A Fast-Learning Sparse Antenna Array

Satish Mulleti*, Chiranjib Saha†, Harpreet S. Dhillon†, Yonina C. Eldar*
*Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel 76100
†Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, USA
Email: mulleti.satish@gmail.com, csaha@vt.edu, hdhillon@vt.edu, yonina.eldar@weizmann.ac.il

***Abstract*—Selecting a sparse subset of antennas to obtain high-resolution direction-of-arrival estimates while circumventing the complexity associated with using a large array is critical in many radar applications. Since this subset selection problem is combinatorial, deep learning has been recently proposed as a possible solution for efficiently solving it. However, the bottleneck in this approach is training data generation, which requires an exhaustive search over all possible subarrays. In this paper, we propose an efficient method for generating training data using ideas from submodular optimization. In particular, we use the log-determinant of the Cramér-Rao lower bound as our cost function due to its submodular structure. It is then minimized through a greedy optimization approach to determine the best subarray. We provide numerical simulations to validate the performance of the proposed array selection strategy. Our simulations show that the proposed approach is ten times faster in training than an exhaustive search method while providing comparable performance.**

I. Introduction

Direction-of-arrival (DOA) estimation is an important problem in applications such as radar and sonar imaging, radio astronomy, acoustic-source localization, and more [1]. The goal is to estimate the directions of far-field targets from the signals received at a set of sensors or an antenna array. For a given operating wavelength of the signals, the angular resolution of DOA estimation is inversely proportional to the number of antenna elements. Hence, ensuring high-resolution requires a large number of elements. There are two practical problems that arise when using a large number of elements. First, the receiver becomes expensive as each array element requires dedicated hardware that typically comprises of a down-converter and an analog-to-digital converter. Second, a large volume of data is generated from a full array which requires computationally expensive processing at the receiver end.

In order to strike a balance between the resolution and both the hardware and computational cost, one may select a sparse array that consists of fewer elements than a full array. A simple and straightforward sparse array choice is to randomly select a few elements from a full array. This is similar, in principle, to designing measurement matrices to achieve compressive sensing [2], [3]. The estimated DOAs from the random measurements may not be optimal. A better choice is to select the subarray that mimimizes some error statistics in the DOA estimation. However, the error statistics depends on the ground-truth DOAs which are not known *a priori*. In the DOA estimation problem, only noisy sensor measurements are available that are, typically, a non-linear function of the DOAs. Hence, it is necessary to consider methods that do not explicitly depend on the unknown parameters.

A supervised learning-based solution is naturally applicable to the problem at hand. In [4], [5], a DL-based approach for sparse antenna selection is developed by treating it as a classification problem, which is solved using a convolutional neural network (CNN). The network is trained with inputs as signals from a full array and output-labels as the desired sparse array. This DL based method was shown to be efficient compared with a random selection strategy. The method does not require *a priori* knowledge of DOAs and works directly with the measurements. However, a bottleneck of the DL-based method lies in the generation of the training data. In [4], [5], the labels are generated by minimizing the CRLB through an exhaustive search method. For large arrays, this approach is impractical. The process of training data generation is equivalent to solving antenna subset selection problem for known DOAs. In this paper, we propose the use of an efficient subset selection method for training data generation.

Prior Art on Subset Selection: Antenna selection is a combinatorial problem where the goal is to select K out of N antenna elements. The selection is achieved via minimizing a predefined cost function that is a measure of accuracy of DOA estimation from the sparse array measurements. In [6]–[8], switching matrix-based array selection method is proposed that minimizes a lower bound on DOA estimation by using a combinatorial search. However, the combinatorial search based methods are not always scalable. For example, for $N = 64$ and $K = 32$, the number of searches required is of the order of 10^{18}.

The selection problem can also be written as a Boolean-convex problem, where the objective function is a measure of estimation accuracy and the constrains are Boolean (can be either one or zero). Joshi and Boyd [9] consider log-determinant of the Cramér-Rao lower bound (CRLB) as the objective function. They proposed a convex relaxation of

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant No. 646804-ERC-COG-BNYQ, the Air Force Office of Scientific Research under grant No. FA9550-18-1-0208, and the Israel Science Foundation under grant No. 0100101. C. Saha and H. S. Dhillon are supported by US NSF under grant CNS-1923807.

978-1-7281-8942-0/20/$31.00 ©2020 IEEE

the optimization problem where the Boolean constraints are relaxed to convex constraints (can take values between zero and one). The relaxed problem is efficiently solved by off-the-shelf optimizer such as interior point methods. Chepuri and Leus [10] considered different functions of the CRLB, such as log-determinant, maximum eigenvalue, and trace, as the objective functions and proposed a relaxed optimization technique that is solved in polynomial time.

Convex relaxation methods may suffer from gradient computation step during each iteration and may not converge in finite amount of iterations. In order to improve the computational aspect, greedy algorithms have been proposed. These approaches takes finite number of iterations for a given size of the ground set and subset and also does not require the gradient computation. In greedy approaches, at each iteration, a new antenna element is added to already chosen elements such that it minimizes a given cost function. Godrich et al. [11] proposed a greedy algorithm for sensor selection where the Cramér-Rao lower bound (CRLB) is used as a performance metric.

Nemhauser et al. [12] showed that the greedy method results in a solution that is closer to the solution by an exhaustive search. In particular, they showed that the value of cost function for a greedy solution is $\left(1-e^{-1}\right)$ of the value of the cost function of an exhaustive search, where e is the Euler's number, provided that the cost function is submodular [13]. A submodular function is a set function that exhibits diminishing marginal gains, that is, adding additional elements results in diminishing benefits. A greedy optimization method with a submodular cost function is called *submodular optimization*. Not surprisingly, submodular optimization has found many applications in wireless networking. For instance, several submodular optimization-based sensor selection methods have been proposed in [14]–[19]. The primary difference between these methods is the choice of the cost functions: [14], [15], [18] use a mutual information (MI)-based cost function, [17] uses frame-potential (FP) of linear measurements, [16] uses maximum a posterior (MAP)-based cost, and [19] uses log-determinant of CRLB.

Contributions: In this paper, we propose an efficient data generation approach for training of the CNN in [4], [5]. To generate the training data, starting from the ground-truth DOAs, we first construct measurements corresponding to the full array which act as an input to the CNN. The computation of the output labels, that is, the sparse arrays, is a sensor selection problem for a given DOA. To this end, we consider a submodular optimization method to estimate the subsets that are subsequently used to train the CNN. Hence, the approach is a combination of submodular optimization and DL, where explicit knowledge of the unknown DOAs is used for training. Once trained, the network can be used to estimate subarrays for any unseen measurements without knowing the DOAs. We compare different submodular cost functions for the DOA estimation problem and show that the log-determinant of CRLB [19] results in a minimum error in estimation of DOAs from the sparse array. We show that the performance of the resulting subarray is comparable to the performance of the CNN that is trained through an exhaustive search. Moreover, for selecting 10 antennas out of 16, the proposed training approach is shown to be 10 times faster than the approach in [4].

The paper is organized as follows. In the next section, we introduce the signal model for DOA estimation and formally define the problem statement. In Section III, we give a brief introduction to submodular optimization and present the proposed fast data labeling approach. Simulation results are presented in Section IV followed by conclusions.

Throughout the paper, we reserve boldface lowercase, boldface uppercase, and calligraphic letters for vectors, matrices, and index sets, respectively. We denote the transpose and Hermitian by $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$, respectively. The number of ways K elements are chosen from N elements is denoted by $\begin{pmatrix} N \\ K \end{pmatrix} = \frac{N!}{K!(N-K)!}$. The expectation operator is denoted by E. The real, imaginary, and angle of any complex numbers are written as $\mathrm{R}e(\cdot)$, $\mathrm{I}m(\cdot)$, and $\angle(\cdot)$, respectively.

II. Signal Model and Problem Formulation

A. Signal Model

Consider an N element antenna linear array system along the x-axis. The position of the n-th element is p_n^x. Assume L targets in the far-field region with respect to the antenna array. The angle of arrival of the signal from the ℓ-th target is denoted by θ_ℓ. We assume that the targets lie on a grid with M points where $M \gg L$.

The n-th array element receives the signal $y_n(t)$. For this setting, the signal received at the full antenna array is given by

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{w}(t), \tag{1}$$

where $\mathbf{y}(t) = [y_1(t),\ y_2(t), \cdots, y_N(t)]^{\mathrm{T}}$. The matrix $\mathbf{A} \in \mathbb{C}^{N\times M}$ is the array steering matrix whose nm-th entry is given by $e^{-\mathrm{j}\frac{2\pi}{\lambda}p_n^x \sin(\theta_m)}$. The functions $\mathbf{x}(t)$ has L nonzero values for each $t \in \mathbb{R}$. The non-zero entries of $\mathbf{x}(t)$ determines the DOAs. In particular, if the support of $\mathbf{x}(t)$ is denoted by a set of integers $\{m_\ell\}_{\ell=1}^L$ then the DOAs of the sources are $\{\theta_{m_\ell}\}_{\ell=1}^L$. The corresponding signals in $\mathbf{x}(t)$, give as $\{x_{m_\ell}(t)\}_{\ell=1}^L$, are narrow band deterministic signals that are transmitted by the source at a wavelength λ. The term $\mathbf{w}(t)$ denotes measurement noise whose samples are assumed to be spatially (across sensors) and temporally, independent and identically distributed, and are circular complex Gaussian random variables with zero mean and variance σ_w^2.

B. Problem Formulation

For simplicity, we omit the time index t from the measurement model (1) and hence write

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}. \tag{2}$$

The DOA estimation problem is equivalent to estimating the support of the sparse vector $\mathbf{x}$,

$$\mathcal{S}_{\mathbf{x}} = \text{support}(\mathbf{x}). \tag{3}$$

As discussed earlier, it is desirable to estimate $\mathcal{S}_{\mathbf{x}}$ from a subset of measurements of $\mathbf{y}$. Let $\mathcal{K}$ be a subset of the set $\mathcal{N} = \{1, 2, \cdots, N\}$ with cardinality $|\mathcal{K}| = K$. Then $\mathbf{y}_{\mathcal{K}}$ denotes K subsamples of $\mathbf{y}$ that are indexed by $\mathcal{K}$.

The objective is to choose the best subset $\mathcal{K}$ out of $Q = \binom{N}{K}$ possible choices such that the error in estimating $\mathcal{S}_{\mathbf{x}}$ from $\mathbf{y}_{\mathcal{K}}$ is minimized. This amounts to solving the following optimization problem

$$\mathcal{K}^{\text{Opt}}_{\mathcal{S}_{\mathbf{x}}} = \underset{\mathcal{K}\in\mathcal{N}}{\text{argmin}} \quad \text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K}) \quad \text{s. t.} \quad |\mathcal{K}| = K, \tag{4}$$

where $\hat{\mathcal{S}}_{\mathbf{x}}$ is an estimation of $\mathcal{S}_{\mathbf{x}}$ from $\mathbf{y}_{\mathcal{K}}$. The cost function $\text{C}()$ is a real-valued function that measures the reconstruction error. The mean-squared error (MSE) $\text{E}\|\mathcal{S}_{\mathbf{x}} - \hat{\mathcal{S}}_{\mathbf{x}}\|_2^2$ is one possible choice of the cost function.

The goal is to solve (4) in order to generate the training data for DL methods.

III. A Fast-Learning Sparse Array

We employ a CNN-based deep network classifier for sparse antenna selection [4]. We train the CNN to find the best sparse array from the covariance matrix of the received signal as in [4], [5]. For fast-training the CNN, we propose to solve (4) by using a submodular-optimization based data generation method.

A. Submodular Function and Greedy Approach

The cost function in (4) depends on the reconstruction algorithm, the ground-truth, $\mathcal{S}_{\mathbf{x}}$, and the subset $\mathcal{K}$. Hence, for a fixed reconstruction algorithm and a given $\mathcal{S}_{\mathbf{x}}$, it is a set function that maps any subset of $\mathcal{N}$ to a real number. The function $\text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K})$ is submodular if it satisfies the property of decreasing marginals. Mathematically, $\text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K})$ is submodular if for every $\mathcal{K}_1 \subseteq \mathcal{K}_2 \subseteq \mathcal{N}$ the cost satisfies the following inequality

$$\text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K}_1 \cup \{i\}) \geq \text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K}_2 \cup \{i\}), \quad \forall i \in \mathcal{N}\backslash\mathcal{K}_2.$$

Further, the cost is monotonic if for every $\mathcal{K}_1 \subseteq \mathcal{K}_2 \subseteq \mathcal{N}$,

$$\text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K}_1) \leq \text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K}_2). \tag{5}$$

In [12], it is shown that if $\text{C}(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K})$ is a monotonic submodular function, then the optimization problem in (4) can be solved in a near-optimal sense by applying a K-step greedy method. Algorithm 1 summarizes this procedure. The MSE is a preferred cost function, however, MSE is not submodular in general [17]. One alternative is to minimize a scalar function of the CRLB that denotes a lower bound on the variance of unbiased estimators. Let $\text{C}_r(\mathcal{S}_{\mathbf{x}}, \mathcal{K})$ denote the CRLB in estimation of $\mathcal{S}_{\mathbf{x}}$ from $\mathbf{y}_{\mathcal{K}}$. Then the cost function

$$\log\det(\text{C}_r(\mathcal{S}_{\mathbf{x}}, \mathcal{K})) \tag{6}$$

Algorithm 1 Greedy Algorithm for Submodular Optimization

Initialize: $\mathcal{K} = \varnothing$
for $k = 1$ to K **do**
  [S1] $i^* = \max_{i\in\mathcal{N}\backslash\mathcal{K}} \; -\text{C}\left(\mathcal{S}_{\mathbf{x}}, \hat{\mathcal{S}}_{\mathbf{x}}, \mathcal{K}\cup\{i\}\right)$
  [S2] $\mathcal{K} \leftarrow \mathcal{K}\cup\{i^*\}$
end for

is submodular and monotone provided that $\mathcal{K} \neq \varnothing$ [19, Theorem VI.2]. Through simulations, we show that the submodular cost function in (6) results in a lower MSE in estimation of DOAs compared with the other choices of submodular costs.

B. A Fast-Training Data Generation

Selection of a K-element sparse subarray from an N-element full array among $Q = \begin{pmatrix} N \\ K \end{pmatrix}$ possible choices is analogous to a classification problem with Q classes. Here each class represents a different choice of subarray. A CNN is trained to achieve this classification [4].

The input to the CNN is an autocorrelation matrix that is computed from the signals received from a full array. Starting from P snapshots of the measurements, $\{\mathbf{y}(t_p)\}_{p=1}^P$, we construct a sample autocorrelation matrix

$$\mathbf{R} = \frac{1}{P}\sum_{p=1}^{P} \mathbf{y}(t_p)\mathbf{y}^{\text{H}}(t_p) \in \mathbb{C}^{N\times N}, \tag{7}$$

where we assume that the targets are not changing their directions over the snapshots. In other words, the vectors $\{\mathbf{x}(t_p)\}_{p=1}^P$ have common support. The matrices $\mathbb{R}e(\mathbf{R})$, $\mathbb{I}m(\mathbf{R})$, and $\angle(\mathbf{R})$ act as input to the CNN and it outputs an estimate of the subset $\mathcal{K}$.

In [4], [5], the output label $\mathcal{K}$ is selected by minimizing the mean CRLB that is computed by averaging the CRLBs of DOAs for the L targets. Minimization is achieved by an exhaustive search over all possible subsets. Then, by using a set of training features-label pairs $((\mathbb{R}e(\mathbf{R}), \mathbb{I}m(\mathbf{R}), \angle(\mathbf{R})), \mathcal{K})$ a network is learned that can be used to solve the optimization problem in (4).

A submodular-cost based data labeling approach is scalable to any array size. We use the cost function in (6) and then by applying Algorithm 1 estimate a subset for a given $\mathcal{S}_{\mathbf{x}}$. Then the CNN is trained with the input-output pairs $((\mathbb{R}e(\mathbf{R}), \mathbb{I}m(\mathbf{R}), \angle(\mathbf{R})), \mathcal{K}_{\text{greedy}})$, where $\mathcal{K}_{\text{greedy}}$ denotes the subset estimated by submodular optimization. During training, we assume that $\mathbf{x}$ is not changing across snapshots. We summarize the steps for generating the training data in Algorithm 2. The training steps are the same as in [4] except for the Step-3. The proposed sparse array selection approach is generic and is applicable to any linear and non-linear measurement model.

The design of the CNN layers is similar to that in [4], [5]. We recall the CNN structure for the sake of completion. The CNN consists of nine layers. The input layer of CNN accepts

Algorithm 2 Fast training data generation for sparse array selection.

Input: Number of antennas N, number of targets L, sparse subarray size K, number of data realizations T, number of DOA angles S, number of snapshots P and σ_w^2.

Output: Training data: Input-output pairs consisting of sample covariances $\mathbf{R}^{(s,i)}$ and output labels $\mathcal{K}^{(s,i)}$ for $s = 1, \ldots, S$ and $i = 1, \ldots, T$.

1: Select a set of L-sparse target DOAs and corresponding amplitudes $\{\mathbf{x}^{(s)}\}_{s=1}^{S}$.

2: For each $\mathbf{x}^{(s)}$, generate P snapshots of the data by using (1) and then compute the autocorrelation matrix $\mathbf{R}^{(s)}$ by using (7).

3: For each $\mathbf{x}^{(s)}$, compute the subarrays $\mathcal{K}^{(s)}$ by using the cost function (6) and Algorithm 1.

4: Repeat steps 1 to 3 for T times by chaining the amplitudes of each $\mathbf{x}^{(s)}$ and generate the input-output pairs $(\mathbf{R}^{(s,i)}, \mathcal{K}^{(s,i)})$ for $s = 1, \cdots, S$ and $i = 1, \cdots, T$.

the two-dimensional inputs $(\mathbb{R}e(\mathbf{R}), \mathbb{I}m(\mathbf{R}), \angle(\mathbf{R}))$ in three real-valued channels. The second, fourth and sixth layers are convolutional layers with 64 filters of size 2×2. To reduce the dimensions, the third and fifth layers are designed as max-pooling that reduce the dimension by half. The next two layers are fully connected layers with 512 units whose 50% are randomly dropped out to minimize overfitting during training. Each convolutional and fully-connected layers are followed by rectified linear units (ReLU) where $\text{ReLU}(a) = \max(a, 0)$. The dimension of the output layer is equal to the number of classes. As shown in [4], the actual number of classes is much smaller than Q.

To train the network, we collect data for S target instances and for T realizations that results in ST data samples. We train the proposed network in MATLAB on a PC with Intel Core i7-6700 CPU. During the training process, 10% of the training data is used for validation. We used the stochastic gradient descent algorithm with momentum for updating the network parameters with learning rate of 0.05 and mini-batch size of 500 samples for 200 epochs.

IV. Simulation Results

We evaluate the performance of the proposed data labeling method and resulting DOA estimation. First, we access the performance of greedy approaches and justify the use of cost function in (6). Then we compare the results of CNNs with different training methods.

Throughout the simulations, we consider one-dimensional linear arrays. The signal-to-noise ratio (SNR) is defined as $10\log_{10}(\|\mathbf{x}\|^2/M\sigma_w^2)$. Once the subsets are computed, orthogonal matching pursuit (OMP) is applied to estimate $\mathbf{x}$ from the measurements from the selected subsets. The normalized MSE $\|\mathbf{x} - \hat{\mathbf{x}}\|^2/\|\mathbf{x}\|^2$ is used as performance measure that is averaged over 500 independent realizations for a given SNR.

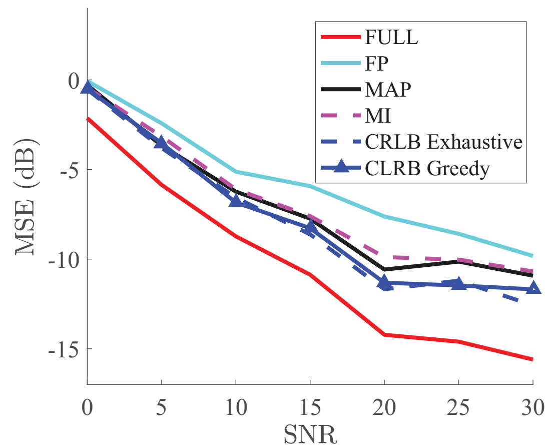

Fig. 1. A comparison of various submodular cost functions in terms of MSE in estimating DOAs from the selected subarrays. CRLB-based cost functions have 1 to 2 dB lower MSE compared to other methods for SNR $\geq$ 20 dB. Both standard (exhaustive) and greedy CRLB methods have comparable errors.

TABLE I

A comparison of average computation time and MSE for 20 dB SNR

Method	Time (in msec.)	MSE (in dB)
FP [17]	63.1	-6.66
MAP [16]	41.2	-6.98
MI [14], [15]	17.5	-6.82
CRLB Exhaustive [4]	510	-9.27
CRLB Greedy [19]	51	-9.18
Random	0.3	-6.56
DL [4], [5]	430	-8.66
DL greedy (proposed)	430	-8.58
Full	-	-11.50

with $N = 16$ elements. The objective is to choose $K = 10$ elements for $L = 3$ targets with $M = 100$.

A. Comparison of Submodular Methods

There are several choices of submodular cost functions such as MI, FP, MAP-based, and CRLB-greedy as in (6). In Fig. 1, we compare the performances of these methods for different SNRs. The objective is to choose $K = 10$ elements from $N = 16$ element array for $L = 3$ targets with $M = 100$. We also compared the error for a full array and an exhaustive search method that minimizes (6). The latter approach is termed as CRLB Exhaustive. It is observed that, among different submodular cost functions, CRLB greedy has the lowest MSE. In Table 1, a comparison of the MSE and computation time is presented for different methods. For CRLB cost, performances of both exhaustive search and greedy method are similar for different SNRs. However, the CRLB greedy method is 10 times faster than CRLB Exhaustive in the present settings.

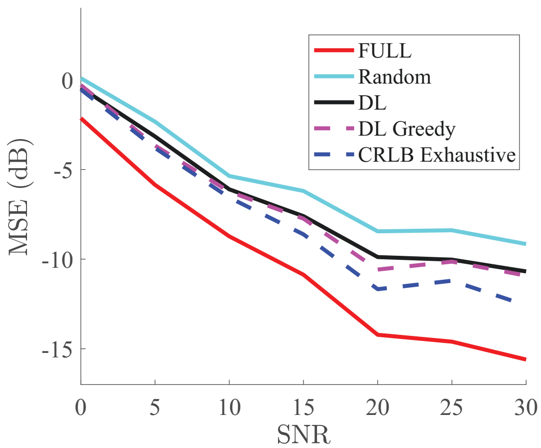

Fig. 2. A comparison of DL-based methods with two different training approaches. Both standard DL method and DL greedy method, trained by using submodualr optimization, have similar errors. DL-based methods are adaptive and perform better than the random selection method.

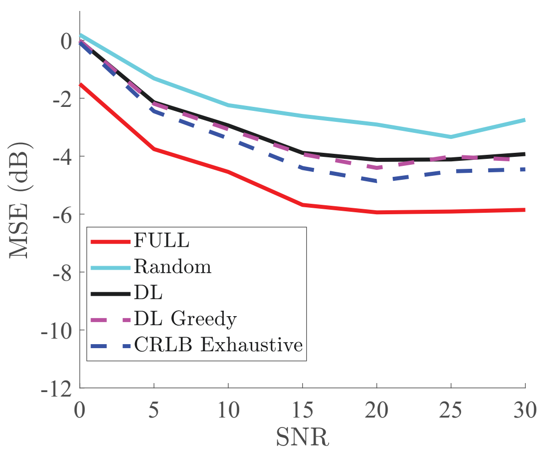

Fig. 3. A comparison of DL-based methods for closely spaced targets. DL-based methods are adaptive and perform better than the random selection method.

B. Comparison of DL-based methods

We consider a full array with $N = 16$ elements, with an objective to choose $K = 10$ elements. The grid size is $M = 100$. We train the proposed CNN structure for $L = 3$ targets whose DOAs are selected uniformly at random. We select $S = 5000$ different target locations and $T = 100$ data realizations with 100 data snapshots. The SNR for training is set to be 20 dB. Two different CNNs are trained. The first network is trained as in [4] where an exhaustive search is used for data labeling. The second network is trained by the data generation process described in Algorithm 1. Once trained, the sparse subsets are estimated for different measurements and subsequently, OMP is applied to estimate the DOAs and corresponding amplitudes from the subsamples.

In Fig. 2, we compare the performance of DL based methods with a random selection method and CRLB Exhaustive. The performance of the DL based technique is similar for both exhaustive and greedy training methods. The latter is computationally efficient and practical for larger arrays. We observe that data-based DL approaches perform better than a random selection strategy. For high SNRs ($\geq$ 20dB), CRLB Exhaustive method results in a 2 dB less error compared with the DL-based methods. In terms of the computation time (See Table 1), the DL based methods are slower compared with the greedy methods. However, both the CRLB Exhaustive method and greedy approaches requires explicit knowledge of the DOAs to estimate the subsets whereas DL-based methods do not require that.

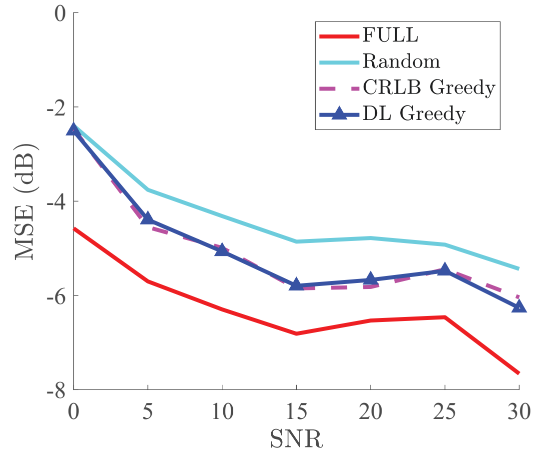

Fig. 4. A comparison of DL-based methods for 10 targets. DL-based methods are adaptive and perform better than the random selection method.

C. Comparison of DL-based methods for closely-spaced targets

In this simulation, we consider two closely spaced targets with $N = 16$, $K = 10$, and $M = 100$. Specifically, we consider the DOAs θ_{m_1} and θ_{m_2} such that $|\sin(\theta_{m_1}) - \sin(\theta_{m_2})| \leq 1/N$. In other words, the spatial frequencies correspond to DOAs are smaller than the Fourier resolution limit. In Fig. 3, we compare the performance of DL based methods with a random selection method, CRLB Exhaustive, and the full array. By using fewer antenna elements, the accuracy of the DL-based methods is reduced by 2 dB compared with the full array. The DL-based methods have 1 dB less error compared with the random selection.

D. Comparison of DL-based methods for a large array

We compare the performance for subset selection methods for a large array and a large number of targets. Specifically, we consider selection of $K = 30$ antennas from $N = 64$. The number of targets is $L = 10$ and we consider a grid size of $M = 200$. A CRLB-exhaustive method for the training data generation amounts to search over choice of the order of 10^{18} and hence can not be used. Hence we compare only the greedy methods that are salable to large arrays. In Fig 4, we compare the performance of the DL-greedy and CRLB-greedy methods with the full and random array. We observe that both CRLB-greedy and DL-greedy methods perform similarly and the DL-greedy method performs better than the random selection.

V. Conclusion

In this paper, we have proposed a submodular optimization-based data generation approach to train a CNN that performs the task of subset selection. Our hybrid approach uses a CRLB-based submodular cost function that is minimized by a computationally efficient greedy algorithm. The resulting sparse arrays are used as output labels to train a CNN with the sample covariance matrix of the received signal as input. DOAs are estimated from the partial measurements by applying the OMP algorithm. The training of the resulting network is 10 times faster than conventional network without any performance degradation.

References

[1] T. E. Tuncer and B. Friedlander, *Classical and Modern Direction-of-Arrival Estimation*. Elsevier Science, 2009.

[2] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*, ser. Compressed Sensing: Theory and Applications. Cambridge University Press, 2012.

[3] A. De Maio, Y. C. Eldar, and A. M. Haimovich, *Compressed Sensing in Radar Signal Processing*. Cambridge University Press, 2019.

[4] A. M. Elbir, K. V. Mishra, and Y. C. Eldar, "Cognitive radar antenna selection via deep learning," *IET Radar, Sonar & Navigation*, vol. 13, no. 6, pp. 871–880, 2019.

[5] A. M. Elbir, S. Mulleti, R. Cohen, R. Fu, and Y. C. Eldar, "Deep-sparse array cognitive radar," in *Int conf. Sampling Theory and Appli. (SampTA)*, 2019, pp. 1–4.

[6] J. Tabrikian, O. Isaacs, and I. Bilik, "Cognitive antenna selection for DOA estimation in automotive radar," in *IEEE Radar Conference*, 2016, pp. 1–5.

[7] D. Mateos-Núñez, M. A. González-Huici, R. Simoni, and S. Brüggenwirth, "Adaptive channel selection for DOA estimation in MIMO radar," in *Int. Workshop on Comput. Adv. in Multi-Sensor Adaptive Process. (CAMSAP)*, 2017.

[8] K. V. Mishra, Y. C. Eldar, E. Shoshan, M. Namer, and M. Meltsin, "A cognitive sub-Nyquist MIMO radar prototype," *IEEE Trans. Aerosp. Electron. Sys.*, vol. 56, no. 2, pp. 937–955, 2019.

[9] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, 2009.

[10] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, 2015.

[11] H. Godrich, A. P. Petropulu, and H. V. Poor, "Sensor selection in distributed multiple-radar architectures for localization: A knapsack problem formulation," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 247–260, 2012.

[12] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions — I," *Math. Program.*, vol. 14, no. 1, pp. 265–294, 1978.

[13] S. Fujishige, *Submodular Functions and Optimization*, ser. ISSN. Elsevier Science, 2005.

[14] G. Shulkind, S. Jegelka, and G. W. Wornell, "Sensor array design through submodular optimization," *IEEE Trans Info. Theory*, vol. 65, no. 1, pp. 664–675, 2019.

[15] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *J. Mach. Learn. Res.*, vol. 9, p. 235–284, Jun. 2008.

[16] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *IEEE Conf. Decision and Control (CDC)*, 2010, pp. 2572–2577.

[17] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," *IEEE Tran Signal Process.*, vol. 62, no. 5, pp. 1135–1146, 2014.

[18] M. Naeem, S. Xue, and D. C. Lee, "Cross-entropy optimization for sensor selection problems," in *Int. Symp. Commun. and Info. Tech.*, 2009, pp. 396–401.

[19] E. Tohidi, M. Coutino, S. P. Chepuri, H. Behroozi, M. M. Nayebi, and G. Leus, "Sparse antenna and pulse placement for colocated mimo radar," *IEEE Tran. Signal Process.*, vol. 67, no. 3, pp. 579–593, 2019.