An Efficient Sparse Representation Algorithm for Direction-of-Arrival Estimation

Lei SUN, Huali WANG, Guangjie XU

College of Communications Engineering, PLA Univ. of Sci.&Tech., Nanjing, China

realmufeng@gmail.com, wanghl2008@gmail.com, moon990@163.com

Abstract. *This paper presents an efficient sparse representation approach to direction-of-arrival (DOA) estimation using uniform linear arrays. The proposed approach constructs the jointly sparse model in real domain by exploiting the properties of centro-Hermitian matrices. Subsequently, DOA estimation is realized via the sparse Bayesian learning (SBL) algorithm. Further, the pruning threshold of SBL is adaptively selected to speed up the basis pruning rate. Simulation results demonstrate that the proposed approach achieves an improved performance and enjoys computational efficiency as compared to the state-of-the-art l_1-norm-based DOA estimators especially in scenarios with small sample size and low signal-to-noise ratio.*

Keywords

Direction-of-arrival (DOA) estimation, uniform linear array, sparse representation, sparse Bayesian learning.

1. Introduction

Direction-of-Arrival (DOA) estimation has been widely used in radar, sonar, wireless communications, and other application fields. A vast number of algorithms have been devised for the DOA estimation problem. Subspace algorithms, such as MUSIC [1], ESPRIT [2], and their variants, are well known to have high resolution capabilities. However, all the subspace algorithms require the exact knowledge of the array signal model and consistent estimate of the noise or signal subspaces. The maximum likelihood (ML) approach is asymptotically optimal under certain regularity conditions [3]. Unfortunately, since the ML objective function is very flat far from the true DOAs, good initial guesses for the DOAs and an exact estimate of the number of the incident sources are required to ensure the global convergence.

Recently, exploiting the inherently spatial sparsity of the incident sources, DOA estimation has been formulated in the sparse representation framework [4]-[11]. The general concept of the sparsity-based DOA estimators is to directly reconstruct the spatial spectrum of the incident sources from multiple measurement vectors (MMV) subject to sparsity constraint. In [4], the global matched filter (GMF) approach is proposed to realize the DOA estimation based on sparsely representing the beamformer samples. The so-called l_1-SVD algorithm in [5] reduces the sample size by means of the singular value decomposition (SVD), and utilizes the l_1-norm penalty to enforce sparsity. The l_1-SVD algorithm has been further investigated in [6]-[8], where weighted l_1-norm penalties are adopted instead. Hyder and Mahata approximated the l_0-norm by a family of Gaussian functions, and proposed the joint $l_{2,0}$ approximation (JLZA) algorithm for direction finding [9]. However, the global convergence of JLZA is not guaranteed since there are lots of parameters have to be properly set. The literature [10] presents a method via sparse representation of array covariance vectors provided that large sample support is available. In the scenario where uncorrelated sources impinge on a uniform linear array (ULA), Xu et al. simplified the MMV model to single measurement vector (SMV) model [11]. Generally speaking, the DOA estimators based on sparse representation outperform conventional approaches particularly in small sample size, low signal-to-noise ratio (SNR) scenarios, and rely less on the priori information of the incident source number. Nonetheless, there are two major drawbacks of the popular l_1-norm-based approaches. First, the computational load of solving the l_1-norm optimization as a second-order cone program (SOCP) is quite expensive for some applications. Second, it is still nebulous to select the optimal regularization parameter.

In this paper, we present an efficient and accurate sparse representation approach to the DOA estimation using a ULA. We construct a real-valued jointly sparse model by taking advantage of the centro-Hermitian property. The sparse Bayesian learning (SBL) algorithm [12]-[14] is adopted to realize the DOA estimation. In addition, we speed up the basis pruning rate of SBL by adaptively setting the pruning threshold according to a rough estimate of the nonzero hyper-parameters. The proposed approach achieves high estimation accuracy with a low computational complexity as compared to several existing l_1-norm-based DOA estimators. Such advantages are demonstrated by simulation results.

The rest of the paper is organized as follows. Section 2 briefly reviews the formulation of the sparsity-based DOA estimators. Section 3 presents the proposed method.

Several numerical simulations are carried out in Section 4 to evaluate the performance of the proposed method. Finally, Section 5 concludes the whole paper.

Notation: Throughout the present work, we use $\mathbf{a}_{m.}$ and $\mathbf{a}_{.n}$ to denote the mth row and nth column of matrix $\mathbf{A}$, respectively. The operators $E(\bullet)$, $(\bullet)^*$, $(\bullet)^T$, $(\bullet)^H$, $(\bullet)^{-1}$, Re(•), Im(•), $diag(\bullet)$, $\|\bullet\|_2$, $\|\bullet\|_F$, and $\mathbf{I}_M$ denote expectation, conjugate, transpose, conjugate transpose, inverse, real part, imaginary part, diagonalization, Euclidean norm, Frobenius norm, and $M\times M$ identity matrix, respectively. j is reserved for the imaginary unit $\sqrt{-1}$.

2. Problem Formulation

Consider a ULA of M sensors receives signals from K far-field uncorrelated narrowband sources. The distance between adjacent sensors is $d = \lambda/2$, where λ is the wavelength of the source. The array observation vector at time t is modeled as

$$\begin{aligned}\mathbf{x}(t) &= \sum_{k=1}^{K} \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t) \\ &= \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad t = 1,2,\ldots,T\end{aligned} \tag{1}$$

where $s_k(t)$ is the baseband waveform of the kth source, $\mathbf{s}(t) = [s_1(t),\ldots,s_K(t)]^T$ is the source waveform vector, $\mathbf{n}(t)$ is the noise vector, T denotes the number of snapshots, the vector $\boldsymbol{\theta} = [\theta_1,\ldots,\theta_K]$ contains the actual DOAs of the incident sources, and the matrix $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1),\ldots,\mathbf{a}(\theta_K)]$ is commonly referred to as the array manifold matrix, whose kth column

$$\mathbf{a}(\theta_k) = \left[1, e^{-j\pi\sin(\theta_k)}, \ldots, e^{-j\pi(M-1)\sin(\theta_k)}\right]^T \tag{2}$$

is the steering vector of the kth source. In this paper, the source waveforms and the additive noise are assumed to be circularly symmetric, independent identically distributed (i.i.d.) complex Gaussian random processes with zero mean, and the noise at each sensor is uncorrelated with each other as well as the incident sources. Under this assumption, we obtain the following array covariance matrix:

$$\mathbf{R}_x = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}(\boldsymbol{\theta})\mathbf{R}_s\mathbf{A}^H(\boldsymbol{\theta}) + \sigma_n^2\mathbf{I}_M \tag{3}$$

where $\mathbf{R}_s = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$ is the source covariance matrix, and σ_n^2 represents the noise variance.

To formulate the DOA estimation in the sparse representation framework, the potential spatial scope of the incident sources is first sampled to form a direction grid $\boldsymbol{\Theta} = [\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_L]$ with $L >> K$ denoting the length of the grid. Correspondingly, the array manifold matrix $\mathbf{A}(\boldsymbol{\theta})$ is generalized to an $M\times L$ over-complete dictionary $\mathbf{B}$ in terms of the direction grid, as

$$\mathbf{B} = [\mathbf{a}(\bar{\theta}_1), \mathbf{a}(\bar{\theta}_2), \ldots, \mathbf{a}(\bar{\theta}_L)]. \tag{4}$$

Note that $\mathbf{B}$ is known and does not depend on $\boldsymbol{\theta}$. Suppose the grid is dense enough so that all elements of $\boldsymbol{\theta}$ lie in (or, close to) the grid, then (1) can be rewritten in an over-complete form as

$$\mathbf{X} = \mathbf{BH} + \mathbf{N} \tag{5}$$

where $\mathbf{X}=[\mathbf{x}(1),\ldots,\mathbf{x}(T)]$, $\mathbf{H}=[\mathbf{h}(1),\ldots,\mathbf{h}(T)]$, $\mathbf{h}(t)$ is a sparse vector whose lth element is nonzero and equal to $s_k(t)$ if source k comes from $\bar{\theta}_l$ and zero otherwise, and $\mathbf{N}=[\mathbf{n}(1),\ldots,\mathbf{n}(T)]$. If the incident sources are fixed over the observation period, it is clear that for all t the nonzero elements of $\mathbf{h}(t)$ occur at the same locations, such that only K rows of $\mathbf{H}$ are nonzero, each corresponding to a different source. The estimation of DOA for all sources then reduces to recoverying and determining the nonzero rows of the sparse matrix $\mathbf{H}$ from the snapshots matrix $\mathbf{X}$, and can be expressed as the following constrained convex optimization problem:

$$\min \|\mathbf{H}\|_{2,1} \quad \text{s.t.} \ \|\mathbf{X} - \mathbf{BH}\|_F \le \beta \tag{6}$$

where β is a fitting error threshold that must be set by user, and the objective $\|\mathbf{H}\|_{2,1}$ is defined as

$$\|\mathbf{H}\|_{2,1} = \sum_{l=1}^{L} \|\mathbf{h}_{l.}\|_2 \tag{7}$$

which is equal to the l_1-norm of the vector $[\|\mathbf{h}_{1.}\|_2,\ldots,\|\mathbf{h}_{L.}\|_2]^T$. The indices of nonzero rows of $\mathbf{H}$ give the DOA estimates. It is worthwhile to note that the recovery performance of $\mathbf{H}$ is heavily dependent on the choice of β.

3. The Proposed Method

In this section, an efficient DOA estimator based on sparse representation is proposed. First, we formulate a real-valued jointly sparse model for the direction finding problem. Next, we develop a solution mechanism to realize the DOA estimation.

3.1 Real-Valued Jointly Sparse Model

Consider the augmented sample matrix $\mathbf{Y}=[\mathbf{X}\ \boldsymbol{\Pi}_M\mathbf{X}^*\boldsymbol{\Pi}_T]$, where $\boldsymbol{\Pi}_M$ is an $M\times M$ exchange matrix with ones on its anti-diagonal and zeros elsewhere. It is shown that $\mathbf{Y}$ is centro-Hermitian, and can be transformed to a real-valued matrix according to [15]

$$\mathbf{Y}_r = \mathbf{Q}_M^H[\mathbf{X}\ \boldsymbol{\Pi}_M\mathbf{X}^*\boldsymbol{\Pi}_T]\mathbf{Q}_{2T} \tag{8}$$

where $\mathbf{Q}$ is a unitary matrix, defined as

$$\mathbf{Q}_{2n+1} = \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{I}_n & \mathbf{0} & j\mathbf{I}_n \\ \mathbf{0}^T & \sqrt{2} & \mathbf{0}^T \\ \boldsymbol{\Pi}_n & \mathbf{0} & -j\boldsymbol{\Pi}_n\end{bmatrix} \tag{9}$$

and the matrix $\mathbf{Q}_{2n}$ can be easily obtained from $\mathbf{Q}_{2n+1}$ by dropping its center row and center column. Without loss of generality, we assume the amount of the sensors $M = 2M'$ to be even. Note that (8) incorporates forward-backward averaging, and therefore the sample size is virtually doubled from T to $2T$. Observe that the over-complete dictionary $\mathbf{B}$ satisfies

$$\mathbf{B\Lambda} = \mathbf{\Pi}_M (\mathbf{B\Lambda})^* \tag{10}$$

where $\mathbf{\Lambda}$ is a $L \times L$ unitary diagonal matrix whose lth diagonal element is given by

$$\Lambda_{ll} = e^{j(M-1)\pi \sin \bar{\theta}_l / 2}, \quad 1 \le l \le L. \tag{11}$$

After some straightforward algebraic manipulations, we can rewrite (8) as below:

$$\begin{aligned} \mathbf{Y}_\mathrm{r} &= \mathbf{Q}_M^H([\mathbf{BH}\ \mathbf{\Pi}_M \mathbf{B}^* \mathbf{H}^* \mathbf{\Pi}_T] + [\mathbf{N}\ \mathbf{\Pi}_M \mathbf{N}^* \mathbf{\Pi}_T])\mathbf{Q}_{2T} \\ &= \sqrt{2}\mathbf{B}_\mathrm{r}[\mathrm{Re}(\mathbf{\Lambda}^*\mathbf{H})\ \ -\mathrm{Im}(\mathbf{\Lambda}^*\mathbf{H})] + \mathbf{N}_\mathrm{r} \\ &= \mathbf{B}_\mathrm{r}\mathbf{H}_\mathrm{r} + \mathbf{N}_\mathrm{r} \end{aligned} \tag{12}$$

where $\mathbf{B}_\mathrm{r} = \mathbf{Q}_M^H \mathbf{B\Lambda} = [\mathbf{a}_\mathrm{r}(\bar{\theta}_1), \ldots, \mathbf{a}_\mathrm{r}(\bar{\theta}_L)]$ with

$$\begin{aligned} \mathbf{a}_\mathrm{r}(\bar{\theta}_l) = \sqrt{2}[&\cos(\pi(M-1)\sin\bar{\theta}_l/2), \ldots, \cos(\pi \sin \bar{\theta}_l / 2), \\ &\sin(\pi(M-1)\sin\bar{\theta}_l/2), \ldots, \sin(\pi \sin\bar{\theta}_l/2)]^T \end{aligned} \tag{13}$$

and the noise term

$$\mathbf{N}_\mathrm{r} = \begin{bmatrix} \mathrm{Re}\{\mathbf{N}_1 + \mathbf{\Pi}_M \cdot \mathbf{N}_2^*\} & -\mathrm{Im}\{\mathbf{N}_1 - \mathbf{\Pi}_M \cdot \mathbf{N}_2^*\} \\ -\mathrm{Im}\{\mathbf{N}_1 + \mathbf{\Pi}_M \cdot \mathbf{N}_2^*\} & \mathrm{Re}\{\mathbf{N}_1 - \mathbf{\Pi}_M \cdot \mathbf{N}_2^*\} \end{bmatrix} \tag{14}$$

with $\mathbf{N}=[\mathbf{N}_1^T,\ \mathbf{N}_2^T]^T$. $\mathbf{H}_\mathrm{r} = \sqrt{2}[\mathrm{Re}(\mathbf{\Lambda}^*\mathbf{H})\ \ -\mathrm{Im}(\mathbf{\Lambda}^*\mathbf{H})]$ shows the same sparsity pattern as $\mathbf{H}$, therefore we can achieve the DOA estimates by locating the nonzero rows of $\mathbf{H}_\mathrm{r}$. The columns of $\mathbf{H}_\mathrm{r}$ and $\mathbf{N}_\mathrm{r}$ still satisfy the i.i.d condition under the circularly complex Gaussian assumption.

Before turning to solve (12), we first analyze $\mathbf{B}_\mathrm{r}$ from the sparse representation viewpoint. It has been shown that the accuracy of the sparse-recovery algorithms is determined by the coherence measure [16], [17] of the dictionary matrix. For the dictionary $\mathbf{B}_\mathrm{r}$, the coherence measure is given by

$$\mu(\mathbf{B}_\mathrm{r}) = \arg\max_{i \ne l} \frac{\left|\mathbf{a}_\mathrm{r}^H(\bar{\theta}_i)\mathbf{a}_\mathrm{r}(\bar{\theta}_l)\right|}{\left\|\mathbf{a}_\mathrm{r}(\bar{\theta}_i)\right\|_2 \left\|\mathbf{a}_\mathrm{r}(\bar{\theta}_l)\right\|_2} \tag{15}$$

where $\mathbf{a}_\mathrm{r}(\bar{\theta}_i)$ and $\mathbf{a}_\mathrm{r}(\bar{\theta}_l)$ denote the ith and lth columns of $\mathbf{B}_\mathrm{r}$, respectively. Since all the columns of $\mathbf{B}_\mathrm{r}$ have a constant Euclidean norm $\sqrt{M}$, we only have to take the numerator term of (15) into account. Then we get that

$$\begin{aligned} &\left|\mathbf{a}_\mathrm{r}^H(\bar{\theta}_i)\mathbf{a}_\mathrm{r}(\bar{\theta}_l)\right| \\ &= \left|e^{-j((M-1)/2)\pi(\sin\bar{\theta}_i - \sin\bar{\theta}_l)}\mathbf{a}^H(\bar{\theta}_i)\mathbf{Q}_M\mathbf{Q}_M^H\mathbf{a}(\bar{\theta}_l)\right| \\ &= \left|\mathbf{a}^H(\bar{\theta}_i)\mathbf{a}(\bar{\theta}_l)\right| \end{aligned} \tag{16}$$

where $\mathbf{a}(\bar{\theta}_i)$ and $\mathbf{a}(\bar{\theta}_l)$ are the ith and lth columns of $\mathbf{B}$, respectively. Considering that the Euclidean norm of each column of $\mathbf{B}$ is also equal to $\sqrt{M}$, it is straightforward to infer that the coherence measure of $\mathbf{B}_\mathrm{r}$ is the same as that of $\mathbf{B}$. As a result, the recovery accuracy of $\mathbf{H}_\mathrm{r}$ is guaranteed.

3.2 DOA Estimation via SBL

We utilize the iterative SBL algorithm to solve the real-valued sparse representation problem (12). The SBL only requires $o(M^2L)$ flops per iteration to solve (12), provided that $\mathbf{Y}_\mathrm{r}$ can be replaced by a $M \times \mathrm{rank}(\mathbf{Y}_\mathrm{r})$ matrix $\bar{\mathbf{Y}}_\mathrm{r}$ such that $\mathbf{Y}_\mathrm{r}\mathbf{Y}_\mathrm{r}^T = \bar{\mathbf{Y}}_\mathrm{r}\bar{\mathbf{Y}}_\mathrm{r}^T$. However, with the additive noise, the rank of $\mathbf{Y}_\mathrm{r}$ is always equal to M, thus the noise component is also reserved in $\bar{\mathbf{Y}}_\mathrm{r}$. Herein, to further mitigate the infection of the additive noise, we exploit the truncated SVD to replace $\mathbf{Y}_\mathrm{r}$:

$$\mathbf{Y}_\mathrm{rsv} = \mathbf{Y}_\mathrm{r}\mathbf{VD} \tag{17}$$

where $\mathbf{Y}_\mathrm{r}$=$\mathbf{ULV}$, $\mathbf{D}$=$[\mathbf{I}_K;\mathbf{0}]$, and $\mathbf{0}$ is a $(2T-K) \times K$ zero matrix. Define $\mathbf{H}_\mathrm{rsv}$=$\mathbf{H}_\mathrm{r}\mathbf{VD}$ and $\mathbf{N}_\mathrm{rsv}$=$\mathbf{N}_\mathrm{r}\mathbf{VD}$, then (12) is simplified to the following form:

$$\mathbf{Y}_\mathrm{rsv} = \mathbf{B}_\mathrm{r}\mathbf{H}_\mathrm{rsv} + \mathbf{N}_\mathrm{rsv}. \tag{18}$$

Obviously, $\mathbf{H}_\mathrm{rsv}$ and $\mathbf{H}_\mathrm{r}$ have the same row support. Like the one did in [8], we exploit the property that

$$\mathbf{n}_\mathrm{rsv} = \mathbf{y}_\mathrm{rsv} - \mathbf{B}_\mathrm{r}\mathbf{h}_\mathrm{rsv} \tag{19}$$

is an additive real Gaussian random vector with zero mean and covariance matrix $\mathbf{R}_\mathrm{rn} = \sigma_n^2\mathbf{I}_M$, where $\mathbf{n}_\mathrm{rsv}$, $\mathbf{y}_\mathrm{rsv}$, $\mathbf{h}_\mathrm{rsv}$ corresponds to the same column in $\mathbf{N}_\mathrm{rsv}$, $\mathbf{Y}_\mathrm{rsv}$, $\mathbf{H}_\mathrm{rsv}$, respectively. Then $\mathbf{y}_\mathrm{rsv}$ can be modeled through a Gaussian likelihood function:

$$p(\mathbf{y}_\mathrm{rsv} \mid \mathbf{h}_\mathrm{rsv}, \sigma_n^2) = \frac{1}{(2\pi\sigma_n^2)^{M/2}}\exp\left(-\frac{\left\|\mathbf{y}_\mathrm{rsv} - \mathbf{B}_\mathrm{r}\mathbf{h}_\mathrm{rsv}\right\|_2^2}{2\sigma_n^2}\right). \tag{20}$$

In a probabilistic perspective, the sparse representation problem can be treated as finding $\mathbf{h}_\mathrm{rsv}$ and σ_n^2 to maximize the posterior probability $p(\mathbf{h}_\mathrm{rsv}, \sigma_n^2 \mid \mathbf{y}_\mathrm{rsv})$ under sparsity constraint. Regarding the SBL algorithm, sparsity is enforced via invoking a Gaussian prior over each element of $\mathbf{h}_\mathrm{rsv}$. By combining each of these priors, we have

$$p(\mathbf{h}_\mathrm{rsv} \mid \boldsymbol{\gamma}) = \prod_{l=1}^{L}\frac{1}{\sqrt{2\pi\gamma_l}}\exp\left(-\frac{(h_l^\mathrm{rsv})^2}{2\gamma_l}\right) \tag{21}$$

where $\boldsymbol{\gamma}$=$[\gamma_1,\ldots,\gamma_L]^T$ is a hyper-parameter vector with γ_l controlling the variance of the corresponding element h_l^rsv. If $\gamma_l = 0$, it means the h_l^rsv will be zero. If $\gamma_l > 0$, it indicates a nonzero h_l^rsv whose magnitude depends on γ_l. Using the Bayesian inference, the posterior distribution over $\mathbf{h}_\mathrm{rsv}$:

$$p(\mathbf{h}_\mathrm{rsv} \mid \mathbf{y}_\mathrm{rsv}, \boldsymbol{\gamma}, \sigma_n^2) = \frac{p(\mathbf{y}_\mathrm{rsv} \mid \mathbf{h}_\mathrm{rsv}, \sigma_n^2)\,p(\mathbf{h}_\mathrm{rsv} \mid \boldsymbol{\gamma})}{p(\mathbf{y}_\mathrm{rsv} \mid \boldsymbol{\gamma}, \sigma_n^2)} \tag{22}$$

turns out to be an analytically multivariate Gaussian distribution with the mean and covariance given by

$$\begin{aligned}\mathbf{u} &= \boldsymbol{\Gamma}\mathbf{B}_{\mathrm{r}}^{T}\boldsymbol{\Sigma}_{\mathrm{yr}}^{-1}\mathbf{y}_{\mathrm{rsv}} \\ \boldsymbol{\Sigma} &= \boldsymbol{\Gamma} - \boldsymbol{\Gamma}\mathbf{B}_{\mathrm{r}}^{T}\boldsymbol{\Sigma}_{\mathrm{yr}}^{-1}\mathbf{B}_{\mathrm{r}}\boldsymbol{\Gamma}\end{aligned} \tag{23}$$

where $\boldsymbol{\Gamma} = diag(\boldsymbol{\gamma})$ and $\boldsymbol{\Sigma}_{\mathrm{yr}} = \sigma_n^2\mathbf{I}_M + \mathbf{B}_{\mathrm{r}}\boldsymbol{\Gamma}\mathbf{B}_{\mathrm{r}}^T$. The posterior mean $\mathbf{u}$ is enlisted as a point estimate of $\mathbf{h}_{\mathrm{rsv}}$. As seen in (23), the recovery of $\mathbf{h}_{\mathrm{rsv}}$ reduces to estimate the unknown $\boldsymbol{\gamma}$ and σ_n^2. By summing up the contributions of the individual column $\mathbf{y}_{.k}^{\mathrm{rsv}}$, we finally get the following marginal log-likelihood function:

$$\begin{aligned}L(\boldsymbol{\gamma}, \sigma_n^2) &\cong \sum_{k=1}^{K}\log\left[p\left(\mathbf{y}_{.k}^{\mathrm{rsv}}\middle|\boldsymbol{\gamma}, \sigma_n^2\right)\right] \\ &\propto K\log\left|\boldsymbol{\Sigma}_{\mathrm{yr}}\right| + \sum_{k=1}^{K}\left(\mathbf{y}_{.k}^{\mathrm{rsv}}\right)^T\boldsymbol{\Sigma}_{\mathrm{yr}}^{-1}\mathbf{y}_{.k}^{\mathrm{rsv}}.\end{aligned} \tag{24}$$

An iterative *type-II* ML procedure [12], using the EM algorithm, is employed to minimize (24) with respect to $\boldsymbol{\gamma}$ and σ_n^2. The E-step is achieved by calculating the posterior moment using (23). For the M-step, the hyper-parameter vector $\boldsymbol{\gamma}$ and the noise variance σ_n^2 are updated:

$$\begin{aligned}\gamma_l^{\mathrm{new}} &= \frac{\left\|\mathbf{u}_{l.}\right\|_2^2}{K(1-\gamma_l^{-1}\Sigma_{ll})}, \quad \forall l = 1,\ldots,L \\ (\sigma_n^2)^{\mathrm{new}} &= \frac{\left\|\mathbf{Y}_{\mathrm{rsv}} - \mathbf{B}_{\mathrm{r}}\mathbf{U}\right\|_F^2}{K(M - L + \sum_{l=1}^{L}\gamma_l^{-1}\Sigma_{ll})}\end{aligned} \tag{25}$$

where the matrix $\mathbf{U} = [\mathbf{u}_{.1}, \mathbf{u}_{.2}, \ldots, \mathbf{u}_{.K}]$ is comprised with the posterior mean vector corresponding to each column of $\mathbf{H}_{\mathrm{rsv}}$, and Σ_{ll} denotes the lth diagonal element of the posterior covariance matrix $\boldsymbol{\Sigma}$. When a convergence criterion is satisfied, we attain the estimates of the sparse $\boldsymbol{\gamma}$ and noise variance σ_n^2. The DOA estimation problem can be immediately accomplished by deciding the locations of the relatively K large values of $\boldsymbol{\gamma}$.

3.3 Discussions

The noise variance balances the recovery sparsity and accuracy of the SBL algorithm. It plays an important role like β in (6). However, it is demonstrated that the explicit learning rule of noise variance in (25) is not robust in strongly noisy cases. Based on the connection between the singular value and eigenvalue, we can get a maximum likelihood estimate (denoted by $\hat{\sigma}_n^2$) of σ_n^2:

$$\hat{\sigma}_n^2 = \frac{1}{2T(M-K)}\sum_{m=K+1}^{M}\lambda_m^2 \tag{26}$$

where λ_m is the mth singular value of $\mathbf{Y}_{\mathrm{r}}$. As the authors recommended in [14], we incorporate fixed $\hat{\sigma}_n^2$ into the update rules of SBL.

Since a sparse prior is invoked over each row of $\mathbf{H}_{\mathrm{rsv}}$, it is observed that many elements of $\boldsymbol{\gamma}$ tend to zero in the iteration process, and only a relatively small set of $\boldsymbol{\gamma}$, which correspond to the incident sources, remain relatively large. A small fixed threshold (e.g., 10^{-5}) is suggested to be set such that, when any element of $\boldsymbol{\gamma}$ is below the threshold, it is pruned from the model. Herein, we propose to adaptively set the threshold according to the average value of the nonzero elements of $\boldsymbol{\gamma}$. When the SBL converges, it is shown that γ_l satisfies

$$\gamma_l = \frac{1}{K}\left\|\mathbf{H}_{l.}^{\mathrm{rsv}}\right\|_2^2 . \tag{27}$$

Note that $\left\|\mathbf{H}_{l.}^{\mathrm{rsv}}\right\|_2^2$ reflects $2TK$ times the average power of the source from $\bar{\theta}_l$ [8]. Straightforwardly, we have

$$\begin{aligned}\sum_{l=1}^{L}\gamma_l &= \frac{1}{K}\sum_{l=1}^{L}\left\|\mathbf{H}_{l}^{\mathrm{rsv}}\right\|_2^2 \\ &= \frac{1}{M}\sum_{m=1}^{M}\lambda_m^2 - 2T\hat{\sigma}_n^2 = 2T\mathbf{P}_{\mathrm{total}}\end{aligned} \tag{28}$$

where $\mathrm{P}_{\mathrm{total}}$ denotes the average total power of all incident sources. Since only K elements of $\boldsymbol{\gamma}$ associated with the incident sources are nonzero, then:

$$\gamma_{\min} \le 2T\mathrm{P}_{\mathrm{total}} / K \le \gamma_{\max} \tag{29}$$

where $\gamma_{\min}$ and $\gamma_{\max}$ denote the minimal and maximal nonzero elements of $\boldsymbol{\gamma}$, respectively. It is rational to assume that the difference between $\gamma_{\min}$ and $\gamma_{\max}$ is modest, thus $\gamma_{\mathrm{a}}=2T\mathrm{P}_{\mathrm{total}}/K$ can be treated as the average value of nonzero elements of $\boldsymbol{\gamma}$. Therefore, we set the small-magnitude elements of $\boldsymbol{\gamma}$ to zero using the thresholding operation:

$$H(\gamma_i) = \begin{cases} 0, & \text{if } 10\log(\gamma_i / \gamma_{\mathrm{a}}) < \tau \\ \gamma_i, & \text{otherwise} \end{cases} \tag{30}$$

after one iteration. As a consequence, elements of $\boldsymbol{\gamma}$ that are more than τ dB down from γ_{a}, as well as the corresponding columns in $\mathbf{B}_{\mathrm{r}}$, are pruned from the model. We shall point out that this simple modification leads to an improved basis pruning rate in practice.

In the proposed method, we use the information about the number of sources K. Practically, this information is often estimated via the classical AIC or MDL criteria [18]. It is known that incorrect estimation of the number of the incident sources degrades the performance of the l_1-SVD algorithm. For the proposed method, this degradation is alleviated, since the forward-backward averaging improves the estimation accuracy of K.

Remark: Considering the computational complexity of the proposed method, it only requires real additions to transform the complex sample matrix into a real one [15], and the calculation of SVD is about $o(M^3)$. Note that $L >> M > K$, then the recovery procedure based on SBL dominates the complexity analysis. Since the computational load of SBL per iteration is fixed, the overall complexity rests with the number of iterations. In moderate SNR cases, the SBL takes less than 25 iterations provided that both the noise variance and the pruning threshold are set as we suggested. Regarding the l_1-SVD algorithm calculated by SOCP software package SeDuMi [19], the computational cost per iteration is $o((KL)^3)$, which denotes complex operations, and the theoretical worst-case bound on the number of iterations is $o((KL)^{0.5})$. For the l_1-ACCV algorithm in [11], the factor K is removed, leading to a considerable reduction compared respect to l_1-SVD. Based on the previous analysis, we can conclude that the proposed method is more efficient than both l_1-SVD and l_1-ACCV.

4. Simulations

In this section, we carry out numerical simulations to evaluate the proposed method. We also compare it with the l_1-SVD in [5], the NSW-l_1 in [6], and the l_1-ACCV in [11]. In all the simulations, a 10-element ULA with half-wave-length spacing is used. The root-mean-square-error (RMSE) of the DOA estimates obtained with 500 Monte Carlo runs is defined as

$$\text{RMSE} = \sqrt{\frac{1}{500K}\sum_{k=1}^{K}\sum_{n=1}^{500}\left(\hat{\theta}_{kn}-\theta_k\right)^2} \tag{31}$$

where $\hat{\theta}_{kn}$ represents the estimate of θ_k in the nth trial. Each narrowband source is generated from a zero mean Gaussian distribution. The additive Gaussian noise is assumed to be white, both spatially and temporally, and uncorrelated with the sources. The SNR is defined as $10\log(\sigma_s^2/\sigma_n^2)$ with σ_s^2 and σ_n^2 denoting the source and noise power, respectively. The number of incident sources is estimated by MDL criterion unless otherwise stated. The spatial grid is uniform in the range −89° to 90° with 1° interval, which means that $L = 180$. We set the pruning threshold of the proposed method to −20 dB. The confidence interval is set to 0.99 for both l_1-SVD and NSW-l_1. The regularization parameter of l_1-ACCV is set according to the suggested value in [11].

Suppose four equal-power uncorrelated sources impinge on the array from [−20.4°, −12°, 15.6°, 30°]. The SNR is −5 dB, and 100 snapshots are collected. Fig. 1 depicts the normalized spatial spectra of various algorithms. Although the DOAs of the first and the fourth signal are not in the spatial grid, it is observed from Fig. 1 that all the algorithms can approximately locate them. However, spurious peaks exist in the spectrum of l_1-SVD.

To see more clearly the capability of the proposed method in resolving closely spaced sources, we plot the bias of the DOA estimates against angle separation between two sources in Fig. 2. We simulate two equal-power uncorrelated sources impinging from −20° and −20°+Δθ, where Δθ varies from 2° to 30°. The SNR is 0 dB and the number of snapshots is 100. It can be seen that the proposed method shows the best performance. The reason is that the SBL provides a tighter approximation to the l_0-norm than the l_1-norm [20], [21].

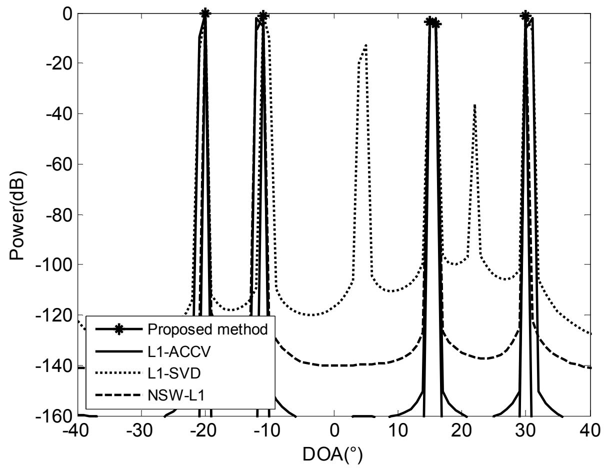

Fig. 1. Spatial spectra obtained by different algorithms.

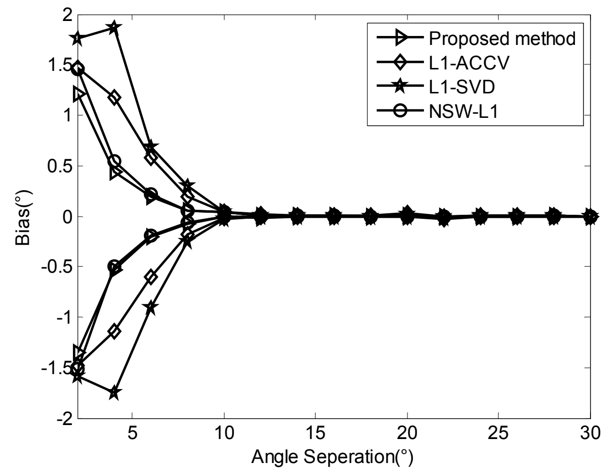

Fig. 2. Bias of the DOA estimates versus angle separation.

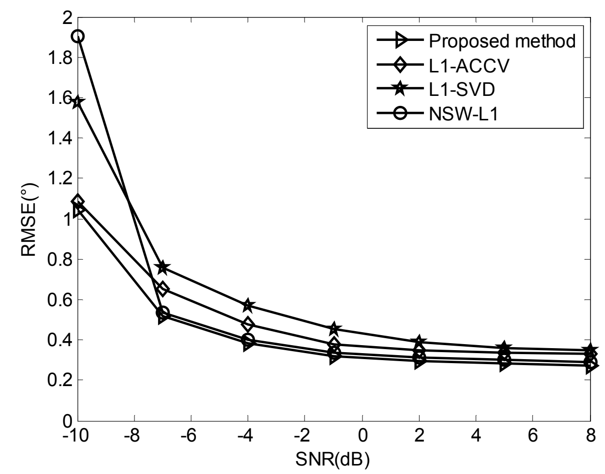

Fig. 3. RMSE of the DOA estimates against input SNR with 200 snapshots.

Next, the performance of these four algorithms is statistically compared for incident sources of various SNRs and sample sizes. Here we simulate four equal-power uncorrelated sources located at [−20.4°, −12°, 15.6°, 30°]. First, we keep the snapshots fixed at 200 and vary the SNR of all sources from −10 dB to 8 dB. Then the SNR is fixed at −5 dB and the number of snapshots is varied from 50 to 300. The RMSE performance is depicted in Fig. 3 and Fig. 4, respectively. It can be seen from the two figures that the proposed method outperforms the other three algorithms in the low SNR and small sample size cases.

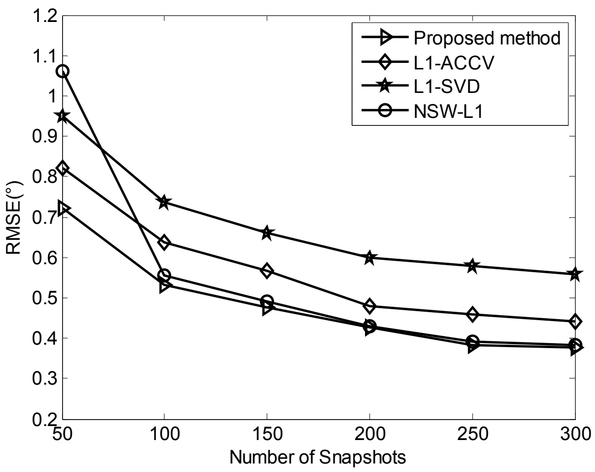

Fig. 4. RMSE of the DOA estimates against the number of snapshots with −5 dB SNR.

Finally, we present an evaluation of the computational complexity using the TIC and TOC instruction in MATLAB. Since the NSW-l_1 performs the same computational complexity as the l_1-SVD, we only choose l_1-SVD for the comparison. Four equal-power uncorrelated sources are simulated to impinge on the array from [−20°, −12°, 10°, 30°].The number of incident sources is assumed to be known beforehand. The SNR is set at −5 dB and 0 dB, and the sample size is set at 50, 100 and 200. For each SNR-sample size pair, the computation time of these three algorithms are averaged over 500 trials, and the results are illustrated in Tab. 1.

SNR (dB)	Sample Size	Computation Time (sec)		
		l_1-SVD	l_1-ACCV	Proposed method
−5	50	0.436	0.275	0.050
	100	0.448	0.276	0.032
	200	0.464	0.278	0.021
0	50	0.463	0.270	0.026
	100	0.484	0.272	0.017
	200	0.506	0.279	0.014

Tab. 1. Computation time comparison of various algorithms.

As can be seen in Tab. 1, the computation time of the proposed method is much smaller than those two l_1-norm based algorithms. Furthermore, the computation time of l_1-SVD and l_1-ACCV increases slightly with sample size. It is worth noting the computation time of the proposed method turns out to be reduced when the SNR and sample size increase. Consequently, the proposed method should be preferred for practical applications.

5. Conclusion

In this paper, we propose a DOA estimator based on SBL with real-valued processing, which exploits the centro-Hermitian property of the ULA. An approach to adaptively choose the pruning threshold is also presented to speed up the basis pruning rate of the SBL. The simulation results show that the proposed method provides an improved performance in comparison with the current state-of-the-art l_1-norm-based DOA estimators with reduced computational complexity. It is worthwhile to note that the proposed approach is not confined to ULA, but can be applied to arrays with centro-symmetric configuration.

Acknowledgements

This work is supported by the National Natural Science Foundation under Grant 61271354 of the People's Republic of China. The authors are grateful to D. Malioutov for providing the source codes of l_1-SVD.

References

[1] SCHMIDT, R. Multiple emitter location and signal parameters estimation. *IEEE Transactions on Antennas and Propagation*, 1986, vol. 34, no. 3, p. 276–280.

[2] ROY, R., KAILATH, K. ESPRIT—Estimation of signal parameter via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1989, vol. 37, no. 7, p. 984–995.

[3] STOICA, P., NEHORAI, A. MUSIC, maximum likelihood, and Cramer–Rao bound. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1989, vol. 37, no. 5, p. 720–741.

[4] FUCHS, J. J. On the application of the global matched filter to DOA estimation with uniform circular arrays. *IEEE Transactions on Signal Processing*, 2001, vol. 49, no. 4, p. 702–709.

[5] MALIOUTOV, D., ÇETIN, M., WILLSKY, A. S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Transactions on Signal Processing*, 2005, vol. 53, no. 8, p. 3010-3022.

[6] ZHENG, C.-D., LI, G., ZHANG, H., WANG, X.-Q. An approach of DOA estimation using noise subspace weighted l_1 minimization. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*. Prague (Czech), 2011, p. 2856–2859.

[7] ZHENG, C.-D., LI, G., LIU Y.-M., WANG, X.-Q. Subspace weighted $l_{2,1}$ minimization for sparse signal recovery. *EURASIP Journal on Advances in Signal Processing 2012*, 2012:98.

[8] XU, X., WEI, X.-H., YE, Z.-F. DOA estimation based on sparse signal recovery utilizing weighted l_1-norm penalty. *IEEE Signal Processing Letters*, 2012, vol. 19, no. 3, p. 155-158.

[9] HYDER, M. M., MAHATA, K. Direction-of-Arrival estimation using a mixed $l_{2,0}$ norm approximation. *IEEE Transactions on Signal Processing*, 2010, vol. 58, no. 9, p. 4646–4655.

[10] YIN, J.-H, CHEN, T.-Q. Direction-of-arrival estimation using a sparse representation of array covariance vectors. *IEEE Transactions on Signal Processing*, 2011, vol. 59, no. 9, p. 4489–4493.

[11] XU, D.-Y, HU, N., YE, Z.-F., BAO, M. The estimate for DOAs of signals using sparse recovery method. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*. Kyoto (Japan), 2012, p. 2573-2576.

[12] TIPPING, M. E. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 2001, vol. 1, p. 211–244.

[13] WIPF, D. P., RAO, B. D. Sparse Bayesian learning for basis selection. *IEEE Transactions on Signal Processing*, 2004, vol. 52, no. 8, p. 4036-4048.

[14] WIPF, D. P., RAO, B. D. An empirical Bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE Transactions on Signal Processing*, 2007, vol. 55, no. 7, p. 3704-3716.

[15] HAARDT, M., NOSSEK, J. A. Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden. *IEEE Transactions on Signal Processing*, 1995, vol. 43, no. 5, p. 1232–1242.

[16] DONOHO, D. L., ELAD, M., TEMLYAKOV, V. Stable recovery of sparse over-complete representations in the presence of noise. *IEEE Transactions on Information Theory*, 2006, vol. 52, no. 1, p. 6-18.

[17] RAUHUT, H., SCHNASS, K., VANDERGHEYNST, P. Compressed sensing and redundant dictionaries. *IEEE Transactions on Information Theory,* 2008, vol. 54, no. 5, p. 2210–2219.

[18] WAX, M., KAILATH, T. Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 1995, vol. 33, no. 4, p. 387-392.

[19] STURM, J. F. Using SeDuMi1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 1999, vol. 11, no. 1, p. 625–653.

[20] WIPF, D. P., RAO, B. D. l_0-norm minimization for basis selection. In *Advances in Neural Information Processing Systems (NIPS17)*, 2005. [Online]. Available at: http://books.nips.cc/papers/files/nips17/NIPS2004_0819.pdf

[21] WIPF, D. P., RAO, B. D. Comparing the effects of different weight distributions on finding sparse representations. In *Advances in Neural Information Processing Systems (NIPS 18),* 2006. [Online]. Available at: http://books.nips.cc/papers/files/nips18/NIPS2005_0158.pdf

About Authors ...

Lei SUN was born in Anhui, China. He received his Bachelor's degree in System Engineering from the PLA University of Science and Technology, Nanjing, in 2007. His research interests include array signal processing and sparse representation. Currently, he is working towards the Ph.D degree in the PLA University of Science and Technology, Nanjing, China.

HuaLi WANG was born in Zhejiang, China. He received his Ph.D. degree in fusee technology from the University of Science and Technology, Nanjing, China, in 1997. His research interests include array signal processing and compressive sensing. Currently, he is a professor with the PLA University of Science and Technology, Nanjing, China. Prof. Wang is a member of the IEEE.

Guangjie XU was born in Jiangsu, China. He received his Bachelor's degree in Electronic Engineering from the PLA University of Science and Technology, Nanjing, in 2009. His research interests include compressive sampling and signal detection. Currently, he is working towards the Ph.D degree in the PLA University of Science and Technology, Nanjing, China.