Correlation Subspaces: Generalizations and Connection to Difference Coarrays

Chun-Lin Liu, *Student Member, IEEE,* and P. P. Vaidyanathan, *Fellow, IEEE*

***Abstract*—Direction-of-arrival (DOA) estimation finds applications in many areas of science and engineering. In these applications, sparse arrays like minimum redundancy arrays, nested arrays, and coprime arrays, can be exploited to resolve $O(N^2)$ uncorrelated sources using N physical sensors. Recently, it has been shown that *correlation subspaces*, which reveal the structure of the covariance matrix, help to improve some existing DOA estimators. However, the bases, the dimension, and other theoretical properties of correlation subspaces remain to be investigated. This paper proposes *generalized correlation subspaces* in one and multiple dimensions. This leads to new insights into correlation subspaces and DOA estimation with prior knowledge. First, it is shown that the bases and the dimension of correlation subspaces are fundamentally related to *difference coarrays*, which were previously found to be important in the study of sparse arrays. Furthermore, generalized correlation subspaces can handle certain forms of prior knowledge about source directions. These results allow one to derive a broad class of DOA estimators with improved performance. It is demonstrated through examples that using sparse arrays and generalized correlation subspaces, DOA estimators with source priors exhibit better estimation performance than those without priors, in extreme cases like low SNR and limited snapshots.**

***Index Terms*—DOA estimation, sparse arrays, difference coarrays, generalized correlation subspaces, discrete prolate spheroidal sequences.**

I. Introduction

Direction-of-arrival (DOA) estimation has been a popular research field in array processing for several decades. This problem aims to estimate the source directions from sensor measurements. It arises in many practical scenarios such as radio astronomy, radar, imaging, and communications [1]–[7]. DOA estimators such as MUSIC [8], ESPRIT [9], MODE [10], and SPICE [11]–[13], to name a few [14]–[16] have been developed for this purpose. These estimators are functional for uniform linear arrays (ULAs). It is known that, for an ULA with N physical sensors, these DOA estimators can identify at most $N-1$ uncorrelated sources [6].

Another family of DOA estimators is applicable to sparse arrays, where the sensors are placed nonuniformly. Some well-known sparse arrays include minimum redundancy arrays [17], nested arrays [18], and coprime arrays [19]–[21]. All these arrays have $O(N^2)$ consecutive elements in the *difference coarray*, with N being the number of physical sensors. Due to this property, sparse arrays can identify $O(N^2)$ uncorrelated sources using N physical sensors and they lead to better spatial resolution than ULAs. DOA estimators for sparse arrays consist of the augmented covariance matrix method [22], [23], positive definite Toeplitz completion [24], [25], spatial smoothing MUSIC (SS MUSIC) [18], [26], [27], dictionary methods [28]–[31], and gridless approaches [32]–[34]. Note that these estimators are also applicable to ULAs.

Recently, the elegant concept of *correlation subspaces* was proposed by Rahmani and Atia [16], to improve DOA estimation in a number of ways. For *any given array geometry*, the correlation subspace is uniquely determined, and imposes some implicit constraints on the structure of the covariance, as we shall see. This subspace can be utilized in denoising the sample covariance matrix. Then the source directions are estimated from the denoised covariance matrix using off-the-shelf DOA estimators, such as the MUSIC algorithm. Note that the correlation subspace depends on the array configurations and prior knowledge about the sources but is independent of the choice of DOA estimators. Hence, a broad class of DOA estimators are applicable to the denoised covariance matrix. However, the explicit expressions for the correlation subspace were not known, so its approximation was computed numerically in [16]. Furthermore, the way in which the correlation subspace is influenced by the array configuration, and by partial knowledge about sources, was not explored.

Inspired by the concept of correlation subspaces introduced in [16], this paper makes a number of new contributions. To analyze the correlation subspace for any array configuration explicitly, we first generalize the definition in [16] to formulate what we call the *generalized correlation subspace*. This makes it possible to incorporate some types of apriori information on source locations, leading to improvements in DOA estimation. Furthermore, we show that the (generalized) correlation subspaces can be uniquely characterized in terms of the difference coarray of the original physical array. In fact we will show how to obtain simple and elegant closed form expressions for the basis vectors of the correlation subspace, in terms of the sensor geometry and the difference coarray geometry. Furthermore, if source directions belong to *apriori known intervals* [1], [5], [34]–[36], then it is shown that the generalized correlation subspace finds close connections to *discrete prolate spheroidal sequences* [37] defined on the difference coarrays. Similar results can be developed in multiple dimensions, and are useful in angle-Doppler estimation [38], [39], angle-delay estimation [40], angle-range estimation [41], 2D DOA estimation [6], and harmonic retrieval [42], [43]. These results not only facilitate the implementation of the denoising framework as in [16] but also offer better understanding of several DOA estimators with prior knowledge about sources.

This work was supported in parts by the ONR grant N00014-15-1-2118, the California Institute of Technology, and the Taiwan/Caltech Ministry of Education Fellowship.

The original work on correlation subspaces [16] did not emphasize any knowledge about the sources except that they are uncorrelated[1]. In our work, we show how to incorporate prior knowledge about source intervals with generalized correlation subspaces. Furthermore, correlation subspaces for 1D and 2D arrays were analyzed numerically in [16] while our work provides closed-form characterizations of generalized correlation subspaces for 1D and multidimensional arrays. Finally, in our work, covariance matrix denoising using the generalized correlation subspace is implemented more efficiently than that in [16].

The outline of this paper is as follows: Section II reviews correlation subspaces. Section III proposes the generalized correlation subspaces while Section IV derives their expressions for unknown and known source intervals. Section V discusses the connections with existing methods while Section VI studies the generalized correlation subspace for multidimensional arrays. Section VII presents several examples and numerical simulations to demonstrate the advantages of the new methods while Section VIII concludes this paper.

Notations: Scalars, vectors, matrices, and sets are denoted by lowercase letters (a), lowercase letters in boldface ($\mathbf{a}$), uppercase letters in boldface ($\mathbf{A}$), and letters in blackboard boldface ($\mathbb{A}$), respectively. An R-dimensional vector $\mathbf{n}$ is denoted by $(n^{(1)}, n^{(2)}, \ldots, n^{(R)})$ or $[n^{(1)}, n^{(2)}, \ldots, n^{(R)}]^T$, where $n^{(r)}$ is the rth coordinate. $\mathbf{A}^T$, $\mathbf{A}^*$, and $\mathbf{A}^H$ are the transpose, complex conjugate, and complex conjugate transpose of $\mathbf{A}$. The Moore-Penrose pseudoinverse of $\mathbf{A}$ is $\mathbf{A}^\dagger$. If $\mathbf{A}$ has full column rank, then $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$. The Kronecker product between $\mathbf{A}$ and $\mathbf{B}$ is written as $\mathbf{A}\otimes\mathbf{B}$. The vectorization operator is defined as $\mathrm{vec}([\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N]) = [\mathbf{a}_1^T, \mathbf{a}_2^T, \ldots, \mathbf{a}_N^T]^T$, where $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N$ are column vectors. For two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \succeq \mathbf{B}$ is equivalent to $\mathbf{A} - \mathbf{B}$ being positive semidefinite. $\mathrm{col}(\mathbf{A})$ stands for the column space of $\mathbf{A}$. The support of a function $f(x)$ is defined as $\mathrm{supp}(f) = \{x : f(x) \neq 0\}$, where x belongs to the domain of f. The indicator function $\mathbf{1}_{\mathbb{A}}(x)$ is one if $x \in \mathbb{A}$ and zero otherwise. $\mathbb{E}[\cdot]$ is the expectation operator. The cardinality of a set $\mathbb{A}$ is denoted by $|\mathbb{A}|$.

The bracket notation [27], [44] is reviewed using the following example. Assume the sensor locations are characterized by an integer set $\mathbb{S} = \{-2, -1, 0\}$. Assume the sensor measurement on $\mathbb{S}$ is denoted by $\mathbf{x}_{\mathbb{S}} = [10, 11, 12]^T$. The square bracket $[\mathbf{x}_{\mathbb{S}}]_i$ represents the ith entry of $\mathbf{x}_{\mathbb{S}}$ while the triangular bracket $\langle \mathbf{x}_{\mathbb{S}} \rangle_n$ denotes the sample value on the support location n. Hence, we have $[\mathbf{x}_{\mathbb{S}}]_1 = 10$, $[\mathbf{x}_{\mathbb{S}}]_2 = 11$, $[\mathbf{x}_{\mathbb{S}}]_3 = 12$, $\langle \mathbf{x}_{\mathbb{S}} \rangle_{-2} = 10$, $\langle \mathbf{x}_{\mathbb{S}} \rangle_{-1} = 11$, and $\langle \mathbf{x}_{\mathbb{S}} \rangle_0 = 12$. Similar notations apply to matrices. For instance, if $\mathbf{A} = \mathbf{x}_{\mathbb{S}}\mathbf{x}_{\mathbb{S}}^T$, then $[\mathbf{A}]_{i,j} = [\mathbf{x}_{\mathbb{S}}]_i[\mathbf{x}_{\mathbb{S}}]_j$ and $\langle \mathbf{A} \rangle_{n_1,n_2} = \langle \mathbf{x}_{\mathbb{S}} \rangle_{n_1} \langle \mathbf{x}_{\mathbb{S}} \rangle_{n_2}$.

II. Review of Correlation Subspaces

Assume that D monochromatic sources impinge on an one-dimensional sensor array[2]. The sensor locations are nd, where n belongs to an integer set $\mathbb{S} \subset \mathbb{Z}$, $d = \lambda/2$, and λ is the wavelength. Let $\theta_i \in [-\pi/2, \pi/2]$ be the DOA of the ith source. The normalized DOA of the ith source is defined as $\bar{\theta}_i = (d/\lambda)\sin\theta_i \in [-1/2, 1/2]$. The measurements on the sensor array $\mathbb{S}$ can be modeled as

$$\mathbf{x}_{\mathbb{S}} = \sum_{i=1}^{D} A_i \mathbf{v}_{\mathbb{S}}(\bar{\theta}_i) + \mathbf{n}_{\mathbb{S}} \quad \in \mathbb{C}^{|\mathbb{S}|}, \tag{1}$$

where A_i is the complex amplitude of the ith source, $\mathbf{v}_{\mathbb{S}}(\bar{\theta}_i) = [e^{j2\pi\bar{\theta}_i n}]_{n\in\mathbb{S}}$ are the steering vectors, and $\mathbf{n}_{\mathbb{S}}$ is the additive noise term. It is assumed that the sources and noise are zero-mean and *uncorrelated*. Namely, let $\mathbf{s} = [A_1, \ldots, A_D, \mathbf{n}_{\mathbb{S}}^T]^T$. Then $\mathbb{E}[\mathbf{s}] = \mathbf{0}$ and $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathrm{diag}(p_1, \ldots, p_D, p_n, \ldots, p_n)$, where p_i and p_n are the power of the ith sources and the noise, respectively.

The covariance matrix of $\mathbf{x}_{\mathbb{S}}$ can be expressed as

$$\mathbf{R}_{\mathbb{S}} = \mathbb{E}[\mathbf{x}_{\mathbb{S}}\mathbf{x}_{\mathbb{S}}^H] = \sum_{i=1}^{D} p_i \mathbf{v}_{\mathbb{S}}(\bar{\theta}_i)\mathbf{v}_{\mathbb{S}}^H(\bar{\theta}_i) + p_n\mathbf{I}. \tag{2}$$

Rearranging the elements in (2) leads to

$$\mathrm{vec}(\mathbf{R}_{\mathbb{S}} - p_n\mathbf{I}) = \sum_{i=1}^{D} p_i \mathbf{c}(\bar{\theta}_i), \tag{3}$$

where the correlation vectors $\mathbf{c}(\bar{\theta}_i)$ are defined as

$$\mathbf{c}(\bar{\theta}_i) \triangleq \mathrm{vec}(\mathbf{v}_{\mathbb{S}}(\bar{\theta}_i)\mathbf{v}_{\mathbb{S}}^H(\bar{\theta}_i)) \quad \in \mathbb{C}^{|\mathbb{S}|^2}. \tag{4}$$

The relation (3) implies

$$\mathrm{vec}(\mathbf{R}_{\mathbb{S}} - p_n\mathbf{I}) \in \mathrm{span}\{\mathbf{c}(\bar{\theta}_i) : i = 1, 2, \ldots, D\} \tag{5}$$

$$\subseteq \mathcal{CS} \triangleq \mathrm{span}\{\mathbf{c}(\bar{\theta}) : -1/2 \leq \bar{\theta} \leq 1/2\}, \tag{6}$$

where the linear span in (6) is defined as the set of *all* vectors of the form $\sum_{p=1}^{P} a_p \mathbf{c}(\bar{\theta}_p)$ where $P \in \mathbb{N}$, $a_p \in \mathbb{C}$, and $-1/2 \leq \bar{\theta}_p \leq 1/2$ [45]. This subspace is called the *correlation subspace*, denoted by $\mathcal{CS}$. Eq. (6) also indicates that $\mathrm{vec}(\mathbf{R}_{\mathbb{S}} - p_n\mathbf{I})$ is constrained in a certain way by $\mathcal{CS}$, and these constraints can be used in designing DOA estimators for improved performance.

It is clear that $\mathcal{CS}$ is a finite-dimensional subspace of $\mathbb{C}^{|\mathbb{S}|^2}$, due to (4). However, the definition of the correlation subspace in (6) is computationally intractable since it involves infinitely many $\mathbf{c}(\bar{\theta})$. The correlation subspace was originally computed by the following definition [16]:

Definition 1: The correlation subspace $\mathcal{CS}$ satisfies

$$\mathcal{CS} = \mathrm{col}(\mathbf{S}), \tag{7}$$

where the correlation subspace matrix $\mathbf{S}$ is defined as

$$\mathbf{S} \triangleq \int_{-\pi/2}^{\pi/2} \mathbf{c}(\bar{\theta})\mathbf{c}^H(\bar{\theta})\mathrm{d}\theta \quad \in \mathbb{C}^{|\mathbb{S}|^2\times|\mathbb{S}|^2}. \tag{8}$$

In Appendix A, we show that this definition is equivalent to our revised definition given in (6). Note that this integral is carried out over the DOA, $\theta \in [-\pi/2, \pi/2]$ and the relation $\bar{\theta} = (\sin\theta)/2$ can be utilized to evaluate (8). According to (8), it can be shown that the correlation subspace matrix $\mathbf{S}$ is Hermitian and positive semidefinite.

[1] A numerical simulation with prior knowledge about source intervals was shown in [16, Section IV-B2] but this idea was not developed further.

[2] For simplicity, we first assume 1D arrays. The multidimensional results will be presented in Section VI.

It was shown in [16] that, the right-hand side of (7) can be simplified further, based on the eigenvectors of $\mathbf{S}$ associated with the nonzero eigenvalues. In particular, let the eigen-decomposition of $\mathbf{S}$ be

$$\mathbf{S} = \underbrace{\begin{bmatrix}\mathbf{Q}_{\mathcal{CS}} & \mathbf{Q}_{\mathcal{CS}^\perp}\end{bmatrix}}_{\mathbf{Q}} \begin{bmatrix}\mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{Q}_{\mathcal{CS}} & \mathbf{Q}_{\mathcal{CS}^\perp}\end{bmatrix}^H, \quad (9)$$

where the diagonal matrix $\mathbf{\Lambda}_1$ contains the positive eigenvalues in the descending order and the columns of $\mathbf{Q}$ consist of the orthonormal eigenvectors. Then, (7) and (9) lead to $\mathcal{CS} = \mathrm{col}(\mathbf{Q}_{\mathcal{CS}})$. Namely, the correlation subspace $\mathcal{CS}$ is the column space of the matrix $\mathbf{Q}_{\mathcal{CS}}$. Eqs. (7), (8), and (9) indicate that the matrix $\mathbf{S}$, its eigenvalues, its eigenvectors, and the correlation subspace *depend purely on the array configuration*.

For any array geometry, the correlation subspace is uniquely determined, and imposes some implicit constraints on the structure of the covariance matrix, as indicated in (6). This leads to a covariance-matrix denoising approach [16]. To begin with, consider finite snapshot sensor measurements $\widetilde{\mathbf{x}}_{\mathbb{S}}(k)$ for $k = 1, \ldots, K$. The sample covariance matrix $\widetilde{\mathbf{R}}_{\mathbb{S}}$ can be estimated by

$$\widetilde{\mathbf{R}}_{\mathbb{S}} = \frac{1}{K}\sum_{k=1}^{K} \widetilde{\mathbf{x}}_{\mathbb{S}}(k)\widetilde{\mathbf{x}}_{\mathbb{S}}^H(k). \quad (10)$$

The algorithm in [16] first denoises the sample covariance matrix $\widetilde{\mathbf{R}}_{\mathbb{S}}$ using the following convex program (P1):

$$\text{(P1)}: \mathbf{R}_{\text{P1}}^\star \triangleq \arg\min_{\mathbf{R}} \ \|\widetilde{\mathbf{R}}_{\mathbb{S}} - p_n\mathbf{I} - \mathbf{R}\|_2^2 \quad (11)$$

$$\text{subject to } (\mathbf{I} - \mathbf{Q}_{\mathcal{CS}}\mathbf{Q}_{\mathcal{CS}}^\dagger)\mathrm{vec}(\mathbf{R}) = \mathbf{0}, \quad (12)$$

$$\mathbf{R} \succeq \mathbf{0}, \quad (13)$$

where the noise power p_n is estimated from the eigenvalues of $\widetilde{\mathbf{R}}_{\mathbb{S}}$ and $\|\cdot\|_2$ denotes the spectral norm of a matrix (i.e., the largest singular value). The cost function in (11) suggests that the matrix $\mathbf{R}_{\text{P1}}^\star$ resembles $\sum_{i=1}^{D} p_i \mathbf{v}_{\mathbb{S}}(\bar{\theta}_i)\mathbf{v}_{\mathbb{S}}^H(\bar{\theta}_i)$ in (2). The constraint (12) ensures that $\mathrm{vec}(\mathbf{R}_{\text{P1}}^\star)$ belongs to the correlation subspace while (13) indicates that $\mathbf{R}_{\text{P1}}^\star$ is positive semidefinite.

The final stage is to perform DOA estimation on $\mathbf{R}_{\text{P1}}^\star$. It was shown in [16] that, the MUSIC algorithm on $\mathbf{R}_{\text{P1}}^\star$ can exhibit better estimation performance than that on $\widetilde{\mathbf{R}}_{\mathbb{S}}$. It should be noted that *the DOA estimators are not restricted to the MUSIC algorithm*. Other estimators, such as ESPRIT, MODE, and SPICE, can also be exploited. This result shows that the structure of the covariance matrix, as specified by the correlation subspace, helps to improve the estimation performance.

Summarizing, the DOA estimator associated with the correlation subspace is composed of the following three steps [16]:

Step 1: Correlation subspace. For a given array geometry, *numerically* calculate the correlation subspace matrix $\mathbf{S}$ and its eigen-decomposition, as in (8) and (9), respectively. Save the matrix $\mathbf{Q}_{\mathcal{CS}}$ for the correlation subspace.

Step 2: Denoising. Given the sensor measurements, evaluate the sample covariance matrix $\widetilde{\mathbf{R}}_{\mathbb{S}}$, as in (10). Then solve the optimization problem (P1). Let the optimal solution be $\mathbf{R}_{\text{P1}}^\star$.

Step 3: DOA estimation. In the simulations of [16], the MUSIC algorithm is applied to $\mathbf{R}_{\text{P1}}^\star$.

Remarks on Step 1: Note that this step needs to be done only once per array. Once the matrix $\mathbf{Q}_{\mathcal{CS}}$ is obtained, it can be used repeatedly in Step 2. To calculate $\mathbf{Q}_{\mathcal{CS}}$, the numerical integration was utilized in [16]. This step is typically done by choosing a dense grid of the parameter θ, which only approximates the integral in (8). Furthermore, the *numerical* eigen-decomposition in (9) introduces perturbations on zero eigenvalues, making it challenging to determine the correlation subspace precisely. It is desirable to mitigate the negative effects caused by numerical computations. It will be shown in Theorem 1 that, the correlation subspace can be fully characterized by simple, elegant, and closed-form expressions.

Remarks on Step 2: The convex optimization problem (P1) can be solved by numerical solvers. However, it requires several numerical iterations to obtain the optimal solution, which could be an issue if real-time processing is needed. To avoid high computational cost, (P1) is approximated by two sub-problems (P1a) and (P1b) in [16]:

$$\text{(P1a)}: \mathbf{R}_{\text{P1a}}^\star \triangleq \arg\min_{\mathbf{R}} \ \|\widetilde{\mathbf{R}}_{\mathbb{S}} - p_n\mathbf{I} - \mathbf{R}\|_F^2 \quad (14)$$

$$\text{subject to } (\mathbf{I} - \mathbf{Q}_{\mathcal{CS}}\mathbf{Q}_{\mathcal{CS}}^\dagger)\mathrm{vec}(\mathbf{R}) = \mathbf{0}, \quad (15)$$

$$\text{(P1b)}: \mathbf{R}_{\text{P1b}}^\star \triangleq \arg\min_{\mathbf{R}} \ \|\mathbf{R} - \mathbf{R}_{\text{P1a}}^\star\|_2^2 \quad (16)$$

$$\text{subject to } \mathbf{R} \succeq \mathbf{0}. \quad (17)$$

In particular, we first compute the solution to (P1a) using the orthogonal projection onto the correlation subspace, $\mathrm{vec}(\mathbf{R}_{\text{P1a}}^\star) = \mathbf{Q}_{\mathcal{CS}}\mathbf{Q}_{\mathcal{CS}}^\dagger\mathrm{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}} - p_n\mathbf{I})$. Then the solution $\mathbf{R}_{\text{P1b}}^\star$ can be obtained explicitly from the eigen-decomposition of $\mathbf{R}_{\text{P1a}}^\star$. It was demonstrated in [16] that this two-step approach can be readily implemented with a moderate degradation in the estimation performance.

It can be seen that we need to estimate the noise power p_n first before solving either (P1) or (P1a). This extra processing is not desirable if the source directions are the only parameters of interest. In Section V, we will present another optimization problem that enjoys good DOA estimation performance without estimating the noise power.

Remarks on Step 3: If $\mathbf{R}_{\text{P1}}^\star \in \mathbb{C}^{|\mathbb{S}|\times|\mathbb{S}|}$, then the MUSIC algorithm on $\mathbf{R}_{\text{P1}}^\star$ can resolve at most $|\mathbb{S}| - 1$ uncorrelated sources, regardless of array configurations [6]. However, sparse arrays can identify more uncorrelated sources than sensors by using augmented covariance matrices [23], positive definite Toeplitz completion [24], [25], and spatial smoothing MUSIC (SS MUSIC) [18], [19], [26], [27]. Furthermore, these approaches usually provide better spatial resolution than the MUSIC algorithm [18], [19], [23]–[27]. Hence, we will use sparse arrays and SS MUSIC in the examples of Section VII.

Furthermore, it was demonstrated numerically in [16] that, prior knowledge about source directions can be embedded into the correlation subspace by changing the intervals of integration in (8). However, the influences of prior knowledge about the correlation subspace and the optimization problem have not been studied in detail. These above points will be addressed by generalized correlation subspaces, as we will present next.

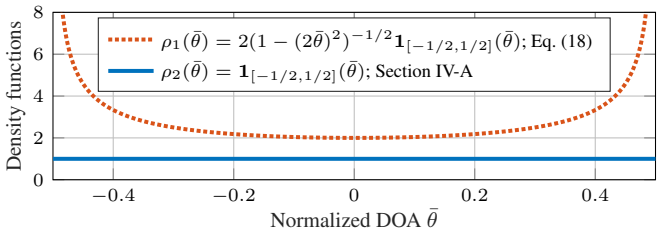


Fig. 1. The density function in (18) (red), and the constant density function in Section IV-A (blue).

III. Generalized Correlation Subspaces

The main difficulty in deriving the closed-form expressions for $\mathcal{CS}$ using (7), (8), and (9) is as follows. Eq. (8) implies that the entries of $\mathbf{S}$ are related to Bessel functions, making it complicated to obtain analytical forms of (9). In order to derive closed-form expressions for $\mathcal{CS}$, we will first propose the generalized correlation subspace $\mathcal{GCS}(\rho)$, which is parameterized by a nonnegative density function $\rho(\bar{\theta})$. Then it will be shown that $\mathcal{GCS}(\rho)$ depends only on the support of $\rho(\bar{\theta})$, and is invariant to its exact shape. Using this property, we can derive simple and closed forms for $\mathcal{CS}$ by selecting some density functions $\rho(\bar{\theta})$ such that 1) $\mathcal{CS} = \mathcal{GCS}(\rho)$ and 2) it is straightforward to derive closed-form expressions for $\mathcal{GCS}(\rho)$.

As a motivating example, let us consider the definition of $\mathbf{S}$ in (8). Since $\bar{\theta} = 0.5\sin\theta$, we have $\mathrm{d}\theta = 2(1-(2\bar{\theta})^2)^{-1/2}\mathrm{d}\bar{\theta}$. Hence, (8) can be rewritten as

$$\mathbf{S} = \int_{-1/2}^{1/2} \mathbf{c}(\bar{\theta})\mathbf{c}^H(\bar{\theta}) \underbrace{\left(2(1-(2\bar{\theta})^2)^{-1/2}\right)}_{\text{the density function}} \mathrm{d}\bar{\theta}. \tag{18}$$

Note that (18) can be regarded as a weighted integral with the density function $2(1-(2\bar{\theta})^2)^{-1/2}$ over $\bar{\theta} \in [-1/2, 1/2]$. Hence, we can generalize the correlation subspace matrix by varying the density function in (18). It is formally defined as

Definition 2: Let the correlation vector $\mathbf{c}(\bar{\theta})$ be defined as in (4). Let $\rho(\bar{\theta})$ be a nonnegative Lebesgue integrable function over the set $[-1/2, 1/2]$. The generalized correlation subspace matrix associated with $\rho(\bar{\theta})$ is defined as

$$\mathbf{S}(\rho) = \int_{-1/2}^{1/2} \mathbf{c}(\bar{\theta})\mathbf{c}^H(\bar{\theta})\rho(\bar{\theta})\mathrm{d}\bar{\theta}. \tag{19}$$

It can be seen that (18) is a special case of Definition 2, with $\rho(\bar{\theta}) = 2(1-(2\bar{\theta})^2)^{-1/2}\mathbf{1}_{[-1/2,1/2]}(\bar{\theta})$. The density function $\rho(\bar{\theta})$ quantifies the importance of $\mathbf{c}(\bar{\theta})\mathbf{c}^H(\bar{\theta})$ in $\mathbf{S}(\rho)$, across different $\bar{\theta}$. This $\rho(\bar{\theta})$ is shown in the dashed curve of Fig. 1. Note that $\rho(\bar{\theta})$ grows rapidly as $\bar{\theta}$ approaches ± 0.5.

Based on Definition 1, the generalized correlation subspace can be defined as follows:

Definition 3: Let $\mathbf{S}(\rho)$ be the generalized correlation subspace matrix associated with $\rho(\bar{\theta})$, as in (19). The generalized correlation subspace is defined as $\mathcal{GCS}(\rho) = \mathrm{col}(\mathbf{S}(\rho))$.

It can be seen from Definition 2 and 3 that the generalized correlation subspaces are parameterized by the density function $\rho(\bar{\theta})$. For any given support of $\rho(\bar{\theta})$, the generalized correlation subspace is invariant to the exact shape of $\rho(\bar{\theta})$ under that support, as indicated by the following lemma:

Lemma 1: Let $\rho_1(\bar{\theta})$ and $\rho_2(\bar{\theta})$ be two nonnegative Lebesgue integrable functions over the set $[-1/2, 1/2]$. If $\mathrm{supp}(\rho_1) = \mathrm{supp}(\rho_2)$, then $\mathcal{GCS}(\rho_1) = \mathcal{GCS}(\rho_2)$.

Proof: See Appendix B. ∎

Corollary 1: Let the density function in (18) be $\rho_1(\bar{\theta}) = 2(1-(2\bar{\theta})^2)^{-1/2}\mathbf{1}_{[-1/2,1/2]}(\bar{\theta})$ and the constant density function be $\rho_2(\bar{\theta}) = \mathbf{1}_{[-1/2,1/2]}(\bar{\theta})$. Then $\mathcal{CS} = \mathcal{GCS}(\rho_1) = \mathcal{GCS}(\rho_2)$.

The density functions $\rho_1(\bar{\theta})$ and $\rho_2(\bar{\theta})$ are illustrated in Fig. 1. It can be observed that these density functions share the same support $[-1/2, 1/2]$. Furthermore, Corollary 1 also enables us to analyze the correlation subspace readily through the generalized correlation subspace $\mathcal{GCS}(\rho_2)$. The details will be developed in Section IV.

IV. Properties of Generalized Correlation Subspaces

In this section, the generalized correlation subspaces for several density functions will be investigated. It will be shown that ***the correlation subspace and the generalized correlation subspace depend on the difference coarray***. Furthermore, we will derive simple, explicit, and computationally tractable representations of the correlation subspace and the generalized correlation subspace in certain cases.

First, it is known from [44] that the correlation vector (4) can be rewritten as:

$$\mathbf{c}(\bar{\theta}_i) = \mathbf{J}\mathbf{v}_{\mathbb{D}}(\bar{\theta}_i), \tag{20}$$

where $\mathbf{v}_{\mathbb{D}}(\bar{\theta}_i) = [e^{j2\pi\bar{\theta}_i m}]_{m\in\mathbb{D}}$ are the steering vectors on the difference coarray. Here the difference coarray $\mathbb{D}$ and the matrix $\mathbf{J}$ [44] are defined as:

Definition 4 (Difference coarray): The difference coarray $\mathbb{D}$ contains the differences between the elements in $\mathbb{S}$, i.e., $\mathbb{D} = \{n_1 - n_2 : \forall\, n_1, n_2 \in \mathbb{S}\}$.

Definition 5 (The matrix $\mathbf{J}$): The binary matrix $\mathbf{J}$ has size $|\mathbb{S}|^2$-by-$|\mathbb{D}|$. The columns of $\mathbf{J}$ satisfy $\langle\mathbf{J}\rangle_{:,m} = \mathrm{vec}(\mathbf{I}(m))$ for $m \in \mathbb{D}$, where $\mathbf{I}(m) \in \{0,1\}^{|\mathbb{S}|\times|\mathbb{S}|}$ is given by

$$\langle\mathbf{I}(m)\rangle_{n_1,n_2} = \begin{cases} 1, & \text{if } n_1 - n_2 = m, \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

Here the bracket notation $\langle\cdot\rangle_{n_1,n_2}$ is defined in Section I.

Example 1: Assume the sensor locations are characterized by an integer set $\mathbb{S} = \{0, 2\}$. According to Definition 4, the difference coarray becomes $\mathbb{D} = \{-2, 0, 2\}$. Next we will evaluate the matrix $\mathbf{J}$, as in Definition 5. First we consider the matrices $\mathbf{I}(m)$ for $m \in \mathbb{D}$ as follows:

$$\mathbf{I}(0) = \begin{matrix} \\ n_1{=}0 \\ 2 \end{matrix}\!\!\begin{matrix} \begin{matrix} n_2{=}0 & 2 \end{matrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{matrix}, \quad \mathbf{I}(2) = \begin{matrix} \\ n_1{=}0 \\ 2 \end{matrix}\!\!\begin{matrix} \begin{matrix} n_2{=}0 & 2 \end{matrix} \\ \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \end{matrix},$$

where n_1 and n_2 are marked in the corresponding rows and columns. Furthermore, due to (21), it can be shown that $\mathbf{I}(-2) = (\mathbf{I}(2))^T$. Hence, the matrix $\mathbf{J}$ can be written as

$$\mathbf{J} = \begin{array}{c} \begin{matrix} \text{vec}(\mathbf{I}(-2)) & \text{vec}(\mathbf{I}(0)) & \text{vec}(\mathbf{I}(2)) \end{matrix} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}, \tag{22}$$

where the first, the second, and the third column of $\mathbf{J}$ correspond to the coarray index $m = -2, 0, 2$, respectively. Finally, we will verify (20) in this example. Starting with (4), the correlation vector is given by

$$\mathbf{c}(\bar{\theta}_i) = \text{vec}\left(\begin{bmatrix} 1 \\ e^{j2\pi\bar{\theta}_i \cdot 2} \end{bmatrix} \begin{bmatrix} 1 & e^{-j2\pi\bar{\theta}_i \cdot 2} \end{bmatrix} \right) = \begin{bmatrix} 1 \\ e^{j4\pi\bar{\theta}_i} \\ e^{-j4\pi\bar{\theta}_i} \\ 1 \end{bmatrix}.$$

Similarly, the quantity $\mathbf{J}\mathbf{v}_{\mathbb{D}}(\bar{\theta}_i)$ can be calculated as

$$\mathbf{J}\mathbf{v}_{\mathbb{D}}(\bar{\theta}_i) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} e^{j2\pi\bar{\theta}_i \cdot (-2)} \\ 1 \\ e^{j2\pi\bar{\theta}_i \cdot 2} \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j4\pi\bar{\theta}_i} \\ e^{-j4\pi\bar{\theta}_i} \\ 1 \end{bmatrix}.$$

This result verifies (20).

Using (20) and Definition 2, the generalized correlation subspace matrix can be expressed in terms of the difference coarray as in Lemma 2:

Lemma 2: The generalized correlation subspace matrix satisfies $\mathbf{S}(\rho) = \mathbf{J}\mathbf{S}_{\mathbb{D}}(\rho)\mathbf{J}^H$, where $\mathbf{J}$ is defined in Definition 5 and $\mathbf{S}_{\mathbb{D}}(\rho)$ is given by

$$\mathbf{S}_{\mathbb{D}}(\rho) = \int_{-1/2}^{1/2} \mathbf{v}_{\mathbb{D}}(\bar{\theta})\mathbf{v}_{\mathbb{D}}^H(\bar{\theta})\rho(\bar{\theta})\mathrm{d}\bar{\theta}. \tag{23}$$

Note that the matrix $\mathbf{S}_{\mathbb{D}}(\rho)$ depends on *the difference coarray*, rather than the physical sensor locations. This property suggests that *the generalized correlation subspace is fundamentally related to the difference coarray*. This is indeed true and we will elaborate this point later. Furthermore, Lemma 2 also allows us to readily analyze the matrix $\mathbf{S}(\rho)$ of size $|\mathbb{S}|^2$-by-$|\mathbb{S}|^2$, by examining a smaller matrix $\mathbf{S}_{\mathbb{D}}(\rho)$ of size $|\mathbb{D}|$-by-$|\mathbb{D}|$. Next, we will consider two simple examples of the density function as follows:

- The density function is a constant over $\bar{\theta} \in [-1/2, 1/2]$. Namely, $\rho(\bar{\theta}) = \mathbf{1}_{[-1/2,1/2]}(\bar{\theta})$.
- The density function is a constant over some known intervals $\mathbb{I}$. That is, $\rho(\bar{\theta}) = \mathbf{1}_{\mathbb{I}}(\bar{\theta})$. This case corresponds to known source intervals.

In both cases, we will present the closed-form expressions of $\mathbf{S}_{\mathbb{D}}(\rho)$ and $\mathbf{S}(\rho)$, from which the generalized correlation subspaces can be analyzed systematically.

A. The constant density function

In this case, the entry of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-1/2,1/2]})$ associated with coarray locations $m_1, m_2 \in \mathbb{D}$ becomes

$$\langle \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-1/2,1/2]}) \rangle_{m_1,m_2} = \int_{-1/2}^{1/2} e^{j2\pi\bar{\theta}(m_1 - m_2)}\mathrm{d}\bar{\theta} = \delta_{m_1,m_2}, \tag{24}$$

since m_1 and m_2 are integers. Substituting (24) into Lemma 2 gives

$$\mathbf{S}(\mathbf{1}_{[-1/2,1/2]}) = \mathbf{J}\mathbf{J}^H. \tag{25}$$

In order to obtain the eigen-decomposition of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$, we invoke the definition of the weight function and a lemma regarding the matrix $\mathbf{J}$:

Definition 6 (Weight function $w(m)$): The weight function $w(m)$ of an array $\mathbb{S}$ is defined as the number of sensor pairs with coarray index m. Namely, $w(m) = |\{(n_1, n_2) \in \mathbb{S}^2 : n_1 - n_2 = m\}|$ for $m \in \mathbb{D}$.

Lemma 3: $\mathbf{J}^H\mathbf{J} = \mathbf{W} \triangleq \text{diag}(w(m))_{m\in\mathbb{D}}$. Namely, $\mathbf{J}$ has orthogonal columns and the norm of the column associated with the coarray index m is $\sqrt{w(m)}$.

Proof: It was proved in [44] that the columns of $\mathbf{J}$ are orthogonal. It suffices to consider the norms of the individual columns of $\mathbf{J}$. For the coarray location $m \in \mathbb{D}$, the norm of the associated column is

$$\begin{aligned} \|\text{vec}(\mathbf{I}(m))\|_2^2 &= \sum_{n_1,n_2\in\mathbb{S}} |\langle \mathbf{I}(m) \rangle_{n_1,n_2}|^2 \\ &= \text{the number of ones in } \mathbf{I}(m) = w(m), \end{aligned}$$

which proves this lemma. ∎

Example 2: Assume the sensor locations are given by $\mathbb{S} = \{0, 2\}$, as in Example 1. Then the weight functions $w(m)$ are given by

$$w(-2) = |\{(0,2)\}| = 1, \quad w(0) = |\{(0,0),(2,2)\}| = 2, \quad w(2) = |\{(2,0)\}| = 1.$$

Hence the matrix $\mathbf{W}$ in Lemma 3 can be written as $\mathbf{W} = \text{diag}(w(-2), w(0), w(2)) = \text{diag}(1, 2, 1)$. Then, Lemma 3 can be verified using (22) and $\mathbf{W}$.

Due to Definition 6 and Lemma 3, the generalized correlation subspace matrix $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$ can be expressed as

$$\mathbf{S}(\mathbf{1}_{[-1/2,1/2]}) = (\mathbf{J}\mathbf{W}^{-1/2})\mathbf{W}(\mathbf{J}\mathbf{W}^{-1/2})^H, \tag{26}$$

where the sizes of these terms are $\mathbf{J}\mathbf{W}^{-1/2} \in \mathbb{C}^{|\mathbb{S}|^2\times|\mathbb{D}|}$ and $\mathbf{W} \in \mathbb{C}^{|\mathbb{D}|\times|\mathbb{D}|}$. Eq. (26) also indicates that the matrix $\mathbf{J}\mathbf{W}^{-1/2}$ corresponds to the orthonormal eigenvectors while the diagonal matrix $\mathbf{W}$ is associated with the eigenvalues. In particular, the *positive eigenvalues* and the associated eigenvectors of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$ are given by

$$\text{Positive eigenvalues of } \mathbf{S}(\mathbf{1}_{[-1/2,1/2]}) = w(m), \tag{27}$$

$$\text{Eigenvectors of } \mathbf{S}(\mathbf{1}_{[-1/2,1/2]}) = \frac{\text{vec}(\mathbf{I}(m))}{\sqrt{w(m)}}, \tag{28}$$

where $m \in \mathbb{D}$. Note that (27) and (28) can be calculated readily from the array geometry using Definition 6 and 5, respectively. Namely, the eigen-decomposition of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$ can be

evaluated without using the numerical integration in Definition 2 and the numerical eigen-decomposition on $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$.

Properties of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$*:* The following list some properties regarding the eigen-structure of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$. Some items are direct consequences of the properties of the weight function $w(m)$, as in [46, Lemma 1].

1) The eigenvalues of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$ are nonnegative integers, due to (26), (27), and [46, Lemma 1-1].
2) The number of nonzero eigenvalues is $|\mathbb{D}|$, which is the size of the difference coarray. Here repeated eigenvalues are counted separately.
3) The largest eigenvalue of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$ is $w(0) = |\mathbb{S}|$ with algebraic multiplicity 1. This is a direct consequence of [46, Lemma 1-4].
4) For any nonzero eigenvalue that is not the largest, the associated algebraic multiplicity is an *even* positive integer, due to [46, Lemma 1-2].
5) The eigenvalue zero has algebraic multiplicity $|\mathbb{S}|^2 - |\mathbb{D}|$.
6) The orthonormal eigenvectors associated with nonzero eigenvalues are given in (28).
7) For all the eigenvalues of $\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})$, the geometric multiplicities achieve the algebraic multiplicities.

Finally, based on (26), the generalized correlation subspace $\mathcal{GCS}(\mathbf{1}_{[-1/2,1/2]})$ becomes

$$\mathrm{col}(\mathbf{S}(\mathbf{1}_{[-1/2,1/2]})) = \mathrm{col}(\mathbf{J}\mathbf{W}^{-1/2}) = \mathrm{col}(\mathbf{J}). \tag{29}$$

The significance of (29) is that, the correlation subspace in (6) can be characterized in closed forms, as in the following theorem:

Theorem 1: Let the matrix $\mathbf{J}$ be defined as in Definition 5. Then the correlation subspace satisfies

$$\mathcal{CS} = \mathrm{col}(\mathbf{J}). \tag{30}$$

Proof: Corollary 1 indicates that $\mathcal{CS} = \mathcal{GCS}(\rho_1) = \mathcal{GCS}(\rho_2)$. The relation $\mathcal{GCS}(\rho_2) = \mathrm{col}(\mathbf{J})$ is due to (29). ∎

This theorem indicates that the correlation subspace is fully characterized by the binary matrix $\mathbf{J}$, which can be readily computed from *sensor locations and the difference coarray* using Definition 5. Namely, to compute the correlation subspace, the numerical integration (8) and the eigen-decomposition (9) can be avoided completely. Due to Theorem 1 and Lemma 3, the dimension of the correlation subspace is given by

Corollary 2: The dimension of the correlation subspace is the size of the difference coarray, i.e., $\dim(\mathcal{CS}) = |\mathbb{D}|$.

B. The constant density function with known source intervals

Here we will study the generalized correlation subspace with known source intervals. This scenario arises in practical applications such as stationary radar and the diagnosis of rotating machines in industrial environments [47].

Table I summarizes the procedure for the generalized correlation subspaces with known source intervals. For a given array configuration $\mathbb{S}$ and source intervals $\mathbb{I}$, we can calculate $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ explicitly using either (32), (38), or (42), as we shall explain later. The eigen-decomposition of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ suggests that $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ can be approximated by $\mathbf{\Psi}_L \cdot \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_L) \cdot \mathbf{\Psi}_L^H$, where

TABLE I
GENERALIZED CORRELATION SUBSPACES WITH KNOWN SOURCE INTERVALS

Input: Array configuration $\mathbb{S}$, source intervals $\mathbb{I}$, and error tolerance δ.
1) Evaluate the matrix $\mathbf{J}$ according to Definition 5.
2) Based on $\mathbb{I}$, compute $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ using either (32), (38), or (42).
3) Numerical eigen-decomposition: $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}}) = \mathbf{\Psi}\mathbf{\Lambda}\mathbf{\Psi}^H$.
a) Eigenvectors: $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{\lvert\mathbb{D}\rvert}]$; $\mathbf{\Psi}^H\mathbf{\Psi} = \mathbf{I}$.
b) Eigenvalues: $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{\lvert\mathbb{D}\rvert})$; $\lambda_1 \geq \cdots \geq \lambda_{\lvert\mathbb{D}\rvert} \geq 0$.
4) Determine a positive integer L using (44).
5) Construct $\mathbf{\Psi}_L = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_L]$.
Output: $\mathcal{GCS}(\mathbf{1}_{\mathbb{I}})$ is approximated by $\mathrm{col}(\mathbf{J}\mathbf{\Psi}_L)$.

the related quantities are given in Table I.[3] This property leads to an approximation of the generalized correlation subspace $\mathcal{GCS}(\mathbf{1}_{\mathbb{I}}) \approx \mathrm{col}(\mathbf{J}\mathbf{\Psi}_L)$. Note that Table I is applicable to a given array $\mathbb{S}$ and given source intervals $\mathbb{I}$. In the following development, we will study the generalized correlation subspaces based on these factors.

1) Hole-free difference coarrays and $\mathbb{I} = [-\alpha/2, \alpha/2]$*:* In this case, the density function is assumed to be

$$\rho(\bar{\theta}) = \mathbf{1}_{[-\alpha/2,\alpha/2]}(\bar{\theta}) = \begin{cases} 1, & \text{if } -\alpha/2 \leq \bar{\theta} \leq \alpha/2, \\ 0, & \text{otherwise,} \end{cases} \tag{31}$$

where $0 < \alpha < 1$. Now we can derive the expression for the generalized correlation subspace matrix. According to Lemma 2, the entry of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ associated with coarray locations $m_1, m_2 \in \mathbb{D}$ is given by

$$\begin{aligned} \langle \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) \rangle_{m_1,m_2} &= \int_{-\alpha/2}^{\alpha/2} e^{j2\pi\bar{\theta}(m_1-m_2)} \mathrm{d}\bar{\theta} \\ &= \alpha \cdot \mathrm{sinc}(\alpha(m_1 - m_2)), \end{aligned} \tag{32}$$

where the normalized sinc function $\mathrm{sinc}(x)$ is 1 for $x = 0$ and $\sin(\pi x)/(\pi x)$ otherwise. The eigen-decomposition of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ is assumed to be

$$\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) = \mathbf{\Psi}\mathbf{\Lambda}\mathbf{\Psi}^H, \tag{33}$$

where the matrices $\mathbf{\Psi}$ and $\mathbf{\Lambda}$ are given by $\mathbf{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{|\mathbb{D}|}]$ and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{|\mathbb{D}|})$, respectively. Here the eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\mathbb{D}|} \geq 0$ and $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_{|\mathbb{D}|}$ are the associated orthonormal eigenvectors.

Note that the matrix $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ depends purely on the difference coarray $\mathbb{D}$ and the parameter α. Next we assume the difference coarray $\mathbb{D}$ is *hole-free*. Namely, $\mathbb{D}$ is composed of consecutive integers. For instance, Fig. 2 depicts array configurations like (a) ULA [6], (b) nested array [18], (c) coprime array [19], and (d) super nested array [21]. It can be observed from the nonnegative part of the difference coarray that (a), (b), and (d) have hole-free difference coarrays while (c) does not. These array configurations will be elaborated in Section VII later.

[3]The symbols $\boldsymbol{\psi}_k$, $\mathbf{\Psi}_L$, and $\mathbf{\Psi}$ are reserved for the eigenvectors of the matrix $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ while the notations $\mathbf{v}_k$ and $\mathbf{V}_L$ represent the discrete prolate spheroidal sequences.

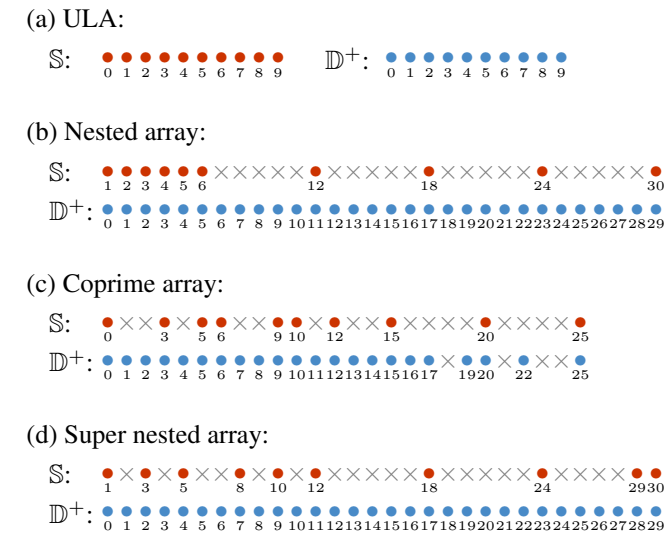


Fig. 2. The sensor locations $\mathbb{S}$ and the nonnegative part of the difference coarrays $\mathbb{D}^+$ for (a) ULA with 10 sensors, (b) the nested array with $N_1 = N_2 = 5$, (c) the coprime array with $M = 3, N = 5$, and (d) the super nested array with $N_1 = N_2 = 5, Q = 2$. Here bullets denote elements in $\mathbb{S}$ or $\mathbb{D}^+$ while crosses represent empty space.

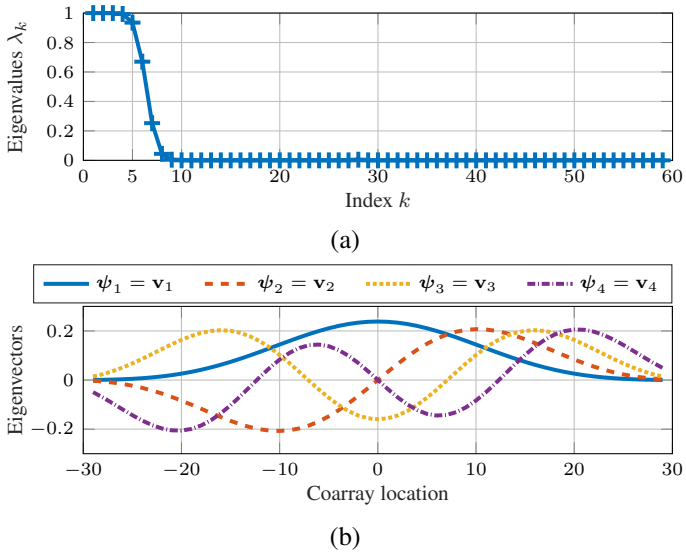


Fig. 3. (a) The eigenvalues and (b) the first four eigenvectors of the generalized correlation subspace matrix $\mathbf{S}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ in Example 3. Here the difference coarray $\mathbb{D} = \{-29, \ldots, 29\}$ and α is 0.1.

For arrays with hole-free difference coarrays, the eigenvectors ψ_k are known to be *discrete prolate spheroidal sequences (DPSS)*. That is, the matrix $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ owns

$$\text{Eigenvalues of } \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) = \lambda_k, \tag{34}$$

$$\text{Eigenvectors of } \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) \triangleq \boldsymbol{\psi}_k = \mathbf{v}_k, \tag{35}$$

where $\mathbf{v}_1, \mathbf{v}_1, \ldots, \mathbf{v}_{|\mathbb{D}|}$ *denote DPSS on the difference coarray* $\mathbb{D}$. Note that DPSS were studied comprehensively in [37], [48] and they arise in various fields such as multitapers [49], time-frequency analysis [50], eigenfilters [51], [52], and MIMO radar [53]. Here several properties of the eigenvalues λ_k and the DPSS $\mathbf{v}_k$ are reviewed briefly using the following example:

Example 3: Consider the super nested array with $N_1 = N_2 = 5$ [21]. The sensor locations are depicted in Fig. 2(d). The difference coarray becomes $\mathbb{D} = \{-29, \ldots, 29\}$ (hole-free) and $|\mathbb{D}| = 59$. We also choose the parameter $\alpha = 0.1$. According to (34) and (35), the eigenvalues λ_k and the eigenvectors are illustrated in Fig. 3. It was shown in [37] that the eigenvalues λ_k are distinct, and the first $\lfloor \alpha|\mathbb{D}| \rfloor$ eigenvalues are close to, but less than one, where $\lfloor \cdot \rfloor$ is the floor function. As the index k exceeds $\lfloor \alpha|\mathbb{D}| \rfloor$, the magnitude of the eigenvalues decays exponentially [37]. This property indicates that the matrix $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ can be well-approximated by a matrix of rank L. Namely,

$$\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) \approx \mathbf{V}_L \cdot \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_L) \cdot \mathbf{V}_L^H, \tag{36}$$

where $\lfloor \alpha|\mathbb{D}| \rfloor \le L \le |\mathbb{D}|$ and $\mathbf{V}_L = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_L]$ consists of the first L DPSS. Note that the exact value of L depends on the approximation error of (36), which will be elaborated in (44) later. In this example, $\lfloor \alpha|\mathbb{D}| \rfloor = 5$, This means that the first five eigenvalues are close to one, as depicted in Fig. 3(a). The DPSS $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ are illustrated in Fig. 3(b). These DPSS can be proved to be orthogonal, real, and unique up to scale [37]. Furthermore, these DPSS satisfy $\langle \mathbf{v}_k \rangle_{-m} = (-1)^{k+1} \langle \mathbf{v}_k \rangle_m$ for $m \in \mathbb{D}$. Namely, they are either even or odd symmetric.

Substituting (36) into Lemma 2, the generalized correlation subspace matrix can be approximated by

$$\mathbf{S}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) \approx \mathbf{J}\mathbf{V}_L \cdot \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_L) \cdot (\mathbf{J}\mathbf{V}_L)^H.$$

Since the matrix $\mathbf{J}\mathbf{V}_L$ has full column rank, the generalized correlation subspace can be characterized by the following theorem:

Theorem 2: Let the density function be $\mathbf{1}_{[-\alpha/2,\alpha/2]}$ for $0 < \alpha < 1$. Assume the difference coarray $\mathbb{D}$ is hole-free. Then we have

$$\mathcal{GCS}(\mathbf{1}_{[-\alpha/2,\alpha/2]}) \approx \mathrm{col}(\mathbf{J}\mathbf{V}_L), \tag{37}$$

where the columns of $\mathbf{V}_L$ contain the first L DPSS on $\mathbb{D}$. The parameter L is obtained by (44), for a given error tolerance δ.

The significance of Theorem 2 is that, the interval information α is embedded in the matrix $\mathbf{V}_L$. The approximation $\mathbf{V}_L$ is constructed by selecting the DPSS associated with the largest L eigenvalues of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[-1/2,1/2]})$. The details in choosing the parameter L will be developed in (43) further. Note that Theorem 2 can be utilized to denoise the sample covariance matrix, as we shall present in Section V.

2) Hole-free difference coarrays and $\mathbb{I} = [\bar{\theta}_{min}, \bar{\theta}_{max}]$*:* Let us consider another scenario where the known interval is not centered around the origin, such as $[\bar{\theta}_{\min}, \bar{\theta}_{\max}]$ with $-0.5 < \bar{\theta}_{\min} < \bar{\theta}_{\max} < 0.5$. In this case, the density function can be

$$\rho(\bar{\theta}) = \mathbf{1}_{[\bar{\theta}_{\min}, \bar{\theta}_{\max}]}(\bar{\theta}) = \begin{cases} 1, & \text{if } \bar{\theta}_{\min} \le \bar{\theta} \le \bar{\theta}_{\max}, \\ 0, & \text{otherwise.} \end{cases}$$

Then, due to Lemma 2, the entry of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[\bar{\theta}_{\min}, \bar{\theta}_{\max}]})$ associated with the coarray location $m_1, m_2 \in \mathbb{D}$ becomes

$$\begin{aligned} &\langle \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[\bar{\theta}_{\min}, \bar{\theta}_{\max}]}) \rangle_{m_1, m_2} \\ &= e^{j2\pi \bar{\theta}_{\text{avg}} m_1} \cdot \alpha \cdot \mathrm{sinc}(\alpha(m_1 - m_2)) \cdot e^{-j2\pi \bar{\theta}_{\text{avg}} m_2}, \end{aligned} \tag{38}$$

where $\bar{\theta}_{\text{avg}} = (\bar{\theta}_{\max} + \bar{\theta}_{\min})/2$ and $\alpha = \bar{\theta}_{\max} - \bar{\theta}_{\min}$ are known. Using similar arguments in Section IV-B1, the matrix $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[\bar{\theta}_{\min},\bar{\theta}_{\max}]})$ owns

$$\text{Eigenvalues of } \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[\bar{\theta}_{\min},\bar{\theta}_{\max}]}) = \lambda_k, \tag{39}$$

$$\text{Eigenvectors of } \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{[\bar{\theta}_{\min},\bar{\theta}_{\max}]}) = \text{diag}(e^{j2\pi\bar{\theta}_{\text{avg}}m})_{m\in\mathbb{D}} \times \mathbf{v}_k, \tag{40}$$

where λ_k are defined in (33) and $\mathbf{v}_k$ are the DPSS. These eigenvectors can be regarded as the modulated version of DPSS. Hence, following similar arguments in Example 3, the generalized correlation subspace can be approximated by

$$\mathcal{GCS}(\mathbf{1}_{[\bar{\theta}_{\min},\bar{\theta}_{\max}]}) \approx \text{col}(\mathbf{J}\cdot\text{diag}(e^{j2\pi\bar{\theta}_{\text{avg}}m})_{m\in\mathbb{D}}\cdot\mathbf{V}_L), \tag{41}$$

where the matrix $\mathbf{J}$ is defined in Definition 5. The matrix $\mathbf{V}_L = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_L]$ includes the first L DPSS, where L will be chosen as in (44).

3) Hole-free difference coarrays and unions of multiple source intervals: In this case, the interval information $\mathbb{I}$ is given by $\cup_{p=1}^{P}[\bar{\theta}_{\min,p}, \bar{\theta}_{\max,p}]$, where $[\bar{\theta}_{\min,p}, \bar{\theta}_{\max,p}]$ for $p = 1, 2, \ldots, P$ are non-overlapping intervals. The density function is assumed to be $\mathbf{1}_{\mathbb{I}}(\bar{\theta})$. Using (38) across all intervals yields

$$\langle \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})\rangle_{m_1,m_2} = \sum_{p=1}^{P} \alpha_p e^{j2\pi\bar{\theta}_{\text{avg},p}(m_1-m_2)}\text{sinc}(\alpha_p(m_1-m_2)), \tag{42}$$

where $\bar{\theta}_{\text{avg},p} = (\bar{\theta}_{\max,p} + \bar{\theta}_{\min,p})/2$ is the centroid of the pth interval and $\alpha_p = \bar{\theta}_{\max,p} - \bar{\theta}_{\min,p}$ is the width of the pth interval. Here the indices $m_1, m_2 \in \mathbb{D}$. Although the entries of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ are sums of modulated sinc functions, the eigenvectors of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ cannot be expressed in terms of DPSS in general. In this case, the generalized correlation subspace $\mathcal{GCS}(\mathbf{1}_{\mathbb{I}})$ has to be evaluated *numerically* using Table I.

4) Difference coarrays with holes: The results in Section IV-A hold true for 1D arrays, regardless of the difference coarrays. However, it is assumed in Section IV-B1 to IV-B3 that the difference coarrays are hole-free. These arrays with hole-free difference coarrays include ULA [6], minimum redundancy arrays [17], nested arrays [18], and super nested arrays [21]. For such arrays, Eqs. (37) and (41) are applicable.

However, for arrays containing holes in the difference coarrays, such as minimum hole arrays [54], coprime arrays [19], and some generalizations of coprime arrays [20], Eqs. (37) and (41) are not applicable in general. It is because the hole-free property of the difference coarray is used to derive (35) and (40). In this case, $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$ can still be computed from (32), (38), and (42) accordingly. Then, the generalized correlation subspaces need to be calculated numerically using Table I.

The choice of the parameter L: Here we will present the details on the parameter L in Section IV-B. Let $\mathbf{\Psi}_\ell\mathbf{\Lambda}_\ell\mathbf{\Psi}_\ell^H$ be the rank-ℓ approximation of $\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})$, where the notations are consistent with those in Table I. Let the matrix $\mathbf{\Lambda}_\ell$ be $\text{diag}(\lambda_1, \ldots, \lambda_\ell)$. Then the relative error $E(\ell)$ is

$$E(\ell) \triangleq \frac{\|\mathbf{\Psi}_\ell\mathbf{\Lambda}_\ell\mathbf{\Psi}_\ell^H - \mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})\|_F^2}{\|\mathbf{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{I}})\|_F^2} = \frac{\sum_{k=\ell+1}^{|\mathbb{D}|}\lambda_k^2}{\sum_{k=1}^{|\mathbb{D}|}\lambda_k^2}. \tag{43}$$

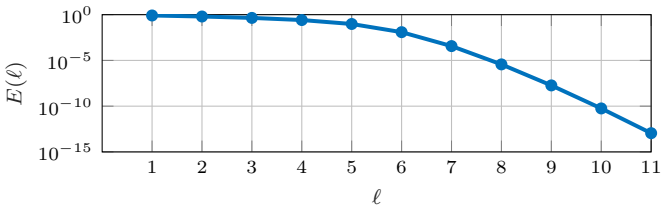


Fig. 4. The dependence of the relative error $E(\ell)$ on the parameter ℓ, where $E(\ell)$ is defined in (43) and the eigenvalues are shown in Fig. 3(a).

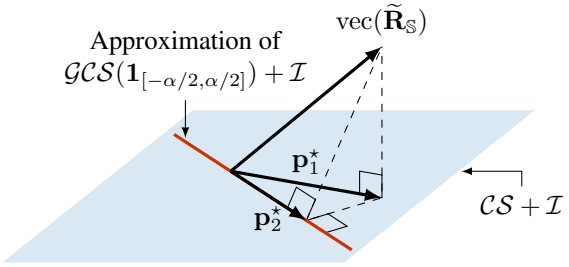


Fig. 5. The geometric interpretation of sample covariance matrix denoising using generalized correlation subspaces (Problem (P2)). The sample covariance matrix is denoted by $\widetilde{\mathbf{R}}_{\mathbb{S}}$. The vectors $\mathbf{p}_1^\star$ and $\mathbf{p}_2^\star$ are the orthogonal projections of $\text{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}})$ onto $\mathcal{CS}+\mathcal{I}$ and onto a subspace that approximates $\mathcal{GCS}(\mathbf{1}_{[-\alpha/2,\alpha/2]})+\mathcal{I}$, respectively. Here $\mathcal{I} = \text{span}(\text{vec}(\mathbf{I}))$ and the sum between subspaces $\mathcal{A}$ and $\mathcal{B}$ is defined as $\mathcal{A}+\mathcal{B} = \{a+b : a\in\mathcal{A}, b\in\mathcal{B}\}$.

For a given error tolerance $0 \le \delta \ll 1$, L is the minimum ℓ such that $E(\ell) < \delta$, i.e.,

$$L = \min_{\ell\in\mathbb{Z},\ 1\le\ell\le|\mathbb{D}|} \ell \quad \text{subject to} \quad E(\ell) < \delta. \tag{44}$$

Then this L is used in computing the generalized correlation subspace.

In particular, the smaller δ is, the larger L is. For instance, Fig. 4 plots the relative error $E(\ell)$, where the eigenvalues are given in Fig. 3(a). It can be seen that, if the error tolerance $\delta = 10^{-5}$, then the parameter $L = 8$. If $\delta = 10^{-10}$, then we have $L = 10$.

V. Connections with Existing Methods

In this section, we will discuss a covariance matrix denoising framework associated with the (generalized) correlation subspace. This method, denoted by problem (P2), can be regarded as a modified version of the optimization problem (P1). This problem (P2) can be solved by simple, closed-form, and computationally tractable expressions, unlike the problem (P1). We also relate it to *redundancy averaging*, which is a well-known processing technique in coarray-based DOA estimators, as shown in Example 4. This idea can be extended to the case of known source intervals, as in Example 5.

To develop some feelings for this method, Fig. 5 demonstrates the main idea of (P2), where the sample covariance matrix $\widetilde{\mathbf{R}}_{\mathbb{S}}$ is defined in (10). For a given element $\text{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}})$, we can calculate its orthogonal projection $\mathbf{p}_1^\star$ onto the subspace $\mathcal{CS}+\mathcal{I}$, where $\mathcal{I} = \text{span}(\text{vec}(\mathbf{I}))$ and the sum of subspaces $\mathcal{A}$ and $\mathcal{B}$ is defined as $\mathcal{A}+\mathcal{B} = \{a+b : a\in\mathcal{A}, b\in\mathcal{B}\}$. Furthermore, if the source interval is known to be $[-\alpha/2, \alpha/2]$, then the projection can be refined as $\mathbf{p}_2^\star$. These projections $\mathbf{p}_1^\star$ and $\mathbf{p}_2^\star$ can be matricized as indefinite Hermitian matrices, to which some existing DOA estimators can be applied.

The rationale for (P2) is based on the following chain of arguments. According to (6), we have $\mathrm{vec}(\mathbf{R}_{\mathbb{S}} - p_n\mathbf{I}) \in \mathcal{CS}$. This result implies that $\mathrm{vec}(\mathbf{R}_{\mathbb{S}})$ can be decomposed as $p_n\mathrm{vec}(\mathbf{I})$ plus a vector in $\mathcal{CS}$. Namely, $\mathrm{vec}(\mathbf{R}_{\mathbb{S}}) \in \mathcal{CS} + \mathcal{I}$, where $\mathcal{I} = \mathrm{span}(\mathrm{vec}(\mathbf{I}))$. Hence, in the finite snapshot scenario, we can find the vector $\mathbf{p}^\star$ in $\mathcal{CS}+\mathcal{I}$ that minimizes the Euclidean distance to $\mathrm{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}})$. More generally, if the prior knowledge about sources is available, as embedded in the generalized correlation subspace $\mathcal{GCS}(\rho)$, then we have $\mathrm{vec}(\mathbf{R}_{\mathbb{S}}) \in \mathcal{GCS}(\rho) + \mathcal{I}$. This idea leads to the following convex program:

$$\text{(P2): } \mathbf{p}^\star \triangleq \arg\min_{\mathbf{p}} \ \|\mathrm{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}}) - \mathbf{p}\|_2^2 \quad \text{subject to} \tag{45}$$

$$(\mathbf{I} - \mathbf{Q}_{\mathcal{GCS}(\rho)+\mathcal{I}}\mathbf{Q}^\dagger_{\mathcal{GCS}(\rho)+\mathcal{I}})\mathbf{p} = \mathbf{0}, \tag{46}$$

where the columns of $\mathbf{Q}_{\mathcal{A}}$ are the bases for the subspace $\mathcal{A}$. The solution to (P2) is given by

$$\mathbf{p}^\star = \mathbf{Q}_{\mathcal{GCS}(\rho)+\mathcal{I}}\mathbf{Q}^\dagger_{\mathcal{GCS}(\rho)+\mathcal{I}}\mathrm{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}}). \tag{47}$$

Note that (47) can be evaluated directly, given the sample covariance matrix $\widetilde{\mathbf{R}}_{\mathbb{S}}$ and the generalized correlation subspace $\mathcal{GCS}(\rho)$. The computational complexity of (47) is much less than solving (P2) numerically. It can be observed that (47) shares similar expressions with the solution to (P1a). The main difference is that, estimating the noise power p_n is required in (P1a), but not in (47).

Next, we will demonstrate some instances of (47) using the generalized correlation subspaces in Section IV.

Example 4: First let us consider the correlation subspace. Recall that $\mathcal{CS} = \mathcal{GCS}(\mathbf{1}_{[-1/2,1/2]}) = \mathrm{col}(\mathbf{J})$, as in Theorem 1. The subspace $\mathcal{CS}+\mathcal{I}$ becomes

$$\mathcal{CS} + \mathcal{I} = \mathrm{col}(\mathbf{J}) + \mathrm{span}(\mathrm{vec}(\mathbf{I})) = \mathrm{col}([\mathbf{J}, \mathbf{J}\mathbf{e}_0]) = \mathrm{col}(\mathbf{J}).$$

Here we use the property that $\mathrm{vec}(\mathbf{I}) = \mathbf{J}\mathbf{e}_0$ [44]. The column vector $\mathbf{e}_0 \in \{0,1\}^{|\mathbb{D}|}$ satisfies $\langle \mathbf{e}_0 \rangle_m = \delta_{m,0}$ for $m \in \mathbb{D}$. Next, according to (47), the orthogonal projection $\mathbf{p}_1^\star$, as shown in Fig. 5, can be written as

$$\mathbf{p}_1^\star = \mathbf{J}\widetilde{\mathbf{x}}_{\mathbb{D}} \quad \in \mathbb{C}^{|\mathbb{S}|^2}, \quad \widetilde{\mathbf{x}}_{\mathbb{D}} \triangleq \mathbf{J}^\dagger \mathrm{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}}) \quad \in \mathbb{C}^{|\mathbb{D}|}. \tag{48}$$

Due to (48) and Lemma 3, the sample value of $\widetilde{\mathbf{x}}_{\mathbb{D}}$ at the coarray location $m \in \mathbb{D}$ is given by

$$\langle \widetilde{\mathbf{x}}_{\mathbb{D}} \rangle_m = \frac{1}{w(m)} \sum_{n_1 - n_2 = m} \langle \widetilde{\mathbf{R}}_{\mathbb{S}} \rangle_{n_1,n_2}, \tag{49}$$

where $n_1, n_2 \in \mathbb{S}$. Eq. (49) was previously known as ***redundancy averaging*** [22], [24], [25] and [27, Definition 3]. The vector $\widetilde{\mathbf{x}}_{\mathbb{D}}$ is known to be the sample autocorrelation vector on the difference coarray, which was used extensively in DOA estimators such as the augmented covariance matrix [22] positive definite Toeplitz completion [24], [25], and SS MUSIC [18], [27]. This example shows that, *redundancy averaging is closely related to* (47)*, which uses the concept of the correlation subspace*.

Example 5: Eq. (47) can be used to derive a large class of DOA estimators with prior knowledge about source directions. According to Section IV-B, if the difference coarray is hole-free and the known source interval is $[-\alpha/2, \alpha/2]$, then the generalized correlation subspace $\mathcal{GCS}(\rho)$ can be approximated by $\mathrm{col}(\mathbf{J}\mathbf{V}_L)$, as in Theorem 2. Hence the subspace $\mathcal{GCS}(\rho)+\mathcal{I}$ is approximated by

$$\mathrm{col}(\mathbf{J}\mathbf{V}_L) + \mathrm{span}(\mathrm{vec}(\mathbf{I})) = \mathrm{col}([\mathbf{J}\mathbf{V}_L, \mathbf{J}\mathbf{e}_0]) = \mathrm{col}(\mathbf{J}\mathbf{U}_L),$$

where $\mathbf{U}_L \triangleq [\mathbf{V}_L, \mathbf{e}_0]$. Then the associated orthogonal projection $\mathbf{p}_2^\star$, as illustrated in Fig. 5, is given by

$$\mathbf{p}_2^\star = (\mathbf{J}\mathbf{U}_L)(\mathbf{J}\mathbf{U}_L)^\dagger \mathrm{vec}(\widetilde{\mathbf{R}}_{\mathbb{S}}) \tag{50}$$

Eq. (50) shows that the orthogonal projection can be evaluated using the matrix $\mathbf{J}$ and the DPSS $\mathbf{V}_L$. If the matrix $\mathbf{J}\mathbf{U}_L$ has full column rank, then Eq. (50) can rewritten as:

$$\mathbf{p}_2^\star = \mathbf{J}\widetilde{\mathbf{y}}_{\mathbb{D}}, \tag{51}$$

where

$$\widetilde{\mathbf{y}}_{\mathbb{D}} = \mathbf{U}_L((\mathbf{W}\mathbf{U}_L)^H\mathbf{U}_L)^{-1}(\mathbf{W}\mathbf{U}_L)^H\widetilde{\mathbf{x}}_{\mathbb{D}}, \tag{52}$$

and the matrix $\mathbf{W}$ is defined in Lemma 3. Here $\widetilde{\mathbf{x}}_{\mathbb{D}}$ is given by (49). Note that (51) shares the same formulation as the first equation of (48). Furthermore, according to (52), the vector $\widetilde{\mathbf{y}}_{\mathbb{D}}$ can be regarded as the denoised version of $\widetilde{\mathbf{x}}_{\mathbb{D}}$, as characterized by the oblique projection operator $\mathbf{U}_L((\mathbf{W}\mathbf{U}_L)^H\mathbf{U}_L)^{-1}(\mathbf{W}\mathbf{U}_L)^H$. Eq. (52) can also be interpreted as ***redundancy averaging with prior knowledge about sources***. Most importantly, coarray-based DOA estimators such as the augmented covariance matrix, positive definite Toeplitz completion, and SS MUSIC can work on the denoised sample autocorrelation vector $\widetilde{\mathbf{y}}_{\mathbb{D}}$, without any modification.

For other cases in Section IV-B, (51) and (52) remains applicable, except that the DPSS $\mathbf{V}_L$ have to be replaced with the eigenvectors $\mathbf{\Psi}_L$.

Covariance matrix denoising using the (generalized) correlation subspace is not limited to the optimization problems presented in this paper. The constraints imposed on the covariance matrices by the (generalized) correlation subspace can be readily applied to state-of-the-art covariance matrix denoising methods and DOA estimators such as SPICE [11]–[13], gridless SPICE [55], and atomic norm denoising [34]–[36], [56]. This idea could lead to new DOA estimators that enjoy improved performance, which could even approach the Cramér-Rao bounds for sparse arrays [44], [57], [58].

VI. Generalized Correlation Subspaces in Multiple Dimensions

The results developed in [16] were not restricted to one dimension. Even though our discussions in this paper were so far restricted to 1D arrays and 1D DOAs, they can also be readily generalized to multiple dimensions as we shall elaborate next. This is useful in many practical scenarios such as angle-Doppler estimation [38], [39], angle-delay estimation [40], angle-range estimation [41], 2D DOA (azimuth and elevation) estimation [6], and harmonic retrieval [42], [43].

Let us consider the data model for the R-dimensional (R-D) case. The R-D sampling locations $\boldsymbol{n} = (n^{(1)}, n^{(2)}, \ldots, n^{(R)})$ are collected by the set $\mathbb{S} \subset \mathbb{Z}^R$ and the harmonic parameters

to be estimated are denoted by $\bar{\boldsymbol{\mu}} = (\bar{\mu}^{(1)}, \bar{\mu}^{(2)}, \ldots, \bar{\mu}^{(R)}) \in [-1/2, 1/2]^R$. Then we have the following data model:

$$\boldsymbol{x}_{\mathbb{S}} = \sum_{i=1}^{D} A_i \boldsymbol{v}_{\mathbb{S}}(\bar{\boldsymbol{\mu}}_i) + \boldsymbol{n}_{\mathbb{S}}, \tag{53}$$

where the *vectors* $\boldsymbol{x}_{\mathbb{S}}$, $\boldsymbol{v}_{\mathbb{S}}(\bar{\boldsymbol{\mu}}_i)$, and $\boldsymbol{n}_{\mathbb{S}}$ represent the measurement, the steering vector, and the noise term, respectively. The entry of the steering vector associated with the sample locations $\boldsymbol{n} \in \mathbb{S}$ is given by $e^{j2\pi \bar{\boldsymbol{\mu}}_i^T \boldsymbol{n}}$. The source amplitudes A_i and the noise term $\boldsymbol{n}_{\mathbb{S}}$ are assumed to be uncorrelated, as in Section II. Then we can define the R-D difference coarray and the matrix $\boldsymbol{J}$ as follows:

Definition 7 (R-D difference coarray): Let $\mathbb{S}$ be a set of R-D sampling locations. The difference coarray is defined as $\mathbb{D} = \{\boldsymbol{n}_1 - \boldsymbol{n}_2 : \forall \boldsymbol{n}_1, \boldsymbol{n}_2 \in \mathbb{S}\}$.

Definition 8 (The matrix $\boldsymbol{J}$ for R-D): Let $\mathbb{S}$ be a set of R-D sampling locations and $\mathbb{D}$ be the R-D difference coarray. The matrix $\boldsymbol{J}$ is an $|\mathbb{S}|^2$-by-$|\mathbb{D}|$ matrix satisfying $\langle \boldsymbol{J} \rangle_{:,\boldsymbol{m}} = \mathrm{vec}(\boldsymbol{I}(\boldsymbol{m}))$. The matrix $\boldsymbol{I}(\boldsymbol{m})$ is given by

$$\langle \boldsymbol{I}(\boldsymbol{m}) \rangle_{\boldsymbol{n}_1, \boldsymbol{n}_2} = \begin{cases} 1, & \text{if } \boldsymbol{n}_1 - \boldsymbol{n}_2 = \boldsymbol{m}, \\ 0, & \text{otherwise}, \end{cases}$$

where $\boldsymbol{m} \in \mathbb{D}$ and $\boldsymbol{n}_1, \boldsymbol{n}_2 \in \mathbb{S}$.

Using Definitions 7 and 8, the correlation vector can be expressed as $\boldsymbol{c}(\bar{\boldsymbol{\mu}}) \triangleq \mathrm{vec}(\boldsymbol{v}_{\mathbb{S}}(\bar{\boldsymbol{\mu}}) \boldsymbol{v}_{\mathbb{S}}^H(\bar{\boldsymbol{\mu}})) = \boldsymbol{J} \boldsymbol{v}_{\mathbb{D}}(\bar{\boldsymbol{\mu}})$, where $\boldsymbol{v}_{\mathbb{D}}(\bar{\boldsymbol{\mu}})$ is the steering vector on the difference coarray. A numerical example is demonstrated in the supplementary document [46, Section II] for clarity. Then, the generalized correlation subspace matrix becomes

$$\boldsymbol{S}(\rho) \triangleq \int_{[-1/2,1/2]^R} \boldsymbol{c}(\bar{\boldsymbol{\mu}}) \boldsymbol{c}^H(\bar{\boldsymbol{\mu}}) \rho(\bar{\boldsymbol{\mu}}) \mathrm{d}\bar{\boldsymbol{\mu}} \tag{54}$$

$$= \boldsymbol{J} \underbrace{\left(\int_{[-1/2,1/2]^R} \boldsymbol{v}_{\mathbb{D}}(\bar{\boldsymbol{\mu}}) \boldsymbol{v}_{\mathbb{D}}^H(\bar{\boldsymbol{\mu}}) \rho(\bar{\boldsymbol{\mu}}) \mathrm{d}\bar{\boldsymbol{\mu}} \right)}_{\boldsymbol{S}_{\mathbb{D}}(\rho)} \boldsymbol{J}^H. \tag{55}$$

Eq. (54) and (55) are analogous to Definition 2 and Lemma 2, respectively. Finally, the generalized correlation subspace $\mathcal{GCS}(\rho) \triangleq \mathrm{col}(\boldsymbol{S}(\rho))$, as in Definition 3. According to (55), *the generalized correlation subspace for R-D depends on the difference coarray*, similar to Lemma 2.

Next we will consider some concrete examples in 2D. They are joint angle-Doppler estimation and 2D DOA estimation.

Example 6 (Joint angle-Doppler estimation): The joint angle-Doppler estimation [38], [39] corresponds to the data model (53) with $R = 2$. It aims to estimate the angle and Doppler frequency from the spatiotemporal measurements $\boldsymbol{x}_{\mathbb{S}}$. The harmonic parameters $\bar{\boldsymbol{\mu}} = (\bar{\mu}^{(1)}, \bar{\mu}^{(2)})$ are related to the DOA and the Doppler frequency as follows:

$$\bar{\mu}^{(1)} = (d/\lambda) \sin\theta, \qquad \bar{\mu}^{(2)} = (T/\lambda) v, \tag{56}$$

where λ is the wavelength, $d = \lambda/2$ is the minimum sensor separation, $\theta \in [-\pi/2, \pi/2]$ is the DOA, T is the sampling interval in the temporal domain, and v is the radial velocity of the source. $\bar{\mu}^{(1)}$ and $\bar{\mu}^{(2)}$ are called the normalized angle and normalized Doppler, respectively. If the sampling interval T is chosen properly, the parameters $(\bar{\mu}^{(1)}, \bar{\mu}^{(2)})$ belongs to $[-1/2, 1/2]^2$. This region is also depicted in Fig. 6(a) for clarity.

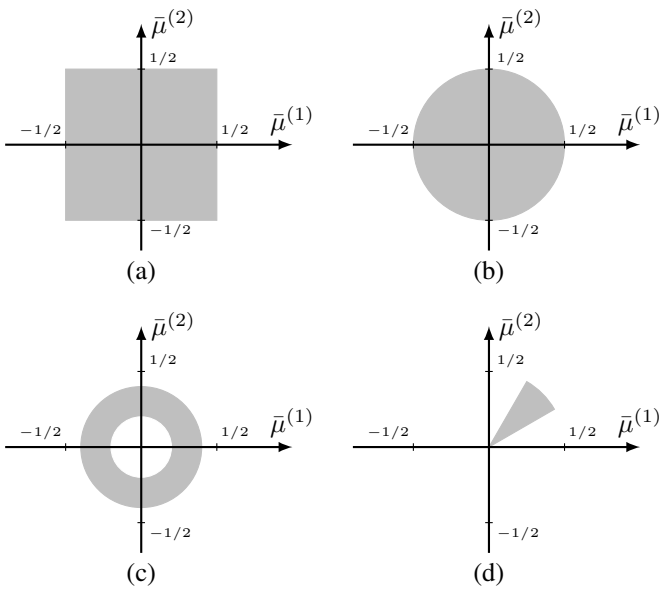


Fig. 6. The visible region (shaded) of (a) angle-Doppler, (b) 2D DOA, (c) 2D DOA with $\theta_{\min} \le \theta \le \theta_{\max}$, and (d) 2D DOA with $\phi_{\min} \le \phi \le \phi_{\max}$.

Next, suppose we choose the density function $\rho(\bar{\boldsymbol{\mu}})$ to be $\mathbf{1}_{[-1/2,1/2]^2}(\bar{\boldsymbol{\mu}})$. The generalized correlation subspace matrix becomes

$$\boldsymbol{S}(\mathbf{1}_{[-1/2,1/2]^2}) = \boldsymbol{J}\boldsymbol{J}^H = (\boldsymbol{J}\boldsymbol{W}^{-1/2}) \boldsymbol{W} (\boldsymbol{J}\boldsymbol{W}^{-1/2})^H,$$

where $\boldsymbol{W} = \mathrm{diag}(w(\boldsymbol{m}))_{\boldsymbol{m} \in \mathbb{D}}$. Here $w(\boldsymbol{m}) = |\{(\boldsymbol{n}_1, \boldsymbol{n}_2) \in \mathbb{S}^2 : \boldsymbol{n}_1 - \boldsymbol{n}_2 = \boldsymbol{m}\}|$ denotes 2D weight function. This result is analogous to that in Section IV-A. It can be inferred that the positive eigenvalues of $\boldsymbol{S}(\mathbf{1}_{[-1/2,1/2]^2})$ are $w(\boldsymbol{m})$ and the associated eigenvectors are $\mathrm{vec}(\boldsymbol{I}(\boldsymbol{m}))/\sqrt{w(\boldsymbol{m})}$. Note that the results in Example 6 can be trivially extended to the R-D case if the parameters $\bar{\boldsymbol{\mu}} \in [-1/2, 1/2]^R$.

Example 7 (2D DOA estimation): Another example of (53) with $R = 2$ is the 2D DOA estimation. The parameters of interest are the *azimuth* $\phi \in [0, 2\pi]$ and the *elevation* $\theta \in [0, \pi/2]$. The relation between the 2D DOA (θ, ϕ) and the harmonic parameters $\bar{\boldsymbol{\mu}} = (\bar{\mu}^{(1)}, \bar{\mu}^{(2)})$ is given by

$$\bar{\mu}^{(1)} = (d_x/\lambda) \sin\theta \cos\phi, \quad \bar{\mu}^{(2)} = (d_y/\lambda) \sin\theta \sin\phi, \tag{57}$$

where $d_x = d_y = \lambda/2$. According to (57), the visible region of $\bar{\boldsymbol{\mu}}$ becomes a disc with radius $1/2$, i.e.,

$$\mathbb{K} = \{\bar{\boldsymbol{\mu}} : \|\bar{\boldsymbol{\mu}}\|_2 \le 1/2\}. \tag{58}$$

The set $\mathbb{K}$ is also depicted in Fig. 6(b). Using (55), the generalized correlation subspace matrix with density function $\rho(\bar{\boldsymbol{\mu}}) = \mathbf{1}_{\mathbb{K}}(\bar{\boldsymbol{\mu}})$ becomes $\boldsymbol{S}(\mathbf{1}_{\mathbb{K}}) = \boldsymbol{J}\boldsymbol{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{K}})\boldsymbol{J}^H$. Here the entry of $\boldsymbol{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{K}})$ associated with coarray location $\boldsymbol{m}_1, \boldsymbol{m}_2 \in \mathbb{D}$ can be derived as [46, Section III]

$$\langle \boldsymbol{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{K}}) \rangle_{\boldsymbol{m}_1, \boldsymbol{m}_2} = \mathrm{jinc}(\|\boldsymbol{m}_1 - \boldsymbol{m}_2\|_2), \tag{59}$$

where the jinc function $\mathrm{jinc}(x)$ is $\pi/4$ for $x = 0$ and $J_1(\pi x)/(2x)$ otherwise [59]. Here $J_1(x)$ is the first-order Bessel function of the first kind [59]. Eq. (59) can be regarded as an extension of (32) since the jinc function is analogous

to the sinc function in the 2D polar coordinate [59]. The eigenvectors of $\boldsymbol{S}_{\mathbb{D}}(\mathbf{1}_{\mathbb{K}})$ are *2D discrete prolate spheroidal sequences (2D DPSS)*. Details about these sequences can be found in [60] and the references therein. Finally the generalized correlation subspace can be approximated by $\mathrm{col}(\boldsymbol{J}\boldsymbol{\Psi}_L)$, where $\boldsymbol{\Psi}_L$ contains the first L 2D DPSS and L is defined in (44).

If the prior knowledge about 2D DOA is available, then the visible region could be an annulus or a circular sector. For instance, if we know a priori that the elevation $\theta_{\min} \leq \theta \leq \theta_{\max}$, then the visible region is depicted in Fig. 6(c). On the other hand, if the the prior knowledge is $\phi_{\min} \leq \phi \leq \phi_{\max}$, then the visible region becomes a circular sector, as illustrated in Fig. 6(d).

VII. Numerical Examples

A. Generalized Correlation Subspaces

In this section, we will consider the following four 1D array configurations: the ULA with 10 sensors, the nested array with $N_1 = N_2 = 5$, the coprime array with $M = 3, N = 5$, and the super nested array with $N_1 = N_2 = 5, Q = 2$, where the notations are in accordance with [18], [21], [26]. The number of sensors is 10 for each array. The sensor locations and the nonnegative part of the difference coarrays for these arrays are depicted in Fig. 2. Since the difference coarray is symmetric [46, Lemma 1-2]., the size of the difference coarray is 19 for ULA, 59 for the nested array, 43 for the coprime array, and 59 for the super nested array.

The following example aims to demonstrate Theorem 1 and (27). The experiment is conducted as follows. We first compute the numerical approximation of $\mathbf{S}(\rho)$, as denoted by $\widetilde{\mathbf{S}}(\rho)$, as follows:

$$\widetilde{\mathbf{S}}(\rho) = \sum_{\ell=-(N_{\mathrm{pt}}-1)/2}^{(N_{\mathrm{pt}}-1)/2} \mathbf{c}(\ell\Delta)\mathbf{c}^H(\ell\Delta)\rho(\ell\Delta) \times \Delta, \qquad (60)$$

where the number of discrete samples is $N_{\mathrm{pt}} = 2^{14} + 1$ and the step size is $\Delta = 1/N_{\mathrm{pt}}$. Then the eigenvalues of $\widetilde{\mathbf{S}}(\rho_1)$ and $\widetilde{\mathbf{S}}(\rho_2)$ are plotted in Fig. 7(a), (c), (e), and (g), where $\rho_1(\bar{\theta})$ and $\rho_2(\bar{\theta})$ are given in Theorem 1. As a comparison, the weight functions for these arrays are plotted in Fig. 7(b), (d), (f), and (h).

The results in Fig. 7 show that, $\mathrm{col}(\widetilde{\mathbf{S}}(\rho_1))$ and $\mathrm{col}(\widetilde{\mathbf{S}}(\rho_2))$ have the same dimension $|\mathbb{D}|$, which is the size of the difference coarray. For example, the nested array (Fig. 7(c)) and the super nested array (Fig. 7(g)) share the same number of nonzero eigenvalues, since they own the same difference coarray, as shown in Fig. 2(b) and (d). Furthermore, the eigenvalues for $\widetilde{\mathbf{S}}(\rho_2)$ should be close to the weight functions, as indicated in (27). This property can be verified, for instance, in Fig. 7(g) and (h), where the eigenvalues $\lambda_2 = \lambda_3 = 4$ and the weight functions $w(2) = w(-2) = 4$.

B. DOA Estimation with Nested Arrays and Prior Source Intervals

Fig. 8 shows the estimation performance as a function of SNR and the number of snapshots. The equal-power and uncorrelated sources have normalized DOAs -0.045, 0, and 0.045. The number of sources D is 3. The array configuration is the nested array with $N_1 = N_2 = 5$ (10 sensors), as depicted in Fig. 2(b). Then the generalized correlation subspaces $\mathcal{GCS}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$ can be evaluated according to Section IV. The error tolerance in (44) is $\delta = 10^{-10}$ so the parameter L becomes 10 and 17 for $\alpha = 0.1$ and 0.2, respectively. In each run, the sensor measurements $\widetilde{\mathbf{x}}_{\mathbb{S}}(k)$ for $k = 1, \ldots, K$ are realized by (1), from which the sample covariance matrix $\widetilde{\mathbf{R}}_{\mathbb{S}}$ can be determined by (10). Then the covariance matrices are denoised according to 1) Problem (P1) and the `cvx` package with perfect knowledge about the noise power; 2) Problem (P2) and (47) without knowing the noise power. Finally the source directions are estimated by the SS MUSIC algorithm [18], [27] on the denoised covariance matrices. The estimation performance is measured in terms of root mean-squared errors (RMSE), defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{D}\sum_{i=1}^{D}(\widehat{\bar{\theta}}_i - \bar{\theta}_i)^2},$$

where $\widehat{\bar{\theta}}_i$ denotes the estimated normalized DOAs. Each sample point is averaged from 1000 Monte-Carlo runs.

It can be seen from Fig. 8 that if the source interval $[-\alpha/2, \alpha/2]$ is known, the RMSEs decrease except at very low SNRs. In this example, we choose $\alpha = 1$, 0.2, and 0.1 in Fig. 8. The following discuss the performances of these estimators:

1) In most cases, the RMSE for (P1) and (P2) decrease with α, for the same SNR and the same number of snapshots. The improvement is significant for low SNR and limited snapshots. This is because smaller source intervals $[-\alpha/2, \alpha/2]$ help to improve the estimation performances.
2) (P1) requires much more time than (P2), provided that the generalized correlation subspace is computed using the results in Section IV. For instance, the computational time for (P1) and $\alpha = 0.1$ in Fig. 8(a) is 9904.5 seconds while that for (P2) and $\alpha = 0.1$ in Fig. 8(a) is 53.7 seconds. The reason is that (P1) is solved numerically using the `cvx` package but (P2) can be implemented efficiently using (47).
3) As shown in Fig. 8(a), empirically, (P2) has better estimation performance than (P1) for the same α in most cases. For instance, if $\alpha = 0.1$ and $\mathrm{SNR} = -10$dB, then the RMSE for (P1) is 4.9×10^{-3} while that for (P2) becomes 1.9×10^{-3}. If $\alpha = 0.1$ and $\mathrm{SNR} = 10$dB, then the RMSEs for (P1) and (P2) are about 8.8×10^{-4} and 6.9×10^{-4}, respectively. These phenomena can also be observed in Fig. 8(b).

C. Computational Complexity

Table II investigates the interplay between the computational complexity and the estimation performance by varying the parameter α and the implementation details. These cases are denoted by Cases A to H, respectively. Here Step 1, which computes the (generalized) correlation subspace, can

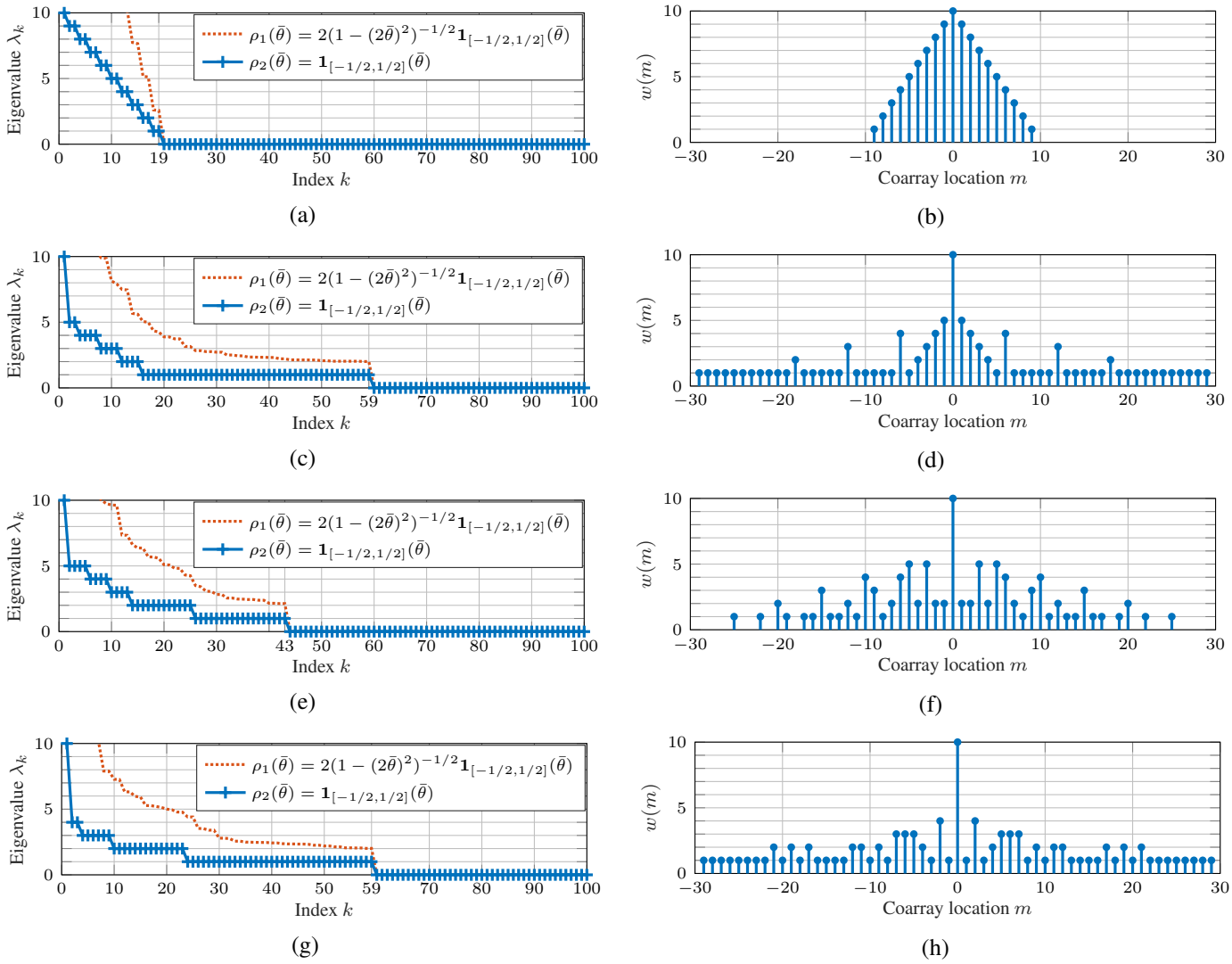


Fig. 7. The eigenvalues of the matrix $\widetilde{\mathbf{S}}(\rho)$ (left) and the weight functions (right) for (a), (b) the ULA with 10 sensors ($|\mathbb{D}| = 19$), (c), (d) the nested array with $N_1 = N_2 = 5$ (10 sensors, $|\mathbb{D}| = 59$), (e), (f) the coprime array with $M = 3, N = 5$ (10 sensors, $|\mathbb{D}| = 43$), and (g), (h) the super nested array with $N_1 = N_2 = 5, Q = 2$ (10 sensors, $|\mathbb{D}| = 59$). Here the matrices $\widetilde{\mathbf{S}}(\rho)$ are given by (60) and the eigenvalues of $\widetilde{\mathbf{S}}(\rho)$ are obtained numerically.

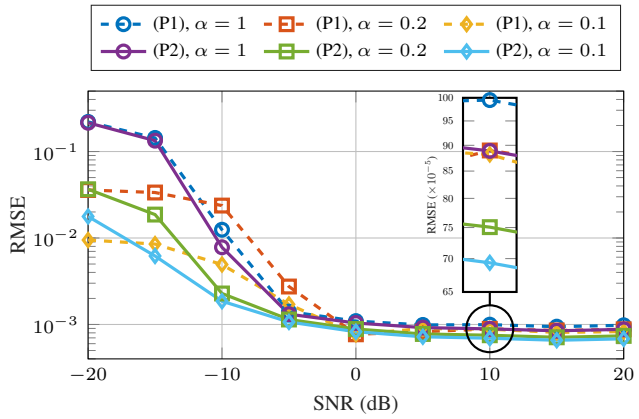
be realized in two ways. One is the numerical integration (60), followed by numerical eigen-decompositions [16]. The number of grids N_{pt} in (60) is $2^{14}+1$ while the density function $\rho(\bar{\theta})$ is $2(1-(2\bar{\theta})^2)^{-1/2}\mathbf{1}_{[-\alpha/2,\alpha/2]}(\bar{\theta})$. The other is to use the proposed analytical expressions, based on Definition 5, Theorem 1, and Table I. Then, in Step 2, either (P1) or (P2) is solved. Finally, Step 3 uses SS MUSIC to estimate the source directions. We assume 100 snapshots and 5dB SNR. The CPU time and the RMSEs in Table II are averaged from 1000 Monte-Carlo runs, on a Ubuntu 16.04 workstation with a Intel Core i7-2600 3.40GHz processor and 8GB RAM. The remaining parameters are given in Section VII-B.

Some observations can be drawn from Table II. First, the proposed analytical expressions lead to almost the same RMSE but much less CPU time than those with numerical approximations. For instance, the CPU time for Step 1 in Case D is about 0.04% of that for Step 1 in Case B. Second, in these examples, (P2) enjoys smaller RMSE and much less computational time than (P1). As an example, the CPU time of Step 2 in Case B costs only 0.007% of that of Step 2 in Case A, while the RMSE is 0.000991 for Case A and 0.000918 for Case B, respectively. Finally, if $\alpha = 0.1$, the proposed analytical approach remains computationally tractable. It can be observed that the CPU time for Step 1 in Case H is approximately 0.2% of that for Step 1 in Case F. These observations show that, for a given α, it is preferable to use the proposed analytical expressions and problem (P2), since they lead to the least total time and the smallest RMSE.

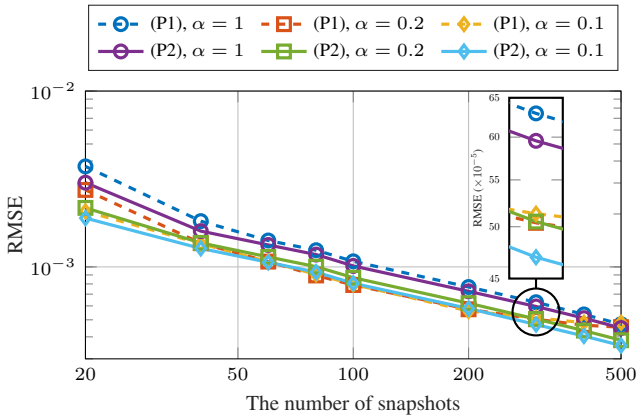
VIII. Concluding Remarks

In this paper, we presented generalized correlation subspaces with applications to DOA estimation for uncorrelated sources and known source intervals. Generalized correlation subspaces have closed-form characterizations in terms of array profiles such as sensor locations and difference coarrays. Furthermore, our theory provides insights to existing DOA estimators and multidimensional sparse arrays with prior knowledge about sources.

In the future, it is of considerable interest to exploit generalized correlation subspaces in other topics of array



(a)



(b)

Fig. 8. The dependence of root mean-squared errors (RMSE) on (a) SNR and (b) the number of snapshots for the optimization problems (P1) and (P2) with generalized correlation subspaces $\mathcal{GCS}(\mathbf{1}_{[-\alpha/2,\alpha/2]})$. There are $D = 3$ equal-power and uncorrelated sources at normalized DOAs -0.045, 0, and 0.045. The array configuration is the nested array with $N_1 = N_2 = 5$ (10 sensors), as depicted in Fig. 2(b). The parameters are (a) 100 snapshots and (b) 0dB SNR. Each data point is averaged from 1000 Monte-Carlo runs.

TABLE II
CPU TIME IN SECONDS AND RMSE FOR SEVERAL CASES

Case	α	Step 1 (time)	Step 2 (time)	Total time (including Step 3)	RMSE
A	1	Numerical (3.288s)	P1 (1.057s)	4.351s	0.000991
B	1	Numerical (3.248s)	P2 (0.00008s)	3.255s	0.000918
C	1	Analytical (0.0013s)	P1 (0.802s)	0.810s	0.000990
D	1	Analytical (0.0013s)	P2 (0.00009s)	0.008s	0.000918
E	0.1	Numerical (3.296s)	P1 (1.118s)	4.421s	0.000903
F	0.1	Numerical (3.271s)	P2 (0.00008s)	3.278s	0.000718
G	0.1	Analytical (0.0066s)	P1 (1.053s)	1.067s	0.000901
H	0.1	Analytical (0.0066s)	P2 (0.00010s)	0.013s	0.000718

† "Analytical" is the new analytical method for Step 1, as in Section IV.
‡ P2 is the proposed method for Step 2, as in Section V.
* Step 3 took between 0.0063 to 0.0068 seconds.

processing, like adaptive beamforming, source detection, and target tracking. Another future topic would be the performance analysis for the proposed approach using generalized correlation subspaces.

APPENDIX A
PROOF OF THE EQUIVALENCE OF (6) AND DEFINITION 1

To begin with, we need the following lemmas:

Lemma 4: Let $\mathbf{M} \in \mathbb{C}^{N\times N}$ be a Hermitian, positive semidefinite matrix. Assume $\mathbf{u} \in \mathbb{C}^N$. Then $\mathbf{u}^H\mathbf{M}\mathbf{u} = \mathbf{0}$ if and only if $\mathbf{u}$ belongs to the null space of $\mathbf{M}$.

Proof: If $\mathbf{u}$ belongs to the null space of $\mathbf{M}$, then $\mathbf{M}\mathbf{u} = \mathbf{0}$ and clearly $\mathbf{u}^H\mathbf{M}\mathbf{u} = 0$. Conversely, if $\mathbf{u}^H\mathbf{M}\mathbf{u} = 0$, then $(\mathbf{M}^{1/2}\mathbf{u})^H(\mathbf{M}^{1/2}\mathbf{u}) = 0$, where the matrix square root exists due to the positive semidefiniteness of $\mathbf{M}$. Hence, $\mathbf{M}^{1/2}\mathbf{u} = \mathbf{0}$, so $\mathbf{M}\mathbf{u} = \mathbf{0}$. ■

Lemma 5: Let $f(x)$ be a real-valued Lebesgue integrable function defined over a measurable set $\mathbb{A}$. Assume that $f(x) \geq 0$ for $x \in \mathbb{A}$ almost everywhere (a.e.). Then $\int_{\mathbb{A}} f(x)\mathrm{d}x = 0$ if and only if $f(x) = 0$ a.e.

Proof: See [61, Chapter V] for details. ■

Next it will be first shown that $\mathcal{CS}^{\perp} = \mathrm{null}(\mathbf{S}^H)$, where $\mathcal{CS}^{\perp}$ denotes the orthogonal complement of $\mathcal{CS}$ and $\mathrm{null}(\mathbf{A})$ represents the null space of $\mathbf{A}$.

First consider any $\mathbf{v} \in \mathrm{null}(\mathbf{S}^H)$. Since $\mathbf{S}$ is a Hermitian positive semidefinite matrix, Lemma 4 implies that $\mathbf{v} \in \mathrm{null}(\mathbf{S}^H)$ is equivalent to $\mathbf{v}^H\mathbf{S}^H\mathbf{v} = 0$. Substituting the definition of $\mathbf{S}$, (8), into $\mathbf{v}^H\mathbf{S}^H\mathbf{v} = 0$ yields

$$\int_{-\pi/2}^{\pi/2} \left|\mathbf{v}^H\mathbf{c}(\bar{\theta})\right|^2 \mathrm{d}\theta = 0. \tag{61}$$

Since the nonnegative function $\left|\mathbf{v}^H\mathbf{c}(\bar{\theta})\right|^2$ is continuous in θ, Lemma 5 indicates that (61) is equivalent to $\left|\mathbf{v}^H\mathbf{c}(\bar{\theta})\right|^2 = 0$ for all $\theta \in [-\pi/2, \pi/2]$. Namely, $\mathbf{v} \in \mathcal{CS}^{\perp}$. Note that all the above arguments are both necessary and sufficient so we have $\mathcal{CS}^{\perp} = \mathrm{null}(\mathbf{S}^H)$.

Finally, the correlation subspace can be obtained by $\mathcal{CS} = \mathcal{CS}^{\perp\perp} = \mathrm{null}(\mathbf{S}^H)^{\perp} = \mathrm{col}(\mathbf{S})$, which proves this observation.

APPENDIX B
PROOF OF LEMMA 1

Let $\rho_1(\bar{\theta})$ and $\rho_2(\bar{\theta})$ be two nonnegative Lebesgue integrable functions with the same support. The support of $\rho_1(\bar{\theta})$ and $\rho_2(\bar{\theta})$ satisfies $\mathrm{supp}(\rho_1) = \mathrm{supp}(\rho_2) = \mathbb{A} \subseteq [-1/2, 1/2]$. The corresponding generalized correlation subspace matrices are denoted by $\mathbf{S}(\rho_1)$ and $\mathbf{S}(\rho_2)$, respectively.

Consider a nonzero and finite-valued vector $\mathbf{u} \in \mathbb{C}^{|\mathbb{S}|^2}$. It will be shown that $\mathbf{u}^H\mathbf{S}(\rho_1)\mathbf{u} = 0$ if and only if $\mathbf{u}^H\mathbf{S}(\rho_2)\mathbf{u} = 0$. Based on Definition 2, the condition $\mathbf{u}^H\mathbf{S}(\rho_1)\mathbf{u} = 0$ is equivalent to the following:

$$\mathbf{u}^H\left(\int_{\mathbb{A}} \mathbf{c}(\bar{\theta})\mathbf{c}^H(\bar{\theta})\rho_1(\bar{\theta})\mathrm{d}\bar{\theta}\right)\mathbf{u} = 0$$

$$\text{if and only if } \int_{\mathbb{A}} |\mathbf{c}^H(\bar{\theta})\mathbf{u}|^2\rho_1(\bar{\theta})\mathrm{d}\bar{\theta} = 0. \tag{62}$$

Due to Lemma 5, Eq. (62) is equivalent to the statement that $|\mathbf{c}^H(\bar{\theta})\mathbf{u}|^2\rho_1(\bar{\theta}) = 0$ for $\bar{\theta} \in \mathbb{A}$ a.e. Since $\bar{\theta}$ belongs

to the support $\mathbb{A}$, multiplying both sides with $\rho_2(\bar{\theta})/\rho_1(\bar{\theta})$ yields $|\mathbf{c}^H(\bar{\theta})\mathbf{u}|^2\rho_2(\bar{\theta}) = 0$ for $\bar{\theta} \in \mathbb{A}$ a.e. Invoking Lemma 5 again gives $\mathbf{u}^H\mathbf{S}(\rho_2)\mathbf{u} = 0$. Combining the above arguments with Lemma 4 leads to $\mathrm{null}(\mathbf{S}(\rho_1)) = \mathrm{null}(\mathbf{S}(\rho_2))$, where $\mathrm{null}(\mathbf{S})$ denotes the null space of $\mathbf{S}$. Taking the orthogonal complement on the null spaces gives $\mathrm{col}(\mathbf{S}(\rho_1)) = \mathrm{col}(\mathbf{S}(\rho_2))$, which proves this theorem.

ACKNOWLEDGMENT

The authors would like to thank Prof. George K. Atia for the discussion on correlation subspaces at the IEEE SAM workshop in 2016. It inspired us to come up with the idea of generalized correlation subspaces. The authors would also like to thank Mr. Mostafa Rahmani and the reviewers for their useful suggestions.

REFERENCES

[1] R. N. Bracewell, *Radio astronomy techniques*. Berlin: Springer, 1962, vol. 54.

[2] S. Haykin, *Array Signal Processing*. Prentice-Hall, 1984.

[3] R. T. Hoctor and S. A. Kassam, "The unifying role of the coarray in aperture synthesis for coherent and incoherent imaging," *Proc. IEEE*, vol. 78, no. 4, pp. 735–752, Apr 1990.

[4] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul 1996.

[5] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. McGraw Hill, 2001.

[6] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Wiley Interscience, 2002.

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

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

[9] 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, Jul 1989.

[10] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1132–1143, Jul 1990.

[11] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, Feb 2011.

[12] P. Stoica and P. Babu, "SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation," *Signal Processing*, 2012.

[13] P. Stoica, D. Zachariah, and J. Li, "Weighted SPICE: A unifying approach for hyperparameter-free sparse estimation," *Digital Signal Processing*, vol. 33, pp. 1–12, October 2014. [Online]. Available: http://dx.doi.org/10.1016/j.dsp.2014.06.010

[14] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185 – 210, 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200498903165

[15] M. Jansson, B. Göransson, and B. Ottersten, "A subspace method for direction of arrival estimation of uncorrelated emitter signals," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 945–956, Apr 1999.

[16] M. Rahmani and G. K. Atia, "A subspace method for array covariance matrix estimation," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, July 2016, pp. 1–5.

[17] A. T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 172–175, 1968.

[18] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, Aug 2010.

[19] 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.

[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, March 2015.

[21] C.-L. Liu and P. P. Vaidyanathan, "Super nested arrays: Linear sparse arrays with reduced mutual coupling–Part I: Fundamentals," *IEEE Trans. Signal Process.*, vol. 64, no. 15, pp. 3997–4012, Aug 2016.

[22] S. Pillai, Y. Bar-Ness, and F. Haber, "A new approach to array geometry for improved spatial spectrum estimation," *Proc. IEEE*, vol. 73, no. 10, pp. 1522–1524, Oct 1985.

[23] S. Pillai and F. Haber, "Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1517–1523, Nov 1987.

[24] Y. Abramovich, D. Gray, A. Gorokhov, and N. Spencer, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays–Part I: Fully augmentable arrays," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2458–2471, Sep 1998.

[25] Y. Abramovich, N. Spencer, and A. 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, Jun 1999.

[26] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," in *Proc. IEEE Dig. Signal Proc. Signal Proc. Educ. Workshop*, Jan 2011, pp. 289–294.

[27] 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, Sept 2015.

[28] Y. D. Zhang, M. G. Amin, and B. Himed, "Sparsity-based DOA estimation using co-prime arrays," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, May 2013, pp. 3967–3971.

[29] P. Pal and P. P. Vaidyanathan, "Pushing the limits of sparse support recovery using correlation information," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 711–726, Feb 2015.

[30] E. BouDaher, F. Ahmad, and M. G. Amin, "Sparsity-based extrapolation for direction-of-arrival estimation using co-prime arrays," in *Proc. SPIE 9857, Compressive Sensing V: From Diverse Modalities to Big Data Analytics*, 2016, p. 98570M. [Online]. Available: http://dx.doi.org/10.1117/12.2225454

[31] Q. Shen, W. Liu, W. Cui, and S. Wu, "Underdetermined DOA estimation under the compressive sensing framework: A review," *IEEE Access*, vol. 4, pp. 8865–8878, 2016.

[32] Z. Tan, Y. Eldar, and A. Nehorai, "Direction of arrival estimation using co-prime arrays: A super resolution viewpoint," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5565–5576, Nov 2014.

[33] 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, June 2014.

[34] A. Koochakzadeh, H. Qiao, and P. Pal, "Compressive spectrum sensing with spectral priors for cognitive radar," in *2016 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, Sept 2016, pp. 100–104.

[35] K. V. Mishra, M. Cho, A. Kruger, and W. Xu, "Spectral super-resolution with prior knowledge," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5342–5357, Oct 2015.

[36] H.-H. Chao and L. Vandenberghe, "Extensions of semidefinite programming methods for atomic decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, March 2016, pp. 4757–4761.

[37] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty – V: The discrete case," *Bell Labs Technical Journal*, vol. 57, no. 5, pp. 1371–1430, 1978.

[38] D. D. Arianandа and G. Leus, "Compressive joint angular-frequency power spectrum estimation," in *Proc. the 21st European Signal Processing Conference (EUSIPCO 2013)*, Marrakech, Morocco, Sep. 2013.

[39] C.-L. Liu and P. P. Vaidyanathan, "Coprime arrays and samplers for space-time adaptive processing," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, Apr. 2015, pp. 2364–2368.

[40] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 405–418, Feb 1998.

[41] S. Qin, Y. D. Zhang, M. G. Amin, and F. Gini, "Frequency diverse coprime arrays with coprime frequency offsets for multitarget localization," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 2, pp. 321–335, March 2017.

[42] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3198–3213, July 2008.

[43] D. Nion and N. Sidiropoulos, "Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5693–5705, Nov 2010.

[44] C.-L. Liu and P. P. Vaidyanathan, "Cramér-Rao bounds for coprime and other sparse arrays, which find more sources than sensors," *Digital Signal Processing*, vol. 61, pp. 43–61, 2017, special Issue on Coprime Sampling and Arrays. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1051200416300264
[45] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, 1985.
[46] http://systems.caltech.edu/dsp/students/clliu/GCS/Supp.pdf.
[47] P. Wirfält, G. Bouleux, M. Jansson, and P. Stoica, "Optimal prior knowledge-based direction of arrival estimation," *IET Signal Processing*, vol. 6, no. 8, pp. 731–742, October 2012.
[48] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge University Press, 1993.
[49] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, Sept 1982.
[50] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 605–612, Jul 1988.
[51] P. P. Vaidyanathan and T. Nguyen, "Eigenfilters: A new approach to least-squares FIR filter design and applications including Nyquist filters," *IEEE Trans. Circuits Syst.*, vol. 34, no. 1, pp. 11–23, Jan 1987.
[52] P. P. Vaidyanathan, *Multirate Systems And Filter Banks*. Pearson Prentice Hall, 1993.
[53] C.-Y. Chen and P. P. Vaidyanathan, "MIMO radar space-time adaptive processing using prolate spheroidal wave functions," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 623–635, Feb 2008.
[54] H. Taylor and S. W. Golomb, "Rulers, Part I," Univ. Southern Calif., Los Angeles, Tech. Rep. 85-05-01, 1985.
[55] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, June 2015.
[56] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec 2013.
[57] A. Koochakzadeh and P. Pal, "Cramér-Rao bounds for underdetermined source localization," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 919–923, July 2016.
[58] 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.
[59] R. N. Bracewell, *The Fourier Transform and its Applications*, 3rd ed. McGraw-Hill, 2000.
[60] F. J. Simons and D. V. Wang, "Spatiospectral concentration in the Cartesian plane," *GEM - International Journal on Geomathematics*, vol. 2, no. 1, pp. 1–36, 2011. [Online]. Available: http://dx.doi.org/10.1007/s13137-011-0016-z
[61] P. R. Halmos, *Measure Theory*, 1st ed., ser. Graduate Texts in Mathematics. Springer-Verlag New York, 1950, vol. 18.

Chun-Lin Liu (S'12) was born in Yunlin, Taiwan, on April 28, 1988. He received the B.S. and M.S. degrees in electrical engineering and communication engineering from National Taiwan University (NTU), Taipei, Taiwan, in 2010 and 2012, respectively. He is currently pursuing the Ph.D. degree in electrical engineering at the California Institute of Technology (Caltech), Pasadena, CA.

His research interests are in sparse array signal processing, sparse array design, tensor signal processing, and statistical signal processing. He received the Best Student Paper Awards at the 41st IEEE International Conference on Acoustics, Speech and Signal Processing, 2016, Shanghai, China, and the 9th IEEE Sensor Array and Multichannel Signal Processing Workshop, 2016, Rio de Janeiro, Brazil. He was also one of the recipients of the Student Paper Award at the 50th Asilomar Conference on Signals, Systems, and Computers, 2016, Pacific Grove, CA.

P. P. Vaidyanathan (S'80–M'83–SM'88–F'91) was born in Calcutta, India on Oct. 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radiophysics and electronics, all from the University of Calcutta, India, in 1974, 1977 and 1979, respectively, and the Ph.D degree in electrical and computer engineering from the University of California at Santa Barbara in 1982. He was a post doctoral fellow at the University of California, Santa Barbara from Sept. 1982 to March 1983. In March 1983 he joined the electrical engineering department of the Calfornia Institute of Technology as an Assistant Professor, and since 1993 has been Professor of electrical engineering there. His main research interests are in digital signal processing, multirate systems, wavelet transforms, signal processing for digital communications, genomic signal processing, radar signal processing, and sparse array signal processing.

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program committee for the 1983 IEEE International symposium on Circuits and Systems, and as the Technical Program Chairman for the 1992 IEEE International symposium on Circuits and Systems. He was an Associate editor for the IEEE Transactions on Circuits and Systems for the period 1985-1987, and is currently an associate editor for the journal IEEE Signal Processing letters, and a consulting editor for the journal Applied and computational harmonic analysis. He has been a guest editor in 1998 for special issues of the IEEE Trans. on Signal Processing and the IEEE Trans. on Circuits and Systems II, on the topics of filter banks, wavelets and subband coders.

Dr. Vaidyanathan has authored nearly 500 papers in journals and conferences, and is the author/coauthor of the four books Multirate systems and filter banks, Prentice Hall, 1993, Linear Prediction Theory, Morgan and Claypool, 2008, and (with Phoong and Lin) Signal Processing and Optimization for Transceiver Systems, Cambridge University Press, 2010, and Filter Bank Transceivers for OFDM and DMT Systems, Cambridge University Press, 2010. He has written several chapters for various signal processing handbooks. He was a recipient of the Award for excellence in teaching at the California Institute of Technology for the years 1983-1984, 1992-93 and 1993-94. He also received the NSF's Presidential Young Investigator award in 1986. In 1989 he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990 he was recepient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommuncations Engineers, India, for his joint paper in the IETE journal. In 2009 he was chosen to receive the IETE students' journal award for his tutorial paper in the IETE Journal of Education. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE SP Transactions, for which the first author (Truong Nguyen) received the Young outstanding author award in 1993. Dr. Vaidyanathan was elected Fellow of the IEEE in 1991. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education. He has given several plenary talks including at the IEEE ISCAS-04, Sampta-01, Eusipco-98, SPCOM-95, and Asilomar-88 conferences on signal processing. He has been chosen a distinguished lecturer for the IEEE Signal Processing Society for the year 1996-97. In 1999 he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal. He is a recepient of the IEEE Signal Processing Society's Technical Achievement Award for the year 2002, and the IEEE Signal Processing Society's Education Award for the year 2012. He is a recipient of the IEEE Gustav Kirchhoff Award (an IEEE Technical Field Award) in 2016, for "Fundamental contributions to digital signal processing." In 2016 he received the Northrup Grumman Prize for excellence in Teaching at Caltech. He is also the recipient of the 2016 Society Award of the IEEE Signal Processing Society.