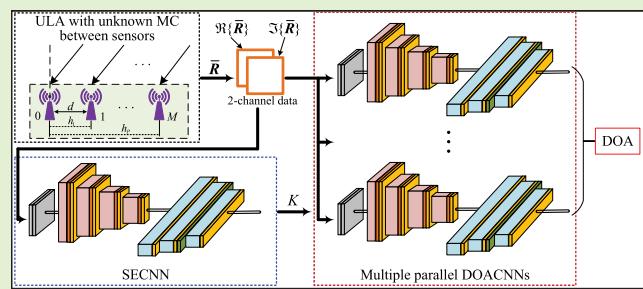


# Effective High-Resolution Off-Grid DOA Estimation With Mutual Coupling via CNN Framework

Huafei Wang<sup>ID</sup>, Member, IEEE, Xianpeng Wang<sup>ID</sup>, Member, IEEE, Xiang Lan<sup>ID</sup>, Member, IEEE, and Ting Su<sup>ID</sup>, Member, IEEE

**Abstract**—Achieving high-resolution direction-of-arrival (DOA) estimation under the condition of mutual coupling (MC) is of great practical importance. Traditional parametric approaches are limited by either the computational complexity or the resolution. While, existing neural network (NN)-based end-to-end methods achieve DOA estimation by spatial spectrum regression task or multilabel classification task, which results in the performance being limited by off-grid errors. Besides, the MC between the sensors may seriously affect the precision and resolution of these approaches. In this article, an NN framework consisting of convolutional NN (CNN) is designed for effective high-resolution end-to-end off-grid DOA estimation with MC. The designed CNN framework consists of source enumeration CNN (SECNN) and multiple parallel DOA estimation CNNs (DOACNN) to enhance its generalization for different numbers of sources. The SECNN is designed to provide knowledge of source numbers for DOACNN, and the DOACNN is to achieve DOA estimation based on the provided knowledge. Both SECNN and DOACNN are trained by the output covariance of certain arrays thus robust to MC. Meanwhile, the DOACNN models off-grid error into labels to achieve effective end-to-end off-grid DOA estimation. As compared to the representative approaches, the proposed method not only can realize high-precision off-grid DOA estimation in the presence of MC but also has noticeable advantages in terms of resolution and efficiency.



**Index Terms**—Array signal processing, convolutional neural network (CNN), high-resolution direction-of-arrival (DOA) estimation, mutual coupling (MC), off-grid error.

## NOMENCLATURE

Italic letters	Scalars.
Bold italic lowercase letters	Vectors.
Bold italic capital letters	Matrices.
$[ \cdot ]^T$	Transpose operator for matrix/vector.
$[ \cdot ]^H$	Hermitian transpose operator for matrix/vector.
$\mathbb{R}^{i \times j}$	$(i \times j)$ dimensional real matrix set.

Received 29 September 2024; accepted 4 November 2024. Date of publication 15 November 2024; date of current version 2 January 2025. This work was supported in part by Hainan Province Science and Technology Special Fund under Grant ZDYF2023GXJS159; and in part by the National Natural Science Foundation of China under Grant 62461020, Grant 62101165, and Grant 62061013. The associate editor coordinating the review of this article and approving it for publication was Dr. Geethu Joseph. (*Corresponding author: Xianpeng Wang.*)

The authors are with the School of Information and Communication Engineering and the State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou 570228, China (e-mail: wong9525@163.com; wxpeng2016@hainanu.edu.cn; xlan@hainanu.edu.cn; suting4190@hainanu.edu.cn).

Digital Object Identifier 10.1109/JSEN.2024.3494884

$\mathbb{C}^{i \times j}$	$(i \times j)$ dimensional complex matrix set.
$\Re\{\cdot\}$	Real component extractor.
$\Im\{\cdot\}$	Imaginary component extractor.
$\text{diag}\{\cdot\}$	Operator for diagonalizing vectors to matrix.
$\otimes$	Convolution operator.
$\lceil \cdot \rceil$	Upward rounding operator.
$  \cdot  $	Modulus of scalar.
$\mathcal{T}\{\cdot\}$	Converters for converting vector to Toeplitz matrix.
$\mathcal{C}_M^m$	Generate combinations with $m$ entries from $M$ entries.

## I. INTRODUCTION

IN ARRAY signal processing, direction-of-arrival (DOA) estimation is a fundamental and longstanding concern that has received scholarly attention for decades [1], [2]. It has been widely applied across numerous fields including wireless communications, navigation, localization, target sensing, etc. [3], [4]. During the long development history of DOA estimation, a large number of outstanding approaches have been proposed [5]. One of the most representative is the

subspace-based approaches, which includes the multiple signals classification (MUSIC) [6], the estimation of parameters by rotational invariant techniques (ESPRIT) [7], and their various variants [8], [9]. However, since the performance of subspace-based approaches is contingent upon the precision of array covariance, which is closely related to the signal-to-noise ratio (SNR) and the number of snapshots, their performance faces degradation at low SNRs and small snapshots. On the other hand, with the development of DOA estimation and the rise of compressed sensing (CS) techniques [10], sparse recovery brings a new perspective to DOA estimation. The representative sparse recovery approaches for DOA estimation include  $\ell_1$ -SVD [11] and sparse Bayesian learning (SBL) [12]. However, the performance of conventional sparse representation approaches is restricted by off-grid error caused by grid mismatch. A sufficiently dense preset grid may be effective in minimizing off-grid errors and thus improving estimation precision, but this comes at the expense of computational complexity. Therefore, the off-grid DOA estimation with low-computational complexity in sparse representation raised much attention, and many approaches have been proposed, such as the refined DOA estimation in [13], off-grid sparse Bayesian inference (OGSBI) method [14], off-grid block SBL (OGBSBL) [15], and root off-grid SBL (ROGSBL) method [16]. However, the above approaches all rely upon an ideal array manifold. In practice, there usually occurs mutual coupling (MC) between sensors due to imperfect sensor calibration [17], [18], which will cause performance degradation or even invalidation for these methods.

For the DOA estimation in the presence of MC, plenty of methods have been proposed [19], [20], [21], [22], [23], [24]. Specifically, a sparse representation approach was presented in [19] by leveraging the Toeplitz structure inherent in the MC matrix (MCM) to handle MC, then using the  $\ell_1$ -norm minimization principle to obtain DOA estimation. However, this method will lead to a diminution in the effective array aperture. Hence, a block sparse representation (BSR) method was proposed in [20], in which the steering vector is re-parameterized to form a block data model so as to avoid the loss of array aperture. Further, inspired by the aforementioned BSR, the block sparse recovery methods based on re-weighted  $\ell_1$ -norm minimization and weighted subspace fitting were respectively presented in [21] and [22] for superior performance. On the other hand, Liu and Zhou introduced a unified SBL method for DOA estimation in the case of array imperfections [23], which is also robust to MC. Similarly, an algorithm called SODMC was proposed in [24] where the parameters including the MC coefficients are estimated by iteratively updating. However, all the above-mentioned methods are model-driven, which depend on an accurate pre-established model, and a complete online optimization procedure is needed for every single estimation. This is time-consuming to some extent and unfriendly to some practical scenarios that require real-time DOA estimation.

Most recently, with the rapid development of artificial intelligence (AI), the DOA estimation based on deep learning

(DL) [25] technique is gradually attracting attention [26], [27], [28]. In contrast to model-driven methods, the DL-based approaches are completely data-driven. They realize DOA estimation through learning features directly from the array output by the powerful nonlinear mapping capabilities of neural networks (NN) without online optimization procedure and pre-built model, which is much time-saving, and the training is to be done offline once and for all. On the other hand, DL-based approaches have greater flexibility in adapting to various environments and tasks, which allows them to be better adapted to different application scenarios. Currently, there are two kinds of DL-based approaches have been investigated: 1) semi-DL approaches [29], [30], [31] and 2) pure-DL approaches [32], [33], [34], [35], [36], [37], [38], [39]. The principal concept of the former approaches involves utilizing NN to reconstruct array covariance, and then following a parametric method to estimate DOAs. While the latter focuses on using NN to realize end-to-end DOA estimation.

Particularly, a novel NN framework is introduced in [32] to achieve frequency estimation with an additional NN for the number of frequencies estimation, which can be directly applied in DOA estimation. In [33], authors construct a deep convolution network using 1-dimensional convolutions for DOA estimation based on the sparse prior, which shows certain performance advantages at low SNRs. Meanwhile, a convolution neural network (CNN) using 2-dimensional convolutions is designed in [34] for DOA estimation in a low SNR scenario. Most recently, an interpretable and efficient beamforming-based DL approach is proposed in [35] for DOA estimation in the case of a single snapshot, which translates minimum power distortionless response (MPDR)-type beamformer into DL to enhance generalization, interpretation, and efficiency. However, the performance of these approaches suffers from off-grid errors caused by grid mismatch. On the other hand, they are grounded upon the assumption of an ideal array manifold, which means they may suffer from performance degradation or even failure when the MC or antenna failure occurs. To tackle this problem, a deep NN consisting of an autoencoder and multiple parallel classifiers was presented in [36] to achieve DOA estimation with robustness to array imperfections. However, it employs a singular SNR for the training and testing phases, which is a practice that may deviate from congruence with practical applications. In [37], authors proposed a DL framework for multi-input multi-output (MIMO) radar robust DOA estimation in the nonideal environments including MC, coherent sources, colored noise, etc. In [38], authors propose a beamspace-based deep NN for the DOA estimation with array imperfections and antenna failure, where the denoising autoencoder is introduced to resist antenna failure and improve DOA estimation accuracy. Meanwhile, a DL framework is designed in [39] to achieve DOA estimation in the case of antenna failures with single snapshot sparse arrays, which shows faster inference times and robust performance in low SNR environments. However, these methods require a known number of sources as a priori knowledge and ignore the

off-grid error. Besides, their performance is not that good when the sources are closely spaced.

Based on the above review, it is evident that the current DL-based end-to-end DOA estimation approaches are either restricted by the off-grid error or have low precision and resolution in the presence of MC. To tackle this issue, therefore, we focus on the high-resolution end-to-end off-grid DOA estimation with MC and without knowledge of source numbers in advance. To achieve this goal, a CNN framework consisting of source enumeration CNN (SECNN) and DOA estimation CNN (DOACNN) is designed in this article. The SECNN estimates the source number from array covariance, while the DOACNN achieves off-grid DOA estimation based on the knowledge of source numbers provided by SECNN. The output covariance of certain arrays is fed into the CNNs for training so as to against the MC. In comparison with contemporary representative approaches, the designed CNN framework demonstrates the capability of attaining effective high-resolution off-grid DOA estimation, while having significant performance advantages. The principal contributions of this article can be succinctly delineated.

- 1) A NN framework consisting of CNNs is designed.

The designed framework is composed of SECNN and DOACNN, which enables it can achieve DOA estimation while providing the knowledge of a number of sources.

- 2) The designed CNNs are robust to the sensor MC and off-grid errors. The covariance of the certain array is fed to train both SECNN and DOACNN, and the off-grid error is modeled into the labels of DOACNN for training, which enables it can realize end-to-end off-grid DOA estimation in the presence of MC.
- 3) Sufficient simulations indicate that the designed framework has an obvious performance advantage as compared to some representative approaches, and shows significant superiority in terms of high-resolution off-grid DOA estimation.

The rest of this article is arranged as. The signal model with MC is introduced in Section II. Then, the designed CNN framework is presented in Section III. In Section IV, the generation of training datasets and corresponding training strategies are given. Afterward, sufficient numerical simulation results are given in Section V to demonstrate the effectiveness and superiority of the designed framework. Finally, the conclusion is made in Section VI based on the simulation results. In addition, the interpretations of notations involved throughout the article are given in Nomenclature for convenience.

## II. SIGNAL MODEL WITH MC

Herein, consider a uniform linear array (ULA) composed of  $M$  sensors as shown in Fig. 1. When  $K$  far-field narrowband sources impinging on it from different directions of  $\theta = \{\theta_1, \theta_2, \dots, \theta_K\}$ , the received signal devoid of MC at the  $t$ th snapshot is articulated as

$$\mathbf{x}(t) = \sum_{k=1}^K \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

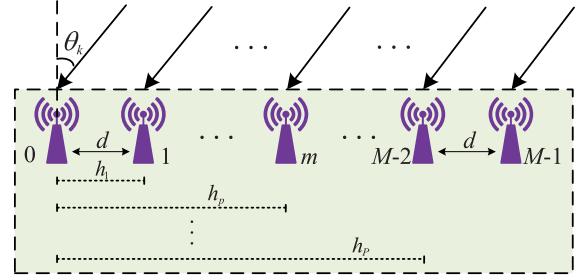


Fig. 1. Schematic of ULA with MC.

where  $t = 1, 2, \dots, T$ ,  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^\top$  and  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_K(t)]^\top$  denote the received signal and incident signal vectors, respectively.  $\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^\top$  represents the vector corresponding to additive noise, which is considered as Gaussian white noise in this article. The steering matrix is  $\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)]$  with steering vector

$$\mathbf{a}(\theta_k) = [\alpha(\theta_k)^0, \alpha(\theta_k)^1, \alpha(\theta_k)^2, \dots, \alpha(\theta_k)^{M-1}]^\top \quad (2)$$

where  $\alpha(\theta_k) = e^{j2\pi d \sin(\theta_k)/\lambda}$ , the variables  $d$  and  $\lambda$ , respectively, symbolize the intersensor spacing and the wavelength of the incident signal.

In the presence of MC interactions among sensors, the expression in (1) turns into

$$\bar{\mathbf{x}}(t) = \bar{\mathbf{A}}\mathbf{s}(t) + \mathbf{n}(t) \quad (3)$$

with  $\bar{\mathbf{A}} = \mathbf{H}[\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)]$  being the destructed steering matrix by sensor MC. The MC in ULA is usually denoted by a matrix (MCM)  $\mathbf{H} \in \mathbb{C}^{M \times M}$ , which is characterized as a symmetric Toeplitz matrix with banded structure [17], i.e.,

$$\mathbf{H} = \mathcal{T}\{[\mathbf{h}, \mathbf{0}]\} = \begin{bmatrix} h_0 & h_1 & \cdots & h_P & \cdots & 0 \\ h_1 & h_0 & h_1 & \cdots & \ddots & \vdots \\ \vdots & h_1 & h_0 & \ddots & \ddots & h_P \\ h_P & \cdots & \ddots & \ddots & h_1 & \vdots \\ \vdots & \ddots & \cdots & h_1 & h_0 & h_1 \\ 0 & \cdots & h_P & \cdots & h_1 & h_0 \end{bmatrix} \quad (4)$$

where  $\mathbf{h} = [h_0, h_1, h_2, \dots, h_P]$  is the MC coefficient vector with  $|h_P| < \dots < |h_2| < |h_1| < |h_0| = 1$ , and  $\mathbf{0}$  denotes a row vector whose all entries are 0.

With MC, the theoretical array covariance can be obtained as

$$\mathbf{R} = \bar{\mathbf{A}}\mathbf{R}_s\bar{\mathbf{A}}^\text{H} + \mathbf{R}_n = \mathbf{H}\mathbf{A}\mathbf{R}_s\mathbf{A}^\text{H}\mathbf{H}^\text{H} + \mathbf{R}_n \quad (5)$$

where  $\mathbf{R}_s = \text{diag}\{\delta_1, \delta_2, \dots, \delta_K\}$  denotes the signal covariance, and  $\mathbf{R}_n$  stands for the noise covariance. However, the above theoretical array covariance is based on  $T \rightarrow \infty$ , which is hardly achieved in practice. Therefore, it is usually approximated by its maximum likelihood estimation, i.e., the

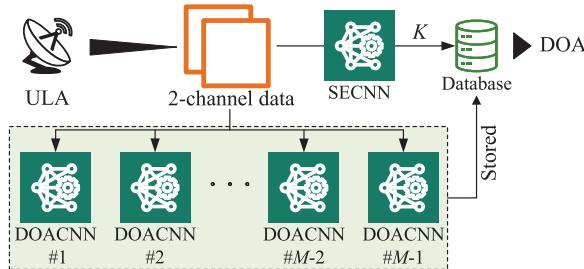


Fig. 2. Overall architecture of the designed CNN framework.

practical array covariance is

$$\bar{\mathbf{R}} = \frac{1}{T} \sum_{t=1}^T \bar{\mathbf{x}}(t) \bar{\mathbf{x}}(t)^H \quad (6)$$

where  $T$  signifies the total count of the snapshots number.

Currently, there already have many traditional model-driven approaches have been proposed based on the  $\bar{\mathbf{R}}$  to achieve DOA estimation in the presence of MC, for example, the approaches in [19] and [22]. These approaches can be viewed as a nonlinear mapping procedure for extracting DOA information  $\bar{\theta}$  from  $\bar{\mathbf{R}}$ , i.e.,  $f : \bar{\mathbf{R}} \rightarrow \bar{\theta}$ . However, they rely on prebuilt models and require a complete online optimization process during each mapping, which is somewhat time-consuming and not applicable to practical application scenarios that require real-time estimation. While, the NN has been proven that it can be more easily realized in real-time estimation by the powerful nonlinear mapping capabilities of trained NNs without online optimization procedures and a prebuilt model, and the training is usually to be done offline once and for all. Besides, the NNs have greater flexibility in adapting to various environments, which allows them to be better adapted to different application scenarios. Therefore, we design a CNN-based framework, which has the nonlinear mapping ability of  $\mathcal{F} : \bar{\mathbf{R}} \rightarrow \bar{\theta}$ , to retrieve DOAs from the array covariance with MC. Meanwhile, the designed framework can achieve high-resolution DOA estimation with robustness to the off-grid error.

### III. DESIGNED CNN-BASED FRAMEWORK FOR OFF-GRID DOA ESTIMATION

The overall architecture of the designed CNN-based framework is illustrated in Fig. 2, where contains two parts: the SECNN and DOACNN. In the proposed framework, multiple parallel DOACNNs are designed to enhance its generalization for different numbers of sources, which therefore leads to its strong dependence on the knowledge of source numbers. However, the source number in practice is usually unknown in advance. Therefore, the SECNN is designed to provide knowledge of source numbers to ensure the effectiveness and generalization of DOACNN. In the designed framework, the DOACNN can be pretrained for the scenarios of different sources number, then fixed and stored in a database. When performing DOA estimation, the source number  $K$  will be first determined by SECNN. Then, the certain trained DOACNN for the scenario of  $K$  sources will

be called directly from the database to facilitate end-to-end DOA estimation.

In the designed framework, both SECNN and DOACNN exhibit analogous structural, but their specific parameters are different to cope with different tasks, which is shown in Fig. 3. Note that the theoretical covariance  $\mathbf{R}$  contains MCM and is based on  $T \rightarrow \infty$ , which makes it impossible to compute in practice for training. Hence, the two-channel tensor data  $\bar{\mathbf{Y}} \in \mathbb{R}^{M \times M \times 2}$  based on practical array covariance  $\bar{\mathbf{R}}$  is fed into CNN, which is

$$\begin{cases} \bar{\mathbf{Y}}(:, :, 1) = \Re\{\bar{\mathbf{R}}\} \\ \bar{\mathbf{Y}}(:, :, 2) = \Im\{\bar{\mathbf{R}}\}. \end{cases} \quad (7)$$

After inputting  $\bar{\mathbf{Y}}$ , it will traverse four convolutional (CON) layers with  $N_{ci}$  ( $i = 1, 2, 3, 4$ ) channels, followed by two fully connected (FC) layers with  $N_{fj}$  ( $j = 1, 2$ ) neurons, then a dropout layer is followed for preventing over-fitting and forcing the network to learn instead of memorizing the data. Finally, the output is obtained in an FC output layer with  $N_{f3}$  nodes. Subsequent to each CON layer, a batch normalization (BN) layer is incorporated to accelerate the training by reducing internal covariate shift [40], then succeeded by an activation (ACT) layer. Likewise, each FC layer is also succeeded by an ACT layer. The kernel of each CON layer is  $\mathcal{K} \in \mathbb{R}^{\kappa \times \kappa \times N_{c(i-1)}}$  with  $\kappa = 3$  and  $N_{c0} = 2$ , and the convolution is performed with stride  $\delta = 1$  and valid padding to maintain the edge information of  $\bar{\mathbf{Y}}$  while reducing the data dimension. Hence, the nonlinear mapping process of the designed CNN can be mathematically expressed as

$$\mathcal{F}(\bar{\mathbf{Y}}) = \mathcal{F}_{f3}(\mathcal{F}_{f2}(\mathcal{F}_{f1}(\mathcal{F}_{c4}(\mathcal{F}_{c3}(\mathcal{F}_{c2}(\mathcal{F}_{c1}(\bar{\mathbf{Y}}))))))) = \mathbf{q} \quad (8)$$

where  $\mathcal{F}_{ci}(\cdot)$  represents the nonlinear function of  $i$ th CON layer,  $\mathcal{F}_{fj}(\cdot)$  denotes the nonlinear function of  $j$ th FC layer, and  $\mathcal{F}_{f3}(\cdot)$  is the nonlinear function of the output layer. The output vector of CNN is denoted by  $\mathbf{q}$ .

During convolutions, the features of the input data will be exacted, and  $N_{ci}$  channels of data will be obtained at each convolution, the data at  $n_{ci}$ th channel after  $i$ th convolution are a 2-D matrix  $\mathbf{G}^{[i, n_{ci}]} \in \mathbb{R}^{\tilde{M}^{[i]} \times \tilde{M}^{[i]}}$  which can be mathematically expressed as [34]

$$\mathbf{G}_{m,n}^{[i, n_{ci}]} = f_{act}^{[ci]} \left\{ \mathcal{B} \left\{ (\bar{\mathbf{Y}}^{[i-1]} \otimes \mathcal{K}^{[i, n_{ci}]})_{m,n} \right\} \right\} \quad (9)$$

where  $i = 1, 2, 3, 4$ ,  $f_{act}^{[ci]}$  stands for the ACT function for  $i$ th CON layer,  $\mathcal{B}\{\cdot\}$  is the BN operator, and the convolution process can be expressed as

$$\begin{aligned} & (\bar{\mathbf{Y}}^{[i-1]} \otimes \mathcal{K}^{[i, n_{ci}]})_{m,n} \\ &= \sum_{i=1}^3 \sum_{j=1}^3 \sum_{t=1}^{N_{c(i-1)}} \mathcal{K}_{i,j,t}^{[i, n_{ci}]} \bar{\mathbf{Y}}_{(m-1)+i, (n-1)+j, t}^{[i-1]} + b_{CON}^{[i, n_{ci}]} \end{aligned} \quad (10)$$

where  $m, n = 1, 2, \dots, \tilde{M}^{[i]}$  with  $\tilde{M}^{[i]} = M^{[i-1]} - 2$  and  $\tilde{M}^{[0]} = M$ . The output of  $(i-1)$ th CON layer is denoted by  $\bar{\mathbf{Y}}^{[i-1]}$  with  $\bar{\mathbf{Y}}^{[0]} = \bar{\mathbf{Y}}$ , and the kernel for  $n_{ci}$ th channel at  $i$ -th CON layer is represented by  $\mathcal{K}^{[i, n_{ci}]}$  with  $n_{ci} = 1, 2, \dots, N_{ci}$ . The bias for  $n_{ci}$ th channel at  $i$ th CON layer is denoted by  $b_{CON}^{[i, n_{ci}]}$ . After four convolutions, the output of the fourth CON

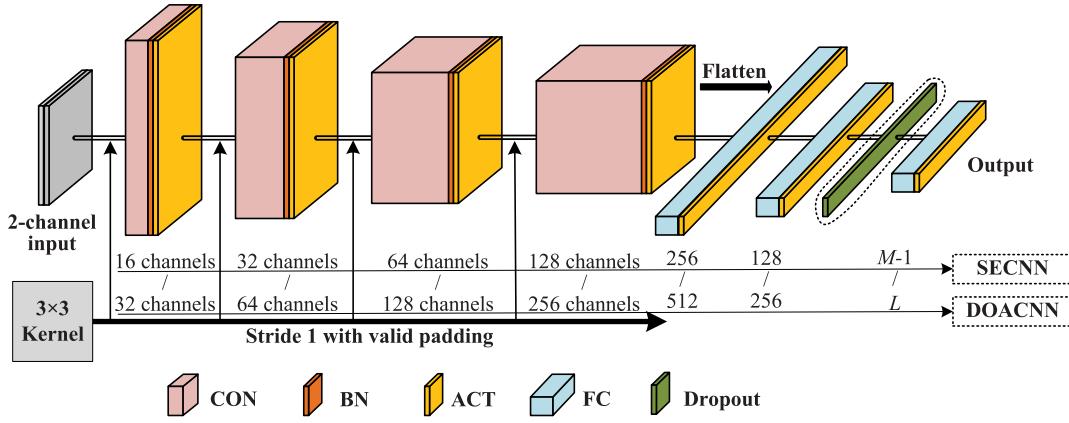


Fig. 3. CNN within the designed framework.

layer is flattened to a vector  $v$  to perform nonlinear mapping by FC layers. The output of  $j$ th FC layer can be expressed as

$$v^{[j]} = f_{act}^{[f_j]} \left\{ W^{[j]} v^{[j-1]} + b_{FC}^{[j]} \right\} \quad (11)$$

where  $j = 1, 2, 3$ ,  $v^{[0]} = v$ ,  $v^{[3]} = q$ ;  $f_{act}^{[f_j]}$  denotes the ACT function for  $j$ th FC layer;  $W^{[j]}$  and  $b_{FC}^{[j]}$  represent the weight matrix and bias vector of the  $j$ th FC layer, respectively. To this end, the output of the CNN can be obtained as

$$q = \mathcal{F}(\bar{Y}) = [q_1, q_2, \dots, q_{N_f}]^T. \quad (12)$$

In the designed CNN-based framework, since the SECNN and DOACNN are for totally different tasks, their output dimension and specific network parameters are slightly different, which is detailed and introduced in Subsections III-A and III-B.

#### A. SECNN

The specific network parameters of SECNN are given in Table I. The rectified linear unit (ReLU), which is  $f_{ReLU}(z) = \max(0, z)$ , is adopted as the ACTs for the CON layers #1 ~ #4 and FC layers #1 ~ #2. For a ULA with  $M$  antennas, the maximum number of sources it can estimate is  $M - 1$ . Therefore, the output of SECNN is a  $(M - 1)$ -dimensional vector. On the other hand, the output layer of the SECNN is followed by a Softmax ACT since it is designed to enumerate sources from  $\bar{Y}$  by multiclass classification task. The Softmax function is expressed as

$$\bar{q}_m = f_{Softmax}(q_m) = \frac{e^{q_m}}{\sum_{m=1}^{M-1} e^{q_m}} \quad (13)$$

where  $q_m$  denotes the  $m$ th entry of the output vector  $q$ . Hence, the output vector of SECNN can be expressed as  $\bar{q} = [\bar{q}_1, \bar{q}_2, \dots, \bar{q}_{M-1}]^T$  with  $\bar{q}_k$  denoting the corresponding probability for the source number is  $k = 1, 2, \dots, M - 1$ . Therefore, the source number  $K$  can be determined eventually by the index corresponding to the max probability.

#### B. DOACNN

The DOACNN is designed to realize end-to-end DOA estimation by regression task. Since the task of DOACNN

TABLE I  
SPECIFIC NETWORK PARAMETERS OF SECNN

Layers	Parameters	Activations
COV #1	$N_{c1} = 16$	$f_{act}^{[c1]} = \text{ReLU}$
COV #2	$N_{c2} = 32$	$f_{act}^{[c2]} = \text{ReLU}$
COV #3	$N_{c3} = 64$	$f_{act}^{[c3]} = \text{ReLU}$
COV #4	$N_{c4} = 128$	$f_{act}^{[c4]} = \text{ReLU}$
FC #1	$N_{f1} = 256$	$f_{act}^{[f1]} = \text{ReLU}$
FC #2	$N_{f2} = 128$	$f_{act}^{[f2]} = \text{ReLU}$
FC #3 (Output)	$N_{f3} = M - 1$	$f_{act}^{[f3]} = \text{Softmax}$

is more complicated than that of SECNN, and it is expected to enable high-resolution off-grid DOA estimation, its number of channels and neurons of CON and FC layers are more than that in SECNN, which is specifically shown in Table II. Besides, to enhance the fitting ability of the DOACNN, the ACTs following CON layers and FC layers are turned to parametric ReLU (PReLU), which is

$$f_{PReLU}(z) = \begin{cases} az, & \text{for } z \leq 0 \\ z, & \text{for } z > 0 \end{cases} \quad (14)$$

where  $z$  represents the output of each layer before activated and  $a$  is a trainable parameter during the training phase. The output of DOACNN is a  $L$ -dimensional  $K$ -sparse vector  $\tilde{q} = [\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_L]^T$ , where  $L$  denotes the total number of spatial grid points. Currently, there are two ways of modeling  $\tilde{q}$ . One of them is to model the non-zero elements in  $\tilde{q}$  as signal power and then achieve end-to-end DOA estimation through regression task [33]. The other one is to model the non-zero elements as 1 and realize end-to-end DOA estimation by multilabel classification [34]. However, both of these two ways are inability to overcome the off-grid error caused by grid mismatch, hence their performance is restricted.

In the proposed DOACNN, the DOA estimation is also realized by regression task, therefore the output layer is followed by a Linear ACT, i.e.,  $\tilde{q} = f_{\text{Linear}}(q) = q$ . While to overcome off-grid errors, the nonzero elements in  $\tilde{q}$  are modeled as being related to off-grid error. Suppose that the preset uniform spatial grid points is  $\vartheta = \{\vartheta_1, \vartheta_2, \dots, \vartheta_L\}$ ,

TABLE II  
SPECIFIC NETWORK PARAMETERS OF DOACNN

Layers	Parameters	Activations
COV #1	$N_{c1} = 32$	$f_{act}^{[c1]} = \text{PReLU}$
COV #2	$N_{c2} = 64$	$f_{act}^{[c2]} = \text{PReLU}$
COV #3	$N_{c3} = 128$	$f_{act}^{[c3]} = \text{PReLU}$
COV #4	$N_{c4} = 256$	$f_{act}^{[c4]} = \text{PReLU}$
FC #1	$N_{f1} = 512$	$f_{act}^{[f1]} = \text{PReLU}$
FC #2	$N_{f2} = 256$	$f_{act}^{[f2]} = \text{PReLU}$
FC #3 (Output)	$N_{f3} = L$	$f_{act}^{[f3]} = \text{Linear}$

and the grid interval is denoted by  $r = \vartheta_{l+1} - \vartheta_l$  with  $l = 1, 2, \dots, L - 1$ . Then, we can intuitively model the nonzero elements in  $\tilde{\mathbf{q}}$  as  $\tilde{q}_{l_k} = \theta_k - \vartheta_{l_k} \in (-(r/2), (r/2))$ , where  $\vartheta_{l_k}$  stands for the grid point nearest to  $\theta_k$ . Nevertheless, in instances where the actual DOAs are very close to the preset grid points, the above modeling may lead to the sparsity of  $\tilde{\mathbf{q}}$  disappear, so as to cause ambiguity during network training. Therefore, to guarantee the sparsity of  $\tilde{\mathbf{q}}$ , its nonzero elements are further modeled as

$$\tilde{q}_{l_k} = \left( \theta_k - \vartheta_{l_k} + \frac{r}{2} \right) \times c \in [1, rc] \quad (15)$$

where  $c$  is a scalar constant. Based on (15), the end-to-end DOA estimation can be ultimately achieved by leveraging the sparsity of  $\tilde{\mathbf{q}}$  and the corresponding value of nonzero elements.

#### IV. TRAINING DATASETS AND STRATEGIES

To realize off-grid DOA estimation in the presence of MC, the SECNN, and DOACNN should be properly trained offline first. The efficacy of these models is significantly influenced by two pivotal factors: the composition of the training dataset and the adopted training strategy. Herein, the training datasets and strategies for SECNN and DOACNN are respectively introduced. During generating training datasets, consider that the spatial range where sources are located is  $\Theta = [-\Phi, \Phi]$ .

##### A. SECNN Training

For the generation of the training dataset for SECNN, consider all possible scenarios of  $K \in \{1, 2, \dots, M - 1\}$ . First, the angle pairs for each scenario of  $K$  is randomly generated from  $\Theta$  for  $1.5 \times 10^4$  times by accurating to  $0.01^\circ$ , while ensuring that the spacing between any two angles in one angle pair is greater than  $1^\circ$ . Then, for each pair of angles, the training input  $\bar{\mathbf{Y}}_{\text{SE}}$  for SECNN is calculated by (6) and (7) with  $\text{SNR} = \{-10, -5, 0, 5, 10, 15\}\text{dB}$  and  $T = 10^4$  to ensure the generalization for lower  $T$ . Therefore, the total size of the training dataset for SECNN is  $\mathcal{D}_{\text{SE}} = (M - 1) \times 6 \times 1.5 \times 10^4$ . For  $d$ th training data, its corresponding label is  $\tilde{\mathbf{q}}^{(d)} = [\tilde{q}_1^{(d)}, \tilde{q}_2^{(d)}, \dots, \tilde{q}_{M-1}^{(d)}]^T$  whose  $K$ th element equals to 1 otherwise 0.

While training SECNN, the training dataset undergoes a random partition, allocating 90% for training purposes and reserving 10% for validation. The training is executed 100 epochs with 128 batch-size, and the Adam optimizer [41] with an initial learning rate 0.001 is utilized for optimizing the

trainable parameters. Since SECNN is to perform multiclass classification task, its loss function is set as categorical cross-entropy, which is

$$\mathcal{L}_{\text{SE}} = -\frac{1}{\mathcal{D}_{\text{SE}}} \left\{ \sum_{d=1}^{\mathcal{D}_{\text{SE}}} \left\{ \sum_{j=1}^{M-1} \tilde{q}_j^{(d)} \ln \left\{ \ddot{q}_j^{(d)} \right\} \right\} \right\} \quad (16)$$

where  $\tilde{q}_j^{(d)}$  denotes the  $j$ th element of ground-truth label  $\tilde{\mathbf{q}}^{(d)}$  corresponding to the  $d$ th data, and  $\ddot{q}_j^{(d)}$  stands for the  $j$ th element of SECNN output  $\ddot{\mathbf{q}}$  by feeding  $d$ th data. To ensure that the training converges to the optimal point, the validation accuracy is monitored to dynamically adjust the learning rate during training. If the accuracy of validation is not reduced in ten contiguous epochs, the learning rate will be reduced by 70%.

##### B. DOACNN Training

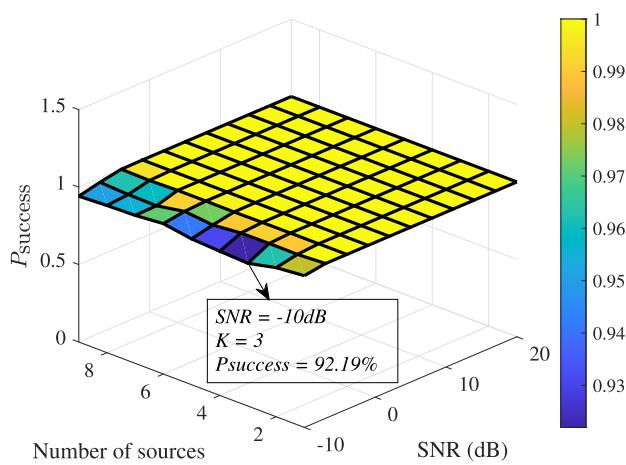
The training dataset for DOACNN is separately generated for each scenario of  $K \in \{1, 2, \dots, M - 1\}$ . Then the DOACNN for each scenario is trained by the corresponding training dataset. In generating training dataset for each scenario of  $K$ , the spatial range  $\Theta = [-\Phi, \Phi]$  is first uniformly discretized by grid interval  $r$  to obtain a spatial grid  $\vartheta = \{-\Phi, \Phi + r, \dots, \Phi - r, \Phi\}$ . For the certain scenario of  $K$ , all possible on-grid angle combinations are first generated from  $\vartheta$ . Then, the off-grid errors for each on-grid angle pair are randomly generated from  $(-r/2, r/2)$  by reserving to two decimal places to generate  $[3 \times 10^5 / \mathcal{C}_L^K]$  off-grid angle pairs. The training input  $\bar{\mathbf{Y}}_{\text{DOA}}$  for DOACNN is also calculated for each  $\text{SNR} = \{-10, -5, 0, 5, 10, 15\}\text{dB}$  and fixed  $T = 10^4$ , which leads to a training dataset size of  $\mathcal{D}_{\text{DOA}} = 6 \times [3 \times 10^5 / \mathcal{C}_L^K]$ . An appropriate  $c = 1/\tau$  is set to generate the ground-truth label of  $d$ th training data  $\tilde{\mathbf{q}}^{(d)} = [\tilde{q}_1^{(d)}, \tilde{q}_2^{(d)}, \dots, \tilde{q}_L^{(d)}]^T$  based on (15).

The DOACNN is trained for 500 epochs with a batch size of 256 to ensure convergence. Other training parameters are the same as that of SECNN except for the loss function. Since DOACNN is designed for regression tasks, introducing mean-squared error (MSE) as its loss function, which is

$$\mathcal{L}_{\text{DOA}} = \frac{1}{\mathcal{D}_{\text{DOA}}} \left\{ \sum_{d=1}^{\mathcal{D}_{\text{DOA}}} \left\{ \frac{1}{L} \sum_{l=1}^L \left| \tilde{q}_j^{(d)} - \ddot{q}_j^{(d)} \right|^2 \right\} \right\} \quad (17)$$

where  $\tilde{q}_j^{(d)}$  denotes the  $j$ th element of ground-truth label  $\tilde{\mathbf{q}}^{(d)}$  corresponding to the  $d$ th data, and  $\ddot{q}_j^{(d)}$  represents the  $j$ th element of DOACNN output  $\ddot{\mathbf{q}}^{(d)}$  by feeding  $d$ th data. During training, the validation MSE is monitored to adjust the learning rate, and the learning reduction strategy is the same as that of training SECNN.

*Remark 1:* When the number of sources increases, the total possible combinations of sources will increase dramatically hence leading to a huge amount of training dataset, which could make training a challenge. To relieve the required time for training in the case of a large number of sources, the training strategy is slightly different. Specifically, the DOACNN for  $K > 2$  is obtained by retraining the DOACNN



**Fig. 4.** Probability of successful enumeration  $P_{\text{success}}$  of SECNN under different SNRs and number of sources.

for  $K = 2$  using the training dataset in the case of  $K > 2$ . In the case of  $K > 2$ , the off-grid errors for all possible on-grid angle combinations are generated once to obtain the corresponding off-grid angle pairs, then randomly choose  $3 \times 10^5$  pairs of off-grid angle from which to calculate the two-channel covariance as the training dataset. Therefore, the total size of the training dataset can be limited to  $3 \times 10^5 \times 6$ , which can effectively reduce the training time and the requirement of hardware in practical implementation meanwhile maintaining the performance of DOACNN in different cases of  $K$ .

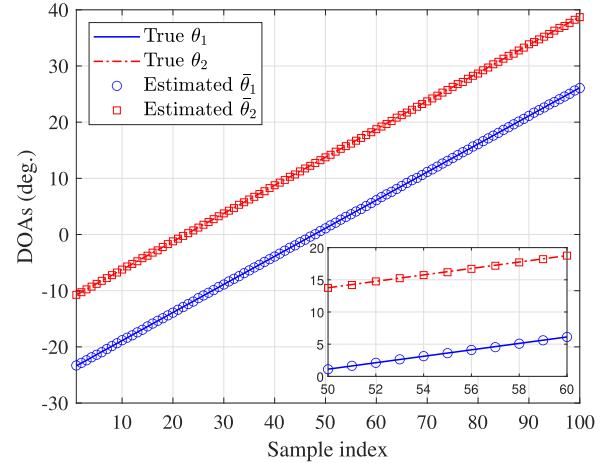
## V. NUMERICAL SIMULATIONS

### A. Simulation Setup

During both the training and testing phases, the number of array antennas and sources are considered as  $M = 10$  and  $K = 2$ , respectively. Since high-resolution DOA estimation is the concern, we set  $\Phi = 60^\circ$  and  $r = 1^\circ$  without otherwise stated, therefore the output dimension of DOACNN is  $L = 121$ . Additionally, to model the sensor MC reasonably, the MC coefficient vector  $\mathbf{h}$  is set according to [36], i.e.,  $\mathbf{h} = [1, \beta^1, \beta^2]$ , where  $\beta = \rho e^{j\pi/3}$ , and  $\rho \in (0, 1)$  denotes the MC coefficient. Unless otherwise specified, the MC coefficients  $\rho$  of a certain array system are assumed to be  $\rho = 0.6$  both in the training and testing phases. The CNNs within the designed framework are trained by utilizing a Windows platform with a 3.2 GHz CPU of Intel Core i9-10920X, 12 GB GPU of NVIDIA GeForce RTX 3090, and 64 GB RAM.

### B. Simulation Results

**1) Effectiveness of SECNN:** In the proposed framework, the effectiveness of DOACNN strongly depends on the knowledge of source numbers provided by the SECNN. Therefore, the effectiveness of SECNN is first evaluated by its probability of successful enumeration  $P_{\text{success}}$  under different SNRs and the number of sources. In each scenario of SNR and the number of sources, the simulation is independently conducted 300 times. To ensure the generality of the results, the ground-truth DOAs

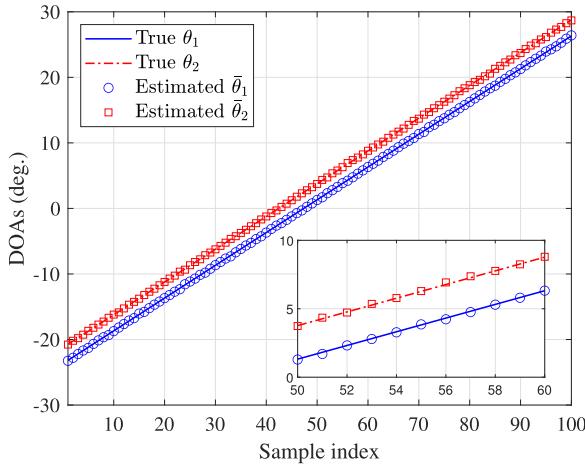


**Fig. 5.** DOA estimation results of DOACNN with well-separated sources.

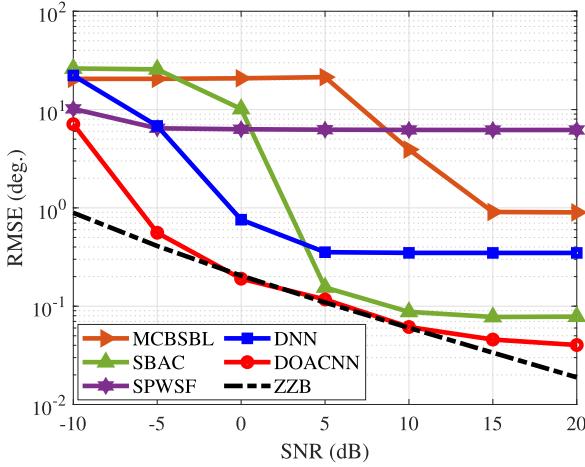
of sources for testing are randomly generated from  $[-60^\circ, 60^\circ]$  with a resolution of  $0.01^\circ$ , and the number of snapshots is fixed at  $T = 500$ . In this case, the corresponding results are depicted in Fig. 4. Observation of the Fig. 4 reveals that the SECNN can achieve effective source number enumeration, and the success probability reaches more than 92% at all cases of SNRs and the number of sources, which means the SECNN can provide sufficiently accurate priori information of sources number thus guarantee the effectiveness and generalization of the subsequent DOACNN.

**2) Effectiveness of DOACNN:** Based on the determined source numbers provided by SECNN, a certain DOACNN for the scenario of  $K$  sources can be called directly from the database for DOA estimation. Herein, we set  $K = 2$  to test the effectiveness of the DOACNN under both well-separated and closely spaced source conditions. In the case of well-separated sources, the initial true DOAs are  $\theta_1 = -23.38^\circ$  and  $\theta_2 = -10.74^\circ$ , which means that the angular interval is  $\Delta\theta = 12.64^\circ$ . While, in the case of closely spaced sources, the initial true DOAs of sources are fixed as  $\theta_1 = -23.18^\circ$  and  $\theta_2 = -20.73^\circ$ , i.e., the angular interval is  $\Delta\theta = 2.45^\circ$ . To test the effectiveness of the DOACNN in different off-grid DOA situations, the ground-truth DOAs are sampled 100 times by increasing the step of  $0.5^\circ$  based on the initial DOAs. The corresponding results are given in Figs. 5 and 6 with  $\text{SNR} = 5 \text{ dB}$  and  $T = 500$ , respectively. As can be seen from the results the proposed DOACNN is effective in achieving off-grid DOA estimation both at well-separated and closely spaced sources, and the precision is obviously very high. On the other hand, the results for closely spaced sources indicate that the proposed DOACNN can effectively achieve high-resolution DOA estimation with high precision.

**3) Statistic Performance Comparison:** To further evaluate the performance of the proposed DOACNN in terms of high-resolution DOA estimation, the statistical performance of different methods is compared in the case of close-spaced sources to demonstrate the performance superiority of the proposed DOACNN. Some representative approaches including the state-of-the-art MCBSBL [3], SPWSF [22], SBAC [23], DNN [36] methods are introduced for



**Fig. 6.** DOA estimation results of DOACNN with closely spaced sources.



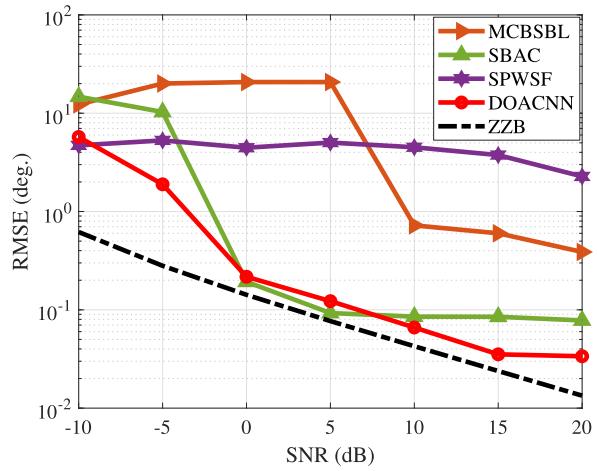
**Fig. 7.** RMSE of different methods under different SNRs with  $K = 2$ .

performance comparison, and the state-of-the-art Ziv-Zakai Bound (ZZB) [42] is introduced as the performance metric. The root-mean-squared-error (RMSE) is introduced as the performance evaluation metric, which is defined as [22]

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{MC}}} \frac{1}{K} \sum_{n_{\text{MC}}=1}^{N_{\text{MC}}} \sum_{k=1}^K (\theta_k - \bar{\theta}_k^{[n_{\text{MC}}]})^2} \quad (18)$$

where  $N_{\text{MC}} = 100$  is the total number of Monte Carlo simulations,  $\theta_k$  denotes the ground-truth DOA of  $k$ th source, and  $\bar{\theta}_k^{[n_{\text{MC}}]}$  stands for the estimated DOA of  $k$ th source at  $n_{\text{MC}}$ th Monte Carlo simulation.

First of all, the RMSE of different methods under different SNRs is compared with  $K = 2$  and the corresponding ground-truth DOAs of sources being  $\boldsymbol{\theta} = [-1.53^\circ, 1.34^\circ]$ . The corresponding results are given in Fig. 7, in which the number of snapshots is set as  $T = 500$ . As clearly illustrated in Fig. 7 that the RMSE of MCB-SBL does not decrease with SNR increasing when  $\text{SNR} \leq 5 \text{ dB}$ , while its performance does not improve further when  $\text{SNR} \geq 15 \text{ dB}$  due to its insufficient resolution. The RMSE of the SPWSF method is greater than  $5^\circ$  in all SNR cases, which indicates its insufficient resolution



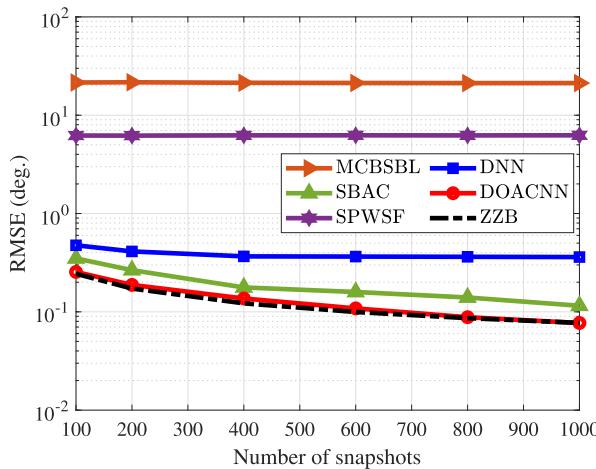
**Fig. 8.** RMSE of different methods under different SNRs with  $K = 4$ .

of DOA estimation. Conversely, the SBAC and DNN have enough resolution to enable them to work well with closely spaced sources when  $\text{SNR} > 0 \text{ dB}$ . However, the proposed DOACNN enjoys the lowest RMSE at all SNR conditions as compared to its rivals, and its RMSE is much closer to the CRB. This result effectively indicates the high-resolution off-grid DOA estimation performance advantage of the proposed method in the case of different SNRs. However, it should be noted that the DOACNN (so does the DNN) exists in the so-called “floor effect,” i.e., the RMSE becomes flattened with the SNR increasing, which is reasonable. Because a NN trained using a certain dataset is usually a biased estimator, which holds for all DL-based estimators [31], [43], not only the proposed network.<sup>1</sup>

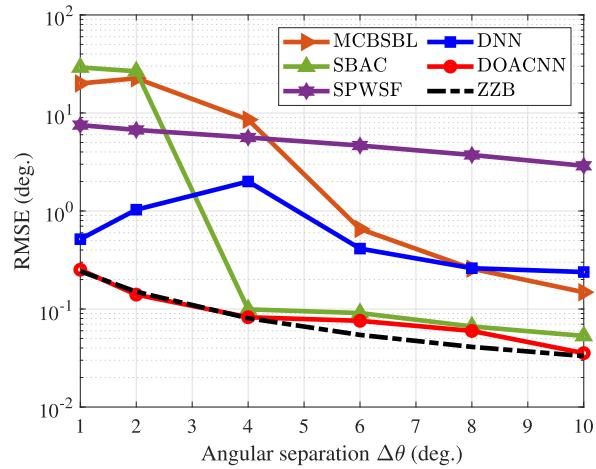
To show the effectiveness and performance superiority of the proposed DOACNN in the case of  $K > 2$ , the RMSE results of different methods under different SNRs with  $K = 4$  are compared in Fig. 8, where the ground-truth DOAs of sources are  $\boldsymbol{\theta} = [-18.15^\circ, -14.13^\circ, 14.22^\circ, 31.64^\circ]$ , and the other simulation parameters are the same as that in Fig. 7. It should be noted that the DNN proposed in [36] is not compared here for a fair comparison, since the training strategy given in [36] is blurry and is applicable only for  $K = 2$ , and the corresponding strategy for  $K > 2$  is not given. As can be seen from Fig. 8, the proposed DOACNN still shows an obvious performance advantage as compared to its rivals.

Afterward, the simulations on RMSE of different methods under different numbers of snapshots are conducted to further illustrate the superiority of the proposed DOACNN in terms of high resolution. The corresponding result is shown in Fig. 9 where the SNR is set as 5 dB for all methods. From

<sup>1</sup>The performance of a neural network is affected by several factors such as the network architecture and the size of the training dataset. Due to the data-driven characteristics of neural networks, the predictions cannot perfectly fit the ground-truth values with a fixed network structure and training dataset. When the training procedure is once done, there must be an error between its predictions and labels (especially for the neural networks based on regression task), resulting in the upper limit of its performance. When the model reaches this upper limit as SNR increases, it is difficult to show further performance improvements. Certainly, the “floor effect” of a neural network can be effectively “delayed” by deeper structure and larger training dataset, but it cannot be completely eliminated.



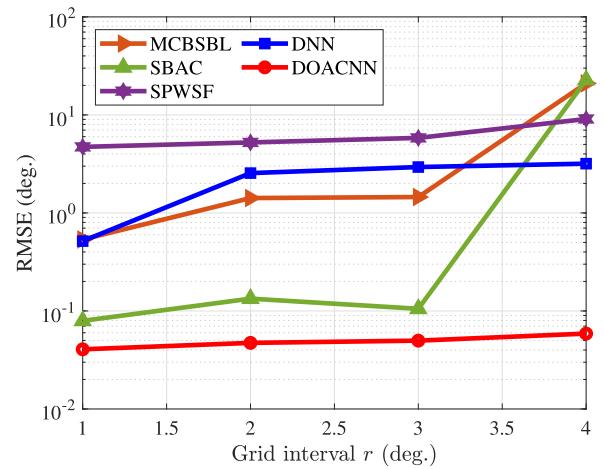
**Fig. 9.** RMSE of different methods under different number of snapshots with  $K = 2$ .



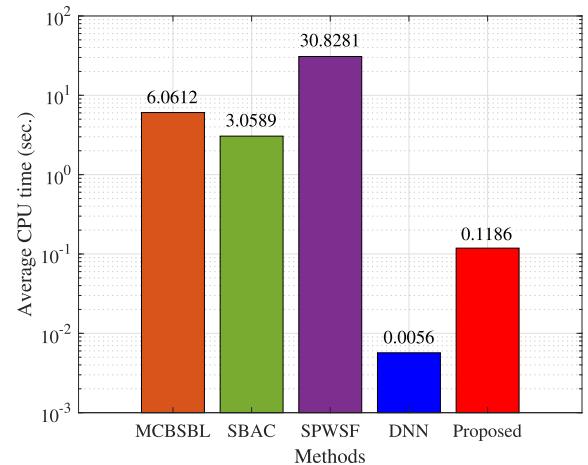
**Fig. 10.** RMSE of different methods under different angular separations with  $K = 2$ .

**Fig. 9**, one can find that the RMSEs of the MCBBSBL and SPWSF methods do not decrease as the number of snapshots increased due to their limitation in terms of angular resolution. Instead, the SBAC, DNN, and proposed DOACNN have enough resolution to enable them to effectively distinguish closely spaced sources in all cases of the number of snapshots. While, the proposed DOACNN exhibits a much smaller RMSE than SBAC and DNN at all numbers of snapshots, which further demonstrates its high-resolution off-grid DOA estimation performance advantage under different numbers of snapshots.

Further, the RMSE of different methods under different angular separations  $\Delta\theta$  are compared to fully illustrate the performance advantages of the proposed DOACNN in terms of resolution. During simulation, the ground-truth DOA of the first source is fixed at  $\theta_1 = -3.61$ , then the true DOA of the second source is set as  $\theta_2 = \theta_1 + \Delta\theta$  with  $\Delta\theta = [1^\circ, 2^\circ, 4^\circ, 6^\circ, 8^\circ, 10^\circ]$ . The result with SNR = 5 dB and  $T = 500$  is depicted in **Fig. 10**. As the result demonstrated, the RMSE of different methods decreases with increasing angular interval, while the proposed DOACNN exhibits consistently lower RMSE than other comparison



**Fig. 11.** RMSE of different methods under different grid intervals with  $K = 2$ .



**Fig. 12.** CPU time required by a single estimation of different methods with  $K = 2$ .

methods across diverse angular separations, especially at small angular separations. The result in **Fig. 10** fully illustrates the performance advantage of the proposed DOACNN for high-resolution off-grid DOA estimation.

Lastly, to show the effectiveness and superiority of the proposed DOACNN in different grid intervals, the results of RMSE of different methods under different grid intervals  $r$  are given in **Fig. 11**. The results are based on that SNR = 5 dB,  $T = 500$  and the ground-truth DOA is  $\theta = [-3.21^\circ, 2.65^\circ]$ . By observing **Fig. 11**, it can be found that the RMSE of the MCBBSBL, SBAC, SPWSF, and DNN methods are all increasing with the increase of grid interval, and the MCBBSBL and SBAC are even failing under the condition of coarse grid interval. On the contrary, the proposed DOACNN not only has a lower RMSE than its rivals for all grid interval conditions, but also its RMSE remains relatively stable as the grid interval increases. This result demonstrates that the proposed DOACNN can achieve effective high-resolution off-grid DOA estimation under different grid intervals while having obvious performance advantages.

**4) Efficiency Comparison:** As claimed in Section I, the DL-based methods are much more time-saving as compared to the traditional model-driven methods. Therefore, the average CPU time required by a single estimation of different methods is finally compared to show the efficiency superiority of the proposed CNN framework. The associated outcomes are illustrated in Fig. 12, which is based on 100 independent simulations and the ground-truth DOAs are  $\theta_1 = -1.53^\circ$  and  $\theta_2 = 1.34^\circ$ . In addition, the SNR and number of snapshots are respectively set as  $\text{SNR} = 5 \text{ dB}$  and  $T = 500$ . Evident from Fig. 12, the CPU time demanded by DL-based methods, including DNN in [36] and the presented CNN framework, is much lower as compared to the alternative model-driven methods. On the other hand, it can be found that the efficiency of the proposed CNN framework is inferior to that of the DNN in [36], which is reasonable since the proposed CNN framework has much more parameters to ensure the precision and resolution of DOA estimation. Nevertheless, despite this, the proposed CNN framework is still able to meet near real-time DOA estimation.

## VI. CONCLUSION

In this article, an NN framework consisting of CNNs is designed for effective high-resolution off-grid DOA estimation in the presence of MC. The designed framework consists of SECNN and DOACNN and is trained by the practical array covariance of certain arrays, which enables it can realize DOA estimation under MC conditions while providing the knowledge of the source number. By modeling off-grid error into labels, the proposed DOACNN in the proposed framework can effectively against grid mismatch so as to realize high-precision off-grid DOA estimation. In addition, the proposed DOACNN exhibits significant performance advantages in terms of high-resolution DOA estimation. Adequate simulation results substantiate the superior performance characteristics of the presented framework, affirming its performance advantage in achieving nearly real-time high-resolution off-grid DOA estimation.

Currently, the effectiveness and superiority of the proposed network have been effectively verified by synthetic data simulations. To make the proposed method more suitable for practical situations, we believe that using experimental data from a physical array system to train the proposed network is a very effective way. But, unfortunately, its practical performance can not be effectively evaluated by experimental data at present, since the limitations in experimental equipment and dataset acquisition, which also are likewise prevalent challenges for current DL-based DOA estimation methods under non-ideal conditions (including MC, non-Gaussian noise, amplitude-phase errors, etc.) to be applied in practice. Therefore, the above limitations will be the focus of our future work, including building a prototype physical array system with MC, dataset collection, and experimental dataset training/testing. Although the proposed network has limitations at present, this does not detract from its theoretical significance and potential utility in the future.

## REFERENCES

- [1] M. Pesavento, M. Trinh-Hoang, and M. Viberg, "Three more decades in array signal processing research: An optimization and structure exploitation perspective," *IEEE Signal Process. Mag.*, vol. 40, no. 4, pp. 92–106, Jun. 2023.
- [2] W. Liu, M. Haardt, M. S. Greco, C. F. Mecklenbräuker, and P. Willett, "Twenty-five years of sensor array and multichannel signal processing: A review of progress to date and potential research directions," *IEEE Signal Process. Mag.*, vol. 40, no. 4, pp. 80–91, Jun. 2023.
- [3] H. Wang, X. Wang, X. Lan, T. Su, and L. Wan, "BSBL-based auxiliary vehicle position analysis in smart city using distributed MEC and UAV-deployed IoT," *IEEE Internet Things J.*, vol. 10, no. 2, pp. 975–986, Jan. 2023.
- [4] J. Cong, X. Wang, C. Yan, L. T. Yang, M. Dong, and K. Ota, "CRB weighted source localization method based on deep neural networks in multi-UAV network," *IEEE Internet Things J.*, vol. 10, no. 7, pp. 5747–5759, Apr. 2023.
- [5] X. Wang, Y. Guo, F. Wen, J. He, and T.-K. Truong, "EMVS-MIMO radar with sparse Rx geometry: Tensor modeling and 2-D direction finding," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 8062–8075, Dec. 2023.
- [6] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.
- [7] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
- [8] B. D. Rao and K. V. S. Hari, "Performance analysis of root-music," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.
- [9] M. Haardt and J. A. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1232–1242, May 1995.
- [10] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [11] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
- [12] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.
- [13] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, "An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 1–11, Oct. 2012.
- [14] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.
- [15] Y. Zhang, Z. Ye, X. Xu, and N. Hu, "Off-grid DOA estimation using array covariance matrix and block-sparse Bayesian learning," *Signal Process.*, vol. 98, pp. 197–201, May 2014.
- [16] J. Dai, X. Bao, W. Xu, and C. Chang, "Root sparse Bayesian learning for off-grid DOA estimation," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 46–50, Jan. 2017.
- [17] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 273–284, Mar. 1991.
- [18] B. Liao, Z.-G. Zhang, and S.-C. Chan, "DOA estimation and tracking of ULAs with mutual coupling," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 891–905, Jan. 2012.
- [19] J. Dai, D. Zhao, and X. Ji, "A sparse representation method for DOA estimation with unknown mutual coupling," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 1210–1213, 2012.
- [20] Q. Wang, T. Dou, H. Chen, W. Yan, and W. Liu, "Effective block sparse representation algorithm for DOA estimation with unknown mutual coupling," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2622–2625, Dec. 2017.
- [21] X. Wang, D. Meng, M. Huang, and L. Wan, "Reweighted regularized sparse recovery for DOA estimation with unknown mutual coupling," *IEEE Commun. Lett.*, vol. 23, no. 2, pp. 290–293, Feb. 2019.
- [22] D. Meng, X. Wang, M. Huang, L. Wan, and B. Zhang, "Robust weighted subspace fitting for DOA estimation via block sparse recovery," *IEEE Commun. Lett.*, vol. 24, no. 3, pp. 563–567, Mar. 2020.

- [23] Z.-M. Liu and Y.-Y. Zhou, "A unified framework and sparse Bayesian perspective for direction-of-arrival estimation in the presence of array imperfections," *IEEE Trans. Signal Process.*, vol. 61, no. 15, pp. 3786–3798, Aug. 2013.
- [24] P. Chen, Z. Cao, Z. Chen, and C. Yu, "Sparse off-grid DOA estimation method with unknown mutual coupling effect," *Digit. Signal Process.*, vol. 90, pp. 1–9, Jul. 2019.
- [25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [26] S. Ge, K. Li, and S. N. B. M. Rum, "Deep learning approach in DOA estimation: A systematic literature review," *Mobile Inf. Syst.*, vol. 2021, pp. 1–14, Sep. 2021.
- [27] J. Fuchs, M. Gardill, M. Lübke, A. Dubey, and F. Lurz, "A machine learning perspective on automotive radar direction of arrival estimation," *IEEE Access*, vol. 10, pp. 6775–6797, 2022.
- [28] S. Xu, Z. Wang, W. Zhang, and Z. He, "End-to-end regression neural network for coherent DOA estimation with dual-branch outputs," *IEEE Sensors J.*, vol. 24, no. 3, pp. 4047–4056, Feb. 2024.
- [29] G. K. Papageorgiou and M. Sellathurai, "Fast direction-of-arrival estimation of multiple targets using deep learning and sparse arrays," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 4632–4636.
- [30] A. Barthelme and W. Utschick, "DoA estimation using neural network-based covariance matrix reconstruction," *IEEE Signal Process. Lett.*, vol. 28, pp. 783–787, 2021.
- [31] X. Wu, X. Yang, X. Jia, and F. Tian, "A gridless DOA estimation method based on convolutional neural network with Toeplitz prior," *IEEE Signal Process. Lett.*, vol. 29, pp. 1247–1251, 2022.
- [32] G. Izacard, S. Mohan, and C. Fernandez-Granda, "Data-driven estimation of sinusoid frequencies," in *Proc. 33rd Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 5127–5137.
- [33] L. Wu, Z. Liu, and Z. 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.
- [34] G. K. Papageorgiou, M. Sellathurai, and Y. C. Eldar, "Deep networks for direction-of-arrival estimation in low SNR," *IEEE Trans. Signal Process.*, vol. 69, pp. 3714–3729, 2021.
- [35] R. Zheng, S. Sun, H. Liu, H. Chen, and J. Li, "Interpretable and efficient beamforming-based deep learning for single-snapshot DOA estimation," *IEEE Sensors J.*, vol. 24, no. 14, pp. 22096–22105, Jul. 2024.
- [36] Z.-M. Liu, C. Zhang, and S. Y. Philip, "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.
- [37] J. Cong, X. Wang, M. Huang, and L. Wan, "Robust DOA estimation method for MIMO radar via deep neural networks," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7498–7507, Mar. 2021.
- [38] Y. Ji, C. Wen, Y. Huang, J. Peng, and J. Fan, "Robust direction-of-arrival estimation approach using beamspace-based deep neural networks with array imperfections and element failure," *IET Radar, Sonar Navigat.*, vol. 16, no. 11, pp. 1761–1778, Nov. 2022.
- [39] R. Zheng, S. Sun, H. Liu, H. Chen, M. Soltanianian, and J. Li, "Antenna failure resilience: Deep learning-enabled robust DOA estimation with single snapshot sparse arrays," 2024, *arXiv:2405.02788*.
- [40] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
- [41] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
- [42] Z. Zhang, Z. Shi, and Y. Gu, "Ziv-zakai bound for DOAs estimation," *IEEE Trans. Signal Process.*, vol. 71, pp. 136–149, 2023.
- [43] Y. Huang, Y. Zhang, J. Tao, C. Wen, G. Liao, and W. Hong, "Off-grid DOA estimation via a deep learning framework," *Sci. China Inf. Sci.*, vol. 66, no. 12, Dec. 2023, Art. no. 222305.



**Huafei Wang** (Member, IEEE) was born in 1995. He received the B.S. and M.S. degrees from Hainan University, Haikou, China, in 2017 and 2020, respectively, where he is currently pursuing the Ph.D. degree in information and communication engineering.

His current research interests include array signal processing, radar signal processing using learning strategy.



**Xianpeng Wang** (Member, IEEE) was born in 1986. He received the M.S. and Ph.D. degrees from the College of Automation, Harbin Engineering University, Harbin, China, in 2012 and 2015, respectively.

He was a full-time Research Fellow at the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2015 to 2016. He is currently a Professor with the School of Information and Communication Engineering, Hainan University, Haikou, China. He is the author of more than 100 papers published in related journals and international conference proceedings and was a reviewer of more than 30 journals. His major research interests include communication system, array signal processing, radar signal processing, compressed sensing, and its applications.



**Xiang Lan** (Member, IEEE) received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China, in 2012, and the M.Sc. and Ph.D. degrees from the Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield, U.K., in 2014 and 2019, respectively.

From 2019 to June 2020, he worked as a Research Associate at the Department of Electronic and Electrical Engineering, The University of Sheffield. He is currently a Lecturer in Information and Communication Engineering, Hainan University, Haikou, China. His research interests include signal processing based on vector sensor arrays (beamforming and DOA estimation with polarized signals) and sparse array processing.



**Ting Su** (Member, IEEE) received the B.S. degree in communication engineering and the Ph.D. degree in electronic science and technology from Nanjing University of Science and Technology, Nanjing, China, in 2006 and 2016, respectively.

She is currently a Postdoctoral Researcher with the Institute of Communications Engineering, Army Engineering University of PLA, Nanjing. Moreover, she is also a Lecturer with Hainan University, Haikou, China. Her research interests include computational electromagnetic, radar signal processing, and wireless communications.