Source Enumeration Method Combining Gerschgorin Circle Transform and Generalized Bayesian Information Criterion in Large-scale Antenna Array

Chuanchuan Wang, Yonghu Zeng, and Liandong Wang

State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System
Luoyang, 471003, China
wangchuan1083@126.com, 471680179@qq.com, 874353112@qq.com

Abstract **─** A new source enumeration method based on gerschgorin circle transform and generalized Bayesian information criterion is devised, for the case that the antenna array observed signals are overlapped with spatial colored noise, and the number of antennas compared with that of snapshots meet the requirement of general asymptotic regime. Firstly, the sample covariance matrix of the observed signals is calculated, and then gerschgorin circle transformation is carried out on the sample covariance matrix. With the help of the more obvious distinction between the transformed signal gerschgorin circle radius and the noise gerschgorin circle radius, the observation statistic used to establish the likelihood function of the information theoretic criterion is constructed, by using the estimated values of the transformed sample covariance matrix's eigenvalues, and according to the idea of corrected Rao's score test, the observed statistics used to establish the likelihood function of the ITC are constructed. Based on the statistics, the source number is estimated by employing the generalized Bayesian information criterion (GBIC). The effectiveness of the proposed method is validated by experiments. Compared with the information theoretic criterion (ITC) methods and gerschgorin circle method (GDE), in Gaussian white noise, at the time $M/N \geq 1$, that is the relationship between the number of antennas and that of snapshots meets the requirement of the general asymptotic regime, the proposed method can accurately estimate the source number with 100% probability, the other methods failed. Compared with the ITC methods based on eigenvalue diagonal loading and GDE, in colored noise, at the time $M/N \geq 1$, the proposed method can accurately estimate the source number with 100% probability, the other methods failed. Compared with the methods based on random matrix theory, in colored noise, the proposed method can estimate the source number with 100% probability, but the estimation of other methods failed. The proposed method has wide applicability, in terms of the relationship between the numbers of antennas and snapshots, it is suitable for both general asymptotic regime and classical asymptotic system, and in terms of noise characteristics, it is suitable for both Gaussian white noise environment and colored noise environment.

Index Terms **─** Colored noise, corrected Rao's score test, general asymptotic regime, Gerschgorin circle transform, source enumeration.

I. INTRODUCTION

The estimation of the emitters' number has important applications in many fields, such as phased array radar, communications, brain imaging, neural networks, speech signal separation and direction of arrival estimation [1-8]. The classical methods for source enumeration are essentially based on the statistical analysis theory of observed data and their moment functions. For example, the hypothesis testing methods and information theoretic criterion (ITC) methods are commonly used for source enumeration, which mainly make use of the statistical distribution of observed data and the statistics of sample eigenvalues [9]. Among the classical source enumeration methods, the hypothesis testing methods include spherical test [10] and eigenvalue detection [11], which are mainly used to construct the observation statistics for hypothesis testing and set the decision threshold by using the statistical distribution law of sample eigenvalues. The ITC methods include Akaike information criterion (AIC) [12], Bayesian information criterion (BIC) [13], minimum description length (MDL) [14] and Predictive description length (PDL) [15], etc., usually assume that the observed data are Gaussian distribution, and then establish a criterion for estimating the number of sources according to the likelihood function of the joint probability distribution of the observed data. The expression of source enumeration is a function of the sample eigenvalues. A new source enumeration method based on higher-order tensors is presented in [16]. All these methods are applicable to Gaussian white noise environment [1-2, 16-17]. The main methods for source enumeration in colored noise environment are gerschgorin circle method [18] and

Submitted On: March 30, 2020
Accepted On: May 17, 2020

ITC methods based on diagonal loading [2, 19]. The performance of ITC methods have been studied in [17], results show that the methods are suitable for small-scale array signals whose sample number is much larger than the number of antennas. The above source enumeration methods are mainly based on the classical asymptotic system, that is, the dimension of the observed data matrix is fixed and the number of snapshots tends to be infinite.

However, in large-scale antenna arrays such as phased array radar and Multiple Input Multiple Output (MIMO) systems, due to the limitation of data storage space and the real-time requirement of signal processing, the observed data is often difficult to meet the condition that the number of snapshots is much larger than that of antennas, and it usually belongs to high-dimensional limited sampling data or even small sampling data. That is, the number of snapshots is in the same order of magnitude as that of antennas, or even less than the number of antennas. As to large-scale array observed data, the proportional relationship between the number of snapshots and that of antennas often does not meet the requirements of classical statistical theory, so the emergence of large-scale array brings new challenges to the classical source enumeration methods [20-21].

At present, the source enumeration in general asymptotic regime is mainly based on random matrix theory, including RMT-AIC method [4], BN-AIC method [5], BIC-variant method [6], LS-MDL method [7], the estimation method based on spike model [22], etc., and these methods are applicable when the number of antennas is less than that of snapshots. As to the estimation method based on spherical test and the estimation method based on modified Rao score test [22], they are applicable when the number of antennas is more than, less than or equal to the number of snapshots. All these methods are not only suitable for source enumeration in general asymptotic regime, but also suitable for classical asymptotic system. However, these methods are only applicable to white noise environment, but fail in colored noise environment [1, 4-7, 22].

Comprehensive analysis shows that at this stage, there is a lack of source enumeration method which is suitable for both classical asymptotic system and general asymptotic regime, whether there is white noise or colored noise environment. Considering that in the actual signal environment, the proportional relationship between the number of antennas and the number of snapshots, and whether the noise of observed signal overlapped by Gaussian white noise or colored noise is unknown, therefore, it is necessary to develop a source enumeration method which is suitable for both classical asymptotic system and general asymptotic regime, and is applicable to both Gaussian white noise and colored noise. In this paper, a source enumeration method based on gerschgorin circle transform and generalized Bayesian information criterion is devised, which does not need to prejudge the relationship between the number of antennas and that of snapshots (applying conditions must be satisfied, the relationship between the number of antenna elements M, the number of sources K, and the number of snapshots N is: $M-K\geq 1$, $K<N$, M can be larger than, equal to or less than N), and whether the observed signal overlapped noise is Gaussian white noise or colored noise. The number of narrowband signal sources such as communications can be blindly estimated in the complex electromagnetic environment.

The remainder of the paper is organized as follows. Section II presents the model of source enumeration problem. Section III gives the proposed source enumeration method. Section IV describes experiment results that validate the proposed method. Finally, the conclusions are drawn in Section V.

II. MATHEMATICAL MODEL OF SOURCE ENUMERATION

Suppose there are far-field signals whose number is K incidenting from the directions $\theta_1, \theta_2, \cdots, \theta_K$ onto an antenna array, and the number of antennas is M. at the sampling time t, the observed signals by the array is expressed as,

$$\boldsymbol{X}(\mathrm{t}) = \sum_{k=1}^{K} a(\theta_k)\, s_k(t) + \boldsymbol{w}(t) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(\mathrm{t}) + \boldsymbol{w}(t), \tag{1}$$

where $\boldsymbol{X}(\mathrm{t}) = [\boldsymbol{X}_1(t), \boldsymbol{X}_2(t), \ldots, \boldsymbol{X}_M(t)]^{\mathrm{T}}$ (the superscript T represents transpose) is the observed signal vector, $a(\theta_k)$ is the array direction vector, $\boldsymbol{A}(\boldsymbol{\theta}) = [a(\theta_1), a(\theta_2), \cdots, a(\theta_K)]$ is the matrix composed of direction vectors, $\theta = [\theta_1, \theta_2, \cdots, \theta_K]^{\mathrm{T}}$ is the incoming wave angle parameter vector of the signals, $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_K(t)]^{\mathrm{T}}$ is the incident signal vector, $\boldsymbol{w}(t) = [w_1(t), w_2(t), \cdots, w_M(t)]^{\mathrm{T}}$ is the additive noise vector, the sampling time is $t = 1,2, \cdots, N$, and N is the number of snapshots. The basic assumptions of the array observed signal model shown in formula (1) are as follows [22]:

(1) The incident signals are narrowband stationary signals independent of each other, which satisfy the mean $\mathrm{E}\{\boldsymbol{s}(t)\} = 0$ and covariance matrix $\mathrm{E}\{\boldsymbol{s}(t)\boldsymbol{s}^H(t)\} = diag\{p_{s_1}, p_{s_2}, \cdots, p_{s_K}\} \triangleq P_{\boldsymbol{s}} \in R^{K\times K}$, where p_{s_k} is the power of the $k-$th signal;

(2) The superimposed noise in the observed signal vector is additive noise (Gaussian white noise or colored noise);

(3) The number of incident signals is less than that of antennas and snapshots at the same time, that is $K < \min(M, N)$;

(4) The incident signals propagate in ideal space, and the antennas have omni-directional consistency.

III. THE PROPOSED SOURCE ENUMERATION METHOD

A. The principle of the proposed method

In practice, the sample data received by antenna array contains noise, and it may not be an ideal Gaussian white noise, but a complex spatial colored noise. In the complex spatial colored noise environment, the noise eigenvalue part of the covariance matrix of the received data will become very divergent and will not vibrate near the noise power like the Gaussian white noise's eigenvalue part. This problem caused by colored noise will invalidate various algorithms for source enumeration using hypothesis testing and ITC. As to the source enumeration method based on gerschgorin circle theorem, and the methods based on eigenvalue diagonal loading combined with ITC are usually only applicable to the classical asymptotic system, that is, the relationship between the number of antennas M and that of snapshots N is: M is fixed and $M/N \ll 1$. While under the general asymptotic regime, the relationship between the number of antennas and that of signal samples is that M and N tend to infinity at the same rate, $M, N \to \infty$ and $M/N \to c \in (0, \infty)$, the above methods usually fail to estimate the source number, regardless of whether the noise is Gaussian white noise or colored noise.

The existing source enumeration methods based on random matrix theory cannot be applied to estimate the source number, in the case of observed signal overlapped with colored noise in general asymptotic regime [1, 4-7, 22]. Through the analysis of the eigenvalues of the observed signals' covariance matrix, it is found that the noise eigenvalues are very divergent in the colored noise environment. As to the source enumeration methods based on gerschgorin circle theorem, they can be used to estimate the source number in Gaussian white noise or colored noise in the classical asymptotic system. When applying these methods, it is necessary to make a special transformation of the observed signals' covariance matrix. And after the transformation, there will be a more obvious distinction between the signal gerschgorin circle radius and the noise gerschgorin circle radius. In order to estimate the source number under the condition of observed signals overlapping with colored noise in general asymptotic regime, with the help of the idea of gerschgorin circle transformation, the sample covariance matrix of the observed signals is calculated first, and makes the gerschgorin circle transformation to the sample covariance matrix, then we get the more obvious distinguishing between the signal gerschgorin circle radius and the noise gerschgorin circle radius after the transformation. According to the idea of corrected Rao's score test (CRST), it can be used to detect the structural characteristics of large-dimensional covariance matrix [23]. The spherical test statistics in CRST can test whether the covariance matrix of the observed data's noise part is proportional to the unit matrix. According to this principle, based on the estimated eigenvalues of the transformed sample covariance matrix, the observed statistics used to establish the likelihood function of the ITC are constructed, and on this basis the source number is estimated by the generalized Bayesian information criterion (GBIC). The conventional BIC yields unsatisfactory results, especially in some difficult conditions, such as small sample sizes, low signal-to-noise ratios (SNRs), close spacing and high correlation between the sources. To improve its performance, Lu et al. [24] proposed a generalized Bayesian information criterion, by incorporating the density of the sample eigenvalues or corresponding statistics.

The proposed method improves the existing source enumeration method based on corrected Rao's score test [22], which can be used not only in the classical asymptotic system, but also in the general asymptotic regime, whether the observed signals are overlapped with Gaussian white noise or colored noise.

B. Specific steps of the proposed method

The specific steps of the proposed method are as follows.

Step 1: assume that the antenna array has M elements, and the observed signals obtained by one measurement can be expressed as $\boldsymbol{X}(\mathrm{t}) = [\boldsymbol{X}_1(t), \boldsymbol{X}_2(t), \dots, \boldsymbol{X}_M(t)]^{\mathrm{T}}$ (superscript T represents transpose). The sampling time is $t = 1,2,\cdots,N$, N is the number of snapshots, and the covariance matrix of the observed signals is calculated as $\boldsymbol{R}(t) = \frac{\boldsymbol{X}(t)\cdot\boldsymbol{X}^H(t)}{N}$.

Step 2: block the sample covariance matrix $\boldsymbol{R}(t)$ as follows:

$$\boldsymbol{R}(t) = \begin{bmatrix} \boldsymbol{R}'(t) & \hat{r} \\ \hat{r}^H & \hat{r}_{MM} \end{bmatrix}.$$

The $M-1$-dimensional square matrix $\boldsymbol{R}'(t)$ is the covariance matrix of the observed data $\boldsymbol{X}'(t)$ obtained by removing the last element of the antenna array. For convenience, the following $\boldsymbol{R}(t)$ and $\boldsymbol{R}'(t)$ will be abbreviated as $\boldsymbol{R}$ and $\boldsymbol{R}'$. Take the characteristic matrix of $\boldsymbol{R}$, and it is recorded as $\boldsymbol{V}$, then we construct a unitary transformation matrix $\boldsymbol{T}$:

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{V} & \boldsymbol{0} \\ \boldsymbol{0}^H & \boldsymbol{1} \end{bmatrix}. \quad (2)$$

The covariance matrix of the observed signals is unitary transformed by the constructed unitary transformation matrix $\boldsymbol{T}$,

$$\boldsymbol{R}_T=\boldsymbol{T}^H\boldsymbol{R}\boldsymbol{T}=\begin{bmatrix}\boldsymbol{V}^H\boldsymbol{R}'\boldsymbol{V} & \boldsymbol{V}^H\hat{\boldsymbol{r}}\\ \hat{\boldsymbol{r}}^H\boldsymbol{V} & \hat{\boldsymbol{r}}_{MM}\end{bmatrix}=\begin{bmatrix}\gamma_1 & 0 & \cdots & 0 & \rho_1\\ 0 & \gamma_2 & \cdots & 0 & \rho_2\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & \gamma_{M-1} & \rho_{M-1}\\ \rho_1^* & \rho_2^* & \cdots & \rho_{M-1}^* & \hat{\boldsymbol{r}}_{MM}\end{bmatrix}. \quad (3)$$

Step 3: write the spectral decomposition of M-dimensional observed data covariance matrix $\boldsymbol{R}$, and $M-1$-dimensional observed data covariance matrix $\boldsymbol{R}'$ as: $\boldsymbol{R}=\sum_{i=1}^{M}\lambda_i\boldsymbol{u}_i\boldsymbol{u}_i^H$, $\boldsymbol{R}'=\sum_{i=1}^{M-1}\gamma_i\boldsymbol{v}_i\boldsymbol{v}_i^H$.

Step 4: do Eigen-Decomposition to the M-dimensional observed data covariance matrix $\boldsymbol{R}$ and the $M-1$-dimensional observed data covariance matrix $\boldsymbol{R}'$, which are respectively expressed as:

$$\boldsymbol{R}=\boldsymbol{U}\boldsymbol{\Sigma}_\lambda\boldsymbol{U}^H, \boldsymbol{R}'=\boldsymbol{V}\boldsymbol{\Sigma}_\lambda'\boldsymbol{V}^H.$$

Divide $\boldsymbol{U},\boldsymbol{V}$ and $\boldsymbol{\Sigma}_\lambda$ into blocks:

$$\boldsymbol{U}=\begin{bmatrix}u_{11} & \cdots & u_{1M}\\ \vdots & \ddots & \vdots\\ u_{M1} & \cdots & u_{MM}\end{bmatrix}=\begin{bmatrix}\boldsymbol{U}' & \boldsymbol{u}_M'\\ \boldsymbol{e}^H & u_{MM}\end{bmatrix}$$

$$\boldsymbol{V}=[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots\ \boldsymbol{v}_{M-1}]$$

$$\boldsymbol{\Sigma}_\lambda=\begin{bmatrix}\boldsymbol{\Sigma}_\lambda' & \cdots & \\ \vdots & \ddots & \vdots\\ & \cdots & \lambda_M\end{bmatrix},$$

where $\boldsymbol{U}'=[\boldsymbol{u}_1'\ \boldsymbol{u}_2'\cdots\boldsymbol{u}_{M-1}']$, $\boldsymbol{u}_i'=\left[\boldsymbol{u}_{1i}\ \boldsymbol{u}_{2i}\cdots\ \boldsymbol{u}_{(M-1)i}\right]^H$ $(i=1,2,\ldots,M)$, $\boldsymbol{e}=\left[\boldsymbol{u}_{M1}\ \boldsymbol{u}_{M2}\ \cdots\ \boldsymbol{u}_{M(M-1)}\right]^H$, $\boldsymbol{\Sigma}_\lambda'=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_{M-1})$.

Step 5: as in formula (2), by applying the unitary transformation matrix $\boldsymbol{T}$, the block matrix $\boldsymbol{U}$ and $\boldsymbol{V}$ to do unitary transformation of $\boldsymbol{\Sigma}_\lambda$, the following can be obtained:

$$\boldsymbol{R}_T=\boldsymbol{T}^H\boldsymbol{U}\boldsymbol{\Sigma}_\lambda\boldsymbol{U}^H\boldsymbol{T}=\begin{bmatrix}\boldsymbol{V} & \boldsymbol{0}\\ \boldsymbol{0}^H & \boldsymbol{1}\end{bmatrix}^H\begin{bmatrix}\boldsymbol{U}' & \boldsymbol{u}_M'\\ \boldsymbol{e}^H & u_{MM}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_\lambda' & \\ & \lambda_M\end{bmatrix}\begin{bmatrix}\boldsymbol{U}' & \boldsymbol{u}_M'\\ \boldsymbol{e}^H & u_{MM}\end{bmatrix}\begin{bmatrix}\boldsymbol{V} & \boldsymbol{0}\\ \boldsymbol{0}^H & \boldsymbol{1}\end{bmatrix}=\begin{bmatrix}\boldsymbol{V}^H\boldsymbol{U}'\boldsymbol{\Sigma}_\lambda'\boldsymbol{U}'^H\boldsymbol{V}+\lambda_M\boldsymbol{V}^H\boldsymbol{u}_M'\boldsymbol{u}_M'^H\boldsymbol{V} & \boldsymbol{V}^H\boldsymbol{U}'\boldsymbol{\Sigma}_\lambda'\boldsymbol{e}+\lambda_M\boldsymbol{V}^H\boldsymbol{u}_M'u_{MM}\\ \boldsymbol{e}^H\boldsymbol{\Sigma}_\lambda'\boldsymbol{U}'^H\boldsymbol{V}+\lambda_M u_{MM}\boldsymbol{u}_M'^H\boldsymbol{V} & \boldsymbol{e}^H\boldsymbol{\Sigma}_\lambda'\boldsymbol{e}+\lambda_M u_{MM}u_{MM}^*\end{bmatrix}. \quad (4)$$

Step 6: similar to the formula (3) in **Step 2**, take the first to the $M-1$ rows and the M column of the formula (4) in **Step 5**, and it is expressed as ρ_i' $(i=1,2,\cdots,M-1)$, then we take its absolute value $|\rho|_i'$ $(i=1,2,\cdots,M-1)$, write it as $r_i=|\rho|_i'$, and it can be regarded as the estimated values of the $M-1$ eigenvalue of the covariance matrix $\boldsymbol{R}'$.

Step 7: as to r_i $(i=1,2,\cdots,M-1)$, ordering that $r_M=r_{M-1}$, expressing r_i and r_M as a sequence $r_i'=r_i$ $(i=1,2,\cdots,M)$, judging whether the values of r_i' are arranged in the order of $r_1'\geq r_2'\geq\cdots\geq r_M'$, if yes, the sequence r_i' is retained and proceed to the next step; if the values of r_i' are arranged in the order of $r_1'\leq r_2'\leq\cdots\leq r_M'$, the values of r_i' would be in reverse order, that is, the values of r_i' would be arranged in the order from the largest to the smallest, and they are expressed as $r_i'^f$, the serial number is $i=1,2,\cdots,M$. For convenience, we express r_i' or $r_i'^f$ as $r_i'^{new}=r_i'$ or $r_i'^f$.

Step 8: according to the eigenvalue sequence $r_i'^{new}$, the modified Rao score test method is introduced to estimate the number of sources. Defining $\widehat{\mathbf{R}}_W^{(k)}=\mathrm{diag}\{r_{k+1}'^{new},\cdots,r_M'^{new}\}$, and $\boldsymbol{T}^{(k)}$ is calculated by the following formula:

$$\boldsymbol{T}^{(k)}=\frac{1}{\sqrt{\tilde{\tau}^{(k)}}}Tr\left[\left(\frac{1}{\hat{\sigma}_k^2}\widehat{\mathbf{R}}_W^{(k)}-\mathbf{I}_{M-k}\right)^2\right]-(M-k)\tilde{c}_N^{(k)}, \quad (5)$$

where, $\tilde{\tau}^{(k)}=2(\tilde{c}_N^{(k)})^2\ (1+2\,\tilde{c}_N^{(k)})$, $\tilde{c}_N^{(k)}=(M-k)/(N-1)$, $\hat{\sigma}_k^2=\frac{1}{M-k}\sum_{i=k+1}^{M}r_i'^{new}$.

Step 9: define the formula of Source Enumeration based on gerschgorin circle theorem and modified Rao score test as following:

$$\mathrm{GDE-CRSTGBIC}(k)=\left(\boldsymbol{T}^{(k)}\right)^2+(k+1)\log N. \quad (6)$$

Step 10: estimate the source number by the following formula:

$$\hat{K}=\mathrm{argmax}\ \mathrm{GDE-CRSTGBIC}(k),\quad k=1,2,\cdots,M-1. \quad (7)$$

IV. EXPERIMENTS AND ANALYSIS

The validation of the proposed method is carried out under the simulation

condition of the DELL9020MT personal computer, Intel (R) Core (TM) i7mur4770 CPU @ 3.40GHz Windows 64-bit operating system, and the simulation software is MATLAB R2010a. In order to fully verify the effectiveness of the proposed method (we name it as GDE-CRSTGBIC), the calculation results of the proposed method and the reference methods are compared, and three groups of tests are carried out.

Experiment 1: Comparison between the proposed method (GDE-CRSTGBIC) and the ITC methods (BIC, AIC, MDL, KIC), gerschgorin circle method (GDE), in the environment of Gaussian white noise. The experimental conditions are set as follows:

1) s_1 is a BPSK signal with a subpulse width of 3×10^{-7}s and a carrier frequency of 10MHz.

2) s_2 is a CW signal with a subpulse width of 1.5×10^{-5}s and a carrier frequency of 10MHz.

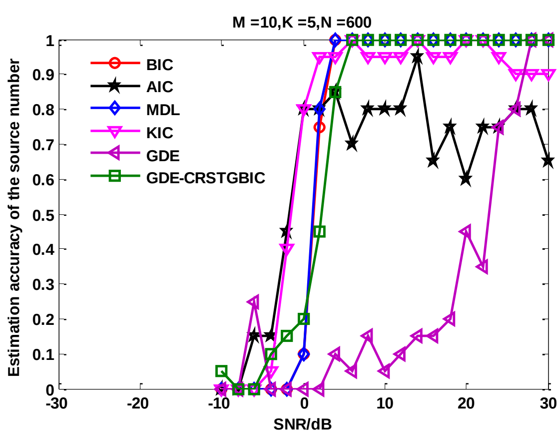
3) s_3 is a LFM signal with a carrier frequency of 10MHz and pulse repetition rate of 0.1MHz.

4) s_4 is a FSK signal with a subpulse width of 10^{-7}s. The carrier frequency varies with the binary baseband signal between 25MHz and 50MHz.

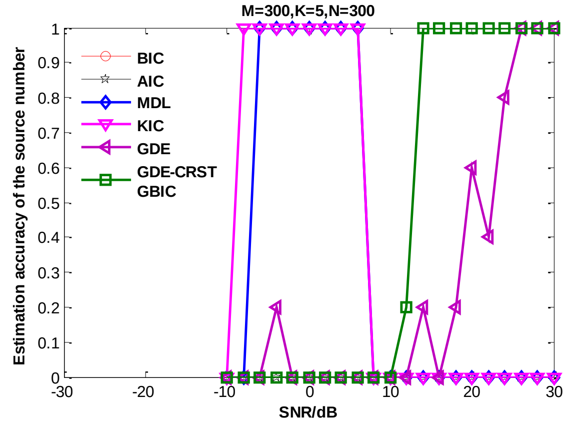
5) s_5 is a MPSK signal with a subpulse width of 4×10^{-7}s and a carrier frequency of 50MHz.

If the number of sources is set as $K = 4$, the source signals are composed of s_1, s_2, s_3 and s_5. If the number of sources is set as K=5, the source signals are composed of s_1~s_5. Set different number of array antenna elements M, and mixing matrix $\boldsymbol{A}$ is generated by random function, the sampling frequency is 120MHz, snapshots is N, the observed signals are overlapped with white Gaussian noise, the variation range of signal-to-noise ratio (SNR) is -10dB~30dB, step size is 2dB, 1000 Monte Carlo simulations are carried out on each SNR. The experimental results are shown in Figs. 1 (a)-(d). In addition, when $M = 340, K = 5, N = 300$, the histogram of the estimated source number at SNR = 15dB is shown in Figs. 2 (a)-(f).

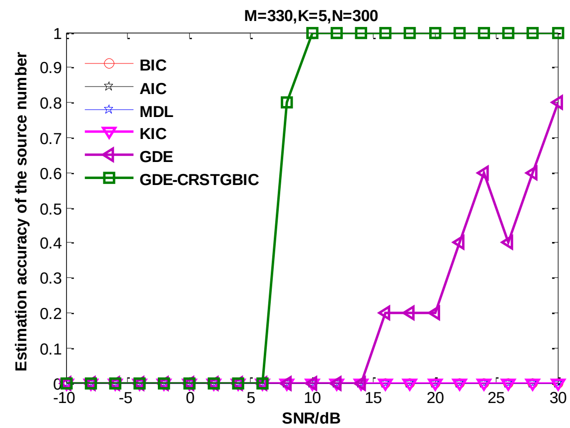
Figure 1 shows the comparison of results by the GDE-CRSTGBIC method and the ITC methods (BIC, AIC, MDL, KIC), gerschgorin circle method (GDE), in the Gaussian white noise environment. As can be seen from Fig. 1 (a), at this time $M/N \ll 1$, the relationship between the number of antennas and that of snapshots meets the requirements of the classical asymptotic system. Under the condition of Gaussian white noise, when the SNR is larger than 3dB, the GDE-CRSTGBIC method, MDL method and BIC method can accurately estimate the source number with 100% probability, but the gerschgorin circle method needs more than 26dB of the SNR to reach 100% probability. In Fig. 1 (b), Fig. 1 (c) and Fig. 1 (d), $\frac{M}{N} \geq 1$, so the relationship between the number of antennas and that of snapshots meets the requirement of the general asymptotic regime. Under the condition of Gaussian white noise, the GDE-CRSTGBIC method can accurately estimate the source number with 100% probability when the SNR is larger than 14dB, 8dB, and 7dB, respectively. Other ITC and GDE methods failed to estimate the source number.



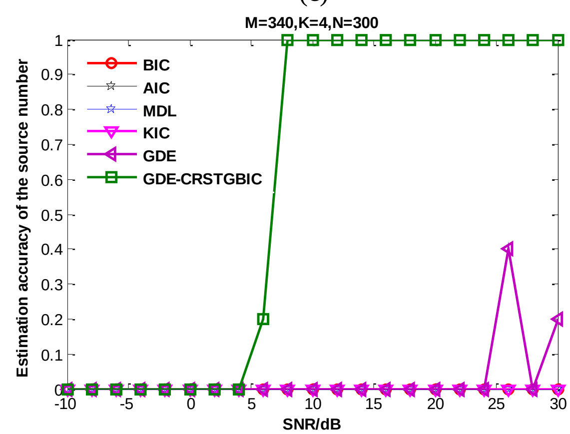
(a)



(b)



(c)

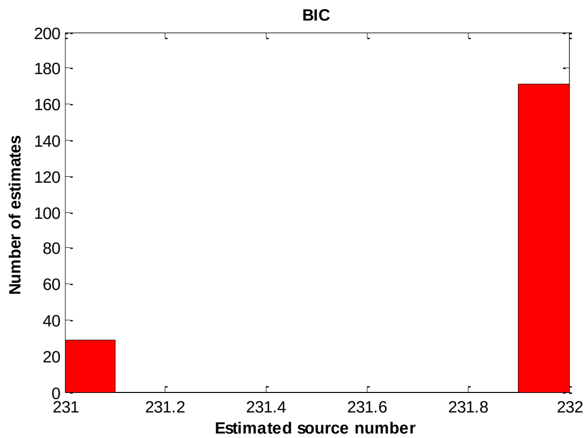


(d)

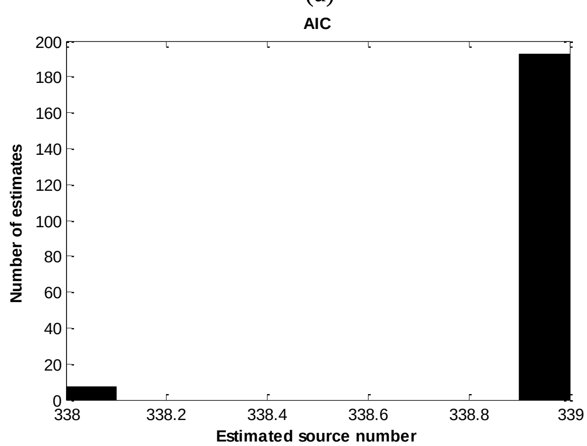
Fig. 1. The source number estimation results by proposed method are compared with those of the ITC methods and the GDE method under the condition of white noise.

Figure 2 shows, when $M = 340, K = 5, N = 300$, that is the relationship between the number of antennas and that of snapshots meets the requirement of the general asymptotic regime. In Gaussian white noise, the

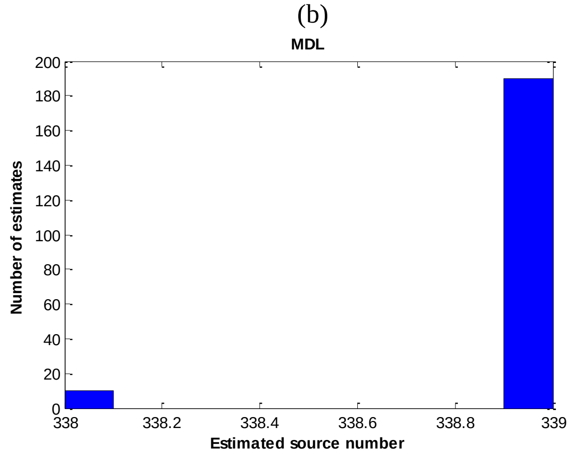
GDE-CRSTGBIC method can accurately estimate the source number with 100% probability when the SNR is 15dB. Other ITC and GDE methods failed to estimate the source number.



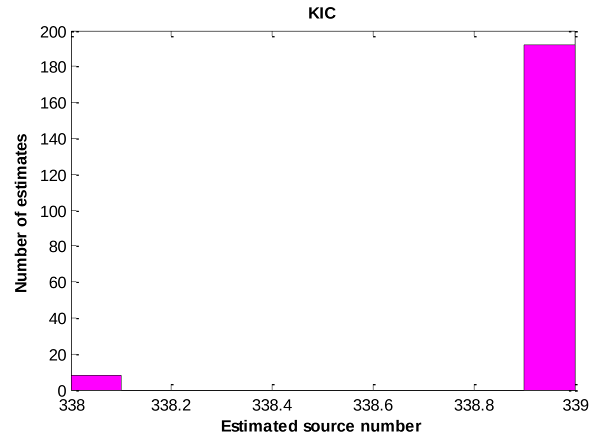
(a)



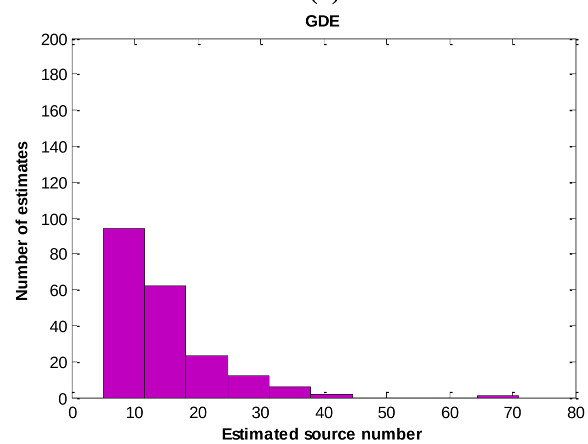
(b)



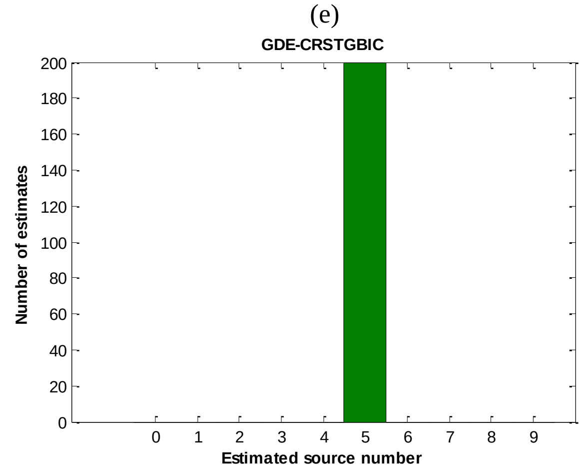
(c)



(d)



(e)



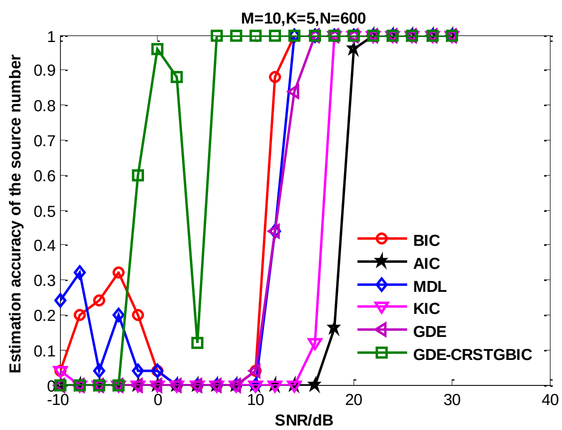
(f)

Fig. 2. Histogram of estimated source number at SNR $=$ 15dB when $M = 340, K = 5, N = 300$.

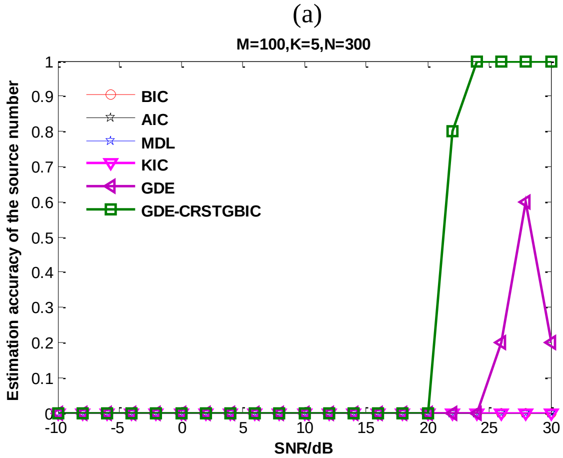
Experiment 2: The proposed method (GDE-

CRSTGBIC) is compared with the ITC methods (BIC, AIC, MDL, KIC) based on eigenvalue diagonal loading and gerschgorin circle method (GDE) in colored noise. The source signals are the same with those in ***Experiment 1***.

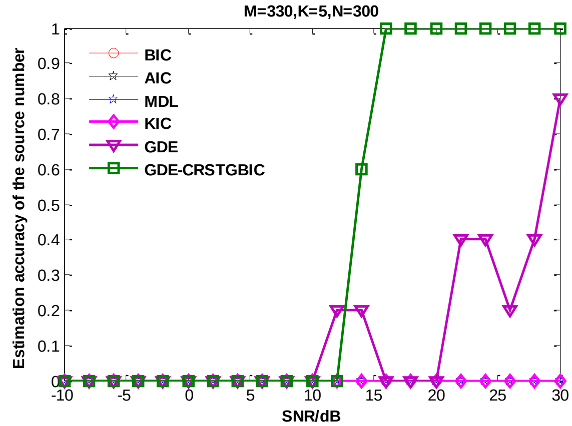
If the number of sources is set as $K = 4$, the source signals are composed of s_1, s_2, s_3 and s_5. If the number of sources is set as K=5, the source signals are composed of s_1~s_5. Set different number of array antenna elements M, mixing matrix $\boldsymbol{A}$ is generated by random function, sampling frequency is 120MHz, snapshots is N, observed signals are overlapped with spatial color noise, the elements of its covariance matrix are expressed as $n_{ik} = \sigma_n^2 0.9^{|i-k|}\exp[(\mathrm{j}(i\text{-}k)\pi/2)]$, $i, k = 1,2,\cdots, M$. σ_n is an adjustable parameter, which is used to set the SNRs of observed signals, the variation range of signal-to-noise ratio (SNR) is -10dB~30dB , step size is 2dB, 1000 Monte Carlo simulations are carried out on each SNR. The experimental results are shown in Figs. 3 (a)-(d). In addition, when $M = 340, K = 5, N = 300$, the histogram of the estimated source number at SNR = 20dB is shown in Figs. 4 (a)-(f).



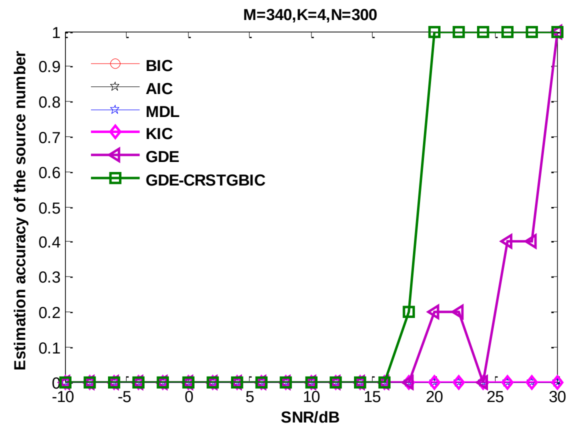
(a)



(b)



(c)



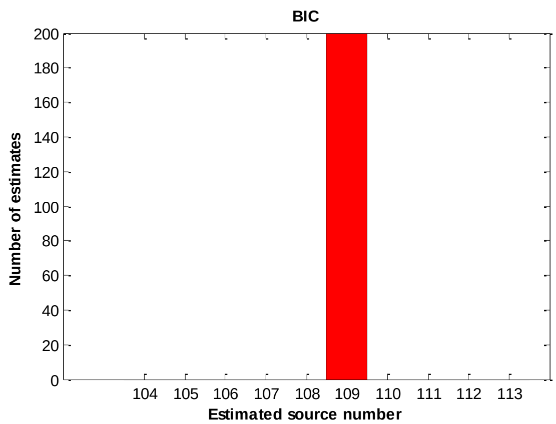
(d)

Fig. 3. Comparison of the results by the proposed method with those of the ITC methods and the GDE method in colored noise environment.

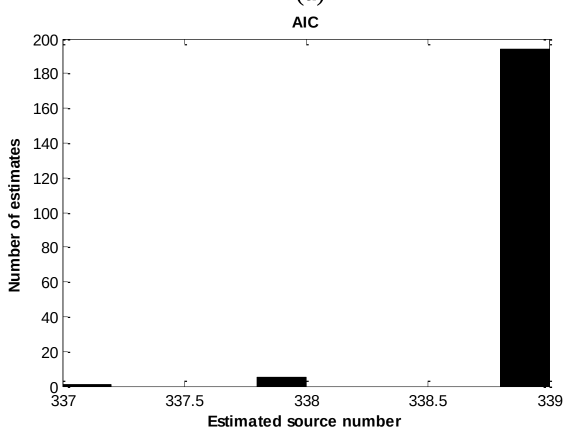
Figure 3 shows the comparison of results between the proposed method (GDE-CRSTGBIC method) and the ITC methods based on eigenvalue diagonal loading (BIC, AIC, MDL, KIC), gerschgorin circle method (GDE), in colored noise environment. In Fig. 3 (a), at this time $M/N \ll 1$, the relationship between the number of antennas and that of snapshots meets the requirements of the classical asymptotic system. Under the condition of colored noise, when the SNR is larger than 5dB, the GDE-CRSTGBIC method can accurately estimate the source number with 100% probability, and other methods need larger SNR. In Fig. 3 (b), Fig. 3 (c) and Fig. 3 (d), $\frac{M}{N} \geq 1$, so the relationship between the number of antennas and that of snapshots meets the requirements of general asymptotic regime. Under the condition of colored noise, when the SNRs are larger than 9 dB , 15 dB and 19 dB respectively, the source number can be estimated accurately with 100%

probability by the proposed method, while other methods failed.

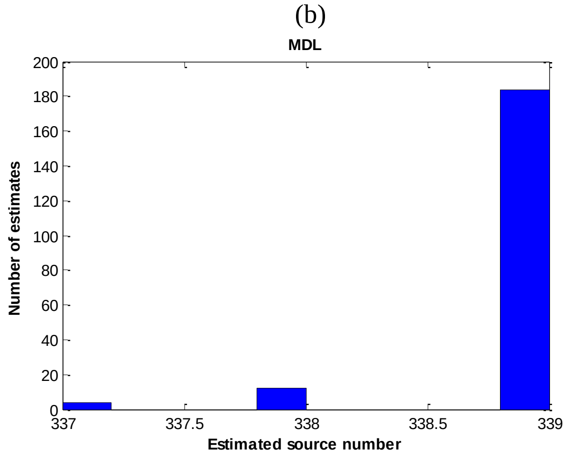
Figure 4 shows, when $M = 340, K = 5, N = 300$, that is the relationship between the number of antennas and that of snapshots meets the requirement of the general asymptotic regime. In colored noise, the GDE-CRSTGBIC method can accurately estimate the source number with 100% probability when the SNR is 20dB. Other ITC and GDE methods failed to estimate the source number.



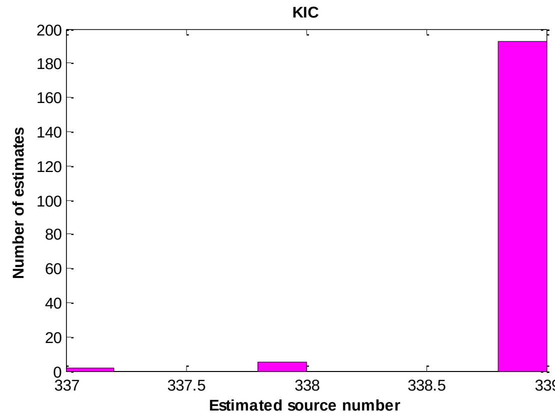
(a)



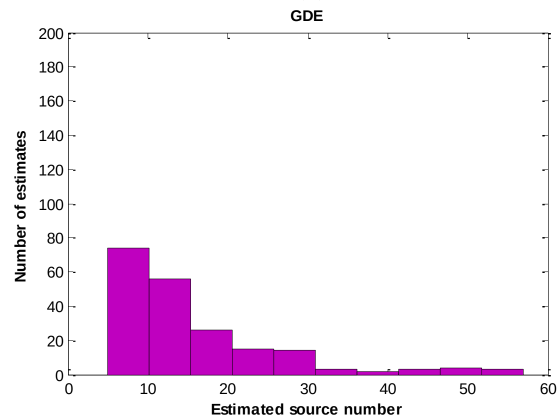
(b)



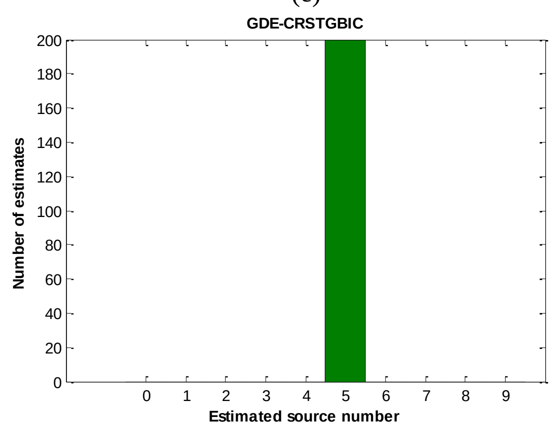
(c)



(d)



(e)



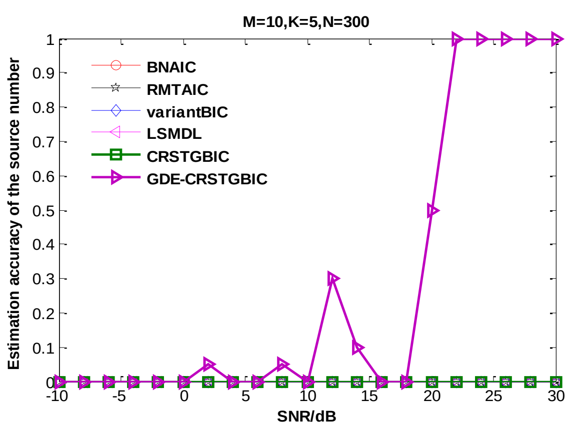
(f)

Fig. 4. Histogram of estimated source number at SNR = 20dB when $M = 340, K = 5, N = 300$.

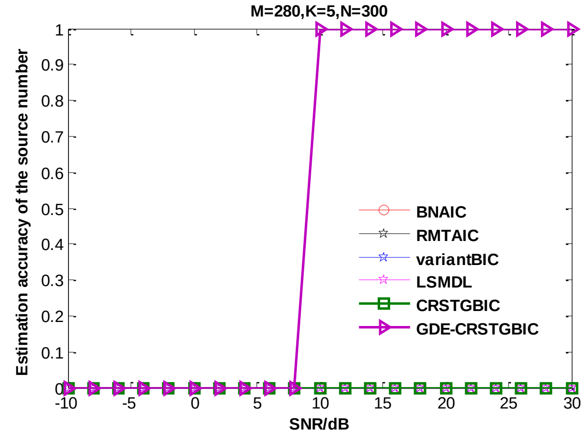
Experiment 3: Comparison of the results between the proposed method and methods based on random matrix theory (BN-AIC, RMT-AIC, BIC-variant, LS-MDL, CRST-GBIC), in colored noise environment.

The source signals used in this experiment are the same as those in ***Experiment 1***, and the number of source signals is 5. Set different number of the antennas M, mixing matrix $\boldsymbol{A}$ is generated by random function, sampling frequency is 120MHz, the number of snapshots is N, observed signals are overlapped with spatial color noise, the elements of its covariance matrix are expressed as $n_{ik} = \sigma_n^2 0.9^{|i-k|}\exp[(\mathrm{j}(i\text{-}k)\pi/2)]$, $i,k = 1,2,\cdots,M$. σ_n is used to set the signal-to-noise ratio (SNR) of observed signals, the variation range of SNR is -10dB~30dB, step size is 2dB, 1000 Monte Carlo simulations are carried out on each SNR. The experimental results are shown in Figs. 5(a)-(d). In addition, when $M = 340, K = 5, N = 300$, the histogram of the estimated source number at SNR = 20dB is shown in Figs. 6 (a)-(f).

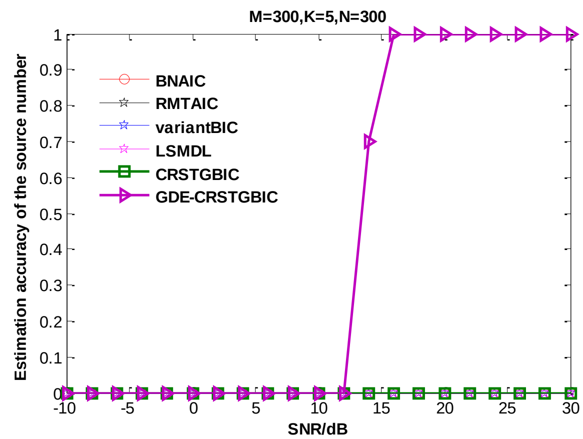
Figure 5 are the comparison of the results between the proposed method (GDE-CRSTGBIC) and the methods based on random matrix theory (BN-AIC, RMT-AIC, BIC-variant, LS-MDL, CRST-GBIC) in colored noise environment. As can be seen from Fig. 5 (a), at this time $M/N \ll 1$, the relationship between the number of antennas and that of snapshots meets the requirements of the classical asymptotic system. Under the condition of colored noise, the GDE-CRSTGBIC method compared with a variety of source enumeration methods based on random matrix theory, the former can accurately estimate the source number with 100% probability when SNR is larger than 23dB, but the other methods fail. In Fig. 5 (b), Fig. 5 (c) and Fig. 5 (d), $\frac{M}{N} \approx$ or ≥ 1, the relationship between the number of antennas and that of snapshots belongs to the classical asymptotic system. In colored noise, the proposed method can estimate the source number with 100% probability when SNRs are larger than 10dB, 15dB and 13dB respectively, but the estimation of other methods failed.



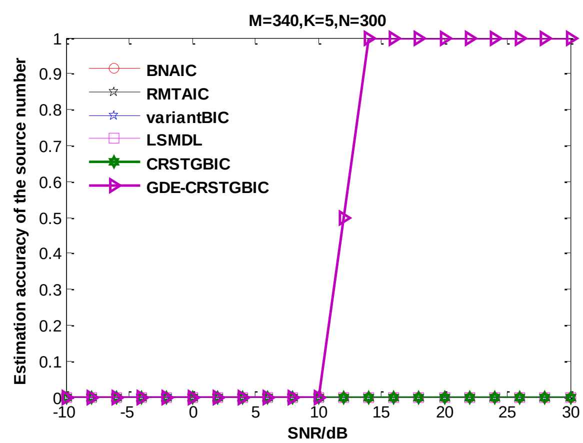
(a)



(b)



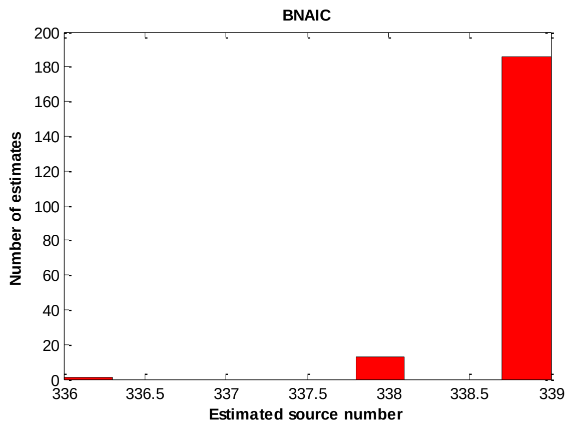
(c)



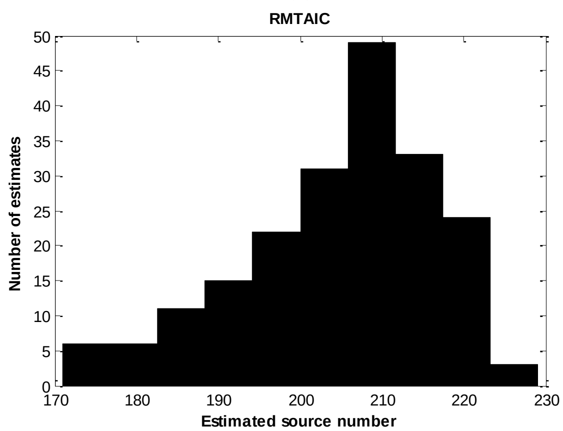
(d)

Fig. 5. Comparison of the results between the proposed method and the methods based on random matrix theory in colored noise environment.

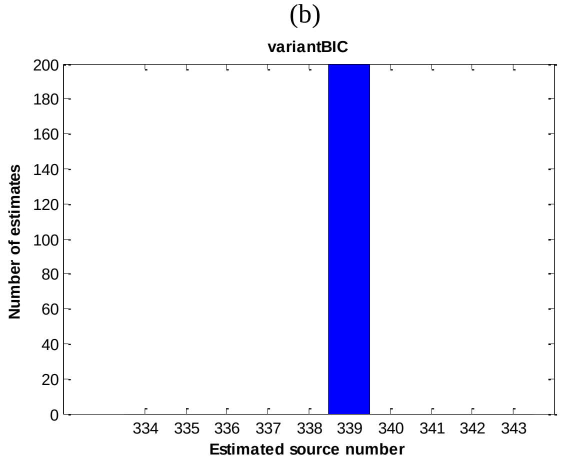
Figure 6 shows, when $M = 340, K = 5, N = 300$, that is the relationship between the number of antennas and that of snapshots meets the requirement of the general asymptotic regime. In colored noise, the GDE-CRSTGBIC method can accurately estimate the source number with 100% probability when the SNR is 20dB. Other RMT methods failed to estimate the source number.



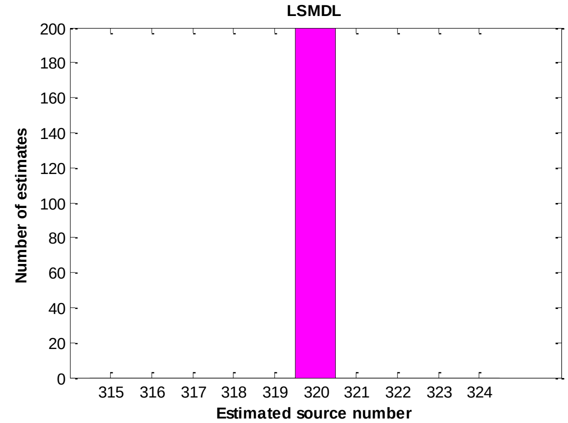
(a)



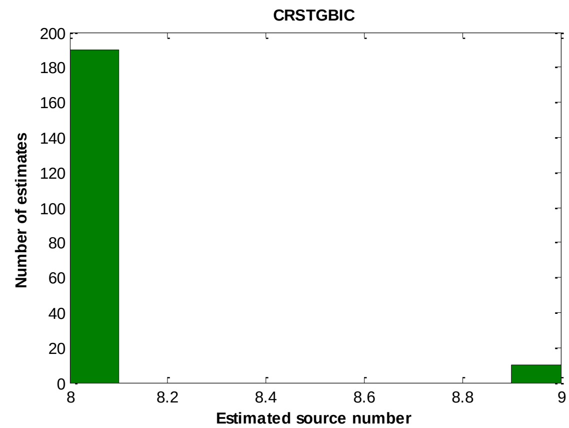
(b)



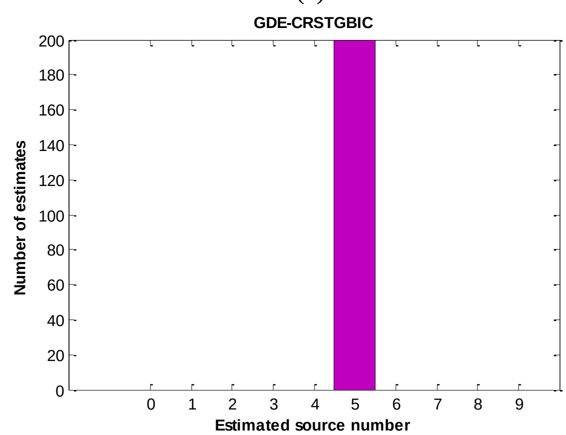
(c)



(d)



(e)



(f)

Fig. 6. Histogram of estimated source number at SNR = 20dB when $M = 340, K = 5, N = 300$.

V. CONCLUSION

A new method is proposed for source enumeration, under the condition that the observed signals are overlapped with spatial colored noise, and the number of antennas and that of snapshots meet the requirements of general asymptotic regime. The proposed method does not need to presuppose or assume the relationship between the number of antennas and that of snapshots, that is, it is applicable to the classical asymptotic system (the number of antennas is fixed and much smaller than the number of snapshots), and also suitable in a general asymptotic regime (the number of antennas is equal to or larger than that of snapshots). At the same time, the proposed method can be used to estimate the source number not only in the Gaussian white noise environment, but also in the colored noise environment. In view of the lack of source enumeration method in the general asymptotic regime, and the observed signals overlapping with colored noise, an effective approach is provided by us.

ACKNOWLEDGMENT

This research was funded by the National Natural Science Foundation of China (Grant No. 61801480).

REFERENCES

[1] H. Asadi and B. Seyfe, "Signal enumeration in Gaussian and non-Gaussian noise using entropy estimation of eigenvalues," *Digital Signal Process.*, vol. 78, pp. 163-174, Mar. 2018.

[2] M. J. Chen, G. Q. Long, and Z. R. Huang, "Source number estimation in the presence of nonuniform noise," *J. Signal Process.*, vol. 34, no. 2, pp. 134-139, Feb. 2018.

[3] C. C. Wang, Y. H. Zeng, W. H. Fu, and L. D. Wang, "Estimation method for an underdetermined mixing matrix based on maximum density point searching," *J. Xidian University*, vol. 46, no. 1, pp. 106-111, Feb. 2019.

[4] R. Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *IEEE Trans. Signal Process*., vol. 56, no. 7, pp. 2625-2638, July 2008.

[5] B. Nadler, "Nonparametric detection of signals by information theoretic criteria: performance analysis and an improved estimator," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2746-2756, May 2010.

[6] L. Huang, Y. Xiao, K. Liu, H. C. So, and J. K. Zhang, "Bayesian information criterion for source enumeration in large-scale adaptive antenna array," *IEEE Trans. Vehicular Technology*, vol. 65, no. 5, pp. 3018-3032, May 2016.

[7] L. Huang and H.C. So, "Source enumeration via MDL criterion based on linear shrinkage estimation of noise subspace covariance matrix," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4806-4821, Oct. 2013.

[8] C. C. Wang and R. Jia, "A source signal recovery method for underdetermined blind source separation based on shortest path," *Applied Computational Electromagnetics Society Journal*, vol. 35, no. 4, pp. 406-414, Apr. 2020.

[9] B. Jiang, A. N. Lu, and J. Xu, "An improved signal number estimation method based on information theoretic criteria in array processing," *2019 IEEE 11th International Conference on Communication Software and Networks*, Chongqing, China, pp. 193-197, June 12-June 15, 2019.

[10] P. Cung, J. Bohme, C. Mecklenbrauker, and A. Hero, "Detection of the number of signals using the Benjamini-Hochberg procedure," *IEEE Trans. Signal Process*., vol. 55, no. 6, pp. 2497-2508, June 2007.

[11] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automatic Control*, vol. AC-19(6), pp. 716-723, 1974.

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

[13] G. Schwarz, "Estimating the dimension of a model," *Annals Statistic*, vol. 6, no. 2, pp. 461-464, 1978.

[14] M. Wax, "Detection and localization of multiple sources via the stochastic signals model," *IEEE Trans. Signal Process.*, vol. 39, no. 11, pp. 2450-2456, Nov. 1991.

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

[16] Y. Xie, K. Xie, and S. L. Xie, "Source number estimation and effective channel order determination based on higher-order tensors," *Circuits, Systems, and Signal Process.*, published online: Apr. 2019.

[17] W. Cheng, Z. S. Zhang, and Z. J. He, "Information criterion-based source enumeration methods with comparison," *J. Xi'an Jiaotong University*, vol. 49, no. 8, pp. 38-44, Aug. 2015.

[18] 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, June 1995.

[19] Q. T. Zhang and K. M. Wong, "Information theoretic criteria for the determination of the number of signals in spatially correlated noise," *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1652-1663, Apr. 1993.

[20] Y. Y. Liu, X. Y. Sun, and G. H. Liu, "Source enumeration in large arrays using corrected Rao's score test and relatively few samples," *2017 25th European Signal Processing Conference (EUSIPCO)*, Kos Kilkis, Greece, pp. 1405-1409, Aug. 28-Sep. 2, 2017.

[21] Z. D. Bai, S. R. Zheng, and D. D. Jiang, *Large Dimensional Statistical Analysis*. Beijing: Higher Education Press, 2012.
[22] Y. Y. Liu, "Applications of large random matrix theory to array signal parameters estimation," Ph.D. dissertation, *Communication Engineering Dep., Jilin Univ.*, Changchun, China, 2017.
[23] D. D. Jiang, "Tests for large-dimensional covariance structure based on Rao's score test," *Journal of Multivariate Analysis*, no. 152, pp. 28-39, Aug. 2016.
[24] Z. H. Lu and A. M. Zoubir, "Generalized Bayesian information criterion for source enumeration in array processing," *IEEE Trans. Signal Process*., vol. 61, no. 6, pp. 1470-1480, Mar. 2013.

Chuanchuan Wang received the B.A., M. S. and Ph.D. degrees from Shi Jia Zhuang, Mechanical Engineering College, China, in 2007, 2009, and 2013, respectively.

Since 2014, he has been a Research Assistant with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System. He is the author of more than 40 articles, and co-authored 21 inventions. His research interests include blind signal processing and effectiveness evaluation theory.

Yonghu Zeng received the B.S., M.S. and Ph.D. degrees from National University of Defense Technology of China, in 1994, 1997, and 2004 respectively. Since 2012, he has been a Researcher with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System. He has co-authored more than 60 articles, 6 books, and 26 inventions.

His research interests include radar signal processing and effectiveness evaluation theory.

Liandong Wang received the B.S., M.S. and Ph.D. degrees from National University of Defense Technology of China, in 1989, 1993, and 2000 respectively. Since 2012, he has been a Researcher with the State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System. He has co-authored more than 100 articles, 10 books, and 30 inventions.

His research interests include radar signal processing, electromagnetic environment effects and effectiveness evaluation theory.