Direction of Arrival Estimation using Compressive Sensing

Raktim Pal, Vikas Sakaray and P Palanisamy

Abstract

In this manuscript, we consider the problem of Direction of Arrival (DOA) estimation of signals using Compressive Sensing (CS). We propose a novel and systematic approach to DOA estimation using a rectangular array by exploiting sparsity in the angular domain. With the help of simulations, the proposed approach is compared with the widely used MUSIC algorithm and its advantages over MUSIC are elaborated. We also extend our proposed model to a generalized 2D array and analyze its performance with an L-shaped array-a specific example of 2D array. Using our proposed model we can greatly reduce the hardware as well as software complexity while still maintaining the desired high resolution.

1 Introduction

Antenna arrays find extensive applications in many fields such as sonar, radar, communications and so on. The direction of arrival estimation of signals impinging on an array of sensors is a fundamental problem in array signal processing and many estimation methods have already been proposed so far [1]. Uniform linear array (ULA) is one such class of arrays which has been extensively studied, but one major disadvantage of ULA is that it can not be used for estimating both the azimuthal and elevation angles of the source. On the contrary, rectangular arrays can be efficiently used to estimate both the azimuthal and elevation angles to accurately define the angular location of the source. To this end, well established methods like MUSIC and ESPRIT have already been used for DOA estimation using rectangular arrays with satisfactory estimation accuracy. On the other hand, compressive sensing [2], [3], [4], an emerging innovative procedure of sensing and acquiring signals provides us with an alternative method for DOA estimation [5], [6]. In this work, we are mainly interested in the direction of arrival estimation of signals impinging on a rectangular array using compressive sensing. CS allows us to take considerably fewer measurements compared to the standard methods, while still maintaining high resolution of estimated parameters. This greatly reduces the hardware and software complexities: the hardware complexity is greatly reduced because of a much smaller number of front end circuit chains, and the software (estimation algorithms) complexity is also reduced because of a smaller dimension of the array data. These are the principal reasons behind the selection of this approach for addressing the DOA problem, but as we go along we will discover other advantages of this method as well. The proposed method is analysed by carrying out simulations in MATLAB under different conditions: in the presence of different levels of noise, spatially correlated or uncorrelated noise and so on, and a comparison with MUSIC is carried out. Furthermore, the model is investigated for its compatibility with different array geometries such as the L-shaped array, and identify if there is a scope to extend it to a generalised 2D array.

The rest of the paper is organized as follows: Section 2 describes the data model. Section 3 explains the compressive sensing framework. Section 4 contains a brief account of the M-FOCUSS algorithm [7] that we have used in our work. Section 5 depicts all the simulation results and a few important discussions. In section 6 we extend our proposed model to an L shaped array and eventually to a generalized 2D array. Finally, section 7 concludes the paper.
2 Data Model

Let the rectangular array has $M_x$ elements along the x-axis and $M_y$ elements along the y-axis. The total number of elements in the array is therefore $M_x M_y$. The inter sensor spacing along the x and y axes are $\Delta_x$ and $\Delta_y$ respectively. The number of uncorrelated sources is assumed to be known (can be determined using Akaike Information Criterion) and is equal to $d$. The $i^{th}$ source has angular location defined by $\theta^{(i)}$ and $\phi^{(i)}$, which are the polar and azimuthal angles respectively. The sources are assumed to be emitting signals at the frequency $c\lambda$.

Each array element is denoted by $(k_x, k_y)$ and $x_{k_x, k_y}(t)$ is the signal sample received by the array element $(k_x, k_y)$ at time instant $t$. At each time instant there are $M_x M_y$ samples for the whole rectangular array.

Let $\chi(t) \in C^{M_x \times M_y}$ represent the matrix containing the received signal values at the sensors in the rectangular array at time instant $t$. $\chi(t)$ can be expressed in terms of the signal values from the $d$ sources as:

$$\chi(t) = \sum_{i=1}^{d} a(\mu_i) a^T(\nu_i) s_i(t) + N(t) \quad (1)$$

where $N(t)$ is the additive noise. $a(\mu_i) \in C^{M_x}$ and $a(\nu_i) \in C^{M_y}$ are defined as:

$$a(\mu_i) = \begin{bmatrix} e^{-j\mu_i} & e^{-j2\mu_i} & \ldots & e^{-j(M_x-1)\mu_i} \end{bmatrix}^T$$

$$a(\nu_i) = \begin{bmatrix} e^{-j\nu_i} & e^{-j2\nu_i} & \ldots & e^{-j(M_y-1)\nu_i} \end{bmatrix}^T$$

These two vectors can be interpreted as array steering vectors of ULA of size $M_x$ and $M_y$ respectively. The expressions for $\mu_i$ and $\nu_i$ are given by:

$$\mu_i = \frac{2\pi}{\lambda} \Delta_x u_i \quad \text{where} \quad u_i = \cos \phi^{(i)} \sin \theta^{(i)}$$

$$\nu_i = \frac{2\pi}{\lambda} \Delta_y v_i \quad \text{where} \quad v_i = \sin \phi^{(i)} \sin \theta^{(i)}$$

For a one dimensional array the data model equation is:

$$X = AS + N \quad (2)$$

where $A$ is the array steering matrix. For a single time snapshot, $X$, $N$ and $S$ are vectors. The DOA estimation algorithms are defined and established based on this data model. However, from equation (1) we can see that $\chi(t)$ is a matrix of size $M_x \times M_y$. Hence, the data model equation is not conformal to the standard DOA data model. In order to make it conformal to the standard data model, it is necessary to map the matrix to an appropriate vector. We define a $vec\{\cdot\}$ operation which maps an $M_x \times M_y$ dimensional matrix to a column vector of size $M_x \times M_y$. In this operation, each column of the matrix starting from the left is selected and stacked together below the previous column to form a vector. $vec\{\cdot\}$ operation can be used on the data matrix $\chi(t)$ to form a data vector $x$ of length $M_x \times M_y$.

$$x(t) = vec\{\chi(t)\}$$

If $x(t)$ is used in equation (1) instead of $\chi(t)$, the left hand side of the equation becomes a column
vector; and therefore appropriate changes need to be made to the right hand side to satisfy the
dimensionality of the equation. We define \( A(\mu_i, \nu_i) \), the steering matrix corresponding to the \( i^{th} \) source as:

\[
A(\mu_i, \nu_i) = a(\mu_i) a^T(\nu_i) \\
A(\mu_i, \nu_i) \in \mathbb{C}^{M_x \times M_y}
\]

We apply the \( \text{vec}\{\cdot\} \) operation on it and convert it to a vector \( a(\mu_i, \nu_i) \) of length \( M_x \times M_y \).

\[
a(\mu_i, \nu_i) = \text{vec}\{A(\mu_i, \nu_i)\}
\]

(3)

Using the \( a(\mu_i, \nu_i) \) corresponding to each of the \( d \) sources, we can define the steering matrix \( A \) for
the array as:

\[
A = [a(\mu_1, \nu_1) \ a(\mu_2, \nu_2) \ \cdots \ a(\mu_d, \nu_d)]
\]

In equation (1), the noise \( N(t) \) is a matrix and we can apply the same \( \text{vec}\{\cdot\} \) operator to convert it
into a noise vector \( w(t) \) of dimension \( M_x \times M_y \). Converting all the necessary matrices in equation
(1) to vectors, we get an equation that is conformal to the single dimensional data model represented
in equation (2). Let us assume that we take \( T \) snapshots of the incoming signals. Then the received
data matrix can be represented as:

\[
X = [x(t_1) \ x(t_2) \ \cdots \ x(t_T)]
\]

The noise matrix for \( T \) snapshots is represented as:

\[
W = [w(t_1) \ w(t_2) \ \cdots \ w(t_T)]
\]

Since, there are \( D \) sources, the signal vector for a single time instant is :

\[
s(t) = [s_1(t) \ s_2(t) \ \cdots \ s_d(t)]^T
\]

The signal matrix for \( T \) time instants is:

\[
S = [s(t_1) \ s(t_2) \ \cdots \ s(t_T)]
\]

From the above definitions, the DOA model equation for the rectangular array can finally be written
as:

\[
X = AS + W
\]

(4)

The equation is of the same nature as equation (2) with the additional extension to multiple time
snapshots.

### 3 Compressive Sensing Framework

Before we cast the DOA problem as a compressive sensing problem, we briefly discuss the compressive
sensing framework. Let \( x \in \mathbb{C}^N \) be an \( N \times 1 \) signal vector which can be expressed as \( x = \Psi f \), where
\( \Psi \) is an \( N \times N \) sparsity basis matrix; and \( f \) is an \( N \times 1 \) sparse vector, \( k \) elements of which are nonzero.
Compressed sensing theory says that \( x \) can be recovered from \( M = K\mathcal{O}(\log n) \) non adaptive linear
projection measurements onto an \( M \times N \) matrix \( \Xi \) that is incoherent with \( \Psi \) and satisfies Restricted
Isometry Property (RIP). Random matrices are the class of matrices that satisfy RIP, incoherence and
\( \Xi \) can therefore be modelled as a random matrix whose elements can be chosen independently from
a Gaussian or Bernoulli distribution. The measurement vector \( y \) can be written as:

\[
y = \Xi x = \Xi \Psi f
\]

3
\( f \) can be efficiently recovered from the measurements \( y \) using various compressive sensing reconstruction methods.

In order to cast this DOA problem as a compressive sensing problem, we need to design a sparse basis for the data model. For this purpose, we can exploit the sparsity in the angular location of the sources and thereby develop a sparse representation of the signal as described below:

For the polar angle \( \theta \in [0, 2\pi) \) we take a finite set consisting of \( N_1 \) angles such that:

\[
\Theta = \{ \theta_1, \theta_2, \ldots, \theta_{N_1} \}
\]

Similarly, we take a finite set of \( N_2 \) angles for the azimuthal angle \( \phi \in [0, \pi) \) such that:

\[
\Phi = \{ \phi_1, \phi_2, \ldots, \phi_{N_2} \}
\]

For each combination of \((\theta_i, \phi_j)\) from the set \( \Theta \times \Phi \), we can define the steering vector of the rectangular array along the same lines as equation (3):

\[
a(\theta_i, \phi_j) = a(\mu_{i,j}, \nu_{i,j})
\]

where \( \mu_{i,j} \) and \( \nu_{i,j} \) are defined as:

\[
\mu_{i,j} = \frac{2\pi}{\lambda} \Delta x \cos \phi_j \sin \theta_i \\
\nu_{i,j} = \frac{2\pi}{\lambda} \Delta y \sin \phi_j \sin \theta_i
\]

Here, \( \mu_{i,j} \) and \( \nu_{i,j} \) play the same role as \( \mu_i \) and \( \nu_i \) play in equation (3). Now, we can define a sparsifying matrix \( \Psi \) as:

\[
\Psi = [a(\theta_1, \phi_1) \cdots a(\theta_1, \phi_{N_2})] \\
[ a(\theta_2, \phi_1) \cdots a(\theta_2, \phi_{N_2}) ] \\
\cdots \cdots \\
[ a(\theta_{N_1}, \phi_1) \cdots a(\theta_{N_1}, \phi_{N_2}) ]
\]

We also define an \( N_1 N_2 \times 1 \) sparse vector \( f(t_n) \)

\[
f(t_n) = \begin{bmatrix}
    f_{\theta_1, \phi_1}(t_n) \\
    \vdots \\
    f_{\theta_{i_1}, \phi_{j_1}}(t_n) \\
    f_{\theta_2, \phi_1}(t_n) \\
    \vdots \\
    f_{\theta_{i_2}, \phi_{j_2}}(t_n) \\
    \vdots \\
    f_{\theta_{i_d}, \phi_{j_d}}(t_n) \\
    \vdots \\
    f_{\theta_{N_1}, \phi_1}(t_n) \\
    \vdots \\
    f_{\theta_{N_1}, \phi_{N_2}}(t_n)
\end{bmatrix}
\]

If a source \( k \) has the angle pair \((\theta_i, \phi_j)\) then \( f_{\theta_i, \phi_j}(t_n) \) takes the value \( s_k(t_n) \) (the signal from the \( k^{th} \) source). Hence, if there are \( d \) sources, only \( d \) elements out of \( N_1 \times N_2 \) elements are nonzero. Rest
of the $N_1 \times N_2 - d$ elements are equal to zero. For $T$ snapshots, the matrix $F$ of sparse vectors is constructed below:

$$F = [f(t_1) \ f(t_2) \ \cdots \ f(t_T)]$$

$f(t_1), f(t_2), \ldots, f(t_T)$ are jointly sparse since they have the non-zero coefficients at the same positions. Hence, $\Psi$ acts as a sparse basis for the received signals and $X$ can therefore be rewritten as:

$$X = \Psi F + W \quad (5)$$

The measured data $Y$ is obtained by multiplying the received signal matrix $X$ with a random matrix $\Xi$ of dimension $M \times MxMy$, where $M$ is the number of compressive measurements that we intend to take.

$$Y = \Xi X$$

Using equation (5) we can rewrite,

$$Y = \Xi \Psi F + \Xi W \quad (6)$$

and $Y$ is an $M \times T$ matrix:

$$Y = [y(t_1) \ y(t_2) \ \cdots \ y(t_T)]$$

We can exploit the joint sparsity of $F$ to recover it from the measurements $Y$ using joint compressive sensing recovery methods such as $M$-FOCUSS - details of which are given in the next section. Let the recovered $F$ be given by:

$$\hat{F} = [\hat{f}(t_1) \ \hat{f}(t_2) \ \cdots \ \hat{f}(t_T)]$$

The angle spectrum function is defined using the recovered $\hat{F}$ as given below:

$$P_y(\theta, \phi) = \frac{1}{T} \sum_{n=1}^{T} |\hat{f}_{\theta,\phi}(t_n)|^2 \quad \text{for} \ \theta \in \Theta \ \text{and} \ \phi \in \Phi \quad (7)$$

If there is a signal source at $(\theta_i, \phi_j)$, then the recovered $\hat{f}_{\theta_i,\phi_j}(t_n)$ will be non zero. Hence, the time average of squared $\hat{f}_{\theta_i,\phi_j}(t_n)$ will result in a non zero value i.e there will be a peak in the angle spectrum. Therefore, peaks in angle spectrum correspond to angular locations of the signal sources.

## 4 M-FOCUSS

As stated previously, we have made use of the M-FOCUSS algorithm for estimating the sparse vectors $\hat{F}$ from the measurements $Y$. It is an iterative algorithm for obtaining sparse solutions to linear inverse problems with multiple measurement vectors (MMV). A brief description of the algorithm is provided in this section. The noiseless MMV problem can be stated as solving the following $L$ undetermined system of equations:

$$Ax^{(l)} = b^{(l)}, \quad l = 1, \cdots, L \quad (8)$$

where $A \in C^{m \times n}$, $m < n$ and often $m \ll n$. It is assumed that $A$ has full row rank. $L$ is the number of measurement vectors and it is usually assumed that $L < m$. Vectors $b^{(l)} \in C^{m}$, $l = 1, \cdots, L$ are the measurement vectors and $x^{(l)} \in C^{n}$, $l = 1, \cdots, L$ are the corresponding source vectors. We can rewrite (8) as:

$$AX = B \quad (9)$$
where $X = [x^{(1)}, \ldots, x^{(L)}]$, and $B = [b^{(1)}, \ldots, b^{(L)}]$. The important assumptions about the desired solution are: $x^{(l)}, l = 1, \ldots, L$ are sparse, i.e. most of the entries are zero; and all the $L$ vectors are assumed to have the same sparsity profile so that the indices of the nonzero entries are independent of $l$. In order to incorporate noise into the model, equation (9) can be modified and written as:

$$AX + N = B$$

where $N = [n^{(1)}, \ldots, n^{(L)}]$. $n^{(l)} \in \mathbb{C}^m$ represent the additive noise. In the presence of noise, an additional complicating factor one has to consider is the trade off between quality of fit e.g., as measured by $\|AX - B\|$ and the sparsity of the solution. The diversity measure that has been used to find sparse solution is:

$$J^{(p,q)}(X) = \sum_{i=1}^{n} (||x[i]||_q)^p, \quad 0 \leq p \leq 1, q \geq 1$$  \hspace{1cm} (10)

where $x[i] = [x^{(1)}[i], x^{(2)}[i], \ldots, x^{(L)}[i]]$ is the $i$th row of $X$, and the row norm is given by $||x[i]||_q = \left(\sum_{l=1}^{L} |x^{(l)}[i]|^q\right)^{1/q}$. Minimization of this diversity measure leads to sparse solutions. For simplicity, we consider $q = 2$ and denote $J^{(p,2)}(X)$ by $J^{(p)}(X)$ i.e.,

$$J^{(p)}(X) = \sum_{i=1}^{n} (||x[i]||_2)^p = \sum_{i=1}^{n} \left(\sum_{l=1}^{L} |x^{(l)}[i]|^2\right)^{p/2}$$  \hspace{1cm} (11)

The minimization of the $J^{(p)}(X)$ measure (11) has been found to lead to a low complexity computational algorithm. The details of the algorithm are summarized as follows:

$$W_{k+1} = \text{diag} \left(c_k[i]^{1-p/2}\right)$$

where $c_k[i] = \left(\sum_{l=1}^{L} (x^{(l)}[i])^2\right)^{1/2}, \quad p \in [0, 2]$

$$Q_{k+1} = A_{k+1}^H (A_{k+1} A_{k+1}^H + \lambda I)^{-1} B$$

where $A_{k+1} = AW_{k+1}$ with $\lambda \geq 0$

$$X_{k+1} = W_{k+1} Q_{k+1}$$

The algorithm is terminated once a convergence criterion has been satisfied e.g.,

$$\frac{\|X_{k+1} - X_k\|_F}{\|X_k\|_F} < \delta$$

where $\delta$ is a user selected parameter. This algorithm can be proven to reduce $J^{(p,q)}(X)$ in each iteration. The parameters in the regularized M-FOCUSS algorithm impacting its performance are the regularization parameter $\lambda$, the parameter $p$ and the initial condition. The modified l-curve method explained in [8],[9] can be used to choose the regularization parameter. The choice of $p$ is dictated by the speed of convergence and the sparsity of the solution generated. In practice, values of $p$ between 0.8 and 1.0 have been found to represent a good compromise between speed of convergence and quality of the generated sparse solution. If no good starting points are available, then the minimum Frobenius norm solution is a good initializer. However, convergence to a solution was fastest when the minimum norm solution was used as the starting point.
5 Simulation Results and Comparison

For simulating the proposed model, we take \(d = 2\) signal sources with angular positions given by \(\theta^{(1)} = 30, \phi^{(1)} = 20\) and \(\theta^{(2)} = 40, \phi^{(2)} = 70\) respectively. The signal power of the sources and the noise power at the sensor elements can be specified depending on the Signal to Noise Ratio. We model a 100 element rectangular array with 10 sensors along each of the x and y directions \((M_x = 10, M_y = 10)\). We take \(T = 15\) snapshots and \(M = 20\) compressive measurements. The random matrix \(\Xi\) has entries drawn from a Gaussian distribution \((\text{mean}=0, \text{variance}=1)\). The joint CS recovery algorithm used is regularized M-FOCUSS with \(p = 0.8\). The angle spectrum as defined by equation (7) is plotted in Figure 1 and SNR in this case is 0 dB. The two peaks correspond to the two sources. The data points in the figure are well defined with the data tips representing the \(\theta, \phi\) and \(P_y(\theta, \phi)\) values as x, y and z respectively. Hence, it can be seen that our proposed method is able to correctly estimate the location of the two sources.

![Figure 1: Angle Spectrum](image)

Let us take two closely situated signal sources described by angular locations \(\theta^{(1)} = 30, \phi^{(1)} = 50\) and \(\theta^{(2)} = 31, \phi^{(2)} = 51\). The plotted angle spectrum spectrum in Figure 2 clearly shows that our compressive sensing approach is able to differentiate between the angular location of the two sources even though they are very closely situated.

We are now interested in analyzing the performance of compressed sensing approach of direction of arrival estimation in the presence of correlated noise. We assume that correlation in the time domain can be overcome by tapping data at optimum time instances. Therefore, we only consider correlation in the spatial domain. Since we are measuring data using a rectangular array of sensors, we require a two dimensional spatial correlation model. The modified Gudmundson’s model has been adopted for this purpose.

Let \(X\) be a Gaussian distributed space-time correlated random field. The correlation is given by the following equation:

\[
E[X(x, y)X(x + \delta_x, y + \delta_y)] = e^{-\frac{d_x^2 + d_y^2}{2d_{corr}^2}}
\]

where \(d_{corr}\) is the decorrelation distance. Now that we have generated the spatially correlated noise, we can use it to assess the performance of the CS DOA algorithm. For simulation purpose, we take the same angular position of the sources as that used in the first simulation (Figure 1) and other parameters are also kept the same; with the only exception that we add spatially correlated noise to
the model instead of uncorrelated noise. But, Figure 3 distinctly shows that practically no difference is made to the performance of the algorithm in the presence of correlated noise when compared to Figure 1, where uncorrelated noise is used. Hence, CS DOA is found to be immune to the presence of spatially correlated noise. The result is not very hard to understand because MFOCUSS does not impose any kind of restriction on the properties of the noise model used.

5.1 Comparison with MUSIC

In this section we compare our proposed method with the conventional MUSIC method and identify the advantages of CS DOA method over MUSIC. For simulating the MUSIC method, we use an identical environment as that used for CS DOA, described by similar signal source locations, time snapshots, size of rectangular array and all other relevant factors. Figure 4 illustrates the results of the MUSIC method with the two highest peaks corresponding to the angular location of the two sources. A direct comparison of Figure 1 and Figure 4 implies that CS DOA method gives sharp and accurate spikes corresponding to the angular location of the sources as opposed to the bell shaped plots.
obtained from MUSIC method. Hence, CS DOA turns out to be a more accurate DOA estimation method compared to MUSIC.

We are also interested in comparing the resolving capability of the two methods. We take the same arrangement that we used for plotting Figure 2 and use the MUSIC algorithm. Figure 5 illustrates the results obtained with MUSIC algorithm and it clearly shows four closely spaced distinct spikes of almost the same height. Hence, MUSIC finds it difficult to distinctly resolve two closely located sources, which CS DOA can do very efficiently.

Figure 4: Angle Spectrum obtained using MUSIC on a rectangular array

Figure 5: Angle Spectrum obtained using MUSIC for two closely situated sources
In the previous section we studied the performance of CS DOA in the presence of spatially corre-
lated noise and found that it works perfectly irrespective of the noise being spatially correlated or not.
On the contrary, it is well known that MUSIC algorithm fails miserably in the presence of spatially
correlated noise and it can also be observed from Figure 6. A direct comparison of Figure 3 and
Figure 6 implies that CS DOA is a better method than MUSIC as far as performance in the presence
of correlated noise is concerned.

5.2 Reduction in Hardware and Software Complexity: A discussion

In this subsection we will try to investigate the capability of our proposed model to reduce the
hardware as well as software complexity compared to other models which do not involve compressive
sensing. For ease of understanding, we will take the same sensor arrangement that we previously used
for simulation.

A typical receiver circuit that is attached to each sensor element consists of low noise amplifiers, anti
aliasing filters and analog to digital converter. A $10 \times 10$ rectangular array using a non compressive
DOA estimation technique uses 100 such circuits. On the other hand, in our proposed model the
analog signals received at the sensor elements is multiplied with a random matrix and this reduces the
dimensionality of the received signal. In the simulations, we multiplied the received signal vector of
length 100 with a random matrix of dimension $20 \times 100$ thus reducing the signal dimension from 100
to 20. Therefore, we need only 20 such receiving circuits corresponding to 20 measurements. Hence, it
can be clearly understood that CS DOA results in a great amount of savings in hardware requirements
when compared to conventional DOA estimation methods.

The same example can be used to explain the ability of CS DOA to reduce the software complexity.
By multiplying with a random matrix, we reduce the dimension of the resulting measurements and
hence the matrices that we have to deal with in the estimation algorithm are comparatively smaller
in size, resulting in lesser computations and therefore reduced software complexity.

6  Extension to an L-Shaped Array and generalized 2D array

An L shaped array can be assumed to be a special case of the rectangular sensor arrangement, with
sensors arranged only along the x and y axes. Hence, the DOA estimation framework that we have
developed so far can be applied to an L shaped array after proper modifications to the originally
developed equations. Let us assume that the L-shaped array has $M_x$ and $M_y$ sensors along the x and
y axis respectively. Following up from the conventions that we have defined for a rectangular array, we can represent the signal \( \chi_l(t_n) \) received by the L shaped array as:

\[
\chi_l(t_n) = \begin{bmatrix}
x_{1.1}(t_n) & x_{1.2}(t_n) & x_{1.3}(t_n) & \cdots & x_{1.M_y}(t_n) \\
x_{2.1}(t_n) & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x_{M_x.1}(t_n) & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

We can write down the vector form of the signal matrix as \( \mathbf{x}_l(t_n) \) as defined below.

\[
\mathbf{x}_l(t_n) = \text{vec}\{\chi_l(t_n)\}
\]

The steering matrix for an L shaped array corresponding to a single source \( i \) can be written as:

\[
A_{l(i)} = \begin{bmatrix}
1 & e^{-j\nu_i} & e^{-j2\nu_i} & \cdots & e^{-j(M_y-1)\nu_i} \\
e^{-j\mu_i} & 0 & 0 & \cdots & 0 \\
e^{-j2\mu_i} & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
e^{-j(M_x-1)\mu_i} & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

The vector form of the steering matrix is:

\[
\mathbf{a}_l(\mu_i, \nu_i) = \text{vec}\{A_{l(i)}\}
\]

Taking into account all the \( d \) sources, the steering matrix for the array can be expressed as:

\[
A_l = \begin{bmatrix}
\mathbf{a}_l(\mu_1, \nu_1) & \mathbf{a}_l(\mu_2, \nu_2) & \cdots & \mathbf{a}_l(\mu_d, \nu_d)
\end{bmatrix}
\]

The sparse basis matrix \( \Psi \) for an L-shaped array can be defined in a way similar to that of a rectangular array.

### 6.1 Simulation

We take two sources located at \( \theta^{(1)} = 30, \phi^{(1)} = 50 \) and \( \theta^{(2)} = 40, \phi^{(2)} = 80 \) and run the proposed DOA estimation algorithm for the L shaped array. The plotted angle spectrum in Figure 7 distinctly shows that the L shaped array can accurately estimate the angular location of the two signal sources.

From the L shaped array discussion, we get the idea that any two dimensional array of arbitrary shape can be assumed to be a special case of the rectangular array. The data model and the CS DOA model for such an array can be suitably designed as it has been done for an L shaped array. Hence, the compressed sensing DOA approach using a rectangular array gives us the freedom to use the same framework for different array geometries. In this context, a significant advantage of the CS DOA method can be identified i.e. if a few sensors of the rectangular array fail, the sensor bed can still be used for DOA estimation after appropriate changes to the CS DOA model in accordance with the position of failed sensors. This would imply robustness of the framework in the event of sensor failures.

### 7 Conclusion

Uniform linear arrays (ULAs) are one of the most widely used class of arrays for Direction of Arrival estimation and have been extensively studied. However, one major disadvantage of ULA is that it cannot be used for estimating both the azimuthal and elevation angles of the source.
In this work, the DOA problem was studied for rectangular arrays, which can be efficiently used to estimate both the azimuthal and elevation angles to accurately define the angular location of the source, thereby overcoming the limitations of ULAs. Compressive Sensing, an emerging innovative procedure of sensing and acquiring signals, was adopted for DOA estimation. The principal motivation this is the reduction in the number of measurements in comparison to conventional methods such as MUSIC and ESPRIT.

The proposed method was analysed by carrying out simulations in MATLAB under different conditions and a comparison between M-FOCUSS and MUSIC was carried out. M-FOCUSS distinctly provided the location of the sources, even when they are closely situated, but MUSIC failed in such a situation. While MUSIC fails to give suitable estimates in the presence of spatially correlated noise, M-FOCUSS is immune to the presence of spatially correlated noise, since it does not impose any conditions on the noise model used. An expression for Cramer Rao Lower Bound was also derived for the proposed model with a single signal source to get the minimum variance of the estimated angles (See Appendix). RMSE of the estimated angles can be compared to the Cramer Rao lower bound to determine the accuracy of the estimates.

The model was investigated for its compatibility with different array geometries such as the L-shaped array. The simulation results show that an L-Shaped array, after some modification to the data and CS DOA models, can be used to accurately estimate the angular location of sources. For a fixed side length of the array, the rectangular array gives a more accurate estimate than the L-Shaped geometry since it has greater number of sensors. However, there is a significant reduction in number of sensors and computation in the case of L-Shaped array compared to rectangular array. Thus, there is a trade-off between number of sensors and accuracy of estimation.

The success of the proposed method with a L-Shaped array led us to extend the model to a generalised two dimensional array since there is no constraint or assumption as far as the geometry of the array is concerned. The data and the CS DOA models can be modified suitably as it has been done for an L shaped array. In other words, any two dimensional array of arbitrary shape can be assumed to be a special case of the rectangular array. Hence, the compressed sensing DOA approach gives us the flexibility and freedom to use the same framework for different array geometries. This would also imply robustness of the model in the event of sensor failures. By modifying the data and CS DOA models suitably, the same array can be used for estimation purposes.
8 Appendix

8.1 Derivation of Cramer Rao Lower Bound (CRLB)

DOA is actually a parameter estimation problem from the received data. The minimum variance of this estimate is given by the Cramer Rao Lower Bound. CRLB indicates the best we can achieve in estimating the parameters from noisy observations.

CRB Theory: Given a length L vector of received signals $x$ dependent on a set of $N$ parameters $\Theta = [\theta_1, \theta_2, \cdots, \theta_N]^T$, corrupted by additive noise,

$$x = v(\Theta) + n$$

where $v(\Theta)$ is a known function of the parameters; then the variance of an unbiased estimate of the $p$-th parameter $\theta_p$ is greater than the Cramer Rao bound

$$\text{var}(\theta_p) \geq J^{-1}_{pp}$$

where $J^{-1}_{pp}$ is the $p$-th diagonal entry of the inverse of the Fischer information matrix $J$ whose $(i,j)^{th}$ element is given by

$$J_{ij} = -E \left\{ \frac{\partial^2 \ln f_X(x|\Theta)}{\partial \theta_i \partial \theta_j} \right\}$$

where $f_X(x|\Theta)$ is the probability density function of the received signal given the parameter $\Theta$. Since the CRLB depends on the inverse of the Fischer information matrix, no parameter can be ignored and hence $\Theta$ should include all the parameters in the model. For the DOA problem with a single source, the data model can be constructed as follows:

$$x = \alpha s(\theta, \phi) + n$$

where $\alpha = ae^{jb}$ and $s(\theta, \phi)$ is the steering vector of the signal whose directions $\theta, \phi$ we are attempting to estimate. The noise vector $n$ is zero mean gaussian with covariance $\sigma^2 I$. Here $a, b, \theta, \phi$ are modeled as unknown but deterministic constants. Hence, the parameter set $\Theta$ consists of all four unknown parameters.

$$\Theta = [a, b, \theta, \phi]^T$$

$$v(\Theta) = \alpha s(\theta, \phi)$$

The pdf of the received signal vector given $\Theta$ can be found out to be:

$$f_x(x|\Theta) = Ce^{-(x-v)^H R^{-1} (x-v)}$$

where $R = \sigma^2 I$ and $C$ is a normalization constant.

$$\ln f_x(x|\Theta) = \ln C - \frac{(x-v)^H (x-v)}{\sigma^2}$$

$$= \ln C + \frac{-x^H x + v^H x + x^H v - v^H v}{\sigma^2}$$

$$= \ln C + \frac{-x^H x + \alpha s^H (\theta, \phi)x + \alpha x^H s(\theta, \phi) - v^H v}{\sigma^2}$$
Note that the first two terms in the final expression are constant with respect to the parameters $\theta_i$. According to equation (14) we are interested in taking derivatives of this expression with respect to $\theta'_i$s; hence the first two terms can be ignored and the final expression can be rewritten as a function of $\Theta$:

$$g(\Theta) = \frac{1}{\sigma^2} [ae^{-jb} s^H(\theta, \phi)x + ae^{jb} s^H s(\theta, \phi) - a^2 s^H(\theta, \phi)s(\theta, \phi)]$$  (15)

In order to exploit certain special properties of the steering vector $s(\theta, \phi)$, it has been constructed with the phase reference at the center of the rectangular array and the number of elements in the array has been assumed to be odd[10].

Using equation (14) and (15), the Fischer information matrix is constructed below.

$$J = -E \begin{bmatrix}
\frac{\partial^2 g}{\partial a^2} & \frac{\partial^2 g}{\partial a \partial b} & \frac{\partial^2 g}{\partial a \partial \theta} & \frac{\partial^2 g}{\partial a \partial \phi} \\
\frac{\partial^2 g}{\partial a \partial b} & \frac{\partial^2 g}{\partial b^2} & \frac{\partial^2 g}{\partial b \partial \theta} & \frac{\partial^2 g}{\partial b \partial \phi} \\
\frac{\partial^2 g}{\partial a \partial \theta} & \frac{\partial^2 g}{\partial b \partial \theta} & \frac{\partial^2 g}{\partial \theta^2} & \frac{\partial^2 g}{\partial \theta \partial \phi} \\
\frac{\partial^2 g}{\partial a \partial \phi} & \frac{\partial^2 g}{\partial b \partial \phi} & \frac{\partial^2 g}{\partial \theta \partial \phi} & \frac{\partial^2 g}{\partial \phi^2}
\end{bmatrix}$$

$g(\Theta)$ defined in equation (15) has been proved to satisfy the following two relations.

$$E \left\{ \frac{\partial^2 g}{\partial a \partial b} \right\} = 0 \quad \text{and} \quad E \left\{ \frac{\partial^2 g}{\partial b \partial a} \right\} = 0$$

$s(\theta, \phi)$ and $s^H(\theta, \phi)$ have been found to satisfy the following relations and they can be used to prove the relations (16) and (17).

$$s^H(\theta, \phi) \frac{\partial s(\theta, \phi)}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial s^H(\theta, \phi)}{\partial \theta} s(\theta, \phi) = 0$$

$$E \left\{ \frac{\partial^2 g}{\partial a \partial \theta} \right\} = 0 \quad \text{and} \quad E \left\{ \frac{\partial^2 g}{\partial \theta \partial a} \right\} = 0$$  (16)

$$E \left\{ \frac{\partial^2 g}{\partial b \partial \theta} \right\} = 0 \quad \text{and} \quad E \left\{ \frac{\partial^2 g}{\partial \theta \partial b} \right\} = 0$$  (17)

$s(\theta, \phi)$ and $s^H(\theta, \phi)$ have also been found to satisfy the following relations and they can be used to prove the relations (18) and (19).

$$s^H(\theta, \phi) \frac{\partial s(\theta, \phi)}{\partial \phi} = 0 \quad \text{and} \quad \frac{\partial s^H(\theta, \phi)}{\partial \phi} s(\theta, \phi) = 0$$

$$E \left\{ \frac{\partial^2 g}{\partial a \partial \phi} \right\} = 0 \quad \text{and} \quad E \left\{ \frac{\partial^2 g}{\partial \phi \partial a} \right\} = 0$$  (18)

$$E \left\{ \frac{\partial^2 g}{\partial b \partial \phi} \right\} = 0 \quad \text{and} \quad E \left\{ \frac{\partial^2 g}{\partial \phi \partial b} \right\} = 0$$  (19)

Furthermore, the following relations also hold for $s(\theta, \phi)$ and can be used to prove relation (20).
\[
\frac{\partial s^H(\theta, \phi)}{\partial \phi} \frac{\partial s(\theta, \phi)}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial s^H(\theta, \phi)}{\partial \theta} \frac{\partial s(\theta, \phi)}{\partial \phi} = 0
\]

\[
E \left\{ \frac{\partial^2 g}{\partial \theta \partial \phi} \right\} = 0 \quad \text{and} \quad E \left\{ \frac{\partial^2 g}{\partial \phi \partial \theta} \right\} = 0
\]

It is clear that all the non-diagonal elements of the Fischer information matrix are zero and hence the only quantities of interest are \( E \left\{ \frac{\partial^2 g}{\partial \theta^2} \right\} \) and \( E \left\{ \frac{\partial^2 g}{\partial \phi^2} \right\} \). \( E \left\{ \frac{\partial^2 g}{\partial \theta^2} \right\} \) can be computed as,

\[
E \left\{ \frac{\partial^2 g}{\partial \theta^2} \right\} = \frac{-2a^2}{\sigma^2} \left( \frac{\partial s^H(\theta, \phi)}{\partial \theta} \frac{\partial s(\theta, \phi)}{\partial \theta} \right)
\]

From equation (13), the CRLB for \( \theta \) can be computed as,

\[
\text{var}(\hat{\theta}) \geq -E \left\{ \frac{\partial^2 g}{\partial \theta^2} \right\}^{-1}
\]

\[
\text{var}(\hat{\theta}) \geq \frac{3\lambda^2 \sigma^2}{2a^2 \pi^2 \cos^2 \theta M_x M_y \left\{ (M_x^2 - 1)(\Delta_x \cos \phi)^2 + (M_y^2 - 1)(\Delta_y \sin \phi)^2 \right\}}
\]

Similarly, we can compute \( E \left\{ \frac{\partial^2 g}{\partial \phi^2} \right\} \) and the CRLB for \( \phi \),

\[
E \left\{ \frac{\partial^2 g}{\partial \phi^2} \right\} = \frac{-2a^2}{\sigma^2} \left( \frac{\partial s^H(\theta, \phi)}{\partial \phi} \frac{\partial s(\theta, \phi)}{\partial \phi} \right)
\]

\[
= \frac{-2a^2 \pi^2}{3 \lambda^2 \sigma^2} M_x M_y \left\{ (M_x^2 - 1)(\Delta_x \sin \phi)^2 + (M_y^2 - 1)(\Delta_y \cos \phi)^2 \right\}
\]

\[
\text{var}(\hat{\phi}) \geq -E \left\{ \frac{\partial^2 g}{\partial \phi^2} \right\}^{-1}
\]

\[
\text{var}(\hat{\phi}) \geq \frac{3\lambda^2 \sigma^2}{2a^2 \pi^2 \sin^2 \theta M_x M_y \left\{ (M_x^2 - 1)(\Delta_x \sin \phi)^2 + (M_y^2 - 1)(\Delta_y \cos \phi)^2 \right\}}
\]

References


