# Deep Autoencoders for DOA Estimation of Coherent Sources using Imperfect Antenna Array

Aya Mostafa Ahmed, Omar Eissa and Aydin Sezgin
Institute of Digital Communication Systems
Faculty of Electrical and Computer Engineering
Ruhr-Universitt Bochum, Germany
Email: {aya.mostafaibrahimahmad; omar.eissa; aydin.sezgin}@rub.de

*Abstract*—In this paper a robust algorithm for DOA estimation of coherent sources in presence of antenna array imperfections is presented. We exploit the current advances of deep learning to overcome two of the most common problems facing the state of the art DOA algorithms (i.e. coherent sources and array imperfections). We propose a deep auto encoder (AE) that is able to correctly resolve coherent sources without the need of spatial smoothing, hence avoiding possible processing overhead and delays. Moreover, we assumed the presence of array imperfections in the received signal model such as mutual coupling, gain/ phase mismatches, and position errors. The deep AE is trained using the covariance matrix of the received signal, where it alleviates the effect of imperfections, and at the same time act as a filters for the coherent sources. The results show significant improvement compared to the methods used in the literature.

*Index Terms*—DOA, Coherent sources, antenna array imperfections, Deep autoencoders, MUSIC

## I. Introduction

Direction-of-arrival (DOA) estimation is a common problem in different fields, including wireless communication, astronomical observation, radar and sonar. There are various challenges facing DOA estimation such as accuracy and precision of estimates in non-ideal scenarios, e.g., multipaths and antennas imperfections. There are many DOA estimation techniques that can be classified to multiple categories. On one hand, there are the conventional methods which depends on the locations of peaks in the spatial spectrum to determine the DOA, i.e., delay and sum beamforming and minimum variance distortionless response (MVDR) [1]. On the other hand there are also the subspace methods e.g., MUltiple SIgnal Classification (MUSIC) [2] which depends on the eigenstructure of the spatial correlation matrix, offering high resolution DOA estimation. There also exists parametric-based approach like the maximum likelihood (ML) technique, which uses a statistical structure for the process of generating data [3]. However, the application of those methods in real life applications is very limited, as it requires the accurate knowledge of the received signal without any errors. Hence, they are limited to specific scenarios where the antenna array measurements are ideal and the sources are not correlated. For example, the MUSIC algorithm has the ability to detect and measure multiple sources at the same time with high precision measurement only under ideal array conditions in a

non multipath environment. However its performance degrades significantly in the existence of array imperfections or when the sources are coherent (fully correlated) due to multi-path. In such scenario, the rank of the signal covariance matrix is less than the number of impinging signals, which significantly affects the estimation performance. To solve this problem, spatial smoothing techniques have been widely used to remove the coherence between sources or de-correlate their signals by reconstructing a full rank covariance matrix before going through the estimation algorithm [4]. However, this adds processing overhead, making it difficult to use in real time applications. Moreover, spatial smoothing only solves the coherence problem, and does not solve other problems like antenna array imperfections which is a common problem in practice. As a matter of fact, those imperfections generally occur in practical applications due to the non-idealities in the antenna array such as mutual coupling, gain / phase gradual changes over time and as well as changes in the antenna locations. For instance, mutual coupling occurs due to interference from nearby antennas during transmitting [5], while gain and phase inconsistency can result from the aging of electrical components or thermal effects. All of those factors change the antenna array response, causing significant degradation in the performance of DOA algorithms like MUSIC [5]. Many approaches in the literature addressed those problems using auto calibration algorithms as in [6], [7]. However, such algorithms require the prior knowledge of the imperfections formulations, which can be challenging in practical applications. Other approaches used the latest advances in machine learning, and deep learning to solve this problem as in [8]. However, the authors in [8] only focused on the imperfections problem, where they proposed a neural network followed by a classifier. However, they didn't take into consideration potential correlation among the sources that can result from multipath.

In this paper we propose a deep neural network framework for DOA estimation that is generic and robust against both array imperfections, along with the existence of coherent sources. The results show that our DNN can correctly estimate the directions of spatially close coherent sources without any prior spatial smoothing techniques. Furthermore, our algorithm showed significantly better results compared to the state of the art methods for coherence removal as spatial smoothing MUSIC algorithm [4], and to similar algorithms for imperfections

as aforementioned approach found in [8].

## II. SYSTEM MODEL

### A. Array Imperfections

Assume that $K$ sources are impinging on a receive array consisting of $M$ antennas, whose DOA are $\theta_1, \ldots, \theta_K$. The received signal is sampled at $N$ uniquely spaced time instants $t_1, \ldots, t_N$ to obtain multiple snapshots gathered in matrix $\mathbf{Z} = [\mathbf{z}(t_1), \ldots, \mathbf{z}(t_N)]$, with

$$\boldsymbol{z}(t_n) = \sum_{k=1}^{K} \boldsymbol{a}(\theta_k) x_k(t_n) + \boldsymbol{w}(t_n), \quad \text{for } n = 1, \ldots, N. \tag{1}$$

$x_k(t_n)$ is the transmit waveform of the $k$ th source, and $\boldsymbol{w}(t_n)$ is the zero-mean Gaussian noise. $\boldsymbol{a}(\theta_k)$ denotes the imperfection free steering vector and is defined as

$$\boldsymbol{a}(\theta_k) = \frac{1}{\sqrt{M}} \left[ 1, e^{-j2\pi \frac{d}{\lambda} \sin \theta_k}, \cdots, e^{-j2\pi \frac{d}{\lambda}(M-1)\sin\theta_k} \right]^T \tag{2}$$

The covariance matrix of the received signal $\boldsymbol{z}$ is

$$\boldsymbol{R} = \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{z}(t_n) \boldsymbol{z}^H(t_n). \tag{3}$$

The model in (1) is the idealistic received signal without any imperfections in the antenna array, which is commonly used in the literature. However, it is quite impractical. Hence, we re-define (1) as

$$\boldsymbol{z}(t_n) = \sum_{k=1}^{K} \boldsymbol{a}(\theta_k, \boldsymbol{e}) x_k(t_n) + \boldsymbol{w}(t_n), \quad n = 1, \ldots, N \tag{4}$$

where $\boldsymbol{a}(\theta_k, \boldsymbol{e})$ is the array response after adding the array imperfections. Here we consider gain and phase errors ($\mathbf{e}_g$ and $\mathbf{e}_p$), antenna position error ($\mathbf{e}_{pos}$), and mutual coupling error ($\mathbf{e}_{mc}$). To this end, the definition of $\boldsymbol{a}(\theta_k, \boldsymbol{e})$ would be as [8]

$$\begin{aligned} \boldsymbol{a}(\theta, \boldsymbol{e}) = (\boldsymbol{I}_M + \alpha_{mc} \boldsymbol{E}_{mc}) \times (\boldsymbol{I}_M + \text{diag}(\alpha_g \boldsymbol{e}_g)) \\ \times \text{diag}(\exp(j\alpha_e \boldsymbol{e}_p)) \times \boldsymbol{a}(\theta, \alpha_{pos} \boldsymbol{e}_{pos}) \end{aligned} \tag{5}$$

$$\boldsymbol{e}_g = [0, \underbrace{0.2, \ldots, 0.2}_{\frac{M}{2}}, \underbrace{-0.2, \ldots, -0.2}_{\frac{M}{2}-1}]^T \tag{6}$$

$$\boldsymbol{e}_p = \left[ 0, \underbrace{-\frac{\pi}{6}, \ldots, -\frac{\pi}{6}}_{\frac{M}{2}}, \underbrace{\frac{\pi}{6}, \ldots, \frac{\pi}{6}}_{\frac{M}{2}-1} \right]^T \tag{7}$$

The position biases are

$$\boldsymbol{e}_{pos} = [0, \underbrace{-0.2, \ldots, -0.2}_{\frac{M}{2}}, \underbrace{0.2, \ldots, 0.2}_{\frac{M}{2}-1}]^T \times d \tag{8}$$

The mutual coupling coefficient vector is

$$\boldsymbol{e}_{mc} = [0, \gamma^1, \ldots, \gamma^{19}]^T \tag{9}$$

where $\gamma = 0.3 e^{j\frac{\pi}{3}}$ is the mutual coupling coefficient between adjacent antennas, $\alpha_i \in [0, 1], i \in \{g, p, pos, mc\}$ is weighting parameter for each error. The choice of the error values in (6),(7),(8) and (9) is system dependent, and can be changed accordingly. $\boldsymbol{E}_{mc}$ is defined as a Toeplitz matrix with parameter vector $\boldsymbol{e}_{mc}$ [8]. In addition to array imperfections, in any real scenario, the received signals would be highly correlated, due to the contribution of multi-paths which makes $\boldsymbol{R}$ rank deficient or singular. Hence, the next section will define the model of coherent sources.

### B. Coherent Sources

In order to generate the multi-path component of each source, we consider having $K$ sources arriving from $K$ directions. At time instant n, there are $k$ transmit signals $x_k(t_n)$, $\forall k = 1, \ldots, K$, which arrive as replica of one of them i.e. $x_1(t_n)$, but phase delayed and magnitude weighted [4]. Hence, the transmit waveform $x_k(t_n)$ in (4) can be redefined as

$$x_k(t_n) = g_k e^{j\phi_k} x_1(t_n), \quad k = 1, \ldots, K, \tag{10}$$

where $g_k$ is the amplitude factor of source $k$ and $\phi_k$ is the phase change of source $k$. Such model will impose rank deficiency on the covariance matrix structure in (3), causing the existing DOA algorithms to fail accordingly. To solve such problem, in the next section, we propose a deep autoencoder to remove the effects of both coherence and array imperfections.

## III. DOA ESTIMATION BASED ON DEEP LEARNING

### A. Deep Neural Network Architecture

A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers. The DNN learns the correct mathematical manipulation to map the input into the output. It has the ability to solve complex nonlinear problems [9]. Here, we propose a DNN architecture which is an autoencoder (AE), where the first hidden layer performs the function of an encoder as it reduces the dimension of the input by extracting the main features of the input. Afterwards, the encoding layer is followed by four hidden layers that help in the decoding process by retrieving the information to restore back the input of the AE. Table I shows the size of all the hidden layers. The output layer consists of six sub-regions where each region is considered as decoder by itself. Therefore, the AE has six decoders and each decoder retrieves specific information from the input, as shown in Fig. 1. The process of encoding-decoding helps decrease the impact of disturbances in the autoencoder input, through denoising the input, retrieving only the useful information. The disturbance is our case is mainly due to array imperfections, noise, and coherent sources.

### B. Learning Scheme

The input of the proposed DNN is the covariance matrix of the received signal as defined by (3). In order to reduce the dimension of the input layer, we consider only the correlated elements in the covariance matrix. By exploiting the symmetry
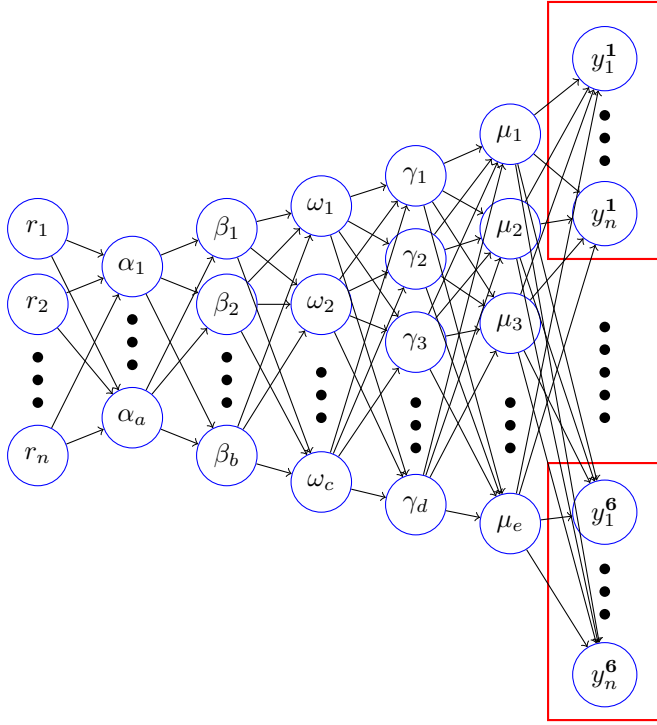
Fig. 1: Proposed autoencoder consisting of 5 hidden layers with hidden layer 1 acts like an encoder and the following layers combined act like a decoder

| Layer | Size |
|---|---|
| Input Layer | $n = 380$ |
| Hidden Layer 1 | $a = 190$ |
| Hidden Layer 2 | $b = 380$ |
| Hidden Layer 3 | $c = 570$ |
| Hidden Layer 4 | $d = 760$ |
| Hidden Layer 5 | $e = 950$ |
| Output Layer | $6 \times 380 = 2280$ |

TABLE I: Sizes of all layers

in the correlation matrix, only the strict upper or strict lower triangular part of the matrix can be considered [8]. In this paper, the off-diagonal upper right matrix of the covariance matrix is considered. For example, if we have $M = 3$ antennas, the covariance matrix will be

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \qquad (11)$$

then the following input is obtained

$$\hat{\boldsymbol{r}} = \begin{bmatrix} r_{12} & r_{13} & r_{23} \end{bmatrix}^T. \qquad (12)$$

Generally, let $\mathbf{R} \in \mathbb{C}^{M \times M}$, then the input vector is $\hat{\boldsymbol{r}} \in \mathbb{C}^{M(M-1)/2}$. Additionally, the input of the DNN must be real valued, hence, $\hat{\boldsymbol{r}}$ is converted from complex to real,

by concatenating $\mathrm{Real}\left\{\hat{\boldsymbol{r}}^T\right\}$ with $\mathrm{Imag}\left\{\hat{\boldsymbol{r}}^T\right\}$ to produce $\mathbf{r} \in \mathbb{C}^n$, where $n = M(M - 1)$.

$$\boldsymbol{r} = \left[\mathrm{Real}\left\{\hat{\boldsymbol{r}}^T\right\}, \mathrm{Imag}\left\{\hat{\boldsymbol{r}}^T\right\}\right]^T. \qquad (13)$$

The proposed autoencoder in Fig.1 decomposes its input into 6 spatial subregions, each spatial subregion is a specific range of angles and all subregions are of the same size. To define these subregions, 7 particular spatial angles were chosen such as

$$\theta^{(1)} < \theta^{(1)} < \cdots < \theta^{(7)} \qquad (14)$$

with constant gaps such that $\theta^{i+1} - \theta^i = \mathrm{constant} \ \forall i, i = 1, \ldots, 7$, and each subregion $j$ is defined as $\left[\theta^j, \theta^{j+1}\right]$ where $j = 1, \ldots, 6$. Therefore, if the input vector of the autoencoder $\boldsymbol{r}_k$ is generated using a signal impinging from source $k$ on the antenna array at angle $\Theta_k$ within the $j$-th subregion, then the output of the $j$-th decoder will be $\boldsymbol{r}_k$, while the output of the other decoders will be zero as there are no signals impinging from those range of angles. The AE is trained to be able to separate multiple signals transmitted from sources located in different subregions impinging onto the array simultaneously. Hence, it is able to decompose the input vector with components belonging to different subregions and extract the information that belongs to every subregion by retrieving it in the related decoders.

### C. Training Process

The data set is constructed by generating $I$ training samples of the covariance vector $\boldsymbol{r}$ corresponding to single-signal scenarios. The data was generated with random angles that spans all subregions. The output of each decoder is determined based on which subregion the generated $\boldsymbol{r}$ belong to. Alternatively, the decoder can be considered as a spatial filter. This filter extracts the covariance vector information that belongs to a particular subregion. In order to build up the training label of the entire output of the AE, the outputs of the six decoders are concatenated as follows

$$\boldsymbol{y} = [\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_6^T]^T$$
$$= \left[\boldsymbol{0}, \ldots, \boldsymbol{0}, \underbrace{\boldsymbol{r}^T(\Theta_k)}_{j \text{ th subregion}}, \boldsymbol{0}, \ldots, \boldsymbol{0}\right]^T, \qquad (15)$$

in which $\boldsymbol{y}_j$ is the output of the $j$ decoder. To train the autoencoder, the squared $l_2$-norm distance between the actual output and the expected one is used as the loss function. That is,

$$\boldsymbol{L}(\Theta_k) = \frac{1}{2}\|\tilde{\boldsymbol{y}}(\Theta_k)\|_2^2$$

where

$$\tilde{\boldsymbol{y}}(\Theta_k) = \boldsymbol{y}(\Theta_k) - \hat{\boldsymbol{y}}(\Theta_k)$$

and $\hat{\boldsymbol{y}}(\Theta_k)$ is the actual output of the autoencoder when $\boldsymbol{r}(\Theta_k)$ is the input. The optimizer used in the training process to minimize the loss function is a RMSProp optimizer [8].

## D. Scanning

After training the network shown in Fig. 1, the training label in (15) is used to estimate the original directions of the correlated sources impinging on the array. This is done through spatial scanning of the output of each decoder (i.e. filter) $\boldsymbol{y}_j$. The scanning phase aims at calculating the gain of each filter in all directions, in which the actual source angles would have large gain values, while the other directions would have much smaller gains. Afterwards, a threshold value is used to select the angles whose gains have peaks surpassing the threshold. The gain response of each filter is obtained by

$$g^{(j)} = \left| \overline{\boldsymbol{r}}^H \left( \Theta_k \right) \overline{\boldsymbol{y}}_j \right|, \quad j = 1, \ldots, 6, \tag{16}$$

where the superscript $(\bullet)^H$ is the conjugate transpose of the matrices and vectors, $\overline{\boldsymbol{r}}^H$ is the estimated value of $\hat{r}$, and $\overline{\boldsymbol{y}}_j$ is the complex version of the output of the $j$-th decoder $\boldsymbol{y}_j$. $\overline{\boldsymbol{y}}_j$ is obtained from $\boldsymbol{y}_j$ by concatenating the first half that represents the real values in $\boldsymbol{y}_j$ with their corresponding imaginary values in the second half of $\boldsymbol{y}_j$, similarly $\overline{\boldsymbol{r}}^H \left( \Theta_k \right)$ is obtained from $\boldsymbol{r}^T(\Theta_k)$ .

## IV. SIMULATION RESULTS

In this section, we carry out simulations to evaluate our proposed DNN. We used the python library *tensorflow* to design and process our DNN. The network is trained on $I = 1200$ samples, with learning rate of $0.001$, while the batch size is 100 and the number of epochs is 1000 epochs. We use a uniform linear array (ULA) of size $M = 20$ elements with spacing $d = \lambda/2$ to predict directions of signals impinging from sources located in the spatial range of $[-60°, 60°]$, which is divided equally into six subregions. The training samples are generated randomly from directions $\Theta_i = -60° + 0.1°, \forall i = 1, i = 1, \ldots, 1200$. The covariance input vector $\boldsymbol{r} \left( \Theta_i \right)$ is generated using $N = 800$ snapshots. To evaluate the performance of our algorithm, we use forward/backward spatial smoothing along with the MUSIC algorithm (SS-MUSIC) in [10], and compared it against our DNN in multipath environment by randomly changing $g_k$ and $\phi_k$ in (10) for every target $k$.

## A. Gain responses of each decoder (i.e filter)

The DNN is tested using a covariance vector obtained from two correlated sources located at $\theta_1 = -15°, \theta_2 = -5°$ respectively, which belongs to subregion 3, i.e. $[-20°, 0°]$, with signal to noise ratio $\text{SNR} = 10$ dB. Those specific directions were chosen, because they are in the same subregion, making them spatially close, hence it would be harder to separate them compared to distant sources. The threshold value to find the peaks in the scanning process is set to $= 0.3$, which is set by experience. Fig. 2 shows the gain obtained from (16) for all the filters. It can be depicted that the spatial gain response of the filter corresponding to the assigned subregion has higher peaks compared to the other filters. It is clear that the filter managed to differentiate between both angles despite the fact they are coherent, and the antenna array suffers from imperfection errors as stated in (5).
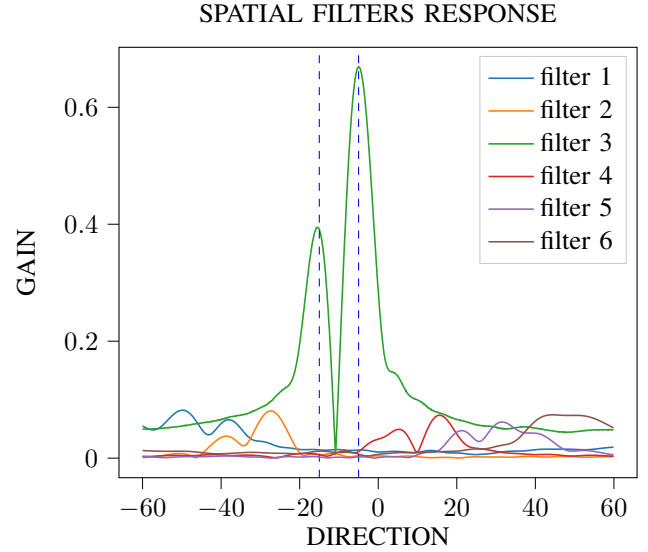


Fig. 2: Spatial filters response at SNR = 10 dB when two coherent signals located at $(\theta_1 = -15°, \theta_2 = -5°)$ are incident to the antenna array. The blue dashed lines are at $-15°$ and $-5°$ respectively
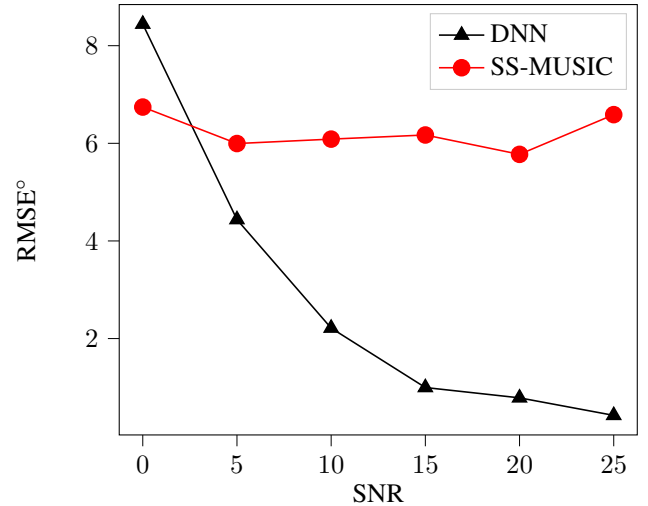


Fig. 3: DOA estimation RMSE in degrees of the proposed DNN method and spatial smoothing MUSIC algorithm for two impinging signals from directions $(\theta_1 = -15°, \theta_2 = -5°)$

## B. Performance against SS-MUSIC

Fig.3 compares DOA estimation performance of our proposed DNN with SS-MUSIC in presence of correlated sources and array imperfections combined. The average root mean square error (RMSE) in degrees is used to measure the accuracy of DOA estimates for various SNR. It can be shown from the figure that the RMSE of our DNN starts high at $\text{SNR} = 0$ dB, then decreases significantly compared to SS-Music as the SNR increases. It can be seen from the figure that the RMSE of SS-MUSIC algorithm is heavily impacted by the presence of imperfections due to the fact that it assumes

ideal steering vector model with no imperfections as in (1)

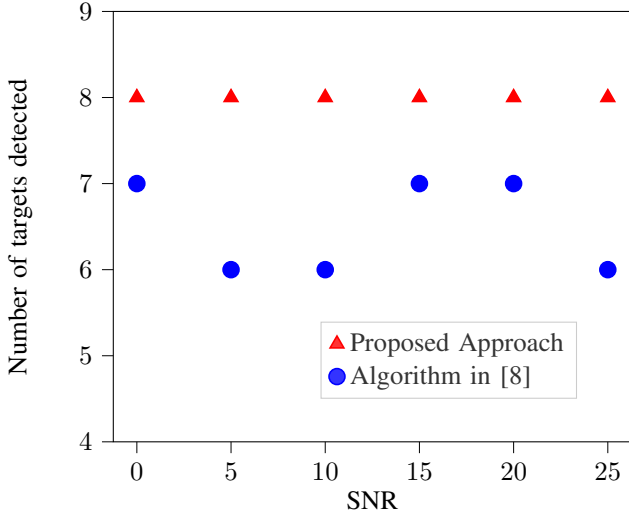*C. Perfomance against algorithm in [8]*



Fig. 4: DOA estimation RMSE of the proposed DNN framework and DNN found in [8] for $K = 8$ coherent sources

Fig. 4 compares the detection performance of our DNN compared to the algorithm in [8], where the authors only considered array imperfections and assumes perfectly uncorrelated sources for DOA estimation. The same training data was used for both algorithms for fair comparison. The figure shows a consistent behavior for our algorithm detecting all 8 targets, however the algorithm in [8] shows lower probability of detection due to its failure to detect all coherent sources.

## V. CONCLUSION

In this paper we presented a framwork for DOA estimation of correlated sources in presence of array imperfections. Our approach was based on Deep AE with 5 hidden layers, one acting as encoder and 4 as decoder. From the simulations we showed that the AE acts as a denoiser, where it could successfully remove the effect of coherence and imperfections producing accurate DOA estimates compared to the commonly used SS-MUSIC. Moreover, we compared our algorithm with the approach in [8] where the authors only deal with imperfections, and our algorithm showed better and more consistent behavior.

## REFERENCES

[1] J. Li, Z. Wang, and P. Stoica, *Robust Adaptive Beamforming*, 2005, ch. 3, pp. 91–200.

[2] S. Theodoridis and R. Chellappa, *Academic Press Library in Signal Processing, Volume 3: Array and Statistical Signal Processing*, 1st ed. Orlando, FL, USA: Academic Press, Inc., 2013.

[3] P. Stoica, J. Li, and H. He, "Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram," *Trans. Sig. Proc.*, vol. 57, no. 3, p. 843858, Mar. 2009.

[4] Qing Chen and Ruolun Liu, "On the explanation of spatial smoothing in music algorithm for coherent sources," in *International Conference on Information Science and Technology*, March 2011, pp. 699–702.

[5] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 3, pp. 273–284, March 1991.

[6] T. Svantesson, "Modeling and estimation of mutual coupling in a uniform linear array of dipoles," in *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99*, vol. 5, March 1999, pp. 2961–2964.

[7] M. Lin and L. Yang, "Blind calibration and DOA estimation with uniform circular arrays in the presence of mutual coupling," *IEEE Antennas and Wireless Propagation Letters*, vol. 5, pp. 315–318, 2006.

[8] Z. Liu, C. Zhang, and P. S. Yu, "Direction-of-arrival estimation based on deep neural networks with robustness to array imperfections," *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 12, pp. 7315–7327, Dec 2018.

[9] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85 – 117, 2015.

[10] U. Pillai and B. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Transactions on Signal Processing*, vol. 37, no. 1, pp. 8–15, 1 1989.