Optimization Of Antenna Array Configuration With A Neural Network

Oded Bialer and Noa Garnett
General Motors - Advanced Technical Center Israel

***Abstract*—The accuracy of angle of arrival (AOA) estimation with an antenna array, depends on the antenna elements positions. In this paper, we introduce a novel method for optimizing the antenna elements position that minimizes the AOAs estimation error in the case of an unknown number of sources and a single array realization (snapshot). The method utilizes a deep neural network (DNN) for estimating the number of sources and their AOAs for a given antenna elements positions, and minimizes the estimation error by jointly optimizing the antenna positions and the DNN parameters. The use of the DNN estimator in this case, enables to calculate the gradient of the estimation error with respect to the antenna elements position, and thus to minimize the estimation error with respect to the antenna positions by gradient descent. The proposed approach is unique because it determines the antenna array configuration that explicitly minimize the AOAs estimation error, while other reference methods use an optimization objective that is implicitly related to the AOA estimation error. We show that the proposed optimization method attains a significant performance advantage in the RMSE of the AOAs estimation compared to reference antenna configuration optimization methods.**

I. Introduction

Estimating the angles of arrival (AOAs) of an unknown number of multiple sources transmitting unknown signals with a single array received signal (single snapshot) is a challenging problem with various applications such as localization, radar, sonar, anti-jamming and wireless communication [1]. A deep neural network (DNN) approach for estimating the AOAs in this case [2] has shown to attains superior performance to signal processing methods due to several reasons. First, good model order determination requires multiple realizations (snapshots) [3]. Second, even if the model order is known, the maximum likelihood (ML) estimator is intractable when the number of sources is large, and feasible approximate ML estimators such as alternating projections (AP) [4], IQML [5], EM [6], and Orthogonal Matching Pursuit (OMP) [7] have performance degradation with respect to ML. Third, super resolution methods for AOAs estimation, such as MUSIC [8], ESPRIT [9], and MVDR [10], require multiple realizations to perform well.

The AOAs estimation performance of an antenna array dependents not only on the estimation method but also on the antenna elements positions. A standard antenna array with uniform spacing of half a wavelength between antennas provides unambiguous AOA estimation. However, the angular resolution is proportional to the array aperture, hence with a small number of antenna elements the standard uniform spacing attains a short aperture with relatively low angular resolution.

One approach for optimizing the antenna elements position is based on the array's beam pattern, which is obtained with the a Bartlett beamformer [11]. In this approach, the antenna elements positions are optimized to reduce the beam-pattern maximal energy outside of a specified main-lobe width [12]-[15]. The main-lobe width determines the desired angular resolution, and minimizing the side-lobes energy mitigates angle ambiguity issues as well as improves the discrimination between multiple sources. This criterion is not optimal in terms of AOA estimation, since it is not derived directly from the AOAs estimator, and therefore, does not optimally minimize the estimation error. Furthermore, the main-lobe width is pre-determines, and not part of the optimization, thus the important tradeoff between the main-lobe width and side-lobe levels is not optimized.

Other approaches for optimizing an array configuration are coprime arrays [16]- [17], and nested arrays [18]. In these approaches, the antenna elements have a unique non-uniform spacing, such that the array spans a large aperture (and thus has high resolution) with a large number of degrees-of-freedom for AOA estimation. These array configurations can provide better AOA estimation accuracy than an array with uniform spacing. However, they are not optimized to explicitly minimize the AOA estimation error in the case of multiple sources and a single received array realization (single snapshot).

In this paper, we introduce a novel method for optimizing the antenna elements position that minimizes the AOAs estimation error in the case of an unknown number of sources and a single array realization (snapshot). The method uses a DNN to estimate the number of sources and their AOAs from the received signal, which is a function of the antenna array configuration. The DNN parameters and the antenna elements positions are jointly optimized to minimize the AOAs estimation error, over a set of examples. The use of the DNN enables to calculate the gradient of the AOA estimation error with respect to the antenna's positions via back-propagation. As a result, the major advantage of using a DNN for optimizing the antenna array configuration is that it enables to determine the antenna elements positions that explicitly minimize the AOA estimation error, while in other methods, the optimization objective is implicitly related to the AOA estimation error. The results show that the proposed array configuration optimization method attains better AOA estimation performance than the performance of a DNN AOA estimator with other optimized reference array configurations.

II. Problem Definition

We consider an array of N linearly spaced antenna elements that are receiving signals from multiple point sources at different angles. The received array signal can be expressed by

$$\boldsymbol{y} = \sum_{m=0}^{M-1} \boldsymbol{a}(\theta_m) s_m + \boldsymbol{v}, \quad (1)$$

where $\boldsymbol{v}$ is the noise vector, M is the number of unknown sources, s_m, θ_m, are the unknown complex signal coefficient and the angle of the m-th source, respectively, and

$$\boldsymbol{a}(\theta) = \begin{bmatrix} e^{\frac{j2\pi}{\lambda} x_1 sin(\theta)} \\ : \\ e^{\frac{j2\pi}{\lambda} x_N sin(\theta)} \end{bmatrix}, \quad (2)$$

is a steering vector for an angle of arrival θ, where λ is the wavelength, and x_n is the n^{th} antenna position with respect to the linear array center point.

The number of sources and their AOAs are estimated from a single realization (snapshot) of the received signal, $\boldsymbol{y}$, which is a function of the antenna elements positions, $x_1, .., x_N$. The problem at hand, is to find the antenna elements positions that yield the maximal probability of accurate detection of the number of sources, and the minimal mean square error (MSE) of the AOAs estimation.

III. Antennas Configuration Optimizing

For estimating the number of sources and their AOAs, we use a DNN because of two important reasons. First, the DNN has showed higher probability of accurate number of sources detection, and lower AOAs MSE compared to other reference signal processing methods [2]. Second, and even more importantly, the DNN enables to minimize the estimation error with respect to the antenna positions, by back-propagating the estimation error, via the DNN, to the received signal, as will be further explained in the following.

A block diagram of the antenna configuration optimization method is shown in Fig. 1. A received signal, $\boldsymbol{y}$, is generated per given antenna configuration, random sources signals, s_m, and random AOAs, θ_m. The received signal is fed into a DNN that estimates the number of sources and their AOAs. The estimation error per each received signal is calculated (with the use of ground truth of the AOAs), and the antennas positions and the DNN weights are optimized to minimize the estimation error over a large set of received signals.

The neural network architecture that was used to estimate the number of sources and their AOAs is depicted in Fig. 2. The DNN was designed for an array of 16 antennas ($N = 16$), and for estimating up to 4 sources ($M \leq 4$). It is straight forward to modify the same DNN architecture for different number of antennas, and different maximal number of possible sources. The input to the network is the array received signal of $2N$ values, which are the real and imaginary components of the array response, $\boldsymbol{y}$. Following the input layer there

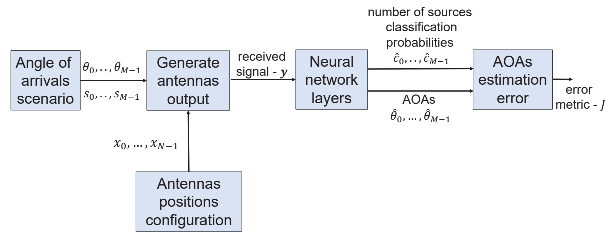

Fig. 1. Block diagram of the antenna elements positions optimization method.

are 8 fully connected (FC) layers, each followed by batch normalization [19] and ReLU nonlinearity. The number of neurons in each layer is indicated next to each FC layer in Fig. 2. Following these FC layers the network splits into five different FC paths. The upper FC path that is shown in the figure, outputs the classification of the number of sources. It has a Softmax [20] layer at the end that outputs four probabilities, $\hat{c}_0, \hat{c}_1, \hat{c}_2, \hat{c}_3$, which are the probabilities of each one of the four classes. Each class corresponds to a different number of sources, between 1-4. The other four paths are the AOAs estimations for each of the four classification options. The estimated number of sources, $\hat{M}$, is the index of the maximal Softmax probability, and the final estimated AOAs are the set of AOAs that correspond to the estimated number of arrivals.

The estimation error metric, that was used for the optimization, combines the classification and the AOAs estimation error, and is given by

$$J(\boldsymbol{w}, \boldsymbol{x}) = \sum_{q=0}^{Q-1} L_c(\hat{c}_0^q, .., \hat{c}_{M-1}^q) + \beta L_\theta(\hat{\theta}_0^q, .., \hat{\theta}_{\hat{M}-1}^q), \quad (3)$$

where Q is the number of examples (scenarios), q is the index of the example, $\boldsymbol{w}$ is a vector of all the network weights, $\boldsymbol{x} = [x_1, .., x_N]^T$ is a vector of the antenna positions, β is a weighting factor, L_c is the cross-entropy classification error term given by

$$L_c(\hat{c}_0^q, .., \hat{c}_{M-1}^q) = -\sum_{m=0}^{M-1} c_m^q \log(\hat{c}_m^q), \quad (4)$$

where $[c_0^q, .., c_{M-1}^q]$ is the classification bits ground truth one-hot vector[1], and L_θ is the AOAs estimation MSE that is given

[1] One-hot vector has zeros in all elements except the element that corresponds to the correct classification index, which has value 1

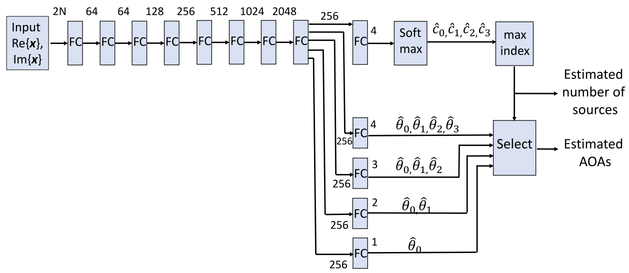

Fig. 2. The neural network architecture with fully connected (FC) layers, followed by Softmax for classification of the number of sources (1-4), and AOAs regression for each one of the classes.

by

$$L_\theta(\hat{\theta}_0^q, .., \hat{\theta}_{\hat{M}-1}^q) = \frac{1}{M+\hat{M}} \left[\sum_{m=0}^{M-1} \min_{i\in[0,..,\hat{M}-1]} \left\{ \left(\hat{\theta}_i^q - \theta_m^q \right)^2 \right\} + \sum_{i=0}^{\hat{M}-1} \min_{m\in[0,..,M-1]} \left\{ \left(\hat{\theta}_i^q - \theta_m^q \right)^2 \right\} \right], \quad (5)$$

where $\hat{\theta}_k^q$, and θ_k^q, are the estimated and ground truth AOAs, respectively, of the k^{th} source index and q^{th} example index. The metric in (5) is a bi-directional MSE calculation, that was chosen to account for the issue of correspondence between the multiple estimated AOAs and the multiple true AOAs. The first summation in the bi-directional MSE is the squared error between each true AOA and its closest estimated AOA, and the second summation in the MSE is the squared error between each estimated AOA and its closest true AOA.

The optimization objective can be expressed by

$$\underset{\boldsymbol{x},\boldsymbol{w}}{\text{argmin}} J(\boldsymbol{w}, \boldsymbol{x}). \quad (6)$$

We solve the optimization in (6) with stochastic gradient descent. In each stochastic gradient descent iteration, a batch of Q new random scenarios of AOAs are produced[2], and Q corresponding received array realizations, $\boldsymbol{y}$, are generated according to (1), with the antennas configuration of the previous iteration. The DNN outputs an estimate of the number of sources, $\hat{M}$, and their AOAs, $\hat{\theta}_0, .., \hat{\theta}_{M-1}$, for each received signal, and the estimation error in (3) is calculated. Then, the DNN weights and the array elements positions are updated as follows

$$\boldsymbol{w}_k = \boldsymbol{w}_{k-1} - \mu_w \frac{\partial J(\boldsymbol{w}, \boldsymbol{x})}{\partial \boldsymbol{w}} \bigg|_{(\boldsymbol{w},\boldsymbol{x})=(\boldsymbol{w}_{k-1},\boldsymbol{x}_{k-1})}, \quad (7)$$

and

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} - \mu_x \frac{\partial J(\boldsymbol{w}, \boldsymbol{x})}{\partial \boldsymbol{x}} \bigg|_{(\boldsymbol{w},\boldsymbol{x})=(\boldsymbol{w}_{k-1},\boldsymbol{x}_{k-1})}, \quad (8)$$

[2]Each scenario has random number of sources, AOAs and signals

where k is the stochastic gradient iteration index, $\boldsymbol{w}_k$, $\boldsymbol{x}_k$ are the network parameters and antennas positions at the k^{th} iteration, respectively, and μ_w, μ_x are their corresponding update step sizes. We obtain the derivative of the cost function $J(\boldsymbol{w}, \boldsymbol{x})$ with respect to $\boldsymbol{x}$, using the chain rule, as follows

$$\frac{\partial J(\boldsymbol{w}, \boldsymbol{x})}{\partial \boldsymbol{x}} = \frac{\partial J(\boldsymbol{w}, \boldsymbol{x})}{\partial \boldsymbol{y}} \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}}, \quad (9)$$

where $\frac{\partial J(\boldsymbol{w},\boldsymbol{x})}{\partial \boldsymbol{y}}$ can be computed by applying the chain rule on the network layers, i.e., applying back propagation [21], and

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \text{diag}\left\{ \frac{\partial y_1}{\partial x_1}, \frac{\partial y_2}{\partial x_2}, .., \frac{\partial y_N}{\partial x_N} \right\}, \quad (10)$$

is a diagonal matrix, with diagonal elements that are given by

$$\frac{\partial y_i}{\partial x_i} = \frac{j2\pi}{\lambda} \sum_{m=0}^{M-1} sin(\theta_m) e^{\frac{j2\pi}{\lambda} x_i sin(\theta_m)} s_m. \quad (11)$$

An intuitive explanation for the iterative optimization method which is described above, is that the antennas positions are updated to reduce the estimation error, and in return, the DNN parameters are updated, to match the estimator to the updated antennas positions. It is important to realize that the use of the neural network as the estimator, enables to obtain the derivative of the cost function, J, with respect to the antenna positions, $\boldsymbol{x}$, which is essential for the optimization. Using other estimators, such as super resolution methods (e.g. MUSIC), maximum likelihood estimator, sparsity methods, etc., would not provide a tractable solution for (8).

The minimization function in (6) is not convex. Although neural networks are known to reach a good solution even for non-convex optimization functions, still, the minimization with respect to the antenna positions may reach a local minimum point. To mitigate this problem, the optimization method was initiated with 100 different antennas' positions configurations that had random spacings. An individual independent optimization was carried out for each pair of initial random antenna configuration and initial random network parameters (weights). This resulted in 100 optimized pairs of antenna configuration and network parameter, from which the pair that attained the lowest metric in (3), was selected as the final optimization result. After the optimization was complete, the selected optimal pair was fixed ("frozen") and used for estimating the number of sources and AOAs during inference time.

IV. Results and Discussion

We trained and tested the proposed antennas' positions optimizations method with simulated realizations according to (1), with 16 antenna elements ($N = 16$). The number of sources per each example was randomly selected between 1-4, with uniform distribution. The complex signals, s_m, in each example had a phase with uniform distribution in the range $(0, 2\pi)$, and amplitude with uniform distribution in the range $(0.5, 1.5)$. We tested three different distribution for the AOAs of the sources, which are detailed below, and optimized the array configuration for each one of the distribution.

The performance of the proposed array configuration optimization method is compared to four reference antenna configurations. The first, was a uniform linear array (ULA) with maximal spacing between antennas that provides unambiguous AOA estimation. The maximal spacing depends on the maximal AOA in the test scenario. The second, was a coprime array [16]- [17], that was composed of two ULAs of 8 elements, that were shifted by 4λ, one with respect to the other. One ULA had 4λ spacing and the other had 4.5λ spacing. The third, was a nested array [18], that had an 8 elements ULA with 0.5λ spacing, followed by another 8 elements ULA with 4λ spacing. As mentioned in Section I, coprime and nested arrays are known to achieve good AOA estimation performance since they span a large aperture with non-uniform spacing that has a relatively large number of degrees-of-freedom for AOA estimation. The fourth reference method was the array configuration optimization method proposed in [12]. In this reference method, a particle swarm optimization algorithm was applied for determining the antenna array configuration that minimizes the maximal side-lobe level in the Bartlett beamforming spectrum, for seven different target angles that uniformly span the sources AOA range in each test scenario. The side-lobe region is the beamforming output that does not include the main-lobe beam-width.

In the figures below, the proposed optimization method in Section III is referred to as 'optimized DNN', the optimization based on side-lobe level minimization is referred to as 'optimized side-lobes', the uniform linear array is referred to as 'ULA', co-prime array as 'co-prime', and nested array as 'nested'.

Each one of the four reference antenna configurations had a separate DNN estimator for the number of sources and their AOAs, which was trained and optimized for the individual antenna configuration. The same DNN architecture that is shown in Fig. 2, was used for the estimator of all reference antenna configurations. However, individual network parameters were trained and optimized for each different antenna configuration. For each antenna configuration, the network was trained with one million batches of size 1000. The examples in each batch had random signal to noise ratio (SNR) in the range 20-40 dB. The performance were tested with 100000 examples, which were not part of the training set.

Figs. 3-4 show the AOAs estimation RMSE (the square root of (5)) and the accuracy in the detection of the number of sources, respectively, of all the tested methods for the case that the sources AOAs (between one to four sources per example) had uniform distribution in the range of $(-85°, 85°)$. It is seen that optimizing the array with a DNN for this scenario achieves a significant performance advantage in the AOAs RMSE with respect to the reference antenna configurations. As for the accuracy in the number of sources estimation, the optimized array achieves modest performance improvement with respect to the reference methods.

Fig. 5 presents the RMSE performance of all the tested

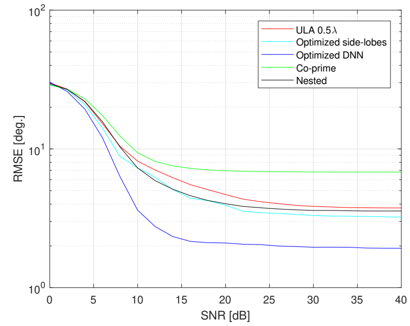

Fig. 3. AOAs estimation RMSE performance for AOAs uniformly distributed in $(-85°, 85°)$. Number of sources per example was random between one to four.

methods for the case that the sources AOAs had a uniform distribution in the range $(-50°, 50°)$, which is a narrower AOA range than in Figs. 3-4. In this case, the uniform linear array spacing was increased to 1.1 times the wavelength, which provided maximal aperture with unambiguous AOA estimation since the maximal AOA was $50°$. The ULA with relatively wide spacing outperformed the nested and coprime arrays. The array optimization method that minimizes the maximal side-lobe level achieved slightly better performance than the ULA, however, the optimized DNN method outperforms all the reference methods.

Fig. 6, shows the RMSE results for the case that the AOAs were uniformly distributed within a cluster of $10°$, and the cluster center was uniformly distributed in the range $(-80°, 80°)$. There are various applications in which the AOAs are clustered, such as local scatterer [22], collocated sources or distributed target [23]. For this scenario, the coprime array achieves best performance from all the reference methods, yet the proposed DNN optimization method outperforms it.

In summary, the results in Figs. 3-6, show that the proposed method achieves lower AOA estimation error compared to all the reference method for various AOAs distributions. The performance advantage is because the proposed method optimizes the antenna configuration to explicitly minimize the AOAs estimation error per each specific distribution of AOAs. Unlike the proposed method, the reference methods are designed to minimize side-lobes or to have a large number of degrees of freedom for AOA estimation. These optimization objectives are related implicitly to the AOA estimation error and to the AOA distribution.

V. Conclusions

A novel method for optimizing the antenna elements positions that minimize the AOAs estimation error, was introduces. The method utilizes a DNN for estimating the number of sources and the sources AOAs for a given antenna elements positions. The use of the DNN in this case enables to calculate the gradient of the estimation error with respect to the antenna elements position, and thus to minimize the estimation error

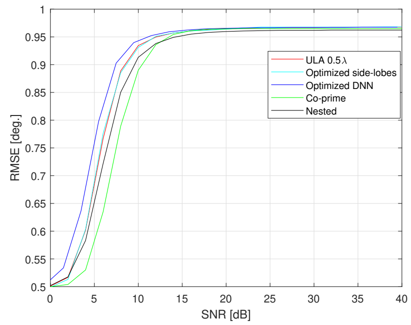

Fig. 4. Accuracy of the number of sources estimation for AOAs that are uniformly distributed $(-85°, 85°)$. Number of sources per example was random between one to four.

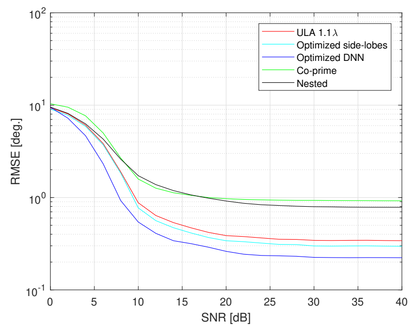

Fig. 5. AOAs estimation RMSE performance for AOAs with uniform distribution in $(-50°, 50°)$. Number of sources per example was random between one to four.

with respect to the antenna positions by gradient descent. The antenna elements positions are updated iteratively to minimize the estimation error, and in return, the DNN adapts the estimator to match to the updated antenna positions. The proposed approach is unique because it optimizes the antenna configuration to explicit minimize the estimation error, while other reference methods use an optimization objective that is only implicitly related to the AOA estimation error. The optimized array showed a significant performance advantage in the RMSE of the AOAs estimation, compared to reference antenna configurations in various scenarios of AOAs distribution.

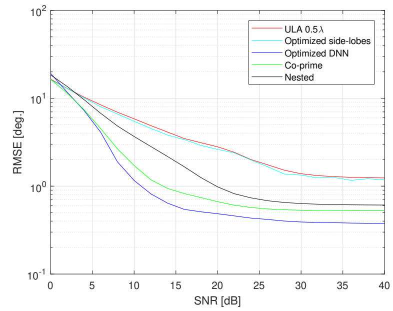

Fig. 6. AOAs estimation RMSE performance for AOAs that are clustered within 10°, and the cluster center is uniformly distributed in $(-80°, 80°)$. Number of sources per example was random between one to four.

REFERENCES

[1] K. Hamid, and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67-94, 1996.

[2] O. Bialer, N. Garnett and T. Tirer, "Performance Advantages of Deep Neural Networks for Angle of Arrival Estimation," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, 2019, pp. 3907-3911.

[3] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. On Acoustics, Speech and Signal Processing*, vol. 33, no. 2, pp. 387-392, 1985.

[4] I. Ziskind, and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 36, pp. 1553-1560, Oct. 1988.

[5] Y. Bresler, and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, pp. 1081-1089, Oct. 1986 .

[6] Feder, M. and Weinstein, E., "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Transactions on acoustics, speech, and signal processing*, Vol. 36, pp. 477-489, Apr. 1988.

[7] T. T. Cai, L. Wang, "Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise," *IEEE Trans. on information theory*, vol. 57, no. 7, July 2011.

[8] R.O Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans. Antennas Propagation*, vol. AP-34, pp.276-280, March 1986.

[9] R. Roy, A. Paulraj and T. Kailath, "Estimation of Signal Parameters via Rotational Invariance Techniques-ESPRIT", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984-995, Jul. 1989.

[10] J. Capon, "High Resolution Frequency-Wavenumber Spectrum Analysis," *Proc. of the IEEE*, vol. 57, pp. 1408-1418, Aug. 1969.

[11] H.L. Van Trees, Optimum Array Processing, Wiley and Sons, 2002.

[12] M. M. Khodier and C. G. Christodoulou, "Linear array geometry synthesis with minimum sidelobe level and null control using particle swarm optimization," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 8, pp. 2674-2679, Aug. 2005.

[13] P. J. Bevelacqua and C. A. Balanis, "Minimum Sidelobe Levels for Linear Arrays," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 12, pp. 3442-3449, Dec. 2007.

[14] A. A. Minasian and T. S. Bird, "Particle Swarm Optimization of Microstrip Antennas for Wireless Communication Systems," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 12, pp. 6214-6217, Dec. 2013.

[15] P. J. Bevelacqua and C. A. Balanis, "Geometry and Weight Optimization for Minimizing Sidelobes in Wideband Planar Arrays," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 4, pp. 1285-1289, April 2009.

[16] 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, Feb. 2011.

[17] 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, March15, 2015.

[18] 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, Aug. 2010.

[19] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *International Conference on Machine Learning*, pp. 448-456. 2015.

[20] I. Goodfellow, Y. Bengio, A. Courville, "Deep learning", The MIT Press, 2016.

[21] D. E. Rumelhart, G. E. Hinton and R. J. Williams, "Learning representations by back-propagating errors", *Nature*, vol. 323, pp. 533-536, 1986

[22] D. Astely and B. Ottersten, "The effects of local scattering on direction of arrival estimation with MUSIC," *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3220-3234, Dec. 1999.

[23] R. Raich, J. Goldberg and H. Messer, "Bearing estimation for a distributed source: modeling, inherent accuracy limitations and algorithms," *IEEE Transactions on Signal Processing*, vol. 48, no. 2, pp. 429-441, Feb. 2000.