Central DOA Estimation Method for Exponential-Type Coherent Distributed Source Based on Fourth-Order Cumulant

Hao Li 🟢, Weijia Cui, Bin Ba 🟢, Haiyun Xu 🟢, and Yankui Zhang 🟢

National Digital Switching System Engineering & Technological Research Center, Zhengzhou, Henan Province 450001, China

Correspondence should be addressed to Hao Li; wcdmalh@163.com

Received 2 May 2020; Accepted 23 June 2020; Published 11 July 2020

Copyright © 2020 Hao Li 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.

The performance of direction-of-arrival (DOA) estimation for sparse arrays applied to the distributed source is worse than that applied to the point source model. In this paper, we introduce the coprime array with a large array aperture into the DOA estimation algorithm of the exponential-type coherent distributed source. In particular, we focus on the fourth-order cumulant (FOC) of the received signal which can provide more useful information when the signal is non-Gaussian than when it is Gaussian. The proposed algorithm extends the array aperture by combining the sparsity of array space domain with the fourth-order cumulant characteristics of signals, which improves the estimation accuracy and degree of freedom (DOF). Firstly, the signal-received model of the sparse array is established, and the fourth-order cumulant matrix of the received signal of the sparse array is calculated based on the characteristics of distributed sources, which extend the array aperture. Then, the virtual array is constructed by the sum aggregate of physical array elements, and the position set of its maximum continuous part array element is obtained. Finally, the center DOA estimation of the distributed source is realized by the subspace method. The accuracy and DOF of the proposed algorithm are higher than those of the distributed signal parameter estimator (DSPE) algorithm and least-squares estimation signal parameters via rotational invariance techniques (LS-ESPRIT) algorithm when the array elements are the same. Complexity analysis and numerical simulations are provided to demonstrate the superiority of the proposed method.

1. Introduction

Direction-of-arrival (DOA) estimation is one of the important research fields in array signal processing [1–3]. At present, the angle parameter estimation technology based on subspace method has been relatively mature, and the typical representatives include multiple signal classification (MUSIC) method [4] and estimation signal parameters via rotational invariance techniques (ESPRIT) method [5], but these algorithms are all based on the point source model. However, in actual communication, a large number of multipath phenomena are caused due to the reflection and scattering of a complex environment, which makes the signal source expanded at a certain angle in space. Under this condition, using the traditional algorithm for point source model, the estimation performance will be seriously deteriorated [6–8]. Therefore, it is essential to study the model and algorithm of the distributed source. Because of the spatial distribution characteristics of the distributed source, the difference between the mathematical models of point source and the distributed source is mainly reflected in the difference of the steering matrix which causes the point source algorithm to be no longer applicable to the distributed source model. The idea of the MUSIC method based on the point source model is applied to a distributed source model by using the principle of subspace orthogonality in literature [9], namely, the DSPE algorithm and LS-ESPRIT algorithm. However, to obtain the central DOA and angle spread, a two-dimensional spectral-peak search is required and the computation complexity is high. In order to reduce the complexity, the algorithm is proposed, where an approximate rotation invariant relation is constructed to solve the central DOA using a special array layout [10–12], which averts spectral-peak search. The paper [13] estimates the rotation invariant matrix by constructing a propagation operator to solve the central DOA without eigenvalue
decomposition and with lower computational complexity. It is proved that, for any centrosymmetric array, the deterministic angular signal distribution function (DADF) vector has symmetric rotation invariance in literature [14]. This property is used to construct a polynomial, and we can obtain DOA estimation by finding the root of it. The paper [15] realizes the central DOA estimation of two-dimensional incoherent distributed sources with low complexity by using three parallel arrays and noncircular characteristics of signals. When it comes to large-scale MIMO, a beam space three parallel arrays and noncircular characteristics of incoherent distributed sources with low complexity by using [15] realizes the central DOA estimation of two-dimensional property is used to construct a polynomial, and we can minimize the angular signal distribution function (DADF) vector is proved that, for any centrosymmetric array, the determinant of the covariance matrix in documents [19]. An efficient approach for estimating the DOA of coherent signals using coprime arrays is proposed in the literature [20]. However, vectorization converts the covariance matrix into a single snapshot receiving a model of a virtual array. The parameters are solved by spatial smoothing, which causes the aperture of the virtual array to be sacrificed. For Gaussian signals, the first-order and second-order statistics can completely describe their statistical characteristics. However, at present, most of the artificial communication signals, like binary phase-shift keying (BPSK) and M-ary amplitude shift keying (MASK), are non-Gaussian [21, 22], and their higher-order cumulants contain information, from which, we can gain more useful information to further improve the estimation performance. In the point source model, a MUSIC-like method is proposed based on the fourth-order cumulant matrix (FCM) [23]. FOC is utilized to expand the array aperture, which effectively improves the estimated number of sources compared with the method based on the second-order statistics. A redundancy-elimination fourth-order cumulant DOA estimation algorithm is proposed, which removes the redundant information of the FOC matrix in the MUSIC-like algorithm and significantly reduces the computational complexity [24, 25]. An algorithm is proposed for MA system identification using the third- and fourth-order cumulants [26]. The above fourth-order cumulant methods are all proposed for point source models, whose performance will deteriorate when applied to distributed source based on sparse arrays.

To solve the above problems, this paper combines the characteristics of exponential distribution with the fourth-order cumulant matrix to improve the performance of DOA estimation. The main contributions of this paper are as follows:

(1) The specific angular distributed form of the distributed source is not considered in the aforementioned algorithms; therefore, in this paper, the exponential-type coherent distributed source is considered, and the corresponding DOA estimation algorithm with high performance is designed by analyzing the characteristics of steering vector and combining with the large aperture characteristics of mutual matrix. At last, a novel parameter estimation algorithm of exponential-type coherent distributed source based on the fourth-order cumulant under sparse array is proposed.

The rest of the paper is organized as follows. In Section 2, the exponential-type coherent distributed source model under sparse array is given. According to the signal model, a DOA estimation algorithm of distributed source based on the fourth-order cumulant is proposed in Section 3. In Section 4, performance simulation experiments are carried out and simulation results are given. Our conclusions and directions for future research are presented in Section 5.

Table 1 shows the key notations and explanations in this paper.

2. Signal Model

The coprime array proposed in the literature [17] is considered in this paper. The coprime linear array is composed of two uniform linear subarrays, whose element numbers are $M_A$ and $M_B$, respectively. $M_A$ and $M_B$ are mutual primes, and the spacing of subarrays is $(M_B \lambda/2)$ and $(M_A \lambda/2)$, respectively, where $\lambda$ represents the wavelength of the carrier. The topological structure of the coprime linear array is shown in Figure 1.

It is assumed that there exist $K$ target distributed sources transmitting plane waves from center directions $\Theta = (\theta_1, \theta_2, \ldots, \theta_K)$ to the measuring array, which contains $M = M_A + M_B - 1$ array elements. $P$ represents the aggregation of the locations of all physical array elements, where $P = \{d_1, d_2, \ldots, d_{M-1}\}$ and $d_m$ represents the distance between the $m + 1$ th array element and the first array element. Then, the signal output vector can be expressed as

$$x(t) = \sum_{k=1}^{K} \int_{-\pi/2}^{\pi/2} a(\theta) s_k(\theta, t; \mu_k) d\theta + n(t),$$

(1)
The angular signal density function of the characteristics of the distributed source is an exponential distribution, namely,

\[ g_k(\theta - \theta_k; \rho_k) = \frac{1}{1 - \rho_k \exp \left( j(\theta - \theta_k) \right)} \]

where \( \rho_k \) is directly related to the angle extension parameter \( \sigma_k \) of the distributed source (0 < \( \rho_k \) ≤ 1). The closer \( \rho_k \) is to 1, the smaller the angle spread is.

In this case, the distributed source generalized direction vector of the exponential distribution can be expressed as

\[ \mathbf{b}(\theta_k, \rho_k) = \left[ \rho_k^{2d_1/\lambda} e^{-j2\pi(d_1\sin\theta_k/\lambda)} \cdots \rho_k^{2d_M/\lambda} e^{-j2\pi(d_M\sin\theta_k/\lambda)} \right] \]

Therefore, the received signal matrix can be expressed as

\[ \mathbf{x}(t) = \mathbf{B}\mathbf{s}(t) + \mathbf{n}(t). \]

### 3. Proposed Algorithm

#### 3.1. Method Introduction.

Generally speaking, the FOC can be defined in two ways, \( \text{cum}(\mathbf{x}_{m1}, \mathbf{x}_{m2}, \mathbf{x}_{m3}, \mathbf{x}_{m4}) \) and \( \text{cum}(\mathbf{x}_{m1}, \mathbf{x}_{m2}, \mathbf{x}_{m3}, \mathbf{x}_{m4}) \), respectively. For distributed sources, because of the particularity of the steering vector, the definition of the conventional fourth-order cumulant, namely, the former, is no longer applicable to the distributed source. Refer to Appendix A for a detailed analysis. Instead, the FOC of the received signal can be defined as the later, namely,

\[ \text{cum}(\mathbf{x}_{m1}, \mathbf{x}_{m2}, \mathbf{x}_{m3}, \mathbf{x}_{m4}) = \left[ \begin{array}{c} \mathbf{x}_{m1} \mathbf{x}_{m2}^* \mathbf{x}_{m3}^* \mathbf{x}_{m4}^* \\
- E[\mathbf{x}_{m1} \mathbf{x}_{m2}^*] E[\mathbf{x}_{m3} \mathbf{x}_{m4}^*] \\
- E[\mathbf{x}_{m1}^* \mathbf{x}_{m2}] E[\mathbf{x}_{m3}^* \mathbf{x}_{m4}] \\
- E[\mathbf{x}_{m1}^* \mathbf{x}_{m2}] E[\mathbf{x}_{m3} \mathbf{x}_{m4}^*] \end{array} \right] \]

where \( \mathbf{x}_m \) represents the received signal of the \( m \)th array element, \( \forall m_1, m_2, m_3, m_4 \in \{1, 2, \ldots, M\} \).

Then, defining \( \mathbf{C}_x \) as an \( M^2 \times M^2 \)-dimensional matrix whose element in \((m_1 - 1)M + m_2 \) th row and \((m_3 - 1)M + m_4 \) th column is \( \text{cum}(\mathbf{x}_{m1}, \mathbf{x}_{m2}, \mathbf{x}_{m3}, \mathbf{x}_{m4}) \), thus, \( \mathbf{C}_x \) can be calculated as follows:

\[ \mathbf{C}_x = E[\mathbf{x} \otimes \mathbf{y}] E[\mathbf{x} \otimes \mathbf{y}]^H \]

where the symbol \( \otimes \) represents the Kronecker product and the symbol \( \mathbf{x} \otimes \mathbf{y} \) represents the mixed-Kronecker product which is defined as follows:

\[ \mathbf{x} \otimes \mathbf{y} = \mathbf{J}^{-1} \left( \mathbf{x} \mathbf{y}^T \right) \]
\[
[Y_{mno} \otimes Z_{pq}]_{\text{mixed}} = \begin{bmatrix}
Y \otimes z_1, Y \otimes z_2, \ldots, Y \otimes z_q
\end{bmatrix}
\]
\[
= \begin{bmatrix}
y_{11} z_{11} \cdots y_{1n} z_{11} \cdots y_{1q} z_{1q} \\
\vdots & \ddots & \ddots & \ddots \\
y_{m1} z_{p1} \cdots y_{mn} z_{p1} \cdots y_{mn} z_{pq}
\end{bmatrix}
\]
(9)

Since it is assumed that \(n(t)\) is Gaussian noise and independent of the signal, therefore, its FOC is identical to zero, which can be substituted into the above equation; then, we can obtain
\[
C_x = (B \otimes B)C_T (B \otimes B)^H,
\]
(10)

where \(C_x\) is the FOC matrix of the signal \(s\) and is \(K \times K\) dimensions. When the signal is independent stationary non-Gaussian, the following formula is given:
\[
C_x = E\left[[s \otimes s] [s \otimes s]^H\right] - E[s \otimes s]E[s \otimes s]^H
\]
\[
- E\left[ss^H\right] \otimes E\left[ss^H\right]
\]
\[
- \left[E\left[ss^H\right] \otimes E\left[ss^H\right]\right]_{\text{mixed}},
\]
(11)

and independent sources have zero cross-cumulants only when \(i = j = k = l\); we have \(\text{cum}(s_i, s_j, s_k, s_l^*) \neq 0\). Hence,
\[
C_x = \begin{bmatrix}
\text{cum}(s_i, s_i^*, s_1, s_1^*) \\
\vdots \\
\text{cum}(s_K, s_K^*, s_K, s_K^*)
\end{bmatrix},
\]
(12)

and we have
\[
C_x = H(\theta, \rho)C_T H(\theta, \rho)^T,
\]
(13)

where
\[
H = \begin{bmatrix}
b(\theta_1, \rho_1) \otimes b(\theta_1, \rho_1), & \ldots, & b(\theta_K, \rho_K) \otimes b(\theta_K, \rho_K)
\end{bmatrix},
\]
(14)

\(H\) can be considered as the steering matrix of the virtual array, whose dimension is extended to \(M^2 \times K\).

Considering the \(((u-1)M + v)\)th row and the \(k\)th column element of the \(h(\theta_k, \rho_k)\), we have
\[
h_{u,v}(\theta_k, \rho_k) = \rho_k^{2(d_u+d_v)\lambda} e^{-\lambda \sigma((d_u+d_v) \sin \theta_k)},
\]
(15)

that is, the delay of distributed source signal from the center arrival angle \(\theta_k\) at \((d_u, d_v)\) relative to the reference position.

It is equivalent to the existence of a virtual array element at \((d_u + d_v)\), where \(d_u, d_v \in \mathbb{P}\).

We define the sum aggregate of physical array elements:
\[
Z = \left\{(d_u + d_v) \mid 1 \leq u, v \leq M\right\},
\]
(16)

which represents the location of all array elements in the extended virtual array.

We take the coprime array in Section 2; for example, when \(M_A = 2, M_B = 3\), the set of positions of physical sensors is \(\mathbb{P} = \{0, 2, 3, 4\}\). Virtual array elements are constructed by FOC used in this paper. The positions of sensor elements and physical sensors are shown in Figure 2.

We can observe that both fourth-order cumulant construction methods can extend the array aperture, and the expansion ability is not the same. Because the traditional construction method of the fourth-order cumulant can construct more virtual arrays, it is widely used in the point source model, but not in the distributed source model. Although the construction method of the fourth-order cumulant used in this paper is not as capable as that of the traditional method in extending the array aperture, it also obtains a larger DOF than the physical array. It is vital that the method used in this paper is applicable to the distributed source model.

The position of the largest continuous part of the virtual array element is denoted by \(\mathbb{P}'\), which contains \(\tilde{M}\) elements. Assume that \(h_{i_m}(\theta, \rho)\) is the \(i_m\)th element of \(h(\theta, \rho)\); then, the transformed virtual array steering vector is
\[
\tilde{h}(\theta, \rho) = \begin{bmatrix}
h_{i_1}(\theta, \rho), & \ldots, & h_{i_{\tilde{M}}}(\theta, \rho)
\end{bmatrix}^T, \quad i_m \in \mathbb{P}',
\]
(17)

which is \(\tilde{M}\)-dimensional column vectors, and the corresponding array that received signal covariance can be expressed as follows:
\[
C_{\text{vula}} = \begin{bmatrix}
(C_X)_{i_1,j_1} \cdots (C_X)_{i_1,j_{\tilde{M}}} \\
\vdots & \ddots & \ddots \\
(C_X)_{i_{\tilde{M}},j_1} \cdots (C_X)_{i_{\tilde{M}},j_{\tilde{M}}}
\end{bmatrix}, \quad i_m, j_m \in \mathbb{P}',
\]
(18)
3.2. Summary of the Method’s Steps. Through the introduction of the above algorithm, the steps of the method are summarized as follows:

1. Use the array receiving the signal to estimate the fourth-order cumulant matrix $C_n$ according to equation (8).

2. Construct the virtual array according to formula (16), and obtain the position set of its maximum continuous part array element.

3. Get rid of the redundancy of $C_n$, and construct the matrix $C_{vula}$ according to equation (18).

4. Conduct the eigenvalue decomposition for $C_{vula}$ and obtain the noise subspace.

5. Apply equation (20) to solve the center DOA and distribution parameters of distributed source.

3.3. Complexity Analysis. According to the introduction and summary of the above algorithm, the following two aspects make up the complexity of this algorithm: the solution of the fourth-order cumulant and the application of the subspace decomposition algorithm. The complexity of the fourth-order cumulant mainly includes formula (14). The complexity is $O(TM^4)$, where $T$ represents the number of snapshots. The subspace decomposition algorithm is an alternative method. The complexity of solving the DOA estimation by using the DSPE method includes the complexity of eigenvalue decomposition and that of a two-dimensional spectral peak search. The former is $O(M^2)$ and the latter is $O(M^2 + \tilde{M}T\rho_{j,f})$. The total complexity is $O(M^2 + (\tilde{M} + \tilde{M}^2)J_\rho J_\rho)$, where $J_\rho$ and $J_\rho$ represent the number of grids in a two-dimensional spectral peak search. $J_\rho = (180/\Delta \theta) + 1$ and $J_\rho = (1/\Delta \rho) + 1$, where $\Delta \theta$ and $\Delta \rho$ represent the step length of researching. LS-ESPRIT algorithm has advantages over the DSPE algorithm in reducing complexity, whose complexity is $O(4\tilde{M}^2T + 8\tilde{M}^2 + \tilde{K}M^2J_\rho J_\rho)$. Table 2 shows the computational complexity elaborately.

To get more intuition about the computational complexity, we experiment on computation complexity with different numbers of array elements and different numbers of the search grid. It can be seen from Figures 3 and 4 that the proposed method increases the computational complexity to some extent but improves the estimation accuracy and degree of freedom, so the complexity is within the accepted range.

4. Simulation Results

In this section, the simulation experiments of the proposed method are given. The root mean square error (RMSE) is regarded as the performance index of the proposed algorithm, and the definition of RMSE, which is the deviation between the estimated DOA and the true DOA, is given as follows:

$$RMSE = \sqrt{\frac{1}{NK} \sum_{j=1}^{N} \sum_{i=1}^{K} (\hat{\theta}_i - \theta_i) \cdot (\hat{\theta}_i - \theta_i)^2},$$

where $\theta_i$ is the true DOA of the distributed source, $\hat{\theta}_i$ is the estimated DOA coordinate of the $i$th distributed source by the $j$th Monte Carlo experiment, and $N$ is the number of Monte Carlo. The Cramer-Rao Lower Bound (CRLB) describes the best performance of the parameter estimation method in theory [27, 28]. We derive the CRLB about parameter vector $\mu$ as

$$CRLB(\mu) = \frac{1}{T}U^H\hat{\mu}U^{-1},$$

where $\hat{\mu} = I - V(V^H\Sigma^{-1}V)^{-1}V^H$. Refer to Appendix B for the detailed derivation of (22). The CRLB derived from the
Simulation 1. Estimate ability of different distribution parameters.

In this part, we draw the parameter coordinates estimated by each Monte Carlo experiment into a scatter diagram and describe the effectiveness of the estimation by its aggregation degree. The subarrays $M_A = 2M_B = 3$. The DOA of the three distributed sources is evenly distributed from $-20^\circ$ to $20^\circ$, and the parameter $\rho = 0.7, \rho = 0.8$, and $\rho = 0.9$, respectively. The step length of searching is $\Delta\theta = \Delta\rho = 0.01$. We conduct simulation experiments under the following three conditions: (1) SNR = 0 dB; (2) SNR = 10 dB; (3) SNR = 20 dB. For 200 independent trials, the simulation results are shown in Figures 5–7, corresponding to the three cases mentioned above, respectively. The black dots are the estimation results of all Monte Carlo experiments and the cross in red represents the true value. We observe that the simulation results in discrete distributions around true value; moreover, with the increase of SNR, the clustering degree of scattering becomes stronger; therefore, the proposed algorithm can give the correct estimations when different distributed sources have different parameters.

Simulation 2. Performance simulation under underdetermined condition.

Underdetermined condition refers to the case where the number of sources is larger than the number of array elements. The proposed algorithm improves the degree of freedom by constructing a virtual array, which can estimate more sources than the number of array elements. The subarrays $M_A = 2M_B = 3$. The DOA of the seven distributed sources is evenly distributed from $-50^\circ$ to $50^\circ$, and the parameter $\rho = 0.8$. The number of snapshots is $T = 10000$ and the step length of searching is $\Delta\theta = \Delta\rho = 0.01$. We conduct simulation experiments under the following two conditions: (1) SNR = 10 dB; (2) SNR = 20 dB. The normalized spectral peak is shown in Figure 8 when the SNR is 10 dB and in Figure 9 when it is 20 dB. In the figure, the red line is the spectral peak search curve, and the blue dashed line represents the true incoming wave direction. We observe that the position of the spectral peak is very close to the real value, so the proposed algorithm has the ability to estimate under undetermined conditions.

Simulation 3. RMSE of DOA estimates versus SNRs.

In order to research the performance of the proposed algorithm under different SNRs, we compare it with the DSPE algorithm and the LS-ESPRIT algorithm. The DOA of the three distributed sources is evenly distributed from $-20^\circ$ to $20^\circ$, and the parameter $\rho = 0.8$. The number of snapshots is set as $T = 2000$. The subarrays $M_A = 2M_B = 3$ in the proposed algorithm. The step length of searching is $\Delta\theta = \Delta\rho = 0.01$. To show the superiority of the proposed algorithm, the number of arrays is set as 4 in the DSPE algorithm and the distance between adjacent sensors is 0.5λ. Figure 10 shows the RMSE of the central DOA estimations with SNR ranging from $-4$ dB to 20 dB based on the 500 Monte Carlo experiments. We observe that, under the same SNR, the estimation accuracy of the proposed method is

Table 2: Complexity comparison.

| Algorithm               | Complexity                        |
|-------------------------|-----------------------------------|
| Uniform-DSPE            | $O(TM^2 + M^3 + (M^2 + M)\rho J)$ |
| Uniform-LS-ESPRIT       | $O(8M^3 + KM^2 J\rho + 4TM^2)$   |
| CLA-FOC-DSPE (proposed) | $O(TM^2 + M^3 + (M^2 + M)\rho J)$ |

Figure 3: Computation complexity with different numbers of array elements.

Figure 4: Computation complexity with different numbers of the search grid.

maximum-likelihood criterion will be plotted in the results of simulation experiments to measure the performance of the proposed method.
higher in contrast with that of the DSPE algorithm and the LS-ESPRIT algorithm when the array elements are the same and closer to the CRLB.

Simulation 4. RMSE of DOA estimates versus the number of snapshots.

In this part, we experiment on the effect of different snapshots on the performance of the algorithm. The SNR is 20 dB, and the other parameters are identical to those in Simulation 3. Figure 11 shows the RMSE of the central DOA estimations with the number of snapshots ranging from 400 to 8000 based on 500 Monte Carlo experiments. We observe that, under the same number of snapshots, the estimation accuracy of the proposed method is higher in contrast with that of the DSPE algorithm and the LS-ESPRIT algorithm when the array elements are the same and closer to the CRLB.

Simulation 5. RMSE of DOA estimates versus different parameters $\rho$.

In this part, we experiment on the effect of different parameters $\rho$ on the performance of the algorithm. The number of snapshots is fixed to 2000 and the SNR is set as 10 dB. The other parameters are the same as those in
Simulation 3. Based on 500 Monte Carlo experiments, the simulation results are shown in Figure 12. We observe that, with the increase of the parameter $\rho$, the performance of RMSE decreases. When the parameter $\rho = 1$, the distributed source model is reduced to the point source model; thus, one conclusion can be drawn; that is, the point source model is a particular case of a distributed source model.

Simulation 6. RMSE of DOA estimates versus different step lengths $\Delta \theta$.

In order to research the performance of the proposed algorithm under different step lengths of searching, we conduct 500 Monte Carlo experiments under the condition of different step lengths $\Delta \theta$ with the number of snapshots being fixed to 2000. The SNR ranges from $-4\, \text{dB}$ to $20\, \text{dB}$ and $\Delta \rho = 0.01$. The other parameters are identical to those in Simulation 3. The simulation results are shown in Figure 13. It is obvious that the smaller the step length $\Delta \theta$, the better the performance, but at the same time, the greater the complexity. Therefore, in practical applications, it is necessary to balance the weight of complexity and accuracy to obtain the best estimation effect.
5. Conclusion

In order to achieve high precision and high DOF estimation of multiple distributed sources, this paper introduces fourth-order coherent cumulant into distributed sources and presents a central DOA estimation method for the exponential-type coherent distributed source based on the fourth-order cumulant. Firstly, an exponential-type coherent distributed source signal model under sparse array is established, and according to the signal model, we combine the sparsity of array space domain with the fourth-order cumulant characteristics of signals. In this way, we extend the array aperture and degree of freedom and improve the estimation accuracy. Then, in order to get rid of redundancy, we get the maximum continuous part of the extended virtual array element. Finally, we can obtain the DOA through the subspace decomposition algorithm. Compared with the contrast algorithm based on the second-order statistics, the proposed algorithm improves the estimation accuracy and the number of estimable sources under the condition of the same number of array elements. Complexity analysis and numerical simulations are provided to demonstrate the superiority of the proposed algorithm.

Appendix

A. Detailed Reason Why the Former FOC Is No Longer Applicable to Distributed Source

The conventional FOC of the received signal can be expressed as

\[
\begin{align*}
\text{cum}(x_m, x_m, x_m, x_m) &= E[x_m x_m x_m x_m] \\
&= E[x_m x_m] E[x_m x_m] - E[x_m x_m] E[x_m x_m] \\
&- E[x_m x_m] E[x_m x_m] \\
&- E[x_m x_m] E[x_m x_m] (A.1)
\end{align*}
\]

where \(x_m\) represents the received signal of the \(m\)th array element. Then, defining \(C_\theta\) as an \(M^2 \times M^2\)-dimensional matrix whose element in \(((m_1 - 1)M + m_2)\)th row and \(((m_2 - 1)M + m_1)\)th column is \(E[x_m x_m x_m x_m]\), thus, \(C_\theta\) can be calculated as follows:

\[
\begin{align*}
C_\theta &= E[xx^H] - E[xx^*] E[xx^*]^H \\
&= E[xx^H] - E[xx^*] E[xx^*]^H + E[xx^H] - E[xx^*] E[xx^*]^H \\
&= C_{\text{mixed}}.
\end{align*}
\]

where the symbol \(\otimes\) represents the Kronecker product.

Since \(n(t)\) is Gaussian noise and independent of the signal, therefore, its FOC is identical to zero, which can be substituted into the above equation; then, we can obtain

\[
C_x = (B \otimes B^*) C_x (B \otimes B^*)^H, \quad (A.3)
\]

where \(C_x\) is the FOC matrix of the signal \(s\) and is \(K \times K\) dimensions. When the signal is independent stationary non-Gaussian, the following formula is given:

\[
\begin{align*}
C_x &= E[ss^* [ss^*]^H] \\
&= E[ss^*] E[ss^*]^H \\
&= E[ss^*] E[ss^*]^H + E[ss^*] E[ss^*]^H \\
&= C_{\text{mixed}}.
\end{align*}
\]

and independent sources have zero cross-cumulants only when \(i = j = k = l\); we have \(\text{cum}(s_i, s_j, s_k, s_l) \neq 0\).

Hence,

\[
C_x = \begin{bmatrix}
\text{cum}(s_i, s_i, s_i, s_i) \\
\vdots \\
\text{cum}(s_K, s_K, s_K, s_K)
\end{bmatrix},
\]

and we have

\[
C_x = (B \otimes B^*) C_x (B \otimes B^*)^H. \quad (A.6)
\]

\(C_x\) is the covariance matrix of the output signal of the virtual array after expansion, where

\[
H = \{b(\theta_1, \rho_1) \otimes b^*(\theta_1, \rho_1), \ldots, b(\theta_K, \rho_K) \otimes b^*(\theta_K, \rho_K)\}. \quad (A.7)
\]

\(H\) can be considered as the steering matrix of the virtual array, whose dimension is extended to \(M^2 \times K\).

Considering the \(((u - 1) M + v)\)th row and the \(k\)th column element of the \(h\) \((\theta_i, \rho_i)\), we have

\[
h_{u,v}(\theta_i, \rho_i) = \rho_k^2 (d_i^2 + d_i^2) e^{j 2\pi (\sin \theta_i, \rho_i))}. \quad (A.8)
\]

From the formula above, we can see that the extended steering vector does not have the form of a distributed source steering vector. So, the traditional construction of the fourth-order cumulant is no longer applicable to the distributed source model.
B. The CRLB Derived from the Maximum-Likelihood Criterion

The process of deriving CRLB from the maximum-likelihood criterion is as follows.

From Sections 2 and 3, we can obtain
\[
C_x = H(\theta, \rho)C_y H(\theta, \rho)^T + \sigma_n^2 I_M, \tag{B.1}
\]
where
\[
H = [b(\theta_1, \rho_1) \otimes b(\theta_1, \rho_1), \ldots, b(\theta_K, \rho_K) \otimes b(\theta_K, \rho_K)],
\]
\[
b(\theta_k, \rho_k) = \left[ \begin{array}{c} 2d_{k1} e^{-j2\pi (d_{11} \sin \theta_k \rho_k) \ldots} d_{k1} e^{-j2\pi (d_{11} \sin \theta_k \rho_k)} \end{array} \right].
\]

And then, the vector \( \mu \) is denoted by
\[
\mu = [\theta^T, \rho^T] \in \mathbb{R}^{2K+1}, \tag{B.3}
\]
where \( \theta = (\theta_1, \theta_2, \ldots, \theta_K) \), \( \rho = (\rho_1, \rho_2, \ldots, \rho_K) \).

The vector \( \nu \) is denoted by
\[
\nu = [\sigma_r, \sigma_n^2] \in \mathbb{R}^{(K+1)\times 1}, \tag{B.4}
\]
where \( \sigma = (\sigma_r, \sigma_1^2, \ldots, \sigma_n^2) \) represents the power of signals, and \( \sigma_n^2 \) represents the power of noise.

Then, we define a vector \( \eta \) as
\[
\eta = [\mu^T, \nu^T]^T \in \mathbb{R}^{(3K+1)\times 1}, \tag{B.5}
\]
which contains all the unknown parameters. The CRLB of \( \eta \) can be calculated from [29]:
\[
\text{CRLB}(\eta) = F^{-1}. \tag{B.6}
\]

\( F \) is the Fisher information matrix (FIM):
\[
[F]_{mn} = T \text{tr} \left( C^{-1} \frac{\partial C}{\partial \eta_m} C^{-1} \frac{\partial C}{\partial \eta_n} \right), \tag{B.7}
\]
\[
\frac{1}{T} F = \left( \frac{\partial l}{\partial \nu} \right)^T \left( C^{-1} \otimes C^{-1} \right) \left( \frac{\partial l}{\partial \nu} \right), \tag{B.8}
\]
where \( l = \text{vec}(C_M) \), and \( T \) is the number of snapshots.

We apply the partition
\[
\left( C^{-(T/2)} \otimes C^{-1/2} \right) \left[ \begin{array}{c} \frac{\partial l}{\partial \mu} \\ \frac{\partial l}{\partial \rho} \end{array} \right] \doteq [U | V], \tag{B.9}
\]
where
\[
U = \left( C^{-(T/2)} \otimes C^{-1/2} \right) \left[ \begin{array}{c} \frac{\partial l}{\partial \mu} \\ \frac{\partial l}{\partial \rho} \end{array} \right] = [U_\theta | U_\rho], \tag{B.10}
\]
\[
V = \left( C^{-(T/2)} \otimes C^{-1/2} \right) \left[ \begin{array}{c} \frac{\partial l}{\partial \sigma_r} \\ \frac{\partial l}{\partial \sigma_n^2} \end{array} \right] = [V_\sigma | V_n].
\]

Then, (B.8) can be rewritten as
\[
\frac{1}{T} F = \begin{bmatrix} U^H & U^H \ V \end{bmatrix}. \tag{B.11}
\]

The \( k \) th column of \( U_\theta, U_\rho, V_\sigma, \) and \( V_n \) is given by
\[
U_\theta (:; k) = \text{vec} \left( C^{-1/2} \frac{\partial C}{\partial \theta_k} C^{-1/2} \right), \tag{B.12}
\]
\[
U_\rho (:; k) = \text{vec} \left( C^{-1/2} \frac{\partial C}{\partial \rho_k} C^{-1/2} \right),
\]
\[
V_\sigma (:, k) = \text{vec}(I_M),
\]
\[
V_n = \text{vec}(R^{-1}). \tag{B.14}
\]

Data Availability

The data, which are produced by simulations, used to support the findings of this study are available from the corresponding author upon request.

Disclosure

The authors claim that the data used in this article are provided by their simulations and this is developed without using any data in a published article to support their results.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The work was supported by the National Natural Science Foundation of China (Grant no. 61401513).

References

[1] H. Krim and M. Viberg, “Two decades of array signal processing: the parametric approach,” IEEE Signal Processing Magazine, vol. 13, pp. 87–94, 1996.
[2] A. Hero, H. Messer, J. Goldberg et al., “Highlights of statistical signal and array processing,” IEEE Signal Processing Magazine, vol. 15, no. 5, pp. 21–64, 1998.
[3] P. Zetterberg and B. Ottersten, “The spectrum efficiency of a base station antenna array system for spatially selective transmission,” IEEE Transactions on Vehicular Technology, vol. 44, no. 3, pp. 651–660, 1995.
[4] P. Stoica and A. Nehorai, “Musical maximum likelihood, and cramer-rao bound,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 37, no. 5, pp. 720–741, 1989.
