SOURCE DETECTION IN THE PRESENCE OF NONUNIFORM NOISE

Saïd Aouada, Abdelhak M. Zoubir

Signal Processing Group
Darmstadt University of Technology
Merckstrasse 25, D-64283, Darmstadt, Germany

Chong Meng Samson See

Center for Communications Systems
DSO National Labs
20 Science Park Drive, Singapore 118230

ABSTRACT

We consider the problem of source number estimation in the presence of unknown spatially nonuniform noise. Successive array element suppression is applied to isolate the contribution of the noise powers and a likelihood function is derived. When it is combined with the appropriately defined penalty function, an MDL-like criterion is defined. Performance of the new criterion is assessed through simulations and it is shown that the method is powerful in a nonuniform noise environment.

1. INTRODUCTION

Source number detection is a critical problem in array signal processing as it is prerequisite for signal parameter estimation algorithms. It finds applications in wireless communications, RADAR, SONAR and seismic exploration [1, 2]. Several detection schemes with a number of variants have been continuously proposed and analyzed in the literature, ranging from methods based on hypothesis testing [3, 4] to information theoretic criteria [5, 6]. Most detectors apply for different assumptions on the data, especially with respect to the correlation among the sources and/or the noise. In all cases, the noise is always assumed to be spatially uniform. This uniformity makes it possible to fully exploit the information embedded in the eigenvalues of the covariance matrix of data and the detectors use test statistics or goodness-of-fit terms which are functions of the eigenvalues. However, when the noise is not uniform, i.e., when the noise powers are different from one sensor to another, most of the detectors fail to perform satisfactorily and few dedicated detection techniques are available. A number of instances where nonuniform noise occurs are mentioned in [7] and a model highlighting its structure was addressed for the problem of estimation where the number of sources is known in advance. A more adapted information theoretic criterion based on Gerschgorin's theorem was proposed in [6] where the detection is not solely based on the eigenvalues. Although it is suggested the the scheme can be applied to nonuniform noise, it can be shown that the same limitations apply due to the misleading contribution of the unordered eigenvalues, resulting in erroneous detection.

In what follows, we propose an alternative detection criterion by deriving a new likelihood function. The method copes with the spatial non-uniformity of the noise by successively eliminating the contribution of single elements from the array. When associated with the appropriate penalty function, the derived likelihood function results in an MDL criterion. The proposed Non-Uniform MDL (NU-MDL) is suitable for both nonuniform noise and the special case of uniform noise.

2. DATA MODEL

Consider an array of M sensors receiving p narrow-band signals from coplanar sources with unknown DOAs, $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_P]^T$, where $(.)^T$ stands for matrix transpose. The number of sources p is to be estimated. The received signal vector at instant i can be modeled as [7, 8]

$$\mathbf{x}(i) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(i) + \mathbf{n}(i), \quad i = 1, \ldots, L \tag{1}$$

where

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_p)] \tag{2}$$

is the $(M \times p)$-dimensional steering matrix, $\mathbf{a}(\theta_q); q = 1, \ldots, p$ are the vectors of the array response to the signal directions, $\mathbf{s}(i)$ is the p-dimensional vector of the source signals and $\mathbf{n}(i)$ is the M-dimensional vector of white sensor noise.

Sensor noise is assumed to be a zero-mean spatially and temporally white Gaussian process with an unknown covariance matrix $\mathbf{Q}$ satisfying the following structure

$$\mathbf{Q} = \mathsf{E}\left\{\mathbf{n}(i)\mathbf{n}^H(i)\right\} = \text{diag}\{\mathbf{q}\} \tag{3}$$

where $(.)^H$ denotes matrix Hermitian transpose, $\mathsf{E}(.)$ stands for expectation, $\mathbf{q} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2]^T$ and diag $\{\mathbf{q}\}$ is a diagonal matrix with elements elements of the vector $\mathbf{q}$.

The source signals and the noise are assumed to be uncorrelated. The array covariance matrix is therefore given by

$$\mathbf{R} = \mathsf{E}\left\{\mathbf{x}(i)\mathbf{x}^H(i)\right\} = \mathbf{A}(\boldsymbol{\theta})\mathbf{R_s}\mathbf{A}^H(\boldsymbol{\theta}) + \mathbf{Q} \tag{4}$$

where $\mathbf{R_s} = \mathsf{E}\{\mathbf{s}(i)\mathbf{s}^H(i)\}$ is the source signal covariance matrix. The received signal waveforms are assumed to be a random zero-mean Gaussian process [8, 7], satisfying $\mathbf{x}(i) \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$.

3. DETECTION SCHEME

3.1. Covariance Matrix Transformation

Similarly to [6], we introduce a unitary covariance matrix transformation based on array element suppression. For simplicity and without loss of generality, we discard the M-th element of the array. The resulting $(M-1) \times p$ dimensional steering matrix is therefore

$$\mathbf{A}_M(\boldsymbol{\theta}) = [\mathbf{a}_M(\theta_1), \mathbf{a}_M(\theta_2), \ldots, \mathbf{a}_M(\theta_p)] \tag{5}$$

where the vectors $\mathbf{a}_M(\theta_p)$ are the same as in (2) with the M-th element removed. Similarly to (4), the covariance matrix of the collected data over the reduced $(M-1)$-element array is given by

$$\mathbf{R}_M = \mathbf{A}_M(\boldsymbol{\theta})\mathbf{R_s}\mathbf{A}_M^H(\boldsymbol{\theta}) + \mathbf{Q}_M \tag{6}$$

where the reduced noise covariance matrix $\mathbf{Q}_M$ is defined as

$$\mathbf{Q}_M = \text{diag}\{\mathbf{q}_M\} \tag{7}$$

with $\mathbf{q}_M = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_{M-1}^2]^T$. Note that

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_M & \mathbf{r} \\ \mathbf{r}^H & r_{MM} \end{bmatrix} \tag{8}$$

where r_{MM} is the (M,M)-th element of $\mathbf{R}$ and

$$\mathbf{r} = \mathbf{A}_M(\boldsymbol{\theta})\mathbf{R_s}\mathbf{b}_M^H \tag{9}$$

with $\mathbf{b}_M$ being the removed M-th row of $\mathbf{A}(\boldsymbol{\theta})$.

The reduced covariance matrix $\mathbf{R}_M$ has the following eigendecomposition

$$\mathbf{R}_M = \mathbf{E}\mathbf{D}\mathbf{E}^H \tag{10}$$

with

$$\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_{M-1}] \tag{11}$$

$$\mathbf{D} = \text{diag}\{\lambda_1, \ldots, \lambda_{M-1}\} \tag{12}$$

where λ_m and $\mathbf{e}_m$, $m = 1, \ldots, M-1$, are the eigenvalues and their corresponding eigenvectors, respectively.

A unitary matrix $\mathbf{U}$ is defined as follows

$$\mathbf{U} = \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{13}$$

where $\mathbf{0}$ is an $(M-1)$-dimensional vector of zero elements and $\mathbf{E}$ is defined in (12). Applying transformation $\mathbf{U}$ to the covariance matrix $\mathbf{R}$ leads to

$$\boldsymbol{\mathcal{R}} = \mathbf{U}^H\mathbf{R}\mathbf{U} \tag{14}$$

$$= \begin{bmatrix} \mathbf{E}^H\mathbf{R}_M\mathbf{E} & \mathbf{E}^H\mathbf{r} \\ \mathbf{r}^H\mathbf{E} & r_{MM} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{D} & \mathbf{c} \\ \mathbf{c}^H & r_{MM} \end{bmatrix} \tag{15}$$

Let $|c_1| \geq |c_2| \geq \ldots \geq |c_{M-1}|$ be the magnitudes of the elements of vector $\mathbf{c}$ in (15). From (9), note that the m-th element c_m, has the following structure

$$c_m = \mathbf{e}_m^H\mathbf{A}_M(\boldsymbol{\theta})\mathbf{R_s}\mathbf{b}_M^H \tag{16}$$

Moreover, due to the fact that the noise subspace is orthogonal to the direction matrix $\mathbf{A}_M(\boldsymbol{\theta})$, the elements c_m satisfy the following

$$c_m \begin{cases} = 0, & \text{if } \mathbf{e}_m \text{ is a noise eigenvalue,} \\ \neq 0, & \text{if } \mathbf{e}_m \text{ is a signal eigenvalue.} \end{cases} \tag{17}$$

Thus, based on the information contained in the elements c_m, $m = 1, \ldots, M-1$, it is possible to separate the noise and signal subspaces. In other words,

$$|c_1| \geq |c_2| \geq \ldots \geq |c_p| \geq |c_{p+1}| = |c_{p+2}| = \ldots = |c_{M-1}| = 0 \tag{18}$$

3.2. Geometric Interpretation

The elements c_m of relation (18) can be interpreted as the projection of the M-th column of $\mathbf{R}$ onto the m-th eigenvector, $\mathbf{e}_m$, of $\mathbf{R}_M$. Also, from Gerschgorin's theorem [10], the first $M-1$ eigenvalues of the transformed covariance matrix $\boldsymbol{\mathcal{R}}$ are the centers of the corresponding Gerschgorin disks, whose radii ρ_m, $m = 1, \ldots, M-1$, are given by the magnitude of the corresponding elements c_m, i.e., $\rho_m = |c_m|$. The value of these radii indicates the multiplicity of the eigenvalues and the subspaces that their eigenvectors span [6, 10]. It is clear from (14) that two distinct subsets of disks are easily identifiable, representing the signal subspace for the first p radii, $|c_1|, \ldots, |c_p|$, and the noise subspace for the smallest and equal $M-1-p$ radii, $|c_{p+1}|, \ldots, |c_{M-1}|$.

3.3. Source Number Estimation

As the noise is not spatially uniform, it is not possible to order the eigenvalues of the covariance matrix of the data for source detection. Instead we use the information provided by the elements of vector $\mathbf{c}$ as it is defined in (15).

Under the Gaussianity assumption of the data, it is easily shown that the stochastic negative Log-Likelihood (LL) function of the observed data is [9]

$$\mathcal{L}(\boldsymbol{\eta}) = L \ln\{\det[\mathbf{R}(\boldsymbol{\eta})]\} + \text{trace}\left\{\mathbf{R}^{-1}(\boldsymbol{\eta})\hat{\mathbf{R}}\right\} \tag{19}$$

where $\boldsymbol{\eta} = [p, \boldsymbol{\theta}^T, \mathbf{p}^T, \mathbf{q}^T]^T$ is the vector of unknown parameters, with $\mathbf{p} = [\Re(\mathbf{R_s}), \Im(\mathbf{R_s})]$ and $\hat{\mathbf{R}}$ is the sample covariance matrix of the data, given by

$$\hat{\mathbf{R}} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{x}(i)\mathbf{x}^H(i) \tag{20}$$

Recalling that the transformation matrix $\mathbf{U}$ is unitary, i.e., $\mathbf{U}\mathbf{U}^H = \mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix, we can replace $\mathbf{R}$ in (19) by the transformed covariance matrix $\boldsymbol{\mathcal{R}}$ defined in (14), and obtain the following modified negative LL cost function

$$\mathcal{L}(p) = L \ln\{\det[\boldsymbol{\mathcal{R}}(p)]\} + \text{trace}\left\{\boldsymbol{\mathcal{R}}^{-1}(p)\hat{\boldsymbol{\mathcal{R}}}\right\} \tag{21}$$

where $\hat{\boldsymbol{\mathcal{R}}} = \mathbf{U}\hat{\mathbf{R}}\mathbf{U}^H$.

From (15), using properties of (2×2)-block matrices, it is straight forward to show that

$$\det\{\boldsymbol{\mathcal{R}}(p)\} = \det\{\mathbf{D}\}\det\left\{r_{MM} - \mathbf{c}^H\mathbf{D}^{-1}\mathbf{c}\right\} \tag{22}$$

Exploiting the fact that matrix $\mathbf{D}$ is diagonal, and from (12) and (18), the above equation becomes

$$\det\{\boldsymbol{\mathcal{R}}(p)\} = \left(\prod_{m=1}^{M-1}\lambda_m\right)\left(r_{MM} - \sum_{m=1}^{p}\frac{|c_m|^2}{\lambda_m}\right) \tag{23}$$

On the other hand, using (14), it is also straight-forward to show that

$$\text{trace}\left\{\boldsymbol{\mathcal{R}}^{-1}(p)\hat{\boldsymbol{\mathcal{R}}}\right\} = \text{trace}\left\{\mathbf{U}^H\mathbf{R}^{-1}(p)\hat{\mathbf{R}}\mathbf{U}\right\}$$

$$= \text{trace}\left\{\mathbf{R}^{-1}(p)\hat{\mathbf{R}}\right\} \approx M \tag{24}$$

Using (23) and (24) and omitting terms independent of p, the negative LL function reduces to

$$\mathcal{L}(p) = L \ln\left(r_{MM} - \sum_{m=1}^{p}\frac{|c_m|^2}{\lambda_m}\right) \tag{25}$$

The obtained LL is a monotonic function of the squared magnitudes of elements c_m, $m = 1, \dots, p$. Note that only these elements provide the necessary information for the estimation of the number of sources p. The corresponding eigenvalues λ_m, $m = 1, \dots, p$, only play the role of weighting factors. It is also important to note that this function is decreasing with p and it describes the goodness-of-fit part for the detection criterion. The number of free parameters in the negative LL function (25) is clearly p for the elements c_m, and p^2 for the signal subspace which translates to the rank condition on $\mathbf{R}$ or $\mathbf{R}_M$ [6] . Following the development in [1], a penalty function can be applied leading to an information theoretic criterion for automatic source number estimation. More specifically, a Minimum Description Length (MDL) criterion can be formulated as follows

$$\text{NU-MDL}_M(p) = \mathcal{L}(p) + P(p) \tag{26}$$

where subscript M corresponds for the M-th removed array element and the penalty function $P(p)$ is defined as

$$P(p) = \frac{1}{2}(p + p^2)\ln\{L\} \tag{27}$$

3.4. Comparison with Gerschgorin MDL

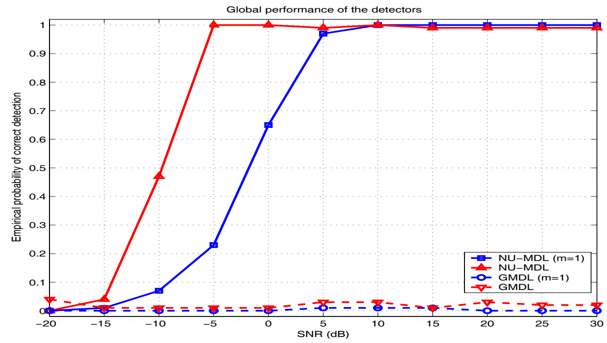
A similar idea, namely the Gerschgorin MDL (GMDL), was presented in [6] for uniform noise. The GMDL criterion applies the same penalty function as (27) and its goodness-of-fit part has the following expression

$$\begin{aligned} \mathcal{L}(p) &= L(M-1-p)\ln\left\{\frac{\left(\prod_{m=p+1}^{M-1}\lambda_m\right)^{1/(M-1-p)}}{\frac{1}{M-1-p}\sum_{m=p+1}^{M-1}\lambda_m}\right\} \\ &\quad -L\ln\left(r_{MM} - \sum_{m=1}^{p}\frac{|c_m|^2}{\lambda_m}\right) \end{aligned} \tag{28}$$

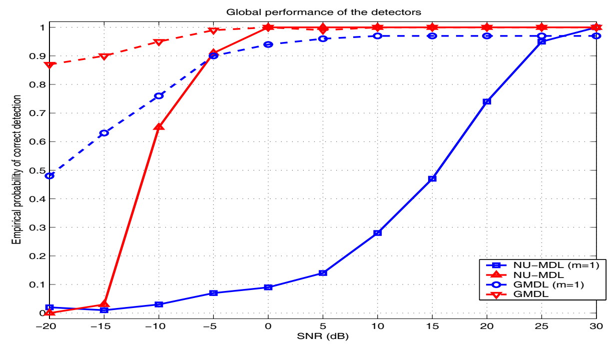
As compared to (25), this function contains an extra part in addition to the contribution of the elements of $\mathbf{c}$ as seen previously. This extra part is a monotonic function of the ratio of the geometric mean of the noise eigenvalues to the corresponding arithmetic mean. When the noise is uniform, the smallest $M-1-p$ eigenvalues in (12) are equal and their corresponding eigenvectors span the noise subspace. This information can be successfully used to separate the noise and the signal subspaces. However, when the noise is nonuniform, the eigenvalues can no longer be used directly for subspace separation. Moreover, application of the GMDL criterion to nonuniform noise results in an erroneous detection which will be confirmed through simulation results. On the other hand, our proposed NU-MDL_M (26) does not select the number of sources through the ordered eigenvalues λ_m, $m = 1, \dots, M-1$, but uses only the elements of $\mathbf{c}$. It is worth noting that although the NU-MDL criterion applies to nonuniform noise and the special case of uniform noise, it is not a generalization of the GMDL criterion.

3.5. Averaged NU-MDL

Since the noise powers are not equal from one sensor to another, accuracy of the NU-MDL_m detector depends on the index of the particular array element to be removed. It is clear that M distinct NU-MDL_m criteria can be obtained from the same array and an improved detector can be formulated by averaging the result over the M detectors as follows



(a) Nonuniform noise.



(b) Uniform noise.

Fig. 1. Comparison between NU-MDL and GMDL vs SNR.

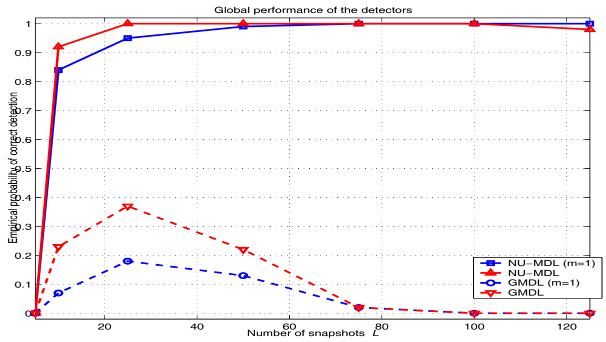
$$\text{NU-MDL}(p) = \frac{1}{M}\sum_{m=1}^{M}\text{NU-MDL}_m(p) \tag{29}$$

The above criterion involves M times the eigen-decomposition of an $(M-1)\times(M-1)$-dimensional covariance matrix and it is of the same order of complexity as the averaged GMDL criterion [6].

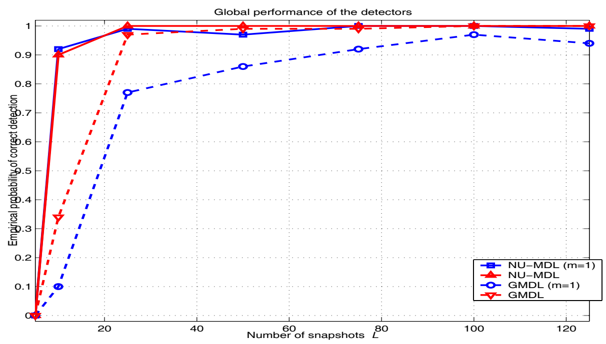
4. SIMULATION RESULTS

In what follows we show the global performance of the NU-MDL detector and compare it to GMDL. For the simulations, a Uniform Linear Array (ULA) is assumed with $M = 8$ sensors. The true number of sources is $p = 2$. In the examples labeled (a), the noise powers are given by $\mathbf{q} = [3.3, 2.6, 5.2, 1.2, 4.1, 5.0, 3.2, 6.0]^T$, therefore the Worst Noise Power Ratio (WNPR) as defined in [7] is WNPR=20. In the examples labeled (b), the noise is uniform over the sensors. All the examples illustrate the empirical probability of correct detection resulting from 200 Monte-Carlo runs.

Figure 1 illustrates the performance with respect to the Signal to Noise Ratio (SNR). The fixed parameters are the number of snapshots $L = 100$ and the angles of arrival $\boldsymbol{\theta} = [0°, 25°]^T$, whereas the SNR is set to vary from -20 dB to 30 dB. It is clear that NU-MDL outperforms GMDL in nonuniform noise as the latter relies on the misleading order of the eigenvalues. In the uniform noise case however, GMDL uses both the Gerschgorin radii and the eigenvalues to retrieve the number of sources, and therefore performs better than NU-MDL. Observe that for relatively high SNR ($>$ 0 dB), NU-MDL applies to uniform noise despite the fact that it is derived for nonuniform noise. For both detectors,

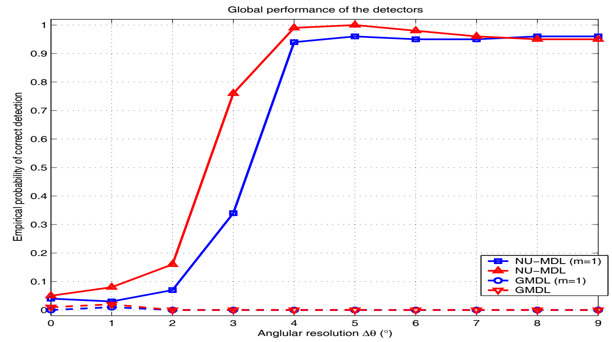


(a) Nonuniform noise.

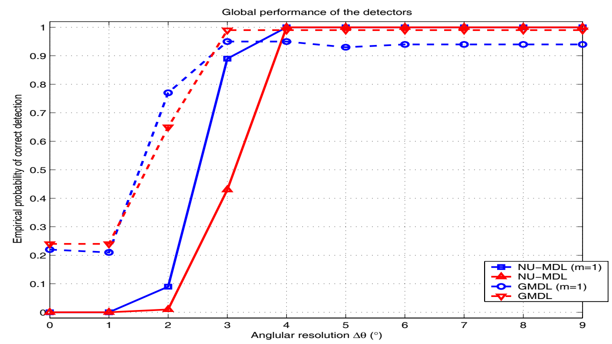


(b) Uniform noise.

Fig. 2. Comparison between NU-MDL and GMDL vs L.



(a) Nonuniform noise.



(b) Uniform noise.

Fig. 3. Comparison between NU-MDL and GMDL vs $\Delta\theta$.

the averaged version exhibits better results over the 'delete-one' detectors.

Figure 2 illustrates the performance with respect to the number of snapshots L which varies from 0 to 120. The fixed parameters are SNR=10 dB and the angles of arrival $\boldsymbol{\theta} = [0^\circ, 25^\circ]^T$. Similar comments as for the previous examples apply.

Figure 3 illustrates the performance with respect to the angular resolution $\Delta\theta$. The fixed parameters are SNR=10dB and the number of snapshots $L = 100$. The first angle of arrival is fixed at $\theta_1 = 0^\circ$ whereas the second one, θ_2, varies from 0° to 9°. The relative performances of NU-MDL and GMDL are similar to the previous examples.

5. CONCLUSION

A source detection algorithm, the Non-Uniform MDL (NU-MDL), has been proposed for a nonuniform noise environment. The detector applies a transformation of the covariance matrix of the data, resulting from array element suppression to cope with different noise powers. Through simulations, we show the high power of the method to detect sources in nonuniform noise. The results show the applicability to the simpler case of uniform noise as compared to the Gerschgorin MDL (GMDL).

6. REFERENCES

[1] M. Wax, T. Kailath, "Detection of Signals by Information Theoretic Criteria", *IEEE Trans. Acoustics, Speech, Signal Processing*, Vol. ASSP-33, No. 2, pp. 387-392, Feb. 1985

[2] A. Zoubir, "Bootstrap Methods for Model Selection", *International Journal of Electronics and Communications*, Vol. 53, No. 6, pp. 386-392, 1999.

[3] W. Chen, K.M. Wong, J.P. Reilly, "Detection of the Number of Signals: A Predicted Eigen-Threshold Approach", *IEEE Trans. Signal Processing*, Vol. 39, No. 5, pp. 1088-1098, May 1991.

[4] R.F. Brcich, A.M.Zoubir, P. Pelin, "Detection of Sources Using Bootstrap Techniques", *IEEE Trans. Signal Processing*, Vol. 50, No. 2, pp.206-215, Feb. 2002.

[5] M. Wax, I. Ziskind, "Detection of the Number of Coherent and Noncoherent Signals in the Presence by the MDL Principle", *IEEE Trans. Acoustics, Speech, Signal Processing*, Vol. ASSP-37, No.8, pp, 1190-1196, Aug. 1989.

[6] H.T. Wu, J.F. Yang, F.K. Chen, "Source Number Estimation Using Transformed Gerschgorin Radii", *IEEE Trans. Signal Processing*, Vol. 43, No. 6, pp. 1325-1333, Jun. 1995.

[7] M. Pesavento and A. B. Gershman, "Maximum-Likelihood Direction-Of-Arrival Estimation in the Presence of Unknown Nonuniform Noise," *IEEE Trans. on Signal Processing*, Vol. 49, No. 7, pp. 1310-11324, Jul. 2001.

[8] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and the Cramer-Rao Bound", *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. 37, No. 5, pp. 720-741, May 1989.

[9] L. L. Scharf, *Statistical Signal Processing : Detection, Estimation and Time Series Analysis*, Addison Wesley, 1990.

[10] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford: Clarendon, 1965.