Exact Unconditional ML Estimation of DOA

Masakiyo Suzuki* and Haihua Chen*
*Graduate School of Engineering
Kitami Institute of Technology
165 Koencho Kitami Hokkaido 090-8507 Japan
Email: masakiyo@mail.kitami-it.ac.jp

***Abstract*—This paper presents an exact formulation of Stochastic or Unconditional Maximum Likelihood (UML) estimation for directions-of-arrival (DOA) finding. In the previous formulation of UML estimation, an important condition is missing. That is the non-negative definiteness of the covariance matrix of signal components without additive noises. Because of the lack of the important condition, inadequate global solution appears in the solution space and global search fails to find adequate solution. We have derived an exact formulation including this important condition. Then the inadequate global solution disappears and global search finds adequate solution.**

I. Introduction

The localization of multiple signal sources by a passive sensor array is of great importance in a wide variety of fields, such as radar, geophysics, radio-astronomy, biomedical engineering, communications, underwater acoustics, and so on. The basic problem in this context is to estimate directions-of-arrival (DOA) of narrow-band signal sources located in the far field of the array. A number of super-resolution techniques have been introduces, such as Conditional or Deterministic Maximum Likelihood (CML) method [1], [2], [3], [6], [9], Unconditional or Stochastic ML (UML) method [6], [8], [7], [9], MUSIC [11], [12], ESPRIT [10], Weighted Subspace Fitting (WSF) [13] and the Bayesian method [14].

The CML, UML, WSF and Bayesian techniques have properties superior to other methods since they can handle coherent signals without any preprocessing, such as the spatial smoothing [12]. They can also handle small number of snapshots, although the Bayesian method [14] is formulated only for a single snapshot. It is known that the UML estimator shows better solutions for coherent signals than the others.

In the previous formulation of UML estimation, an important condition is missing. That is the non-negative definiteness of the covariance matrix of signal components without additive noises. Because of the lack of the important condition, inadequate global solution appears in the solution space and global search fails to find adequate solution. We have derived an exact formulation including this important condition. Then the inadequate global solution disappears and global search finds adequate solution.

II. Problem Formulation

Consider an array composed of p sensors with arbitrary locations and arbitrary directional characteristics, and assume that q narrow-band source, centered around a known frequency, say ω_0, impinge on the array from distinct directions θ_1, θ_2, ..., θ_q, respectively.

Using complex envelope representation, the p-dimensional vector received by the array can be expressed as

$$\boldsymbol{x}(t) = \sum_{k=1}^{q} \boldsymbol{a}(\theta_k) s_k(t) + \boldsymbol{n}(t), \tag{1}$$

where $s_k(t)$ is the k-th signal received at a certain reference point. $\boldsymbol{n}(t)$ is a p-dimensional noise vector. $\boldsymbol{a}(\theta)$ is the "steering vector" of the array towards direction θ, which is represented as

$$\boldsymbol{a}(\theta) = [a_1(\theta)e^{-j\omega_0\tau_1(\theta)}, ..., a_p(\theta)e^{-j\omega_0\tau_p(\theta)}]^T \tag{2}$$

where $a_i(\theta)$ is the amplitude response of the i-th sensor to a wave-front impinging from the direction θ. $\tau_i(\theta)$ is the propagation delay between the i-th sensor and the reference point. The superscript T denotes the transpose of a matrix.

In the matrix notation, (1) can be rewritten as

$$\boldsymbol{x}(t) = \boldsymbol{A}(\Theta)\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{3}$$
$$\boldsymbol{A}(\Theta) = [\ \boldsymbol{a}(\theta_1)\ \boldsymbol{a}(\theta_2)\ \cdots\ \boldsymbol{a}(\theta_q)\ ], \tag{4}$$
$$\boldsymbol{s}(t) = [\ s_1(t)\ s_2(t)\ \cdots\ s_q(t)\ ]^T, \tag{5}$$
$$\Theta = \{\ \theta_1\ \theta_2\ \cdots\ \theta_q\ \}. \tag{6}$$

Suppose that the received vectors $\boldsymbol{x}(t)$ is sampled at N time instants t_1, t_2, ..., t_N and define the matrix of the sampled data as

$$\boldsymbol{X} = [\ \boldsymbol{x}(t_1)\ \boldsymbol{x}(t_2)\ \cdots\ \boldsymbol{x}(t_N)\ ]. \tag{7}$$

The problem of DOA finding is to be stated as follows. Given the sampled data $\boldsymbol{X}$, obtain a set of estimated directions

$$\hat{\Theta} = \{\ \hat{\theta}_1\ \hat{\theta}_2\ \cdots\ \hat{\theta}_q\ \}. \tag{8}$$

of θ_1, θ_2, ..., θ_q.

III. ML Estimation

In this section, an exact formulation of UML estimation of DOA is derived.

To solve the problem of ML estimation of DOA, we make the following assumptions.

A1) The array configuration is known and any p steering vectors for different p directions are linearly independent.

A2) $\boldsymbol{n}(t_i)$ are statistically independent samples from a complex Gaussian random vector with zero mean and the covariance matrix $\sigma^2 \boldsymbol{I}_p$, where $\boldsymbol{I}_p$ is a $p \times p$ identity matrix.

A3) $\boldsymbol{s}(t_i)$ are statistically independent samples from a complex Gaussian random vector with zero mean and a certain covariance matrix $\boldsymbol{S}$ with $\text{rank}\{\boldsymbol{S}\} = r$, where $r \leq q$. In the case of $r < q$, the signals are coherent or fully correlated which happens, e.g., in specular multi-path propagation. $\boldsymbol{s}(t_i)$ are independent of $\boldsymbol{n}(t_j)$ for any i and j and satisfy

$$\text{rank}\{[\ \boldsymbol{s}(t_1)\ \boldsymbol{s}(t_2)\ \cdots\ \boldsymbol{s}(t_N)\ ]\} = r. \tag{9}$$

A4) q is known.

A5) p, q and r satisfy the condition that a unique solution of DOA exists in the noise-free case. When the direction θ is expressed by a single real parameter, the sufficient condition of the uniqueness is given by $q < 2rp/(2r+1)$ and the necessary condition is given by $q \leq 2rp/(2r+1)$ [15].

A. Stochastic Model

According to the assumptions A1) to A5), $\boldsymbol{x}(t)$ is a p-dimensional complex Gaussian random vector with zero mean and the covariance matrix $\boldsymbol{R}$,

$$\boldsymbol{R} = E\{\boldsymbol{x}(t)\boldsymbol{x}^H(t)\} = \boldsymbol{A}(\Theta)\boldsymbol{S}\boldsymbol{A}^H(\Theta) + \sigma^2 \boldsymbol{I}_p \tag{10}$$

where the superscript H denotes the Hermitian transpose of a matrix. The probability density function of $\boldsymbol{X}$ is given as

$$f(\boldsymbol{X}) = \left(\frac{1}{\pi^p \det\{\boldsymbol{R}\}}\right)^N \exp\left\{-\sum_{n=1}^{N} \boldsymbol{x}^H(t_n)\boldsymbol{R}^{-1}\boldsymbol{x}(t_n)\right\} \tag{11}$$

The covariance matrix $\boldsymbol{R}$ is parametrized by Θ, $\boldsymbol{S}$ and σ^2. Θ indicates a set of directions. $\boldsymbol{S}$ is a non-negative Hermitian matrix with $\text{rank}\{\boldsymbol{S}\} = r$. σ^2 is non-negative real number.

The log-likelihood function of unknown parameters Θ, $\boldsymbol{S}$ and σ^2 for given $\boldsymbol{X}$ is defined as

$$\begin{aligned} L(\Theta, \boldsymbol{S}, \sigma^2) &= -N \ln\det\{\boldsymbol{R}\} - \sum_{n=1}^{N} \boldsymbol{x}^H(t_n)\boldsymbol{R}^{-1}\boldsymbol{x}(t_n) \\ &= -N \left(\ln\det\{\boldsymbol{R}\} + \text{tr}\{\boldsymbol{R}^{-1}\tilde{\boldsymbol{R}}\}\right) \end{aligned} \tag{12}$$

where a constant term is ignored and $\tilde{\boldsymbol{R}}$ is the sample covariance matrix defined by

$$\tilde{\boldsymbol{R}} = \frac{1}{N}\boldsymbol{X}\boldsymbol{X}^H. \tag{13}$$

Using a square root matrix of a non-negative definite matrix[1], the $p \times q$ matrix $\boldsymbol{V}_S(\Theta)$ composed of the orthonormal system of the signal subspace spanned by $\boldsymbol{A}(\Theta)$ is represented as

$$\boldsymbol{V}_S(\Theta) = \boldsymbol{A}(\Theta)\left(\boldsymbol{A}^H(\Theta)\boldsymbol{A}(\Theta)\right)^{-H/2} \tag{14}$$

and define the unitary matrix

$$\boldsymbol{G}(\Theta) = [\boldsymbol{V}_S(\Theta)\ \boldsymbol{V}_N(\Theta)] \tag{15}$$

where $\boldsymbol{V}_N(\Theta)$ is a $p \times (p-q)$ matrix composed of the orthonormal system of the noise subspace which is an orthogonal complement of the signal subspace. Then, the covariance matrix $\boldsymbol{R}$ can be represented as

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{V}_S(\Theta)\boldsymbol{P}\boldsymbol{V}_S^H(\Theta) + \sigma^2 \boldsymbol{I}_p \\ &= \boldsymbol{G}(\Theta)\begin{bmatrix} \boldsymbol{R}_{SS} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma^2 \boldsymbol{I}_{p-q} \end{bmatrix}\boldsymbol{G}^H(\Theta) \end{aligned} \tag{16}$$

where $\boldsymbol{P} = \left(\boldsymbol{A}^H(\Theta)\boldsymbol{A}(\Theta)\right)^{H/2}\boldsymbol{S}\left(\boldsymbol{A}^H(\Theta)\boldsymbol{A}(\Theta)\right)^{1/2}$ and $\boldsymbol{R}_{SS} = \boldsymbol{P} + \sigma^2 \boldsymbol{I}_q$.

Let $\lambda_1, \lambda_2, ..., \lambda_q$ be eigenvalues of $\boldsymbol{R}_{SS}$. Since $\text{rank}\{\boldsymbol{S}\} = \text{rank}\{\boldsymbol{P}\} = r$, it holds that $\lambda_k > \sigma^2$ for $k = 1, 2, ..., r$ and $\lambda_{r+1} = \lambda_{r+2} = ... \lambda_q = \sigma^2$. Let $\boldsymbol{v}_k$ be eigenvectors corresponding to λ_k for $k = 1, 2, ..., q$. Define p-dimensional vectors

$$\boldsymbol{u}_k = [\boldsymbol{v}_k^T\ 0\ 0\ ...\ 0]^T \tag{17}$$

The model of the covariance matrix $\boldsymbol{R}$ with $\text{rank}\{\boldsymbol{S}\} = \text{rank}\{\boldsymbol{P}\} = r$ is given as follows.

$$\boldsymbol{R} = \boldsymbol{G}(\Theta)\left(\sum_{k=1}^{r}(\lambda_k - \sigma^2)\boldsymbol{u}_k\boldsymbol{u}_k^H + \sigma^2 \boldsymbol{I}_p\right)\boldsymbol{G}^H(\Theta). \tag{18}$$

The set of unknown parameters is $\{\ \Theta, \lambda_1, \lambda_2, ...\ \lambda_r, \boldsymbol{v}_1, \boldsymbol{v}_2, ...\ \boldsymbol{v}_r, \sigma^2\ \}$, where Θ is a set of directions and σ^2 is a non-negative real value as mentioned above. Furthermore the following conditions are imposed on the parameters: λ_k is a real value and satisfies $\lambda_k > \sigma^2$ for $k = 1, 2, ..., r$ and $\boldsymbol{v}_1, \boldsymbol{v}_2, ...\ \boldsymbol{v}_r$ satisfy

$$\boldsymbol{v}_i^H \boldsymbol{v}_j = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}. \tag{19}$$

[1]For a non-negative definite matrix $\boldsymbol{B}$, the square root matrix $\boldsymbol{B}^{1/2}$ is defined as a matrix $\boldsymbol{C}$ which satisfies $\boldsymbol{B} = \boldsymbol{C}\boldsymbol{C}^H$. The following notations are used, $(\boldsymbol{B}^{1/2})^H = \boldsymbol{B}^{H/2}$, $(\boldsymbol{B}^{1/2})^{-1} = \boldsymbol{B}^{-1/2}$, $((\boldsymbol{B}^{1/2})^H)^{-1} = ((\boldsymbol{B}^{1/2})^{-1})^H = \boldsymbol{B}^{-H/2}$, and we have $(\boldsymbol{B}^{-1})^{1/2} = \boldsymbol{B}^{-H/2}$.

B. Likelihood Function

Using the model of $\boldsymbol{R}$ in (18), the inverse of the matrix $\boldsymbol{R}$ is represented as

$$\boldsymbol{R}^{-1} = \boldsymbol{G}(\Theta)\left(\sum_{k=1}^{r}(\frac{1}{\lambda_k} - \frac{1}{\sigma^2})\boldsymbol{u}_k\boldsymbol{u}_k^H + \frac{1}{\sigma^2}\boldsymbol{I}_p\right)\boldsymbol{G}^H(\Theta). \tag{20}$$

The log-likelihood function in (12) for a fixed r is rewritten as

$$\begin{aligned} & L_r(\Theta, \lambda_1, ..., \lambda_r, \boldsymbol{v}_1, ..., \boldsymbol{v}_r, \sigma^2) \\ & = -N\left(\ln\left\{\lambda_1\lambda_2...\lambda_r\left(\sigma^2\right)^{p-r}\right\} + \sum_{k=1}^{r}\frac{l_k}{\lambda_k}\right. \\ & \qquad \left. + \frac{1}{\sigma^2}\left\{\mathrm{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{k=1}^{r} l_k\right\}\right) \end{aligned} \tag{21}$$

where

$$\begin{aligned} l_k &= \boldsymbol{u}_k^H\boldsymbol{G}^H(\Theta)\tilde{\boldsymbol{R}}\boldsymbol{G}(\Theta)\boldsymbol{u}_k \\ &= \boldsymbol{v}_k^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k \quad \text{for } k = 1, 2, ..., r \end{aligned} \tag{22}$$

$$\tilde{\boldsymbol{R}}_{SS}(\Theta) = \boldsymbol{V}_S^H(\Theta)\tilde{\boldsymbol{R}}\boldsymbol{V}_S(\Theta). \tag{23}$$

C. Maximization with Respect to λ_1, λ_2, ... λ_r and σ^2

Given Θ and an orthonormal system $\{\ \boldsymbol{v}_1,\ \boldsymbol{v}_2,\ ...,\ \boldsymbol{v}_r\ \}$ of an r-dimensional subspace in q-dimensional complex Euclid space $\mathcal{C}^q$, we consider the maximization of L_r in (21) with respect to λ_1, λ_2, ..., λ_r and σ^2 under the conditions $\lambda_k \geq \sigma^2$ for $k = 1, 2, ..., r$ instead of $\lambda_k > \sigma^2$.

We assume with no loss of generality that $\boldsymbol{v}_k$ for $k = 1$, 2, ..., r are ordered so that $l_1 \geq l_2 \geq ... \geq l_r$ where l_k are defined in (22) and define

$$\sigma_0^2 = \frac{1}{p}\mathrm{tr}\{\tilde{\boldsymbol{R}}\} \tag{24}$$

$$\sigma_k^2 = \frac{1}{p-k}\left(\mathrm{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{i=1}^{k} l_i\right) \quad \text{for } k\text{=1, 2, ..., } r. \tag{25}$$

Let ρ be an index in { 0, 1, ..., r } which satisfies one of followings

$$\begin{array}{ll} \sigma_\rho^2 \geq l_1 & \text{for } \rho = 0 \\ l_\rho > \sigma_\rho^2 \geq l_{\rho+1} & \text{for } \rho = 1, 2, ..., r-1\ . \\ l_r > \sigma_\rho^2 & \text{for } \rho = r \end{array} \tag{26}$$

The maximum likelihood estimators of λ_1, λ_2, ..., λ_r and σ^2 for fixed Θ and $\{\ \boldsymbol{v}_1,\ \boldsymbol{v}_2,\ ...,\ \boldsymbol{v}_r\ \}$ under the conditions $\lambda_k \geq \sigma^2$ for $k = 1, 2, ..., r$ are obtained as follows:

$$\lambda_k = l_k \quad \text{for } k = 1, 2, ..., \rho \tag{27}$$

$$\lambda_k = \sigma_\rho^2 \quad \text{for } k = \rho+1, \rho+2, ..., r \tag{28}$$

$$\sigma^2 = \sigma_\rho^2. \tag{29}$$

Of course, (27) is ignored if $\rho = 0$ and also (28) is ignored if $\rho = r$.

1) Proof of the uniqueness of ρ: The uniqueness of the index ρ is proved as follows. Since σ_k^2 depends on l_k as shown in

$$\sigma_k^2 = \frac{1}{p-k}\left((p-k+1)\sigma_{k-1}^2 - l_k\right), \tag{30}$$

the following equivalence in inequalities is derived.

$$\sigma_k^2 < l_k \quad \Leftrightarrow \quad \sigma_{k-1}^2 < l_k \tag{31}$$

$$\sigma_k^2 > l_k \quad \Leftrightarrow \quad \sigma_{k-1}^2 > l_k \tag{32}$$

$$\sigma_k^2 = l_k \quad \Leftrightarrow \quad \sigma_{k-1}^2 = l_k \tag{33}$$

$$\text{for } k = 1, 2, ..., r.$$

If $l_k > \sigma_k^2 \geq l_{k+1}$ is not true for $k = 1, 2, ..., r-1$, then we have two cases that $l_k \geq l_{k+1} > \sigma_k^2$ or $\sigma_k^2 \geq l_k \geq l_{k+1}$ for $k = 1, 2, ..., r-1$ because of $l_k \geq l_{k+1}$. It follows from the former case that $l_r > \sigma_r^2$ or $\rho = r$. From the later case, it follows that $\sigma_0^2 \geq l_1$ or $\rho = 0$. Therefore there exists at least one index of ρ which satisfies one condition in (26).

Assuming $\rho > 0$ and $l_\rho > \sigma_\rho^2$, then $l_\rho > \sigma_{\rho-1}^2$ follows from (31) and $l_{\rho-1} > \sigma_{\rho-1}^2$ follows from $l_{\rho-1} \geq l_\rho$. Applying the same procedure recursively, we have

$$l_{k+1} > \sigma_k^2 \quad \text{for } k = 0, 1, ..., \rho-1. \tag{34}$$

This indicates that $\sigma_k^2 \geq l_{k+1}$ does not hold for $k = 0$, 1, ..., $\rho-1$ if $l_\rho > \sigma_\rho^2$ for $\rho > 0$.

Assuming $\rho < r$ and $\sigma_\rho^2 \geq l_{\rho+1}$, then $\sigma_{\rho+1}^2 \geq l_{\rho+1}$ follows from (32) and (33) and $\sigma_{\rho+1}^2 \geq l_{\rho+2}$ follows from $l_{\rho+1} \geq l_{\rho+2}$. Applying the same procedure recursively, we have

$$\sigma_k^2 \geq l_k \quad \text{for } k = \rho+1, \rho+2, ..., r. \tag{35}$$

This indicates that $l_k > \sigma_k^2$ does not hold for $k = \rho+1$, $\rho+2$, ..., r if $\sigma_\rho^2 \geq l_{\rho+1}$ for $\rho < r$.

Therefore the index ρ which satisfies one condition in (26) is unique.

2) Proof of (27), (28) and (29): The function

$$h(\lambda) = -\left(\ln\lambda + \frac{l}{\lambda}\right) \tag{36}$$

has a single peak at $\lambda = l$. Consider the case that the domain of λ is restricted to $\sigma^2 \leq \lambda$. If $\sigma^2 \leq l$, the maximum value of $h(\lambda)$ is obtained at $\lambda = l$. If $l \leq \sigma^2$, the maximum value of $h(\lambda)$ is obtained at $\lambda = \sigma^2$.

Assuming that

$$l_\rho > \sigma^2 \geq l_{\rho+1}, \tag{37}$$

and maximizing L_r with respect to λ_1, λ_2, ..., λ_r under the conditions $\lambda_k \geq \sigma^2$ for $k = 1, 2, ..., r$, we can readily obtain (27) and

$$\lambda_k = \sigma^2 \quad \text{for } k = \rho+1, \rho+2, ..., r. \tag{38}$$

Substituting (38) into (21), L_r in (21) becomes equivalent to L_ρ as follows.

$$L_r(\Theta, \lambda_1, ..., \lambda_r, \boldsymbol{v}_1, ..., \boldsymbol{v}_r, \sigma^2)|_{\lambda_k=\sigma^2 \text{ for } k=\rho+1,\rho+2,...,r}$$
$$= -N\left(\ln\left\{\lambda_1\lambda_2...\lambda_\rho\left(\sigma^2\right)^{p-\rho}\right\} + \sum_{k=1}^{\rho}\frac{l_k}{\lambda_k} + \frac{1}{\sigma^2}\left\{\mathrm{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{k=1}^{\rho} l_k\right\}\right)$$
$$= L_\rho(\Theta, \lambda_1, ..., \lambda_\rho, \boldsymbol{v}_1, ..., \boldsymbol{v}_\rho, \sigma^2) \tag{39}$$

Maximizing L_ρ with respect to σ^2, we can readily obtain $\sigma^2 = \sigma_\rho^2$ or (29). It is consistent with the assumption (37).

3) Maximum Log-Likelihood Function: Substituting (27), (28) and (29) into into (21), we have the formulation of the log-likelihood function after maximizing with respect to λ_1, λ_2, ... λ_r and σ^2 as follows.

$$L_r(\Theta, \boldsymbol{v}_1, ..., \boldsymbol{v}_r) = L_\rho(\Theta, \boldsymbol{v}_1, ..., \boldsymbol{v}_\rho) = -N\ln\left\{\lambda_1\lambda_2...\lambda_\rho\left(\sigma^2\right)^{p-\rho}\right\} \tag{40}$$

Finally in this subsection, we have to note that there is no ML solution for the model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = r$ unless $\rho = r$, because the solutions in (28) belong to the marginal set of the solution space $\lambda_k > \sigma^2$ for $k = \rho+1, \rho+2 ..., r$ and the marginal set does not included in the solution space. Instead the ML solution for each model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = k$ for $k = 0, 1, ..., \rho$ is obtained as

$$\lambda_i = l_i \quad \text{for } i = 1, 2, ..., k \tag{41}$$
$$\sigma^2 = \sigma_k^2. \tag{42}$$

D. Maximization with Respect to $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ... $\boldsymbol{v}_r$

Next we consider the maximization of L_r with respect to $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_r$ assuming that $l_k > \sigma_r^2$ for $k = 1, 2, ..., r$.

Introducing Lagrange's multipliers to realize the constraints in (19) in a new criterion, taking derivatives of the new criterion with respect to unknown real parameters in $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_r$ and making the derivatives equal to zero, then we can obtain a set of r equations in the complex form as follows.

$$\left(\frac{1}{\sigma_r^2} - \frac{1}{l_k}\right)\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k = \alpha_{k1}\boldsymbol{v}_1 + \alpha_{k2}\boldsymbol{v}_2 + ... + \alpha_{kr}\boldsymbol{v}_r \tag{43}$$

$$\alpha_{ki} = \bar{\alpha}_{ik} \qquad \text{for } k, i = 1, 2, ..., r. \tag{44}$$

where α_{ki} is a complex number determined by Lagrange's multipliers and the bar indicates the complex conjugate.

Because of (19) and (22), multiplying (43) by $\boldsymbol{v}_i^H$ from the left, we have

$$\alpha_{ki} = \left(\frac{1}{\sigma_r^2} - \frac{1}{l_k}\right)\boldsymbol{v}_i^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k. \tag{45}$$

From (44), we also have

$$\alpha_{ki} = \bar{\alpha}_{ki} = \left(\frac{1}{\sigma_r^2} - \frac{1}{l_i}\right)\left(\boldsymbol{v}_k^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_i\right)^H = \left(\frac{1}{\sigma_r^2} - \frac{1}{l_i}\right)\boldsymbol{v}_i^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k. \tag{46}$$

1) In the Case of $l_k \neq l_i$: From (45) and (46), we have $\alpha_{ki} = 0$ if $l_k \neq l_i$ for $k \neq i$. Therefore if $l_k \neq l_i$ for all combinations of k and i that $k \neq i$, then we have the eigenequation

$$\left(\frac{1}{\sigma_r^2} - \frac{1}{l_k}\right)\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k = \alpha_{kk}\boldsymbol{v}_k \quad \text{for } k = 1, 2, ..., r. \tag{47}$$

Therefore $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_r$ must be eigenvectors of the matrix $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ and orthogonal each other. The Hermitian form $l_k = \boldsymbol{v}_k^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k$ in (22) is an eigenvalue of the eigenvector $\boldsymbol{v}_k$ for $k = 1, 2, ..., r$.

2) In the Case of $l_1 = l_2 = ... = l_\mu$: Let $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_r$ be the solutions which maximize L_r and assume that $l_1 = l_2 = ... = l_\mu (= l_0)$ for a certain μ that $\mu \leq r$. Then the maximum log-likelihood function is rewritten as

$$L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) = -N\ln\left\{(l_0^\mu l_{\mu+1}...l_r\left(\sigma_r^2\right)^{p-r}\right\}$$
$$\sigma_r^2 = \frac{1}{p-r}\left(\mathrm{tr}\{\tilde{\boldsymbol{R}}\} - \left(\mu l_0 + \sum_{k=\mu+1}^{r} l_k\right)\right). \tag{48}$$

and the equations (43) and (44) are reduced as follows.

$$\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k = \alpha'_{k1}\boldsymbol{v}_1 + \alpha'_{k2}\boldsymbol{v}_2 + ... + \alpha'_{k\mu}\boldsymbol{v}_\mu \tag{49}$$
$$\alpha'_{kk} = \boldsymbol{v}_k^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k = l_0 \tag{50}$$
$$\alpha'_{ki} = \boldsymbol{v}_i^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k = \bar{\alpha}'_{ik} \tag{51}$$
$$\text{for } k, i = 1, 2, ..., \mu.$$

Define the following 2×2 matrix $\boldsymbol{A}_{12}$

$$\begin{bmatrix}\boldsymbol{v}_1^H \\ \boldsymbol{v}_2^H\end{bmatrix}\tilde{\boldsymbol{R}}_{SS}(\Theta)[\boldsymbol{v}_1\ \boldsymbol{v}_2] = \begin{bmatrix} l_0 & \alpha'_{12} \\ \bar{\alpha}'_{12} & l_0\end{bmatrix} = \boldsymbol{A}_{12}. \tag{52}$$

The matrix $\boldsymbol{A}_{12}$ has two eigenvalues, i.e., $l'_1 = l_0 - |\alpha_{12}|$ and $l'_2 = l_0 + |\alpha_{12}|$. Let $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ be the unit eigenvectors corresponding to l'_1 and l'_2, respectively. Define

$$\boldsymbol{v}'_1 = [\boldsymbol{v}_1\ \boldsymbol{v}_2]\boldsymbol{e}_1 \text{ and } \boldsymbol{v}'_2 = [\boldsymbol{v}_1\ \boldsymbol{v}_2]\boldsymbol{e}_2. \tag{53}$$

Then we have

$$\boldsymbol{v}_1'^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}'_1 = \boldsymbol{e}_1^H\tilde{\boldsymbol{A}}_{12}\boldsymbol{e}_1 = l'_1 = l_0 - |\alpha_{12}|, \tag{54}$$
$$\boldsymbol{v}_2'^H\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}'_2 = \boldsymbol{e}_2^H\tilde{\boldsymbol{A}}_{12}\boldsymbol{e}_2 = l'_2 = l_0 + |\alpha_{12}|. \tag{55}$$

If it holds that $l'_1 \geq \sigma_r^2$, then we readily obtain the following inequality.

$$L_r(\Theta, \boldsymbol{v}'_1, \boldsymbol{v}'_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r)$$
$$= -N\ln\left\{(l_0^2 - |\alpha'_{12}|^2)l_0^{\mu-2}l_{\mu+1}...l_r\left(\sigma_r^2\right)^{p-r}\right\}$$
$$\geq L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r). \tag{56}$$

The equal sign holds iff $|\alpha'_{12}| = 0$. Therefore we have $|\alpha'_{12}| = 0$ because $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_r$ should give the maximum of L_r.

Although l'_2 is greater than σ_r^2, l'_1 may not. In the case that $l'_1 \le \sigma_r^2$, the log-likelihood function is rewritten as

$$L_r(\Theta, \boldsymbol{v}'_1, \boldsymbol{v}'_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) = L_{r-1}(\Theta, \boldsymbol{v}'_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r)$$
$$= -N \ln \left\{ (l_0 + |\alpha'_{12}|) l_0^{\mu-2} l_{\mu+1} ... l_r \left(\sigma'^2_{r-1} \right)^{p-r+1} \right\} \quad (57)$$

where

$$\sigma'^2_{r-1} = \frac{1}{p-r+1} \left((p-r)\sigma_r^2 + l'_1 \right). \quad (58)$$

Because l'_1 is written in the Hermitian form of the non-negative definite matrix $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ as shown in (54), l'_1 has a non-negative real value. Under the conditions $0 \le l'_1 \le \sigma_r^2$ in addition to $\sigma_r^2 < l_0$, we can derive the following inequality

$$L_{r-1}(\Theta, \boldsymbol{v}'_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) > L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r). \quad (59)$$

This conflicts with the assumption that $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_r$ are the solutions which maximize L_r. Therefore $l'_1 \ge \sigma_r^2$ and $|\alpha_{12}|$ vanishes.

From the same discussion as above, we obtain $\alpha_{ki} = 0$ for all combinations of k and i that $k \ne i$ in { 1, 2, ..., μ }. Then the equations (49) are rewritten as

$$\tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_k = l_0 \boldsymbol{v}_k \qquad \text{for } k = 1, 2, ..., \mu. \quad (60)$$

Therefore l_0 becomes an eigenvalue of $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ with μ multiplicity and $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ ... $\boldsymbol{v}_\mu$ are the corresponding unit eigenvectors orthogonal each other.

3) Selection of Eigenvalues: Let $l_1 \ge l_2 \ge ... \ge l_q$ be the eigenvalues of $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ for a certain fixed Θ. As well as the definition in (25), we define

$$\sigma_k^2 = \frac{1}{p-k} \left(\text{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{i=1}^{k} l_i \right) \quad \text{for } k\text{=1, 2, ..., } q. \quad (61)$$

and σ_0^2 as in (24). Also as well as the definition of ρ in (26), we define η be an index in { 0, 1, ..., q } which satisfies one of followings

$$\begin{array}{ll} \sigma_\eta^2 \ge l_1 & \text{for } \eta = 0 \\ l_\eta > \sigma_\eta^2 \ge l_{\eta+1} & \text{for } \eta = 1, 2, ..., q-1 \\ l_q > \sigma_\eta^2 & \text{for } \eta = q \end{array} . \quad (62)$$

It is apparent that there is no ML solution for the model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = r$ if $r > \eta$. In the case of $r \le \eta$, the maximum log-likelihood function of the model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = r$ is given as

$$L_r(\Theta) = -N \ln \left\{ l_1 l_2 ... l_r \left(\sigma_r^2 \right)^{p-r} \right\} \quad (63)$$

where σ_r^2 is defined in (61). In other words, { l_1, l_2, ..., l_r } is the best selection of all choices of r eigenvalues from { l_1, l_2, ..., l_q }. It is proved as follows.

Let $l_{k_1}, l_{k_2}, ..., l_{k_{r-1}}$ be a certain choice of $r-1$ eigenvalues from all eigenvalues. We consider how to select the r-th eigenvalue form remaining $p - r + 1$ eigenvalues. We assume that

$$l_{k_i} > \sigma_{r-1}^2(k_1, k_2, ..., k_{r-1}) \quad \text{for } i = 1, 2, ..., r-1, \quad (64)$$

$$\sigma_{r-1}^2(k_1, k_2, ..., k_{r-1}) = \frac{1}{p-r+1} \left(\text{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{i=1}^{r-1} l_{k_i} \right), \quad (65)$$

and define

$$L_r(\lambda) = -N \ln \left\{ l_{k_1} l_{k_2} ... l_{k_{r-1}} \lambda \left(\sigma_r^2(\lambda) \right)^{p-r} \right\}, \quad (66)$$

$$\sigma_r^2(\lambda) = \frac{1}{p-r} \left((p-1+1)\sigma_{r-1}^2(k_1, k_2, ..., k_{r-1}) - \lambda \right). \quad (67)$$

The curve of $L_r(\lambda)$ as a function of λ has a single valley within the domain of $0 \le \lambda \le (p-1+1)\sigma_{r-1}^2(k_1, k_2, ..., k_{r-1})$ and becomes minimal at $\lambda = \sigma_{r-1}^2(k_1, k_2, ..., k_{r-1})$.

Let l_{k_r} be the largest eigenvalue of remaining $p - r + 1$ eigenvalues. In order to maximize $L_r(\lambda)$ by substituting one of remaining eigenvalues into λ, the assignment $\lambda = l_{k_r}$ is the best, where $l_{k_r} > \sigma_{r-1}^2(k_1, k_2, ..., k_{r-1})$ is guaranteed by the assumption $r \le \eta$.

Removing the smallest eigenvalue from l_{k_1}, l_{k_2}, ..., l_{k_r} and adding the largest eigenvalue of the remaining eigenvalues, a better selection of r eigenvalues is obtained. Iterating the same procedure at most r times, the best selection of r eigenvalues is obtained as l_1, l_2, ..., l_r.

E. ML Estimation of Θ

At first, dependence of variables on Θ is explicitly expressed as follows: $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ in (23), its eigenvalues $l_1(\Theta)$, $l_2(\Theta)$, ..., $l_q(\Theta)$ and eigenvectors $\boldsymbol{v}_1(\Theta)$, $\boldsymbol{v}_2(\Theta)$, ..., $\boldsymbol{v}_q(\Theta)$, also $\sigma_1^2(\Theta)$, $\sigma_2^2(\Theta)$, .., $\sigma_q^2(\Theta)$ in (61) and $\eta(\Theta)$ determined in (62).

The maximum log-likelihood functions $L_0(\Theta)$, $L_1(\Theta)$, ..., $L_{\eta(\Theta)}(\Theta)$ defined by (63) for any fixed Θ satisfy the relationships

$$L_0(\Theta) < L_1(\Theta) < ... < L_{\eta(\Theta)}(\Theta), \quad (68)$$

which follow from the property of $L_r(\lambda)$ in (66) and the conditions $l_r(\Theta) > \sigma_r^2(\Theta)$ for $r = 1, 2, ..., \eta(\Theta)$.

If no condition is imposed on the rank of $\boldsymbol{S}$ or $\boldsymbol{P}$, $L_{\eta(\Theta)}(\Theta)$ gives the maximum value of the log-likelihood function for fixed Θ. Therefore, the ML estimation $\hat{\Theta}$ of Θ is determined as follows.

$$\hat{\Theta} = \arg\max_{\Theta} L(\Theta), \quad (69)$$

$$L(\Theta) = L_{\eta(\Theta)}(\Theta). \quad (70)$$

To find $\hat{\Theta}$, multivariate non-linear optimization techniques should be used. The ML estimations $\hat{\boldsymbol{P}}$ and $\hat{\boldsymbol{R}}$ of $\boldsymbol{P}$ and $\boldsymbol{R}$ are given as

$$\hat{\boldsymbol{P}} = \sum_{k=1}^{\eta(\hat{\Theta})} \left(l_k(\hat{\Theta}) - \sigma^2_{\eta(\hat{\Theta})}(\hat{\Theta}) \right) \boldsymbol{v}_k(\hat{\Theta}) \boldsymbol{v}_k^H(\hat{\Theta}), \tag{71}$$

$$\hat{\boldsymbol{R}} = \boldsymbol{V}_S(\hat{\Theta}) \hat{\boldsymbol{P}} \boldsymbol{V}_S^H(\hat{\Theta}) + \sigma^2_{\eta(\hat{\Theta})}(\hat{\Theta}) \boldsymbol{I}_p. \tag{72}$$

It is guaranteed that the estimated covariance matrix of signal components

$$\hat{\boldsymbol{S}} = \left(\boldsymbol{A}^H(\hat{\Theta}) \boldsymbol{A}(\hat{\Theta}) \right)^{-1/2} \hat{\boldsymbol{P}} \left(\boldsymbol{A}^H(\hat{\Theta}) \boldsymbol{A}(\hat{\Theta}) \right)^{-H/2} \tag{73}$$

is non-negative definite.

When the rank of $\boldsymbol{S}$ or $\boldsymbol{P}$ is restricted to be a certain fixed r in { 1, 2, ..., q }, first we define the solution space Ω_r of Θ in which the ML function of the model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = r$ can be defined. That is

$$\Omega_r = \{\Theta \mid r \leq \eta(\Theta)\} = \{\Theta \mid l_r(\Theta) > \sigma_r^2(\Theta)\}. \tag{74}$$

The ML estimation $\hat{\Theta}_r$ of Θ for the model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = r$ should be searched as

$$\hat{\Theta}_r = \arg\max_{\Theta \in \Omega_r} L_r(\Theta). \tag{75}$$

However, because of the condition $l_r(\Theta) > \sigma_r^2(\Theta)$ in (74), the marginal set Γ_r = { $\Theta \mid l_r(\Theta) = \sigma_r^2(\Theta)$ and $\Theta \in \overline{\Omega}_r$ } is not included in Ω_r, where $\overline{\Omega}_r$ is the closure of Ω_r. Define

$$\overline{\Theta}_r = \arg\max_{\Theta \in \overline{\Omega}_r} L_r(\Theta). \tag{76}$$

In the case that $\overline{\Theta}_r$ belongs to Γ_r, $\hat{\Theta}_r$ does not exist. It happens occasionally. If $\hat{\Theta}_r$ exists, then The ML estimations $\hat{\boldsymbol{P}}_r$ and $\hat{\boldsymbol{R}}_r$ of $\boldsymbol{P}$ and $\boldsymbol{R}$ for the model of rank$\{\boldsymbol{S}\}$ = rank$\{\boldsymbol{P}\} = r$ are given as

$$\hat{\boldsymbol{P}}_r = \sum_{k=1}^{r} \left(l_k(\hat{\Theta}_r) - \sigma_r^2(\hat{\Theta}_r) \right) \boldsymbol{v}_k(\hat{\Theta}_r) \boldsymbol{v}_k^H(\hat{\Theta}_r) \tag{77}$$

$$\hat{\boldsymbol{R}}_r = \boldsymbol{V}_S(\hat{\Theta}_r) \hat{\boldsymbol{P}}_r \boldsymbol{V}_S^H(\hat{\Theta}_r) + \sigma_r^2(\hat{\Theta}_r) \boldsymbol{I}_p \tag{78}$$

and it is guaranteed that the following estimated covariance matrix of signal components is non-negative definite.

$$\hat{\boldsymbol{S}}_r = \left(\boldsymbol{A}^H(\hat{\Theta}_r) \boldsymbol{A}(\hat{\Theta}_r) \right)^{-1/2} \hat{\boldsymbol{P}}_r \left(\boldsymbol{A}^H(\hat{\Theta}_r) \boldsymbol{A}(\hat{\Theta}_r) \right)^{-H/2} \tag{79}$$

IV. Simulations

In this section, comparisons of the proposed formulation of UML and the previous formulation [6], [7], [8], [9], In the incoherent case, the previous formulation is written as follows.

$$\hat{\Theta}_C = \arg\max_{\Theta} L_C(\Theta) \tag{80}$$

where

$$L_C(\Theta) = -N \ln L_S(\Theta) L_N(\Theta) \tag{81}$$

$$L_S(\Theta) = \det\{\tilde{\boldsymbol{R}}_{SS}(\Theta)\} \quad L_N(\Theta) = \det\{\boldsymbol{R}_{NN}(\Theta)\} \tag{82}$$

$$\boldsymbol{R}_{NN}(\Theta) = \hat{\sigma}^2(\Theta) \boldsymbol{I}_{p-q} \tag{83}$$

$$\hat{\sigma}^2(\Theta) = \frac{1}{p-q} \left\{ \mathrm{tr}\{\tilde{\boldsymbol{R}}\} - \mathrm{tr}\{\tilde{\boldsymbol{R}}_{SS}(\Theta)\} \right\} \tag{84}$$

The estimations of $\boldsymbol{P}$ and $\boldsymbol{R}$ are given as

$$\hat{\boldsymbol{P}}_C = \tilde{\boldsymbol{R}}_{SS}(\hat{\Theta}_C) - \sigma^2(\hat{\Theta}_C) \boldsymbol{I}_q \tag{85}$$

$$\hat{\boldsymbol{R}}_C = \boldsymbol{V}_S(\hat{\Theta}_C) \hat{\boldsymbol{P}}_C \boldsymbol{V}_S^H(\hat{\Theta}_C) + \sigma^2(\hat{\Theta}_C) \boldsymbol{I}_p. \tag{86}$$

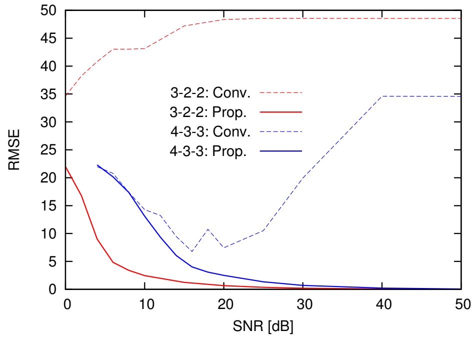

Fig. 1. Comparisons of RMSE for proposed and conventional formulations of UML DOA estimation. A uniform linear array of omni-directional sensors with the sensor space of half wavelength is assumed. The three figures $p-q-r$ in the graph represent the number of sensors, the number of impinging signals and the number of independent signals in the impinging signals. The directions of arrival are $0°$ and $8°$ for $q = 2$ and $0°$, $8°$ and $16°$ for $q = 3$. The number of sample vectors (snapshots) is 100. Optimal DOA's for both formulations are searched by Alternating Maximization algorithm with a sequence of one-dimensional global search.

In Fig. 1, the root mean squares errors (RMSE) of the proposed estimation of DOA in (70) and the conventional estimation in (80) are depicted. The scenario of the simulation is shown in the figure caption.

From Fig. 1, it is found that the conventional formulation fails to find DOA. The maximization of $L_C(\Theta)$ is associated with each minimization of $L_S(\Theta)$ or $L_N(\Theta)$. The local solutions of $\hat{\Theta}_C$ associated with the local minimum of $L_S(\Theta)$ are inadequate DOA, since $\hat{\boldsymbol{P}}_C$ has negative eigenvalues. When one of such local solutions becomes global solution, the estimation of DOA fails. This is the reason of the failure of the conventional formulation. While such problem never happen in the proposed formulation.

V. Conclusions

The estimation of the covariance matrix of signal components in the previous formulation of UML estimation becomes non-negative definite in the condition of high SNR or a large number of samples but is not guaranteed to be non-negative definite although it must be non-negative definite.

This paper present an exact formulation of the UML estimation in which the estimation of the covariance matrix of the signal components is guaranteed to be non-negative definite.

REFERENCES

[1] Y. Bresler and A. Macovski, “Exact maximum likelihood parameter estimation of superimposed exponential signals in noise,” IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, No. 5, pp. 1081-1089, Oct. 1986.

[2] I. Ziskind and M. Wax, “Maximum likelihood localization of multiple sources by alternating projection,” IEEE Trans. Acoust., Speech, Signal Processing, vol. 36, No. 10, pp. 1553-1560, Oct. 1988.

[3] P. Stoica and K. C. Sharman, “Maximum likelihood methods for direction-of-arrival estimation,” IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. 38. No. 7. pp. 1132–1143, July 1990.

[4] M. Suzuki, H. Sanada, N. Nagai: “Irreducible Form for AP Algorithm for Detecting the Number of Coherent Signals Based on the MDL Principle,” in Proceedings of ICASSP2001, vol. V, pp.2865-2868, May 2001.

[5] M. Suzuki: “An Efficient Algorithm for Bearing Estimation Based on Newton’s Method Using a Uniform Linear Array,” in Proceedings of ISCIT2001, Nov. 2001.

[6] P. Stoica and A. Nehoria, ”Performance study of conditional and unconditional direction-of-arrival estimation” IEEE Trans. Acoust., Speech. and Signal Processing, Vol. 38, No. 10, pp. 1783-1795, Oct. 1990.

[7] P. Stoica, Fellow, B. Ottersten, M. Viberg, R. L. Moses, ”Maximum Likelihood Array Processing for Stochastic Coherent Sources,” IEEE Trans. Signal Processing, Vol. 44, No. 1, pp. 96–105, Jan 1996.

[8] M. Wax, ”Detection and Localization of Multiple Sources via the Stochastic Signals Model” IEEE Trans. Signal Processing, Vol. 39, No. 11, pp. 2450-2456, Nov. 1991.

[9] Harry L. van Trees, “Optimum array processing, Part IV of Detection, estimation and modulation theory,” John Wiley and Sons, Inc., New York, 2002.

[10] R. Roy and T. Kailath, “ESPRIT-Estimation of Signal Parameters Via Rotational Invariance Techniques,” IEEE Trans. Acoust., Speech. Signal Processing. vol. 37, no 7, 984–995, July 1989.

[11] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propagat.,* vol. AP-34, No. 3, pp. 276-280, Mar. 1986.

[12] R. T. Williams, S. Prasad, A. K. Mahalanabis and L. H. Sibul, “An improved spatial smoothing technique for bearing estimation in a multipath environment,” IEEE Trans. Acoust., Speech, Signal Processing, vol. 36, No. 4, pp. 425-432, Apr. 1988.

[13] M. Viberg, B. Ottersten, and T.Kailath, “Detection and estimation in sensor array using weighted subspace fitting,” IEEE Trans. Acoust., Speech, Signal Processing, vol. 39, No. 11, pp. 2436-2449, Nov. 1991.

[14] B. M. Radich and K. M. Buckley, “Single-snapshot DOA estimation and source number detection,” IEEE Signal Processing Letters, vol. 4, no. 4, pp. 109–111, April 1997.

[15] M. Wax and I. Ziskind, “On unique localization of multiple sources by passive sensor arrays,” IEEE Trans. Acoust., Speech, Signal Processing, vol. 37, No. 7, pp. 996-1000, Jul. 1989.