Enhancing Off-Grid One-Bit DOA Estimation with Learning-Based Sparse Bayesian Approach for Non-Uniform Sparse Array

Yunqiao Hu[†], Shunqiao Sun[†], and Yimin D. Zhang[§]

[†]Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL 35487
[§]Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122

***Abstract*—This paper tackles the challenge of one-bit off-grid direction of arrival (DOA) estimation in a single snapshot scenario based on a learning-based Bayesian approach. Firstly, we formulate the off-grid DOA estimation model, utilizing the first-order off-grid approximation, incorporating one-bit data quantization. Subsequently, we address this problem using the Sparse Bayesian based framework and solve iteratively. However, traditional Sparse Bayesian methods often face challenges such as high computational complexity and the need for extensive hyperparameter tuning. To balance estimation accuracy and computational efficiency, we propose a novel Learning-based Sparse Bayesian framework, which leverages an unrolled neural network architecture. This framework autonomously learns hyperparameters through supervised learning, offering more accurate off-grid DOA estimates and improved computational efficiency compared to some state-of-the-art methods. Furthermore, the proposed approach is applicable to both uniform linear arrays and non-uniform sparse arrays. Simulation results validate the effectiveness of the proposed framework.**

***Index Terms*—Sparse arrays, off-grid, DOA estimation, Bayesian approach**

I. Introduction

The problem of direction of arrival (DOA) estimation is fundamentally important in sensor array signal processing with widely application in radar, sonar, navigation, and wireless communications [1–5]. Most super-resolution DOA estimation algorithms such as MUltiple SIgnal Classification (MUSIC) [6] and Estimation of Signal Parameters via Rotational Invariant Techniques (ESPRIT) [7] have been primarily developed and applied to uniform linear arrays (ULAs), where sensor elements are arranged in a straight line with equal spacing, typically half the signal wavelength. However, in practical applications, achieving higher resolution with ULA requires a larger aperture and thus an increased number of array elements, significantly raising hardware costs [4]. Furthermore, ULAs are susceptible to mutual coupling effects, which can degrade DOA estimation performance [8]. To address this problem, sparse linear arrays (SLAs) have been used over the past few decades to achieve desired apertures with fewer active elements. Some SLA configurations, such as the minimum redundancy array (MRA) [9], the nested array [10], the co-prime array [11], and the generalized coprime array configurations [12], have been well studied and analyzed in the past decades. While subspace-based methods, such as MUSIC and weighted subspace fitting [13, 14], and covariance matrix-based compressive sensing methods [12, 15] can be applied to SLAs, they require a high number of snapshots to achieve accurate covariance matrix estimation, making them impractical in snapshot-limited scenarios commonly encountered in automotive applications [5].

In antenna array systems, high-precision, high-sampling-rate analog-to-digital converters (ADCs) are costly and power-intensive [16]. Low-bit or even one-bit quantization offers a cost-effective solution to simplify sampling hardware and increase sampling rates [17]. Consequently, low-resolution ADC signal processing has drawn significant research interest, with research on one-bit DOA estimation becoming one of the key focuses. Subspace-based methods exploiting one-bit quantized data, along with performance analyses, are presented for both ULAs and SLAs [1], [18, 19].

In the past decade, methods leveraging sparse reconstruction [20] and compressed sensing (CS) [21] have emerged to address one-bit DOA estimation, leading to various estimators such as the *Binary Iterative Hard Thresholding* (BIHT) algorithm [22], the *Bayesian Compressed Sensing* (BCS) algorithm [23], *Sparse Learning via Iterative Minimization* (SLIM), and *Iterative Adaptive Approaches* (IAA) [16]. These techniques operate effectively with SLAs in single-snapshot scenarios. All the above methods are considered on-grid approaches, as they determine DOAs by peak searching over a fixed discrete angular spectrum.

For off-grid source signals, grid refinement is often necessary to maintain estimation accuracy, necessitating denser grids and increased computational complexity. In CS-based methods, denser grids also lead to higher correlations among dictionary atoms, thereby degrading algorithm performance. To this end, gridless and off-grids estimation methods have been proposed to address an one-bit off-grid DOA estimation problem. Gridless methods, such as *Atomic Norm Minimization* (ANM) [24], estimate DOAs without grid division but

The work of Y. Hu and S. Sun was supported in part by U.S. National Science Foundation (NSF) under Grants CCF-2153386 and ECCS-2033433. The work of Y. D. Zhang was supported in part by NSF under Grant ECCS-2236023.

require computationally intensive semidefinite programming (SDP). On the other hand, off-grid methods, such as *off-grid iterative reweighted* (OGIR) algorithm [25], alternatively refine on-grid spectrum estimates and grid gap estimates for improved estimation accuracy. Off-grid methods generally eliminate the need for dense grid division making them more flexible and applicable for real DOA estimation tasks compared to gridless methods. However, these methods often require hundreds of iterations to achieve satisfacory results, with each iteration involving computationally matrix inversion.

Recently, model-based deep learning has gained traction in the signal processing community [26], including DOA estimation research [27–29]. Such hybrid approaches combine the interpretability of classical array signal models with the representation power of deep neural networks, thus better addressing limitations of traditional methods in handling high-dimensional, noisy, or complex data.

In this paper, we propose a novel learning-based sparse Bayesian approach to tackle the problem of one-bit off-grid DOA estimation. We choose the sparse Bayesian framework for its robustness to noise and superior reconstruction accuracy, especially in single-snapshot scenarios. We first formulate the one-bit off-grid model using a first-order grid approximation. By applying the maximum a posteriori (MAP) criterion and incorporating sparse signal priors, we construct an iterative minimization problem, which is mapped to a neural network architecture, employing convolutional neural networks (CNNs) to replace matrix inversion operations and deep neural networks (DNNs) for off-grid updates. Simulation results demonstrate that the proposed method outperforms state-of-the-art algorithms in terms of accuracy, providing more precise off-grid DOA estimates across various signal-to-noise ratio (SNR) scenarios, using only a single snapshot of data.

II. Problem Formulation

A. Signal Model

Consider a scenario involving K narrowband, far-field source signals, denoted as $s_k(t)$ for $k = 1, \ldots, K$, arriving at an N-element SLA from directions $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_K]^{\mathrm{T}}$, where $(\cdot)^{\mathrm{T}}$ represents transpose. The array signal model with one-bit quantizated data is expressed as:

$$\begin{aligned}\mathbf{y}(t) &= \operatorname{csgn}\left(\sum_{k=1}^{K}\mathbf{a}(\theta_k)s_k(t) + \mathbf{n}(t)\right) \\ &= \operatorname{csgn}\left(\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t)\right), \quad t = 1, \cdots, T,\end{aligned} \tag{1}$$

where $\mathbf{y}(t)$ is the received signal vector, $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_K)]$ is the array manifold matrix, $\mathbf{s}(t)$ is the source signal vector, and $\mathbf{n}(t)$ is the complex Gaussian noise vector. The complex sign function is defined as $\operatorname{csgn}(\cdot) = \operatorname{sign}(\Re(\cdot)) + j\operatorname{sign}(\Im(\cdot))$, where $\operatorname{sign}(\cdot)$ returns value in $\{1, -1\}$, $\Re(\cdot)$ and $\Im(\cdot)$ respectively return the real and imaginary parts of a complex number. Each column in array manifold matrix $\mathbf{A}(\boldsymbol{\theta})$ corresponds to a steering vector, given for the k-th signal as:

$$\mathbf{a}(\theta_k) = \left[1, e^{j\frac{2\pi d_2}{\lambda}\sin\theta_k}, \ldots, e^{j\frac{2\pi d_N}{\lambda}\sin\theta_k}\right]^{\mathrm{T}}, \tag{2}$$

where d_n specifies the spacing between the n-th element and the first element. This paper focuses on estimating the signal DOAs, $\boldsymbol{\theta}$, using a single-snapshot data vector $\mathbf{y}$. With T set to 1, the model simplifies to:

$$\mathbf{y} = \operatorname{csgn}\left(\mathbf{A}(\boldsymbol{\theta})\mathbf{s} + \mathbf{n}\right). \tag{3}$$

B. One-Bit Off-Grid Single-Measurement Vector Model

To estimate $\boldsymbol{\theta}$ from (3), we reformulate it as a single-measurement vector model:

$$\mathbf{y} = \operatorname{csgn}(\mathcal{A}(\tilde{\boldsymbol{\theta}})\mathbf{x} + \mathbf{n}). \tag{4}$$

where $\mathcal{A} = [\mathbf{a}(\tilde{\theta}_1), \mathbf{a}(\tilde{\theta}_2), \ldots, \mathbf{a}(\tilde{\theta}_M)] \in \mathbb{C}^{N\times M}$ is the dictionary matrix, $\Theta = \{\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_M\}$ are the discretized angle grids and $\mathbf{x} = [x_1, x_2, \ldots, x_M]^T$ are sparse coefficients to be estimated.

Under this model, M grid points serve as the basis for sparse signal representation, and the signals are assumed to align with K grid points, i.e., $\theta_k \in \Theta$, $k = 1, \cdots, K$. This approach is known as the on-grid model. However, in practice, true DOAs often don't align perfectly with the predefined grid, i.e., $\theta_k \notin \Theta$. Assuming the grid is sufficiently dense, the true DOA θ_k lies near the fixed grid point $\tilde{\theta}_{n_k}$, $n_k \in \{1, 2, \ldots, M\}$. Using the first-order Taylor expansion, the true DOA can then be approximated as:

$$\theta_k = \tilde{\theta}_{n_k} + (\theta_k - \tilde{\theta}_{n_k}), \tag{5}$$

where $(\theta_k - \tilde{\theta}_{n_k})$ represents the off-grid gap. The steering vector $\mathbf{a}(\theta_k)$ can be approximated as:

$$\mathbf{a}(\theta_k) = \mathbf{a}(\tilde{\theta}_{n_k}) + \mathbf{b}(\tilde{\theta}_{n_k})(\theta_k - \tilde{\theta}_{n_k}), \tag{6}$$

where $\mathbf{b}(\tilde{\theta}_{n_k}) = \frac{\partial \mathbf{a}(\theta)}{\partial \theta}|_{\tilde{\theta}_{n_k}}$ is the first-order derivative. By incorporating the approximation error into the measurement noise, the measurement model can be reformulated as:

$$\mathbf{y} = \operatorname{csgn}\left(\mathbf{C}(\boldsymbol{\beta})\mathbf{x} + \mathbf{n}\right), \tag{7}$$

where $\mathbf{C}(\boldsymbol{\beta}) = \mathcal{A} + \mathcal{B}\ \operatorname{diag}(\boldsymbol{\beta})$ is the approximation dictionary with $\operatorname{diag}(\cdot)$ denoting diagonal matrix, $\mathcal{B} = [\mathbf{b}(\tilde{\theta}_1), \mathbf{b}(\tilde{\theta}_2), \ldots, \mathbf{b}(\tilde{\theta}_M)]$, and $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_M]^{\mathrm{T}}$ are off-grid gaps defined as:

$$\beta_n = \begin{cases}\theta_k - \theta_{n_k}, & \text{if } n = n_k, k \in \{1, 2, ..., K\}, \\ 0, & \text{otherwise.}\end{cases} \tag{8}$$

III. Algorithm Framework

A. Sparse Bayesian Formulation

To establish a sparse Bayesian framework for the estimation of one-bit off-grid DOA estimation, we first introduce a probabilistic model to quantify the probability of $\mathbf{x}$ given the input $\mathbf{y}$. According to (7), the posterior probability of $\mathbf{x}$, under the MAP criterion, is determined by the likelihood function

$p(\mathbf{y}|\mathbf{x};\boldsymbol{\beta})$ and the prior probability density function (PDF) $p(\mathbf{x})$. The likelihood function is given by [30]:

$$p(\mathbf{y}|\mathbf{x};\boldsymbol{\beta}) = \prod_{m=1}^{M} \Phi\left(\frac{\Re(y_m)\,\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\,\mathbf{x})}{\sigma/\sqrt{2}}\right) \cdot \Phi\left(\frac{\Im(y_m)\,\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\,\mathbf{x})}{\sigma/\sqrt{2}}\right), \tag{9}$$

where $\Phi(\cdot)$ denotes the cumulative density function of the standard normal distribution, y_m is the m-th element in $\mathbf{y}$, $\boldsymbol{c}_m^{\mathrm{T}}$ is the m-th row vector in $\mathbf{C}$. For the convenience of calculation, let $\hat{\mathbf{x}} = \frac{\sqrt{2}}{\sigma}\mathbf{x}$, and (9) can be reformulated as:

$$p(\mathbf{y}|\hat{\mathbf{x}};\boldsymbol{\beta}) = \prod_{m=1}^{M} \Phi(\Re(y_m)\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}}))\Phi(\Im(y_m)\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}})). \tag{10}$$

A suitable prior PDF for $\hat{\mathbf{x}}$, such as a Laplacian prior [31] or exponential distribution [16], should be selected to encourage sparsity. In this paper, we select the prior PDF as:

$$p(\hat{\mathbf{x}}) = \prod_{i=1}^{M} \exp\left(-\frac{\lambda\,|\hat{x}_i|^{\alpha}}{\alpha}\right), \quad 0 < \alpha \leq 1, \tag{11}$$

where λ is a parameter. As α approaches 0, $p(\hat{\mathbf{x}})$ reaches its maximum at $\hat{\mathbf{x}} = 0$, enforcing sparsity on the signal. The off-grid gaps $\boldsymbol{\beta}$ follow a uniform distribution, $p(\boldsymbol{\beta}) \sim U\left(-\frac{r}{2}, \frac{r}{2}\right)$ [32], where r denotes the grid interval size. With Bayes rule, the MAP estimator is given as:

$$\left\{\hat{\mathbf{x}}^*, \boldsymbol{\beta}^*\right\} = \arg\min_{x,\beta} -\ln p(\mathbf{y}|\hat{\mathbf{x}};\boldsymbol{\beta}) - \ln p(\hat{\mathbf{x}}) - \ln p(\boldsymbol{\beta}). \tag{12}$$

B. Updating Formula

By substituting (10) and (11) into equation (12), we obtain the following cost function to be minimized:

$$\mathcal{L} = \sum_{m=1}^{M} \left\{-\ln\Phi\left(\Re(y_m)\,\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}})\right) - \ln\Phi\left(\Im(y_m)\,\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}})\right) + \sum_{i=1}^{N} \frac{\lambda\,|x_i|^{\alpha}}{\alpha} + \text{const}\right\}. \tag{13}$$

Since the object function in (13) is non-convex, we apply convex relaxation to simplify it. Specifically, with the majorization-minimization (MM) principle, we can find the upper bound of the first two terms in (13) as [30]:

$$\begin{aligned}
&\sum_{m=1}^{M}\left\{-\ln\Phi(\Re(y_m)\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}})) - \ln\Phi(\Im(y_m)\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}}))\right\} \\
&\leq \sum_{m=1}^{M} \frac{1}{2}(\Re(y_m)\,\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}}))^2 + \frac{1}{2}(\Im(y_m)\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}}))^2 \\
&- \Re(v_m^t)\Re(y_m)\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}}) - \Im(v_m^t)\Im(y_m)\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta})\hat{\mathbf{x}}) + c'.
\end{aligned} \tag{14}$$

Fig. 1: Overall architecture of the Unrolled Network.

where $v_m^t = \left[\Re(y_m)\Re(\tilde{v}_m^t)\right] + j\left[\Im(y_m)\Im(\tilde{v}_m^t)\right]$, $\tilde{v}_m^t = d_m^t - \mathrm{I}'(d_m^t)$, $d_m^t = \Re(y_m)\Re(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta}^t)\hat{\mathbf{x}}^t) + j\Im(y_m)\Im(\boldsymbol{c}_m^{\mathrm{T}}(\boldsymbol{\beta}^t)\hat{\mathbf{x}}^t)$, and c' is a constant. Here, superscript $(\cdot)^t$ indicates variables from the tth iteration. In addition, function $\mathrm{I}'(x)$ is defined as

$$\mathrm{I}'(x) = -\frac{\exp\left(-\Re(x)^2/2\right)}{\sqrt{2\pi}\Phi(\Re(x))} - j\frac{\exp\left(-\Im(x)^2/2\right)}{\sqrt{2\pi}\Phi(\Im(x))}. \tag{15}$$

The third term in (13) can be smoothly approximated as:

$$\sum_{i=1}^{N} \frac{\lambda\,|x_i|^{\alpha}}{\alpha} \approx \frac{\lambda}{\alpha}\sum_{i=1}^{N}\left(|x_i|^2 + \eta\right)^{\frac{\alpha}{2}}. \tag{16}$$

where $\eta > 0$ is a small constant, typically set to 10^{-6}. By substituting (14) and (16) into (13), we obtain a new minimization problem defined as:

$$\left\{\hat{\mathbf{x}}^*, \boldsymbol{\beta}^*\right\} = \arg\min_{\hat{x},\beta} \frac{1}{2}\left\|\mathbf{C}(\boldsymbol{\beta})\hat{\mathbf{x}} - \mathbf{v}^t\right\|_2^2 + \frac{\lambda}{\alpha}\sum_{i=1}^{N}\left(|x_i|^2 + \eta\right)^{\frac{\alpha}{2}} + c', \tag{17}$$

where $\mathbf{v}^t = [v_1^t, \cdots, v_M^t]^{\mathrm{T}}$. Then, we use iterative update strategy to estimate $\mathbf{x}$ and $\boldsymbol{\beta}$ as:

$$\hat{\mathbf{x}}^{t+1} = \left[\mathbf{C}^{\mathrm{H}}(\boldsymbol{\beta}^t)\mathbf{C}(\boldsymbol{\beta}^t) + \lambda\boldsymbol{\Lambda}(\hat{\mathbf{x}}^t)\right]^{-1}\mathbf{C}^{\mathrm{H}}(\boldsymbol{\beta}^t)\mathbf{v}^t, \tag{18}$$

$$\boldsymbol{\beta}^{t+1} = \Re\left((\mathcal{B}^{\mathrm{H}}\mathcal{B})^*\hat{\mathbf{x}}^{t+1}(\hat{\mathbf{x}}^{t+1})^{\mathrm{H}}\right)^{-1} \cdot \Re\left(\mathrm{diag}\left((\hat{\mathbf{x}}^t)^*\right)\mathcal{B}^{\mathrm{H}}\left(\mathbf{v}^t - \mathcal{A}\hat{\mathbf{x}}^t\right)\right), \tag{19}$$

where

$$\boldsymbol{\Lambda}\left(\hat{\mathbf{x}}^t\right) = \mathrm{diag}\left(\left(|\hat{x}_1|^2 + \eta\right)^{\frac{\alpha}{2}-1}, \cdots, \left(|\hat{x}_N|^2 + \eta\right)^{\frac{\alpha}{2}-1}\right).$$

IV. Neural Network Derived from the Algorithm

In this section, we follow the algorithm unrolling paradigm to design the network architecture by mapping the iteration steps (18)–(19) to customized layers.

A. Network Architecture

The network consists of the following three main blocks: Initialization Block, Unrolled Block 1, and Unrolled Block 2, as shown in Figure 1.

1) Initialization Block: The initialization block simply performs the following operation on the input vector:

$$\hat{\mathbf{x}}^0 = \mathbf{C}^{\mathrm{H}}(\boldsymbol{\beta}^0)\mathbf{y}, \tag{20}$$

where $\boldsymbol{\beta}^0 = \mathbf{0}$, which means that no initial off-grid gaps are assumed.

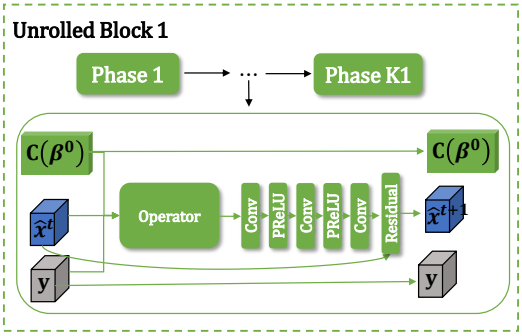


Fig. 2: Detail architecture of Unrolled Block 1.

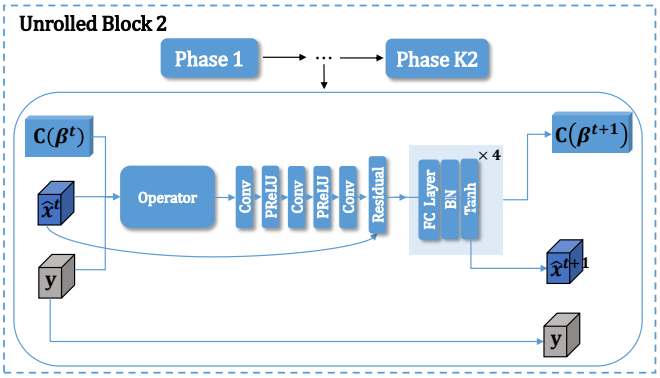


Fig. 3: Detail architecture of Unrolled Block 2.

2) Unrolled Block 1: Unrolled Block 1 consists of K_1 unrolled phases. The details of one unrolled phase are illustrated in Figure 2. Unrolled Block 1 is designed to refine the spectrum in fixed grids. Therefore, the grid gaps are set to $\boldsymbol{\beta}^0 = \mathbf{0}$, and grid update is not considered in this block. The operation is defined as:

$$\tilde{\mathbf{v}}^t = \mathbf{C}^{\mathrm{H}}(\boldsymbol{\beta}^0)\mathbf{v}^t, \tag{21}$$

where $\boldsymbol{\beta}^0 = \mathbf{0}$, $\mathbf{v}^t = \left[\Re\left(\mathbf{y}\right) \odot \Re(\tilde{\mathbf{v}}^t)\right] + j\left[\Im(\mathbf{y}) \odot \Im(\tilde{\mathbf{v}}^t)\right]$, $\tilde{\mathbf{v}}^t = \mathbf{D}^t - \mathrm{I}'(\mathbf{D}^t)$, and $\mathbf{D}^t = \Re\left(\mathbf{y}\right) \odot \Re(\mathbf{C}(\boldsymbol{\beta}^0)\hat{\mathbf{x}}^t\Big) + j\Im(\mathbf{y}) \odot \Im\left(\mathbf{C}(\boldsymbol{\beta}^0)\hat{\mathbf{x}}^t\right)$, $\odot$ is Hadamard product. According to (18), matrix inversion is used to compute the estimate. However, the computational complexity of this operation increases with matrix size, making network training challenging. To address this, we replace matrix inversion in (18) with convolutional layers. In addition, each convolutional layer is followed by a Parametric Rectified Linear Unit (PReLU) activation layer [33] to introduce nonlinearity into the network. A residual connection is incorporated into each unrolled phase to give the output $\hat{\mathbf{x}}^{t+1}$.

3) Unrolled Block 2: Unrolled Block 2 comprises K_2 unrolled phases and is designed to update both the angle spectrum coefficients and the off-grid gaps, as shown in Figure 3. Each phase follows the same operations as in (21), with convolutional layers identical to those in Unrolled Block 1. For off-grid gap updates, we replace the formula in (19) with four Fully Connected (FC) layers, each followed by batch normalization and Tanh activation. The FC layers take the absolute values of the estimated signal spectrum, $|\hat{\mathbf{x}}^{t+1}|$, as input and output the updated off-grid gaps corresponding to the on-grid spectra. The output is normalized to a single grid interval $[-\frac{r}{2}, \frac{r}{2}]$.

B. Generating Datasets

We generate training datasets using two SLA configurations with half-wavelength element spacing: an 18-element array at $\frac{\lambda}{2}[0, 1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19]$ and a 10-element array at $\frac{\lambda}{2}[0, 3, 4, 5, 6, 7, 11, 16, 18, 19]$. The number of signal sources is set to 2, and the array field of view (FOV) is defined as $[-60^\circ, 60^\circ]$, discretized with a fixed grid size of 2°. The off-grid gaps corresponding to target sources are randomly generated following a uniform distribution $U\left(-1^\circ, 1^\circ\right)$. Reflection coefficients for DOA sources are generated as random complex numbers, with their real and imaginary parts uniformly distributed in $U\left(0.5, 1\right)$. Denoting the ground truth of the n-th DOA as G_n, the signals are labeled according to

$$\mathbf{s}_n^* = \begin{cases} |s_k|, & \text{if } \theta_k = G_n, \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

For the labeling of off-grid gaps:

$$\beta_n^* = \begin{cases} |\beta_k|, & \text{if } \theta_k = G_n, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

We randomly generate $100,000$ samples across input SNR levels ranging between 0 dB and 30 dB in 5 dB increments. 90% of the dataset is used for training and the remaining 10% is used for validation.

C. Training Approach

Our training process is divided into two stages. In the first stage, we train only Unrolled Block 1 while keeping the parameters of Unrolled Block 2 frozen. This training process is performed with a batch size of 64 for 100 epochs. The objective of this stage is to generate angle spectra with low sidelobe levels, facilitating the subsequent training of Unrolled Block 2. The loss function used in this stage is the Binary Cross-Entropy (BCE) loss, defined as:

$$\mathcal{L}_1(\hat{\mathbf{x}}, \mathbf{s}^*) = -\frac{1}{N}\sum_{i=1}^{N}\left[s_i^* \cdot \log\hat{x}_i + (1 - s_i^*) \cdot \log(1 - \hat{x}_i)\right]. \tag{24}$$

In the second stage, the parameters of Unrolled Block 1 are frozen, and only Unrolled Block 2 is trained. In this stage, we apply a combination of the mean squared error (MSE) loss and the BCE loss. The total loss function is defined as:

$$\mathcal{L}_2(\hat{\mathbf{x}}, \mathbf{s}^*; \hat{\boldsymbol{\beta}}, \boldsymbol{\beta}^*) = -\frac{1}{N}\sum_{i=1}^{N}\left[s_i^* \cdot \log\hat{x}_i + (1 - s_i^*) \cdot \log(1 - \hat{x}_i)\right] + \frac{1}{N}\sum_{i=1}^{N}[\hat{\beta}_i - \beta_i^*]^2. \tag{25}$$

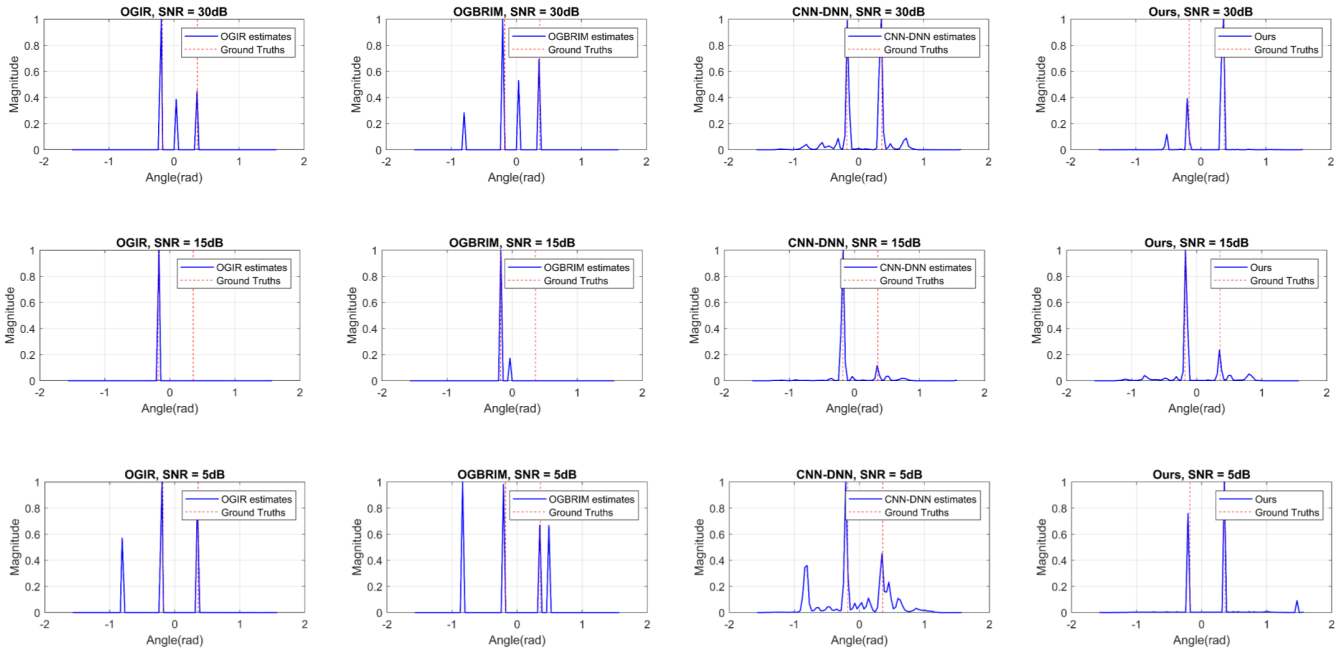


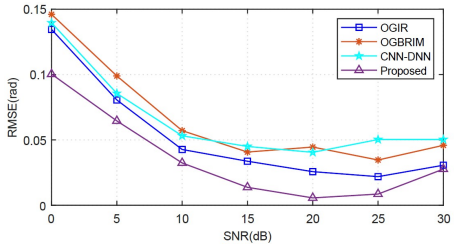
Fig. 4: Spectral outputs for 10-element SLA.

V. Performance Evaluation

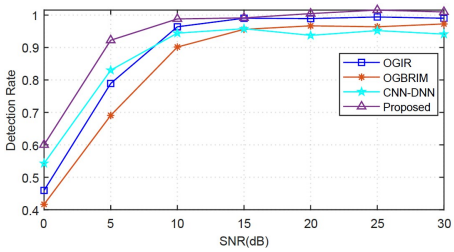
In this section, we evaluate the performance of the proposed method in terms of detection rate and root mean square error (RMSE). The angular space between -90° and 90° is discretized into 91 fixed grids, each with the grid size of 2°. An angle estimation is considered successful if the absolute errors of all estimated DOAs are within a threshold of 0.5°. Otherwise, it is considered a failure. The detection rate is defined as $\frac{N_s}{N_t}$, where N_t is the total number of Monte Carlo tests and N_s is the number of successful tests. The RMSE is calculated as $\text{RMSE} = \sqrt{\frac{1}{N_s K}\sum_{t=1}^{N_s}||\hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta}^*||_2^2}$, where $\hat{\boldsymbol{\theta}}_t$ represents the estimated DOA vector in the tth test round. A total of $1,024$ Monte Carlo trials are conducted for testing. The evaluation considers two off-grid target sources, $\boldsymbol{\theta}^* = [-10.28^\circ, 20.56^\circ]$, tested with the two SLA configurations described in Section IV-B. For comparison, we include the OGIR algorithm [25], representing algorithm-based methods, and the CNN-DNN network [34], representing network-based approaches. Additionally, the algorithm framework proposed in this paper, referred to as "OGBRIM", is used as a baseline for comparison.

We first compare the RMSE and detection rate for DOA estimation using an 18-element SLA and a 10-element SLA. As illustrated in Figure 5 for the 18-element SLA, the proposed method achieves the lowest RMSE and the highest detection rate, showcasing superior estimation performance. In Figure 6, while all methods exhibit performance degradation with the sparser 10-element SLA, the proposed method consistently maintains lower RMSE and higher detection rates, highlighting its robustness to array sparsity.

Additionally, we compare the spectra outputs of different



(a) RMSE vs. inpput SNR



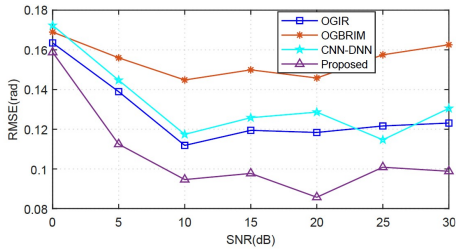
(b) Detection rate vs. input SNR

Fig. 5: 18-element SLA.

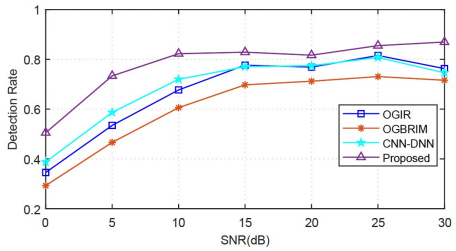
methods for the 10-element SLA, and the results are depicted in Figure 4. It is observed that the proposed method not only suppresses spurs effectively but also resolves signals across various input SNR levels. In contrast, other algorithms, such as OGIR and OGBRIM, tend to produce false estimates or miss targets.

VI. Conclusion

In this paper, we proposed a novel learning-based sparse Bayesian approach for one-bit off-grid DOA estimation. The



(a) RMSE vs. input SNR



(b) Detection rate vs. input SNR

Fig. 6: 10-element SLA.

method effectively combines the advantages of traditional Bayesian modeling with modern neural network architectures to achieve robust and accurate off-grid angle estimation. Simulation results demonstrated that the proposed approach outperforms state-of-the-art methods across a range of SNR scenarios. Future work will explore the generalization capabilities of the proposed method to enhance its reliability and extend the approach to handle real-world radar datasets.

REFERENCES

[1] S. Sedighi, M. Soltanalian, B. Ottersten *et al.*, "On the performance of one-bit DoA estimation via sparse linear arrays," *IEEE Transactions on Signal Processing*, vol. 69, pp. 6165–6182, 2021.

[2] S. Sun, A. P. Petropulu, and H. V. Poor, "MIMO radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 98–117, 2020.

[3] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, 2007.

[4] S. M. Patole, M. Torlak, D. Wang, and M. Ali, "Automotive radars: A review of signal processing techniques," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 22–35, 2017.

[5] S. Sun and Y. D. Zhang, "4D automotive radar sensing for autonomous vehicles: A sparsity-oriented approach," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 879–891, 2021.

[6] R. O. Schmidt, *A signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*. Ph.D. Dissertation, Stanford University, 1982.

[7] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[8] H. Singh, H. Sneha, and R. Jha, "Mutual coupling in phased arrays: A review," *International Journal of Antennas and Propagation*, vol. 2013, no. 348123, pp. 1–23, 2013.

[9] C.-Y. Chen and P. P. Vaidyanathan, "Minimum redundancy MIMO radars," in *Proceedings of 2008 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2008, pp. 45–48.

[10] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, 2010.

[11] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, 2011.

[12] S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1377–1390, 2015.

[13] C.-L. Liu and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray music," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1438–1442, 2015.

[14] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 11, pp. 2436–2449, 1991.

[15] C. Zhou, Y. Gu, X. Fan, Z. Shi, G. Mao, and Y. D. Zhang, "Direction-of-arrival estimation for coprime array via virtual array interpolation," *IEEE Transactions on Signal Processing*, vol. 66, no. 22, pp. 5956–5971, 2018.

[16] X. Shang, J. Li, and P. Stoica, "Weighted SPICE algorithms for range-Doppler imaging using one-bit automotive radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 4, pp. 1041–1054, 2021.

[17] O. Bar-Shalom and A. Weiss, "DOA estimation using one-bit quantized measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 868–884, 2002.

[18] K. Yu, Y. D. Zhang, M. Bao, Y.-H. Hu, and Z. Wang, "DOA estimation from one-bit compressed array data via joint sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1279–1283, 2016.

[19] X. Huang and B. Liao, "One-bit MUSIC," *IEEE Signal Processing Letters*, vol. 26, no. 7, pp. 961–965, 2019.

[20] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[21] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[22] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2082–2102, 2013.

[23] F. Li, J. Fang, H. Li, and L. Huang, "Robust one-bit Bayesian compressed sensing with sign-flip errors," *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 857–861, 2015.

[24] C. Zhou, Z. Zhang, F. Liu, and B. Li, "Gridless compressive sensing method for line spectral estimation from 1-bit measurements," *Digital Signal Processing*, vol. 60, pp. 152–162, 2017.

[25] L. Feng, L. Huang, Q. Li, Z.-Q. He, and M. Chen, "An off-grid iterative reweighted approach to one-bit direction of arrival estimation," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 6, pp. 8134–8139, 2023.

[26] N. Shlezinger, J. Whang, Y. C. Eldar, and A. G. Dimakis, "Model-based deep learning," *Proceedings of the IEEE*, vol. 111, no. 5, pp. 465–499, 2023.

[27] J. P. Merkofer, G. Revach, N. Shlezinger, T. Routtenberg, and R. J. G. van Sloun, "DA-MUSIC: Data-driven DoA estimation via deep augmented MUSIC algorithm," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 2, pp. 2771–2785, 2024.

[28] R. Zheng, H. Liu, S. Sun, and J. Li, "Deep learning based computationally efficient unrolling IAA for direction-of-arrival estimation," in *Proceedings of 2023 European Signal Processing Conference (EUSIPCO)*, 2023, pp. 730–734.

[29] R. Zheng, S. Sun, H. Liu, H. Chen, and J. Li, "Interpretable and efficient beamforming-based deep learning for single-snapshot DOA estimation," *IEEE Sensors Journal*, vol. 24, no. 14, pp. 22 096–22 105, 2024.

[30] J. Ren, T. Zhang, J. Li, and P. Stoica, "Sinusoidal parameter estimation from signed measurements via majorization–minimization based RELAX," *IEEE Transactions on Signal Processing*, vol. 67, no. 8, pp. 2173–2186, 2019.

[31] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.

[32] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, 2013.

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1026–1034.

[34] H. Chung, H. Seo, J. Joo, D. Lee, and S. Kim, "Off-grid DOA estimation via two-stage cascaded neural network," *Energies*, vol. 14, no. 228, pp. 1–11, 2021.