Journal of Marine Science and Engineering

MDPI

Article

A Deep Shrinkage Network for Direction-of-Arrival Estimation with Sparse Prior

Lei Zhou [1,2,3], Shihong Zhou [1,*], Yubo Qi [1], Lixin Wu [1,2] and Yannick Benezeth [3]

[1] State Key Laboratory of Acoustics and Marine Information, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China; zhoulei2021@mail.ioa.ac.cn (L.Z.); qyb@mail.ioa.ac.cn (Y.Q.); wlx@mail.ioa.ac.cn (L.W.)
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Laboratory ImViA, EA7535, University of Burgundy, 21078 Dijon, France; yannick.benezeth@u-bourgogne.fr
* Correspondence: shih_zhou@mail.ioa.ac.cn

Abstract: Direction-of-arrival (DOA) estimation of underwater multipath signals plays a indispensable role in both military and civilian underwater applications. Despite its importance, accurately estimating DOA under multipath conditions is challenging due to the proximity of paths in the spatial domain. Current methods struggle with this problem in passive detection scenarios. To address these limitations, this study proposes a deep learning (DL)-based DOA estimation framework leveraging sparse representation. First, the approach models the array covariance matrix as an undersampled linear measurement of the spatial spectrum. Then, a super-resolution deep shrinkage reconstruction network (SDSR-Net) is designed to map the sparse representation of the covariance matrix directly to the DOA. The network integrates a shrinkage module as nonlinear transformation layers, promoting sparsity and enhancing the discrimination of features. Simulations and experimental evaluations validated the effectiveness of the proposed method, showing that the DOA estimation accuracy was significantly improved and able to achieve a resolution of 0.2° in the spatial spectrum. Compared with existing methods, SDSR-Net achieved superior performance by effectively utilizing a sparsity prior, maintaining a high-resolution performance at signal-to-noise ratios higher than −10 dB. This work contributes a robust and efficient solution to DOA estimation challenges in underwater environments.

Keywords: direction-of-arrival estimation; deep learning; sparse representation

check for updates

Received: 4 January 2025
Revised: 20 January 2025
Accepted: 20 January 2025
Published: 22 January 2025

Citation: Zhou, L.; Zhou, S.; Qi, Y.; Wu, L.; Benezeth, Y. A Deep Shrinkage Network for Direction-of-Arrival Estimation with Sparse Prior. *J. Mar. Sci. Eng.* **2025**, *13*, 206. https://doi.org/10.3390/jmse13020206

Copyright: © 2025 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 a crucial topic in various fields such as radar, sonar, communications, and other signal processing applications [1–4]. It plays a vital role in array signal processing, where accurate localization of sources is essential for numerous applications. DOA estimation methods are generally categorized into three primary approaches: beamforming methods, subspace-based methods, and sparsity-based methods. However, in the context of underwater acoustic channels, DOA estimation presents significant challenges. The underwater environment is characterized by its complex, time–space–frequency-varying nature, which includes multipath effects, transmission fading, Doppler shifts, and propagation delays [5]. These factors make accurate DOA estimation even more challenging, particularly in environments with high interference or significant multipath propagation.

Previous methods have a relatively good performance in situations where the signal-to-noise ratio (SNR) is high, but the performance of DOA degrades when the SNR is low or

when faced with multipath scenarios. Many DOA estimation algorithms based on sparse representation and compressed sensing (CS) theories have been developed, leveraging the sparsity of signals in the spatial domain to achieve high-resolution DOA estimation [6,7]. Sparsity-based methods can provide accurate DOA estimates in the presence of coherent signals, but they are computationally intensive, making them impractical for real-time or resource-constrained applications. Moreover, in low SNR scenarios, the noise signals arriving from various directions disrupt the sparsity of the target multipath signals, significantly impairing the accuracy and resolution of compressed sensing algorithms. Recently, several solutions have been proposed utilizing deep learning (DL) [8–10]. These solutions employ various types of architectures; however, these networks primarily utilize fully connected layers and estimate source locations on coarse spatial grids.

Building on the above techniques, this paper proposes a novel approach for DOA estimation in underwater multipath environments using a super-resolution deep shrinkage reconstruction network (SDSR-Net). This method utilizes sparse signal recovery from quantized measurements and adapts thresholds to varying noise conditions, thereby improving the DOA estimation accuracy, particularly in low SNR scenarios. The proposed SDSR-Net is formulated as a multi-task model to estimate both the number of sound source paths and their corresponding arrival directions, achieving high-resolution DOA estimation across different angular intervals.

The rest of this paper is organized as follows. Section 2 introduces the related works about DOA estimation. Section 3 introduces the array signal model and formulates the problem using the sparse representation theory underlying the proposed method. Section 4 details the structure of the proposed DOA estimation neural network based on sparse features, along with the experimental setup. Section 5 presents the simulation and experimental results to validate the performance of the proposed SDSR-Net. Finally, Section 6 concludes the paper.

2. Related Work

Beamforming techniques, particularly conventional beamforming (CBF), are widely used for DOA estimation, due to their robustness against signal mismatch. However, CBF suffers from limitations in resolution, especially when dealing with side lobe leakage, due to a finite number of spatial sampling points determined by the Nyquist spatial theorem (NST). The minimum variance distortionless response (MVDR) algorithm [11] offers better resolution by minimizing the interference from undesired directions. However, MVDR's performance significantly degrades when the signals are highly correlated, which is a common challenge when estimating the arrival angles of coherent signals. Additionally, both CBF and MVDR struggle to separate closely spaced multipath signals, particularly when small-aperture arrays are used.

Subspace-based methods such as multiple signal classification (MUSIC) [12,13] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [14,15] are popular alternatives for high-resolution DOA estimation. These methods exploit the orthogonality between the signal and noise subspaces to achieve better resolution. However, their performance deteriorates under low SNR conditions, and they are also less effective at resolving closely located sources. In underwater environments, traditional subspace methods often fail to provide accurate DOA estimates in the presence of multipath signals. Therefore, several subspce-based methods combined with deep learning have been proposed for DOA estimation [16,17].

More recently, the application of CS-based techniques for DOA estimation has gained attention, especially by exploiting the sparsity of signals in the spatial domain. There are several CS-based methods, such as L1 singular value decomposition (L1-SVD) [6],

orthogonal matching pursuit (OMP), and sparse Bayesian learning (SBL) [18]. L1-SVD introduces a CS-based framework that reformulates the nonconvex l_0-norm optimization problem into a convex l_1-norm minimization problem, improving computational efficiency. Meanwhile, SBL employs a Bayesian framework to incorporate sparsity by using a sparse prior distribution for the signal [19–21]. While sparsity-based methods offer improved resolution over beamforming and subspace methods, their performance is hindered at low SNRs, where noise disrupts the sparsity of the target signal. In recent years, deep-learning-based approaches have emerged as a promising avenue for DOA estimation, particularly in complex environments [22,23]. Refs. [24,25] introduced three neural network models specifically designed for DOA estimation with covariance reconstruction. Similarly, Ref. [26] proposed a deep neural network (DNN) framework for super-resolution DOA estimation. The algorithms proposed in [27,28] are based on multipath DOA estimation problems. However, these networks estimate source locations on coarse spatial grids.

3. Problem Formulation

3.1. Signal Model

Assume that the uniform linear array (ULA) consists of M elements, with an inter-element spacing of d. Let there be L independent narrowband signals incident on the array from directions $\theta_1, \theta_2, \cdots, \theta_L$. Using the first array element as a reference, each narrowband signal is represented as $s_1(n), s_2(n), \cdots, s_L(n)$. The signal received by the m-th array element can then be expressed as

$$y_m(n) = \sum_{l=1}^{L} s_l(n) e^{-j2\frac{\pi}{\lambda}(m-1)d\sin\theta_l} + \delta_m(n), \tag{1}$$

where $\lambda = c/f$ denotes the wavelength of the signals at carrier frequency f, and $\delta_m(n)$ represents the noise at the m-th array element. The received signal can be expressed as

$$\boldsymbol{y}(n) = \boldsymbol{A}(\theta)\boldsymbol{s}(n) + \boldsymbol{\delta}(n), \tag{2}$$

where $\boldsymbol{s}(n) = [s_1(n), \cdots, s_L(n)]^{\mathrm{T}} \in \mathbb{C}^{L\times 1}$ represents the signal vector, $[\cdot]^{\mathrm{T}}$ denotes the transposition operator, and $\boldsymbol{\delta}(n) = [\delta_1(n), \cdots, \delta_M(n)]^{\mathrm{T}} \in \mathbb{C}^{M\times 1}$ denotes the noise vector. The matrix $\boldsymbol{A}(\theta) = [\boldsymbol{a}(\theta_1), \cdots, \boldsymbol{a}(\theta_L)] \in \mathbb{C}^{M\times L}$ is the direction matrix, where $\boldsymbol{a}(\theta_i)$ is the direction vector for the i-th signal, which is expressed as

$$\boldsymbol{a}(\theta_i) = [1, e^{-j\frac{2\pi}{\lambda}d\sin\theta_i}, \cdots, e^{-j\frac{2\pi}{\lambda}(M-1)d\sin\theta_i}]^{\mathrm{T}}. \tag{3}$$

By collecting T time-sampling snapshots, Equation (2) can be reformulated as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{S} + \boldsymbol{\delta}, \tag{4}$$

where $\boldsymbol{Y} = [\boldsymbol{y}(0), \ldots, \boldsymbol{y}(T-1)]$, $\boldsymbol{A} = \boldsymbol{A}(\theta)$, $\boldsymbol{S} = [\boldsymbol{s}(0), \ldots, \boldsymbol{s}(T-1)]$, and $\boldsymbol{\delta} = [\boldsymbol{\delta}(0), \ldots, \boldsymbol{\delta}(T-1)]$. The objective is to estimate the DOA of each narrowband signal from the received signal $\boldsymbol{Y}$.

In the case of a single narrowband signal incident on the array $(L = 1)$, the problem is referred to as single-source DOA estimation. Under this scenario, $\boldsymbol{A}(\theta) = [\boldsymbol{a}(\theta_1)] \in \mathbb{C}^{M\times 1}$ and $\boldsymbol{s}(n) = [\boldsymbol{s}_1(n)]^{\mathrm{T}} \in \mathbb{C}^{1\times 1}$. When multiple narrowband signals are incident on the array simultaneously $(L > 1)$, the problem is referred to as multi-source DOA estimation.

Traditional DOA estimation methods typically rely on a covariance matrix for estimation. Assuming that each noise component is independently and identically distributed additive white Gaussian noise (AWGN), the array covariance matrix can be expressed as

$$\boldsymbol{R}_{yy} = \mathbb{E}[\boldsymbol{y}(n)\boldsymbol{y}^{\mathrm{H}}(n)] = \boldsymbol{A}\sigma_s^2\boldsymbol{A}^{\mathrm{H}} + \sigma_n^2\boldsymbol{I}, \tag{5}$$

where $\sigma_s^2 = \mathbb{E}[\boldsymbol{s}(n)\boldsymbol{s}^{\mathrm{H}}(n)]$ denotes the signal power, $\sigma_n^2 = \mathbb{E}[\boldsymbol{\delta}(n)\boldsymbol{\delta}^{\mathrm{H}}(n)]$ denotes the noise power, $\boldsymbol{I}$ is the identity matrix, and $[\cdot]^{\mathrm{H}}$ represents the conjugate transpose. The mapping between $\boldsymbol{R}_{yy}$ and $\boldsymbol{\theta}$ is given by

$$f(\theta_1, \theta_2, \ldots, \theta_L) = \boldsymbol{R}_{yy}. \tag{6}$$

3.2. Sparse Representation

To construct a sparse measurement matrix for compressed sensing, the spatial plane $[-90°, 90°]$ is discretized into a set of directions $\Theta = [\Phi_1, \Phi_2, \cdots, \Phi_L]^{\mathrm{T}}$, where $L \gg K$, with a sampling interval of $\Delta\Phi$. The true signal source direction θ lies within Φ, ensuring minimal quantization error. The array-received signal can then be expressed in an overcomplete form:

$$y(t) = \sum_{l=1}^{L} \boldsymbol{a}(\phi_l)\overline{s_l}(t) + n(t) \tag{7}$$

where $\bar{s}_l(t) = s_k(t)$ if $\Phi_l = \theta_k$, and $\bar{s}_l(t) = 0$ otherwise. The spatial covariance matrix is given by

$$\boldsymbol{R}_{yy} = \mathbb{E}[\boldsymbol{y}(n)\boldsymbol{y}^{\mathrm{H}}(n)] = \sum_{l=1}^{L} \eta_l \boldsymbol{a}(\Phi_l)\boldsymbol{a}^{\mathrm{H}}(\Phi_l) + \sigma_n^2\boldsymbol{I} \tag{8}$$

where $\eta_l = \mathbb{E}[\bar{s}_l(t)\bar{s}_l^{\mathbf{H}}(t)]$ represents the signal energy in the lth direction, and $\eta = [\eta_1, \eta_2, \cdots, \eta_L]^{\mathrm{T}}$ is nonzero only at the true source locations, making the spatial spectrum η sparse. DOA estimation is achieved by recovering η from $\boldsymbol{R}_{yy}$.

The mth row of $\boldsymbol{R}_{yy}$ can be expressed as

$$x_m = \boldsymbol{A_m}\boldsymbol{\eta} + \sigma_n^2\boldsymbol{e_m} \tag{9}$$

where $\boldsymbol{A_m} = [\boldsymbol{a}(\Phi_1)\boldsymbol{a}^{\mathrm{H}}(\Phi_1)\boldsymbol{e_m}, \cdots, \boldsymbol{a}(\Phi_L)\boldsymbol{a}^{\mathrm{H}}(\Phi_L)\boldsymbol{e_m}]$, and $\boldsymbol{e_m}$ is an $M \times 1$ vector with the mth element set to 1 and others to 0. Expanding Equation (9), the observed signal after sparse representation becomes

$$\tilde{\boldsymbol{X}} = \tilde{\boldsymbol{A}} \cdot \boldsymbol{\eta} + \sigma_n^2\tilde{\boldsymbol{e}} \tag{10}$$

where $\tilde{\boldsymbol{X}} = [x_1; x_2; \cdots; x_M]$, $\tilde{\boldsymbol{A}} = [A_1; A_2; \cdots; A_M]$, $\tilde{\boldsymbol{e}} = [e_1; e_2; \cdots; e_M]$. Equation (10) demonstrates that recovering the spatial spectrum η from $\tilde{\boldsymbol{X}}$ is a classic sparse linear inverse problem. The relationship between $\tilde{\boldsymbol{X}}$ and $\boldsymbol{\eta}$ can be expressed as

$$\boldsymbol{\eta} = f^{-1}(\tilde{\boldsymbol{A}}^{\mathrm{T}}\tilde{\boldsymbol{X}}) \tag{11}$$

While various sparsity-driven methods exist to solve this problem, they often face challenges such as a high computational complexity and instability under low signal-to-noise ratio (SNR) conditions. To address these limitations, we propose a deep convolutional network that directly learns the mapping from $\tilde{\boldsymbol{X}}$ to $\boldsymbol{\eta}$. This approach leverages the representational power of deep learning to enable accurate and efficient DOA estimation, even in challenging scenarios.

3.3. Theory of the Proposed SDSR-Net

Leveraging the powerful nonlinear processing capabilities of deep learning, a DL-based method is introduced to construct a covariance matrix from the array output covariance matrix. Through network training, a mapping model between the sparse representation of the covariance matrix and the sparse multipath DOA is iteratively optimized.

We propose a SDSR network for estimating sparse target arrival angles, with the network structure depicted in Figure 1. The proposed DOA estimation method utilizes simulated data to learn the mapping relationship between the covariance matrix and the DOA, enabling effective DOA estimation in two steps: (1) Substitute the received signal into Equations (8)–(10) to compute $\tilde{A}$ and $\tilde{X}$. (2) Use $\tilde{A}^{\mathrm{T}}\tilde{X}$ as the input to the SDSR network. The network extracts the target spatial features through multiple layers and outputs the target angle. The mapping model between the input of SDSR and the target angle is represented by Equation (11).

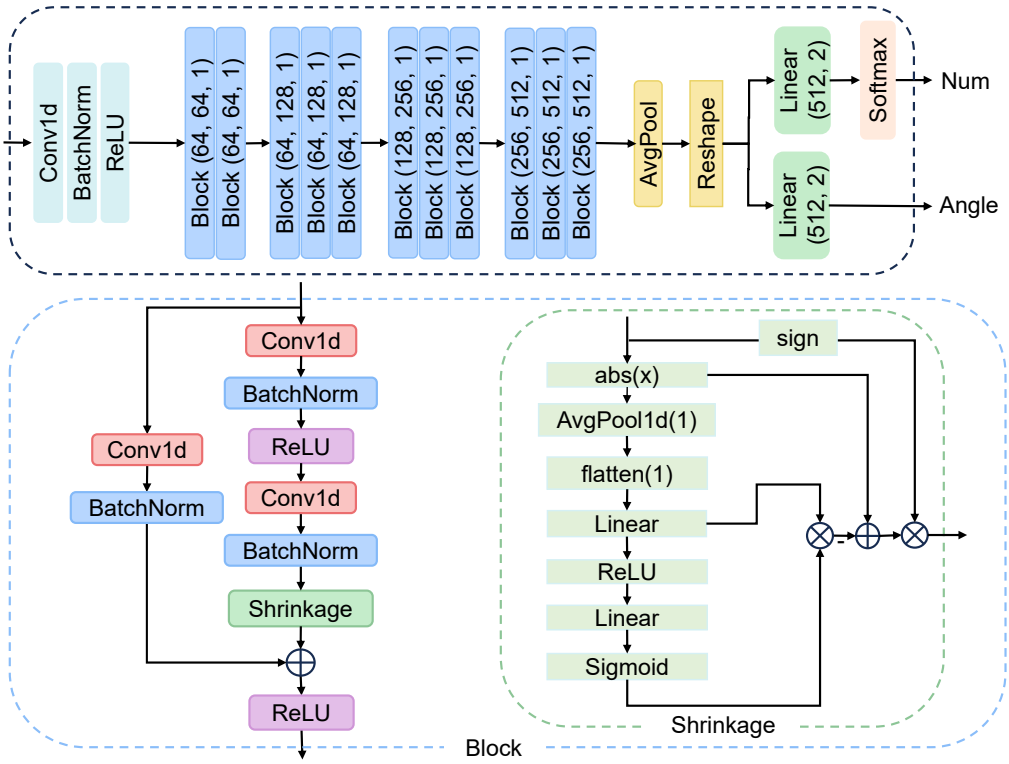

Figure 1. Overall architecture of the proposed SDSR-Net.

To recover the sparse spatial vector $\boldsymbol{\eta}$ from the covariance matrix $\tilde{X}$ in Equation (10), the problem of estimating $\boldsymbol{\eta}$ is formulated as the following optimization problem:

$$\min_{\boldsymbol{\eta}} \|\boldsymbol{\eta}\|_1 + \lambda\|[\tilde{X} \odot \tilde{A}\boldsymbol{\eta}]_-\| - 1, \quad s.t. \|\boldsymbol{\eta}\|_2 = 1 \tag{12}$$

Here, $[\cdot]_-$ denotes the negative function, defined as $[z]_- = \max(0, -z)$, and $\odot$ represents the element-wise product.

The above problem is solved using an iterative method. The first step involves applying gradient descent to minimize the cost associated with the barrier function. By computing the gradient of $\|\boldsymbol{\eta}\|_1$ and $\|[\tilde{X} \odot \tilde{A}\boldsymbol{\eta}]_-\| - 1$ with respect to $\boldsymbol{\eta}$, the gradient at the k-th iteration is given by

$$g(\boldsymbol{\eta})_k = \tilde{A}^{\mathrm{T}}(\tilde{A}\boldsymbol{\eta}_{k-1} - \tilde{X}). \tag{13}$$

Before applying gradient descent, the gradient is projected onto a unit sphere centered at the origin, to enforce the constraints.

$$\overline{g}(\boldsymbol{\eta})_k = g(\boldsymbol{\eta})_k \langle g(\boldsymbol{\eta})_k, \boldsymbol{\eta}_{k-1} \rangle \boldsymbol{\eta}_{k-1} \tag{14}$$

Consequently, the gradient descent step can be expressed as

$$\overline{\boldsymbol{\eta}}_k = \boldsymbol{\eta}_{k-1} - \tau \overline{g}(\boldsymbol{\eta})_k \tag{15}$$

where τ is the descent step size.

The second step involves a shrinkage operation that applies a soft-thresholding function to reduce the magnitudes of the non-zero components of the sparse signal. The shrinkage operator is defined as follows:

$$\boldsymbol{\eta}_k = \text{sign}(\overline{\boldsymbol{\eta}}_k) \odot \max(|\overline{\boldsymbol{\eta}}_k| - \frac{\tau}{\lambda}, 0). \tag{16}$$

The sparsity level of $\boldsymbol{\eta}$ is enhanced through the application of soft-thresholding shrinkage. By iteratively applying this operator, the vector progressively becomes sparser. At each iteration, the estimates are normalized to ensure they satisfy unit-norm constraints. In the proposed network, sparsification is achieved via a shrinkage module that adaptively learns the soft threshold. Multiple residual modules further refine the sparsity, improve resolution, and suppress noise.

4. Proposed Method

4.1. Architecture of the SDSR-Net

The overall architecture of the proposed SDSR network is illustrated in Figure 1. The input data are first processed by a 1D convolutional (Conv1d) layer with a kernel size of 3, a stride of 1, and padding of 1. Based on the output of the Conv1d layer, the DOA features are further extracted through a batch normalization (BatchNorm) layer and rectified linear unit (ReLU) activation. The extracted features are then passed through several deep residual shrinkage blocks. As shown at the bottom of Figure 1, the residual blocks are the core components of this network. Each block consists of three Conv1d layers, three BatchNorm layers, and two ReLU layers. The key feature of residual blocks is the skip connection, which distinguishes them from conventional convolutional networks. In typical convolutional networks, the gradients of the cross-entropy error are back-propagated layer-by-layer. However, with skip connections, the gradients flow more effectively to the earlier layers, closer to the input layer, enabling more efficient parameter updates.

Furthermore, rather than manually designing the threshold value, deep learning allows learning it automatically. Consequently, the combination of soft thresholding and deep learning presents a promising approach for eliminating noise-related information and constructing highly discriminative features. As shown in Equation (16), a shrinkage module was designed to adaptively learn the soft threshold and remove redundant information from the input data based on this threshold. In traditional signal denoising algorithms, determining an appropriate threshold value can be challenging, and its optimal value often varies across different scenarios. To address this issue, the thresholds in the SDSR-Net are automatically determined within the deep architecture, eliminating the need for manual intervention. The output of the deep residual shrinkage block is followed by an average pooling (AvgPool) layer and a reshape layer. Subsequently, the data are passed through two branching networks. One branch extracts the target angles through a linear layer, with its output representing the estimated angle rather than the spatial spectrum, in order to

mitigate the off-grid problem. The other branch processes the features through a linear layer and then computes the probability of the number of target angles using a softmax function.

It is important to note that the proposed method directly derives the angle value of the target from the input spatial feature, rather than computing the entire spatial spectrum. This approach significantly reduces the model's learning complexity. Furthermore, it avoids the off-grid problem commonly encountered in many DOA estimation methods that rely on dividing the computational grid.

4.2. Dataset

To train the proposed SDSR-Net, simulated data from a 16-element uniform linear array (ULA) were used. The direction-of-arrival was defined on a grid with a spatial resolution of 1° and a maximum angle of $\theta_{\max} = 90^\circ$, forming a grid set $D = \{0^\circ, 1^\circ, \ldots, 90^\circ\}$, which included $|D| = 91$ grid points. One of the primary advantages of the proposed SDSR-Net is its ability to infer the number of signal sources, as the task is framed as a multi-class classification problem. Consequently, the training dataset was divided into two parts: (1) Two-source data: This subset modeled multi-path signals from the sea surface. (2) Single-source data: This subset modeled single-path signals, helping to mitigate overfitting to multi-path data.

The number of signal sources was denoted as K. For the two-source data, pairs of DOAs were generated from all possible combinations of D, resulting in $\binom{|D|}{K}$ pairs. A covariance matrix for each pair was computed as the input to the model, with the corresponding angles used as labels. For the two-source dataset, 4095 samples were generated. For the single-source dataset, 91 pairs of DOAs were generated, but to balance the sample count, the data were resampled to obtain 3000 samples. This resulted in a total of 7095 training samples.

The training signal-to-noise ratio (SNR) ranged from -10 dB to 10 dB. Since the actual SNR of the received signals was unknown, the model was trained using a range of SNR values. The training data were synthesized by calculating a covariance matrix based on on-grid DOAs, while the test dataset used off-grid DOAs. The angular spacing of the training grid was 1°, whereas the simulated arrival angles for testing were generated with finer spacing. Specifically, the first path's arrival angle was defined as $D_1 = 0^\circ : 0.2^\circ : 80^\circ$, and the angular difference between the first and second paths was $\Delta\theta = 0.2^\circ : 0.2^\circ : 10^\circ$. This resulted in a total of 20,050 pairs of arrival angles for testing. The simulation parameters are listed in Table 1.

Table 1. Simulation parameter settings.

Parameter	Value
Frequency (Hz)	100
Array interval (m)	7.5
Snapshot	1
Number of elements	16
Reference sound speed (m/s)	1500
SNR (dB)	−10–1

4.3. Training Setup

The training batch size was set to 256, and the model was trained for 500 epochs. For optimization, we employed the Adam optimizer with an initial learning rate of 1.0×10^{-4}, along with the settings $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1.0 \times 10^{-8}$. The model was implemented using PyTorch 2.2.2 and trained on an NVIDIA RTX 4090D GPU for efficiency.

We optimized the trainable parameters of the SDSR-Net by using the training dataset and loss function. The final output of the model included the number of target paths,

$Num \in \{(0.5, 0.5), (1, 0), (0, 1)\}$, which corresponded to 0, 1, or 2 paths, respectively, and the corresponding angles of arrival, $Angle \in R^{1\times 2}$. If Num is $[0.5, 0.5]$, this indicates that there is no target, and the corresponding $Angle$ is labeled as $[-1, -1]$. If Num is $[1, 0]$, this indicates a single target, and the corresponding $Angle$ is $[\theta_1, -1]$; if Num is $[0, 1]$, this indicates two targets, and the corresponding $Angle$ is $[\theta_1, \theta_2]$. Therefore, we used different loss functions for the Num and $Angle$ estimations. The mean squared error (MSE) loss was used for angle estimation, while for Num estimation, which can be treated as a classification task, we applied cross-entropy loss.

5. Results and Analysis

In this section, we present extensive simulation results to verify and evaluate the performance of our proposed DOA estimation model under various scenarios. The evaluation includes three main aspects: (1) a comparison of the proposed network with existing DOA estimation algorithms, (2) an analysis of the DOA estimation accuracy for signals arriving at different spatial angles and spatial intervals, and (3) an assessment of the model's performance under varying SNR conditions.

For both training and testing, a ULA with 16 elements, spaced at half-wavelength intervals, was employed. This section provides an intuitive comparison of the DOA estimation results against baseline methods, including (1) classical beamforming (CBF), (2) minimum variance distortionless response (MVDR), (3) multiple signal classification (MUSIC), (4) orthogonal matching pursuit (OMP) [29], (5) sparse Bayesian learning (SBL) [19], (6) deep convolutional network (DCN) [30], (7) DeepFPC [31], (8) DA-MUSIC [17], and (9) SubspaceNet [16]. Among these, (1)–(3) are traditional methods, (4)–(5) are compressive sensing (CS)-based methods, and (6)–(9) are deep-learning-based methods.

5.1. Comparison of Different DOA Estimation Methods

In this subsection, we evaluate the proposed SDSR-Net with two sources positioned at three different angle intervals: 2.2°, 5.2°, and 8.9°, respectively. The angle of arrival for the first source varied from 0° to 80°, while the DOA of the second source spanned from 2.2° to 82.2°, from 5.2° to 85.2°, and from 8.9° to 88.9°, respectively, all sampled in steps of 0.2°. Each scenario used samples with an SNR of 10 dB, and for each angle interval, 401 pairs of angles were collected for DOA estimation. These results enable a performance comparison of the proposed SDSR-Net against existing methods. The corresponding outcomes are presented in Figures 2 and 3.

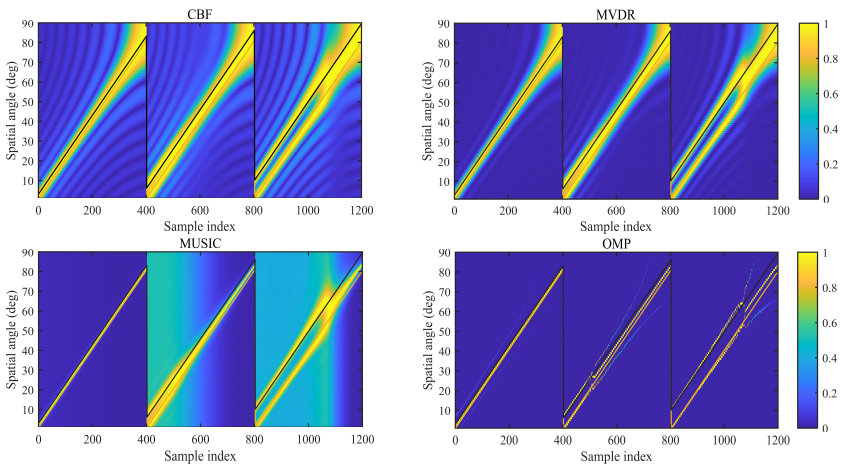

Figure 2. *Cont.*

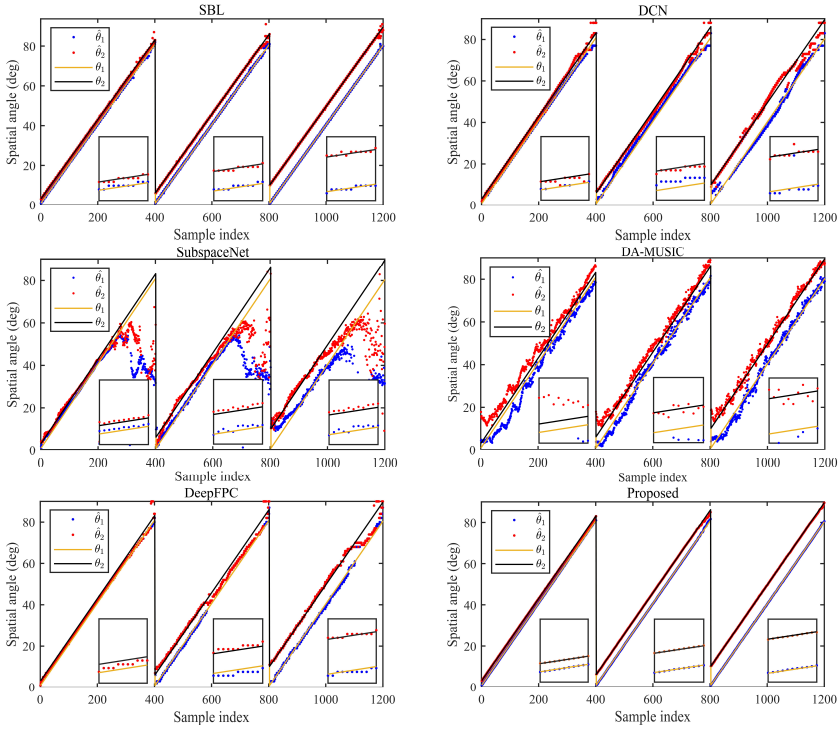

Figure 2. Comparison of DOA estimation methods with two off-grid sources with the angle intervals of 2.2°, 5.2°, and 8.9°, respectively.

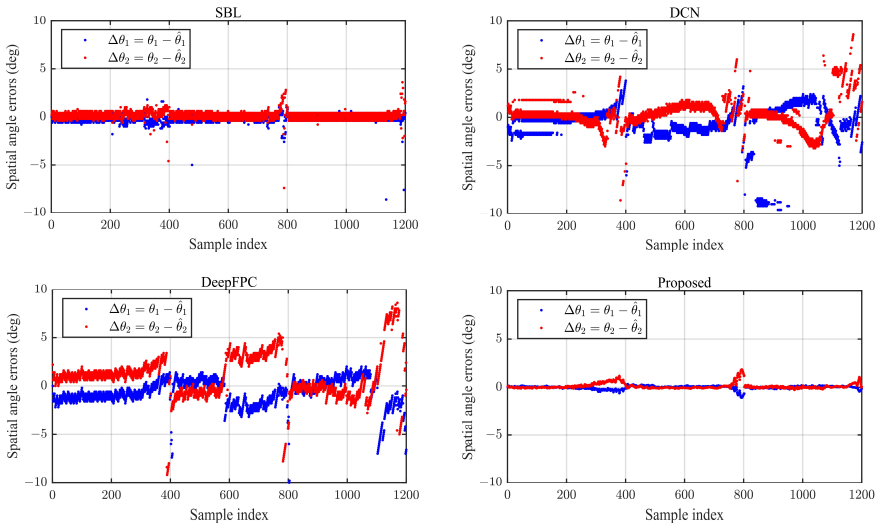

Figure 3. Comparison of the corresponding errors of off-grid sources with angle intervals of 2.2°, 5.2°, and 8.9°, respectively.

Figure 2 presents the DOA estimation results across the evaluated scenarios using conventional methods, CS-based algorithms, and DL-based models. The straight lines in the figure represent the true DOA values for the two sources. The results reveal that the conventional algorithms, including CBF, MVDR, and MUSIC, suffered from wide beamwidths, leading to a poor spatial resolution. Compared with the high-resolution methods such as SBL, DCN, SubspaceNet, DA-MUSIC, and DeepFPC, the proposed SDSR-Net demonstrated superior performance in DOA estimation. However, SBL was constrained by significant off-grid issues, limiting its estimates to predefined grid points. When the signal angle exceeded 60°, the performance of the SubspaceNet declined significantly. Furthermore, the DCN and DeepFPC methods showed higher estimation errors compared with SDSR-Net, particularly for signals arriving at larger angles.

Figure 3 shows the estimation errors of four high-resolution algorithms compared to the true values, offering deeper insights into their performance. It is evident that the overall errors of the proposed SDSR-Net were the smallest, while the other deep-learning-based methods showed larger errors. Moreover, all methods displayed significant errors near the edges of the angular interval. Notably, the CNN and DNN algorithms exhibited substantial error variance, with deviations reaching up to $\pm 10^\circ$. Outside the edge regions, the error for SBL was constrained within $\pm 0.8^\circ$, while the proposed SDSR-Net achieved markedly lower errors, limited to within $\pm 0.2^\circ$.

5.2. Performance at Different SNRs

This subsection evaluates the statistical performance of the proposed SDSR-Net using Monte Carlo trials, comparing its performance with CBF, MVDR, MUSIC, OMP, SBL, DeepFPC, and DCN. The SNR of the signal ranged from -10 dB to 20 dB in 2 dB intervals, to assess the DOA estimation capabilities of each method. Root mean square error (RMSE) was employed as the performance metric:

$$RMSE = \sqrt{\frac{1}{D \cdot K} \sum_{d=1}^{D} \sum_{k=1}^{K} (\theta_k^d - \hat{\theta}_k^d)^2} \tag{17}$$

where D, K, θ_k^d, $\hat{\theta}_k^d$ represents the number of Monte Carlo trails, number of angles, the true value of the k-th angle, and the estimated value of the k-th angle. A smaller value of $RMSE$ indicates that the angle predicted by the model was closer to the true value.

To ensure the test dataset was distinct from the training dataset, off-grid angles were used for testing. The angle θ was defined to range from 0.5° to 60.5° in 1° increments. To differentiate the test set further, the angle of the first source was specified as $\theta_1 = \theta + \mathrm{randn}(1, 61)$. For the second source, 20 distinct angle intervals were sampled, ranging from 0.5° to 10° in 0.5° steps. A total of 100 Monte Carlo trials were conducted, generating 122,000 signals. The average performance results on the test dataset are illustrated in Figure 4.

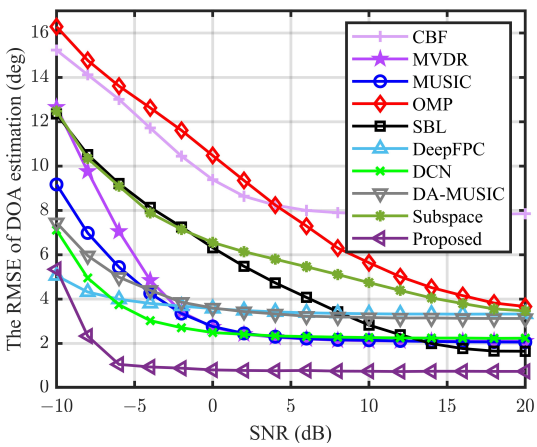

Figure 4. The RMSE of each method versus SNR.

Figure 4 presents the RMSE of DOA estimation as a function of SNR. The proposed SDSR network consistently achieved the lowest RMSE across the SNR range from -10 to 20 dB. This superior performance is attributed to the network's ability to effectively separate target features from noise. At high SNRs, the performance of the CS-based algorithms, such as the OMP and SBL methods, declined with the SNR, due to the limitations of their computational grids and signal sparsity. In contrast, the errors of the other methods remained relatively stable. In the SNR range between -5 dB and 0 dB, the performance of

all methods, except the SDSR network, significantly deteriorated as the SNR decreased. Remarkably, the proposed SDSR network maintained a robust performance even in low SNR scenarios, with noticeable degradation only when the SNR dropped below −5 dB. Even in such cases, it achieved the smallest DOA estimation errors among all evaluated methods.

To gain deeper insights into the DOA estimation performance of the different methods, we selected two sets of test data with varying angular intervals. The first set comprised angles of $(43.55°, 47.05°)$ with an angular interval of 3.5°, while the second set included angles of $(37.15°, 43.65°)$ with a wider angular interval of 6.5°. The normalized bearing recordings for each method are illustrated in Figures 5 and 6.

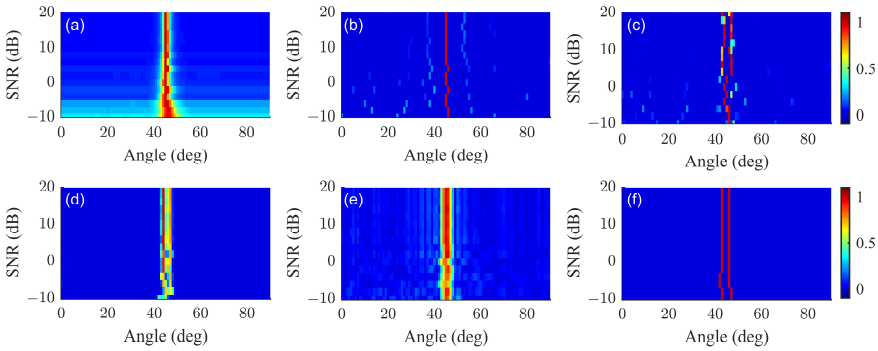

Figure 5. DOA estimation in the inter-source angle 3.5° case with a SNR range of −10 dB to 20 dB. (**a**) MUSIC, (**b**) OMP, (**c**) SBL, (**d**) DCN, (**e**) DeepFPC, (**f**) SDSR.

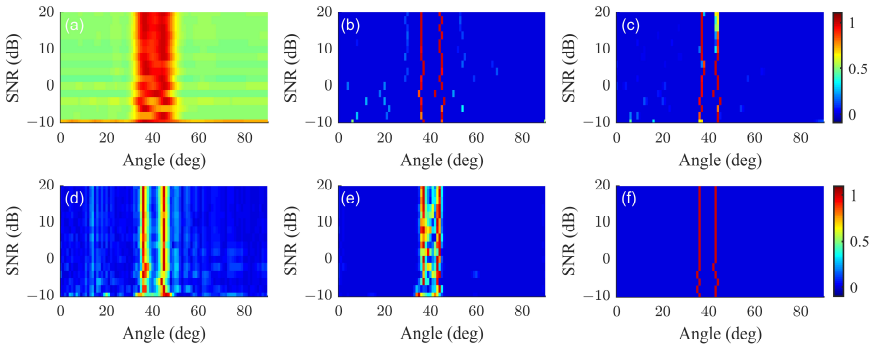

Figure 6. DOA estimation in the inter-source angle 6.5° case with a SNR range of −10 dB to 20 dB. (**a**) MUSIC, (**b**) OMP, (**c**) SBL, (**d**) DCN, (**e**) DeepFPC, (**f**) SDSR.

From Figure 5, it is evident that only SBL, DCN, and the proposed SDSR successfully achieved separation at an angular interval of 3.5°. However, both SBL and DCN struggled to resolve the two signal estimations effectively under low SNR conditions. In contrast, the SDSR method demonstrated a more distinct separation of the two targets in the spatial domain compared to DCN. Figure 6 illustrates that, at an angular interval of 6.5°, all methods were capable of accurate DOA estimation at high SNRs. The OMP achieved a resolution comparable to DeepFPC but exhibited less stability as the SNR decreased. The SDSR method consistently provided high-precision DOA estimation for both small and large angular interval targets, whereas the other methods failed to achieve such precision for small angular intervals, due to resolution limitations and the impact of the low SNR conditions.

5.3. Performance at Different Angle Separations

This subsection evaluates the performance of the proposed SDSR network under varying angular intervals between the two sources and compares the RMSE across each interval with the OMP, SBL, and DCN methods. The experiments were conducted at an

SNR of 10 dB, with angular intervals ranging from 0.2° to 10° in steps of 0.2°. For each scenario, the RMSE was computed as the average over 401 samples.

Figure 7 illustrates the RMSE of the DOA estimation for the two sources. The angle of the first source was 20°, the angle of the second source was larger than the first source. The results show that OMP, SBL, and the proposed SDSR methods achieved similar RMSEs at large angle intervals. However, the performance of OMP and SBL degraded rapidly as the angle interval decreased. Notably, while both the DCN and SDSR methods could accurately estimate the DOAs in scenarios with small angle intervals, the SDSR method consistently achieved a lower RMSE.

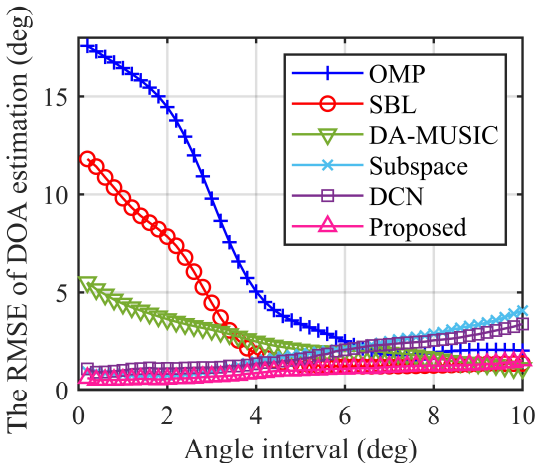

Figure 7. Performance of OMP, SBL, DCN, DA-MUSIC, SubspaceNet, and the proposed SDSR network on two sources with angle intervals from 0.2° to 10°.

To further evaluate the performance, we compared the effect of the angle on the performance of the proposed method against other methods, using the same angle interval. The angle interval was set to 4°, with an SNR of 10 dB. The angle of the first source ranged from 0° to 80°, with a step size of 0.2°. For each value of θ, the RMS error was averaged over 50 samples. The results are shown in Figure 8.

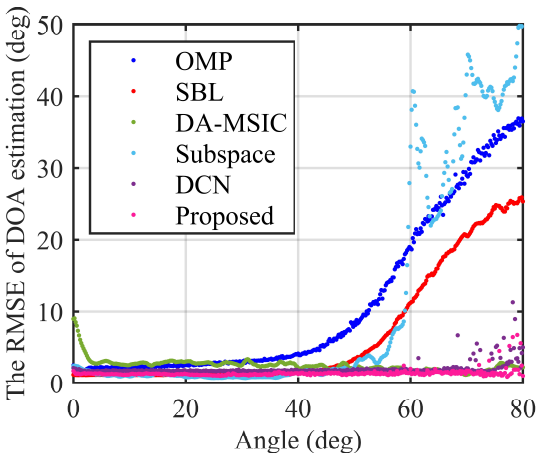

Figure 8. Performance of OMP, SBL, DCN, DA-MUSIC, SubspaceNet, and the proposed SDSR network on two sources with an angle interval of 4°.

Figure 8 illustrates that all methods exhibited error growth in the angular edge region. However, when the estimated angle exceeded 60°, the RMSE of the compressed sensing-based OMP and SBL algorithms increased rapidly, while the deep-learning-based algorithms maintained relatively stable errors. In the range of $[0°, 80°]$, the RMSE of our proposed SDSR network remained small, ranging between $[1.01°, 6.74°]$, while the

RMSE of the DCN method ranged between $[1.66°, 11.29°]$. In the angular edge region, the OMP and SBL methods performed worse, with RMSE values between $[1.61°, 36.9°]$ and $[1.13°, 25.91°]$, respectively.

To further demonstrate the DOA estimation performance of the proposed method, Figure 9 shows a spatial spectrum comparison in one-source and two-source scenarios.

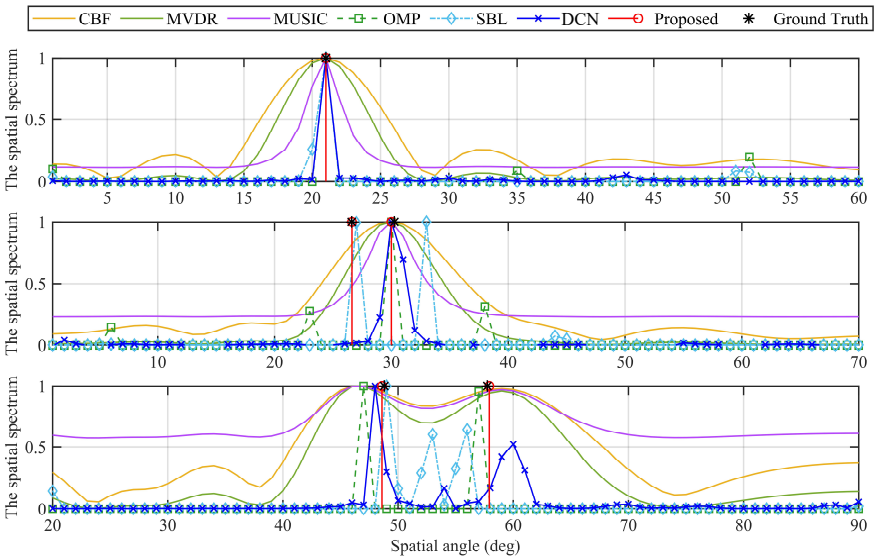

Figure 9. Spatial spectra in one-source and two-source scenarios with DOA estimation methods. First row: one-source scenario. Second row: two-source with a small angle interval of 3.6°. Third row: two-source with a large angle interval of 9°.

This experiment evaluated the performance of the various algorithms in three DOA estimation scenarios. In the first case, there was a single source, as shown in the first row of Figure 9, with a true DOA of 21°. In this case, all of the DOA estimation methods successfully estimated the true value of the target angle, though the conventional methods exhibited varying levels of beamwidth. However, these errors were avoided by the high-resolution algorithms based on compressed sensing and deep learning, all of which achieved a high-accuracy DOA estimation of the target source. The second case involved two targets with a very small angle interval, as shown in the second row of Figure 9. The true DOA angles were 26.6° and 30.2°, respectively. In this case, all methods except SBL and our proposed method could only estimate one target, missing the other one. The third case involved two targets with a large angle interval, as shown in the third row of Figure 9. The true DOA angles were 48.8° and 57.8°. In this scenario, the conventional methods lacked sufficient resolution, leading to errors in both angles.

Figure 9 provides a clear visual representation of the spatial spectrum for all three cases, highlighting the differences in DOA estimation accuracy between the proposed SDSR network and the conventional methods. In the one-source scenario, the methods exhibited varying degrees of resolution, with the SDSR network showing superior performance in terms of accuracy and noise suppression. In the two-source scenario with small angle intervals, the SDSR network effectively resolved both sources, even with closely spaced sources, while the other methods struggled with accurate estimation, due to limited resolution or grid-related issues. The SDSR method also performed well in scenarios with larger angle intervals, while the conventional methods failed to accurately estimate the DOAs, due to limited resolution. This comparison emphasizes the effectiveness of the SDSR method in dealing with complex DOA estimation challenges, particularly in challenging cases where conventional methods struggle.

5.4. Experimental Results

To evaluate the DOA estimation performance of our proposed method in a real-world deep-sea environment, experimental observed signals were chosen for testing. The depth of the experimental area was approximately 4200 m, and a 16-element VLA spaced at 7.5 m was deployed near the seafloor. The first hydrophone in the VLA was positioned at a depth of 4022.5 m. The sampling rate of the observed signals was 16 kHz, with the sources situated at a depth of 200 m. The target approached the VLA from a distance, moved above the VLA, and then moved away from it. The initial and final horizontal distances between the target and the VLA were 7.02 km and 15.85 km, respectively.

To obtain the true values of the arrival paths in this real environment, the acoustic toolbox Bellhop [32] was used in Matlab to simulate the underwater acoustic multipath environment. The water sound speed was calculated using the Munk profile, and the environment parameter settings are listed in Table 2. The simulation setting is depicted in Figure 10. The eigenrays included the direct path, and the once- and twice-reflected paths by the surface and bottom of the water.

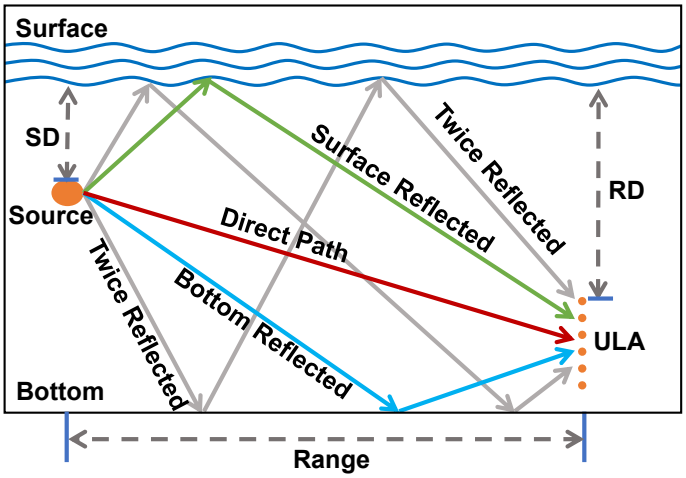

Figure 10. The setting of the simulation scenario (SD = 200 m, RD = 4022.5 m, and Water Depth = 4200 m).

Table 2. Environment parameter settings.

Parameter	Value
Frequency (Hz)	100
Water Depth (m)	4200
Source Depth (m)	200
Receive Depth (m)	4022.5
Array element number	16
Array depth interval (m)	7.5
Bottom Sound Speed (m/s)	1600
Bottom density (g/cm^3)	1.6
Bottom attenuation coefficient (dB/λ)	0.15

In Bellhop, we could obtain information about the number of paths, and the amplitude and time delay of each path by setting the marine environmental parameters.

The eigenrays of the two sources that connected the source and receiver are shown in the left of Figure 11. It can be seen that the signals emitted by the source traveled through multipath propagation before impinging on the array. The underwater channel impulse responses of the source generated by Bellhop are depicted in the right of Figure 11. It is shown that the direct (D) paths were the strongest, the once-reflected paths were the second strongest, and the multiple-reflected paths had severe energy attenuation. In our simulations, the paths from -90° to 0° were ignored and only the D and the SR paths were considered, which dominated the strongest paths.

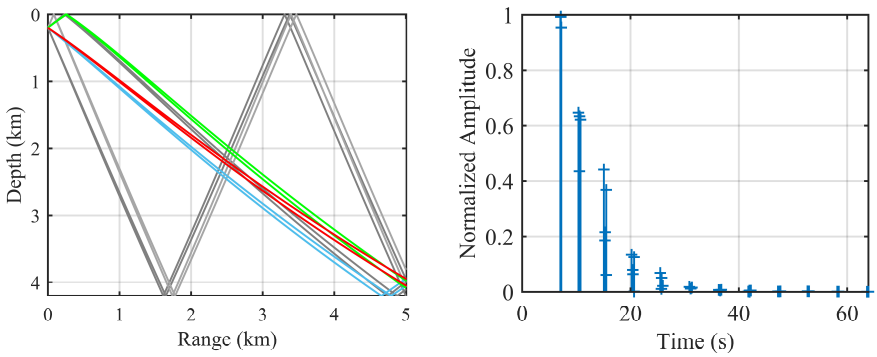

Figure 11. The eigenrays and underwater channel impulse responses of the simulation scenario (SD = 200 m, RD = 4022.5 m, and Range = 10 km).

Figure 12 presents the DOA estimation results for real data processed using conventional methods, CS-based methods, and DL-based methods, respectively. These results allow a direct comparison of the performance of the proposed SDSR method in a challenging, realistic underwater acoustic setting, demonstrating its ability to effectively estimate DOAs in dynamic and complex environments.

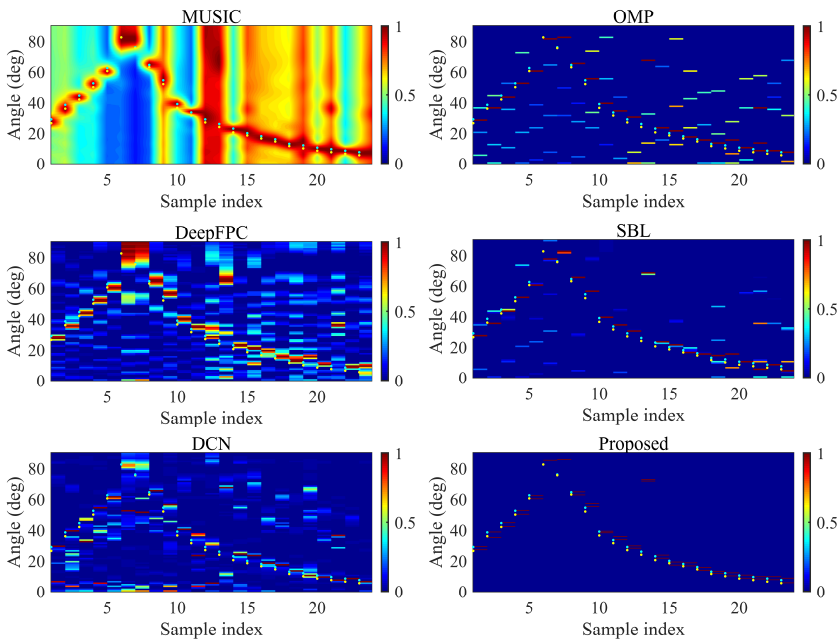

Figure 12. DOA estimation of MUSIC, OMP, DeepFPC, SBL, DCN, and proposed SDSR network on the experimentally observed data. The blue dots in the figures represent the true value of the angle of the SR path, and the yellow dots represent the true value of the angle of the D path.

From Figure 12, it is illustrated that all methods were able to estimate the spatial trajectory of the target during its motion. However, it is worth noting that, when the target angle lay in the direction of the VLA endfire, all methods struggled to discriminate between the direct (D) path and the surface reflection (SR) path. In this case, besides the sound from the D path, there was another energetic sound arriving at the VLA from the SR path. Despite this, the DOA estimation results from the conventional methods and compressed sensing-based methods failed to capture the multipath information.

In contrast, the deep-learning-based methods successfully show the separation of the multipath signals. Among these methods, our proposed SDSR method exhibited the highest accuracy in estimating the DOAs of both the direct and surface reflection paths. This demonstrates the effectiveness of the SDSR model in handling complex scenarios

where multipath interference occurs, particularly in underwater acoustic environments where the presence of such multipaths can significantly complicate the estimation task.

6. Conclusions

In this paper, we proposed an efficient sparsity-based DOA estimation algorithm. The core idea involves transforming the DOA estimation problem into a sparse linear inverse problem using a spatially overcomplete formulation. We then described the structure and training procedure of the SDSR network, which significantly improved the spatial resolution in DOA estimation and demonstrated robust performance under low SNR conditions.

Compared to conventional iterative-based sparse recovery algorithms, the SDSR-Net requires only feedforward calculations, enabling real-time direction finding, which is crucial for many practical applications. Additionally, the deep residual shrinkage network integrates the shrinkage modules as trainable components, allowing it to automatically adjust the thresholds, without needing expert knowledge of signal processing. This feature not only simplifies the system but also enhances the model's adaptability to various scenarios.

The learning and generalization abilities of the shrinkage module contribute to its competitive, and often superior, performance in DOA estimation, especially under challenging conditions such as low SNR of -10 dB or angle separations as small as 0.2°. The results from both simulations and real-world experiments clearly demonstrated the superiority of the proposed method over existing techniques.

Author Contributions: Methodology, L.Z.; Formal analysis, L.Z.; Data curation, L.Z., S.Z. and Y.Q.; Writing—original draft, L.Z.; Writing—review & editing, S.Z., L.W., Y.B. and Y.Q.; Supervision, S.Z., L.W. and Y.B.; Funding acquisition, S.Z. All authors have read and agreed to the published version of this manuscript.

Funding: This research was supported by the National Key Research and Development Program of China (No. 2021YFC3101403), the National Natural Science Foundation of China (Grant No. 12174419) and the China Scholarship Council.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare that they have no known conflicts of interest that could have appeared to influence the work reported in this paper.

References

1. Shi, J.; Hu, G.; Zhang, X.; Sun, F.; Zheng, W.; Xiao, Y. Generalized co-prime MIMO radar for DOA estimation with enhanced degrees of freedom. *IEEE Sens. J.* **2017**, *18*, 1203–1212. [CrossRef]
2. Nielsen, U.; Dall, J. Direction-of-arrival estimation for radar ice sounding surface clutter suppression. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5170–5179. [CrossRef]
3. Lee, H.; Ahn, J.; Kim, Y.; Chung, J. Direction-of-arrival estimation of far-field sources under near-field interferences in passive sonar array. *IEEE Access* **2021**, *9*, 28413–28420. [CrossRef]
4. Huang, H.; Yang, J.; Huang, H.; Song, Y.; Gui, G. Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system. *IEEE Trans. Veh. Technol.* **2018**, *67*, 8549–8560. [CrossRef]
5. Yang, T. Characteristics of underwater acoustic communication channels in shallow water. In Proceedings of the OCEANS 2011 IEEE-Spain, Santander, Spain, 6–9 June 2011; pp. 1–8.
6. Malioutov, D.; Cetin, M.; Willsky, A.S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.* **2005**, *53*, 3010–3022. [CrossRef]
7. Yin, J.; Chen, T. Direction-of-arrival estimation using a sparse representation of array covariance vectors. *IEEE Trans. Signal Process.* **2011**, *59*, 4489–4493. [CrossRef]

8. 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]
9. Zhang, H.; Pour, S.Z.; Yan, H.; Liu, P.; Arigong, B. A Low-Cost Monopulse Receiver with Enhanced Estimation Accuracy Via Deep Neural Network. *arXiv* **2024**, arXiv:2411.17734. [CrossRef]
10. Qin, Y. Deep Networks for Direction of Arrival Estimation with Sparse Prior in Low SNR. *IEEE Access* **2023**, *11*, 44637–44648. [CrossRef]
11. Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE* **1969**, *57*, 1408–1418. [CrossRef]
12. Schmidt, R. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **1986**, *34*, 276–280. [CrossRef]
13. Belouchrani, A.; Amin, M.G. Time-frequency MUSIC. *IEEE Signal Process. Lett.* **1999**, *6*, 109–110. [CrossRef]
14. Roy, R.; Kailath, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 984–995. [CrossRef]
15. Lin, J.; Ma, X.; Yan, S.; Hao, C. Time-frequency multi-invariance ESPRIT for DOA estimation. *IEEE Antennas Wirel. Propag. Lett.* **2015**, *15*, 770–773. [CrossRef]
16. Shmuel, D.H.; Merkofer, J.P.; Revach, G.; Van Sloun, R.J.G.; Shlezinger, N. SubspaceNet: Deep Learning-Aided Subspace Methods for DoA Estimation. *IEEE Trans. Veh. Technol.* **2024**, 1–15. [CrossRef]
17. Merkofer, J.P.; Revach, G.; Shlezinger, N.; Routtenberg, T.; van Sloun, R.J.G. DA-MUSIC: Data-Driven DoA Estimation via Deep Augmented MUSIC Algorithm. *IEEE Trans. Veh. Technol.* **2024**, *73*, 2771–2785. [CrossRef]
18. Liu, Z.M.; Huang, Z.T.; Zhou, Y.Y. An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 1–11. [CrossRef]
19. Gerstoft, P.; Mecklenbräuker, C.F.; Xenaki, A.; Nannuru, S. Multisnapshot sparse Bayesian learning for DOA. *IEEE Signal Process. Lett.* **2016**, *23*, 1469–1473. [CrossRef]
20. Liang, G.; Li, C.; Qiu, L.; Shen, T.; Hao, Y. State-updating-based DOA estimation using sparse Bayesian learning. *Appl. Acoust.* **2022**, *192*, 108719. [CrossRef]
21. Park, Y.; Meyer, F.; Gerstoft, P. Sequential sparse Bayesian learning for time-varying direction of arrival. *J. Acoust. Soc. Am.* **2021**, *149*, 2089–2099. [CrossRef]
22. Xiang, H.; Chen, B.; Yang, T.; Liu, D. Improved de-multipath neural network models with self-paced feature-to-feature learning for DOA estimation in multipath environment. *IEEE Trans. Veh. Technol.* **2020**, *69*, 5068–5078. [CrossRef]
23. 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]
24. Barthelme, A.; Utschick, W. DoA Estimation Using Neural Network-Based Covariance Matrix Reconstruction. *IEEE Signal Process. Lett.* **2021**, *28*, 783–787. [CrossRef]
25. Alam, A.M.; Ayna, C.O.; Biswas, S.; Rogers, J.T.; Ball, J.E.; Gurbuz, A.C. Deep Learning-Based Direction-of-Arrival Estimation with Covariance Reconstruction. In Proceedings of the 2024 IEEE Radar Conference (RadarConf24), Denver, CO, USA, 6–10 May 2024; pp. 1–6. [CrossRef]
26. Ma, J.; Wang, M.; Chen, Y.; Wang, H. Deep Convolutional Network-Assisted Multiple Direction-of-Arrival Estimation. *IEEE Signal Process. Lett.* **2024**, *31*, 576–580. [CrossRef]
27. Yu, J.; Wang, Y. Deep Learning-Based Multipath DoAs Estimation Method for mmWave Massive MIMO Systems in Low SNR. *IEEE Trans. Veh. Technol.* **2023**, *72*, 7480–7490. [CrossRef]
28. Jiarun, Y.; Yafeng, W. MDTCNet: Multi-Task Classifications Network and TCNN for Direction of Arrival Estimation. *China Commun.* **2024**, *21*, 1–19. [CrossRef]
29. Donoho, D.L.; Tsaig, Y.; Drori, I.; Starck, J.L. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **2012**, *58*, 1094–1121. [CrossRef]
30. Wu, L.; Liu, Z.M.; Huang, Z.T. Deep convolution network for direction of arrival estimation with sparse prior. *IEEE Signal Process. Lett.* **2019**, *26*, 1688–1692. [CrossRef]
31. Xiao, P.; Liao, B.; Deligiannis, N. Deepfpc: A deep unfolded network for sparse signal recovery from 1-bit measurements with application to doa estimation. *Signal Process.* **2020**, *176*, 107699. [CrossRef]
32. Porter, M.B. *The Bellhop Manual and User's Guide: Preliminary Draft*; Tech. Rep.; Heat, Light, and Sound Research, Inc.: La Jolla, CA, USA, 2011; Volume 260.

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.