JOINT DIRECTION-OF-ARRIVAL AND FREQUENCY ESTIMATION WITHOUT SOURCE ENUMERATION

Cheng Qian[⋆] *Lei Huang*[⋆] *Yunmei Shi*[⋆] *H. C. So*[†]

[⋆]Department of Electronic and Information Engineering
Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China
[†]Department of Electronic Engineering
City University of Hong Kong, Hong Kong, China

ABSTRACT

Joint estimation of the directions-of-arrival (DOAs) and frequencies of multiple signals is addressed in this paper. By constructing a set of joint diagonalization matrices, two cost functions that do not require *a priori* information of the source number are devised for DOA and frequency estimation in a separate manner. This enables us to estimate DOAs and frequencies via two *one-dimensional* search steps in their corresponding spatial and frequency domains. Thus, the tremendous *two-dimensional* search required in the standard approaches can be avoided. Simulation results demonstrate the effectiveness of the proposed approach.

Index Terms— Direction-of-arrival (DOA) estimation, frequency estimation, joint diagonalization.

1. INTRODUCTION

Joint direction-of-arrival (DOA) and frequency estimation of multiple signals is an important problem in spatial-temporal radio channel measurement [1]-[6]. A precise estimation of DOAs and frequencies for the signals of interest can help to provide better channel information in support of improved link quality.

To tackle this issue, various estimators have been developed in the literature. ESPRIT-like algorithms [4]-[6] have been proposed to balance the estimation accuracy and computational complexity. In [5], a joint angle and frequency estimation (JAFE) algorithm has been suggested, in which the observation data are preprocessed by the temporal-spatial smoothing technique, and then the conventional ESPRIT scheme is employed to estimate the DOAs and frequencies. With the numbers of antennas and samples increase, there exists overlapping elements between the two sides of the rotational invariance equations. This is the reason why the conventional ESPRIT which is based on least squares technique suffers performance degradation [6]-[7]. To overcome this problem, a structured least squares based ESPRIT (SLS-ESPRIT) algorithm [6] has been devised. The SLS-ESPRIT approach has a better performance accuracy compared to the conventional ESPRIT algorithms. However, it relies on the *a priori* knowledge of number of signals. As a matter of fact, the number of signals is usually unknown to the receiver in practice, and its estimate [8]-[10] is used instead. Moreover, the probability of successfully detecting the source number is still low when the signal-to-noise ratio (SNR) and sample size are smaller than a certain threshold [11].

In this paper, we propose a new approach for joint DOA and frequency estimation which does not need the source number information. Our proposal employs a set of data matrices which have the same joint diagonalization structure to construct two cost functions, that is, one of them is for DOA estimation and the other one is for frequency estimation. As a result, the DOA and frequency parameters are determined by using two *one-dimensional* peak search procedures separately.

The work described in this paper was supported by the National Natural Science Foundation of China under Grants 61222106 and 61171187.

2. PROBLEM FORMULATION

Consider a uniform linear array (ULA) of M omnidirectional sensors. There are P $(P < M)$ uncorrelated narrowband sinusoids impinging on the array from distinct directions $\{\theta_1 \ \cdots \ \theta_P\}$ in the far field. Then the $M \times 1$ observation vector at the receiver is

$$\begin{aligned}\mathbf{x}(t) &= \sum_{p=1}^{P} \alpha_p \mathbf{a}_p e^{j\omega_p t} + \mathbf{n}(t) \\ &= \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \qquad (1)\end{aligned}$$

where α_p and ω_p are the complex amplitude and frequency of the pth signals, $\mathbf{s}(t) = [\alpha_1 e^{j\omega_1 t} \ \cdots \ \alpha_P e^{j\omega_P t}]^T$ is the source signal vector with $(\cdot)^T$ being the transpose, $\mathbf{n}(t)$ is the additive white Gaussian noise vector with mean zero and covariance $\sigma_n^2 \mathbf{I}_M$. Here, $\mathbf{I}_M$ is the $M \times M$ identity matrix. The $\mathbf{A} = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_P]$ is the array manifold with

$$\mathbf{a}_p = \left[1 \ e^{j2\pi \sin(\theta_p) d/\lambda} \ \ldots \ e^{j2\pi(M-1)\sin(\theta_p)d/\lambda}\right]^T \qquad (2)$$

being the pth steering vector. Here, λ is the carrier wavelength and $d = \lambda/2$ is the interelement spacing.

3. PROPOSED ALGORITHM

3.1. Data Processing

It is easy to verify that

$$\begin{aligned}\mathbf{x}(t+k) &= \mathbf{A}\mathbf{s}(t+k)+\mathbf{n}(t+k)\\ &= \mathbf{A}\mathbf{\Phi}^k\mathbf{s}(t)+\mathbf{n}(t+k)\end{aligned} \tag{3}$$

where

$$\mathbf{\Phi} = \text{diag}\{e^{j\omega_1} \ \cdots \ e^{j\omega_P}\}. \tag{4}$$

To proceed, we form a $Mm \times 1$ data vector

$$\begin{aligned}\mathbf{y}_i &= \left[\mathbf{x}^T(i) \ \cdots \ \mathbf{x}^T(i+m-1)\right]^T\\ &= \mathbf{B}\mathbf{s}(i)+\mathbf{n}_i, \ \ i=0 \ \cdots \ N-m\end{aligned} \tag{5}$$

where $\mathbf{B} = \mathbf{C} \odot \mathbf{A}$ with $\odot$ being the Khatri-Rao product, $\mathbf{n}_i = [\mathbf{n}^T(i) \ \cdots \ \mathbf{n}^T(i+m-1)]^T$, N is the number of samples and

$$\mathbf{C} = \begin{bmatrix} 1 & \cdots & 1\\ e^{j\omega_1} & \cdots & e^{j\omega_P}\\ \vdots & \ddots & \vdots\\ e^{j(m-1)\omega_1} & \cdots & e^{j(m-1)\omega_P}\end{bmatrix}. \tag{6}$$

We express $\mathbf{y}_i$ as a $M \times m$ data matrix, denoted by $\mathbf{Y}_i$, which has the form of

$$\mathbf{Y}_i = \begin{bmatrix} y_1 & y_{M+1} & \cdots & y_{M(m-1)+1}\\ y_2 & y_{M+2} & \cdots & y_{M(m-1)+2}\\ \vdots & \vdots & \ddots & \vdots\\ y_M & y_{2M} & \cdots & y_{Mm}\end{bmatrix} \tag{7}$$

The (i,j) entry of $\mathbf{Y}_i$ is represented as $[\mathbf{Y}]_{i,j} = y_{i+(j-1)M}$. It is obvious that the vectorization of $\mathbf{Y}_i$ is equivalent to $\mathbf{y}_i$ such that we have[1]

$$\mathbf{Y}_i = \mathbf{A}\mathbf{\Lambda}_i\mathbf{C}^T + \mathbf{N}_i \tag{8}$$

where $\mathbf{\Lambda}_i = \text{diag}\{\alpha_1 e^{j\omega_1 i} \ \cdots \ \alpha_P e^{j\omega_P i}\}$ and $\mathbf{N}_i$ is obtained similarly to $\mathbf{Y}_i$, i.e., by reshaping $\mathbf{n}_i$ into a $M \times m$ matrix.

3.2. Joint DOA and Frequency Estimation without Source Number Information

3.2.1. DOA Estimation

In the absence of noise, $\mathbf{Y}_i$ can be written as

$$\mathbf{Y}_i = \mathbf{A}\mathbf{\Lambda}_i\mathbf{C}^T = \sum_{p=1}^{P} s_{i,p}\mathbf{a}_p\mathbf{c}_p^T \tag{9}$$

where $s_{i,p}$ is the pth element of $\mathbf{s}(i)$ and $\mathbf{c}_p$ is the pth column of $\mathbf{C}$. It follows from (9) that for the jth source, there always exists a vector $\mathbf{b}_j \in \mathbb{C}^m$ that is orthogonal to the range space spanned by the remaining $(P-1)$ steering vectors except $\mathbf{c}_j$, i.e.,

$$\mathbf{c}_p^T\mathbf{b}_j = \begin{cases}\mathbf{c}_j^T\mathbf{b}_j, & p=j\\ 0, & p\neq j.\end{cases} \tag{10}$$

Substituting (10) into (9) yields

$$\mathbf{Y}_i\mathbf{b}_j = \sum_{p=1}^{P} s_{i,p}\mathbf{a}_p\mathbf{c}_p^T\mathbf{b}_j = g_i\mathbf{a}_j \tag{11}$$

where $g_i = s_{i,j}\mathbf{c}_j^T\mathbf{b}_j$.

From (11), we confirm that if θ_j is one of the true DOAs, there always exists a scalar g_i that makes $\mathbf{Y}_i\mathbf{b}_j$ and $\mathbf{a}_j$ parallel. Since (11) holds true for $0 \leq i \leq N-m$, we try to minimize the total distance between the $(N-m+1)$ equations in (11). To this end, similar to [12]-[15], we construct the following optimization problem

$$\begin{aligned}&\min_{\theta} \ \mathbf{J}(\theta,\mathbf{g},\mathbf{b}) = \sum_{i=0}^{N-m} ||\mathbf{Y}_i\mathbf{b} - g_i\mathbf{a}||^2\\ &\text{s. t.} \ \ ||\mathbf{g}|| = 1\end{aligned} \tag{12}$$

where $||\cdot||$ is the Euclidean norm, $\mathbf{a}$ is the steering vector with parameter θ to be optimized, $\mathbf{g} = [g_0 \ \cdots \ g_{N-m}]^T \in \mathbb{C}^{N-m+1}$ and the constraint $||\mathbf{g}|| = 1$ is used to avoid the trivial solution of (12), i.e., $\mathbf{g} = \mathbf{0}_{N-m+1}$ and $\mathbf{b} = \mathbf{0}_m$ [12]-[13] with $\mathbf{0}_m$ being the zero vector.

Since $\mathbf{b}$ and $\mathbf{g}$ are unknown parameters, it is difficult to optimize (12) by searching for the DOAs directly. To circumvent this issue, we attempt to simplify (12), so that it is not affected by $\mathbf{b}$ and $\mathbf{g}$. Let

$$\mathbf{F}_\theta = \sum_{i=0}^{N-m} \mathbf{Y}_i^H\mathbf{Y}_i \in \mathbb{C}^{m\times m} \tag{13}$$

$$\mathbf{G}_\theta = \left[\mathbf{Y}_0^H\mathbf{a} \ \cdots \ \mathbf{Y}_{N-m}^H\mathbf{a}\right] \in \mathbb{C}^{m\times(N-m+1)}. \tag{14}$$

The cost function in (12) can be rewritten as

$$\mathbf{J}(\theta,\mathbf{g},\mathbf{b}) = \mathbf{b}^H\mathbf{F}_\theta\mathbf{b} - \mathbf{b}^H\mathbf{G}_\theta\mathbf{g} - \mathbf{g}^H\mathbf{G}_\theta^H\mathbf{b} + M. \tag{15}$$

By using the method of Lagrange multiplier, we have [15]

$$\begin{aligned}\mathbf{J}(\theta,\mathbf{g},\mathbf{b}) = \ &\mathbf{b}^H\mathbf{F}_\theta\mathbf{b} - \mathbf{b}^H\mathbf{G}_\theta\mathbf{g} - \mathbf{g}^H\mathbf{G}_\theta^H\mathbf{b} + M\\ &+ \lambda(||\mathbf{g}|| - 1).\end{aligned} \tag{16}$$

For fixed θ and $\mathbf{g}$, we differentiate (16) with respect to $\mathbf{b}$ and then set the so-obtained expression to zero, yielding

$$\mathbf{b}_{\text{opt}} = \mathbf{F}_\theta^\dagger\mathbf{G}_\theta\mathbf{g} \tag{17}$$

where $(\cdot)^\dagger$ is the pseudo-inverse. Substituting (17) back into (12), the optimization problem is reduced to

$$\min_{\theta} \ \mathbf{J}(\theta,\mathbf{g}) = M - \mathbf{g}^H\mathbf{G}_\theta^H\mathbf{F}_\theta^\dagger\mathbf{G}_\theta\mathbf{g} \tag{18}$$

[1] If $\mathbf{B}$ is a diagonal matrix, we have $\text{vec}\{\mathbf{ABC}\} = (\mathbf{C}^T \odot \mathbf{A})\mathbf{b}$ where $\mathbf{B}$ is diagonal and $\mathbf{b}$ contains the diagonal elements of $\mathbf{B}$.

Minimizing (18) equals to maximizing $\mathbf{g}^H\mathbf{G}_\theta^H\mathbf{F}_\theta^\dagger\mathbf{G}_\theta\mathbf{g}$. It is easy to verify that the maximum of $\mathbf{g}^H\mathbf{G}_\theta^H\mathbf{F}_\theta^\dagger\mathbf{G}_\theta\mathbf{g}$ is achieved if and only if $\mathbf{g}$ is the eigenvector corresponding to the maximum eigenvalue of $\mathbf{G}_\theta^H\mathbf{F}_\theta^\dagger\mathbf{G}_\theta$ [15]. Therefore, (18) can be further simplified as

$$\min_{\theta} \quad \mathbf{J}(\theta) = M - \max \text{eig}\left\{\mathbf{G}_\theta^H\mathbf{F}_\theta^\dagger\mathbf{G}_\theta\right\} \tag{19}$$

where $\max \text{eig}(\cdot)$ denotes the maximum eigenvalue of a matrix. Thus, we can obtain the pseudo output power spectrum

$$P(\theta) = \frac{1}{M - \max \text{eig}\left\{\mathbf{G}_\theta^H\mathbf{F}_\theta^\dagger\mathbf{G}_\theta\right\}}. \tag{20}$$

Given the search range, the DOAs are selected as the angles corresponding to the highest local maxima of $P(\theta)$.

3.2.2. Frequency Estimation

Since the frequency information is contained in $\mathbf{C}$, it follows from (9) that $\mathbf{Y}_i^T$ spans the same range space of $\mathbf{C}$, i.e.,

$$\text{span}\{\mathbf{Y}_i^T\} = \text{span}\{\mathbf{C}\}. \tag{21}$$

Hence, taking transpose of $\mathbf{Y}_i$ yields

$$\mathbf{Y}_i^T = \mathbf{C}\mathbf{\Lambda}_i\mathbf{A}^T = \sum_{p=1}^{P} s_{i,p}\mathbf{c}_p\mathbf{a}_p^T \tag{22}$$

There is a vector $\mathbf{e}_j \in \mathbb{C}^M$ such that

$$\mathbf{a}_p^T\mathbf{e}_j = \begin{cases} \mathbf{a}_j^T\mathbf{e}_j, & p = j \\ 0, & p \neq j. \end{cases} \tag{23}$$

Substituting (23) into (22) yields

$$\mathbf{Y}_i^T\mathbf{e}_j = \sum_{p=1}^{P} s_{i,p}\mathbf{c}_p\mathbf{a}_p^T\mathbf{e}_j = h_i\mathbf{c}_j \tag{24}$$

where $h_i = s_{i,j}\mathbf{a}_j^T\mathbf{e}_j$.

Utilizing the similar manipulations as those in DOA estimation, we construct the following optimization problem

$$\begin{aligned} &\min_{\omega} \ \mathbf{J}(\omega, \mathbf{h}, \mathbf{e}) = \sum_{i=0}^{N-m} ||\mathbf{Y}_i^T\mathbf{e} - h_i\mathbf{c}||^2 \\ &\text{s. t.} \ ||\mathbf{h}|| = 1. \end{aligned} \tag{25}$$

Define

$$\mathbf{F}_\omega = \sum_{i=0}^{N-m} \mathbf{Y}_i^*\mathbf{Y}_i^T \in \mathbb{C}^{M\times M} \tag{26}$$

$$\mathbf{G}_\omega = \left[\mathbf{Y}_0^*\mathbf{c} \ \cdots \ \mathbf{Y}_{N-m}^*\mathbf{c}\right] \in \mathbb{C}^{M\times(N-m+1)} \tag{27}$$

where $(\cdot)^*$ is the conjugate. The objective function in (25) is further simplified as

$$\min_{\omega} \ \mathbf{J}(\omega) = m - \max \text{eig}\left\{\mathbf{G}_\omega^H\mathbf{F}_\omega^\dagger\mathbf{G}_\omega\right\}. \tag{28}$$

Our goal is to estimate the frequencies by searching over a given frequency range. Hence, the pseudo output power spectrum for frequency estimation is given as

$$P(\omega) = \frac{1}{m - \max \text{eig}\left\{\mathbf{G}_\omega^H\mathbf{F}_\omega^\dagger\mathbf{G}_\omega\right\}}. \tag{29}$$

The frequencies corresponding to the highest local maxima of $P(\omega)$ are selected as the estimated frequencies.

3.3. Pairing Procedure

When there exists more than one targets, a pairing process is needed to group $\hat{\theta}_i$ and $\hat{\omega}_i$. To begin, we set $\mathbf{Y} = [\mathbf{y}_0 \ \cdots \ \mathbf{y}_{N-m}]$. According to (5), in the absence of noise, $\mathbf{Y}$ can be expressed as $\mathbf{Y} = \mathbf{B}\cdot[\mathbf{s}(0) \ \cdots \ \mathbf{s}(N-m)]$. Then, let $\mathbf{U}_s = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_P]$ be the signal subspace where $\mathbf{u}_p$ is the pth eigenvector that is corresponding to the pth largest eigenvalue associated with $\mathbf{Y}$. Define

$$\boldsymbol{\gamma}(\omega_m, \theta_n) = \mathbf{c}_m \otimes \mathbf{a}_n, \ m, n = 1 \ \cdots \ P \tag{30}$$

where $\otimes$ is the Kronecker product. For the mth frequency, we vary $\hat{\theta}$ and calculate $\mathbf{t}_m = [t_1 \ \cdots \ t_P]$ where $t_i = \sum_{p=1}^{P} |\mathbf{u}_p^H\boldsymbol{\gamma}(\omega_m, \theta_i)|^2$. With the results in [16], the DOA corresponding to the mth frequency is determined by choosing the DOA associated with the maximum t_i in $\mathbf{t}_m$.

The proposed method for joint DOA and frequency estimation is summarized in Table I.

4. SIMULATION RESULTS

The performance of the proposed algorithm is examined in this section. We consider a ULA of $M = 8$ sensors successively separated by a half-wavelength. The noise is white Gaussian process with zero mean and unit variance.

In the first example, we assume that seven signals with equal powers arrive at the ULA from angles $\theta_1 = -50°$, $\theta_2 = -33.3°$, $\theta_3 = -16.7°$, $\theta_4 = 0°$, $\theta_5 = 16.7°$, $\theta_6 = 33.3°$ and $\theta_7 = 50°$. Meanwhile, their frequencies are $\omega_1 = 0.2\pi$, $\omega_2 = 0.3\pi$, $\omega_3 = 0.4\pi$, $\omega_4 = 0.5\pi$, $\omega_5 = 0.6\pi$, $\omega_6 = 0.7\pi$ and $\omega_7 = 0.8\pi$. We set $N = 50$ and $m = 20$. The α is randomly generated from Gaussian distribution. The SNR is set to be 10 dB. Fig. 1(a) displays ten spatial spectra for DOA and frequency estimation. It is observed that the proposed method successfully estimates the seven DOAs and frequencies. We now consider a case of two closely spaced signals with frequency $\omega_1 = 0.4\pi$ and $\omega_2 = 0.6\pi$ located at $\theta_1 = -2°$ and $\theta_2 = 2°$. The other parameters are the same as those in Fig. 1(a). It is shown in Fig. 1(b) that there are two distinct peaks in both DOA and frequency estimation pseudo spectra.

In the second example, we consider a case when $M = 8$, $N = 40$, $m = 16$ and SNR $= 10$ dB. There are $P = 7$ signals

Table 1. Pseudo-code of proposed algorithm

(a) DOA Estimation

Step 1: Use (7) to construct $\{\mathbf{Y}_i\}_{i=0}^{N-m}$.

Step 2: Employ (13) and (14) to construct $\mathbf{F}_\theta$ and $\mathbf{G}_\theta$, respectively.

Step 3: Utilize (20) to form the pseudo spectrum $P(\theta)$.

Step 4: Estimate the DOAs by searching for the peaks of $P(\theta)$.

(b) Frequency Estimation

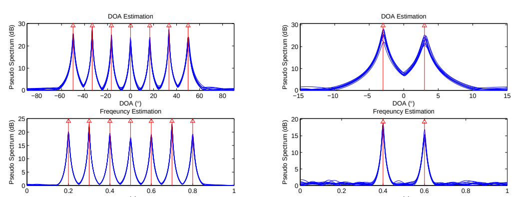
Step 1: Take the transpose of $\{\mathbf{Y}_i\}_{i=0}^{N-m}$.

Step 2: Construct $\mathbf{F}_\omega$ and $\mathbf{G}_\omega$ based on (26) and (27).

Step 3: Utilize (29) to obtain the pseudo spectrum $P(\omega)$.

Step 4: Estimate the frequencies by searching for the peaks of $P(\omega)$.

(c) Pair the DOA and frequency estimates using the procedure in Section 3.3.



(a) $P = 7$ (b) $P = 2$

Fig. 1. Pseudo output power spectrum. Vertical lines show the true DOAs and frequencies. ($M = 8$, $m = 20$, $N = 50$, SNR$= 10$ dB)

with frequencies being 0.56π, -0.3π, 0, -0.8π, 0.7π, -0.4π and 0.3π and DOAs being $-23°$, $-45°$, $1°$, $19°$, $30°$, $-9°$ and $50°$, respectively. It is seen from Fig. 2 that the proposed algorithm can correctly pair the frequencies and DOAs.

In the third example, the mean square error (MSE) performance of the proposed scheme is compared with that of the ESPRIT [4], JAFE [5] and SLS-ESPRIT [6] methods as a function of SNR. Consider a scenario where there are two signals with frequencies $\omega_1 = 0.2\pi$ and $\omega_2 = 0.56\pi$ located at $\theta_1 = 1°$ and $\theta_2 = 10°$, respectively. We set $\alpha_1 = 0.4 + 0.5j$, $\alpha_2 = 0.9 - 0.7j$, $N = 40$ and $m = 16$. The SNR is varied from 0 dB to 30 dB. Furthermore, we assume that the number of signals is known to the ESPRIT, JAFE and SLS-ESPRIT algorithms. It is shown in Fig. 3 that for DOA estimation, the proposed and the SLS-ESPRIT schemes achieve almost the same performance and both of them outperform the ESPRIT and JAFE algorithms. For frequency estimation, however, the

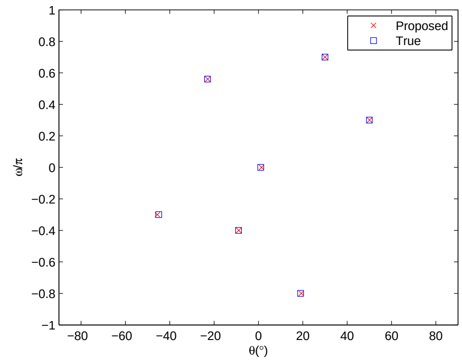


Fig. 2. DOA and frequency estimation for seven targets. ($M = 8$, $m = 16$, $N = 40$, SNR $= 10$ dB, $P = 7$)

performance of the proposed method is a little bit inferior to its counterparts. Compared to the ESPRIT, JAFE and SLS-ESPRIT algorithms which rely on the *a priori* knowledge of source number, the main advantage of the proposed approach is that it does not need the source number information. Therefore, it is much more attractive for practical applications.

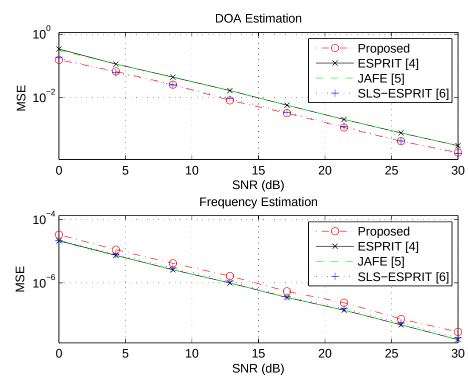


Fig. 3. MSE versus SNR. ($M = 8$, $m = 16$, $N = 40$, $P = 2$)

5. CONCLUSION

In this paper, we have devised a joint DOA and frequency estimation algorithm based on the joint diagonalization structure of a set of transformed data matrices. The most favorable advantage of the proposed scheme is that it does not require the source number information. Such an advantage is highly desirable for practical applications since detection of the source number is usually a very difficult task. Moreover, extensive simulation results demonstrate that the proposed approach is able to provide comparable estimation performance with the state-of-the-art methods.

6. REFERENCES

[1] J. D. Lin, W. H. Fang, Y. Y. Wang and J. T. Chen, "FSF MUSIC for joint DOA and frequency estimation and its performance analysis," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4529-4542, 2006.

[2] M. Djeddou, A. Belouchrani and S. Aouada, "Maximum likelihood angle-frequency estimation in partially known correlated noise for low-elevation targets," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3057-3064, 2005.

[3] M. Viberg and P. Stoica, "A computationally efficient method for joint direction finding and frequency estimation in colored noise," *Proc. IEEE Conf. Record of the Thirty-Second Asilomar Conf., Signals, Systems and Computers*, vol. 2, pp. 1547-1551, Pacific Grove, USA, 1988.

[4] A. N. Lemma, A. van der Veen and E. F. Deprettere, "Joint angle-frequency estimation using multiresolution," *Proc. IEEE Int. Conf., Acoustics, Speech, and Signal Processing*, vol. 4, pp. 1957-1960, Seattle, USA, 1998.

[5] A. N. Lemma, A. van der Veen and E. F. Deprettere, "Analysis of joint angle-frequency estimation using ESPRIT," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1264-1283, 2003.

[6] C. Qian, L. Huang, Y. Shi, and H. C. So, "Joint angle and frequency estimation using structured least squares," *Proc. Intern. Conf. Acoust., Speech, and Signal Process. (ICASSP 2014)*, pp. 2996-3000, Florence, Italy, 2014.

[7] M. Haardt. "Structured least squares to improve the performance of ESPRIT-type algorithms," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 792-799, 1997.

[8] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acous. Speech and Signal Process.*, vol. 33, no. 2, pp. 387-392, 1985.

[9] L. Huang and H. C. So, "Source enumeration via MDL criterion based on linear shrinkage estimation of noise subspace covariance matrix," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4806-4821, 2013.

[10] Q. Cheng, P. Pal, M. Tsuji and Y. Hua, "An MDL algorithm for detecting more sources than sensors using outer-products of array output," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6438-6453, 2014.

[11] P. M. Djuric, "A model selection rule for sinusoids in white Gaussian noise," *IEEE Trans. Signal Process.*, vol. 44, no. 7, pp. 1744-1751, 1996.

[12] W. J. Zeng, X. L. Li and X. D. Zhang, "Direction-of-arrival estimation based on the joint diagonalization structure of multiple fourth-order cumulant matrices," *IEEE Signal Process. Letters*, vol. 16, no. 3, pp. 164-167, 2009.

[13] W. J. Zeng and X. L. Li, "High-resolution multiple wideband and nonstationary source localization with unknown number of sources," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3125-3136, 2010.

[14] C. Qian, L. Huang, W.-J. Zeng and H. C. So, "Direction-of-arrival estimation for coherent signals without knowledge of source number," *IEEE Sensors Journal*, vol. 14, no. 9, pp. 3267-3273, Sep. 2014.

[15] C. Qian, L. Huang, Y. Xiao and H. C. So, "Localization of coherent signals without source number knowledge in unknown spatially correlated Gaussian noise," *Signal Process.*, vol. 111, pp. 170-178, 2015.

[16] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2267-2280, 1992.