SOURCE NUMBER ESTIMATION IN NON-GAUSSIAN NOISE

G. V. Anand and P. V. Nagesha

Department of Electronics and Communication Engineering, PES University, Bangalore 560085, India
gargeshwari.anand@gmail.com, nagesh.uwa@gmail.com

ABSTRACT

In this paper a new method of source number estimation in non-Gaussian noise is presented. The proposed signal subspace identification (SSI) method involves estimation of the array signal correlation matrix and determining the number of positive eigenvalues of the estimated correlation matrix. The SSI method is applied to the problem of estimating the number of plane wave narrowband signals impinging on a uniform linear array. It is shown that the performance of the SSI method in non-Gaussian heavy-tailed noise is significantly better than that of the widely used minimum description length (MDL) method and the recently proposed entropy estimation of eigenvalues (EEE) method based on random matrix theory.[1]

Index Terms—Non-Gaussian noise, noise variance estimation, signal subspace identification, source number estimation

1. INTRODUCTION

High-resolution direction-of-arrival (DOA) estimation is a problem of great interest in several applications such as radar, sonar, wireless communication, biomedical engineering, etc. High-resolution DOA estimation techniques such as MUSIC and ESPRIT require prior knowledge of the number of sources. Methods of source enumeration based on the classical information theoretic criteria include the Akaike information criterion (AIC) [1], minimum description length (MDL) [2], Bayesian information criterion (BIC) [3], predictive description length (PDL) [4], Gerschgorin disk estimator (GDE) [5], and several variants of these. Other methods include bootstrap techniques [6, 7], random matrix theory [8] and entropy estimation of eigenvalues (EEE) [9]. In most of these methods [1-4, 6, 8], noise is assumed to be Gaussian and white. In this paper, we present a new method of source number estimation in white noise with an unknown probability distribution. The proposed method is based on a comparison of an estimate of the noise variance with estimates of eigenvalues of the array data correlation matrix. All the estimates are obtained from a sufficiently large number of snapshots of the array data vector.

The paper is organized as follows. The method of source number estimation is described in Section 2. Simulation results are presented in Section 3 to illustrate the performance of the method under different conditions. Conclusions are presented in Section 4.

2. SOURCE NUMBER ESTIMATION

Let the signals received from J uncorrelated narrowband sources be measured by an N-sensor array, with $N > J$. Consider L snapshots of the N-dimensional data vector $\boldsymbol{x}(t)$:

$$\boldsymbol{x}(t) = \boldsymbol{s}(t) + \boldsymbol{w}(t) = \sum_{j=1}^{J} \boldsymbol{a}(\emptyset_j) p_j(t) + \boldsymbol{w}(t); \quad t = 0, \ldots, L-1;\ L > N. \qquad (1)$$

In (1), $\boldsymbol{s}(t)$ is the array signal vector, $p_j(t); j = 1, \ldots, J; t = 0, \ldots. L-1$ are the complex amplitudes of the received signals modeled as mutually uncorrelated complex Gaussian random variables with variances σ_j^2, $\boldsymbol{a}(\emptyset_j) = \left[1\ e^{i\pi\cos\emptyset_j} \ldots e^{i(N-1)\pi\cos\emptyset_j}\right]^{\mathrm{T}}$ is the steering vector associated with the j^{th} source located at an unknown direction $\emptyset_j$, and $\boldsymbol{w}(t)$ is the noise vector independent of the signals, with zero mean and correlation matrix $\boldsymbol{R}_w = \sigma^2 \boldsymbol{I}$. Let the eigenvalues of the data correlation matrix $\boldsymbol{R}_x = \mathrm{E}[\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{H}}(t)]$, arranged in non-decreasing order, be denoted by $\lambda_{x,1}, \ldots, \lambda_{x,N}$, and let σ^2 be the noise variance; so that we have $\lambda_{x,n} \geq \lambda_{x,k}$ for $n < k$, $\lambda_{x,n} > \sigma^2$ for $n \leq J$, and $\lambda_{x,n} = \sigma^2$ for $n > J$. If the actual eigenvalues and noise variance are replaced by their estimates $\{\hat{\lambda}_{x,n}; n = 1, \ldots, N\}$ and $\hat{\sigma}^2$, obtained from a sufficiently large number of snapshots of $\boldsymbol{x}(t)$, we expect the following relations to hold: $\hat{\lambda}_{x,n} > \hat{\sigma}^2$ for $n \leq J$ and $\hat{\lambda}_{x,n} \leq \hat{\sigma}^2$ for $n > J$. Hence an estimate of J may be obtained as $\hat{J} =$ the largest integer k such that $\hat{\lambda}_{x,k} > \hat{\sigma}^2$.

An estimate of the data covariance matrix $\boldsymbol{R}_x$ is given by $\widehat{\boldsymbol{R}}_x = \frac{1}{L}\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}$, where $\boldsymbol{X} = [\boldsymbol{x}(0) \ldots \boldsymbol{x}(L-1)]$. The noise

[1] This work was supported by Naval Research Board, Defence R&D Organisation, Ministry of Defence, Government of India, through a grant for Project no. NRB-261/SSB/11-12.

variance may be estimated by extending a procedure proposed by Bioucas-Dias and Nascimento [5] for estimating the noise correlation matrix $\boldsymbol{R}_w$. Let $\boldsymbol{Y} = \boldsymbol{X}^{\mathrm{T}}$. The n^{th} column of $\boldsymbol{Y}$, denoted by $\boldsymbol{y}_n = [y_n(0) \dots y_n(L-1)]^T$, is the noisy signal waveform measured at the n^{th}sensor. Let $\boldsymbol{Y}_{\partial_n} = [\boldsymbol{y}_1 \dots \boldsymbol{y}_{n-1}\boldsymbol{y}_{n+1} \dots \boldsymbol{y}_N]$ be the matrix obtained by removing the n^{th} column from $\boldsymbol{Y}$. In view of the high correlation between the elements of the signal vector $\boldsymbol{s}(t)$ for all t, we write the following regression relation

$$\boldsymbol{y}_n = \boldsymbol{Y}_{\partial_n}\boldsymbol{\alpha}_n + \boldsymbol{\eta}_n; n = 1, \dots, N, \tag{2}$$

where $\boldsymbol{\alpha}_n$ is the $(N-1)$-dimensional regression vector and $\boldsymbol{\eta}_n$ is the modelling error vector. For each $n \in \{1, \dots, N\}$, the least squares estimator of $\boldsymbol{\alpha}_n$ is given by

$$\widehat{\boldsymbol{\alpha}}_n = \left[\boldsymbol{Y}_{\partial_n}^{\mathrm{H}}\boldsymbol{Y}_{\partial_n}\right]^{-1}\boldsymbol{Y}_{\partial_n}^{\mathrm{H}}\boldsymbol{y}_n. \tag{3}$$

The vector $\widehat{\boldsymbol{\xi}}_n = \boldsymbol{Y}_{\partial_n}\widehat{\boldsymbol{\alpha}}_n$ may be considered to be the linear prediction of the signal waveform $\boldsymbol{\xi}_n = [s_n(0) \dots s_n(L-1)]^{\mathrm{T}}$ at the n^{th} sensor derived from the noisy signal waveform measurements at the other sensors. An estimate of the noise waveform $\boldsymbol{\eta}_n = [w_n(0) \dots w_n(L-1)]^{\mathrm{T}}$ is therefore given by

$$\widehat{\boldsymbol{\eta}}_n = \boldsymbol{y}_n - \widehat{\boldsymbol{\xi}}_n = \boldsymbol{y}_n - \boldsymbol{Y}_{\partial_n}\widehat{\boldsymbol{\alpha}}_n; n = 1, \dots, N. \tag{4}$$

The corresponding estimate of the noise correlation matrix $\boldsymbol{R}_w$ is

$$\widetilde{\boldsymbol{R}}_w = \frac{1}{L}[\widehat{\boldsymbol{\eta}}_1 \dots \widehat{\boldsymbol{\eta}}_N]^{\mathrm{H}}[\widehat{\boldsymbol{\eta}}_1 \dots \widehat{\boldsymbol{\eta}}_N]. \tag{5}$$

In the present context, we have assumed that $\boldsymbol{R}_w = \sigma^2\boldsymbol{I}$. Let $\{\tilde{\lambda}_{w,n}; n = 1, \dots, N\}$ be the eigenvalues of $\widetilde{\boldsymbol{R}}_w$ arranged in non-ascending order. Define

$$\mu_w = \frac{1}{N}\sum_{n=1}^{N}\tilde{\lambda}_{w,n} = \frac{1}{N}tr\{\widetilde{\boldsymbol{R}}_w\}. \tag{6}$$

Ideally, the mean eigenvalue μ_w may be considered to be an estimate of noise variance σ^2, and those eigenvalues of $\widehat{\boldsymbol{R}}_x$ which are larger than μ_w may be considered to be the signal eigenvalues. However, $\widetilde{\boldsymbol{R}}_w$ and μ_w are biased estimates of $\boldsymbol{R}_w$ and σ^2 for the following reasons. (1) The estimates are derived from finite data. (2) $\widehat{\boldsymbol{\xi}}_n$ is the linear prediction of signal waveform $\boldsymbol{\xi}_n$ obtained from the noisy signal waveforms $\{\boldsymbol{y}_m; m \neq n\}$. Since the energy of $\boldsymbol{y}_n$ is larger than that of $\boldsymbol{\xi}_n$, we may expect the energy of $\widehat{\boldsymbol{\xi}}_n$ to overestimate the energy of $\boldsymbol{\xi}_n$. Hence the energy of $\widehat{\boldsymbol{\eta}}_n = \boldsymbol{y}_n - \widehat{\boldsymbol{\xi}}_n$ would underestimate the energy of the noise waveform $\boldsymbol{\eta}_n$. Consequently, the mean eigenvalue μ_w underestimates the noise variance σ^2. Underestimation of σ^2 leads to an overestimation of the number of sources J, particularly so when $J = 0$ or 1. Hence we propose the following procedure for estimating the noise variance and the number of sources.

Define

$$\mu_{x,k} = \frac{1}{N-k}\sum_{n=k+1}^{N}\hat{\lambda}_{x,n}\,; k = 0,1, \dots, N-1, \tag{7}$$

Since we have $\hat{\lambda}_{x,n} \geq \hat{\lambda}_{x,k}$ for $n < k$, it follows that $\mu_{x,0} > \mu_{x,1} > \cdots > \mu_{x,N-1}$. If there are J sources, $\mu_{x,J}$ should ideally be equal to μ_w. If J is known, the following estimator of σ^2 may be used to avoid underestimation: $\hat{\sigma}^2 = \tilde{\lambda}_{w,1} + \theta_0(\mu_{x,J} - \mu_w)$, where θ_0 is the hard thresholding operator defined as

$$\theta_0(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

Since the value of J is not known, we may define

$$\tilde{\lambda}_{w,1}^{(k)} = \tilde{\lambda}_{w,1} + \theta_0(\mu_{x,k} - \mu_w); k = 0,1, \dots, N-1, \tag{9}$$

and proceed to estimate the number of sources as the smallest integer k such that $\tilde{\lambda}_{w,1}^{(k)} \geq \hat{\lambda}_{x,k+1}$.

The procedure described above provides a good estimate of the number of sources J if $J \geq 2$. But the probability of overestimation is high if $J \leq 1$. If $J = 0$, both $\widehat{\boldsymbol{R}}_x$ and $\widetilde{\boldsymbol{R}}_w$ are estimates of the noise covariance matrix. But the eigenvalues of $\widetilde{\boldsymbol{R}}_w$ are significantly smaller than those of $\widehat{\boldsymbol{R}}_x$ when $J = 0$, and consequently $\sigma^2 - \mu_w$ is very large in this case. This deficiency can be overcome by adopting the refinement procedure described below.

Consider a set of random matrices

$$\boldsymbol{B}(q) = [\boldsymbol{\beta}_1(q) \dots \boldsymbol{\beta}_L(q)]; q = 1, \dots, Q, \tag{10}$$

where $\boldsymbol{\beta}_l(q) = [\beta_{l1}(q) \dots \beta_{lN}(q)]^{\mathrm{T}}$, and $\{\beta_{ln}(q); l = 1, \dots, L; n = 1, \dots, N; q = 1, \dots, Q\}$ are independent and identically distributed random variables having approximately the same distribution as the noise at the sensor outputs. Define the sample correlation matrix $\boldsymbol{R}_\beta(q) = \frac{1}{L}\boldsymbol{B}(q)\boldsymbol{B}(q)^{\mathrm{H}}$. Use the matrix $\boldsymbol{B}(q)$ to construct another correlation matrix $\boldsymbol{R}_\gamma(q)$ corresponding to the matrix $\widetilde{\boldsymbol{R}}_w$ in (5), using the regression-based procedure described there. Let the eigenvalues of $\boldsymbol{R}_\beta(q)$ and $\boldsymbol{R}_\gamma(q)$ be denoted respectively by $\lambda_{\beta,1}(q), \dots, \lambda_{\beta,N}(q)$, and $\lambda_{\gamma,1}(q), \dots, \lambda_{\gamma,N}(q)$. Equalize the means of the two sets of eigenvalues by defining

$$\hat{\lambda}_{\gamma,n}(q) = \lambda_{\gamma,n}(q) + \tfrac{1}{N}\textstyle\sum_{m=1}^{N}\{\lambda_{\beta,m}(q) - \lambda_{\gamma,m}(q)\}; \quad n = 1, \dots, N\,; q = 1, \dots, Q \tag{11}$$

Define

$$\Delta\lambda_{max} = \max_{n,q}\{\lambda_{\beta,n}(q) - \hat{\lambda}_{\gamma,n}(q)\}, \tag{12}$$

$$\hat{\lambda}_{w,1}^{(k)} = \begin{cases} \tilde{\lambda}_{w,1} + \Delta\lambda_{max},; k = 0,1 \\ \tilde{\lambda}_{w,1}; k \geq 2 \end{cases}. \tag{13}$$

Replace (9) by the following equation

$$\tilde{\lambda}_{w,1}^{(k)} = \hat{\lambda}_{w,1}^{(k)} + \theta_0(\mu_{x,k} - \mu_w); k = 0,1,\dots,N-1. \tag{14}$$

The modified expression for $\tilde{\lambda}_{w,1}^{(k)}$ given in (14) can be used to obtain estimates of J, σ^2 and $\boldsymbol{R}_w$ and $\boldsymbol{R}_s$:

$$\hat{J} = \text{the smallest integer } k \text{ such that } \tilde{\lambda}_{w,1}^{(k)} \geq \hat{\lambda}_{x,k+1}, \tag{15}$$

$$\hat{\sigma}^2 = \tilde{\lambda}_{w,1}^{(\hat{J})}, \ \widehat{\boldsymbol{R}}_w = \tilde{\lambda}_{w,1}^{(\hat{J})}\boldsymbol{I}, \quad \widehat{\boldsymbol{R}}_s = \widehat{\boldsymbol{R}}_x - \widehat{\boldsymbol{R}}_w. \tag{16}$$

Extensive simulations indicate that variation of either the noise probability density function (PDF) or the noise variance does not have much influence on the value of the parameter $\Delta\lambda_{max}$ defined in (12). Hence the proposed source enumeration method is robust to uncertainties in the noise PDF. The subspace spanned by the eigenvectors $\{\hat{\lambda}_{x,n}; n = 1,\dots,\hat{J}\}$ is identified as the signal subspace. Obviously, $\hat{J}$ may also be interpreted as the number of positive eigenvalues of $\widehat{\boldsymbol{R}}_s$. We shall designate the proposed method of source number estimation as the signal subspace identification (SSI) method.

3. SIMULATION RESULTS

We consider a uniform linear array of N sensors with J incoherent narrowband plane waves impinging from directions $\{\emptyset_1,\dots,\emptyset_J\}$ with respect to the array axis. The number of sensors is $N = 15$ in all examples, unless otherwise stated. We have compared the performance of the proposed SSI method with those of the MDL and EEE [9] methods for different noise distributions. The measure of performance is the probability of correct estimation defined as $P_c = P(\hat{J} = J)$.

In the first experiment, we compared the performance of SSI, MDL and EEE methods in Laplacian noise. Figure 1 shows the plots of P_c versus SNR for $L = 300$ snapshots. We considered three sources with equal SNR, located at $\emptyset_1 = 50°, \emptyset_2 = 60°$, and $\emptyset_3 = 70°$. For the same directions, plots of P_c versus L for a fixed SNR are shown in Fig. 2. It is seen that the SSI method provides the best performance. For the SSI method, P_c is very close to 1 for moderate values of SNR and L. The MDL method performs poorly for SNR less than -10dB. Figure 3 shows plots of P_c versus number of sensors N, at -12 dB SNR. It is seen that the performance of SSI keeps improving as N is increased, while the performance of EEE tends to saturate. Figure 4 illustrates the effect of variation of the number of sources J on the performance of the SSI method. At low SNR, P_c increases as J is reduced, as expected. For all the methods, $P_c = 1$ when $J = 0$.

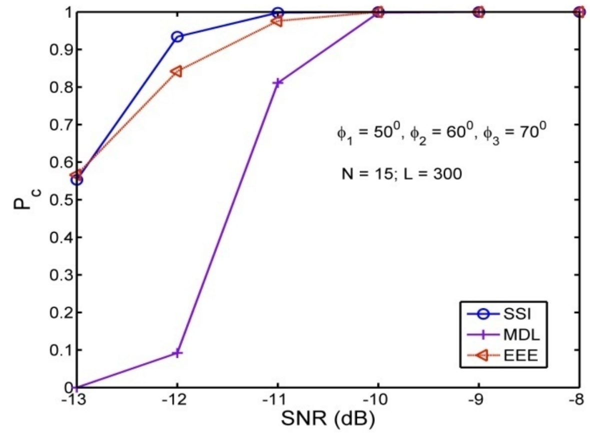

Fig. 1 Performance comparison of SSI, MDL and EEE for Laplacian noise. $N = 15, L = 300$. Three sources at 50^0, 60^0, and 70^0.

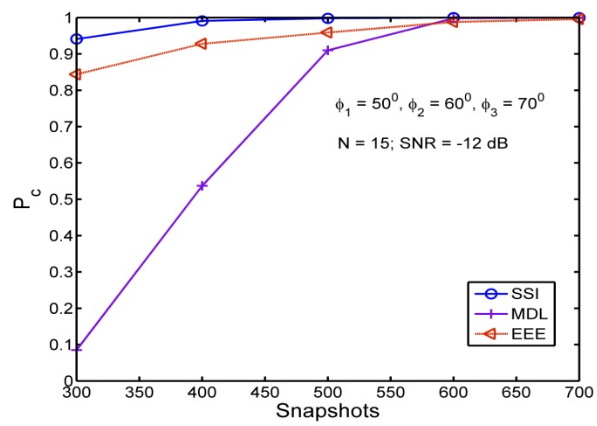

Fig. 2 Performance comparison of SSI, MDL and EEE for Laplacian noise. $N = 15, SNR = -12\ dB$. Three sources at 50^0, 60^0, and 70^0.

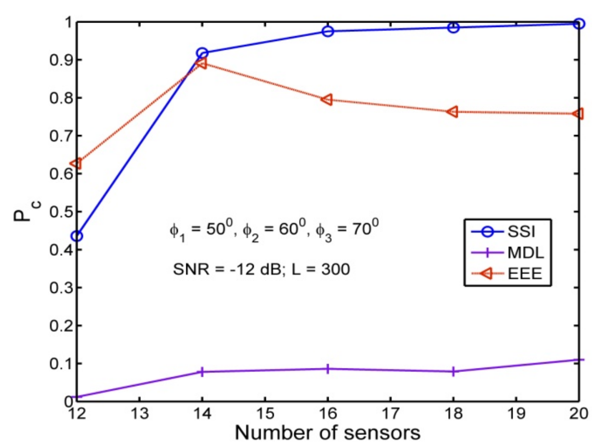

Fig. 3 Performance comparison of SSI, MDL and EEE for Laplacian noise. $N = 15, L = 300$, $SNR = -12\ dB$. Three sources at 50^0, 60^0, and 70^0.

In the next experiment, we investigated the effect of increasing non-Gaussianity of noise on the performance of the three methods. For this purpose, we considered the family of generalized Gaussian (GG) distributions and the family of 2-component zero-mean Gaussian mixture (GM) distributions to model noise. The PDF of GG noise with unit variance is given by $f_{GG}(x) = B(p)\exp[-C(p)|x|^p]$, $p > 0$, where $C(p) = p\left(\frac{\Gamma(3/p)}{\Gamma(1/p)}\right)^{p/2}$, $\Gamma(.)$ is the gamma function, and $B(p)$ is a normalization constant. The PDF is heavy-tailed for $p < 2$, and the heaviness of the tail increases as p is reduced. Noise is Gaussian for $p = 2$, and Laplacian for $p = 1$. Figure 5 shows plots of P_c versus the GG noise exponent p, for all the methods under consideration. The performance of SSI is seen to be better than that of the other methods. Figure 6 shows plots of P_c versus SNR for SSI for different noise distributions, viz. Gaussian, Laplacian, GG with $p = 0.5$, GM with ($\alpha_2/\alpha_1 = 0.1$, $\sigma_2^2/\sigma_1^2 = 100$), and GM with ($\alpha_2/\alpha_1 = 0.1$, $\sigma_2^2/\sigma_1^2 = 1000$). Here, α_1 and α_2 denote the probabilities of the GM components, and σ_1^2 and σ_2^2 are the respective variances. It is seen that, in all cases, P_c is very close to 1 for the SSI method if SNR exceeds -10 dB.

In another experiment, we compared the resolving capability of the methods in Laplacian noise, by considering two sources with a small angular separation. Figure 7 shows plots of P_c versus angular separation, when one source is at a fixed bearing of $\emptyset_1 = 50°$. Once again, it is seen that the SSI method outperforms the other two.

4. CONCLUSION

In this paper we presented a new method of source number estimation in non-Gaussian noise. The first step in the SSI method involves estimation of the noise correlation matrix. The noise variance is then estimated using a heuristic approach. Finally, the array signal correlation matrix is estimated and the number of sources is determined as the number of positive eigenvalues of the estimated signal correlation matrix. The method was applied to the problem of estimating the number of plane wave narrowband signals impinging on a uniform linear array. It was shown that the performance of the SSI method in non-Gaussian heavy-tailed noise is better than that of the recently proposed EEE method, and significantly better than that of the widely used MDL method.

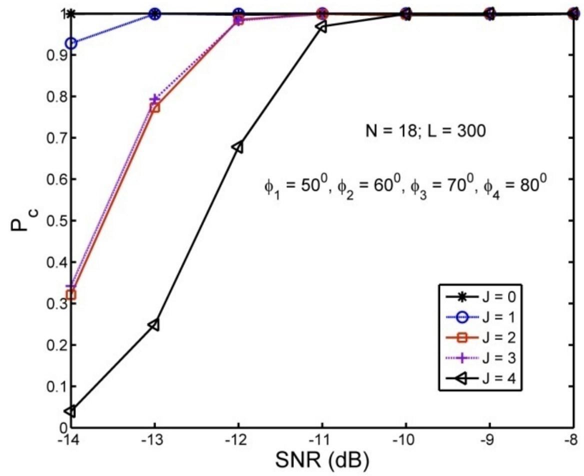

Fig. 4 Performance of SSI in Laplacian noise for different number of sources. $N = 18, L = 300$.

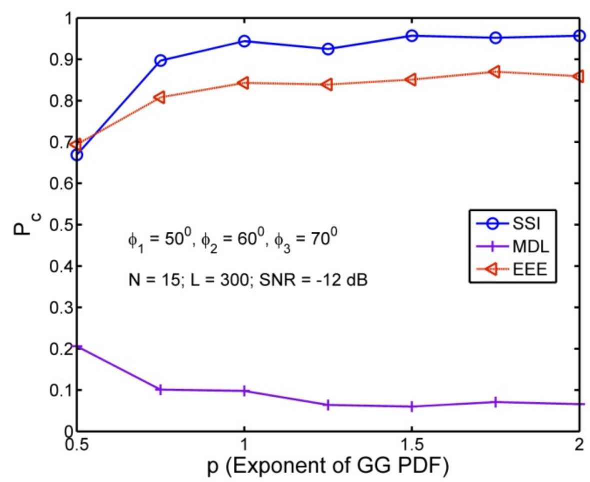

Fig. 5 Performance comparison of SSI, MDL and EEE for different values of exponent p of GG noise PDF. $L = 300, SNR = -12\ dB$. Three sources at 50^0, 60^0, and 70^0.

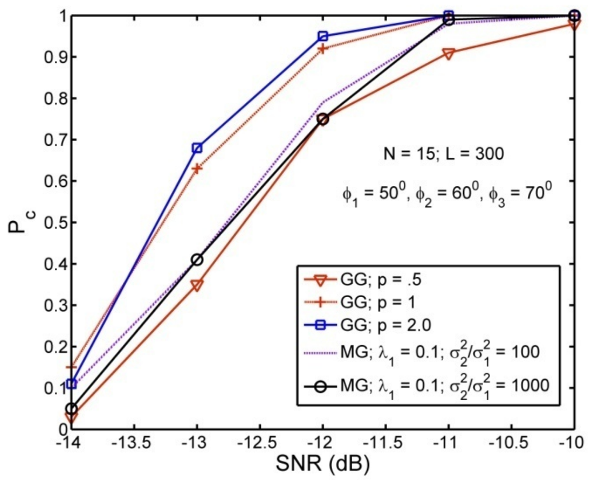

Fig. 6 Performance of SSI for different noise distributions. $N = 15, L = 300$. Three sources at 50^0, 60^0, and 70^0.

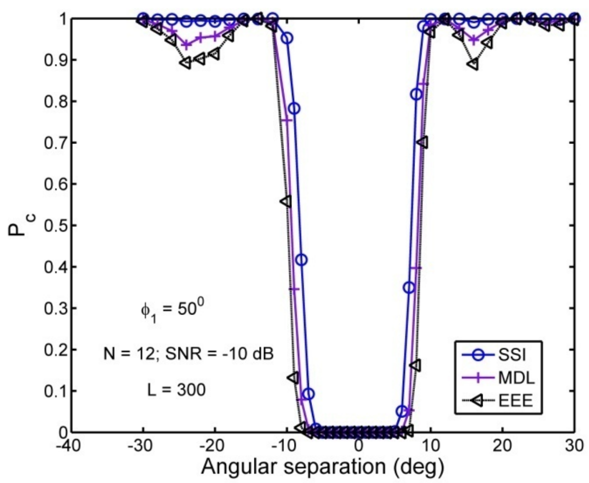

Fig. 7. Comparison of resolving capability of SSI, MDL and EEE. Plots of probability of correct estimation (P_c) vs. angular separation between two sources. One source is at 50°. $N = 12, L = 300, SNR = -10\ dB$.

REFERENCES

[1] H. Akaike, "A new look at the statistical model identification", *IEEE Trans. on Automatic Control*, vol. AC-19, no. 6, pp. 716- 723, December 1974.

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

[3] G. Schwarz, "Estimating the dimension of a model", *Annals of Statistics*, vol. 6, pp. 461-464, 1978.

[4] S. Valaee and P. Kabal, "An information theoretic approach to source enumeration in array signal processing", *IEEE Trans. on Signal Processing*, vol. 52, no. 5, pp. 1171-1178, May 2004.

[5] H. T. Wu, "Source number estimators using transformed Gerschgorin radii", *IEEE Trans. on Signal Processing*, vol. 43, no. 6, pp. 1325-1333, June 1995.

[6] R. F. Brcich, A. M. Zoubir, and P. Pelin, "Detection of sources using bootstrap techniques", *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 206-215, 2002.

[7] Z. Lu, and A. M Zoubir, "Source number estimation in impulsive noise environments using bootstrap techniques and robust statistics", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011.

[8] S. Kritchman and B. Nadler, "Nonparametric detection of the number of signals: hypothesis testing and random matrix theory", *IEEE Trans. on Signal Processing*, vol. 57, no. 10, pp.3930-3941, October 2009.

[9] H. Asadi and B. Seyfe, "Source number estimation via entropy estimation of eigenvalues (EEE) in Gaussian and non-Gaussian noise", ArXiv e-prints:1311.6051.

.