Exact Formulation and Efficient Algorithms for Stochastic

ML Estimation of DOA

Information and Communication Engineering

Kitami Institute of Technology

Haihua Chen

March 2011

TABLE OF CONTENTS

Chapter 1

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. In mobile, radio and space communications, the DOA estimation plays an important role for MIMO (Multiple-Input Multiple-Output) system, radio surveillance, and so on.

A number of super-resolution techniques have been introduced, such as Maximum likelihood (ML) method [1, 2, 3, 7, 9, 8, 10, 4, 5, 6, 11, 12, 13, 14] MUSIC [18, 20], ESPRIT [17], Weighted Subspace Fitting (WSF) [15, 14] and the Bayesian method [21].

The ML, WSF and Bayesian techniques have properties superior to other methods since they can handle coherent signals without any preprocessing, such as the spatial smoothing [20]. They can also handle small number of snapshots, although the Bayesian method [21] is formulated only for a single snapshot.

ML estimators are useful not only for DOA finding but also the enumeration of the number of signals based on the Minimum Description Length (MDL) model selection. There have been proposed two kinds of ML estimators for DOA finding. That is, Conditional or Deterministic ML (DML) [1, 2, 3, 7, 8, 4, 5, 6]. and Unconditional or Stochastic ML (SML) [7, 9, 8, 10, 6, 11, 12, 13, 14]. Difference between them lies in their models of signals. The SML shows better estimation for coherent signals than the DML if adequate local solutions of DOA are searched [12, 13].

The conventional SML estimation [7, 9, 8] are formulated without an important con-

dition that the covariance matrix of signal components must be non-negative definite. A likelihood function can not be evaluated exactly for all sets of directions in the context of the conventional SML. This involves theoretical difficulty in applying the conventional SML to others, such as the MDL model selection.

Bresler [11] has proposed an ML estimator of the signal covariance matrix for fixed directions using the model of the SML estimation. It guarantees the non-negative definiteness of the the estimated signal covariance matrix.

Stoica et al. [14] and Harry [8] pointed out the fact that the conventional SML does not guarantee the non-negative definiteness of the estimated signal covariance matrix. To overcome this defect, Stoica et al. have derived an ML estimator based on the model of the SML under the condition that a large number of samples are available [14]. The asymptotic approximation of the SML leads to the same criterion as that of the WSF.

Wang et al. [6] have proposed a technique to solve a global maximization of "compressed likelihood function" based on Monte Carlo importance sampling. It deals with stochastic signals. Although the non-negative definiteness of the estimated covariance matrix of signals is guaranteed, the model is different from that used in the SML and the variance of noises is estimated separately from DOA parameters in [6], while jointly with DOA parameters in the SML.

In this paper, first it is revealed that the conventional SML has three problems due to the lack of the condition. 1) Solutions in the noise-free case are not unique. 2) Global solution in the noisy case becomes ambiguous occasionally. 3) There exist situations that any local solution does not satisfy the condition of the non-negative definiteness.

Next we propose an exact formulation of the SML estimation of DOA to evaluate a likelihood function exactly for any possible set of directions. The proposed formulation

can be utilized without any theoretical difficulty. The three problems of the conventional SML are solved by the proposed exact SML estimation. Although the exact SML is formulated separately from [11], the estimation of the signal covariance matrix for fixed directions is equivalent to that in [11].

A performance study of exact SML is also done with comparison to some other super-resolution techniques such as: MUSIC (Spectral-MUSIC, Root-MUSIC), WSF and DML. We reveal the following properties of exact SML. 1) The exact SML can always success in DOA finding when a unique solution exists. 2) The WSF or the asymptotic approximation of the SML [14] is sensitive to the rank of the signal covariance matrix, which is required to be known, while the exact SML does not. 3) The exact SML shows the best Root-Mean-Square-Error (RMSE) performance in the threshold region (when SNR is low or snapshots is small). When the observation condition is good (high SNR or large samples), WSF shows the best RMSE performance, and exact SML's is a litter worse than that of WSF. 4) The computation cost of exact SML is the highest when the most conventional technique, Alternating Minimization (AM) method is used.

Since the computational cost of exact SML is very high, it is difficult to apply this technique in real applications. To reduce the computational cost possibly with less loss of resolution, we propose the following three efficient algorithms.

First, we use the solution of DML as the initial value and apply a local search method to find the optimal or suboptimal solution of SML. The reason is that the solution of DML is unique. More-over, the DML estimation asymptotically achieves the exact SML solution in large-samples for incoherent signals. But we should note that this method sometimes can not find the optimal solution of SML especially in the threshold region. In simulations, we have confirmed that the solution of this local search method is identical to that of the exact SML when DOA is resolvable. Since local search method

is used, computational cost can be greatly reduced.

Then, Based on the local search method, an efficient version of the Alternating Minimization (AM) algorithm called EAM is proposed. This algorithm can reduce the computational cost greatly without any impact of resolution. The main idea consists in dividing the SML criterion into two components. One is independent of a single variable parameter, and is held fixed, when the parameter varies. Another part is depending on the single variable parameter and changes with the varying of the parameter. So that we only need to calculate the variable part in each alternating process.

Furthermore, when the sensor array is assumed to be uniform linear, an irreducible form of EAM is derived using polynomial forms. We call it IAM. This form can avoid the numerical instability in calculation of SML criterion and the computational cost is further greatly reduced.

Simulations are also shown in each stage of our research to demonstrate the effectiveness and efficiency of our proposed criterions and algorithms.

Chapter 2

Problem Formulation

2.1   Introduction

In this chapter, we give the problem formulation of DOA finding at first, and then we introduce the Deterministic Maximum-Likelihood (DML) and Stochastic ML (SML) estimations of DOA finding. We also introduce the constraints for the SML estimation, and the conventional SML estimation which can not meet one of the constraints. Finally, the most conventional technique, Alternating Minimization (AM) algorithm, for the SML estimation is also described.

2.2   DOA Finding

Consider an array composed of p sensors with arbitrary locations and arbitrary directional characteristics, and assume that q narrow-band sources, 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

$$x(t) = \sum_{i=1}^{q} a(\theta_k) s_k(t) + n(t), \tag{2.1}$$

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

θ, which is represented as

$$\boldsymbol{a}(\theta) = [a_1(\theta)e^{-j\omega_0\tau_1(\theta)}, ..., a_p(\theta)e^{-j\omega_0\tau_p(\theta)}]^T \tag{2.2}$$

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, (2.1) can be rewritten as

$$\boldsymbol{x}(t) = \boldsymbol{A}(\Theta)\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{2.3}$$

$$\boldsymbol{A}(\Theta) = [\, \boldsymbol{a}(\theta_1)\ \boldsymbol{a}(\theta_2)\ \cdots\ \boldsymbol{a}(\theta_q)\,], \tag{2.4}$$

$$\boldsymbol{s}(t) = [\, s_1(t)\ s_2(t)\ \cdots\ s_q(t)\,]^T, \tag{2.5}$$

$$\Theta = \{\, \theta_1\ \theta_2\ \cdots\ \theta_q\,\}. \tag{2.6}$$

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

$$\boldsymbol{X} = [\, \boldsymbol{x}(t_1)\ \boldsymbol{x}(t_2)\ \cdots\ \boldsymbol{x}(t_M)\,]. \tag{2.7}$$

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

$$\hat{\Theta} = \{\, \hat{\theta}_1\ \hat{\theta}_2\ \cdots\ \hat{\theta}_q\,\}. \tag{2.8}$$

of $\theta_1, \theta_2, ..., \theta_q$.

The model (2.1) or (2.3) can be rewritten as follows

$$\boldsymbol{x}(t) = \boldsymbol{V}_S(\Theta)\boldsymbol{x}_S(t) + \boldsymbol{V}_N(\Theta)\boldsymbol{x}_N(t) \tag{2.9}$$

where $\boldsymbol{V}_S(\Theta)$ is a $p \times q$ matrix of which columns form an orthonormal system of the signal subspace, that is spanned by $\{\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), ..., \boldsymbol{a}(\theta_q)\}$, $\boldsymbol{V}_N(\Theta)$ is a $p \times (p-q)$

matrix of which columns form an orthonormal system of the noise subspace, which is the orthogonal complement of the signal subspace. $x_S(t)$ is the signal component of $x(t)$ in the signal subspace, and $x_N(t)$ is the noise component of $x(t)$ in the noise subspace.

2.3   ML Estimation

In this section, brief descriptions of the DML estimation[2], the SML estimation, the conventional SML estimation[9] and the conventional solving technique, Alternating Minimization (AM) method are shown.

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

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

A2) $n(t_i)$ are statistically independent samples from a complex Gaussian random vector with zero mean and the covariance matrix $\sigma^2 I_p$, where σ^2 is an unknown parameter, I_p is a $p \times p$ identity matrix.

A3) $s(t_i)$ satisfy the condition

$$\text{rank}[\, s(t_1)\, s(t_2)\, \cdots\, s(t_N)\,] = r(\leq q). \tag{2.10}$$

In the case of $r < q$, the signals are coherent or fully correlated which happens, e.g., in specular multi-path propagation.

A4) p, q is known. r is unknown.

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 $q \leq 2rp/(2r+1)$ [22].

2.3.1   Deterministic ML (DML)

The DML estimator is derived by imposing the following assumption on the signals in addition to A1)-A5),

A6D) $s(t_i)$ are unknown deterministic parameters.

The DML criterion according to [2] is given by

$$\hat{\Theta}_{\text{DML}} = \arg \min_{\Theta} L_{\text{DML}}(\Theta), \tag{2.11}$$

$$L_{\text{DML}}(\Theta) = \text{tr}\{\tilde{\boldsymbol{R}}_{NN}(\Theta)\}, \tag{2.12}$$

where

$$\tilde{\boldsymbol{R}}_{NN}(\Theta) = \boldsymbol{V}_N^H(\Theta)\tilde{\boldsymbol{R}}\boldsymbol{V}_N(\Theta), \tag{2.13}$$

$$\tilde{\boldsymbol{R}} = \frac{1}{M}\boldsymbol{X}\boldsymbol{X}^H, \tag{2.14}$$

where H denotes the Hermitian conjugate.

$\boldsymbol{V}_N(\Theta)$ is defined as above. $\tilde{\boldsymbol{R}}$ is the sample covariance matrix of sampled data. The $(p-q) \times (p-q)$ matrix, $\tilde{\boldsymbol{R}}_{NN}(\Theta)$, corresponds to the sample covariance matrix for the components of $\boldsymbol{x}(t)$ in the noise subspace.

2.3.2   Stochastic ML (SML)

The SML [7, 9, 8] estimation of DOA is formulated under the following assumption.

A6) $s(t_i)$ are statistically independent samples from a complex Gaussian random vector with zero mean and a certain covariance matrix S with rank$\{S\} = r$.

According to the assumptions A1) to A6), $x(t)$ is modeled as a p-dimensional complex Gaussian random vector with zero mean and the covariance matrix R,

$$R = E\{x(t)x^H(t)\} = A(\Theta)S A^H(\Theta) + \sigma^2 I_p \tag{2.15}$$

where S is the signal covariance matrix and it must be non-negative definite.

The joint density function of the sampled data X is given by [8]

$$f(X) = \prod_{i=1}^{M} \frac{1}{\det[\pi R]} \times \exp\{x(t_i)^H R^{-1} x(t_i)\} \tag{2.16}$$

where $\det[\cdot]$ represents the determinant. Taking the log function of (2.16) and ignoring the constant term, the log-likelihood function of the model R in (2.15) with respect to the unknown parameters Θ, S and σ^2 is

$$L(\Theta, S, \sigma^2) = -M\left(\ln\det\{R\} + \frac{1}{N}\sum_{i=1}^{N} x(t_i)^H R^{-1} x(t_i)\right) \tag{2.17}$$

The second term of (2.17) can be written into a trace form of matrix, then (2.17) becomes,

$$L(\Theta, S, \sigma^2) = -M\left(\ln\det\{R\} + \text{tr}\{R^{-1}\tilde{R}\}\right) \tag{2.18}$$

where the sample covariance matrix $\tilde{R}$ is given as above.

The SML estimation of DOA is to be stated as the problem to find Θ which maximizes (2.18) under the conditions

C1) $\Theta = \{ \theta_1 \ \theta_2 \ \cdots \ \theta_q \} \in U$. U is a set of all possible Θ determined by the array configuration.

C2) S is non-negative definite.

C3) σ^2 is a non-negative real number.

The Model of R in (2.15) can be rewritten as

$$R = V_S(\Theta) P V_S^H(\Theta) + \sigma^2 I \tag{2.19}$$

$$= G(\Theta) \begin{bmatrix} R_{SS} & 0 \\ 0 & \sigma^2 I \end{bmatrix} G^H(\Theta), \tag{2.20}$$

$$G(\Theta) = [V_S(\Theta) \ V_N(\Theta)], \tag{2.21}$$

where $V_S(\Theta)$ and $V_N(\Theta)$ are defined as above; $P = T(\Theta)S T^H(\Theta)$ and $R_{SS} = P + \sigma^2 I_q$ where $T(\Theta)$ is a $q \times q$ non-singular matrix which satisfies $A(\Theta) = V_S(\Theta)T(\Theta)$. The matrix P seems to be a function of Θ. However, since it is unknown, it can be dealt with a matrix independent of Θ.

Then the unknown parameters of the likelihood function of model (2.19) and (2.20) becomes Θ, P or R_{SS} and σ^2. Correspondingly P or $(R_{SS} - \sigma^2 I)$ should also be non-negative definite.

2.3.3 Conventional SML

The estimation of conventional SML according to [8] is like follows.

(1) Given Θ.

(2) Evaluate $A(\Theta)$.

(3) Obtain an orthogonal system $V_N(\Theta) = A(\Theta)T(\Theta)$.

(4) Calculate $\tilde{R}_{SS}(\Theta) = V_N^H(\Theta)\hat{R}V_N(\Theta)$.

(5) Obtain $\tilde{\sigma}^2(\Theta) = \frac{1}{p-q}\left\{\mathrm{tr}\{\tilde{R}\} - \mathrm{tr}\{\tilde{R}_{SS}(\Theta)\}\right\}$.

(6) Evaluate the likelihood function

$$L_c(\Theta) = -M \ln\left(\det\{\tilde{R}_{SS}(\Theta)\} \cdot (\tilde{\sigma}^2(\Theta))^{p-q}\right).$$

The conventional SML estimation is the problem to obtain

$$\hat{\Theta}_c = \arg\max_{\Theta} L_c(\Theta). \tag{2.22}$$

The other parameters of σ^2, R_{SS}, P and S are obtained as

$$\hat{\sigma}_c^2 = \tilde{\sigma}^2(\hat{\Theta}_c), \quad \hat{R}_{SS_c} = \tilde{R}_{SS}(\hat{\Theta}_c), \tag{2.23}$$

$$\hat{P}_c = \hat{R}_{SS_c} - I\hat{\sigma}_c^2, \quad \hat{S}_c = T^{-1}(\hat{\Theta}_c)\hat{P}_cT^{-H}(\hat{\Theta}_c). \tag{2.24}$$

In the estimation above, the condition C2) is ignored. As a result, $\hat{P}_c$ equivalent to $\hat{S}_c$ equivalent to $\left(\hat{R}_{SS_c} - \hat{\sigma}_c^2 I\right)$ are all not guaranteed to be non-negative definite.

2.3.4 AM Algorithm for the SML Estimation

Many techniques for solving the DML estimation have been proposed, such as Alternating Projection (AP) algorithm [2] and its irreducible version [23], MODE [3], and gradient methods [15, 24]. According to numerous simulations for uniform linear arrays, all of these techniques show good properties of convergence and provide global solutions.

While for solving the SML estimation, only the Alternating Minimization (AM) algorithm has been used [9]. The AM method is a popular iterative technique for solving

a nonlinear multivariate minimization problem with a multi-modal criterion. The AM algorithm can be applied to the SML estimation of DOA in the following manner.

Let $L_k(\hat{\Theta}^{(k)})$ be a cost function of $L_c(\Theta)$ for which the signal number is assumed to be k instead of q, where $\hat{\Theta}^{(k)} = \{\hat{\theta}_1, \ \hat{\theta}_2, \cdots, \hat{\theta}_k\}$.

[Initialization Phase]

First assuming a single signal, k=1, find $\hat{\theta}_1$ minimizing $L_1(\hat{\Theta}^{(1)})$ by one-dimensional global search with respect to $\hat{\theta}_1$. Next, assuming two signals, k=2, and fixing $\hat{\theta}_1$ at the value obtained for a single signal, find $\hat{\theta}_2$ minimizing $L_2(\hat{\Theta}^{(2)})$ by one-dimensional global search with respect to $\hat{\theta}_2$. Continue in this fashion until all the initial values for $\hat{\theta}_k, k = 1, 2, ..., q$ are computed.

[Convergence Phase]

Repeat the following updating process until all parameters are converged. At each updating process, let one parameter, say $\hat{\theta}_k$, be variable and let all other parameters be held fixed. Find $\hat{\theta}_k$ minimizing the criterion $L_q(\hat{\Theta}^{(q)})$ by one-dimensional global search with respect to $\hat{\theta}_k$. Change the index k of the parameter to be updated into $(k \mod q) + 1$.

Although a global minimum is not guaranteed in the AM algorithm, global solutions can be obtained in most cases because of one-dimensional global searches performed in each update process.

2.4 Conclusions

In this chapter, we have introduced what are DML and SML estimations. We also introduced the conventional SML estimation which is formulated by ignoring an important condition of SML estimation. The conventional solving technique for SML estimation, AM algorithm, is also described. So in next chapter we will discuss what is the problem of the conventional SML estimation.

Chapter 3

Problems in Conventional SML Estimation

3.1 Introduction

In this chapter, we discuss the problems in conventional SML estimation. We show that because the condition C2) is ignored in the conventional SML, the following three problems occur.

1. Solutions in the noise-free case are not unique.

2. Global solution in the noisy case becomes ambiguous occasionally.

3. There exist situations that any local solution does not satisfy the condition C2).

3.2 Problems in Conventional SML Estimation

In this section, first we have an important definition of solution space. Define

$$\mathcal{S}_{\tilde{R}} = \{\Theta \mid \Theta \in U \text{ and } (\tilde{R}_{SS}(\Theta) - \tilde{\sigma}^2(\Theta)I) \geq 0\} \tag{3.1}$$

in which the local solution becomes the exact SML solution. We call this area the ML solution space for the conventional SML. The constraint condition C2) is violated outside $\mathcal{S}_{\tilde{R}}$. Therefore a solution outside $\mathcal{S}_{\tilde{R}}$ can be considered to be inadequate in the sense of ML.

3.2.1 The Loss of the Uniqueness in the Noise Free Case

First, we deal with the noise-free case, i.e., $\sigma^2 = 0$. Because of the assumption A5), a unique solution of DOA exists. As proved in appendix D, the DML estimator have a unique solution in the noise-free case. Therefore, the DML estimator provides the true directions of arrival Θ.

The SML estimation is the problem to find $\hat{\Theta}$ which maximize:

$$L_c(\Theta) = -M \ln \left(\det\{\tilde{\boldsymbol{R}}_{SS}(\Theta)\} \cdot (\tilde{\sigma}^2(\Theta))^{p-q} \right). \tag{3.2}$$

Define

$$L_c(\Theta) = L_S(\Theta) L_N(\Theta), \tag{3.3}$$

where,

$$L_S(\Theta) = \det\{\tilde{\boldsymbol{R}}_{SS}(\Theta)\}, \tag{3.4}$$

$$L_N(\Theta) = (\tilde{\sigma}^2(\Theta))^{p-q}. \tag{3.5}$$

In the noise-free case, $L_c(\Theta) = 0$ holds when $L_S(\Theta) = 0$ or $L_N(\Theta) = 0$. Since $L_N(\Theta) = 0$ is equivalent to $L_{\text{DML}}(\Theta) = 0$, a set of the true directions Θ is one of solutions.

Next, we consider the question if there exists $\hat{\Theta}$ such that $L_S(\hat{\Theta}) = 0$. Let $\boldsymbol{P}^{1/2}$ be a square-root matrix of a non-negative definite Hermitian matrix $\boldsymbol{P}$, such that $\boldsymbol{P}^{1/2} \boldsymbol{P}^{H/2} = \boldsymbol{P}$ where $\boldsymbol{P}^{H/2} = (\boldsymbol{P}^{1/2})^H$ and define the notations, $\boldsymbol{P}^{-1/2} = (\boldsymbol{P}^{1/2})^{-1}$ and $\boldsymbol{P}^{-H/2} = (\boldsymbol{P}^{H/2})^{-1}$. Then we have

$$\boldsymbol{V}_S(\hat{\Theta}) = \boldsymbol{A}(\hat{\Theta})\{\boldsymbol{A}^H(\hat{\Theta})\boldsymbol{A}(\hat{\Theta})\}^{-H/2}, \tag{3.6}$$

and

$$\det \boldsymbol{R}_{SS}(\hat{\Theta}) = \frac{|\det\{\boldsymbol{A}^H(\hat{\Theta})\boldsymbol{A}(\Theta)\}|^2 \det \boldsymbol{\Sigma}}{\det\{\boldsymbol{A}^H(\hat{\Theta})\boldsymbol{A}(\hat{\Theta})\}}, \tag{3.7}$$

where

$$\boldsymbol{\Sigma} = \frac{1}{M}\boldsymbol{S}\boldsymbol{S}^H. \tag{3.8}$$

$L_S(\hat{\Theta}) = 0$ holds when

$$\det \boldsymbol{\Sigma} = 0 \quad \text{or} \quad \det\{\boldsymbol{A}^H(\hat{\Theta})\boldsymbol{A}(\Theta)\} = 0 \tag{3.9}$$

holds. In the former case, i.e., the case of $\eta < q$, $L_c(\hat{\Theta}) \equiv 0$ holds for any $\hat{\Theta}$. This indicates that the SML estimation fails completely. If σ^2 is not zero but very small, then the same failure does not happen. Instead, the later case, i.e., $\det\{\boldsymbol{A}^H(\hat{\Theta})\boldsymbol{A}(\Theta)\} = 0$ becomes a dominant factor which makes $L_S(\hat{\Theta})$ or $L_c(\hat{\Theta})$ small.

The remaining question is that there exists $\hat{\Theta}$ such that $\det\{\boldsymbol{A}^H(\hat{\Theta})\boldsymbol{A}(\Theta)\} = 0$. From here, we consider the case of uniform linear arrays. The steering vector for a uniform linear array composed of omnidirectional sensors is represented as

$$\boldsymbol{a}(\theta) = [\,1\ e^{-j\phi(\theta)}\ \cdots\ e^{-j(p-1)\phi(\theta)}]^T, \tag{3.10}$$
$$\phi(\theta) = \frac{2\pi\Delta}{\lambda}\sin\theta. \tag{3.11}$$

where λ is the wavelength of signals impinging on the array, and Δ is the sensor spacing between adjacent two sensors. As a necessary condition that a unique direction θ is determined by the phase parameter ϕ, $\Delta \leq \lambda/2$ is imposed on the array configuration.

Numerical solutions of (3.9) are shown in Fig. 4.1, where the sensor spacing $\Delta = \lambda/2$. In Fig. 4.1.(c) and (d), $\hat{\theta}_3$ is held fixed to -60°. Other parameters of simulations are shown in figure caption. For any other fixed $\hat{\theta}_3$, the similar results to Fig. 4.1.(c) and (d) are obtained.

16

(a) $p = 3$, $q = 2$, $\theta_1 = 0°$ and $\theta_2 = 8°$.

(b) $p = 4$, $q = 2$, $\theta_1 = 0°$ and $\theta_2 = 8°$.

(c) $p = 4$, $q = 3$, $\theta_1 = 0°$, $\theta_2 = 8°$, $\theta_3 = 16°$, and $\hat{\theta}_3 = -60°$.

(d) $p = 6$, $q = 3$, $\theta_1 = 0°$, $\theta_2 = 8°$, $\theta_3 = 16°$, and $\hat{\theta}_3 = -60°$.
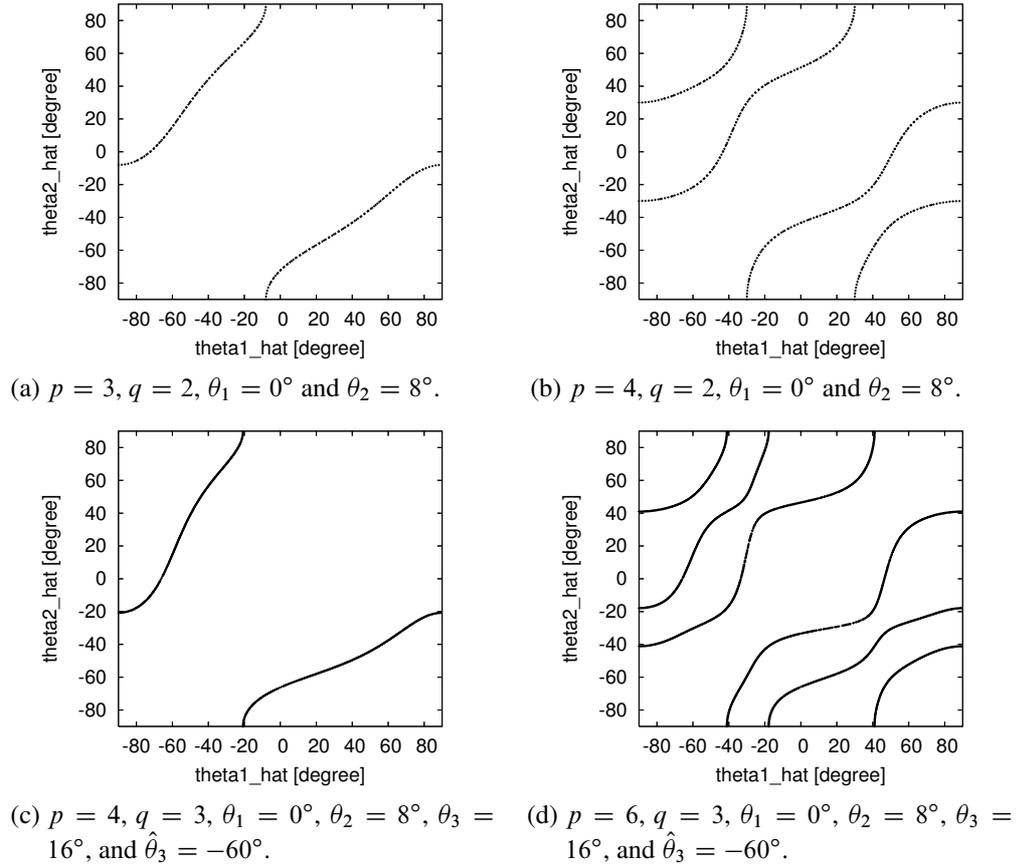
Figure 4.1: Orbits of $(\hat{\theta}_1, \hat{\theta}_2)$ for which $L_S(\hat{\Theta}) = 0$.

From Fig. 4.1, it can be confirmed that there exist infinite number of $\hat{\Theta}$ such that $\det\{A^H(\hat{\Theta})A(\Theta)\} = 0$ for uniform linear arrays. We have to note that these are solutions of the SML estimation but extremely far from the true vales Θ. From the discussion above, it can be stated that the SML estimation is not unique in the noise-free case.

3.2.2 Ambiguity of Global Solutions in Noisy Case

In this and the next subsection, we discuss these problems using examples shown in Fig. 4.2. The scenarios of simulations are described in captions. In Fig. 4.2, a cross represents a local solution. A shaded area represents the solution space of SML as we

(a) The point A is the local solution closest to the true DOA while B is the global solution where SNR = 30 dB.

(b) There is no local solution in the ML solution space where SNR = 0 dB.

(c) There is no local solution in the ML solution space where SNR = 0 dB.

Figure 4.2: Samples in the conventional SML estimation with uniform linear arrays of omni-directional sensors. $p = 3, q = 2, r = 2, N = 10$ in (a) and (b), $p = 4, q = 2, r = 1, N = 100$ in (c). The true DOA are $0°$ and $8°$.

defined above.

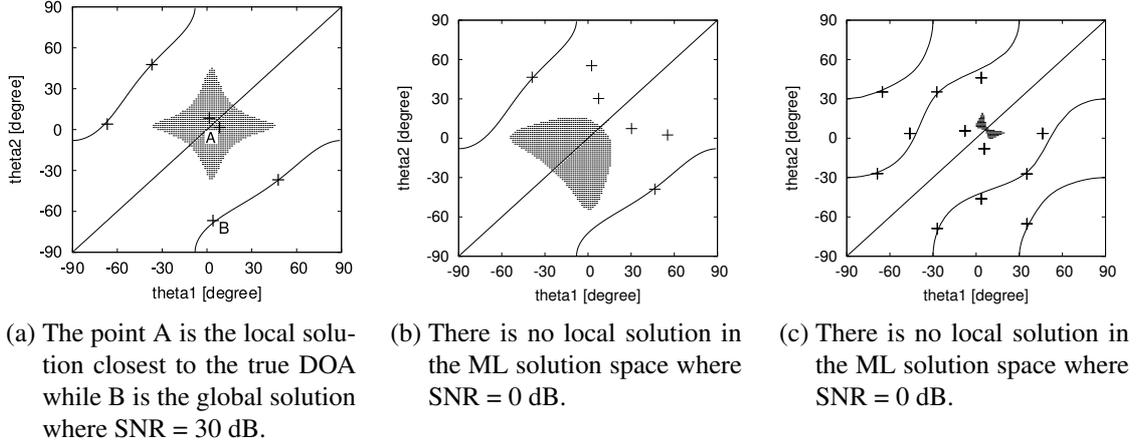In the case that noises exist, global search for the conventional SML usually results in finding a global solution of conventional SML. Fig. 4.2 (a) is a failure sample in which the adequate solution of SML is not found.

In Fig. 4.2 (a), the point A is the adequate solution of SML because it is close to the true DOA (0 and 8 degree) and it locates in the solution space. While, the point B is the inadequate solution of SML because it is far from the true DOA and it locates out of the solution space. However, the global solution is point B. According to 30 independent trials with different noise samples in the same scenario as in Fig. 4.2 (a), global solutions appear around B at 15 times and around A at other 15 times.

Although a unique global solution is found with probability one due to the existence of noises, it is not always the closest local solution to the true DOA or is occasionally inadequate. Therefore, the global solution of the conventional SML is ambiguity in this

case. As a result, it fails in DOA finding in this case.

The inadequate global solutions appear along the solid curves as shown in Fig. 4.1 (a). The ambiguity of global solutions comes from the loss of the uniqueness of the conventional SML solutions in the noise-free case. Furthermore, we have found empirically that this kind of problems occur when signals are incoherent and r is close to q.

3.2.3 No Local Solutions in the ML Solution Space

Fig. 4.2 (b) and (c) show the cases that all local solutions do not satisfy the condition of C2). In other words, no local solution is found in the ML solution space of the shaded area. Fig. 4.2 (b) is the case of incoherent signals with low SNR and Fig. 4.2 (c) is the case of coherent signals with comparatively high snapshots.

The conventional SML provides a formulation of the exact SML as long as Θ belongs to $\mathcal{S}_{\tilde{R}}$. A question arises that whether it is sufficient to find Θ in $\mathcal{S}_{\tilde{R}}$ using the conventional SML criterion in $L_c(\Theta)$. Unfortunately the answer is negative if no local solution exists in $\mathcal{S}_{\tilde{R}}$. This remedy for the conventional SML provides only a kind of suboptimal solution since there exists an optimal or exact solution outside $\mathcal{S}_{\tilde{R}}$ as shown in the next section.

Outside $\mathcal{S}_{\tilde{R}}$, there does not exist any SML solution of which the model has rank$\{S\} = q$. In order to derive an exact formulation of the SML estimation, we need to introduce a stochastic model with rank$\{S\} = r < q$.

3.3 Conclusions

In this chapter we have shown the three problems of conventional SML estimation. The reason is that the conventional SML is formulated without considering the condition C2).

Due to the problems we discussed above, the conventional SML fails in DOA finding occasionally. More importantly, because the adequate solution of SML is not found, the ML likelihood function can not be evaluated correctly. This involves theoretical difficulty in applying the conventional SML to others, such as the MDL model selection. Therefore, the exact formulation of SML is needed.

Chapter 4

Formulation of Exact SML

4.1 Introduction

In this chapter, we give an exact formulation of SML estimation. In the exact formulation the condition C2) which is ignored in conventional SML estimation is considered. Then a performance study of the exact SML is also done with comparison to some other techniques.

4.2 Formulation of Exact SML

The main idea of our proposed exact formulation is how to guarantee the condition C2) while maximizing the log-likelihood function (2.18). This is also the main difference between our proposed SML estimation and the conventional SML estimation.

Our maximization of (2.18) is like follows.

(1) Define a new representation of SML model for covariance matrix R which has a consideration of rank$\{S\} = r$, and r is unkown.

(2) Maximize the likelihood function under the application of new representation of R and consider that the condition C2) is already satisfied.

(3) Show what we should do when the condition C2) is violated.

4.2.1 New Representation of SML model for R

First we assume that rank$\{S\}$ = rank$\{P\}$ = $r \leq q$ for the stochastic model in (2.15), (2.19) or (2.20). Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_q$ be eigenvalues of R_{SS} and their corresponding eigenvectors be $v_1, v_2, ..., v_q$. Using the spectral representation, the new reprsentation of covariance matrix R is shown by the following R_r.

$$R_r = G(\Theta)\left(\sum_{k=1}^{r}(\lambda_k - \sigma^2)u_k u_k^H + \sigma^2 I_p\right)G^H(\Theta) \tag{4.1}$$

$$u_k = [v_k^T \ 0 \ 0 \ ... \ 0]^T \ \text{for } k = 1, 2, ..., r \tag{4.2}$$

where unknown parameters are $\{ \Theta, \lambda_1, \lambda_2, ..., \lambda_r, v_1, v_2, ... v_r, \sigma^2 \}$ and the constraint condition C2) is replaced as follows.

C2') $\lambda_1 \geq ... \geq \lambda_r > \sigma^2$

The following additional constraint is imposed on the model in (4.1).

C4) $\{ v_1, v_2, ... v_r \}$ is an orthonormal system of an r-dimensional complex Euclid space $\mathcal{C}^r$.

Implicitly we have $\lambda_{r+1} = ... = \lambda_q = \sigma^2$ and $\{ v_{r+1}, ... v_q \}$ compose an orthonormal system of the orthogonal complement of span$\{ v_1, v_2, ... v_r \}$.

4.2.2 Likelihood Function

The inverse of the matrix $\boldsymbol{R}_r$ is represented as

$$\boldsymbol{R}_r^{-1} = \boldsymbol{G}(\Theta) \left(\sum_{k=1}^r (\frac{1}{\lambda_k} - \frac{1}{\sigma^2}) \boldsymbol{u}_k \boldsymbol{u}_k^H + \frac{1}{\sigma^2} \boldsymbol{I}_p \right) \boldsymbol{G}^H(\Theta). \tag{4.3}$$

Substituting $\boldsymbol{R}_r$ in (4.1) and $\boldsymbol{R}_r^{-1}$ in (4.3) into (2.18), we have the log-likelihood function as follows.

$$\begin{aligned} L_r(&\Theta, \lambda_1, ..., \lambda_r, \boldsymbol{v}_1, ..., \boldsymbol{v}_r, \sigma^2) \\ &= -M \left(\ln \left\{ \lambda_1 \lambda_2 ... \lambda_r \left(\sigma^2 \right)^{p-r} \right\} + \sum_{k=1}^r \frac{l_k(\Theta, \boldsymbol{v}_k)}{\lambda_k} \right. \\ &\qquad \left. + \frac{1}{\sigma^2} \left\{ \operatorname{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{k=1}^r l_k(\Theta, \boldsymbol{v}_k) \right\} \right) \end{aligned} \tag{4.4}$$

where

$$\begin{aligned} l_k(\Theta, \boldsymbol{v}_k) &= \boldsymbol{u}_k^H \boldsymbol{G}^H(\Theta) \tilde{\boldsymbol{R}} \boldsymbol{G}(\Theta) \boldsymbol{u}_k \\ &= \boldsymbol{v}_k^H \tilde{\boldsymbol{R}}_{SS}(\Theta) \boldsymbol{v}_k \quad \text{for } k = 1, 2, ..., r \end{aligned} \tag{4.5}$$

and $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ is given as in step (4) in conventional SML.

The SML estimation is the maximization problem to find Θ, $\lambda_1,..., \lambda_r$, $\boldsymbol{v}_1, ..., \boldsymbol{v}_r, \sigma^2$ which maximize the likelihood function in (4.4) under the constraints of C1), C2'), C3) and C4).

In subsection 5.3, 5.4 and 5.5, we assume that the condition C2') is already satisfied in order to make the derivation process simple; in subsection 5.6 we discuss what we should do when the condition C2') is violated; at last we show the exact SML criterion for DOA finding with two different cases.

4.2.3 Maximization by $\lambda_1, \lambda_2, \dots \lambda_r$ and σ^2

Given Θ and an orthonormal system $\{ \boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r \}$ we consider the maximization of L_r in (4.4) with respect to $\lambda_1, \lambda_2, \dots, \lambda_r$ and σ^2.

Since the function

$$h(\lambda) = -\left(\ln \lambda + \frac{l}{\lambda} \right) \tag{4.6}$$

has a single peak at $\lambda = l$, it is easy to get that

$$\lambda_k = l_k(\Theta, \boldsymbol{v}_k) \quad \text{for } k = 1, 2, \dots, r \tag{4.7}$$

$$\sigma^2 = \sigma_r^2 = \frac{1}{p-r} \left(\text{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{i=1}^{r} l_i(\Theta, \boldsymbol{v}_i) \right). \tag{4.8}$$

and $l_k(\Theta, \boldsymbol{v}_k)$ are given by (4.5).

Then the likelihood function is rewritten as

$$L_r(\Theta, \boldsymbol{v}_1, \dots, \boldsymbol{v}_r) = -M \ln \left\{ \lambda_1 \lambda_2 \dots \lambda_r \left(\sigma^2 \right)^{p-r} \right\} \tag{4.9}$$

where λ_k and σ^2 are given above.

4.2.4 Maximization by $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_r$

Fixing $\boldsymbol{\Theta}$, then maximize $L_r(\Theta, \boldsymbol{v}_1, \dots, \boldsymbol{v}_r)$ in (4.15) with respect to $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_r$. We prove that these vectors become eigenvectors of $\tilde{\boldsymbol{R}}_{SS}(\Theta)$.

Introducing Lagrange's multipliers under the condition C4) in a new criterion, taking derivatives of the new criterion with respect to unknown real parameters in $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots,$ $\boldsymbol{v}_r$ and making the derivatives equal to zero, then we can obtain a set of r equations in

the complex form as follows.

$$\left(\frac{1}{\sigma^2} - \frac{1}{l_k(\Theta)}\right)\tilde{R}_{SS}(\Theta)v_k$$

$$= \alpha_{k1}v_1 + \alpha_{k2}v_2 + ... + \alpha_{kr}v_r \tag{4.10}$$

$$\alpha_{ki} = \bar{\alpha}_{ik} \qquad \text{for } k, i = 1, 2, ..., r. \tag{4.11}$$

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

Multiplying (4.10) by v_i^H from the left, we have

$$\alpha_{ki} = \left(\frac{1}{\sigma^2} - \frac{1}{l_k(\Theta)}\right) v_i^H \tilde{R}_{SS}(\Theta)v_k. \tag{4.12}$$

And we also have

$$\alpha_{ki} = \bar{\alpha}_{ki} = \left(\frac{1}{\sigma^2} - \frac{1}{l_i(\Theta)}\right) \left(v_k^H \tilde{R}_{SS}(\Theta)v_i\right)^H$$

$$= \left(\frac{1}{\sigma^2} - \frac{1}{l_i(\Theta)}\right) v_i^H \tilde{R}_{SS}(\Theta)v_k. \tag{4.13}$$

Next we prove that $v_1, v_2, ...v_r$ are eigenvectors of $\tilde{R}_{SS}(\Theta)$. For simplicity we use l_k to replace $l_k(\Theta)$ in the proving process.

In the Case of $l_k \neq l_i$

From (4.12) and (4.13), 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

$$\left(\frac{1}{\sigma_r^2} - \frac{1}{l_k}\right) \tilde{R}_{SS}(\Theta)v_k = \alpha_{kk}v_k \quad \text{for } k = 1, 2, ..., r. \tag{4.14}$$

Therefore $v_1, v_2, ..., v_r$ must be eigenvectors of the matrix $\tilde{R}_{SS}(\Theta)$ and orthogonal each other. The Hermitian form $l_k = v_k^H \tilde{R}_{SS}(\Theta)v_k$ in (4.5) is an eigenvalue of the eigenvector v_k for $k = 1, 2, ..., r$.

In the Case of $l_1 = l_2 = ... = l_\mu$

In the case that $l_1 = l_2 = ... = l_\mu$, we can also derive that $\alpha_{ki} = 0$ for $k \neq i$ as shown in appendix (B) using appendix (A). Therefore $l_1(\Theta), l_2(\Theta), ..., l_r(\Theta)$ are eigenvalues of $\tilde{R}_{SS}(\Theta)$.

After that we have the likelihood function as

$$L_r(\Theta) = -M \ln \left\{ \lambda_1 \lambda_2 ... \lambda_r \left(\sigma^2\right)^{p-r} \right\} \tag{4.15}$$

where

$$\lambda_k = l_k(\Theta) \quad \text{for } k = 1, 2, ..., r \tag{4.16}$$

$$\sigma^2 = \sigma_r^2 = \frac{1}{p-r} \left(\text{tr}\{\tilde{R}\} - \sum_{i=1}^{r} l_i(\Theta) \right). \tag{4.17}$$

and $l_1(\Theta), l_2(\Theta), ..., l_r(\Theta)$ are eigenvalues of $\tilde{R}_{SS}(\Theta)$.

4.2.5 Optimal Selection of Eigenvalues

Assuming that there exist a selection of eigenvalues in (4.17) that satisfies the condition C2'), It is mathematically easy to know that the optimal selection is obtained as the selection of largest r eigenvalues.

Let $\tilde{\lambda}_1(\Theta) \geq \tilde{\lambda}_2(\Theta) \geq ... \geq \tilde{\lambda}_q(\Theta)$ be eigenvalues of the matrix $\tilde{R}_{SS}(\Theta)$ and $\tilde{v}_1(\Theta), \tilde{v}_2(\Theta), ..., \tilde{v}_q(\Theta)$ be their corresponding eigenvectors. Unknown parameters are

represented by functions of Θ as

$$\lambda_k = \tilde{\lambda}_k(\Theta) \quad \text{for } k = 1, 2, ..., r \tag{4.18}$$

$$\boldsymbol{v}_k = \tilde{\boldsymbol{v}}_k(\Theta) \quad \text{for } k = 1, 2, ..., r \tag{4.19}$$

$$\sigma^2 = \tilde{\sigma}_r^2(\Theta) = \frac{1}{p-r} \left(\text{tr}\{\tilde{\boldsymbol{R}}\} - \sum_{i=1}^{r} \tilde{\lambda}_i(\Theta) \right). \tag{4.20}$$

The final form of the likelihood function is written as follows.

$$L_r(\Theta) = -M \ln \left\{ \tilde{\lambda}_1(\Theta)\tilde{\lambda}_2(\Theta)...\tilde{\lambda}_r(\Theta) \left(\tilde{\sigma}_r^2(\Theta) \right)^{p-r} \right\} \tag{4.21}$$

4.2.6   Applying the Condition C2')

Define the ML solution space for the model of $\boldsymbol{R}_r$ as

$$\Omega_r = \{\Theta \mid \Theta \in \boldsymbol{U} \text{ and } \tilde{\lambda}_r(\Theta) > \tilde{\sigma}_r^2(\Theta)\}. \tag{4.22}$$

As well as the conventional SML, the log-likelihood function in (4.21) has meaning if and only if the constraint condition C2') or $\tilde{\lambda}_r(\Theta) > \tilde{\sigma}_r^2(\Theta)$ is satisfied. The model with rank$\{S\} = r$ has a likelihood function only for $\Theta \in \Omega_r$.

Because of the use of normal distributions, the likelihood function in (4.21) is continuous. When Θ approaches to the border of Ω_r from its inside, both $\tilde{\lambda}_k(\Theta)$ and $\tilde{\sigma}_r^2(\Theta)$ approach to a certain value keeping the relation $\tilde{\lambda}_k(\Theta) > \tilde{\sigma}_r^2(\Theta)$. Finally they have an identical value on the border. Then the matrix $\boldsymbol{S}$ can not keep the rank r. Outside the border, if the condition C2') is relaxed to $\tilde{\lambda}_r(\Theta) \geq \tilde{\sigma}_r^2(\Theta)$, the solution becomes $\tilde{\lambda}_k(\Theta) = \tilde{\sigma}_\rho^2(\Theta)$ for $k = \rho, \rho + 1 ... r$ with certain ρ less than r. Therefore, the model with rank$\{S\} = r$ does not have a likelihood function outside Ω_r. Instead a certain model with rank$\{S\} < r$ has.

Furthermore if the model with rank$\{S\} = r$ has a likelihood function for a certain Θ, also the mode with rank$\{S\} < r$ has a likelihood function for the same Θ because $\tilde{\lambda}_{r-1}(\Theta) > \tilde{\sigma}^2_{r-1}(\Theta)$ is readily derived from $\tilde{\lambda}_r(\Theta) > \tilde{\sigma}^2_r(\Theta)$.

From the discussion above, we can state that, for each $\Theta \in U$, there exists the maximum rank of S that the stochastic model has a likelihood function. Let it be represented by $\eta(\Theta)$. $\eta(\Theta)$ is determined in the following way.

- If $\tilde{\lambda}_1(\Theta) \le \tilde{\sigma}^2_1(\Theta)$, then $\eta(\Theta) = 0$.

- If $\tilde{\lambda}_q(\Theta) > \tilde{\sigma}^2_q(\Theta)$, then $\eta(\Theta) = q$.

- If $\tilde{\lambda}_k(\Theta) > \tilde{\sigma}^2_k(\Theta)$ and $\tilde{\lambda}_{k+1}(\Theta) \le \tilde{\sigma}^2_{k+1}(\Theta)$ for a certain $k = 1, 2, ..., q-1$, then $\eta(\Theta) = k$.

Next, we compare the values of $L_{r-1}(\Theta)$ and $L_r(\Theta)$ under the condition of $\tilde{\lambda}_r(\Theta) > \tilde{\sigma}^2_r(\Theta)$. Let $\tilde{\lambda}_r(\Theta) = (1 + \alpha)\tilde{\sigma}^2_r(\Theta)$ and $n = p - r + 1$, where $\alpha > 0$. Then we have

$$L_r(\Theta) - L_{r-1}(\Theta) = N \ln \left(\frac{\left(1 + \dfrac{\alpha}{n}\right)^n}{1 + \alpha} \right) > 0. \tag{4.23}$$

It follows from (4.23) that

$$L_0(\Theta) < L_1(\Theta) < ... < L_{\eta(\Theta)}(\Theta). \tag{4.24}$$

4.2.7 Exact SML Criterion

The exact SML estimation of DOA has different criterion when the rank condition of S is imposed or not.

The first case: If no condition is imposed on the rank of S, $L_{\eta(\Theta)}(\Theta)$ gives the maximum value for fixed Θ among all possible log-likelihood functions with different

rank of $\boldsymbol{S}$. Therefore, the exact ML estimations are obtained as follows.

$$\hat{\Theta}_e = \arg \max_{\Theta} L_{\eta(\Theta)}(\Theta), \tag{4.25}$$

$$\hat{\lambda}_{ke} = \tilde{\lambda}_k(\hat{\Theta}_e) \quad \text{for } k = 1, 2, ..., \eta(\hat{\Theta}_e) \tag{4.26}$$

$$\hat{\boldsymbol{v}}_{ke} = \tilde{\boldsymbol{v}}_k(\hat{\Theta}_e) \quad \text{for } k = 1, 2, ..., \eta(\hat{\Theta}_e) \tag{4.27}$$

$$\hat{\sigma}_e^2 = \tilde{\sigma}_{\eta(\hat{\Theta}_e)}^2(\hat{\Theta}_e). \tag{4.28}$$

Furthermore the SML estimation of the signal covariance matrix is given as follows.

$$\hat{\boldsymbol{R}}_{SS_e} = \sum_{k=1}^{\eta(\hat{\Theta}_e)} \hat{\lambda}_{ke} \hat{\boldsymbol{v}}_{ke} \hat{\boldsymbol{v}}_{ke}^H, \quad \hat{\boldsymbol{P}}_e = \hat{\boldsymbol{R}}_{SS_e} - \hat{\sigma}_e^2 \boldsymbol{I}, \tag{4.29}$$

$$\hat{\boldsymbol{S}}_e = \boldsymbol{T}^{-1}(\hat{\Theta}_e) \hat{\boldsymbol{P}}_e \boldsymbol{T}^{-H}(\hat{\Theta}_e), \tag{4.30}$$

where the non-negative definiteness of $\hat{\boldsymbol{S}}_e$ is guaranteed because of the definition of $\eta(\Theta)$.

The second case: If we try to find an exact SML estimation of the model with rank$\{\boldsymbol{S}\} = r$ for certain fixed r, the solution can be obtained as long as the SML solution space Ω_r includes a local maximum of $L_r(\Theta)$. The exact SML estimations have the same forms as in (4.25) to (4.30) except for replacing $\eta(\Theta)$ with r. If the SML solution space Ω_r does not include any local maximum of $L_r(\Theta)$, there does not exist the likelihood SML estimation for the fixed r of rank$\{\boldsymbol{S}\}$.

Furthermore, we should note that if we impose that $\eta(\Theta) = q$, our proposed exact SML criterion is equivalent to the conventional SML criterion.

4.3 Performance Study of Exact SML

In this section, some simulation results are shown to demonstrate properties of the proposed exact SML estimation of DOA by comparing other techniques.

In the simulations, the array configuration is a uniform linear array composed of omni-directional sensors, of which steering vector is represented as

$$\boldsymbol{a}(\theta) = [\, 1 \; e^{-j\phi(\theta)} \; \cdots \; e^{-j(p-1)\phi(\theta)}]^T, \tag{4.31}$$

$$\phi(\theta) = \frac{2\pi\Delta}{\lambda} \sin\theta. \tag{4.32}$$

where λ is the wavelength of signals impinging on the array, Δ is the sensor spacing between adjacent two sensors. As a necessary condition that a unique direction θ is determined by the phase parameter ϕ, $\Delta \leq \lambda/2$ is imposed on the array configuration. In our simulations, $\Delta = \lambda/2$.

The SNR is defined as

$$\mathrm{SNR}_k = 10\log_{10}\frac{\mathrm{E}[|s_k(t)|^2]}{\sigma^2}. \tag{4.33}$$

All signals are assumed to have the identical SNR. The Root-Mean-Square-Error (RMSE) is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{qN}\sum_{k=1}^{q}\sum_{n=1}^{N}|\hat{\theta}_{k,n} - \theta_k|^2}, \tag{4.34}$$

where $\hat{\theta}_{k,n}$ is the estimation of θ_k at the n-th trial.

The stochastic Carmer-Rao lower bound (CRLB) [8, 16] is also done in the simulation.

4.3.1 Property of Global Solutions

Fig. 5.1 shows the solution property of exact SML. The conditions are written in the captions. In Fig. 5.1, the solid lines that $\det\{\tilde{\boldsymbol{R}}_{SS}(\Theta)\} = 0$ in noise-free case are canceled.
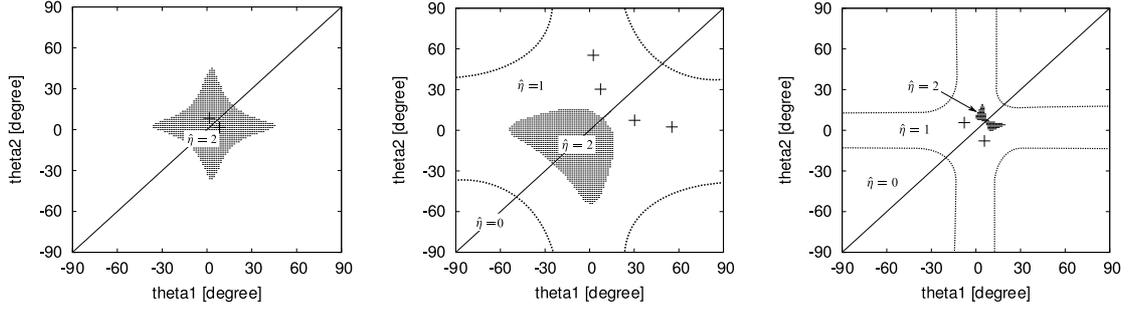
(a) The exact solution locates in the shadow space $\hat{\eta} = 2$.

(b) The exact solution locates in the space $\hat{\eta} = 1$.

(c) The exact solution locates in the space $\hat{\eta} = 1$

Fig. 5.1: Samples in the exact SML estimation. The scenarios are the same as in Fig. 4.1 respectively. Crosses represent the local solutions of exact SML. The shadow represents the space $\hat{\eta} = 2$; In (b) and (c) the space which locates out of the shadow and inside of the dotted lines represents $\hat{\eta} = 1$; the other space represents $\hat{\eta} = 0$.
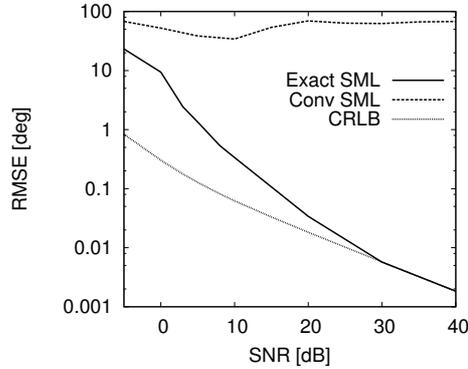

Fig. 5.2: Property of global solutions of the exact and conventional SML and local search for the conventional SML.

Comparing Fig. 4.1 and Fig. 5.1, we can find the following facts. (1) All the inadequate solutions around the solid lines in Fig. 4.1 vanish in the exact SML estimation. (2) The local solution located in the shadow area in Fig. 4.1 (a) is the exact SML solution which is the same as in Fig. 5.1 (a). (3) In conventional SML estimation, there are no SML solution in Fig. 4.1 (b) and (c). While in exact SML estimation, the exact SML solution can be found in the space $\hat{\eta} = 1$ as shown in Fig. 5.1 (b) and (c). This indicates that there is no SML solution with the model rank$\{S\} = 2$ and the solution exists in the model rank$\{S\} = 1$ in these two cases.

31

Fig. 5.2 shows the comparison of RMSE between the conventional SML and the proposed exact SML. The scenario of the simulation is the same as Fig. 4.1 (a) and (b) and 30 independent trials are fulfilled.

In Fig. 5.2, "Conv SML" and "exact SML" show the results from the conventional SML and the proposed exact SML, respectively. It is found from Fig. 5.2 that RMSE of the conventional SML is very large even when the SNR is very high. This indicates that there exist inadequate global solutions in the conventional SML. It is also found from Fig. 5.2 that the proposed exact SML provides good solutions.

4.3.2   Sensitivity to the Rank of S in the Estimation Model

The WSF estimation needs to know the rank of S which is used for its formulation. Let r_w represent the rank of S in the model of the WSF estimation. In this subsection, we investigate the sensitivity of DOA estimation to r_w.

Fig. 5.3 shows RMSE of the WSF with different r_w. The scenario is $p = 6$, $q = 4$, $\theta_1 = 0°$, $\theta_2 = 10°$, $\theta_3 = 20°$, $\theta_4 = 30°$, rank$\{S\} = r = 3$ and $N = 300$. Three cases that $r_w = 2$, $r_w = 3$ and $r_w = 4$ are assumed and 30 independent trials are fulfilled.

It is found from Fig. 5.3 that the WSF shows excellent estimation if the true rank of S is given, i.e., $r_w = 3$. However in both cases that r_w is larger and less than the true rank, the estimation performance becomes worse. The WSF is sensitive to r_w. If the true rank is not known, r_w should be estimated by using some proper technique. Otherwise $r_w = q$ is used.

In the proposed exact SML, the rank of S in the model is determined as $\eta(\Theta)$ automatically. In order to compare the exact SML with the WSF, we introduce the upper
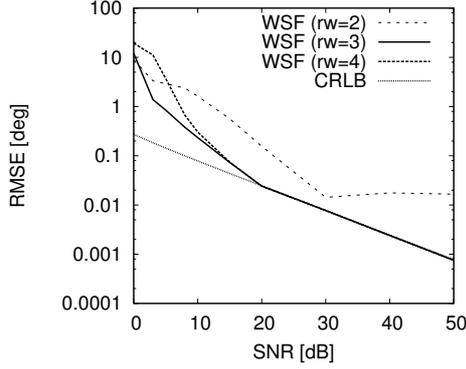
Fig. 5.3: Sensitivity of WSF to the rank of S in the estimation model.

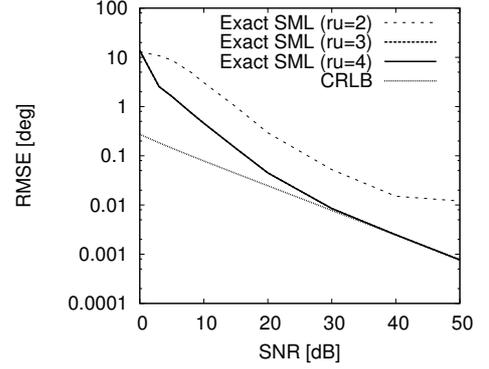Fig. 5.4: Sensitivity of the exact SML to the rank of S in the estimation model.

limit of the rank of S in the model of the SML estimation. Let r_u represents the upper limit and r_u is used as follows. The SML criterion $L_{r_u}(\Theta)$ is used instead of $L_{\eta(\Theta)}(\Theta)$ if $\eta(\Theta) \geq r_u$. Otherwise $L_{\eta(\Theta)}(\Theta)$ is used. In the case that $r_u = q$, the exact SML estimation with $L_{\eta(\Theta)}(\Theta)$ is fulfilled.

Fig. 5.4 shows the simulation results for $r_u = 2$, $r_u = 3$ and $r_u = 4$. The scenario is the same as in Fig. 5.3. Difference in estimation performance can not be seen, between the cases that $r_u = 3$ and $r_u = 4$, while the case $r_u = 2$ shows worse performance. From this result, we can say that, as long as the upper limit of the rank of S in the model is not less than the true rank, the exact SML is not sensitive to the rank. The use of $L_{\eta(\Theta)}(\Theta)$ is the best selection.

4.3.3 Comparison of RMSE

In Fig. 5.5, we show comparisons of RMSE for several DOA estimation methods. Fig. 5.6 is a enlarge figure of Fig. 5.5 when SNR is less than 20 dB.

The scenario of the simulation is the same as in Fig. 5.3 and Fig. 5.4 and the rank S is assumed to be unknown.
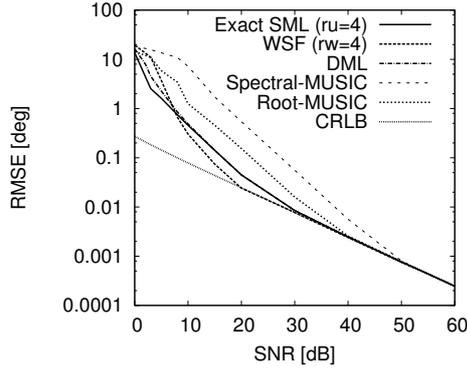
Fig. 5.5: RMSE comparison of exact SML, WSF, DML, Spectral-MUSIC and Root-MUSIC when the rank of S is unknown.
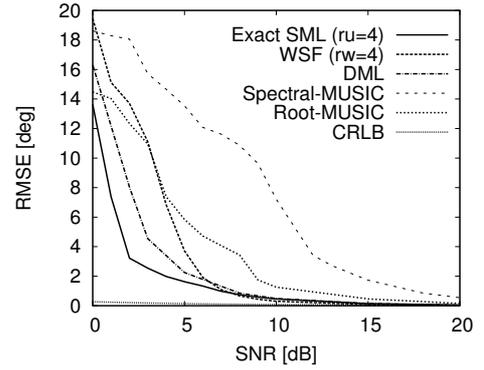
Fig. 5.6: The enlarge figure of Fig.5.5 when SNR is less than 20 dB.

We pick up the methods of the exact SML, the DML, the WSF with $r_w = q = 4$, the Spectral-MUSIC [18] and the Root-MUSIC [19] estimation. In order to apply the Spectral-MUSIC and the Root-MUSIC estimation to coherent signal models, a spatial smoothing technique [20] is used. In Fig. 5.5, these are labeled by their name.

we can find that the RMSE of all methods including the Spectral MUSIC get closer to CRLB as SNR becomes larger. In the scenario of this simulation, signals are coherent. Therefore SNR needs to be very high so that the RMSE of the MUSIC approaches to CRLB.

From Fig. 5.6, we can find that the estimation performance of the exact SML is the best in the threshold region (when SNR is less than about 10 dB), and when the condition is good its performance is a litter worse than that of WSF as shown in Fig. 5.6. The same results are observed in other simulations.
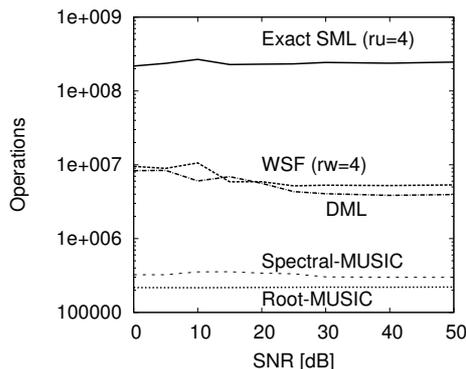
Fig. 5.7: Computational load of exact SML, WSF, DML, Spectral-MUSIC and Root-MUSIC.

4.3.4 Comparison of Operations

In fig. 5.7, we show comparison of the average number of floating operations of all of addition, subtraction, multiplication and division of the methods picked up in Fig. 5.5. The scenario is also the same as in Fig. 5.5. In Fig. 5.7, they are labeled by their name. From Fig. 5.7, we can find that there is no advantage of the proposed method in operations since eigen-decomposition is involved in each step of iterations.

4.4 Conclusions

In this chapter, we gave the exact formulation of SML estimation and then a performance study of exact SML is also done. Based on the above simulations, we can find that the ML-based methods require large computational cost compared with Spectral-MUSIC and Root-MUSIC. If signals are incoherent and observation condition is good, Root-MUSIC give almost equivalent RMSEs to ML-based methods. However, ML-based methods especially our proposed exact SML show much better performance in estimation accuracy than Root-MUSIC in the case of coherent signals and in the threshold region, such as low SNR or small number of samples. Furthermore, the proposed

35

exact SML has following advantages. First it provides the exact global solution of SML, which can become a standard for evaluating estimation accuracy of other methods, especially approximated SML methods. Second, the proposed exact SML method does not have any theoretic problem in applying the ML solution to other techniques such as the model selection technique based on MDL principle.

Chapter 5

Efficient Algorithms for the SML Estimation

5.1   Introduction

As we have discussed above, the exact SML estimation requires high computational cost. It is difficult to exploit efficient algorithm for the exact SML criterion. To reduce computational cost possibly with less loss of resolution. We propose three efficient algorithms to find the optimal or suboptimal solution of SML.

At first, we propose a local search method together with the DML estimation as initialization. Then based on the local search method, we propose an effective version of AM (EAM) algorithm to find the solution of SML. At last the irreducible form of EAM, called IAM is also proposed. The validity of the algorithms are also shown by simulations.

5.2   Use of Local Search Method

Through the discussion in chapter 4 and 5 (especially, form Fig. 4.2 and Fig. 5.1), we have known that, the conventional SML fails in DOA finding occasionally because the global solution is ambiguity. The global solution of conventional SML is sometimes not adequate. However, the adequate solution of SML do exist in the local solutions of the conventional SML estimation. Therefore, we wonder whether we can find a method which can find the adequate solution of SML from the conventional SML estimation. The motivation we do this is that it is much easier to exploit efficient algorithms based

on the conventional SML criterion.

Since global searches have no chance to get the adequate solution of the SML, local search method may have a good chance to get the adequate solution. The remaining problem is how to determine the initial value.

In the noise-free case, it can be proved that the DML estimation gives unique solution which is identical to the true DOA. The proof is shown in Appendix D. This solution is also one of the solutions of the SML estimation. Therefore the solution of the DML is the adequate solution of the SML in the noise-free case.

In the noisy case, if σ^2 is enough small, or SNR is enough high, the solution of DML can be enough close to the true DOA. So the solution of the DML estimation provides an excellent initial value. Therefore, we propose the algorithm that a local search is applied to the conventional SML criterion using the solution of DML estimation as an initial value.

The local search is fulfilled by replacing the one-dimensional global search with any one-dimensional local search in the update process of the convergence phase of the AM algorithm.

If the SNR is enough high, the proposed algorithm is successful. The level of SNR considered to be enough high is not clear. Since the inadequate solutions of SML are located far from the true DOA, it is expected that the initial value obtained by the DML estimation are still valid when SNR is large some extent.
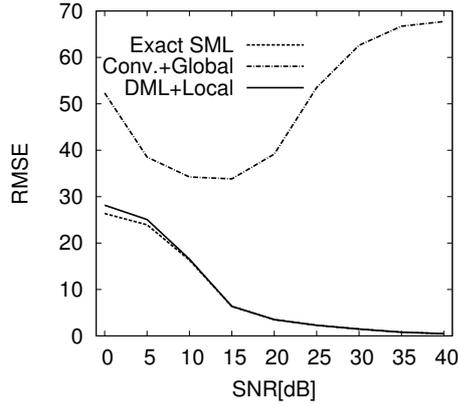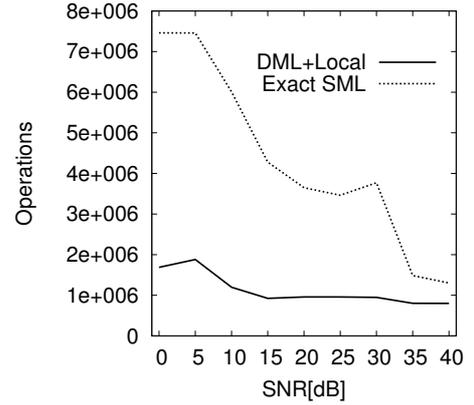
| (a) Comparison of RMSE. | (b) Comparison of operations. |

Fig. 6.1: the scenario is the same as Fig. 5.2, "DML+Local": proposed method, "Conv.+Global": the AM algorithm for the coventional SML estimation, "Exact UML": exact UML estimation with global search, RMSE: Root-Mean-Square-Error.

5.2.1 Validity of the Local Search Method

Simulation results are shown in Fig. 6.1. The scenarios and the labels are explained in the caption.

In Fig. 6.1 (a), we can find the following facts: 1) The global search for conventional SML estimation fails completely in DOA finding in this case. 2) When the observation condition is good (when SNR is greater than about 10dB), the RMSE of the proposed local search algorithm is the same as that of the exact SML estimation. 3) In the threshold region (SNR is less than 10dB), the RMSE is a litter worse than that of the exact SML estimation.

The second fact holds because the adequate solution of SML is correctly searched, and this solution is the optimal solution and is the same as that of exact SML. The third fact happens because the cases like Fig. 4.2 (b) or (c) appears. The adequate solution of SML is not found in this case. The solution of the proposed method searches the suboptimal solution of SML in this case. This case only occurs in the threshold region.

However, in simulations we have confirmed that when the DOA finding is resolvable (for example, in Fig. 6.1 the two signals are located in 0 and 10 degree, so if RMSE is less than about 5, we consider that the DOA is resolvable.), the proposed local search method can show the same RMSE performance of exact SML.

In Fig. 6. 1(b), the computational cost of the local search method is much lower that of the exact SML estimation.

5.3 An Efficient Version of AM Algorithm (EAM) for the SML Estimation

In the previous section, we have shown that the proposed local search algorithm is effective and efficient. However, when p and or q becomes larger, the computational load still seems to be high. Therefore, we try to establish a more efficient algorithm.

In this section, we propose an efficient version of the AM algorithm. We call it the EAM algorithm. The main idea of this algorithm consists in dividing the conventional SML criterion into two components. One depends on a variable parameter and the other component is independent of the variable parameter. To simplify the derivation, From now on, the˜of a matrix is omitted.

In each updating process, let θ_l be a variable parameter and define

$$\Theta_l = \{\theta_1, \theta_2, ..., \theta_{l-1}, \theta_{l+1}, ..., \theta_q\} \tag{5.1}$$

$$A_l = [a(\theta_1), ..., a(\theta_{l-1}), a(\theta_{l+1}), ..., a(\theta_q)] \tag{5.2}$$

V_{S_l} : an orthonormal system of the subspace spanned by $\{A_l\}$.

V_{N_l} : an orthonormal system of the orthogonal complement of the

subspace spanned by $\{A_l\}$

$$P_{A_l} = A_l\{A_l^H A_l\}^{-1} A_l^H = V_{S_l} V_{S_l}^H \tag{5.3}$$

$$P_{A_l}^{\perp} = I - P_{A_l} = V_{N_l} V_{N_l}^H \tag{5.4}$$

$$R_{S_l} = V_{S_l}^H R V_{S_l} \tag{5.5}$$

$$R_{N_l} = V_{N_l}^H R V_{N_l} \tag{5.6}$$

$$v_l(\theta_l) = P_{A_l}^{\perp} a(\theta_l) / \parallel P_{A_l}^{\perp} a(\theta_l) \parallel \tag{5.7}$$

Note that, when the value of θ_l changes, only $v_l(\theta_l)$ varies and all others above are fixed. From these definitions, we have

$$\mathrm{span}\{V_{N_l}\} = \mathrm{span}\{v_l(\theta_l)\} \oplus \mathrm{span}\{V_N(\Theta)\} \tag{5.8}$$

$$\mathrm{span}\{V_S(\Theta)\} = \mathrm{span}\{v_l(\theta_l)\} \oplus \mathrm{span}\{V_{S_l}\} \tag{5.9}$$

where $\oplus$ represents the direct sum of subspaces. It follows from (5.8) and (5.9) that there exist a $(p - k + 1) \times (p - k + 1)$ unitary matrix T_1 and a $k \times k$ unitary matrix T_2 which satisfy

$$V_{N_l} = [v_l(\theta_l) \; V_N(\Theta)] \, T_1 \tag{5.10}$$

$$V_S(\Theta) = [v_l(\theta_l) \; V_{S_l}] \, T_2 \tag{5.11}$$

Substituting (5.10) into (5.6), we have

$$R_{N_l}$$

$$= T_1^H \begin{bmatrix} v_l^H(\theta_l) R v_l(\theta_l) & v_l^H(\theta_l) R V_N(\Theta) \\ V_N^H(\Theta) R v_l(\theta_l) & V_N^H(\Theta) R V_N(\Theta) \end{bmatrix} T_1$$

$$= T_1^H \begin{bmatrix} v_l^H(\theta_l) R v_l(\theta_l) & v_l^H(\theta_l) R V_N(\Theta) \\ V_N^H(\Theta) R v_l(\theta_l) & R_{NN}(\Theta) \end{bmatrix} T_1 \tag{5.12}$$

Taking the trace of both sides in (5.12), we have

$$
\begin{aligned}
\text{tr}\{\boldsymbol{R}_{N_l}\} &= \text{tr}\{\boldsymbol{R}_{NN}(\Theta)\} + \boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{v}_l(\theta_l) \\
&= \text{tr}\{\boldsymbol{R}_{NN}(\Theta)\} + \frac{\boldsymbol{a}^H(\theta_l)\boldsymbol{V}_{N_l}\boldsymbol{R}_{N_l}\boldsymbol{V}_{N_l}^H\boldsymbol{a}(\theta_l)}{\boldsymbol{a}^H(\theta_l)\boldsymbol{V}_{N_l}\boldsymbol{V}_{N_l}^H\boldsymbol{a}(\theta_l)}
\end{aligned}
\tag{5.13}
$$

Substituting (5.11) into (5.5), we have

$$
\begin{aligned}
&\boldsymbol{R}_{SS}(\Theta) \\[2mm]
&= \boldsymbol{T_2}^H
\begin{bmatrix}
\boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{v}_l(\theta_l) & \boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{V}_{S_l} \\[2mm]
\boldsymbol{V}_{S_l}^H\boldsymbol{R}\boldsymbol{v}_l(\theta_l) & \boldsymbol{V}_{S_l}^H\boldsymbol{R}\boldsymbol{V}_{S_l}
\end{bmatrix}
\boldsymbol{T_2} \\[4mm]
&= \boldsymbol{T_2}^H
\begin{bmatrix}
\boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{v}_l(\theta_l) & \boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{V}_{S_l} \\[2mm]
\boldsymbol{V}_{S_l}^H\boldsymbol{R}\boldsymbol{v}_l(\theta_l) & \boldsymbol{R}_{S_l}
\end{bmatrix}
\boldsymbol{T_2}
\end{aligned}
\tag{5.14}
$$

Taking the determinant of both sides in (5.14), we have

$$
\begin{aligned}
&\det\{\boldsymbol{R}_{SS}(\Theta)\} \\[2mm]
&= \det\{\boldsymbol{R}_{S_l}\}\left(\boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{v}_l(\theta_l) - \boldsymbol{v}_l^H(\theta_l)\boldsymbol{R}\boldsymbol{V}_{S_l}\boldsymbol{R}_{S_l}^{-1}\boldsymbol{V}_{S_l}^H\boldsymbol{R}\boldsymbol{v}_l^H(\theta_l)\right) \\[2mm]
&= \det\{\boldsymbol{R}_{S_l}\}\frac{\boldsymbol{a}^H(\theta_l)\boldsymbol{V}_{N_l}\boldsymbol{W}\boldsymbol{V}_{N_l}^H\boldsymbol{a}(\theta_l)}{\boldsymbol{a}^H(\theta_l)\boldsymbol{V}_{N_l}\boldsymbol{V}_{N_l}^H\boldsymbol{a}(\theta_l)}
\end{aligned}
\tag{5.15}
$$

where

$$
\boldsymbol{W} = \boldsymbol{R}_{N_l} - \boldsymbol{V}_{N_l}^H\boldsymbol{R}\boldsymbol{V}_{S_l}\boldsymbol{R}_{S_l}^{-1}\boldsymbol{V}_{S_l}^H\boldsymbol{R}\boldsymbol{V}_{N_l}
\tag{5.16}
$$

Define

$$
\boldsymbol{u}(\theta_l) = \boldsymbol{V}_{N_l}^H\boldsymbol{a}(\theta_l)
\tag{5.17}
$$

then the final form of the proposed algorithm, EAM algorithm, can be derived from (5.13) and (5.15) as follows,

$$
\text{tr}\{\boldsymbol{R}_{NN}(\Theta)\} = \text{tr}\{\boldsymbol{R}_{N_l}\} - \frac{\boldsymbol{u}^H(\theta_l)\boldsymbol{R}_{N_l}\boldsymbol{u}(\theta_l)}{\parallel \boldsymbol{u}(\theta_l)\parallel^2}
\tag{5.18}
$$

$$\det\{\boldsymbol{R}_{SS}(\Theta)\}$$
$$= \det\{\boldsymbol{R}_{S_l}\}\frac{\boldsymbol{u}^H(\theta_l)\boldsymbol{W}\boldsymbol{u}(\theta_l)}{\|\boldsymbol{u}(\theta_l)\|^2} \tag{5.19}$$

In (5.17), (5.18), and (5.19), all quantities except for $\boldsymbol{u}(\theta_l)$ are fixed and can be computed before starting the one-dimensional search with respect to θ_l. Therefore main computations of each step in the one-dimensional search are a product of the matrix $\boldsymbol{V}_{N_l}^H$ and the vector $\boldsymbol{a}(\theta_l)$ in (5.17) and evaluation of two Hermitian forms in (5.18) and (5.19).

5.3.1 Efficiency of the EAM Algorithm

In this subsection, we compare the computational complexity of the EAM algorithm, the local search algorithm proposed in section 6.2 and the exact SML estimation.

In simulation, The label "Exact SML" represents the original AM algorithm for exact SML estimation. "DML+Local" represents the local search method proposed in section 6.2. "EAM" represents the EAM algorithm proposed in section 6.3.

Simulation results are shown in Fig. 6.2. The scenario is shown in the caption of each figure.

From these figures, we can find that "DML+Local" is more efficient than "Exact SML" as well as we shown in subsection 6.2.1. "EAM" is more efficient than "DML+Local" as we expected. According to many simulations, we have found that the greater p, q are, more efficient of the EAM algorithm.
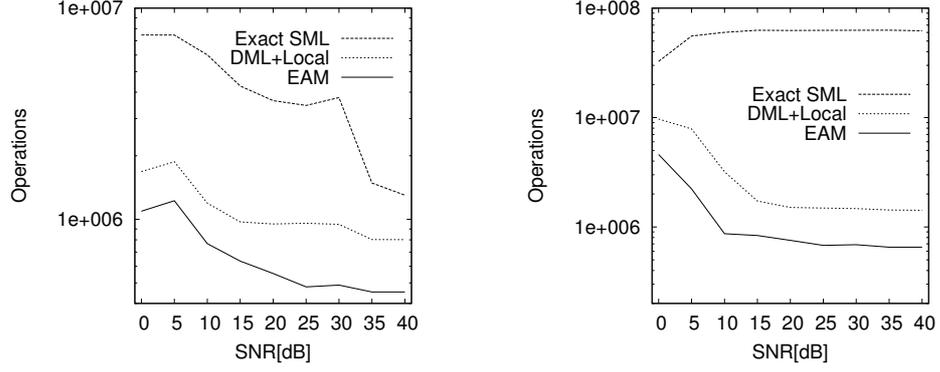
(a) $p = 3$, $q = 2$, $\eta = 2$, $\theta_1 = 0°$, $\theta_2 = 8°$, $M = 200$ and 30 independent trials.

(b) $p = 4$, $q = 3$, $\eta = 2$, $\theta_1 = 0°$, $\theta_2 = 8°$, $\theta_3 = 16°$, $M = 200$ and 30 independent trials.

Figure 6.2: Efficiency of the EAM algorithm.

5.4 Irreducible Form of EAM Algorithm (IAM)

Using the above AM and EAM algorithms in real calculation of the conventional SML estimation , we found that it becomes numerical unstable, when more than one direction parameters are going to have an identical value. It is caused by the fact that the conventional SML criterion with duplicated directions is indefinite. To solve this problem and reduce the computational cost more, in this section we propose the irreducible form of the EAM algorithm by applying a uniform linear array.

5.4.1 Numerical Instability in AM and EAM Criterion

In this subsection, we show the reason of the numerical Instability in AM and EAM algorithm for the conventional SML estimation.

Define the subspace

$$U(\Theta) = span\{a(\theta_1)a(\theta_2)...a(\theta_q)\}. \tag{5.20}$$

44

The linear combination

$$\frac{\boldsymbol{a}(\theta_j) - \boldsymbol{a}(\theta_i)}{\theta_j - \theta_i} \tag{5.21}$$

always belong to $\boldsymbol{U}(\Theta)$. Even when $\theta_j \to \theta_i$, we have $\dim\{\boldsymbol{U}(\Theta)\,|_{\theta_i \to \theta_j}\} = q$, where $\dim\{.\}$ represents the dimension of the subspace.

On the other hand, when $\theta_j = \theta_i$, we have $\dim\{\boldsymbol{U}(\Theta)\,|_{\theta_i = \theta_j}\} < q$. This implies that the criterion $L_c(\Theta)$ has discontinuous points in the parameter space.

Next, we consider computation of the EAM criterion $\boldsymbol{v}_l(\theta_l)$ in eq. (6.5). When the variable parameter θ_l approaches the value of a fixed parameter θ_i, $\boldsymbol{v}_l(\theta_l)$ vanishes and both the numerator and denominator in eq. (6.5) becomes zero. Then $\boldsymbol{v}_l(\theta_l)$ becomes indefinite. Therefore the computation of $\boldsymbol{v}_l(\theta_l)$ becomes numerical unstable. This can be verified in the simulation section

In the case that the DOA can be solved, the numerical instability does not occur, since each parameter in the convergence phase of the EAM criterion comes apart from others. However, at the threshold region, when more than one signal approach an identical value, the numerical instability becomes significant. In practice, when this case happens, the sequence of DOA obtained in the convergence phase of the EAM criterion shows oscillation. That is because the estimated bearings can not converge well due to the numerical instability and would oscillate around that identical value. Therefore, the oscillation makes the condition for terminating iterations complex and it requires extra computation. This will be shown in simulation section.

5.4.2 Irreducible Form of Efficient AM Criterion (IAM)

In this section, we present an improvement of EAM criterion using a uniform linear array. We derive an irreducible form of the efficient AM criterion (IAM) for UML estimation which is numerical stable and more efficient.

The array configuration is uniform linear array composed of omnidirectional sensors, of which steering vector is represented as

$$\boldsymbol{a}(\theta) = [\, 1 \; e^{-j\phi(\theta)} \; \cdots \; e^{-j(p-1)\phi(\theta)}]^T, \tag{5.22}$$

$$\phi(\theta) = \frac{2\pi\Delta}{\lambda} \sin\theta. \tag{5.23}$$

where λ is the wavelength of signals impinging on the array, Δ is the sensor spacing between adjacent two sensors. As a necessary condition that a unique direction θ is determined by the phase parameter ϕ, $\Delta \leq \lambda/2$ is imposed on the array configuration. In this paper, $\Delta = \lambda/2$.

Using the uniform linear array, we derive the irreducible form of eq. (6.18) and (6.19). Define

$$f_l(\theta) = \frac{\boldsymbol{u}^H(\theta_l)\boldsymbol{R}_{N_l}\boldsymbol{u}(\theta_l)}{\|\,\boldsymbol{u}(\theta_l)\,\|^2}, \quad \tilde{f}_l(\theta) = \frac{\boldsymbol{u}^H(\theta_l)\tilde{\boldsymbol{R}}_{N_l}\boldsymbol{u}(\theta_l)}{\|\,\boldsymbol{u}(\theta_l)\,\|^2} \tag{5.24}$$

which are the varying parts in eq. (6.18) and (6.19).

First we derive the irreducible form of the the eq. (6.18). Substituting eq. (6.4) and (6.17) into $f_l(\theta)$, we get

$$f_l(\theta) = \frac{\boldsymbol{a}^H(\theta_l)\boldsymbol{P}_{A_l}^{\perp}\boldsymbol{R}\boldsymbol{P}_{A_l}^{\perp}\boldsymbol{a}(\theta_l)}{\boldsymbol{a}^H(\theta_l)\boldsymbol{P}_{A_l}^{\perp}\boldsymbol{a}(\theta_l)}. \tag{5.25}$$

Using the uniform linear array defined above, the steering vector can be represented like follows,

$$\zeta_p(z) = [1 \; z \ldots z^{p-1}], \tag{5.26}$$

Then $f_l(\theta)$ can be rewritten into the form of a rational function

$$f_l(\theta) = \frac{\zeta_p(z)\boldsymbol{P}_{A_l}^{\perp}\boldsymbol{R}\boldsymbol{P}_{A_l}^{\perp}\zeta_{p*}(z)}{\zeta_p(z)\boldsymbol{P}_{A_l}^{\perp}\zeta_{p*}(z)} \tag{5.27}$$

$$= \frac{N(z)}{D(z)}\Big|_{z=e^{j\theta}}$$

where $\zeta_{p*}(z)$ is the para-conjugate of $\zeta_p(z)$ defined as

$$\zeta_{p*}(z) = \zeta_p^H(1/z^*) \tag{5.28}$$

$$= [1\ z^{-1}\ldots z^{-(p-1)}]^T,$$

and the superscript $*$ is the complex conjugate of a complex number.

Let θ_l be variable and all other parameters are held fixed as well as in section 6.3. Both the polynomials N(z) and D(z) have double zeros at $z = \theta_i$ in the complex z-plane, since it holds $\zeta_p(e^{j\theta_i})\boldsymbol{P}_{A_l}^{\perp} = 0$. Without canceling these common zeros, $f_l(\theta)$ is indefinite at $\theta = \theta_i$.

The irreducible form of the EAM criterion of $f_l(\theta)$ can be derived by canceling these common zeros. First, we define the polynomial $W_l(z)$ having zeros at $z = e^{j\theta_i}$, $i = 1, 2\ldots, l-1, l+1, \ldots, q$,

$$W_l(z) = \prod_{i=1,i\neq l}^{q} (z - e^{j\theta_i}) \tag{5.29}$$

$$= w_0 + w_1 z + \cdots + w_{q-1}z^{q-1}.$$

Using the coefficients of $W_k(z)$, define the following $p \times (p - q + 1)$ matrix as

$$
W_l(z) = \begin{bmatrix}
w_0 & 0 & \cdots & 0 \\
w_1 & w_0 & \cdots & 0 \\
\vdots & \vdots & \ddots & 0 \\
w_{q-1} & w_{q-2} & \ddots & w_0 \\
0 & w_{q-1} & \ddots & w_1 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & w_{q-1}
\end{bmatrix}
\tag{5.30}
$$

Then we have $a^H(\theta_i)W_k = 0$ for $i = 1, \ldots, k-1, k+1, \ldots, q$. Since the column vectors in W_l are all orthogonal to $a(\theta_i)$, the projection matrix $P_{A_l}^{\perp}$ can be written as

$$
P_{A_l}^{\perp} = W_l G_l^{-1} W_l^H
\tag{5.31}
$$

$$
G_l = W_l^H W_l
\tag{5.32}
$$

Using the expressions

$$
\zeta_p(z)W_l = [W_l(z)\ zW_l(z) \ldots z^{p-q}W_l z] = W_l(z)\zeta_{p-q}(z)
\tag{5.33}
$$

$$
\zeta_{p-q}(z) = [1 \ z \ldots z^{p-q}],
\tag{5.34}
$$

the irreducible form of $f_l(\theta)$ is derived as

$$
f_l(\theta) = \frac{\zeta_{p-q}(z)N_l\zeta_{(p-q)*}(z)}{\zeta_{p-q}(z)D_l\zeta_{(p-q)*}(z)} \Big|_{z=e^{j\theta}}
\tag{5.35}
$$

where

$$
N_l = G_l^{-1} W_l^H R W_l G_l^{-1} = \{n_{i,j}\}_{i,j=0}^{p-q}
$$

$$
D_l = G_l^{-1} = \{d_{i,j}\}_{i,j=0}^{p-q}
\tag{5.36}
$$

As for the irreducible form of $\tilde{f}_l(\theta)$, it can be derived like this form similarly,

$$
\tilde{f}_l(\theta) = \frac{\zeta_{p-q}(z)\tilde{N}_l\zeta_{(p-q)*}(z)}{\zeta_{p-q}(z)D_l\zeta_{(p-q)*}(z)} \Big|_{z=e^{j\theta}}
\tag{5.37}
$$

where

$$\tilde{N}_l = N_l - G_l^{-1} W_l^H R V_{S_l} R_{S_l}^{-1} V_{S_l}^H R W_l G_l^{-1} \tag{5.38}$$

$$= \{\tilde{n}_{i,j}\}_{i,j=0}^{p-q}$$

Define the following polynomials

$$n_l(z) = n_0 + n_1 z + \cdots + n_{p-q} z^{p-q} \tag{5.39}$$

$$d_l(z) = d_0 + d_1 z + \cdots + d_{p-q} z^{p-q} \tag{5.40}$$

where

$$n_0 = \sum_{i=0}^{p-q} n_{i,i} \quad n_m = 2 \sum_{i=m}^{p-q} n_{i,i-m} \tag{5.41}$$

$$d_0 = \sum_{i=0}^{p-q} d_{i,i} \quad d_m = 2 \sum_{i=m}^{p-q} d_{i,i-m} \quad m = 1, 2 \ldots, p - q \tag{5.42}$$

and similarly define $\tilde{n}_l(z)$ and to calculate the matrix $\tilde{N}_l$ which realize the same function as $n_l(z)$.

Then we have the final form of $f_l(\theta)$ and $\tilde{f}_l(\theta)$ are shown like follows,

$$f_l(\theta) = \frac{\mathrm{Re}\{n_l(e^{j\theta})\}}{\mathrm{Re}\{d_l(e^{j\theta})\}}, \quad \tilde{f}_l(\theta) = \frac{\mathrm{Re}\{\tilde{n}_l(e^{j\theta})\}}{\mathrm{Re}\{d_l(e^{j\theta})\}} \tag{5.43}$$

where $\mathrm{Re}\{\}$ represents the real part of the complex value.

Therefore the irreducible form of efficient AM criterion (IAM) is shown like this.

$$\mathrm{tr}\{R_{NN}\} = \mathrm{tr}\{R_{N_l}\} - \frac{\mathrm{Re}\{n_l(e^{j\theta})\}}{\mathrm{Re}\{d_l(e^{j\theta})\}}, \tag{5.44}$$

$$\det\{R_{SS}\} = \det\{R_{S_l}\} \frac{\mathrm{Re}\{\tilde{n}_l(e^{j\theta})\}}{\mathrm{Re}\{d_l(e^{j\theta})\}}. \tag{5.45}$$
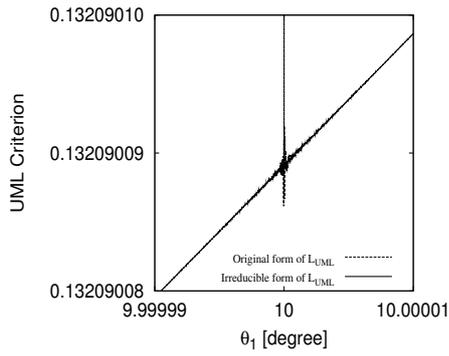
Fig. 6.3: Numerical instability in AM criterion and stability in IAM criterion.
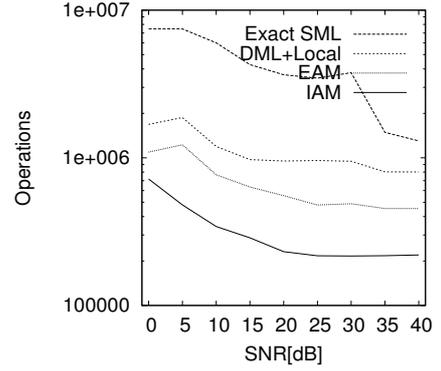
Fig. 6.4: Average amount of operations for calculating AM, EAM and IAM criterions.
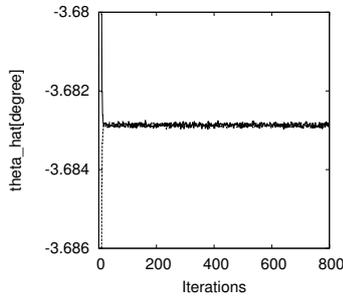
(a) The iteration does not stop until it reaches the maximum number, 800, using AM method.
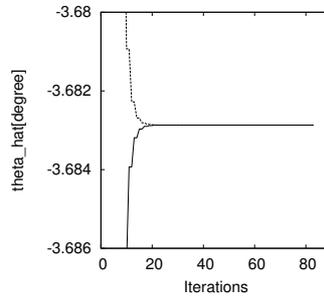
(b) The iteration stops at about 85 using IAM criterion.
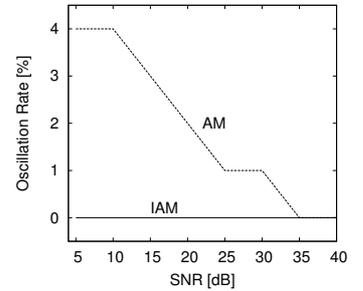
(c) Oscillation rate of AM and IAM criterions.

Fig. 6.5: Oscillation in AM criterion while not in IAM.

5.4.3 Validity of IAM Algorithm

In this section, we show some simulations to demonstrate that the IAM criterion is instability and more efficient than the AM and EAM criterions for the conventional SML estimation.

In Fig. 6.3, the scenario is $p = 3, q = 2, \eta = 2$, SNR=10, $M = 100$. The two true sources are located at 0 and 8 degree. The estimated bearing θ_2 is fixed at 10 degree, while θ_1 varies from 9.99999 to 10.00001. The dashed line represents the value of SML with AM criterion, while the solid line represents the IAM criterion. It shows clearly

that numerical instability occurs when the AM criterion is used. Especially, the value changes violently around the point $\theta_2 = \theta_1 = 10$ because it is indefinite. As for the IAM criterion, we can find that the value becomes monotonic and it is numerical stable.

Due to the numerical Instability, oscillation may happen as Fig.6.5 shows. The scenario is the same to Fig. 6.3. In Fig. 6.5 (a) and (b), the estimated two bearings θ_1 and θ_2 converge to an identical value represented by the solid line and the dashed line. The iteration is stopped when the variation of each bearing is less than 1e-05 or when the iteration reaches the maximum number, 800. When the AM criterion is used, the iteration does not stop until it reaches the maximum number. As for the IAM criterion, it converges well for less than 100 iterations. Fig. 6.5 (c) shows the oscillation rate of the two criterions in 100 independent trials. We can find that there is no oscillation when the IAM criterion is used.

At last, let us see the average amount of operations of exact SML, local search method, EAM and IMA criterions represented by "ExactSML", "DML+Local", "EAM" and "IAM" respectively in Fig. 6. 4. The scenario is the same to Fig. 6.3. It shows clearly that the IAM is the most efficient, while the exact SML is the worst.

5.5 Conclusions

In this chapter, we have proposed three algorithms to find the adequate solution of SML with less loss of resolution and low computational cost. First we proposed a local search method together with DML estimation as initialization to find the optimal or suboptimal solution of the SML. The criterion is based on the conventional SML because it is more easier to exploit efficient algorithms for its calculation. Simulation results are shown to proved that when DOA is resolvable, the proposed local search method can reach the

same solution of the exact SML. Then based on the local search method, an efficient version of AM algorithm, called EAM algorithm, is proposed. It is shown that the EAM algorithm is much efficient than the AM algorithm. However, one problem happens in calculation of conventional suing the AM or the EAM algorithms. That is the numerical instability happens when two parameters are going to have a same value. To solve this problem and reduce computational cost more, we proposed the irreducible form the EAM algorithm. We call it IAM algorithm. Simulation results are also shown that the IAM is numerical stable and more efficient than the EAM algorithm.

Chapter 6

Conclusions

In this paper, we address the issue of stochastic ML estimation of DOA. First we have the basic problem formulation of DOA and introduce the DML and SML estimation. Then the conventional SML estimation is shown. As some other literature also pointed out that the conventional SML estimation are formulated without considering an important condition that the estimated signal covariance matrix must be non-negative definite. Due to this reason, the following problems happens for the conventional SML estimation. 1) Solutions in the noise-free case are not unique. 2) Global solution in the noisy case becomes ambiguous occasionally. 3) There exist situations that any local solution does not satisfy that important condition. As a result, conventional SML estimation fails in DOA finding occasionally. More importantly, because the adequate solution of SML is not found, the ML likelihood function can not be evaluated correctly. This involves theoretical difficulty in applying the conventional SML to others, such as the MDL model selection.

To solve these problems, an exact formulation of SML is given. The exact SML can solve the problems above. It also shown that it has good resolution in the threshold region and it does not need the information of the rank of the signal covariance matrix. Furthermore, the exact SML has following advantages. First it provides the exact global solution of SML, which can become a standard for evaluating estimation accuracy of other methods, especially approximated SML methods. Second, the proposed exact SML method does not have any theoretic problem in applying the ML solution to other techniques such as the model selection technique based on MDL principle.

The only problem of exact SML is that it require high computational cost. To find

the adequate solution of SML with less loss of resolution and low computational cost. we proposed three efficient algorithms. First we proposed a local search method together with DML estimation as initialization to find the optimal or suboptimal solution of the SML. The criterion is based on the conventional SML because it is more easier to exploit efficient algorithms and the adequate solution of SML do exit in the local solutions of the conventional SML. Simulation results are shown to proved that when DOA is resolvable, the proposed local search method can reach the same solution of the exact SML. Then based on the local search method, an efficient version of AM algorithm, called EAM algorithm, is proposed. It is shown that the EAM algorithm is much efficient than the AM algorithm. However, one problem happens in calculation of conventional suing the AM or the EAM algorithm. That is the numerical instability happens when two parameters are going to have a same value. To solve this problem and reduce computational cost more, we proposed the irreducible form the EAM algorithm. We call it IAM algorithm. Simulation results are also shown that the IAM is numerical stable and more efficient than the EAM algorithm.

However, we have to note that until now the SML estimation can not be applied in practice because the computational cost is still very high. So the next step of my research is to exploit more efficient algorithms for the SML estimation. We have been considering the gradient method such as the Newton method for the SML estimation. Exciting results are expected.

Appendix A

Maximization of Likelihood Function with Respect to $\lambda_1, \lambda_2, \ldots \lambda_r$ and σ^2 When $l_r(\Theta, v_r) < \sigma^2$

First, we assume with no loss of generality that v_k for $k = 1, 2, \ldots, r$ are ordered so that $l_1(\Theta, v_1) \geq l_2(\Theta, v_2) \geq \ldots \geq l_r(\Theta, v_r)$.

When $l_r(\Theta, v_r) < \sigma^2$, the solution in (30) and (31) violate the contiotion of $\sigma^2 < \lambda_r$.

To solve the maximization of $L_r(\Theta, \lambda_1, \ldots, \lambda_r, v_1, \ldots, v_r, \sigma^2)$ with respect to λ_1, $\lambda_2, \ldots \lambda_r$ and σ^2 in this case, we consider the maximization of the function $h(\lambda)$ in (29) when the domain of λ is restricted to $\sigma^2 \leq \lambda$. The solution of this probles is abtained as follow.

- 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 $l_r(\Theta, v_r) < \sigma^2$, the maximization of $L_r(\Theta, \lambda_1, \ldots, \lambda_r, v_1, \ldots, v_r, \sigma^2)$ with respect to λ_r with the condition $\sigma^2 \leq \lambda_r$ results in $\lambda_r = \sigma^2$. The maximization problem with respect to $\lambda_1, \lambda_2, \ldots \lambda_r$ and σ^2 is changed into the maximization problem of

$$L_r(\Theta, \lambda_1, \ldots, \lambda_{r-1}, \sigma^2, v_1, \ldots, v_r, \sigma^2) \tag{A.1}$$

$$= L_{r-1}(\Theta, \lambda_1, \ldots, \lambda_{r-1}, v_1, \ldots, v_{r-1}, \sigma^2) \tag{A.2}$$

with respect to $\lambda_1, \lambda_2, \dots \lambda_{r-1}$ and σ^2. Therefore we have

$$\lambda_k = l_k(\Theta, \boldsymbol{v}_k) \quad \text{for } k = 1, 2, \dots, r-1 \tag{A.3}$$

$$\lambda_r = \sigma^2 \tag{A.4}$$

$$\sigma^2 = \sigma_{r-1}^2 = \frac{(p-r)\sigma_r^2 + l_r(\Theta, \boldsymbol{v}_r)}{p-r+1} \tag{A.5}$$

where the condition C2') is relaxed as

C2') $\lambda_1 \geq \dots \geq \lambda_r \geq \sigma^2$.

Appendix B

Maximization of Likelihood Function with Respect to $v_1, v_2, ...\ v_r$ in the Case of $l_1(\Theta, v_1)$ $= l_2(\Theta, v_2) = ...\ l_\mu(\Theta, v_\mu)$

In this appendix, we use the notation that l_k instead of $l_k(\Theta, v_k)$. Let $v_1, v_2, ..., v_r$ be the solutions which maximize $L_r(\Theta, v_1, ..., v_r)$ in (29) 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

$$L_r(\Theta, v_1, v_2, v_3, ..., v_r) = -M \ln \left\{ (l_0^\mu l_{\mu+1}...l_r \left(\sigma_r^2\right)^{p-r} \right\}$$

$$\sigma_r^2 = \frac{1}{p-r} \left(\text{tr}\{\tilde{R}\} - \left(\mu l_0 + \sum_{k=\mu+1}^{r} l_k \right) \right). \tag{B.1}$$

The equations (4.10) and (4.11) are reduced as follows.

$$\tilde{R}_{SS}(\Theta)v_k = \alpha'_{k1}v_1 + \alpha'_{k2}v_2 + ... + \alpha'_{k\mu}v_\mu \tag{B.2}$$

$$\alpha'_{kk} = v_k^H \tilde{R}_{SS}(\Theta)v_k = l_0 \tag{B.3}$$

$$\alpha'_{ki} = v_i^H \tilde{R}_{SS}(\Theta)v_k = \bar{\alpha}'_{ik} \tag{B.4}$$

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

Define the following 2×2 matrix A_{12}

$$\begin{bmatrix} v_1^H \\ v_2^H \end{bmatrix} \tilde{R}_{SS}(\Theta)[v_1 \ v_2] = \begin{bmatrix} l_0 & \alpha'_{12} \\ \bar{\alpha}'_{12} & l_0 \end{bmatrix} = A_{12}. \tag{B.5}$$

The matrix A_{12} has two eigenvalues, i.e., $l'_1 = l_0 - |\alpha_{12}|$ and $l'_2 = l_0 + |\alpha_{12}|$. Let e_1 and e_2 be the unit eigenvectors corresponding to l'_1 and l'_2, respectively. Define

$$v'_1 = [v_1 \ v_2]e_1 \text{ and } v'_2 = [v_1 \ v_2]e_2. \tag{B.6}$$

57

Then we have

$$\boldsymbol{v}_1^{\prime H} \tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_1^{\prime} = \boldsymbol{e}_1^H \tilde{\boldsymbol{A}}_{12}\boldsymbol{e}_1 = l_1^{\prime} = l_0 - |\alpha_{12}|, \tag{B.7}$$

$$\boldsymbol{v}_2^{\prime H} \tilde{\boldsymbol{R}}_{SS}(\Theta)\boldsymbol{v}_2^{\prime} = \boldsymbol{e}_2^H \tilde{\boldsymbol{A}}_{12}\boldsymbol{e}_2 = l_2^{\prime} = l_0 + |\alpha_{12}|. \tag{B.8}$$

If it holds that $l_1^{\prime} \geq \sigma_r^2$, then we readily obtain the following inequality.

$$
\begin{aligned}
&L_r(\Theta, \boldsymbol{v}_1^{\prime}, \boldsymbol{v}_2^{\prime}, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) \\
&= -M \ln \left\{ (l_0^2 - |\alpha_{12}^{\prime}|^2) l_0^{\mu-2} l_{\mu+1}...l_r \left(\sigma_r^2\right)^{p-r} \right\} \\
&\geq L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r).
\end{aligned}
\tag{B.9}
$$

The equal sign in (B.9) holds iff $|\alpha_{12}^{\prime}| = 0$. If $|\alpha_{12}^{\prime}| \neq 0$, the inequality in (B.9) contracts the assumption that $\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_r$ are solution that maximize $L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r)$. Therefore we have $|\alpha_{12}^{\prime}| = 0$.

Although $l_2^{\prime}$ is greater than σ_r^2, $l_1^{\prime}$ may not. In the case that $l_1^{\prime} \leq \sigma_r^2$, using the result described in appendix A, the log-likelihood function is obtained as follows.

$$
\begin{aligned}
L_r(\Theta, \boldsymbol{v}_1^{\prime}, \boldsymbol{v}_2^{\prime}, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) &= L_{r-1}(\Theta, \boldsymbol{v}_2^{\prime}, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) \\
&= -M \ln \left\{ (l_0 + |\alpha_{12}^{\prime}|) l_0^{\mu-2} l_{\mu+1}...l_r \left(\sigma_{r-1}^{\prime 2}\right)^{p-r+1} \right\}
\end{aligned}
\tag{B.10}
$$

where

$$\sigma_{r-1}^{\prime 2} = \frac{1}{p - r + 1} \left((p - r)\sigma_r^2 + l_1^{\prime} \right). \tag{B.11}$$

Because $l_1^{\prime}$ is written in the Hermitian form of the non-negative definite matrix $\tilde{\boldsymbol{R}}_{SS}(\Theta)$ as shown in (B.7), $l_1^{\prime}$ has a non-negative real value. Under the conditions $0 \leq l_1^{\prime} \leq \sigma_r^2$ in addition to $\sigma_r^2 < l_0$, we can derive the following inequality

$$L_{r-1}(\Theta, \boldsymbol{v}_2^{\prime}, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) > L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r). \tag{B.12}$$

The inequality in (B.12) is proved in appendix C.

This conflicts with the assumption that v_1, v_2, ..., v_r are the solutions which maximize L_r. Therefore $l_1' \geq \sigma_r^2$ and $|\alpha_{12}|$ vanishes.

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

$$\tilde{R}_{SS}(\Theta)v_k = l_0 v_k \qquad \text{for } k = 1, 2, ..., \mu. \tag{B.13}$$

Therefore l_0 becomes an eigenvalue of $\tilde{R}_{SS}(\Theta)$ with μ multiplicity and v_1, v_2 ... v_μ are the corresponding unit eigenvectors orthogonal each other.

Appendix C

Proof of B.12

From (B.12) we have

$$
\begin{aligned}
L_{r-1}(\Theta, \boldsymbol{v}_2', \boldsymbol{v}_3, ..., \boldsymbol{v}_r) &- L_r(\Theta, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, ..., \boldsymbol{v}_r) \\
&= N \ln \left\{ \frac{l_0^\mu l_{\mu+1}...l_r \left(\sigma_r^2\right)^{p-r}}{(l_0 + |\alpha_{12}'|)l_0^{\mu-2}l_{\mu+1}...l_r \left(\sigma_{r-1}'^2\right)^{p-r+1}} \right\} \\
&= N \ln \left\{ \frac{l_0}{(l_0 + |\alpha_{12}'|} \frac{l_0}{\sigma_{r-1}'^2} \left(\frac{\sigma_r^2}{\sigma_{r-1}'^2}\right)^{p-r} \right\}.
\end{aligned}
\tag{C.1}
$$

We have the followings from above,

$$
\sigma_r^2 = \frac{1}{p-r} \left((p-r+1)\sigma_{r-1}'^2 - (l_0 - |\alpha_{12}'|) \right).
\tag{C.2}
$$

Define

$$
\alpha = \frac{l_0 - |\alpha_{12}'|}{\sigma_{r-1}'^2} \qquad \beta = \frac{l_0 - |\alpha_{12}'|}{l_0}.
\tag{C.3}
$$

Then it follows from the relation $l_0 > \sigma_r^2 \geq \sigma_{r-1}'^2 \geq l_0 - |\alpha_{12}'| \geq 0$, that

$$
0 \leq \alpha \leq 1 \qquad 0 \leq \beta < 1
\tag{C.4}
$$

and

$$
\alpha = \frac{l_0}{\sigma_{r-1}'^2}\beta \qquad \frac{l_0}{\sigma_{r-1}'^2} > 1
\tag{C.5}
$$

$$
\alpha > \beta.
\tag{C.6}
$$

Using these definitions, we have

$$
l_0 + |\alpha_{12}'| = l_0 + l_0 - \beta l_0 = (2 - \beta)l_0,
\tag{C.7}
$$

$$
\sigma_{r-1}'^2 = \beta l_0 / \alpha,
\tag{C.8}
$$

$$
\sigma_r'^2 = \frac{1}{p-r}(p-r+1-\alpha)\sigma_{r-1}'^2.
\tag{C.9}
$$

Then substitute these equations into (C.1) and because of $0 \leq \alpha \leq 1$, $(p - r) \geq 1$, we have

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

$$= N \ln \left\{ \frac{\alpha}{(2-\beta)\beta}(1 + (1-\alpha)/(p-r))^{p-r} \right\}$$

$$\geq N \ln \left\{ \frac{\alpha}{(2-\beta)\beta}(1 + (1-\alpha)) \right\}$$

$$= N \ln \left\{ \frac{\alpha(2-\alpha)}{\beta(2-\beta)} \right\}. \tag{C.10}$$

From the inequality in (C.6), we have

$$\frac{\alpha(2-\alpha)}{\beta(2-\beta)} > 1 \tag{C.11}$$

Appendix D

Uniqueness of the DML Estimation in Noise-Free Case

In assumption A5), we assume that $q < 2\eta p/(2\eta + 1)$ and $M \geq \eta$. These conditions guarantee that a unique solution of DOA exists in the noise-free case [22]. In other words, the following equation has a unique solution of Θ to the observation data X when p, q, M and η satisfy the assumption A5).

$$X = A(\Theta)S \tag{D.1}$$

On the other hand, in the noise-free case, the solution $\hat{\Theta}$ of DML estimator obviously makes the DML estimator to the minimum value 0, i.e.,

$$L_{\text{DML}}(\hat{\Theta}) = \text{tr}\{\boldsymbol{R}_{NN}(\hat{\Theta})\} = 0. \tag{D.2}$$

Next, we demonstrate that the solution $\hat{\Theta}$ of equation (D.1) is identical to the solution of equation (D.2).

$Proof$: Assumption A1) guarantees that the matrix $A(\Theta)$ has full rank. Since it holds

$$X = A(\Theta)S, \tag{D.3}$$

then we have

$$A^H(\Theta)X = A^H(\Theta)A(\Theta)S. \tag{D.4}$$

And since $A(\Theta)$ is not singular, we get

$$S = (A^H(\Theta)A(\Theta))^{-1}A^H(\Theta)X. \tag{D.5}$$

Substituting the equation (D.5) into equation (D.1), we have

$$X = A(\Theta)(A^H(\Theta)A(\Theta))^{-1}A^H(\Theta)X. \tag{D.6}$$

Here, define

$$\boldsymbol{P}_{A(\Theta)} = A(\Theta)(A^H(\Theta)A(\Theta))^{-1}A^H(\Theta), \tag{D.7}$$

where $\boldsymbol{P}_{A(\Theta)}$ is the projection matrix onto the signal subspace spanned by $A(\Theta)$. From the definition of $\boldsymbol{V}_S(\Theta)$ and $\boldsymbol{V}_N(\Theta)$, we have

$$\boldsymbol{P}_{A(\Theta)} = \boldsymbol{V}_S(\Theta)\boldsymbol{V}_S^H(\Theta), \tag{D.8}$$

$$\boldsymbol{P}_{A(\Theta)}^\perp = \boldsymbol{I} - \boldsymbol{P}_{A(\Theta)} = \boldsymbol{V}_N(\Theta)\boldsymbol{V}_N^H(\Theta), \tag{D.9}$$

where $\boldsymbol{P}_{A(\Theta)}^\perp$ is the projection matrix onto the noise subspace which is the orthogonal complement of the signal subspace. Hence, the solution $\hat{\Theta}$ of equation (D.1) is identical to the solution of the following equation.

$$X = \boldsymbol{P}_{A(\Theta)}X. \tag{D.10}$$

Then, we have

$$(\boldsymbol{I} - \boldsymbol{P}_{A(\Theta)})X = \boldsymbol{P}_{A(\Theta)}^\perp X = 0. \tag{D.11}$$

On the other hand, from the definition of $\boldsymbol{R}_{NN}(\hat{\Theta})$, it is obvious that equation (D.2) is identical to the following equation.

$$\boldsymbol{V}_N^H(\hat{\Theta})X = 0 \tag{D.12}$$

Multiply $\boldsymbol{V}_N(\hat{\Theta})$ in each side of equation (2.31), then we have

$$\boldsymbol{V}_N(\hat{\Theta})\boldsymbol{V}_N^H(\hat{\Theta})X = \boldsymbol{P}_{A(\hat{\Theta})}^\perp X = 0, \tag{D.13}$$

which is the same as equation (D.11).

Therefore, the solution $\hat{\Theta}$ of equation (D.1) is identical to the solution of equation (D.2). it demonstrates that the DML estimation has a unique solution in the noise-free case under the assumption A5).

BIBLIOGRAPHY

[1] Y. Bresler and A. Macovski, "exact maximum exactlikelihood 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, Jul. 1990.

[4] J. C. Pei and J. F. Bohme, "Recursive EM and SAGE-Inspired Algorithms With Application to DOA Estimation" IEEE Trans. on Signal Processing, VoL. 53, No. 8, PP. 2664-2677, Aug. 2005.

[5] Y. C. Chih and Y. T. Su, "Maximum likelihood DOA Estimation Based on the Cross-Entropy Method," Proc. ISIT 2006, pp. 851-855, 2006.

[6] H. Wang, S. Kay and S. Saha, "An Importance Sampling Maximum Likelihood Direction of Arrival Estimator" IEEE Trans. on Signal Processing, VoL. 56, No. 10, PP. 5082-5092, Oct. 2008.

[7] P. Stoica and A. Nehorai, "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.

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

[9] 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.

[10] J. C. Pei and J. F. Bohme, "Recursive EM algorithm for stochastic ML DOA estimation," Proc. ICASSP 2002, III-3029 - III-3032, 2002.

[11] Y. Bresler, "Maximum likelihood Estimation of a Linearly Structured Covariance with Application to Antenna Array Processing," Porc. 4th ASSP Workshop on Spectrum Estimation and Modeling, minnessota, pp. 172-175, Aug. 1988.

[12] H. Chen and M. Suzuki, "An Algorithm for Finding the Best Solution of Stochastic ML Estimation of DOA," 7th International Symposium on Communications and Information Technologies (ISCIT 2007) F3C-5, pp.1060-1065, Oct. 2007.

[13] M. Suzuki and H. Chen, "exact Unconditional ML Estimation of DOA," 2009 APSIPA Annual Summit and Conference APSIPA ASC 2009, WA-L3-4, pp.876-882, Oct. 2009 (Sapporo).

[14] P. Stoica, B. Otterstern, M. Viberg and R. L. Moses, "Maximum likelihood Array Processing for Stochastic Coherent Sources," IEEE Trans. on Signal Processing, VoL. 44, No. 1, PP. 96-105, Jan. 1996.

[15] 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.

[16] A. J. Weiss and B. Friedlander, "On the Cramer-Rao bound for direction finding of correlated signals," IEEE Trans. Signal Processing, Vol. SP-41, No. 1, pp. 495-499, Jan. 1993.

[17] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters Via Rotational Invariance Techniques," IEEE Trans. Acoust., Speech. Signal Processing, vol. 37, no 7, pp. 984-995, Jul. 1989.

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

[19] B. D. RAO and K. V. S. Hari, "Performance Analysis of Root-Music," IEEE Trans. Acoust., Speech. Signal Processing, vol. 37, no 12, pp. 1939-1949, Dec. 1989.

[20] 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.

[21] 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, Apr. 1997.

[22] 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.

66

[23] 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.

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

ACKNOWLEDGEMENTS

The study has been accomplished while the author is in master and doctor courses of Computer Sciences, Graduate School of Engineering, Kitami Institute of Technology, at the research of Array Signal Processing since Oct. 2005.

I would like to acknowledge so many people for helping me during my study in Kitami Institute of Technology.

First and foremost I would like to express my great appreciation to my supervisor, Professor Masakiyo Suzuki, for his valuable supervision and advice on study. His words of encouragement, quiet urgings and careful amending of all of my writings will never be forgotten. Throughout my study in Kitami he encouraged me to develop independent thinking and research skills. All these will become my great wealth during all my life.

Next I would like to thank Professor Eisaka. Thanks for helping me dealing with the procedure for my Ph.D.

I extend thanks to my colleagues and friends. I have received so much kind help not only in study but also aboard the studying life.

Finally, I'd like to thank my family. Without their supports and encouragement, I could not accomplish the master and doctor course in Japan. The most important is that thank them for teaching me how to make a life and to be a person in the society.

LIST OF PAPERS

Journal Paper:

[1] H. Chen and M. Suzuki, "Exact Formulation for Stochastic ML Estimation of DOA," IEICE Trans. Fundamentals, vol. E93-A, no. 11, pp. 2141-2152, Nov. 2010.

International Conference:

[1] H. Chen and T. Teng, M. Zhang and M. Suzuki, "An Efficient Algorithm for Detection and Localization of Multiple Sources via the Stochastic Model," in Proceedings of International Workshop on Modern Science and Technology 2006 (IWMST 2006), pp. 298-304 (May 2006, Wuhan).

[2] H. Chen and M. Suzuki, "Non-Uniqueness of Unconditional ML Estimation of DOA and Its Best Solution," in Proceedings of 2007 International Conference on Telecommunications, Industry and Regulatory Development (ICTIR'07), pp. 6-11 (Aug. 2007, Bangkok).

[3] H. Chen and M. Suzuki, "An Algorithm for Finding the Best Solution of Stochastic ML Estimation of DOA," in Proceedings of 7th International Symposium on Communications and Information Technologies (ISCIT 2007) F3C-5, pp.1060-1065 (Oct. 2007 Sydney).

[4] H. Chen and M. Suzuki, "Properties of Criterion for Unconditional ML Estimation of DOA," in Proceedings of 2007 International Workshop on Smart Info-Media Systems in Bangkok (SISB 2007), pp. 1-6 (Nov. 2007 Bangkok).

[5] H. Chen and M. Suzuki, "On Non-Uniqueness of Stochastic ML Estimation of DOA and Its Best Solution," in Proceedings of 2007 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2007), pp.169-172 (Nov-Dec. 2007 Xiamen).

[6] H. Chen, and M. Suzuki, "Efficient Algorithms for DOA Estimation of Incoherent

Signals Based on Unconditional ML Criterion," in Proceedings of International Workshop on Modern Science and Technology 2008 (IWMST 2008), F032, pp. 399-408 (Nov. 2008, Harbin).

[7] H. Chen, and M. Suzuki, "Efficient algorithms for optimal and suboptimal unconditional ML estimation of DOA" in Proceedings of 2009 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS 2009), pp. 469-472 (Jan. 2009 Kanazawa).

[8] T. Teng, H. Chen, M. Zhang and M. Suzuki, "Gradient Method for ML and WSF Bearing Estimator Using a Uniform Linear Array," in Proceedings of International Workshop on Modern Science and Technology 2006 (IWMST 2006), pp.403-409 (May 2006, Wuhan).

[9] M. Zhang, H. Chen, T. Teng and M. Suzuki, "A Design of Information Theoretic Criteria for Detecting the Number of Incoherent Signals," in Proceedings of International Workshop on Modern Science and Technology 2006 (IWMST 2006), pp. 358-365 (May 2006, Wuhan).

[10] M. Suzuki, M. Zhang, H. Chen and T. Teng, "A Design Method of Criteria for Detecting the Number of Signals," in Proceedings of International Symposium on Communications And Information Technologies 2006 (ISCIT 2006), F1A-1 (Oct. 2006, Bangkok).

[11] M. Suzuki, M. Zhang, H. Chen and T. Teng, "A Design of Information Theoretic Criteria for Detecting the Number of Incoherent Signals," in Proceedings of 2006 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2006), FAM2-4-2, pp. 991-994 (Dec. 2006, Yonago).

[12] M. Suzuki and H. Chen, "Exact Unconditional ML Estimation of DOA" in Proceedings of 2008 International Workshop on Smart Info-Media Systems in Bangkok (SISB 2008), pp. 41-46 (Dec. 2008 Bangkok).

IEICE Joint Convention Record:

[1] H. Chen, T. Teng, M. Zhang and M. Suzuki, "An Efficient Iterative Technique for ML Estimation of DOA Based on Stochastic Signal Model," 2006 Joint Convention Record, The Hokkaido Chapters of the Institutes of Electrical and Information Engineers, Japan, 182 (Oct. 2006, Muroran).

[2] H. Chen and M. Suzuki, "Non-Uniqueness of SML Estimation of DOA and an algorithm for Its Best Solution," 2007 Joint Convention Record, The Hokkaido Chapters of the Institutes of Electrical and Information Engineers, Japan, 12, Hokkaido University (Oct. 2007 Sapporo).

[3] H. Chen and M. Suzuki, "Exact Unconditional ML Criterion for DOA Estimation," 2008 Joint Convention Record, The Hokkaido Chapters of the Institutes of Electrical and Information Engineers, Japan, 46, Hokkaido University (Oct. 2008 Sapporo).

[4] H. Chen and M. Suzuki, "Local Search method in Conventional Criterion of UML Estimation of DOA," 2009 Joint Convention Record, The Hokkaido Chapters of the Institutes of Electrical and Information Engineers, Japan, 13, Kitami Institute of Technology (Oct. 2009 Kitami).

[5] H. Chen, Y. Maeda, and M. Suzuki, "Performance Study on Exact Stochastic ML Estimation of DOA," 2010 Joint Convention Record, The Hokkaido Chapters of the Institutes of Electrical and Information Engineers, Japan, 11, Kitami Institute of Technology (Oct. 2010 Sapporo).

Local Conference and Symposium in Japan:

[1] H. Chen, T. Teng, M. Zhang and M. Suzuki, "An Efficient Alternating Minimization Algorithm for Signal Number Detection Based on the Stochastic Signal Model," IEICE Technical Report, Vol. 106, No. 112, CAS2006-18, pp. 61-65 (June 2006, Kitami).

[2] H. Chen and M. Suzuki, "Convergence of AM Algorithm for ML Estimator of DOA

Based on Stochastic Signal Model," in Proc. the 20th Signal Processing Symposium, A8-1 (Nov, 2006, Kyoto).

[3] H. Chen and M. Suzuki, "Properties of Criteria for ML Estimation of DOA Based on Stochastic Signal Model" in Proc. The 29th Symposiumon Information Theory and Its Applications (SITA 2006), 11.3, pp.165-168 (Nov. 2006, Hakodate).

[4] H. Chen and M. Suzuki, "Uniqueness of ML Estimation of DOA Based on Stochastic Signal Model," IEICE Technical Report, Vol. 107, No. CAS-1, VLD-1, SIP-1, pp.91-96 (June 2007, Sapporo).

[5] H. Chen and M. Suzuki, "Global and Local Solutions to Stochastic ML Estimation of DOA," IEICE Tech. Report, Vol. 107, No. SIP-235, SIS-237, SP-239, pp. 17-22 (Sep. 2007 Nagoya).

[6] H. Chen, M. Suzuki, "Ambiguity in Exact Unconditional ML Criterion for DOA Estimation," IEICE Technical Report, vol. IEICE-108, no. IEICE-EA-68, IEICE-SIP-70, IEICE-EA2008-6, IEICE-SIP2008-15, pp. 1-6, (2008-5, Kannsai).

[7] H. Chen, M. Suzuki, "Avoidance of Numerical Unstability in Exact Unconditional ML Estimation of DOA," IEICE Technical Report, vol. IEICE-108, no. IEICE-CAS-105, IEICE-VLD-107, IEICE-SIP-109, IEICE-CAS2008-29, IEICE-VLD2008-42, IEICE-SIP2008-63, pp. 61-66, (2008-6, Sapporo).

[8] H. Chen, M. Suzuki, "Irreducible Form of AM Criterion for Unconditional ML Estimation of DOA," IEICE Technical Report, vol. IEICE-108, no. IEICE-SIS-85, IEICE-SIS2008-7, pp. 35-40, (2008-6, Asahigawa).

[9] M. Suzuki, T. Teng, M. Zhang and H. Chen, "A Criterion for Bearing Estimation Using a Sensor Array," IEICE Technical Meeting, Vol.105, No.427, SIS2005-54, pp. 7-10 (Nov. 2005, Hitachi).

[10] T. Teng, H. Chen, M. Zhang and M. Suzuki, "An Exact Newton Method for DOA Estimation Using a Uniform Linear Array," IEICE Technical Report, Vol. 106, No. 112,