**Research Article**

**A Design Method of Sparse Array with High Degrees of Freedom Based on Fourth-Order Cumulants**

*Fengtong Mei, Daming Wang, Chunxiao Jian, Yinsheng Wang, and Weijia Cui*

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

Correspondence should be addressed to Weijia Cui; cuilink_work@sina.com

Received 13 March 2021; Revised 31 May 2021; Accepted 11 June 2021; Published 2 July 2021

Academic Editor: Xiaofei Zhang

Copyright © 2021 Fengtong Mei 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.

Recently, the design of sparse linear array for direction of arrival (DOA) estimation of non-Gaussian signals has attracted considerable interest due to the fact that the fourth-order difference coarray offered by non-Gaussian significantly increases the aperture of a virtual linear array, which improves the performance of DOA estimation. In this paper, a super four-level nested array (S-FL-NA) configuration based on fourth-order cumulants (FOC) is proposed. The S-FL-NA consists of uniform linear arrays which have different interelement spacing. The proposed array configuration is designed based on interelement spacing, which, for a given number of sensors, is uniquely determined by a closed-form expression. We also derive the closed-form expression for the degrees of freedom (DOFs) of the proposed array. The optimal distribution of the number of sensors in each uniform linear array of the proposed array is given for an arbitrary number of sensors. Compared with the existing sparse arrays, the proposed array can provide a higher number of degrees of freedom and a larger physical array aperture. In addition, to improve the calculation speed of the fourth-order cumulant matrix, we simplify the FOC matrix by removing some redundancy. Numerical simulations are conducted to verify the superiority of the S-FL-NA over other sparse arrays.

1. Introduction

Direction-of-arrival (DOA) estimation is one of the important research fields in array signal processing [1–6]. For uniform linear array (ULA), the traditional estimation method such as multiple signal classification (MUSIC) [7] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [8] can only resolve \( M - 1 \) sources under the condition of \( M \) physical sensors. However, the underdetermined conditions in which the number of target sources is larger than that of array sensors are very common in the real world.

To solve this problem, many sparse array structures have been proposed to increase the number of degrees of freedom (DOFs), especially for underdetermined conditions. The minimum redundant array (MRA) [9] is a well-known sparse array that maximizes the number of consecutive lags based on difference coarray. However, due to the fact that the closed-form expression of the configuration of MRA does not exist, it is difficult for array design and precise DOF analysis. The nested arrays (NA) [10] and coprime arrays (CPA) [11] are the most notable sparse arrays presented recently. They both consist of two uniform linear arrays, which make it possible to express the structure of NA and CPA analytically. In [12], coprime arrays with compressed interelement spacing (CACIS) and coprime arrays with displaced subarrays (CADiS) are proposed, which increases the DOF and improves estimation accuracy greatly. Based on these arrays, many underdetermined DOA estimation approaches have been proposed, such as spatial smoothing subspace MUSIC (SS-MUSIC) [13] and compressed sensing algorithm [14–16].

The designs above and estimation methods of the sparse arrays are mainly based on second-order statistics. For non-Gaussian signals, the fourth-order cumulants can also be used to achieve DOA estimation [17, 18] because the fourth-order cumulants of non-Gaussian signals are not zero, while those of Gaussian signals are zero. In [19], DOA estimation based on the fourth-order cumulant (FOC) has been exploited for underdetermined conditions. The method of
virtual array concept based on FOC is proposed in [20]. Compared with second-order statistics, utilizing the fourth-order cumulant can improve the degrees of freedom and eliminate the Gaussian noise. Therefore, it is attractive to exploit how to construct a sparse array based on the fourth-order cumulant. The \(2q\)-th-order difference coarray concept and \(2q\)-level nested arrays are proposed in [21] based on the \(2q\)-th-order cumulants [22, 23]. A four-level NA (FL-NA) concept to the case of FOC is investigated as a special case of \(2q\)-level nested arrays when \(q\) is equal to 2. By adding to the NA and CPA a third uniform linear array, E-NA and E-CPA are proposed based on FOC in [24, 25], respectively. In [26], an expanding and shift scheme is discussed by considering fourth-order difference coarrays as the sum set of two second-order difference lags, leading to the expanding and shifted nested array (EAS-NA-NA) with more DOFs. A two-level NA based on FOC is developed in [27], without a gap to break the consecutiveness of coarray. An enhanced four-level nested array (E-FL-NA) is proposed [28] by adding two extra subarrays to the NA, achieving a higher number of the consecutive coarray lags than the aforementioned sparse arrays. Although these array structures increase DOFs greatly, the inner structure of their virtual arrays has not been fully analyzed in the researches given above. Therefore, they are not optimum, and further improvement is possible.

In this paper, a novel four-level nested array structure is proposed based on FOC, which is composed of four sparse ULAs, which provides the largest aperture virtual uniform linear array compared to the existing array structures with a given number of sensors. The closed-form expressions for the physical sensor location and the virtual aperture are derived, and the optimal distribution of sensors to each of the sparse ULAs is also given. Then, to increase the calculation speed of the FOC matrix, we simplify the FOC matrix by removing some redundancy of that.

Throughout this paper, we make use of the following notations. Matrices and vectors are represented by capital letters and lower letters in boldface, respectively. Given a matrix \(A\), we use \(A^T\), \(A^H\), and \(A^*\) to denote the transpose, the Hermitian transpose, and the conjugate of \(A\), respectively. We use \(\otimes\) and \(\odot\) to denote the Kronecker product and the Khatri–Rao product, respectively. The statistical expectation is denoted by \(\mathbb{E}(\cdot)\), and \(\text{vec}(\cdot)\) is the vectorization operation.

The rest of this paper is organized as follows. In Section 2, the signal model is presented. Section 3 describes the novel sparse array structure based on fourth-order difference coarrays and compares the DOFs of the proposed array with those of FL-NA, E-CPA, E-FL-NA, and EAS-NA-NA. Section 4 describes the simplified FOC matrix. Section 5 presents numerical examples, and Section 6 summarizes this paper.

2. Cumulant-Based Fourth-Order Difference Coarray

2.1. Signal Model. Consider a sparse linear array with \(M\) physical sensors; the unit spacing is equal to half wavelength \((\lambda/2)\), denoted by \(d\). The set of sensor positions can be expressed as

\[
\mathcal{S} = \{q_1 d, q_2 d, \ldots, q_M d\},
\]

where \(q_i\) is an integer. Suppose that there are \(K\) far-field uncorrelated narrow-band sources denoted by \(s(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T\) impinging on the array from directions \(\theta = [\theta_1, \theta_2, \ldots, \theta_K]\), and the steering vector at direction \(\theta_k\) is given by

\[
a(\theta_k) = [e^{-j\beta_k d \sin \theta_1}, e^{-j\beta_k d \sin \theta_2}, \ldots, e^{-j\beta_k d \sin \theta_K}]^T,
\]

where \(\beta_k = 2\pi d \sin \theta_k / \lambda\). The received data of the \(m\)-th sensor can be written as

\[
x_m(t) = \sum_{k=1}^{K} s_k(t)e^{-j\beta_k d \sin \theta_k} + n_m(t), \quad t \in \{1, 2, \ldots, T\},
\]

where \(n_m(t)\) is independent and identically distributed (i.i.d.) additive Gaussian noise with power \(\sigma_n^2\) and it is independent of the sources. \(T\) is the number of snapshots. We further assume that all signals are non-Gaussian with zero mean, and then the fourth-order cumulants of them are not zero.

2.2. Cumulant-Based Fourth-Order Difference Coarray. There are several definitions of FOC in the literature. In this paper, we utilize the following definition for the calculation of FOC:

\[
C_x(m, n, u, v) = \text{cum}(x_m, x_n, x_u, x_v)
\]

\[
= \mathbb{E}(x_m, x_n, x_u, x_v) - \mathbb{E}(x_m, x_n)\mathbb{E}(x_u, x_v) - \mathbb{E}(x_m, x_u)\mathbb{E}(x_n, x_v) - \mathbb{E}(x_m, x_v)\mathbb{E}(x_n, x_u), \tag{4}
\]

where \(x_m, x_n, x_u,\) and \(x_v\) denote the received signals at the \(m\)-th, \(n\)-th, \(u\)-th, and \(v\)-th sensors, respectively, and \(\text{cum}(\cdot)\) is the cumulant operator. Substituting (3) into (4) gives

\[
C_x(m, n, u, v) = \sum_{k=1}^{K} \gamma_k e^{-2\pi j (\eta_m - \eta_n) + (\eta_u - \eta_v)} \sin \theta_k, \tag{5}
\]

where \(\gamma_k = \text{cum}(s_k, s_k, s_k, s_k)\). Consider an \(M^2 \times M^2\) matrix denoted by \(C_x\); if the element in row \((m - 1)M + u\) and column \((n - 1)M + v\) of the matrix is equal to \(C_x\), the matrix can be calculated by the following formula:

\[
C_x = B(\theta)C_{k}\mathbb{B}^H(\theta), \tag{6}
\]

where

\[
B(\theta) = \begin{bmatrix} \mathbf{b}(\theta_1), \mathbf{b}(\theta_2), \ldots, \mathbf{b}(\theta_K) \end{bmatrix},
\]

\[
\mathbf{b}(\theta_k) = a(\theta_k) \otimes \mathbf{a}^*(\theta_k), \tag{7}
\]

\[
C_{k} = \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_K).
\]

\(B(\theta)\) is called the equivalent array manifold of FOC. In order to increase the array degrees of freedom, \(C_x\) can be vectorized as follows:
we can obtain all the properties of the set. We denote the array with elements located at the location set as the virtual array based on second-order difference coarray stage. By permutation invariance, it can also be seen as the difference coarray of the virtual array generated by the second-order sum coarray. Set \( C \) is centrally symmetric about the origin. Therefore, by analyzing its nonnegative part only, we can obtain all the properties of the set. We denote the nonnegative part of \( C \) as \( C^+ \).

According to (10), a nested array consists of two uniform linear arrays, where the first subarray has \( N_1 - 1 \) sensors with an interspacing \( d \) and the second subarray has \( N_2 \) sensors with the interspacing \( N_1 d \). A topological structure of NA is shown in Figure 1, and the sensor positions of a nested array can be written as

\[
S = \{ n_1 d | n_1 = 1, 2, \ldots, N_1 - 1 \} \cup \{ n_2 N_1 d | n_2 = 1, 2, \ldots, N_2 \}.
\]

Therefore, the second-order difference coarray of a NA can be expressed as

\[
Z_1 = \{ z | -N_1 N_2 + 1 \leq z \leq N_1 N_2 - 1 \},
\]

and the sum coarray of that is given by

\[
Z_2 = Z_{21} \cup Z_{22},
\]

where

\[
Z_{21} = \{ z | 0 \leq z \leq N_1 N_2 + 1 \},
\]

\[
Z_{22} = \{ z | z = 2N_1 | N_2 + 2 \leq z \leq 2N_2 \}.
\]

Note that \( Z_1 \) contains the consecutive range from \(-N_1 N_2 + 1\) to \(N_1 N_2 - 1\), and \( Z_{21} \) represents the consecutive range of \( Z_2 \) from 2 to \(N_1 N_2 + N_1\), while \( Z_{22} \) denotes the sparse range with interelement spacing \( N_1 \) from \((N_1 + 2)N_1\) to \(2N_1 N_2\).

### 3. The Proposed Four-Level Nested Array Structure

In this section, to maximize the number of consecutive lags in the virtual array based on FOC, we propose a novel sparse array composed of four uniform linear arrays. Then, the optimizing distribution of the number of sensors of the four ULAs is derived and the comparison of DOFs for the proposed array with other sparse arrays is made, which shows that the proposed array can provide a higher number of DOFs and a larger physical array aperture.

#### 3.1. The Super Four-Level Nested Array

**Proposition 1.** The super four-level nested array (S-FL-NA) is composed of four sparse uniform linear arrays, which is shown in Figure 2. The locations of the S-FL-NA are given by

\[
\begin{align*}
S & = S_1 \cup S_2 \cup S_3 \cup S_4, \\
S_1 & = \{ n_1 d | n_1 = 1, 2, \ldots, N_1 - 1 \}, \\
S_2 & = \{ n_2 N_1 d | n_2 = 1, 2, \ldots, N_2 \}, \\
S_3 & = \{ n_1 N_2 d + n_3 D d | n_3 = 1, 2, \ldots, N_3 - 1 \}, \\
S_4 & = \{ n_1 N_2 + (N_4 - 1)D d + n_4 N_4 d | n_4 = 1, 2, \ldots, N_4 - 1 \},
\end{align*}
\]

where \( D = 3N_1 N_2 - 2 \). The set of the consecutive lags of the fourth-order difference coarray for the S-FL-NA is given by

\[
C_c = \{ \eta | -M_{\text{max}} \leq \eta \leq M_{\text{max}} \},
\]

where \( M_{\text{max}} = 3N_1 N_2 N_4 - 2N_3 N_4 - N_1 N_2 \). For simplicity, we use \( M_s = 2M_{\text{max}} + 1 \) to represent the number of consecutive lags.

**Proof.** The S-FL-NA can be seen as an array composed of two subarrays denoted by \( S_{12} \) and \( S_{34} \), which can be expressed as

\[
\begin{align*}
S_{12} & = S_1 \cup S_2, \\
S_{34} & = S_3 \cup S_4 \cup \{ N_1 N_2 \}.
\end{align*}
\]

Obviously, \( S_{13} \) is the same nested array as (11). \( S_{34} \) is an expanded nested array, where the first subarray has \( N_3 - 1 \) sensors starting from the position \( N_1 N_2 \) with the interelement spacing \( D d \) and the second subarray has \( N_4 \) sensors with the interelement spacing \( N_1 D d \).

Now, we consider analyzing the fourth-order difference coarray of the S-FL-NA using the two sets, given by

\[
\begin{align*}
C_1 & = \{ (q_m - q_u) | q_m, q_u \in S_{12}; q_m, q_u \in S_{34} \}, \\
C_2 & = \{ (q_m + q_v) | q_m, q_v \in S_{12}; q_m, q_v \in S_{34} \},
\end{align*}
\]

where \( C_1 \) represents the second-order sum coarray of the virtual array generated by the second-order difference coarray stage and \( C_2 \) denotes the second-order difference coarray of the virtual array generated at the second-order sum coarray stage, when \( q_m, q_u \in S_{12} \) and \( q_m, q_v \in S_{34} \). The relationship between the two sets and \( C \) can be expressed as
Hence, the second-order sum coarray of $S_{34}$ is expanded and sensor locations are translated to the right.

where

$$Z_3 = \{z \mid -N_4N_3 + 1 \leq z \leq N_3N_4 - 1\}.$$  \hfill (21)

Substituting (11) and (21) into (18) yields

$$C_1 = \{z_1 + z_2 \mid z_1 \in Z_1; z_2 \in Z_3\}.$$  \hfill (22)

Therefore, we get

$$C_1 = \bigcup_{i=-N_1N_4+1}^{N_1N_4-1} \mathcal{Y}_i,$$  \hfill (23)

where

$$\mathcal{Y}_i = \{z + iD \mid -N_1N_2 + 1 \leq z \leq N_1N_2 - 1\}.$$  \hfill (24)

Obviously, each $\mathcal{Y}_i$ is a continuous set of integers.

Second, we investigate the inner structure of $C_2$. For $S_{34}$, it is generated by a nested array whose interelement spacing is expanded and sensor locations are translated to the right. Hence, the second-order sum coarray of $S_{34}$ is given by

$$Z_4 = Z_{41} \cup Z_{42},$$  \hfill (25)

where

$$\left\{ \begin{array}{l} Z_{41} = \{z + 2N_1N_2 \mid 0 \leq z \leq N_1N_2 + N_3 - 2\}, \\ Z_{42} = \{(z N_3 - 2) + 2N_1N_2 \mid 0 \leq z \leq 2N_4\} \end{array} \right.$$  \hfill (26)

Note that the set $Z_{41}$ is an ULA with interelement spacing $Dd$. Substituting (14) and (25) into (19) gives

$$C_2 = \{z_2 - z_4 \mid z_1 \in Z_2; z_2 \in Z_4\}.$$  \hfill (27)

Next, we focus on $Z_{23}$ and $Z_{41}$. The difference set between $Z_{23}$ and $Z_{41}$ is given by

$$C'_2 = \{z_2 - z_4 \mid z_1 \in Z_{23}; z_2 \in Z_{41}\} = \bigcup_{i=0}^{N_1N_4+N_3-1} \mathcal{W}_i,$$  \hfill (28)

where

$$\mathcal{W}_i = \{z + iD \mid N_2 N_1 - N_1 \leq z \leq 2N_1 N_2 - 2\}.$$  \hfill (29)

Obviously, each $\mathcal{W}_i$ is a set with continuous integers.

By analyzing $\mathcal{Y}_i$ and $\mathcal{W}_i$, the following results can be given:

$$\max\{\mathcal{Y}_i\} - \min\{\mathcal{W}_i\} = iD + N_1N_2 - 1 - (iD + N_1N_2 - N_1) = N_1 - 1,$$  \hfill (30)

indicating that $\mathcal{Y}_i \cup \mathcal{W}_i$ is a consecutive integer set.

Figure 2: Topological structure of S-FL-NA, consisting of four sparse uniform linear arrays.

To maximize the consecutive lags in set $C_1 \cup C'_2$, we consider constructing a new consecutive set $\mathcal{Y}_{i+1} \cup \mathcal{W}_i \cup \mathcal{Y}_{i+1}$.

Hence, the following relationship should be satisfied:

$$\min\{\mathcal{Y}_{i+1}\} = \max\{\mathcal{W}_i\} + 1.$$  \hfill (31)

Then we get the following equation:

$$(i + 1)D - N_1N_2 + 1 = iD + 2N_1N_2 - 2 + 1.$$  \hfill (32)

By solving (32), the value of $D$ can be expressed as

$$D = 3N_1N_2 - 2.$$  \hfill (33)

As a result, the set of consecutive integers generated by the combination of $C_1$ and $C'_2$ can be expressed as

$$C' = \bigcup_{i=0}^{N_1N_4+N_3-1} \{Y_i \cup W_i\}.$$  \hfill (34)

The max value of $C'$ is equal to the max value of $W_{N_1N_2-1}$, i.e.,

$$\max\{C'\} = 3N_4N_3N_2N_1 - 2N_3N_4 - N_2N_1.$$  \hfill (35)

For $C$ is centrally symmetric about the origin and $C' \in C$, the proof is finished.

The following remarks are the comparison between S-FL-NA and FL-NA [21], EAS-NA-NA [26], and E-FL-NA [28].

**Remark 1.** According to (11), the S-FL-NA can be seen as a combination of two nested arrays. The configuration of NA1 is $(N_1-1, N_2)$. NA2 is an array that is configured as $(N_3-1, N_4)$ and then shifts $N_1N_3$ to the right, so the first element of NA2 coincides with the last one of NA1. Therefore, the number of sensors of S-FL-NA is $M = \sum_{i=1}^{N-1} N_m - 3$, which is the same as the other three arrays.

**Remark 2.** FL-NA, EAS-NA-NA, E-FL-NA, and S-FL-NA all consist of four uniform linear arrays and have similar subarray 1 and subarray 2, however, the third subarrays of them have different interelement spacing. The interelements spacing of the third subarray of FL-NA is $N_1N_2$, and those of EAS-NA-NA and E-FL-NA are $2N_1N_2 - 1$, while that of S-FL-NA is $3N_1N_2 - 2$.

**Remark 3.** EAS-NA-NA and S-FL-NA can be seen as a sparse array composed of a nested array with sensor positions in $S_{12}$ plus the second array with sensor positions in $S_{34}$. However, the design of EAS-NA-NA is to maximize the number of consecutive lags in $C_1$, while S-FL-NA is to maximize the number of consecutive lags in $C_1 \cup C_2$. If not
considering $C_2$, the S-FL-NA will be reduced to EAS-NA-NA by making $\mathcal{Y}_1$ be adjacent to $\mathcal{Y}_{i+1}$.

To take full advantage of the array proposed above, we take the virtual uniform linear array part in $z$, and then the observation vector of the virtual ULA can be expressed as follows:

$$\tilde{z} = \tilde{B}_d\Phi^\dagger.$$  

(36)

We divided the virtual ULA into $M_v$ overlapping uniform subarrays of size $M_v$, where $M_v = M_{\text{max}} + 1$. The output of the $i$-th subarray is given by

$$\tilde{z}_i = \tilde{B}_{di}\Phi^\dagger\Pi_i,$$

(37)

where

$$\Phi = \begin{pmatrix} e^{-jn_1\sin \theta_1} & \cdots & e^{-jn_{M_v}\sin \theta_1} \\ \vdots & \ddots & \vdots \\ e^{-jn_1\sin \theta_k} & \cdots & e^{-jn_{M_v}\sin \theta_k} \end{pmatrix},$$

$$\tilde{b}(\theta_i) = \begin{pmatrix} 1 \\ \vdots \\ \tilde{b}(\theta_i) \end{pmatrix} \in \mathbb{C}^{M_v \times 1},$$

(38)

By averaging the covariance matrices of the smooth subarrays, we can obtain

$$\mathbf{R} = \frac{1}{M_v} \sum_{i=1}^{M_v} \tilde{z}_i\tilde{z}_i^H.$$  

(39)

The eigenvalue decomposition of $\mathbf{R}$ is carried out as follows:

$$\mathbf{R} = \mathbf{U}_s\Sigma_s^\dagger \mathbf{U}_s^H + \mathbf{U}_n\Sigma_n^\dagger \mathbf{U}_n^H.$$  

(40)

Since noise subspace and the signal subspace are orthogonal and the space formed by the steering vector and the signal subspace belong to the same space, the noise subspace and the steering vector are also orthogonal. Hence, the spatial spectrum of DOA is given by

$$G(\theta) = \frac{1}{(b(\theta))^H \mathbf{U}_n \Sigma_n^\dagger \mathbf{U}_n^H b(\theta)}.$$  

(41)

Then the DOAs can be obtained by the peak search method which divides the angle $0-2\pi$ into small parts and searches the $k$ maximum points of $G(\theta)$ to form the DOAs.

3.2. Optimal Sensor Allocation Problem. In this section, we address the problem of the optimal sensor distribution to maximize the degrees of freedom for a fixed number of physical sensors. According to the above proposition, the number of the consecutive virtual lags of S-FL-NA is $2M_{\text{max}} + 1$. Hence, the degrees of freedom optimization problem can be formulated as

$$\max(3N_1N_2N_3N_4 - 2N_1N_4 - N_2N_1),$$

subject to $N_1 + N_2 + N_3 + N_4 = M + 3$.  

(42)

In other words, what we are interested in is how to configure $(N_1, N_2, N_3, N_4)$ to make virtual ULA with the largest possible number of lags under a fixed number of sensors. We define integers $m$ and $n$ to be the quotient and remainder of $M + 3$ modulo 4, i.e., $M + 3 = 4m + n$, $0 \leq n \leq 3$. In terms of $m$ and $n$, the solution to (22) can then be completely characterized as given by the following proposition.

Proposition 2. One solution to the optimization problem in (42) can be expressed as

$$\begin{cases} \begin{aligned} N_1 &= N_2 = N_3 = N_4 = m, & \text{if } n = 0, \\
N_1 &= m + 1, N_2 = N_3 = N_4 = m, & \text{if } n = 1, \\
N_1 &= N_2 = m + 1, N_3 = N_4 = m, & \text{if } n = 2, \\
N_1 &= N_2 = N_3 = m + 1, N_4 = m, & \text{if } n = 3. \end{aligned} \end{cases}$$  

(43)

The length of the corresponding consecutive lags can be written as

$$L_{c} = \begin{cases} \begin{aligned} 6m^4 - 6m^2 + 1, & \text{if } n = 0, \\
6m^4 + 6m^2 - 2m^2 - 2m + 1, & \text{if } n = 1, \\
6m^4 + 12m^2 - 4m - 1, & \text{if } n = 2, \\
6m^4 + 18m^2 - 12m^2 - 2m - 1, & \text{if } n = 3. \end{aligned} \end{cases}$$  

(44)

where $m = \lfloor (M + 3)/4 \rfloor$.

Proof. See Appendix A.  

\[ \square \]

Example 1. For the illustrated purpose, Figure 3 shows an example of the proposed array. We assume that the number of physical sensors $M$ is 6. By using formula (42), we can get $N_1 = 3, N_2 = 2, N_3 = 2, N_4 = 1$. For $S_{12}$, its second-order difference coarray is $\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$ and second-order sum coarray is $\{2, 3, 4, 5, 6, 7, 8, 9\} \cup \{12\}$. For $S_{34}$, its second-order difference coarray and sum coarray are $\{-48, -32, -16, 0, 16, 32, 48\}$ and $\{12, 28, 44, 60, 76\} \cup \{108\}$, respectively.

Therefore, $C^+_1 = Y_0 \cup Y_1 \cup Y_2 \cup Y_3$, where $Y_0 = \{0, 1, 2, 3, 4, 5\}, Y_1 = \{11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21\}, Y_2 = \{27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37\}$, and $Y_3 = \{43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53\}$; and $C^+_2 = W_0 \cup W_1 \cup W_2 \cup W_3$, where $W_0 = \{3, 4, 5, 6, 7, 8, 9, 10\}$, $W_1 = \{19, 20, 21, 22, 23, 24, 25, 26\}$, $W_2 = \{35, 36, 37, 38, 39, 40, 41, 42\}$, and $W_3 = \{51, 52, 53, 54, 55, 56, 57, 58\}$. By combination of $Y_0 \cup W_0 \cup Y_1 \cup W_1 \cup Y_2 \cup W_2 \cup Y_3 \cup W_3$, we can obtain $C^+$ having 59 consecutive lags.
3.3. Comparison for Configurations Based on the Fourth-Order Difference Coarray. The sparse array based on fourth-order difference coarray proposed in this paper effectively increases the consecutive virtual aperture. The virtual aperture of the proposed array is larger than those of EAS-NA-NA and E-FL-NA, and the number of consecutive lags is more at the same time. For comparison between S-FL-NA and the E-FL-NA, we have the following corollary.

**Corollary 1.** Given the same number of physical sensors, the potential DOFs provided by our proposed array can be more than E-FL-NA.

*Proof.* Given \( \sum_{m=1}^{4} N_m - 3 \) sensors, the number of consecutive coarray lags provided by E-FL-NA is \( M_E = 4N_1N_2N_3N_4 + 2N_1N_2N_4 - 2N_1N_2N_3 - 2N_4 + 1 \), which is proved in [28]. For S-FL-NA, this number is \( M_S = 6N_1N_2N_3N_4 - 4N_1N_4 - 2N_2N_4 + 1 \). In the same configuration \((N_1, N_2, N_3, N_4)\), the difference between \( M_E \) and \( M_S \) is given by

\[
\Delta M = M_S - M_E = 2N_4(N_1N_3 - 1)(N_4 - 1) > 0.
\]

To ensure the existence of each level, \( N_3 \geq 2 \). \( \square \)

The number of consecutive lags at the fourth-order difference coarray stage for different arrays is listed in Table 1. It can be observed that S-FL-NA has the largest number of DOFs compared with FL-NA, E-CPA, EAS-NA-NA, and E-FL-NA. Note that the parameters \( N_1 \) and \( N_2 \) of E-CPA are coprime because E-CPA is an extension of coprime array.

### 4. Simplified FOC Matrix for Low Complexity

#### 4.1. Simplified FOC Matrix.** According to (6), we know that \( C_x \) is an \( M^2 \times M^2 \) matrix, which has great computational complexity. In the section, we propose a simplified fourth-order cumulant matrix denoted by \( \tilde{C}_x \) to reduce the computational complexity.

The FOC matrix \( C_x \) after vectorization is equivalent to the observation data of a virtual ULA with the manifold \( B_d \) which is an \( M^2 \) – dimensional vector. However, in the above estimation algorithm based on S-FL-NA, we only make use of the virtual ULA part with \( M_S \) lags, implying that the virtual steering vector \( B_d \) and the FOC matrix \( C_x \) include redundancy. Therefore, we think that it is only necessary to calculate the cumulant matrix containing virtual ULA parts to perform the subsequent vectorization, without calculating the entire FOC matrix. Accordingly, a simplified FOC matrix based on S-FL-NA is proposed.

First, we rewrite the FOC matrix \( C_x \) as

\[
C_x = \begin{pmatrix}
C_{x}^{1,1} & C_{x}^{1,2} & \cdots & C_{x}^{1,M} \\
C_{x}^{2,1} & C_{x}^{2,2} & \cdots & C_{x}^{2,M} \\
\vdots & \vdots & \ddots & \vdots \\
C_{x}^{M,1} & C_{x}^{M,2} & \cdots & C_{x}^{M,M}
\end{pmatrix}
\] (46)

Obviously, \( C_x \) consists of \( M \times M \) blocks, each of which is an FOC submatrix with size \( M \times M \). Specifically, the \((p,q)\)th block \( C_{x}^{p,q} \) is given by

\[
C_{x}^{p,q} = \text{cum}(x_p, x_q^*, x^*, x^T).
\] (47)

By analyzing the structure of \( C_x \) for S-FL-NA, we find that the virtual ULA part is contained in the following matrix blocks:

\[
\tilde{C}_{x}^{p,q} = \{C_{x}^{p,q}|0 \leq p, q \leq r; M - r \leq p, q \leq M\},
\] (48)

where \( r = \lceil M/4 \rceil \). Then we reconstruct a simplified FOC matrix \( \tilde{C}_x \), which is given by

\[
\tilde{C}_x = [\tilde{C}_{x}^{1,1}; \tilde{C}_{x}^{1,2}; \tilde{C}_{x}^{1,3}; \tilde{C}_{x}^{1,4}],
\] (49)

where
The complexity of calculating $C_x$ is $O(3M^4(T + 1) + 2M^2T + 2M^3 + M^2G_0)$. The complexity of the simplified FOC matrix calculation is $O((4r^2 - 4Mr/M^2)(3M^4(T + 1) + 2M^2T) + 2M^3 + M^2G_0)$. Table 2 shows the comparison of the complexity of the two methods.

Figure 4(a) shows that the complexity of calculating $C_x$ is higher than that of calculating $\bar{C}_x$, and when the number of array elements is small, the difference between the two is not big. As the number of array elements increases, the advantage of $\bar{C}_x$ becomes larger.

Figure 4(b) exhibits the changing curve of the ratio of $\bar{C}_x$ and $C_x$ as the number of array elements increases. It can be observed that although the complexity of $\bar{C}_x$ has been reduced to a certain extent, it will not be lower than 75 percent of $C_x$. The 75 percent ratio can be achieved when $M$ is an integer multiple of 4.

### 5. Simulation Results

In this section, simulation results are present to demonstrate the performance of the proposed array structure. First, the number of consecutive lags and physical aperture of five kinds of sparse array are compared, and then MUSIC spectra are simulated to show the number of distinguishable sources. Finally, the root mean square error (RMSE) results of the proposed array under different SNR and different number of snapshots are simulated, respectively, and compared with different arrays. The RMSE is defined as

$$
\text{RMSE} = \sqrt{\frac{1}{Q_mK} \sum_{m=1}^{Q_m} \sum_{k=1}^{K} (\hat{\theta}_k(m) - \theta_k)^2},
$$

where $Q_m$ is the number of Monte Carlo simulations, $K$ is the number of source targets, and $\hat{\theta}_k(m)$ is the $i$-th source DOA in the $m$-th Monte Carlo simulation. The simulation conditions are listed in Table 3.

#### 5.1. The Number of Consecutive Lags and Physical Aperture

In the first set of simulations, comparisons about the number of consecutive lags and physical aperture are carried out to show the superiority of the proposed array. We chose FL-NA [21], E-CPA [25], EAS-NA-NA [26], and E-FL-NA [28] as the comparison arrays here and set the number of sensors $M$ varying from 8 to 20. The number of consecutive lags versus the number of sensors is illustrated in Figure 5(a). It can be observed that the number of consecutive lags of S-FL-NA is significantly higher than those of the other four sparse arrays. Figure 5(b) shows the physical aperture with respect to the number of sensors. It is evident that the S-FL-NA has a larger physical aperture compared to the other sparse arrays. The above results indicate that the proposed array holds the promising potential to achieve better DOA estimation performance compared to the other arrays.

#### 5.2. MUSIC Spectra

For the second set of simulations, we consider 35 independent sources distributed between $-60^\circ$ and $60^\circ$, impinging on the sparse linear arrays consisting of 6 sensors, $(N_1, N_2, N_3, N_4)$ valued by $(3, 2, 2, 2)$. The parameters set are 10 dB SNR and 100000 snapshots to make sure it is sufficient to calculate the FOC matrix. Figure 6 shows the MUSIC spectra for four kinds of 6-element sparse arrays. We can observe that, by using FL-NA, EAS-NA-NA, and E-FL-NA, we are not able to resolve all the 35 sources. On the other hand, all the sources are successfully resolved if the proposed array is used, illustrating the improved
resolution capability and a higher number of DOFs offered by the proposed array.

5.3. The RMSE of DOA Estimation. The next simulations consider the RMSE performance versus the input SNR, as well as the number of snapshots. We assume that $K = 6$ signals from directions listed in Table 3 impinging on the sparse arrays. As such, physical apertures of all the arrays hold: $23 (FL - NA) \leq 35 (E - CPA) \leq 38 (EAS - NA - NA) \leq 44 (E - FL - NA) \leq 53 (S - FL - NA)$. Likewise, the number of consecutive lags for all the arrays satisfies $71 (FL - NA) \leq 85 (E - CPA) \leq 87 (EAS - NA - NA) \leq 97 (E - FL - NA) \leq 117 (S - FL - A)$. The fixed parameter setting is 0 dB SNR and 7500 snapshots. The RMSEs of DOA estimates with respect to the SNR and the number of snapshots are shown in Figure 7. It can be observed that, along with the increase of SNR and snapshots, all the RMSEs decrease and $S-FL-NA$ yields the lowest RMSE among all the arrays. This can be attributed to a larger physical aperture and a higher consecutive number of lags of $S-FL-NA$ compared with $FL-NA$, $E-CPA$, $EAS-NA-NA$, and $E-FL-NA$.

The Cramer–Rao lower bound (CRLB) is the lower bound of the unbiased estimation variance, which is used to

| Algorithms                  | Computational complexity                                      |
|-----------------------------|-------------------------------------------------------------|
| Unsimplified                | $O(3M^4(T + 1) + 2M^2T + 2M^3 + M_6^2G_{6})$                 |
| Simplified                  | $O((4r^2 - 4Mr)(3M^4(T + 1) + 2M^2T)/M^2 + 2M_4^1 + M_6^2G_{6})$ |

Table 2: The computational complexity of algorithms based on unsimplified and simplified FOC matrix.

| Simulation parameters         | Value |         |
|------------------------------|-------|---------|
| Antenna number               | $M = 6$  |
| Sensor locations of FL-NA    | 0, 1, 2, 5, 11, 23 |
| Sensor locations of E-CPA    | 0, 2, 3, 4, 6, 9, 35 |
| Sensor locations of EAS-NA-NA| 0, 1, 2, 5, 16, 38 |
| Sensor locations of E-FL-NA  | 0, 1, 2, 6, 17, 44 |
| Sensor locations of S-FL-NA  | 0, 1, 2, 5, 21, 53 |
| Source number                | $D = 7$  |
| DOAs                         | $-40^\circ, -26.7^\circ, -13.3^\circ, 0^\circ, 13.3^\circ, 26.7^\circ, 40^\circ$ |
| Carrier frequency            | $f = 2.1$ GHz |
| Speed of light               | $c = 3 \times 10^8$ m/s |
| Monte Carlo times            | $Q_m = 200$ |

Table 3: Simulation conditions for the experiments.

Figure 4: Comparison of computational complexity between $C_x$ and $\tilde{C}_x$ with respect to the number of sensors. (a) The complexity of $C_x$ and $\tilde{C}_x$. (b) The ratio of complexity between $\tilde{C}_x$ and $C_x$.
Figure 5: Comparison of the number of consecutive lags and physical aperture for proposed array and other sparse arrays. (a) The curves of the number of consecutive lags. (b) Physical aperture.

Figure 6: Continued.
measure the deviation of the DOA estimation. Based on [29, 30], we simulate the CRLB of the five kinds of sparse array. Figure 8 shows the results of the CRLB of the DOA estimation of the different sparse arrays. We can observe that the CRLB of S-FL-NA are lower than those of the other arrays, and the accuracy of DOA estimation is improved obviously.

5.4. The RMSE Based on Two Kinds of FOC Matrices. In order to verify the effectiveness of the method based on the simplified FOC matrix proposed in Section 4, we simulate its RMSE and compare it with the method using the prototype FOC matrix. The results are shown in Figure 8. It is evident that, under different signal-to-noise ratios and snapshots, the performance of the method based on $C_x$ is close to that
based on $C_x$, showcasing that the proposed method improves the calculation speed while ensuring accuracy.

6. Conclusion

In this paper, a four-level nested array composed of four ULAs has been proposed to maximize the number of consecutive lags based on the fourth-order difference coarray. The closed-form expressions for the physical sensor location and the virtual aperture are specifically derived. Given a fixed number of physical sensors, the optimal distribution of each subarray is also provided. In the future, it will be interesting to design a sparse array based on FOC to reduce the effect of mutual coupling. In the simulation part, numerical simulation results validate that the sparse array proposed in this paper can detect more sources and improve the DOA estimation accuracy because of the higher number of consecutive lags.

Appendix

Proof of Proposition 2. Let us start with a simple optimization problem:

$$\max (3UV),$$

s.t. $U + V = \tilde{M}$. 

(A.1)
When $U$ and $V$ are nonnegative integers, we apply the arithmetic mean-geometric mean (AM-GM) inequality to the optimization problem. The solution of (A.1) is $U = [M/2]$. If we modify the above optimization problem (A.1) to be

$$\text{max}(3UV - 2V - U),$$

s.t. $U + V = M$,

(A.2)

the solution to the optimization problem in (A.2) is also $U = [M/2]$. Therefore, by setting $U = N_1N_2$ and $V = N_3N_4$, formula (42) can be simplified as

$$\text{max}(3N_1N_2N_3N_4),$$

s.t. $N_1 + N_2 + N_3 + N_4 = M + 3$.

(A.3)

According to the proof of Theorem 3 in [21], we consider any pair integers of $(N_1, N_2, N_3, N_4)$, denoted by $N_u$ and $N_v$. If $N_u - N_v \geq 2$, we can create a new pair of integers $N_u'$ and $N_v'$, where $N_u' = N_u - 1$ and $N_v' = N_v - 1$, such that

$$N_u'N_v' = N_uN_v + (N_u + N_v - 1) > N_uN_v.$$ (A.4)

Under the condition that $N_u' + N_v' = N_uN_v$, $N_u'N_v'$ and $N_v'$ instead of $N_u$ and $N_v$ can strictly increase the objective function, as long as $N_u - N_v \geq 2$. Therefore, in the optimal solution $(N_1, N_2, N_3, N_4)$, the difference between any two integers cannot be more than one. So, the solution of (42) can be given by

$$\begin{cases} 
(m, m, m, m), & \text{if } n = 0, \\
(m + 1, m, m, m), & \text{if } n = 1, \\
(m + 1, m + 1, m, m), & \text{if } n = 2, \\
(m + 1, m + 1, m + 1, m), & \text{if } n = 3.
\end{cases}$$ (A.5)

By substituting (A.5) into (16), we obtain the number of the corresponding consecutive lags expressed by (44).

Data Availability

The authors claim that the data used in this article are provided by their simulations and the data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors confirm that the funding did not lead to any conflicts of interest regarding the publication of this manuscript.

Acknowledgments

This 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 research: the parametric approach,” IEEE Signal Processing Magazine, vol. 13, no. 4, pp. 67–94, 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] A. Das, “Real-valued sparse Bayesian learning for off-grid direction-of-arrival (DOA) estimation in ocean acoustics,” IEEE Journal of Oceanic Engineering, vol. 46, pp. 1–11, 2020.

[4] S. K. Yadav and N. V. George, “Fast direction-of-arrival estimation via coarray interpolation based on truncated nuclear norm regularization 2020,” IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 12, pp. 1–115, 2020.

[5] E. Soubies, A. Chinatto, P. Lazzabul, J. M. T. Romano, and L. Blanc-Féraud, “Direction-of-arrival estimation through exact continuous $\ell_2$, 0-norm relaxation,” IEEE Signal Processing Letters, vol. 28, pp. 16–20, 2021.

[6] H. Xiang, B. Chen, M. Yang, S. Xu, and Z. Li, “Improved direction-of-arrival estimation method based on LSTM neural networks with robustness to array imperfections,” Applied Intelligence, vol. 51, pp. 1–14, 2021.

[7] R. Schmidt and R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” IEEE Transactions on Antennas and Propagation, vol. 34, no. 3, pp. 276–280, 1986.

[8] R. Roy and T. Kailath, “ESPRIT—estimation of signal parameters via rotational invariance techniques,” IEEE Transactions on Acoustics Speech and Signal Processing, vol. 37, no. 7, pp. 984–995, 1989.

[9] A. Moffet, “Minimum-redundancy linear arrays,” IEEE Transactions on Antennas and Propagation, vol. 16, no. 2, pp. 172–175, 1968.

[10] P. Pal and P. P. Vaidyanathan, “Nested arrays: a novel approach to array processing with enhanced degrees of freedom,” IEEE Transactions on Signal Processing, vol. 58, no. 8, pp. 4167–4181, 2010.

[11] P. P. Vaidyanathan and P. Pal, “Sparse sensing with co-prime samplers and arrays,” IEEE Transactions on Signal Processing, vol. 59, pp. 573–586, 2010.

[12] S. Qin, Y. D. Zhang, and M. G. Amin, “Generalized coprime array configurations for direction-of-arrival estimation,” IEEE Transactions on Signal Processing, vol. 63, no. 6, pp. 1377–1390, 2015.

[13] P. Pal and P. P. Vaidyanathan, “Coprime sampling and the music algorithm,” in Proceedings of the Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE), pp. 289–294, Sedona, AZ, USA, January 2011.

[14] Y. D. Zhang, M. G. Amin, and B. Himed, “Sparsity-based DOA estimation using co-prime arrays,” in Proceedings of the 2013 IEEE International Conference on Acoustics, pp. 3967–3971, Vancouver, Canada, May, 2013.

[15] Q. Shen, W. Liu, W. Cui, and S. Wu, “Underdetermined DOA estimation under the compressive sensing framework: a review,” IEEE Access, vol. 4, 2016.

[16] Q. Shen, W. Cui, W. Liu, S. Wu, Y. D. Zhang, and M. G. Amin, “Underdetermined wideband DOA estimation of off-grid sources employing the difference co-array concept,” Signal Processing, vol. 130, pp. 299–304, 2017.

[17] S. Lie, A. R. Leyman, and Y. H. Chew, “Fourth-order and weighted mixed order direction-of-arrival estimators,” IEEE Signal Processing Letters, vol. 13, no. 11, pp. 691–694, 2006.

[18] H. Wan and B. Liao, “Fourth-order direction finding in antenna arrays with partial channel gain/phase calibration,” IEEE Journal of Oceanic Engineering, vol. 46, pp. 1–115, 2020.

[19] M. C. Dogan and J. M. Mendel, “Applications of cumulants to array processing I: Aperture extension and array calibration,” IEEE Transactions on Signal Processing, vol. 43, no. 5, pp. 1200–1216, 1995.
[20] P. Chevalier and A. Ferreol, “On the virtual array concept for the fourth-order direction finding problem,” IEEE Transactions on Signal Processing, vol. 47, no. 9, pp. 2592–2595, 1999.

[21] P. Pal and P. P. Vaidyanathan, “Multiple level nested array: an efficient geometry for $2q$th order cumulant based array processing,” IEEE Transactions on Signal Processing, vol. 60, no. 3, pp. 1253–1269, 2012.

[22] P. Chevalier, A. Ferreol, and L. Albera, “High-resolution direction finding from higher order statistics: the $2rm q$-MUSIC algorithm,” IEEE Transactions on Signal Processing, vol. 54, no. 8, pp. 2986–2997, 2006.

[23] G. Birot, L. Albera, and P. Chevalier, “Sequential high-resolution direction finding from higher order statistics,” IEEE Transactions on Signal Processing, vol. 58, no. 8, pp. 4144–4155, 2010.

[24] Q. Shen, W. Liu, W. Cui, and S. Wu, “Extension of nested arrays with the fourth-order difference co-array enhancement,” in Proceedings of the 2016 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 2991–2995, Shanghai, China, March 2016.

[25] Q. Shen, W. Liu, W. Cui, and S. Wu, “Extension of co-prime arrays based on the fourth-order difference co-array concept,” IEEE Signal Processing Letters, vol. 23, no. 5, pp. 615–619, 2016.

[26] J. Cai, W. Liu, R. Zong, and Q. Shen, “An expanding and shift scheme for constructing fourth-order difference coarrays,” IEEE Signal Processing Letters, vol. 24, no. 4, pp. 480–484, 2017.

[27] A. Ahmed, Y. D. Zhang, and B. Himed, “Effective nested array design for fourth-order cumulant-based DOA estimation,” in Proceedings of the 2017 IEEE Radar Conference (RadarConf), pp. 0998–1002, Seattle, WA, USA, May 2017.

[28] Q. Shen, W. Liu, W. Cui, S. Wu, and P. Pal, “Simplified and enhanced multiple level nested arrays exploiting high-order difference co-arrays,” IEEE Transactions on Signal Processing, vol. 67, no. 13, pp. 3502–3515, 2019.

[29] P. Stoica and A. Nehorai, “MUSIC maximum likelihood and Cramer-Rao bound,” IEEE Transactions on Acoustics Speech and Signal Processing, vol. 37, no. 5, pp. 720–741, 1989.

[30] C.-L. Liu and P. P. Vaidyanathan, “Cramér-Rao bounds for coprime and other sparse arrays which find more sources than sensors,” Digital Signal Processing, vol. 61, pp. 43–61, 2017.