

DOA Estimation Using Compressed Sparse Array

Muran Guo, *Student Member, IEEE*, Yimin D. Zhang, *Senior Member, IEEE*, and Tao Chen

Abstract—Sparse arrays, such as nested arrays and coprime arrays, can achieve a high number of degrees of freedom (DOFs) for direction of arrival (DOA) estimation with a reduced number of antennas. On the other hand, the compressive measurement method provides an effective way to reduce the number of front-end circuit chains. In this paper, we generalized current works on the two categories of methods to a compressed sparse array (CSA) scheme which combines the compressive measurement method and the sparse array together to significantly reduce the system complexity. After introducing the proposed scheme, the Cramér-Rao bound (CRB) of the proposed CSA scheme is derived. We then determine the corresponding existing conditions of the CRB, based on which the number of DOFs is derived and examined for the first time. It is proved that, for a CSA which compressed the output of an L -element sparse array to $M < L$ chains, a higher number of DOFs is obtained as compared to that of the M -element array with the same sparse structure. Furthermore, the DOA estimation accuracy using the M -chain CSA is higher than that using the M -element sparse array due to the extended array aperture. Numerical simulations verify the superiority of the proposed CSA scheme.

Index Terms—Compressive sensing, Cramér-Rao bound, DOA estimation, information theory, sparse array.

I. INTRODUCTION

DIRECTION of arrival (DOA) estimation is an important field in array signal processing, and uniform linear array (ULA) is a popular array structure for this purpose [1], [2]. It is well known that, in order to achieve a high DOA estimation accuracy, a large ULA aperture is required. In this case, the number of the array elements and, subsequently, the number of front-end circuit chains that are connected with the antennas, also increase, thus leading to a high system hardware complexity and cost. Two categories of methods have been proposed to overcome the drawbacks involved in ULA, namely, the sparse array structure [3], [4], [5], [6], [7] and the compression matrix method [8], [9], [10].

Sparse array structures, such as nested arrays [3] and coprime arrays [4], [5], [6], are attractive because they provide effective means to estimate the direction of $\mathcal{O}(L^2)$ sources with only $\mathcal{O}(L)$ antennas. The auto-correlation information of the

received signal vector is treated as a virtual received signal on the difference coarray in sparse array-based DOA estimation methods. In this case, DOA estimation based on sparse arrays is similar to the estimation of DOAs of coherent sources, which yields a rank-deficient covariance matrix. As such, for subspace-based DOA estimation algorithms, such as multiple signal classification (MUSIC) [11] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [12], it is essential to exploit the spatial smoothing technique in order to recover the rank of the covariance matrix of the coarray [13], [14]. Furthermore, for some sparse arrays, such as coprime arrays, there exist holes in the coarray. Because only the consecutive lags in the coarray can be exploited to perform spatial smoothing, the number of degrees of freedom (DOFs) is decreased due to the existence of these holes. To address this drawback, several coarray interpolation methods have been proposed [15], [16], [17]. Another simple and effective way to estimate DOAs is using the algorithms based on compressive sensing (CS) [18]. CS-based DOA estimation algorithms can utilize the full DOFs offered by all unique auto-correlation lags without performing spatial smoothing or coarray interpolation.

The compression matrix method, first proposed in [8], [9], is a more general approach to reduce the system complexity. It introduces a complex compression matrix $\Phi \in \mathbb{C}^{M \times L}$ at the antenna outputs, where M is the number of circuit chains and L is the number of antennas with $M < L$. For an arbitrary compression matrix Φ , the multiplication of Φ with the received signal vector is essentially a linear combination which effectively reduces the dimension of the signal vector from L to M . The compression matrix Φ is commonly designed by choosing its entries from independent and identically distributed (i.i.d.) random entries. In addition, Φ is a time-invariant matrix, which means that Φ keeps unchanged during an estimation period. In this case, however, the compression operation results in a loss in the Fisher information matrix (FIM). When a ULA is used as the receive antenna array, such loss can be characterized by a factor of $(L - M)/L$ [19]. To avoid the information loss, optimized design for the compression matrix Φ has been proposed [10], [20], [21], where a ULA is used as the receive antenna array. In this case, to obtain a high estimation accuracy, a massive number of physical antennas is still need. As such, the complexity and cost of the antenna part remain high, although the number of front-end circuit chains is effectively reduced. To solve the problem, the coprime array was exploited as the receive antenna array in [22]. However, reference [22] only proposed the basic configuration of using the coprime array as the receive array. Some strict assumptions were made and the corresponding performance was not analyzed. In addition, the compression matrix Φ is randomly drawn from

The work of M. Guo was supported by the China Scholarship Council (CSC) for his stay at the Temple University. The work of Y. D. Zhang was supported in part by the National Science Foundation (NSF) under grant AST-1547420. The work of T. Chen was supported by the National Natural Science Foundation of China (No. 61571146) and the Fundamental Research Funds for the Central Universities (HEUCFP201769).

M. Guo is with the College of Information and Communication Engineering, Harbin Engineering University, Harbin, 150001, China, and is also with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA (e-mail: guomuran@hrbeu.edu.cn).

Y. D. Zhang is with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA (e-mail: ydzhang@temple.edu).

T. Chen is with the College of Information and Communication Engineering, Harbin Engineering University, Harbin, 150001, China (e-mail: chentao@hrbeu.edu.cn).

a Gaussian distribution, thus resulting in a degradation on the estimation accuracy. Moreover, the switched-element system proposed in [23] introduced a similar structure but with a time-variant binary selection matrix to reduce the number of circuit chains. There are two key differences between [23] and the compression method. First, the selection matrix in [23] consists of binary integers (1 and 0) to switch on and off the antennas, whereas the compression matrix used in our method generally has full complex-valued entries. Second, the selection matrix is time-variant based on the different subarray structures, whereas the compression matrix keeps unchanged during each estimation period.

In this paper, we develop a general compressive measurement scheme, referred to as the compressed sparse array (CSA), which combines the compression matrix and the sparse array together. The CSA takes an arbitrary sparse array as the receive antenna array, and then exploits a combining network to compress the dimension of the received signal vector from L to M where $M < L$. Hence, the numbers of both the physical antennas and the front-end circuit chains are significantly reduced. We also propose an improved CS-based DOA estimation approach by which, unlike the approach proposed in [22], the knowledge on the noise power is no longer required. In addition, Φ is optimized in this paper by maximizing the Shannon mutual information between the compressed measurements and the DOAs when the *a priori* probability distribution of the DOAs is available. The Cramér-Rao bound (CRB) for DOA estimation using CSA and the corresponding existing conditions are derived to analyze the performance and determine the number of DOFs acquired by the CSA. An interesting result is that, for a given number of circuit chains, the proposed CSA scheme involves a higher number of receive antennas than the classical sparse array structure. As a result, the proposed CSA scheme can obtain a higher number of DOFs and a better DOA estimation accuracy than the classical sparse array structure.

To clearly show the contributions of this paper, the difference between [22] and this paper is highlighted as follows.

- a) The most important difference is that the performance of CSA in terms of the CRB and the corresponding existence conditions are analyzed in this paper. Based on the CRB existence conditions, the number of the achieved DOFs is analyzed for the first time.
- b) In [22], only the coprime array is considered, while in this paper we consider a general class of sparse linear array structures. Furthermore, we compare the performance when different sparse array configurations are used.
- c) The compression matrix Φ is randomly selected from Gaussian distribution in [22], thus causing a relatively high information loss in the compression operation. In this paper, Φ is optimized based on the *a priori* probability distribution, thereby improving the DOA estimation accuracy.
- d) The noise power is assumed to be known for the DOA estimation approach proposed in [22], whereas in this paper, this assumption is not necessary for the proposed DOA estimation approach.
- e) In [22], the compression operation is assumed to be noise-free, which is a strict assumption. In practice, the com-

pression operation is realized by introducing a combining network which is often implemented in analog circuits consisting of phase shifters and summators. Additional noise will occur when the signals go through the combining network. Thus, the model proposed in this paper is more general.

This paper is organized as follows. In Section II, we introduce the system model of the CSA and review a high-resolution DOA estimation approach. The proposed CS-based DOA estimation approach and the optimization of the compression matrix Φ are described in Section III. Some preliminaries for the CRB are reviewed in Section IV. Then, in Section V, we derive the expression of the CRB for the normalized spatial frequencies based on the proposed CSA scheme and the corresponding conditions under which the CRB exists, and the number of DOFs is analyzed based on the existence conditions. Simulation results are presented in Section VI to examine the effectiveness. Section VII concludes this paper.

Notations: we use the lower-case letter (e.g., a), lower-case bold letter (e.g., \mathbf{a}), and upper-case bold letter (e.g., \mathbf{A}) to represent the scalars, vectors and matrices, respectively. The superscripts $*$, T and H denote the complex conjugate, the transpose and the complex conjugate transpose, respectively. In addition, $\text{vec}(\cdot)$ and $\mathbb{E}(\cdot)$ are used to represent the vectorization and expectation operations. The diagonal matrix whose diagonal entries are given in \mathbf{a} is expressed by $\text{diag}(\mathbf{a})$. Moreover, $\text{tr}(\mathbf{A})$ means the trace of matrix \mathbf{A} . $j = \sqrt{-1}$ is the unit imaginary, and \mathbf{I}_L is the $L \times L$ identity matrix. \otimes and \circ are used to represent the Kronecker product and Khatri-Rao product (column-wise Kronecker product). For instance, the Khatri-Rao product between two matrices $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_c]$ and $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_c]$ is given as $\mathbf{A} \circ \mathbf{B} = [\mathbf{a}_1 \otimes \mathbf{b}_1, \dots, \mathbf{a}_c \otimes \mathbf{b}_c]$. $[a]$ means the maximum integer that is lower than or equal to a . We use the triangle bracket $\langle \mathbf{x}_{\mathbb{S}} \rangle_n$ to represent the value corresponding to the support $n \in \mathbb{S}$. For example, let $\mathbf{x}_{\mathbb{S}} = \{2, 3, 4\}$ and $\mathbb{S} = \{-1, 0, 1\}$. Then, we have $\langle \mathbf{x}_{\mathbb{S}} \rangle_{-1} = 2$, $\langle \mathbf{x}_{\mathbb{S}} \rangle_0 = 3$ and $\langle \mathbf{x}_{\mathbb{S}} \rangle_1 = 4$.

II. THE PROPOSED CSA SCHEME

The proposed CSA scheme is introduced in this section. First, the system model of the CSA is set up. Then, based on this system model, we review a high resolution DOA estimation approach, named compressive sensing minimum variance distortionless response (CS-MVDR), which has been proposed in [22].

A. System Model

Consider Q uncorrelated far-field narrowband sources with directions $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_Q]^T$ impinging on a sparse array consisting of L omni-directional antennas with locations $\{d_1 d_0, d_2 d_0, \dots, d_L d_0\}$, where $d_0 = \lambda/2$ is the unit inter-element spacing with λ denoting the wavelength of the sources. In addition, $\mathbb{S} = \{d_1, d_2, \dots, d_L\}$ is an integer set indicating the physical antenna positions in terms of d_0 .

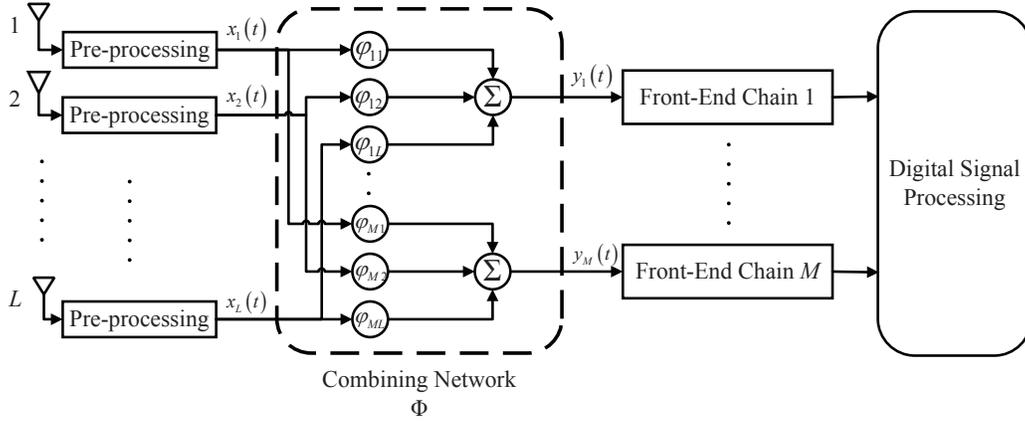


Fig. 1. System model.

Without loss of generality, we set $d_1 = 0$. Then, the received signal vector $\mathbf{x}_S(t) \in \mathbb{C}^L$ is expressed as

$$\begin{aligned} \mathbf{x}_S(t) &= [\mathbf{a}_S(\omega_1), \dots, \mathbf{a}_S(\omega_Q)]\mathbf{s}(t) + \mathbf{v}(t) \\ &= \mathbf{A}_S\mathbf{s}(t) + \mathbf{v}(t), \end{aligned} \quad (1)$$

where $\mathbf{a}_S(\omega_q) = [e^{j2\pi d_1 \omega_q}, \dots, e^{j2\pi d_L \omega_q}]^T$ is the steering vector of the q -th source and $\omega_q = (d_0/\lambda) \sin \theta_q$ is the spatial frequency of the q -th source with $q = 1, 2, \dots, Q$. \mathbf{A}_S is referred to as the manifold matrix of \mathbb{S} , and $\mathbf{s}(t)$ is the complex amplitude vector of the sources. In addition, t is the time index and $\mathbf{v}(t) \in \mathbb{C}^L$, referred as the signal noise, is the white Gaussian noise vector involved in the received signals. It is noted that $\mathbf{v}(t)$ is assumed to be uncorrelated at different antennas, and the yielding covariance matrix of $\mathbf{v}(t)$ is thus diagonal.

Define the difference coarray corresponding to \mathbb{S} as

$$\mathbb{D} = \{\mu - \nu | \mu, \nu \in \mathbb{S}\}. \quad (2)$$

Then, the corresponding manifold matrix of the difference coarray is

$$\mathbf{A}_D = [\mathbf{a}_D(\omega_1), \dots, \mathbf{a}_D(\omega_Q)]. \quad (3)$$

Denote $\{d_1^{(D)}, \dots, d_{|\mathbb{D}|}^{(D)}\}$ as the lags in coarray \mathbb{D} in the ascent order, where $|\mathbb{D}|$ represents the cardinality of \mathbb{D} . Then, $\mathbf{a}_D(\omega_q) = [e^{j2\pi d_1^{(D)} \omega_q}, \dots, e^{j2\pi d_{|\mathbb{D}|}^{(D)} \omega_q}]^T$ is the steering vector of the coarray associated with the q -th source. As demonstrated in [24], the relationship between \mathbf{A}_S and \mathbf{A}_D is expressed as

$$\mathbf{A}_S^* \circ \mathbf{A}_S = \mathbf{J}\mathbf{A}_D, \quad (4)$$

where \mathbf{J} is the binary matrix as defined in Definition 1.

Definition 1. [24] Denote \mathbb{B} as the binary set which consists of 0 and 1. The i -th column of $\mathbf{J} \in \mathbb{B}^{|\mathbb{S}|^2 \times |\mathbb{D}|}$ is defined as

$$\langle \mathbf{J} \rangle_{:,i} = \text{vec}(\mathbf{I}(i)), \quad i \in \mathbb{D},$$

where $\mathbf{I}(i)$ is given by

$$\langle \mathbf{I}(i) \rangle_{n_1, n_2} = \begin{cases} 1 & \text{if } n_1 - n_2 = i, \\ 0 & \text{otherwise,} \end{cases}$$

where $n_1, n_2 \in \mathbb{S}$.

As proposed in [8], [9], a combining network is inserted after the antenna outputs to decrease the number of required front-end circuit chains, thus reducing the hardware complexity and cost. The system model is depicted in Fig. 1. While it is similar to that in [10], it is noted that, reference [10] assumed a ULA, whereas the receive antenna array in Fig. 1 is a sparse array. The function of the combining network can be described as a compression matrix $\Phi \in \mathbb{C}^{M \times L}$. Then, the output signal vector of the combining network, denoted as $\mathbf{y}(t) \in \mathbb{C}^M$, is expressed as

$$\begin{aligned} \mathbf{y}(t) &= \Phi \mathbf{x}_S(t) + \mathbf{w}(t) \\ &= \Phi \mathbf{A}_S \mathbf{s}(t) + \mathbf{n}(t), \end{aligned} \quad (5)$$

where $\mathbf{w}(t) \in \mathbb{C}^M$ is the white Gaussian measurement noise vector induced by the combining network, and $\mathbf{n}(t) = \Phi \mathbf{v}(t) + \mathbf{w}(t)$ is the total noise.

Assume that the sources follow a stochastic model [25] and are uncorrelated to each other. In addition, $\mathbf{v}(t)$ and $\mathbf{w}(t)$ are assumed to be mutually independent white Gaussian noise, which are also independent of the sources, with mean zero and covariance matrices $p_v \mathbf{I}_L$ and $p_w \mathbf{I}_M$, respectively. Then, we can express the covariance matrix of $\mathbf{y}(t)$ as

$$\begin{aligned} \mathbf{R}_{\mathbf{y}\mathbf{y}} &= \mathbb{E}[\mathbf{y}(t)\mathbf{y}^H(t)] \\ &= \Phi \mathbf{A}_S \mathbf{R}_{ss} \mathbf{A}_S^H \Phi^H + p_v \Phi \Phi^H + p_w \mathbf{I}_M, \end{aligned} \quad (6)$$

where $\mathbf{R}_{ss} = \text{diag}([p_1, p_2, \dots, p_Q])$ is the covariance matrix of the sources with p_q denoting the power of the q -th source.

In practice, $\mathbf{R}_{\mathbf{y}\mathbf{y}}$ is estimated using a finite number of snapshots. In this case, the sample covariance matrix $\tilde{\mathbf{R}}_{\mathbf{y}\mathbf{y}}$ is expressed as

$$\tilde{\mathbf{R}}_{\mathbf{y}\mathbf{y}} = \frac{1}{T} \sum_{t=1}^T \mathbf{y}(t)\mathbf{y}^H(t), \quad (7)$$

where T is the number of snapshots.

B. High Resolution DOA Estimation Approach

The CS-MVDR beamformer was proposed in [21] to achieve high-resolution DOA estimation with a reduced number of front-end circuit chains. In this section, we briefly

introduce the CS-MVDR algorithm. It is noted that coarrays are not used in CS-MVDR. More specifically, the CS-MVDR algorithm estimates the DOAs by directly applying the covariance matrix of received signal vector. Thus, CS-MVDR does not take the advantages of coarrays to increase the number of DOFs.

Given T samples of the compressed received signal, we first compute the corresponding covariance matrix $\tilde{\mathbf{R}}_{\mathbf{y}\mathbf{y}}$ using (7). Then, the source DOAs can be estimated by searching the peaks of the following spatial spectrum [21]:

$$P_{\text{CS-MVDR}}(\omega) = \frac{1}{L} \frac{(\Phi \mathbf{a}_S(\omega))^H (\Phi \mathbf{a}_S(\omega))}{(\Phi \mathbf{a}_S(\omega))^H \tilde{\mathbf{R}}_{\mathbf{y}\mathbf{y}}^{-1} (\Phi \mathbf{a}_S(\omega))}. \quad (8)$$

The values of ω corresponding to the first Q peaks of $P_{\text{CS-MVDR}}(\omega)$ indicate the estimated spatial frequencies, which are uniquely associated with the DOAs.

When compression is not performed, i.e., Φ is an identity matrix, (8) is simplified as

$$P_{\text{CS-MVDR}}(\omega) = \frac{1}{\mathbf{a}_S^H(\omega) \tilde{\mathbf{R}}_{\mathbf{y}\mathbf{y}}^{-1} \mathbf{a}_S(\omega)}, \quad (9)$$

where the fact that $\mathbf{a}_S^H(\omega) \mathbf{a}_S(\omega) = L$ is utilized. In this case, the CS-MVDR degenerates to the standard MVDR algorithm.

III. IMPROVED APPROACHES FOR DOA ESTIMATION

In this section, we first propose the CS-based DOA estimation approach with improved DOF, where the assumption on the noise power is no longer required. Then, the optimization procedure of the compression matrix Φ for the CSA scheme is proposed, where the maximum mutual information criterion is utilized.

A. DOA Estimation Approach with Improved DOF

The approach proposed in this subsection is based on the coarray approach, and the CS method is used to resolve more sources than the number of antennas. Vectorizing (6), we obtain

$$\begin{aligned} \mathbf{r}_{\mathbf{y}\mathbf{y}} &= \text{vec}(\mathbf{R}_{\mathbf{y}\mathbf{y}}) \\ &= (\Phi^* \otimes \Phi)(\mathbf{A}_S^* \circ \mathbf{A}_S) \mathbf{p} \\ &\quad + p_v (\Phi^* \otimes \Phi) \text{vec}(\mathbf{I}_L) + p_w \text{vec}(\mathbf{I}_M), \end{aligned} \quad (10)$$

where the property $\text{vec}(\mathbf{A}\mathbf{X}\mathbf{B}) = (\mathbf{A}^T \otimes \mathbf{B})\text{vec}(\mathbf{X})$ is utilized [26]. Denote $\Psi = [(\Phi^* \otimes \Phi)(\mathbf{A}_S^* \circ \mathbf{A}_S), (\Phi^* \otimes \Phi)\text{vec}(\mathbf{I}_L), \text{vec}(\mathbf{I}_M)]$ and $\mathbf{b} = [\mathbf{p}^T, p_v, p_w]^T$, where $\mathbf{p} = [p_1, p_2, \dots, p_Q]^T$ denotes the source power vector. Then, (10) can be rewritten as

$$\mathbf{r}_{\mathbf{y}\mathbf{y}} = \Psi \mathbf{b}. \quad (11)$$

By dividing the entire spatial domain into a grid denoted by $\omega^{(g)} = [\omega_1^{(g)}, \dots, \omega_G^{(g)}]^T$, where $G \gg Q$, we can obtain a sparse vector \mathbf{b}° consisting of G entries, in which only the Q entries corresponding to the true DOAs are non-zero. Thus, the corresponding sensing matrix $\Psi^{(g)}$ and the sparse vector \mathbf{b}° are given by $\Psi^{(g)} = [(\Phi^* \otimes \Phi)((\mathbf{A}_S^{(g)})^* \circ \mathbf{A}_S^{(g)}), (\Phi^* \otimes \Phi)\text{vec}(\mathbf{I}_L), \text{vec}(\mathbf{I}_M)]$ and $\mathbf{b}^\circ = [\mathbf{p}^{\circ T}, p_v, p_w]^T$, respectively, where $\mathbf{A}_S^{(g)} = [\mathbf{a}_S(\omega_1^{(g)}), \dots, \mathbf{a}_S(\omega_G^{(g)})]$ is the manifold matrix

with respect to $\omega^{(g)}$. Thus, \mathbf{b}° can be recovered by solving the following constrained minimization problem

$$\tilde{\mathbf{b}}^\circ = \arg \min_{\mathbf{b}^\circ} \|\mathbf{b}^\circ\|_0 \quad \text{s.t.} \quad \|\mathbf{r}_{\mathbf{y}\mathbf{y}} - \Psi^{(g)} \mathbf{b}^\circ\|_2 < \epsilon, \quad (12)$$

where $\|\mathbf{b}^\circ\|_0$ is the l_0 norm of \mathbf{b}° , namely, the number of non-zero entries in \mathbf{b}° , and ϵ is the user-defined parameter which is determined mainly based on the error between the sample covariance matrix and real covariance matrix. Eq. (12) is a typical CS problem and many existing CS algorithms can be applied to obtain the optimal $\tilde{\mathbf{b}}^\circ$. In this paper, we use the least absolute shrinkage and selection operator (LASSO) algorithm [27], whose objective function is given by

$$\tilde{\mathbf{b}}^\circ = \arg \min_{\mathbf{b}^\circ} \left[\frac{1}{2} \|\mathbf{r}_{\mathbf{y}\mathbf{y}} - \Psi^{(g)} \mathbf{b}^\circ\|_2 + \eta \|\mathbf{b}^\circ\|_1 \right], \quad (13)$$

where $\|\cdot\|_1$ and $\|\cdot\|_2$ represent the l_1 and l_2 norms, respectively, and η is the regularization parameter to balance the error and the sparsity. Hence, the DOAs can be estimated by exploiting the support of the first G entries in the obtained optimal solution $\tilde{\mathbf{b}}^\circ$. Different from the approach proposed in [22], the noise powers are combined with the source powers here. Hence, the noise powers are not required to be known in this approach.

B. Optimization of Compression Matrix

In [22], each entry of the compression matrix Φ is drawn from an i.i.d. random Gaussian distribution $\mathcal{CN}(0, 1/(M + L - 1))$. As pointed out in [19], randomly selecting Φ results in information loss. In many scenarios, on the other hand, the prior knowledge of the DOA distribution is known, thus making it possible to optimize Φ to obtain a more accurate estimation. Hereby, we optimize the compression matrix Φ based on the maximum mutual information criterion [20], [21] for the proposed CSA scheme. For simplicity, only the stochastic signal model, where the source signals are assumed to follow Gaussian distribution, is considered in the following optimization. However, as demonstrated in [20], the compression matrix Φ under both deterministic and stochastic signal models can be optimized based on the maximum mutual information criterion.

Note that generally the prior distribution is with respect to the DOA θ rather than to $\omega = (d_0/\lambda) \sin \theta$. Thus, we use θ to optimize the compression matrix in this subsection. Denote $f(\theta)$ as the probability density function (pdf) of θ . Discretize the spatial domain into K equally divided bins with a width of $\Delta\bar{\theta}$, and let $p_k^\circ = f(\bar{\theta}_k) \Delta\bar{\theta}$, where $\sum_{k \in \mathcal{K}} p_k^\circ = 1$ and $\mathcal{K} = \{1, 2, \dots, K\}$. In addition, $\bar{\theta}_k$ is the nominal angle corresponding to the k -th bin. Then, the pdf of the compressed measurement vector is approximated as [21]

$$f(\mathbf{y}) \approx \sum_{k \in \mathcal{K}} p_k^\circ f(\mathbf{y}|\bar{\theta}_k). \quad (14)$$

For the k -th angular bin at $\bar{\theta}_k$, the compressed measurement vector is

$$\mathbf{y}(t)|_{\theta=\bar{\theta}_k} = \Phi[\mathbf{a}_S(\bar{\theta}_k)s(t) + \mathbf{v}(t)] + \mathbf{w}(t). \quad (15)$$

Since the source signal $s(t)$ follows the Gaussian distribution $\mathcal{CN}(0, p_s)$ whereas $\mathbf{v}(t)$ and $\mathbf{w}(t)$ are white Gaussian, $\mathbf{y}(t)|_{\theta=\bar{\theta}_k}$ also follows the Gaussian distribution with $\mathcal{CN}(\mathbf{0}, \mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k)$, where $\mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k$ is expressed as

$$\mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k = \Phi(p_s \mathbf{a}_S(\bar{\theta}_k) \mathbf{a}_S^H(\bar{\theta}_k) + p_v \mathbf{I}_L) \Phi^H + p_w \mathbf{I}_M. \quad (16)$$

The gradient of the Shannon mutual information $I(\mathbf{y}; \theta)$ between the compressed measurement vector $\mathbf{y}(t)$ and the DOA θ with respect to the compression matrix Φ is defined as

$$\nabla_{\Phi} I(\mathbf{y}; \theta) = \nabla_{\Phi} h(\mathbf{y}) - \nabla_{\Phi} h(\mathbf{y}|\theta), \quad (17)$$

where $\nabla_{\Phi}\{\cdot\}$ represents the gradient with respect to Φ . In addition, $h(\mathbf{y}) = -\mathbb{E}\{\log[f(\mathbf{y})]\}$ and $h(\mathbf{y}|\theta) = -\mathbb{E}\{\log[f(\mathbf{y}|\theta)]\}$ represent the differential entropy and the conditional differential entropy, respectively, of the compressed measurement vector $\mathbf{y}(t)$.

Following the derivations in [21] and using (16), we obtain the gradient $\nabla_{\Phi} I(\mathbf{y}; \theta)$ as shown in (18), where $|\cdot|$ means the determinant of the matrix.

Once the gradient is obtained, we search the optimal compression matrix Φ by iteratively updating the following formulation,

$$\Phi^{(\xi+1)} = \Phi^{(\xi)} + \kappa \nabla_{\Phi^{(\xi)}} I(\mathbf{y}; \theta), \quad (19)$$

where ξ represents the ξ -th iteration and $\kappa > 0$ is the step size. It is noted that, as a gradient-based algorithm, the above iteration may fall into local optima. This problem may be effectively avoided by using several initial values and then selecting the best result.

IV. PRELIMINARIES ON THE CRB

The CRB for the parameters to be estimated is of great importance in evaluating the performance of a system in general, since the covariance of any unbiased estimator under the same system model is lower bounded by the CRB. For a random vector \mathbf{x} with an arbitrary pdf $p(\mathbf{x}, \alpha)$, where α is a real-valued parameter vector, the FIM is defined as [28]

$$\text{FIM}(\alpha) = -\mathbb{E} \left\{ \frac{\partial^2 \ln p(\mathbf{x}, \alpha)}{\partial \alpha \partial \alpha^T} \right\}. \quad (20)$$

The FIM is guaranteed to be positive semidefinite [29]. If the FIM is positive definite, i.e., FIM is invertible, the CRB is defined as the inverse matrix of FIM:

$$\text{CRB}(\alpha) = \text{FIM}^{-1}(\alpha). \quad (21)$$

Furthermore, for a random vector \mathbf{x} that follows a complex Gaussian distribution with mean zero and covariance matrix \mathbf{C} , the (m, n) -th entry of the FIM can be expressed as [28]

$$[\text{FIM}(\alpha)]_{m,n} = \text{tr} \left(\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \alpha_m} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \alpha_n} \right). \quad (22)$$

From the system model, we define the following real-valued parameter vector,

$$\alpha = [\omega^T \quad \rho^T \quad p_v \quad p_w]^T, \quad (23)$$

where $\omega = [\omega_1, \dots, \omega_Q]^T$ and $\rho = [\rho_1, \dots, \rho_Q]^T$. Under the previous assumptions, each snapshot of the signal vector $\mathbf{y}(t)$ follows the Gaussian distribution with mean zero and covariance matrix $\mathbf{R}_{\mathbf{y}\mathbf{y}}$. By stacking the snapshots on top of the other, we have the following complex Gaussian distribution

$$[\mathbf{y}(1)^T, \mathbf{y}(2)^T, \dots, \mathbf{y}(T)^T]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_T \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}}). \quad (24)$$

Substituting (24) into (22), the (m, n) -th entry of the FIM exploiting the T i.i.d. measurement vectors is expressed as

$$[\text{FIM}(\alpha)]_{m,n} = T \text{tr} \left(\mathbf{R}_{\mathbf{y}\mathbf{y}}^{-1} \frac{\partial \mathbf{R}_{\mathbf{y}\mathbf{y}}}{\partial \alpha_m} \mathbf{R}_{\mathbf{y}\mathbf{y}}^{-1} \frac{\partial \mathbf{R}_{\mathbf{y}\mathbf{y}}}{\partial \alpha_n} \right). \quad (25)$$

Given the following properties [26]

$$\begin{aligned} \text{tr}(\mathbf{X}\mathbf{Y}) &= (\text{vec}(\mathbf{X})^H)^H \text{vec}(\mathbf{Y}), \\ \text{vec}(\mathbf{X}\mathbf{Y}\mathbf{Z}) &= (\mathbf{Z}^T \otimes \mathbf{X}) \text{vec}(\mathbf{Y}), \end{aligned} \quad (26)$$

(25) can be further simplified as

$$\begin{aligned} [\text{FIM}(\alpha)]_{m,n} &= T \left[\text{vec} \left(\frac{\partial \mathbf{R}_{\mathbf{y}\mathbf{y}}}{\partial \alpha_m} \right) \right]^H (\mathbf{R}_{\mathbf{y}\mathbf{y}}^{-T} \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}}^{-1}) \text{vec} \left(\frac{\partial \mathbf{R}_{\mathbf{y}\mathbf{y}}}{\partial \alpha_n} \right) \\ &= T \left[(\mathbf{R}_{\mathbf{y}\mathbf{y}}^T \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}})^{-\frac{1}{2}} \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial \alpha_m} \right]^H \left[(\mathbf{R}_{\mathbf{y}\mathbf{y}}^T \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}})^{-\frac{1}{2}} \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial \alpha_n} \right], \end{aligned} \quad (27)$$

where $\mathbf{r}_{\mathbf{y}\mathbf{y}}$ is defined in (10). Since the DOAs are the parameters of interest, the parameter vector α can be divided into $[\omega^T \quad \rho^T \quad p_v \quad p_w]^T$. Thus, (27) is rewritten as

$$\text{FIM}(\alpha) = T \begin{bmatrix} \mathbf{G}^H \\ \Delta^H \end{bmatrix} \begin{bmatrix} \mathbf{G} & \Delta \end{bmatrix}, \quad (28)$$

where \mathbf{G} and Δ are defined as

$$\mathbf{G} = (\mathbf{R}_{\mathbf{y}\mathbf{y}}^T \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}})^{-\frac{1}{2}} \begin{bmatrix} \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial \omega_1}, \dots, \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial \omega_Q} \end{bmatrix}, \quad (29)$$

$$\nabla_{\Phi} I(\mathbf{y}; \theta) \approx \frac{\sum_{k \in \mathcal{K}} p_k^{\diamond} \left| \frac{\mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k}{p_v} \right|^{-1} \left[\frac{\mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k}{p_v} \right]^{-1} \Phi \left(\frac{p_s}{p_v} \mathbf{a}_S(\bar{\theta}_k) \mathbf{a}_S^H(\bar{\theta}_k) + \mathbf{I}_L \right)}{\sum_{k \in \mathcal{K}} p_k^{\diamond} \left| \frac{\mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k}{p_v} \right|^{-1}} - \sum_{k \in \mathcal{K}} p_k^{\diamond} \left[\frac{\mathbf{R}_{\mathbf{y}\mathbf{y}}|\bar{\theta}_k}{p_v} \right]^{-1} \Phi \left(\frac{p_s}{p_v} \mathbf{a}_S(\bar{\theta}_k) \mathbf{a}_S^H(\bar{\theta}_k) + \mathbf{I}_L \right), \quad (18)$$

$$\mathbf{\Delta} = (\mathbf{R}_{\mathbf{y}\mathbf{y}}^T \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}})^{-\frac{1}{2}} \begin{bmatrix} \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial p_1}, \dots, \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial p_Q}, \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial p_v}, \frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial p_w} \end{bmatrix}. \quad (30)$$

If the FIM is invertible, the CRB for $\boldsymbol{\omega} = [\omega_1, \dots, \omega_Q]^T$ can be given as the inverse of the Schur complement of the block $\mathbf{\Delta}^H \mathbf{\Delta}$ of the FIM [30]

$$\text{CRB}(\boldsymbol{\omega}) = \frac{1}{T} (\mathbf{G}^H \mathbf{\Pi}_{\mathbf{\Delta}}^{\perp} \mathbf{G})^{-1}, \quad (31)$$

where $\mathbf{\Pi}_{\mathbf{\Delta}}^{\perp} = \mathbf{I} - \mathbf{\Delta}(\mathbf{\Delta}^H \mathbf{\Delta})^{-1} \mathbf{\Delta}^H$.

V. PERFORMANCE ANALYSIS OF THE PROPOSED CSA SCHEME

In this section, we first give the existence conditions of the CRB based on the proposed CSA scheme and analyze the number of DOFs of the proposed CSA. Then, the improvement of the CSA on the number of DOFs is intuitively demonstrated by using different sparse arrays as the receive array. Finally, the CRB expression is derived.

A. Analysis on the Number of DOFs

From (31), we observe that the existence of the CRB, i.e., the nonsingularity of the FIM, is equivalent to the nonsingularity of $\mathbf{\Delta}^H \mathbf{\Delta}$ and $\mathbf{G}^H \mathbf{\Pi}_{\mathbf{\Delta}}^{\perp} \mathbf{G}$ (Lemma 1 in [24]). To obtain the condition under which the CRB exists, we propose the following two lemmas. Note that \mathbf{J} is a binary matrix defined in Definition 1.

Lemma 1. *Let $\mathbf{V}_0 = \mathbf{\Phi}^* \otimes \mathbf{\Phi}$ and $\mathbf{W}_{\mathbb{D}} = [\mathbf{V}_0 \mathbf{J} \mathbf{A}_{\mathbb{D}}, \mathbf{V}_0 \mathbf{J} \mathbf{e}_0, \text{vec}(\mathbf{I}_M)]$ with \mathbf{e}_0 satisfying $\text{vec}(\mathbf{I}_L) = \mathbf{J} \mathbf{e}_0$. If $\text{rank}(\mathbf{W}_{\mathbb{D}}) = Q + 2$, then $\mathbf{\Delta}^H \mathbf{\Delta}$ is nonsingular.*

Proof. See Appendix A. \square

Lemma 2. *Define the augmented coarray manifold matrix as $\mathbf{V}_{\mathbb{D}} = [\mathbf{V}_0 \mathbf{J} \text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}}, \mathbf{W}_{\mathbb{D}}]$. If $\mathbf{V}_{\mathbb{D}}$ is full column rank, i.e., $\text{rank}(\mathbf{V}_{\mathbb{D}}) = 2Q + 2$, then $\mathbf{G}^H \mathbf{\Pi}_{\mathbf{\Delta}}^{\perp} \mathbf{G}$ is nonsingular.*

Proof. See Appendix B. \square

It is clear that Lemma 1 always holds when Lemma 2 is satisfied. Hence, Lemma 2 is the rank condition for the existence of the CRB based on CSA. From Lemma 2, we can find that the rank condition depends on the following factors: the compressive matrix $\mathbf{\Phi}$, the coarray lags \mathbb{D} , the spatial frequencies $\boldsymbol{\omega}$, \mathbf{e}_0 , and $\text{vec}(\mathbf{I}_M)$. Thus, we have the following corollaries on the rank condition.

Corollary 1. *If $M^2 > |\mathbb{D}|$, then, for any choice of Q distinct DOAs, $\text{rank}(\mathbf{V}_{\mathbb{D}}) < 2Q + 2$, i.e., the FIM is singular, when $Q > (|\mathbb{D}| - 1)/2$.*

If $M^2 \leq |\mathbb{D}|$, then, for any choice of Q distinct DOAs, $\text{rank}(\mathbf{V}_{\mathbb{D}}) < 2Q + 2$, i.e., the FIM is singular, when $Q > \lfloor (M^2 - 2)/2 \rfloor$.

Proof. See Appendix C. \square

Corollary 1 gives the maximum number of DOFs that the CSA can achieve. It is worth noting that, for some values of Q distinct DOAs, the FIM is nonsingular when Q approaches

$(|\mathbb{D}| - 1)/2$ with $M^2 > |\mathbb{D}|$ or $\lfloor (M^2 - 2)/2 \rfloor$ with $M^2 \leq |\mathbb{D}|$. In contrast, for some other values of the DOAs, the FIM may be singular before Q achieves the number of DOFs in Corollary 1. To show the superiority of the CSA in terms of the number of DOFs, we provide the following theorem.

Theorem 1. *Consider a CSA with M front-end circuit chains based on an L -element sparse array with an arbitrary linear configuration, where $L > M$. Let N be the number of overlapping lags. If $N > 2$, the CSA can obtain a higher number of DOFs than the M -element sparse array with the same configuration.*

Proof. See Appendix D. \square

We will explain the meaning of "the same configuration" in next subsection by taking some sparse arrays with nested structure as instance. From Theorem 1, we can straightforwardly have the following corollary.

Corollary 2. *If the number of front-end circuit chains M is higher than 3 and $L > M$, then, for any sparse linear array configuration, the number of DOFs of the M -chain CSA is always higher than that of the M -element sparse array with the same configuration.*

Proof. For sparse arrays, the number of overlapping lags in the 0-th position is $M - 1$. Thus, if $M > 3$, we have $N > 2$. According to Theorem 1, if $L > M$, the number of DOFs for CSA is always higher than that of the M -element sparse array. \square

From the previous corollaries, we can find that the improvement on the number of DOFs of the CSA is closely related to the number of unique lags in the coarray, i.e., the redundancy of the sparse array. A better improvement can be obtained when the array structure has a higher redundancy. However, it does not mean that a higher number of DOFs can be obtained by using sparse arrays which have higher redundancy. The reason is that the number of DOFs of a CSA also depends on the number of DOFs of the sparse array before compression as indicated by Corollary 1.

B. Examples of Sparse Arrays with Nested Structure

The sparse array with nested structure has a hole-free coarray, which is an important aspect in designing a sparse array structure. In addition, unlike the minimum redundancy array (MRA) [31], the nested sparse array structures have a closed-form expression on the antenna positions. Thus, we show the superiority of the proposed CSA scheme on the number of DOFs by considering the ULA and the following three nested structures, i.e., the nested coprime array with compressed inter-element spacing (N-CACIS) [6], the nested coprime array with displaced subarrays (N-CADiS) [6], and the nested array [3]. The corresponding array configurations are depicted in Fig. 2. To maximum the number of DOFs, the values of N_1 and N_2 are selected as

$$\begin{cases} N_1 = L/2, & N_2 = L/2, & \text{for even } L, \\ N_1 = (L - 1)/2, & N_2 = (L + 1)/2, & \text{for odd } L, \end{cases} \quad (32)$$

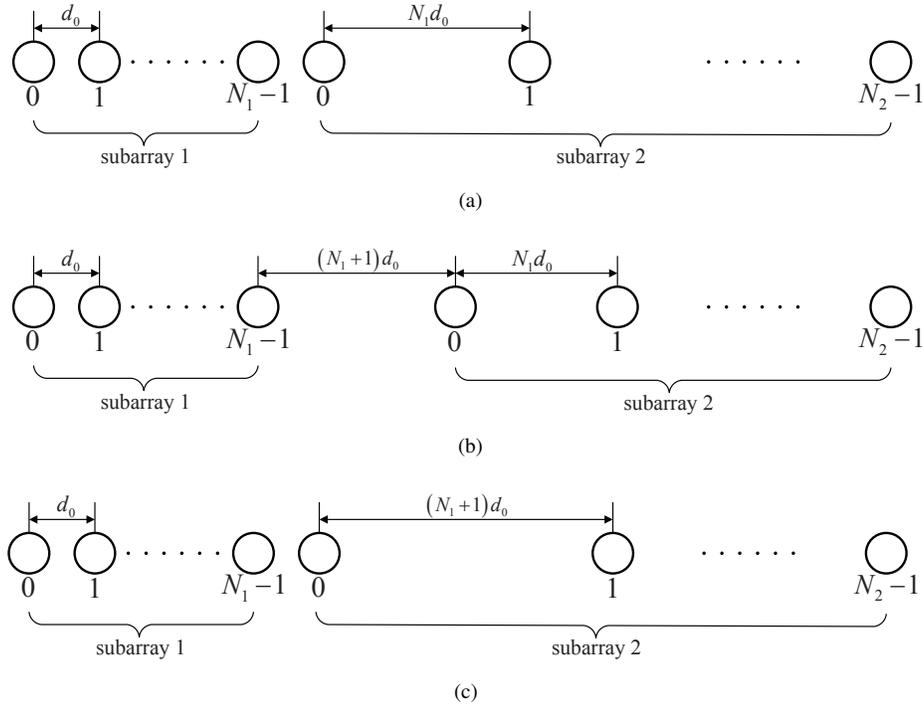


Fig. 2. Array configurations of (a) N-CACIS, (b) N-CADiS, and (c) nested array.

where $L \geq 3$. Note that, when $L = 2$, all the sparse array configurations are same, namely, two antennas with unit interelement spacing. As mentioned before, the sparse array with the same configuration means that, for example, if N-CACIS is exploited, then, the number of antennas in each subarray of both the L -element N-CACIS and the M -element N-CACIS is determined as indicated in (32) and the antennas are placed under the N-CACIS configuration as shown in Fig. 2(a).

To quantitatively illustrate the superiority of the proposed CSA, we summarize the improvement on the number of DOFs of a CSA based on the N-CACIS, N-CADiS, nested array, and ULA for odd and even values of L in Table I and Table II, respectively, where $L \geq 3$ is the total number of physical antennas. Note that the redundancy for all array configurations is 1 when $L = 2$. Fig. 3 gives the redundancy versus L for $2 \leq L \leq 14$. It is clear that the ULA has the highest redundancy. When L is even, the N-CADiS has lowest redundancy, whereas when L is odd, the N-CADiS and nested arrays have the same lowest redundancy.

To show the improvement more clearly, we plot in Fig. 4 the number of DOFs for the ULA and the N-CADiS, with different values of M ranging from 2 to 14. The ULA-based M -chain CSA and N-CADiS-based M -chain CSA exploit the ULA and N-CADiS with 14 antennas as the receive array, respectively. It is observed that the M -chain CSA can obtain a higher number of DOFs than the M -element sparse array. For the ULA, its DOFs are not lost even when the number of chains is compressed from 14 to 6 while for the N-CADiS, the number of DOFs is limited by M when $M \leq 10$. However, in this case, the CSA based on the N-CADiS still obtains the highest number of DOFs.

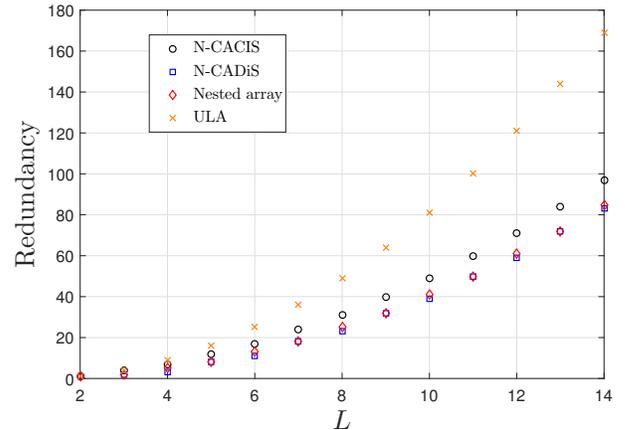


Fig. 3. Redundancy versus L .

C. The CRB Expression

From Lemma 1 and Lemma 2, we can obtain the expression of the CRB for the normalized spatial frequencies exploiting CSA, which is given in Theorem 2.

Theorem 2. *If the rank of the augmented coarray manifold matrix $\mathbf{V}_{\mathbb{D}}$ is $2Q + 2$, then the CRB for the normalized spatial frequencies $\boldsymbol{\omega} = [\omega_1, \dots, \omega_Q]^T$ is expressed as*

$$\text{CRB}(\boldsymbol{\omega}) = \frac{1}{4\pi^2 T} (\mathbf{G}_0^H \boldsymbol{\Pi}_{\mathbf{M}_0 \mathbf{V}_{\mathbb{D}}} \mathbf{G}_0)^{-1}, \quad (33)$$

where

$$\mathbf{G}_0 = \mathbf{M}_0 \mathbf{V}_0 \mathbf{J} \text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}} \text{diag}([p_1, \dots, p_Q]), \quad (34)$$

TABLE I
NUMBER OF DOFS OF CSA BASED ON POPULAR SPARSE ARRAY CONFIGURATIONS FOR ODD L

Array configuration	Unique lags	Redundancy	DOF of M -chain CSA
N-CACIS [6]	$\frac{L^2 + 1}{2}$	$\frac{L^2 - 1}{2}$	$\frac{L^2 - 1}{4}, M^2 > \frac{L^2 + 1}{2}; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq \frac{L^2 + 1}{2}$
N-CADiS [6]	$\frac{L^2 + 2L - 1}{2}$	$\frac{L^2 - 2L + 1}{2}$	$\frac{L^2 + 2L - 3}{4}, M^2 > \frac{L^2 + 2L - 1}{2}; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq \frac{L^2 + 2L - 1}{2}$
Nested Array [3]	$\frac{L^2 + 2L - 1}{2}$	$\frac{L^2 - 2L + 1}{2}$	$\frac{L^2 + 2L - 3}{4}, M^2 > \frac{L^2 + 2L - 1}{2}; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq \frac{L^2 + 2L - 1}{2}$
ULA	$2L - 1$	$L^2 - 2L + 1$	$L - 1, M^2 > L^2 - 2L + 1; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq L^2 - 2L + 1$

TABLE II
NUMBER OF DOFS OF CSA BASED ON POPULAR SPARSE ARRAY CONFIGURATIONS FOR EVEN L

Array configuration	Unique lags	Redundancy	DOF of M -chain CSA
N-CACIS [6]	$\frac{L^2 + 2}{2}$	$\frac{L^2 - 2}{2}$	$\frac{L^2}{4}, M^2 > \frac{L^2 + 2}{2}; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq \frac{L^2 + 2}{2}$
N-CADiS [6]	$\frac{L^2 + 2L + 2}{2}$	$\frac{L^2 - 2L - 2}{2}$	$\frac{L^2 + 2L}{4}, M^2 > \frac{L^2 + 2L + 2}{2}; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq \frac{L^2 + 2L + 2}{2}$
Nested Array [3]	$\frac{L^2 + 2L - 2}{2}$	$\frac{L^2 - 2L + 2}{2}$	$\frac{L^2 + 2L - 4}{4}, M^2 > \frac{L^2 + 2L - 2}{2}; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq \frac{L^2 + 2L - 2}{2}$
ULA	$2L - 1$	$L^2 - 2L + 1$	$L - 1, M^2 > L^2 - 2L + 1; \lfloor \frac{M^2 - 2}{2} \rfloor, M^2 \leq L^2 - 2L + 1$

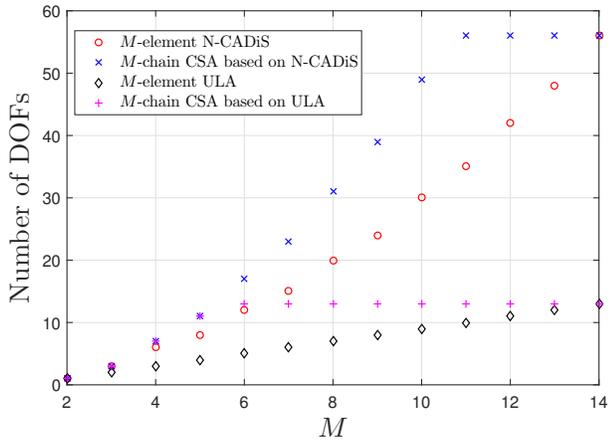


Fig. 4. Number of DOFs versus M .

$$\mathbf{M}_0 = (\mathbf{R}_{\mathbf{y}\mathbf{y}}^T \otimes \mathbf{R}_{\mathbf{y}\mathbf{y}})^{-\frac{1}{2}}, \quad (35)$$

$$\mathbf{V}_0 = (\mathbf{\Phi}^* \otimes \mathbf{\Phi}), \quad (36)$$

$$\mathbf{W}_{\mathbb{D}} = [\mathbf{V}_0 \mathbf{J} \mathbf{A}_{\mathbb{D}}, \mathbf{V}_0 \mathbf{J} \mathbf{e}_0, \text{vec}(\mathbf{I}_M)], \quad (37)$$

$$\mathbf{R}_{\mathbf{y}\mathbf{y}} = \mathbf{\Phi} \mathbf{A}_{\mathbb{S}} \mathbf{R}_{\mathbb{S}} \mathbf{A}_{\mathbb{S}}^H \mathbf{\Phi}^H + p_v \mathbf{\Phi} \mathbf{\Phi}^H + p_w \mathbf{I}_M, \quad (38)$$

$$\mathbf{A}_{\mathbb{D}} = [\mathbf{a}_{\mathbb{D}}(\omega_1), \dots, \mathbf{a}_{\mathbb{D}}(\omega_Q)], \quad (39)$$

$$\text{vec}(\mathbf{I}_L) = \mathbf{J} \mathbf{e}_0. \quad (40)$$

Here, p_q is the power of the q -th source, $q = 1, 2, \dots, Q$, p_v is the power of the signal noise, and p_w is the power of the measurement noise.

Proof. It is straightforward to prove (33) by substituting (44), (47), and (48) into (31). \square

Overall, the superiority of the proposed CSA scheme is summarized as follows:

- Given $L > M > 3$, according to Corollary 2, the M -chain CSA based on an L -element sparse array can always obtain a higher number of DOFs than the M -element sparse array with the same structure. This conclusion applies to all linear array configurations, including the original and generalized coprime arrays, nested arrays and the MRA. Furthermore, since an M -chain CSA is based on the L -element sparse array where $L > M$, the corresponding receive array aperture is always larger than that of the M -

- element sparse array where the number of circuit chains is the same, thus enabling a higher estimation accuracy than directly using the M -element sparse array, provided that the sources are resolvable under both structures.
- b) The hardware complexity is reduced by exploiting the proposed CSA implementation, since the number of front-end circuit chains is substantially decreased. If the prior knowledge on DOAs is known, the DOA estimation performance loss can be reduced by optimizing the compression matrix Φ ;
 - c) Compared with using a ULA, using a sparse array can obtain a much higher number of DOFs and a much larger aperture with the same number of antennas. For the same array aperture, using the sparse array can lead to a much lower cost from the receive array aspect.

VI. SIMULATION RESULTS

Throughout this section, SNR_1 and SNR_2 are defined in terms of the signal power p_s as

$$\text{SNR}_1 = \frac{p_s}{p_v}, \quad \text{SNR}_2 = \frac{p_s}{p_w},$$

respectively. All sources are assumed to have the same power. In addition, the sources are uncorrelated with each other and independent to both the signal noise and the measurement noise. The root mean square error (RMSE) of the estimated spatial frequency is defined as

$$\text{RMSE} = \sqrt{\frac{1}{IQ} \sum_{i_0=1}^I \sum_{q=1}^Q (\omega_q - \tilde{\omega}_q^{(i_0)})^2},$$

where I is the number of Monte Carlo trials and $\tilde{\omega}_q^{(i_0)}$ is the estimated normalized spatial frequency of the q -th source in the i_0 -th trial. The N-CACIS is used as the receive antenna array. Unless otherwise specified, the entries of the compression matrix Φ are independently drawn from a complex random Gaussian distribution $\mathcal{CN}(0, 1)$. It is noted that the following simulations aim at evaluating the performance of the proposed CSA architecture. More specifically, we compare the proposed CSA, consisting of L antennas and M circuits where $L > M$, with two conventional N-CACIS sparse array structures, one with M elements and the other with L elements. The values of M and L are separately determined in each simulation. As previously analyzed, it is expected that the proposed CSA achieves a lower system complexity than the L -element N-CACIS by slightly sacrificing the estimation accuracy. On the other hand, the propose CSA is expected to outperform the M -element N-CACIS in terms of both the number of DOFs and the estimation accuracy aspects due to the extension of the receive array.

A. Number of DOFs

In this subsection, the M -chain CSA is designed based on a 10-element N-CACIS. The LASSO algorithm is used to estimate the spatial frequencies ω which is set as $-0.49 + 0.99(q-1)/Q$ with $q = 1, 2, \dots, Q$.

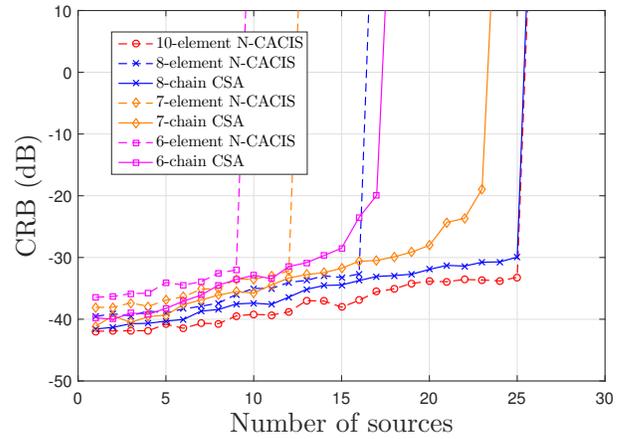


Fig. 5. The CRB versus the number of sources. SNR_1 is 0 dB and the number of snapshots is 5,000.

We first consider the CRBs for both N-CACIS and CSA. The CRB of the normalized spatial frequencies for the N-CACIS is plotted as discussed in [24]. Similar results of the CRB based on the sparse arrays can be found in [32] and [33]. Meanwhile, the expression of the CRB for the normalized spatial frequencies based on the CSA is derived in Theorem 2. For simplicity, we only verify the scenarios where M equals 6, 7, and 8. SNR_1 and SNR_2 are set as 0 dB and 20 dB, respectively, and the number of snapshots is 5,000. The results are plotted in Fig. 5. It is clear that, for N-CACIS with 6, 7, and 8 antennas, the number of DOFs is respectively 9, 12, and 16, while 17, 23, and 25 sources can be resolved for the CSA with 6, 7, and 8 front-end circuit chains. The result is consistent with Corollary 1.

Then, we examine the spatial spectra in the corresponding scenarios. In addition, 15, 20, and 23 uncorrelated far-field narrowband sources are considered for $M = 6, 7, \text{ and } 8$, respectively. The estimated spatial spectra are shown in Fig. 6. The red dash lines show the true spatial frequencies. We can find that the M -element N-CACIS cannot resolve the sources, while the M -element CSA correctly estimated the directions of the sources. The results confirm that the CSA can estimate more sources than the N-CACIS when they assume the same number of front-end circuit chains, M .

B. Estimation Accuracy

In this subsection, we examine the DOA estimation accuracy of CSA. Three configurations are taken into account, namely, the N-CACIS with 10 antennas, the N-CACIS with 8 antennas, and the CSA with 8 chains compressed from 10-element N-CACIS. Limited by the number of DOFs available to the 8-element N-CACIS, only 10 sources are considered, and their directions are set as $-0.49 + 0.99(q-1)/Q$ with $q = 1, 2, \dots, Q$. The LASSO algorithm is used to estimate the DOAs, and 500 Monte Carlo trials are conducted to calculate the RMSE. Then RMSE versus SNR_1 and the number of snapshots are plotted in Fig. 7 and Fig. 8, respectively, where SNR_2 is set to 40 dB. In Fig. 7, 2,000 snapshots are used, and η is set as 0.2. In Fig. 8, 20 dB SNR_1 is used, and η

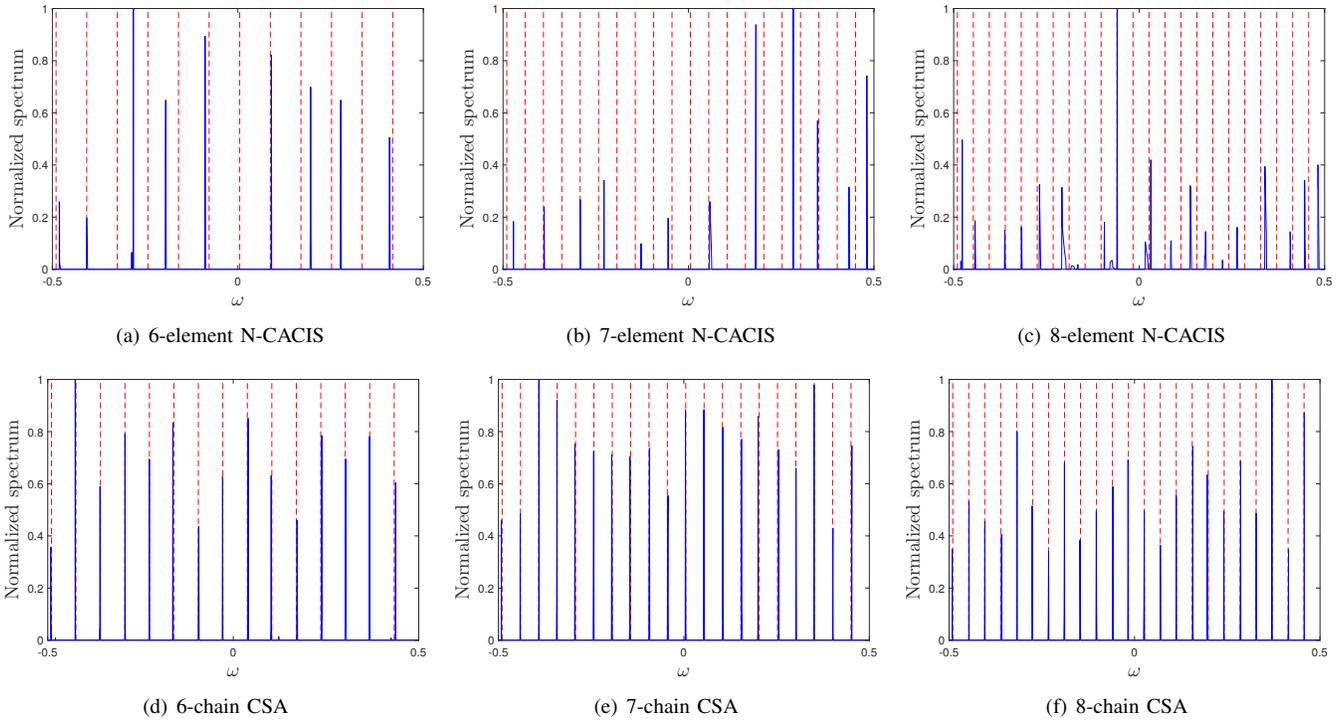


Fig. 6. The spatial spectra under 0 dB SNR₁ and 5,000 snapshots: (a) 15 sources with 6-element N-CACIS; (b) 20 sources with 7-element N-CACIS; (c) 23 sources with 8-element N-CACIS; (d) 15 sources with 6-chain CSA; (e) 20 sources with 7-chain CSA; (f) 23 sources with 8-chain CSA.

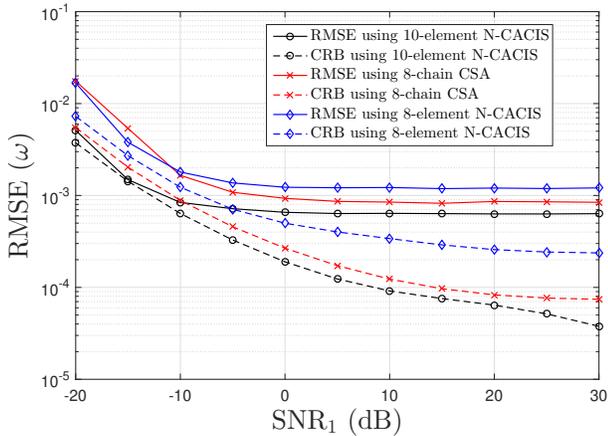


Fig. 7. The RMSE versus SNR₁. SNR₂ is 40 dB and the number of snapshots is 2,000.

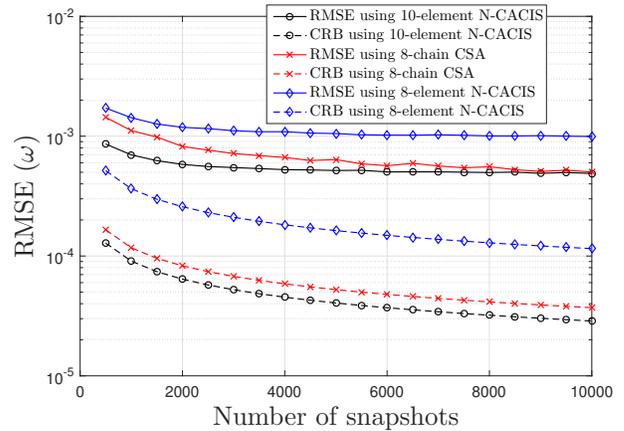


Fig. 8. The RMSE versus the number of snapshots. SNR₁ is 20 dB and SNR₂ is 40 dB.

is set as 0.15 for the 10-element N-CACIS and 0.2 for the 8-element N-CACIS and the 8-chain CSA. In this case, the 8-chain CSA based on 10-element N-CACIS achieves a higher accuracy as compared to the 8-element N-CACIS. Thus, given the number of front-end circuit chains, both the number of DOFs and the estimation accuracy can be improved by using the CSA structure.

C. Optimization of Compression Matrix

The effectiveness of the optimization procedure for Φ is examined in this subsection. It is noted that θ in degree is used here since the prior information is given for θ instead

of ω . To consider a more general scenario, we assume two sources θ_1 and θ_2 which both follow the Gaussian distribution $\mathcal{N}(0, 5^2)$. Furthermore, the interval between the two sources is constrained to be larger than 2° . The 12-element N-CACIS is compressed to 6 chains here, and the yielding compression ratio is 2. The CS-MVDR is used to estimated the DOAs and no coarray operation is performed. For the optimization procedure, Φ is optimized through 2000 steps with a step size of 0.01. In addition, as demonstrated in [20], we can optimize Φ in the high SNR₁ region when the exact input SNR₁ is unknown. Here Φ is optimized by using 60 dB SNR₁, and 500 Monte Carlo trials are conducted to calculate the RMSE.

Given 100 snapshots and 50 dB SNR₂, the RMSE versus

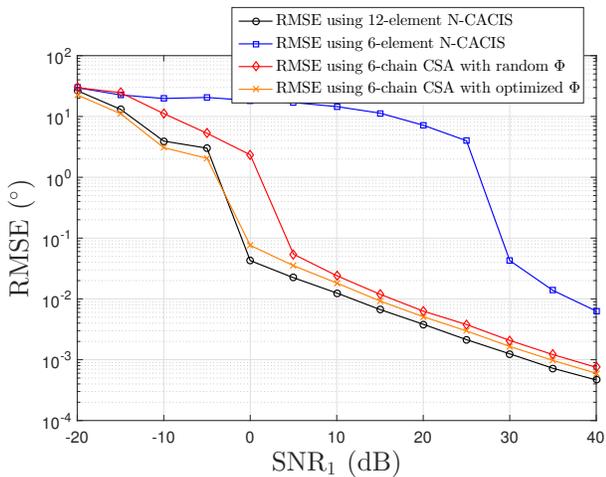


Fig. 9. The RMSE versus SNR_1 . SNR_2 is 50 dB, and the number of snapshots is 100.

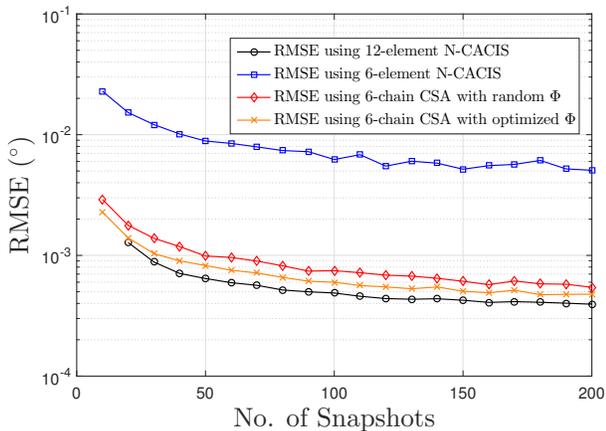


Fig. 10. The RMSE versus the number of snapshots. SNR_1 is 40 dB and SNR_2 is 50 dB.

SNR_1 is plotted in Fig. 9. Furthermore, the RMSE versus the number of snapshots is plotted in Fig. 10, where SNR_1 is 40 dB and SNR_2 is 50 dB. It is evident that using optimized Φ can obtain a higher accuracy than using randomly selected Φ . Furthermore, both optimized and randomly selected Φ can have a much better estimation performance than using the 6-element N-CACIS. Note in Fig. 10 that the estimation is failed using the 12-element N-CACIS when the number of snapshots is 10. One important reason that leads to the gap between the 6-chain CSA using the optimized Φ and the 12-element N-CACIS is the constraint on the interval between the two sources, i.e., the sources do not exactly follow the Gaussian distribution due to such constraint. Thus, there is still some information loss even when the optimized Φ is used.

D. Angular Resolution

In this subsection, the angular resolution capability for two closely spaced sources is evaluated. A 12-element N-CACIS is compressed to 6 chains. CS-MVDR is performed to estimate the DOAs. Since no prior knowledge is given here, every entry in Φ is randomly selected from an independently and identi-

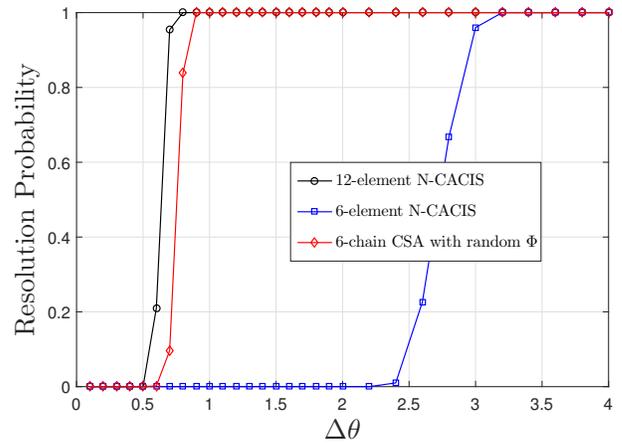


Fig. 11. The RMSE versus the number of snapshots. Both SNR_1 and SNR_2 are 20 dB.

cally distributed complex Gaussian distribution $\mathcal{CN}(0, 1)$. Two sources located on $30^\circ - \Delta\theta$ and $30^\circ + \Delta\theta$ are considered. The two sources are supposed to be correctly resolved when there are two peaks and the estimated DOAs $\hat{\theta}_1$ and $\hat{\theta}_2$ satisfy $|\hat{\theta}_1 - \theta_1| < \Delta\theta$ and $|\hat{\theta}_2 - \theta_2| < \Delta\theta$. In this simulation, SNR_1 and SNR_2 are respectively set to 10 dB and 20 dB, and 100 snapshots are used. $\Delta\theta$ ranges from 0.1° to 4° with step size 0.1° from 0.1° to 2° and step size 0.2° from 2° to 4° . The resolution probability is computed through 500 Monte Carlo trials. The simulation result is shown in Fig. 11. As expected, the resolution probability of 6-chain CSA is slightly worse than that of 12-element N-CACIS without compression. However, the resolution capability of the 6-chain CSA is much better than that of the 6-element N-CACIS.

VII. CONCLUSION

We developed a generalized structure for compressed sparse array, which consists of an arbitrary sparse array and a combining network, in order to achieve effective DOA estimation using a reduced number of front-end chains. By analyzing the Cramr-Rao bound and its existence conditions, we derived the analytical expression of the achievable number of degrees of freedom. It is revealed that, given the same number of front-end chains, the number of DOFs as well as the DOA estimation accuracy can be increased by using more sparse array antennas. The analytical expressions developed in this paper enabled analytical assessment of the achievable number of DOFs and optimized design of the sparse array and the combining network. We further optimized the compression matrix used in the combining network to minimize the information loss in the process of data combining in order to achieve a high DOA estimation accuracy.

APPENDIX A PROOF OF LEMMA 1

Here we first derive the expression for Δ . The partial derivative of $\mathbf{r}_{\mathbf{y}\mathbf{y}}$ with respect to p_q is given by

$$\frac{\partial \mathbf{r}_{\mathbf{y}\mathbf{y}}}{\partial p_q} = \mathbf{V}_0 \mathbf{J}_{\mathbf{a}_{\mathbb{D}}}(\omega_q). \quad (41)$$

Similarly, we have

$$\frac{\partial \mathbf{r}_{yy}}{\partial p_v} = \mathbf{V}_0 \mathbf{J} \mathbf{e}_0, \quad (42)$$

$$\frac{\partial \mathbf{r}_{yy}}{\partial p_w} = \text{vec}(\mathbf{I}_M). \quad (43)$$

Denote $[\mathbf{V}_0 \mathbf{J} \mathbf{A}_{\mathbb{D}}, \mathbf{V}_0 \mathbf{J} \mathbf{e}_0, \text{vec}(\mathbf{I}_M)]$ as $\mathbf{W}_{\mathbb{D}}$. Then, Δ is expressed as

$$\Delta = (\mathbf{R}_{yy}^T \otimes \mathbf{R}_{yy})^{-\frac{1}{2}} \mathbf{W}_{\mathbb{D}}. \quad (44)$$

It is easily to verify that the dimension of $\mathbf{W}_{\mathbb{D}}$ is $M^2 \times (Q+2)$ and the rank of Δ equals the rank of $\mathbf{W}_{\mathbb{D}}$.

For $\Delta^H \Delta$, we have the following rank properties

$$\begin{aligned} \text{rank}(\Delta^H \Delta) &\geq \text{rank}(\Delta^H) + \text{rank}(\Delta) - (Q+2) \\ &= 2\text{rank}(\Delta) - (Q+2), \end{aligned} \quad (45)$$

$$\begin{aligned} \text{rank}(\Delta^H \Delta) &\leq \max\{\text{rank}(\Delta^H), \text{rank}(\Delta)\} \\ &= \text{rank}(\Delta). \end{aligned} \quad (46)$$

Thus, if Δ has a full column rank, i.e., $\text{rank}(\mathbf{W}_{\mathbb{D}}) = Q+2$, then $\text{rank}(\Delta^H \Delta) = Q+2$, indicating that $\Delta^H \Delta$ is nonsingular.

APPENDIX B PROOF OF LEMMA 2

Combining (10) and (29) yields

$$\begin{aligned} \mathbf{G} &= (\mathbf{R}_{yy}^T \otimes \mathbf{R}_{yy})^{-\frac{1}{2}} \mathbf{V}_0 \mathbf{J} \left[p_1 \frac{\partial \mathbf{a}_{\mathbb{D}}(\omega_1)}{\partial \omega_1}, \dots, p_Q \frac{\partial \mathbf{a}_{\mathbb{D}}(\omega_Q)}{\partial \omega_Q} \right] \\ &= j2\pi (\mathbf{R}_{yy}^T \otimes \mathbf{R}_{yy})^{-\frac{1}{2}} \mathbf{V}_0 \mathbf{J} \text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}} \mathbf{P}, \end{aligned} \quad (47)$$

where $\mathbf{P} = \text{diag}([p_1, \dots, p_Q])$. For notational simplicity, denote $(\mathbf{R}_{yy}^T \otimes \mathbf{R}_{yy})^{-\frac{1}{2}}$ as \mathbf{M}_0 . Then, we can obtain

$$\frac{\mathbf{G}^H \Pi_{\Delta}^{\perp} \mathbf{G}}{4\pi^2} = \mathbf{G}_0^H \Pi_{\mathbf{M}_0 \mathbf{W}_{\mathbb{D}}}^{\perp} \mathbf{G}_0, \quad (48)$$

where $\mathbf{G}_0 = \mathbf{M}_0 \mathbf{V}_0 \mathbf{J} \text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}} \mathbf{P}$. Since the projection matrix $\Pi_{\mathbf{M}_0 \mathbf{W}_{\mathbb{D}}}^{\perp}$ is Hermitian and idempotent [34], an arbitrary vector \mathbf{u}_0 satisfies the following inequality

$$\|\Pi_{\mathbf{M}_0 \mathbf{W}_{\mathbb{D}}}^{\perp} \mathbf{G}_0 \mathbf{u}_0\|_2^2 \geq 0. \quad (49)$$

It is easily to find that the equality in (49) holds if and only if $\Pi_{\mathbf{M}_0 \mathbf{W}_{\mathbb{D}}}^{\perp} \mathbf{G}_0 \mathbf{u}_0 = \mathbf{0}$, which indicates that $\mathbf{G}_0 \mathbf{u}_0$ lies in the column space of $\mathbf{M}_0 \mathbf{W}_{\mathbb{D}}$, namely, $\mathbf{G}_0 \mathbf{u}_0 = \text{col}(\mathbf{M}_0 \mathbf{W}_{\mathbb{D}})$. Thus, there exists a vector \mathbf{v}_0 satisfying

$$\mathbf{G}_0 \mathbf{u}_0 = \mathbf{M}_0 \mathbf{W}_{\mathbb{D}} \mathbf{v}_0. \quad (50)$$

Since \mathbf{M}_0 is positive definite, the following equality is satisfied:

$$\mathbf{V}_0 \mathbf{J} \text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}} \mathbf{P} \mathbf{u}_0 - \mathbf{W}_{\mathbb{D}} \mathbf{v}_0 = 0. \quad (51)$$

Rewrite (51) into a matrix form as

$$[\mathbf{V}_0 \mathbf{J} \text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}} \quad \mathbf{W}_{\mathbb{D}}] \begin{bmatrix} \mathbf{P} \mathbf{u}_0 \\ -\mathbf{v}_0 \end{bmatrix} = 0. \quad (52)$$

Hence, if $\mathbf{V}_{\mathbb{D}} = [(\Phi^* \otimes \Phi) \text{diag}(\mathbb{D}), \mathbf{W}_{\mathbb{D}}]$ has a full column rank, (52) holds if and only if $\mathbf{P} \mathbf{u}_0$ is a zero vector, indicating that \mathbf{u}_0 is a zero vector.

Combining (52) and (49), for an arbitrary non-zero vector \mathbf{u}_0 , if the rank of $\mathbf{V}_{\mathbb{D}}$ is $2Q+2$, (49) is simplified to

$$\|\Pi_{\mathbf{M}_0 \mathbf{W}_{\mathbb{D}}}^{\perp} \mathbf{G}_0 \mathbf{u}_0\|_2^2 > 0. \quad (53)$$

As such, $\mathbf{G}^H \Pi_{\Delta}^{\perp} \mathbf{G}$ is positive definite and nonsingular, which completes the proof.

APPENDIX C PROOF OF COROLLARY 1

Rewrite $\mathbf{V}_{\mathbb{D}}$ as $\mathbf{V}_{\mathbb{D}} = [\mathbf{V}_0 \mathbf{J} [\text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}}, \mathbf{A}_{\mathbb{D}}, \mathbf{e}_0], \text{vec}(\mathbf{I}_M)]$ and denote $\hat{\mathbf{V}}_{\mathbb{D}} = [\text{diag}(\mathbb{D}) \mathbf{A}_{\mathbb{D}}, \mathbf{A}_{\mathbb{D}}, \mathbf{e}_0]$. The dimension of $\hat{\mathbf{V}}_{\mathbb{D}}$ is $|\mathbb{D}| \times (2Q+1)$. Thus, if $M^2 > |\mathbb{D}|$, then, for any choice of Q distinct DOAs, the following inequalities hold when $Q > (|\mathbb{D}| - 1)/2$,

$$\begin{aligned} \text{rank}(\mathbf{V}_{\mathbb{D}}) &\leq \text{rank}(\mathbf{V}_0 \mathbf{J} \hat{\mathbf{V}}_{\mathbb{D}}) + 1 \\ &\leq \text{rank}(\hat{\mathbf{V}}_{\mathbb{D}}) + 1 \\ &\leq |\mathbb{D}| + 1 \\ &< 2Q + 2. \end{aligned} \quad (54)$$

If $M^2 \leq |\mathbb{D}|$, then, for any choice of Q distinct DOAs, $\mathbf{V}_{\mathbb{D}}$ become a fat matrix when $Q > [(M^2 - 2)/2]$, indicating that $\text{rank}(\mathbf{V}_{\mathbb{D}}) < 2Q + 2$.

APPENDIX D PROOF OF THEOREM 1

We should first note that the M -element sparse array needs to have the same design with the L -element sparse array. For example, if the L -element sparse array is a CACIS [6], then, the M -element sparse array also follows the CACIS design. Denote the number of unique lags in the L -element sparse array and the M -element sparse array as $|\mathbb{D}_L|$ and $|\mathbb{D}_M|$, respectively. Due to the property of sparse arrays, $|\mathbb{D}_L|$ is always larger than $|\mathbb{D}_M|$, if $L > M$.

When $M^2 > |\mathbb{D}|$, according to Corollary 1, the number of DOFs for an M -chain CSA is $(|\mathbb{D}_L| - 1)/2$, while the number of DOFs for an M -element sparse array is $(|\mathbb{D}_M| - 1)/2$. Thus, the M -chain CSA can obtain a higher number of DOFs in contrast with the M -element sparse array.

When $M^2 \leq |\mathbb{D}|$, according to Corollary 1, the number of DOFs for the M -chain CSA is $\lfloor (M^2 - 2)/2 \rfloor$. Then, the number of DOFs is expressed as

$$\text{DOF}_{\text{CSA}} = \begin{cases} \frac{M^2 - 3}{2}, & M \text{ is odd,} \\ \frac{M^2 - 2}{2}, & M \text{ is even.} \end{cases} \quad (55)$$

For an M -element sparse array with N overlapping lags in the corresponding coarray, the number of unique lags is

$$|\mathbb{D}_M| = M^2 - N. \quad (56)$$

We should note that the N overlapping lags consist of two parts, i.e., the N_0 overlapping lags on the 0-th position and the N_1 overlapping lags on other positions, where $N = N_0 + N_1$.

Due to the symmetric property of the coarray in terms of the 0-th position, N_0 is even if M is odd, otherwise N_0 is odd. Furthermore, N_1 is always even for any value of M . As a result, N is even if M is odd while N is odd if M is even, indicating that $|\mathbb{D}_M|$ is always odd for any choice of M . The number of DOFs for the M -element sparse array is expressed as [24]

$$\text{DOF}_{\text{SA}} = \frac{|\mathbb{D}_M| - 1}{2} = \frac{M^2 - N - 1}{2}. \quad (57)$$

Comparing (55) with (57), it is straightforward that if $N > 2$, the number of DOFs of the CSA is higher than that of the sparse array.

REFERENCES

- [1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*. New York, NY: Wiley, 2002.
- [2] S. Chandran, *Advances in Direction-of-Arrival Estimation*. Norwood, MA: Artech House, 2006.
- [3] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, April 2010.
- [4] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Oct. 2010.
- [5] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," in *Proc. IEEE Digital Signal Process. Signal Process. Educ. Workshop*, Sedona, AZ, March 2011, pp. 289–294.
- [6] S. Qin, Y. D. Zhang, and M. G. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1377–1390, March 2015.
- [7] A. Ahmed, Y. D. Zhang, and B. Himed, "Effective nested array design for fourth-order cumulant-based DOA estimation," in *Proc. IEEE Radar Conf.*, Seattle, WA, May 2017, pp. 998–1002.
- [8] Y. Wang, G. Leus, and A. Pandharipande, "Direction estimation using compressive sampling array processing," in *Proc. IEEE Stat. Signal Process. Workshop*, Cardiff, UK, Aug. 2009, pp. 626–629.
- [9] J. F. Gu, W. P. Zhu, and M. N. S. Swamy, "Compressed sensing for DOA estimation with fewer receivers than sensors," in *Proc. IEEE Int. Symp. Circuits and Systems*, Rio de Janeiro, Brazil, May 2011, pp. 1752–1755.
- [10] M. Ibrahim, V. Ramireddy, A. Lavrenko, J. König, F. Römer, M. Landmann, M. Grossmann, G. D. Galdo, and R. S. Thomä, "Design and analysis of compressive antenna arrays for direction of arrival estimation," *Signal Process.*, vol. 138, pp. 35–47, Sep. 2017.
- [11] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, March 1986.
- [12] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.
- [13] J. E. Evans, J. R. Johnson, and D. F. Sun, "High resolution angular spectrum estimation techniques for terrain scattering analysis and angle of arrival estimation," in *Proc. 1st ASSP Workshop Spectral Estimation*, Hamilton, ON, Canada, Aug. 1981, pp. 134–139.
- [14] C.-L. Liu, and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1438–1442, March 2015.
- [15] C.-L. Liu, P. P. Vaidyanathan, and P. Pal, "Coprime coarray interpolation for DOA estimation via nuclear norm minimization," in *Proc. 2016 IEEE Int. Symp. Circuits Syst.*, Montreal, QC, Canada, May 2016, pp. 2639–2642.
- [16] M. Guo, T. Chen, and B. Wang, "An improved DOA estimation approach using coarray interpolation and matrix denoising," *Sensors*, vol. 17, no. 5, pp. 1140, May 2017.
- [17] T. Chen, M. Guo, and L. Guo, "A direct coarray interpolation approach for direction finding," *Sensors*, vol. 17, no. 9, pp. 2149, Sep. 2017.
- [18] Y. D. Zhang, M. G. Amin, and B. Himed, "Sparsity-based DOA estimation using co-prime arrays," in *Proc. IEEE ICASSP*, Vancouver, BC, Canada, May 2013, pp. 3967–3971.
- [19] P. Pakrooh, A. Pezeshki, L. L. Scharf, D. Cochran, S. D. Howard, "Analysis of Fisher information and the Cramér-Rao bound for nonlinear parameter estimation after random compression," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6423–6428, Aug. 2015.
- [20] Y. Gu and N. A. Goodman, "Information-theoretic compressive sensing kernel optimization and Bayesian Cramér-Rao bound for time delay estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4525–4537, May 2017.
- [21] Y. Gu, Y. D. Zhang, N. A. Goodman, "Optimized compressive sensing-based direction-of-arrival estimation in massive MIMO," in *Proc. IEEE ICASSP*, New Orleans, LA, March 2017, pp. 3181–3185.
- [22] C. Zhou, Y. Gu, Y. D. Zhang, Z. Shi, T. Jin, and X. Wu, "Compressive sensing based coprime array direction-of-arrival estimation," *IET Commun.*, vol. 11, no. 11, pp. 1719–1724, Sep. 2017.
- [23] W. Wu, C. Cooper, and N. Goodman, "Switched-element direction finding," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 3, pp. 1209–1217, July 2009.
- [24] C.-L. Liu and P. P. Vaidyanathan, "Cramér-Rao bounds for coprime and other sparse arrays, which find more sources than sensors," *Digital Signal Process.*, vol. 61, pp. 43–61, Feb. 2017.
- [25] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.
- [26] G. H. Golub and C. F. Van Loan, *Matrix Computations, Third Edition*. Baltimore, MD: Johns Hopkins Univ. Press, 1996.
- [27] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Statistical Society, Series B*, vol. 58, no. 1, pp. 267–288, Feb. 1996.
- [28] P. Stoica, and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice Hall, 2005.
- [29] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Upper Saddle River, NJ: Prentice Hall, 1993.
- [30] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: A textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, May 2001.
- [31] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 172–175, March 1968.
- [32] M. Wang and A. Nehorai, "Coarrays, MUSIC, and the Cramér-Rao bound," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 933–946, Nov. 2016.
- [33] A. Koochakzadeh and P. Pal, "Cramér-Rao bounds for underdetermined source localization," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 919–923, May 2016.
- [34] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1987.

Muran Guo (S'18) received the B.S. degree from Harbin Engineering University, Harbin, China, in 2014. From 2016 to 2017, he was a visiting student at the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, USA. Currently, he is an MD-PhD student at the College of Information and Communication Engineering, Harbin Engineering University, Harbin, China. His research interests include direction-of-arrival estimation, compressive sensing, matrix completion, and Cramér-Rao bound analysis.

Yimin D. Zhang (SM'01) received his Ph.D. degree from the University of Tsukuba, Tsukuba, Japan, in 1988. He joined the faculty of the Department of Radio Engineering, Southeast University, Nanjing, China, in 1988. He served as a Director and Technical Manager at the Oriental Science Laboratory, Yokohama, Japan, from 1989 to 1995, and a Senior Technical Manager at the Communication Laboratory Japan, Kawasaki, Japan, from 1995 to 1997. He was a Visiting Researcher at the ATR Adaptive Communications Research Laboratories, Kyoto, Japan, from 1997 to 1998. From 1998 to 2015, he was with the Villanova University, Villanova, PA, where he was a Research Professor at the Center for Advanced Communications, and was the Director of the Wireless Communications and Positioning Laboratory and the Director of the Radio Frequency Identification (RFID) Laboratory. Since 2015, he has been with the Department of Electrical and Computer Engineering, College of Engineering, Temple University, Philadelphia, PA, where is an Associate Professor. Dr. Zhang's publications include more than 100 journal articles, 13 book chapters, and more than 200 conference papers in the areas of statistical signal and array processing, including compressive sensing, machine learning, convex optimization, nonstationary signal and timefrequency analysis, MIMO systems, radar imaging, directional finding, target localization and tracking, wireless and cooperative networks, and jammer suppression, with applications to radar, communications, and satellite navigation. He received the 2016 IET Radar, Sonar & Navigation Premium Award and the 2017 IEEE Aerospace and Electronic Systems Society Harry Rowe Mimno Award.

Dr. Zhang is an Associate Editor for the IEEE Transactions on Signal Processing and an editor for the Signal Processing journal. He was an Associate Editor for the IEEE Signal Processing Letters during 2006–2010, and an Associate Editor for the Journal of the Franklin Institute during 2007–2013. Dr. Zhang is a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society, and a Technical Co-Chair of the 2018 IEEE Sensor Array and Multichannel Signal Processing Workshop.

Tao Chen received the B.S. degree in Information Engineering, the M.S. degree in Signal and Information Processing, and the Ph.D. degree in Communication and Information Systems from Harbin Engineering University, Harbin, China, in 1997, 2001, and 2004, respectively. He is currently a Professor with the College of Information and Communication Engineering, Harbin Engineering University, Harbin, China. His research interests include direction-of-arrival estimation, wideband signal processing, and array signal processing.