A Robust ℓ_1 Penalized DOA Estimator

Ashkan Panahi and Mats Viberg

Signal processing group, Signals and Systems department

Chalmers university of technology, Gothenburg, Sweden

Abstract—The SPS-LASSO has recently been introduced as a solution to the problem of regularization parameter selection in the complex-valued LASSO problem. Still, the dependence on the grid size and the polynomial time of performing convex optimization technique in each iteration, in addition to the deficiencies in the low noise regime, confines its performance for Direction of Arrival (DOA) estimation. This work presents methods to apply LASSO without grid size limitation and with less complexity. As we show by simulations, the proposed methods loose a negligible performance compared to the Maximum Likelihood (ML) estimator, which needs a combinatorial search We also show by simulations that compared to practical implementations of ML, the proposed techniques are less sensitive to the source power difference.

I. INTRODUCTION

The Least Absolute Shrinkage and Selection Operator (LASSO) is a method of choosing a small set of bases among a large collection, best representing a set of data linearly. It is based on ℓ_1 regularization of an Ordinary Least Square (OLS), which always gives a sparse solution based on the conic nature of the cost function [1]. The consistency of such a technique, under the assumption of the "best collection", is well discussed in an asymptotic case that both the data dimension and the size of collection increase to infinity [2].

A much different attempt has been made by applying LASSO into a finite dimensional data set with an asymptotically large set of basis dictated by a physical model. A well known example is in [3], where the LASSO technique is applied to the problem of estimating Direction Of Arrivals (DOAs) which can be expressed and solved in a linear regression fashion by discretizing the DOA parameter space. This method gained more attention since the solution could be found robustly, independent of the choice of the initial values in the numerical optimization technique due to the convex nature of the cost function.

Unlike the statistical regression application of LASSO, the choice of the Regularization Parameter (RP) is critical in the current application. This parameter implements the trade off between model precision and model order. There are numerous ways to estimate the RP when the true number of sources is unknown [4], [5]. However, the estimation quality is always improved by choosing a smaller RP value with the same model order. On the other hand, finding the smallest such parameter is not straightforward, since the implementation of LASSO by convex programming techniques [6] is computationally costly, and can not be performed for a fine search over possible RP values. A stagewise solution is found in [7] for the real regression case, by the observing that the homotopy path of

LASSO solutions is piecewise linear [8]. Later, in [9] the idea was generalized to the complex problems by introducing a different LASSO optimality condition which is called Singular Point Selection (SPS)-LASSO. Decreasingly in RP, the SPS-LASSO follows the points in the homotopy path, in which a new regressor is born. These points are known as the candidate points.

The SPS-LASSO technique is not suitable for the DOA estimation problem, especially in low SNR cases, due to the fact that the solution support moves continuously between two singular points. Furthermore, the discretized nature of the problem introduces an additional quantization error. Accordingly, in this work, we introduce an alternative recursive solution, by keeping only the estimate support at each iteration and modifying the optimality conditions to avoid a discretized space. The resulting algorithm does not have the initialization problems of NonLinear Least Squares (NLLS) and is faster than convex programming techniques. We further simplify the algorithm to get a faster solution. We show by simulations that the faster version is also convergent to the optimal point with very high probability. The results show that the proposed optimal point is very close to the ML global minimum point, so that we loose a negligible performance compared to ML. Note that ML is very costly to be implemented when the model order grows. We further compare LASSO to other approximations of ML such as Space alternating Generalized EM (SAGE [10]).

II. PROBLEM STATEMENT

This work concerns the problem of DOA estimation with an array of m sensors, receiving one snapshot of narrowband signals from n far sources. Due to the far-field model of the transmitting wave, the received signal pattern is mostly defined by phase shifts at each sensor so that the received signal vector $\mathbf{x}$ can be written as ([11])

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{a}(\theta_i)s_i + \mathbf{n} = \mathbf{A}(\boldsymbol{\theta})\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{s} = [s_1 \ s_2 \ \dots s_n]^T$, $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \dots \theta_n]$, and $\mathbf{n}$ are the source waveform, DOA, and measurement noise vectors respectively. Furthermore, $\mathbf{a}(\theta)$ is the steering vector, which represents the phase shift operations at different sensors corresponding to the DOA θ. The problem is to estimate the DOA vector $\boldsymbol{\theta}$ given the measurement vector $\mathbf{x}$ assuming an uncorrelated, circularly symmetric, centered Gaussian noise vector $\mathbf{n}$.

A. Conventional Solutions

Under the above statistical assumptions of the noise vector, the deterministic ML estimator is given by ([11])

$$(\hat{\boldsymbol{\theta}}, \hat{\mathbf{s}}) = \underset{(\boldsymbol{\theta}, \mathbf{s})}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}\|_2^2, \tag{2}$$

where $\boldsymbol{\theta} \in [0 \ \pi]^n$. The solution of 2 could be found by noting its corresponding Karush-Kuhn-Tucker (KKT) conditions ([12])

$$\forall \theta_i \in \boldsymbol{\theta} : \begin{cases} \mathbf{a}^H(\theta_i)\hat{\mathbf{n}} = 0 \\ \Re(s_i^* \mathbf{d}^H(\theta_i)\hat{\mathbf{n}}) = 0. \end{cases} \tag{3}$$

where $\hat{\mathbf{n}} = \mathbf{x} - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}$ and $\mathbf{d}(\theta) = \frac{\partial \mathbf{a}(\theta)}{\partial \theta}$. The system (3) does not have a unique solution although the ML global optimum is unique. However, starting from a sufficiently close point, and following a Newton recursive algorithm, the solution may converge to the global optimum. We introduce Algorithm 1 as a typical such solution in which we write the complex waveforms in polar coordinates as $s_i = r_i e^{j\alpha_i}$ and $\mathbf{J}_{\mathrm{ML}}$ is given in (4) where $\boldsymbol{\Delta}_1$ and $\boldsymbol{\Delta}_2$ are diagonal matrices who's diagonal elements are given by $\mathbf{d}^H(\theta_i)\hat{\mathbf{n}}$ and $\mathbf{c}^H(\theta_i)\hat{\mathbf{n}}$ with $\mathbf{c}(\theta) = \frac{\mathrm{d}\mathbf{d}(\theta)}{\mathrm{d}\theta}$, respectively. Furtheremore, $\mathbf{D}(\boldsymbol{\theta}) = [\mathbf{d}(\theta_1) \ \mathbf{d}(\theta_2) \ldots \mathbf{d}(\theta_n)]$ and $\mathbf{S}$ and $\boldsymbol{\Gamma}$ are diagonal matrices who's diagonal elements are s_i and $e^{j\alpha_i}$ respectively. More details of such methods could be found in [13].

Algorithm 1 ML estimator Newton solver

$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta}_0$ and $\mathbf{s} \leftarrow \mathbf{s}_0$.
while not converging **do**
    $\hat{\mathbf{n}} \leftarrow \mathbf{x} - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}$
    $\boldsymbol{\eta}_{\mathrm{ML}} \leftarrow \left[\Re(\mathbf{A}^H(\boldsymbol{\theta})\hat{\mathbf{n}})^T \ \Im(\mathbf{A}^H(\boldsymbol{\theta})\hat{\mathbf{n}})^T \ \Re(\mathbf{D}^H(\boldsymbol{\theta})\hat{\mathbf{n}})^T\right]^T$
    $\boldsymbol{\eta}_0 \leftarrow \mathbf{J}_{\mathrm{ML}}^{-1}\boldsymbol{\eta}_{\mathrm{ML}}$
    for i=1:n **do**
        $\theta_i \leftarrow \theta_i + \eta_{0,i}$
        $r_i \leftarrow r_i + \eta_{0,n+i}$
        $\alpha_i \leftarrow \alpha_i + \eta_{0,2n+i}$
    end for
end while

The convergence could be checked simply by thresholding the difference in the sequence of estimates or the sequence of the cost function. Although the Algorithm 1 is naive in practice, due to the inversion instability of the Jacobian matrix $\mathbf{J}$, we introduce it as a base of development toward our proposed method. As we have already explained, the direct ML realization is costly when the desired model order is high. Furtheremore, such method is not stable due to the .

There has been a variety of approximate, low-cost solutions to ML such as Expectation-Maximization (EM) algorithms and Space Alternating Generalized EM (SAGE [10]) as well as RELAX [14] all more or less dependent on the choice of the initial values. As another alternative, the possibility of estimating DOAs using ℓ_1 penalized Ordinary Least Squares (OLS) has been known and discussed for almost a decade ([3]). This method solves the problem of local minima by minimizing an approximate convex cost function which is

independent of the initial values. Assume a discretization $\boldsymbol{\theta}^g = \{\theta_1^g, \theta_2^g, \ldots, \theta_N^g\}$ of $[0 \ \pi]$. According to [3], the solution of (2) could be approximated by first solving

$$\hat{\mathbf{s}}^g = \underset{s^g}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{x} - \mathbf{A}^g \mathbf{s}^g\|_2^2 + \lambda\|\mathbf{s}^g\|_1, \tag{5}$$

where $\mathbf{A}^g = \mathbf{A}(\boldsymbol{\theta}^g)$ and λ is a regularization parameter [4], and next introducing $\hat{\boldsymbol{\theta}}$ as the elements in $\boldsymbol{\theta}^g$ with nonzero corresponding $\hat{\mathbf{s}}^g$ elements.

The parameter λ sets the compromise between the level of sparseness and model fit, and is usually hard to determine analytically. On the other hand, realizing LASSO in (2) for a big set of λ values is infeasible. In [9], an alternative solution is introduced by the SPS-LASSO stagewise algorithm, which only needs LASSO realizations at a recursively determined finite set of candidate λ values. This could be done by observing the following optimality conditions of LASSO [9]

$$\forall \theta_i^g \in \boldsymbol{\theta}^g : \begin{cases} |\mathbf{a}^H(\theta_i^g)\hat{\mathbf{n}}| \leq \lambda \\ \theta_i^g \in I \Rightarrow \mathbf{a}^H(\theta)\hat{\mathbf{n}} = \lambda \frac{\hat{s}_i}{|\hat{s}_i|} \end{cases}, \tag{6}$$

where $I = \{\theta_i^g | \hat{s}_i^g \neq 0\}$ and $\hat{\mathbf{n}} = \mathbf{x} - \mathbf{A}^g \mathbf{s}$. Let us define the following function:

1) Marginalized Source Attractor (G): Given the quadruple (I, r, α, λ) and a point θ, the function $G(I, r, \alpha, \phi; \theta)$ gives the convergence point (r', α', λ') of the Newton iterative solution of the following system of equations starting from (r, α, λ) with $s_i = r_i e^{j\alpha_i}$.

$$\mathbf{a}^H(\theta_i)\hat{\mathbf{n}}' = \lambda' e^{j\alpha_i'}$$
$$|\mathbf{a}^H(\theta_i)\hat{\mathbf{n}}'| = \lambda'. \tag{7}$$

where $\hat{\mathbf{n}}' = \mathbf{x} - \mathbf{A}(I)\mathbf{s}'$. The G function could also be computed faster using an iterative algorithm as in [9], but we will not express it here. The algorithm could be written as Algorithm 2 where G_λ denotes the λ element of the function G. Note that it is assumed that between each two candidate points in the solution path of LASSO (as a function of λ) the DOA estimate is constant.

Algorithm 2 SPS LASSO

$\theta_0 \leftarrow \underset{\theta^g \in \boldsymbol{\theta}^g}{\operatorname{argmax}} |\mathbf{a}^H(\theta^g)\mathbf{x}|$
$I \leftarrow \{\theta_0\}$ and $\lambda \leftarrow |\mathbf{a}^H(\theta_0)\mathbf{x}|$
$r \leftarrow 0$ and $e^{j\alpha} \leftarrow \frac{\mathbf{a}^H(\theta_0)\mathbf{x}}{|\mathbf{a}^H(\theta_0)\mathbf{x}|}$
counter $\leftarrow 0$.
while counter $< n$ **do**
    $\theta_1 \leftarrow \underset{\theta^g \in \boldsymbol{\theta}^g}{\operatorname{argmax}} G_\lambda(I, r, \alpha, \lambda; \theta^g)$
    $I \leftarrow I \cup \theta_1$
    $(r, \alpha, \lambda) \leftarrow G(I, r, \alpha, \lambda; \theta_1)$
    $r_{\theta_1} \leftarrow 0$ and $e^{j\alpha(\theta_1)} \leftarrow \frac{\mathbf{a}^H(\theta_0)\mathbf{x}}{|\mathbf{a}^H(\theta_0)\mathbf{x}|}$
    counter $\leftarrow$ counter+1.
end while

The performance of SPS-LASSO is still limited by the size of the grid which may not be increased arbitrarily due to the complexity of convex optimization techniques [6].

$$\mathbf{J}_{\text{ML}} = \begin{bmatrix} \Re(\boldsymbol{\Delta_1} - \mathbf{A}^H(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}) & -\Re(\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}) & \Im(\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\mathbf{S}) \\ \Im(\boldsymbol{\Delta_1} - \mathbf{A}^H(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}) & -\Im(\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}) & -\Re(\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\mathbf{S}) \\ \Re(\boldsymbol{\Delta_2} - \mathbf{D}^H(\boldsymbol{\theta})\mathbf{D}(\boldsymbol{\theta})\mathbf{S}) & -\Im(\mathbf{D}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}) & -\Re(\mathbf{D}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\mathbf{S}) \end{bmatrix}$$
$$\mathbf{j} = \begin{bmatrix} -\mathbf{D}(\boldsymbol{\theta})\mathbf{S} & -\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma} & -j\mathbf{A}(\boldsymbol{\theta})\mathbf{S} \end{bmatrix} \tag{4}$$

Furthermore, SPS-LASSO works incorrectly in very high SNR regimes, where the desired regularization parameter is extremely small and the approximately fixed DOAs assumption does not hold.

To overcome such difficulties, in this work we redefine the parameter space as the space of low dimensional non sparse vectors. However, inspired by the optimality conditions of LASSO, we introduce slightly modified feasible conditions which do not depend on any discretization. We demonstrate the details of two algorithms to find the unique optimal point of these new conditions.

III. CONTINUOUS LASSO

Inspired by SPS-LASSO, we look for a finite subset $I = \{\theta_1, \theta_2, \ldots, \theta_r\} \subset [0 \ \pi]$ and a complex function $s_i = s(\theta_i)$ on I satisfying

$$\forall \theta \in [0 \ \pi] : \begin{cases} |\mathbf{a}^H(\theta)\hat{\mathbf{n}}| \leq \lambda \\ \theta = \theta_i \in I \Rightarrow \mathbf{a}^H(\theta)\hat{\mathbf{n}} = \lambda \frac{s_i}{|s_i|} \end{cases}, \tag{8}$$

where $\hat{\mathbf{n}} = \mathbf{x} - \mathbf{A}(I)\mathbf{s}$. We also write $s = re^{j\alpha}$ where (r, α) are the polar coordinates of the function s. The existence and uniqueness of such set as well as consistency conditions of such method is proved but will not be presented in this work.

Finding such a set is guaranteed by the homotopy rule and the similarities to convex LASSO based estimation. Note that finding the solution for a fixed λ by a gradient descent algorithm is not possible due to the infinite dimensional nature of the problem. Thus, we assume that λ is variable. Note that for the values of $\lambda > \max_{\theta} |\mathbf{a}^H(\theta)\mathbf{x}|$ the solution is given by $I = \emptyset$. Note that conditions (8) imply that each $\theta_i \in I$ is a global maximum point for the function $f(\theta) = |\mathbf{a}^H(\theta)\hat{\mathbf{n}}|$. Thus, its derivative is zero which after some manipulations results in the condition

$$\Re(s_i^* \mathbf{d}^H(\theta_i)\hat{\mathbf{n}}) = 0. \tag{9}$$

The equation (9) and the second line of (8) might be compared to the ML optimality conditions in (3). The main idea in this work is to relax the assumption that DOAs are fixed in the smooth homotopy pieces. Roughly speaking, this means that the marginalized source attractor G should be modified to include DOA changes imposed by (9) as well as source changes. Accordingly, we modify G as follows:

2) Local Marginalized Attractor (F): : Given a quadruple (I, r, α, λ) and a point θ, the function $F(I, r, \phi, \lambda; \theta)$ gives the convergence point quadruple $(I', r', \alpha', \lambda')$ starting from (I, r, α, λ) and following the Newton recursive solution of the

system of equations

$$\mathbf{a}^H(\theta_i')\hat{\mathbf{n}}' = \lambda' \frac{\mathbf{s}_i'}{|\mathbf{s}_i'|}$$
$$\Re\left(\frac{\mathbf{s}_i'^*}{|\mathbf{s}_i'|}\mathbf{d}^H(\theta_i')\hat{\mathbf{n}}'\right) = 0$$
$$|\mathbf{a}^H(\theta)\hat{\mathbf{n}}'| = \lambda', \tag{10}$$

where $\hat{\mathbf{n}}' = \mathbf{x} - \mathbf{A}(I')\mathbf{s}'$. Note that F is obtained from G by relaxing the DOA parameters and imposing (9). The G function could be found using a Newton algorithm similar to Algorithm 1 by substituting $\mathbf{J}_{\text{ML}}$ and $\boldsymbol{\eta}_{\text{ML}}$ with $\mathbf{J}_{\text{LMA}}$ and $\boldsymbol{\eta}_{\text{LMA}}$ respectively where

$$\boldsymbol{\eta}_{\text{LMA}} = \begin{bmatrix} \boldsymbol{\eta}_{\text{LEA}} \\ |\mathbf{a}^H(\theta)\hat{\mathbf{n}}|^2 - \lambda^2 \end{bmatrix}, \tag{11}$$

and

$$\mathbf{J}_{\text{LMA}} = \begin{bmatrix} & \Re(e^{j\alpha_1}) \\ & \Re(e^{j\alpha_2}) \\ & \vdots \\ & \Re(e^{j\alpha_n}) \\ \mathbf{J}_{\text{LEA}} & \Im(e^{j\alpha_1}) \\ & \Im(e^{j\alpha_2}) \\ & \vdots \\ & \Im(e^{j\alpha_n}) \\ \Re\left(\mathbf{a}^H(\theta)\mathbf{j}\hat{\mathbf{n}}^H\mathbf{a}(\theta)\right) & -2\lambda \end{bmatrix}, \tag{12}$$

with $\mathbf{j}$ given in (4),

$$\boldsymbol{\eta}_{\text{LEA}} = \boldsymbol{\eta}_{\text{ML}} - \lambda[\Re(e^{j\alpha}) \ \Im(e^{j\alpha}) \ 0]^T, \tag{13}$$

and

$$\mathbf{J}_{\text{LEA}} = \mathbf{J}_{\text{ML}} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \lambda\Im(\boldsymbol{\Gamma}) \\ \mathbf{0} & \mathbf{0} & -\lambda\Re(\boldsymbol{\Gamma}) \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{14}$$

The proposed C-LASSO algorithm is summarized as 3. below.

The proposed algorithm gives the precise optimal solution of (8) which as we show by simulations is a proper robust approximation of the ML estimator. However, the searching steps might still be costly for certain fast applications. In this case, we introduce a different method of solving (8) which we call C-LASSO$_h$.

The idea in this modification is that the long jumps in the LASSO path by the attractor F could be substituted by a smoother development of λ. Although it may take more steps to achieve the next singular point, the simplicity of computations at each step compensates the complicated process at searching steps of C-LASSO. Note that the conditions in (8) imply that the spectrum $f(\theta) = |\mathbf{a}^H(\theta)\hat{\mathbf{n}}|$ is bounded by λ and

Algorithm 3 C-LASSO

$\theta_0 \leftarrow \underset{\theta}{\operatorname{argmax}} |\mathbf{a}^H(\theta)\mathbf{x}|$
$I \leftarrow \{\theta_0\}$ and $\lambda \leftarrow |\mathbf{a}^H(\theta_0)\mathbf{x}|$
$r \leftarrow 0$ and $e^{j\alpha} \leftarrow \frac{\mathbf{a}^H(\theta_0)\mathbf{x}}{|\mathbf{a}^H(\theta_0)\mathbf{x}|}$
counter $\leftarrow 0$.
while counter $< n$ **do**
    $\theta_1 \leftarrow \underset{\theta}{\operatorname{argmax}} F_\lambda(I, r, \alpha, \lambda; \theta)$
    $(I, r, \alpha, \lambda) \leftarrow F(I, r, \alpha, \lambda; \theta_1)$
    $I \leftarrow I \cup \theta_1$
    $r_{\theta_1} \leftarrow 0$ and $e^{j\alpha(\theta_1)} \leftarrow \frac{\mathbf{a}^H(\theta_1)\mathbf{x}}{|\mathbf{a}^H(\theta_1)\mathbf{x}|}$
    counter $\leftarrow$ counter+1.
end while

Algorithm 4 C-LASSO$_h$

$\theta_0 \leftarrow \underset{\theta}{\operatorname{argmax}} |\mathbf{a}^H(\theta)\mathbf{x}|$
$I \leftarrow \{\theta_0\}$ and $\lambda \leftarrow |\mathbf{a}^H(\theta_0)\mathbf{x}|$
$r \leftarrow 0$ and $e^{j\alpha} \leftarrow \frac{\mathbf{a}^H(\theta_0)\mathbf{x}}{|\mathbf{a}^H(\theta_0)\mathbf{x}|}$
counter $\leftarrow 0$.
while counter $< n$ **do**
    $\hat{\mathbf{n}} \leftarrow \mathbf{x} - \mathbf{A}(I)\mathbf{s}$
    $\theta_1 \leftarrow \underset{\theta}{\operatorname{argmax}} |\mathbf{a}^H(\theta)\hat{\mathbf{n}}|$ s.t. $\theta \notin I$ is a local maximum.
    $p \leftarrow |\mathbf{a}^H(\theta_1)\hat{\mathbf{n}}|$
    if $p = \lambda$ **then**
        $I \leftarrow I \cup \theta_1$
        $r_{\theta_1} \leftarrow 0$ and $e^{j\alpha(\theta_1)} \leftarrow \frac{\mathbf{a}^H(\theta_1)\mathbf{x}}{|\mathbf{a}^H(\theta_1)\mathbf{x}|}$
        counter $\leftarrow$ counter+1.
    else
        $\lambda \leftarrow \mu\lambda + (1-\mu)p$
        $(I, r, \alpha) \leftarrow H(I, r, \alpha, \lambda)$
    end if
end while

---

achieves this value at the DOA estimates. As a graphical view, assume the graph $y = f(\theta)$ and the line $y = \lambda$ as the limit line touching the graph from above at DOA estimates. Decreasing λ gradually, the limit line decreases and the graph gets more compressed. Roughly speaking, some points in the graph resist decreasing forming a peak which eventually touches the limit line. At this RP value a new DOA at the new touching point is introduced, which later follows decreasing with the limiting line. Thus, we can check our distance to the next singular point when decreasing λ by looking at the highest peak point different to the estimates. If this peak is smaller than λ we can continue our pass with a local change of parameters. However, at the next singular point the new touching point should be added to the active solutions. Otherwise, the extra peak will be more than λ which indicates a wrong solution and a need to increase λ. Accordingly, we first introduce the following function:

3) Local Equilevel Attractor (H): Given a quadruple (I, r, α, λ), the function H gives the convergence point (I', r', α') of the Newton iterative solution of the system of equations given by the second line of (8) and (9), i.e.

$$\mathbf{a}^H(\theta_i')\hat{\mathbf{n}}' = \lambda \frac{s_i'}{|s_i'|}$$
$$\Re(e^{-j\alpha_i'}\mathbf{d}^H(\theta_i')\hat{\mathbf{n}}') = 0. \tag{15}$$

The function H can be computed fast using the Newton algorithm similar to the ML solution in Algorithm 1 substituting $\mathbf{J}_{\text{ML}}$ and η_{ML} by $\mathbf{J}_{\text{LEA}}$ and η_{LEA} in (14) and (13) respectively. The C-LASSO$_h$ algorithm can then be expressed as in Algorithm 4, where μ is a parameter setting the compromise between the speed and the probability of convergence.

IV. NUMERICAL RESULTS

Assuming a standard Uniform Linear Array , we compared the result of applying C-LASSO to the single snapshot DOA model with the ML and RELAX estimator. The RELAX method alternates between estimating different subsets of DOAs using the previous estimate of the complement subset. In this work we use a singleton subset at each iteration to estimate which is often the case in practice. The ML estimator is implemented by first an exhaustive search and then the local
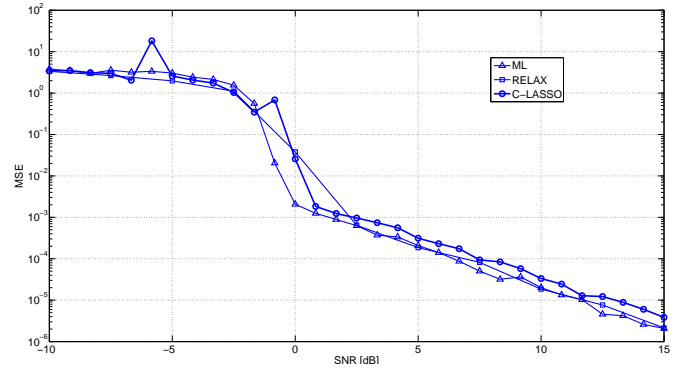

Fig. 1. Comparing MSE Vs SNR for 100 trials of single-snapshot data generated from a half wavelength ULA of m=15 sensors and 2 sources separated by the electrical angle $\frac{4\pi}{m}$.

Newton method solving (3). Figure 1 shows the estimation Mean Squared Error (MSE) for these three techniques in the case of two sources with $s_1 = s_2 = 1$ and separation $\Delta\phi = \frac{4\pi}{m}$, when the number of sensors m is 15. As expected, LASSO has an estimation error level higher than MSE due to its bias. The RELAX estimate gives closer solution in this case. However, the C-LASSO reaches the threshold region at slightly lower SNR compared to RELAX. Figure 2 shows the error result of LASSO compared to RELAX fixing SNR to 10 dB and varying the separation in terms of electrical angle in the previous scenario. It shows a typical behavior of LASSO in the high SNR case, in which very close sources are absorbed to one source so that the LASSO solution path does not contain any point with correct model order. This is shown as the "undefined region" in Figure 2. However, the LASSO solution shows instabilities until separation reaches the fundamental resolution [15].
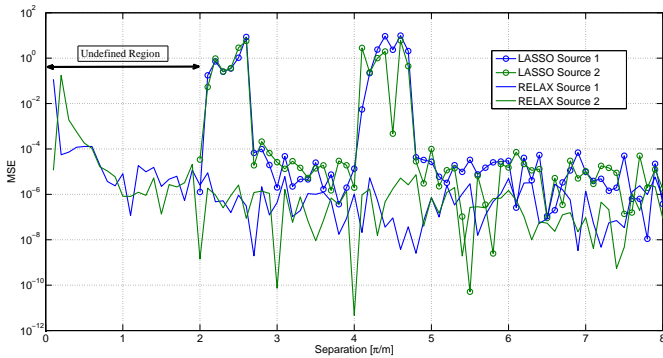
In another experiment, we compared the C-LASSO$_h$ and

Fig. 2. Comparing MSE Vs separation in electrical angles for 100 trials of single-snapshot data generated from a half wavelength ULA of m=15 sensors. SNR=10 dB.
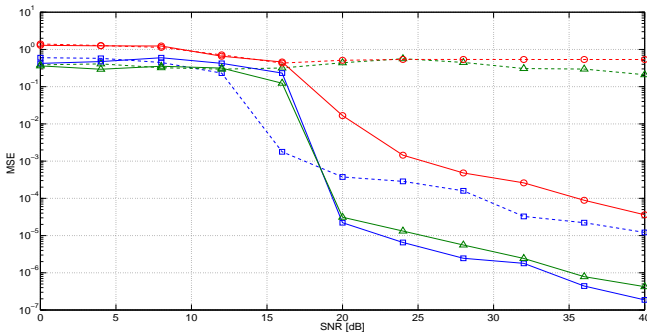

Fig. 3. C-LASSO performance compared to RELAX. The square, triangle, and circle markers show the DOA estimates of the first, second and third source respectively while the dashed lines show the result for RELAX.

RELAX performance in a more complicated case of three close sources at electrical angles $\left[\frac{-5\pi}{m} \; 0 \; \frac{3.5\pi}{m}\right]$ with $s_1 = s_2 = 1$ and $s_3 = 0.1j$. The SNR is measured by observing s_1 and n_1 when $\mu = 0.8$. The scenario is hard for the RELAX algorithm in the one snapshot case due to the different levels of sources. The results are shown in Figure 3. As can be seen, the RELAX technique can only resolve the reference signal s_1, while too wide dynamic range makes it hard to estimate the DOAs of the second and the third sources respectively. However, one may note the higher threshold SNR of C-LASSO$_h$. Note also that the realization of C-LASSO$_h$ took 60 times longer time than RELAX using MATLAB programming.

V. Conclusion

In this work, we introduced a LASSO realization technique, C-LASSO, in the one snapshot case and an improvement C-LASSO$_h$ from the complexity point of view which do not need any discretization. They are much faster than the convex programming realizations, such as the interior point technique considering the problem of RP selection. These algorithms are based on the idea of following the optimality point path of a set of generalized optimality conditions derived from the LASSO original ones.

The results show that in lower SNRs and more sources case, C-LASSO dominates ML from complexity point of view with-

out loosing much performance. Although there exists other DOA estimation techniques approximating the ML solution, we showed by simulation that the LASSO algorithm is more robust to the problem of wide dynamic range . It can be shown, while neglected in this work, that the LASSO bias is linear with the noise level independent of the true DOAs and sources.

The C-LASSO algorithms drawback is its computational time due to one dimensional search steps and the parameter μ in C-LASSO$_h$, which could not be decreased arbitrarily due to the convergence problem. However, sacrificing some performance, one may confine the search to a sufficiently fine grid neglecting the fine tuning step which can speed up the algorithm. Furthermore, while we have not presented the mathematical details here, LASSO encounters a consistency problem in the one snapshot case of the sources separated less than a fundamental resolution (see [15] for a similar argument).

Finally, it should be noted that similar to the Group-LASSO (G-LASSO) formalism, C-LASSO could be adapted to multiple-snapshot model. It is expected that due to the good performance of LASSO in high SNRs, the grouped C-LASSO also provides a robust technique.

References

[1] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B, (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[2] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289 –1306, april 2006.

[3] D. Malioutov, M. Cetin, and A. Willsky, "Source localization by enforcing sparsity through a laplacian prior: an svd-based approach," *IEEE Workshop on Statistical Signal Processing*, pp. 573 – 576, sept.-1 oct. 2003.

[4] A. Panahi and M. Viberg, "Maximum aposteriory based regularization parameter selection," in *International Conference on Acoustics, Speech, and Signal Processing*, 2011.

[5] T. Park and G. Casella, "The bayesian lasso," *Journal of the American Statistical Association*, vol. 103, p. 681 686, 2008.

[6] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx, Apr. 2011.

[7] B. Efron, T. Hastie, L. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, pp. 407–499, 2004.

[8] M. R. Osborne, B. Presnell, and B. Turlach, "A new approach to variable selection in least squares problems," 1999.

[9] A. Panahi and M. Viberg, "Fast candidate points selection in the lasso path," *Signal Processing Letters, IEEE*, vol. 19, no. 2, pp. 79 –82, feb. 2012.

[10] J. Fessler and A. Hero, "Space-alternating generalized expectation-maximization algorithm," *Signal Processing, IEEE Transactions on*, vol. 42, no. 10, pp. 2664 –2677, oct 1994.

[11] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *Signal Processing Magazine, IEEE*, vol. 13, no. 4, pp. 67 –94, jul 1996.

[12] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms; 3rd ed.* Newark, NJ: Wiley, 2006.

[13] D. Starer and A. Nehorai, "Newton algorithms for conditional and unconditional maximum likelihood estimation of the parameters of exponential signals in noise," *Signal Processing, IEEE Transactions on*, vol. 40, no. 6, pp. 1528 –1534, jun 1992.

[14] Z.-S. Liu, J. Li, and P. Stoica, "Relax-based estimation of damped sinusoidal signal parameters," *Signal Processing*, vol. 62, no. 3, pp. 311 – 321, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168497001321

[15] A. Panahi and M. Viberg, "On the resolution of the lasso-based doa estimation method," in *Smart Antennas (WSA), 2011 International ITG Workshop on*, feb. 2011, pp. 1 –5.