution, we require arrays with a large number of antennas which is not always feasible. In addition, having a Radio Frequency (RF) chain comes with high inefficiencies, as mixers, power amplifiers, and either digital-to-analog or analog-to-digital converters increase both the implementation

costs and the power consumption [32]–[34]. Therefore, it is fundamental to develop solutions that do not require a high number of antennas.

Generally speaking, hardware is more expensive and less scalable than software: if we need more receiver antennas to get better DOA estimation, this increases the production and deployment costs. Furthermore, the physical hardware is not flexible to adjust to the requirements. Therefore, to reduce the number of RF chains without a loss in the array aperture, data-driven methods are an effective solution to reduce costs and achieve better performance. In particular, we can extend a large aperture with small Degree-of-Freedom (DOF) applied Deep Learning (DL) to obtain better DOA estimation.

Since only a small number of channels need to be sampled and digitized, the hardware complexity remains comparably low while covering a large aperture, yielding to a higher flexibility than traditional fixed channels. These approaches fall in the domain of data-driven approaches, where solutions make no assumption on the underlying physical array structure but try to learn a mapping between different domains based on the observed data.

With the rapid improvement of Artificial Intelligence in various fields such as image, speech and expression recognition processing [35]–[38], researchers have proposed many DL methods to achieve DOA estimation efficiently [39]–[44]. These approaches directly learn the nonlinear relationship between the array outputs and sources positions thanks to their non-linearity, adaptive learning capability, and high generalization. Based on these premises, methods in [45]–[51] also make no pre-assumptions on the physics-driven models. In [45], the authors use a Support Vector Regression model to learn the mapping relationship among real and imaginary parts of array outputs and DOAs, whereas the authors in [46] and [47] develop a Long Short-Term Memory scheme and deep Convolutional Neural Network (CNN) respectively for super-resolution DOA estimation with phase enhancement learning. These DL-based methods have been confined to the case of ULA and/or resolve up to $M-1$ sources with an M-element array. Hence, some DOA estimation methods based on the inherent sparsity in the angular domain has raised considerable attention. Both in [48] and [49], the authors use classification models that divide the spatial space into many discrete subregions, and then use the classifier networks to estimate DOAs in each region. The aim of [48] and [49] is to integrate Deep Neural Network (DNN) group against unknown array imperfections and design CNN for multi-speaker DOA estimation. A Deep Convolution Network (DCN) that learns the inverse transformation from large training dataset can be used to recover the clear power spectrum of signals [50]. The authors first convert the DOA estimation problem into a sparse linear inverse problem by introducing a spatially overcomplete formulation. The overcomplete formulation is given as input to the DCN. On the other hand, the overcomplete known labels is used as output for the training process. The idea is that, after training, the DCN is able to reconstruct the clean overcomplete formulation starting from the noisy received signal. However, this approach suffers from leakage effects when the sources are off the grid, and can not effectively resolve the sources located in the same beam. To deal with the issue of grid mismatch, the authors of [51] propose two classification learning models integrating the angle separation to solve the problem of coherent DOA estimation. However, this algorithm is only applicable for two signals. State-of-the-art data-driven approaches cannot properly estimate the DOA from a number of sources that is larger than the number of antennas at the receiver array. This represents a fundamental limit, as the number of transmitters is likely to be larger than the number of antennas considering the ever-increasing number of users in beyond fifth-generation (5G) networks. Furthermore, the dependency on a specific array physical structure limits the applicability of the solution to the different cases.

In this paper, we propose a novel data-driven Line-of-Sight (LOS) DOA estimation algorithm achieving good estimation performance while reducing the deployment and operational costs. In contrast to existing methods, we consider the special case where all sources share the same normalized power spectrum [7], which is the case for many practical aerial targets in LOS scenarios. Compared to other state-of-the-art estimators that resort to the physical structure of the receiver array to perform DOA estimation, our data-driven method allows to remove the assumptions on the underlying array structure and to detect a number of sources larger than the number of receiver antennas. Our work is based on the intuition that, considering the general case where the number of sources exceeds the number M of antennas at the receiver array, the columns of the array steering matrix are linearly correlated, so the covariance matrix is full-ranked and the number of spanned subspaces is M. Therefore, by extending the covariance matrix over a larger subspace thanks to a sampling approach, we are able to separate the signals coming for the different sources and hence identify them. The significant contributions of our work can be outlined as follows.

1) We propose a sampling approach thanks to which we can increase the dimensionality of the covariance matrix to increase the distinguishability of the multiple sources. We refer to this matrix as the extended correlation matrix. Our approach obtains the same effect of an increased number of antennas, but without incurring in higher deployment and operational costs.
2) We propose a neural network structure able to map the received signals to the extended correlation matrix. Thanks to our approach, we reduce the number of antennas required to estimate DOAs and avoid relying on specific physical structures.
3) Via extensive simulation, we demonstrate the superiority of our proposed method in terms of DOA estimation accuracy and generalization capability versus the Signal-to-Noise Ratio (SNR) and number of snapshots, especially when considering closely located spatial sources.

The rest of the paper is organized as follows. In Section II, we introduce the signal model for DOA estimation and its sparse spatial spectrum representation same as for the references [50], [51]. Section III describes the DNN framework and interprets how it expends the virtual channels. Section IV

carries out simulations to demonstrate the predominance of the proposed method in solving multiple sources. Conclusions are provided in Section V.

Notations: We use lower-case (upper-case) bold characters to denote vectors (matrices). $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{\mathrm{T}}$ represent the Hermitian and transpose operators. Furthermore, $\mathbb{E}[\cdot]$ denotes the expectation operator, $\mathrm{vec}(\cdot)$ is the vectorization operator that turns a matrix into a vector by stacking all columns on top of the another. $\otimes$ and $\odot$ are used to represent the Kronecker product and Khatri-Rao product (column-wise Kronecker product). In particular, $\boldsymbol{I}_M$ denotes the $M \times M$ identity matrix, whereas $\mathbf{1}_M$ and $\mathbf{0}_M$ respectively denote the N-element column vectors of all ones and all zeros. Additionally, $j = \sqrt{-1}$ represents the imaginary unit, and $\|\cdot\|$ denotes l_2 norm.

II. Problem Formulation

In this section, we describe the general problem. In particular, in Section II-A, we outline the problem statement and the general signal model. In Section II-B, we take the spatial sparsity of the incident signals into consideration, and present the sparse signal model.

A. General Signal Model

We consider a ULA composed by M isotropic elements with inter-element spacing d, whose antenna positions are given by $d_1, d_2, \cdots, d_M$. The ULA receives L far-field uncorrelated narrowband LOS signals $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_L(t)]^{\mathrm{T}}$ coming from the L distinct directions $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]^{\mathrm{T}}$. As common in the literature on DOA estimation [7], [23], [51], we assume that all transmitters are located at the same distance from the receiver. This allows us to avoid having a per-transmitter SNR value, but a single value for all sources. The baseband data vector $\boldsymbol{y}(t)$ received at time instant t with a total of T snapshots can be modeled as

$$\boldsymbol{y}(t) = \sum_{l=1}^{L} \boldsymbol{a}(\theta_l) s_l(t) + \boldsymbol{n}(t), \ t = 1, 2, \cdots, T, \tag{1}$$

where $t \in \mathbb{N}^+$ denotes the index of the segment at time instant t. We assume that the elements of the noise vector $\boldsymbol{n}(t) \sim \mathcal{CN}(\mathbf{0}_M, \sigma_n^2 \boldsymbol{I}_M)$ are independent and identically distributed (i.i.d.) complex Additive White Gaussian Noise (AWGN), and that they are uncorrelated from the impinging sources. We denote as $\boldsymbol{a}(\theta_l) \in \mathbb{C}^{(M,1)}$ the array steering vector corresponding to angle θ_l, which can be written as

$$\boldsymbol{a}(\theta) = \left[1, e^{-j\frac{2\pi}{\lambda} d sin(\theta)}, \cdots, e^{-j\frac{2\pi}{\lambda}(M-1) d sin(\theta)}\right]^{\mathrm{T}}, \tag{2}$$

where $\lambda = v/f$ is the carrier wavelength, and f and v are the frequency and propagation speed, respectively.

Using the matrix notation, the vector of the received signals $\boldsymbol{y}(t)$ can be expressed compactly as

$$\boldsymbol{y}(t) = \boldsymbol{A}(\theta)\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{3}$$

where $\boldsymbol{A}(\theta) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_L)] \in \mathbb{C}^{M \times L}$ is the array steering matrix.

Our objective is to estimate the number and DOAs $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]$ of the signals impinging on the receiver antenna array without making any prior assumption on the number of transmitters and without exploiting any particular antenna geometry. We particularly target the case where $L > M$, hence focusing on the scenarios where state-of-the-art solutions are unable to resolve a number of transmitters higher than the number of receiver antennas.

We notice that our model uniquely considers a LOS channel without accounting for possible signal reflections. However, focusing on the field of DOA estimation, we can apply our LOS method to aerial targets, and use the LOS scenario and AWGN channel model to design DOA estimation algorithms as common in the literature [7], [12]. We will focus on the non-LOS scenario in our future works, especially targeting radar systems on the ground or vehicle radar.

B. Signal Sparse Representation

Let us denote as $\boldsymbol{\Phi} = [\phi_1, \phi_2, \cdots, \phi_K]^{\mathrm{T}}$ the set of discrete directions sampled from the potential space of DOAs with constant sampling interval $\Delta\phi = \phi_{k+1} - \phi_k \, \forall k = 1, \cdots, K-1$. If K is large enough, the real DOAs are contained in set $\boldsymbol{\Phi}$ with moderately small quantization errors. Based on the directions in $\boldsymbol{\Phi}$, the output $\boldsymbol{y}(t)$ of the receive antenna array can be reformulated in an overcomplete form as [51]

$$\boldsymbol{y}(t) = \sum_{k=1}^{K} \boldsymbol{a}(\phi_k) \bar{s}_k(t) + \boldsymbol{n}(t), \ t = 1, 2, \cdots, T, \tag{4}$$

where $\bar{s}_k(t) = s_l(t)$ if $\phi_k = \theta_l$ (with small quantization error), otherwise, $\bar{s}_k(t) = 0$. The spatial correlation matrix $\boldsymbol{R}$ of the received signal can be computed as

$$\boldsymbol{R} = \mathbb{E}\left[\boldsymbol{y}(t)\boldsymbol{y}^{\mathrm{H}}(t)\right] = \sum_{k=1}^{K} \eta_k \boldsymbol{a}(\phi_k)\boldsymbol{a}^{\mathrm{H}}(\phi_k) + \sigma_n^2 \boldsymbol{I}_M, \tag{5}$$

where $\eta_k = \mathbb{E}[\bar{s}_k(t)\bar{s}_k^{\mathrm{H}}(t)]$ represents the source power on the kth direction ϕ_k. By denoting the spatial spectrum as $\boldsymbol{\eta} = [\eta_1, \eta_2, \cdots, \eta_K]^{\mathrm{T}}$, we notice it only has nonzero elements at the true signal DOAs, i.e., the spatial spectrum is sparse. Then the mth column of $\boldsymbol{R}$ can be reformulated as [51]

$$\begin{aligned} \boldsymbol{c}_m &= \sum_{k=1}^{K} \eta_k \boldsymbol{a}(\phi_k)\boldsymbol{a}^{\mathrm{H}}(\phi_k)\boldsymbol{e}_m + \sigma_n^2 \boldsymbol{I}_M \boldsymbol{e}_m \\ &= \left[\boldsymbol{a}(\phi_1)\boldsymbol{a}^{\mathrm{H}}(\phi_1)\boldsymbol{e}_m, \cdots, \boldsymbol{a}(\phi_K)\boldsymbol{a}^{\mathrm{H}}(\phi_K)\boldsymbol{e}_m\right]\boldsymbol{\eta} + \sigma_n^2 \boldsymbol{e}_m \\ &= \boldsymbol{A}_m \boldsymbol{\eta} + \sigma_n^2 \boldsymbol{e}_m, \end{aligned} \tag{6}$$

where $\boldsymbol{A}_m(:, k) = \boldsymbol{a}(\phi_k)\boldsymbol{a}^{\mathrm{H}}(\phi_k)\boldsymbol{e}_m$, and $\boldsymbol{e}_m$ is an $M \times 1$ all-zero column vector except for a 1 at the mth location. We denote the vectorized form of $\boldsymbol{R}$ as

$$\begin{aligned} \boldsymbol{c} = \mathrm{vec}(\boldsymbol{R}) &= \left[\boldsymbol{c}_1^{\mathrm{T}}, \boldsymbol{c}_2^{\mathrm{T}}, \cdots, \boldsymbol{c}_M^{\mathrm{T}}\right]^{\mathrm{T}} \\ &= [\boldsymbol{A}_1; \boldsymbol{A}_2; \cdots; \boldsymbol{A}_M]\boldsymbol{\eta} + \sigma_n^2 [\boldsymbol{e}_1; \boldsymbol{e}_2; \cdots; \boldsymbol{e}_M;] \\ &= \tilde{\boldsymbol{A}}\boldsymbol{\eta} + \sigma_n^2 \vec{\mathbf{1}}_n. \end{aligned} \tag{7}$$

where $\vec{\mathbf{1}}_n = \mathrm{vec}(\boldsymbol{I}_M) = [\boldsymbol{e}_1; \boldsymbol{e}_2; \cdots; \boldsymbol{e}_M]$, and $\tilde{\boldsymbol{A}}$ can be represented as

$$\tilde{\boldsymbol{A}} = \boldsymbol{A}^* \odot \boldsymbol{A} = [\boldsymbol{a}_1^* \otimes \boldsymbol{a}_1, \boldsymbol{a}_2^* \otimes \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K^* \otimes \boldsymbol{a}_K], \tag{8}$$

where with kth element being in the form

$$\boldsymbol{a}_k^* \otimes \boldsymbol{a}_k = \Big[e^{(j\frac{2\pi}{\lambda}(d_1-d_1)sin(\phi_k)}, e^{j\frac{2\pi}{\lambda}(d_2-d_1)sin(\phi_k)}, \cdots e^{j\frac{2\pi}{\lambda}(d_i-d_j)sin(\phi_k)}, \cdots, e^{j\frac{2\pi}{\lambda}(d_M-d_M)sin(\phi_k)}\Big]^T, \quad (9)$$
$$i = 1, \cdots, M, \quad j = 1, \cdots, M.$$

Comparing (3) and (6), the new vector $\boldsymbol{c}$ can be regarded as a single-snapshot received data vector corresponding to a single-snapshot source vector $\boldsymbol{\eta}$, where $\sigma_n^2\vec{\mathbf{1}}_n$ becomes a deterministic term. The distinct columns of $\tilde{\boldsymbol{A}}$ act as the virtual array manifold of an extended array aperture. Clearly, the purpose of sparse representation is to recover the continuous-valued DOAs $\{\theta_k\}_{k=1}^K$, i.e., the reconstructed sparse spatial spectrum $\boldsymbol{\eta}$ can be expressed as follows

$$\boldsymbol{\eta} \approx \tilde{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{c}. \quad (10)$$

As a result, the $\odot$ product can generate the so-called difference co-array (investigated in detail in [21], [22], [52]), which can increase the DOF of ULA to $2(M-1)$. This advantage allows us to estimate more than M signals with only M antennas.

Figures 1(a) and 1(b) show the reconstructed spatial spectrum (10) of $L = 3$ sources and $L = 11$ sources, respectively.

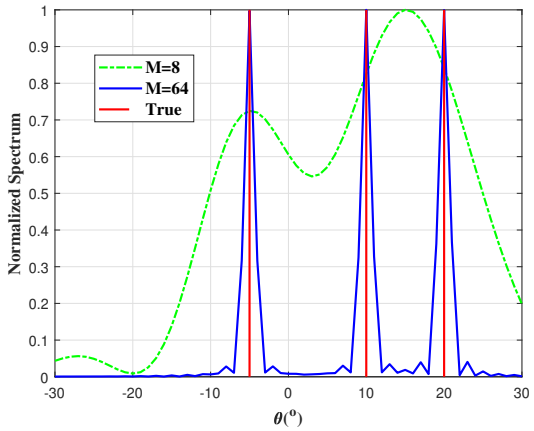
We observe that, when the adjacent angular signals get closer and L increases, the spatial spectrum obtained with $M = 8$ antennas has no peak along the true directions. Hence, the algorithm is not able to detect multiple closely spatial sources. On the other hand, increasing the number of antennas allows for improved distinguishability of the different sources, however coming at the price of increased hardware complexity, implementation costs, and power consumption. To overcome the limitations imposed by a reduced number of antenna elements and starting from $\boldsymbol{R}$, we use a regression learning model to recover an $\boldsymbol{\eta}$ extended over the space obtained with a number $M' > M$ of antennas. Notice that, this requires that the angle of signals are not less than a beamwidth [50]. Accordingly, we aim to resolve the multiple closely sources while improving the DOA estimation precision.

III. Learning Networks for DOA Estimation

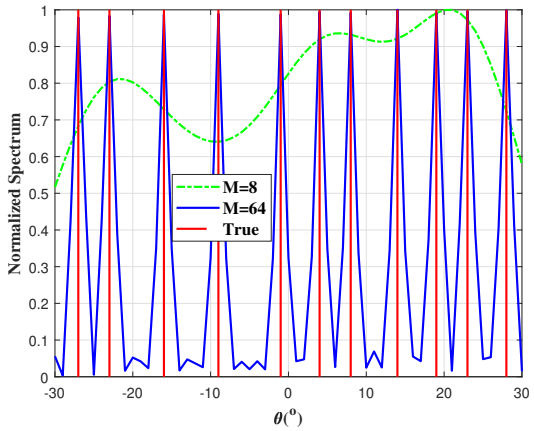
In this section, we present our method for super-resolution DOA estimation. We first present the intuition behind our approach in Section III-A. Then, we present our learning-based approach in Section III-B.

A. Intuition Behind Our Learning-Based Approach

The main purpose of our work is to allow for the estimation of a number of sources higher than the number of receiver antennas. To achieve our goal, it is fundamental to obtain a signal representation able to separate signal contributions from different transmitters. Thus, the non-linear mapping between the low dimension covariance matrix and its extended version in the an overcomplete form as in (4) allows us to project the phases obtained from the covariance matrix into a higher dimensional space. In this way, we are able to increase the resolution of the received signal and distinguish a number of sources higher than the number of receiver antennas. When the DOA are known, we can obtain the extended representation space (4). However, when these are not known, we do not know the space in which we need to project the received signal. Our DNN-based approach serves this purpose, i.e., learning the mapping between received signal and its extended representation.



(a) Spatial spectrum with $L = 3$.



(b) Spatial spectrum with $L = 11$.

Fig. 1. Reconstructed spatial spectrum of $M = 8$ and $M = 64$. (a) $L = 3$. (b) $L = 11$.

B. DNN Learning Framework

The formulation derived in the previous section allow us to formulate a novel methodology to estimate a number of DOA larger than the number of receiver antenna elements. The idea is to learn the mapping from the received signals spatial covariance matrix to the extended covariance matrix. Notice that we do not obtain the DOAs estimates out of the learning algorithm, as selecting the number of outputs constraints the number of sources we are able to estimate. Therefore, thanks to our approach, we do not need to make any assumption on the number of users. We propose a supervised DL approach, with the proposed DNN structure shown in Fig. 2. The supervised learning approach operates in two phases:

- Learning Phase: Starting from a set of known signals and their true DOA, we compute their correlation matrix and their reconstructed spatial spectrum in (10). These two matrices represent the input and the output of the DNN, respectively. In this way, the DNN can learn the inherent mapping between the received signal and their extended spatial spectrum.
- Testing Phase: We use a second set of known signals (testing set) to verify the effectiveness of the learning structure in estimating DOA of unknown locations.

The variability of the DNN input is significantly affected by the variability of signal waveforms. To limit it, we follow the guidelines of [45], [48], [51]. We compute the array covariance matrix and reformulate the off-diagonal upper right matrix elements as an input vector to the DNN [48] as

$$\hat{\boldsymbol{r}} = [\text{real}(\tilde{\boldsymbol{r}}), \text{imag}(\tilde{\boldsymbol{r}})], \tag{11}$$

with

$$\tilde{\boldsymbol{r}} = [\boldsymbol{R}_{1,2}, \boldsymbol{R}_{1,3}, \cdots, \boldsymbol{R}_{1,M}, \boldsymbol{R}_{2,3}, \cdots, \boldsymbol{R}_{2,M}, \cdots, \boldsymbol{R}_{M-1,M}], \tag{12}$$

where $\boldsymbol{R}_{i,j}$ represents the (i,j)th element of $\boldsymbol{R}$. Before feeding it to the DNN, $\tilde{\boldsymbol{r}}$ should be normalized to prevent neurons inactivation. The Gaussian normalized $\hat{\boldsymbol{r}}$ is given by

$$\boldsymbol{r} = \frac{\hat{\boldsymbol{r}} - \boldsymbol{\mu}_{\boldsymbol{r}}}{\boldsymbol{\sigma}_{\boldsymbol{r}}}, \tag{13}$$

where $(\boldsymbol{\mu}_{\boldsymbol{r}}, \boldsymbol{\sigma}_{\boldsymbol{r}})$ are the statistical mean and standard deviation of $\hat{\boldsymbol{r}}$, respectively. We denote the corresponding desired covariances matrix as $\boldsymbol{R}'$, and $\boldsymbol{r}'$ is the desired concatenated real/imaginary parts off-diagonal upper right covariance matrix. The output of the DNN is the extended semi-manufactured covariance matrix $\boldsymbol{r}'$. During the training stage, we are able to construct this matrix thanks to our knowledge on the number of sources and their exact location. During testing instead, we need to reconstruct DOA starting from the DNN output. Consequently, we reconstructed the covariance matrix $\boldsymbol{R}'$ according to the real/imaginary part in (11). Ultimately, the DOA estimates are obtained via amplitude interpolation within the most significant peaks of the reconstructed spectra by (9) and (10).

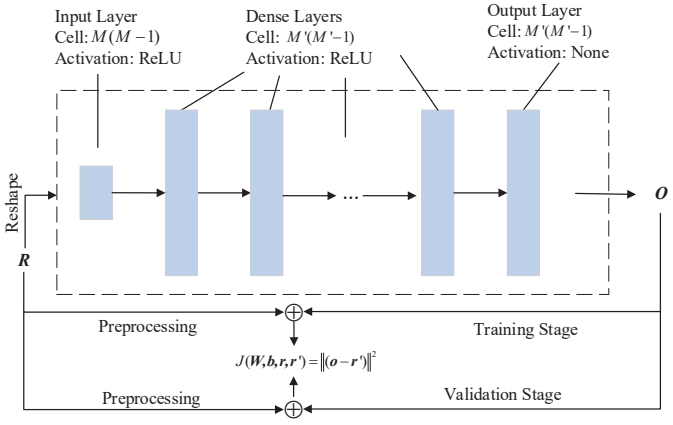


Fig. 2. Supervised learning framework of the DNN.

We first feed the DNN with the input vectors $\boldsymbol{r}$ to extract the principal components in the original input, and then enlarge it to the extended space via hidden layers, i.e., $(\boldsymbol{r}, \boldsymbol{r}')$ are the training sample set. The number of hidden layers and the number of neurons in each layer are chosen to reach the trade-off between the nonlinear capacity and the overfitting risk of the DNN. The outputs of the pth layer are formulated as follows

$$\boldsymbol{o}^p = \begin{cases} \sigma(\boldsymbol{z}^p) = \sigma(\boldsymbol{W}^p \times \boldsymbol{o}^{(p-1)} + \boldsymbol{b}^p),\ p = 1, 2, \cdots, P-1; \\ \boldsymbol{W}^p \times \boldsymbol{o}^{(p-1)} + \boldsymbol{b}^p,\ p = P; \end{cases} \tag{14}$$

with

$$\sigma(x) = \text{ReLU}(x) = \begin{cases} x, & x \geq 0; \\ 0, & x < 0. \end{cases} \tag{15}$$

where $\boldsymbol{W}^p$ and $\boldsymbol{b}^p$ represent the weight matrix and bias vector corresponding to the pth layer, P is the number of dense layers, and $\sigma(\cdot)$ is the nonlinear activation functions. We adopt the Rectified Linear Unit (ReLU) as nonlinear activation. To ensure fast convergence of the learning algorithm in the training phase, we use the least mean square error (MSE) between the output $\boldsymbol{o}^P$ and the desired $\boldsymbol{r}'$ as the loss function

$$J(\boldsymbol{W}, \boldsymbol{b}, \boldsymbol{r}, \boldsymbol{r}') = ||\boldsymbol{o}^P - \boldsymbol{r}'||^2, \tag{16}$$

Notice that MSE represents a suitable criterion for our learning objective. Indeed, we want to compute an output matrix that resembles as much as possible the target matrix, thus lowering the reconstruction error. The weight matrices and bias vectors are then updated based on the backpropagated gradients [53] of the loss function with respect to the variables

$$(\boldsymbol{W}, \boldsymbol{b}) = \min_{\boldsymbol{W}, \boldsymbol{b}} \frac{1}{D} \sum_{i=1}^{D} J(\boldsymbol{W}, \boldsymbol{b}, \boldsymbol{r}, \boldsymbol{r}'), \tag{17}$$

where D denotes the batch size. The gradients can be computed by mathematical derivations under the recursion relations. The weight and bias vectors are then updated using their gradients as follow

$$\begin{aligned} \boldsymbol{W}^p &= \boldsymbol{W}^p - \mu(\frac{\partial J(\boldsymbol{W}, \boldsymbol{b}, \boldsymbol{r}, \boldsymbol{r}')}{\partial \boldsymbol{W}^p}); \\ \boldsymbol{b}^p &= \boldsymbol{b}^p - \mu(\frac{\partial J(\boldsymbol{W}, \boldsymbol{b}, \boldsymbol{r}, \boldsymbol{r}')}{\partial \boldsymbol{b}^p}); \end{aligned} \tag{18}$$

where μ is the learning rate. Once trained, the network can be used to predict $\boldsymbol{r}'$ that corresponds to a new measurement vector $\boldsymbol{o}^P$ in near real time. An updating algorithm called adaptive moment estimation (Adam) is used to optimize the parameters of the neural networks [54].

IV. Numerical Results

In this section, we show via numerical results the superiority of our method over state-of-art machine learning-based DOA estimation methods [50], [51]. We first outline our methodology in Section IV-A. We then present results obtained in a simplified scenario in Section IV-B, and focus on root mean squared error over estimates in Sections IV-C and IV-D. We then show the effect of multiple sources in Section IV-E.

A. Methodology

We use machine learning framework Tensorflow v.2.7.0, and compute the gradients using its embedded tools. The DNN training is based on Python 3.7 and Adam optimizer with 1000 epochs. In order to prevent over-fitting, we use dropout with 0.95 ratio at each nonlinear layer. In the following simulations, we use a ULA with $M = 8$ physical antenna to receive signals, and apply an extended $M' = 64$ dimensional virtual channel to estimate DOAs. Moreover, we assume $d = 0.05m$, and $\lambda = 0.1m$.

To optimize our model, we investigate different number layers and batch size values. We generate data set with a fixed snapshot value of 200 and varying SNR from -10 dB to 20 dB. We collect 5000 samples on training stage and 1000 samples on testing stage. We set the input size to 56 as the length of sampled covariance matrix $\boldsymbol{R}$, and the output size to 4096 as the length of extended sampled matrix. TABLE I shows the MSE of the training loss for different number of layers and batch size. From results, we see that a higher number of layer or larger batch size is not a good approach. Indeed, having too many layers may cause the neural network to be over-representative with respect to the actual mapping between the covariance matrix and the extended space. This leads to worst performance.

TABLE I
LOSS (MSES) OF TRAINING PROCESS

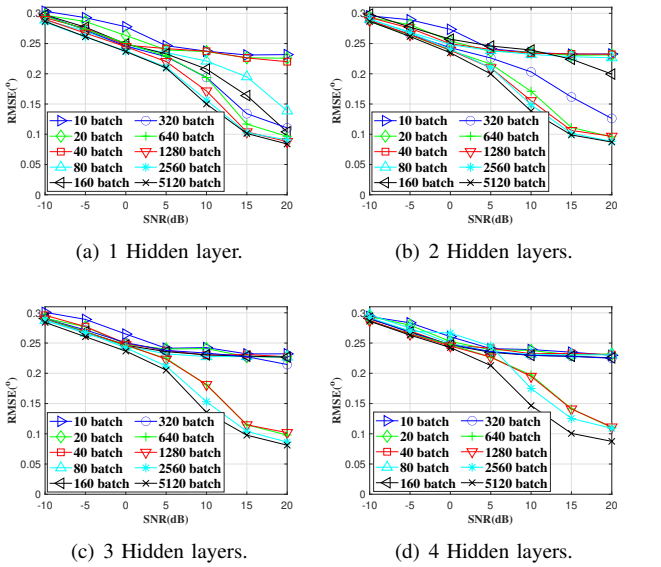
Layers	Batch Size									
	10	20	40	80	160	320	640	1280	2560	5120
1	120.42	108.03	103.32	94.68	86.37	84.37	**79.18**	75.43	73.10	72.05
2	112.15	111.55	105.37	102.40	104.68	89.89	**79.03**	73.79	76.12	72.58
3	119.75	108.68	108.66	107.64	104.73	102.80	**80.05**	78.91	71.98	70.07
4	106.61	116.60	106.65	104.91	100.23	105.75	85.40	85.10	81.98	74.66
5	118.77	108.88	116.46	104.57	102.40	102.99	103.38	100.56	98.07	80.77
6	158.00	119.26	113.94	100.47	103.48	105.73	101.74	101.95	93.51	100.01
7	150.51	157.47	117.72	108.36	104.19	106.27	104.36	102.68	101.41	100.80
8	170.38	151.29	146.30	148.87	153.07	115.87	105.14	102.52	101.77	101.57

Figure 3 shows the DOA estimation RMSE vs. SNR for different batch size, whereas Fig. 4 shows the DOA estimation RMSE vs. SNR for different number of hidden layers. We consider SNR values from -10 dB to 20 dB with a step of 5 dB. By looking at results, we observe that we achieve the best performance when the number of layers is approximately three. We instead notice that a higher batch size generally brings advantages up to a certain extent. Indeed, doubling it from 2560 to 5120 does not bring considerable advantages.

Another primary metric is the time complexity. TABLE II gives the comparison between loss and time complexity for different numbers of hidden layers and batch size. We see that time complexity grows with the number of hidden layers and batch size. In addition, batch size has higher impact on time complexity than the number of layers.

Thus, considering time and previous results, we select the most effective batch size as 640. From Fig. 4(a), we see that in almost all the cases the best number of hidden layers is 2 when the batch size is 640. Therefore, we use these values in successive evaluations.

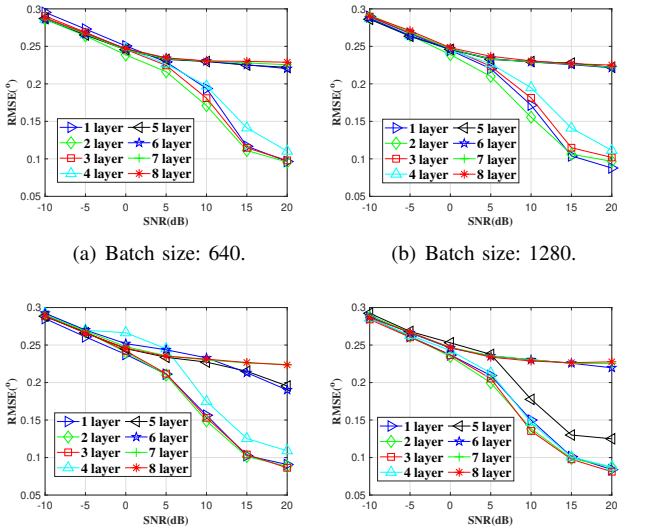
To identify the optimal size K of the spatial grid, we further conducted extended simulations. We divide the spatial directions from $-10°$ to $10°$ into 201, 67, 41, 29, 23 and 19 grids. Figure 5 shows the estimation RMSE for different K.



(a) 1 Hidden layer. (b) 2 Hidden layers.

(c) 3 Hidden layers. (d) 4 Hidden layers.

Fig. 3. RMSE of DOA estimation versus SNR under different number of hidden layers.



(a) Batch size: 640. (b) Batch size: 1280.

(c) Batch size: 2560. (d) Batch size: 5120.

Fig. 4. RMSE of DOA estimation versus SNR under different batch size.

Based on these results, we use $K = 201$ in the following investigations.

B. Spatial Spectrum and DOA Estimates

In this section, we compare our approach with two data-driven methods [50], [51] and MUSIC algorithm [55]. We firstly conduct experiments on two signals (i.e., $L = 2$) to validate the feasibility of DOA estimation on virtual channels by DNN.[1] We consider signals impinging on the receiver array from directions $-0.7°$ and $4.0°$, respectively. We consider 20 dB SNR and 20 snapshots. Figure 6(a) shows the original

[1]Considering that the beamwidth of RF channels is about $12.7°$, when two signals impinging from the spatial scope are less than $12.7°$ apart, then the sparse spatial spectrum algorithm in Section II is not available.

TABLE II
TIME COMPLEXITY OF TRAINING PROCESS (MIN)

Layers	Batch Size									
	10	20	40	80	160	320	640	1280	2560	5120
1	10.01	9.93	10.35	10.15	10.26	10.50	11.12	12.32	14.62	18.09
2	10.46	10.46	10.38	10.66	10.77	11.09	11.75	13.09	15.28	18.04
3	10.52	10.54	10.83	10.90	11.02	10.97	11.80	12.77	14.54	18.68
4	10.85	10.67	10.91	10.82	10.70	10.78	11.46	12.88	14.61	18.74
5	10.40	10.41	10.48	10.56	10.99	10.93	11.24	12.78	14.68	19.19
6	10.65	10.84	10.78	10.82	10.91	11.01	12.86	13.76	14.91	19.94
7	10.57	10.79	10.64	10.85	10.91	11.16	11.70	13.04	15.41	21.12
8	11.18	11.20	10.98	10.90	11.03	11.36	11.90	13.31	16.11	21.97

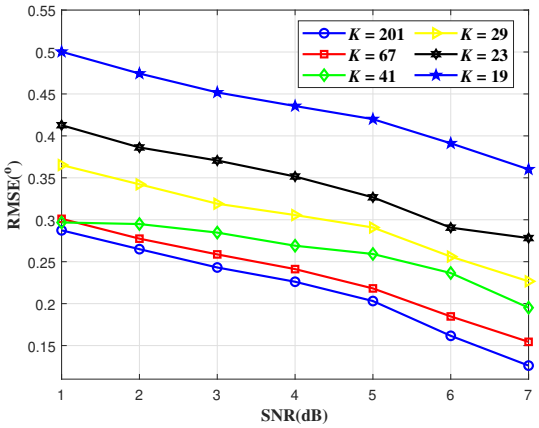
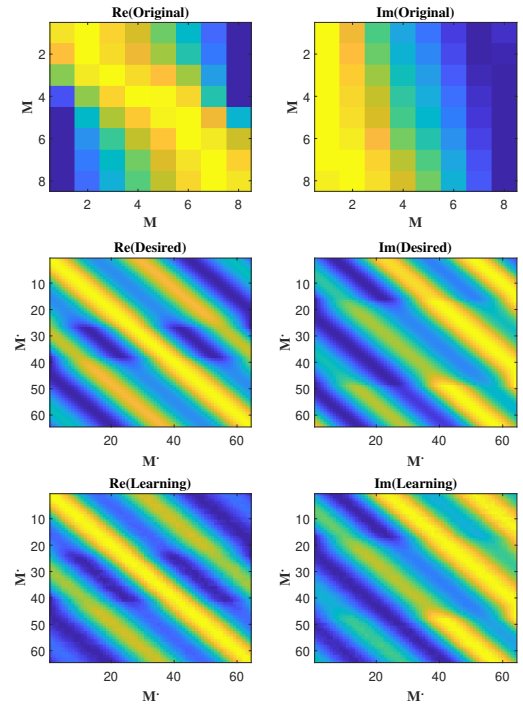
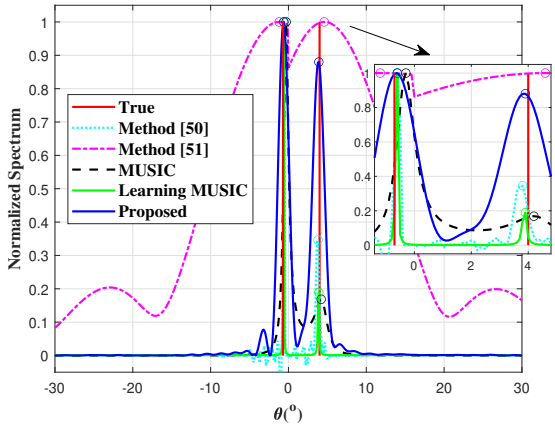


Fig. 5. RMSE of different spatial grids.



(a) Covariance matrices.



(b) Spatial spectrum.

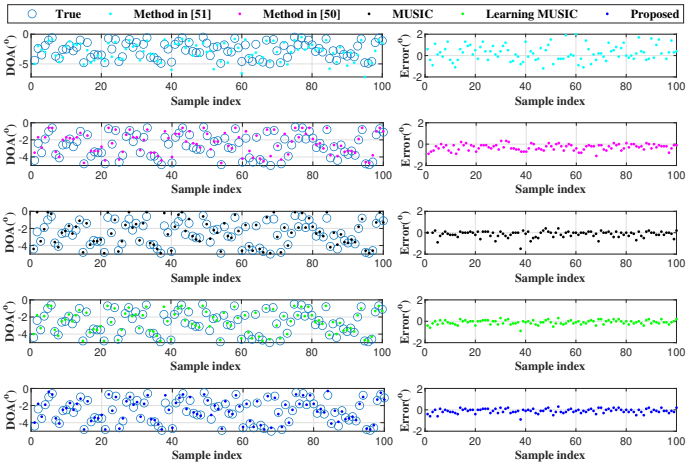
Fig. 6. DOA estimation of two sources.

(i.e., $M = 8$), desired (i.e., $M = 64$), and learned (i.e., $M' = 64$) covariance matrices. We observe that the original real/imaginary parts are polluted by noise, whereas the desired one shows the benefits of using a larger number of antennas. After DNN training, we see that the reconstructed covariance matrix is close to the desired covariance matrix. Observing the corresponding spatial spectra in Fig. 6(b), the two close signals can affect the DOA estimation accuracy in methods [50] and [51], whereas our approach provides precise estimation and provides hence superior performance. We also notice the superiority compared to MUSIC. Lastly, we apply MUSIC on top of our learning-based framework. We see that it precisely locates the two sources. However, it comes at the cost of prior knowledge on the number of sources to estimate.

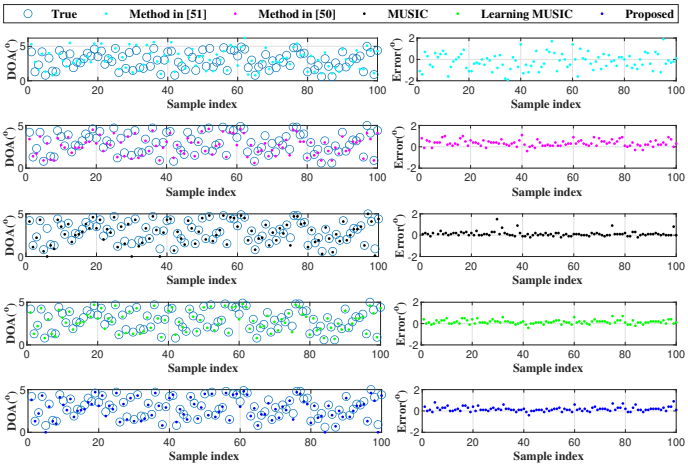
To evaluate the DOA estimation accuracy, we perform 10000 runs and compute the Root Mean Squared Error (RMSE) of DOA estimates. The RMSE can be obtained as

$$\text{RMSE} = \sqrt{\frac{1}{HK}\sum_{h=1}^{H}||\hat{\boldsymbol{\theta}}^h - \boldsymbol{\theta}||^2}, \tag{19}$$

where $\hat{\boldsymbol{\theta}}^h$ is the estimation results in the hth test, $\boldsymbol{\theta}$ is the true source direction set, H is the number of Monte-Carlo simulations and K is the number of signals. To simplify results visualization, Fig. 7 shows 100 DOA estimates and the corresponding estimation error of two sources, where the first source has random direction θ_1 from the set $[-5°, -0.5°]$, whereas the second source has random direction θ_2 from the set $[0.5°, 5°]$. The circles represent the true DOAs, and the points represent the results of algorithms. We observe that the DOA estimates obtained via our approach well match the true values and most of the estimation errors are smaller than $0.5°$. Methods in [50] and [51] fail to always recover the spatial spectrum since they are not robust enough when two signals are close to each other. Especially in method [50], the small spikes in the spatial spectrum degrade the DOA estimation precision. Looking at results we see that some of the signal realizations incur in worst estimation performance. However, this is uniquely due to the randomness of the signals impinging on the receiver array. Results show that our learning based method achieves the better estimation performance. Compared to MUSIC, our method achieves comparable results. However, compared to learning MUSIC, our method does not require



(a) θ_1



(b) θ_2

Fig. 7. DOA estimates of two sources and corresponding errors.

the number of sources as preliminary knowledge. Therefore, in later results, we focus on the sparse reconstructed spectrum rather than the constructed MUSIC spectrum.

To fully compare methods, we record the time needed for every methods in TABLE III. The estimation time represents the time needed to recombine covariance matrix, post processing, or spectral peak searching. We notice that our methods requires a higher time compared to [51] and [50]. However, it comes with the great benefit of better estimation and does not need prior knowledge on the number of sources to locate.

TABLE III
COMPUTATIONAL COMPLEXITY

	Train time (s)	Test time (s)	Estimation time (s)
Method [50]	115.1059	0.0297	13.7025
Method [51]	229.8324	0.0414	0.04203
Learning MUSIC	533.78	0.4398	65.1950
Proposed Method	533.78	0.4398	44.0019

C. RMSE versus SNR

In order to verify the performance of the proposed method, we investigate the DOA estimation accuracy vs. SNR. The SNR of the training dataset is varied from -10 dB to 20 dB with a step of 5 dB. We consider two different scenarios. In the first scenario, the SNR of the validation and training datasets are equal. That is, the SNR of validation is matched by the training datasets. In the second scenario, the SNR of the validation dataset is different from that of the training dataset, and is varied from -12 dB to 18 dB with a step of 5 dB, i.e. the SNR of validation is mismatched. In both scenarios, we fix the number of snapshots to 300. We use the average RMSE of all incident signals to calculate the statistic performance of all methods. As a benchmark value, we use the Cramer-Rao lower Bound (CRB) [56]. The CRB has been widely used to investigate the performance limit of unbiased parameter estimators. Figure 8 shows the results for scenario 1. The improved performance of the proposed estimator at low and moderate SNR is evident compared to data-driven methods in [50], [51].

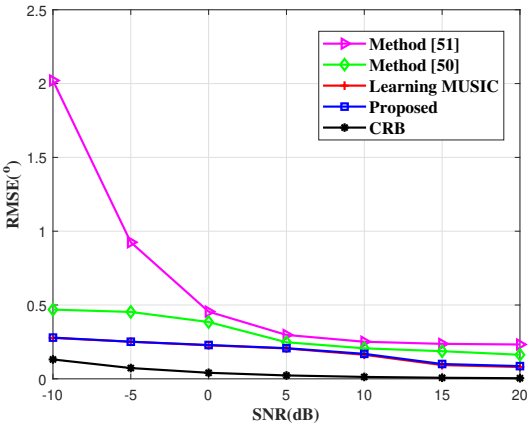


Fig. 8. RMSE vs. SNR with matched training-testing SNR.

To analyze the generalization capability of our algorithm to SNR, Fig. 9 shows the statistical RMSE of DOA estimation for scenario 2. We can see that the performance of our proposed method outperforms existing algorithms even though there is an SNR mismatch between training dataset and validation dataset. Our proposed method hence has higher generalization capability with respect to SNR mismatch.

To better show the impact of mismatched SNR on the estimation RMSE, we generate testing data with a SNR difference of $[0:5:30]$ dB with respect to the 20 dB baseline data. The 0 dB in Fig.10 represents the situation where the test data matches the training data, whereas 30 dB denotes a 30 dB difference between training and testing data. As the mismatched SNR increases, the method has worse performance. However, we notice that our proposal maintains the best performance for each mismatched SNR value.

D. RMSE versus Snapshot

In this section we test the generalization capability with respect to the number of snapshots. We consider two different

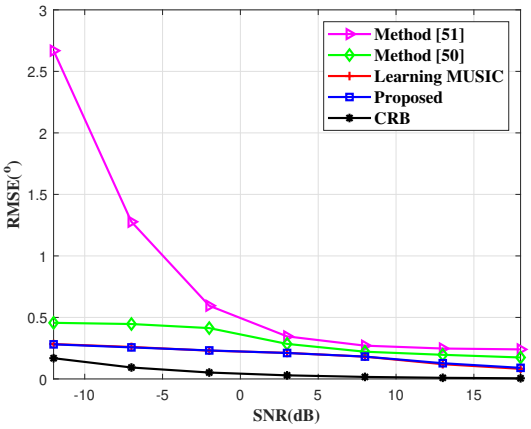


Fig. 9. RMSE vs. SNR with mismatched training-testing SNR.

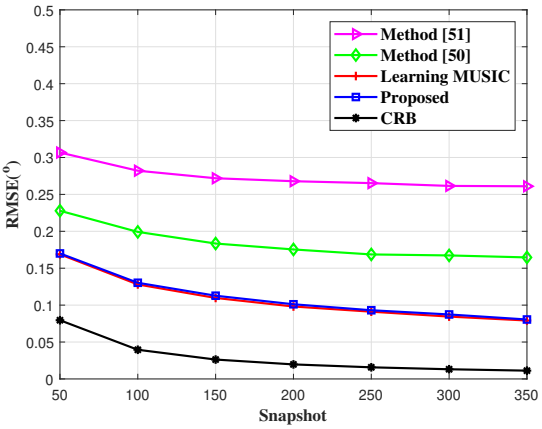


Fig. 11. RMSE vs. SNR with matched training-testing number of snapshots.

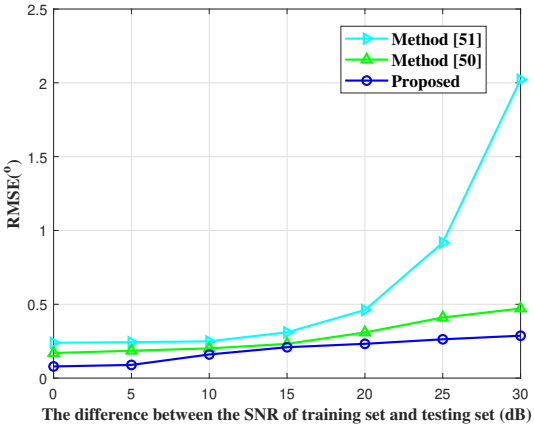


Fig. 10. The difference between the SNR of the training set and the SNR of the testing set.

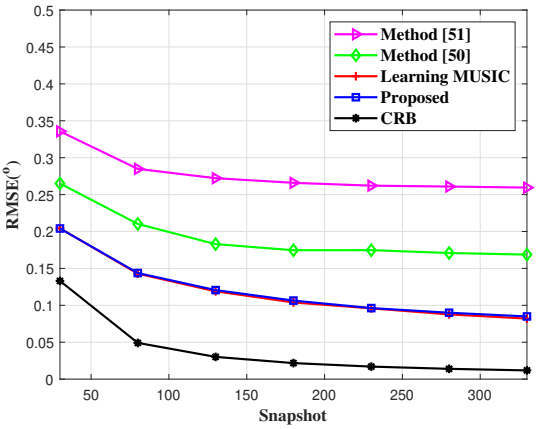


Fig. 12. RMSE vs. SNR with mismatched training-testing number of snapshots.

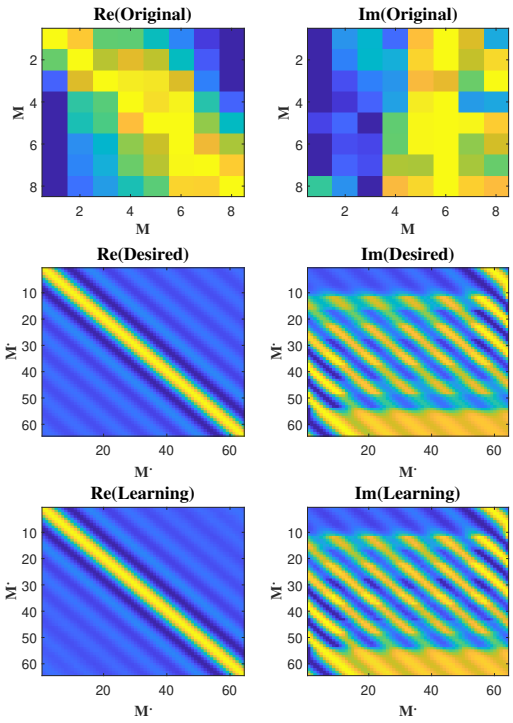
scenarios, where in both scenarios the SNR is 10 dB. In the first scenario, the number of snapshots of the validation dataset matches with that of the training dataset, i.e., goes from 50 to 350 with a step of 50. As shown in Fig. 11, thanks to virtual channels, our proposed method shows a smaller RMSE than methods in [50], [51]. The corresponding RMSE values are smaller than 0.2° as the number of snapshot is increasing.

In the second scenario, the number of snapshot of the validation dataset mismatches with that of the training dataset, and is varied from 30 to 330 with a step of 50. Figure 12 shows the RSME vs. number of snapshots in the second scenario. We observe that performance are comparable with those in Fig. 11. Our proposed method hence also has generalization capability with respect to s mismatched number of snapshots.

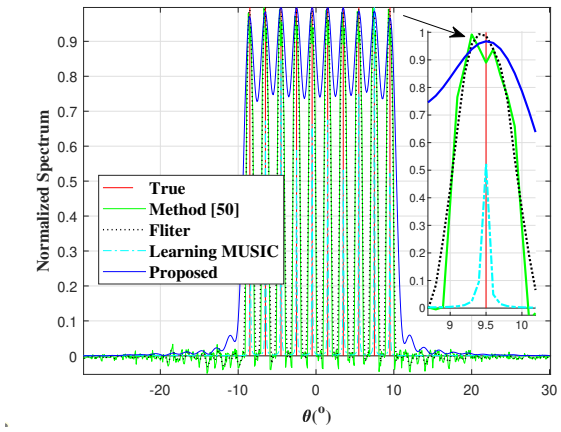
E. Multiple Sources

In this section, we consider multiple closely located sources. The main drawback of the method in [51] lies in the limitation of extending to the case of more than two targets. Thus, we neglect [51]. We first reconstruct the spectrum of multiple closely located targets. Figure 13(a) shows the covariance matrices, where the upper figure is from physical channels, the middle figure is our desired covariance matrix, and the lower figure is the one obtained via our method.

We consider signals impinging form the spatial scope of $[-30^\circ, 30^\circ]$. We consider a grid of sectors with size 0.1°, i.e., $K = 601$ sectors in total with $\theta_1 = -30^\circ, \theta_2 = -29.9^\circ, \cdots, \theta_K = 30^\circ$. For the training set, the source number L is set as $\{5, 10, 15, 20, 25\}$ and the angular separation between targets set as 2°. For $L = 5$, the corresponding first direction are randomly generated in the range of $[-4^\circ, -3^\circ]$. For lager number of sources ($L = 10, 15, 20, 25$), the first direction are randomly chosen in the range of $[-9^\circ, -8^\circ]$, $[-14^\circ, -13^\circ]$, $[-19^\circ, -18^\circ]$, and $[-24^\circ, -23^\circ]$, respectively. The SNR is set as 0 dB, and the number of snapshots is fixed at 20. Then 20000 groups of data set are used for training, and another 10000 groups are chosen for validation. We compare our proposed method with the method in [50]. We see that the original real/imaginary parts are polluted by noise, whereas thanks to DNN the distortion in the learned one is mitigated and resembles the desired one. In addition, Fig. 13(b) shows the spectrums of 10 signals from directions in $[-8.5^\circ, 9.5^\circ]$ with a step of 2°, where red lines indicate the true signal locations. We see that there are undistinguishable peaks of



(a) Covariance matrix



(b) Spatial spectrum

Fig. 13. DOA estimation of ten sources.

green lines due to the close targets location. To obtain clear peaks out of method [50], we implement Arithmetic Mean Filter (AMF) [57] on the output vector of spatial spectrum $\boldsymbol{\eta}$

$$\gamma_l = \frac{1}{Q} \sum_{q=l-\frac{Q}{2}}^{l+\frac{Q}{2}} \gamma_q, \tag{20}$$

where Q denotes the number of averaged points. After smoothing, it obtains obvious peak around the actual signal. Our approach instead does not require any smoothing filter.

DOA estimation can be obtained by searching the peaks in the overcomplete spatial spectrum, and this can be obtained by setting a threshold value above which signals represent peaks. Our work focuses on developing DOA estimation algorithms under optimal threshold selection. However, there are various imperfections in array systems and neural networks due to non-ideal antenna manufacture and neural network design that cause the output normalized power spectrum of all sources not necessarily being the same. Figure 14 illustrates the spectrum of 5 signals, while the method [50] detects 10 targets rather than 5 targets. The detectable number of targets mismatch the true number of targets in method [50].

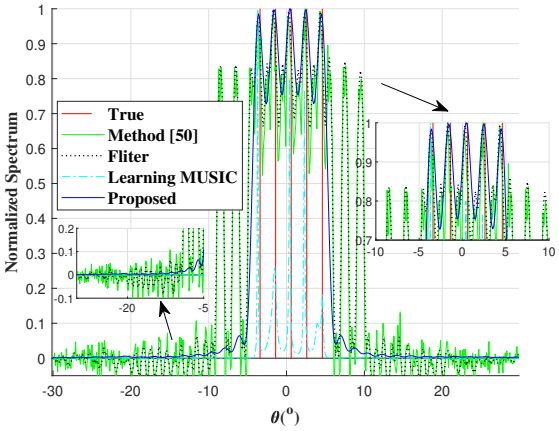


Fig. 14. The mismatched number of signals.

To avoid biases due to fixed inter-source spacing, Fig. 15 shows the spatial spectrum with $L = 10$, where signals are distributed over random locations. Also in this case, we see that method [50] detects several fake peaks, while ours do not. Thus, our method provides better results also when sources are randomly separated in space.

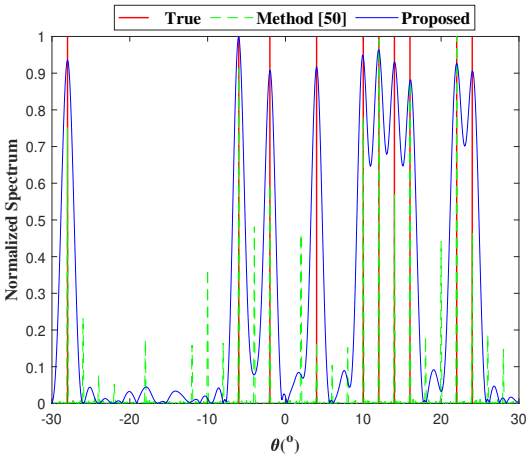
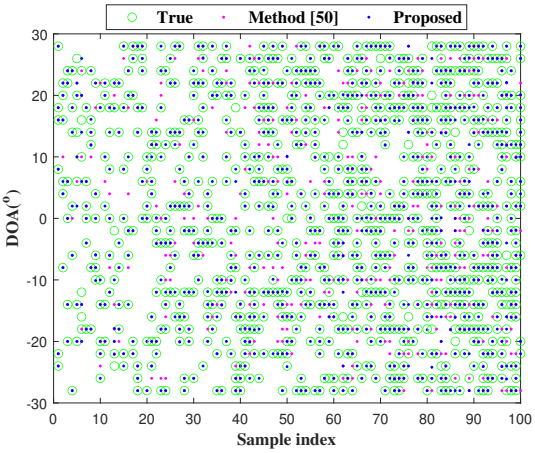


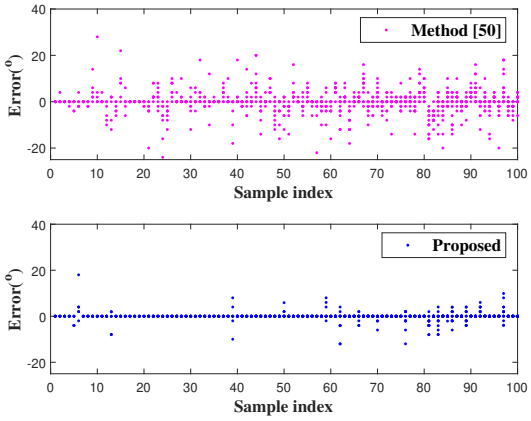
Fig. 15. Spatial spectrum of ten random sources.

Figure 16(a) shows the estimation of 100 DOA randomly distributed in space, together with the corresponding estimation error.

We see that our proposed method produces more precise DOA estimates for different directions compared to other methods. Therefore, we see its robustness and generalization to



(a) DOA estimation.



(b) Estimation error with random sources.

Fig. 16. DOA estimates of random distributed sources.

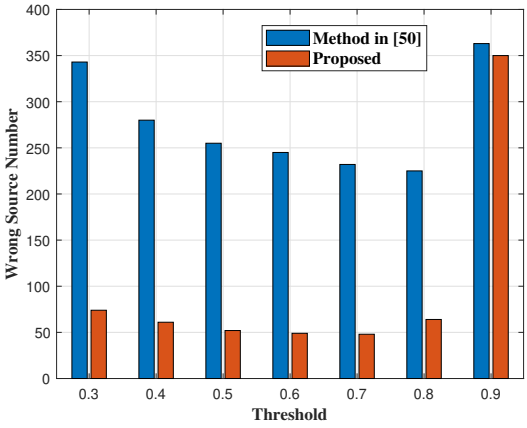
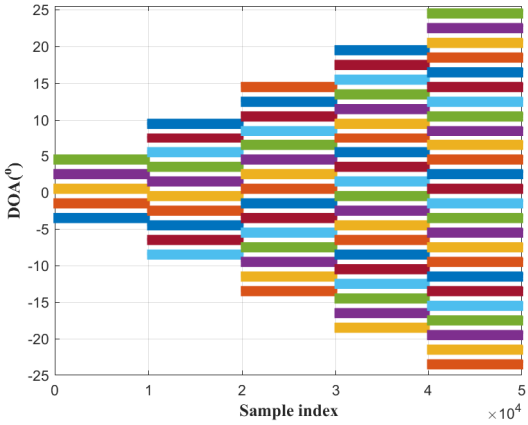
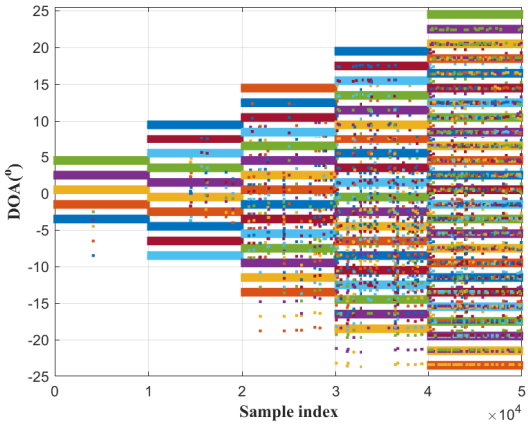


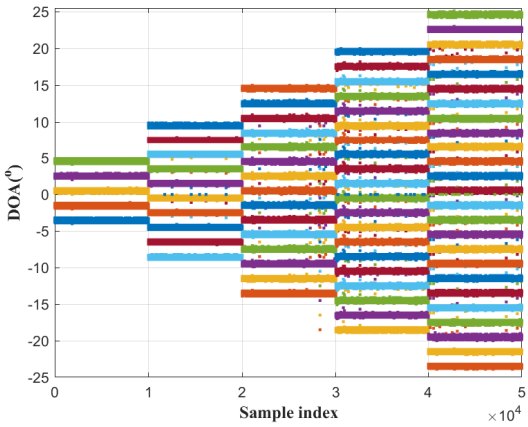
Fig. 17. Wrong estimates under different thresholds.



(a) True DOA



(b) Method in [50]



(c) Our method

Fig. 18. DOA estimates to different sources number.

random sources distances. To assess the performance obtained by our method under different threshold values, we evaluate the difference between the real number of sources and the estimated one. Figure 17 shows the number of wrongly estimated sources vs. the threshold value and compares the results obtained with our method with that in [50]. We see that the estimation accuracy is highest for threshold 0.8 for the method in [50], while for our proposed method, the optimal threshold is 0.7. Furthermore, in this simulation, we still observe that the robustness of our proposed methods is better than the method in [50]. Furthermore, we notice that at threshold 0.9 both methods attain worst performance. This is due to the fact that this threshold value is too high, and causes the estimation

methods to miss some peaks.

In Fig. 18, we show how the proposed method behaves when the testing data contains a different number of signals from the training data. The 50000 long validation set is divided into 5 segments, each of which contains 5, 10, 15, 20, and 25 signals, respectively. Figure 18(a) shows the true directions, whereas Fig. 18(b) shows the DOA estimation results from method [50]. We consider closely located sources instead of randomly distributed ones to highlight the separation capability of our method. We see that the performance of the method in [50] suffers significant degradation when $L > 5$ and the spatial angular separation is small between multiple targets. This is caused by the inherent hardware property. On the other hand, the DOA estimates of our proposed method well match their true values in Fig. 18(c). This implicates that our method can resolve more sources than the number of degrees of freedom provided by the physical antennas and the virtual difference co-array.

We also investigate the RMSE performance of our method considering a larger number of closely located sources. The comparison results between the proposed estimator and the method in [50] are provided in Fig. 19. We see that the proposed method provides low RMSE for a varying number of sources compared to method in [50].

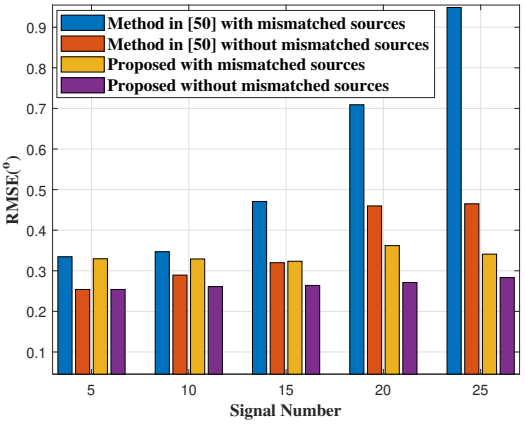


Fig. 19. RMSE of estimation versus different sources number.

We observe that performance degradation is common in method [50] when the number of source is higher than the DOF. Due to the large array aperture, the DOA estimation with sparse prior based on [50] faces the challenge of extracting enough features from small angle separation. This leads to the number of peaks of spatial spectrum being not equal to the true number of sources. After we eliminate the mismached number of sources, the RMSE of method in [50] decreases 0.5° directly. Both before and after eliminating the wrong sources, our proposed method has been shown to adapt well to multiple sources and consistently offers better performance than the method in [50] to estimate the angle of closely spaced targets. However, in practice, it is unknown if the estimated number of sources equals the true number of sources. These results illustrate the importance of flexibility in the array design provided by our deep learning framework under the condition that the physical structure of the array is not been changed.

V. Conclusion

In this paper, we proposed a generalized DNN framework for DOA estimation of LOS scenarios that achieves higher number of DOFs by extending virtual channels, which can make up for the drawbacks of spatial recovery algorithms and some data-driven methods. Compared with state-of-art data-driven methods, the proposed method aims to provide a large aperture with a reduced hardware complexity and allowing reconfigurability. We first convert the DOA estimation problem into a sparse representation method. Then the architecture designing and learning procedure of DNN are illustrated. Note that DNN-based learning models usually involve complicated nonlinear transformations of the parameters obtained through nonconvex optimization. Lastly, numerical simulations indicate that the proposed method can even resolve more sources than the number of antenna and achieve effective DOA estimation using a reduced number of circuit chains. Also, the algorithm equally outperforms excellent in the accuracy and generalization capability in all considered scenarios, when the SNR is small, or alternatively, when the number of snapshots is small.

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996, doi: 10.1109/79.526899.

[2] L. Wan, Y. Sun, L. Sun, Z. Ning, and J. J. P. C. Rodrigues, "Deep learning based autonomous vehicle super resolution DOA estimation for safety driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4301–4315, 2021, doi: 10.1109/TITS.2020.3009223.

[3] R. Akter, V.-S. Doan, T. Huynh-The, and D.-S. Kim, "RFDOA-Net: An efficient ConvNet for RF-based DOA estimation in UAV surveillance systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 12 209–12 214, 2021, doi: 10.1109/TVT.2021.3114058.

[4] X. Lin, X. Zhang, L. He, and W. Zheng, "Multiple emitters localization by UAV with nested linear array: System scheme and 2D-DOA estimation algorithm," *China Communications*, vol. 17, no. 3, pp. 117–130, 2020, doi: 10.23919/JCC.2020.03.010.

[5] M. Wang, F. Gao, S. Jin, and H. Lin, "An overview of enhanced massive MIMO with array signal processing techniques," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 886–901, 2019, doi: 10.1109/JSTSP.2019.2934931.

[6] M. Ibrahim, V. Ramireddy, A. Lavrenko, J. König, F. Römer, M. Landmann, M. Grossmann, G. Del Galdo, and R. S. Thomä, "Design and analysis of compressive antenna arrays for direction of arrival estimation," *Signal Processing*, vol. 138, pp. 35–47, 2017, doi: 10.1016/j.sigpro.2017.03.013.

[7] M. Guo, Y. D. Zhang, and T. Chen, "DOA estimation using compressed sparse array," *IEEE Transactions on Signal Processing*, vol. 66, no. 15, pp. 4133–4146, 2018, doi: 10.1109/TSP.2018.2847645.

[8] D. Park, E. Yang, S. Ahn, and J. Chun, "Adaptive beamforming for low-angle target tracking under multipath interference," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2564–2577, 2014, doi: 10.1109/TAES.2014.130185.

[9] Z. Zheng and S. Mu, "Two-dimensional DOA estimation using two parallel nested arrays," *IEEE Communications Letters*, vol. 24, no. 3, pp. 568–571, 2020, doi: 10.1109/LCOMM.2019.2958903.

[10] X. Dai, X. Zhang, and Y. Wang, "Extended DOA-matrix method for DOA estimation via two parallel linear arrays," *IEEE Communications Letters*, vol. 23, no. 11, pp. 1981–1984, 2019, doi: 10.1109/LCOMM.2019.2939245.

[11] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, no. 3, pp. 276–280, 1986, doi: 10.1109/TAP.1986.1143830.

[12] P. Gupta and M. Agrawal, "Design and analysis of the sparse array for DOA estimation of noncircular signals," *IEEE Transactions on Signal Processing*, vol. 67, no. 2, pp. 460–473, 2019, doi: 10.1109/TSP.2018.2883035.

[13] S. Ge, C. Fan, J. Wang, and X. Huang, "Low-complexity one-bit DOA estimation for massive ULA with a single snapshot," *Remote Sensing*, vol. 14, no. 14, p. 3436, 2022, doi: 10.3390/rs14143436.

[14] Y. Tan, K. Wang, L. Wang, and H. Wen, "Efficient FFT based multi source DOA estimation for ULA," *Signal Processing*, vol. 189, p. 108284, 2021, doi: 10.1016/j.sigpro.2021.108284.

[15] N. Wang, P. Agathoklis, and A. Antoniou, "A new DOA estimation technique based on subarray beamforming," *IEEE Transactions on Signal Processing*, vol. 54, no. 9, pp. 3279–3290, 2006, doi: 10.1109/TSP.2006.877653.

[16] S. Qiu, W. Sheng, X. Ma, and T. Kirubarajan, "A maximum likelihood method for joint DOA and polarization estimation based on manifold separation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2481–2500, 2021, doi: 10.1109/TAES.2021.3059094.

[17] P. Forster, P. Larzabal, and E. Boyer, "Threshold performance analysis of maximum likelihood DOA estimation," *IEEE Transactions on Signal Processing*, vol. 52, no. 11, pp. 3183–3191, 2004, doi: 10.1109/TSP.2004.836463.

[18] M. Meller and K. Stawiarski, "On DOA estimation for rotating arrays using stochastic maximum likelihood approach," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5219–5229, 2020, doi: 10.1109/TSP.2020.3022207.

[19] C.-L. Liu and P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1438–1442, 2015, doi: 10.1109/LSP.2015.2409153.

[20] F.-G. Yan, L. Shuai, J. Wang, J. Shi, and M. Jin, "Real-valued root-MUSIC for DOA estimation with reduced-dimension EVD/SVD computation," *Signal Processing*, vol. 152, pp. 1–12, 2018, doi: 10.1016/j.sigpro.2018.05.009.

[21] F. Wang, Z. Tian, G. Leus, and J. Fang, "Direction of arrival estimation of wideband sources using sparse linear arrays," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4444–4457, 2021, doi: 10.1109/TSP.2021.3094718.

[22] A. Ahmed and Y. D. Zhang, "Generalized non-redundant sparse array designs," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4580–4594, 2021, doi: 10.1109/TSP.2021.3100977.

[23] G. Qin, M. G. Amin, and Y. D. Zhang, "DOA estimation exploiting sparse array motions," *IEEE Transactions on Signal Processing*, vol. 67, no. 11, pp. 3013–3027, 2019, doi: 10.1109/TSP.2019.2911261.

[24] M. Viberg and A. Swindlehurst, "Analysis of the combined effects of finite samples and model errors on array processing performance," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3073–3083, 1994, doi: 10.1109/78.330367.

[25] P. Stoica, Z. Wang, and J. Li, "Extended derivations of MUSIC in the presence of steering vector errors," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1209–1211, 2005, doi: 10.1109/TSP.2004.842201.

[26] J. Kim, H. J. Yang, B. W. Jung, and J. Chun, "Blind calibration for a linear array with gain and phase error using independent component analysis," *IEEE Antennas and Wireless Propagation Letters*, vol. 9, pp. 1259–1262, 2010, doi: 10.1109/LAWP.2010.2104132.

[27] A. Liu, G. Liao, C. Zeng, Z. Yang, and Q. Xu, "An eigenstructure method for estimating DOA and sensor gain-phase errors," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5944–5956, 2011, doi: 10.1109/TSP.2011.2165064.

[28] W. Wang, R. Wu, J. Liang, and H. C. So, "Phase retrieval approach for DOA estimation with array errors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2610–2620, 2017, doi: 10.1109/TAES.2017.2706878.

[29] X. Zhang, Z. He, X. Zhang, and Y. Yang, "DOA and phase error estimation for a partly calibrated array with arbitrary geometry," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 497–511, 2020, doi: 10.1109/TAES.2019.2915422.

[30] S. Zhang, A. Ahmed, Y. D. Zhang, and S. Sun, "Enhanced DOA estimation exploiting multi-frequency sparse array," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5935–5946, 2021, doi: 10.1109/TSP.2021.3122292.

[31] F. Wu, F. Cao, X. Ni, C. Chen, Y. Zhang, and J. Xu, "L-shaped sparse array structure for 2-D DOA estimation," *IEEE Access*, vol. 8, pp. 140 030–140 037, 2020, doi: 10.1109/ACCESS.2020.3012685.

[32] A. Brighente, J. Gambini, and S. Tomasin, "Modular hybrid beamforming for mmWave fixed wireless access," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 5145–5158, 2020, doi: 10.1109/TCOMM.2020.2992470.

[33] J. Wang, W.-X. Sheng, Y.-B. Han, and X.-F. Ma, "Adaptive beamforming with compressed sensing for sparse receiving array," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 823–833, 2014, doi: 10.1109/TAES.2014.120532.

[34] I. Aboumahmoud, A. Muqaibel, M. Alhassoun, and S. Alawsh, "A review of sparse sensor arrays for two-dimensional direction-of-arrival estimation," *IEEE Access*, vol. 9, pp. 92 999–93 017, 2021, doi: 10.1109/ACCESS.2021.3092529.

[35] B. Brkljač, M. Janev, R. Obradović, D. Rapaić, N. Ralević, and V. Crnojević, "Sparse representation of precision matrices used in GMMs," *Applied intelligence*, vol. 41, no. 3, pp. 956–973, 2014, doi: 10.1007/s10489-014-0581-6.

[36] Y. Ma, X. Wang, and L. Wei, "Multi-level spatial and semantic enhancement network for expression recognition," *Applied Intelligence*, pp. 1–14, 2021, doi: 10.1007/s10489-021-02254-0.

[37] A. Sinha and M. Rathi, "COVID-19 prediction using AI analytics for South Korea," *Applied Intelligence*, pp. 1–19, 2021, doi: 10.1007/s10489-021-02352-z.

[38] O. Appel, F. Chiclana, J. Carter, and H. Fujita, "Successes and challenges in developing a hybrid approach to sentiment analysis," *Applied Intelligence*, vol. 48, no. 5, pp. 1176–1188, 2018, doi: 10.1007/s10489-017-0966-4.

[39] A. H. El Zooghby, C. G. Christodoulou, and M. Georgiopoulos, "A neural network-based smart antenna for multiple source tracking," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 5, pp. 768–776, 2000, doi: 10.1109/8.855496.

[40] M. Pastorino and A. Randazzo, "A smart antenna system for direction of arrival estimation based on a support vector regression," *IEEE transactions on antennas and propagation*, vol. 53, no. 7, pp. 2161–2168, 2005, doi: 10.1109/TAP.2005.850735.

[41] H. Xiang, B. Chen, M. Yang, and C. Li, "Altitude measurement based on characteristics reversal by deep neural network for VHF radar," *IET Radar, Sonar & Navigation*, vol. 13, no. 1, pp. 98–103, 2018, doi: 10.1049/iet-rsn.2018.5121.

[42] K. Terabayashi, R. Natsuaki, and A. Hirose, "Ultrawideband direction-of-arrival estimation using complex-valued spatiotemporal neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 9, pp. 1727–1732, 2014, doi: 10.1109/TNNLS.2014.2313869.

[43] A. Randazzo, M. A. Abou-Khousa, M. Pastorino, and R. Zoughi, "Direction of arrival estimation based on support vector regression: Experimental validation and comparison with MUSIC," *IEEE Antennas and Wireless propagation letters*, vol. 6, pp. 379–382, 2007, doi: 10.1109/LAWP.2007.903491.

[44] Y. Gao, D. Hu, Y. Chen, and Y. Ma, "Gridless 1-b DOA estimation exploiting SVM approach," *IEEE Communications Letters*, vol. 21, no. 10, pp. 2210–2213, 2017, doi: 10.1109/LCOMM.2017.2723359.

[45] L.-L. Wu and Z.-T. Huang, "Coherent SVR learning for wideband direction-of-arrival estimation," *IEEE Signal Processing Letters*, vol. 26, no. 4, pp. 642–646, 2019, doi: 10.1109/LSP.2019.2901641.

[46] H. Xiang, B. Chen, M. Yang, S. Xu, and Z. Li, "Improved direction-of-arrival estimation method based on LSTM neural networks with robustness to array imperfections," *Applied Intelligence*, pp. 1–14, 2021, doi: 10.1007/s10489-020-02124-1.

[47] H. Xiang, B. Chen, T. Yang, and D. Liu, "Phase enhancement model based on supervised convolutional neural network for coherent DOA estimation," *Applied Intelligence*, vol. 50, no. 8, pp. 2411–2422, 2020, doi: 10.1007/s10489-020-01678-4.

[48] Z.-M. Liu, C. Zhang, and S. Y. Philip, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7315–7327, 2018, doi: 10.1109/TAP.2018.2874430.

[49] S. Chakrabarty and E. A. Habets, "Multi-speaker DOA estimation using deep convolutional networks trained with noise signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 1, pp. 8–21, 2019, doi: 10.1109/JSTSP.2019.2901664.

[50] L. Wu, Z.-M. Liu, and Z.-T. Huang, "Deep convolution network for direction of arrival estimation with sparse prior," *IEEE Signal Processing Letters*, vol. 26, no. 11, pp. 1688–1692, 2019, doi: 10.1109/LSP.2019.2945115.

[51] H. Xiang, B. Chen, M. Yang, and S. Xu, "Angle separation learning for coherent DOA estimation with deep sparse prior," *IEEE Communications Letters*, vol. 25, no. 2, pp. 465–469, 2020, doi: 10.1109/LCOMM.2020.3032733.

[52] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, "Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 1–12, 2013, doi: 10.1109/TWC.2013.071113.121305.
[53] P. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990, doi: 10.1109/5.58337.
[54] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014, doi: 10.48550/arXiv.1412.6980.
[55] M. H. Gruber, "Statistical digital signal processing and modeling," *Technometrics*, vol. 39, no. 3, pp. 335–336, 1997, doi: 10.1080/00401706.1997.10485128.
[56] P. Stoica, E. Larsson, and A. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Processing Letters*, vol. 8, no. 5, pp. 148–150, 2001, doi: 10.1109/97.917699.
[57] K. A. Oyesina and M. A. Ogunlade, "Noise reduction using arithmetic mean filtering (A comparison study of application to different noise types)," *International Journal of Science and Research (IJSR)*, vol. 6, no. 2, pp. 1–4, 2017, doi: 10.21275/ART2017774.

Saiqin Xu received the B.S. degree in the Computer Science and Technology, Anhui Normal University, China, in 2019. She is currently working toward the Ph.D. degree with National Laboratory of Radar Signal Processing, Xidian University, Xi'an, China. Since October 2022, she has been a Visiting Student with SPRITZ Security and Privacy Research Group, Department of Mathematics, University of Padova, Italy, with the support of China Scholarship Council. Her research interests include signal processing, parameter estimation, radar systems engineering, artificial neural networks and security and privacy for wireless vehicular communications.

Alessandro Brighente is assistant professor at the University of Padova. He received his Ph.D. degree in Information Engineering from the University of Padova in Feb. 2021. He was visiting researcher at Nokia Bell Labs, Stuttgart and University of Washington, Seattle in 2019 and 2022, respectively. He has been involved in European projects and industrial projects with the University of Padova. He served as TPC for several conferences, including Globecom and VTC. He is guest editor for IEEE Transactions on Industrial Informatics. His current research interests include security and privacy in cyber-physical systems, vehicular networks, blockchain, and physical layer security.

Baixiao Chen was born in Anhui, China, in 1966. He graduated from the Metallurgy College of East China, Ganzhou, China, in 1987 and worked there from July 1987 to July 1991. He received the master's degree in circuit and system, and the Ph.D. degree in signal and information processing from Xidian University, Xi'an, China, in 1994 and 1997, respectively. He was with the Metallurgy College of East China from July 1987 to July 1991. Since 1997, he has been a Faculty Member with the National Laboratory of Radar Signal Processing, Xi'an, China. He was a Lecturer and an Associate Professor from October 1997 to May 1999 and from June 1996 to May 2003, respectively. In 2006, he was selected into "New Century elitist Support Program" of the Ministry of Education. He is currently a Professor of Signal and Information Processing with Xidian University. His general research interests include radar signal processing, new radar system design, array signal processing, precision guidance, etc.

Mauro Conti is Full Professor at the University of Padua, Italy. He is also affiliated with TU Delft and University of Washington, Seattle. He obtained his Ph.D. from Sapienza University of Rome, Italy, in 2009. After his Ph.D., he was a Post-Doc Researcher at Vrije Universiteit Amsterdam, The Netherlands. In 2011 he joined as Assistant Professor at the University of Padua, where he became Associate Professor in 2015, and Full Professor in 2018. He has been Visiting Researcher at GMU, UCLA, UCI, TU Darmstadt, UF, and FIU. He has been awarded with a Marie Curie Fellowship (2012) by the European Commission, and with a Fellowship by the German DAAD (2013). His research is also funded by companies, including Cisco, Intel, and Huawei. His main research interest is in the area of Security and Privacy. In this area, he published more than 450 papers in topmost international peer-reviewed journals and conferences. He is Editor-in-Chief for IEEE Transactions on Information Forensics and Security, Area Editor-in-Chief for IEEE Communications Surveys & Tutorials, and has been Associate Editor for several journals, including IEEE Communications Surveys & Tutorials, IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Information Forensics and Security, and IEEE Transactions on Network and Service Management. He was Program Chair for TRUST 2015, ICISS 2016, WiSec 2017, ACNS 2020, CANS 2021, and General Chair for SecureComm 2012, SACMAT 2013, NSS 2021 and ACNS 2022. He is Fellow of the IEEE, Senior Member of the ACM, and Fellow of the Young Academy of Europe.

Xiancheng Cheng received the B.S. degree in electrical and information engineering from XidiChenn University, Xi'an, China, in 2016. He is currently work-ing toward the Ph.D. degree in signal and information processing with the School of Electronic Engineering, Xidian University, Xi'an, China. Since September 2019, he has been a Visiting Student with the Department of Electrical Engineering and Computer Science, Syra-cuse University, Syracuse, NY, USA, with the support of China Scholarship Council. His current research interests include data fusion, wireless sensor network, target detection and tracking.

Dongchen Zhu received the B.S. degree in electrical and information engineering from Xidian University, Xi'an, China, in 2017. He is currently working toward the Ph.D. degree in signal and information processing with the School of Electronic Engineering, Xidian University, Xi'an, China. His current research interests include array signal processing, polarization sensitive array, parameter estimation, etc.