



Article

Gridless DOA Estimation Method for Arbitrary Array Geometries Based on Complex-Valued Deep Neural Networks

Yuan Cao , Tianjun Zhou and Qunfei Zhang *

School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China; chaya66@mail.nwpu.edu.cn (Y.C.); zhoutianjun@mail.nwpu.edu.cn (T.Z.)

* Correspondence: zhangqf@nwpu.edu.cn

Abstract: Gridless direction of arrival (DOA) estimation methods have garnered significant attention due to their ability to avoid grid mismatch errors, which can adversely affect the performance of high-resolution DOA estimation algorithms. However, most existing gridless methods are primarily restricted to applications involving uniform linear arrays or sparse linear arrays. In this paper, we derive the relationship between the element-domain covariance matrix and the angular-domain covariance matrix for arbitrary array geometries by expanding the steering vector using a Fourier series. Then, a deep neural network is designed to reconstruct the angular-domain covariance matrix from the sample covariance matrix and the gridless DOA estimation can be obtained by Root-MUSIC. Simulation results on arbitrary array geometries demonstrate that the proposed method outperforms existing methods like MUSIC, SPICE, and SBL in terms of resolution probability and DOA estimation accuracy, especially when the angular separation between targets is small. Additionally, the proposed method does not require any hyperparameter tuning, is robust to varying snapshot numbers, and has a lower computational complexity. Finally, real hydrophone data from the SWellEx-96 ocean experiment validates the effectiveness of the proposed method in practical underwater acoustic environments.

Keywords: gridless direction of arrival (DOA) estimation; complex-valued deep neural network; arbitrary array geometry; high resolution



Citation: Cao, Y.; Zhou, T.; Zhang, Q. Gridless DOA Estimation Method for Arbitrary Array Geometries Based on Complex-Valued Deep Neural Networks. *Remote Sens.* **2024**, *16*, 3752. <https://doi.org/10.3390/rs16193752>

Academic Editors: Xuebo Zhang, Lei Zhang, Jingyue Lu, Haiyong Xu, Marc Pinto and Farman Ali

Received: 11 September 2024

Revised: 6 October 2024

Accepted: 7 October 2024

Published: 9 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Direction of arrival (DOA) estimation is one of the most fundamental problems in array signal processing, which obtains azimuthal information about targets by sampling the signal in the spatial domain [1,2]. Conventional DOA estimation methods are mainly based on beamforming, subspace decomposition, or sparse reconstruction. Conventional beamforming (CBF) is the simplest beamforming technique and has been widely used in practice due to its very low complexity and high robustness. However, the resolution of CBF can only be enhanced by expanding the array aperture and cannot be improved by increasing the number of snapshots or the signal-to-noise ratio (SNR) [3]. To enhance the performance of CBF when the array aperture is limited, superdirectivity beamforming methods, such as Capon, have been proposed. These methods greatly enhance the resolution of small-aperture arrays, but at the cost of increased sensitivity to errors, which affects robustness [4]. Subspace decomposition is another classical class of DOA estimation methods, of which the MUSIC [5] and ESPRIT [6] algorithms are the most representative. The advantage of such methods is their excellent performance when the number of snapshots is sufficient. However, if the snapshot number is insufficient, the signal covariance matrix cannot be accurately estimated, leading to a significant decline in algorithm performance [7].

Subsequently, sparse reconstruction DOA estimation methods have been proposed based on signal sparsity. These methods are generally divided into three categories based on the signal model: on-grid, off-grid, and gridless approaches [8]. On-grid methods

discretize the continuous angular domain into a finite set of points, assuming that the actual DOAs coincide with these discrete points [9–12]. While these methods enable the direct application of compressed sensing techniques, they are susceptible to grid mismatch errors. Off-grid methods, though they estimate DOAs within the continuous angular domain, still rely on the discretized grid model [13–16]. To alleviate grid mismatch, these methods employ strategies such as first-order Taylor approximations of the steering vector near the grid [14] or adaptive grid adjustments [15]. However, their performance is contingent on the accuracy of these approximations, and they often introduce non-convex optimization problems, complicating the solution process. In contrast, gridless methods perform DOA estimation directly in the continuous angular domain without grid discretization. These approaches typically rely on techniques like Toeplitz covariance matrix reconstruction or trigonometric polynomial rooting [8,17]. Initially, gridless methods were limited to uniform linear arrays (ULAs) and sparse linear arrays (SLAs), but they have since been extended to coprime arrays through interpolation techniques [18]. Recent developments have further broadened their applicability to arbitrary array geometries by representing the steering vector as a Fourier series [19–23]. Despite these advancements, gridless methods often require semi-definite programming, which can result in high computational complexity. Additionally, in methods that use convex optimization to solve for the spatial spectrum and obtain DOA estimates, hyperparameter tuning presents a significant challenge. Improper hyperparameter settings can severely degrade the performance of these algorithms [9].

In recent years, deep learning (DL) has been extensively applied to DOA estimation tasks due to its non-linear capabilities and feature-learning potential [24]. Compared to sparse reconstruction methods, DL-based approaches offer several advantages. First, DOA estimation using a trained deep network bypasses the need for high-dimensional matrix inversion and multiple iterations, leading to a significant reduction in computational complexity. Second, the process does not require hyperparameter tuning, which helps avoid performance degradation due to improper parameter settings. Currently, most DL methods frame the DOA estimation problem as a multi-label classification task, where the deep network outputs the probability of targets being present in different directions. However, this approach is still fundamentally based on the discrete grid model [25–28]. While increasing the grid resolution can theoretically reduce grid mismatch errors, it also increases the output dimensionality of the network, making the training process more difficult. Although gridless DL-based DOA estimation methods have also been explored, similar to gridless sparse reconstruction approaches, they are typically limited to ULA, SLA [29–31], or coprime arrays [32].

In this paper, we first derive the Toeplitz structure of the angular-domain covariance matrix for arbitrary array geometries. Then, we design and train a complex-valued deep neural network (CDNN) to reconstruct this Toeplitz matrix from the sample covariance matrix (SCM) and subsequently employ the Root-MUSIC method to achieve gridless DOA estimation for arbitrary array geometries. The main contributions of this paper are as follows: (1) We represent the steering vector as a Fourier series and derive the Toeplitz structure of the angular-domain covariance matrix for arbitrary array geometries. Furthermore, we establish the relationship between this matrix and the element-domain covariance matrix, providing a theoretical foundation for gridless DOA estimation using SCM in arbitrary array geometries. (2) We propose a novel approach that utilizes CDNN to reconstruct the Toeplitz matrix from SCM and leverages Root-MUSIC for gridless DOA estimation. This method avoids grid mismatch errors, eliminates the need for hyperparameter tuning, and delivers strong performance with low computational complexity. (3) We analyze the impact of the Toeplitz matrix's dimensionality on DOA estimation performance.

The rest of this paper is organized as follows: Section 2 presents the DOA estimation problem model and derives the Toeplitz structure of the angular-domain covariance matrix from the element-domain covariance matrix. Section 3 describes the training of the CDNN for the subarray of horizontal linear array sorting (HLAS) in the Swellex-96 experiment, to reconstruct the Toeplitz structure of the angular-domain covariance matrix. In Section 4, the

performance of the proposed method is assessed through numerical simulations. Section 5 demonstrates the algorithm's effectiveness using real data obtained from sea trials. Finally, Section 6 provides the conclusions and discussions of this paper.

2. Problem Formulation

In this section, we present the mathematical model for gridless DOA estimation. Specifically, Section 2.1 introduces the general model for the DOA estimation problem in a two-dimensional (2D) plane. Section 2.2 derives the Toeplitz structure of the angular-domain covariance matrix for arbitrary array geometries in the DOA estimation problem.

2.1. Signal Model

In a 2D plane, we consider the scenario where K far-field narrowband plane waves with DOAs $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K]$ impinge on an array of M omnidirectional sensors. The mutual coupling effects between sensors are not considered. It is assumed that $K < M$ and the number of sources K is known. The received snapshots of the array can be expressed as

$$\mathbf{y}_l = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}_l + \mathbf{n}_l, \quad (1)$$

where $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)] \in \mathbb{C}^{M \times K}$ is the steering matrix of the array corresponding to the directions $\boldsymbol{\theta}$. $\mathbf{x}_l \in \mathbb{C}^K$ and $\mathbf{n}_l \in \mathbb{C}^M$ are the signal vector and noise vector at the l th snapshot, respectively, and $l = 1, 2, \dots, L$ is the snapshot index. For an array with arbitrary geometry, the steering vector can be expressed as

$$\mathbf{a}(\theta) = [e^{j2\pi f r_1 \cos(\theta - \varphi_1)/c}, e^{j2\pi f r_2 \cos(\theta - \varphi_2)/c}, \dots, e^{j2\pi f r_M \cos(\theta - \varphi_M)/c}]^T, \quad (2)$$

where the superscript T denotes the transpose operator, f is the center frequency of the narrowband signal, and (r_m, φ_m) are the polar coordinates of the m th array element, with c being the speed of waves.

The covariance matrix of the received signal in the element domain can be written as

$$\mathbf{R}_y = \mathbf{A}\mathbf{R}_x\mathbf{A}^H + \mathbf{R}_n, \quad (3)$$

where the superscript H denotes the conjugate transpose operator, and \mathbf{R}_x and \mathbf{R}_n are the covariance matrices of the source signals and noise, respectively. For uncorrelated sources, $\mathbf{R}_x = \text{diag}(\alpha_1^2, \alpha_2^2, \dots, \alpha_K^2)$, where α_k^2 is the power of the k th source. For simplicity, we abbreviate $\mathbf{A}(\boldsymbol{\theta})$ as \mathbf{A} . In practical scenarios with limited observation data, the SCM $\widehat{\mathbf{R}}_y = \frac{1}{L} \sum_{l=1}^L \mathbf{y}_l \mathbf{y}_l^H$ is often used to estimate \mathbf{R}_y . When the noise is Gaussian, $\widehat{\mathbf{R}}_y$ provides an unbiased estimate of \mathbf{R}_y .

2.2. Angular-Domain Covariance Matrix

In a 2D plane, the steering vector $\mathbf{a}(\theta)$ exhibits periodicity with respect to θ , with a period of 2π . Therefore, it can be expanded into a Fourier series in terms of θ , given by

$$[\mathbf{a}(\theta)]_m = \sum_{n=-\infty}^{\infty} b_{nm} e^{jn\theta} \approx \sum_{n=-N}^N b_{nm} e^{jn\theta}, \quad (4)$$

where $[\mathbf{a}(\theta)]_m$ represents the m th element of $\mathbf{a}(\theta)$, n is the Fourier series index, and b_{nm} is the Fourier coefficient. When $b_{Nm} \approx 0$, the Fourier series can be truncated at $n = N$, and the resulting expression in matrix form is

$$\mathbf{a}(\theta) = \mathbf{B}^H \mathbf{f}(\theta), \quad (5)$$

where $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_M] \in \mathbb{C}^{(2N+1) \times M}$ is the Fourier coefficient matrix, $\mathbf{b}_m = [b_{-Nm}, \dots, b_{0m}, \dots, b_{Nm}]^H$ is the coefficient vector, and $\mathbf{f}(\theta) = [e^{-jN\theta}, \dots, 1, \dots, e^{jN\theta}]^T$. Therefore, the steering matrix \mathbf{A} can be expressed as

$$\mathbf{A} = \mathbf{B}^H \mathbf{F}, \quad (6)$$

where $\mathbf{F} = [\mathbf{f}(\theta_1), \mathbf{f}(\theta_2), \dots, \mathbf{f}(\theta_K)]$. In the absence of noise, substituting (6) into (3) yields

$$\widehat{\mathbf{R}}_y = \mathbf{B}^H \mathbf{F} \mathbf{R}_x \mathbf{F}^H \mathbf{B} = \mathbf{B}^H \mathbf{T}(\mathbf{u}) \mathbf{B}, \quad (7)$$

where $\mathbf{T}(\mathbf{u}) = \mathbf{F} \mathbf{R}_x \mathbf{F}^H$ is a Hermitian Toeplitz matrix, and the vector $\mathbf{u} \in \mathbb{C}^{2N+1}$ represents the first row of the matrix $\mathbf{T}(\mathbf{u})$. According to [33], the Fourier coefficients $b_{nm} = j^n J_n(2\pi f r_m / c) e^{-jn\varphi_m}$, where $J_n(\bullet)$ denotes the n th-order Bessel function of the first kind. We refer to $\mathbf{T}(\mathbf{u})$ as the angular-domain covariance matrix. It is evident that $\mathbf{T}(\mathbf{u})$ has a rank of K and is positive semi-definite. According to the Vandermonde decomposition theorem for semi-definite Toeplitz matrices, \mathbf{F} and \mathbf{R}_x can be uniquely recovered from $\mathbf{T}(\mathbf{u})$, enabling gridless DOA estimation [8,34].

It is important to note that cross-correlation between signals, as well as between signals and noise, causes the SCM to deviate from the true element-domain covariance matrix [35,36]. This cross-correlation introduces inaccuracies in the signal subspace, resulting in degraded performance of subspace-based algorithms [36]. By recovering a rank-deficient Toeplitz matrix $\mathbf{T}(\mathbf{u})$, the influence of signal correlation on the signal subspace can be mitigated, thereby improving the performance of subspace algorithms. However, directly reconstructing $\mathbf{T}(\mathbf{u})$ from $\widehat{\mathbf{R}}_y$ is challenging. Therefore, we propose using a CDNN to learn the mapping between $\widehat{\mathbf{R}}_y$ and $\mathbf{T}(\mathbf{u})$, and subsequently employing the Root-MUSIC algorithm [37], which obtains DOA estimates by utilizing the roots of a polynomial derived from $\mathbf{T}(\mathbf{u})$. This approach aims to achieve gridless DOA estimation and improve performance for arbitrary array geometries.

3. Angular-Domain Covariance Matrix Reconstruction Network

In this section, we propose a CDNN to model the mapping between the SCM $\widehat{\mathbf{R}}_y$ and the angular-domain covariance matrix $\mathbf{T}(\mathbf{u})$ using the subarray composed of elements #1 to #8 of the HLAS in the Swellex-96 experiment [38]. Section 3.1 introduces the CDNN, Section 3.2 provides the array geometry and parameter settings, Section 3.3 describes the dataset used for training the network, and Section 3.4 presents the proposed network framework along with the training process and results.

3.1. Complex-Valued Deep Neural Network

CDNNs [39] can capture the phase information between the real and imaginary components of complex numbers, which are often crucial in signal processing tasks. As a result, CDNNs have been widely applied in these tasks [40–42], achieving better performance compared to real-valued deep neural networks (RDNNs). To achieve better performance in covariance matrix reconstruction with fewer parameters, we employ CDNN. The principles of CDNN are similar to those of RDNN. In the case of a fully connected neural network, it consists of an input layer, an output layer, and multiple cascaded hidden layers. Within each hidden layer, the input undergoes a linear transformation followed by a nonlinear activation function. This process can be expressed as

$$\mathbf{o} = \sigma(\mathbf{z}) = \sigma(\mathbf{W}\mathbf{o}' + \mathbf{b}), \quad (8)$$

where \mathbf{W} and \mathbf{b} are the complex weight matrix and bias vector, respectively, \mathbf{o}' and \mathbf{o} are the complex outputs of the previous layer and the current layer, respectively, and $\sigma(\bullet)$

represents the complex nonlinear activation function. The complex matrix multiplication in real space can be expressed as

$$\begin{bmatrix} \Re(\mathbf{z}) \\ \Im(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} \Re(\mathbf{W}) & -\Im(\mathbf{W}) \\ \Im(\mathbf{W}) & \Re(\mathbf{W}) \end{bmatrix} \bullet \begin{bmatrix} \Re(\mathbf{o}') \\ \Im(\mathbf{o}') \end{bmatrix} + \begin{bmatrix} \Re(\mathbf{b}) \\ \Im(\mathbf{b}) \end{bmatrix}, \quad (9)$$

where $\Re(\bullet)$ and $\Im(\bullet)$ denote the real and imaginary parts of the complex number, respectively. According to (8) and (9), a complex-valued layer can be expressed as

$$\mathbf{h}(\mathbf{o}) = (\mathbf{h}_r(\mathbf{o}_r) - \mathbf{h}_i(\mathbf{o}_i)) + j(\mathbf{h}_r(\mathbf{o}_i) + \mathbf{h}_i(\mathbf{o}_r)), \quad (10)$$

where $\mathbf{o}_r = \Re(\mathbf{o})$, $\mathbf{o}_i = \Im(\mathbf{o})$, $\mathbf{h}_r(\mathbf{x}) = \Re(\mathbf{W})\mathbf{x} + \Re(\mathbf{b})$, $\mathbf{h}_i(\mathbf{x}) = \Im(\mathbf{W})\mathbf{x} + \Im(\mathbf{b})$. Thus, CDNN is an extension of RDNN, and a CDNN of the same scale has twice the number of parameters as an RDNN.

3.2. Array Geometry and Parameter Settings

We use the subarray composed of elements #1 to #8 from HLAS in Swellex-96 to generate a dataset for training the CDNN for angular-domain covariance matrix reconstruction. With element #5 as the origin of the Cartesian coordinate system, the coordinates of the 8-element array are shown in Table 1. The array is approximated as a linear array with unequal element spacing, with element #1 aligned in the direction of azimuth $\theta = 0$. The DOA is defined as the angle between the end-fire direction of the #1 element and the direction of the incoming source, with the counterclockwise direction considered positive, ranging from 0° to 180° . The center frequency of the narrowband far-field plane wave is set to $f = 50$ Hz, and the speed of sound in water is assumed to be $c = 1490$ m/s. The maximum element spacing in the linear array is $d_{\max} = 0.21\lambda$, and the minimum element spacing is $d_{\min} = 0.12\lambda$.

Table 1. Array Coordinates.

Element No.	#1	#2	#3	#4	#5	#6	#7	#8
x (m)	11.93	9.26	6.50	3.41	0	-3.75	-7.87	-12.43
y (m)	-10.4	-8.24	-5.78	-3.06	0	3.40	7.34	11.62

According to [19,21], when the wavelength of the narrowband signal is fixed, the error between the steering vector approximated by (5) and the true steering vector is related to the distance of the array elements from the coordinate origin, as well as the truncation order N . Figure 1a shows the magnitude of the Fourier coefficients for different element distances when $f = 50$ Hz, $\theta = 0$, and $\varphi_m = 0$, with the Fourier coefficients expressed in decibels, i.e., $20 \lg|b_n|$. Figure 1b shows the relationship between the steering vector error and truncation order for different azimuth angles in the array specified in Table 1, with the steering vector error also expressed in decibels, i.e., $20 \lg|\mathbf{a}(\theta) - \mathbf{B}^H \mathbf{f}(\theta)|$. We find that the magnitude of the Fourier coefficients decreases with increasing order, so higher truncation orders reduce the approximation error in (4) and (5), making $\widehat{\mathbf{R}}_y$ closer to the true element-domain covariance matrix. At the same time, the dimensionality of $\mathbf{T}(\mathbf{u})$ increases with higher truncation orders, theoretically enhancing the performance of DOA estimation. However, this increase in dimensionality also raises the complexity of the network by expanding the number of output parameters, making the network more challenging to train. Therefore, we set the truncation orders $N = 4$, $N = 7$ and $N = 10$ in the following experiments and analyze the effect of truncation order on the performance of CDNN in reconstructing $\mathbf{T}(\mathbf{u})$ through simulation results.

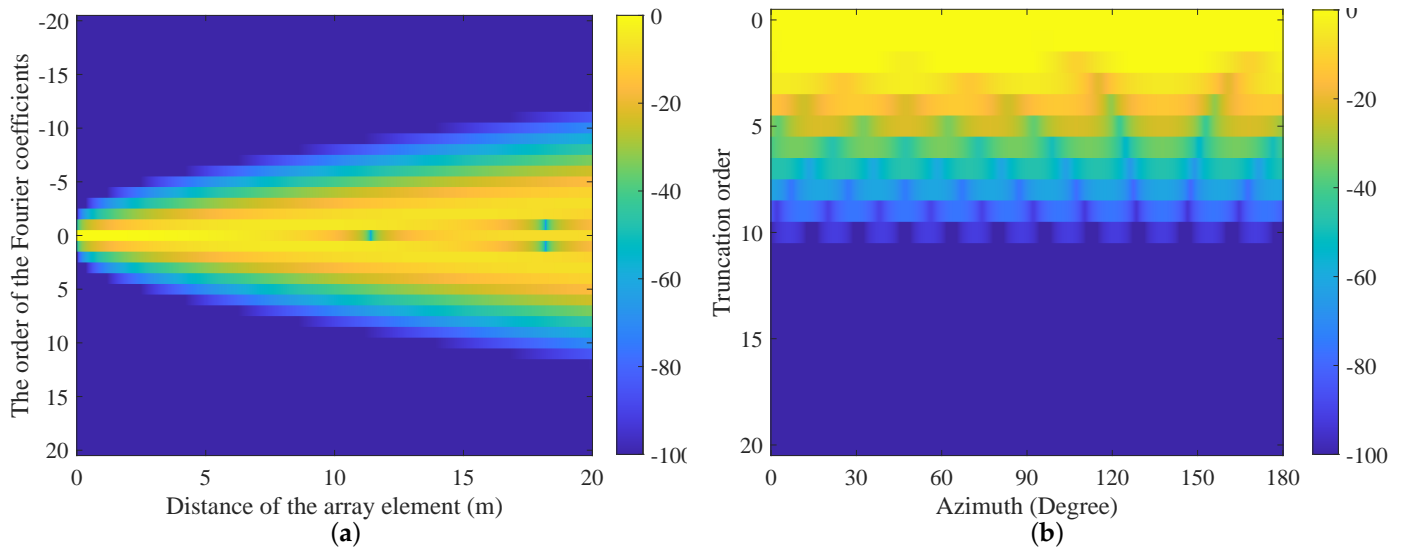


Figure 1. (a) Magnitude of Fourier coefficients at various orders. (b) Relationship between array steering vector error and N .

3.3. Dataset

We use MATLAB to generate the dataset for training the CDNN according to the array geometry and parameters provided in Section 3.2. The number of incoherent sources in the dataset is denoted as $K = \{2, 3\}$, with data corresponding to $K = 2$ and $K = 3$ each comprising half of the dataset. In the data numbered i , the DOA θ_{ik} for the k th incoherent source are randomly generated within the range of $[0, \pi]$, following an independent uniform distribution. The source amplitudes $\alpha_{i1} = 1$, and $\alpha_{ik}, k \neq 1$ follow independent uniform distributions over $[0, 1]$. To enhance the robustness of the CDNN in real noise environments, we introduce spatially and temporally uncorrelated Gaussian white noise into the data, with the SNR uniformly distributed over $[0, 20]$ dB. For multiple sources, the SNR is defined

as $\frac{\sum_{k=1}^K \alpha_k^2}{\sigma^2}$, where σ^2 represents the noise power. Substituting the above parameters into (1), we obtain the array snapshot signal $\mathbf{Y}_i = [\mathbf{y}_{i1}, \mathbf{y}_{i2}, \dots, \mathbf{y}_{iL}]$, where $\mathbf{y}_{il} \sim CN(\mathbf{0}, \mathbf{A}_i \mathbf{R}_{xi} \mathbf{A}_i^H + \frac{\sum_{k=1}^K \alpha_{ik}^2}{SNR_i} \mathbf{I})$, $\mathbf{R}_{xi} = \text{diag}(\alpha_{i1}^2, \dots, \alpha_{iK}^2)$, $\mathbf{A}_i = [\mathbf{a}(\theta_{i1}), \dots, \mathbf{a}(\theta_{iK})]$, and \mathbf{I} is the identity matrix. To enhance the randomness of the dataset, we set the number of snapshots for i th data sample as $L_i = \text{mod}(i, 10) \times 10$, where $\text{mod}(x, y)$ represents the remainder when dividing x by y .

We model the mapping from $\widehat{\mathbf{R}}_y$ to $\mathbf{T}(\mathbf{u})$ as a regression task, with $\widehat{\mathbf{R}}_y$ as the input and $\mathbf{T}(\mathbf{u})$ as the target label. The input data $\widehat{\mathbf{R}}_y$ is the SCM of the array's snapshot signal \mathbf{Y} , which is a complex-valued matrix of dimension $M \times M$. For the i th input data $\widehat{\mathbf{R}}_{y_i} = \mathbf{Y}_i \mathbf{Y}_i^H / L_i$, the corresponding label is the ideal angular-domain covariance matrix $\mathbf{T}(\mathbf{u})_i = \mathbf{F}_i^H \mathbf{R}_{xi} \mathbf{F}_i$, where $\mathbf{F}_i = [\mathbf{f}(\theta_{i1}), \mathbf{f}(\theta_{i2}), \dots, \mathbf{f}(\theta_{iK})]$ is a complex-valued matrix of dimension $(2N + 1) \times (2N + 1)$. Therefore, the i th data pair in the dataset is represented as $(\widehat{\mathbf{R}}_{y_i}, \mathbf{T}(\mathbf{u})_i)$, forming a dataset $\varepsilon_{tra} = \left\{ (\widehat{\mathbf{R}}_{y_1}, \mathbf{T}(\mathbf{u})_1), (\widehat{\mathbf{R}}_{y_2}, \mathbf{T}(\mathbf{u})_2), \dots, (\widehat{\mathbf{R}}_{y_{E_{tra}}}, \mathbf{T}(\mathbf{u})_{E_{tra}}) \right\}$ of size $E_{tra} = 120,000$ to be used as the training set. Similarly, we generate a validation set ε_{val} of size $E_{val} = 10,000$ for model validation during training. Additionally, we construct a test set ε_{test} of size $E_{test} = 10,000$ to evaluate the performance of the trained CDNN.

3.4. Network Architecture and Training

As shown in Figure 2, the proposed angular-domain covariance matrix reconstruction CDNN consists of input, output, and dense layers. Since the input SCM is a Hermitian

matrix, its upper triangular elements cover all the matrix information. Therefore, the input layer takes a vector $\hat{\mathbf{r}}_y \in \mathbb{C}^{M(M+1)/2}$ composed of the upper triangular part of the SCM. The output layer produces a vector $\mathbf{u} \in \mathbb{C}^{2N+1}$, composed of the first row elements of $\mathbf{T}(\mathbf{u})$. After post-processing $f_p(\mathbf{u}) = \mathbf{T}(\mathbf{u})$, the reconstructed angular-domain covariance matrix is obtained. The CDNN can be represented as

$$\widehat{\mathbf{T}}(\mathbf{u}) = f_{CDNN}(\widehat{\mathbf{R}}_y) = f_p(f_7(f_6 \cdots f_1(\hat{\mathbf{r}}_y))), \quad (11)$$

where $\{f_i(\bullet)\}_{i=1,2,\dots,6}$ is followed by Complex ReLU (CReLU) as the activation function. In $\{f_i(\bullet)\}_{i=1,2,3,4}$, Complex LayerNorm (CLayerNorm) is inserted between the fully connected layers and CReLU for normalization. A Dropout layer with a dropout rate of 0.3 is applied after $f_2(\bullet)$ to mitigate overfitting. The CReLU and CLayerNorm functions can be expressed as

$$\text{CRelu}(\mathbf{x}) = \text{Relu}(\Re(\mathbf{x})) + j\text{Relu}(\Im(\mathbf{x})), \quad (12)$$

$$\text{CNorm}(\mathbf{x}) = \frac{\mathbf{x} - \mu}{\sqrt{\sigma^2 - \varepsilon}}, \quad (13)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_D]^T$ is the output of a certain layer in the network, $\mu = \frac{1}{D} \sum_{d=1}^D x_d$, $\sigma^2 = \frac{1}{D} \sum_{d=1}^D |x_d - \mu|^2$, and ε is a small value to prevent division by zero in (13).

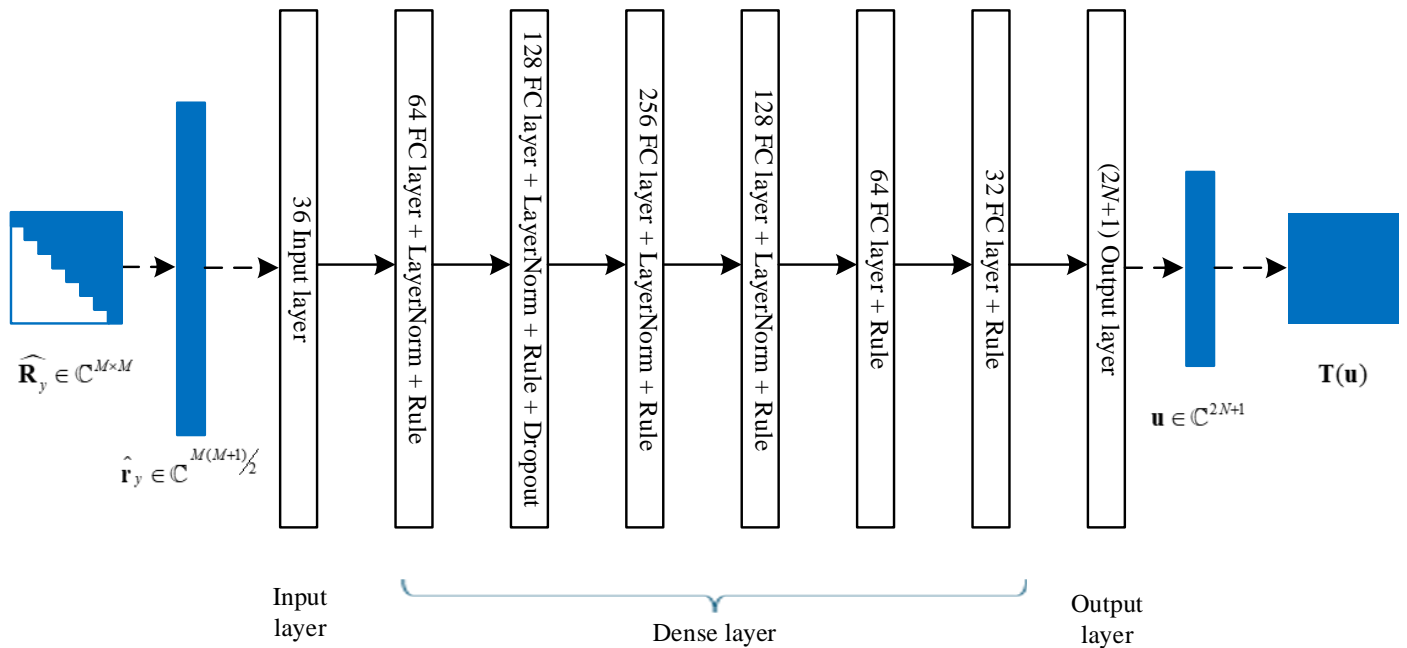


Figure 2. Angular-domain covariance matrix reconstruction network architecture.

Since $\mathbf{T}(\mathbf{u})$ is expressed in terms of power, we chose the mean absolute error between $f_{CDNN}(\widehat{\mathbf{R}}_{y_i})$ and $\mathbf{T}(\mathbf{u})_i$ as the loss function during network training to ensure that the network's output closely approximates the true angular-domain covariance matrix. We use the Adam optimizer [43] to optimize the network parameters, with an initial learning rate of 0.01 and parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$. Additionally, a 20-epoch warmup is introduced to stabilize network convergence. The training process adopts mini-batch gradient descent with a batch size of 1024. At the beginning of each epoch, the training data are shuffled to ensure that the distribution of each mini-batch is random. Then, the data are divided into multiple mini-batches, with each mini-batch undergoing forward propagation, loss computation, backpropagation, and parameter updates. To improve training efficiency

and prevent overfitting, we introduce ReduceLRonPlateau and Early Stopping techniques. If the loss on the validation set does not decrease within 10 epochs, ReduceLRonPlateau reduces the learning rate by half, allowing the optimizer to search the parameter space more finely. If the loss on the validation set does not improve for 20 consecutive epochs, Early Stopping terminates training to prevent overfitting on the training set.

Figure 3 illustrates the changes in training loss, validation loss, and learning rate over epochs for different values of $N = 4$, $N = 7$ and $N = 10$. Table 2 summarizes the training results of the three CDNN models, including the final losses on the training and validation sets, as well as the losses on the test set after training. Based on the results shown in Table 2 and Figure 3, it can be observed that all three CDNN models converge well and do not exhibit significant overfitting. Additionally, since we only varied the output dimension across these three CDNN models, we find that while all models achieve good training performance, the CDNN with a higher output dimension tends to exhibit larger average errors on the test set.

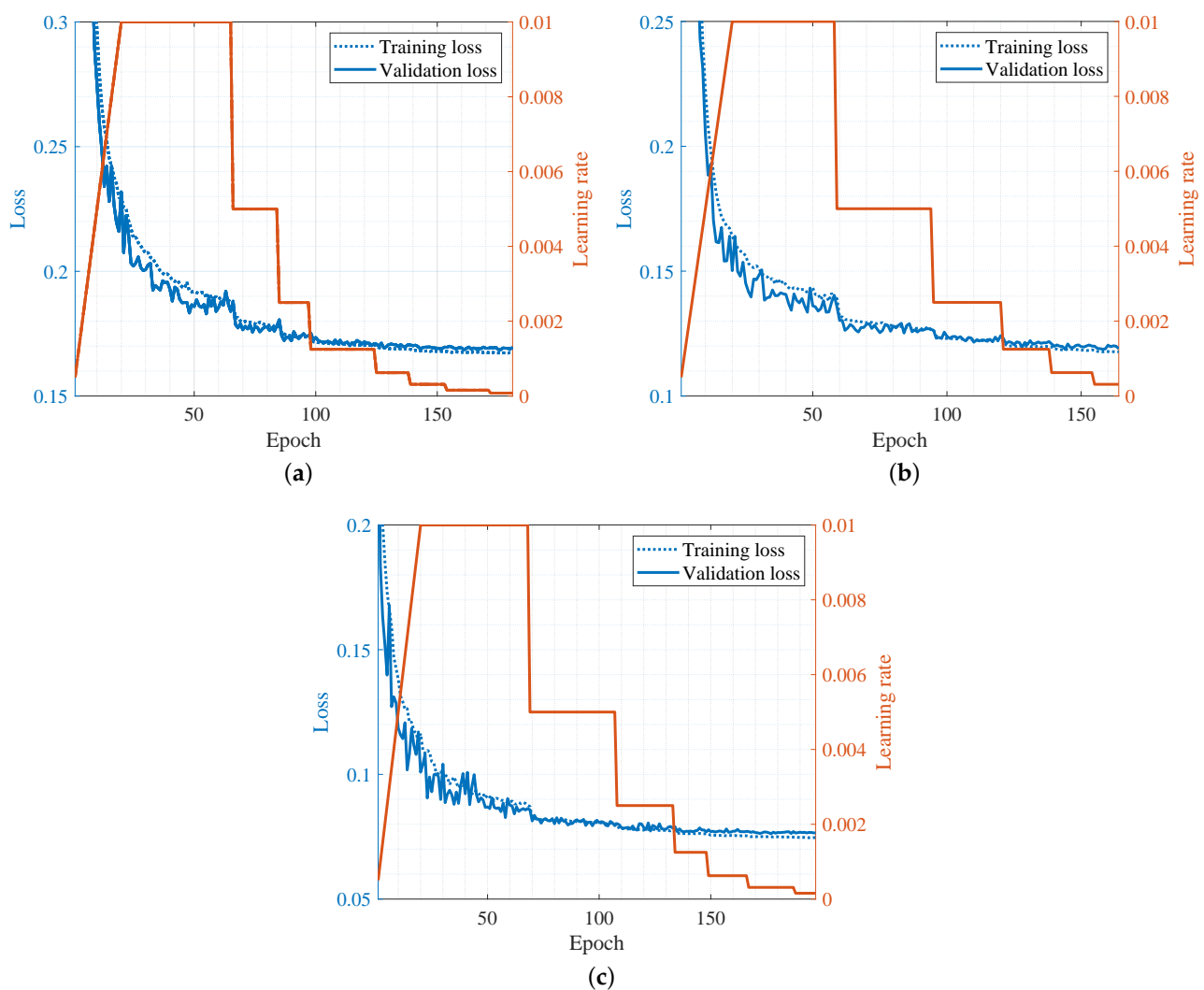


Figure 3. Training loss, validation loss, and learning rate variation with epochs for the following: (a) $N = 4$. (b) $N = 7$. (c) $N = 10$.

Table 2. CDNN training results.

	Training Loss	Validation Loss	Test Loss
$N = 4$	0.0750	0.0762	0.0766
$N = 7$	0.1185	0.1187	0.1197
$N = 10$	0.1676	0.1685	0.1692

4. Simulation

In this section, we will analyze the performance of our proposed method through numerical simulations and compare it with existing DOA estimation methods such as MUSIC [5], SPICE [12], and SBL [44]. First, we compare the gridless DOA estimation performance of CDNN models with truncation orders $N = 4$, $N = 7$ and $N = 10$ under different SNR conditions. Then, we present the DOA estimation results of these methods for varying numbers of sources within the 0° to 180° range. Finally, we evaluate the impact of the SNR, target angular separation, and the number of snapshots on the performance of the aforementioned algorithms in uncorrelated source scenarios through Monte Carlo simulations.

We consider K uncorrelated narrowband acoustic sources with equal power at a center frequency of $f = 50$ Hz, whose DOAs are $\theta = [\theta_1, \theta_2, \dots, \theta_K]$, assuming that the number of sources is known. In the Monte Carlo simulations, the number of sources is set to $K = 2$. The performance of DOA estimation is evaluated using resolution probability (RP) and root mean square error (RMSE) [9,45,46]. The Cramer–Rao Bound (CRB) provides the lower bound of RMSE for unbiased estimators, serving as a reference for the DOA estimation algorithms studied in this paper [47]. For grid-based methods, DOA estimates are obtained from the peaks of the spatial spectrum. Specifically, the spatial spectrum obtained from the DOA estimation algorithm is first normalized, and then peaks with power greater than -10 dB are identified, where each peak corresponds to a target. If the algorithm identifies two targets and the DOA estimates $\hat{\theta} = [\hat{\theta}_1, \hat{\theta}_2]$ satisfy $|\hat{\theta}_1 - \theta_1| + |\hat{\theta}_2 - \theta_2| < |\theta_1 - \theta_2|$, the algorithm is considered to have successfully resolved the targets. To ensure the accuracy of the simulations, we only simulate the RMSE of the algorithm for cases where $RP \geq 0.5$.

The RMSE of DOA estimation is defined as $\sqrt{\frac{1}{2T} \left[\sum_{t=1}^T (\hat{\theta}_{1t} - \theta_1)^2 + \sum_{t=1}^T (\hat{\theta}_{2t} - \theta_2)^2 \right]}$, where T is the number of successful resolutions, and $\hat{\theta}_t = [\hat{\theta}_{1t}, \hat{\theta}_{2t}]$ is the estimate of θ in the t th successful simulation. Additionally, the number of Monte Carlo simulations is set to 1000. For SPICE and SBL, the number of iterations is set to 10 and 1000, respectively. For the MUSIC algorithm, we divide the azimuth angle grid at 0.1° intervals, while for SPICE and SBL, the grid is divided at 1° intervals.

4.1. Impact of N on DOA Estimation Performance

Figure 4 illustrates the relationship between RMSE and RP with respect to the SNR for the gridless DOA estimation performance of CDNNs with truncation orders $N = 4$, $N = 7$, and $N = 10$ under different scenarios: $\theta = [85^\circ, 95^\circ]$, $\theta = [80^\circ, 100^\circ]$ and $\theta = [70^\circ, 110^\circ]$. In the simulation, the noise is modeled as a zero-mean Gaussian white random process, assumed to be spatially and temporally uncorrelated. The SNR is defined as $\frac{\alpha_1^2 + \alpha_2^2}{\sigma^2}$, where α_1^2 , α_2^2 and σ^2 represent the power of the two sources and the noise at the array elements, respectively. The number of snapshots is set to 50. As shown in Figure 4d,f, even when the angular separation between two targets is large and the SNR is high, the CDNN with $N = 10$ still occasionally fails to distinguish the two targets. Additionally, as depicted in Figure 4b, when the angular separation is small, the RP for $N = 10$ deteriorates more significantly as the SNR decreases. Conversely, as seen in Figure 4c,e, the CDNN with $N = 4$ exhibits higher robustness to Gaussian white noise when the target angular separations are 20° and 40° , providing better performance at low SNR levels while maintaining similar performance to $N = 7$ and $N = 10$ at higher SNRs. However, as shown in Figure 4a,

when the target angular separation decreases to 10° and the SNR exceeds 5 dB, the DOA estimation error for the CDNN with $N = 4$ becomes the largest. In summary, we observe that while $N = 10$ provides better resolution for closely spaced targets, its DOA estimation performance becomes unstable when the target angular separation is large and the SNR is high. The instability observed at $N = 10$ is due to the network needing to fit outputs of a higher dimensionality, which can lead to higher fitting errors. On the other hand, $N = 4$ delivers superior DOA estimation performance under conditions of larger angular separation and a lower SNR, but its resolution is limited when the angular separation is small. To balance the DOA estimation performance across varying angular separations, we choose the CDNN with $N = 7$, where the corresponding angular-domain covariance matrix $\mathbf{T}(\mathbf{u})$ is a 15×15 Toeplitz matrix.

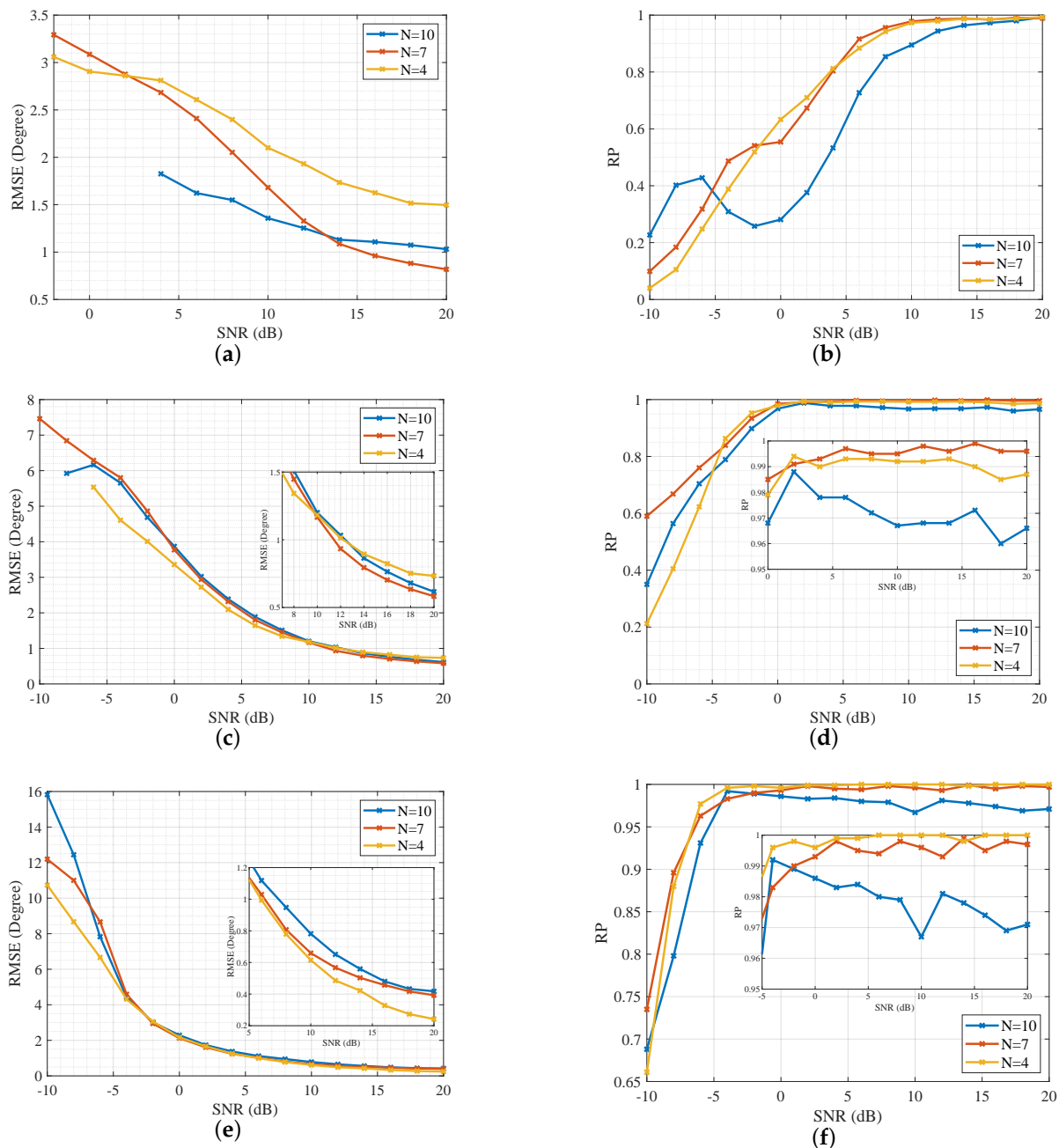


Figure 4. DOA estimation performance of CDNNs with truncation orders $N = 4$, $N = 7$, and $N = 10$ at different target angular separations. (a) RMSE and (b) RP for $\theta = [85^\circ, 95^\circ]$. (c) RMSE and (d) RP for $\theta = [80^\circ, 100^\circ]$. (e) RMSE and (f) RP for $\theta = [70^\circ, 110^\circ]$.

4.2. DOA Estimation Results

In this section, we present the DOA estimation results using the methods under study for scenarios with $K \in \{1, 2, 3\}$ sources within an angular range of from 0° to 180° . When the number of sources is 2 or 3, we set the angular separation between adjacent sources as $\Delta\theta = 40^\circ$, meaning that all source DOAs are $\theta = [\theta_1, \theta_1 + \Delta\theta, \dots, \theta_1 + (K - 1)\Delta\theta]$. We consider the case where θ satisfying $\theta_1 \in [0^\circ : 0.1^\circ : 180^\circ - (K - 1)\Delta\theta]$, with the SNR set to 10 dB, and with the noise being spatiotemporally uncorrelated Gaussian white noise. The number of snapshots is set to 50. The DOA estimation results for various methods are shown in Figure 5, and the RMSE for different angular ranges is provided in Table 3. When the number of sources $K = 1$, the results in Table 3 indicate that the proposed method and MUSIC achieve more accurate DOA estimates than SPICE and SBL. Comparing Figure 5a,d,g,j, it can be seen that even when dividing the grid densely with a separation of 1° , grid mismatch errors still severely affect the performance of SPICE and SBL. When the number of sources $K = 2$ and $K = 3$, and all sources are within the angular range $[45^\circ, 135^\circ]$, all four methods can obtain relatively accurate DOA estimation results. However, when the range of source DOAs is confined to $[0^\circ, 180^\circ]$, the proposed method shows a significant RMSE advantage over the other three methods. According to the simulation results shown in Figure 5, when all sources are far from the end-fire direction, all four methods can effectively distinguish targets and achieve accurate DOA estimation. However, when a source is close to the end-fire direction, the DOA estimation performance of the other three methods deteriorates significantly, making it difficult to distinguish adjacent targets or leading to large DOA estimation errors. In contrast, the performance degradation of the proposed method is relatively mild.

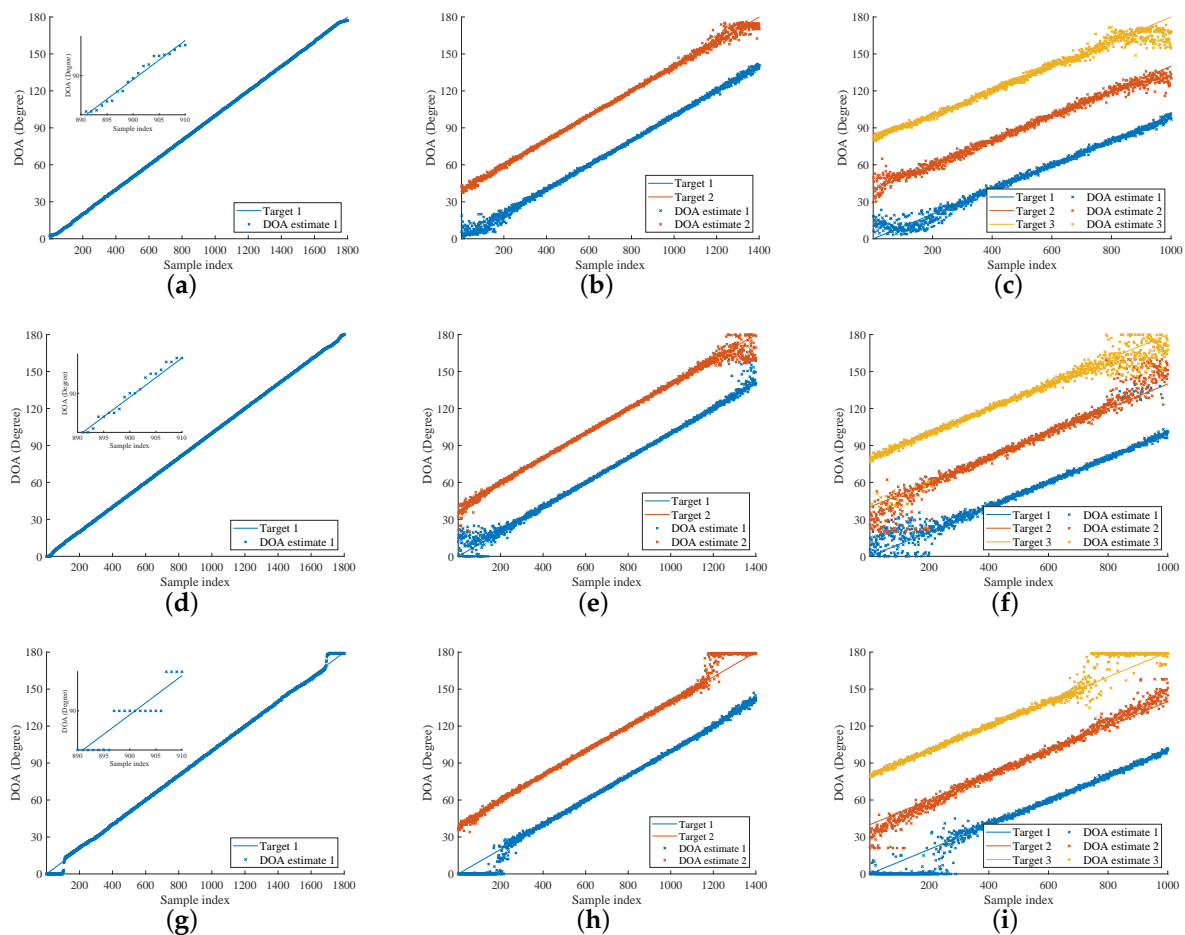


Figure 5. Cont.

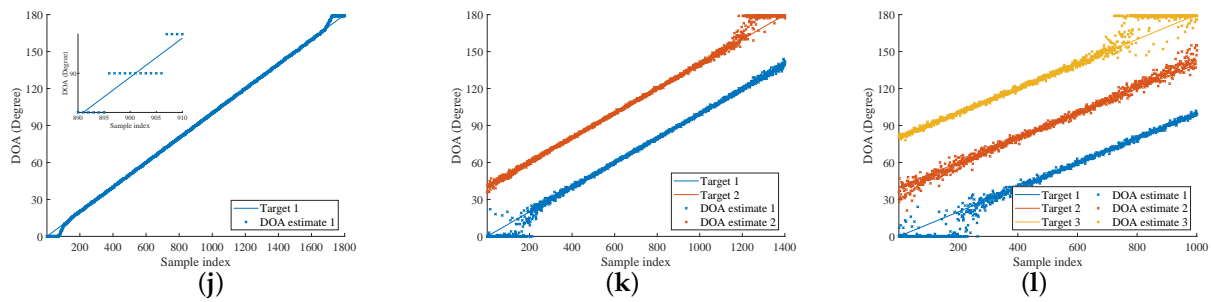


Figure 5. DOA estimation results. The proposed method, source numbers (a) $K = 1$, (b) $K = 2$, and (c) $K = 3$. MUSIC, source numbers (d) $K = 1$, (e) $K = 2$, and (f) $K = 3$. SPICE, source numbers (g) $K = 1$, (h) $K = 2$, and (i) $K = 3$. SBL, source numbers (j) $K = 1$, (k) $K = 2$, and (l) $K = 3$.

Table 3. RMSE of DOA estimation.

K	The Range of $\theta_k, \forall k = 1, 2, \dots, K$	Proposed Method	MUSIC	SPICE	SBL
1	$[45^\circ, 135^\circ]$	0.1278	0.1098	0.2676	0.2678
	$[0^\circ, 180^\circ]$	0.2239	0.1885	1.0060	0.6724
2	$[45^\circ, 135^\circ]$	0.9051	0.8279	0.9340	0.9146
	$[0^\circ, 180^\circ]$	2.0294	2.9906	3.9232	3.2688
3	$[45^\circ, 135^\circ]$	2.4844	2.2639	2.1821	2.0246
	$[0^\circ, 180^\circ]$	5.0632	9.1194	8.6687	7.1800

4.3. Relationship between Algorithm Performance and SNR

We analyze the relationship between algorithm performance and the SNR under two scenarios: $\theta = [80^\circ, 100^\circ]$ and $\theta = [70^\circ, 110^\circ]$. The other definitions and settings in the simulation are the same as those in Section 4.1, and the simulation results are presented in Figure 6. When $\theta = [80^\circ, 100^\circ]$, as shown in Figure 6a, the proposed method achieves a lower RMSE than the other three methods when the SNR is below 14 dB. Additionally, as illustrated in Figure 6b, the proposed method maintains an RP above 0.6 for distinguishing the two sources across the SNR range of -10 to 20 dB. In contrast, the RP of the other three methods rapidly decreases with a decreasing SNR, falling below 0.2 at an SNR of 0 dB, indicating that they almost fail to distinguish the two sources. For the scenario with $\theta = [70^\circ, 110^\circ]$, Figure 6c,d show that in the SNR range of -2 to 12 dB, the proposed method achieves comparable or lower RMSE compared to the other methods, while all methods maintain nearly perfect RP values close to 1, successfully distinguishing the two targets. When the SNR drops below -4 dB, although the RMSE of the proposed method exceeds that of SBL and MUSIC, it still achieves the highest RP for source resolution, with RP remaining above 0.7 even when the SNR decreases to -10 dB.

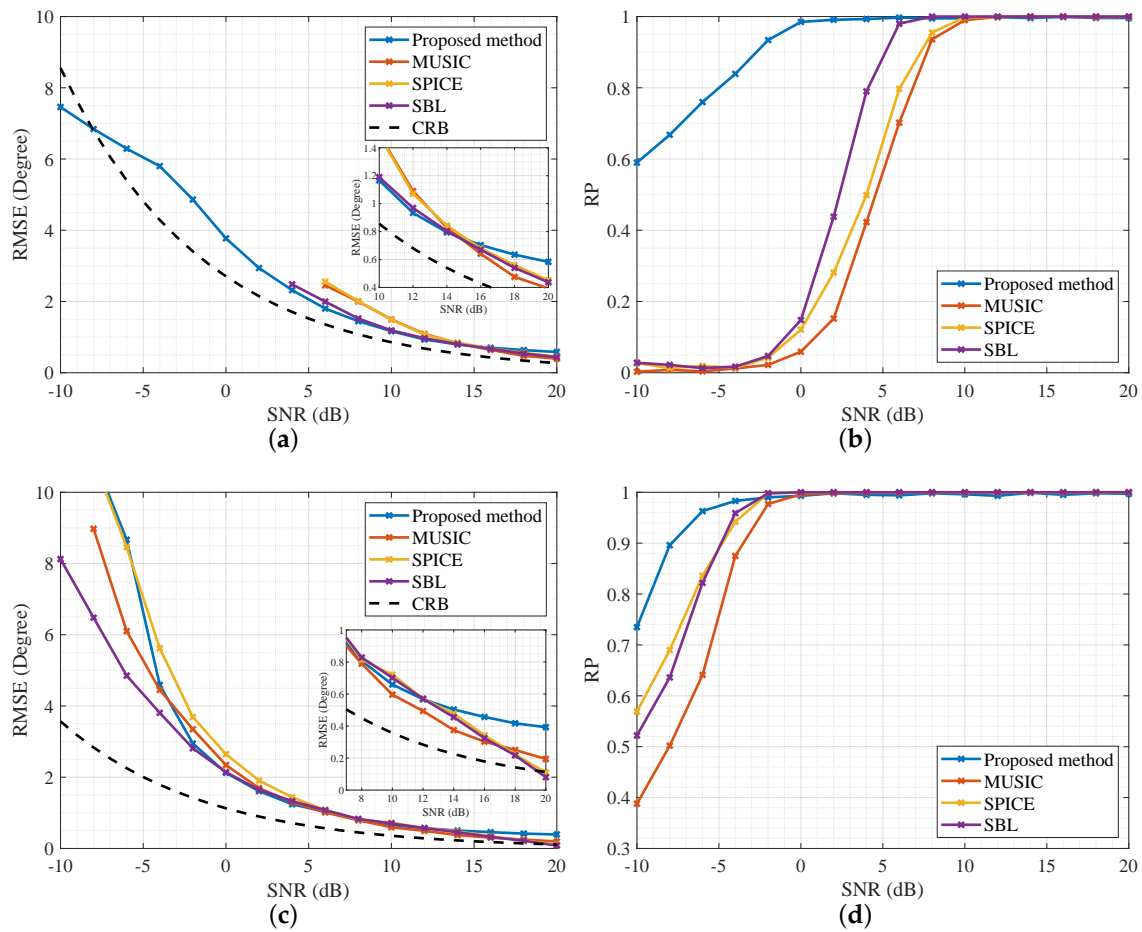


Figure 6. Relationship between SNR and both RMSE and RP under spatio-temporal Gaussian white noise conditions. (a) RMSE and (b) RP for $\theta = [80^\circ, 100^\circ]$. (c) RMSE and (d) RP for $\theta = [70^\circ, 110^\circ]$.

4.4. Relationship between Algorithm Performance and Angle Separation

We set $\theta_1 = 90^\circ - \Delta\theta/2$, $\theta_2 = 90^\circ + \Delta\theta/2$, where $\Delta\theta$ represents the angular separation between the two sources. The SNR in the simulation is set to 10 dB, and other conditions remain the same as in Section 4.1. The relationship between the angle separation and the DOA estimation performance, as measured by RP and RMSE, is shown in Figure 7. When the angular separation between the two sources is between 20° and 50° , the proposed method shows similar RMSE values compared to the other three methods, with the RMSE difference among the four methods being less than 0.2° . Additionally, all four methods are able to distinguish between the two sources with an RP close to 1. However, when the angular separation between the two sources is less than 20° , the proposed method exhibits the lowest DOA estimation RMSE and can still differentiate between the two sources with an RP close to 1. In contrast, the RP of the other three methods decreases rapidly with decreasing angular separation. When the angular separation is reduced to 10° , the RP of all three existing methods drops below 0.2, making them almost unable to distinguish between the sources. The simulation results demonstrate that the proposed method outperforms the existing methods in DOA estimation when the angular separation is small.

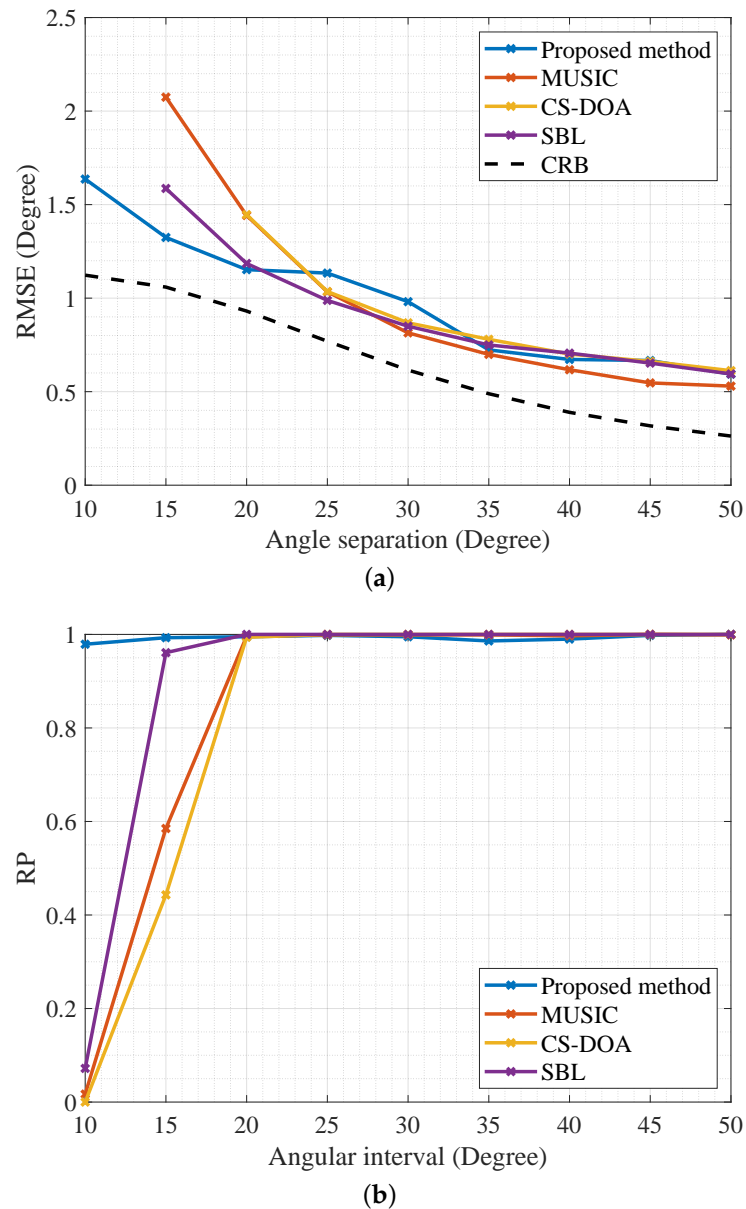


Figure 7. Relationship between RP and RMSE with respect to angle separation. (a) RMSE. (b) RP.

4.5. Impact of Snapshot Number on Algorithm Performance

In this section, we analyze the impact of the number of snapshots on the algorithm's performance. Figure 8a,b illustrate the RMSE and RP of the proposed method and the three existing methods for DOA estimation under different snapshot numbers. In the simulation, we set $\theta = [78^\circ, 102^\circ]$ and the SNR to 10 dB, with other conditions being the same as in Section 4.1. As shown in Figure 8, under the simulation conditions, when the number of snapshots exceeds 50, the proposed method exhibits the highest RMSE, while all four methods are able to distinguish the two sources with an RP close to 1. When the number of snapshots is less than 30, the RMSE of MUSIC and SPICE increases, and their RP decreases rapidly as the number of snapshots decreases. When the number of snapshots falls below 10, the RP of MUSIC and SPICE drops below 0.5, indicating a high probability of failure in accurate DOA estimation. At this point, the proposed method can still distinguish the two sources with an RP close to 1, and it also achieves lower RMSE compared to the existing three methods. Based on the simulation results, we conclude that the proposed method outperforms the other three existing methods when the number of snapshots is low, demonstrating higher adaptability to fewer snapshots.

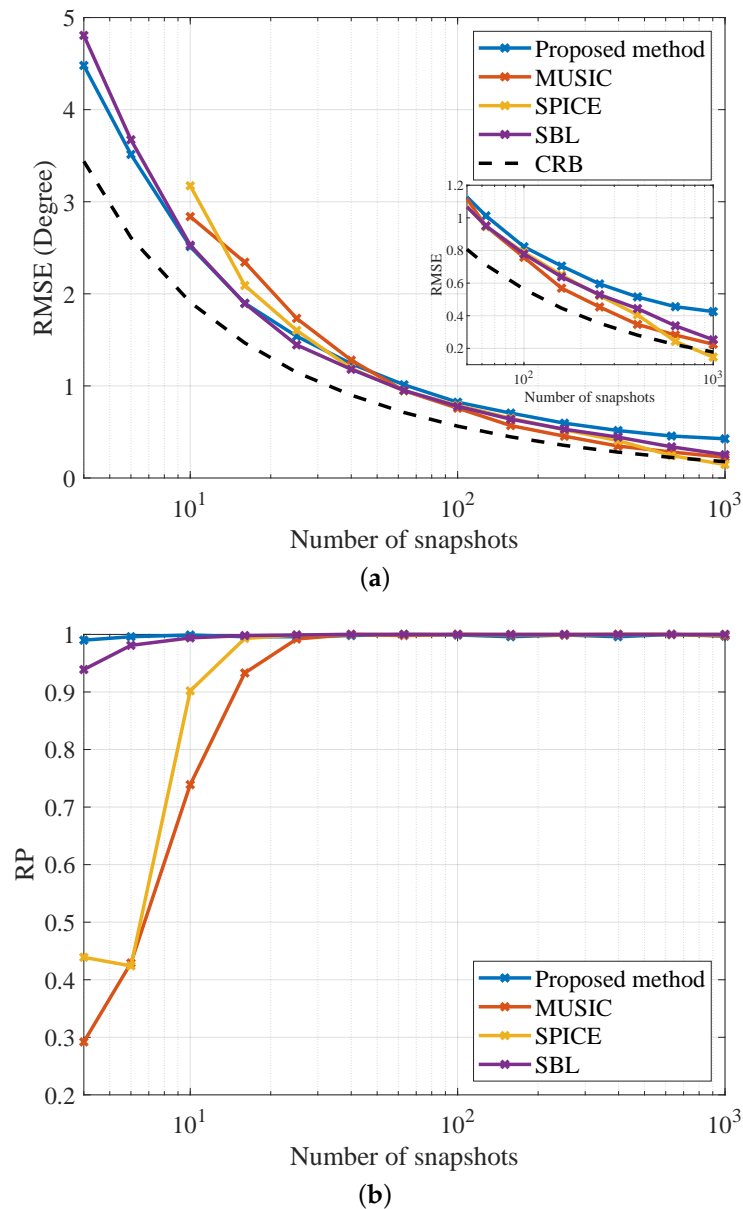


Figure 8. Algorithm performance under different snapshot conditions. (a) RMSE. (b) RP.

5. Swellex-96 Event S59 Experimental Results

In this section, we validate the effectiveness of the proposed method using experimental data from the Swellex-96 Event S59. The SWellEx-96 experiment [38] was conducted in shallow waters approximately 12 km off the tip of Point Loma near San Diego, California, from 10 to 18 May 1996. As shown in Figure 9, the event scenario depicts the source's trajectory in blue and the interferer's trajectory in red. The markers on the tracks represent GPS-recorded coordinates at 5-minute intervals, and the entire experiment lasted for 65 min. During the experiment, the source continuously emitted signals at various frequencies, with the higher power frequencies being 49, 64, 79, 94, 112, 130, 148, 166, 201, 235, 283, 338, and 388 Hz. The HLA subarray used in our study had a sampling frequency of $f_s = 3276.8$ Hz, and the subarray formed by elements #1 to #8 effectively met the far-field condition during the experiment.

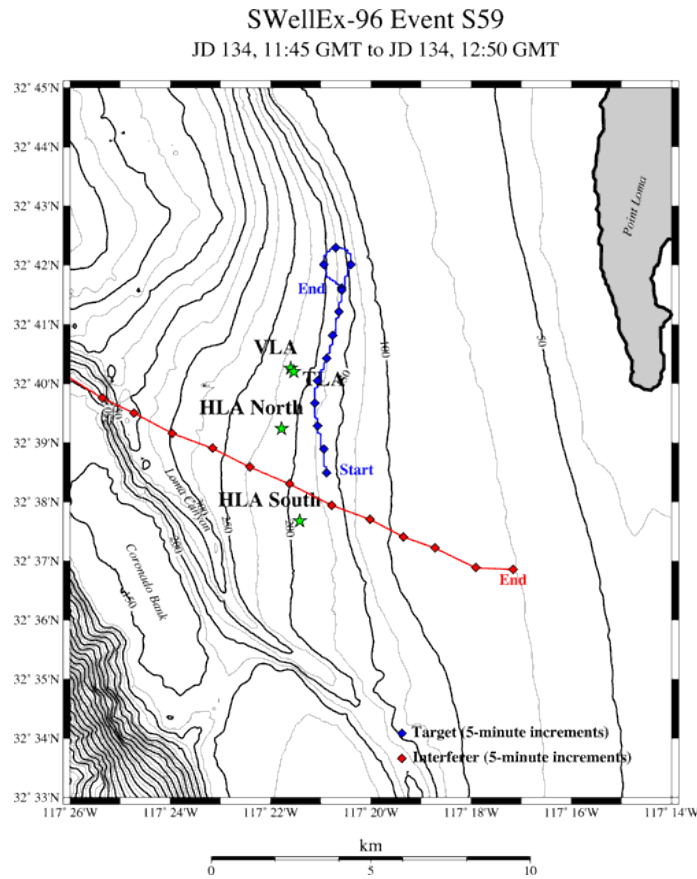


Figure 9. Schematic of the SwellEx-96 Event S59 experiment scenario [38].

The data received by the 8-element subarray are denoted as $\mathbf{Y} = [\mathbf{Y}(:,1), \mathbf{Y}(:,2), \dots, \mathbf{Y}(:,8)]$, where the column vector $\mathbf{Y}(:,m)$ represents the data received by the m th array element. We divided the data into $Q = 361$ segments, each starting at time $t_{start} \in \{t_1, t_2, \dots, t_Q\} = [0s : 10s : 3600s]$ and lasting for a duration $T_{seg} = 30s$. The data received by the m th array element during the q th segment can be expressed as $\bar{\mathbf{y}}_{mq} = \mathbf{Y}(\lfloor f_s t_q \rfloor + 1 : \lfloor f_s (t_q + T_{seg}) \rfloor, m)$, where $\lfloor \bullet \rfloor$ denotes the floor function. Each segment was further divided into $L = 100$ snapshots, with each snapshot containing $N_{snap} = 2048$ sampling points. The duration of each snapshot was $T_{snap} = 0.625 s$, and the overlap between consecutive snapshots was $\eta = 53\%$. Therefore, the l th snapshot within $\bar{\mathbf{y}}_{mq}$ is represented as $\tilde{\mathbf{y}}_{mql} = \bar{\mathbf{y}}_{mq}(\lfloor (l-1)(1-\eta)f_s T_{snap} \rfloor + 1 : \lfloor (l-1)(1-\eta)f_s T_{snap} \rfloor + N_{snap})$. We will apply different DOA estimation methods based on the snapshot model described above to perform narrowband DOA estimation at a frequency of $f_c = 50 \text{ Hz}$ for each data segment, thereby obtaining the DOA trajectories.

Figure 10 presents the bearing time records (BTRs) obtained using different DOA estimation methods based on the SwellEx-96 Event S59 data, along with the GPS-recorded tracks of the source and interferer. As shown in Figure 10b, for CBF, prior to 1800 s, due to the small angular separation between the source and the interferer, the two targets cannot be distinguished. After 2400 s, although the angular separation between the source and interferer increases and two peaks appear in the CBF spectrum, high sidelobes severely affect the performance of DOA estimation. As shown in Figure 10d–f, the three existing methods demonstrate better resolution compared to CBF, allowing the source and interferer tracks to be distinguished. However, all three methods exhibit high levels of pseudospectra around the DOA track, which negatively impact target trajectory identification. Moreover, for SPICE and SBL, the interferer's DOA is close to the end-fire direction of the array,

making it difficult to identify the interferer prior to 1800 s in most cases. As shown in Figure 10c, the proposed method directly achieves gridless DOA estimation, avoiding the influence of pseudospectra on BTR trajectories. Throughout the entire BTR process, the proposed method consistently identifies both the source and interferer with a high probability, making it more effective for trajectory determination. Figure 11 illustrates the results of processing the Swellex-96 Event S59 data using different methods. The RMSE was calculated by interpolating the source trajectory with its GPS angles, and the data were processed using an AMD Ryzen 76800H CPU. The results show that the proposed method significantly outperforms the other three methods in terms of RP while maintaining RMSE values similar to the other methods. Additionally, the proposed method exhibits complexity comparable to MUSIC and is significantly more efficient than SPICE and SBL, which require iterative operations.

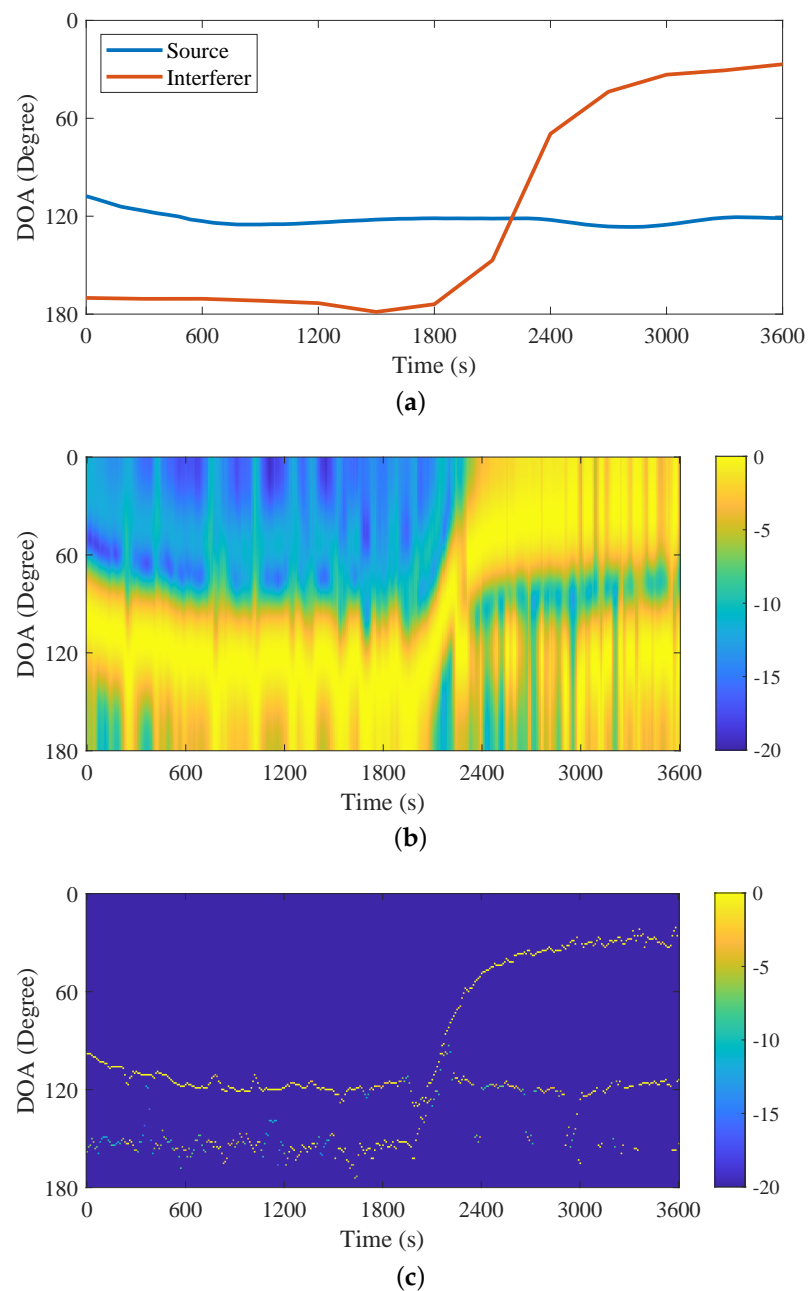


Figure 10. Cont.

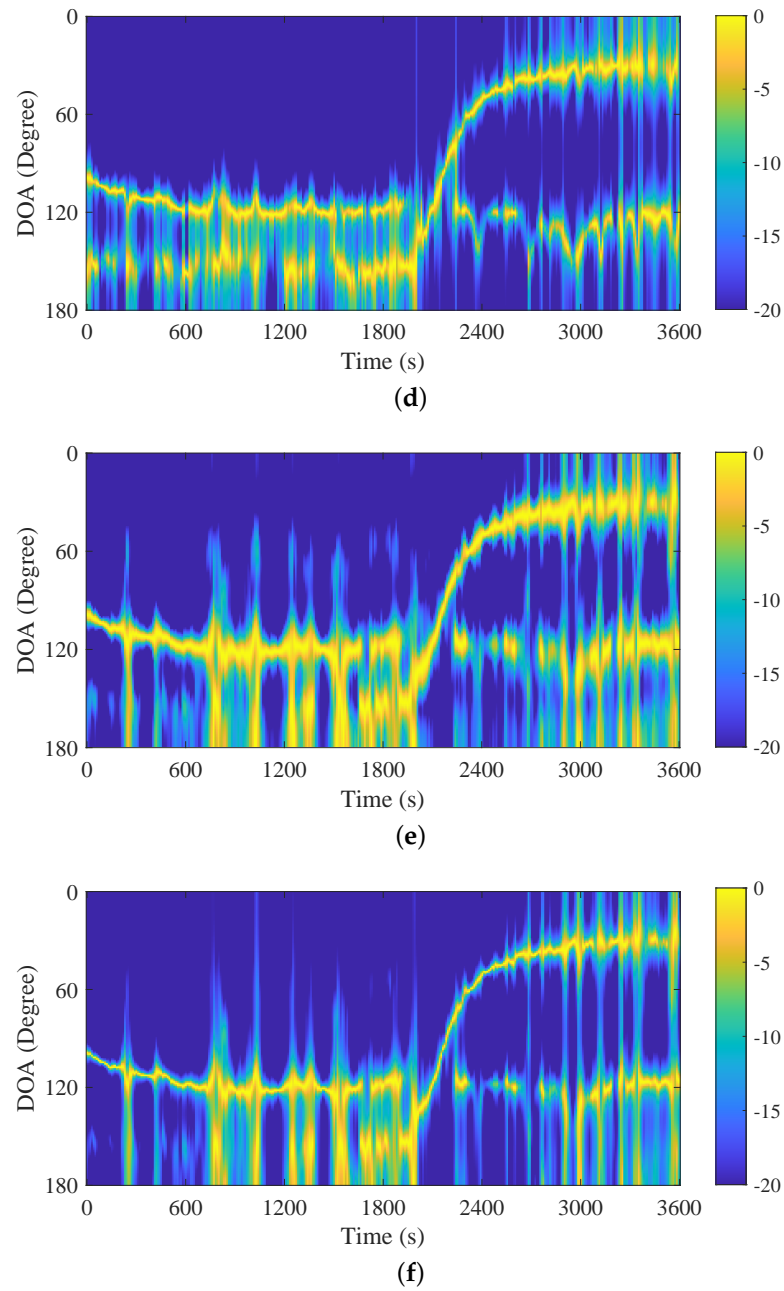


Figure 10. BTR results using different methods. (a) GPS. (b) CBF. (c) Proposed method. (d) MUSIC. (e) SPICE. (f) SBL.

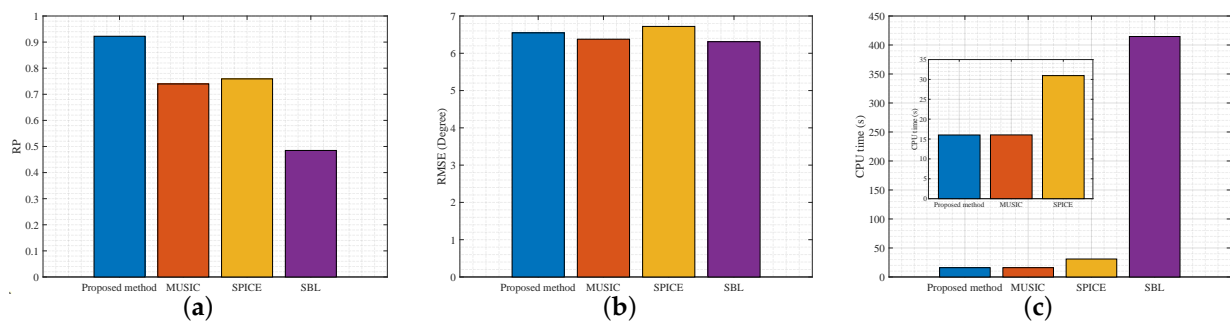


Figure 11. Processing results of Swellex-96 Event S59 data using different methods. (a) RP. (b) RMSE. (c) CPU time.

6. Conclusions and Discussion

In this paper, we propose a novel gridless DOA estimation method for arbitrary array geometries based on a CDNN. First, by expanding the steering vector into a Fourier series, the relationship between the element-domain covariance matrix and the angular-domain-covariance matrix for arbitrary arrays is derived. Then, a CDNN is designed and trained to reconstruct the Toeplitz-structured angular-domain covariance matrix from the SCM. Finally, Root-MUSIC is applied to the reconstructed angular-domain covariance matrix to obtain gridless DOA estimates. Our proposed method demonstrates superior resolution and DOA estimation accuracy compared to existing methods, particularly when target angles are closely spaced. The proposed method does not require the setting of any hyperparameters and exhibits high adaptability to different snapshot counts. Additionally, its lower computational complexity makes it well suited for applications that demand real-time performance or deployment on low-power platforms with limited computational resources. In this paper, we only analyze the performance of the proposed method for single-frequency narrowband signals. In our future work, we will continue to optimize the network structure and training process to make the proposed method applicable to multiple frequencies without needing prior information about the number of sources, thereby enhancing its practical utility in real-world scenarios, such as providing accurate angular estimates for localization and object tracking.

Author Contributions: Conceptualization, Y.C. and T.Z.; methodology, T.Z.; writing—review and editing, Y.C., T.Z. and Q.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (62371393, 62271397) and the Fundamental Research Funds for the Central Universities (D5000240231).

Data Availability Statement: Data are contained within this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Godara, L.C. Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations. *Proc. IEEE* **1997**, *85*, 1195–1245. [[CrossRef](#)]
2. Massa, A.; Rocca, P.; Oliveri, G. Compressive Sensing in Electromagnetics—A Review. *IEEE Antennas Propag. Mag.* **2015**, *57*, 224–238. [[CrossRef](#)]
3. Trees, H.L.V. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*; Wiley: Hoboken, NJ, USA, 2004; pp. 17–79.
4. Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE* **1969**, *57*, 1408–1418. [[CrossRef](#)]
5. Schmidt, R. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **1986**, *34*, 276–280. [[CrossRef](#)]
6. Roy, R.; Paulraj, A.; Kailath, T. ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise. *IEEE Trans. Acoust., Speech, Signal Process.* **1986**, *34*, 1340–1342. [[CrossRef](#)]
7. Thomas, J.K.; Scharf, L.L.; Tufts, D.W. The probability of a subspace swap in the SVD. *IEEE Trans. Signal Process.* **1995**, *43*, 730–736. [[CrossRef](#)]
8. Yang, Z.; Li, J.; Stoica, P.; Xie, L. Chapter 11 - Sparse methods for direction-of-arrival estimation. In *Array, Radar and Communications Engineering*; Chellappa, R., Theodoridis, S., Eds.; Academic Press: Cambridge, MA, USA, 2018; pp. 509–581.
9. Zhang, G.; Liu, K.; Sun, S.; Fu, J.; Wang, J. DOA estimation method for underwater acoustic signals based on two-dimensional power distribution (TPD) for few element array. *Appl. Acoust.* **2021**, *184*, 108352. [[CrossRef](#)]
10. Zhang, Y.; Zhang, L.; Han, J.; Ban, Z.; Yang, Y. A new DOA estimation algorithm based on compressed sensing. *Cluster Comput.* **2019**, *22*, 895–903.
11. Li, X.; Ma, X.; Yan, S.; Hou, C. Single snapshot DOA estimation by compressive sampling. *Appl. Acoust.* **2013**, *74*, 926–930. [[CrossRef](#)]
12. Stoica, P.; Babu, P.; Li, J. SPICE: A Sparse Covariance-Based Estimation Method for Array Processing. *IEEE Trans. Signal Process.* **2011**, *59*, 629–638. [[CrossRef](#)]
13. Tan, Z.; Yang, P.; Nehorai, A. Joint Sparse Recovery Method for Compressed Sensing with Structured Dictionary Mismatches. *IEEE Trans. Signal Process.* **2014**, *62*, 4997–5008. [[CrossRef](#)]
14. Wu, X.; Zhu, W.; Yan, J. Direction of Arrival Estimation for Off-Grid Signals Based on Sparse Bayesian Learning. *IEEE Sens. J.* **2014**, *16*, 2004–2016. [[CrossRef](#)]

15. Wang, Q.; Yu, H.; Li, J.; Chen, F. Adaptive Grid Refinement Method for DOA Estimation via Sparse Bayesian Learning. *IEEE J. Ocean Eng.* **2023**, *48*, 806–819. [[CrossRef](#)]
16. Jagannath, R.; Hari, K.V.S. Block Sparse Estimator for Grid Matching in Single Snapshot DoA Estimation. *IEEE Signal Process. Lett.* **2013**, *11*, 1038–1041. [[CrossRef](#)]
17. Xenaki, A.; Gerstoft, P. Grid-free compressive beamforming. *J. Acoust. Soc. Am.* **2015**, *137*, 1923–1935. [[CrossRef](#)]
18. Zhou, C.; Gu, Y.; Fan, X.; Shi, Z.; Mao, G.; Zhang, Y. Direction-of-Arrival Estimation for Coprime Array via Virtual Array Interpolation. *IEEE Trans. Signal Process.* **2018**, *66*, 5956–5971. [[CrossRef](#)]
19. Chu, Z.; Liu, Y.; Yang, Y.; Yang, Y. A preliminary study on two-dimensional grid-free compressive beamforming for arbitrary planar array geometries. *J. Acoust. Soc. Am.* **2021**, *149*, 3751–3757. [[CrossRef](#)]
20. Raj, A.G.; McClellan, J.H. Super-resolution DOA Estimation for Arbitrary Array Geometries Using a Single Noisy Snapshot. In Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019.
21. Raj, A.G.; McClellan, J.H. Single Snapshot Super-Resolution DOA Estimation for Arbitrary Array Geometries. *IEEE Signal Process. Lett.* **2019**, *26*, 119–123.
22. Yang, Y.; Chu, Z.; Yin, S. Two-dimensional grid-free compressive beamforming with spherical microphone arrays. *Mech. Syst. Signal Proc.* **2022**, *169*, 108642. [[CrossRef](#)]
23. Mahata, K.; Hyder, M.M. Grid-less T.V minimization for DOA estimation, Signal Processing. *Signal Process.* **2017**, *132*, 155–164. [[CrossRef](#)]
24. Grumiaux, P.; Kitić, S.; Girin, L.; Guérin, A. A survey of sound source localization with deep learning methods. *J. Acoust. Soc. Am.* **2022**, *152*, 107–151. [[CrossRef](#)]
25. Papageorgiou, G.K.; Sellathurai, M.; Eldar, Y.C. Deep Networks for Direction-of-Arrival Estimation in Low SNR. *IEEE Trans. Signal Process.* **2021**, *69*, 3714–3729. [[CrossRef](#)]
26. Liu, Y.; Chen, H.; Wang, B. DOA estimation based on CNN for underwater acoustic array. *Appl. Acoust.* **2021**, *172*, 107594. [[CrossRef](#)]
27. Nie, W.; Zhang, X.; Xu, J.; Guo, L.; Yan, Y. Adaptive Direction-of-Arrival Estimation Using Deep Neural Network in Marine Acoustic Environment. *IEEE Sens. J.* **2023**, *23*, 15093–15105. [[CrossRef](#)]
28. Liang, C.; Liu, M.; Li, Y.; Wang, Y.; Hu, X. LDnADMM-Net: A Denoising Unfolded Deep Neural Network for Direction-of-Arrival Estimations in A Low Signal-to-Noise Ratio. *Remote Sens.* **2024**, *16*, 554. [[CrossRef](#)]
29. Wu, X.; Yang, X.; Jia, X.; Tian, F. A Gridless DOA Estimation Method Based on Convolutional Neural Network with Toeplitz Prior. *IEEE Signal Process. Lett.* **2022**, *29*, 1247–1251. [[CrossRef](#)]
30. Gao, S.; Ma, H.; Liu, H.; Yang, J.; Yang, Y. A Gridless DOA Estimation Method for Sparse Sensor Array. *Remote Sens.* **2023**, *15*, 5281. [[CrossRef](#)]
31. Wu, X.; Wang, J.; Yang, X.; Tian, F. A Gridless DOA Estimation Method Based on Residual Attention Network and Transfer Learning. *IEEE Trans. Veh. Technol.* **2024**, *73*, 9103–9108. [[CrossRef](#)]
32. Cui, Y.; Yang, F.; Zhou, M.; Hao, L.; Wang, J.; Sun, H.; Kong, A.; Qi, J. Gridless Underdetermined DOA Estimation for Mobile Agents with Limited Snapshots Based on Deep Convolutional Generative Adversarial Network. *Remote Sens.* **2024**, *16*, 626. [[CrossRef](#)]
33. Huang, G.; Benesty, J.; Chen, J. On the Design of Frequency-Invariant Beampatterns with Uniform Circular Microphone Arrays. *IEEE-ACM Trans. Audio Speech Lang.* **2017**, *25*, 1140–1153. [[CrossRef](#)]
34. Tang, G.; Bhaskar, B.N.; Shah, P.; Recht, B. Compressed Sensing Off the Grid. *IEEE Trans. Inf. Theory* **2013**, *59*, 7465–7490. [[CrossRef](#)]
35. Wu, X.; Zhu, W.; Yan, J. A Toeplitz Covariance Matrix Reconstruction Approach for Direction-of-Arrival Estimation. *IEEE Trans. Veh. Technol.* **2017**, *66*, 8223–8237. [[CrossRef](#)]
36. Pal, P.; Vaidyanathan, P.P. A Grid-Less Approach to Underdetermined Direction of Arrival Estimation Via Low Rank Matrix Denoising. *IEEE Signal Process. Lett.* **2014**, *21*, 737–741. [[CrossRef](#)]
37. Rao, B.D.; Hari, K.V.S. Performance analysis of Root-Music. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 1939–1949. [[CrossRef](#)]
38. The SWelLEX-96 Experiment. 1996. Available online: <http://swellex96.ucsd.edu/> (accessed on 5 September 2024).
39. Trabelsi, C.; Bilaniuk, O.; Serdyuk, D.; Subramanian, S.; Santos, J.F.; Mehri, S.; Rostamzadeh, N.; Bengio, Y.; Pal, C.J. Deep Complex Networks. *arXiv* **2017**, arXiv:1705.09792.
40. Mohammadzadeh, S.; Nascimento, V.H.; Lamare, R.C.; Hajarolasvadi, N. Robust Beamforming Based on Complex-Valued Convolutional Neural Networks for Sensor Arrays. *IEEE Signal Process. Lett.* **2022**, *29*, 2108–2112. [[CrossRef](#)]
41. Zhang, Y.; Zeng, R.; Zhang, S.; Wang, J.; Wu, Y. Complex-Valued Neural Network with Multistep Training for Single-Snapshot DOA Estimation. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 1–5. [[CrossRef](#)]
42. Fan, Z.; Tu, Y.; Lin, Y.; Shi, Q. Class-Incremental Learning for Recognition of Complex-Valued Signals. *IEEE Trans. Cogn. Commun. Netw.* **2024**, *10*, 417–428. [[CrossRef](#)]
43. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
44. Xenaki, A.; Boldt, J.B.; Christensen, M.G. Sound source localization and speech enhancement with sparse Bayesian learning beamforming. *J. Acoust. Soc. Am.* **2018**, *143*, 2912–3921. [[CrossRef](#)]

45. Zhou, T.; He, Z.; Shi, Q.; Lin, C.; Zhang, S. Multisnapshot High-Resolution Gridless DOA Estimation for Uniform Circular Arrays. *IEEE Signal Process. Lett.* **2024**, *31*, 1705–1709. [[CrossRef](#)]
46. Liu, Z.; Zhang, Y.; Wang, W.; Li, X.; Li, H.; Shi, W.; Ali, W. Infinite Weighted p-Norm Sparse Iterative DOA Estimation via Acoustic Vector Sensor Array under Impulsive Noise. *J. Mar. Sci. Eng.* **2023**, *11*, 1798. [[CrossRef](#)]
47. Stoica, P.; Nehorai, A. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 720–741. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.