

Article

Underdetermined DOA Estimation Using MVDR-Weighted LASSO

Amgad A. Salama, M. Omair Ahmad * and M. N. S. Swamy

Department of Electrical and Computer Engineering, Concordia University, Montreal, PQ H3G 1M8, Canada; am_adeli@encs.concordia.ca (A.A.S.); swamy@ece.concordia.ca (M.N.S.S.)

* Correspondence: omair@ece.concordia.ca; Tel.: +1-514-848-2424 (ext. 3075)

Academic Editor: Wilmar Hernandez

Received: 10 June 2016; Accepted: 13 September 2016; Published: 21 September 2016

Abstract: The direction of arrival (DOA) estimation problem is formulated in a compressive sensing (CS) framework, and an extended array aperture is presented to increase the number of degrees of freedom of the array. The ordinary least square adaptable least absolute shrinkage and selection operator (OLS A-LASSO) is applied for the first time for DOA estimation. Furthermore, a new LASSO algorithm, the minimum variance distortionless response (MVDR) A-LASSO, which solves the DOA problem in the CS framework, is presented. The proposed algorithm does not depend on the singular value decomposition nor on the orthogonality of the signal and the noise subspaces. Hence, the DOA estimation can be done without a priori knowledge of the number of sources. The proposed algorithm can estimate up to $((M^2 - 2)/2 + M - 1)/2$ sources using M sensors without any constraints or assumptions about the nature of the signal sources. Furthermore, the proposed algorithm exhibits performance that is superior compared to that of the classical DOA estimation methods, especially for low signal to noise ratios (SNR), spatially-closed sources and coherent scenarios.

Keywords: adaptable LASSO; sparse array; direction of arrival estimation; compressive sensing; sensor array processing

1. Introduction

Array signal processing plays an important role in many applications, including sonar, radar, seismology and radio astronomy. Direction of arrival (DOA) deals with the problem of determining the number and locations of multiple sources using an antenna array. Many of the recent algorithms deal with DOA estimation of spatiotemporal electromagnetic waves emanating from multiple sources. Beamforming can be considered as a very early attempt at DOA estimation. This is based on mechanically steering an array to scan all possible angles in a sector of interest followed by measuring the output power of the array. The power spectrum will have a peak when the angle given by the mechanical steering corresponds to the DOA of one of the incoming source signals [1,2]. However, classical beamforming techniques fail to identify closely-separated source signals. Capon's beamformer can overcome this problem; it maximizes the signal to noise ratio by using all of the available degrees of freedom to form a beam towards the interested signal source and at the same time makes null all of the other directions [3]. The linear prediction method minimizes the sensor array mean output power assuming that the weight on a selected sensor is unity [4]. Use of the maximum likelihood (ML) technique has been investigated for DOA estimation in [5]. ML leads to highly accurate DOA estimation results; however, ML methods are the most time consuming and complex DOA estimation algorithms among all of the DOA estimation techniques.

MUSIC [6] and ESPRIT [7] are two of the most studied algorithms in DOA estimation. The MUSIC algorithm is based on the eigen decomposition of the covariance matrix of the received data. Unlike conventional DOA techniques, it is easy to implement MUSIC with high resolution. However,

it requires the number of sources to be known in advance. If the exact number of sources is unknown, this will lead to a failure of the DOA estimation.

ESPRIT essentially investigates the shift-invariance properties of the sensor array and is based on the assumption that the array is composed of two identical subarrays. The DOA can be estimated by studying the shift-invariance properties of the output of each sub-array. Unfortunately, like the MUSIC algorithm, ESPRIT also uses the eigenvalue decomposition and, hence, requires the number of sources to be known in advance [8,9].

Increasing the number of sensor array elements leads to an enhancement in the array gain and directivity, thus improving the array's overall performance. On the other hand, increasing the number of array elements increases the complexity of the feeding network and the time needed to process the received data. Researchers have focused recently on the pursuit of light weight, small sized and compact antenna arrays to satisfy the requirements of present day applications like the Internet of Things (IoT) and unmanned vehicles [10]. Physically increasing the number of antenna array elements is therefore not an option. Recent literature investigating the virtual array concept has shown it to enhance the performance of a sensor array [11–15].

Minimum redundancy arrays (MRAs) and augmented covariance matrices (ACMs) have been investigated for virtually extending the sensor array aperture by increasing the available degrees of freedom [16–18]. However, there is no closed form expression for the locations of the sensors in the virtual array and for the number of achievable degrees of freedom. Furthermore, ACM is not positive semi-definite for a finite number of snapshots, and a transformation was suggested in [19] to overcome this problem. Higher order cumulants were used to increase the available degrees of freedom and consequently extend the array aperture. However, using the fourth order cumulants, Gaussian source signals cannot be estimated. The Khatri-Rao (KR) product was suggested for the DOA estimation of quasi-stationary source signals in [12,13].

Co-prime arrays have been proposed in the literature [20–23] to increase the degrees of freedom in the array. For these arrays, the number of the degrees of freedom can be determined, and closed-form expressions for the locations of the sensors in the corresponding virtual arrays can be obtained. However, the corresponding virtual arrays are not uniform. Nested arrays have been investigated for the first time in [14]. These arrays, in addition to all of the advantages provided by the co-prime arrays, result in virtual arrays that are uniform and provide degrees of freedom that are even higher [23]. For example, using six elements, the nested array provides 23 degrees of freedom in contrast to only 17 provided by the co-prime array.

Compressive sensing (CS) has evolved rapidly in recent years and has found multiple applications in various fields, such as medicine [24], ultrasound imaging [25] and radar detection [26]. Malioutov et al. [27] investigated the DOA estimation performance using CS with respect to the signal to noise ratio, the number of sources and the coherence of the sources' signals. Xenaki et al. [28] studied the DOA estimation using CS with coherent arrivals, single-snapshot data and different array geometries. It has been proven in [28] that CS does not require the arrivals to be incoherent. Furthermore, single or multiple snapshots can be used. In terms of resolution, the performance of the CS-based DOA is superior to that of the minimum variance distortionless response (MVDR) [29].

The least square (LS) method is based on minimizing the residual squared error. LS estimates have low bias and large variance; therefore, LS estimates suffer from low prediction accuracy, which can be enhanced through setting some coefficients to zero. Furthermore, LS estimates can be improved by subset selection and ridge regression; however, both of these have some drawbacks. Subset selection can provide models that are interpretable. However, small changes in the data can lead to very different models being selected. Ridge regression is more stable, but is not an interpretable model [30].

The least absolute shrinkage and selection operator (LASSO) [30] minimizes the residual sum of squares (subject to the sum of the absolute values of the coefficients being less than a constant), and it is popular for solving CS problems. The LASSO technique combines the advantages of both subset selection and ridge regression. Zou [31] proposed a new version of LASSO, whereby adaptable

weights are used for penalizing different coefficients in the ℓ_1 penalty function. Panahi et al. [32] discussed the resolution of the LASSO-based DOA estimation; it has been noted that the LASSO-based DOA estimation is better than that of the traditional beamforming techniques. It should be pointed out that the above-mentioned techniques use uniform arrays, and some of them require the number of signal sources to be known in advance, especially when singular value decomposition (SVD) of the covariance matrix of the received data is performed [27].

In this paper, we propose a new algorithm for DOA estimation, the MVDR A-LASSO, which combines the benefits of the virtual array concept for extended array aperture [12,14] along with CS. In Section 2, we introduce the co-array principle and apply it to a sparse array. The CS framework is presented in Section 3. In Section 4, a modified version of the adaptable LASSO, MVDR A-LASSO, is proposed in order to further improve the performance. In Section 5, the performance of the proposed MVDR A-LASSO is studied using simulations, and finally, conclusions are drawn in Section 6.

Notations

Superscript H denotes conjugate transpose; superscript $*$ denotes conjugation without transpose; and T denotes the transpose operation. The symbol \odot denotes the Khatri-Rao (KR) product [33] between two matrices of appropriate sizes.

2. Difference Co-Array

Consider a linear array (LA), uniform or non-uniform, consisting of M elements. Let d_i denote the i -th element position in the array. Let us assume that there are L narrowband, far-field sources with angles-of-arrival (AOA) (θ_l) and powers (σ_l^2), $l = 1, 2, \dots, L$. It is also assumed that the source signals are uncorrelated with one another. Let $\mathbf{a}(\theta_l) \in \mathbb{C}^{M \times 1}$ be the steering vector corresponding to AOA (θ_l), whose i -th element is $e^{-jk_0 d_i \cos(\theta_l)}$, where $k_0 = 2\pi/\lambda$ is the wavenumber and λ is the wavelength of the propagating waves. Let the vector $\mathbf{s}(t) = [s_1(t) \ s_2(t) \ \dots \ s_L(t)]^T$, where $\mathbf{s} \in \mathbb{C}^{L \times 1}$ represent the source signals. Then, the output of LA can be written as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{A} = [\mathbf{a}(\theta_1) \ \mathbf{a}(\theta_2) \ \dots \ \mathbf{a}(\theta_L)]$, $\mathbf{A} \in \mathbb{C}^{M \times L}$ is the array manifold matrix and $\mathbf{n}(t) \in \mathbb{C}^{M \times 1}$ is an additive white Gaussian noise (AWGN) that is uncorrelated with the source signals. One can obtain the covariance matrix of the received signals as [34]:

$$\begin{aligned} \mathbf{R}_{xx} &= E[\mathbf{x}\mathbf{x}^H] \\ &= \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H + \sigma_n^2\mathbf{I} \\ &= \mathbf{A} \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_l^2 & & \\ & & & \ddots & \\ & & & & \sigma_L^2 \end{bmatrix} \mathbf{A}^H + \sigma_n^2\mathbf{I} \end{aligned} \quad (2)$$

where σ_l^2 , $l = 1, \dots, L$, correspond to the power of the source signals, \mathbf{I} is the identity matrix of size $(M \times M)$ and σ_n^2 is the noise power. One can now vectorize $\mathbf{R}_{xx} \in \mathbb{C}^{M \times M}$ as [12,13]:

$$\begin{aligned} \mathbf{V} &= \text{vec}(\mathbf{R}_{xx}) = \text{vec} \left[\sum_{l=1}^L \sigma_l^2 (\mathbf{a}(\theta_l)\mathbf{a}^H(\theta_l)) \right] + \sigma_n^2\mathbf{1} \\ &= (\mathbf{A}^* \odot \mathbf{A})\mathbf{p} + \sigma_n^2\mathbf{1} \end{aligned} \quad (3)$$

where $\mathbf{p} \in \mathbb{C}^{L \times 1} = [\sigma_1^2 \sigma_2^2 \dots \sigma_L^2]^T$ and $\mathbf{1} \in \mathbb{C}^{M \times M} = [\mathbf{e}_1^T \mathbf{e}_2^T \dots \mathbf{e}_M^T]^T$ with $\mathbf{e}_i \in \mathbb{C}^{M \times 1}$ being a column vector of zeros except for a one at the i -th position. Comparing Equations (3) and (1), we can see that $\mathbf{V} \in \mathbb{C}^{M^2 \times 1}$ in Equation (3) can be considered as the output of an array with a manifold $(\mathbf{A}^* \odot \mathbf{A})$, \mathbf{p} representing the equivalent source signals and the noise given by $\sigma_n^2 \mathbf{1}$ being deterministic. The distinct rows of $(\mathbf{A}^* \odot \mathbf{A})$ form the virtual array (VA); the locations of whose distinct elements are given by the set:

$$D = \{d_i - d_j\}, \quad \forall i, j = 1, 2, \dots, M \tag{4}$$

where d_i is the position vector of the i -th sensor in the original array. This array is known as the difference co-array [14].

We now assume that the original array is a two-level nested array [14] for which M is even, and each level contains $M/2$ sensor elements. In such a case, the VA is a uniform linear array (ULA) consisting of \bar{M} sensors, $\bar{M} = (M^2/2 + M - 1)$, which are located from $-(\bar{M} - 1)d/2$ to $(\bar{M} - 1)d/2$, where d is the distance between two adjacent sensors [14]. It should be noted that the equivalent source signal vector \mathbf{p} (for the difference co-array) contains the power of the sources $\sigma_l^2, l = 1, \dots, L$. Therefore, they act like fully-correlated sources. A spatial smoothing technique was suggested by Pal et al. [14] to overcome this problem of correlated sources. However, by using the CS technique, we will no longer need spatial smoothing or any preprocessing scheme, and CS will be able to detect the source signals.

3. Compressive Sensing Framework

Since the sources are assumed to be located in the far field, they can be considered as point sources; hence, the sources become sparse in space. Let Ω denote the set of all possible source locations, $\{\bar{\theta}_n\}_{n=1}^N$, N denoting a grid that covers Ω , with $N \gg L$. Let:

$$\bar{\mathbf{s}}(t) = [\bar{\sigma}_1 \bar{\sigma}_2 \dots \bar{\sigma}_n \dots \bar{\sigma}_N]^T \tag{5}$$

where $\bar{\mathbf{s}} \in \mathbb{C}^{N \times 1}$, and:

$$\Phi = [\bar{\mathbf{a}}(\bar{\theta}_1) \bar{\mathbf{a}}(\bar{\theta}_2) \dots \bar{\mathbf{a}}(\bar{\theta}_n) \dots \bar{\mathbf{a}}(\bar{\theta}_N)]$$

$$= \begin{bmatrix} e^{jk_0 d (-(\bar{M}-1)/2) \cos \bar{\theta}_1} & e^{jk_0 d (-(\bar{M}-1)/2) \cos \bar{\theta}_2} & \dots & e^{jk_0 d (-(\bar{M}-1)/2) \cos \bar{\theta}_N} \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \ddots & 1 \\ \vdots & \vdots & & \vdots \\ e^{jk_0 d ((\bar{M}-1)/2) \cos \bar{\theta}_1} & e^{jk_0 d ((\bar{M}-1)/2) \cos \bar{\theta}_2} & \dots & e^{jk_0 d ((\bar{M}-1)/2) \cos \bar{\theta}_N} \end{bmatrix} \tag{6}$$

where $\Phi \in \mathbb{C}^{\bar{M} \times N}$ and $\bar{\mathbf{a}}(\bar{\theta}_n) \in \mathbb{C}^{\bar{M} \times 1}$ is the steering vector of the VA corresponding to the AOA ($\bar{\theta}_n$). Then, the received signal at the \bar{m} -th sensor is:

$$y_{\bar{m}}(t) = \phi_{\bar{m}} \bar{\mathbf{s}}(t) + \bar{n}_{\bar{m}}(t), \quad \bar{m} = 1, 2, \dots, \bar{M} \tag{7}$$

where $\phi_{\bar{m}}$ is the \bar{m} -th row of Φ . The n -th element of $\bar{\mathbf{s}}(t)$, $\bar{s}_n(t)$, is nonzero only if ($\bar{\theta}_n = \theta_l$), and in that case, $\bar{\sigma}_n = \sigma_l$. Then, Equation (7) can be rewritten as:

$$\mathbf{y}(t) = \Phi \bar{\mathbf{s}}(t) + \bar{\mathbf{n}}(t) \tag{8}$$

where $\mathbf{y} \in \mathbb{C}^{\bar{M} \times 1}$ and $\bar{\mathbf{n}} \in \mathbb{C}^{\bar{M} \times 1}$. In accordance with conventional DOA estimation, the technique is to estimate the signal energy as a function of the source location showing peaks corresponding to the source locations. Since the sources are point sources and their number is small, the spatial spectrum

is sparse. Hence, we can solve this problem by regularizing it to favour sparse signal fields using LASSO [30]. The LASSO minimization is defined as:

$$\min_s \left\| \mathbf{y} - \sum_{n=1}^N \phi_n \bar{s}_n \right\|_2^2 + \tau \sum_{n=1}^N |\bar{s}_n| \quad (9)$$

where ϕ_n is the n -th element of $\phi_{\bar{m}}$ and \bar{s}_n is the n -th element of $\bar{\mathbf{s}}$. Equation (9) can be rewritten as:

$$\hat{\mathbf{s}}_{lasso} = \min_s \|\mathbf{y} - \Phi \bar{\mathbf{s}}\|_2^2 + \tau \|\bar{\mathbf{s}}\|_1 \quad (10)$$

where τ is a nonnegative regularization parameter. The first term in Equation (9) is the ℓ_2 norm, while the second is an ℓ_1 penalty function, which is very important for the success of LASSO. LASSO shrinks the coefficients toward zero, as the regularization parameter τ increases. This parameter, τ , controls the relative importance between the sparsity of the solution (ℓ_1 -norm term) and the fitness to the measurements (ℓ_2 -norm term). However, the ℓ_1 -norm penalty associated with LASSO tends to produce biased estimates for large coefficients [35], thus degrading the estimation accuracy. Zou [31] proposed a new version of LASSO, the adaptable LASSO (A-LASSO), wherein adaptable weights are used for penalizing the coefficients in the ℓ_1 -norm term iteratively. Furthermore, Zou [31] suggests using the ordinary least squares (OLS) solution as the initial weights to construct the adaptable weights in the adaptable LASSO first iteration. We shall refer to this as OLS A-LASSO. It should be mentioned that the ℓ_1 penalization approach is also known as basis pursuit [36].

4. Modified LASSO for DOA Estimation

One can notice from Equation (9) that the regularization parameter, τ , penalizes the coefficients equally in the ℓ_1 -norm term. Therefore, the LASSO estimates could be biased [35] and result in reducing the solution accuracy. In order to overcome this deficiency, we apply the A-LASSO in the DOA estimation problem for the first time. Hence, the A-LASSO minimizes:

$$\|\mathbf{y} - \Phi \bar{\mathbf{s}}\|_2^2 + \tau \sum_{n=1}^N w_n |\bar{s}_n| \quad (11)$$

where w_n is the n -th element of the weight vector, $\mathbf{w} \in \mathbb{C}^{N \times 1}$. Let, $\hat{\mathbf{s}}$ be the initial estimate for $\bar{\mathbf{s}}$. Now, choosing any weight factor, γ , where $\gamma > 0$, and defining the weight vector as $\hat{\mathbf{w}} = [\hat{w}_1 \hat{w}_2 \dots \hat{w}_N]^T$, where:

$$\hat{w}_n = \frac{1}{|\hat{s}_n|^\gamma} \quad n = 1, 2, \dots, N \quad (12)$$

the A-LASSO is given by:

$$\hat{\mathbf{s}}^{(k)} = \min_s \|\mathbf{y} - \Phi \bar{\mathbf{s}}\|_2^2 + \tau_k \sum_{n=1}^N \hat{w}_n |\bar{s}_n| \quad (13)$$

where k is the iteration number and \hat{w}_n is the n -th element of the weight vector, $\hat{\mathbf{w}}$. The minimization in Equation (13) corresponds to a convex optimization problem; it does not have multiple local minima, and its global minimizer can easily be found. The A-LASSO is ℓ_1 penalized, so any efficient algorithm that can solve the conventional LASSO should also be able to solve the adaptable version. The least angle regression (LARS) algorithm [37] is utilized to solve the A-LASSO using the following steps:

1. Let the initial estimate for $\bar{\mathbf{s}}$ be $\hat{\mathbf{s}}$.
2. Find $\hat{\mathbf{w}}$, where the n -th element of $\hat{\mathbf{w}}$, \hat{w}_n , is given by $\hat{w}_n = \frac{1}{|\hat{s}_n|^\gamma}$, $n = 1, 2, \dots, N$.
3. Define $\bar{M} \times N$ matrix Φ^* , such that its (\bar{m}, n) -th element is given by $\phi_{\bar{m}n} / \hat{w}_n$, where $\bar{m} = 1, 2, \dots, \bar{M}$ and $n = 1, 2, \dots, N$.
4. Solve the LASSO problem as:

$$\hat{\mathbf{s}}^* = \min_{\mathbf{s}} \|\mathbf{y} - \Phi^* \mathbf{s}\|_2^2 + \tau_k \|\mathbf{s}\|_1$$

5. Calculate $\hat{\mathbf{s}}^{(k)} = \hat{\mathbf{s}}_n^*/\hat{w}_n$, $n = 1, 2, \dots, N$.

The steps from 2 to 5 are repeated until convergence to a predefined residual, R , is obtained or when the chosen number of iterations is reached. The computational cost is of the order $O(KN^2)$, where K is the total number of iterations, which is of the same order as the computation of a single OLS minimization. Figure 1 summarizes the above steps from 1 to 5. The efficient path algorithm makes the A-LASSO an attractive method for practical applications [31].

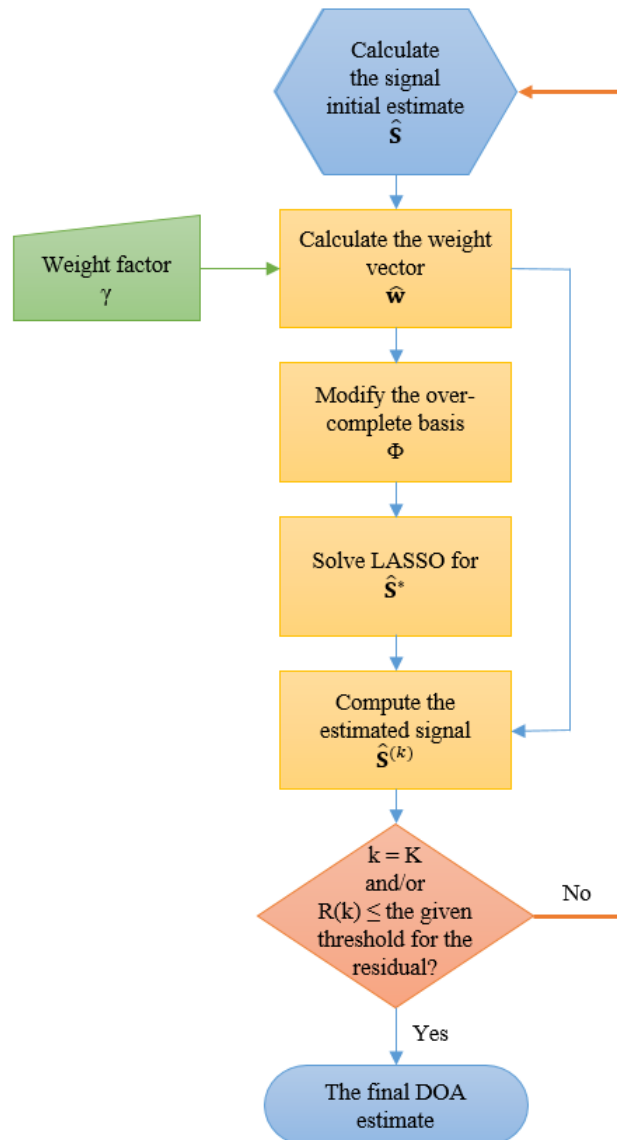


Figure 1. A flowchart of the algorithm for adaptable (A)-LASSO-based DOA estimation.

For the uniqueness of the sparse solution, the spark of matrix Φ , defined as the smallest number of columns from Φ that are linearly dependent [38], must be investigated. Hence, the above algorithm can identify a unique L -sparse solution only if $L < \text{Spark}[\Phi]/2$, $\text{Spark}[\cdot]$ denoting the spark of a matrix [38]. Since any set of $(M^2/2 + M - 1)$ columns of Φ is linearly independent, $\text{Spark}(\Phi) = (M^2/2 + M)$. Hence, the algorithm can identify L -sparse solutions only if $L < \text{Spark}(\Phi)/2$. That is, our algorithm can detect up to $(M^2/2 + M - 2)/2$ sources using an array of M sensors.

4.1. OLS A-LASSO

Previously, a vector of ones is assumed as the initial signal estimate for $\bar{\mathbf{s}}$. However, as proven from the simulations, a vector of ones is not the appropriate guess for the signal to be estimated, especially since there is no relation between the vector of ones and the signal to be estimated. Furthermore, multiplying that vector by a factor, β , will affect the regularized solution (the same effect as that of changing τ , the regularization parameter). Even more, if we try to push the algorithm to the limits (by choosing a small β), we find spurious peaks along with the genuine peaks. Therefore, we will use the OLS solution as the initial signal estimate for $\bar{\mathbf{s}}$, with the expectation that this modification leads to better results and that the OLS A-LASSO solution converges faster than that of the A-LASSO.

We assume replacing \hat{s}_n in Equation (12) with \hat{s}_{OLSn} , where \hat{s}_{OLSn} is the n -th element of $\hat{\mathbf{s}}_{OLS} \in \mathbb{C}^{N \times 1}$ and is given by:

$$\hat{\mathbf{s}}_{OLS} = \min_{\mathbf{s}} \|\mathbf{y} - \Phi \mathbf{s}\|_2^2 \tag{14}$$

The minimization in Equation (14) is known as the ordinary least square minimization. The computational cost for Equation (14) is of order $O(N^2)$. It should be mentioned that the number of source signals is not required to be known in advance for OLS A-LASSO. However, we do not use OLS for DOA estimation. We use it in OLS A-LASSO only as an initial guess for the signal to be estimated (not a stand-alone DOA estimation technique). Furthermore, OLS gives nonzero estimates to all of the coefficients (compared to LASSO minimization) and does not favour sparse signals as in the case of LASSO minimization.

4.2. MVDR A-LASSO

The MVDR technique uses the available degrees of freedom to form a beam in the look direction and, at the same time, nulling the output in all of the other directions. Thus, for a particular DOA, MVDR uses all, but one of the degrees of freedom to minimize the array output while using the remaining ones to constrain a unity gain in the look direction according to the following optimization [2]:

$$\min_{\mathbf{z}} \mathbf{z}^H \mathbf{R}_{ss} \mathbf{z}, \quad \text{subject to} \quad \mathbf{z}^H \mathbf{a}_1(\bar{\theta}) = 1 \tag{15}$$

where $\mathbf{z} \in \mathbb{C}^{(M^2/4+M/2) \times 1}$ is the MVDR beamformer weight vector, \mathbf{R}_{ss} is the spatially-smoothed (SS) covariance matrix, which we will now obtain, and $\mathbf{a}_1 \in \mathbb{C}^{(M^2/4+M/2) \times 1}$ is the steering vector of the array whose SS covariance matrix is \mathbf{R}_{ss} . It should be noted that we are not able to use the covariance matrix \mathbf{R}_{xx} in Equation (2), since this contains information only about the real sensor array. Furthermore, the received source signals are represented as the deterministic vector \mathbf{p} in Equation (3). Therefore, we are not able to use \mathbf{V} of Equation (3) directly for MVDR, since the resultant covariance matrix is rank defective. However, in our case, it is required to construct the covariance matrix of the virtual array. Hence, we perform spatial smoothing on \mathbf{V} to construct a full rank covariance matrix for the virtual array. Assuming a two-level nested array containing M even sensors, the distinct elements of vector \mathbf{V} in Equation (3), $\bar{\mathbf{V}} \in \mathbb{C}^{\bar{M} \times 1}$ can be rewritten as:

$$\bar{\mathbf{V}} = [\bar{v}_1 \ \bar{v}_2 \ \dots \ \bar{v}_{\bar{m}} \ \dots \ \bar{v}_{\bar{M}}]^T, \quad \bar{M} = (M^2/2 + M - 1) \tag{16}$$

Then, the covariance matrix of the virtual array can be obtained as follows. Let \mathbf{R} be the Toeplitz matrix:

$$\mathbf{R} = \begin{bmatrix} \bar{v}_{\frac{(M-1)}{2}+1} & \bar{v}_{\frac{(M-1)}{2}+2} & \dots & \bar{v}_{\bar{M}} \\ \bar{v}_{\frac{(M-1)}{2}} & \bar{v}_{\frac{(M-1)}{2}+1} & \dots & \bar{v}_{\bar{M}-1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{v}_1 & \bar{v}_2 & \dots & \bar{v}_{\frac{(M-1)}{2}+1} \end{bmatrix} \tag{17}$$

where $\mathbf{R} \in \mathbb{C}^{M^2/4+M/2 \times M^2/4+M/2}$. Forward-backward (FB) SS [34] is applied to \mathbf{R} to obtain the spatial smoothed covariance matrix, \mathbf{R}_{ss} . It should be noted that FBSS is used here only in establishing a full rank covariance matrix. The resulting n -th element of the weight vector $\mathbf{w}_{MVDR} \in \mathbb{C}^{N \times 1}$ is given by [2]:

$$w(\bar{\theta}_n) = \frac{1}{\mathbf{a}_1^H(\bar{\theta}_n) \mathbf{R}_{ss}^{-1} \mathbf{a}_1(\bar{\theta}_n)}, \quad n = 1, 2, \dots, N \quad (18)$$

which is also known as the scalar output power for a single steering direction [39]. The computational complexity of the MVDR algorithm [40] is as shown in Table 1. It is noted that the MVDR-based DOA estimation technique does not require the number of the source signals to be known in advance. Furthermore, the MVDR DOA estimation method performance is better than that of the conventional beamforming. In addition, assuming that ($N \gg \bar{M}$), the computational complexity of obtaining the MVDR weights is less than that of obtaining the OLS one.

Table 1. Computational complexity of the minimum variance distortionless response (MVDR) algorithm.

Operation	Computation	Cost
Inverse Covariance Matrix	\mathbf{R}_{ss}^{-1}	$O(\bar{M}^3)$
Beamformer Weight	$z = \frac{\mathbf{R}_{ss}^{-1} \mathbf{a}_1(\bar{\theta}_n)}{\mathbf{a}_1^H(\bar{\theta}_n) \mathbf{R}_{ss}^{-1} \mathbf{a}_1(\bar{\theta}_n)}$	$O(2\bar{M}^2 + 3\bar{M})$
Beamformer Sum	$w(\bar{\theta}_n) = [\mathbf{a}_1^H(\bar{\theta}_n) \mathbf{R}_{ss}^{-1} \mathbf{a}_1(\bar{\theta}_n)]^{-1}$	$O(\bar{M}N)$

It should be noted that our proposed algorithm does not depend on the orthogonality of the signal subspaces nor on implementing singular value decomposition (SVD) on the sensor array data. Therefore, it can perform DOA estimation without knowing the number of source signals in advance. On the other hand, subspace-based techniques, such as MUSIC and ESPRIT, cannot estimate the DOA without a priori knowledge of the number of source signals. Furthermore, it is known from the literature that the MUSIC algorithm is superior to the ESPRIT algorithm [41–43]. However, for the sake of evaluating our proposed algorithm in comparison with MUSIC, we assume that the number of signal source is to be known a priori.

4.3. Wideband MVDR A-LASSO DOA

Recently, wideband wireless applications have dominated the area of wireless communications. Orthogonal frequency division multiplexing (OFDM) transmission and ultra-wide band (UWB) systems are the most-used techniques in that field. DOA of wideband signals is required in such applications. However, there are only a few techniques that are applicable to the wideband scenario [44,45]. When DOA of wideband source signals needs to be estimated, we cannot represent the delays of the signals by simple phase shifts. An easy way to overcome this problem is to decompose the wideband signals into several narrowband ones and apply the DOA scheme for each of the narrowband signals. A filter-bank (or Fourier transform) consisting of Q filters could be used to decompose the wideband signal into the narrowband ones, where each filter $h_q(t)$ has a narrowband around the central frequency f_q , $q = 1, 2, \dots, Q$. Thus, our DOA estimation is essentially equivalent to estimating the DOA of Q narrowband signals, given by:

$$\mathbf{y}_q(t) = \Phi(f_q) \bar{\mathbf{s}}_q(t) + \mathbf{n}_q(t) \quad q = 1, 2, \dots, Q \quad (19)$$

The over-complete steering matrix, $\Phi(f_q)$, is computed for each filter centre-frequency, f_q .

$$f_q = \frac{\check{f}_q + \hat{f}_q}{2} \quad (20)$$

where \check{f}_q and \hat{f}_q are the lower and higher cut-off frequencies of the filters. It should be noted that in [45], the KR product is used to virtually extend the array aperture, thus resulting in an increase of the available degrees of freedom. Consequently, the algorithm proposed in [45] can detect more source signals than the number of sensors used. However, our proposed technique can detect more source signals than that of [45]. Moreover, the technique proposed in [45] cannot detect correlated source signals (unless they are decorrelated before DOA estimation); but, our proposed technique can identify correlated source signals. In addition, as will be demonstrated in the last experiment of Section 6, our proposed technique can estimate the DOAs of the source signals even for a wideband scenario with a very much smaller number of snapshots compared to that required with the technique in [45].

4.4. Selecting the Regularization Parameter

Choosing the regularization parameter, τ , is an important issue for the success of LASSO minimization Equation (10). The regularization parameter controls the trade-off between the data fidelity ($\|\mathbf{y} - \Phi\mathbf{s}\|_2^2$) and the prior information ($\|\mathbf{s}\|_1$). The discrepancy principle (DP), cross-validation (CV), generalized cross-validation (GCV) and L-curve method are some of the existing regularization parameter selection methods. The regularization parameter in DP is chosen so that the sum of squares of the weighted residuals is equal to the mean of a chi-square distribution [46–48]. CV selects the regularization parameter that minimizes the mean square error, while GCV selects the value of the regularization parameter that minimizes the GCV function, which is a leave-one-out CV function for large-scale problems [48,49]. The L-curve criterion is based on a log-log plot of the corresponding values of the solution norms and the residuals for a range of values of the regularization parameter [29,50,51]. From Figures 2a and 3a, it is seen that as the value of the regularization parameter τ is increased, the significance of the A-LASSO estimates shifts from large non-sparse estimates to smaller sparse estimates. In other words, a small value of τ leads to an under-regularized estimate, whereas a large value results into an over-regularized estimate. Therefore, suitable values of τ are those lying in the knee of the L-curve. We consider the beginning of the knee to correspond the value of τ at which the solution norm starts to decrease and the end of the knee to correspond to the value of τ at which the residual norm does not significantly change. These are the two red stars on the L-curve. We empirically determine a segment of this knee corresponding to which all of the τ values provide satisfactory estimates. We have chosen the value of the regularization parameter τ to be the midpoint of this segment. It should be mentioned that there are different methods to select a suitable value for the regularization parameter in the literature [46–49]. However, it has been shown in [52] that the L-curve [29,50,51] method gives a good estimation of the regularization parameter. In order to illustrate how τ is chosen, we consider two source signals from DOAs of 60° and 120° to be impinging a six-sensor two-level nested array with the sampling grid being uniform from 1° to 180° , in increments of 1° , and an SNR of 10 dB; the corresponding L-curve plot is as shown in Figure 2a. Selecting τ to be between 1.39 and 2.19, the resulting DOA estimation is as shown in Figure 2b. Lowering the SNR to be 0 dB, the results for the same specifications are shown in Figure 3. In this case, a suitable value for τ is between 1.86 and 2.58. From Figure 2b, it can be seen that we are able to identify correctly the two source signals, even at low SNR conditions.

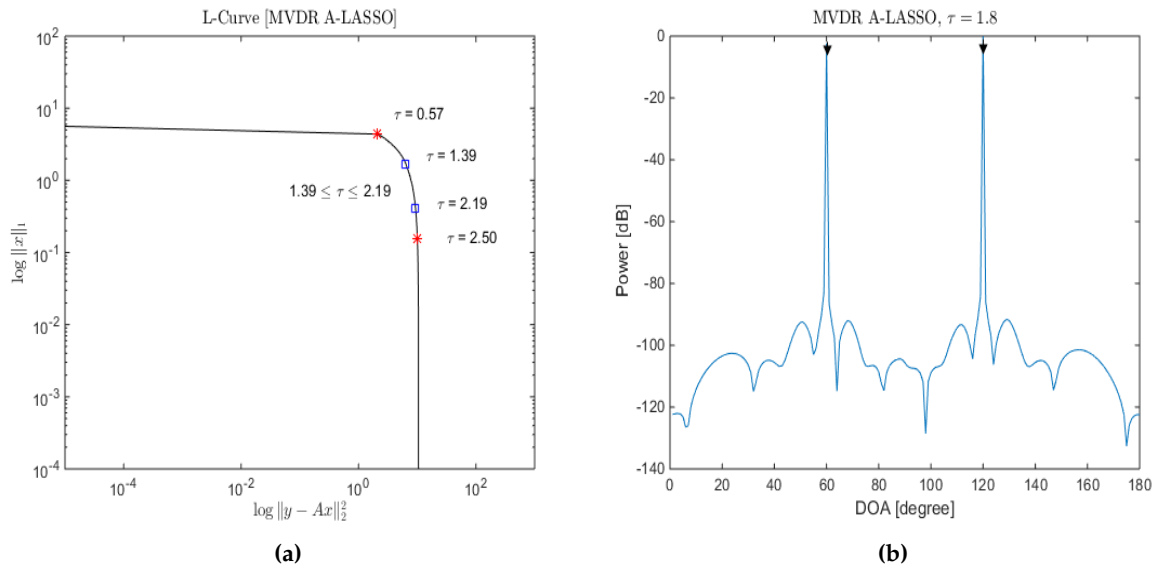


Figure 2. (a) The data residual $\|y - \Phi^* \bar{s}\|_2^2$ versus the solution ℓ_1 -norm linear scale on a log-log scale (L-curve); (b) DOA estimation for two source signals; τ was selected using L-curve, in the MVDR A-LASSO problem (SNR = 10 dB).

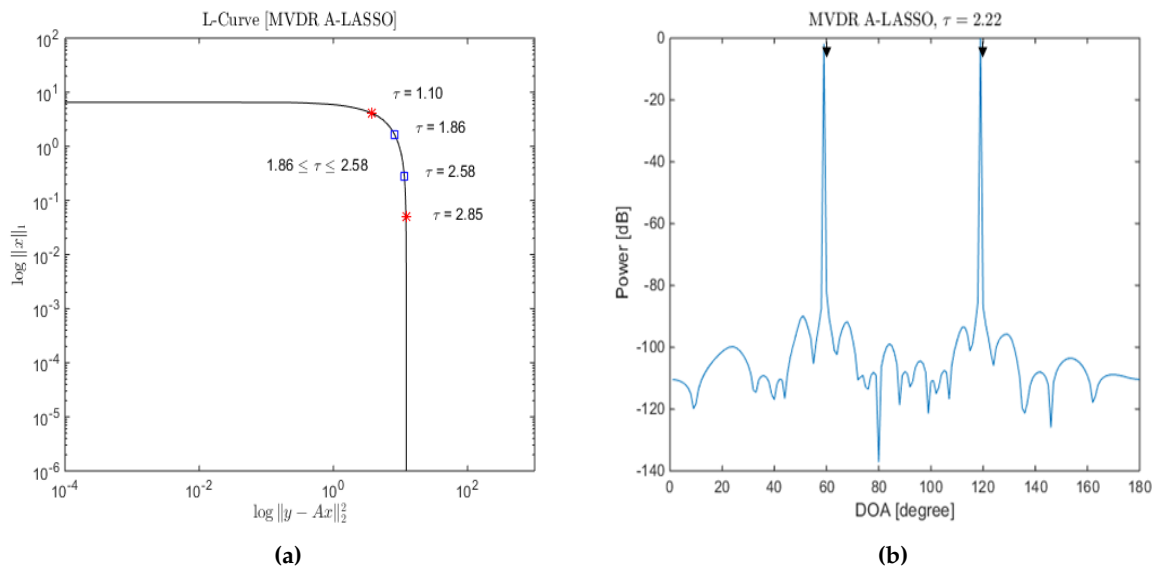


Figure 3. (a) The data residual $\|y - \Phi^* \bar{s}\|_2^2$ versus the solution ℓ_1 -norm linear scale on a log-log scale (L-curve); (b) DOA estimation for two source signals; τ was selected using L-curve, in the MVDR A-LASSO problem (SNR = 0 dB).

5. Simulation Results

Consider a sparse linear two-level nested array, for which M is even, consisting of $M = 6$ elements, as shown in Figure 4. Investigating the array output by applying Equations (1) to (3) and extracting the equivalent distinct virtual elements from the virtual array manifold ($\mathbf{A}^* \odot \mathbf{A}$), one can see that the virtual uniform linear array (ULA) contains $\bar{M} = 23$ elements, as shown in Figure 4. It should be noted that the resultant virtual is a ULA [14]. The sampling grid $\theta_n \in [1^\circ : 180^\circ]$ that covers Ω is chosen to be of 1° step, except for the twelfth simulation and $d = \lambda/2$, where λ is the wavelength of the propagating waves.

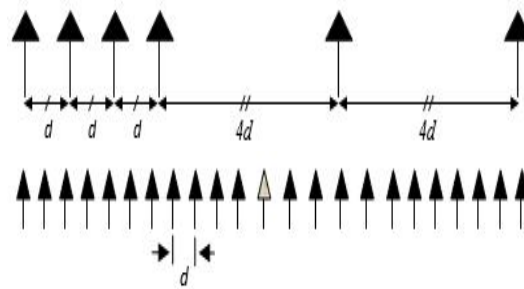


Figure 4. The proposed sparse (**upper**) and the virtual co-array (**lower**).

All of the simulated source signals are assumed to be equi-power and uncorrelated with one another or with the noise, except in the fifth simulation, where the sources are assumed to be correlated. The weight parameter, γ , is set to 0.5 in all of the simulations, except for the eleventh simulation. The total number of simulations, N_{sim} , is set to $N_{sim} = 100$ for each observation point except in the case of the eleventh simulation, where it is set to 10. For each simulation, the regularization parameter, τ , is selected based on the idea of the L-curve [29,50,51]. For the wideband scenario, two chirp signals are assumed for the wideband simulations, and the filter-bank is assumed to contain 10th order Butterworth bandpass filters.

The CVX toolbox [53,54] for convex optimization that is available within the MATLAB environment is used for examining the performance of the proposed A-LASSO algorithms. It uses semi-definite quadratic-linear programming (SDPT3) [55] to obtain the global solution for the optimization problem.

The root mean square error (RMSE) is used as the performance measure:

$$RMSE = \frac{1}{L} \sum_{l=1}^L \sqrt{\frac{1}{N_{sim}} \sum_{n=1}^{N_{sim}} (\hat{\theta}_{l,n} - \theta_l)^2} \quad (21)$$

where $\hat{\theta}_{l,n}$ is the estimate of the DOA angle θ_l of the n -th Monte Carlo trial.

5.1. Narrowband Signal Sources

5.1.1. Investigations of LASSO-Based Algorithms

In the first simulation, we study the effect of the initial vectors on the performance of the three LASSO algorithms, namely classical LASSO (for which $\gamma = 0$), OLS A-LASSO and MVDR A-LASSO. For the latter two, we assume $\gamma = 0.5$. We consider two source signals impinging on the sparse array from the DOA of 60° and 120° . For SNR of 0 dB, 10 snapshots and one iteration, the results are as shown in Figure 5. It can be seen that the MVDR A-LASSO yields a performance better than that of the classical LASSO, as well as that of the OLS A-LASSO. It may be mentioned that by increasing the number of iterations, OLS A-LASSO can be made to yield a performance similar to that of MVDR A-LASSO. This superior performance of the MVDR A-LASSO can be attributed to the initial weights used, compared to the least square weights used for the OLS A-LASSO, as will be seen later in Simulation 6. It should be noted that classical LASSO uses equal initial weights and has the poorest performance.

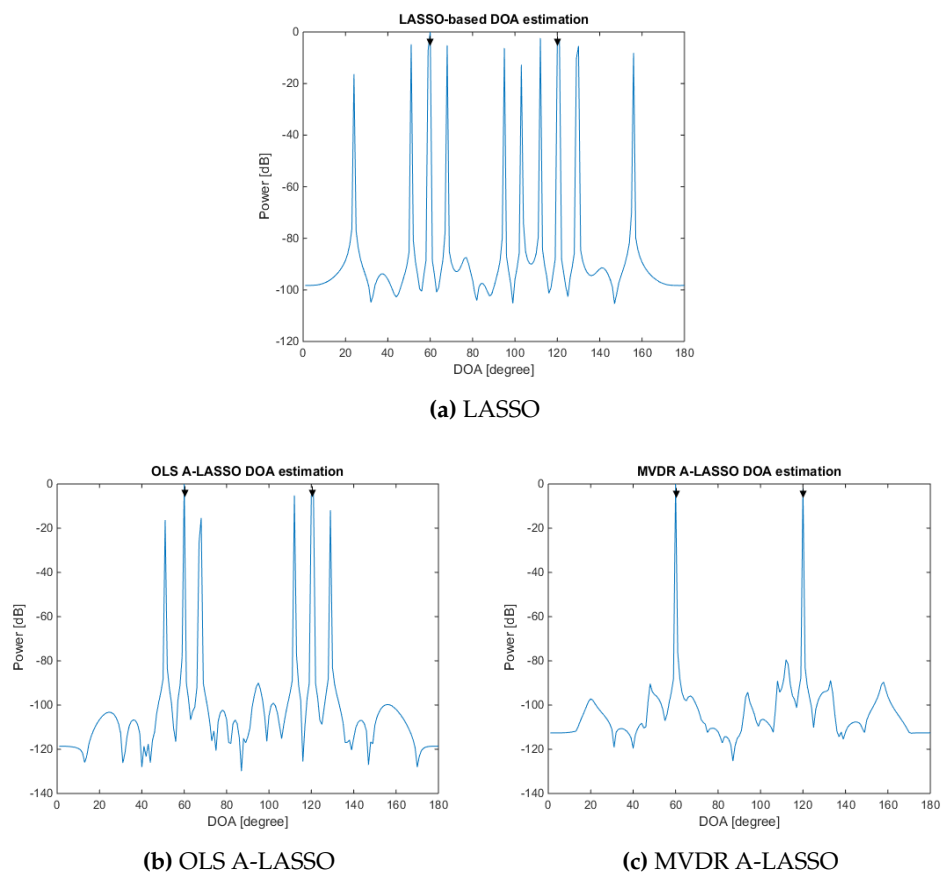


Figure 5. Performance of LASSO, OLS A-LASSO and MVDR A-LASSO, for two source signals at DOAs 60° and 120°, 10 snapshots, SNR = 0 dB and one iteration. (a) LASSO; (b) OLS A-LASSO; and (c) MVDR A-LASSO.

In the second simulation, we investigate the performance of the proposed MVDR A-LASSO algorithm as we vary SNR and compare it with that of LASSO and OLS A-LASSO. Two source signals are assumed to impinge on the sparse array from DOA of 60° and 120°. The performances of the proposed MVDR A-LASSO algorithm, along with that of the conventional LASSO and OLS A-LASSO are shown in Figure 6. It is clear from this figure that the MVDR A-LASSO algorithm outperforms both the LASSO and OLS A-LASSO algorithms for all SNR.

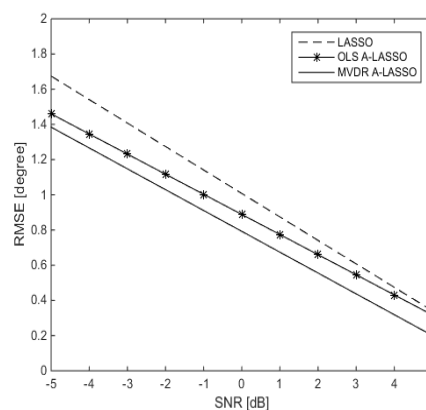


Figure 6. Performance of the three LASSO algorithms versus SNR, for two source signals at DOAs 60° and 120°, 10 snapshots and after one iteration of the MVDR A-LASSO and OLS A-LASSO algorithms.

Assuming now that L (the number of source signals) is known, we compare the performance of the LASSO algorithms with that of the MVDR algorithm and that of MUSIC. For that purpose, two ULAs, one consisting of six elements and another consisting of 23 elements, are used to evaluate the performance of MUSIC, while only the ULA with six elements is considered for the MVDR algorithm. However, for the three LASSO algorithms, the real array used is as shown in Figure 4, namely with six elements. The performance of the various algorithms as SNR is varied is shown in Figure 7. It is observed from the figure that all three LASSO algorithms outperform the MUSIC algorithm, as well as the MVDR algorithm, even when 23 elements are used in the array.

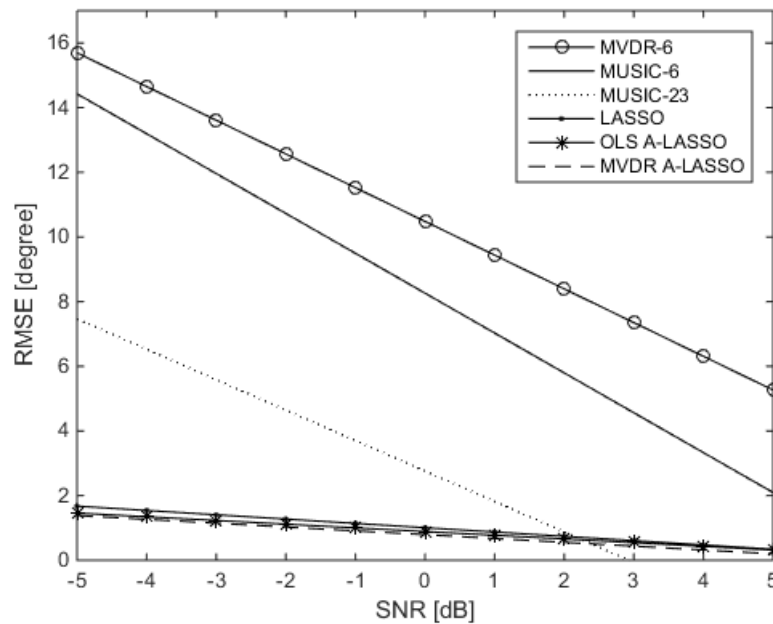


Figure 7. Performance of the LASSO algorithms as SNR is varied in comparison with that of MVDR and MUSIC algorithms, for two source signals at DOAs 60° and 120° , 10 snapshots and after one iteration of the MVDR A-LASSO and OLS A-LASSO algorithms.

We investigate, in the third simulation, the capabilities of the proposed algorithms in detecting the sources even when the number of sources exceeds the number of physical array elements. In other words, the proposed algorithm is for an underdetermined DOA scenario. For that purpose, two ULAs, one consisting of six elements and another consisting of 23 elements, are used to evaluate the performance of MVDR, while only the ULA with 23 elements is considered for the MUSIC algorithm. However, for the three LASSO algorithms, the real array used is as shown in Figure 4, namely with six elements. Let 11 source signals impinge the array from uniformly-distributed DOAs over $\theta = [30^\circ, 150^\circ]$. The snapshots number is chosen to be 70, and SNR is set to be -5 dB. All of the LASSO algorithms can easily identify 11 peaks (even after just one iteration of the MVDR A-LASSO and OLS A-LASSO algorithms), as seen from Figure 8, while MVDR and MUSIC fail to identify the source signals.

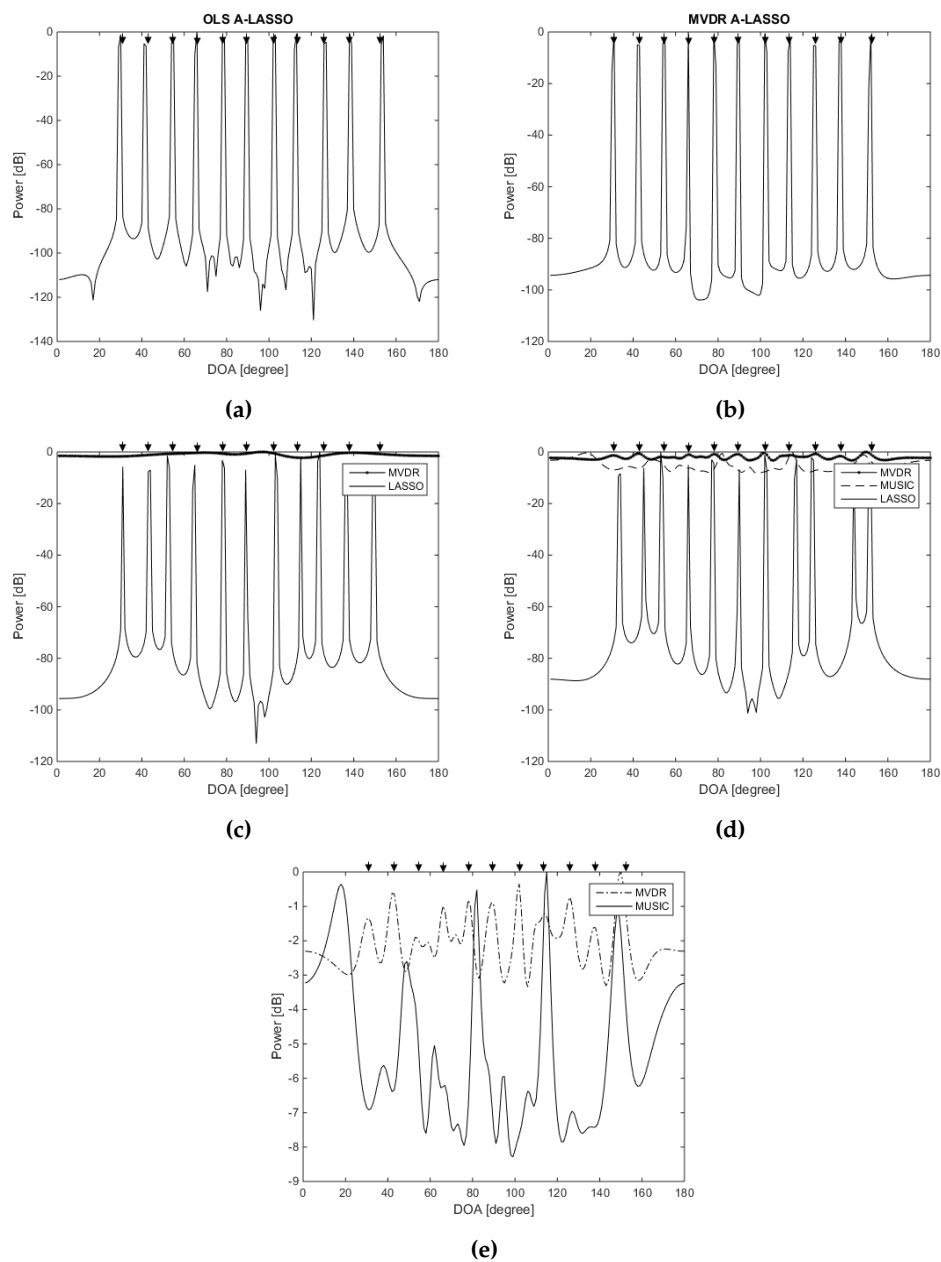


Figure 8. DOA estimation when the number of sources is more than the number of sensors: (a) After one iteration of OLS A-LASSO; (b) After one iteration of MVDR A-LASSO; (c) Classical LASSO and MVDR using a six-element array; (d) Classical LASSO, MVDR and MUSIC using a 23-element array; (e) MVDR and MUSIC using a 23-element array.

In the fourth simulation, we examine the resolution of the proposed adaptive algorithms in comparison with that of MVDR and MUSIC. Two spatially-correlated equi-power signals are assumed to impinge on the array from the DOAs of 85° and 95° . The SNR is set to 15 dB. Figure 9 illustrates the results. Two peaks can easily be identified in the case of the three LASSO algorithms, while in the case of MVDR and MUSIC algorithms, the two peaks are merged into one.

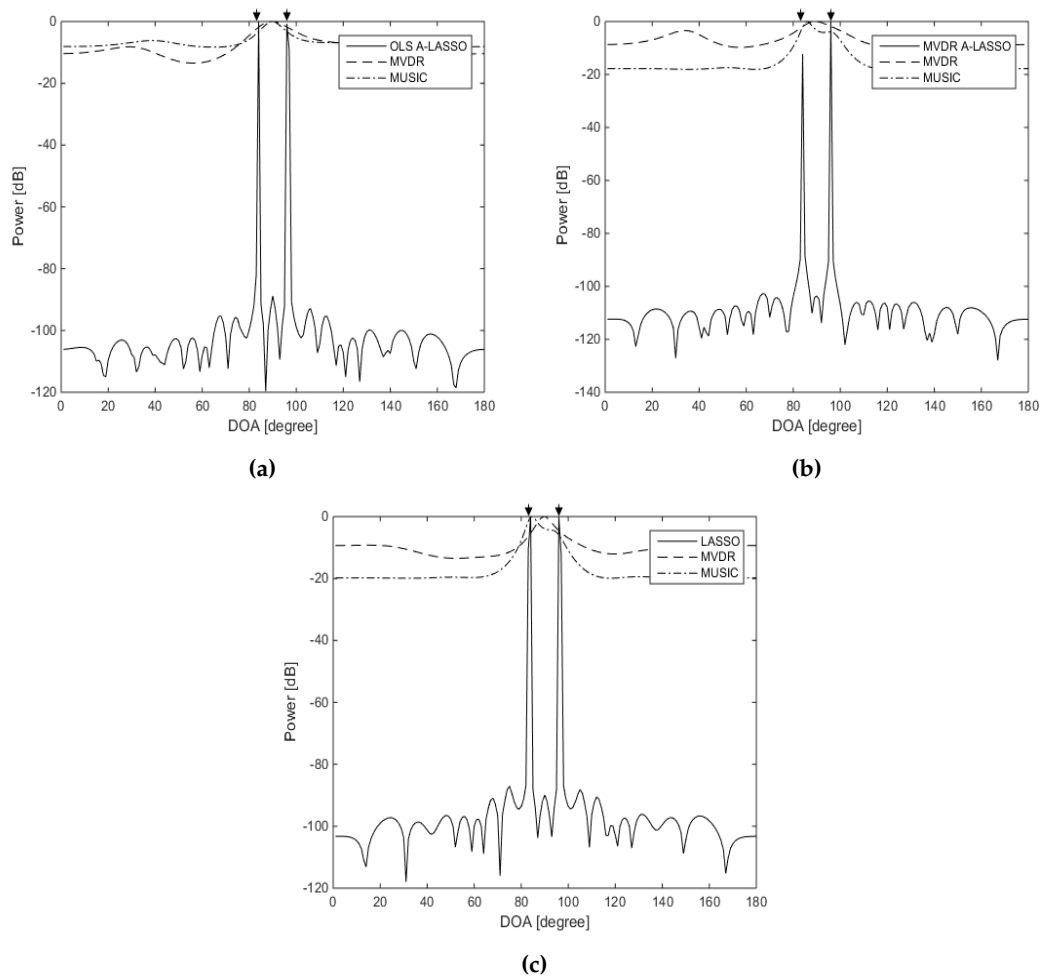


Figure 9. DOA estimation for spatially-closed two-source signals using LASSO algorithms, for two source signals at DOAs 85° and 95° , 10 snapshots, SNR = 15 dB and one iteration of the MVDR A-LASSO and OLS A-LASSO algorithms. (a) OLS A-LASSO after the first iteration; (b) MVDR A-LASSO after the first iteration; and (c) the classical LASSO algorithm.

In the fifth simulation, we examine the performance of our proposed algorithms for the detection of correlated source signals. Two fully-correlated (coherent) source signals are assumed to impinge on the array from directions of 60° and 100° with SNR set to 15 dB. Figure 10 shows that the three LASSO algorithms can resolve the two sources, revealing the capability of the algorithms in detecting correlated source signals. It is also clear that MVDR and MUSIC fail to distinguish the two sources.

From the above simulations, it is seen that the performance of the A-LASSO-based DOA estimation is superior to that of the MVDR (for which L is not required to be known) and that of MUSIC (for which L must be known in advance). Further, we also conclude that the performance of the classical LASSO is inferior to that of the two A-LASSO schemes, even though it exhibits a performance better than that of MVDR and MUSIC. In view of these results, we will not consider MVDR, MUSIC or the classical LASSO algorithm further in our study.

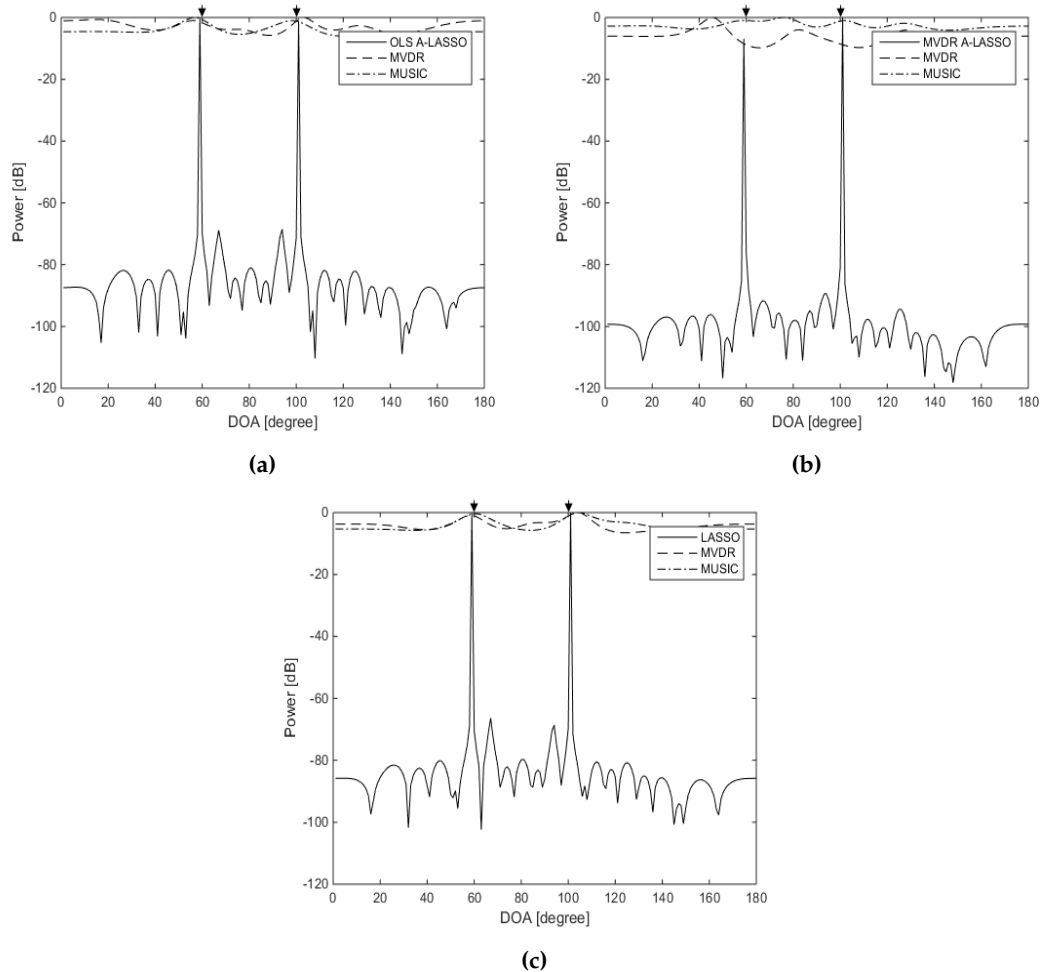


Figure 10. DOA estimation for two correlated source signals using LASSO algorithms, for two source signals at DOAs 60° and 100° , 10 snapshots, SNR = 15 dB and one iteration of the MVDR A-LASSO and OLS A-LASSO algorithms. (a) OLS A-LASSO after the first iteration; (b) MVDR A-LASSO after the first iteration; and (c) the classical LASSO algorithm.

5.1.2. Investigations of A-LASSO Algorithms

In the sixth simulation, we test the performance of the OLS A-LASSO and MVDR A-LASSO algorithms in a low SNR situation. For this purpose, we consider two source signals with DOAs of 60° and 120° , set SNR to 0 dB and snapshots to 10. The results for the two A-LASSO algorithms are shown in Figure 11. It can be seen from this figure that MVDR A-LASSO can detect the source signals after the first iteration itself, while OLS A-LASSO needs more iterations to be able to eliminate all of the false peaks. This can be explained by looking at the initial weights for both the OLS A-LASSO and MVDR A-LASSO algorithms. Figure 11d illustrates the initial weights for OLS A-LASSO and MVDR A-LASSO; it can be seen that the weights using the MVDR A-LASSO algorithm are relatively smooth compared to those of OLS A-LASSO. Furthermore, it can be seen that the OLS A-LASSO weight consists of many peaks that affect its performance and lead to false source signal peaks.

The DOA performance is investigated after five and 15 iterations for both the OLS A-LASSO and MVDR A-LASSO in the seventh simulation. Two signal sources are assumed to be impinging the array from DOAs of 60° and 120° , while the SNR changes. The snapshot number is chosen to be 10; the results are shown in Figure 12. It can be seen from the figure that, in terms of RMSE, MVDR A-LASSO outperforms the OLS A-LASSO algorithms.

Based on the results of the previous simulations (Figures 5, 6, 11 and 12), it is clear that MVDR A-LASSO outperforms OLS A-LASSO. Hence, OLS A-LASSO will not be considered in the rest of the paper.

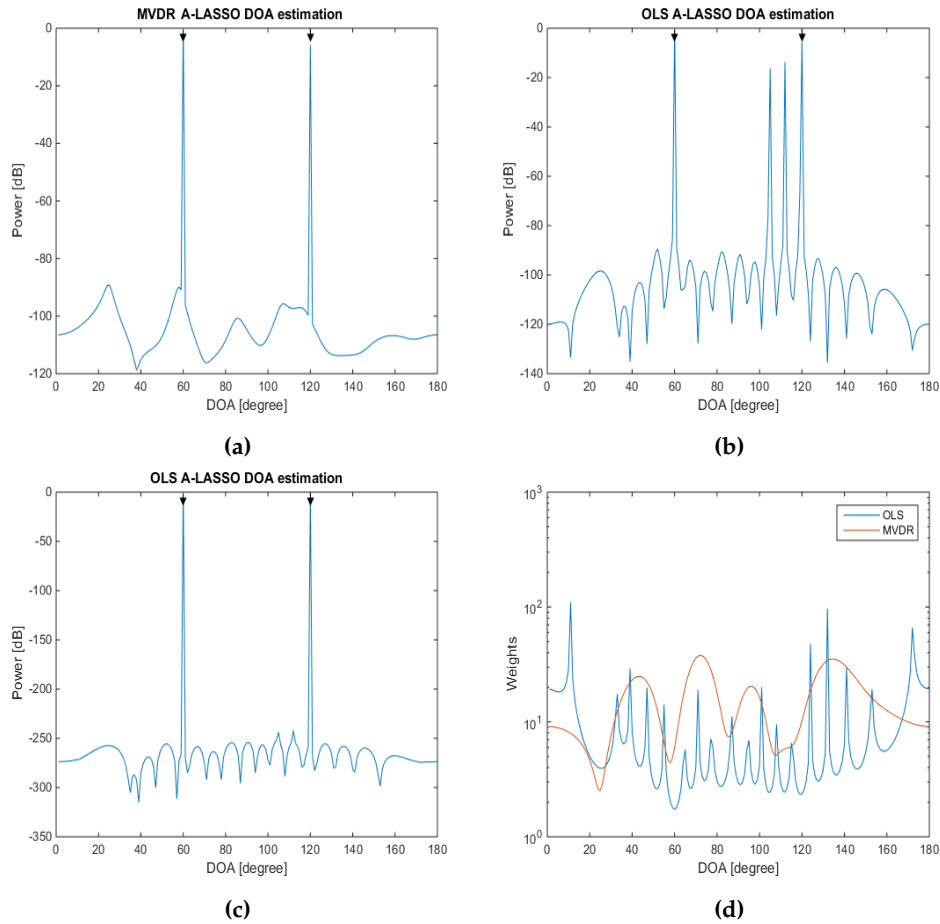


Figure 11. DOA estimation using A-LASSO algorithms, for two source signals at DOAs 60° and 120° , 10 snapshots, SNR = 0 dB. (a) MVDR A-LASSO after the first iteration; (b) OLS A-LASSO after the first iteration; (c) OLS A-LASSO after five iterations; and (d) initial weights of the two algorithms.

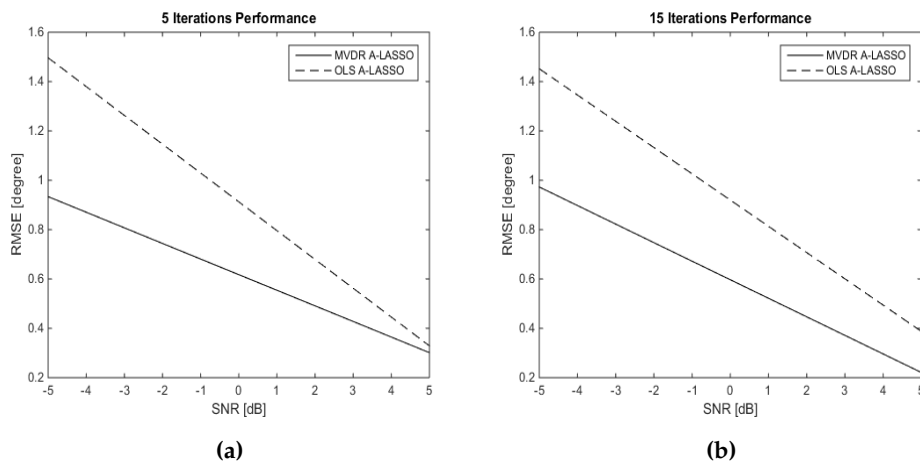


Figure 12. DOA estimation of two source signals at DOAs 60° and 120° , and 10 snapshots (a) after five iterations and (b) after 15 iterations, using the A-LASSO algorithms.

5.1.3. Investigations of the MVDR A-LASSO Algorithm

The eighth simulation investigates the DOA estimation using MVDR A-LASSO as we increase the number of iterations. Let two source signals with DOAs of 60° and 120° impinge on the array, and let the SNR be -5 dB and 50 snapshots be used. The results for the first five iterations are shown in Figure 13. It is seen from the figure that, after the first iteration, fake source signal peaks appear. As the algorithm runs, the weights corresponding to the fake source signals become very large, whereas those of the actual source signals remain constant. Hence, the weights corresponding to the false source signals damp the false peaks, while those of the real source signals remain constant. As a consequence, as the number of iterations increases, it is clear that only the real source signal peaks remain. Furthermore, it can be observed that the sidelobe ratio (SLR) after the fifth iteration is more than twice that after the first iteration.

We now examine the performance of the proposed MVDR A-LASSO algorithm at low and very low SNR situations in the ninth simulation. The same settings as in the previous simulation are used, except that the SNR is set to -10 dB and -15 dB. The snapshot number is set to 150 for the first case and 200 for the second one. The results are as shown in Figure 14 for SNR of -10 dB and Figure 15 for that of SNR set to be -15 dB.

It can be seen from Figures 14 and 15 that the two signals can be identified after only five iterations, even for very low SNR conditions. However, more snapshots are needed in this situation. Thus, it is a trade-off between SNR and the number of snapshots required so that the DOA of the source signals can be correctly identified. It is further observed that even after three iterations, we are able to identify the two source signals.

In the tenth simulation, we study the effect of changing the number of snapshots on the performance of MVDR A-LASSO. We consider two source signals arriving from DOAs of 60° and 120° with SNR changing from -5 dB to 5 dB. The results are shown in Figure 16. It can be observed from this figure that increasing the number of snapshots leads to an enhancement in the performance, in terms of RMSE. In other words, it is a trade-off between the number of snapshots and the RMSE. For low and very low SNR, the number of snapshots has to be increased, for better performance.

The eleventh simulation involves the investigation of the effect of changing γ . Two source signals impinging the array from DOAs of 60° and 120° are considered, the SNR being set to -5 dB while changing γ . Figure 17 shows the residual, R , in terms of the absolute value of $(\|\mathbf{y} - \Phi\bar{\mathbf{s}}_{k+1}\|_2^2 - \|\mathbf{y} - \Phi\bar{\mathbf{s}}_k\|_2^2)$, where k is the iteration number, for 10 simulations. From this figure, it is clear that increasing the number of snapshots decreases the residuals for all three cases of $\gamma = 0.25, 0.5$ and 0.75 . Furthermore, the residual after the end of the first iteration using $\gamma = 0.5$ is smaller than when $\gamma = 0.25$ or 0.75 , even when the number of snapshots is increased. From Figure 17b, it can be seen that, at the end of the second iteration, the residual corresponding to $\gamma = 0.25$ is smaller than that corresponding to $\gamma = 0.5$ or $\gamma = 0.75$. However, for none of these values of γ have we achieved convergence by the end of the second iteration. At the end of Iteration 3, the convergence of the residual is achieved only for the weight factor $\gamma = 0.5$. However, since the residual for $\gamma = 0.5$ at the end of the first iteration is less than that of $\gamma = 0.25$, the signal sources are identified at the end of the first iteration (Figure 18e), while that of $\gamma = 0.25$ needs more iterations (Figure 18a–c). The effect of γ on the DOA estimation, using the same previous source signals, with SNR chosen to be -5 dB is shown in Figure 18. In this case, 50 snapshots are used, while γ assumes values of 0.25, 0.50 and 0.75. It can be seen from this figure that $\gamma = 0.5$ provides the right signal sources at the end of the first iteration with a smaller residual than for $\gamma = 0.25$ or 0.75 . On the other hand, selecting $\gamma = 0.25$ leads to more fake source signals being detected, and as a consequence, more iterations are needed to identify the real sources, while using $\gamma = 0.75$ leads to misidentifying one of the source signals.

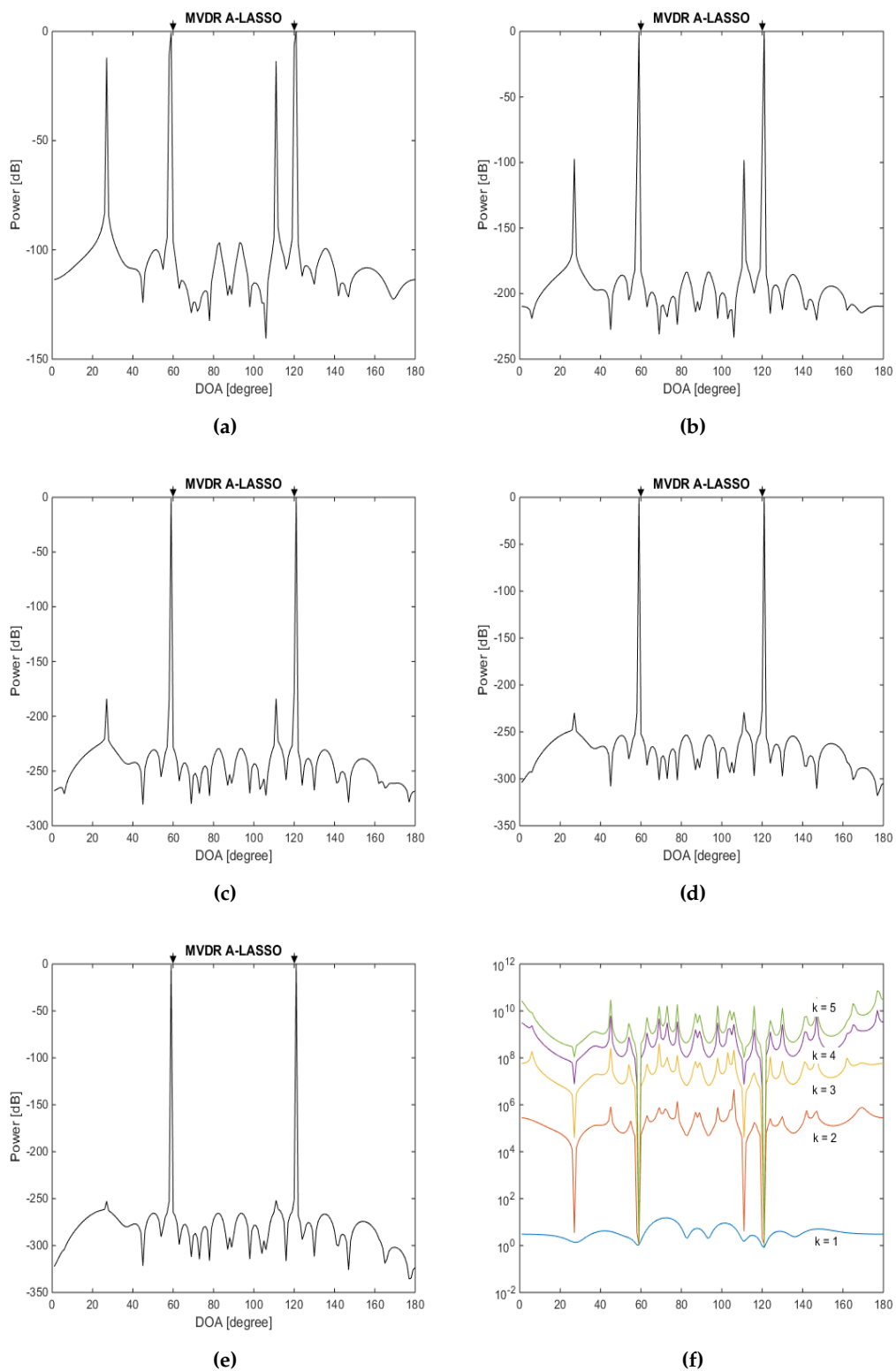


Figure 13. DOA estimation in the case of two source signals at DOAs 60° and 120° , SNR = -5 dB, 50 snapshots using MVDR A-LASSO algorithm. (a–e) after 1 – 5 iterations; and (f) MVDR A-LASSO weights as the number of iterations k varies.

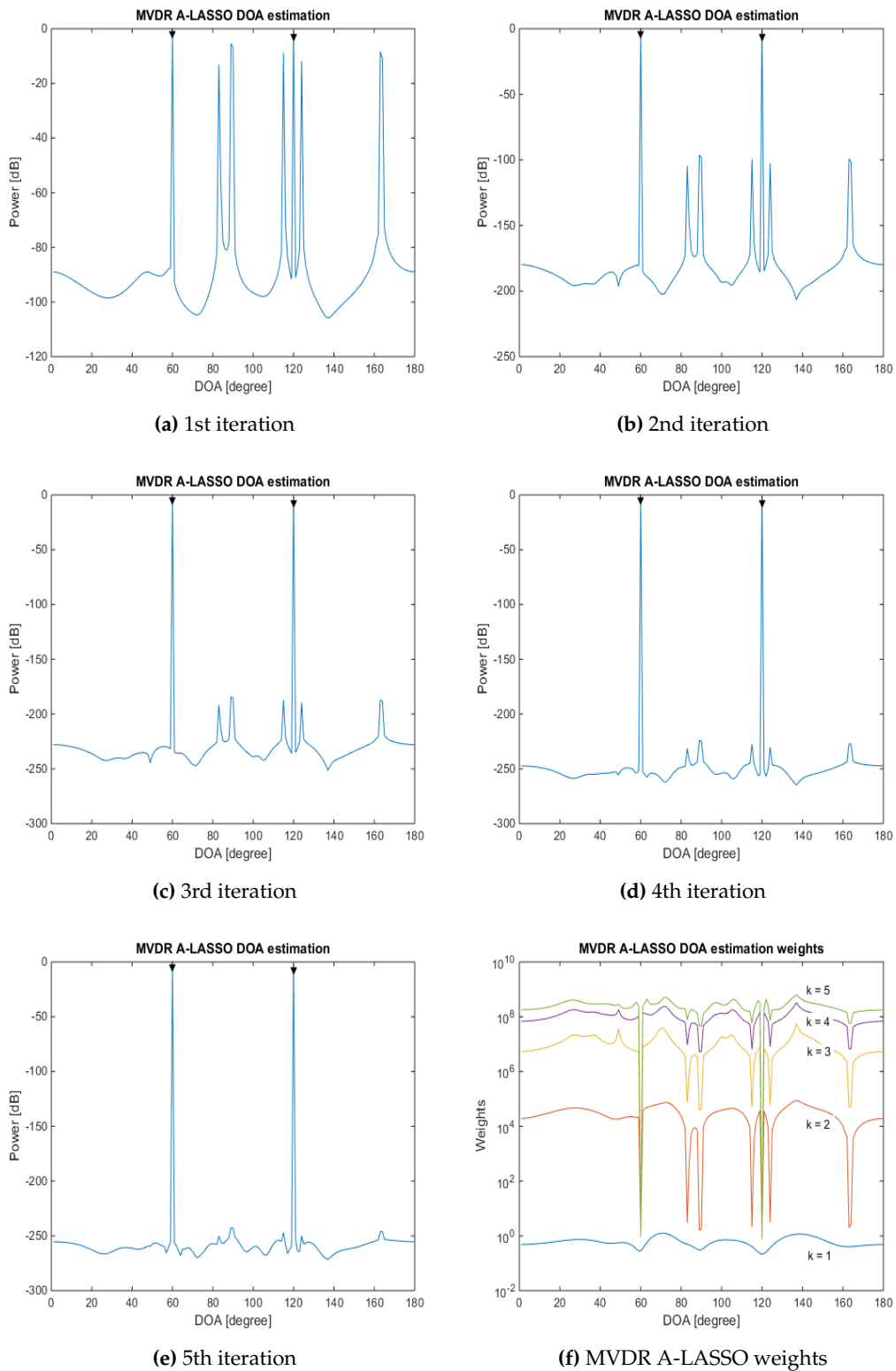


Figure 14. DOA estimation in the case of two source signals at DOAs 60° and 120° , SNR = -10 dB, 150 snapshots, using the MVDR A-LASSO algorithm. (a) to (e), after one to five iterations; and (f) MVDR A-LASSO weights as the number of iterations k varies.

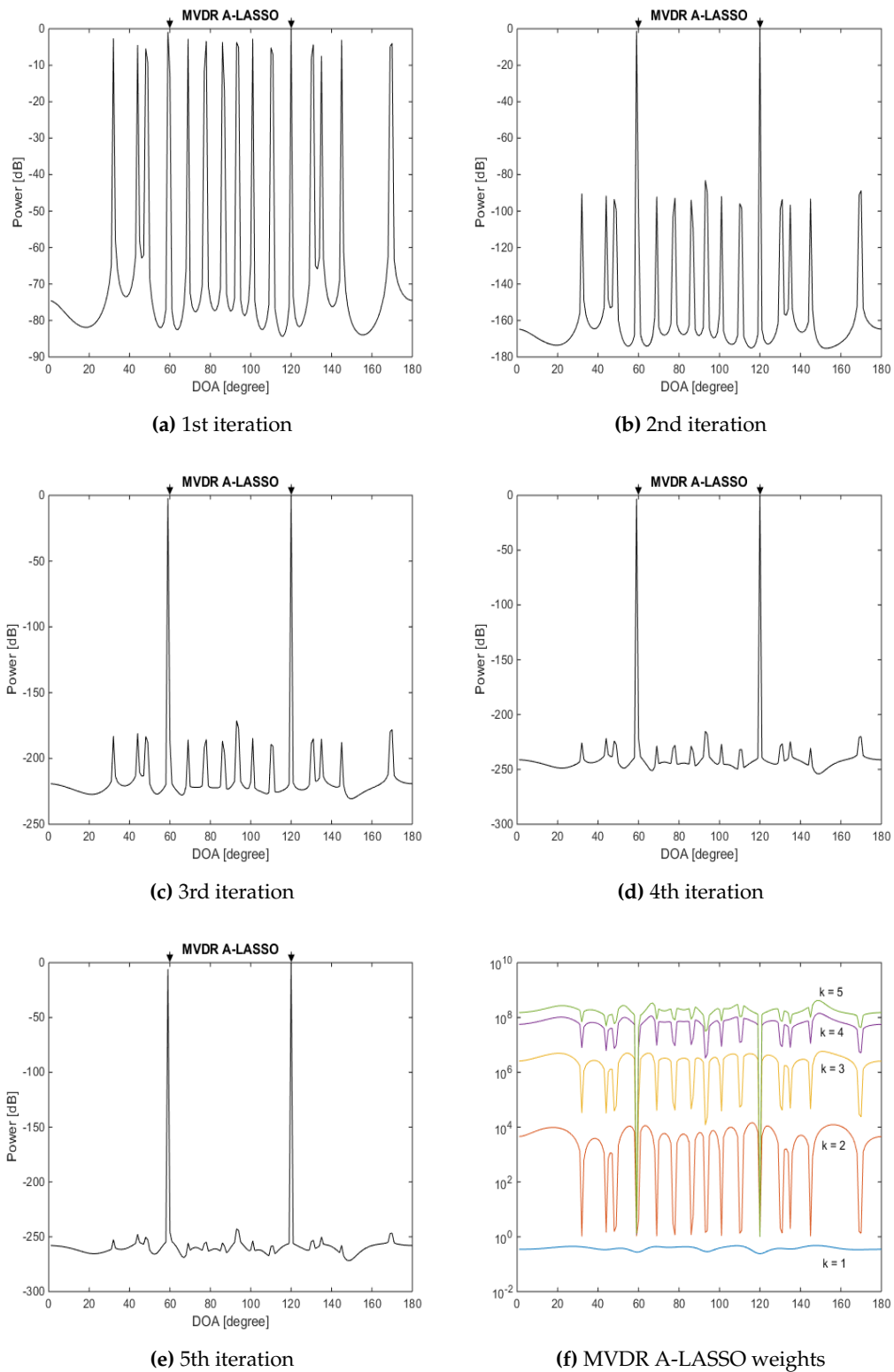


Figure 15. DOA estimation, two source signals at DOAs 60° and 120° , SNR = -15 dB, 200 snapshots, using the MVDR A-LASSO algorithm. (a) to (e) after one to five iterations; and (f) MVDR A-LASSO weights as the number of iterations k varies.

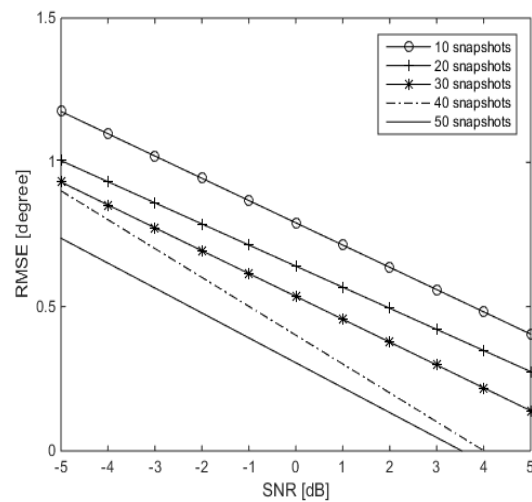


Figure 16. MVDR A-LASSO DOA estimation performance versus the number of snapshots, two source signals with DOAs 60° and 120° , $\gamma = 0.5$.

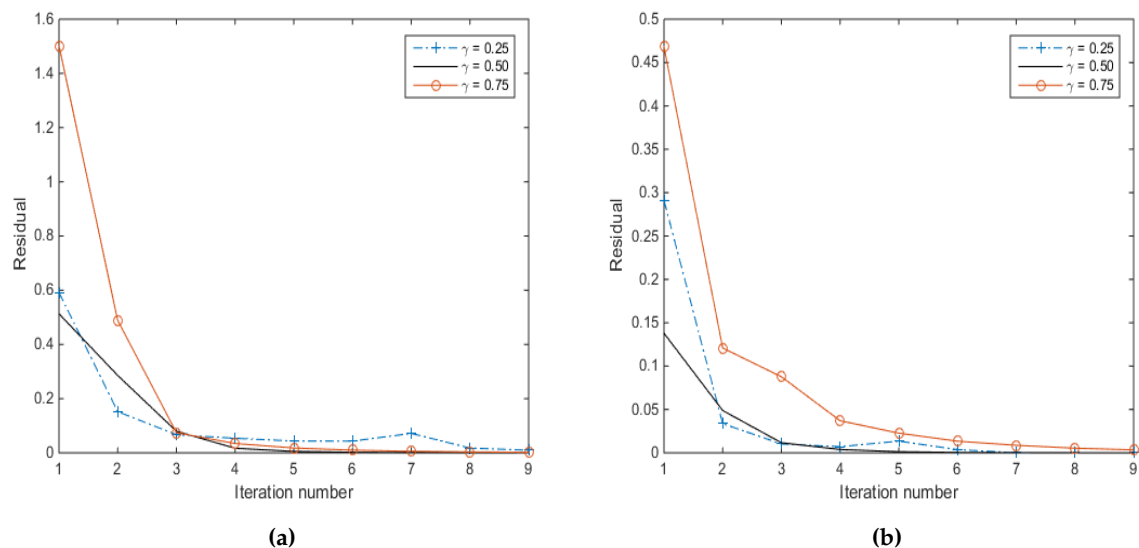


Figure 17. The residual for two source signals at DOAs 60° and 120° , with $\gamma = 0.25, 0.5$ and 0.75 , 10 iterations, SNR = -5 dB using the MVDR A-LASSO algorithm, (a) 10 snapshots and (b) 50 snapshots.

The twelfth simulation involves the investigation of the effect of varying the angular separation between the source signals. Consider two source signals, the first one held fixed at DOA of 60° , while the second one with DOA ranging from 61° to 100° with steps of 1° . The SNR is set to be 10 dB; 10 snapshots are considered for the simulation; 100 trials for each point; and a sampling grid $\bar{\theta}_n \in [1^\circ : 180^\circ]$ chosen to be of 0.1° steps. Figure 19 illustrates the DOA estimation error as a function of the angular separation between the two source signals. It can be seen from this figure that the DOA estimation error is less than 2° for an angular separation $< 5^\circ$. This DOA estimation error is reduced to $< 0.4^\circ$ for an angular separation $\geq 6^\circ$.

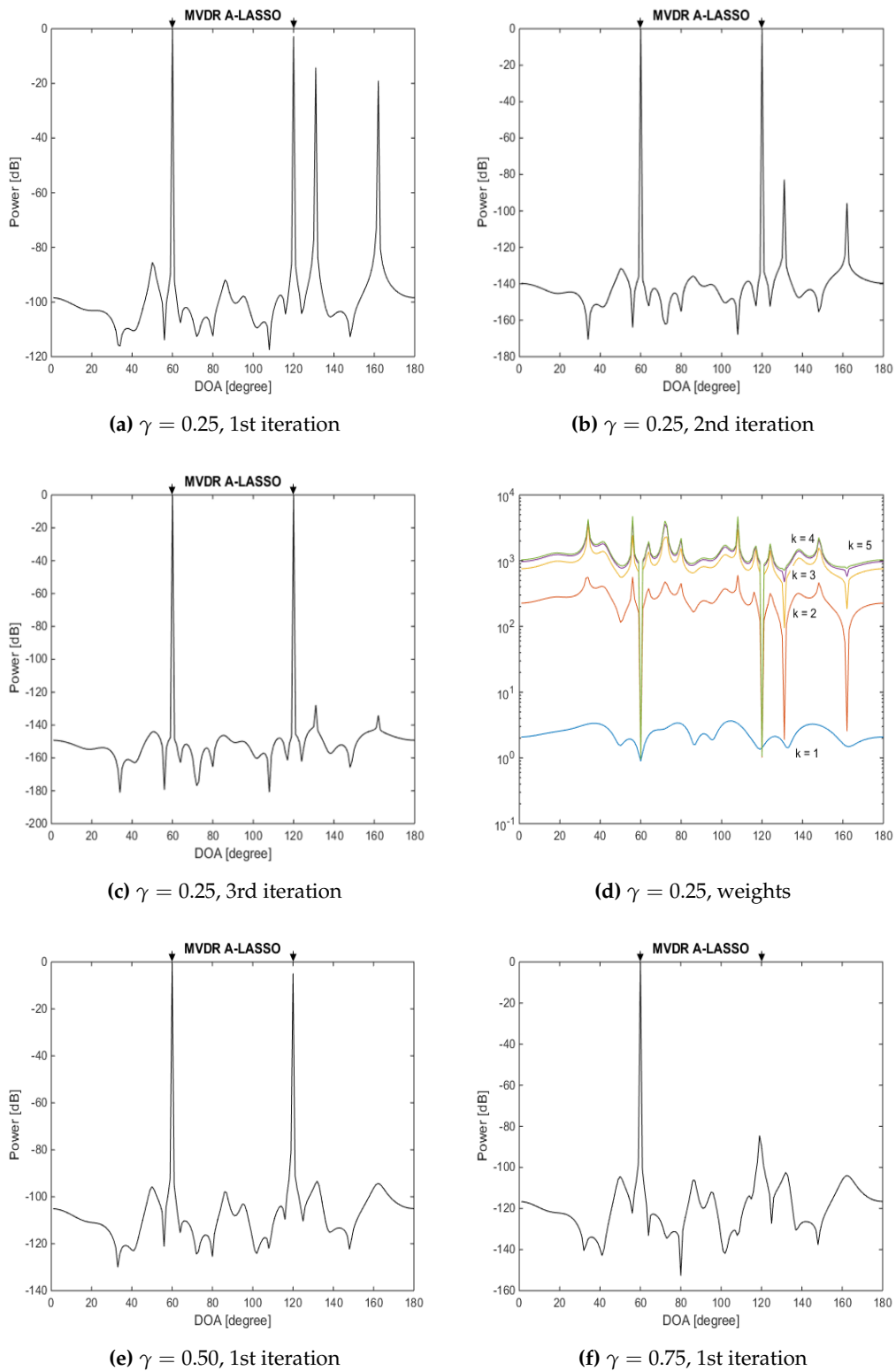


Figure 18. DOA estimation of two source signals at DOAs 60° and 120° , 50 snapshot, SNR = -5 dB, using the MVDR A-LASSO algorithm, (a)–(d) $\gamma = 0.25$; (e) $\gamma = 0.50$; and (f) $\gamma = 0.75$.

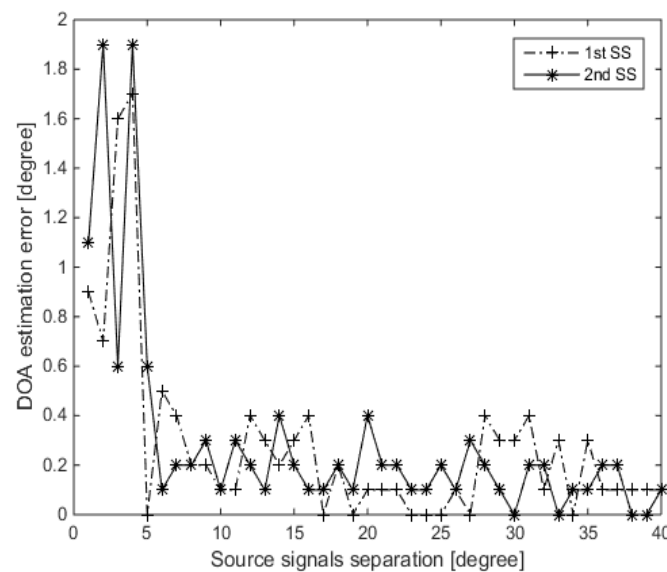


Figure 19. DOA estimation error for two sources as a function of separation between the two sources, SNR = 10 dB, 10 snapshots and one iteration of MVDR A-LASSO.

5.2. Wideband Signal Sources

The last simulation is to investigate the DOA estimation in a wideband source signal scenario. Two chirp source signals, with frequency spanning from 250 to 500 Hz, are assumed to be impinging the array from DOAs of 60° and 120° , with the SNR being set to 30 dB and the number of snapshots being 10. Figures 20 and 21 show the wideband DOA estimation results assuming two ULAs containing six sensors and 23 sensors, respectively. Here, we compare the performance of the proposed MVDR A-LASSO with that of both the conventional beamforming and MUSIC algorithm in a wideband scenario. It can be seen that the conventional beamforming suffers from wide beams so that the source signals can be merged if they are close to each other (Figures 20c,d and 21c,d). Increasing the number of sensors to 23 (the same number of sensors in the virtual array) leads to narrower beams (Figure 21c,d). However, conventional beamforming still has wider beams. Using the MUSIC algorithm, we cannot identify the two source signals using a six-sensor ULA (Figure 20e,f); however, we can identify them using an array with 23 sensors (Figure 21e,f). Using MVDR A-LASSO (Figure 20a,b), we can easily identify the source signals using the sensor array shown in Figure 4. We notice that the sidelobe ratio of the MVDR A-LASSO is superior to that of both the conventional beamforming and MUSIC algorithm. Furthermore, MVDR A-LASSO has a sharper beam compared to that of both the conventional beamforming and MUSIC algorithm. This simulation shows that MVDR A-LASSO is applicable to wideband scenarios, as well as to narrowband situations.

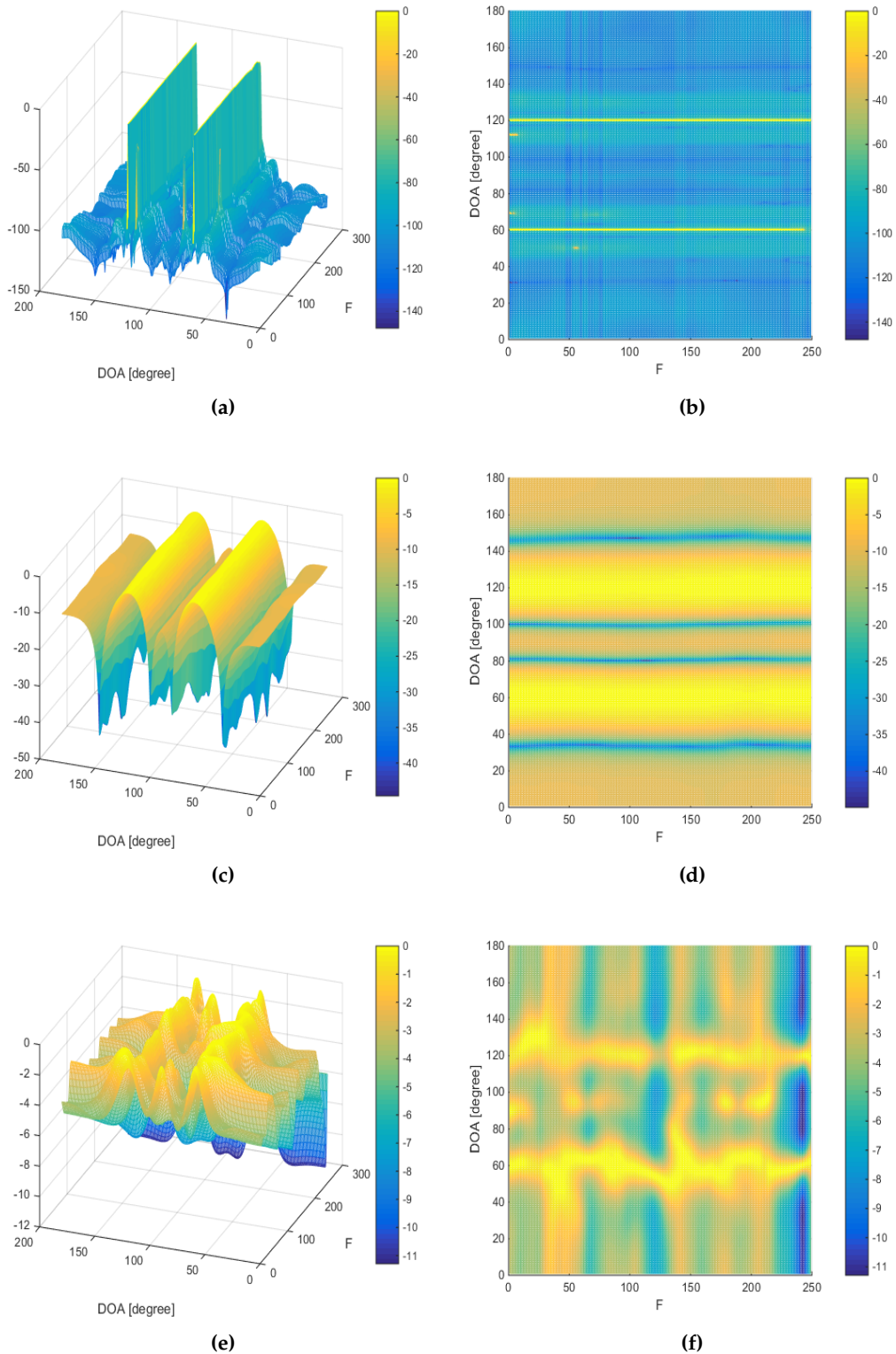


Figure 20. Wideband DOA estimation, two chirp source signals at DOAs 60° and 120° , $\gamma = 0.5$, 10 snapshots, using uniform linear array (ULA) containing six sensors for conventional beamforming and the MUSIC algorithm. (a,b) The MVDR A-LASSO algorithm (after the first iteration); (c,d) Conventional beamforming; and (e,f) the MUSIC algorithm.

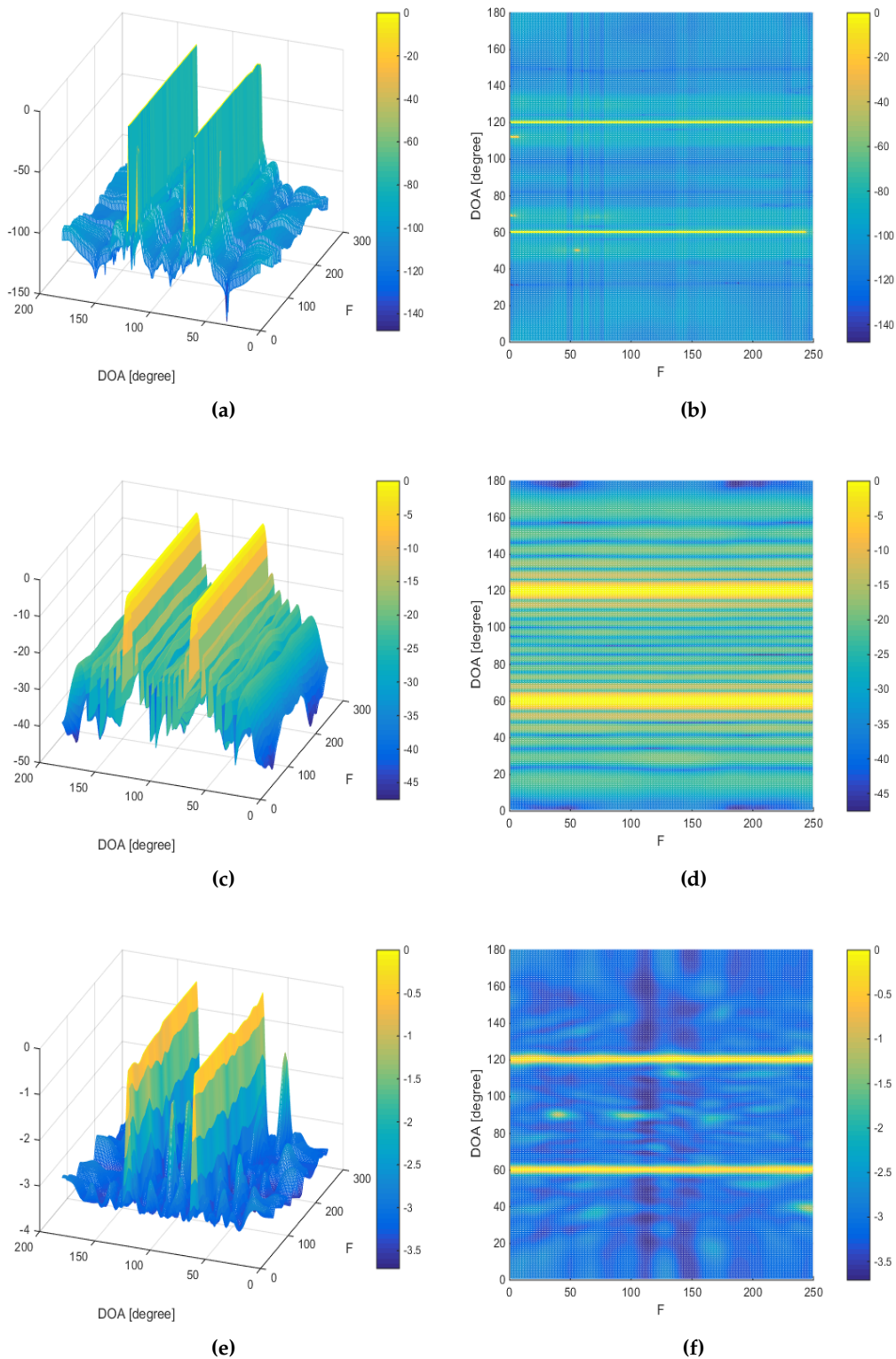


Figure 21. Wideband DOA estimation, two chirp source signals at DOAs 60° and 120°, $\gamma = 0.5$, 10 snapshots, using ULA containing 23 sensors for conventional beamforming and MUSIC algorithm. (a,b) The MVDR A-LASSO algorithm (after the first iteration); (c,d) Conventional beamforming; and (e,f) the MUSIC algorithm.

6. Conclusions

This paper has presented a novel technique using the compressive sensing framework on a sparse linear array. The ordinary least square adaptable LASSO (OLS A-LASSO) has been applied for the first time for the DOA estimation problem. Further, we have proposed a new adaptable LASSO algorithm, the MVDR A-LASSO, for the DOA estimation problem. The proposed A-LASSO algorithm outperforms the classical LASSO, OLS A-LASSO and the classical DOA estimation techniques. It does not require any a priori knowledge about the number of source signals. The proposed algorithm is able to perform DOA estimation using a smaller number of snapshots and is able to estimate correlated source signals and spatially-related source signals. Moreover, even when the number of source signals is known, it outperforms the MUSIC algorithm. Our proposed algorithm can identify $((M^2 - 2)/2 + M - 1)/2$ source signals using M sensors and has a high resolution. Using the proposed technique, we are able to eliminate any spurious peaks and identify only the actual source peaks. Further, it has been shown that, using the proposed MVDR A-LASSO, the source signals can be identified with a lesser number of iterations than that using OLS A-LASSO. Therefore, the computational cost of the MVDR A-LASSO is lower than that of the OLS A-LASSO. Further, the simulations have shown that MVDR A-LASSO is suitable for wideband scenarios, as well. Finally, it is worth mentioning that the proposed scheme for DOA estimation is applicable to other types of array structures, but when the scheme is applied to other structures, it may not provide all of the advantages that it provides to the nested structure.

Acknowledgments: This work is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Regroupement Stratégique en Microélectronique de Québec (ReSMiQ) and a scholarship from the Egyptian Government to the first author.

Author Contributions: The work was realized with the collaboration of all of the authors. All of the authors participated in conceiving of the main ideas of the paper and in the design and implementation of the algorithms contained therein, as well as in organizing and writing the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khmou, Y.; Safi, S.; Frikel, M. Comparative study between several direction of arrival estimation methods. *J. Telecommun. Inf. Technol.* **2014**, *1*, 41–48.
2. Chen, Z.; Gokeda, G.; Yu, Y. Overview of Basic DOA Estimation Algorithms. In *Introduction to Direction-of-Arrival Estimation*; Artech House: Norwood, MA, USA, 2010; pp. 31–64.
3. Capon, J. High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE* **1969**, *57*, 1408–1418.
4. Makhoul, J. Linear prediction: A tutorial review. *Proc. IEEE* **1975**, *63*, 561–580.
5. Ziskind, I.; Wax, M. Maximum likelihood localization of multiple sources by alternating projection. *IEEE Trans. Acoust. Speech Signal Proc.* **1998**, *36*, 1553–1560.
6. Schmidt, R.O. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* **1986**, *34*, 276–280.
7. Roy, R.; Kailath, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Proc.* **1989**, *37*, 984–995.
8. Sidiropoulos, N.D.; Bro, R.; Giannakis, G.B. Parallel factor analysis in sensor array processing. *IEEE Trans. Signal Proc.* **2000**, *48*, 2377–2388.
9. Nion, D.; Sidiropoulos, N.D. A PARAFAC-based technique for detection and localization of multiple targets in a MIMO radar system. In *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, 19–24 April 2009; pp. 2077–2080.
10. Guizzo, E. How Google'S Self-Driving Car Works. Available online: <http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works> (accessed on 16 August 2016).
11. Liang, J. Joint azimuth and elevation direction finding using cumulant. *IEEE Sens. J.* **2009**, *9*, 390–398.

12. Ma, W.-K.; Hsieh, T.-H.; Chi, C.-Y. DOA estimation of quasi-stationary signals via Khatri-Rao subspace. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 2165–2168.
13. Ma, W.-K.; Hsieh, T.-H.; Chi, C.-Y. DOA estimation of quasi-stationary signals with less sensors than sources and unknown spatial noise covariance: A Khatri-Rao subspace approach. *IEEE Trans. Signal Proc.* **2010**, *58*, 2168–2180.
14. Pal, P.; Vaidyanathan, P.P. Nested arrays: A novel approach to array processing with enhanced degrees of freedom. *IEEE Trans. Signal Proc.* **2010**, *58*, 4167–4181.
15. Zhu, C.; Wang, W.-Q.; Chen, H.; So, H.C. Impaired Sensor Diagnosis, Beamforming, and DOA Estimation With Difference Co-Array Processing. *IEEE Sens. J.* **2015**, *15*, 3773–3780.
16. Pillai, S.U.; Bar-Ness, Y.; Haber, F. A new approach to array geometry for improved spatial spectrum estimation. *Proc. IEEE* **1985**, *73*, 1522–1524.
17. Pillai, S.; Haber, F. Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix. *IEEE Trans. Acoust. Speech Signal Proc.* **1987**, *35*, 1517–1523.
18. Moffet, A. Minimum-redundancy linear arrays. *IEEE Trans. Antennas Propag.* **1968**, *16*, 172–175.
19. Abramovich, Y.I.; Gray, D.A.; Gorokhov, A.Y.; Spencer, N.K. Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. Fully augmentable arrays. *IEEE Trans. Signal Proc.* **1998**, *46*, 2458–2471.
20. Zhang, Y.D.; Amin, M.G.; Himed, B. Sparsity-based DOA estimation using co-prime arrays. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 3967–3971.
21. Adhikari, K.; Buck, J.R.; Wage, K.E. Beamforming with extended co-prime sensor arrays. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 4183–4184.
22. Shen, Q.; Liu, W.; Cui, W.; Wu, S.; Zhang, Y.D.; Amin, M.G. Low-complexity direction-of-arrival estimation based on wideband co-prime arrays. *IEEE/ACM Trans. Audio Speech Lang. Proc.* **2015**, *23*, 1445–1456.
23. Vaidyanathan, P.P.; Pal, P. Sparse sensing with co-prime samplers and arrays. *IEEE Trans. Signal Proc.* **2011**, *59*, 573–586.
24. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* **2007**, *58*, 1182–1195.
25. Wagner, N.; Eldar, Y.C.; Friedman, Z. Compressed beamforming in ultrasound imaging. *J. Abbr.* **2012**, *60*, 4643–4657.
26. Herman, M.A.; Strohmer, T. High-resolution radar via compressed sensing. *IEEE Trans. Signal Proc.* **2009**, *57*, 2275–2284.
27. Malioutov, D.; Çetin, M.; Willsky, A.S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Proc.* **2005**, *53*, 3010–3022.
28. Xenaki, A.; Gerstoft, P.; Mosegaard, K. Compressive beamforming. *J. Acoust. Soc. Am.* **2014**, *136*, 260–271.
29. Gerstoft, P.; Xenaki, A.; Mecklenbräuker, C.F. Single and multiple snapshot compressive beamforming. *J. Acoust. Soc. Am.* **2015**, *138*, 2003–2014.
30. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodol.)* **1996**, *58*, 267–288.
31. Zou, H. The adaptive lasso and its oracle properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429.
32. Panahi, A.; Viberg, M. On the resolution of the LASSO-based DOA estimation method. In Proceedings of the 2011 International ITG Workshop on Smart Antennas (WSA), Aachen, Germany, 24–25 February 2011; pp. 1–5.
33. Liu, S.; Trenkler, G. Hadamard, Khatri-Rao, Kronecker and other matrix products. *Int. J. Inf. Syst. Sci.* **2008**, *4*, 160–177.
34. Van Trees, H.L. Adaptive Beamformers. In *Optimum Array Processing, Part IV*; WILEY: New York, NY, USA, 2004; pp. 710–916.
35. Fan, J.; Li, R. Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.* **2001**, *96*, 1348–1360.
36. Chen, S.S.; Donoho, D.L.; Saunders, M.A. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.* **1998**, *20*, 33–61.

37. Efron, B.; Hastie, T.; Johnstone, I.; Tibshirani, R. Least angle regression. *Ann. Stat.* **2004**, *32*, 407–499.
38. Elad, M. The Quest for a Dictionary. In *Sparse and Redundant Representations*; Editor, F., Meditor, A., Eds.; Springer: New York, NY, USA, 2010; pp. 227–246.
39. Sinha, P.; George, A.D.; Kim, K. Parallel algorithms for robust broadband MVDR beamforming. *J. Comput. Acoust.* **2002**, *10*, 69–96.
40. Balakrishnan, S.; Ong, L.T. GNU radio based digital beamforming system: BER and computational performance analysis. In Proceedings of the 2015 23rd European Signal Processing Conference (EUSIPCO), Nice, France, 31 August–4 September 2015; pp. 1596–1600.
41. Lavate, T. B.; Kokate, V.K.; Sapkal, A.M. Performance analysis of MUSIC and ESPRIT DOA estimation algorithms for adaptive array smart antenna in mobile communication. In Proceedings of the 2010 2nd International Conference on Computer and Network Technology (ICCNT), Bangkok, Thailand, 23–25 April 2010; pp. 308–311.
42. Priyadarshini, M.P.; Vinutha, R. Comparative performance analysis of MUSIC and ESPRIT on ULA. In Proceedings of the 2012 International Conference on Radar, Communication and Computing (ICRCC), Tiruvannamalai, India, 21–22 December 2012; pp. 120–124.
43. Mujahid, U.; Mukhtar, M.; Ahmed, J.; Rehman, A.; Abbas, M.; Shahid, U. Spectral estimation for smart antenna system. In Proceedings of the 2013 3rd International Conference on Computer, Control & Communication (IC4), Karachi, Pakistan, 25–26 September 2013; pp. 1–5.
44. Wax, M.; Shan, T.-J.; Kailath, T. Spatio-Temporal Spectral Analysis by Eigenstructure Methods. *IEEE Trans. Acoust. Speech Signal Proc.* **1984**, *32*, 817–827.
45. Yamada, H.; Shirai, S.; Yamaguchi, Y. DOA estimation of wideband signals using extended virtual array. In Proceedings of the 2014 3rd Asia-Pacific Conference on Antennas and Propagation (APCAP), Harbin, China, 26–29 July 2014; pp. 802–805.
46. Morozov, V.A. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.* **1966**, *7*, 414–417.
47. Rust, B.W. Truncated Singular Value Decomposition Solution to Discrete Ill-Posed Problems with Ill-determined Numerical Rank. *SIAM J. Sci. Stat. Comput.* **1990**, *11*, 503–518.
48. Vogel, C.R. Regularization Parameter Selection Methods. In *Computational Methods for Inverse Problems*; SIAM: Philadelphia, PA, USA, 2002; pp. 97–127.
49. Wahba, G. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* **1977**, *14*, 651–667.
50. Hansen, P.C.; Jensen, T.K.; Rodriguez, G. An adaptive pruning algorithm for the discrete L-curve criterion. *J. Comput. Appl. Math.* **2007**, *198*, 483–492.
51. Hansen, P.C.; O’Leary, D.P. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.* **1993**, *14*, 1487–1503.
52. Leming, Q.; Partha, R. Determination of regularization parameter using L-curve by the LARS-LASSO algorithm. In Proceedings of the 2006 Joint Statistical Meetings, Seattle, WA, USA, 6–10 August 2006.
53. Grant, M.; Boyd, S. CVX: Matlab Software for Disciplined Convex Programming, Version 2.1. Available online: <http://cvxr.com/cvx> (accessed on 20 March 2014).
54. Grant, M.; Boyd, S. Graph implementations for nonsmooth convex programs. In *Recent Advances in Learning and Control*; Blondel, V., Boyd, S., Kimura, H., Eds.; Lecture Notes in Control and Information Sciences; Springer: Heidelberg, Germany, 2008; pp. 95–110.
55. Tütüncü, R.H.; Toh, K.C.; Todd, M.J. Solving semidefinite-quadratic-linear programs using SDPT3. *Math. Program.* **2003**, *95*, 189–217.

