Direction of Arrival Estimation for Complex Sources through ℓ_1 Norm Sparse Bayesian Learning

Hua Bai, *Student Member, IEEE*, Marco F. Duarte, *Senior Member, IEEE*, and Ramakrishna Janaswamy, *Fellow, IEEE*

***Abstract*—In this letter, Laplace distribution is used to model the prior for the direction of arrival (DoA) of sources. In order to incorporate the real and imaginary part of the received signal, we propose a method that pairwise estimates the hyperparameters for parts of the signal coefficients. In addition, we propose a multi-task algorithm to extend the application of our method to the situation where multiple measurements are available. Non-uniform linear arrays are used to demonstrate the validity and advantages of the proposed method including its improved efficiency and accuracy compared with state-of-art DoA estimation methods.**

***Index Terms*—non-uniform array, Laplace prior, complex sources, sparse Bayesian learning**

I. Introduction

DOA estimation using linear arrays is a topic of active research. The general DoA problem consists of using an M-element antenna array with known geometry to estimate the bearing of N far-field incident sources, where $M > N$. During the past decades, several methods have been introduced and applied to the DoA estimation, such as MUSIC algorithm, ESPRIT and maximum likelihood approach [1]–[5]. In [6] and [7], sparse Bayesian learning (SBL), also known as relevance vector machine (RVM) [8] was used. When using RVM, the DoA estimation problem is solved by converting it to an optimization problem with ℓ_2 norm regularization, where a Gaussian prior with zero mean is selected for the sources to be estimated. Compared with Gaussian prior, Laplace priors are known to enforce the sparsity constraint more heavily [9].

The contributions of this letter are as follows. First, the DoA problem is formulated as an optimization problem with ℓ_1 norm penalty and solved with the fast marginal likelihood maximization algorithm in order to reduce the computational complexity compared with [10]. Second, the real and imaginary part of the complex signal are separated to yield a real-valued problem; a pairwise algorithm is proposed to leverage the relationship between the real and imaginary part of complex sources during the estimation. Third, in order to extend the applicability of the approach to the situation where the multiple snapshots are available, we propose a multi-task ℓ_1 norm SBL method. The robustness of the proposed method is improved with multiple snapshots, especially under noisy environments. The simulation result demonstrates that the proposed approach improves the efficiency and estimation accuracy.

II. DOA problem formulation

A non-uniform linear array, shown in Fig. 1, is comprised of M identical elements with the ith element located at distance l_i from the origin. The incoming sources, which are mutually independent, are denoted by s_j, j = 1,2,...,N, and the incident angles of the incoming sources are denoted by $\boldsymbol{\theta} = \{\theta_1, ..., \theta_N\}$, where $\theta_i \in$ [-90°, 90°], i = 1,2,...,N. Independent complex white noise is also present in the element observations, denoted by n_i, i = 1,...,M. For the far-field sources, the approximate phase variation of the received signals for the jth source at two elements separated by Δl is equal to $k\Delta l \sin\theta_j$, where $k = \frac{2\pi}{\lambda}$ and λ is the wavelength.

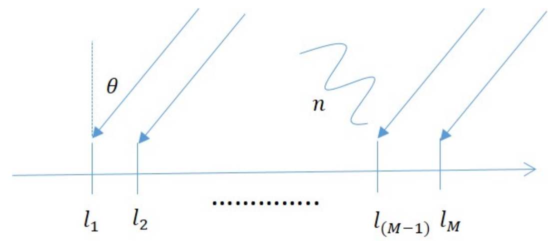


Fig. 1: M-element nonuniform linear array

The received signal of the ith element, y_i, is written as

$$y_i = \sum_{j=1}^{N} e^{-jkl_i \sin\theta_j} s_j + n_i, i = 1, 2, ..., M. \quad (1)$$

In matrix form (1) is written as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{n}, \quad (2)$$

where $\boldsymbol{y} = [y_1, y_2....y_M]^T \in \mathbb{C}^{M\times1}$, $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ denotes the measurement with $\boldsymbol{A}(m, n) = e^{-jkl_m \sin\theta_n}$, $\boldsymbol{s} = [s_1, s_2...s_N]^T \in \mathbb{C}^{N\times1}$ and $\boldsymbol{n} = [n_1, n_2, ..., n_M]^T \in \mathbb{C}^{M\times1}$ is subject to a circularly symmetric complex normal distribution with zero mean and covariance matrix $\Gamma = 2\sigma^2\mathbf{I} \in \mathbb{R}^{M\times M}$, σ^2 is unknown. We allow for the position l_m to change with time, and matrix $\boldsymbol{A}$ will change accordingly. ℓ_1 norm automatically generates a sparse solution [11] for the source estimation problem we have in mind

$$\widetilde{\boldsymbol{s}} = \underset{\boldsymbol{s}}{\operatorname{argmax}} \left\{-||\boldsymbol{y} - \boldsymbol{A}\boldsymbol{s}||_2^2 - \lambda||\boldsymbol{s}||_1\right\}, \quad (3)$$

where $||\cdot||_p$ represents ℓ_p norm. A method with computational complexity $\mathcal{O}(K^3)$ was proposed in [10] to estimate $\boldsymbol{s}$, where K denotes the size of the sampling grid and is typically greater than N. By choosing the appropriate prior for $\boldsymbol{s}$, we can estimate $\boldsymbol{s}$ under the Bayesian learning framework and the computational complexity can be reduced sharply by estimating the sources on the grid individually [12]. According to Bayes' rule, we have the posterior probability in logarithmic terms

$$\log(\text{posterior}) \propto \log(\text{likelihood}) + \log(\text{prior}). \tag{4}$$

Assuming that the received signal is contaminated with white noise, the likelihood of the received signal is

$$P(\boldsymbol{y}|\boldsymbol{s},\sigma^2) = \mathcal{CN}(\boldsymbol{y}|\boldsymbol{A}\boldsymbol{s}, 2\sigma^2\boldsymbol{I}), \tag{5}$$

where $\mathcal{CN}(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma})$ represents complex normal distribution on $\boldsymbol{y}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Its logarithm is

$$\log P(\boldsymbol{y}|\boldsymbol{s},\sigma^2) = -\frac{1}{2\sigma^2}||\boldsymbol{y}-\boldsymbol{A}\boldsymbol{s}||_2^2 \tag{6}$$

where the constant terms are ignored. Choosing the Laplace prior for the sources, we can obtain the prior distribution in logarithmic form by

$$\log P(s_i) = -\lambda |s_i| \ , \tag{7}$$

where the constant terms are ignored and λ is the scalar parameter. The source can be estimated by maximizing the posterior probability. However the number and directions of actual sources are unknown. To facilitate this the sampling grid which is a vector consisting of the possible angles $\tilde{\boldsymbol{\theta}} = [\tilde{\theta}_1, ..., \tilde{\theta}_K] \in \mathbb{R}^{K\times 1}$ is created. The angles of the actual sources are assumed to be in the grid. The source corresponding to $\tilde{\boldsymbol{\theta}}$ is denoted by $\tilde{\boldsymbol{s}} \in \mathbb{C}^{K\times 1}$, the only non-zero entries of $\tilde{\boldsymbol{s}}$ appear at angles the sources arrive from.

III. COMPLEX SOURCE ESTIMATION

In [6], the real and imaginary part of sources are separated during estimation. Namely, (2) is rewritten as

$$\bar{\boldsymbol{y}} = \bar{\boldsymbol{A}}\bar{\boldsymbol{s}} + \bar{\boldsymbol{n}}, \tag{8}$$

where $\bar{\boldsymbol{y}} = \begin{bmatrix}\Re(\boldsymbol{y})\\ \Im(\boldsymbol{y})\end{bmatrix}$, $\bar{\boldsymbol{s}} = \begin{bmatrix}\Re(\tilde{\boldsymbol{s}})\\ \Im(\tilde{\boldsymbol{s}})\end{bmatrix}$, $\bar{\boldsymbol{n}} = \begin{bmatrix}\Re(\boldsymbol{n})\\ \Im(\boldsymbol{n})\end{bmatrix}$ and

$$\bar{\boldsymbol{A}} = \begin{bmatrix}\Re(\widetilde{\boldsymbol{A}}) & -\Im(\widetilde{\boldsymbol{A}})\\ \Im(\widetilde{\boldsymbol{A}}) & \Re(\widetilde{\boldsymbol{A}})\end{bmatrix}, \tag{9}$$

where $\widetilde{\boldsymbol{A}} \in \mathbb{C}^{M\times K}$ is computed similarly to $\boldsymbol{A}$ but using all angles in the grid. The likelihood becomes

$$P(\bar{\boldsymbol{y}}|\bar{\boldsymbol{s}},\sigma^2) = \mathcal{N}(\bar{\boldsymbol{y}}|\bar{\boldsymbol{A}}\bar{\boldsymbol{s}}, \sigma^2\boldsymbol{I}), \tag{10}$$

where $\mathcal{N}(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma})$ represents normal distribution on $\boldsymbol{y}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Before the sources are estimated, a *prior* distribution is assigned to the unknown sources. Since $K \gg N$, $\bar{\boldsymbol{s}}$ is a vector with high degree of sparsity.

A. Bayesian learning with Laplace prior

In RVM, a Gaussian prior is used for the prior distribution over the source. In this letter, the Laplace prior is used because it enforces the sparsity constraint more heavily compared with Gaussian prior [9]. The Laplace prior is known to be non-conjugate with the Gaussian distribution [13], the likelihood in (10) is shown to be Gaussian which prevents the marginal likelihood to be expressed in a closed form with a Laplace prior. In order to overcome the drawback, we adopt the hierarchical model of [14]. In the first stage,

$$P(\bar{\boldsymbol{s}}|\boldsymbol{\alpha}) = \mathcal{N}(\bar{\boldsymbol{s}}|0,\boldsymbol{\alpha}), \tag{11}$$

where $\boldsymbol{\alpha} = \text{diag}(\alpha_1, \alpha_2 ... \ \alpha_{2K})$, α_i denotes the variance for $\bar{s}_i$. In the second stage, we have

$$P(\alpha_i|\lambda) = \frac{\lambda}{2}\exp\left(-\frac{\lambda\alpha_i}{2}\right), i = 1, 2, ...2K. \tag{12}$$

In the third stage, Jeffrey's hyperprior is applied,

$$P(\lambda) \propto \frac{1}{\lambda}. \tag{13}$$

The prior distribution of the sources is given by

$$P(\bar{\boldsymbol{s}}|\lambda) = \prod_{i=1}^{2k}\int_0^\infty P(\bar{s}_i|\alpha_i)P(\alpha_i|\lambda)d\alpha_i. \tag{14}$$

Using the identity

$$\int_0^\infty e^{-uy^2 - v/y^2}dy = \frac{\sqrt{\pi}}{2\sqrt{u}}e^{-2\sqrt{uv}}, \tag{15}$$

we have the following distribution for the sources:

$$P(\bar{\boldsymbol{s}}|\lambda) = \frac{\lambda^K}{2^{2K}}\exp(-\sqrt{\lambda}\sum_{i=1}^{2K}|\bar{s}_i|), \tag{16}$$

where Laplace prior is applied to each signal coefficient. The posterior distribution over $\bar{\boldsymbol{s}}$ is

$$P(\bar{\boldsymbol{s}}|\bar{\boldsymbol{y}},\boldsymbol{\alpha},\sigma^2) \propto P(\bar{\boldsymbol{y}}|\bar{\boldsymbol{s}},\sigma^2)P(\bar{\boldsymbol{s}}|\boldsymbol{\alpha}) = \mathcal{N}(\bar{\boldsymbol{s}}|\boldsymbol{\mu},\boldsymbol{\Sigma}), \tag{17}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given by [8]

$$\boldsymbol{\mu} = \sigma^{-2}\boldsymbol{\Sigma}\bar{\boldsymbol{A}}^T\bar{\boldsymbol{y}}, \tag{18}$$

$$\boldsymbol{\Sigma} = (\boldsymbol{\alpha}^{-1} + \bar{\boldsymbol{A}}^T\bar{\boldsymbol{A}}/\sigma^2)^{-1}. \tag{19}$$

In order to compute (18) and (19), we need to estimate the hyperparamters $\boldsymbol{\alpha}$ defined in (11), λ in (13) and the noise variance σ^2. Their value can be estimated by maximizing the marginal likelihood which is given by

$$P(\bar{\boldsymbol{y}}|\boldsymbol{\alpha},\lambda,\sigma^2) = \int P(\bar{\boldsymbol{y}}|\bar{\boldsymbol{s}},\sigma^2)P(\bar{\boldsymbol{s}}|\boldsymbol{\alpha})P(\boldsymbol{\alpha}|\lambda)P(\lambda)d\bar{\boldsymbol{s}}. \tag{20}$$

The logarithm of marginal likelihood is [8]

$$\begin{aligned}\log P(\bar{\boldsymbol{y}}|\boldsymbol{\alpha},\lambda,\sigma^2) = &-\frac{1}{2}\left[2M\log 2\pi + \log|\boldsymbol{C}| + \bar{\boldsymbol{y}}^T\boldsymbol{C}^{-1}\bar{\boldsymbol{y}}\right]\\ &-\frac{\lambda}{2}\sum_{i=1}^{2M}\alpha_i + (2M-1)\log\lambda,\end{aligned} \tag{21}$$

where

$$\boldsymbol{C} = \sigma^2\boldsymbol{I} + \bar{\boldsymbol{A}}\boldsymbol{\alpha}\bar{\boldsymbol{A}}^{\boldsymbol{T}}. \tag{22}$$

Setting the derivatives of (21) with respect to hyperparameters to zero, we can obtain the optimal hyperparameters [9]

$$\alpha_i^{new} = \frac{1}{-\frac{1}{2\lambda} + \sqrt{\frac{1}{4\lambda^2} + \frac{\mu_i^2 + \Sigma_{i,i}}{\lambda}}}, i = 1, 2, ...2K, \tag{23}$$

$$\lambda^{new} = \frac{2M-1}{\sum_{i=1}^{2K}\alpha_i/2}, \tag{24}$$

$$(\sigma^2)^{new} = \frac{\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{\mu}\|^2}{2M - \sum_{i=1}^{2K}(1-\alpha_i\boldsymbol{\Sigma}_{(i,i)})}. \tag{25}$$

In summary, we obtain initial estimates of the parameters $\boldsymbol{\alpha}$ and σ^2, and our iterations update these estimates using (23)-(25); then we use (18) and (19) to update the estimated mean and variance of the sources. The estimated results $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ will be returned when the marginal likelihood converges.

B. Pairwise fast marginal likelihood maximization

The computational complexity of the method in Sec. III-A is $\mathcal{O}(K^3)$ due to of computation of the inverse of $\boldsymbol{\Sigma}$. In order to reduce the computational complexity, the source can be estimated in the "constructive" way [12]: it starts with selecting source from one direction, then sources from more diretions are added iteratively, the size of $\boldsymbol{\Sigma}$ increases from 1 and stopping at a number which is much smaller than K [12]. In addition, the relationship between the real and imaginary parts of source are taken into account. We propose to estimate hyperparameters in pairs based on the belief that the non-zero entries in the real part suggests that the corresponding imaginary entries are non-zero with high probability and vice versa.

1) Single snapshot: The term $\boldsymbol{C}$ in (22) is rewritten as

$$\boldsymbol{C} = \sigma^2 \boldsymbol{I} + \sum_{j\neq i} \alpha_j^{-1} \bar{\boldsymbol{A}}_j (\bar{\boldsymbol{A}}_j)^T + \alpha_i^{-1} \bar{\boldsymbol{A}}_i (\bar{\boldsymbol{A}}_i)^T, \tag{26}$$

where $\bar{\boldsymbol{A}}_i$ denotes the ith column of $\bar{\boldsymbol{A}}$. We define $\boldsymbol{C}_{-i} \equiv \sigma^2\boldsymbol{I} + \sum_{j\neq i} \alpha_j^{-1} \bar{\boldsymbol{A}}_j (\bar{\boldsymbol{A}}_j)^T$ which separates terms independent of α_i from terms dependent on α_i. With (26), the marginal likelihood (21) can also be rewritten to include $\mathcal{L}(\alpha_{-i})$, which is independent on α_i, and $\mathcal{L}(\alpha_i)$, which is dependent on α_i [8]. $\mathcal{L}(\alpha_i)$ is given by

$$\mathcal{L}(\alpha_i) = \frac{1}{2}\left[\log \frac{1}{1+\alpha_i p_i} + \frac{(q_i)^2 \alpha_i}{1+\alpha_i p_i} - \lambda \alpha_i\right], \tag{27}$$

where $p_i = (\bar{\boldsymbol{A}}_i)^T (\boldsymbol{C}_{-i})^{-1} \bar{\boldsymbol{A}}_i$ and $q_i = (\bar{\boldsymbol{A}}_i)^T (\boldsymbol{C}_{-i})^{-1} (\bar{\boldsymbol{y}})$. Instead of updating all of $\boldsymbol{\alpha}$, the maximum $\mathcal{L}_c(\alpha_i + \alpha_{i+K}) \equiv \mathcal{L}(\alpha_i) + \mathcal{L}(\alpha_{i+K}), i = 1, ...K$, which includes real and imaginary parts *jointly*, is selected at each iteration, and the associated α_i and α_{i+K} are recomputed iteratively. The optimal α_i is given by [12]

$$\alpha_i = \begin{cases} \frac{-Y+\sqrt{Y^2-4XZ}}{2X}, & \text{if } (q_i)^2 - p_i > \lambda \\ 0, & \text{if } (q_i)^2 - p_i \leq \lambda, \end{cases} \tag{28}$$

where $X = \lambda p_i^2$, $Y = p_i^2 + 2\lambda p_i$, $Z = \lambda + p_i - q_i^2$.

The set $\mathcal{S}$ is used to represent the estimated directions for incoming sources. Initially, $\mathcal{S}$ is empty, then $\mathcal{S}$ is iteratively updated when $\boldsymbol{\alpha}$ is updated iteratively. If either α_i or α_{i+K} is non-zero, we add the angle associated with α_i and α_{i+K} to the set $\mathcal{S}$. If both of them are 0, we remove the angle from $\mathcal{S}$.

2) Multiple snapshots: Multiple snapshots helps reduce the correlation between the columns of the measurement matrix, and thus the difficulty in recovering the source is decreased. The effect of multiple snapshots on the Gram matrix (i.e., $\sum_{l=1}^{L} (\bar{\boldsymbol{A}}^{\boldsymbol{l}})^{\boldsymbol{T}} \bar{\boldsymbol{A}}^{\boldsymbol{l}} / L$) is presented in Fig. 2 and Fig. 3, in which a 12-element linear array is considered, the location of element l_m is subject to uniform distribution

$$l_m \sim U[(m-1)*\Delta s, (m-1)*\Delta s + r]. \tag{29}$$

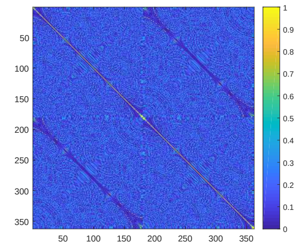


Fig. 2: Gram matrix with one snapshot.

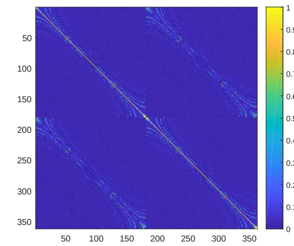


Fig. 3: Gram matrix with 100 snapshots.

When multiple independent measurements are available and the incoming sources are assumed to be unchanged, the marginal likelihood with respect of α_i is written as

$$\mathcal{L}(\alpha_i) = \frac{1}{2}\sum_{l=1}^{L}\left[\log \frac{1}{1+\alpha_i p_i^l} + \frac{(q_i^l)^2 \alpha_i}{1+\alpha_i p_i^l} - \lambda \alpha_i\right], \tag{30}$$

where L is the number of snapshots, the superscript l stands for the lth snapshot, $p_i^l = (\bar{\boldsymbol{A}}_i^l)^T (\boldsymbol{C}_{-i}^l)^{-1} \bar{\boldsymbol{A}}_i^l$, $q_i^l = (\bar{\boldsymbol{A}}_i^l)^T (\boldsymbol{C}_{-i}^l)^{-1} (\bar{\boldsymbol{y}})^l$ and $\bar{\boldsymbol{A}}^l$ denotes the measurement matrix at the lth snapshot. Setting the derivative of (30) with respect to the hyperparameter α_i equal to 0, we have

$$\sum_{l=1}^{L} \frac{-p_i^l + (q_i^l)^2 - \lambda - \alpha_i\left[(p_i^l)^2 + 2\lambda p_i^l\right] - \alpha_i^2 \lambda (p_i^l)^2}{(1+\alpha_i p_i^l)^2} = 0. \tag{31}$$

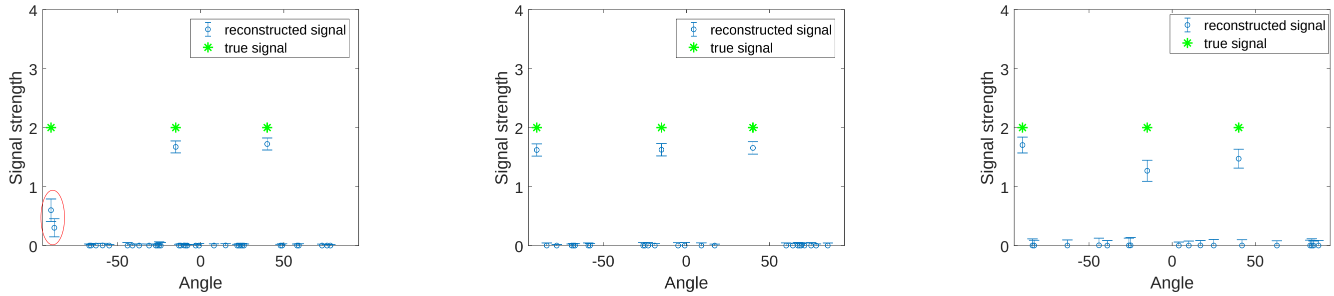
With the approximation $p_i^l \gg \alpha_i^{-1}$ [15] the optimal hyperparameters are given by

$$\alpha_i = \begin{cases} \frac{-Y_m+\sqrt{Y_m^2-4X_m Z_m}}{2X_m}, & \text{if } Q_m - P_m > \lambda R_m \\ 0, & \text{if } Q_m - P_m \leq \lambda R_m, \end{cases} \tag{32}$$

where $X_m = L\lambda$, $Y_m = L + 2\lambda P_m$, $Z_m = \lambda R_m + P_m - Q_m$, $P_m = \sum_{l=1}^{L} \frac{1}{p_i^l}$, $R_m = \sum_{l=1}^{L} \frac{1}{(p_i^l)^2}$ and $Q_m = \sum_{l=1}^{L} (\frac{q_i^l}{p_i^l})^2$. It can be seen when $L = 1$, (32) is equivalent to (28).

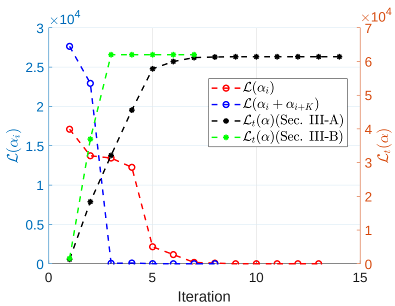
IV. Numerical Results

In this section, the proposed method from III-B is tested with two configurations of linear arrays and $K = 181$. The first configuration is described in (29), and we refer to it as a *normal* non-uniform array. For the second configuration, twelve elements are grouped in four clusters and each cluster has three elements whose positions change with time. The distance between neighboring cluster is 100 wavelengths. We call this the *cluster* configuration. The latter is encountered in real life such as distributed ground communication with mobile user clusters. The signal-to-noise ratio (SNR) is defined as $\text{SNR} = 10 \log_{10} \frac{\|\boldsymbol{y_p}\|_\infty^2}{2\sigma^2}$, where $\boldsymbol{y_p} = \boldsymbol{As} \in \mathbb{C}^{M\times 1}$ stands for the signal vector received without noise. Fig. 4a presents the estimated sources with standard ℓ_1 norm SBL method of Sec. III-A while the results in Fig. 4b are estimated with the method of Sec. III-B. We see that if the relationship between the real and imaginary parts is ignored, one spurious source is included in the estimate which is highlighted in Fig. 4a with a red circle. The estimated results with the cluster array are

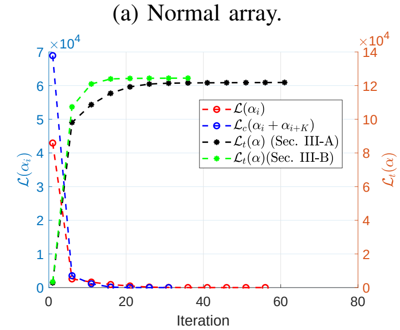


(a) (b) (c)

Fig. 4: Estimated signal direction and strength, SNR = 10 dB, $N = 3, L = 100, M = 12$. (a) Normal array with general ℓ_1 norm algorithm [9], (b) normal array with pairwise algorithm (III-B) and (c) cluster array with pairwise algorithm (III-B).



(a) Normal array.



(b) Cluster array.

Fig. 5: The change of marginal likelihood vs. iteration. Left Y-axis: the likelihood at one iteration (30), right Y-axis: the sum of the likelihood until the current iteration, $\mathcal{L}(\boldsymbol{\alpha})_t = \sum_{i=1}^{t} \mathcal{L}(\alpha^i)$, α^i is the hyperparameter estimate at ith iteration.

plotted in Fig. 4c. The marginal likelihood versus iterations for the example in Fig. 4a is shown in Fig. 5a. The number of required iterations for the pairwise algorithm is around half of the required number of iterations for the standard ℓ_1 norm SBL method under the same convergence criterion. The cluster configuration is also tested with SNR = 10 dB; the results are plotted in Fig. 5b.

The comparison between the proposed method (III-B) with standard RVM [8], RVM with fast algorithm [12] and standard ℓ_1 norm SBL [9] is listed in Table I in terms of accuracy, execution time and number of nonzero (NNZ) elements. The number of actual sources is $N = 3$ in the experiment. We check the strongest N signals of the estimated result which are denoted by $\widehat{\boldsymbol{S}} = \{\widehat{S}(\hat{\theta}_1), \widehat{S}(\hat{\theta}_2), ..., \widehat{S}(\hat{\theta}_N)\}$, $\widehat{S}(\hat{\theta}_i)$ denotes the estimated source from direction $\hat{\theta}_i$. If the estimated source $\widehat{S}(\hat{\theta}_i)$ meets all of the following criteria, it is counted as being predicted correctly

- the estimated angle matches the direction of the actual source, $\hat{\theta}_i \in \boldsymbol{\theta}$
- the estimated signal strength is at least half of the strength of the actual signal, $|\widehat{S}(\hat{\theta}_i)|^2 \geq 0.5 \times |\tilde{S}(\hat{\theta}_i)|^2$, $\tilde{S}(\hat{\theta}_i)$ denotes the actual source at $\hat{\theta}_i$
- the estimated signal strength is at least twice the strongest strength of the estimated sources from wrong directions, $|\tilde{S}(\hat{\theta}_i)|^2 \geq 2 \cdot \max\{\tilde{S}(\hat{\theta}_w) : \hat{\theta}_w = \mathcal{S} - \boldsymbol{\theta}\}$.

From Table I, it can be been seen our proposed method achieves higher prediction accuracy and shorter execution time compared with the existing methods, and a higher degree of sparsity is obtained with our approach.

Accuracy (%)/ time (s) /NNZ	L = 1, SNR = 10dB	L = 10, SNR = 10dB	L = 100, SNR = 10dB
RVM [8]	18.7/4.47/362	27.7/4.91/362	31.5/5.1/362
fast RVM [12]	7.35/0.05/10.6	16.7/0.57/73.5	36.3/2.55/165
ℓ_1 norm [9]	8.7/0.07/19	17.4/0.17/43.2	35.1/0.87/**47.3**
pair ℓ_1 norm	**18.7/0.05/18.3**	**44.1/0.15/42.8**	**70.8/0.45**/52.3

TABLE I: Prediction accuracy and execution time with SNR = 10 dB, NNZ: number of nonzero elements.

V. Conclusion

In this paper, we propose an approach to estimate DoA of complex signals and the approach that has the capability to handle multiple snapshots. The proposed method is compared with existing DoA methods to show its better performance such as higher prediction accuracy and shorter execution time.

VI. Acknowledgement

We thank Prof. Robert W. Jackson for helpful discussions.

References

[1] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[2] F. Yan, M. Jin, and X. Qiao, "Low-complexity DOA estimation based on compressed MUSIC and its performance analysis," *IEEE Trans. Signal Process.*, vol. 61, no. 8, pp. 1915–1930, Apr. 2013.

[3] X. Zhuang, X. Cui, M. Lu, and Z. Feng, "Low-complexity method for DOA estimation based on ESPRIT," *Journal of Systems Engineering and Electronics*, vol. 21, no. 5, pp. 729–733, Oct. 2010.

[4] T. B. Lavate, V. K. Kokate, and A. M. Sapkal, "Performance analysis of MUSIC and ESPRIT DOA estimation algorithms for adaptive array smart antenna in mobile communication," in *Second International Conference on Computer and Network Technology*, Apr. 2010, pp. 308–311.

[5] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3038–3044, Jul. 2008.

[6] M. Carlin, P. Rocca, G. Oliveri, F. Viani, and A. Massa, "Directions-of-arrival estimation through Bayesian compressive sensing strategies," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3828–3838, Jul. 2013.

[7] J. Dai and H. C. So, "Sparse bayesian learning approach for outlier-resistant direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 744–756, Feb. 2018.

[8] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, Jun. 2001.

[9] S. D. Babacan *et al.*, "Fast Bayesian compressive sensing using Laplace priors," in *ICASSP*, Taipei, 2009, pp. 2873 – 2876.

[10] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[11] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, pp. 797 – 829, Jun. 2006.

[12] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *International Workshop on Artificial Intelligence and Statistics*, Key West, FL, 2003.

[13] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory*. Wiley, 2000.

[14] M. A. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 1150–1159, Sep. 2003.

[15] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, pp. 92–106, Jan. 2009.