Direction-of-Arrival Estimation for Coprime Array via Virtual Array Interpolation

Chengwei Zhou, *Member, IEEE*, Yujie Gu, *Senior Member, IEEE*, Xing Fan, Zhiguo Shi, *Senior Member, IEEE*, Guoqiang Mao, *Fellow, IEEE*, and Yimin D. Zhang, *Senior Member, IEEE*

***Abstract*—Coprime arrays can achieve an increased number of degrees-of-freedom by deriving the equivalent signals of a virtual array. However, most existing methods fail to utilize all information received by the coprime array due to the non-uniformity of the derived virtual array, resulting in an inevitable estimation performance loss. To address this issue, we propose a novel virtual array interpolation-based algorithm for coprime array direction-of-arrival (DOA) estimation in this paper. The idea of array interpolation is employed to construct a virtual uniform linear array such that all virtual sensors in the non-uniform virtual array can be utilized, based on which the atomic norm of the second-order virtual array signals is defined. By investigating the properties of virtual domain atomic norm, it is proved that the covariance matrix of the interpolated virtual array is related to the virtual measurements under the Hermitian positive semi-definite Toeplitz condition. Accordingly, an atomic norm minimization problem with respect to the equivalent virtual measurement vector is formulated to reconstruct the interpolated virtual array covariance matrix in a gridless manner, where the reconstructed covariance matrix enables off-grid DOA estimation. Simulation results demonstrate the performance advantages of the proposed DOA estimation algorithm for coprime arrays.**

***Index Terms*—Atomic norm, coprime array, direction-of-arrival estimation, gridless Toeplitz matrix reconstruction, virtual array interpolation.**

I. Introduction

DIRECTION-OF-ARRIVAL (DOA) estimation is a fundamental problem in array signal processing applications including radar, sonar, acoustics, speech, and wireless communications [2–7]. According to the Nyquist sampling theorem, uniform linear array (ULA) is the most commonly adopted array geometry for DOA estimation due to its regular structure and well-developed techniques. Recently, the systematically designed sparse arrays including coprime arrays and nested arrays have attracted noticeable attention, owing to their superior performance to the ULA's [8, 9]. In particular, the sparse arrays provide a larger array aperture than the ULA with the same number of sensors to improve the resolution. More importantly, the sparse arrays enable to break through the limitation of the degrees-of-freedom (DOFs). For instance, the coprime array can resolve up to $\mathcal{O}(MN)$ sources with only $M+N-1$ physical sensors. Motivated by these advantages, a series of efforts have been made to exploit the coprime array for DOA estimation [10], adaptive beamforming [11], and spectrum sensing [12].

In order to exploit the DOF superiority offered by the coprime array, an augmented virtual array can be constructed by computing the difference co-array of the coprime array, and its equivalent virtual array signal statistics can be implemented to perform DOA estimation [13–18]. Since the derived virtual array contains more virtual sensors, the limitation in DOFs constrained by the number of physical sensors is overcome. Among the virtual array-based methods, the spatial smoothing technique [17] is the most popular one, which requires a ULA-based signal model for DOA estimation. Unlike the nested arrays or the minimum redundancy arrays that yield a contiguous virtual array structure, the difference co-arrays derived from coprime arrays usually have holes, indicating that a coprime array is, in general, a partially augmentable array [19]. Hence, it is difficult to operate the derived signal statistics corresponding to the non-uniform virtual array. To address this issue, a common solution is to extract the maximum contiguous part from the non-uniform virtual array to form a virtual ULA. However, the virtual ULA generated in such a way is at the expense of reducing the achievable DOFs and virtual array aperture, since the discontiguous virtual sensors are discarded and, as a result, the information contained in the virtual array is not fully utilized. Although the generalized coprime array configurations [20, 21] and the coprime planar arrays [22] enable further expansion of the number of contiguous virtual sensors, the estimation performance loss is inevitable because the discontiguous virtual sensors are discarded.

To avoid the estimation performance loss, the sparse signal reconstruction algorithm [15] exploits all the non-uniform virtual array signals for DOA estimation, where the sparsity of the sources is considered. However, since all the equivalent virtual signals vectorized from the sample covariance matrix are included, the repeated virtual sensors lead to a high computational cost. In addition, the sparse signal reconstruction is formulated as a basis pursuit denoising problem, where the predefined sampling grids with a specific sampling interval results

The work of C. Zhou, X. Fan, and Z. Shi was supported in part by National Natural Science Foundation of China (No. 61772467), Zhejiang Provincial Natural Science Foundation of China (No. LR16F010002), 973 Project (No. 2015CB352503) and the Fundamental Research Funds for the Central Universities (No. 2017XZZX009-01). Part of the results in this paper was presented at IEEE 85th Vehicular Technology Conference, Sydney, Australia, June 2017 [1]. *(Corresponding authors: Y. Gu and Z. Shi.)*

C. Zhou is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China, and is also with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, Zhejiang 310027, China. (e-mail: zhouchw@zju.edu.cn).

Y. Gu and Y. D. Zhang are with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122, USA. (e-mail: guyujie@hotmail.com).

X. Fan and Z. Shi are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China. (e-mail: shizg@zju.edu.cn).

G. Mao is with School of Computing and Communications, University of Technology Sydney, Sydney, NSW 2007, Australia.

in an inherent basis mismatch for DOA estimation. In order to make a full use of the non-uniform virtual array signals, a high-resolution DOA estimation algorithm is proposed in [23] by utilizing multiple frequencies to fill the missing elements in the difference co-array, permitting the exploitation of the full DOFs offered by the derived non-uniform virtual array. More recently, a gridless DOA estimation algorithm is proposed by interpolating the coprime co-array through nuclear norm minimization [24], based on which the minimum number of virtual sensors required to complete the interpolated virtual array covariance matrix is reported in [25]. Although the abovementioned nuclear norm minimization-based algorithms are regularization-free, the retrieval of the covariance matrix is based on the matrix completion principle, indicating that the virtual signals corresponding to the non-uniform virtual array remain unchanged in the optimized covariance matrix. Since the virtual signals are obtained from the sample covariance matrix, the deviation caused by the finite snapshots may affect the estimation accuracy of the completed covariance matrix. The positive semi-definite (PSD) structure of the covariance matrix is introduced in [26] for the design of an interpolated covariance matrix using a nuclear norm, where a unified framework for analyzing the co-array extrapolation error is presented.

On the other hand, there has been an increasing interest in utilizing the Toeplitz structure of a ULA-based covariance matrix for both physical array-based and virtual array-based DOA estimation [27–31]. A regularization-free approach for matrix recovery from compressed sketches is based on the Toeplitz structure and the low-rank property [27], where the stability of the Toeplitz matrix estimation is investigated in terms of both denoising and prediction aspects. The gridless DOA estimation is implemented by reconstructing a low-rank Toeplitz covariance matrix of the signals received at a physical array [28], which offers a more accurate estimation performance than the non-Toeplitz approaches. In addition, the maximum-likelihood estimation of the Toeplitz-structured co-array covariance matrix is proposed for localizing more sources than sensors [29]. Apparently, the deterministic Toeplitz structure and the related statistical characteristics provide a more accurate approximation for covariance matrix estimation, promoting the feasibility of matrix recovery as well as improving the estimation performance.

In this paper, we propose a novel virtual array interpolation-based DOA estimation algorithm. By reconstructing the covariance matrix of the interpolated virtual array, all the derived virtual sensors are efficiently utilized. Specifically, we first interpolate the non-uniform virtual array to be a uniform one by filling in additional nominal sensors, such that all the virtual sensors in the derived virtual array can be fully utilized for DOA estimation. Then, we define the atomic norm of the interpolated virtual array signals based on an ideal assumption. Different from the atomic norm of the physical array received signals [32–34], the derived equivalent virtual array signals belong to the second-order statistics, which means that only the real-valued signal power is available rather than the complex-valued signal waveforms. Thus, we divide the equivalent virtual signals into multiple virtual measurements, and represent them by atoms through incorporating the phase offsets among virtual measurements. By investigating the properties of the defined atomic norm, the covariance matrix of the interpolated virtual array is proved to follow a Hermitian PSD Toeplitz structure, and its relationship with the equivalent virtual signals is further established. In practice, with the partial correlation observations as the reference, an atomic norm minimization problem is formulated to seek the atomic decomposition of interpolated virtual array signals with the minimal number of atoms for the reconstruction of the Toeplitz covariance matrix in a gridless manner, and the reconstructed covariance matrix is capable of estimating off-grid DOAs. Simulation results demonstrate the superiorities of the proposed DOA estimation algorithm in terms of resolution, estimation accuracy, achievable DOFs, and computational complexity.

The main contributions of this paper can be summarized as follows:

- We define the atomic norm of the second-order virtual signals corresponding to an interpolated virtual ULA, which contains all the virtual sensors in the derived non-uniform virtual array.
- We relate the virtual signals to the Toeplitz covariance matrix of the interpolated virtual array according to the properties of the proposed atomic norm, and transform the atomic norm minimization problem to the selection of a virtual measurement vector.
- We reconstruct a Hermitian PSD Toeplitz covariance matrix corresponding to the interpolated virtual array in a gridless manner, and provide the theoretical performance analyses.

The rest of this paper is organized as follows. In Section II, we present the coprime array signal model. We then propose a virtual array interpolation-based DOA estimation algorithm in Section III, and present theoretical analysis in Section IV. We demonstrate the simulation results in Section V. Finally, we make our conclusions in Section VI.

Notations: We use lower-case and upper-case boldface characters to respectively represent vectors and matrices throughout this paper. The superscripts $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^*$ denote the transpose, conjugate transpose, and complex conjugation, respectively. $(\cdot)^{-1}$, $\mathrm{Tr}(\cdot)$, and $\mathrm{rank}(\cdot)$ respectively denote the inverse, the trace, and the rank of a matrix. The notation $\mathrm{E}[\cdot]$ denotes the statistical expectation, $\mathrm{vec}(\cdot)$ stands for the vectorization operator that sequentially stacks each column of a matrix, and $\mathrm{diag}(\cdot)$ represents a diagonal matrix with the corresponding elements on its diagonal. $\otimes$ and $\circ$ denote the Kronecker product and the Hadamard product, respectively. The curled inequality symbol $\succeq$ denotes matrix inequality. $\|\cdot\|_2$ and $\|\cdot\|_F$ denote the Euclidean norm and Frobenius norm, respectively. $|\mathcal{S}|$ represents the cardinality of a set $\mathcal{S}$. For a given vector $\boldsymbol{a}$, $\langle \boldsymbol{a} \rangle_\ell$ stands for the ℓ-th element in $\boldsymbol{a}$. Finally, $\boldsymbol{I}$ denotes the identity matrix with an appropriate dimension.

II. Coprime Array Signal Model

Consider a pair of sparse ULAs as shown in Fig. 1(a), where the top sparse ULA consists of M sensors spaced Nd apart,

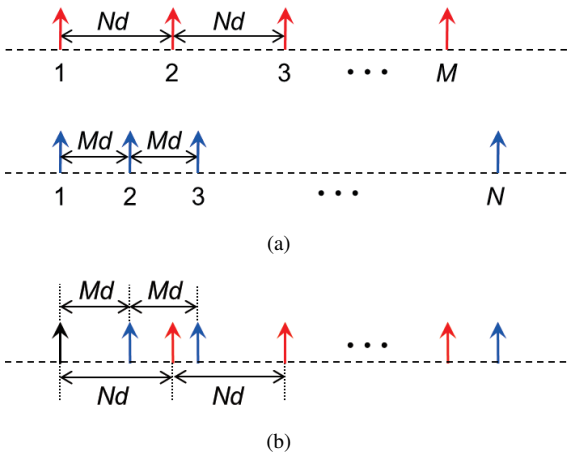


Fig. 1. Illustration of a coprime array. (a) A coprime pair of sparse ULAs; (b) Coprime array configuration.

whereas the bottom one consists of N sensors spaced Md apart. Here, M and N are coprime integers, and d is a half-wavelength, i.e., $d = \lambda/2$. Aligning these two sparse ULAs with the first sensor as the reference yields a coprime array configuration as shown in Fig. 1(b). Due to the coprimality, the sensors do not overlap except the reference one; hence, the coprime array consists of $M + N - 1$ sensors in total.

Assuming K far-field, narrowband and uncorrelated sources impinging from directions $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_K]^{\mathrm{T}}$, the received signals of the coprime array can be modeled as

$$\boldsymbol{x}(t) = \sum_{k=1}^{K} \boldsymbol{a}(\theta_k) s_k(t) + \boldsymbol{n}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{1}$$

where $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_K)] \in \mathbb{C}^{(M+N-1)\times K}$ denotes the coprime array steering matrix, $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_K(t)]^{\mathrm{T}}$ denotes the signal waveform vector, and $\boldsymbol{n}(t) \sim \mathcal{CN}(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I})$ represents the independent and identically distributed (i.i.d.) zero-mean additive white Gaussian noise vector. Here, σ_n^2 denotes the noise power. The k-th column of $\boldsymbol{A}$ represents the steering vector of the k-th source signal as

$$\boldsymbol{a}(\theta_k) = \left[1, e^{-j\frac{2\pi}{\lambda} u_2 d \sin(\theta_k)}, \cdots, e^{-j\frac{2\pi}{\lambda} u_{M+N-1} d \sin(\theta_k)}\right]^{\mathrm{T}}, \tag{2}$$

where $u_l d$, $l \in \{2, 3, \cdots, M+N-1\}$, denotes the position of the l-th sensor in the coprime array, and $j = \sqrt{-1}$ is the imaginary unit.

The covariance matrix of the coprime array received signals $\boldsymbol{x}(t)$ can be expressed as

$$\boldsymbol{R}_{\boldsymbol{x}} = \mathrm{E}\left[\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{H}}(t)\right] = \sum_{k=1}^{K} p_k \boldsymbol{a}(\theta_k)\boldsymbol{a}^{\mathrm{H}}(\theta_k) + \sigma_n^2 \boldsymbol{I}, \tag{3}$$

where p_k denotes the power of the k-th source signal. Considering the fact that the exact covariance matrix $\boldsymbol{R}_{\boldsymbol{x}}$ is unavailable in practice, it can be approximated by its sample version

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{H}}(t), \tag{4}$$

where T denotes the number of snapshots. Here, the sample covariance matrix $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$ is the maximum-likelihood estimator of $\boldsymbol{R}_{\boldsymbol{x}}$, and it will converges to $\boldsymbol{R}_{\boldsymbol{x}}$ when T trends to infinity under the stationarity and ergodicity assumptions [35].

III. Virtual Array Interpolation-based DOA Estimation Algorithm

In this section, we elaborate a novel virtual array interpolation-based DOA estimation algorithm. First, the derived non-uniform virtual array is transformed to a virtual ULA through array interpolation. The atomic norm of multiple virtual measurements is then investigated based on an ideal assumption, from which the relationship between the second-order equivalent virtual signals and the covariance matrix of the interpolated virtual array is established. Finally, based on the partial correlation observations corresponding to the derived non-uniform virtual array, we formulate an atomic norm minimization problem for Toeplitz covariance matrix reconstruction, which enables to estimate off-grid DOAs with the utilization of the entire information contained in the derived non-uniform virtual array.

A. Array Interpolation for Virtual ULA

Sparse arrays such as a coprime array enable an increased number of DOFs by handling the equivalent virtual array signals. By vectorizing the covariance matrix $\boldsymbol{R}_{\boldsymbol{x}}$, we have the equivalent virtual array signals as

$$\boldsymbol{y}_v = \mathrm{vec}(\boldsymbol{R}_{\boldsymbol{x}}) = \boldsymbol{A}_v \boldsymbol{p} + \sigma_n^2 \boldsymbol{i}, \tag{5}$$

where $\boldsymbol{A}_v = \left[\boldsymbol{a}^*(\theta_1) \otimes \boldsymbol{a}(\theta_1), \boldsymbol{a}^*(\theta_2) \otimes \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}^*(\theta_K) \otimes \boldsymbol{a}(\theta_K)\right] \in \mathbb{C}^{(M+N-1)^2 \times K}$, $\boldsymbol{p} = [p_1, p_2, \cdots, p_K]^{\mathrm{T}}$, and $\boldsymbol{i} = \mathrm{vec}(\boldsymbol{I})$. Here, the steering matrix $\boldsymbol{A}_v$ corresponds to an augmented array with the virtual sensors located at $\mathcal{S}d$, where

$$\mathcal{S} = \{u_m - u_n \,|\, m, n = 0, 1, \cdots, M+N-1\}. \tag{6}$$

By removing the repeated elements in $\mathcal{S}$, a proper subset $\mathcal{S}_{\mathcal{V}} \subsetneq \mathcal{S}$ is obtained, and the virtual sensors located at $\mathcal{S}_{\mathcal{V}} d$ constitute a virtual array. For the systematic coprime array configuration shown in Fig. 1(b), the elements in $\mathcal{S}_{\mathcal{V}}$ can also be obtained by picking all the unique values from its difference co-array as

$$\mathcal{S}_{\mathcal{V}} \subsetneq \{\pm(Mn - Nm) \,|\, m = 0, 1, \cdots, M-1, \; n = 0, 1, \cdots, N-1\}. \tag{7}$$

Accordingly, the equivalent virtual signals of the derived virtual array $\mathcal{S}_{\mathcal{V}}$ can be obtained by selecting the corresponding elements from $\boldsymbol{y}_v$ as

$$\bar{\boldsymbol{y}}_v = \bar{\boldsymbol{A}}_v \boldsymbol{p} + \sigma_n^2 \bar{\boldsymbol{i}}, \tag{8}$$

where $\bar{\boldsymbol{A}}_v \in \mathbb{C}^{|\mathcal{S}_{\mathcal{V}}| \times K}$ denotes the steering matrix of the derived virtual array $\mathcal{S}_{\mathcal{V}}$, and $\bar{\boldsymbol{i}}$ denotes a sub-vector of $\boldsymbol{i}$ corresponding to the selected positions of the virtual sensors of $\mathcal{S}_{\mathcal{V}}$ in $\mathcal{S}$.

Since the coprime array is a partially augmentable array, its difference co-array contains several missing elements, which are referred to as holes [8, 19], leading to

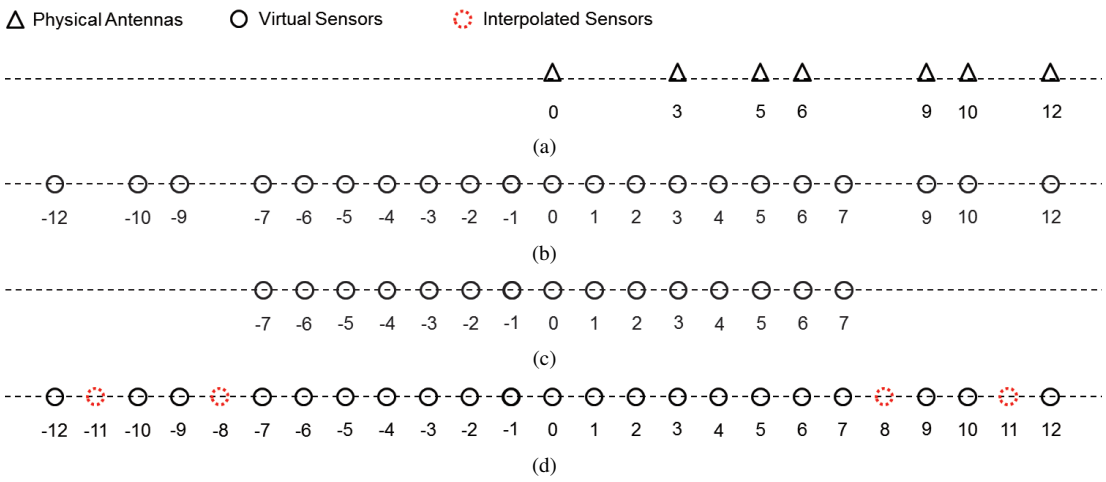


Fig. 2. Illustration of various array representations with an example of $M = 3$ and $N = 5$. (a) Coprime array; (b) $\mathcal{S}_{\mathcal{V}}$, virtual array derived from the difference co-array of the coprime array; (c) $\mathcal{S}_{\mathcal{C}}$, contiguous part of the virtual array; (d) $\mathcal{S}_{\mathcal{I}}$, interpolated virtual array.

a non-uniform virtual array geometry. To have an intuitive understanding, we illustrate a coprime array configuration with $M = 3$ and $N = 5$ in Fig. 2(a). Obviously, we have the derived non-uniform virtual array $\mathcal{S}_{\mathcal{V}} = \{-12, -10, -9, -7, -6, -5, \cdots, 5, 6, 7, 9, 10, 12\}$ as depicted in Fig. 2(b), where the missing elements $\{-11, -8, 8, 11\}$ are the holes. Since the non-uniformity results in difficulties in the subsequent statistical signal processing, a common solution is to pick up the maximum contiguous part $\mathcal{S}_{\mathcal{C}}$ from $\mathcal{S}_{\mathcal{V}}$ while discarding the discontiguous part $\mathcal{S}_{\mathcal{V}} - \mathcal{S}_{\mathcal{C}}$ [13, 17]. Although the obtained virtual ULA $\mathcal{S}_{\mathcal{C}}$ as shown in Fig. 2(c) is easy to operate, part of the information received by the coprime array is apparently ignored due to the discarded discontiguous virtual sensors, resulting in the performance degradation.

In order to make full use of the information involved in the non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$, we introduce the idea of array interpolation to fill in the holes in $\mathcal{S}_{\mathcal{V}}$ with nominal sensors. As such, a virtual ULA $\mathcal{S}_{\mathcal{I}}$ with $2M(N-1)+1$ sensors is constructed as shown in Fig. 2(d), where all the virtual sensors in $\mathcal{S}_{\mathcal{V}}$ are included. Here, the nominal sensor contains two meanings: First, the array interpolation is implemented in the virtual domain, and the nominal sensors exist in a mathematical sense rather than in a physical existence; Second, based on the fact that there is no *a priori* information regarding the equivalent virtual signals corresponding to the holes, we may naturally regard the interpolated nominal sensors as the nonfunctional sensors and set the corresponding virtual signals in these positions to zero. Hence, the $|\mathcal{S}_{\mathcal{I}}|$-dimensional interpolated virtual array signals $\boldsymbol{y}_{\mathcal{I}}$ can be initialized as

$$[\boldsymbol{y}_{\mathcal{I}}]_i = \begin{cases} [\bar{\boldsymbol{y}}_v]_i, & i \in \mathcal{S}_{\mathcal{V}}, \\ 0, & i \in \mathcal{S}_{\mathcal{I}} - \mathcal{S}_{\mathcal{V}}, \end{cases} \tag{9}$$

where $[\cdot]_i$ denotes the virtual signal of the virtual sensor at position id. Clearly, there are two kinds of sensors in the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$, namely, virtual sensors and nominal sensors. The virtual signals in $\boldsymbol{y}_{\mathcal{I}}$ corresponding to the derived virtual sensors in the non-uniform virtual array are consistent with those in $\bar{\boldsymbol{y}}_v$, whereas the remaining elements corresponding to the nominal sensors are zero. The *de facto* ULA-based virtual signals $\boldsymbol{y}_{\mathcal{I}}$ enable the potential utilization of the mature DOA estimation techniques designed for the ULA with all the information contained in $\bar{\boldsymbol{y}}_v$. Therefore, we prefer to perform DOA estimation using the equivalent signals of the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$ rather than using those of the contiguous virtual array $\mathcal{S}_{\mathcal{C}}$ as mentioned earlier. To effectively utilize the interpolated virtual ULA, it is necessary to retrieve the unknown virtual signals corresponding to the interpolated nominal sensors, which are represented as zeros in the initialization of virtual array interpolation (9).

B. Atomic Norm of Multiple Virtual Measurements

In order to overcome the basis mismatch problem, the gridless methods provide a novel insight for statistical signal processing, where the atomic norm is one of the most important mathematical tools to utilize the signal characteristics without pre-defined sampling grids. The analysis for the atomic norm in the virtual domain begins with an ideal scenario assumption, where the interpolated virtual array signals are assumed to be accurate. More specifically, the number of available snapshots for calculating the sample covariance matrix trends to infinity, and the coprime array received signals are noise-free. Besides, the virtual signals of the interpolated nominal sensors in $\mathcal{S}_{\mathcal{I}}$ are assumed to be precise rather than to be zeros as we did in (9). Similar to (5), the ideal virtual signals of the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$ can be modeled as

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{v}(\theta_k) p_k = \boldsymbol{V}\boldsymbol{p}, \tag{10}$$

where $\boldsymbol{V} = [\boldsymbol{v}(\theta_1), \boldsymbol{v}(\theta_2), \cdots, \boldsymbol{v}(\theta_K)] \in \mathbb{C}^{|\mathcal{S}_{\mathcal{I}}| \times K}$ denotes the steering matrix of the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$. Although the interpolated virtual array signal $\boldsymbol{y}$ modeled in (10) has a similar structure as the coprime array received signals $\boldsymbol{x}(t)$ in a noise-free case, $\boldsymbol{y}$ is actually a *second-order* statistics derived from the correlation of the *first-order* coprime array received signals. While $\boldsymbol{y}$ behaves like a single snapshot, the rank-

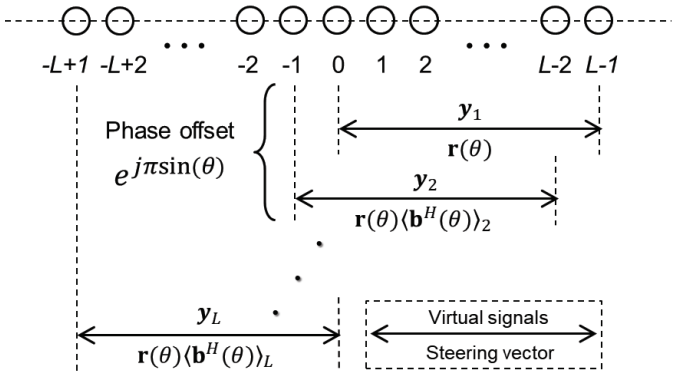


Fig. 3. Phase offsets among the virtual measurements of each sub-array.

deficiency problem of the corresponding correlation statistics makes it difficult to identify multiple sources.

In view of this, we divide the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$ into $L = (|\mathcal{S}_{\mathcal{I}}| + 1)/2$ overlapping sub-arrays with L contiguous virtual sensors for each sub-array as shown in Fig. 3. Since $\mathcal{S}_{\mathcal{I}}$ is symmetric to the zeroth position, the value of L is always an integer. Accordingly, the equivalent virtual signals of each sub-array can be obtained by dividing vector $\boldsymbol{y}$ into L sub-vectors $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_L\}$ as

$$\boldsymbol{y}_\ell = \sum_{k=1}^{K} \boldsymbol{v}_\ell(\theta_k) p_k = \boldsymbol{V}_\ell \boldsymbol{p}, \quad \ell = 1, 2, \cdots, L, \tag{11}$$

where $\boldsymbol{V}_\ell = [\boldsymbol{v}_\ell(\theta_1), \boldsymbol{v}_\ell(\theta_2), \cdots, \boldsymbol{v}_\ell(\theta_K)] \in \mathbb{C}^{L\times K}$. Here,

$$\boldsymbol{v}_\ell(\theta_k) = \Big[e^{-j\pi v_{L-\ell+1}\sin(\theta_k)}, e^{-j\pi v_{L-\ell+2}\sin(\theta_k)}, \cdots, e^{-j\pi v_{2L-\ell}\sin(\theta_k)}\Big]^{\mathrm{T}} \tag{12}$$

denotes the steering vector of the ℓ-th virtual sub-array corresponding to the k-th signal, where v_i denotes the i-th element in $\mathcal{S}_{\mathcal{I}}$. Collecting these L sub-vectors, $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_L] \in \mathbb{C}^{L\times L}$ yields the equivalent virtual signals of the interpolated virtual sub-array in an L virtual measurements manner, and is referred to as the virtual measurements in the sequel.

To analyze the atomic norm of $\boldsymbol{Y}$, an atom with the same dimension is required to represent $\boldsymbol{Y}$. We would like to emphasize the striking differences between the virtual domain atoms and those in the physical domain [32–34]. On one hand, for each source θ_k, the virtual signals $\boldsymbol{y}_\ell$ formulated in (11) belong to the second-order statistics, containing the real-valued source power p_k, unlike the first-order received signals $\boldsymbol{x}(t)$ containing the complex-valued signal waveform $s_k(t)$. On the other hand, for each virtual measurement, the displacement among the L sub-arrays of the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$ creates the phase offsets, which can be utilized to characterize the difference among the L sub-arrays in the virtual domain. While $\boldsymbol{Y}$ contains the second-order virtual measurements received by the sub-arrays as illustrated in Fig. 3, a series of atoms can be formulated to represent $\boldsymbol{Y}$ based on the steering vector of a certain sub-array and the phase offsets among the L sub-arrays. In particular, we set the first sub-array of $\mathcal{S}_{\mathcal{I}}$ shown in Fig. 3 as the reference virtual array, whose steering vector can be calculated by setting $\ell = 1$ in (12) as

$$\boldsymbol{r}(\theta) = \boldsymbol{v}_1(\theta) = \Big[e^{-j\pi v_L \sin(\theta)}, e^{-j\pi v_{L+1}\sin(\theta)}, \cdots, e^{-j\pi v_{2L-1}\sin(\theta)}\Big]^{\mathrm{T}}. \tag{13}$$

Accordingly, the phase offsets between the L sub-arrays and the reference virtual array can be expressed as

$$\boldsymbol{b}(\theta) = \Big[1, e^{-j\pi\sin(\theta)}, \cdots, e^{-j\pi(L-1)\sin(\theta)}\Big]^{\mathrm{T}}, \tag{14}$$

and the steering vector of the ℓ-th sub-array can be represented as

$$\boldsymbol{v}_\ell(\theta) = \boldsymbol{r}(\theta)\langle\boldsymbol{b}^{\mathrm{H}}(\theta)\rangle_\ell, \quad \ell = 1, 2, \cdots, L. \tag{15}$$

Therefore, the virtual measurements $\boldsymbol{Y}$ can be regarded as L second-order virtual signal snapshots received by the reference virtual array, behaving like the snapshots of the first-order signal waveforms. The phase offsets contained in $\boldsymbol{b}(\theta)$ characterize the difference among each virtual measurement. As such, $\boldsymbol{Y}$ contains all the information in the interpolated virtual array since all the elements in $\boldsymbol{y}$ are included.

Based on $\boldsymbol{r}(\theta)$ and $\boldsymbol{b}(\theta)$, an atom for representing $\boldsymbol{Y}$ can be defined as

$$\boldsymbol{B}(\theta) = \boldsymbol{r}(\theta)\boldsymbol{b}^{\mathrm{H}}(\theta), \tag{16}$$

where $\boldsymbol{B}(\theta) \in \mathbb{C}^{L\times L}$ with $\theta \in [-90^\circ, 90^\circ]$, and the corresponding atom set is

$$\mathcal{A} = \{\boldsymbol{B}(\theta) | \, \theta \in [-90^\circ, 90^\circ]\}. \tag{17}$$

With the specifically defined atom set $\mathcal{A}$, the smallest number of atoms for the representation of the virtual measurements $\boldsymbol{Y}$ can be defined as

$$\|\boldsymbol{Y}\|_{\mathcal{A},0} = \inf_{K}\left\{\boldsymbol{Y} = \sum_{k=1}^{K} p_k \boldsymbol{B}(\theta_k),\ p_k \geq 0\right\}, \tag{18}$$

where $\inf$ denotes the infimum. While performing the atomic decomposition of $\boldsymbol{Y}$ with the minimal number of atoms, i.e., minimizing (18), is an NP-hard problem, we introduce the atomic norm convex relaxation as [1]

$$\begin{aligned}\|\boldsymbol{Y}\|_{\mathcal{A}} &= \inf\{h > 0 : \boldsymbol{Y} \in h\,\mathrm{conv}(\mathcal{A})\} \\ &= \inf\left\{\sum_k p_k \,\middle|\, \boldsymbol{Y} = \sum_k p_k \boldsymbol{B}(\theta_k),\ p_k \geq 0\right\},\end{aligned} \tag{19}$$

where $\mathrm{conv}(\mathcal{A})$ denotes the convex hull of the atom set $\mathcal{A}$. Further, we have the following theorem for the equivalent representation of $\|\boldsymbol{Y}\|_{\mathcal{A}}$:

Theorem 1: The atomic norm of the virtual measurements $\boldsymbol{Y}$ defined in (19) can be represented in an equivalent semi-definite programming (SDP) form as

$$\|\boldsymbol{Y}\|_{\mathcal{A}} = \inf_{\boldsymbol{z}\in\mathbb{C}^L, \boldsymbol{W}\in\mathbb{C}^{L\times L}}\left\{\frac{1}{2L}\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) + \frac{1}{2L}\mathrm{Tr}(\boldsymbol{W}) \,\middle|\, \begin{bmatrix} \mathcal{T}(\boldsymbol{z}) & \boldsymbol{Y} \\ \boldsymbol{Y}^{\mathrm{H}} & \boldsymbol{W}\end{bmatrix} \succeq \boldsymbol{0}\right\}, \tag{20}$$

[1]Note that the atomic norm is formally defined as the gauge function of $\mathcal{A}$ [36]. Since the atom set $\mathcal{A}$ for representing the multiple virtual measurements $\boldsymbol{Y}$ is not centrally symmetric, our analysis is based on the underlying convex geometry of $\mathrm{conv}(\mathcal{A})$, and $\|\cdot\|_{\mathcal{A}}$ is referred to as the atomic norm of the set $\mathcal{A}$ with an abuse of terminology [37].

where $\mathcal{T}(\boldsymbol{z})$ denotes a Hermitian Toeplitz matrix with vector $\boldsymbol{z}$ as its first column.

Proof: See Appendix A. ■

Based on the atomic norm of the virtual measurements $\boldsymbol{Y}$ defined in (19) and its equivalent SDP form in (20), we have the following two corollaries regarding the properties of $\mathcal{T}(\boldsymbol{z})$, whose conclusions can be utilized for the subsequent DOA estimation.

Corollary 1: Denoting $\boldsymbol{z}^\star$ as the optimum solution to (20), the Hermitian PSD Toeplitz matrix $\mathcal{T}(\boldsymbol{z}^\star)$ is the covariance matrix of signals received by the reference virtual array, and can be obtained from the principal square root of $\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{H}}$.

Proof: According to (52), the Hermitian PSD Toeplitz matrix $\mathcal{T}(\boldsymbol{z}^\star)$ can be represented as

$$\left[\mathcal{T}\left(\boldsymbol{z}^\star\right)\right]^2 = \left[\sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k)\right]^2. \tag{21}$$

Since the reference virtual array consists of the virtual sensors located at $\{0, d, 2d, \cdots, (L-1)d\}$ as shown in Fig. 3, its corresponding steering vector can be described as

$$\boldsymbol{r}(\theta_k) = \left[1, e^{-j\pi\sin(\theta_k)}, \cdots, e^{-j\pi(L-1)\sin(\theta_k)}\right]^{\mathrm{T}}, \tag{22}$$

which coincides with the phase offset vector $\boldsymbol{b}(\theta_k)$ in (14), i.e., $\boldsymbol{r}(\theta_k) = \boldsymbol{b}(\theta_k)$. Thus, (21) can be rewritten as

$$\begin{aligned}\left[\mathcal{T}\left(\boldsymbol{z}^\star\right)\right]^2 &= \left(\sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k)\right)\left(\sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k)\right)^{\mathrm{H}} \\ &= \underbrace{\sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{b}^{\mathrm{H}}(\theta_k)}_{\boldsymbol{Y}} \underbrace{\sum_{k=1}^{K} p_k \boldsymbol{b}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k)}_{\boldsymbol{Y}^{\mathrm{H}}}.\end{aligned} \tag{23}$$

Therefore, the Hermitian PSD Toeplitz matrix $\mathcal{T}(\boldsymbol{z}^\star)$ is the covariance matrix of the signals received by the reference virtual array, containing all the information of the interpolated virtual array. In addition, it is clear from (23) that the interpolated virtual array covariance matrix $\mathcal{T}(\boldsymbol{z}^\star)$ can be obtained from the principal square root of $\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{H}}$. ■

Corollary 2: The first column of the Hermitian PSD Toeplitz matrix $\mathcal{T}(\boldsymbol{z}^\star)$ is equivalent to the virtual signals (second-order statistics) of the reference virtual array.

Proof: Since $\mathcal{T}(\boldsymbol{z}^\star)$ satisfies a Hermitian Toeplitz structure, its first column $\boldsymbol{z}^\star$ can be calculated as

$$\boldsymbol{z}^\star = \sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\left\langle \boldsymbol{b}(\theta_k)\right\rangle_1 = \sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k). \tag{24}$$

Based on the equivalence relationship between $\boldsymbol{r}(\theta_k)$ and $\boldsymbol{v}_1(\theta_k)$ as established in (13), it is clear that $\boldsymbol{z}^\star$ is equal to the virtual signals of the reference virtual array $\boldsymbol{y}_1$ as defined in (11), thus verifying the corollary. ■

The defined atomic norm (19) provides a gridless model tailored for the second-order virtual measurements, and the associated corollaries and the equivalent SDP form reveal the important relationship between the interpolated virtual array covariance matrix and the second-order virtual signals. Since the involved directional parameters are continuously parameterized with an atomic norm, the incorporation of the abovementioned techniques enables to estimate off-grid DOAs in the virtual domain.

C. DOA Estimation via Toeplitz Matrix Reconstruction

Corollary 1 provides a relationship between the virtual measurements $\boldsymbol{Y}$ and the Hermitian Toeplitz matrix $\mathcal{T}(\boldsymbol{z}^\star)$ containing the signal information of the interpolated virtual array, while Corollary 2 relates the reference virtual array signals to the interpolated virtual array covariance matrix $\mathcal{T}(\boldsymbol{z}^\star)$, such that $\mathcal{T}(\boldsymbol{z}^\star)$ can be directly constructed from $\boldsymbol{y}_1$. However, both $\boldsymbol{Y}$ and $\mathcal{T}(\boldsymbol{z}^\star)$ are derived based on an ideal assumption, i.e., infinite snapshots, noise-free signal model, and the accurate virtual signals for the interpolated nominal sensors. According to the initial array interpolation process (9), the virtual signals corresponding to the interpolated nominal sensors in $\mathcal{S}_{\mathcal{I}}$ are initialized to be zeros in $\boldsymbol{y}_{\mathcal{I}}$. Also, the number of snapshots is finite, and the noise term in (5) is non-negligible. Therefore, we focus on the reconstruction of the interpolated virtual array covariance matrix $\mathcal{T}(\boldsymbol{z}^\star)$ based on the initialized interpolated virtual array signals $\boldsymbol{y}_{\mathcal{I}}$ under the gridless framework of the defined atomic norm for virtual measurements (20).

In practice, the virtual signals of the L sub-arrays in Fig. 3 are obtained by collecting the corresponding elements in $\boldsymbol{y}_{\mathcal{I}}$ (9) rather than those in $\boldsymbol{y}$ (10). Similar to (11), by dividing the initialized interpolated virtual array signals $\boldsymbol{y}_{\mathcal{I}}$ into L sub-vectors as $\tilde{\boldsymbol{y}}_\ell, \ell = 1, 2, \cdots, L$, we have the multiple virtual measurements as $\tilde{\boldsymbol{Y}} = [\tilde{\boldsymbol{y}}_1, \tilde{\boldsymbol{y}}_2, \cdots, \tilde{\boldsymbol{y}}_L]$. According to Corollary 1, the following equation holds

$$\tilde{\boldsymbol{R}}_v^2 = \sum_{\ell=1}^{L} \tilde{\boldsymbol{y}}_\ell\, \tilde{\boldsymbol{y}}_\ell^{\mathrm{H}} = \tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^{\mathrm{H}}, \tag{25}$$

where $\tilde{\boldsymbol{R}}_v \in \mathbb{C}^{L\times L}$ denotes the reference virtual array covariance matrix, containing all information in $\boldsymbol{y}_{\mathcal{I}}$. Since there exist several zero elements in $\tilde{\boldsymbol{Y}}$ resulting from the interpolated nominal sensors, the deviation accumulates for each element in $\tilde{\boldsymbol{R}}_v^2$ during the summation process in (25). Hence, $\tilde{\boldsymbol{R}}_v$ cannot be directly calculated from the principal square root of $\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^{\mathrm{H}}$. Encouragingly, the relationship established in Corollary 2 enables to directly construct $\tilde{\boldsymbol{R}}_v$ according to the initialized virtual signals of the reference virtual array as

$$\tilde{\boldsymbol{R}}_v = \begin{bmatrix} \langle \boldsymbol{y}_{\mathcal{I}}\rangle_L & \langle \boldsymbol{y}_{\mathcal{I}}\rangle_{L+1}^* & \cdots & \langle \boldsymbol{y}_{\mathcal{I}}\rangle_{2L-1}^* \\ \langle \boldsymbol{y}_{\mathcal{I}}\rangle_{L+1} & \langle \boldsymbol{y}_{\mathcal{I}}\rangle_L & \cdots & \langle \boldsymbol{y}_{\mathcal{I}}\rangle_{2L-2}^* \\ \vdots & \vdots & \ddots & \vdots \\ \langle \boldsymbol{y}_{\mathcal{I}}\rangle_{2L-1} & \langle \boldsymbol{y}_{\mathcal{I}}\rangle_{2L-2} & \cdots & \langle \boldsymbol{y}_{\mathcal{I}}\rangle_L \end{bmatrix}, \tag{26}$$

i.e., $\tilde{\boldsymbol{R}}_v = \mathcal{T}\left(\tilde{\boldsymbol{y}}_1\right)$. Due to the fact that $\boldsymbol{y}_{\mathcal{I}}$ contains several zero elements, the diagonals in $\tilde{\boldsymbol{R}}_v$ corresponding to the positions of holes in the reference virtual array are zeros. It is clear that the Hermitian Toeplitz matrix $\tilde{\boldsymbol{R}}_v$ formulated in (26) is the reference virtual array covariance matrix because it satisfies the equivalent condition in (25). Considering that the interpolated virtual array $\mathcal{S}_{\mathcal{I}}$ is symmetric to the zeroth position, the complex-valued virtual signals of the symmetrical

pair are mutually conjugate. Therefore, on the premise of Hermitian Toeplitz structure, the conclusion in Corollary 2 is consistent with the relationship between the coprime co-array covariance matrix and the second-order virtual signals established in [38].

On the other hand, since the additive noise in (1) is independent of the source signals, its correlation term $\sigma_n^2\boldsymbol{I}$ does not destruct the Hermitian Toeplitz structure of the interpolated virtual array covariance matrix. As such, this special covariance matrix structure can be exploited as *a priori* information for DOA estimation, permitting us focusing on the reference virtual array to investigate the covariance matrix corresponding to the interpolated virtual array. Accordingly, we define a binary vector $\boldsymbol{g} \in \mathbb{R}^L$ to distinguish nominal sensors and virtual sensors in the reference virtual array, whose element is 0 for the nominal sensor positions and 1 otherwise.

The reconstruction of the interpolated virtual array covariance matrix $\mathcal{T}(\boldsymbol{z}^\star)$ starts with the analysis of the atomic norm of ideal virtual measurements $\boldsymbol{Y}$, such that the parameters contained in $\boldsymbol{Y}$ can be continuously represented. According to the definition of $\|\boldsymbol{Y}\|_{\mathcal{A}}$ in (19), a natural objective to describe $\boldsymbol{Y}$ is performing the atomic decomposition of $\boldsymbol{Y}$ with the minimal number of atoms, i.e., minimizing $\|\boldsymbol{Y}\|_{\mathcal{A}}$. Alternatively, the relationships between $\boldsymbol{Y}$, $\boldsymbol{z}^\star$, and $\mathcal{T}(\boldsymbol{z}^\star)$ revealed in Corollaries 1 and 2 enable to reconstruct the interpolated virtual array covariance matrix $\mathcal{T}(\boldsymbol{z}^\star)$ by focusing on $\boldsymbol{z}^\star$, since the Hermitian Toeplitz covariance matrix depends on its first column. Based on the relationship established in (24), $\boldsymbol{z}^\star$ is equivalent to the virtual signals of the reference virtual array, and its corresponding atom set can thus be defined with respect to the steering vector of the reference virtual array as

$$\mathcal{A}_r = \left\{\boldsymbol{r}(\theta)|\theta \in [-90^\circ, 90^\circ]\right\}. \tag{27}$$

Note that the atom set $\mathcal{A}_r$ for $\boldsymbol{z}^\star$ relates to the atom set $\mathcal{A}$ for $\boldsymbol{Y}$ defined in (17) by setting the phase offset term in $\mathcal{A}$ as $\langle \boldsymbol{b}(\theta)\rangle_1 = 1$, which is consistent with the fact revealed in Corollary 2 that $\boldsymbol{z}^\star$ is actually the first column of matrix $\boldsymbol{Y}$. Accordingly, the atomic norm of the variable $\boldsymbol{z}$ can be represented as

$$\|\boldsymbol{z}\|_{\mathcal{A}_r} = \inf\left\{\sum_k p_k : \boldsymbol{z} = \sum_k p_k \boldsymbol{r}(\theta_k),\ p_k \geq 0\right\}. \tag{28}$$

By comparing the definitions of $\|\boldsymbol{Y}\|_{\mathcal{A}}$ in (19) and $\|\boldsymbol{z}\|_{\mathcal{A}_r}$ in (28), it is demonstrated that minimizing $\|\boldsymbol{Y}\|_{\mathcal{A}}$ can be equivalently transformed to minimizing $\|\boldsymbol{z}\|_{\mathcal{A}_r}$ by defining the atom set $\mathcal{A}_r$ based on $\mathcal{A}$. Therefore, we focus on minimizing the atomic norm of $\boldsymbol{z}$ for the reconstruction of the interpolated virtual array covariance matrix, which can be viewed as performing the atomic decomposition of $\boldsymbol{z}$ with the minimal number of atoms according to (28).

Taking the matrix $\tilde{\boldsymbol{R}}_v$ in (26), which contains the partial correlation observations collected from the derived non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$, as the reference, an atomic norm minimization problem can be formulated for Toeplitz covariance matrix reconstruction as

$$\begin{aligned} \min_{\boldsymbol{z}\in\mathbb{C}^L} \quad & \|\boldsymbol{z}\|_{\mathcal{A}_r} \\ \text{subject to} \quad & \left\|\mathcal{T}(\boldsymbol{z}) \circ \boldsymbol{G} - \tilde{\boldsymbol{R}}_v\right\|_F^2 \leq \varepsilon, \\ & \mathcal{T}(\boldsymbol{z}) \succeq \boldsymbol{0}, \end{aligned} \tag{29}$$

where $\boldsymbol{G} = \boldsymbol{g}\boldsymbol{g}^{\mathrm{T}} \in \mathbb{R}^{L\times L}$ is a binary matrix to distinguish the zero (interpolated) and non-zero (derived) statistics in $\tilde{\boldsymbol{R}}_v$ due to the initial virtual array interpolation, such that the non-zero elements in $\tilde{\boldsymbol{R}}_v$ are comparable with the corresponding elements in the reconstructed covariance matrix $\mathcal{T}(\boldsymbol{z})$. Here, ε is a threshold to restrict the fitting error between the non-zero elements in $\tilde{\boldsymbol{R}}_v$ and those in the reconstructed covariance matrix $\mathcal{T}(\boldsymbol{z})$ projected onto $\boldsymbol{G}$. The PSD constraint on $\mathcal{T}(\boldsymbol{z})$ follows from its definition in (20). Alternatively, the optimization problem (29) can be reformulated as

$$\begin{aligned} \min_{\boldsymbol{z}\in\mathbb{C}^L} \quad & \frac{1}{2}\left\|\mathcal{T}(\boldsymbol{z}) \circ \boldsymbol{G} - \tilde{\boldsymbol{R}}_v\right\|_F^2 + \tau\|\boldsymbol{z}\|_{\mathcal{A}_r} \\ \text{subject to} \quad & \mathcal{T}(\boldsymbol{z}) \succeq \boldsymbol{0}, \end{aligned} \tag{30}$$

where τ is a regularization parameter to balance the fitting error and the atomic norm term.

Furthermore, with the PSD constraint $\mathcal{T}(\boldsymbol{z}) \succeq \boldsymbol{0}$, if $r = \operatorname{rank}\left(\mathcal{T}(\boldsymbol{z})\right) \leq L-1$, the Hermitian Toeplitz matrix $\mathcal{T}(\boldsymbol{z})$ can be uniquely decomposed via Vandermonde decomposition as [39]

$$\mathcal{T}(\boldsymbol{z}) = \sum_{k=1}^{r} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k), \tag{31}$$

which leads to

$$\operatorname{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) = L\sum_{k=1}^{r} p_k. \tag{32}$$

Based on the definition of atom set $\mathcal{A}_r$ in (27), the relationship between the atomic norm term $\|\boldsymbol{z}\|_{\mathcal{A}_r}$ and the trace of $\mathcal{T}(\boldsymbol{z})$ can be established as

$$\|\boldsymbol{z}\|_{\mathcal{A}_r} = \operatorname{Tr}\left(\mathcal{T}(\boldsymbol{z})\right)/L. \tag{33}$$

Hence, the optimization problem (30) can be equivalently rewritten as

$$\begin{aligned} \min_{\boldsymbol{z}\in\mathbb{C}^L} \quad & \frac{1}{2}\left\|\mathcal{T}(\boldsymbol{z}) \circ \boldsymbol{G} - \tilde{\boldsymbol{R}}_v\right\|_F^2 + \mu \operatorname{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) \\ \text{subject to} \quad & \mathcal{T}(\boldsymbol{z}) \succeq \boldsymbol{0}, \end{aligned} \tag{34}$$

where $\mu = \tau/L$. The equivalent problem (34) is convex, and can be solved using standard and highly efficient software tools based on the interior point methods. With the solution $\hat{\boldsymbol{z}}$ of (34), the interpolated virtual array covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}})$ can be effectively reconstructed with the Hermitian PSD Toeplitz structure. The incorporation of the binary matrix $\boldsymbol{G}$ in (34) constraints the fitting error between the non-zero entries in $\tilde{\boldsymbol{R}}_v$ and the optimized covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}})$ that are projected onto $\boldsymbol{G}$, i.e., $\mathcal{T}(\hat{\boldsymbol{z}}) \circ \boldsymbol{G}$, such that all the available correlation observations corresponding to the non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$ are utilized for denoising. Meanwhile, the proposed optimization problem also simultaneously

Algorithm 1 Virtual Array Interpolation-based DOA Estimation

1: **Input:** Coprime array received signals $\{\boldsymbol{x}(t)\}_{t=1}^{T}$.
2: **Output:** $\hat{\theta}_k, k = 1, 2, \cdots, K$.
3: **Initialize:** $\hat{\boldsymbol{R}}_{\boldsymbol{x}}$, $\bar{\boldsymbol{y}}_v$, L.
4: Derive the second-order virtual array signals $\bar{\boldsymbol{y}}_v$ by (8);
5: Initialize the interpolated virtual array signals $\boldsymbol{y}_{\mathcal{I}}$ via (9);
6: Construct the interpolated virtual array covariance matrix $\tilde{\boldsymbol{R}}_v$ according to (26);
7: Define a binary vector $\boldsymbol{g}$ to distinguish the sensors in the reference virtual array;
8: Solve (34), the equivalent version of the formulated atomic norm minimization problem (29);
9: Calculate $\boldsymbol{f}_{\mathrm{MUSIC}}$ (35) for DOA estimation.

recovers the remaining unknown entries corresponding to the interpolated nominal sensors, which are initialized as zeros in $\tilde{\boldsymbol{R}}_v$, i.e., $\hat{\boldsymbol{Z}} = \mathcal{T}(\hat{\boldsymbol{z}}) - \mathcal{T}(\hat{\boldsymbol{z}}) \circ \boldsymbol{G}$.

Since the reconstructed $\mathcal{T}(\hat{\boldsymbol{z}})$ corresponds to the reference virtual array (i.e., a ULA), existing DOA estimation methods including MUltiple SIgnal Classification (MUSIC)-based [17, 38, 40, 41], Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT)-based [18, 42, 43], and a series of sparsity-based techniques [9, 22, 44, 45] can be incorporated into the virtual domain for unambiguous DOA estimation. For instance, here we present the MUSIC spatial spectrum as

$$\boldsymbol{f}_{\mathrm{MUSIC}}(\theta) = \frac{1}{\boldsymbol{r}^{\mathrm{H}}(\theta)\boldsymbol{N}_{\mathcal{T}(\hat{\boldsymbol{z}})}\boldsymbol{N}_{\mathcal{T}(\hat{\boldsymbol{z}})}^{\mathrm{H}}\boldsymbol{r}(\theta)}, \tag{35}$$

where $\boldsymbol{N}_{\mathcal{T}(\hat{\boldsymbol{z}})}$ denotes the noise subspace of $\mathcal{T}(\hat{\boldsymbol{z}})$. It is worth mentioning that the solution of problem (34), i.e., $\mathcal{T}(\hat{\boldsymbol{z}})$, may not be low-rank in the noisy finite-snapshot cases. As such, the noise subspace $\boldsymbol{N}_{\mathcal{T}(\hat{\boldsymbol{z}})}$ is obtained by collecting the eigenvectors corresponding to the $L-K$ smallest eigenvalues of $\mathcal{T}(\hat{z})$ for the general case, where the number of sources K is assumed known *a priori* with $K < L$. The DOA estimation can be obtained by searching the peaks of $\boldsymbol{f}_{\mathrm{MUSIC}}(\theta)$.

The proposed virtual array interpolation-based DOA estimation algorithm is summarized in Algorithm 1 and has the following key advantages. First, all the information contained in the non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$ is effectively utilized via array interpolation. Second, the reconstructed covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}})$ is strictly Hermitian Toeplitz, which follows the ideal covariance matrix structure of a ULA. Third, with the atomic norm minimization of the equivalent virtual signals, the formulated optimization problem is capable of reconstructing the interpolated virtual array covariance matrix in a gridless manner, where the basis mismatch problem can be avoided. Note that the proposed algorithm is suitable for all kinds of partially augmentable arrays, and the difference due to the diverse virtual array geometries is reflected in the representation of $\boldsymbol{g}$ and $\tilde{\boldsymbol{R}}_v$.

IV. Performance Analysis

A. Covariance Matrix Reconstruction Performance

We analyze the performance of the proposed Toeplitz covariance matrix reconstruction problem depicted in (34). According to (31), the reconstructed covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}}) \in \mathbb{C}^{L\times L}$ behaves like the covariance matrix of signal $\boldsymbol{s}(t)$ received by the L sensors in the reference virtual array, and the virtual array interpolation is thus realized. Considering that matrix $\tilde{\boldsymbol{R}}_v$ taken as the reference in (34) contains the additive noise term as depicted in (26), the theoretical interpolated virtual array covariance matrix can be defined as

$$\mathcal{T}(\tilde{\boldsymbol{z}}) = \mathcal{T}(\boldsymbol{z}^{\star}) + \sigma_n^2\boldsymbol{I} \tag{36}$$

with its first column described as $\tilde{\boldsymbol{z}} = \sum_{k=1}^{K} p_k\boldsymbol{r}(\theta_k) + \sigma_n^2\bar{\boldsymbol{i}}_{\boldsymbol{0}}$. Here, $\bar{\boldsymbol{i}}_{\boldsymbol{0}} \in \mathbb{R}^L$ denotes a vector containing the elements in $\bar{\boldsymbol{i}}$ that correspond to the reference virtual array, whose elements are zeros except a unit element corresponding to the zeroth sensor position. Accordingly, the theoretical covariance matrix corresponding to the partial correlation observations in $\tilde{\boldsymbol{R}}_v$ can be represented as $\mathcal{T}(\tilde{\boldsymbol{z}}) \circ \boldsymbol{G}$. Based on the above preliminaries, we have the following theorem regarding the reconstruction performance of the proposed algorithm:

Theorem 2: There exists a positive constant C such that the regularization parameter

$$\mu \geq \frac{\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\right)}{\sqrt{T}} \tag{37}$$

is sufficient to guarantee that the reconstruction performance of (34) as

$$\left\|\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G} - \mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right\|_F \leq \mu + \sqrt{\mu^2 + 2\mu L\left(\sum_{k=1}^{K} p_k + \sigma_n^2\right)} \tag{38}$$

with the probability at least $1 - 2\mathrm{e}^{-2C\sqrt{T}}$.

Proof: See Appendix B. ■

In particular, when the regularization parameter μ equals to $\mathrm{Tr}(\mathcal{T}(\tilde{\boldsymbol{z}}))/\sqrt{T}$, the reconstruction performance corresponding to the non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$, i.e., the denoising part in (34), follows

$$\left\|\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G} - \mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right\|_F \leq \frac{1+\sqrt{1+2\sqrt{T}}}{\sqrt{T}}\,\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\right), \tag{39}$$

where $\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\right) = L\left(\sum_{K=1}^{K} p_k + \sigma_n^2\right)$ according to (32) and (36). Therefore, the performance of the denoising part relates to the number of snapshots T and the trace of $\mathcal{T}(\tilde{\boldsymbol{z}})$.

In addition to the denoising part, the remaining recovered part $\hat{\boldsymbol{Z}}$ corresponding to the interpolated nominal sensors also influences the reconstruction performance. The reconstruction performance between $\mathcal{T}(\hat{\boldsymbol{z}})$ and $\mathcal{T}(\tilde{\boldsymbol{z}})$ follows

$$\left\|\mathcal{T}(\hat{\boldsymbol{z}}) - \mathcal{T}(\tilde{\boldsymbol{z}})\right\|_F \geq \left\|\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G} - \mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right\|_F, \tag{40}$$

where the equivalence holds if and only if the recovered part $\hat{\boldsymbol{Z}}$ is consistent with its theoretical version $\tilde{\boldsymbol{Z}} = \mathcal{T}(\tilde{\boldsymbol{z}}) - \mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}$. On the other hand, since there are no available correlation observations corresponding to the interpolated nominal sensors in $\mathcal{S}_{\mathcal{I}} - \mathcal{S}_{\mathcal{V}}$ and the corresponding elements in $\tilde{\boldsymbol{R}}_v$ are initialized

to be zeros, the deviation term $\|\mathcal{T}(\hat{\boldsymbol{z}})-\mathcal{T}(\tilde{\boldsymbol{z}})\|_F$ can be further elaborated with respect to the recovered part $\hat{\boldsymbol{Z}}$ and $\tilde{\boldsymbol{Z}}$ in an explicit expression as

$$\begin{aligned}
&\left\|\mathcal{T}(\hat{\boldsymbol{z}})-\mathcal{T}(\tilde{\boldsymbol{z}})\right\|_F \\
&= \left\|\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G}+\hat{\boldsymbol{Z}}-\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}-\tilde{\boldsymbol{Z}}\right\|_F \\
&\le \left\|\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G}-\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right\|_F + \left\|\hat{\boldsymbol{Z}}-\tilde{\boldsymbol{Z}}\right\|_F \\
&\le \mu+\sqrt{\mu^2+2\mu L\left(\sum_{k=1}^{K}p_k+\sigma_n^2\right)}+\left\|\hat{\boldsymbol{Z}}-\tilde{\boldsymbol{Z}}\right\|_F. \qquad (41)
\end{aligned}$$

Since the binary matrix $\boldsymbol{G}$ is determined by the structure of $\mathcal{S}_{\mathcal{V}}$, both $\hat{\boldsymbol{Z}}$ and $\tilde{\boldsymbol{Z}}$ are deterministic as long as the systematic designed coprime array is deployed. Meanwhile, the reconstruction accuracy of $\mathcal{T}(\hat{\boldsymbol{z}})$ will be numerically validated in the second simulation example.

B. DOA Estimation Performance

The performance evaluation of the proposed DOA estimation algorithm follows the stochastic Cramér-Rao bound (CRB) [46], which is the inversion of the Fisher information matrix $\mathbf{FIM}$. Although the reconstructed matrix $\mathcal{T}(\hat{\boldsymbol{z}})$ behaves like the covariance matrix of a virtual array consisting of L sensors, the performance is still determined by the original coprime array because there is no additional information added during the initialization process (9). Thus, with the full utilization of the discontiguous virtual array $\mathcal{S}_{\mathcal{V}}$, the Fisher information matrix for the proposed algorithm is a function of the coprime array covariance matrix $\boldsymbol{R}_{\boldsymbol{x}}$, and the $(\tilde{i},\tilde{j})$-th element can be represented as

$$\mathbf{FIM}_{\tilde{i},\tilde{j}} = T\,\mathrm{Tr}\left[\boldsymbol{R}_{\boldsymbol{x}}^{-1}\frac{\partial\boldsymbol{R}_{\boldsymbol{x}}}{\partial\xi_{\tilde{i}}}\boldsymbol{R}_{\boldsymbol{x}}^{-1}\frac{\partial\boldsymbol{R}_{\boldsymbol{x}}}{\partial\xi_{\tilde{j}}}\right], \qquad (42)$$

where $\xi_{\tilde{i}}$ and $\xi_{\tilde{j}}$ denote the elements in the deterministic parameter vector $\boldsymbol{\xi}$.

Nevertheless, when the number of sources exceeds the number of physical sensors, the abovementioned Fisher information matrix is singular, resulting in the stochastic CRB inapplicable. In view of this, we follow the vectorization process as in (5) and transform the Fisher information matrix into a virtual array-based form as [47]

$$\mathbf{FIM}=T\left[\mathrm{vec}\left(\frac{\partial\boldsymbol{R}_{\boldsymbol{x}}}{\partial\boldsymbol{\xi}}\right)\right]^{\mathrm{H}}\left(\boldsymbol{R}_{\boldsymbol{x}}^{\mathrm{T}}\otimes\boldsymbol{R}_{\boldsymbol{x}}\right)^{-1}\left[\mathrm{vec}\left(\frac{\partial\boldsymbol{R}_{\boldsymbol{x}}}{\partial\boldsymbol{\xi}}\right)\right], \qquad (43)$$

which keeps nonsingular within a much broader range of conditions. Hence, it overcomes the model mismatch issue of the stochastic CRB, and presents a lower bound for the estimation error even when the number of sources is larger than the number of physical sensors.

In our case, the deterministic parameter vector is defined by

$$\boldsymbol{\xi}=\left[\boldsymbol{\theta}^{\mathrm{T}},\boldsymbol{p}^{\mathrm{T}},\sigma_n^2\right]^{\mathrm{T}}. \qquad (44)$$

Accordingly, the Fisher information matrix can be specified as

$$\mathbf{FIM}=T\left[\frac{\partial\boldsymbol{y}_v}{\partial\boldsymbol{\xi}}\right]^{\mathrm{H}}\left(\boldsymbol{R}_{\boldsymbol{x}}^{\mathrm{T}}\otimes\boldsymbol{R}_{\boldsymbol{x}}\right)^{-1}\left[\frac{\partial\boldsymbol{y}_v}{\partial\boldsymbol{\xi}}\right], \qquad (45)$$

where

$$\frac{\partial\boldsymbol{y}_v}{\partial\boldsymbol{\xi}}=\left[\frac{\partial\boldsymbol{y}_v}{\partial\theta_1},\cdots,\frac{\partial\boldsymbol{y}_v}{\partial\theta_K},\frac{\partial\boldsymbol{y}_v}{\partial p_1},\cdots,\frac{\partial\boldsymbol{y}_v}{\partial p_K},\frac{\partial\boldsymbol{y}_v}{\partial\sigma_n^2}\right] \qquad (46)$$

with

$$\begin{aligned}
\frac{\partial\boldsymbol{y}_v}{\partial\theta_k}&=p_k\left[\frac{\partial\boldsymbol{a}^*(\theta_k)}{\partial\theta_k}\otimes\boldsymbol{a}(\theta_k)+\boldsymbol{a}^*(\theta_k)\otimes\frac{\partial\boldsymbol{a}(\theta_k)}{\partial\theta_k}\right],\\
\frac{\partial\boldsymbol{y}_v}{\partial p_k}&=\boldsymbol{a}^*(\theta_k)\otimes\boldsymbol{a}(\theta_k),\\
\frac{\partial\boldsymbol{y}_v}{\partial\sigma_n^2}&=\boldsymbol{i}. \qquad (47)
\end{aligned}$$

Therefore, the CRB for the k-th source can be obtained as

$$\mathrm{CRB}(\theta_k)=\left[\mathbf{FIM}^{-1}\right]_{k,k} \qquad (48)$$

for $1\le k\le K$.

V. Simulation Results

In our simulations, we choose the pair of coprime integers $M=3$ and $N=5$ to deploy the coprime array, which yields a total number of $M+N-1=7$ physical sensors located at $\{0,3d,5d,6d,9d,10d,12d\}$. The proposed virtual array interpolation-based DOA estimation algorithm is compared to several recently reported DOA estimation algorithms using coprime arrays, namely, the Covariance Matrix Sparse Reconstruction (CMSR) algorithm [13], the Sparse Signal Reconstruction (SSR) algorithm [15], the Nuclear Norm Minimization (NNM) algorithm [24], the nuclear norm minimization with PSD constraint (NUC-PSD) algorithm [26], and the Maximum Entropy (ME) algorithm [26]. The sampling interval of the pre-defined sampling grids is selected to be 0.1° for the CMSR algorithm and the SSR algorithm. The regularization parameter μ for the CMSR algorithm, the SSR algorithm, and the proposed algorithm is empirically set to be 0.25 (except for one simulation in Fig. 11, where μ is varied), and the tuning parameters for constraining the optimized solution within the feasible sets in the NUC-PSD algorithm and the ME algorithm are optimally selected as recommended in [26]. The optimization problems are solved using the CVX [48].

In the first example, we compare the resolution by assuming two closely spaced uncorrelated sources from the directions $\theta_1=-0.5^\circ$ and $\theta_2=0.5^\circ$, whose signal-to-noise ratios (SNRs) are 30 dB. The normalized spatial spectra are compared in Fig. 4 with the number of snapshots $T=500$. It is observed from Fig. 4(a) that the estimation results of the CMSR method deviate from the actual source directions. It is because that the CMSR algorithm only incorporates the contiguous part of the virtual array, i.e., $\mathcal{S}_{\mathcal{C}}$ in Fig. 2(c), and the reduced array aperture in the virtual domain affects the resolution. Since all the virtual signals vectorized from the sample covariance matrix are utilized in the SSR algorithm, its spatial spectrum presented in Fig. 4(b) has sharper peaks and more accurate estimates than the CMSR algorithm. The ME algorithm and the NUC-PSD algorithm fail to identify the two closely spaced sources. In contrast, both the NNM algorithm and the proposed algorithm are capable of resolving both peaks in the actual source directions.

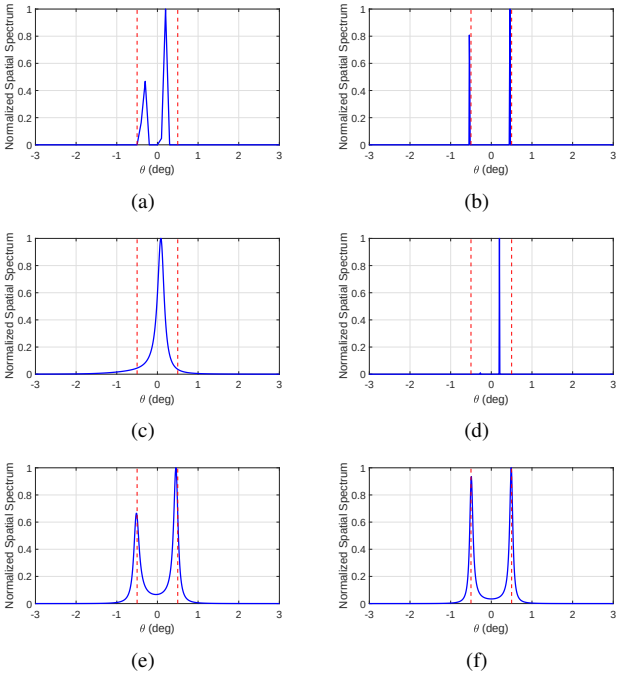


Fig. 4. Resolution comparison in terms of the normalized spatial spectrum with the number of snapshots $T = 500$. The vertical dashed lines denote the actual directions of the incident sources. (a) CMSR algorithm; (b) SSR algorithm; (c) ME algorithm; (d) NUC-PSD algorithm; (e) NNM algorithm; (f) Proposed algorithm.

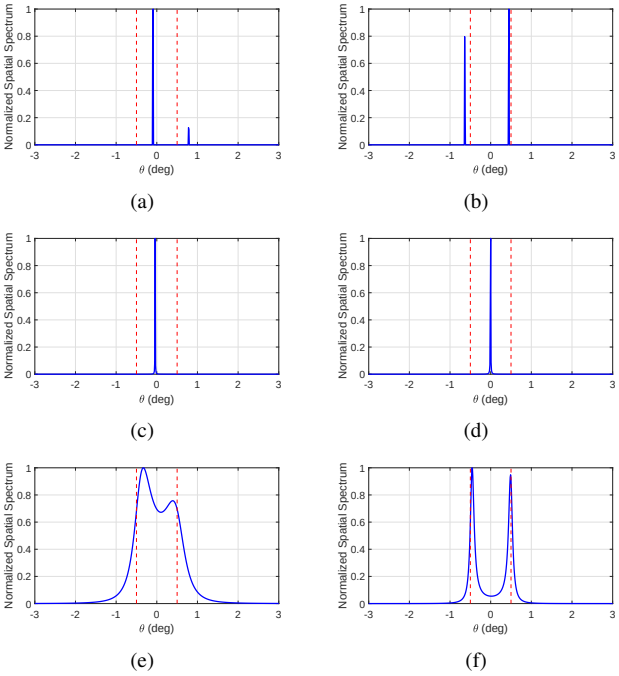


Fig. 5. Resolution comparison in terms of the normalized spatial spectrum with the number of snapshots $T = 100$. The vertical dashed lines denote the actual directions of the incident sources. (a) CMSR algorithm; (b) SSR algorithm; (c) ME algorithm; (d) NUC-PSD algorithm; (e) NNM algorithm; (f) Proposed algorithm.

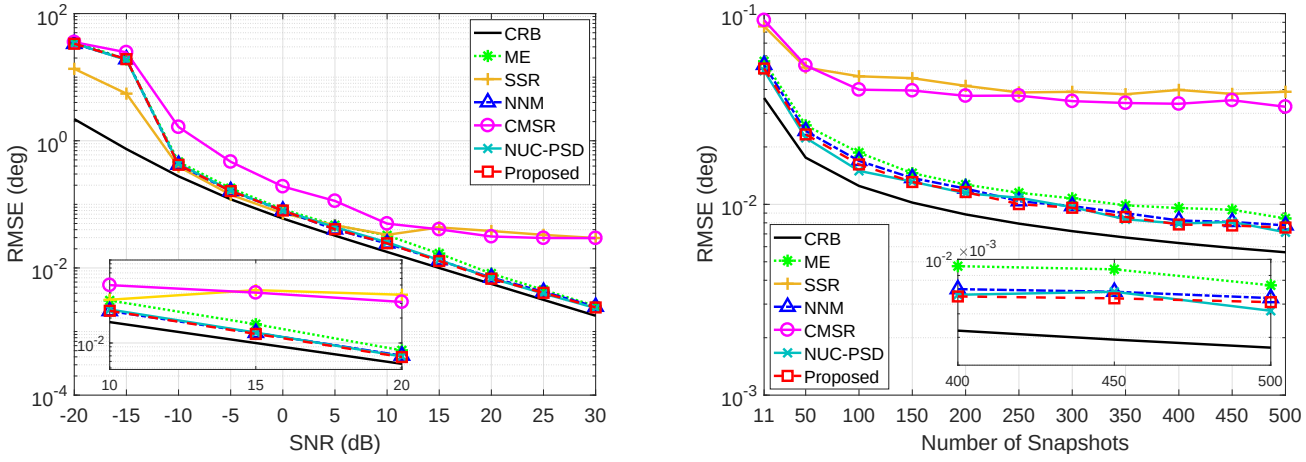
When the number of snapshots reduces to $T = 100$, as depicted in Fig. 5, all the tested algorithms suffer performance degradation to some extent due to the limited number of signal samples. It is observed from Fig. 5(e) that the peaks in the spatial spectrum of the NNM algorithm are no longer as sharp as those obtained from the number of snapshots $T = 500$ as shown in Fig. 4(e), and the DOA estimates deviate from the actual source directions. It is because that the NNM algorithm recovers the covariance matrix of the interpolated virtual array based on the principle of matrix completion, which means that the partial correlation observations corresponding to the non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$ are retained in the optimized covariance matrix corresponding to the interpolated virtual array. Since the equivalent virtual signals are obtained from the sample covariance matrix, there exists an inherent bias due to the finite snapshots. Different from the NNM algorithm, the proposed algorithm retrieves the covariance matrix of the interpolated virtual array through covariance matrix reconstruction, where the partial correlation observations contained in $\tilde{\boldsymbol{R}}_v$ are taken as the reference. Therefore, the elements in the reconstructed covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}})$ may not be the same ones as those in $\tilde{\boldsymbol{R}}_v$. It is evident from Fig. 5(f) that the proposed algorithm achieves a better resolution than the others.

In the second example, we compare the root mean square error (RMSE) of each algorithm in Fig. 6. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{KQ}\sum_{k=1}^{K}\sum_{q=1}^{Q}\left(\hat{\theta}_k(q) - \theta_k\right)^2}, \tag{49}$$

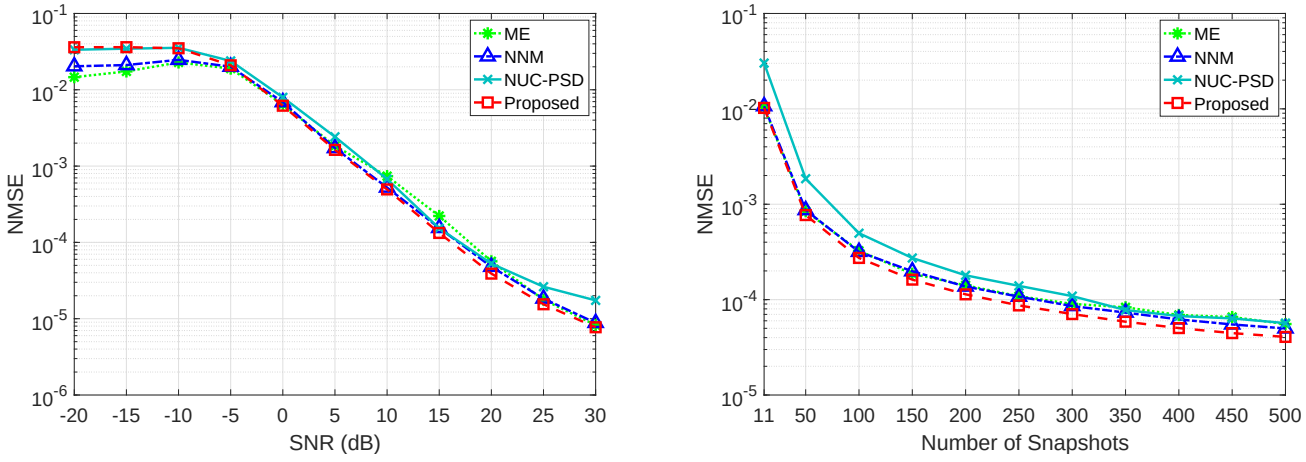
where $\hat{\theta}_k(q)$ denotes the estimated DOA of the k-th source θ_k in the q-th Monte Carlo trial, and Q denotes the number of Monte Carlo trials. The direction of the incident source is randomly generated from the Gaussian distribution $\mathcal{N}(0°, 1°)$, and changes from trial to trial but remains fixed from snapshot to snapshot. The number of snapshots is fixed at $T = 500$ when the SNR varies, whereas the SNR is fixed at 20 dB when the number of snapshots varies. The DOAs are estimated by finding the spectrum peaks for the CMSR algorithm and the SSR algorithm, whereas the root-MUSIC [41] is performed on the optimized covariance matrix of the other gridless algorithms. The Cramér-Rao bound (CRB) (48) is also plotted. For each data point, the RMSE is calculated from $Q = 500$ Monte Carlo trials.

It is shown in Fig. 6(a) that the RMSE curves of both the CMSR algorithm and the SSR algorithm become relatively flat when the SNR is larger than 10 dB. The reason lies in that the fixed sampling interval for the pre-defined sampling grids leads to an inherent basis mismatch, which limits the estimation accuracy. In contrast, the gridless algorithms do not require pre-defined sampling grids, and their estimation performance are not limited by the sampling interval. Therefore, their RMSE are consistent with that manifested in the CRB when the SNR increases. Also, it is observed from Fig. 6(a) that the gridless algorithms, especially the NNM algorithm, the NUC-PSD algorithm, and the proposed algorithm, achieve a quite similar RMSE performance. Similar performance comparison can also be found in Fig. 6(b), where the number of snapshots is varied. As pointed out in the first example, the NNM algorithm and the proposed algorithm formulate the optimization problems



(a) (b)

Fig. 6. RMSE performance comparison with single incident source. (a) RMSE versus SNR with the number of snapshots $T = 500$; (b) RMSE versus the number of snapshots with SNR = 20 dB.



(a) (b)

Fig. 7. NMSE performance comparison of the optimized covariance matrix in each algorithm. (a) NMSE versus SNR with the number of snapshots $T = 500$; (b) NMSE versus the number of snapshots with SNR = 20 dB.

based on different principles, namely, matrix completion and matrix reconstruction, respectively. It is shown in Fig. 6(b) that the RMSE of the proposed algorithm is slightly smaller than that of the NNM algorithm.

In order to demonstrate the estimation accuracy of the reconstructed covariance matrix of the proposed algorithm, we compare with three tested virtual array interpolation-based algorithms by evaluating their optimized Hermitian Toeplitz covariance matrices. Considering that the elements in a Hermitian Toeplitz matrix are determined by its first column, here we define the normalized mean square error (NMSE) as

$$\mathrm{NMSE} = \frac{\mathrm{E}\left[\|\tilde{\boldsymbol{z}} - \hat{\boldsymbol{z}}\|_2^2\right]}{\|\tilde{\boldsymbol{z}}\|_2^2}, \tag{50}$$

where $\hat{\boldsymbol{z}}$ is the first column of the estimated Hermitian Toeplitz matrix $\mathcal{T}(\hat{\boldsymbol{z}})$. It is demonstrated in Fig. 7(a) that, in such a case, the proposed covariance matrix reconstruction-based algorithm outperforms the matrix completion-based NNM algorithm in terms of the estimation accuracy of the reconstructed covariance matrix when the SNR is higher than 15 dB. It is because that the proposed algorithm utilizes the partial correlation observations as the reference for the covariance matrix reconstruction, whereas the NNM algorithm keeps the correlation observations fixed in the optimized covariance matrix. The NMSE performance versus the number of snapshots shown in Fig. 7(b) also verifies the superiority of the proposed algorithm.

In the third example, we compare the available DOFs of the tested algorithms. Assume that there are seven uncorrelated equal-power incident sources uniformly distributed in $[-50°, 50°]$ with SNR = 30 dB and $T = 500$. It can be seen from Fig. 8 that all the tested algorithms exhibit peaks around the actual directions by using only seven physical sensors, where the DOF superiority using the coprime array is demonstrated. When the number of incident sources increases to nine, some targets are apparently missed by the CMSR algo-

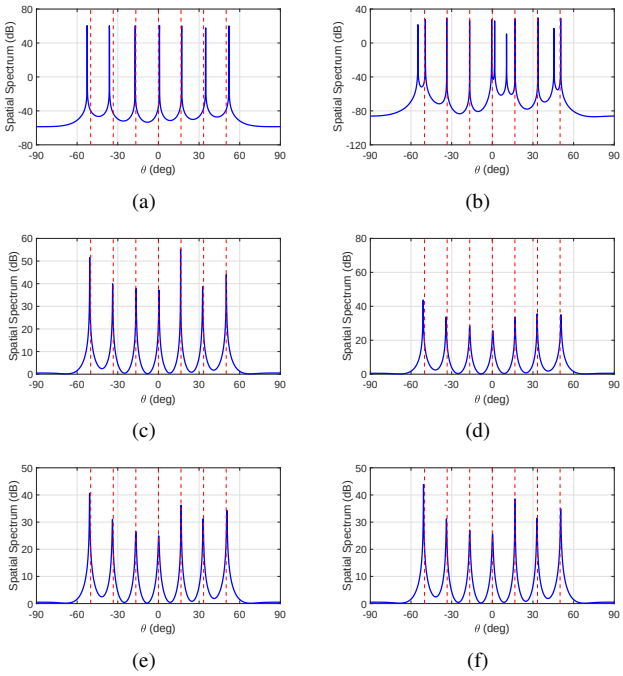


Fig. 8. DOFs capability comparison in terms of the spatial spectrum, number of sources $K = 7$. The vertical dashed lines denote the actual directions of the incident sources. (a) CMSR algorithm; (b) SSR algorithm; (c) ME algorithm; (d) NUC-PSD algorithm; (e) NNM algorithm; (f) Proposed algorithm.

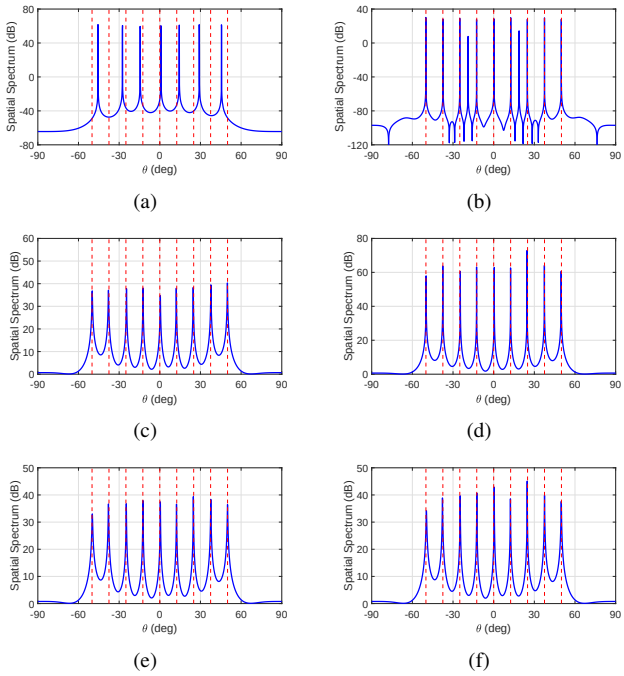


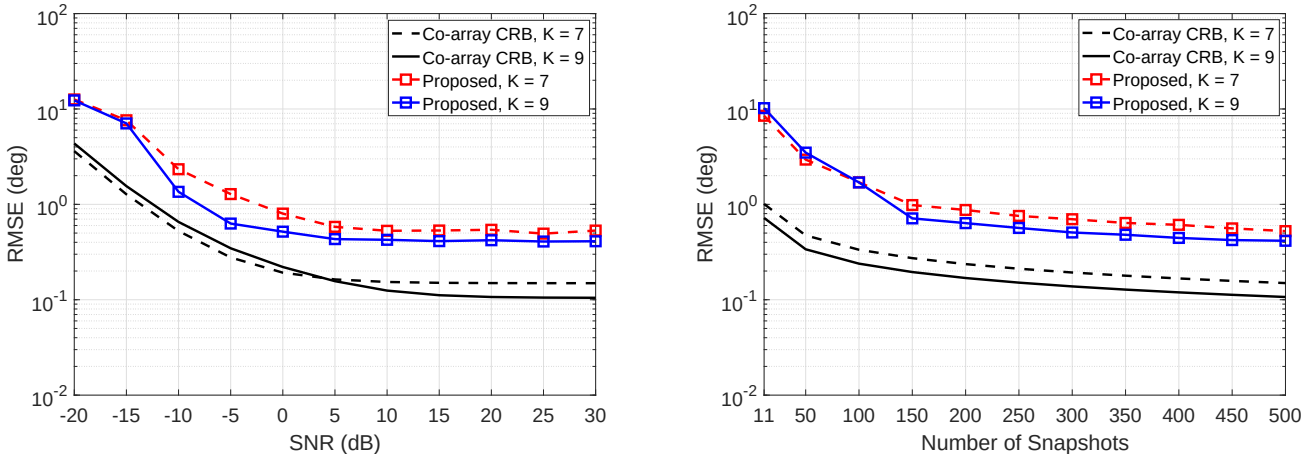
Fig. 9. DOFs capability comparison in terms of the spatial spectrum, number of sources $K = 9$. The vertical dashed lines denote the actual directions of the incident sources. (a) CMSR algorithm; (b) SSR algorithm; (c) ME algorithm; (d) NUC-PSD algorithm; (e) NNM algorithm; (f) Proposed algorithm.

rithm as illustrated in Fig. 9(a), because the CMSR algorithm picks only the contiguous part of the difference co-array $\mathcal{S}_{\mathcal{C}}$ to use the spatial smoothing technique. Hence, the maximum achievable number of DOFs for the CMSR algorithm is seven. It is observed from Fig. 8(b) and Fig. 9(b) that there are several spurious peaks in the obtained sparse spatial spectrum of the SSR algorithm. In contrast, all the virtual array interpolation-based algorithms are capable of identifying the nine sources as exhibited in Figs. 9(c) to 9(f), demonstrating that the DOFs offered by the non-uniform virtual array can be obtained via virtual array interpolation.

In the fourth example, we evaluate the RMSE of the proposed algorithm in Fig. 10 in the case that the number of sources is equal to or greater than the number of physical sensors, where the sources are uniformly distributed in $[-50^\circ, 50^\circ]$ with the number of sources $K = 7$ and $K = 9$, respectively. The DOAs are estimated from the spatial spectra presented in Fig. 8 and Fig. 9 with the spectrum peak search interval of $\Delta\theta = 0.1^\circ$, and the RMSE is calculated from 500 Monte Carlo trials for each scenario. The virtual array-based CRB (48) is also presented in Fig. 10 as the reference. It is observed from Fig. 10(a) that, with the increase of the SNR, the CRB converges to a constant when the SNR is larger than 5 dB rather than keeps decreasing as in Fig. 6(a). This is the typical *saturation* behavior [49]. Although the CRB for 7 sources is lower than that for 9 sources when the SNR is relatively small, the CRB for 9 sources converges to a lower one when SNR is larger than 5 dB. It is demonstrated in Fig. 10 that the RMSE of the proposed algorithm has a similar trend with the CRB, and exhibits the saturation behavior as the CRB in the asymptotic performance region.

In the fifth example, we compare the RMSE performance with respect to the regularization parameter μ in the case of SNR = 30 dB and $T = 500$. The other parameters are the same as those in the second example. Three algorithms are considered, namely, the SSR algorithm, the CMSR algorithm, and the proposed algorithm. According to the comparison results shown in Fig. 11, it is clear that the change of regularization parameter μ does not affect the RMSE performance of the proposed algorithm, whereas the RMSE curves of the other two algorithms are fluctuant when μ varies. Therefore, the proposed algorithm is robust to the regularization parameter, which is comparable to the regularization-free NNM algorithm. In addition, the proposed algorithm has the lowest RMSE among the three tested algorithms.

In the last example, we compare the computational complexity measured by the computation time for 100 Monte Carlo trials on an Intel(R) Core(TM) i7-7600U CPU, 16G RAM laptop, where the sampling/searching interval is varied. According to Fig. 12, the computational complexities of the CMSR algorithm and the SSR algorithm increase exponentially when the sampling interval decreases. This is because the pre-defined dense sampling grids dramatically increase the computational cost when solving the corresponding optimization problems. In contrast, since the gridless algorithms formulate the optimization problem without pre-defined sampling grids, their computational costs are insensitive to the selection of the sampling interval. The ME algorithm has a higher computational complexity than the other gridless algorithms because it adopts a two-stage optimization process, whereas the NNM algorithm, the NUC-PSD algorithm, and the proposed algorithm have the same order of computational



(a) (b)

Fig. 10. RMSE performance comparison with multiple incident sources. (a) RMSE versus SNR with the number of snapshots $T = 500$; (b) RMSE versus the number of snapshots with SNR = 20 dB.

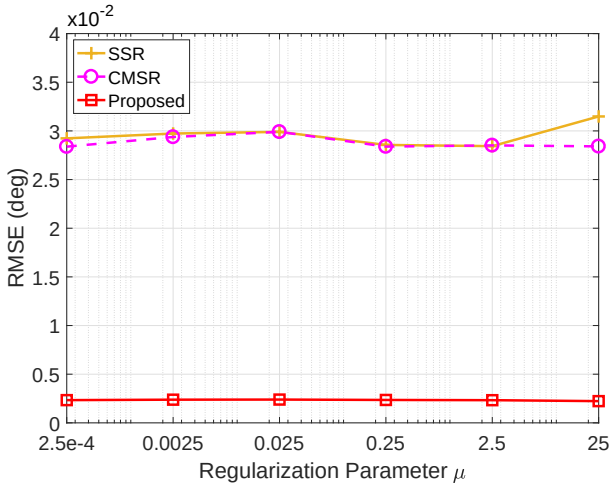


Fig. 11. RMSE performance comparison with different regularization parameter μ.

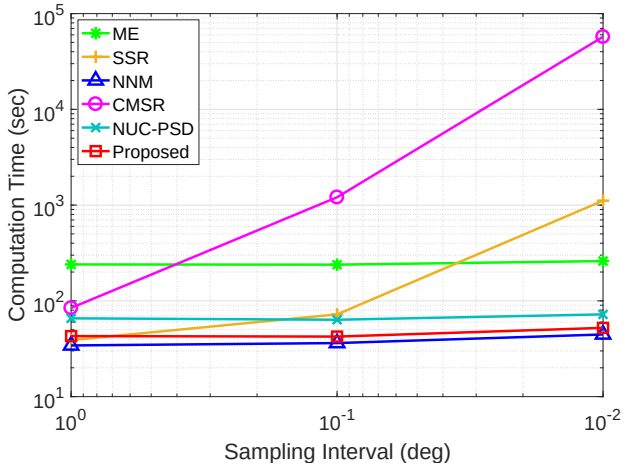


Fig. 12. Computation time comparison with different sampling interval.

complexity. In particular, it is demonstrated in Fig. 12 that the proposed algorithm has a lower computational complexity than the NUC-PSD algorithm, and has a slightly higher computational complexity than the NNM algorithm. Therefore, the efficiency of the proposed algorithm is validated.

VI. Conclusions

In this paper, we proposed a novel virtual array interpolation-based coprime array DOA estimation algorithm. The equivalent virtual signals corresponding to the non-uniform difference co-array are derived, and the nominal sensors are interpolated to generate a virtual ULA, where all information of the coprime array received signals is included. The atomic norm of multiple virtual measurements is investigated, based on which the relationship between the covariance matrix of the interpolated virtual array and the virtual signals is established. An optimization problem is formulated by minimizing the atomic norm of a virtual measurement vector in a gridless manner, where the reconstructed Toeplitz covariance matrix is utilized for DOA estimation with an increased number of DOFs. The effectiveness of the proposed algorithm is verified through simulation comparisons with existing algorithms.

Although the proposed DOA estimation algorithm is for a systematic coprime array, the virtual array interpolation-based technique proposed in this paper is applicable to a general class of partially augmentable arrays, and the implementation in different sparse array configurations is straightforward.

Appendix A
Proof of Theorem 1

We prove Theorem 1 from two aspects, namely, the lower bound of $\|\boldsymbol{Y}\|_{\mathcal{A}}$ and the upper bound of $\|\boldsymbol{Y}\|_{\mathcal{A}}$.

Part A: The lower bound of $\|\boldsymbol{Y}\|_{\mathcal{A}}$.

Proof: Assume that the atomic decomposition of $\boldsymbol{Y}$ is

$$\boldsymbol{Y} = \sum_{k=1}^{K} p_k \boldsymbol{B}(\theta_k). \tag{51}$$

According to the Vandermonde decomposition lemma [39], there exists a variable vector $\boldsymbol{z}$ satisfying

$$\mathcal{T}(\boldsymbol{z}) = \sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k). \tag{52}$$

Accordingly, we have

$$\begin{aligned}
&\left[\begin{array}{cc} \mathcal{T}(\boldsymbol{z}) & \boldsymbol{Y} \\ \boldsymbol{Y}^{\mathrm{H}} & \boldsymbol{W} \end{array}\right] \\
&= \left[\begin{array}{cc} \sum\limits_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k) & \sum\limits_{k=1}^{K} p_k \boldsymbol{B}(\theta_k) \\ \sum\limits_{k=1}^{K} p_k \boldsymbol{B}^{\mathrm{H}}(\theta_k) & \sum\limits_{k=1}^{K} p_k \boldsymbol{b}(\theta_k)\boldsymbol{b}^{\mathrm{H}}(\theta_k) \end{array}\right] \\
&= \sum_{k=1}^{K} p_k \left[\begin{array}{c} \boldsymbol{r}(\theta_k) \\ \boldsymbol{b}(\theta_k) \end{array}\right] \left[\begin{array}{cc} \boldsymbol{r}^{\mathrm{H}}(\theta_k) & \boldsymbol{b}^{\mathrm{H}}(\theta_k) \end{array}\right] \succeq \boldsymbol{0}.
\end{aligned} \tag{53}$$

Based on the definitions of $\boldsymbol{r}(\theta)$ in (13) and $\boldsymbol{b}(\theta)$ in (14), we have

$$\mathrm{Tr}\left(\boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k)\right) = \mathrm{Tr}\left(\boldsymbol{b}(\theta_k)\boldsymbol{b}^{\mathrm{H}}(\theta_k)\right) = L, \tag{54}$$

thus we can obtain that

$$\frac{1}{2L}\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) + \frac{1}{2L}\mathrm{Tr}\left(\boldsymbol{W}\right) = \sum_{k=1}^{K} p_k = \|\boldsymbol{Y}\|_{\mathcal{A}}. \tag{55}$$

Therefore,

$$\|\boldsymbol{Y}\|_{\mathcal{A}} \geq \inf_{\boldsymbol{z}\in\mathbb{C}^L, \boldsymbol{W}\in\mathbb{C}^{L\times L}} \left\{ \frac{1}{2L}\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) + \frac{1}{2L}\mathrm{Tr}(\boldsymbol{W}) \middle| \left[\begin{array}{cc} \mathcal{T}(\boldsymbol{z}) & \boldsymbol{Y} \\ \boldsymbol{Y}^{\mathrm{H}} & \boldsymbol{W} \end{array}\right] \succeq \boldsymbol{0} \right\}. \tag{56}$$

□

Part B: The upper bound of $\|\boldsymbol{Y}\|_{\mathcal{A}}$.

Proof: Assume that the PSD constraint in (20) holds. With the Vandermonde decomposition, the PSD Toeplitz matrix $\mathcal{T}(\boldsymbol{z})$ can be represented as

$$\mathcal{T}(\boldsymbol{z}) = \boldsymbol{D}\boldsymbol{C}\boldsymbol{D}^{\mathrm{H}}, \tag{57}$$

where $\boldsymbol{D} \in \mathbb{C}^{L\times K}$ is a Vandermonde matrix, and $\boldsymbol{C} = \mathrm{diag}(c_k)$ with $c_k \geq 0$ [32]. Since $\|\boldsymbol{r}(\theta_k)\|_2 = \sqrt{L}$, we have

$$\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) = \mathrm{Tr}\left(\sum_{k=1}^{K} p_k \boldsymbol{r}(\theta_k)\boldsymbol{r}^{\mathrm{H}}(\theta_k)\right) = L\,\mathrm{Tr}(\boldsymbol{C}). \tag{58}$$

Consequently, each $\boldsymbol{r}(\theta_k), k = 1, 2, \cdots, K$, lies in the range spanned by $\boldsymbol{D}$, and $\boldsymbol{Y}$ can be rewritten as

$$\boldsymbol{Y} = \boldsymbol{D}\boldsymbol{E}, \tag{59}$$

where the matrix $\boldsymbol{E} \in \mathbb{C}^{K\times L}$ contains both the power and phase offsets information among the virtual measurements. Denoting $\boldsymbol{e} = [e_1, e_2, \cdots, e_K]^{\mathrm{T}}$ as the first column of $\boldsymbol{E}$, the PSD matrix $\boldsymbol{W}$ can then be represented as

$$\boldsymbol{W} = \boldsymbol{E}^{\mathrm{H}}\boldsymbol{Q}\boldsymbol{E}, \tag{60}$$

where $\boldsymbol{Q}$ is also a PSD matrix. Since the matrix

$$\begin{aligned}
&\left[\begin{array}{cc} \mathcal{T}(\boldsymbol{z}) & \boldsymbol{Y} \\ \boldsymbol{Y}^{\mathrm{H}} & \boldsymbol{W} \end{array}\right] \\
&= \left[\begin{array}{cc} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{E}^{\mathrm{H}} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{C} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{Q} \end{array}\right] \left[\begin{array}{cc} \boldsymbol{D}^{\mathrm{H}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{E} \end{array}\right] \succeq \boldsymbol{0},
\end{aligned} \tag{61}$$

according to the Schur complement lemma, we have

$$\left[\begin{array}{cc} \boldsymbol{C} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{Q} \end{array}\right] \succeq \boldsymbol{0}, \tag{62}$$

and

$$\boldsymbol{Q} \succeq \boldsymbol{C}^{-1}. \tag{63}$$

Here, we can naturally obtain that

$$\mathrm{Tr}\left(\boldsymbol{W}\right) = \mathrm{Tr}\left(\boldsymbol{E}^{\mathrm{H}}\boldsymbol{Q}\boldsymbol{E}\right) \geq \mathrm{Tr}\left(\boldsymbol{E}^{\mathrm{H}}\boldsymbol{C}^{-1}\boldsymbol{E}\right). \tag{64}$$

Further, we have

$$\mathrm{Tr}\left(\boldsymbol{E}^{\mathrm{H}}\boldsymbol{C}^{-1}\boldsymbol{E}\right) = \mathrm{Tr}\left(\boldsymbol{C}^{-1}\boldsymbol{E}\boldsymbol{E}^{\mathrm{H}}\right) = L\sum_{k} c_k^{-1}|e_k|^2. \tag{65}$$

Substituting (65) into (64) yields

$$\mathrm{Tr}(\boldsymbol{W}) \geq L\sum_{k} c_k^{-1}|e_k|^2. \tag{66}$$

With (58) and (66), we have

$$\begin{aligned}
&\frac{1}{2L}\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) + \frac{1}{2L}\mathrm{Tr}\left(\boldsymbol{W}\right) \\
&\qquad \geq \frac{1}{2}\mathrm{Tr}\left(\boldsymbol{C}\right) + \frac{1}{2}\sum_{k} c_k^{-1}|e_k|^2 \\
&\qquad = \frac{1}{2}\sum_{k} c_k + \frac{1}{2}\sum_{k} c_k^{-1}|e_k|^2 \\
&\qquad \geq \sqrt{\sum_{k} c_k \sum_{k} c_k^{-1}|e_k|^2} \\
&\qquad \geq \sum_{k} |e_k| \geq \|\boldsymbol{Y}\|_{\mathcal{A}}.
\end{aligned} \tag{67}$$

Therefore,

$$\|\boldsymbol{Y}\|_{\mathcal{A}} \leq \inf_{\boldsymbol{z}\in\mathbb{C}^L, \boldsymbol{W}\in\mathbb{C}^{L\times L}} \left\{ \frac{1}{2L}\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) + \frac{1}{2L}\mathrm{Tr}(\boldsymbol{W}) \middle| \left[\begin{array}{cc} \mathcal{T}(\boldsymbol{z}) & \boldsymbol{Y} \\ \boldsymbol{Y}^{\mathrm{H}} & \boldsymbol{W} \end{array}\right] \succeq \boldsymbol{0} \right\}. \tag{68}$$

□

Combining (56) and (68), we can draw the conclusion that

$$\|\boldsymbol{Y}\|_{\mathcal{A}} = \inf_{\boldsymbol{z}\in\mathbb{C}^L, \boldsymbol{W}\in\mathbb{C}^{L\times L}} \left\{ \frac{1}{2L}\mathrm{Tr}\left(\mathcal{T}(\boldsymbol{z})\right) + \frac{1}{2L}\mathrm{Tr}(\boldsymbol{W}) \middle| \left[\begin{array}{cc} \mathcal{T}(\boldsymbol{z}) & \boldsymbol{Y} \\ \boldsymbol{Y}^{\mathrm{H}} & \boldsymbol{W} \end{array}\right] \succeq \boldsymbol{0} \right\}, \tag{69}$$

which can be viewed as an equivalent expression of $\|\boldsymbol{Y}\|_{\mathcal{A}}$ defined in (19). ■

APPENDIX B
PROOF OF THEOREM 2

Proof: The fitting error between the observed correlation observations in $\tilde{\boldsymbol{R}}_v$ and the corresponding elements in the reconstructed covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G}$ can be expressed as

$$\begin{aligned}&\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2\\&=\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}+\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2\\&=\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F^2+\left\|\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2\\&\quad+2\left\langle\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G},\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\rangle_F,\end{aligned}\tag{70}$$

where $\langle\cdot,\cdot\rangle_F$ denotes the Frobenius inner product. Since $\hat{\boldsymbol{z}}$ is the optimal solution to (34), we have

$$\begin{aligned}&\left\|\mathcal{T}(\hat{\boldsymbol{z}})\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2-\left\|\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2\\&\qquad\le2\mu\,\mathrm{Tr}\left(\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\right)-2\mu\,\mathrm{Tr}\left(\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\right).\end{aligned}\tag{71}$$

Combining the Cauchy-Schwarz inequality on the Frobenius inner product term as

$$\begin{aligned}&\left|\left\langle\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G},\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\rangle_F\right|\\&\quad\le\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F\left\|\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F,\end{aligned}\tag{72}$$

(70) can be reformulated as

$$\begin{aligned}&\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F^2\\&=\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2-\left\|\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F^2\\&\quad-2\left\langle\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G},\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\rangle_F\\&\le2\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F\left\|\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F\\&\quad+2\mu\,\mathrm{Tr}\left(\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\right)-2\mu\,\mathrm{Tr}\left(\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\right).\end{aligned}\tag{73}$$

Before proceeding, we discuss the selection of the regularization parameter μ in (34), where the following bound established in [27] is utilized.

Lemma 1 [27]: Let $\{\boldsymbol{x}(t),t=1,2,\cdots T\}$ be zero mean i.i.d. Gaussian random vectors distributed as $\boldsymbol{x}(t)\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{R}_{\boldsymbol{x}})$. Then,

$$\mathbb{P}\left\{\left\|\boldsymbol{R}_{\boldsymbol{x}}-\hat{\boldsymbol{R}}_{\boldsymbol{x}}\right\|_F\ge\frac{\mathrm{Tr}\left(\boldsymbol{R}_{\boldsymbol{x}}\right)}{\sqrt{T}}\right\}\le2\mathrm{e}^{-2C\sqrt{T}},\tag{74}$$

where $\mathbb{P}(\cdot)$ denotes the probability.

By generalizing the relationship revealed in (74) to the virtual domain, we have

$$\left\|\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}-\tilde{\boldsymbol{R}}_v\right\|_F\le\frac{\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right)}{\sqrt{T}}\tag{75}$$

with probability at least $1-2\mathrm{e}^{-2C\sqrt{T}}$. Without loss of generality, the selection of the regularization parameter μ can be related to the fitting error between the reference covariance matrix $\tilde{\boldsymbol{R}}_v$ and its theoretical version $\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}$ as

$$\mu\ge\frac{\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right)}{\sqrt{T}}=\frac{\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\right)}{\sqrt{T}}=\frac{L}{\sqrt{T}}\left(\sum_{k=1}^{K}p_k+\sigma_n^2\right),\tag{76}$$

where the equation $\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\circ\boldsymbol{G}\right)=\mathrm{Tr}\left(\mathcal{T}(\tilde{\boldsymbol{z}})\right)$ holds due to the fact that the zeroth position is always included in the derived non-uniform virtual array $\mathcal{S}_{\mathcal{V}}$ for the sparse arrays.

On the other hand, since the reconstructed covariance matrix $\mathcal{T}(\hat{\boldsymbol{z}})$ is a PSD matrix, we have $\mathrm{Tr}\left(\mathcal{T}(\hat{\boldsymbol{z}})\right)\ge0$. Then, the relationship in (73) continues as

$$\begin{aligned}&\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F^2\\&\qquad\le2\mu\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F+2\mu\,\mathrm{Tr}\left(\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\right).\end{aligned}\tag{77}$$

Denoting $\delta=\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\|_F$, the factorization of the quadratic inequality (77) yields

$$\left(\delta-\mu-\sqrt{\mu^2+2\mu\mathrm{Tr}(\mathcal{T}(\tilde{\boldsymbol{z}}))}\right)\left(\delta-\mu+\sqrt{\mu^2+2\mu\mathrm{Tr}(\mathcal{T}(\tilde{\boldsymbol{z}}))}\right)\le0.\tag{78}$$

Hence, we have

$$\begin{aligned}\left\|\mathcal{T}\left(\hat{\boldsymbol{z}}\right)\circ\boldsymbol{G}-\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\circ\boldsymbol{G}\right\|_F&\le\mu+\sqrt{\mu^2+2\mu\,\mathrm{Tr}\left(\mathcal{T}\left(\tilde{\boldsymbol{z}}\right)\right)}\\&=\mu+\sqrt{\mu^2+2\mu L\left(\sum_{k=1}^{K}p_k+\sigma_n^2\right)},\end{aligned}\tag{79}$$

which establishes the relationship described in (38). ■

ACKNOWLEDGMENT

The authors would like to thank the associate editor Prof. Romain Couillet and the anonymous reviewers for their helpful comments and suggestions. We also thank Dr. Xiaohuan Wu, Dr. Chun-Lin Liu, Dr. Yuanxin Li and Heng Qiao for their helpful communications during the preparation and revision of this paper.

REFERENCES

[1] X. Fan, C. Zhou, Y. Gu, and Z. Shi, "Toeplitz matrix reconstruction of interpolated coprime virtual array for DOA estimation," in *Proc. IEEE VTC2017-Spring*, Sydney, Australia, June 2017.

[2] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*. New York, NY: John Wiley & Sons, 2002.

[3] 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, Sep. 2017.

[4] C. Yang, L. Feng, H. Zhang, S. He, and Z. Shi, "A novel data fusion algorithm to combat false data injection attacks in networked radar systems," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 1, pp. 125–136, Mar. 2018.

[5] C. Shi, F. Wang, M. Sellathurai, J. Zhou, and S. Salous, "Power minimization based robust OFDM radar waveform design for radar and communication systems in coexistence," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1316–1330, Mar. 2018.

[6] X. Wu, W.-P. Zhu, and J. Yan, "A high-resolution DOA estimation method with a family of nonconvex penalties," *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 4925–4938, June 2018.

[7] Y. Gu and A. Leshem, "Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3881–3885, July 2012.

[8] 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, Feb. 2011.

[9] P. Pal, "Correlation awareness in low-rank models: Sampling, algorithms, and fundamental limits," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 56–71, July 2018.

[10] C. Zhou, Z. Shi, Y. Gu, and X. Shen, "DECOM: DOA estimation with combined MUSIC for coprime array," in *Proc. WCSP*, Hangzhou, China, Oct. 2013.

[11] C. Zhou, Y. Gu, S. He, and Z. Shi, "A robust and efficient algorithm for coprime array adaptive beamforming," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1099–1112, Feb. 2018.

[12] S. Qin, Y. D. Zhang, M. G. Amin, and A. M. Zoubir, "Generalized coprime sampling of Toeplitz matrices for spectrum estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 81–94, Jan. 2017.

[13] Z. Shi, C. Zhou, Y. Gu, N. A. Goodman, and F. Qu, "Source estimation using coprime array: A sparse reconstruction perspective," *IEEE Sensors J.*, vol. 17, no. 3, pp. 755–765, Feb. 2017.

[14] P. Pal and P. P. Vaidyanathan, "A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 737–741, Jun. 2014.

[15] Y. D. Zhang, M. G. Amin, and B. Himed, "Sparsity-based DOA estimation using co-prime arrays," in *Proc. IEEE ICASSP*, Vancouver, Canada, May 2013, pp. 3967–3971.

[16] 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, Aug. 2017.

[17] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," in *Proc. IEEE DSP/SPE Workshop*, Sedona, AZ, Jan. 2011, pp. 289–294.

[18] C. Zhou and J. Zhou, "Direction-of-arrival estimation with coarray ESPRIT for coprime array," *Sensors*, vol. 17, no. 8, p. 1779, Aug. 2017.

[19] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays – Part II: Partially augmentable arrays," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1502–1521, June 1999.

[20] 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, Mar. 2015.

[21] J. Shi, G. Hu, X. Zhang, F. Sun, W. Zheng, and Y. Xiao, "Generalized co-prime MIMO radar for DOA estimation with enhanced degrees of freedom," *IEEE Sensors J.*, vol. 18, no. 3, pp. 1203–1212, Feb. 2018.

[22] J. Shi, G. Hu, X. Zhang, F. Sun, and H. Zhou, "Sparsity-based two-dimensional DOA estimation for coprime array: From sum-difference coarray viewpoint," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5591–5604, Nov. 2017.

[23] E. BouDaher, Y. Jia, F. Ahmad, and M. G. Amin, "Multi-frequency co-prime arrays for high-resolution direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3797–3808, July 2015.

[24] C.-L. Liu, P. P. Vaidyanathan, and P. Pal, "Coprime coarray interpolation for DOA estimation via nuclear norm minimization," in *Proc. IEEE ISCAS*, Montréal, Canada, May 2016, pp. 2639–2642.

[25] S. M. Hosseini and M. A. Sebt, "Array interpolation using covariance matrix completion of minimum-size virtual array," *IEEE Signal Process. Lett.*, vol. 24, no. 7, pp. 1063–1067, July 2017.

[26] H. Qiao and P. Pal, "Unified analysis of co-array interpolation for direction-of-arrival estimation," in *Proc. IEEE ICASSP*, New Orleans, LA, Mar. 2017, pp. 3056–3060.

[27] H. Qiao and P. Pal, "Gridless line spectrum estimation and low-rank Toeplitz matrix compression using structured samplers: A regularization-free approach," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2221–2236, May 2017.

[28] X. Wu, W.-P. Zhu, and J. Yan, "A Toeplitz covariance matrix reconstruction approach for direction-of-arrival estimation," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 8223–8237, Sep. 2017.

[29] H. Qiao and P. Pal, "On maximum-likelihood methods for localizing more sources than sensors," *IEEE Signal Process. Lett.*, vol. 24, no. 5, pp. 703–706, May 2017.

[30] H. Qiao and P. Pal, "Generalized nested sampling for compressing low rank Toeplitz matrices," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1844–1848, Nov. 2015.

[31] X. Wu, W.-P. Zhu, and J. Yan, "A fast gridless covariance matrix reconstruction method for one- and two-dimensional direction-of-arrival estimation," *IEEE Sensors J.*, vol. 17, no. 15, pp. 4916–4927, Aug. 2017.

[32] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[33] G. Tang, B. N. Bhaskar, and B. Recht, "Near minimax line spectral estimation," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 499–512, Jan. 2015.

[34] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016.

[35] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction)," *Ann. Math. Statist.*, vol. 34, no. 1, pp. 152–177, Mar. 1963.

[36] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1970.

[37] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Found. Comput. Math.*, vol. 12, no. 6, pp. 805–849, Dec. 2012.

[38] 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, Sep. 2015.

[39] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice Hall, 2005.

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

[41] B. D. Rao and K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.

[42] 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.

[43] H. Chen, C. Hou, W.-P. Zhu, W. Liu, Y.-Y. Dong, Z. Peng, and Q. Wang, "ESPRIT-like two-dimensional direction finding for mixed circular and strictly noncircular sources based on joint diagonalization," *Signal Process.*, vol. 141, pp. 48–56, Dec. 2017.

[44] X. Wu, W.-P. Zhu, J. Yan, and Z. Zhang, "Two sparse-based methods for off-grid direction-of-arrival estimation," *Signal Process.*, vol. 142, pp. 87–95, Jan. 2018.

[45] X. Wu, W.-P. Zhu, and J. Yan, "Direction of arrival estimation for off-grid signals based on sparse Bayesian learning," *IEEE Sensors J.*, vol. 16, no. 7, pp. 2004–2016, Apr. 2016.

[46] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound: Further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 12, pp. 2140–2150, Dec. 1990.

[47] C.-L. Liu and P. P. Vaidyanathan, "Cramér-Rao bounds for coprime and other sparse arrays, which find more sources than sensors," *Digit. Signal Process.*, vol. 61, pp. 43–61, Feb. 2017.

[48] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[49] M. Wang and A. Nehorai, "Coarrays, MUSIC, and the Cramér-Rao bound," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 933–946, Feb. 2017.