Research Article

An Improved Method for Parametric Spatial Spectrum Estimation in Internet of Things

Yong Liu, Jingya Zhao, Qinghua Zhu, and Yanqiu Wang

Beijing Polytechnic, Beijing 100176, China

Correspondence should be addressed to Yong Liu; 403229043@qq.com

Received 27 March 2021; Revised 10 April 2021; Accepted 20 April 2021; Published 5 May 2021

Academic Editor: Xin Liu

Copyright © 2021 Yong Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

When massive numbers of wireless IoT devices are being deployed, cognitive spectrum management is critical to satisfy the explosive broadband requirements of IoT applications. Heterogeneous the target of the spatial spectrum estimation which is part of array signal processing is to achieve the evaluation of space signal parameters and source location, which result in that the spatial spectrum estimation becomes the most basic content of the array signal processing. It needs a method by which the large dimensional array still has its consistency. Therefore, this paper studies an improved large dimensional array parameterized spatial spectrum estimation method based on Pisarenko method, named G-Pisarenko method. Firstly, an improved estimator about the logarithm of the covariance matrix of a certain bilinear form is analyzed which is based on the theory of large dimension random matrix. We can find out a relatively better method, i.e., MW method. The method will become the primitive method for us to improve. Then, aimed at the relating covariance matrix in MW, we use an improved large dimensional array estimation method which can improve the logarithm of the covariance matrix estimation. Finally, we compare the improved method and the original method by simulation, and it can be seen the clear advantage of G-Pisarenko method when the sample number and observed dimensions are in the same order of magnitude.

1. Introduction

The development of Industrial Internet of Things (IIoT) has been limited due to the shortage of spectrum resources. Also, the traditional ground industrial Internet of Things (IIoT) cannot supply wireless interconnections anywhere due to its small-scale communication coverage, meanwhile the Internet of Things (IoT) is facing the shortage of spectrum resources due to the rapid growth of IoT terminals and big data services. Fifth generation (5G) network owns sufficient spectrum resources and supplies large data volume business, which can help to expand the communication resources of the IoT by combing IoT with 5G network [1–3].

Also known as spatial signal processing, array signal processing utilizes some properties of signals in space to process signals [4]. The specific implementation process is to put multiple sensors in different directions in space to form an array, then utilize the specific array of the sensors to receive and process signals, and extract interested signals and related parameters. At the same time, noise and other interference information are suppressed and eliminated. At present, this processing method has occupied a very important position in the field of signal processing, and it is also of great significance in the fields of radar and sonar [5]. Array signal processing has many advantages, which are strong anti-interference ability, higher gain, strong spatial superresolution, and flexible beam control. Therefore, array signal processing occupies a very important position in the current research [6, 7].

The earliest spatial spectrum estimation algorithm is called conventional beamforming (CBF) method, which is based on the Fourier spectrum estimation method in time domain. It mainly replaces the data processed in the traditional time domain with the data received by each array element in the spatial domain. It is also called Bartlett beamforming method, but it has a fatal disadvantage. Namely, the resolution of the array angle would be limited by its physical aperture, and the target located in a beam width space could be often indistinguishable. Subsequently, there are many high-resolution estimation methods that
extend nonlinear signal processing technology to spatial domain. There are Burge’s maximum entropy method (MEM), Pisarenko’s method, and Capon’s minimum variance method (MVM) [8–10]. All these methods have a common premise. Namely, they are based on the continuous distribution of signal sources in space, but they are not satisfied in practice. Therefore, there are great limitations. Later, there would be algorithms to mathematically decompose the data received by the array, such as MUSIC algorithm, which divides the received data into two subspaces of mutually orthogonal signal and noise. Hereafter, the spatial spectrum peak is constructed to improve the resolution [11].

However, no matter what method is used for spatial spectrum estimation, multivariate statistical knowledge is needed. Moreover, in multivariate statistical analysis, there are two kinds of limit results which are quite different. One is that the assumed dimension is very small and the sample capacity is relatively much larger than the dimension. At this time, the theory of satisfaction is called classical limit theory. The other is the large-dimensional limit theory. When the dimension is very high, it might even fail according to the situation described by the classical limit theory. Therefore, the large-dimensional limit theory is proposed. At this moment, it is necessary to combine the large-dimensional random matrix with the spatial spectrum estimation to find an improved estimation method, which could satisfy both the dimension and the observed value geometry. With the maturity of spatial spectrum estimation technology, spatial spectrum estimation based on large array would inevitably become the mainstream of development, and the application of random matrix theory in spectral estimation is also a great progress and breakthrough.

Random matrix theory (RMT) refers to the matrix, whose elements are real, complex, or quaternion of random variables, while the analytical and statistical properties of eigenvalues and eigenvectors are mainly researched in random matrix theory. In the recent decades, with the rapid development of computer science and technology, large-scale data analysis is becoming more and more important. There are statistical problems of massive data in the fields of biological microarray data, wireless communication network, financial stock analysis, spatial spectrum estimation, and so on. However, the classical statistical method is based on the case that the dimension is much larger than the sample capacity. This is not satisfied when the dimension sample capacity tends to infinity at the same time, so the large dimension random matrix theory is used.

At the earliest, random matrix theory was researched by Wishart and Hsu. The main work was the joint distributions of matrix elements and eigenvalues in multivariate statistical analysis. Subsequently, physicists introduced the theory into atomic physics and thought that a random Hermitian matrix could replace the Shrodinger operator, and its eigenvalues were used to describe the observed energy level. Nowadays, there are more and more kinds of random matrix, and the common basic research object is the limit spectrum distribution. This is helpful to research the more precise properties of matrix eigenvalues, such as the increase and decrease of eigenvalues and the distribution of maximum and minimum eigenvalues. At present, many scholars have conducted more in-depth research.

Recently, based on the technique of contour integral in random matrix theory, an improved estimator for determining the logarithm of covariance matrix in bilinear form has been proposed [12]. This newly demonstrated estimator is applicable not only to high sample capacity (such as traditional estimator) but also to the case that both observed dimension and sample capacity increase at the same rate. In practical application, sample capacity and observed dimension always have the same order of magnitude. In this case, this feature shows very good performance. Better fitting estimation could be obtained through using this estimator in spatial spectrum estimation.

In conclusion, array signal processing has already become a very important branch of signal processing field. Thereinto, the spatial spectrum estimation technology could be used to estimate some parameters of signal source position or signal spatial domain, so it becomes the most basic content of array signal processing. At present, many spatial spectrum estimation methods have been proposed, but they have a common drawback. It is not considered that the number of samples and the observed dimension (the number of array elements) are in the same order of magnitude, but only the number of samples is much larger than the observed dimension under the assumption. Therefore, it is urgent to research an estimation method that could still be consistent in the case of large arrays. This is also the important research content in this paper.

2. Theoretical Analysis

Through the analysis of traditional Pisarenko algorithm, it is found that these estimates have a common feature, which is only for the case that the number of samples is far greater than the observed dimension. When the number of samples and the number of sensors is at the same level, there are large errors in all estimates, but the actual situation is like this. In this section, an improved estimator about the logarithm of the covariance matrix of a certain bilinear form is analyzed. This new estimator is applicable not only to high sample capacity (such as traditional estimator) but also to observed dimension and sample capacity increasing at the same rate. This estimate value is based on the theory of large dimension random matrix.

2.1. Mathematical Modeling. A set of observations [13] with a definite zero mean of \( M \times N \) dimensions \((M,N)\) are considered. That is \( \{y_1, y_2, \cdots, y_N\} \) and that is to say \( y_n \in \mathbb{C}^M \), \( n = 1, \cdots, N \). It is assumed that these observations are equally distributed with zero mean and form a covariance matrix \( \mathbf{R}_m \) together. \( \lambda_1^{(M)} \leq \lambda_2^{(M)} \leq \cdots \leq \lambda_M^{(M)} \) and \( \{e_k^{(M)}\}, \) \( k = 1, \cdots, M \) are used to represent the eigenvalues and related eigenvectors of matrix \( \mathbf{R}_m \). \( \log \mathbf{R}_m \) is used to represent the logarithm of matrix \( \mathbf{R}_m \). Then, that is

\[
\log \mathbf{R}_m = \sum_{k=1}^{M} \log \left( \lambda_k^{(M)} e_k^{(M)} (e_k^{(M)})^H \right).
\]
The logarithm of the covariance matrix plays an important role in many signal processing applications, such as spectrum estimation, parameter identification, and spectrum representation. For example, in the spectrum estimation, Pisarenko proposed the following form of spectrum estimation: \[ P_h(f) = h^{-1}(s_M^H(f) h(R_m) s_M(f)). \]

In this formula, \( h() \) is a definite reversible function with inverse \( h^{-1}(\cdot) \). \( s_M(f) \) is a column vector of \( M \times 1 \), and its \( k \)th value is equal to the value of \( M^{-1/2} \exp [j2\pi kf] \). In the spectrum estimation, when there is \( h() = \log h() \), there would be power estimators:

\[ P_h(f) = \exp (s_M^H(f) \log R_M s_M(f)). \] (2)

Namely, the power depends on the quadratic form of the covariance matrix logarithm of the observations.

In practice, the real covariance matrix is unknown and must be estimated from the samples. The common method is to construct a sample covariance matrix, such as \( R_M = (1/N) \sum_{n=1}^N y_n y_n^H \), and then, \( \log R_M \) is estimated from \( \log R_M \). If there is a process from \( N \) to infinity, the estimated value of \( R_M \) is always satisfied. It could be assumed that \( \log R_M \) is infinitely close to \( \log R_M \), when \( N \) is large enough. However, in practice, the number of the sample \( N \) may not be very large compared with the actual dimension of the matrix \( R_m \), so the \( \log R_M \) accuracy is quite low in this case.

To solve this problem, a generalized progressive form would be focused on. Namely, \( M \) and \( N \) could be considered as large but comparable numbers. The consistent estimation of definite bilinear form of \( \log R_M \) would be considered, when \( M \) and \( N \) increase to infinity at the same rate. More specifically, we would think that the numerical estimates like \( a_M^H \log R_M b_M \) are consistent in this new progressive form, if \( a_M \) and \( b_M \) are two universal \( M \times 1 \) column vectors with the same unit standard. If \( a_M \) and \( b_M \) are, respectively, determined to be the \( i \)-th column and the \( j \)-th column in the unit matrix \( M \times M \), the value of \( a_M^H \log R_M b_M \) would be equal to the value of the \((i, j)\) position in the matrix \( \log R_m \). In some applications (such as spectrum power estimation and spectrum coefficient determination), \( a_M = b_M = s_M(f) \) is a more meaningful thing. Thereinto, \( s_M(f) \) is defined above. In this case, the main target is to estimate the value of \( s_M^H(f) \log R_M s_M(f) \).

2.2. Method of Estimation. The number of samples is regarded as a function of the observed dimension. That is \( N = N(M) \). \( C_m = M/N(M) \) is introduced. It is strictly assumed that the number of samples is higher than the observed dimension. That is \( 0 < C_m < 1 \). The following assumptions [14] are made:

(i) (Hypothesis 1). According to the zero mean Gauss law of covariance matrix, the set \( \{ \text{Re}(y_m), \text{Im}(y_m), n = 1, \ldots, N \} \) of \( M \times 1 \) dimension vectors is distributed independently and identically

(ii) (Hypothesis 2). The minimum eigenvalue of \( R_m \) (or the maximum eigenvalue) is not equal to zero or infinite in \( M \)

(iii) (Hypothesis 3). The \( \{ C_m \} \) sequence is not equal to 0 or 1

The main conclusion under this hypothesis is that the estimator of the universal \( \log R_m \) bilinear form is generally applicable, when \( M \) and \( N \) tend to infinity at the same rate. The above definition of sample covariance is considered.

\[ \lambda_1^{(M)} < \lambda_2^{(M)} < \cdots < \lambda_M^{(M)} \] and \( \{ \tilde{\alpha}_k^{(M)}, k = 1, \ldots, M \} \) are used to represent its eigenvalues and corresponding eigenvectors, respectively.

Theorem 1. The both unit standards could be considered to be the certain column vector sequences \( a_M \) and \( b_M \) of \( M \times 1 \). \( M, N \to \infty \) could be satisfied. There is \( \hat{\alpha}_m^{(M)} = a_M^H \log R_m b_M \), when all three hypotheses are satisfied, Thereinto:

\[ \hat{\alpha}_m^{(M)} = \sum_{m=1}^M \hat{\alpha}_m^{(M)} (e^{\lambda_m^{(M)}})^H b_M. \] (3)

The coefficient is

\[ I_m^{(M)} = 1 + \left( 1 + \sum_{k=1}^M \frac{\hat{\lambda}_k^{(M)}}{\lambda_k^{(M)} - \lambda_m^{(M)}} - \sum_{k=1}^M \frac{\hat{\mu}_k^{(M)}}{\lambda_k^{(M)} - \mu_m^{(M)}} \right) \log \hat{\lambda}_m^{(M)} + \sum_{r=1}^M \frac{\hat{\lambda}_r^{(M)}}{\lambda_r^{(M)} - \lambda_m^{(M)}} \log \hat{\mu}_r^{(M)} \] (4)

Thereinto, the value of \( \hat{\mu}_1^{(M)} < \cdots < \hat{\mu}_M^{(M)} \) is equal to the value \( \mu \) in the following equation:

\[ I = \frac{1}{N} \sum_{r=1}^M \frac{\hat{\lambda}_r^{(M)}}{\lambda_r^{(M)} - \mu}. \] (5)

It is worth pointing out here that the new estimate value could converge to the classical value, if \( M \) is taken as a fixed value and \( N \) tends to infinity. Actually, in the observation in equation (5), when \( N \) tends to infinity, \( I_m^{(M)} \to \log (\hat{\lambda}_m^{(M)}) \) could be obtained in the expression, in which \( \hat{\mu}_m^{(M)} \to \lambda_m^{(M)} \) and \( (\hat{\lambda}_m^{(M)} - \mu_m^{(M)}) \to \hat{\lambda}_m^{(M)} \) could apply this progressivity with \( I_m^{(M)} \) in formula (3).

2.3. Theorem Proving. The proof of the theorem is based on the contour integral method. The eigenvalue of the covariance matrix is consistent with the estimate value of the eigenvector, when it is proved that \( M \) and \( N \) tend to infinity and the value of \( M/N \) is a constant. Firstly, the estimated quantity
needs to be expressed in the form of complex contour integral [15] of real covariance matrix decomposition function. This is defined as a matrix \((\mathbf{R}_M - 2\mathbf{I}_M)^{-1}\), \(z \in C^* = \{ z \in C : \text{Im}(z) > 0 \}\). Thereinto, the contour of the integral is parameterized by a very convenient function \(\omega_M(z)\), which would be described in the following content. This parameterization allows us to use a definite algebraic expression of consistent estimation of decomposition function \((\mathbf{R}_M - 2\mathbf{I}_M)^{-1}\) of sample covariance matrix to express the integrand function. Finally, the expression could be obtained through using the classical Cauchy residue theorem.

### 2.3.1. Formulation of Contour Integral and Parameterization of Contour

From the observed value, Cauchy residue theorem could be used to express \(\log \mathbf{R}_M \mathbf{b}_M\) as an integral form:

\[
a_M^H \log \mathbf{R}_M \mathbf{b}_M = \frac{1}{2\pi i} \int_{C} a_M^H (\mathbf{R}_M - \omega I_M)^{-1} \mathbf{b}_M \log \omega d\omega. \tag{6}
\]

Thereinto, \(C^*\) includes the clockwise direction of all \(\mathbf{R}_M\) eigenvalues oriented contour, and \(\log(\cdot)\) is a holomorphic logarithmic branch above \(C^* \setminus R^*\). The contour is regarded as a curve in \(C = \{ \omega_M(z), z \in \partial R^- \}\). For arbitrary \(y > 0\), \(\partial R\) could satisfy:

\[
\eta_{\text{inf}} = \inf \{ 1 - \sqrt{\frac{y}{\lambda_1}} \lambda_1^{(M)} \},
\]

\[
\eta_{\text{sup}} = \sup \{ 1 + \sqrt{\frac{y}{\lambda_2}} \lambda_2^{(M)} \},
\]

for the rectangle \(R = [\eta_{\text{inf}}, \eta_{\text{sup}}] \times [-y, y]\’s periphery, \(\omega_M(z)\) is a result of the following polynomial equation:

\[
z = \omega_M(z) \left( 1 - \frac{1}{N} \sum_{m=1}^M \frac{\lambda_m^{(M)}}{\lambda_m^{(M)} - \omega_M(z)} \right). \tag{8}
\]

In particular, seen in Figure 1, it is worth noting that \(\omega_M(z)\) is the only root which makes both \(\text{Im}(\omega_M(z)) \geq 0\) and \((1/N) \sum_{m=1}^M |\lambda_m^{(M)}/\lambda_m^{(M)} - \omega_M(z)|^2\) true, if there is \(\text{Im}(z) > 0\). If there is \(\text{Im}(z) < 0\), one take \(\omega_M(z) = \omega_M^*(z)\), that the contour does not intersect the negative real axis could be deduced because \(\text{Re}(\omega_M) < 1\).

The parameterization of the contour \(C\) is used, and formula (8) is rewritten as follows:

\[
\frac{1}{2\pi i} \int_{R} a_M^H (\mathbf{R}_M - \omega_M(z) I_M)^{-1} \mathbf{b}_M \log [\omega_M(z)] \omega_M(z) dz. \tag{9}
\]

Thereinto, \(\omega_M^*(z)\) is the complex derivative of \(\omega_M(z)\). That 0 does not belong to the rectangle \(R\) could be seen (because of \(\text{Re}(\omega_M) < 1\), and there is periphery distance between of the distance 0 of \(\omega_M(z)\) and the eigenvalue of \(\mathbf{R}_M\). The both holomorphic functions \(m_M(z)\) and \(q_M(z)\) could be defined as follows:

\[
m_M(z) = \frac{1 - \zeta_M}{\zeta_M \omega_M(z)}, \tag{10}
\]

\[
q_M(z) = \frac{\omega_M(z)}{z} a_M^H (\mathbf{R}_M - \omega_M(z) I_M)^{-1} \mathbf{b}_M.
\]

makes,

\[
a_M^H \log \mathbf{R}_M \mathbf{b}_M = \frac{1}{2\pi i} \int_{\partial R} q_M(z) \left( \frac{1 - \zeta_M + z^2 \zeta_M m_M(z)}{1 - \zeta_M - z \zeta_M m_M(z)} \right) \times \log \left[ \frac{z}{1 - \zeta_M - z \zeta_M m_M(z)} \right] dz. \tag{11}
\]

Thereinto, \(q_M^*(z)\) is the derivative of \(m_M(z)\).

#### 2.3.2. Consistent Estimation of Integrand

Under Hypothesis 1 and Hypothesis 3 [16], \(m_M(z) \approx \hat{m}_M(z)\) and \(q_M(z) \approx \hat{q}_M(z)\) are almost applied to all \(z \in \partial R\), when \(M\) and \(N\) tend to infinity. Thereinto, \(\hat{m}_M(z)\) and \(\hat{q}_M(z)\) are defined as follows:

\[
\hat{m}_M(z) = \frac{1}{M} \text{tr} \left[ (\mathbf{R}_M - 2\mathbf{I}_M)^{-1} \right],
\]

\[
\hat{q}_M(z) = a_M^H (\mathbf{R}_M - 2\mathbf{I}_M)^{-1} \mathbf{b}_M.
\]

According to the convergence theorem, the results in the relevant data could be reused, and the random quantities \(\hat{m}_M(z)\) and \(\hat{q}_M(z)\) in the above integral sum could replace \(q_M(z)\). Moreover, when \(M\) and \(N\) tend to infinity, it is certain that there could be:

\[
a_M^H \log \mathbf{R}_M \mathbf{b}_M = \sum_{m=1}^M I_m^{(M)} a_m^H \left( \omega_m^{(M)} \right) \mathbf{b}_M. \tag{13}
\]

Thereinto,

\[
I_m^{(M)} = \frac{1}{2\pi i} \int_{\partial R} \frac{1}{\lambda_m^{(M)} - z} \frac{1 - \zeta_M + z^2 \zeta_M \hat{m}_M(z)}{1 - \zeta_M - z \zeta_M \hat{m}_M(z)} \times \log \left[ \frac{z}{1 - \zeta_M - z \zeta_M \hat{m}_M(z)} \right] dz. \tag{14}
\]

Therefore, it is necessary that \(I_m^{(M)}\) in (14) could also be expressed as \(I_m^{(M)}\) in (4), if the proof of Theorem 1 could be obtained.

#### 2.3.3. Processing the Produced Integral

Unfortunately, the nonisolated singularity could be appeared for the above integral in \(R\) (due to the appearance of logarithms). Moreover, the integral could not be processed with the conventional complex integration method. To solve this, the following real-valued function is defined:

\[
I_m^{(M)}(\theta) = \frac{1}{2\pi i} \int_{\partial R} \frac{1}{\lambda_m^{(M)} - z} \frac{1 - \zeta_M + z^2 \zeta_M \hat{m}_M(z)}{1 - \zeta_M - z \zeta_M \hat{m}_M(z)} \times \log \left[ \frac{z}{1 - \theta(1 + z \hat{m}_M(z))} \right] dz. \tag{15}
\]

Wireless Communications and Mobile Computing
This function could prove that \( c^1(0, c_M) \) could be well-defined. In addition, \( \Lambda_m^{(M)} = 1^{(M)}(c_M) \) and \( \Lambda_m^{(M)}(0) \) could be easily solved by Cauchy residue theorem. Actually, the integral of \( \Lambda_m^{(M)}(0) \) is meromorphic, has single and plural poles at the place \( z = \hat{\lambda}_m^{(M)}, k \neq m \), has double and plural poles at the place \( \hat{\lambda}_m^{(M)} \) [17], and has single and plural poles at the place \( \hat{\mu}_k^{(M)} \). In \( k = 1, \cdots, M \), the algebraic property could be utilized to know that \( \Lambda_m^{(M)}(0) \) is equal to \( \Lambda_m^{(M)} \) defined in (4).

The first three terms of the above equation are a simple primitive function for \( \theta \). Thereinto, the appearance of the function \( \hat{\nu}_r^{(M)}(\theta) \) makes the last term difficult to be processed. To simplify the last term, the variables in the above formula could be changed to make \( \theta \rightarrow \hat{\nu}_r^{(M)}(\theta) \). Afterwards, the polynomial parameters produced in \( \hat{\nu}_r^{(M)}(\theta) \) are expressed as the sum of simple fractions. The derived original function and the value \( \Lambda_m^{(M)}(0) \) derived above are used to obtain:

\[
\frac{d \Lambda_m^{(M)}(\theta)}{d \theta} = \frac{1}{2 \pi i} \left( \frac{1}{\pi} \sum_{k=1}^{M} \frac{\hat{\lambda}_k^{(M)}}{z - \hat{\lambda}_k^{(M)}} - \frac{1}{\pi} \sum_{k=1}^{M} \frac{\hat{\mu}_k^{(M)}}{z - \hat{\mu}_k^{(M)}} \right) \times \left( 1 - \frac{(1/N) \sum_{k=1}^{M} \left( \hat{\lambda}_k^{(M)} / (\hat{\lambda}_k^{(M)} - z) \right) \left( \hat{\lambda}_k^{(M)} / (\hat{\lambda}_k^{(M)} - z) \right) \right) \right) \frac{1 - \epsilon_M + (1/N) \sum_{k=1}^{M} \left( \hat{\lambda}_k^{(M)} / (\hat{\lambda}_k^{(M)} - z) \right) \left( \hat{\lambda}_k^{(M)} / (\hat{\lambda}_k^{(M)} - z) \right) \right)dz.
\]

(16)

Finally, \( \theta = \epsilon_M \) is taken, and some algebraic identities are used to obtain the value as shown in Theorem 1.

It could be seen that this method is indeed a good estimate, and it still has good properties in large-dimensional arrays.

3. G-Pisarenko Algorithm Design

In this section, the aforementioned logarithmic improved estimator is applied to the traditional Pisarenko algorithm to propose a new algorithm, namely, G-Pisarenko algorithm.

3.1. Signal Model. The array used in this paper is a linear array. With the origin as the reference point, the array element position is assumed to be \( x_k(k = 1, 2, \cdots, M) \). When the angle of the signal incidence is \( \theta_i = \sigma_i \sin \theta / c \), there is \( r_{ki} = (x_k \sin \theta / c)(k = 1, \cdots, N) \). When the array is a uniform linear array, there is only a relative delay difference among the different array elements, which receive signals in the same direction. At this moment, the first array element is regarded as the reference array element, and there is \( r_{ki} = ((m - 1)d \)
\[ a_i(\omega_0) = \begin{bmatrix} 1 \\ \exp \left( -j2\pi \frac{d}{\lambda} \sin \theta_1 \right) \\ \vdots \\ \exp \left( -j2\pi (M-1) \frac{d}{\lambda} \sin \theta_N \right) \end{bmatrix}, \quad i = 1, 2, \ldots, N. \]

\[ A = [a_1(\omega_0), a_2(\omega_0), \ldots, a_N(\omega_0)], \]

\[ X(t) = AS(t) + N(t). \]

Thereinto, Theorem 1 in Section 2.2 is combined to obtain the following conclusion:

Since \( \sigma \) is the determined sequence of column vectors, \( M, N \to \infty \) could be satisfied; there is \( \widetilde{D}_R^{(M)} \approx \alpha^* log R_M \alpha \), when all three assumptions in Section 2 are satisfied. Thereinto,

\[ \widetilde{D}_R^{(M)} = \alpha^* \left[ \sum_{m=1}^{M} I_m^{(M)} \zeta_m^{(M)} \left( \sigma_m^{(M)} \right)^* \right] \alpha. \]

The coefficient is:

\[ I_m^{(M)} = 1 + \sum_{k=1}^{M} \frac{\lambda_k - \lambda_m}{\lambda_k - \lambda_m} - \sum_{k=1}^{M} \frac{\gamma_k^{(M)} - \gamma_m^{(M)}}{\lambda_k - \lambda_m} \log ^\dagger \lambda_m \\
+ \sum_{r=1}^{M} \frac{\rho_r^{(M)} - \rho_m^{(M)}}{\lambda_m - \lambda_r} \log \rho_r^{(M)} + \sum_{r=1}^{M} \frac{\tilde{\rho}_r^{(M)} - \rho_m^{(M)}}{\lambda_m - \tilde{\rho}_r^{(M)}} \log \tilde{\rho}_r^{(M)}. \]

Thereinto, the value of \( \tilde{\mu}_1^{(M)} < \cdots < \tilde{\mu}_M^{(M)} \) is equal to the value \( \mu \) in the following formula:

\[ 1 = \frac{1}{N} \sum_{r=1}^{M} \frac{\tilde{\lambda}_r^{(M)}}{\lambda_r} - \mu. \]

The two expressions of \( \widetilde{D}_R^{(M)} \) are compared. In the case of large dimensions, \( \sum_{m=1}^{M} I_m^{(M)} \sigma_m^{(M)} \left( \sigma_m^{(M)} \right)^* \) could be used to represent \( log R_M \alpha \).

\[ \sigma_{G-Pisarenko} = \exp \left\{ \alpha^* \sum_{m=1}^{M} I_m^{(M)} \sigma_m^{(M)} \left( \sigma_m^{(M)} \right)^* \alpha \right\}. \]

According to equation (23), the relevant simulations are performed, and the following simulation results are obtained.

### 4. Simulation Results and Analysis

#### 4.1. Spectrum Peak Search Results

In order to illustrate the validity of the proposed estimates, an array processing procedure with three signal sources is considered. The array element is composed of a uniform array. The distance between the array elements is half a wavelength. All three signal sources are subject to the standard of Gaussian distribution. The superimposed noise is a random process with a mean value being zero and a variance being one. The arrival direction of the signal source being relative to the periphery of the array is fixed to [-40, 0, 40] degree. The simulations are performed in the following cases:

The first case is \( M = 10 \) and \( N = 6000 \). Namely, the number of samples is much larger than the observation dimension (the number of arrays), and it is run once. At this moment, the spectrum peak search results of traditional Pisarenko and G-Pisarenko methods are shown in Figures 2 and 3.

It could be observed that at this moment, the both have the spectrum peaks at -40 degrees, 0 degree, and 40 degrees. However, there would be more serious pseudo peaks under the traditional method. In comparison, there are obvious peak values in the improved method. If the angular interval is very small, the estimation result would be inaccurate. This is also one problem in the traditional Pisarenko’s MW method. There are obvious peak values in the improved method, when the angular interval is small.
This experiment is repeated 50 times, and its results are shown in Figures 4 and 5.

It could be seen that in the traditional method, there are serious pseudo peaks, and the graphics are more chaotic around 0 degree. At -40 degrees and 40 degrees, its effect is not as clear as that of the improvement method also.

The second case is \( M = 10 \) and \( N = 20 \). The number of samples and the value of the observation dimension are smaller, but the size is in the same order of magnitude. At the same time, it is a constant, which satisfies \( M/N \) and is greater than 0 and less than 1. At this moment, the simulation results are shown in Figures 6 and 7.

It could be observed that in the traditional method, not only the pseudo peaks exist but also the values of the pseudo peaks are relatively larger in this case, which is similar to case 1. Its difference is not as obvious as that of the improved method. Therefore, the improved method still has its advantages in this case.

This experiment is repeated 50 times, and its simulation results are shown in Figures 8 and 9.

The third case is \( M = 10 \) and \( N = 20 \). The number of samples and the value of the observation dimension are larger, but the size is in the same order of magnitude. At the same time, it is a constant, which satisfies \( M/N \) and is greater than 0 and less than 1. Since the amount of calculation is especially large when the values \( M \) and \( N \) are larger, 100 and 200 are used to roughly fit the case of large dimension. At this moment, the simulation results are shown in Figures 10 and 11.

It could be observed that the pseudo peak values in Figure 9 are very close, and the distortion is serious. The spectrum peaks in Figure 10 are obvious. Due to the large amount of calculation, the simulation results after 50 repetitions would not be discussed here, but it could be inferred that there is not much difference between 50 times and 1 time. The estimated spectrum peaks are still more obvious in the improved algorithm.

4.2. Comparison for Root Mean Square Error. Next, the root mean square errors generated in the traditional method and
the improved method could be compared, when the different values are taken for the signal-to-noise ratio. Here, the arrival direction of the signal source could be fixed to the degree \(\frac{-40, 0, 40}{C_{138}}\), which is relative to the periphery of the array. The signal-to-noise ratio is between -15 and 20. When running 100 times, the simulation results are as follows:

**The first case:** when there are \(M = 20\) and \(N = 60\), the simulation results at this moment are shown in Figure 12. It could be found that when the number of array elements and the number of samples are in the same order of magnitude, the root mean square error of the G-Pisarenko method is much lower than that of the traditional Pisarenko method. As the signal-to-noise ratio increases, the mean square error of the improved method gradually approaches zero.

**The second case:** when there are \(M = 20\) and \(N = 600\), the simulation results at this moment are shown in Figure 13. It could be found that when the number of array elements is relatively small, the root mean square error of the G-Pisarenko method is still much lower than that of the traditional Pisarenko. As the signal-to-noise ratio increases, the root mean square error of the improved method gradually approaches zero. The spectrum estimation effect is very well.
5. Summary

In this paper, an improved Pisarenko parameterized spatial spectrum estimation method, namely, the G-Pisarenko method, is proposed based on random matrix theory. From the simulation, it could be found that when $M$ and $N$ values are very small, the spectrum peaks of the improved algorithm are more obvious, and there is no redundant pseudo peak. When the $M$ and $N$ values are larger, the spectral peaks in the traditional algorithm are chaotic, but the performance is still well in the improved algorithm. From the perspective of root mean square error, the error of the improved algorithm is much smaller than that of the traditional algorithm. So, this is a very good improved estimation. According to the principle of the improved estimation, the methods in this paper could be used in other algorithms to achieve an improved feature decomposition algorithm for large-dimensional arrays. The scheme can be applied to the Internet of things and play an effective improvement.

However, at the same time, it is not difficult to find that during the spectrum peak search, a large amount of calculation could be generated in the method proposed in this paper. The complexity of the calculation is so high that it is impossible to discuss the situation where the values of $M$ and $N$ tend to be infinite. Therefore, this paper also needs to be improved on this basis. A simpler and more effective algorithm could be found. In addition, the situations researched in this paper are based on the ideal situations of many assumptions, which could only be ideally approximated in practical occasions. The effectiveness needs to be improved.

Data Availability

The (data type) data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

This paper has clear and legal intellectual property rights or property rights.

Acknowledgments

This work was supported by the Beijing City Board of education project (No. KM202010858005 and 2020Z006-KXZ).

References

[1] X. Liu and X. Zhang, "NOMA-based resource allocation for cluster-based cognitive industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5379–5388, 2020.

[2] X. Liu, X. Zhai, W. Lu, and C. Wu, "QoS-guarantee resource allocation for multibeam satellite industrial Internet of Things with NOMA," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2052–2061, 2021.

[3] X. Liu and X. Zhang, "Rate and energy efficiency improvements for 5G-based IoT with simultaneous transfer," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 5971–5980, 2019.
[4] W. Wang, J. Du, and W. Wang, "Improved MUSIC method for spatial spectrum estimation of ultrasound sources," in 2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pp. 898–902, Honolulu, HI, USA, 2017.

[5] Y. Sun, X. Yang, L. Guo, and T. Long, "Experimental array signal processing demonstration system by utilizing microphone array," in 2016 CIE International Conference on Radar (RADAR), pp. 1–5, Guangzhou, China, 2016.

[6] H. Zhang, H. Zhang, W. Liu, K. Long, J. Dong, and V. C. M. Leung, "Energy efficient user clustering, hybrid precoding and power optimization in terahertz MIMO-NOMA systems," IEEE Journal on Selected Areas in Communications, vol. 38, no. 9, pp. 2074–2085, 2020.

[7] X. Liu, C. Sun, M. Zhou, C. Wu, B. Peng, and P. Li, “Reinforcement learning-based multistorey double-threshold spectrum sensing with Bayesian fusion for industrial big spectrum data,” IEEE Transactions on Industrial Informatics, vol. 17, no. 5, pp. 3391–3400, 2021.

[8] P. Wright, M. C. Parker, and A. Lord, “Minimum- and maximum-entropy routing and spectrum assignment for flex-grid elastic optical networking [invited],” IEEE/OSA Journal of Optical Communications and Networking, vol. 7, no. 1, pp. A66–A72, 2015.

[9] M. Kagalenko, "Modified Pisarenko spectrum for frequency analysis," in 2018 7th Mediterranean Conference on Embedded Computing (MECO), pp. 1–4, Budva, Montenegro, 2018.

[10] D. F. Gomes, A. J. Zimbico, J. M. Maia et al., “Sparse arrays method with minimum variance for high quality image in ultrasound ultrafast imaging,” in 2019 IEEE SENSORS, pp. 1–4, Montreal, Canada, 2019.

[11] M. Wu and Y. Wang, "A feature selection algorithm of music genre classification based on ReliefF and SFS," in 2015 IEEE/A- CIS 14th International Conference on Computer and Information Science (ICIS), pp. 539–544, Las Vegas, NV, USA, 2015.

[12] X. Mestre, F. Rubio, and P. Vallet, "Improved estimation of the logarithm of the covariance matrix," in 2012 IEEE 7th Sensor Array and Multichannel Signal Processing Workshop (SAM), pp. 377–380, Hoboken, NJ, USA, 2012.

[13] P. Stoica and N. Sandgren, "Cepstrum thresholding scheme for nonparametric estimation of smooth spectra," in 2006 IEEE Instrumentation and Measurement Technology Conference Proceedings, pp. 874–878, Sorrento, Italy, 2006.

[14] V. Pisarenko, "On the estimation of spectra by means of nonlinear functions of the covariance matrix," Geophysical Journal International, vol. 28, no. 5, pp. 511–531, 1972.

[15] P. Stoica, J. Li, and X. Tan, "On spatial power spectrum and signal estimation using the pisarenko framework," IEEE Transactions on Signal Processing, vol. 56, no. 10, pp. 5109–5119, 2008.

[16] M. Lagunas, P. Stoica, and M. Rojas, "ARMA parameter estimation: revisiting a cepstrum-based method," in 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3685–3688, Las Vegas, NV, USA, 2008.

[17] C. I. Ryrnes, P. Enqvist, and A. Lindquist, "Cepstral coefficients, covariance lags, and pole-zero models for finite data strings," IEEE Transactions on Signal Processing, vol. 49, no. 4, pp. 677–693, 2001.

[18] X. Liu, Q. Sun, W. Lu, C. Wu, and H. Ding, "Big-data-based intelligent spectrum sensing for heterogeneous spectrum communications in 5G," IEEE Wireless Communications, vol. 27, no. 5, pp. 67–73, 2020.