Direction of Arrival Estimation : A LASSO Based Approach

Pranav Chimote
The Bradley Department of ECE
Virginia Tech
Blacksburg, US
pchimote@vt.edu

Avinash Kondeppady
The Bradley Department of ECE
Virginia Tech
Blacksburg, US
avinashk@vt.edu

***Abstract*—Direction of Arrival Estimation has been a common problem, appearing in various fields of signal processing and communications. Classical techniques used to tackle this problem include Beamforming, non-parametric methods such as periodograms and finally parametric methods such as MUSIC [5] and ESPRIT [6]. FFT beamformers are known to have accuracy and resolution issues while other spectral estimation techniques need prior knowledge such as the number of sources in order to estimate the angle of arrival accurately. In this report, we focus on establishing the direction of arrival problem as an optimization problem followed by discussing some techniques that can be applied to solve the problem efficiently. By taking a compressed sensing approach, the methods discussed does not require any prior knowledge of the signal and they are capable of accurate estimate even in situations where classical techniques fail such as low signal to noise ratios, spatially close objects or coherent signals.**

I. Introduction

Array signal processing is a widely used concept, playing a major role in fields like radar processing, sonar, and radio astronomy. Direction of Arrival deals with the problem of determining the number and locations of various emitters of signals. The earliest attempt at solving these, was the classical technique of beamforming. This method involved mechanically steering the antenna beam pattern to all possible angles within a region of interest. The power spectrum will have a peak corresponding to the angles of one of the incoming signals. However, this method suffers from being able to distinguish between closely spaced signals, implying a poor resolution. The Capon Beamformer was able to tackle this problem to a certain extent, by trying to maximise the signal to noise ratio at the required angles while simultaneously trying to induce nulls at every other. However, the resolution of the Capon Beamformer depends on the length of the antenna array and is therefore restricted.

The most heavily studied algorithms in the field of DOA are MUSIC [5] and ESPRIT [6]. Both of these techniques fall under the parametric estimation category and have made remarkable contributions to the field. However both these techniques are based on the Eigen Value Decomposition (EVD) of the Auto-correlation matrix in order to distinguish between the signal space and the noise space. In a similar manner as Principal Component Analysis, the signal space will have larger eigen values corresponding to a larger contribution to the power of the received signal. However, prior knowledge is required about the number of emitters/sources. Unfortunately in most practical applications of radar and sonar, the number of targets is unknown. In such cases these algorithms fail to perform efficiently. Furthermore, to develop an efficient auto-correlation matrix, multiple snapshots of the incoming signal is required.

In the recent past, with the advent of complex antenna design patterns, and the concept of virtual antenna, compressed sensing (CS) has shown better results at solving the issue of DOA. Compressed sensing is the technique of solving under-determined linear systems pertaining to sparse signal processing applications. Most DOA scenarios fall under this category, the reason for which will be discussed in the later sections. By recognising this fact, we analyse the usage of the LASSO technique along with the Least squares (LS) technique to provide a refined estimate of the DOA. Furthermore, we combine a classical technique known as Minimum Variance Distortionless Response (MVDR) along with the LASSO to improve our estimate.

II. Signal Model

We consider a linear array (LA), consisting of M elements, Let d_i denote the i-th element position in the array.

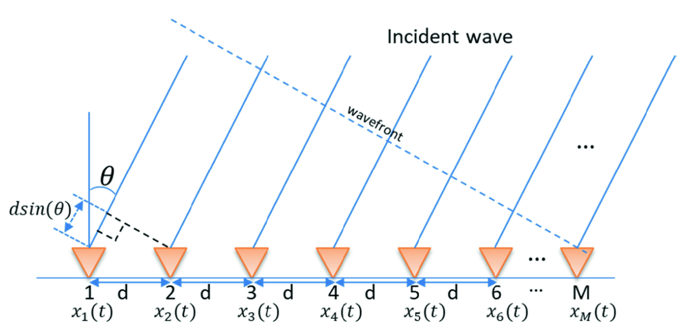

Fig. 1. Linear Antenna Array

let us assume that there are L narrow-band, far-field sources with Direction of arrival (DOA) (θ_l) and powers (σ_l^2), l=1,2,. . . ,L. It can also be assumed that the source signals are uncorrelated with one another. Let $\mathbf{a}(\theta_l) \in C^{M\times 1}$be the steering vector corresponding to DOA (θ_l), whose i-th element is $e^{-jk_o d_i sin(\theta_l)}$, where $k_o = \frac{2\pi}{\lambda}$ is the wave number and λ is the wavelength of the propagating waves. Let the vector $\mathbf{s}$(t) = $[s_1(t)s_2(t)\dots s_L(t)]^\top$, where s $\in$ $C^{L\times 1}$ represent the source signals. The output of LA can be written as:

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

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

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, $\{\overline{\Omega}_n\}_{n=1}^N$, N denoting a grid that covers Ω, with N >> L.

Let:

$$\overline{s}(t) = [\overline{\sigma}_1 \quad \overline{\sigma}_2 \quad \dots \quad \overline{\sigma}_N]^\top \tag{2}$$

where $\overline{s} \in C^{N\times 1}$, and:

$$\Phi = [\overline{\mathbf{a}}(\overline{\theta}_1)\overline{\mathbf{a}}(\overline{\theta}_2) \quad \dots\overline{\mathbf{a}}(\overline{\theta}_n) \quad \dots\overline{\mathbf{a}}(\overline{\theta}_N)] \tag{3}$$

$$\begin{bmatrix} 1 & \dots & 1 \\ e^{-jk_o d_i sin(\theta_l)} & \dots & e^{-jk_o d_i sin(\theta_N)} \\ \vdots & \ddots & \vdots \\ e^{-jk_o (M-1) d_i sin(\theta_l)} & \dots & e^{-jk_o (M-1) d_i sin(\theta_N)} \end{bmatrix}$$

where $\Phi \in C^M$ and $\overline{a}(\overline{\theta}_n) \in C^{M\times 1}$ is the steering vector for the array. The received signal at the $\tilde{m}-th$ sensor is:

$$\mathbf{y}_{\overline{m}}(t) = \phi_{\overline{m}}\overline{\mathbf{s}}(t) + \overline{n}_{\overline{m}}, \quad \overline{m} = 1, 2, \dots, M \tag{4}$$

where $\phi_{\overline{m}}$ is the m-th row of Φ. The n-th row of $\overline{s}$(t), $\overline{s}_n(t)$ is non zero only if ($\overline{\theta}_n = \theta_l$), and in that case, $\overline{\sigma}_n = \sigma_l$. Thus Equation (4) can be rewritten as:

$$\mathbf{y}(t) = \Phi\overline{\mathbf{s}}(t) + \overline{\mathbf{n}}(t) \tag{5}$$

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. The locations of the sources can be obtained by ordinary least squares method (OLS), which leads us to the following optimization problem:

$$\min \left|\left| y - \sum_{n=1}^{N} \phi_n \overline{s}_n \right|\right|^2 \tag{6}$$

The problem with this approach is that this will lead to non-zero estimates for all the coefficients. Since the sources are point sources and their number is small (L << N), we aim to make the spatial spectrum as sparse as possible thus eliminating the minuscule values which basically represent noise.

III. The LASSO Method

Since we aim for sparsity in our optimization problem, we can solve it by regularizing it to favour sparse signal fields using LASSO. The minimization problem in Equation (6) gets redefined as:

$$\min \left|\left| y - \sum_{n=1}^{N} \phi_n \overline{s}_n \right|\right|^2 + \tau \sum_{n=1}^{N} |\overline{s}_n| \tag{7}$$

where ϕ_n is the n-th element of $\phi_{\overline{m}}$ and $\overline{s}_n$ is the n-th element of $\overline{s}$. Equation(7) can be rewritten as:

$$\hat{s}_{lasso} = min_s ||\mathbf{y} - \Phi\overline{\mathbf{s}}||_2^2 + \tau||\overline{\mathbf{s}}||_1 \tag{8}$$

where τ is a non-negative regularization parameter. The first term in Equation (8) 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, thus degrading the estimation accuracy. Zou [**?**] 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.

A. Adaptable LASSO

The regularization parameter in Equation(7) equally penalizes the coefficients of the l_1 norm. Thus this may lead to an inherent bias in these estimates resulting in reduced accuracy. In order to overcome this deficiency, we apply the A-LASSO in the DOA estimation problem [3]. Hence, the A-LASSO minimizes:

$$\hat{s}_{lasso} = min_s ||\mathbf{y} - \Phi\overline{\mathbf{s}}||_2^2 + \tau \sum_{n=1}^{N} w_n |\overline{\mathbf{s}}_n| \tag{9}$$

where w_n is the n-the element of the vector. Let $\hat{s}$ be the initial estimate of $\overline{s}$. Now choosing any weight factor γ, where $\gamma > 0$ and defining the weight vector as $\hat{\mathbf{w}} = [w_1 w_2 \dots w_n]^\top$, where:

$$\hat{w} = \frac{1}{|s_n|^\gamma} \qquad n = 1, 2, \dots, N \tag{10}$$

The A- LASSO is given by:

$$\hat{\mathbf{s}}^k = min_s ||\mathbf{y} - \Phi\overline{\mathbf{s}}||_2^2 + \tau_k \sum_{n=1}^{N} w_n |\overline{\mathbf{s}}_n| \tag{11}$$

where k is the iteration number. The iteration repeats itself until the number of sparse sources are below a certain threshold value $\hat{l}$

The minimization in Equation(11) 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 is utilized to solve the A-LASSO using the following steps:

Algorithm 1: A-LASSO

initialize $\overline{s} \leftarrow \hat{s}$

repeat

$\hat{w} \leftarrow \frac{1}{|\hat{s}^n|}$

$\Phi^*(\overline{m}, n) \leftarrow \phi_{(\overline{m},n)}/\hat{w}_n$

$\hat{s}^* \leftarrow min_s||\mathbf{y} - \Phi^*\overline{\mathbf{s}}||_2^2 + \tau_k||\overline{\mathbf{s}}||_1$

$\hat{s}^k \leftarrow \hat{s}^*/\hat{w}_n$

until number of non zero elements in $\overline{s} <= \hat{l}$

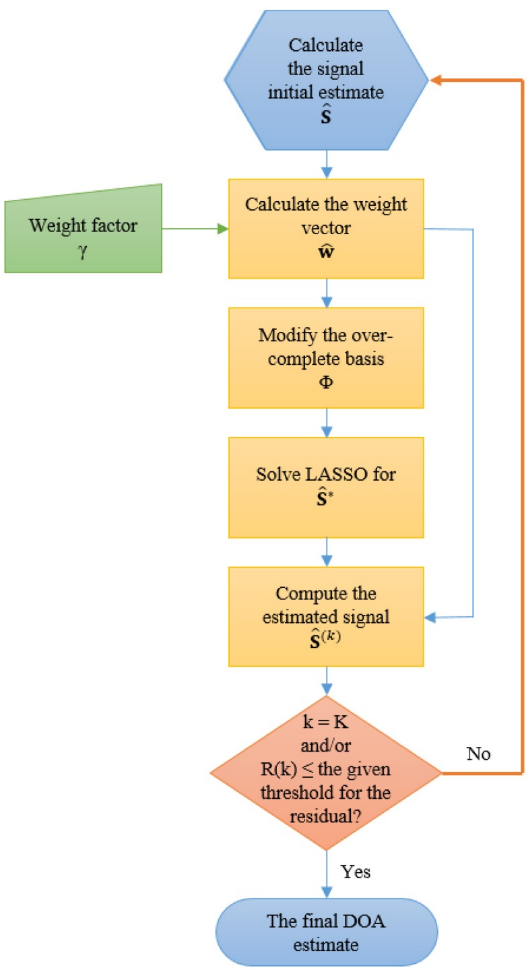

Fig. 2. Flow Chart of A-LASSO algorithm

B. OLS LASSO

We always assume a vector of ones as the initial estimate for $\overline{s}$. However, it has been proven from prior simulations that 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. Even more, if we try to push the algorithm to the limits (by choosing a small τ), we find spurious peaks along with the genuine peaks. Therefore, we will use the OLS solution as the initial signal estimate for $\overline{s}$, 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 $\hat{s}_n$ with $\hat{s}_{ols_n}$, which is the n-th element of $\hat{s}_{ols}$.

$$\hat{s}_{ols} = min_s||\mathbf{y} - \Phi\overline{s}||_2^2 \tag{12}$$

The number of source signals is not required to be known in advance for OLS A-LASSO. It should be noted that we do not use OLS for standalone DOA estimation. We simply use it as an initial guess in OLS A-LASSO

IV. Modifying the Weights of LASSO

In this section we discuss a better approach to modify the weights of the A-LASSO technique. To explain our approach, we would have to first dive into a classical beamforming technique explained below.

A. Beamforming and Minimum Variance Distortionless Response

Every antenna pattern has a particular radiation pattern associated with it. Beamformers aim at creating a radiation pattern by combining signals from different antennae with different weights. They do so by compensating for the phase shifts occurring at each of the receiver antenna.

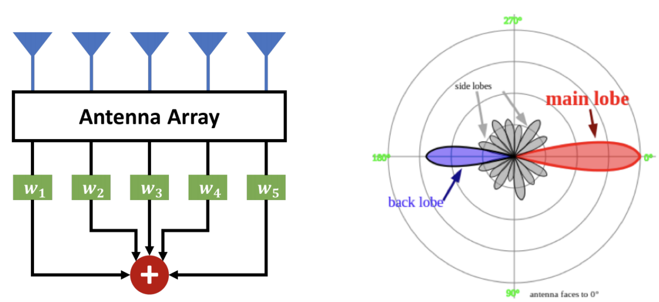

Fig. 3. A beamformer and an Antenna pattern

As shown in the diagram above, the weights are multiplied in order to steer the beam. Depending on the weights of the beamformer, the main lobe of the antenna pattern can be directed at different angles. The output can be expressed as the following :

$$\mathbf{y}(t) = \sum_{n=0}^{N-1} w_n^* x_n = \mathbf{w}^H \mathbf{x} \tag{13}$$

These techniques are implemented by running the above equation for different weight vectors corresponding to different steering angles. Thus it is easy to observe that in this case w = $\mathbf{a}(\theta_i)$ where i would keep varying depending the angle in question. The power spectrum of the spectrum is then analysed and the angle with the maximum power is determined as the DOA. We can observe the resemblance between beamforming techniques and the ordinary least squares (OLS) method. In OLS, we basically multiply the vector **x** by the column vectors of Φ. Then we try to analyse which of these vectors has tried to best minimise the error.

The Minimum Variance Distortionless Response (MVDR) beamformer [4] behaves in a different manner however. The key idea in MVDR is that we want to maintain the signals from the desired direction while suppressing others. The weights in this section are calculated via an optimisation equation given as follows :

$$\begin{aligned} \min_{w} \quad & \mathbf{Y}\mathbf{Y}^H = \min_{w} \mathbf{w}^H \mathbf{x}\mathbf{x}^H \mathbf{w} \quad = \min_{w} \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^H \mathbf{a}(\theta) = 1 \end{aligned} \tag{14}$$

The term R_{xx} denotes the auto-correlation matrix of the input. The constraint on the optimisation equation suggests that the signals from $\mathbf{a}(\theta)$ are maintained while the minimisation equation aims at reducing the variance of the output. Just as in the case of the beamformer, the algorithm is applied to every value of $\mathbf{a}(\theta_i)$ in the desired region of interest. The above equation can be easily solved using the Lagrangian multiplier as shown below :

$$\mathcal{L}(\mathbf{w}, \lambda) = \mathbf{w}^H \mathbf{R}_{xx} \mathbf{w} + \lambda(\mathbf{w}^H \mathbf{a}(\theta) - 1)$$

Taking the derivative w.r.t w

$$\frac{\partial \mathcal{L}}{\partial w} = \mathbf{w}\mathbf{R}_{xx} + \lambda \mathbf{a}(\theta)$$

Applying the constraint equation we get,

$$\mathbf{w}_{opt}(\theta) = \frac{\mathbf{R}_{xx}^{-1}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{R}_{xx}^{-1}\mathbf{a}(\theta)}$$

And the power spectrum is

$$\mathbf{P}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\mathbf{R}_{xx}^{-1}\mathbf{a}(\theta)}$$

The MVDR algorithm is known to be better at identifying spatially close by objects have a superior performance as compared to the conventional beamformer

B. MVDR LASSO Technique

Using the MVDR algorithm we can now modify the weights of the A-LASSO algorithm replacing step [2] in Algorithm [1], to achieve a superior performance. This modification will result in sharper peaks thereby increasing the resolution. It is important to notice that in this set up, we do not need to know any prior knowledge of the number of sources. The autocorrelation matrix must be a full rank matrix, hence it is advisable to perform spatial smoothing technique like forward-backward spatial smoothing as mentioned in [**?**].

V. Implementation Details

A. The Complex Lasso

The conventional LASSO algorithm as suggested in [2], can handle real values only. However, in our signal model, we deal with the analytic signals which are complex variables. Therefore, we need to modify the algorithm to accommodate for this difference. The LASSO solution is given by :

$$\min \left\| y - \sum_{n=1}^{N} \phi_n \bar{s}_n \right\|^2 + \tau \sum_{n=1}^{N} |\bar{s}_n| \tag{15}$$

where the L1 norm constraint $\|\bar{\mathbf{s}}\|_1$ is replaced by $\sqrt{a^2 + b^2}$ where a and b represent the real and imaginary parts of the vector **s**. Since the L1 norm constraint is also a convex function, the LASSO algorithm still holds true. This change leads to modifying the algorithm mentioned in [5]. The soft threshold in LASSO are now updated given by the following equation :

$$S(x, \lambda) = e^{iarg(x)}(|x| - \gamma)_+ \quad for\, x \in \mathbb{C}, \gamma \in \mathbb{R} \tag{16}$$

Applying these changes, ensures that the elements in the source vector **s** are sparse, such that both the real and complex components are both simultaneously driven down to zero.

B. Choosing the Regularisation Parameter

The performance of the LASSO problem is heavily dependent on the choice of the regularisation parameter τ. This regularisation parameter controls the trade off between model fitting and the L1 norm constraint. A higher τ would increase the error of the model but would ensure a lot of the values in the vector (s) would be zero. While s would be sparse, the DOA would have considerable errors. Thus, choosing a suitable value of τ directly affects the accuracy of our estimator. In this project, we use a cross validation method to select the regularisation parameter. We generate multiple snapshots of the data, and run the LASSO algorithm with varying parameters of τ to drive down the L2 norm error, while also achieving a relatively sparse solution. The performance of these results is then compared against a cross validation set to observe the overall performance with the chosen lambda. This is important to observe that for different SNR values, the lambda values are different and the cross validation test must be repeated in such cases.

VI. Results and Inferences

In this section, we discuss the results and conclusions drawn from our project. We ran the OLS A-LASSO and the MVDR A-LASSO algorithms on a Uniform Linear Array (ULA) of 10 sensor elements with sources at angles [-70, -30, 30, 60] with an SNR of 15dB. The number of snapshots was 300, and we used 75% of the number of snapshots for finding the τ estimate and the remaining 25% as a cross-validation set. For an SNR of 5dB, we found a τ value ranging between 0.1 - 0.25 led to an acceptable sparse estimate. The figures for the normalised power spectrum are shown below :

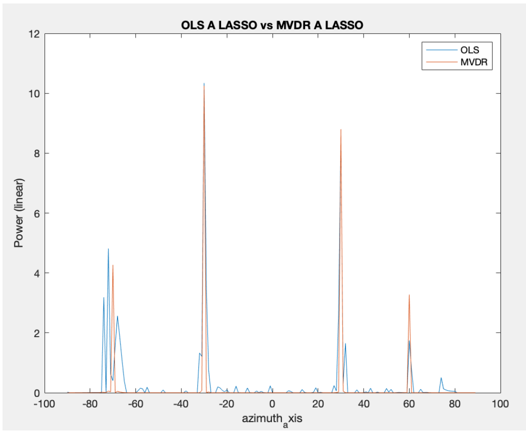

Fig. 4. OLS A LASSO versus MVDR A LASSO

The MVDR A LASSO produces a relatively more sparse estimate as compared to the OLS A LASSO. Moreover the OLS A LASSO took a higher number of iterations, around 6 iterations as compared to MVDR A LASSO which took just 3 iterations. The time taken for the OLS LASSO was roughly 14.2 seconds, while the MVDR A LASSO took 4.5 seconds on MATLAB. The reason for this is quite apparent when viewing the power spectrum for the OLS solution and the MVDR beamformer

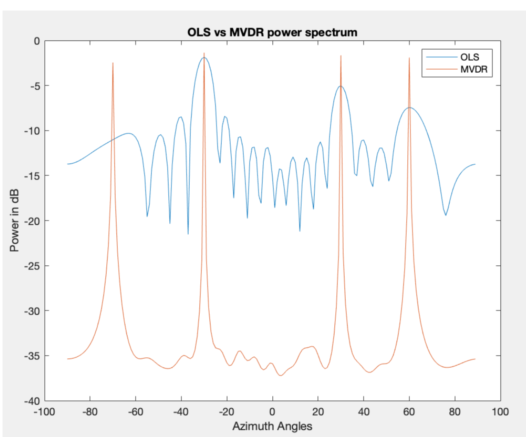

Fig. 5. Performance comparison of OLS solution vs MVDR

The MVDR beamformer has a lower noise floor as compared to the OLS solution. Therefore the weights are much higher for the undesired steering vectors in the MVDR case as opposed to the OLS solution. Thus the MVDR A LASSO is able to give a better, sparse estimate quicker than the OLS A LASSO. Furthermore, since the peaks are easily distinguished in the MVDR beamformer, the solution is more sparse as well.

We also ran our proposed algorithm on the same inputs on a similar array configuration but with a different SNR of 5 dB. On running the same cross validation test, we noticed the τ parameter should vary between 0.8 - 1.75 in this case. For this scenario, we observe that the estimate is sparse, the peaks of the MVDR A LASSO are still sharp and sparse. However, the OLS A LASSO is not as accurate as the MVDR A LASSO

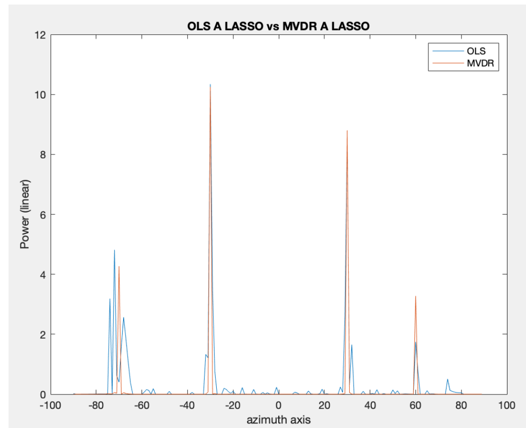

Fig. 6. OLS A LASSO versus MVDR A LASSO for SNR = 5 dB

Our results show that, the MVDR A LASSO outperforms the OLS A LASSO in both accuracy as well as computational time. Our claims are corroborated based on results.

VII. Future Work

The proposed algorithm here may not work as well when we have closely spaced sources because the steering vectors on an exhaustive grid may be correlated with another. To have a better resolution, the LASSO algorithm can be replaced with the Least Angle Regression Solution (LARS) as mentioned in [6]. Furthermore, in low SNR cases, the MVDR beamformer is known to have a poor performance. To combat this issue, the weights of the Adaptive LASSO can be replaced with that of a finer estimate such as the MUSIC algorithm.

References

[1] Salama A.A., Ahmad M.O., Swamy M.N.S. Underdetermined DOA Estimation Using MVDR-Weighted LASSO. Sensors. 2016;16:1549. doi: 10.3390/s16091549.

[2] Tibshirani R. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Methodol.) 1996;58:267–288.

[3] Zou H. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc. 2006;101:1418–1429. doi: 10.1198/016214506000000735.

[4] Chen Z., Gokeda G., Yu Y. Introduction to Direction-of-Arrival Estimation. Artech House; Norwood, MA, USA: 2010. Overview of Basic DOA Estimation Algorithms; pp. 31–64.

[5] A. Maleki, L. Anitori, Z. Yang and R. G. Baraniuk, "Asymptotic Analysis of Complex LASSO via Complex Approximate Message Passing (CAMP)," in IEEE Transactions on Information Theory, vol. 59, no. 7, pp. 4290-4308, July 2013

[6] Efron B., Hastie T., Johnstone I., Tibshirani R. Least angle regression. Ann. Stat. 2004