Robust Estimation of the Number of Sources Using an MMSE-based MDL Method

Lei Huang, Teng Long and Erke Mao
Department of Electronic Engineering, Beijing Institute of Technology, Beijing, 100081, P. R. China.
E-mail: lhuang8sasp@hotmail.com

Keywords: Sensor array processing, Source number estimation, Minimum description length (MDL), Eigenvalue decomposition (EVD).

Abstract: In this paper, using the mutual information, we bridge the probability density function and the minimum mean-square error (MMSE) between the observed data and the desired signal, and then employ the MMSE to construct an MMSE-based MDL criterion for accurate source enumeration. The presented numerical results demonstrate that the proposed method is superior to the existing MDL methods in detection performance.

I. Introduction

In practical applications of array processing, such as radar and sonar, it is desired to super-resolve the incident signals and then efficiently suppress the interference signals and the background noise. As a consequence, an important preliminary step is to determine the number of sources. In the community of array processing, source enumeration and direction-of-arrival (DOA) estimation have been extensively investigated [1]-[5].

In this paper, an MDL method based on minimum mean square error (MMSE) is addressed. By exploiting the mutual information, we bridge the probability density function and the minimum mean-square error (MMSE) between the observed data and the desired signal, and then use the MMSE to construct an MMSE-based MDL estimator (mMDLE) for source number. Unlike the classical MDL method, the proposed method suggests to use the MMSEs to describe the code length of the observed data. Meanwhile, since the proposed method does not use the arithmetic and geometric means of the smallest eigenvalues, it is superior to the classical MDL method in robustness, in particular when both the number of sensors and the number of sources are large and the noise power levels are nonuniform.

II. Problem Formulation

Consider an uniform linear array (ULA) of M sensors receiving p ($p < M$) narrow-band sources from distinct directions $\theta_1, \cdots, \theta_p$. Assume that the incident signals are narrow-band and that the sources and the array are in the same plane. In the sequel, the ℓth snapshot vector of the array output can be written as

$$\begin{aligned} \boldsymbol{x}_M(t_\ell) &= [x_1(t_\ell), x_2(t_\ell), \cdots, x_M(t_\ell)]^T \\ &= \boldsymbol{A}_M(\boldsymbol{\theta})\, \boldsymbol{s}(t_\ell) + \boldsymbol{n}(t_\ell), \end{aligned} \tag{1}$$

where $\boldsymbol{A}_M(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \cdots, \boldsymbol{a}(\theta_p)]$, $\boldsymbol{s}(t_\ell) = [s_1(t_\ell), \cdots, s_p(t_\ell)]^T$, $\boldsymbol{n}(t_\ell) = [n_1(t_\ell), \cdots, n_M(t_\ell)]^T$, $(\cdot)^T$ is the transpose operation and the steering vector is given by

$$\boldsymbol{a}(\theta_i) = \left[1, e^{j2\pi d \sin(\theta_i)/\lambda}, \cdots, e^{j2\pi d(M-1)\sin(\theta_i)/\lambda}\right]^T, \tag{2}$$

where λ denotes the wavelength and d is the inter-sensor spacing. The source waveform $s_i(t_\ell)$ $(i = 1, \cdots, p)$ is assumed to be a jointly stationary, statistically uncorrelated, zero-mean complex Gaussian random process, and p represents the *unknown* number of sources. The sensor noise $\boldsymbol{n}(t_\ell)$ is assumed to be an ergodic, zero-mean, spatially and temporally white complex Gaussian process with the covariance matrix $\sigma_n^2 \boldsymbol{I}_M$. In addition, the noise is presumed to be uncorrelated with the sources.

Given the observed data $\boldsymbol{x}(t_\ell) = \boldsymbol{B}\boldsymbol{x}_M(t_\ell)$ and the desired signal $d(t_\ell) = \boldsymbol{e}^T \boldsymbol{x}_M(t_\ell)$ where $\boldsymbol{e} = [1, 0, \cdots, 0]^T$ and $\boldsymbol{B} = [\boldsymbol{0}_{M-1} \vdots \boldsymbol{I}_{M-1}]$. it is nature to use the classical Wiener filter (WF) to obtain the optimal estimate of the desired signal in the MMSE sense. The Wiener filtering problem is depicted in Fig. 1, where $\boldsymbol{w}_{WF}$ is the M-dimensional Wiener filter, $\hat{d}(t_\ell)$ is the scalar estimate of $d(t_\ell)$ and $e(t_\ell)$ is the scalar error signal. The classical Wiener filtering problem is to minimize

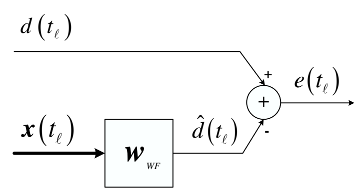

Fig. 1. The classical Wiener filter

the mean-square error between the reference signal $d(t_\ell)$ and its estimate

$$\hat{d}(t_\ell) = \boldsymbol{w}^H \boldsymbol{x}(t_\ell), \tag{3}$$

where $\boldsymbol{w}$ is a linear filter, yielding the classical Wiener filter

$$\boldsymbol{w}_{WF} = \boldsymbol{R}_x^{-1} \boldsymbol{r}_{xd}, \tag{4}$$

where $\boldsymbol{R}_x = E\left[\boldsymbol{x}(t_\ell)\boldsymbol{x}^H(t_\ell)\right]$. Consequently, the error signal is given by

$$e(t_\ell) = d(t_\ell) - \hat{d}(t_\ell) = d(t_\ell) - \boldsymbol{w}_{WF}^H \boldsymbol{x}(t_\ell). \tag{5}$$

Denote the eigenvalue decomposition (EVD) of $\boldsymbol{R}_x$ as

$$\boldsymbol{R}_x = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H, \tag{6}$$

where $\boldsymbol{\Lambda} = \text{diag}\{[\lambda_1, \cdots, \lambda_M]\}$, $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_M]$, $\lambda_1 \geq \cdots \geq \lambda_{p+1} = \cdots = \lambda_M$ are the eigenvalues and $\boldsymbol{u}_i$ $(i = 1, \cdots, M)$ are the corresponding eigenvectors. In the sequel, the minimum mean-square error (MMSE) can be written as

$$\text{MMSE} = E\left[|e(t_\ell)|^2\right] = \sigma_d^2 - \boldsymbol{r}_{xd}^H \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^H \boldsymbol{r}_{xd}, \tag{7}$$

where $\sigma_d^2 = E\left[d(t_\ell)d^*(t_\ell)\right] = \sigma_d^2$ and

$$\begin{aligned} \boldsymbol{r}_{xd} &= E\left[\boldsymbol{x}(t_\ell)d^*(t_\ell)\right] \\ &= \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{\beta}, \end{aligned} \tag{8}$$

where

$$\boldsymbol{A}_M(\boldsymbol{\theta}) = \left(\begin{array}{c} \boldsymbol{A}(\boldsymbol{\theta}) \\ \text{last column} \end{array}\right)$$

and $\boldsymbol{\beta}$ is nonzero. Substituting (8) into (7) and noticing that $|\boldsymbol{r}_{xd}^H \boldsymbol{u}_i|^2 = 0$ $(i = p+1, \cdots, M)$ yields

$$\xi^{(1)} > \xi^{(2)} > \cdots > \xi^{(p+1)} = \cdots = \xi^{(M)}, \tag{9}$$

where

$$\xi^{(k)} \triangleq \sigma_d^2 - \sum_{i=1}^{k} \frac{|\boldsymbol{r}_{xd}^H \boldsymbol{u}_i|^2}{\lambda_i} \tag{10}$$

is the kth MMSE.

III. Proposed MMSE-Based MDL Estimator

For the complex Gaussian processes $\boldsymbol{X} = \{\boldsymbol{x}(t_\ell)\}_{\ell=1}^N$, the probability density function (PDF) can be given by

$$f(\boldsymbol{X}|\boldsymbol{\Theta}) = \prod_{\ell=1}^{N} \frac{1}{\pi^M \det(\boldsymbol{R}_x)} \exp\{-\boldsymbol{x}^H(t_\ell)\boldsymbol{R}_x\boldsymbol{x}(t_\ell)\}, \tag{11}$$

where $\boldsymbol{\Theta}$ is an unknown parameter vector. In the sequel, the mutual information between the reference signal $d(t_\ell)$ and the observed data $\boldsymbol{x}(t_\ell)$ can be expressed as

$$I(d, \boldsymbol{x}) = H(d) - H(d|\boldsymbol{x}), \tag{12}$$

where $H(z) \triangleq -\int f(z)\log f(z)dz$ is the entropy of the signal process $z(t_\ell)$. Note that

$$H(d|\boldsymbol{x}) = H(d, \boldsymbol{x}) - H(\boldsymbol{x}). \tag{13}$$

It follows from (12) and (13) that

$$I(d, \boldsymbol{x}) = H(d) + H(\boldsymbol{x}) - H(d, \boldsymbol{x}). \tag{14}$$

For the complex Gaussian processes $d(t_\ell)$ and $\boldsymbol{x}(t_\ell)$, the corresponding entropies are, respectively, given by

$$H(d) = \log\{\pi e \sigma_d^2\} \tag{15}$$

$$H(\boldsymbol{x}) = \log\left\{(\pi e)^M \det\left(\boldsymbol{R}_x\right)\right\}. \tag{16}$$

Notice that the covariance matrix of $\boldsymbol{y}(t_\ell) \triangleq [d(t_\ell), \boldsymbol{x}^T(t_\ell)]^T$ can be written as

$$\boldsymbol{R}_y \triangleq E[\boldsymbol{y}(t_\ell)\boldsymbol{y}^H(t_\ell)] = \left(\begin{array}{cc} \sigma_d^2 & \boldsymbol{r}_{xd}^H \\ \boldsymbol{r}_{xd} & \boldsymbol{R}_x \end{array}\right). \tag{17}$$

In the sequel, the joint entropy of the $d(t_\ell)$ and $\boldsymbol{x}(t_\ell)$ can be calculated by

$$\begin{aligned} &H(d, \boldsymbol{x}) \\ &= \log\left\{(\pi e)^{M+1}\det\left(\boldsymbol{R}_y\right)\right\} \\ &= \log(\pi e) + \log\left\{(\pi e)^M\left(\det(\boldsymbol{R}_x)\times\det(\sigma_n^2 - \boldsymbol{r}_{xd}^H\boldsymbol{R}_x^{-1}\boldsymbol{r}_{xd})\right)\right\} \\ &= \log(\pi e) + \log\left\{(\pi e)^M\det(\boldsymbol{R}_x)\right\} + \log\left(\sigma_d^2 - \boldsymbol{r}_{xd}^H\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^H\boldsymbol{r}_{xd}\right). \end{aligned} \tag{18}$$

Substituting (15), (16) and (18) into (14) yields

$$\begin{aligned} I(d, \boldsymbol{x}) &= \log(\sigma_d^2) - \log\left(\sigma_d^2 - \boldsymbol{r}_{xd}^H\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^H\boldsymbol{r}_{xd}\right) \\ &= \log(\sigma_d^2) - \log\left(\sigma_d^2 - \sum_{i=1}^{M}\frac{|\boldsymbol{r}_{xd}^H\boldsymbol{u}_i|^2}{\lambda_i}\right). \end{aligned} \tag{19}$$

Denoting the *supposed* number of sources by k, substituting (10) into (19), and considering that $|\boldsymbol{r}_{xd}^H\boldsymbol{u}_i| = 0$ $(i = k+1, \cdots, M)$, we obtain

$$I^{(k)}(d, \boldsymbol{x}) = \log(\sigma_d^2) - \log\left(\xi^{(k)}\right). \tag{20}$$

It is shown in [7] that the Kullback-Leibler (K-L) divergence between two PDFs can be defined by

$$D\left(f_1 \| f_2\right) = \int_z f_1(z)\log\frac{f_1(z)}{f_2(z)}dz. \tag{21}$$

Meanwhile, it is indicated in [12] that the mutual information and the K-L divergence have the following equality

$$\begin{aligned} I^{(k)}(d, \boldsymbol{x}) &= D^{(k)}\left(f\left(\boldsymbol{y}\right) \| \prod_{i=1}^{M+1} f\left(y_i\right)\right) \\ &= \sum_{i=1}^{M+1} H\left(y_i\right) - H^{(k)}\left(\boldsymbol{y}\right) \\ &= \sum_{i=1}^{M+1}\log\left(\pi e \sigma_{y_i}^2\right) - H^{(k)}\left(\boldsymbol{y}\right). \end{aligned} \tag{22}$$

It follows from (20) and (22) that

$$H^{(k)}\left(\boldsymbol{y}\right) = \log\left(\xi^{(k)}\right) + \varrho, \tag{23}$$

$\varrho \triangleq \sum_{i=1}^{M+1}\log\left(\pi e \sigma_{y_i}^2\right) - \log(\sigma_d^2)$. On the other hand, for $N \to \infty$, it follows from [7] that

$$H^{(k)}\left(\boldsymbol{y}\right) = -\frac{1}{N}\log\left(f(\boldsymbol{Y}^{(k)}|\boldsymbol{\Theta})\right), \tag{24}$$

where $\boldsymbol{Y}^{(k)} = \{\boldsymbol{y}(t_\ell)\}_{\ell=1}^N$. In the sequel, substituting the ML estimates of $\xi^{(k)}$ and $\boldsymbol{\Theta}$, i.e., $\hat{\xi}^{(k)}$ and $\hat{\boldsymbol{\Theta}}$, into (23) and (24), and omitting the terms independent of $\boldsymbol{\Theta}$, we attain the minimum description length (MDL) encoding the observed data $\boldsymbol{y}(t_\ell)$ as

$$\begin{aligned} \mathcal{L}\{\boldsymbol{y}(t_\ell)\} &= -\log f(\boldsymbol{Y}^{(k)}|\hat{\boldsymbol{\Theta}}) + \frac{1}{2}K\log N \\ &= N\log\left(\hat{\xi}^{(k)}\right) + \frac{1}{2}K\log N, \end{aligned} \tag{25}$$

where K equals the number of free parameters in $\boldsymbol{\Theta}$. It follows from (7) and (10) that the calculation of the MMSEs involves the projection of the k-dimensional signal subspace and k largest eigenvalues λ_i $(i = 1, \cdots, k)$, which implies that the number of the free adjustable parameters is $K = k^2 + k$. Therefore, the number of sources can be yielded by the MMSE-based MDL estimator (mMDLE):

$$\hat{p}_{\mathrm{mMDLE}} = \arg \min_{k=1,\cdots,M-1} \mathrm{mMDLE}(k), \tag{26}$$

where

$$\mathrm{mMDLE}(k) = N \log\left(\hat{\xi}^{(k)}\right) + \frac{1}{2}(k^2 + k) \log N, \tag{27}$$

and $\hat{p}_{\mathrm{mMDLE}}$ denotes the estimated number of sources.

Remark A: It is indicated in (10) that when there is no incident signal, all the MMSEs are equal to σ_d^2; otherwise, all the MMSEs are lower than σ_d^2. In the sequel, we may use the MMSEs to determine whether the sources exist or not. Noticing that the calculated MMSEs are the $O(\sqrt{\log\log N/N})$ estimates of the true MMSEs, we can define an adaptive detector for determining whether there exist the incident sources or not:

$$\left| \frac{1}{M} \sum_{m=1}^{M} \hat{\xi}^{(m)} - \hat{\sigma}_d^2 \right| \lessgtr_{\mathcal{H}_1}^{\mathcal{H}_0} \gamma, \tag{28}$$

where $\gamma = C\sqrt{\log\log N/N}$, C is a positive constant number, and $\mathcal{H}_0$ and $\mathcal{H}_1$ denote the cases of source absence and, respectively, source presence. If $\mathcal{H}_0$ is accepted, we can conclude that there is no incident source; otherwise $\mathcal{H}_1$ is accepted, we then proceed to apply the proposed estimator in (27) to accurately obtain the number of sources.

Remark B: Following the results of Wu *et al* [6], an NU-MDL method [8] was proposed to estimate the number of sources. The minimum description length of NU-MDL method is given as

$$\mathcal{L}\{\boldsymbol{x}(t)\} = N \log\left(r_{MM} - \sum_{i=1}^{k} \frac{|c_i|^2}{\lambda_i} \right) + \frac{1}{2}(k^2 + k) \log N, \tag{29}$$

where $r_{MM} \triangleq E[|x_M(t)|^2]$, $c_i \triangleq |\boldsymbol{r}_{xd}^H \boldsymbol{u}_i|$ $(i = 1, \cdots, k)$ are the so-called Gerschgorin radii which satisfy $c_1 \geq \cdots \geq c_k \geq c_{k+1} = \cdots = c_{M-1} = 0$, and the eigenvalues are arranged in a decreasing order as $\lambda_1 \geq \cdots \geq \lambda_k \geq \lambda_{k+1} = \cdots = \lambda_{M-1}$. It is worth stressing that the mMDLE method does not use the *ordered* $|\boldsymbol{r}_{xd}^H \boldsymbol{u}_i|$ $(i = 1, \cdots, k)$ in the calculation of the MMSE whereas the NU-MDL method requires $c_1 \geq \cdots \geq c_k \geq c_{k+1} = \cdots = c_{M-1} = 0$, as is the essential difference of the proposed method from the NU-MDL method while they take the similar formulations. Actually, the nonnegative term $|c_k|^2/\lambda_k$ is the well-known cross-spectral (CS) energy of the cross-spectral metric (CSM) [13] for adaptive reduced-rank filtering. If the CS energy terms are arranged in a decreasing order and then used to calculate the MMSE, one may obtain the faster convergence of the MMSE. This is why the CS energy can be used in the CSM to obtain the reduced rank of the observed space in the adaptive reduced-rank filtering. However, when the CS energy is employed to determine the number of sources, i.e., the rank of the principal eigen-subspace, the CS energy terms should be calculated by directly using the eigen-pairs $\{\boldsymbol{u}_i, \lambda_i\}$ instead of the *ordered* Gerschgorin radii so that the CS energy terms are not in a decreasing order. Otherwise, the estimated number of sources is the rank of the CS subspace which is generally less than the true number of sources. Therefore, the NU-MDL method tends to underestimate the number of sources, particularly for the large system where both the number of sensors and the number of sources are large.

IV. Numerical Results

As noted in Remark B, the NU-MDL method in general underestimates the number of sources because it requires the *ordered* Gerschgorin radii, equivalently selecting the cross-spectral subspace. This becomes more true for the large system where both the number of sensors and the number of sources are large. To demonstrate that the proposed mMDLE method offers the consistency while the NU-MDL method is inconsistent in the large system, we assume that the ULA is of 16 sensors, which received five equal-power sources. Meanwhile, we assume that the sensor noise is a stationary, spatially and temporally white, Gaussian random process which is uncorrelated with the sources.

The empirical probabilities of correct detection versus the number of snapshots are demonstrated in Fig. 2. We can observe that the mMDLE and cMDL methods can correctly detect the sources. The NU-MDL method, however, can not correctly estimate the source number even when the number of snapshots become large enough. The empirical probabilities of correct detection versus the number of sources are shown in Fig. 3. It can be observed that as the number of sources creases, the probability of correct detection of the NU-MDL method converges to zero, thereby indicating that it can not provide a reliable estimate of the number of sources. On the contrary, both proposed mMDLE method and the cMDL method offer the consistency for the scenario of large system.

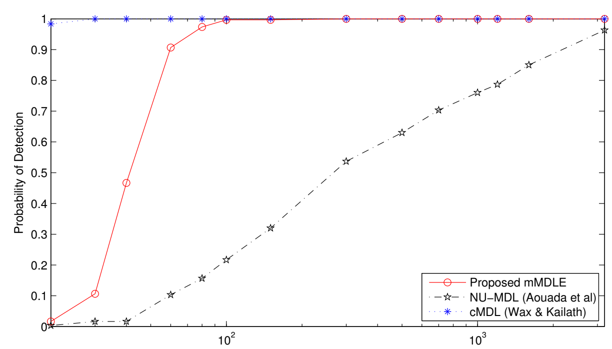

Fig. 2. Probability of correct detection versus number of snapshots for the case of spatially and temporally white noise. SNR = 0dB, $[\theta_1, \cdots, \theta_5] = [0^\circ, 5.5^\circ, 63^\circ, -12.2^\circ, 44.4^\circ]$ and $C = 2$. 300 trials.

In the following, we consider the nonuniform noise, whose power levels are given as

$$[\sigma_1^2, \cdots, \sigma_{16}^2] = [1.78\ 0.50\ 1.60\ 0.85\ 1.67\ 1.52\ 0.74$$
$$0.88\ 2.00\ 1.66\ 1.77\ 1.05\ 1.52\ 1.15\ 1.48\ 1.52]$$

Fig. 4 demonstrates the empirical probabilities of correct detection versus the number of snapshots. Observe that the mMDLE method can correctly detect the sources. The cMDL

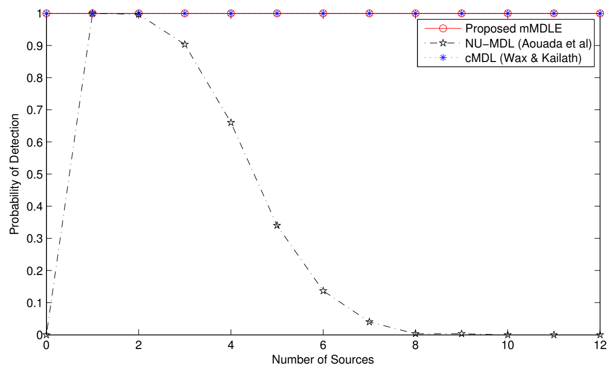

Fig. 3. Probability of correct detection versus number of sources for the case of spatially and temporally white noise. The number of sources varies from 0 to 12, SNR $=$ 10dB, $N = 600$, $M = 16$ and $C = 2$. 300 trials.

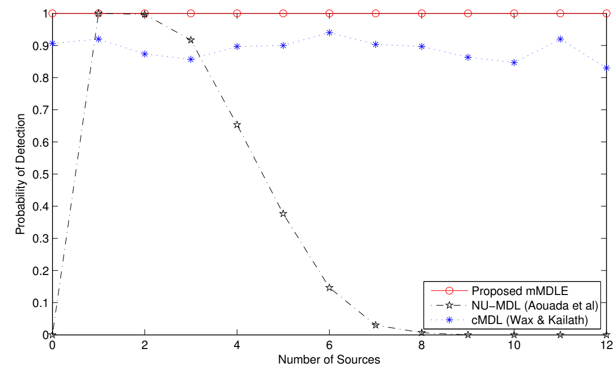

Fig. 5. Probability of correct detection versus number of sources for the case of nonuniform noise. The number of sources varies from 0 to 12, SNR $=$ 10dB, $N = 600$, $M = 16$ and $C = 2$. 300 trials.

and NU-MDL methods, however, fail to correctly estimate the source number even when the number of snapshots infinitely increases. The empirical probabilities of correct detection versus the number of sources are illustrated in Fig. 5. It can be observed that as the number of sources creases, the probability of correct detection of the NU-MDL method converges to zero, thereby indicating that it can not provide a reliable estimate of the number of sources. Meanwhile, the cMDL method can not correctly detect the sources either no matter what the number of sources is. However, the proposed mMDLE method can retain its consistency even for the scenario of nonuniform noise, as demonstrated in Fig. 5.

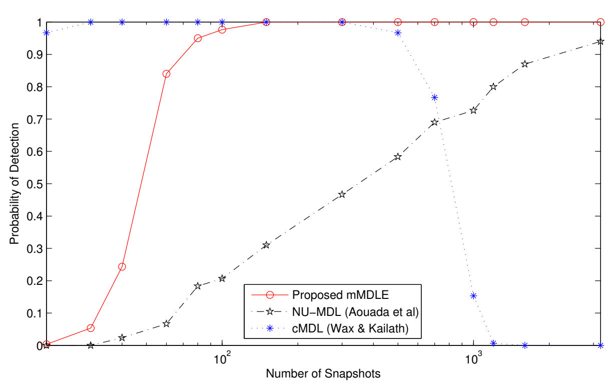

Fig. 4. Probability of correct detection versus number of snapshots for the case of nonuniform noise. SNR $=$ 0dB, $[\theta_1, \cdots, \theta_5] = [0°, 5.5°, 63°, -12.2°, 44.4°]$ and $C = 2$. 300 trials.

V. Conclusion

An MMSE-based MDL estimator for source number has been addressed in this paper. Using the MMSE-based MDL criterion, the proposed method can yield the robust estimation of the number of sources in the large system. Meanwhile, we show that the NU-MDL method can not correctly estimate the number of sources in the large system due to the use of the *ordered* Gerschgorin radii while the proposed mMDLE and the classical MDL methods retain their consistency. However, as the nonuniform noise becomes more strong, the mMDLE method might not provide the robust estimate of the number of sources anymore. Developing more robust method for source enumeration may be our future efforts.

Acknowledgment

The work was supported by the Natural Science Foundation of China (NSFC) under Grant No. 60702068.

References

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

[2] M. Wax and I. Ziskind, "Detection of the number of coherent signals by the MDL principle," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 8, pp. 1190-1196, Aug. 1989.

[3] L. Huang, S. Wu, D. Feng, and L. Zhang, "Low complexity method for signal subspace fitting," *Electron. Lett.*, vol. 40, pp. 847C848, Jul. 2004.

[4] L. Huang, S. Wu, D. Feng, and L. Zhang, "Computationally efficient direction-of-arrival estimation based on partial a priori knowledge of signal sources,"*EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1-7.

[5] L. Huang, D. Feng, L. Zhang, and S. Wu, "Computationally efficient method of signal subspace fitting for direction-of-arrival estimation ," *IEICE Trans. Commu.*, vol. E88-B, no. 8, 2005, pp. 3408-3415.

[6] H. T. Wu, J. F. Yang, and F. K. Chen, "Source number estimators using transformed Gerschgorin radii," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1325-1333, 1995.

[7] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 716-723, Dec. 1974.

[8] S. Aouada, A. M. Zoubir, and C. M. S. See, "Source detection in the presence of nonuniform noise," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, pp. 165-168, May 2004.

[9] L. Huang, T. Long, and S. Wu, "Source enumeration for high-resolution array processing using improved Gerschgorin radii without eigendecomposition," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 5916-5925, Dec. 2008.

[10] L. Huang and S. Wu, "Low-complexity MDL method for accurate source enumeration," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 581-584, Sep. 2007.

[11] L. Huang, S. Wu and X. Li, "Reduced-rank MDL method for source enumeration in high-resolution array processing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5658-5667, Dec. 2007.

[12] M. Novey and T. Adali, "Complex ICA by negentropy maximization," *IEEE Trans. Neural Networks*, vol. 19, no. 4, pp. 596-609, Apr. 2008.

[13] J. S. Goldstein and I. S. Reed, "Subspace selection for partially adaptive sensor array processing," *IEEE Trans. Aerosp. and Electron. Syst.* vol. 33, no. 2, pp. 539-544, Apr. 1997.