Received October 21, 2020, accepted October 26, 2020, date of publication November 4, 2020, date of current version November 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3035793

Intelligent Signal Detection Under Spatially Correlated Noise

LEI WANG[1], JIANG XUE[1], (Senior Member, IEEE), JOHN THOMPSON[2], AND JIA YU[3]

[1]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China
[2]School of Engineering, Institute for Digital Communications, The University of Edinburgh, Edinburgh EH9 3JL, U.K.
[3]Research Institute of 5G Communication and Big Data, Hangzhou 310001, China

Corresponding author: Jiang Xue (x.jiang@xjtu.edu.cn)

This work was supported in part by the Science and Technology Commission of the Central Military Commission, in part by the National Defense Foundation of China, in part by the National Defense Science and Technology Innovation Special Zone Project Sub-project under Grant 19-163-11-ZD-099-001-01-02-12, in part by the Science and Technology Research and Development Program of Shaanxi Provincial Science and Technology Department under Grant 019GY-041, and in part by the China University S&T Innovation Plan Guided by the Ministry of Education, Subproject of the Artificial Intelligence Algorithm Strategy Research Project.

ABSTRACT Detecting the signal of the antenna array is a major problem in theoretical research and practical application. In this paper, several new methods are given for the number of signals detection at first, secondly, a new method called Principal Component Analysis of Signal Estimation (PCASE) will be introduced which can simultaneously detect the number of signals and the direction of arrival. In recent decades, the signal detection method based on the information theory criterion has been widely studied. The problem has been adequately solved under the assumption of uncorrelated white noise. However, considering the actual situation of wireless communication, the noise is spatially correlated or the noise information is unknown. In this case, traditional methods such as Akaike's information criteria (AIC), minimum descriptive length (MDL) and sparse and parameter approach (SPA) will often lead to a wrong estimation. Therefore, this paper introduces an improved eigenvalue correction method for the number of signals, and applies it to two new methods: the improved eigenvalue gradient method (Im-EGM) and the improved eigen-increment stop rule (Im-EISR), and studies a new estimation algorithm based on signal cancellation (SC). In addition, previous algorithms for estimating the direction of arrival (such as MUSIC, ESPRIT) require the number of known signals to estimate the direction of arrival. Therefore, this paper proposes a new method called PCASE, which can estimate the number of signals and the direction of arrival at the same time. This method combines the SPA and the Principal Component Analysis method (PCA) in machine learning. Compared with the existing methods, the accuracy of these new methods is verified by Monte Carlo simulation.

INDEX TERMS Correlated noise, eigen-increment stopping rule, Gerschgorin disk estimator, principal component analysis, direction of arrival.

I. INTRODUCTION

Array signal processing is one of the critical technology for wireless signal processing. The purpose of array signal processing is to obtain some parameters of signal, such as the number of signals, direction of arrival and frequency, by processing the signal received by sensor array. In the research of array signal processing, it is important and realistic to determine the number of signals and signal locations. The direction estimation algorithms widely used in radar, sonar and mobile communication, such as MUSIC proposed by Schmidt [1], ESPRIT proposed by Roy and Kailath [2], all need to know the exact number of signals. If the estimated number of signals are not consistent with the actual number of signals, the performance of the above algorithm will decline sharply or even completely fail. Thus, this orientation direction has attracted the attention of many researchers in recent decades [3]–[16].

Eigenvalues and eigenvectors of the observed covariance matrix provides an effective method in multiple applications, such as pattern classification, econometrics, statistical inference and signal processing [3]. The well known

The associate editor coordinating the review of this manuscript and approving it for publication was Ahmed Farouk.

methods of Akaike's Information Criterion (AIC) and Minimum Descriptive Length (MDL) were proposed in [17] and [18], respectively. Both AIC and MDL criteria work perfectly for the scenario of uncorrelated signal/noise and use the same computational function but with different penalty terms. To improve the AIC and MDL criteria, a new statistical approach was introduced by Wax and Kailath in [19], which does not require any subjective threshold settings and the estimation of signals number is determined by minimization of AIC and MDL criterion. Thereafter, the traditional AIC and MDL criteria have been analyzed and improved in [13], [20].

However, most of the frequency spectrum of noise is mainly non-white low frequency spectrum, the white Gaussian noise passing through the channel is affected by the channel frequency and becomes colored, that is, the noise with uneven power spectral density function, called colored noise, which is correlated noise. The traditional criteria for estimating the number of signals are based on information theoretic criteria and perform well under the assumption of white Gaussian noise, but they always fail to detect the correct number of signals when the noise is spatially correlated. It should be noted that when the noise is low-frequency band white noise, it also can be correlated. Solving this problem has triggered enormous research interests in recent years.

Assuming the noise has a band covariance matrix structure, the signal estimation problem was studied in [4] and the spectral matrices were computed by using delayed blocks to eliminate the noise influence in [6]. Considering the correlated noise, new detection approaches based on Gerschgorin radii, called the Gerschgorin disk estimator (GDE), were introduced in [5]. Lu used the Gerschgorin disk estimator to estimate the number of sources for the minimal redundant array in [9]. Furthermore, an improved Gerschgorin disk estimator for source enumeration, which is robust to spatial non-uniform noise, was proposed by [10], and the method of applying the Gerschgorin disk estimator to blind source estimation was proposed by [21]. However, it is not easy to determine a threshold of the radii for the GDE algorithm to separate the signals and noise in practice. Using the analysis of the difference between different eigenvalues, the estimator based on the eigen-increment was introduced in [7], while the eigenvalue gradient methods (EGM) were investigated based on the difference or ratio of different eigenvalues in [8], [11]. The traditional $K-$means clustering algorithm was applied to estimate the number of signals in colored noise field in [12]. However, based on our tests, most of the methods mentioned above do not provide acceptable accuracy estimation results.

In addition, various methods have been proposed and analyzed for the far-field narrow-band signal location problem [1], [22]–[25], also known as direction of arrival (DOA) estimation, for array observation snapshots in recent decades. Parameter estimation is one of the main methods of DOA estimation. Recently, parameter estimation is carried out by studying the subspace of covariance matrix, which is implemented as a large snapshot of maximum likelihood (ML) method in the case of uncorrelated signal (see [26], [27]). However, when the number of snapshots is small or there is a correlation between signals, it also requires the number of signals, and there is no reliable performance. Moreover, the traditional direction finding algorithms (such as MUSIC and ESPRIT) require an accurate number of direction finding signals to estimate DOA. Therefore, an exact discretization-free method which is a sparse and parametric approach (SPA) proposed for uniform and sparse linear arrays in [28]. In the case of uncorrelated signals, SPA is a large number snapshot implementation of ML estimation, and the number of snapshots is consistent. SPA is suitable for any number of snapshots and is robust to signal dependencies, especially it does not require known user parameters. However, all the above methods are based on the assumption of white Gaussian noise. Therefore, assuming spatial correlation of noise, the introduction of Principal Component Analysis (PCA) method in [29] to deal with correlated noise can well solve the problem. Hence, DOA estimation under noise correlation is a practical problem to be solved in signal processing.

In this paper, the problem of how to reliably detect the number of co-channel signals in the presence of spatially correlated noise is investigated and new algorithms and methods are introduced. The main contributions of this paper are as follows:

- Two new approaches, improved Eigenvalue Grads Method (Im-EGM) and improved Eigen-Increment Stopping Rule (Im-EISR) will be introduced based on an improvement of eigenvalue correction. Increasing the difference between the eigenvalues will improve the accuracy of the signal detection algorithm.
- A new estimation algorithm is investigated based on Signal Cancellation (SC), which the robust threshold SC can be set for particular direction and frequency after the short time of training to achieve high estimation accuracy.
- A new method of principal component analysis of signal estimation (PCASE) is proposed, which combines SPA and PCA to solve the problem of spatial correlation noise. Different from other methods, this method can simultaneously detect the number of signals and estimate the DOA.

The rest of this paper is structured as follows: in section II, we describe the system model. Several new algorithms for detecting signals are presented, and the validity is verified by numerical experiments in section III. In section IV, we present the PCASE for detecting signal and estimating direction of arrival. Finally, section V concludes this paper.

Notation: In this paper, lowercase boldface letters represent vectors, while uppercase boldface letters represent matrices. $\mathbf{A}^T$, $\mathbf{A}^H$ and $\overline{\mathbf{A}}$ indicate the matrix transpose, conjugate transpose and Hermitian of $\mathbf{A}$, respectively. $\mathbb{C}^{m\times n}$ and tr($\mathbf{A}$) are denoted as the vector space of all $m \times n$ complex matrices and the trace of a matrix $\mathbf{A}$, respectively. $||\cdot||_F$, $j = \sqrt{-1}$ and diag($\mathbf{x}$) represents the Frobenius form operator, imaginary unit and a diagonal matrix with $\mathbf{x}$ as the principal.

Furthermore, $\mathbf{A}-\mathbf{B} \succeq \mathbf{0}$ means matrix $\mathbf{A}-\mathbf{B}$ is a positive semidefinite and $\mathbf{x} \succeq \mathbf{0}$ means $x_j \geq 0$ for all j.

II. SYSTEM MODEL

In this section, we introduce two kinds of antenna arrays, the uniform circular array (UCA) and uniform linear array (ULA). UCA is suitable for estimating the number of signals, while ULA is suitable for simultaneously detecting the number of signals and estimate the DOA. First, we give the form of the observation model

$$\mathbf{Y} = \mathbf{AS} + \mathbf{E}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{C}^{N\times L}$ represents the received signal matrix, the matrix $\mathbf{S} \in \mathbb{C}^{M\times L}$ and $\mathbf{E} \in \mathbb{C}^{N\times L}$ represent the transmitted signal and noise, respectively. L is the number of snapshots, M is the number of signals and N is the number of sensors, respectively.

The theoretical covariance matrix is defined as

$$\begin{aligned}\mathbf{R}_{yy} &= E[\mathbf{YY}^H] \\ &= \mathbf{AR}_{ss}\mathbf{A}^H + \mathbf{R}_{nn} \\ &= \Psi + \mathbf{R}_{nn},\end{aligned} \tag{2}$$

where $\mathbf{R_{ss}} = E[\mathbf{SS}^H]$ is the signal covariance matrix, which is diagonal in the case of uncorrelated signals.[1] The noise covariance matrix is $\mathbf{R_{nn}} = E[\mathbf{EE}^H]$. And denote the sample covariance:

$$\widehat{\mathbf{R}}_{yy} = \frac{1}{L}\sum_{i=1}^{L}\mathbf{y}_i\mathbf{y}_i^H. \tag{3}$$

A. CIRCULAR ARRAY

We consider the antenna array system equipped with 8 receive antennas located in a circle with a radius of d meters as shown in Figure 1. The time delay is given by

$$\tau_n = \frac{\mathbf{r}_n^T\mathbf{u}}{c} = \frac{1}{c}[x_n\cos\theta\cos\phi + y_n\sin\theta\cos\phi + z_n\sin\phi], \tag{4}$$

where c is the speed of light, θ and ϕ are defined in Figure 1,

$$\begin{aligned}\mathbf{r}_n &= [x_n, y_n, z_n]^T; \\ \mathbf{u}(\theta,\phi) &= [\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi]^T.\end{aligned} \tag{5}$$

The wave number vector is defined by

$$\mathbf{k}(\theta,\phi) = \frac{\omega}{c}\mathbf{u}(\theta,\phi) = \frac{\omega}{f\lambda}\mathbf{u}(\theta,\phi) = \frac{2\pi}{\lambda}\mathbf{u}(\theta,\phi), \tag{6}$$

where ω is temporal angular frequency of the wave, f and λ are the frequency and wavelength of light, respectively. The nth element of steering vector is given by

$$\begin{aligned}a_n(\theta,\phi) &= e^{-j\frac{2\pi}{\lambda}[x_n\cos\theta\cos\phi + y_n\sin\theta\cos\phi + z_n\sin\phi]} \\ &= e^{-j\mathbf{r}_n^T\mathbf{k}(\theta,\phi)}.\end{aligned} \tag{7}$$

The received signal is given by

$$\mathbf{y}(t) = \mathbf{A}(\theta,\phi)\,\mathbf{s}(t) + \mathbf{n}(t), \tag{8}$$

where the columns of matrix $\mathbf{A}(\theta,\phi)$ are the steering vectors.

[1]We assume that the signals are in our future work.

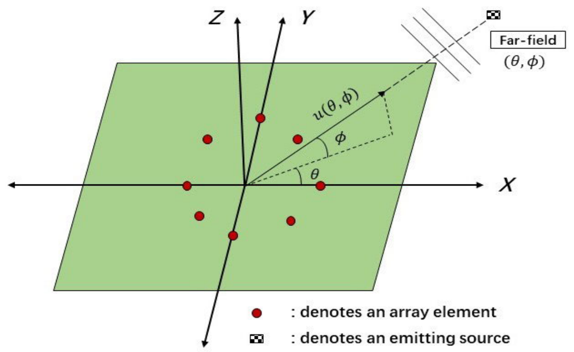

FIGURE 1. System model.

B. UNIFORM LINEAR ARRAYS

Considering the ULA with M narrow-band far-field signals impacting the sensor from direction $\theta_m \in [-90°, 90°)$, $m = 1, 2, \ldots, M$, the spacing of the sensor is not less than $\frac{\lambda}{2}$. The DOA estimation problem is to estimate the direction vector $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_M]^T$. Denote $\vartheta_m = \frac{\sin(\theta_m)+1}{2} \in [0, 1)$, $m = 1, 2, \ldots, M$. As the relation $\boldsymbol{\theta} \leftrightarrow \boldsymbol{\vartheta}$ is one-to-one, so the estimate of $\boldsymbol{\vartheta} = [\vartheta_1, \ldots, \vartheta_M]^T$ is equivalent to $\boldsymbol{\theta}$. $\boldsymbol{\theta}$ can be obtained by the inverse solution $\boldsymbol{\vartheta}$. If we adjust the spacing of the antennas to always keep the wavelength $\frac{\lambda}{2}$, then $\boldsymbol{\vartheta}$ is the frequency parameter.

For an N-element ULA, the mth element of steering vector is given by

$$\mathbf{a}(\vartheta_m) = [1, e^{i2\pi\vartheta_m}, \ldots, e^{i2(N-1)\pi\vartheta_m}]^T. \tag{9}$$

The array manifold matrix $\mathbf{A}(\vartheta) = [a(\vartheta_1), \ldots, a(\vartheta_M)]$ contains the steering vectors.

According to [22], [30], a simple phase shift is used to represent the time delay of different sensors, and an observation model is obtained:

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\vartheta})\,\mathbf{s}(t) + \mathbf{n}(t), \tag{10}$$

where t indexes the snapshot.

Based on the above assumptions, in order to better estimate the effective parameters. Let

$$\Psi = \mathbf{A}(\boldsymbol{\vartheta})\mathrm{diag}(\mathbf{p})\mathbf{A}^H(\boldsymbol{\vartheta}), \tag{11}$$

where $\mathbf{p} = [p_1, p_2, \ldots, p_N]^T$ is the signal power parameter. It can be seen that $\Psi \succeq 0$. So Ψ is a Toeplitz matrix determined by $2N-1$ complex numbers, which can be written as $\Psi = T(\mathbf{u})$ for some $\mathbf{u} \in \mathbb{C}^M$,

$$T(\mathbf{u}) = \begin{bmatrix} u_1 & u_2 & \cdots & u_N \\ \bar{u}_2 & u_1 & \cdots & u_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{u}_N & \bar{u}_{N-1} & \cdots & u_1 \end{bmatrix}. \tag{12}$$

III. DETECTING THE NUMBER OF SIGNALS

This section presents several algorithms for detecting the number of signals.

The additive white Gaussian noise covariance matrix is $\mathbf{R}_{nn} = \mathbf{E}\{\mathbf{n}(t)\ \mathbf{n}^H(t)\} = \sigma^2\mathbf{I}$, with noise variance σ^2. With the independence assumption and applying traditional AIC and MDL criteria, the number of signals can be estimated which is very important for setting the partition of the eigenvectors of the covariance matrix $\mathbf{R}_{yy}$ in MUSIC, ESPRIT, etc.

However, if the noise is correlated, the expression of $\mathbf{R}_{yy}$ is rewritten as

$$\begin{aligned}\mathbf{R}_{yy} &= \mathbf{A}(\theta,\phi)\mathbf{R}_{ss}\mathbf{A}^H(\theta,\phi) + \Sigma,\\ &= \sum_{i=1}^{M}(\lambda_i)\mathbf{v}_i\mathbf{v}_i^H + \sum_{j=1}^{N}\sigma_j^2\mathbf{n}_j\mathbf{n}_j^H,\end{aligned} \tag{13}$$

where $\Sigma = \mathbf{E}\{\mathbf{n}(t)\ \mathbf{n}^H(t)\}$, λ_i and $\mathbf{v}_i$, $i = 1, \ldots, M$ are the ordered eigenvalues and eigenvectors of $\mathbf{A}(\theta,\phi)\mathbf{R}_{ss}\mathbf{A}^H(\theta,\phi)$, σ_j^2 and $\mathbf{n}_j, j = 1, \ldots, N$ are the eigenvalues and eigenvectors of Σ.

As the following, two new methods of detecting the number of signals are introduced, which are Eigenvalue Correction Improvement and Signal Cancellation.

A. EIGENVALUE CORRECTION IMPROVEMENT

Considering the existing methods, the differences between the eigenvalues of $\widehat{\mathbf{R}}_{yy}$ are not obvious enough for detection, because of the noise correlation. The idea of eigenvalue correction is to try to update the eigenvalues of the covariance matrix $\widehat{\mathbf{R}}_{yy}$ to enlarge the difference between the eigenvalues of the signals and noise.

If we can estimate the matrix $\mathbf{R}_{yy}$ by considering the off-diagonal elements in Σ and deal with the mixture of signals and noise, the signal number can be estimated directly by using the log-likelihood function F as [19],

$$F = -L\log\det(\mathbf{R}_{yy}) - \mathrm{tr}\left[(\mathbf{R}_{yy})^{-1}\widehat{\mathbf{R}}_{yy}\right], \tag{14}$$

However, it is very hard to separate information of signal and noise from $\mathbf{R}_{yy}$. Because of that, we provide new applicable approaches and improved methods based on the eigen-decomposition of $\widehat{\mathbf{R}}_{yy}$.

Assuming e_i, $i = 1, \ldots, N$, to be the correction value of the eigenvalues and order the eigenvalues of $\widehat{\mathbf{R}}_{yy}$ in ascending order as: $\lambda_{min} = l_1 \le l_2 \le \cdots \le l_N = \lambda_{max}$. On the one hand, considering physical meaning of correction value, we need to add white noise to arrays, such that

$$e_i \le \lambda_{max}, \quad i = 1, 2, \ldots, N; \tag{15}$$

and on the other hand, in order to smoothen colored noise effectively,

$$e_i \ge \lambda_{min}, \quad i = 1, 2, \ldots, N. \tag{16}$$

Based on the idea of eigenvalues correction, we define a new correction method as followings:

- Order the eigenvalues in ascending order: $l_1 \le l_2 \le \cdots \le l_N$.
- Define the correction factor as

$$e_1 = \sqrt{l_1}, \tag{17}$$

$$e_i = \sqrt{\sum_{j=1}^{i} l_j}, \quad i = 2, \ldots, N. \tag{18}$$

- Update the corrected eigenvalues as

$$c_i = l_i + e_i, \quad i = 1, 2, \ldots, N. \tag{19}$$

- Rearrange c_i in descending order: $c_1 \ge c_2 \ge \cdots \ge c_N$.

According to this correction, the difference between eigenvalues of the signals and noise can be enlarged and it will contribute to improve the estimators. This is because, by performing the eigenvalues correction as above, large value corrections are added to enlarge the bigger eigenvalues (normally they represent signals) and small value corrections are put to the smaller eigenvalues. Increased difference between eigenvalues will improve the accuracy of signal detection algorithms. Two algorithms based on the corrected eigenvalues are introduced in the following.

1) IMPROVED EIGENVALUE GRADS METHOD (IM-EGM)

Considering the traditional eigenvalue grads method based on the eigenvalues correction mentioned above, the number of signals can be estimated as follows:

Firstly, we calculate the average grads of all corrected eigenvalues as

$$\overline{c} = \frac{c_1 - c_N}{N-1}. \tag{20}$$

Secondly, each grad between adjacent eigenvalues is calculated as

$$\overline{c}_i = (c_i - c_{i+1}), \quad i = 1, \ldots, N-1. \tag{21}$$

Finally, comparing $\overline{c}_i$ with $\overline{c}$ from $i = 1$ to $N-1$. If $\overline{c}_i \le \overline{c}$, the algorithm finishes and the estimate number of signals is equal to $i-1$. The computational complexity of the core step of the method is $O(N^3 + N^2 + 7N)$.

2) IMPROVED EIGEN-INCREMENT STOPPING RULE

In this subsection, a stopping rule is provided for the grads of eigenvalues to improve the accuracy of estimation.

A single eigen-increment stopping rule (EISR) is defined as

$$EISR = f(N,L)\frac{C_{EISR}}{(1+\sqrt{C_{EISR}})^2}, \tag{22}$$

where C_{EISR} is the estimation of the signal power and is given by

$$C_{EISR} = \frac{c_1 - c_N}{N}, \tag{23}$$

and $f(N,L)$ is a coefficient function with the following features [7]:

1) $f(N,L) = 1$ when $L \to \infty$
2) $f(N,L)$ increases as L reduces
3) $f(N,L)$ decreases as N increases.

Using the EISR, the detection criterion can be described as follows: if

$$\hat{c}_i = (c_i - c_{i+1}) \geq EISR, \quad i = N-1, N-2, \ldots, 1, \quad (24)$$

it stops and the estimate of the number of signals is i. The computational complexity of the core step of this method is $O(N^3 + N^2 + 3N)$.

B. SIGNAL CANCELLATION (SC)

The observed covariance matrix $\widehat{\mathbf{R}}_{yy}$ includes the mixture information of signals and spatially correlated noise. The idea of signal cancellation is to subtract the signals based on their eigenvalues and eigenvectors which are larger than the eigenvalues and eigenvectors of noise, and stop at particular step when only the noise is left.

Obviously, after subtracting all the signals, the covariance matrix $\widehat{\mathbf{R}}_{yy}$ should only have the noise, and this can be measured by comparing the Frobenius norm of covariance matrix at each step.

Assuming there are d signals, the algorithm can be expressed as

$$\widehat{\mathbf{R}}^j_{yy} = \widehat{\mathbf{R}}_{yy} - l_j \mathbf{u}_j \mathbf{u}_j^H, \quad j = 1, 2, \ldots, d, d+1, \ldots, N, \quad (25)$$

where $\mathbf{u}_j$, $j = 1, 2, \ldots, d, d+1, d+2, \ldots, N$ are the eigenvectors of $\widehat{\mathbf{R}}_{yy}$ according to the eigenvalues $l_1 \geq l_2 \geq \cdots \geq l_N$.

It should be noticed that when $j > d$, we can derive the covariance matrix $\widehat{\mathbf{R}}^j_{yy}$ by subtracting only one item of noise. Furthermore, considering the practical properties of noise, the difference of the generalized Frobenius norm of $\widehat{\mathbf{R}}^{d+1}_{yy}$ and $\widehat{\mathbf{R}}^{d+2}_{yy}$ is very small. To cope with the wide value range of the real data, we normalize their Frobenius norms by

$$N(\widehat{\mathbf{R}}^1_{yy}) = 1,$$
$$N(\widehat{\mathbf{R}}^j_{yy}) = \frac{||\widehat{\mathbf{R}}^j_{yy}||_F}{||\widehat{\mathbf{R}}^1_{yy}||_F}, \quad j = 2, \ldots, N. \quad (26)$$

Assuming that the stopping threshold is SC, the algorithm will stop when

$$N(\widehat{\mathbf{R}}^j_{yy}) - N(\widehat{\mathbf{R}}^{j+1}_{yy}) \leq SC. \quad (27)$$

Then the estimated number of signals is $j-1$.

Another advantage of signal cancellation method is that the robust threshold SC[2] can be set for particular direction and frequency after short time of training to achieve high estimation accuracy.

C. NUMERICAL RESULTS

In this section, first, simulation results are presented by showing the successful detection probability of new estimation methods. The system is simulated as shown in Figure 1. Eight antennas are uniformly located on the circle with radius 3.5 m. In all simulations, we assume that there are two independent signal sources impinging from 45° and 90° respectively, but the noise is spatial correlated. Parameter Settings are shown in Table 1. We assume that the correlation matrix of noise is a Symmetric Toeplitz matrices (or called as centrosymmetric matrix), and the correlation matrix element is $\alpha^{|i-j|}$, $i, j = 1, \ldots, N$. The correlation of noise is generated by correlation matrix and controlled by the parameter α, smaller value of α means less correlation between the noise and the correlation is stronger when α increases. The optimal parameter settings are shown in Table 2. After determining the two values, the value with the probability reaching 1 is selected according to the probability of successfully estimating the number of signals. The best value is greater than or equal to this value.

TABLE 1. Parameter setting.

Signals(M)	Sensors(N)	Snapshots(L)	α
2	8	512	0.5,0.8,1

TABLE 2. Best value comparison of SNR, sensors, and snapshots.

	Kmeans	GDE	Im-EGM	Im-EISR	SC
SNR (dB)	with N=8 and L=512				
	5	-1	-1	8	0
Sensors (N)	with SNR=0 dB and L=512				
	8	9	5	7	8
Snapshots (L)	with SNR=0 dB and N=8				
	150	100	100	100	100

Figure 2 presents the success estimation probability of *EGM* algorithm, where *EGM* denotes the estimation based on original eigenvalues, and *Im-EGM* denotes the corrected eigenvalues based estimator. The noise correlation parameter α varies from 0.5 to 1. For the spatial correlated noise scenario, EGM works well when SNR is greater than 0 dB, while there is no obvious difference that can be seen between *EGM* and *Im-EGM*. Generally speaking, the Im-EGM performs similarly to EGM.

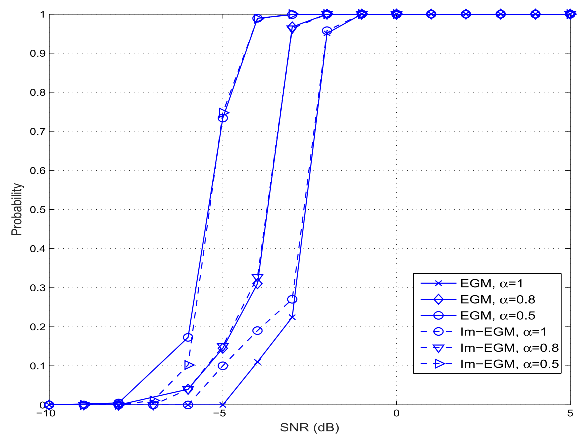

FIGURE 2. Success Estimation Probability of signals detection based on EGM.

[2]The threshold SC can be set by the experience based on different scenarios.

The estimation algorithm based on EISR is shown in Figure 3, where the estimator based on original eigenvalues is denoted as *EISR* and *Im-EISR* represents the estimation using corrected eigenvalues. The correlation between noise is varying from weak $\alpha = 0.5$ to strong $\alpha = 1$. The estimator based on EISR can provide accurate detection when SNR is greater than 5 dB and noise correlation is weak, while it requires high SNR to provide acceptable detection results when the noise correlation increases. It worth to notice that *Im-EISR* works better than *EISR* when there are weak spatial correlations between noises.

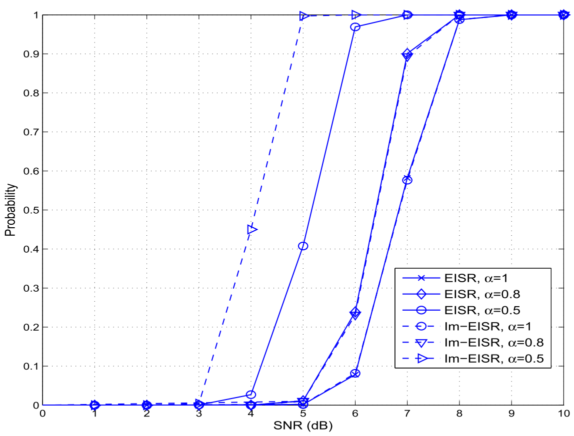

FIGURE 3. Success estimation probability of signals detection based on EISR.

Simulation results of success estimation probability of signal detection by applying the signal cancellation algorithm is illustrated in Figure 4. First of all, comparing to *EGM*, *Im-EGM*, *EISR* and *Im-EISR*, the simulation results show that the estimation algorithm based on signal cancellation works better in low SNR regime, but the success estimation probability increases slowly with SNR.

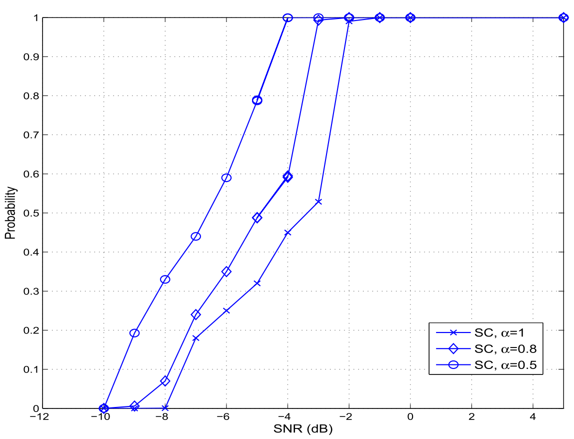

FIGURE 4. Success estimation probability of signals detection applying SC.

The simulation results of new estimation algorithms are compared with each other and existing well known methods GDE [10] and $K-$means [12] in Figure 5. The noise correlation parameter is $\alpha = 1$ for all the curves. It can be seen that the estimation algorithm based on signal cancellation performs the best in low SNR regime, but the successful probability of GDE method increases more quickly when SNR increases. The $K-$means algorithm and estimation based on *Im-EISR* are worse than the others. Comparing to the original criteria, effective methods have been found in the correlated noise environment and they are verified by the simulation.

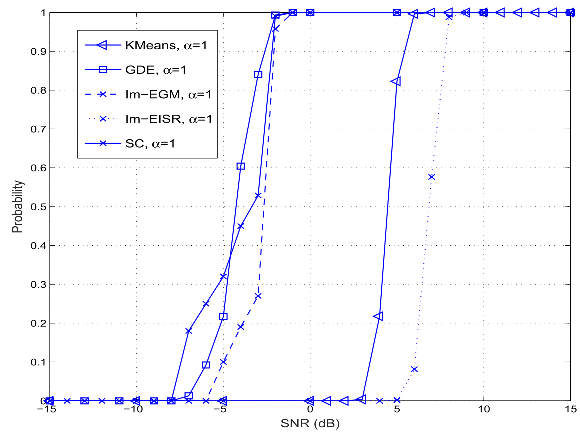

FIGURE 5. Comparison of success estimation probability of different estimators.

IV. PCASE

The methods shown in section III can only improve the accuracy of signal number estimate, in this section, the PCASE algorithm will be introduced simultaneously to detect the number of signal sources and estimate DOA. Based on the popular machine learning method PCA [29] and the basic idea of SPA (see in [28]), we study the new estimation method PCASE. The flow chart is shown in Figure 6, which is described in detail below.

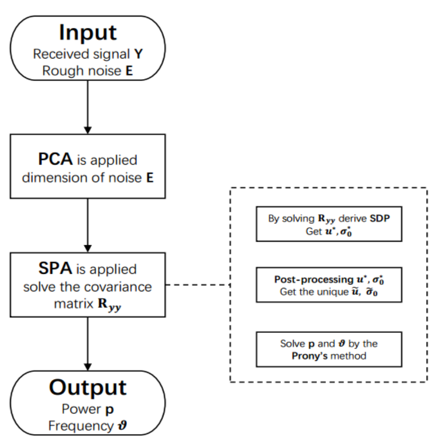

FIGURE 6. PCASE flowchart.

Remark 1: In this paper, we assume that rough information of the noise is known, which mean it is not necessary to know the exact information of noise.[3] Herein, we apply the pre-whitening technique by using the information of noise at different times slot, but adjacent frequencies and bandwidths. At least, it is applicable and could improve the detection accuracy, which is verified by the simulations.

A. PRINCIPAL COMPONENT ANALYSIS (PCA)

Principal component analysis (PCA) is a technology widely used in dimensionality reduction, lossy data compression, feature extraction and data visualization, and is an unsupervised linear method. PCA uses orthogonal transformation to transform a group of variables which may be correlated into a group of linearly uncorrelated variables. If the noise is correlated, the covariance matrix of the noise is not a diagonal matrix. Therefore, in order to obtain independent noise, PCA was used to reduce the noise dimension.

The spatially and temporarily white Gaussian noise covariance matrix is $\mathbf{R}_{nn} = E\{\mathbf{n}(t_1)\mathbf{n}^H(t_2)\} = \text{diag}(\boldsymbol{\sigma})\delta_{t_1,t_2}$, with noise variance parameter is $\boldsymbol{\sigma} = [\sigma_1, \ldots, \sigma_N]^T \in R_+^M$ and $\delta_{t1,t2}$ is a delta function that equals 1 if $t_1 = t_2$ or 0 otherwise. And the date snapshots are uncorrelated with each other.

Now, if the noises are correlated, $\mathbf{R}_{nn}$ does not have this form. In order for $\mathbf{R}_{nn}$ to have this form, let's do the following for noise $\mathbf{E}$:

- Centralize all samples of $\mathbf{E}$:
$$\mathbf{e}(t_n) \leftarrow \mathbf{e}(t_n) - \bar{\mathbf{e}};$$
- Calculate the noise covariance matrix $\mathbf{EE}^H$;
- Eigenvalue decomposition of the covariance matrix $\mathbf{EE}^H$:
- Take the eigenvalue vector corresponding to the maximum K eigenvector $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_K$.

The final form is

$$\begin{aligned}\widehat{\mathbf{E}} &= \mathbf{WB}\\ &= [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K][b_{t1}, b_{t2}, \cdots, b_{tK}], \end{aligned} \tag{28}$$

where w_i is the eigenvector corresponding to the largest K eigenvalues of $\mathbf{S}_0 = \frac{1}{L}\sum_{n=1}^{L}(\mathbf{e}(t_n) - \bar{\mathbf{e}})(\mathbf{e}(t_n) - \bar{\mathbf{e}})^T$, $\bar{\mathbf{e}} = \frac{1}{L}\sum_{n=1}^{L}\mathbf{e}(t_n)$. $\mathbf{B}$ is the coefficient matrix, and its covariance matrix should be the diagonal matrix, which can be set as $\text{diag}(\boldsymbol{\sigma}_0)$, $\boldsymbol{\sigma}_0 = [\sigma_{01}, \ldots, \sigma_{0K}]^T \in R_+^K$.

The expression of $\mathbf{R}_{nn}$ is rewritten as

$$\mathbf{R}_{nn} = \mathbf{W}diag(\boldsymbol{\sigma}_0)\mathbf{W}^H.$$

In this way, the number of parameters does not increase when solving SDP problems in the next subsection. SDP was explained in the next page.

B. SPARSE AND PARAMETRIC APPROACH (SPA)

Based on the established covariance fitting criterion and convex optimization, SPA performs parameter estimation in the continuous range. It is a sparse parameter method without discretization. Unlike existing parameterization methods, SPA method can detect the number of signals and estimate the DOA. The method consists of the following parts.

1) SEMI-DEFINTE PROGRAMMING (SDP)

When both $\mathbf{R}_{yy}$ and $\widehat{\mathbf{R}}_{yy}$ are invertible, in order to estimate the unknown parameters, we consider the following covariance fitting criterion (see [23], [31], [32]):

$$f(\boldsymbol{\vartheta}, \mathbf{p}, \boldsymbol{\sigma}_0) = \left\| \mathbf{R}_{yy}^{-\frac{1}{2}}(\widehat{\mathbf{R}}_{yy} - \mathbf{R}_{yy})\widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}} \right\|_F^2, \tag{29}$$

where $\mathbf{R}_{yy}^{-1}$ is present with the noise and $\sigma_{0j} > 0$ for $j = 1, 2, \ldots, N$, and f is the distance between $\mathbf{R}_{yy}$ and $\widehat{\mathbf{R}}_{yy}$.

Remark 2: The covariance fitting criterion makes use of the assumption that the information signal is not correlated, thus obtaining the expression of $\mathbf{R}_{yy}$ in (2). However, the theoretical explanation proposed in [23], [31] proves that the criterion is robust to the correlation of signals. Therefore, the proposed method maintains this robustness under the same criterion.

It can be derived with some simple algebraic manipulations that

$$\begin{aligned} f &= tr\left[(\mathbf{R}_{yy}^{-\frac{1}{2}}(\widehat{\mathbf{R}}_{yy} - \mathbf{R}_{yy})\widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}})(\mathbf{R}_{yy}^{-\frac{1}{2}}(\widehat{\mathbf{R}}_{yy} - \mathbf{R}_{yy})\widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}})^H\right]\\ &= tr(\mathbf{R}_{yy}^{-1}\widehat{\mathbf{R}}_{yy}) + tr(\widehat{\mathbf{R}}_{yy}^{-1}\mathbf{R}_{yy}) - 2N. \end{aligned} \tag{30}$$

According to equation (12), the structure of $\mathbf{R}_{yy}$ under constraint conditions $T(\mathbf{u}) \succeq 0$ and $\boldsymbol{\sigma}_0 \succeq 0$ is obtained as

$$\mathbf{R}_{yy} = T(\mathbf{u}) + \mathbf{W}\text{diag}(\boldsymbol{\sigma}_0)\mathbf{W}^H. \tag{31}$$

Under the condition of semi-positive definite $T(\mathbf{u})$, the distance f is expected to be the minimum, and the problem can be converted into the following optimization:

$$\begin{aligned} &\min_{\mathbf{u},\boldsymbol{\sigma}_0 \succeq 0} f(\boldsymbol{\vartheta}, \mathbf{p}, \boldsymbol{\sigma}_0)\\ &s.t.\ T(\mathbf{u}) \succeq \mathbf{0}. \end{aligned} \tag{32}$$

The number of $2N$ has been expressed in $\mathbf{R}_{yy}$, which will lead to the redundancy of $\mathbf{R}_{yy}$. However, the redundancy problem will not affect the currently estimated $\mathbf{R}_{yy}$. The handling of this problem is described in the next subsection. So the above optimization problem can be written as

$$\begin{aligned} &\min_{\mathbf{u},\boldsymbol{\sigma}_0 \succeq 0} tr(\mathbf{R}_{yy}^{-1}\widehat{\mathbf{R}}_{yy}) + tr(\widehat{\mathbf{R}}_{yy}^{-1}\mathbf{R}_{yy})\\ &s.t.\ T(\mathbf{u}) \succeq \mathbf{0}. \end{aligned} \tag{33}$$

After a series of derivations (see in [28]), the above optimization problem is finally converted into the following SDP problem

$$\begin{aligned} &\min_{\mathbf{X},u,\boldsymbol{\sigma}_0 \succeq 0} tr(\mathbf{X}) + tr(\widehat{\mathbf{R}}_{yy}^{-1}\mathbf{R}_{yy})\\ &s.t.\ \begin{bmatrix} \mathbf{X} & \widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}} & \\ \widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}} & \mathbf{R}_{yy} & \\ & & T(\mathbf{u}) \end{bmatrix} \succeq \mathbf{0}. \end{aligned} \tag{34}$$

where $\mathbf{X} \succeq \widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}}\mathbf{R}^{-1}\widehat{\mathbf{R}}_{yy}^{-\frac{1}{2}}$.

[3]The exact information of noise means we know all the information of noise at the time of doing detection. The rough information of noise means we only have to know the information of noise around detection frequency at the time before the detection

Therefore, the problem in (33) can be expressed as an SDP problem and thus is convex, so the SDP problems can be iteratively solved using the standard CVX toolbox, a Matlab package for specifying and solving convex programs ([33], [34]). We can get the estimated value $\widetilde{\mathbf{R}}_{yy} = T(\mathbf{u}^*) + \mathbf{W}\text{diag}(\boldsymbol{\sigma}_0^*)\mathbf{W}^H$ of $\mathbf{R}_{yy}$, where $\mathbf{u}^*$ and $\boldsymbol{\sigma}_0^*$ are the solutions of the SDP problem.

2) POSTPROCESSING

The signal and noise in the estimated covariance matrix can be separated by the post-processing method, so that the signal part can be represented by as few signals as possible on the basis of the minimum description length principle. Since there is redundancy on the diagonal of $\mathbf{R}_{yy}$, and the effect of redundancy is double, the solution of SDP problem cannot be guaranteed to be unique, so it cannot be directly used as the final estimation of noise variance. Therefore, the solution needs to be processed by using post-processing.

$\widetilde{R}_{yy}$ is expressed as follows

$$\widetilde{\mathbf{R}}_{yy} = T(\widetilde{\mathbf{u}}) + \mathbf{W}\text{diag}(\widetilde{\boldsymbol{\sigma}}_0)\mathbf{W}^H, \tag{35}$$

where $T(\widetilde{\mathbf{u}}) = A(\widetilde{\boldsymbol{\theta}})\text{diag}(\widetilde{\mathbf{p}})A^H(\widetilde{\boldsymbol{\theta}}) \succeq \mathbf{0}$ and $\widetilde{\boldsymbol{\sigma}}_0 \succeq \mathbf{0}$ are the estimate of Ψ and nosie covariance. The $\widetilde{\mathbf{u}}$ and $\widetilde{\boldsymbol{\sigma}}$ are satisfing that

$$\widetilde{\mathbf{u}} = \mathbf{u}^* - \begin{bmatrix} \delta \\ \mathbf{0} \end{bmatrix}, \quad \widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma}^* + \delta\mathbf{I}, \tag{36}$$

this is one form of decomposition, but it actually lists all the possible forms. Based on $T(\widetilde{\mathbf{u}}) = T(\mathbf{u}^*) - \delta I \succeq \mathbf{0}$, and $\delta = \lambda_{min}(T(\mathbf{u}^*))$ is the smallest eigenvalue of $T(\mathbf{u}^*)$, so the decomposition is unique. In this way, the solution of SDP is unique. In this condition, it is possible to conclude that the final $(\mathbf{u}, \boldsymbol{\sigma}_0)$ estimates have good statistical properties.

3) SOLVING ϑ AND p

The next step is to estimate parameters $\widetilde{\vartheta}$ and $\widetilde{p}$ for given $T(\widetilde{\mathbf{u}})$, which is based on the classical Vandermonde decomposition for semi-definite Toeplitz matrices ([28], [35], [36]).

According to Vandermonde decomposition, $\widetilde{\boldsymbol{\vartheta}}$ and $\widetilde{\mathbf{p}}$ can be uniquely determined by $T(\widetilde{\mathbf{u}})$. In practice, $\widetilde{\boldsymbol{\vartheta}}$ and $\widetilde{\mathbf{p}}$ can be obtained by the following method. Given $T(\widetilde{\mathbf{u}}) = A(\widetilde{\boldsymbol{\vartheta}})diga(\widetilde{\mathbf{p}})A^H(\widetilde{\boldsymbol{\vartheta}})$. It's easy to prove that

$$\begin{bmatrix} A(\widetilde{\boldsymbol{\vartheta}}) \\ \overline{A}_{2,\ldots,M}(\widetilde{\boldsymbol{\vartheta}}) \end{bmatrix} \widetilde{\mathbf{p}} = \begin{bmatrix} \widetilde{\widetilde{\mathbf{u}}} \\ \widetilde{\mathbf{u}}_{2,\ldots,M} \end{bmatrix}. \tag{37}$$

where $\widetilde{\mathbf{p}} \succ 0$, and $\overline{\mathbf{A}}_{2,\ldots,M}(\widetilde{\boldsymbol{\vartheta}})$ just takes the first row of the matrix $\mathbf{A}_{\vartheta}$. According to [37], Prony' s method can be used to solve $\widetilde{\boldsymbol{\vartheta}}$ and $\widetilde{\mathbf{p}}$ in this kind of equations. For details of the process, please refer to [37].

C. SIMULATION RESULTS

In this section, we illustrate the performance of the proposed PCASE method. In simulation, we consider $M = 2$ uncorrelated signals with power $\mathbf{p}^\circ = [3, 1]^T$ and frequency vector $\boldsymbol{\vartheta}^\circ = [0.10, 0.50]^T$. Each signal is generated with constant amplitude and random phase, according to the general settings in [31]. The SNR defined as $10log_{10}\frac{min(p_j^\circ)}{\sigma_0^\circ}$ (in dB). Parameter Settings are shown in Table 3.

TABLE 3. Parameter setting.

Signals(M)	Sensors(N)	Snapshots(L)	α	Threshold
2,3	8,30	512	0.5,0.7,0.9	0.1,0.4,0.7

In the first case, we consider an array with 8 sensors. Figure 7 shows the PCASE algorithm's probability of success estimation.[4] The noise correlation parameter α ranges from 0.5 to 0.9. It can be seen from the figure that PCASE can provide acceptable test results when the SNR is less than 0 dB, regardless of the strong or weak noise correlation. When the SNR is greater than 5 dB, 100% accurate detection results can be provided.

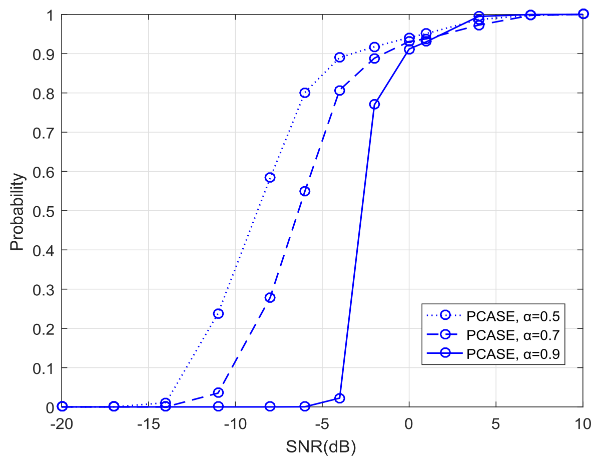

FIGURE 7. Success estimation probability of signals detection applying PCASE.

As shown in Figure 8, we compared the simulation results of three algorithms: the PCASE, Im-EISR, SC, GDE and the existing well known methods K-means. The noise correlation parameter of all curves is $\alpha = 0.5$. It is easy to see that in the case of low SNR, the estimation algorithm based on PCASE has the best estimation effect, but the PCASE method's probability of success grows slowly when SNR increase. However, PCASE can not only detect the number of signal sources, but also estimate the Direction of arrival. Therefore, PCASE is effective in detecting the number of signals.

In following, we show that PCASE can detect the number of signal sources and estimate DOA ($\theta_m = arcsin(2\vartheta_m - 1)$) at the same time. Considering $M = 3$ and three signals have $\mathbf{p}^\circ = [3, 2, 1]^T$ and $\boldsymbol{\vartheta}^\circ = [0.10, 0.15, 0.50]^T$. In Figure 9, when $N = 8$ and $\alpha = 0.5$, the simulation result shows that PCASE is more effective and accurate than SPA under correlated noise.

[4]Our simulation is different from spectrum perception under the premise of assuming a signal, so the results show the probability of success estimation.

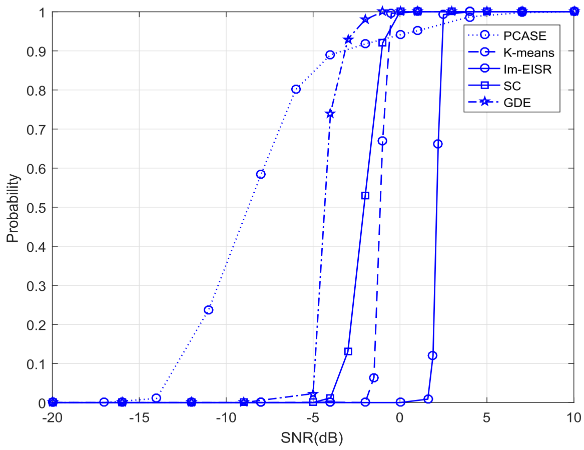

FIGURE 8. Comparison of success estimation proability of different estimators.

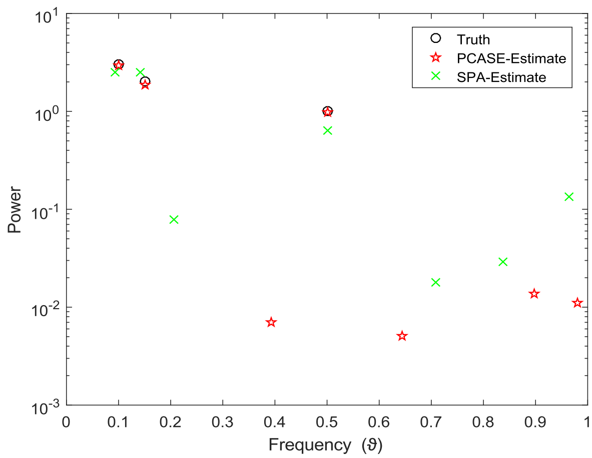

FIGURE 9. Frequency and power estimates of PCASE for estimating $M = 3$ uncorrelated signals when $N = 8$ and $L = 512$.

Considering large scale antenna array, we increased the number of receiving antennas and the noise correlation coefficient, which a ULA with $N = 30$ is used to receive the signals and $\alpha = 0.9$. The simulation result is shown in Figure 10. The SPA approach does not work well in this case. However, our method can accurately detect the number of signals and give their directions. In addition, We have regressed our experimental results with a function ($y = a * x^b$), which is able to classify the signal and noise into two distinct categories. Therefore, the validity of our method is verified.

According to the linear regression curve in Figure 10, the boundary value and the middle of this curve are selected as thresholds to judge the number of signals. Figure 11 shows false alarm rates under different thresholds. As shown in the figure, false alarm rate will increase with the decrease of threshold value. When we select the appropriate threshold value, false alarm rate will be reduced.

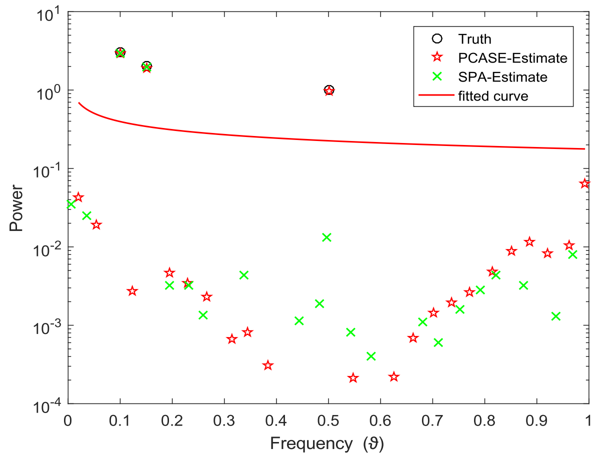

FIGURE 10. Frequency and power estimates of PCASE for estimating $M = 3$ uncorrelated signals when $N = 30$ and $L = 512$.

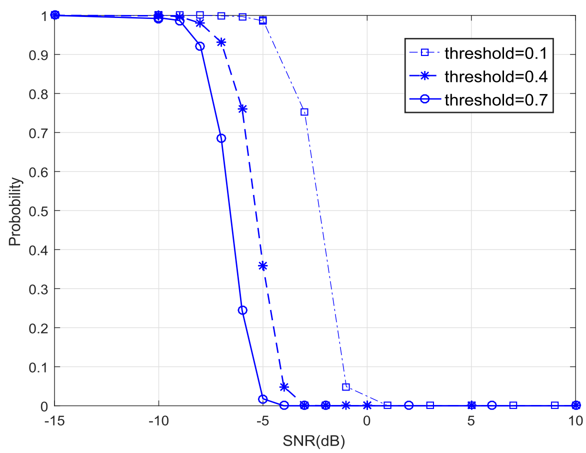

FIGURE 11. False alarm rate under different thresholds.

V. CONCLUSION

In this paper, the problem of detecting signals with correlated noise has been investigated. The traditional criteria of AIC and MDL were reviewed and the reason of their failure in correlated noise environment was discussed. Furthermore, new criteria were introduced, such as *Im-EGM*, *Im-EISR* and *SC*. *Im-EGM* and *Im-EISR* can improve the accuracy of the signal detection algorithm by increasing the difference between the eigenvalues. *SC* can set a robust threshold for a specific direction and frequency when the training time is short, and achieve high estimation accuracy. The simulations show that all the methods perform well when SNR is greater than 5 dB. At low SNR regime, the estimation based on SC works better than the others. For future work, the SC method will be improved based on robust threshold setting. In addition, we also propose a method that can simultaneously detect the number of signals and estimate DOA. This method combines machine learning and sparse parameter method. Simulation results show that this method can not only detect the number of signal sources accurately, but also estimate the DOA of signal sources. Compared with other methods for detecting the number of signals, this method is effective in the case of low SNR. Therefore, the new method is a powerful tool for detection. We will be considered to extend our work and try to derive effective methods to study coherent signal estimation in our future work.

REFERENCES

[1] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[2] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[3] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5113–5129, Nov. 2008.

[4] J.-J. Fuchs, "Estimation of the number of signals in the presence of unknown correlated sensor noise," *IEEE Trans. Signal Process.*, vol. 40, no. 5, pp. 1053–1061, May 1992.

[5] 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, Jun. 1995.

[6] P. Fabry, C. Serviere, and J. L. Lacoume, "Improving signal subspace estimation for blind source separation in the context of spatially correlated noises," in *Proc. 9th IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Sep. 1998, pp. 2377–2380.

[7] O. Hu, F. Zheng, and M. Faulkner, "Detecting the number of signals using antenna array: A single threshold solution," in *Proc. 5th Int. Symp. Signal Process. Appl. (ISSPA)*, Aug. 1999, pp. 905–908.

[8] J. Luo and Z. Zhang, "Using eigenvalue grads method to estimate the number of signal source," in *Proc. 5th Int. Conf. Signal Process., 16th World Comput. Congr. (WCC ICSP)*, Aug. 2000, pp. 223–225.

[9] Z. Lu, M. Gao, and H. Jiang, "Source number estimation for minimum redundancy arrays with Gerschgorin disk estimator," in *Proc. IEEE 11th Int. Conf. Signal Process.*, Oct. 2012, pp. 311–314.

[10] Z. Liu, Z. Lu, Z. Huang, and Y. Zhou, "Improved Gerschgorin disk estimator for source enumeration with robustness against spatially non-uniform noise," *IET Radar, Sonar Navigat.*, vol. 5, no. 9, pp. 952–957, Dec. 2011.

[11] Q. Zhang, Y. Yin, and J. Huang, "Detecting the number of sources using modified EGM," in *Proc. IEEE Region Conf. (TENCON)*, Nov. 2006, pp. 1–4.

[12] C. Zhang, X. Si, and J. Xie, "A method for determining the number of signals based on modified K-means clustering," in *Proc. IEEE Int. Conf. Inf. Autom.*, Harbin, China, Jun. 2010, pp. 20–23.

[13] 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.

[14] Y. Zhou, D. Wang, T. Pei, and S. Tian, "Robust estimation fusion in wireless senor networks with outliers and correlated noises," *Int. J. Distrib. Sensor Netw.*, vol. 10, no. 4, pp. 393–802, Apr. 2014.

[15] D. Ciuonzo, "On time-reversal imaging by statistical testing," *IEEE Signal Process. Lett.*, vol. 24, no. 7, pp. 1024–1028, Jul. 2017.

[16] D. Ciuonzo, G. Romano, and R. Solimene, "Performance analysis of time-reversal MUSIC," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2650–2662, May 2015.

[17] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.

[18] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.

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

[20] Q. Wu and D. R. Fuhrmann, "A parametric method for determining the number of signals in narrow-band direction finding," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1848–1857, Aug. 1991.

[21] L. Yang, H. Zhang, J. Li, H. Yang, and Y. Cai, "Blind source enumeration based on Gerschgorin disk estimator and virtual array extension," in *Proc. 8th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Yangzhou, China, Oct. 2016, pp. 1–4.

[22] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[23] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35–47, Jan. 2011.

[24] F. Vincent, O. Besson, and E. Chaumette, "Approximate unconditional maximum likelihood direction of arrival estimation for two closely spaced targets," *IEEE Signal Process. Lett.*, vol. 22, no. 1, pp. 86–89, Jan. 2015.

[25] Y. Ma, K. Deng, and Z. Ding, "Direction-of-arrival estimation for coherently distributed sources via symmetric uniform linear array," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Oct. 2017, pp. 1–5.

[26] R. Schmidt, "A signal subspace approach to multiple emitter location spectral estimation," Ph.D. dissertation, Dept. Eng., Electron. Elect., Stanford Univ., Stanford, CA, USA, 1981.

[27] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer–Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[28] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4959–4973, Oct. 2014.

[29] C. M. Bishop, *Pattern Recognition and Machine Learning* (Information Science and Statistics). New York, NY, USA: Springer-Verlag, 2006.

[30] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Pearson/Prentice-Hall, 2005.

[31] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, Feb. 2011.

[32] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digit. Signal Process.*, vol. 8, no. 3, pp. 185–210, Jul. 1998.

[33] M. Grant and S. Boyd. (Sep. 2013). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*. [Online]. Available: http://cvxr.xom/cvx

[34] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (Lecture Notes in Control and Information Sciences), V. Blondel, S. Boyd, and H. Kimura, Eds. London, U.K.: Springer, 2008, pp. 95–110.

[35] C. Carathéodory and L. R. Fejér, "Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landau'schen Satz," *Rendiconti del Circolo Matematico di Palermo*, vol. 32, no. 1, pp. 218–239, 1911.

[36] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. London, U.K.: Univ. of California Press, 1958.

[37] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 31–40, Mar. 2008.

LEI WANG received the B.S. degree in applied mathematics and the M.S. degree in computational mathematics from Xinjiang University, in 2015 and 2018, respectively. She is currently pursuing the Ph.D. degree in mathematics and statistics with Xi'an Jiaotong University. Her research interest includes wireless communication.

JIANG XUE (Senior Member, IEEE) received the B.S. degree in information and computing science from Xi'an Jiaotong University, Xi'an, China, in 2005, the M.S. degree in applied mathematics from Lanzhou University, China, in 2008, the M.S. degree in applied mathematics from Uppsala University, Sweden, in 2009, and the Ph.D. degree in electrical and electronic engineering from ECIT, Queens University of Belfast, U.K., in 2012. From 2013 to 2017, he was a Research Fellow with The University of Edinburgh, U.K. Since 2017, he has been as a Professor with the National Engineering Laboratory for Big Data Analytics, Xi'an International Academy for Mathematics and Mathematical Technology, School of Mathematics and Statistics, Xi'an Jiaotong University. His main interests include machine learning and intelligent wireless communication, performance analysis of general multiple antenna systems, stochastic geometry, and cooperative communication.

JOHN THOMPSON is currently a Professor of signal processing and communications with the School of Engineering, The University of Edinburgh. He specializes in antenna array processing, cooperative communications systems, and energy efficient wireless communications. He has published in excess of 300 articles on these topics, including one hundred journal article publications. He is also the Project Coordinator for the EU Marie Curie International Training Network Project ADVANTAGE, which studies how communications and power engineering can provide future smart grid systems. He was an elected Member-at-Large of the Board of Governors of the IEEE Communications Society, from 2012 to 2014, the second largest IEEE Society. He was elevated to the Fellow of the IEEE for contributions to multiple antenna and multi-hop wireless communications, in 2016. He is an Editor of the Green Communications and Computing Series that appears regularly in the *IEEE Communications Magazine*. He is also a Distinguished Lecturer on the topic of energy efficient communications and smart grid for the IEEE Communications Society, from 2014 to 2015.

JIA YU received the B.Eng. degree in electronic information engineering and the M.S. degree in electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2007 and 2009, respectively, and the Ph.D. degree in digital communication from The University of Edinburgh, U.K, in 2017. From 2009 to 2013, he was a Wireless Communication Engineer with Huawei Technologies Company, Shenzhen, China. He is currently working as a Research Fellow with the Institute of 5G Communication and Big Data, Hangzhou. His main interests include machine learning, wireless communication, and signal processing.