Underdetermined High-Resolution DOA Estimation: A 2\( \rho \)-th Order Source-Signal/Noise Subspace Constrained Optimization

Jin Ho Choi, Student Member, IEEE, and Chang D. Yoo, Senior Member, IEEE

Abstract—For estimating the direction of arrival (DOAs) of non-stationary source signals such as speech and audio, a constrained optimization problem (COP) that exploits the spatial diversity provided by an array of sensors is formulated in terms of a noise-eliminated local \( 2\rho \)-th order cumulant matrix. The COP solution provides a weight vector to the look direction such that it is constrained to the \( 2\rho \)-th order source-signal subspace when the look direction is in alignment with the true DOA; otherwise, it is constrained to the \( 2\rho \)-th order noise subspace. This weight vector is incorporated into the spatial spectrum to determine the degree of orthogonality between itself and either the \( 2\rho \)-th order source-signal subspace when the number of sources is unknown, or the \( 2\rho \)-th order noise subspace when the number of sources is known. For a uniform linear array (ULA) of \( M \) sensors, the spatial spectrum for known number of sources can theoretically be shown to identify up to \( 2\rho(M - 1) \) sources. Realizing the difficulty in identifying stationarity in the received sensor signals, the estimate of the noise-eliminated local \( 2\rho \)-th order cumulant matrix is marginalized over various possible stationary segmentations, for a more robust DOA estimation. In this paper, we focus on the use of local second and fourth order cumulants (\( \rho = 1, 2 \)), and the proposed algorithms when \( \rho = 1 \) outperformed the KR subspace-based algorithms and also the 4-MUSIC for globally non-stationary, non-Gaussian synthetic data and also for speech/audio in various adverse environments. We verified that the identifiability for \( \rho = 2 \) is improved by two-folds compared to that for \( \rho = 1 \) with an ULA.

Index Terms—DOA estimation, non-stationary source signal, constrained optimization, spatial spectrum.

I. INTRODUCTION

An array of sensors with arbitrary geometry which includes the uniform circular array (UCA), uniform linear array (ULA) and uniform planar array (UPA) can be used to estimate the direction of arrivals (DOAs) of source signals under various adverse conditions. The DOA estimation has applications in many fields including wireless data communication, radar/sonar and biomedical signal detection, speech/audio recognition [2], [4], [5], [7], [8], [14], [17], [18], [20], [21], [23], [27]–[30]. In radar and sonar systems, the obstacles (objects) such as cars, birds, airplanes, submarines and missiles can be tracked by estimating their DOAs, and in speech recognition, the DOAs of speakers can be used to enhance recognition performance.

Thus far, many high-resolution DOA estimation algorithms that are capable of resolving sources situated close to one another under low signal-to-noise ratio (SNR) have been proposed. These algorithms can be grouped into two categories: (1) maximum-likelihood (ML) based algorithms [7], [14], [21] and (2) subspace-based algorithms such as the multiple signal classification (MUSIC) algorithm, the estimation signal parameter via a rotational invariant technique (ESPRIT), subspace fitting [5], [8], [10], [14] and various other algorithms recently reported [26]–[31], [33]. As mentioned in previous literatures in [7], [14], [21], the ML-based algorithms yield optimal solutions, however, are computationally expensive for they typically require nonlinear and multidimensional optimization procedures. In comparison, many subspace-based algorithms, where the directions are estimated by searching an one-dimensional spatial spectrum, are computationally inexpensive and perform relatively well [5], [14], [26]–[29]. Furthermore, the wide-band extension of the narrow-band subspace-based algorithm achieves good performance with modest increase in computation by simple subspace manipulation in frequency domain [4], [19], [26], [28], [29]. For these reasons, to track time-variant DOAs of moving sources in an on-line manner with low computational cost, the subspace-based algorithms are often preferred over the ML-based algorithms.

Each of the subspace-based algorithms mentioned above has been proposed based on a number of assumptions and criteria: source statistical characteristics regarding its stationarity and distribution, the number of sources and sensors, the availability of the number of sources and prior knowledge on the number of statistical independent sources [5], [14], [26]–[30], [33]. The conventional MUSIC algorithm [5] which is based on the second order (SO) cumulants assumes the source to be stationary, and the number of sources to be known \( a \) priori and be strictly less than the number of sensors. Many variants of the MUSIC algorithm [14] have been proposed to outperform the MUSIC proposed in [5]. Recently, Zeng et al. [29] proposed an algorithm based on multiple SO cumulant matrices that generalizes to the wide-band MUSIC. For non-stationary sources, the algorithm can accurately estimate the DOAs when the number of sources is either less than or equal to the number of sensors.
The algorithm does not require the number of sources to be known a priori. In [27], Zeng et al. proposed a high-resolution DOA estimation algorithm based on the joint diagonalization of multiple fourth order (FO) cumulant matrices when the unknown number of sources is less than or equal to the number of sensors. Here, the source is assumed to be stationary and non-Gaussian.

When the number of sources is larger than the number of sensors, often referred to as the underdetermined case, the above SO and FO cumulant-based algorithms have difficulty in accurately estimating the DOAs. To resolve this difficulty, the DOA estimation algorithms based on higher-order (HO) cumulants (higher than second order) organized in terms of a virtual array response matrix-constructed by taking the Kronecker products of the original array response vector [11], [16], [22], [25], [33]-have been proposed, and the HO cumulant-based algorithms such as the 4-MUSIC have shown to possess better source identifiability and higher resolution compared to the algorithms based on the SO cumulant such as the MUSIC. In the HO cumulant-based algorithms, the source is assumed to be stationary and non-Gaussian. While these algorithms have certain advantages [11], [13], [16], [22], [25], [33], they require larger sample size for accurately estimating the HO cumulants and involve longer computation time for estimation, compared to the SO cumulant-based algorithms. All of the HO cumulant algorithms mentioned above require the number of sources to be known a priori.

For non-stationary sources in underdetermined case, Ma et al. [28] proposed two Khatiri-Rao (KR) subspace-based algorithms: (1) the KR-MUSIC that requires the number of sources to be known a priori and (2) the KR-Capon that does not. Using the local SO cumulants of the sources, the KR-MUSIC can identify up to 2(M − 1) sources where M is the number of sensors in the ULA. Each sensor is assumed to share identical characteristics such as radiating pattern and polarization [22], [25]. Based on the same ULA setting, the identifiability of the KR-MUSIC is identical to that of the 4-MUSIC in [16], [22], [25]. The KR subspace-based algorithms have the ability to estimate the DOAs in underdetermined situation in the presence of unknown spatially correlated stationary noises, without having to resort to the HO cumulants. The KR subspace-based algorithms have been shown to perform better than those algorithms such as the 4-MUSIC for the non-stationary sources such as speech.

This paper proposes underdetermined high-resolution DOA estimation algorithms for both known and unknown number of globally non-stationary sources that are locally stationary such as speech and audio. The proposed algorithms are derived from a novel constrained optimization problem (COP) formulated in terms of a noise-eliminated local 2pth-order cumulant matrix by exploiting the spatial diversity provided by an array of sensors. When p = 1, the proposed algorithms outperformed the KR subspace-based algorithms and also the conventional fourth order cumulant-based algorithm such as the 4-MUSIC in various adverse environments. We verified that the identifiability for p = 2 is improved by two-folds compared to that for p = 1 with an ULA.

The remainder of this paper is organized as follows. Section II describes signal model, 2pth-order cumulant matrix and virtual array response matrix. Section III proposes two types of underdetermined high-resolution DOA estimation algorithms based on a 2pth-order source-signal/noise subspace constrained optimization, using a noise-eliminated local 2pth-order cumulant matrix. Section IV evaluates the proposed algorithms for narrow-band non-stationary source signals generated from generalized Gaussian distribution, and also for wide-band non-stationary source signals such as speech and audio in determined and underdetermined cases, and Section V concludes the paper.

II. SIGNAL MODEL, 2pTH-ORDER CUMULANT MATRIX AND VIRTUAL ARRAY RESPONSE MATRIX

The statistical relationship between the received sensor signals and source signals is defined in terms of the 2pth-order statistic. For this, a 2pth-order cumulant matrix of the received sensor signals is defined in terms of the virtual array response matrix of the DOAs, a 2pth-order cumulant matrix of source signals and that of noise signals. The statistical relationship is derived from a signal model based on five assumptions.

A. Signal model and five assumptions

Consider an arbitrary array of M sensors such that the nth sensor is located at (x_m,y_m) in a two-dimensional space, the far-field scenario such that the size of the sensor array aperture is much smaller than the distance from sources to the sensor array [22], [28], [29] is assumed. When I wide-band non-stationary sources (s_i(t) i = 0,...,I − 1) located at distinct directions impinge on the array where I > M, the received sensor signal x_m(t) at the mth sensor can be modeled as,

\[ x_m(t) = \sum_{i=0}^{I-1} \alpha_i s_i(t - \tau_{mi}) + z_m(t), \quad m = 0, ..., M - 1. \] (1)

where \( \alpha_i \) and \( \tau_{mi} \) are respectively the attenuation factor of the ith source and the propagation time delay from the first \( (m - 0) \) sensor to the mth sensor of the ith source. Here, \( z_m(t) \) is a stationary noise at the mth sensor. Taking the short-time discrete Fourier transform (STDFT) of \( x_m(t) \), assuming sampling rate \( f_s \) [29], the kth frequency bin component of the mth sensor at time \( n \) is given as

\[ X_{m,k}[n] = \sum_{\tau=0}^{I-1} x_m[n + \tau] \exp(-j \frac{2\pi k}{N} \tau), \]

\[ = \sum_{i=0}^{I-1} \alpha_i S_{i,k}[n] \exp(-j \frac{2\pi k f_s \tau_{mi}}{N}) + Z_{m,k}[n], \]

\[ m = 0, ..., M - 1, \quad k = 0, ..., N - 1 \] (2)

where \( u[n] \) (0 \( \leq n \leq L_u - 1 \) and \( N \) are respectively the window sequence and the number of DFT points. Here, \( x_m[n] = x_m \left( \frac{n}{f_s} \right) \); furthermore, \( S_{i,k}[n] \) and \( Z_{m,k}[n] \) are the kth STDFT coefficients of \( s_i \left( \frac{n}{f_s} \right) \) and \( z_m \left( \frac{n}{f_s} \right) \), respectively. Here, \( \tau_{mi} \) is given as

\[ \tau_{mi} = \tau_m(\theta_i), \] (3)

with

\[ \tau_m(\theta) = \frac{(x_m - x_o) \cos(\theta) + (y_m - y_o) \sin(\theta)}{c}. \]
where $\theta_i$ is the $i$th source DOA and $c$ is the source velocity (see [25]). Assuming negligible differences in the distances from the sources to any one of the sensors, the attenuation factors of all the sources are assumed to be equal such that $a_i = 1, \forall i$. According to (2) and (3), the array response vector of source from direction $\theta$ is given as

$$a_k(\theta) = [a_{0,k}(\theta), a_{1,k}(\theta), \ldots, a_{M-1,k}(\theta)]^T$$

where $a_{m,k}(\theta) = \exp \left( -j \frac{2\pi}{N} f_s r_m(\theta) \right)$.

In deriving the proposed algorithms, the following assumptions are considered:

A1) For any $\{i,k\} \in \{i = 0, \ldots, I - 1, k = 1, \ldots, N\}$, $S_{i,k}[n]$ is zero-mean, globally non-stationary and can be either Gaussian or non-Gaussian.

A2) For any $k \in \{1, \ldots, N\}$, $\{S_{i,k}[n] : i = 0, \ldots, I - 1\}$ are statistically independent.

A3) For any $(m,k) \in \{m = 0, \ldots, M - 1, k = 1, \ldots, N\}$, $Z_{m,k}[n]$ is zero-mean, stationary and can be either Gaussian or non-Gaussian.

A4) For any $k \in \{1, \ldots, N\}$, $\{Z_{m,k}[n] : m = 0, \ldots, M - 1\}$ can be either spatially correlated or uncorrelated.

A5) For any $k \in \{1, \ldots, N\}$ and $(i,m) \in \{i = 0, \ldots, I - 1, m = 0, \ldots, M - 1\}$, $S_{i,m}[n]$ and $Z_{m,k}[n]$ are statistically independent.

B. Statistical Relationship in Terms of the $2\text{nd}$-Order Cumulant

Before deriving the statistical relationship between the received sensor signals and source signals in terms of the $2\text{nd}$-order cumulant, the $2\text{nd}$-order cumulant is defined in terms of the $2\text{nd}$-order moment which is defined below.

For a given random vector $\mathbf{v} = [V_0, V_1, \ldots, V_L]$, the associated $2\text{nd}$-order moment is defined as

$$\mathcal{M}(\mathbf{v}, g^L_p, S) = E \left( \prod_{i \in S} V_{g_i}^L \right)$$

where $g^L_p = [g_0, g_1, \ldots, g_{p-1}]$, $0 \leq g < L$.

Here, $S$ and $E(\cdot)$ denote a subset of the set of all the elements of $\mathbf{v}$ and the expectation. According to [22], [25], $\epsilon_{i-2q+1} = -1$, $\epsilon_{j-2q+1} = 1$ where $V^1 = V_a, V^2 = (V_a)^*$ and $*$ denotes the conjugate operator. The $2\text{nd}$-order cumulant can be mathematically represented based on the Leonov-Shiryaev formula [6], [22], [25] as

$$\mathcal{K}(\mathbf{v}, g^L_p) = \text{Cum}_\varepsilon[V_{g_0}, V_{g_1}, \ldots, V_{g_{p-1}}, V_{g_p}]$$

$$= \sum_{p=1}^{2p} (-1)^{p-1} (p-1)! \sum_{S} \mathcal{M}(\mathbf{v}, g^L_p, S')$$

where $(S'_1, S'_2, \ldots, S'_{p})$ describes the $i$th partition of $\{V_{g_0}, V_{g_1}, \ldots, V_{g_p}, V_{g_2}, \ldots, V_{g_p}\}$ with $p$ subsets. For instance, when $p = 2$ where $\{V_{g_0}, V_{g_1}, V_{g_2}, V_{g_3}\}$, the sub-sets are as follows for $p = 2$: $S_{2,1} = \{x_1, x_2\}$ and $S_{2,2} = \{x_3, x_4\}$. $S_{2,1} = \{x_1, x_2\}$ and $S_{2,2} = \{x_3, x_4\}$. $S_{2,1} = \{x_1, x_2\}$ and $S_{2,2} = \{x_3, x_4\}$. $S_{2,1} = \{x_1, x_2\}$ and $S_{2,2} = \{x_3, x_4\}$.

Let the received sensor signal vector at the $k$th frequency bin and time $n$ be

$$\mathbf{x}_k[n] = [X_{0,k}[n], X_{1,k}[n], \ldots, X_{M-1,k}[n]]^T$$

For $n_t < n < n_{t+1} - 1$, $\mathbf{x}_k[n]$ is locally-stationary where $n_t$ is the starting time marker for the $t$th stationary segment. The associated average moment can be mathematically represented as

$$\mathcal{M}(\mathbf{x}_k[n_t], g^M_p, S) = \sum_{n_t = 1}^{n_{t+1} - 1} \mathcal{M}(\mathbf{x}_k[n], g^M_p, S).$$

For a reliable estimate of $\mathcal{M}(\mathbf{x}_k[n_t], g^M_p, S)$, the stationary duration $(n_{t+1} - n_t)$ should not be overly short. Then, the associated average moment is defined as

$$\mathcal{K}(\mathbf{x}_k[n_t], g^M_p) = \sum_{p=1}^{2p} (-1)^{p-1} (p-1)! \sum_{S} \mathcal{M}(\mathbf{x}_k[n_t], g^M_p, S').$$

Here, the average cumulant $\mathcal{K}(\mathbf{x}_k[n_t], g^M_p)$ based on the five assumptions discussed in Section II-A can be mathematically represented as

$$\mathcal{K}(\mathbf{x}_k[n_t], g^M_p) = \sum_{i=0}^{I-1} a_{g_0,k}(\theta_i) a_{g_1,k}(\theta_i) \cdots a_{g_{2p-2,k}(\theta_i) a_{g_{2p-1,k}(\theta_i) a_{g_2,k}}}.$$

$$\text{for } s_k[n] = [S_{0,k}[n], S_{1,k}[n], \ldots, S_{I-1,k}[n]],$$

$$\mathbf{z}_k[n] = [Z_{0,k}[n], Z_{1,k}[n], \ldots, Z_{M-1,k}[n]].$$

and $1_{2p}$ is defined as the $2p$ length index vector whose elements are all ones.

C. Statistical relationship in terms of a $2\text{nd}$-order cumulant matrix and virtual array response matrix

A $2\text{nd}$-order cumulant matrix can be constructed from the $2\text{nd}$-order cumulant $\mathcal{K}(\mathbf{x}_k[n_t], g^M_p)$ defined in (10). Note that the source signals which include speech and audio are assumed to be globally non-stationary as assumed in [28] such that

$$\mathcal{K}(\mathbf{s}_k[n_t], i \cdot 1_{2p}) \neq \mathcal{K}(\mathbf{s}_k[n_{t+1}], i \cdot 1_{2p})$$

and

$$\mathcal{K}(\mathbf{s}_k[n_t], i \cdot 1_{2p}) \neq \mathcal{K}(\mathbf{s}_k[n_{t+1}], j \cdot 1_{2p}), \text{ for } i \neq j$$

Then, (13) is used to construct the virtual array response matrix

$$\mathcal{K}(\mathbf{x}_k[n_t], g^M_p) \neq \mathcal{K}(\mathbf{x}_k[n_{t+1}], g^M_p).$$

Let the total ordered set of possible $g^M_p$ be $\Omega(g^M_p)$, then

$$|\Omega(g^M_p)| = (2p)^p.$$
The cumulants $\overline{K}(x_k[n_t], g_p^M)$ for different values of $g_p^M$ and $n_t$ can be organized into a single cumulant matrix $K_{x_k}^{(p)} \in \mathbb{C}^{M \times t \times h}$ such that

$$[K_{x_k}^{(p)}]_{d,n_t} = \overline{K}(x_k[n_t], g_p^M)$$

with a set of starting time markers for all locally-stationary segments $s = \{n_t\}_{t=1,\ldots,t}$ where $[M]_{i,j}$ denotes the $(i,j)$th element of $M$.

According to (10),

$$K_{x_k}^{(p)} \in \mathbb{C}^{M \times t \times h} = A_k^{(p)} D_{x_k}^{(p)} + K_{x_k}^{(p)}$$

where

$$[D_{x_k}^{(p)}]_{i,n_t} = \overline{K}(s_k[n_t], \cdot \cdot \cdot 1_{2p}),$$

$$[K_{x_k}^{(p)}]_{d,n_t} = \overline{K}(z_k[n_t], g_p^M),$$

$$A_k^{(p)} \in \mathbb{C}^{M \times t \times I} = \left[\alpha_k^{(p)}(\theta_0), \ldots, \alpha_k^{(p)}(\theta_{I-1})\right],$$

and

$$a_k^{(p)}(\theta) = (a_k^{(p)}(\theta) \otimes a_k^{(p)}(\theta))^\otimes p.$$  

Here,

$$(u)^\otimes p = u \otimes \cdots \otimes u$$

where $\otimes$ denotes the $p$-fold Kronecker product, and it is assumed that the number of the locally-stationary segments $b$ is larger than the number of sources $I$ such that $b > I$.

D. Properties of the $2p$th-order cumulant matrix $K_{x_k}^{(p)}$

In case of stationary sources and stationary noises,

$$\text{rank}(K_{x_k}^{(p)}) = 1$$

since

$$[K_{x_k}^{(p)}]_{d,n_t} = \left[K_{x_k}^{(p)}\right]_{d,n_t}$$

with

$$[K_{x_k}^{(p)}]_{d,n_t} = \left[K_{x_k}^{(p)}\right]_{d,n_t+1}$$

where $\text{rank}(M)$ denotes the rank of $M$.

The proposed algorithms assume non-stationary sources satisfying (11), (12) and (13) and stationary noises satisfying (20), and therefore, (18) does not apply. In (17), assuming that $\text{rank}(A_k^{(p)}) = I$, the following statements can be made.

(a) From (11) and (12), all $I$ row vectors $\{[D_{x_k}^{(p)}]_{i,n_t}\}_{i=0,\ldots,t-1}$ are linearly independent.

(b) From (20), $[K_{x_k}^{(p)}]_{i,n_t} = c$, $F_{x_k,j}^{(p)}$ for any $(i,j)$ where $c$ is a constant.

(c) From (13), any row vector $[K_{x_k}^{(p)}]_{d,n_t}$ is linearly independent of all $I$ row vectors $\{[D_{x_k}^{(p)}]_{i,n_t}\}_{i=0,\ldots,t-1}$.

Detailed discussion on above statements regarding $K_{x_k}^{(p)}$ when $p = 1$ can be found in [28].

Note that in practice, the larger the number of data samples, larger the number of locally-stationary segments $b$. This will generally lead to closer approximation of the above three statements.

From (c) (see [3]),

$$\text{rank}(K_{x_k}^{(p)}) = \text{rank}(A_k^{(p)} D_{x_k}^{(p)} + K_{x_k}^{(p)}),$$

$$= \text{rank}(A_k^{(p)} D_{x_k}^{(p)}) + \text{rank}(K_{x_k}^{(p)}).$$

From (a), $\text{rank}(D_{x_k}^{(p)}) = I$ with $b > I$ and from (b), $\text{rank}(A_k^{(p)}) = I$ and thus,

$$\text{rank}(K_{x_k}^{(p)}) = \text{rank}(A_k^{(p)} D_{x_k}^{(p)}) + 1,$$

$$= \min\{\text{rank}(A_k^{(p)}), \text{rank}(D_{x_k}^{(p)})\} + 1,$$

$$= I + 1.$$  

Note that for correlated sources, (22) is invalid.

III. UNDERDETERMINED HIGH-RESOLUTION DOA ESTIMATION: $2p$TH-ORDER SOURCE-SIGNAL/NOISE SUBSPACE CONSTRAINED OPTIMIZATION

Two DOA estimation algorithms for known and unknown number of sources are derived from a COP that exploits the spatial diversity provided by the array of sensors. The COP is expressed in terms of a noise-eliminated local $2p$th order cumulant matrix obtained from $K_{x_k}^{(p)}$ defined in (17). The algorithms when $p > 1$ give higher identifiability than previously proposed DOA estimation algorithms, for a given number of sensors and array geometry.

A. A noise-eliminated local $2p$th-order cumulant matrix

For stationary noises, $[K_{x_k}^{(p)}]_{d,n_t} = [K_{x_k}^{(p)}]_{d,n_t+1}$, thus,

$$K_{x_k}^{(p)} = c_k^{(p)} 1_b^T$$

where $c_k^{(p)} \in \mathbb{R}^{M \times 1}$ is a column vector such that $[c_k^{(p)}]_d = [K_{x_k}^{(p)}]_{d,n_t}$ and $1_b$ is a $b$-length column vector whose elements are all one. Therefore, $K_{x_k}^{(p)}$ can be eliminated by projecting $K_{x_k}^{(p)}$ onto the orthogonal complement projection matrix

$$P = I_b - (1/b) 1_b 1_b^T$$

where $I_b$ is $b \times b$ identity matrix such that

$$K_{x_k}^{(p)} P = A_k^{(p)} D_{x_k}^{(p)} P + K_{x_k}^{(p)} P = A_k^{(p)} D_{x_k}^{(p)} P$$

where

$$K_{x_k}^{(p)} P - c_k^{(p)} 1_b^T (I_b - (1/b) 1_b 1_b^T),$$

$$- c_k^{(p)} 1_b^T (1/b) c_k^{(p)} (1_b^T 1_b - b) 1_b^T,$$

$$0_{M \times r \times b}$$

and $0_{M \times r \times b} \in \mathbb{R}^{M \times r \times b}$ is the zero matrix. Note that when $p > 1$, for Gaussian noises, $K_{x_k}^{(p)} = 0_{M \times r \times b}$.

Here,

$$\text{rank}(K_{x_k}^{(p)} P) = \text{rank}(A_k^{(p)} D_{x_k}^{(p)} P) - I$$

Denote $[M]_{i,j}$ as the $i$th row vector of $M$.  

and

\[ \mathcal{R} \left( K^{(p)}_{\mathbf{n}_k} \mathbf{P} \right) = \mathcal{R} \left( A^{(p)}_{k} D^{(p)}_{\mathbf{n}_k} \mathbf{P} \right) - \mathcal{R} \left( A^{(p)}_{k} \right) \]  

(27)

where \( \mathcal{R}(\cdot) \) denotes the range space \([2] \).

A noise-eliminated local 2th-order cumulant matrix is defined as

\[ D^{(p)}_{\mathbf{n}_k} = \left( K^{(p)}_{\mathbf{n}_k} \mathbf{P} \right) \left( \left( K^{(p)}_{\mathbf{n}_k} \mathbf{P} \right)^H \right) = A^{(p)}_{k} \left( D^{(p)}_{\mathbf{n}_k} \mathbf{P} \mathbf{P}^H \left( D^{(p)}_{\mathbf{n}_k} \right)^H A^{(p)}_{k} \right)^H, \]  

(28)

such that

\[ \text{rank} \left( D^{(p)}_{\mathbf{n}_k} \right) = \text{rank} \left( A^{(p)}_{k} \right) \]  

(29)

and

\[ \mathcal{R} \left( D^{(p)}_{\mathbf{n}_k} \right) = \mathcal{R} \left( A^{(p)}_{k} \right) \]  

(30)

where \( D^{(p)}_{\mathbf{n}_k} \) is symmetric and satisfies \( \text{rank}(D^{(p)}_{\mathbf{n}_k}) = 1 \). Note that for a fixed number of sensors and array geometry, when \( p > 1 \), the maximum rank of \( K^{(p)}_{\mathbf{n}_k} \) defined in (17) is larger than that of the cumulant matrix constructed by the SO virtual array in [28] and also that of the cumulant matrix constructed by the HO virtual array of the 4-MUSIC in [22]. This allows the proposed algorithms to estimate a larger number of the DOAs (higher identifiability).

### B. A COP for high-resolution DOA estimation

A COP expressed in terms of \( C^{(p)}_{\mathbf{n}_k} \) defined in (28) is formulated to maximize the square of the array gain such that the inner product of the weight vector \( \mathbf{w}^{(p)}_{\mathbf{k}} \) and virtual array response vector \( \mathbf{a}^{(p)}_{\mathbf{k}}(\theta) \) to the look direction \( \theta \), under the constraint that equates the sum of squares of the inner products of \( \mathbf{w}^{(p)}_{\mathbf{k}} \) and the eigenvectors of \( \mathcal{R}(C^{(p)}_{\mathbf{n}_k}) \), and the inner products of \( \mathbf{w}^{(p)}_{\mathbf{k}} \) and the eigenvectors of \( N(C^{(p)}_{\mathbf{n}_k}) \) to a value that is dependent on the look direction \( \theta \). Here, the strengths of the eigenvectors are determined by a user set parameter discussed later in Section III-B. The notations \( \mathcal{R}(\cdot) \) and \( N(\cdot) \) respectively denote the range and null space.

Furthermore, the constraint is conditioned on whether the number of sources \( I \) is known or not, according to the parameter that determines the strengths of the eigenvectors, \( \mathbf{w}^{(p)}_{\mathbf{k}} \) can be constrained to (1) \( \mathcal{R}(C^{(p)}_{\mathbf{n}_k}) \), (2) both \( \mathcal{R}(C^{(p)}_{\mathbf{n}_k}) \) and \( N(C^{(p)}_{\mathbf{n}_k}) \), and (3) \( N(C^{(p)}_{\mathbf{n}_k}) \).

Thus, the COP is formulated in terms of maximization criterion \( J_{\text{COP}}(\theta) \) in estimating \( \mathbf{w}^{(p)}_{\mathbf{k}} \) as follows:

\[ J_{\text{COP}}(\theta) = \max_{\mathbf{w}^{(p)}_{\mathbf{k}}} \| \mathbf{a}^{(p)}_{\mathbf{k}}(\theta) \|^2 \]  

\[ \text{subject to} \]  

\[ \| \mathbf{w}^{(p)}_{\mathbf{k}} \|_2^2 \mathbf{B}^{(p)}_{\mathbf{k}} \mathbf{w}^{(p)}_{\mathbf{k}} = (c^{(p)}_{\mathbf{k}}) \]  

(31)

\[ \text{subject to} \]  

\[ \mathbf{B}^{(p)}_{\mathbf{k}} \mathbf{w}^{(p)}_{\mathbf{k}} = (c^{(p)}_{\mathbf{k}}) \]  

(32)

where

\[ \mathbf{B}^{(p)}_{\mathbf{k}} = \begin{cases} \mathbf{U}^{(p)}_{\mathbf{s}_k} (\Sigma^{(p)}_{\mathbf{s}_k} + \alpha^{(p)}_{k} \mathbf{I})_{\mathbf{n}_k}^{-1} \mathbf{U}^{(p)}_{\mathbf{n}_k}^H, \quad \text{for known } I, \\ \mathbf{C}^{(p)}_{\mathbf{s}_k} + \alpha^{(p)}_{k} \mathbf{I}_M, \quad \text{for unknown } I. \end{cases} \]  

(33a)

(33b)

(34)

where the columns of \( \mathbf{U}^{(p)}_{\mathbf{s}_k} \) and \( \mathbf{U}^{(p)}_{\mathbf{n}_k} \) are respectively the eigenvectors of \( \mathcal{R}(C^{(p)}_{\mathbf{n}_k}) \) and \( \mathcal{N}(C^{(p)}_{\mathbf{n}_k}) \), and \( \mathbf{I}_M \) is an identity matrix.

The eigenvalue decomposition (EVD) of \( C^{(p)}_{\mathbf{n}_k} \) relates \( \mathbf{U}^{(p)}_{\mathbf{s}_k} \) and \( \mathbf{U}^{(p)}_{\mathbf{n}_k} \) to \( C^{(p)}_{\mathbf{n}_k} \) such that

\[ \mathbf{C}^{(p)}_{\mathbf{n}_k} = \mathbf{U}^{(p)}_{\mathbf{s}_k} \Sigma^{(p)}_{\mathbf{s}_k} \mathbf{U}^{(p)}_{\mathbf{n}_k}^H. \]  

(34)

In the above COP, \( \mathbf{w}^{(p)}_{\mathbf{k}}(\theta) \) is the weight vector of the \( k \)th frequency bin to the look direction \( \theta \) and the constraint (32) is conditioned on whether the number of sources \( I \) is known or not. Note that (33a) and (33b) lead to the same COP for \( \mathbf{C}^{(p)}_{\mathbf{n}_k} = \mathbf{u}^{(p)}_{\mathbf{s}_k} \Sigma^{(p)}_{\mathbf{s}_k} \mathbf{u}^{(p)}_{\mathbf{n}_k}^H, \) but for known \( I \) an EVD can be performed.

In (35) of (33a) or (33b), \( \alpha^{(p)}_{k} > 0 \) determines the strengths (the square roots of eigenvalues) of the eigenvectors of both \( \mathcal{R}(C^{(p)}_{\mathbf{n}_k}) \) and \( \mathcal{N}(C^{(p)}_{\mathbf{n}_k}) \), and in (32), \( c^{(p)}_{\mathbf{k}} \) is an arbitrary nonzero coefficient that determines the sum of squares of the inner products of \( \mathbf{w}^{(p)}_{\mathbf{k}} \) and the eigenvectors with their strengths, given as a function of \( \theta \). Therefore, satisfying (32) given \( \alpha^{(p)}_{k} \) and \( c^{(p)}_{\mathbf{k}} \) of (31) can be obtained as a linear weighted sum of the eigenvectors to the look direction \( \theta \). Here, the optimal weight vector is given as,

\[ \mathbf{w}^{(p)}_{\mathbf{k}}(\theta) = \mathbf{B}^{(p)}_{\mathbf{k}}^{-1} \mathbf{a}^{(p)}_{\mathbf{k}}(\theta) \]  

(35)

where

\[ \mathbf{B}^{(p)}_{\mathbf{k}} = \left[ \begin{array}{c} \mathbf{U}^{(p)}_{\mathbf{s}_k} \mathbf{U}^{(p)}_{\mathbf{n}_k} \left[ \Sigma^{(p)}_{\mathbf{s}_k} + \alpha^{(p)}_{k} \mathbf{I} \right]^{-1} \mathbf{U}^{(p)}_{\mathbf{n}_k}^H \\ \mathbf{U}^{(p)}_{\mathbf{n}_k} \end{array} \right] \]  

(36)

The proof is provided in Appendix A.

Regardless of whether the number of sources \( I \) is known or not,

\[ \mathbf{B}^{(p)}_{\mathbf{k}} = \left( \begin{array}{c} \mathbf{U}^{(p)}_{\mathbf{s}_k} \mathbf{U}^{(p)}_{\mathbf{n}_k} \left[ \Sigma^{(p)}_{\mathbf{s}_k} + \alpha^{(p)}_{k} \mathbf{I} \right]^{-1} \mathbf{U}^{(p)}_{\mathbf{n}_k}^H \\ \mathbf{U}^{(p)}_{\mathbf{n}_k} \end{array} \right) \]  

(37)

\[ \mathbf{B}^{(p)}_{\mathbf{k}} = \left( \begin{array}{c} \mathbf{U}^{(p)}_{\mathbf{s}_k} \mathbf{U}^{(p)}_{\mathbf{n}_k} \left[ \Sigma^{(p)}_{\mathbf{s}_k} + \alpha^{(p)}_{k} \mathbf{I} \right]^{-1} \mathbf{U}^{(p)}_{\mathbf{n}_k}^H \\ \mathbf{U}^{(p)}_{\mathbf{n}_k} \end{array} \right) \]  

(37)
Here, $\cdot^{-1}$ denotes the matrix inverse. Inserting (37) into (35), where

$$
(w_{k}^{(p)})_{\theta,\text{opt}} = \beta_k^{(p)} \left( U_{s,k}^{(p)} W_{s,k}^{(p)} e_k^{(p)}(\theta) + \frac{U_{n,k}^{(p)} e_k^{(p)}(\theta)}{\alpha_k^{(p)}} \right)
$$

(38)

and

$$
\beta_k^{(p)} = \sqrt{\left( e_k^{(p)}(\theta)^H W_{s,k}^{(p)} e_k^{(p)}(\theta) + \frac{\| e_k^{(p)}(\theta) \|^2}{\alpha_k^{(p)}} \right)^{-1}}
$$

Here, $W_{s,k}^{(p)}$ is a diagonal matrix with $i$th diagonal element $W_{s,k}^{(p)}_{ii} = 1/(\sigma_{s,k}^2 + \alpha_k^{(p)})$.

Depending on whether the look direction $\theta$ is along the DOAs of the sources or not, $(w_{k}^{(p)})_{\theta,\text{opt}}$ will lie one of the following spaces: (1) $\mathcal{R}(C_{s,k}^{(p)})$, (2) both $\mathcal{R}(C_{s,k}^{(p)})$ and $\mathcal{N}(C_{s,k}^{(p)})$, and (3) $\mathcal{N}(C_{s,k}^{(p)})$.

1) When $\theta = \theta_i$ for $i = 0, \ldots, I - 1$: $e_k^{(p)}(\theta_i)$ is null, thus

$$
(w_{k}^{(p)})_{\theta_i,\text{opt}} = \beta_k^{(p)} \left( U_{s,k}^{(p)} W_{s,k}^{(p)} e_k^{(p)}(\theta_i) \right)
$$

(39)

where

$$
\beta_k^{(p)} = \sqrt{\left( e_k^{(p)}(\theta_i)^H W_{s,k}^{(p)} e_k^{(p)}(\theta_i) \right)^{-1}}
$$

Therefore, $(w_{k}^{(p)})_{\theta_i,\text{opt}} \in \mathcal{R}(C_{s,k}^{(p)})$.

2) When $\theta \neq \theta_i$ for $i = 0, \ldots, I - 1$: $\alpha_k^{(p)} \gg \max(\Sigma_{s,k}^{(p)})$

where $\max(\cdot)$ selects the largest value,

$$
(w_{k}^{(p)})_{\theta,\text{opt}} = \beta_k^{(p)} \left( U_{s,k}^{(p)} W_{s,k}^{(p)} e_k^{(p)}(\theta) \right)
$$

(40)

where

$$
\beta_k^{(p)} = \sqrt{\left( e_k^{(p)}(\theta)^H W_{s,k}^{(p)} e_k^{(p)}(\theta) \right)^{-1}}
$$

from

$$
U_{s,k}^{(p)} e_k^{(p)}(\theta) + U_{n,k}^{(p)} e_k^{(p)}(\theta) = \alpha_k^{(p)} e_k^{(p)}(\theta).
$$

(41)

Therefore, $(w_{k}^{(p)})_{\theta,\text{opt}} \in \mathcal{R}(C_{s,k}^{(p)}) \cup \mathcal{N}(C_{s,k}^{(p)})$.

(ii) $0 < \alpha_k^{(p)} \ll \min(\Sigma_{s,k}^{(p)})$ where $\min(\cdot)$ selects the smallest value, not zero,

$$
(w_{k}^{(p)})_{\theta,\text{opt}} = \beta_k^{(p)} \left( U_{n,k}^{(p)} e_k^{(p)}(\theta) \right)
$$

(42)

Therefore, $(w_{k}^{(p)})_{\theta,\text{opt}} \in \mathcal{N}(C_{s,k}^{(p)})$.

C. The spatial spectra of the proposed algorithms for known and unknown $I$

With an appropriate $\alpha_k^{(p)}$ (discussed in Section III-D) such that $(w_{k}^{(p)})_{\theta,\text{opt}} \in \mathcal{R}(C_{s,k}^{(p)})$ when $\theta = \theta_i$ and $(w_{k}^{(p)})_{\theta,\text{opt}} \in \mathcal{N}(C_{s,k}^{(p)})$ when $\theta \neq \theta_i$, the optimal weight vector $(w_{k}^{(p)})_{\theta,\text{opt}}$ is incorporated into the spatial spectrum to determine the degree of orthogonality of $(w_{k}^{(p)})_{\theta,\text{opt}}$ to either $\mathcal{R}(C_{s,k}^{(p)})$ or $\mathcal{N}(C_{s,k}^{(p)})$.

For known $I$, a spatial spectrum to measure the orthogonality of $(w_{k}^{(p)})_{\theta,\text{opt}}$ to $\mathcal{R}(C_{s,k}^{(p)})$ is considered, while for unknown $I$, a spatial spectrum to measure the orthogonality of $(w_{k}^{(p)})_{\theta,\text{opt}}$ to $\mathcal{R}(C_{s,k}^{(p)})$ is considered.

Here, $(w_{k}^{(p)})_{\theta,\text{opt}}$ is given as either (39), (40) or (42).

1) The spatial spectrum for known $I$: The constrained 2$p$th-order KR-MUSIC (c-$2p$-KR-MUSIC) spatial spectrum is proposed as

$$
P_{c-2p-\text{KR-MUSIC}}(\theta) = \left( \sum_k \left( (w_{k}^{(p)})_{\theta,\text{opt}} U_{n,k}^{(p)} \right)^{\text{H}} U_{n,k}^{(p)} \right)^{-1}
$$

(43)

where

$$
\rho = 1 \quad \text{and} \quad \alpha_k^{(p)} \gg \max(\Sigma_{s,k}^{(p)})
$$

2) The spatial spectrum for unknown $I$: The constrained 2$p$th-order KR-Capon (c-$2p$-KR-Capon) spatial spectrum is proposed as

$$
P_{c-2p-\text{KR-Capon}}(\theta) = \sum_k \left( (w_{k}^{(p)})_{\theta,\text{opt}} G_{s,k}^{(p)} \left( w_{k}^{(p)} \right)_{\theta,\text{opt}} \right)^{\text{H}} G_{s,k}^{(p)} \left( w_{k}^{(p)} \right)_{\theta,\text{opt}}^{-1}
$$

(44)

where

$$
\gamma_k^{(p)} = \sqrt{\left( (w_{k}^{(p)})_{\theta,\text{opt}} \right)^{\text{H}} G_{s,k}^{(p)} \left( w_{k}^{(p)} \right)_{\theta,\text{opt}}}
$$

is just a constant. When $(c_k^{(p)})_{\theta} = \alpha_k^{(p)}$, this spectrum is identical to that of the KR-MUSIC.

2) The spatial spectrum for unknown $I$: The constrained 2$p$th-order KR-Capon (c-$2p$-KR-Capon) spatial spectrum is proposed as

$$
P_{c-2p-\text{KR-Capon}}(\theta) = \sum_k \left( (w_{k}^{(p-1)})_{\theta,\text{opt}} G_{s,k}^{(p-1)} \left( w_{k}^{(p-1)} \right)_{\theta,\text{opt}} \right)^{\text{H}} G_{s,k}^{(p-1)} \left( w_{k}^{(p-1)} \right)_{\theta,\text{opt}}^{-1}
$$

(45)

where

$$
\gamma_k^{(p-1)} = \sqrt{\left( (w_{k}^{(p-1)})_{\theta,\text{opt}} \right)^{\text{H}} G_{s,k}^{(p-1)} \left( w_{k}^{(p-1)} \right)_{\theta,\text{opt}}}
$$

is just a constant. When $(c_k^{(p-1)})_{\theta} = \alpha_k^{(p-1)}$, this spectrum is identical to that of the KR-Capon.
obtained by inserting

\( (c_k^{(p)})_\theta = 1/\sqrt{\{(a_k^{(p)}(\theta))^H(B_k^{(p)})^{-1}a_k^{(p)}(\theta)\}} \)

into \( \hat{p}_k^{(p)} \) defined in (35). This spectrum is identical to that of the KR-Capon.

By searching the look direction \( \theta \), the DOAs are estimated as the local peaks of the spatial spectra of the c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon.

D. The choice of \( \alpha_k^{(p)} \) and \( (c_k^{(p)})_\theta \) for the c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon

When \( \alpha_k^{(p)} = \epsilon \) where \( \epsilon \) is small, both c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon lead to high-resolution DOA estimation.

In practice, the estimate of \( \hat{c}_{x_k}^{(p)} \) will not satisfy (29) and (30). In other words, \( \hat{c}_{x_k}^{(p)} \) and \( \hat{c}_{x_k}^{(p)} \) is such that an EVD of \( \hat{c}_{x_k}^{(p)} \) is given in the following form (compare with (34))

\[
\hat{c}_{x_k}^{(p)} = [\hat{U}_{x,k}^{(p)}, \hat{U}_{n,k}^{(p)}] \begin{bmatrix}
\tilde{\Sigma}_{x_k}^{(p)} & 0 \\
0 & \tilde{\Sigma}_{x_k}^{(p)}
\end{bmatrix} \begin{bmatrix}
\hat{U}_{x,k}^{(p)} \\
\hat{U}_{n,k}^{(p)}
\end{bmatrix}^H.
\]

(47)

Here, the elements of \( \tilde{\Sigma}_{x_k}^{(p)} \) are assumed to be smaller than the elements of \( \tilde{\Sigma}_{x_k}^{(p)} \) and

\[
\hat{U}_{x,k}^{(p)} = U_{x,k}^{(p)} + \Delta U_{x_k}^{(p)},
\]

\[
\hat{U}_{n,k}^{(p)} = U_{n,k}^{(p)} + \Delta U_{n_k}^{(p)}
\]

(48)

(49)

where \( \Delta U_{x_k}^{(p)} \in \mathcal{N}(U_{x_k}^{(p)}) \) and \( \Delta U_{n_k}^{(p)} \in \mathcal{N}(U_{n_k}^{(p)}) \) are the estimation errors of \( U_{x_k}^{(p)} \) and \( U_{n_k}^{(p)} \), respectively. Generally, the larger the data sample size, smaller \( \Delta U_{x_k}^{(p)} \) and \( \Delta U_{n_k}^{(p)} \). Overall, the estimation error in \( \hat{c}_{x_k}^{(p)} \) affects \( (w_k^{(p)})_\theta^{opt} \) such that when \( \theta = \theta_i \), \( (w_k^{(p)})_\theta^{opt} \) has a small estimation error component in \( \mathcal{N}(U_{x_k}^{(p)}) \), and when \( \theta \neq \theta_i \), even though \( \alpha_k^{(p)} \rightarrow 0 \), \( (w_k^{(p)})_\theta^{opt} \) has a small estimation error component in \( \mathcal{R}(U_{x_k}^{(p)}) \). Thus, the spatial spectra of c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon are affected by the estimation error.

Taking the estimation error into account, \( \alpha_k^{(p)} \) is determined such that \( (w_k^{(p)})_\theta^{opt} \) in the spatial spectra (43) and (45) gives high-resolution DOA estimation: high spatial spectra values when \( \theta = \theta_i \) and low spatial spectra values when \( \theta \neq \theta_i \), as much as possible. In Appendix B, a derivation is given to determine \( \alpha_k^{(p)} \) for the following spatial spectra of the c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon with the estimation error.

1) The c-2\( \rho \)-KR-MUSIC:

\[
\mathbf{P}_{c-2\rho-KR-MUSIC}(\theta) = \left( \sum_k \left\| \left( (w_k^{(p)})_\theta^{opt} \right)^H \hat{U}_{x,k}^{(p)} \right\|^2 \right)^{-1} \left( (w_k^{(p)})_\theta^{opt} \right)^H \hat{U}_{x,k}^{(p)}.
\]

(50)

Considering the effect of \( \Delta U_{x_k}^{(p)} \) in \( \hat{c}_{x_k}^{(p)} \) from Appendix B-A, \( (w_k^{(p)})_\theta^{opt} \) should be close to \( \mathcal{R}(A_k^{(p)}) \) when \( \theta = \theta_i \), and move closer to \( \mathcal{N}(A_k^{(p)}) \) when \( \theta \neq \theta_i \). Thus, \( \alpha_k^{(p)} \) is set to

\[
\alpha_k^{(p)} = (1 + \epsilon) \max \left( \tilde{\Sigma}_{x_k}^{(p)} \right)
\]

(51)

where \( \epsilon \) is a small positive value.

2) The c-2\( \rho \)-KR-Capon:

\[
\mathbf{P}_{c-2\rho-KR-Capon}(\theta) = \sum_k \left( (w_k^{(p)})_\theta^{opt} \right)^H \hat{c}_{x_k}^{(p)} \left( (w_k^{(p)})_\theta^{opt} \right). \]

(52)

Considering the effects of \( \Delta U_{x_k}^{(p)} \) and \( \Delta U_{n_k}^{(p)} \) in \( \hat{c}_{x_k}^{(p)} \) from Appendix B-B, \( (w_k^{(p)})_\theta^{opt} \) should be close to \( \mathcal{R}(A_k^{(p)}) \) when \( \theta = \theta_i \), and move closer to \( \mathcal{N}(A_k^{(p)}) \) when \( \theta \neq \theta_i \). Thus, \( \alpha_k^{(p)} \) is set to

\[
\alpha_k^{(p)} = (1 - \epsilon) \min \left( \tilde{\Sigma}_{x_k}^{(p)} \right)
\]

(53)

where \( 0 \leq \epsilon < 1 \).

To satisfy \( |(w_k^{(p)})_\theta^{opt}|_2 = 1 \), \( \forall \theta \), in (32),

\[
(\alpha_k^{(p)})_\theta = \frac{\left( (w_k^{(p)})_\theta^{opt} \right)^H (B_k^{(p)})^{-1} (\alpha_k^{(p)})_\theta}{\left( (w_k^{(p)})_\theta^{opt} \right)^H (B_k^{(p)})^{-2} (\alpha_k^{(p)})_\theta}.
\]

(54)

In Section IV, it is verified that \( (w_k^{(p)})_\theta^{opt} \) with an appropriate \( \alpha_k^{(p)} \) in the c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon leads to better performance than other high-resolution DOA estimation algorithms.

Note that in the proposed algorithms, \( (w_k^{(p)})_\theta^{opt} \) for the c-2\( \rho \)-KR-MUSIC and c-2\( \rho \)-KR-Capon can be obtained as the unit norm eigenvector, corresponding to non-zero eigenvalue of \( (B_k^{(p)})^{-1} (a_k^{(p)}(\theta) (a_k^{(p)}(\theta))^H \) as shown by the following equation

\[
(B_k^{(p)})^{-1} a_k^{(p)}(\theta) (a_k^{(p)}(\theta))^H (w_k^{(p)})_\theta^{opt} = \lambda_k^{(p)} (w_k^{(p)})_\theta^{opt}.
\]

(55)

This is derived in (61) of Appendix A.

E. Estimation of \( \mathbf{C}_{x_k}^{(p)} \) to enhance the performance of DOA estimation

For a set of starting time markers of all locally-stationary segments \( n_t \), \( n_t \) defined in (28) can be estimated based on the empirical (time-average) estimate of \( \mathbf{C}_{x_k}^{(p)} \) defined in (17); however, for non-stationary source signals such as speech and audio, \( \mathbf{C}_{x_k}^{(p)} \) is unknown and difficult to determine. A fixed set value for \( n_t \), \( \forall \theta \) in \( \mathbf{N} \), as we often use in practice, can not lead to an accurate DOA estimation. For this reason, the estimate of \( \mathbf{C}_{x_k}^{(p)} \) is obtained by marginalizing over all possible set of segment-lengths for local stationarity \( L = \{l_t\}_t=1,...,L \) where \( l_t = n_{t+1} - n_t \), as

\[
\mathbf{E}_L (\hat{c}_{x_k}^{(p)} L ).
\]

(56)

Here, \( l_t \sim p(l_t) \) and \( \hat{c}_{x_k}^{(p)} \) is the empirical estimate of \( \mathbf{C}_{x_k}^{(p)} \) given \( L \). Instead of using \( \hat{c}_{x_k}^{(p)} \), (56) is considered in the COP to enhance the accuracy and robustness of the DOA estimation.
Algorithm The proposed algorithms

Step 1) Obtain \(J\) possible sets of segment-lengths for local stationarity \(\mathbf{L}\) and corresponding estimates \(\mathbf{C}_{\chi, s}^{(p)}\) \(\mathbf{L}\) as the time-average. Average \(\mathbf{C}_{\chi, s}^{(p)}\) \(\mathbf{L}\) over \(J\) possible sets for \(\mathbf{E}_{1}\) \((\mathbf{C}_{\chi, s}^{(p)} \mathbf{L})\) in (56): for given \(\mathbf{L}\), \(\mathbf{C}_{\chi, s}^{(p)}\) is obtained using the empirical estimate of \(\mathbf{K}_{\chi, s}^{(p)}\) that is obtained by the empirical estimate of \(\mathbf{K}_{\chi, s}[n]\) in (9) expressed by the empirical estimate of \(\mathbf{M}(\mathbf{x}_{k}[n]),\mathbf{M},\mathbf{S}\) in (8) such that

\[
\mathbf{M}(\mathbf{x}_{k}[n],\mathbf{M},\mathbf{S})_{\text{emp}} = \sum_{n_{t}+1}^{n_{t}+1} \prod_{i \in S} X_{\chi, k}^{(p)}[n]
\]

Step 2) for \(\theta\) do

Solve (55) for \((\mathbf{w}_{k}^{(p)})_{\text{opt}}\), which is obtained as the unit norm eigenvector corresponding to non-zero eigenvalue of \((\mathbf{B}_{\chi, s}^{(p)})^{-1}\mathbf{a}_{k}^{(p)}(\theta)\mathbf{a}_{k}^{(p)}(\theta)^{H}\) using 1) \(\mathbf{B}_{\chi, s}^{(p)}\) in (33a) and \(\alpha_{k}^{(p)}\) in (51), or 2) \(\mathbf{B}_{\chi, s}^{(p)}\) in (33b) and \(\alpha_{k}^{(p)}\) in (53).

Incorporate \((\mathbf{w}_{k}^{(p)})_{\text{opt}}\) to

1) \(\mathbf{p}_{c} = 2\rho\text{-KR MUSIC}(\theta)\) in (50) or
2) \(\mathbf{p}_{c} = 2\rho\text{-KR Capon}(\theta)\) in (52).

end for

Step 3) Determine the DOAs corresponding to the local maxima of the proposed spatial spectra.

F. Identifiability of the c-2\(\rho\)-KR-MUSIC for an ULA

Given a ULA of \(M\) sensors equally spaced by \(d_{\chi}\) in a two-dimensional space located at \((x_{m} = m d_{\chi}, m = 0, \ldots, M - 1)\) with the array response vector \(\mathbf{a}_{k}(\theta)\) defined in (4), the virtual sensors of the virtual array response vector \(\mathbf{a}_{k}^{(p)}(\theta)\) of \(\mathbf{A}_{k}^{(p)}\) defined in (17) (see [22]) are located at \((x_{n} = n d_{\chi}, y_{m} = m d_{\theta})\), \(n = -\rho(M - 1), \ldots, 1, 0, 1, \ldots, \rho(M - 1)\). Examples are available in [16], [22], [25], [30], [33]. It is generally known that \(V\) number of different virtual sensors gives the DOA identifiability of \(V - 1\) for the 2\(\rho\)-MUSIC in [22], [25] and the KR-MUSIC [28]. Here, the number of different virtual sensors is \(2\rho(M - 1) + 1\) and therefore, the identifiability of the c-2\(\rho\)-KR-MUSIC based on the ULA [28], [33], which is a function of order \(\rho\) and \(M\), is given as

\[
I(\rho, M) = 2\rho(M - 1).
\] (57)

When \(\rho > 1\), the identifiability of c-2\(\rho\)-KR-MUSIC is greater than that of the KR-MUSIC which is \(2(M - 1)\) when \(\rho = 1\) [28]. Note that for an arbitrary geometry array such as the UCA and UPA, the c-2\(\rho\)-KR-MUSIC can give higher identifiability than that of the ULA [16], [22], [25], [30], [33].
The PoS is defined as the probability of estimating all the DOAs of sources and \( \text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \mathbf{E}(\theta_i - \hat{\theta}_i)^2} \) where \( \theta_i \) and \( \hat{\theta}_i \) denote the true and estimated DOAs. Here it is assumed there are \( I \) sources.

Experiments are performed using an ULA with distance \( d_x \) apart on 300 Monte Carlo trials at each SNR with the per-trial synthesized sources such that the sources for all the trials are different from one another.

A. Narrow-band non-stationary source signals generated from generalized Gaussian distribution (GGD)

Narrow-band non-stationary source signals are synthetically generated from GGD (see [28]). Here, the \( k \)th frequency SNR \( = 10 \log_{10} \left( \frac{P_{x,k}}{\sigma_{m,k}^2} \right) \) where \( P_{x,k} = \mathbf{E}(\| S_{i,k} \|)^2 \) and \( \sigma_{m,k}^2 = \mathbf{E}(\| Z_{m,k} \|)^2 \). The sources are generated such that \( S_{i,k}[n] = S_{i,k}[n] + jS_{i,k}[n](n = 0, \ldots, L_s - 1) \) where \( S_{i,k}[n] \) and \( S_{i,k}[n] \) are sampled independently from two different GGD. For instance, the GGD for \( S_{i,k}[n] \) is described as

\[
P \left( S_{i,k}[n] \mid \sqrt{\frac{P_{x,k}}{2}}, \beta \right) = \frac{\nu(\beta)}{\sqrt{P_{x,k}}} \cdot \exp \left\{ -c(\beta) \frac{S_{i,k}[n]}{\sqrt{P_{x,k}/2}} \right\}^{2/(1+\beta)}
\]

(58)

where \( c(\beta) = \frac{1}{\Gamma\left(\frac{1}{2}(1+\beta)\right)} \left( \frac{1}{\Gamma\left(\frac{1}{2}(1+\beta)\right)} \right)^{1/(1+\beta)} \), \( \nu(\beta) = \Gamma \left( \frac{3}{2}(1+\beta) \right)^{1/(1+\beta)} \Gamma \left( \frac{1}{2}(1+\beta) \right)^{3/2} \), and \( \Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)dt \) is the Gamma function. Here, \( \sqrt{P_{x,k}/2} > 0 \) is the standard deviation and \( \beta \) controls the distance between the GGD and a normal distribution. When \( \beta = 1 \), it is a Laplace distribution, when \( \beta = 0 \), it is a normal distribution and when \( \beta = -1 \), it is a uniform distribution. In our experiments, we use \( \{ \beta = 1, 0.7, 0.7, 1 \} \) for both super- and sub- Gaussian distribution classes. Non-stationary sources are synthetically generated from time-varying

\[
U(0, 1).
\]

For stationary duration \( l_s \sim U(L_{\text{low}}, L_{\text{upp}}) \), where \( L_{\text{low}} \) and \( L_{\text{upp}} \) are the smallest and largest stationary durations. In our experiment, \( L_{\text{low}} = 300 \) and \( L_{\text{upp}} = 700 \) are used.

Thus, a procedure to generate samples of \( S_{i,k}[n] \) is as follows.

The GGD is defined using a fixed \( \beta \) and \( \sqrt{P_{x,k}/2} \) sampled from \( U(0, 1) \), and \( S_{i,k}[n] \) and \( S_{i,k}[n] \) are sampled independently from the two identical GGDs based on a simple sampling method called importance sampling [24] to obtain \( l_s \) data samples where \( l_s \) is sampled from \( U(L_{\text{low}}, L_{\text{upp}}) \). This procedure is done recursively for \( t = 1, \ldots, b \), until \( S_{i,k}[n] \) and \( S_{i,k}[n] \) have \( L_s \) data samples. Fig. 1 shows an illustration of \( \{ S_{i,k}[n] \} \ n = 0, 1, 2, 3 \} \) for \( \beta = 1 \). Note that a set of \( I \) cyclostationary source signals with a common period \( P \) can be generated by a procedure described above. Samples of \( S_{i,k}[n] \) are generated from a GGD defined in (58) with a fixed \( \beta \) and standard deviation \( \sqrt{P_{x,k}/2} \) that varies after every \( l_s \) samples within each period, sampling from \( U(0, 1) \). It is assumed that there are \( b \) locally-stationary segments within each period where \( b > I \). The sequence of variation in \( \sqrt{P_{x,k}/2} \) within each period is remembered and repeated every periods. Therefore, \( S_{i,k}[n] \) is locally-stationary for \( l_s = l, \forall t \) samples and cyclostationary with period \( P \). The proposed algorithms are applicable whenever (22) holds, and it turns out (22) holds for the cyclostationary source signals satisfying the five assumptions and the condition for the number of locally-stationary segments \( b \): the \( D_{\text{gg}}(\theta) \) in (17) is repeated concatenation of source statistics of one period, and \( \text{rank}(D_{\text{gg}}(\theta)) = I \).

Based on the experimental settings and specifications given in Table II, the proposed algorithms were evaluated.

First, an experiment was conducted to give performance comparison according to the number of data samples \( L_s \) in Table II for the proposed algorithms, KR-MUSIC, KR-Capon and 4-MUSIC with a fixed \( \lambda \) and SNR. Fig. 2 shows that the RMSEs versus the number of segments \( b \) of the proposed algorithms, KR-MUSIC, KR-Capon and 4-MUSIC when \( (M, I) = (2, 2), \theta_0 = 40^\circ, \theta_1 = 70^\circ, \forall t, l_s = 512 \) and SNR = 10 dB. Over certain number of data samples such as \( L_s = l_s \times b \) given a segment-length \( l_s \), the c-2-KR-MUSIC and the c-2-KR-Capon can give better RMSEs than the 4-MUSIC. Irrespective of the number of data samples, the c-2-KR-MUSIC and c-2-KR-Capon can outperform the KR-MUSIC and KR-Capon, respectively.

The following experiments are carried out based on Table II.

1) Determined case: I. High-resolution capability

Fig. 3 shows the spatial spectra of KR-Capon, c-2-KR-Capon, 4-KR-Capon and c-4-KR-Capon when \( (M, I) = (2, 2), \theta_0 = 40^\circ, \theta_1 = 42^\circ \) and SNR = 20 dB. The spatial spectra show that the c-2-KR-Capon and c-4-KR-Capon can produce better DOA resolution than the KR-Capon and 4-KR-Capon, respectively.
Fig. 2. The RMSEs versus the number of segments $b$ of the proposed algorithms, KR-MUSIC, KR-Capon and 4-MUSIC when $(M, I) = (2, 2)$, $\theta_0 = 40^\circ$, $\theta_1 = 70^\circ$, $\nu_t = 512$ and SNR = 10 dB.

Fig. 3. The spatial spectra of KR-Capon, c-2-KR-Capon, 4-KR-Capon and c-4-KR-Capon when $(M, I) = (2, 2)$, $\theta_0 = 40^\circ$, $\theta_1 = 42^\circ$ and SNR = 20 dB.

Fig. 4. The spatial spectra of the KR-MUSIC, c-2-KR-MUSIC, 4-KR-MUSIC, c-4-KR-MUSIC and 4-MUSIC when $(M, I) = (2, 2)$, $\theta_0 = 40^\circ$, $\theta_1 = 42^\circ$ and SNR = 15 dB.

Fig. 5. The RMSEs versus SNR for the proposed algorithms, the previous KR subspace-based algorithms and the 4-MUSIC.

Fig. 6. The RMSEs versus SNR for the proposed algorithms, the previous KR subspace-based algorithms and the 4-MUSIC.

Section III-D, compared to those of KR-MUSIC, KR-Capon and 4-MUSIC. Here, the values of all the spatial spectra are normalized between 0 and 1.}

**II. Pos and RMSE versus SNR**

Fig. 5 shows the RMSEs versus SNR for the proposed algorithms, the previous KR subspace-based algorithms and the 4-MUSIC when $(M, I) = (2, 2)$, $\theta_0 = 40^\circ$ and $\theta_1 = 70^\circ$ with the PoS equals to 1. At any SNR, the RMSEs of the c-2-KR-Capon, c-4-KR-Capon, c-2-KR-MUSIC and c-4-KR-MUSIC (dashed lines) are lower than those of the KR-Capon, 4-KR-Capon, KR-MUSIC, 4-KR-MUSIC (solid lines), respectively. The proposed algorithms when $\rho = 1$ outperform the proposed algorithms when $\rho = 2$; this might result from the fact that the estimates of local $2\rho$th-order cumulants for $\rho > 1$ have higher variance than those of local SO cumulants, given a short segment-length from experimental results in [16], [22], [25], [32].

Under identical experimental setup, the c-2-KR-MUSIC-M, c-2-KR-Capon-M, c-4-KR-MUSIC-M and c-4-KR-Capon-M are evaluated (see in Table 1) Proposed algorithms for $\rho = 1, 2$; -M denotes use of $E_e$ ($\mathbb{C}_{\mathbb{K}}^{(\rho)}$) instead of $\mathbb{C}_{\mathbb{K}}^{(\rho)}$ for $\rho = 1, 2$. Fig. 6 shows the RMSEs versus SNR for
the proposed algorithms, the previous KR subspace-based algorithms and the 4-MUSIC. The c-2-KR-MUSIC-M, c-2-KR-Capon-M, c-4-KR-MUSIC-M and c-4-KR-Capon-M outperform the KR-MUSIC, KR-Capon, 4-KR-MUSIC and 4-KR-Capon, respectively. Using $E_i \left( \hat{\theta}_{x_a}^{(p)} \right)$ increases the RMSE margins between the dashed lines (the RMSEs for the c-2-KR-MUSIC-M, c-2-KR-Capon-M, c-4-KR-MUSIC-M and c-4-KR-Capon-M) and the solid lines (the RMSEs for the KR-MUSIC, KR-Capon, 4-KR-MUSIC and 4-KR-Capon). This fact is realized by comparing the margins in Fig. 5 with those in Fig. 6. This comparison shows that $E_i \left( \hat{\theta}_{x_a}^{(p)} \right)$ is indeed effective in providing more accurate DOAs.

2) Underdetermined case: I. Identifiability

Fig. 7 shows the spatial spectra of 4-KR-MUSIC, c-4-KR-MUSIC, 4-KR-Capon and c-4-KR-Capon when $(M, I) = (2, 4), \theta_0 = 40^\circ, \theta_1 = 70^\circ, \theta_2 = 110^\circ, \theta_3 = 150^\circ$ and SNR = 20 dB. The spatial spectra of 4-KR-MUSIC, c-4-KR-MUSIC, 4-KR-Capon and c-4-KR-Capon have four local peaks and therefore, can identify up to 2$p(M-1)$ sources when $p = 2$, while the KR-Capon and KR-MUSIC cannot estimate the DOAs: the identifiability of the KR-MUSIC and KR-Capon is only $2(M - 1)$ [28]. Therefore, given a fixed number of sensors and array geometry, the 4-KR-MUSIC, 4-KR-Capon, 4-KR-Capon and c-4-KR-Capon can estimate the DOAs of sources more than the KR-MUSIC, KR-Capon and also the 4-MUSIC.

II. High-resolution capability

Fig. 8 shows the spatial spectra of 4-KR-MUSIC, c-4-KR-MUSIC, 4-KR-Capon and c-4-KR-Capon when $(M, I) = (2, 3), \theta_0 = 40^\circ, \theta_1 = 55^\circ, \theta_2 = 100^\circ$ and SNR = 20 dB. The spatial spectra show that the c-4-KR-MUSIC and c-4-KR-Capon can produce better DOA resolution than the 4-KR-MUSIC and 4-KR-Capon, respectively.

II. PoS and RMSE versus SNR

Fig. 9 shows the RMSEs versus SNR for the 4-KR-MUSIC, c-4-KR-MUSIC-M, 4-KR-Capon and c-4-KR-Capon-M when $(M, I) = (2, 3), \theta_0 = 40^\circ, \theta_1 = 70^\circ$ and $\theta_2 = 110^\circ$ with the PoS equals to 1. The c-4-KR-MUSIC-M and c-4-KR-Capon-M outperform the 4-KR-MUSIC and 4-KR-Capon, respectively.

B. Wide-band non-stationary source signals: speech and audio

This section considers DOA estimation for wide-band source signals such as speech and audio. Fig. 10 shows an illustration of speech/audio $\{s_i[n], i = 0, 1, 2, 3\}$ sampled at $f_s = 16$ kHz. The STDFTs of received sensor signals are processed for each frequency bin by the proposed algorithms [26]. Table III describes the experimental settings and specifications of the proposed algorithms. Note that in Table III, different number of the DFT points $N$ can improve the performance of the proposed algorithms [26].

I) Determined case: I. High-resolution capability

Fig. 11 shows the spatial spectra of KR-Capon, c-2-KR-Capon, 4-KR-Capon and c-4-KR-Capon when $(M, I) = (2, 2), \theta_0 = 30^\circ, \theta_1 = 37^\circ$ and SNR = 30 dB. The c-2-KR-Capon and c-4-KR-Capon can estimate the DOAs of two sources explicitly, but the KR-Capon and 4-KR-Capon cannot. In Fig. 12, the left figure shows the spatial spectra of KR-MUSIC and c-2-KR-MUSIC when $(M, I) = (2, 2), \theta_0 = 30^\circ, \theta_1 = 45^\circ$, and SNR = 30 dB, the right figure shows the spatial spectra of 4-KR-MUSIC and c-4-KR-MUSIC when $(M, I) = (2, 2)$,
TABLE III
THE EXPERIMENTAL SETTINGS AND SPECIFICATIONS FOR THE PROPOSED ALGORITHMS

<table>
<thead>
<tr>
<th>Spec.</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>the length of a source signal</td>
<td>( L_y = 200 (l_x \times 400) )</td>
</tr>
<tr>
<td>the length of window ( w[n] )</td>
<td>( L_w = 256 )</td>
</tr>
<tr>
<td>the window shift for ST-DFT</td>
<td>4 samples</td>
</tr>
<tr>
<td>( I_x ), ( I_y ) in estimating ( C_{M,k}^{(\rho)} )</td>
<td>( L_s = 200 )</td>
</tr>
<tr>
<td>( \alpha_{\phi}^{(\rho)} )</td>
<td>( \max(\Sigma_{\phi}, \Sigma_{\psi}) ) for c-2(p)-KR-MUSIC-M and c-4(p)-KR-Capon-M</td>
</tr>
<tr>
<td>( \alpha_{\phi}^{(\rho)} )</td>
<td>( 0.1 \min(\Sigma_{\phi}, \Sigma_{\psi}) ) for c-2(p)-KR-Capon-M</td>
</tr>
</tbody>
</table>

Fig. 11. The spatial spectra of KR-Capon, c-2-KR-Capon, 4-KR-Capon and c-4-KR-Capon when \((M, I) = (2, 2), \theta_0 = 30^\circ, \theta_1 = 37^\circ\) and SNR = 30 dB.

Fig. 12. The spatial spectra of KR-MUSIC and c-2-KR-MUSIC when \((M, I) = (2, 2), \theta_0 = 30^\circ, \theta_1 = 37^\circ\), and \( \theta_2 = 50^\circ \) and SNR = 25 dB in the right figure, and the spatial spectra of 4-MUSIC and c-2-KR-MUSIC when \( \theta_0 = 30^\circ, \theta_1 = 44^\circ \) and SNR = 25 dB in the bottom figure.

\( \theta_0 = 30^\circ, \theta_1 = 50^\circ \) and SNR = 25 dB, and the bottom figure shows the spatial spectra of 4-MUSIC and c-2-KR-MUSIC when \( \theta_0 = 30^\circ, \theta_1 = 44^\circ \) and SNR = 25 dB. The c-2-KR-MUSIC and c-4-KR-MUSIC can estimate the DOAs of two sources with more confidence than the KR-MUSIC and 4-KR-MUSIC, respectively. The c-2-KR-MUSIC can also give better DOA resolution than the 4-MUSIC. Note that the c-2\(p\)-KR-MUSIC and c-2\(p\)-KR-Capon have better DOA resolution capability than the 2\(p\)-KR-MUSIC and 2\(p\)-KR-Capon when \( \rho = 1, 2 \), respectively.

II. PoS and RMSE versus SNR

When \((M, I) = (2, 2), \theta_0 = 36^\circ \) and \( \theta_1 = 76^\circ \), Fig. 13 shows the PoS versus SNR for the proposed algorithms, the previous KR subspace-based algorithms and the 4-MUSIC. Each PoS of c-2\(p\)-KR-MUSIC-M and c-2\(p\)-KR-Capon-M is higher than that of 2\(p\)-KR-MUSIC and 2\(p\)-KR-Capon when \( \rho = 1, 2 \) (especially in the case where the SNR is low, and also for the proposed algorithms when \( \rho = 2 \)). Fig. 14 shows the RMSEs versus SNR for the proposed algorithms, the previous KR subspace-based algorithms and the 4-MUSIC. Note that the empty markers for some algorithms in Fig. 14 signify that the PoS equals to 0, and the RMSEs can not be estimated. The c-2\(p\)-KR-MUSIC-M and c-2\(p\)-KR-Capon-M outperform the 2\(p\)-KR-MUSIC and 2\(p\)-KR-Capon when \( \rho = 1, 2 \), respectively.

2) Underdetermined case: I. Identifiability

The experiments to evaluate the identifiability of the proposed algorithms when \( \rho = 2 \), are discussed in Section IV-A-2 for narrow-band non-stationary source signals and omitted for wide-band non-stationary source signals.

II. High-resolution capability

Fig. 15 shows the spatial spectra of 4-KR-MUSIC, c-4-KR-MUSIC, 4-KR-Capon and c-4-KR-Capon when \((M, I) = (2, 3), \theta_0 = 30^\circ, \theta_1 = 50^\circ, \theta_2 = 110^\circ \)
TABLE IV

<table>
<thead>
<tr>
<th>Proposed algorithms</th>
<th>Advantages</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-2-KR-MUSIC(-M)</td>
<td>• better resolution and RMSE than KR-MUSIC</td>
<td>• no guarantee to be better than 4-MUSIC</td>
</tr>
<tr>
<td></td>
<td>and 4-MUSIC over certain number of data samples</td>
<td>under certain number of data samples</td>
</tr>
<tr>
<td>c-2-KR-Capon(-M)</td>
<td>• better resolution and RMSE than KR-Capon</td>
<td>• inability for stationary sources compared to 4-MUSIC</td>
</tr>
<tr>
<td></td>
<td>and 4-MUSIC over certain number of data samples</td>
<td></td>
</tr>
<tr>
<td>(c)-4-KR-MUSIC(-M)</td>
<td>• higher identifiability than KR-MUSIC and 4-MUSIC</td>
<td>• no guarantee to be higher than 4-MUSIC</td>
</tr>
<tr>
<td></td>
<td>over certain number of data samples</td>
<td>under certain number of data samples</td>
</tr>
<tr>
<td>(c)-4-KR-Capon(-M)</td>
<td>• higher identifiability than KR-Capon and 4-MUSIC</td>
<td>• inability for stationary sources compared to 4-MUSIC</td>
</tr>
<tr>
<td></td>
<td>over certain number of data samples</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 15. The spatial spectra of 4-KR-MUSIC, c-4-KR-MUSIC, 4-KR-Capon and c-4-KR-Capon when \((M, I) = (2, 3), \theta_0 = 30^\circ, \theta_1 = 50^\circ, \theta_2 = 110^\circ\) and \(\text{SNR} = 20\ \text{dB}\).

Fig. 16. The PoS versus SNR for the 4-KR-MUSIC, c-4-KR-MUSIC-M, 4-KR-Capon and c-4-KR-Capon-M.

III. PoS and RMSE versus SNR

When \((M, I) = (2, 3), \theta_0 = 40^\circ, \theta_1 = 70^\circ\) and \(\theta_2 = 100^\circ\), Fig. 16 shows the PoS versus SNR for the 4-KR-MUSIC, c-4-KR-MUSIC-M, 4-KR-Capon and c-4-KR-Capon-M. Each PoS of c-4-KR-MUSIC-M and c-4-KR-Capon-M is much higher than that of 4-KR-MUSIC and 4-KR-Capon (especially in the case where the SNR is low). Fig. 17 shows the RMSEs versus SNR for the 4-KR-MUSIC, c-4-KR-MUSIC-M, 4-KR-Capon and c-4-KR-Capon-M. The c-4-KR-MUSIC-M and c-4-KR-Capon-M outperform the 4-KR-MUSIC and 4-KR-Capon, respectively.

Therefore, it is demonstrated empirically that the c-2p-KR-MUSIC-M and c-2p-KR-Capon-M when \(\rho = 1, 2\) are useful in estimating the DOAs for non-stationary and non-Gaussian source signals such as speech and audio for determined and underdetermined cases. Table IV as a summarization of Section IV, gives the advantages and the weaknesses of the proposed algorithms with respect to KR-Capon, KR-MUSIC and 4-MUSIC.

V. CONCLUSION

In this paper, two types of high-resolution DOA estimation algorithms, the c-2p-KR-MUSIC(-M) and c-2p-KR-Capon(-M) are proposed. When \(\rho = 1\) and under certain parameter setting in the constraint, the spatial spectra of c-2p-KR-MUSIC and the c-2p-KR-Capon are identical to those of KR-MUSIC and KR-Capon, respectively. Using the space diversity provided by an ULA of \(M\) sensors, the c-2p-KR-MUSIC can theoretically be shown to identify up to \(2p(M-1)\) sources and when \(\rho > 1\), an improvement for the identifiability over the KR-MUSIC and 4-MUSIC whose identifiability is \(2(M-1)\), can be obtained. Experimental results demonstrate that when \(\rho = 1\), the c-2p-KR-MUSIC(-M) and c-2p-KR-Capon(-M) can give higher-resolution DOA estimation and better RMSEs than the previous KR subspace-based algorithms, the KR-MUSIC, KR-Capon and the 4-MUSIC, and when \(\rho = 2\), the c-2p-KR-MUSIC(-M) and c-2p-KR-Capon(-M) respectively perform better than the 2p-KR-MUSIC and 2p-KR-Capon for globally non-stationary, non-Gaussian synthetic and real speech/audio sources under various adverse conditions. We verified that the identifiability for \(\rho = 2\) is improved by two-folds compared to that for \(\rho = 1\) for the ULA.
The above COP (31) with (32), can be solved using the Lagrange multiplier $\lambda^{(p)}_k$ and the following Lagrangian is given as

$$L \left( \lambda^{(p)}_k, (w_k^{(p)})_\theta \right) = \left( w_k^{(p)} \right)_\theta^H a_k^{(p)}(\theta) \left( a_k^{(p)}(\theta) \right)_\theta^H \left( w_k^{(p)} \right)_\theta - \lambda^{(p)}_k \left( \left( w_k^{(p)} \right)_\theta^H B_k^{(p)} \left( w_k^{(p)} \right)_\theta - \left( c_k^{(p)} \right)_\theta \right)$$

where $\lambda^{(p)}_k > 0$. Taking the partial derivative of $L(\lambda^{(p)}_k, (w_k^{(p)})_\theta)$ with respect to $(w_k^{(p)})_\theta$

$$\frac{\partial L}{\partial (w_k^{(p)})_\theta} = 2 a_k^{(p)}(\theta) \left( a_k^{(p)}(\theta) \right)_\theta^H \left( w_k^{(p)} \right)_\theta - 2 \lambda^{(p)}_k B_k^{(p)} \left( w_k^{(p)} \right)_\theta$$

Setting the above gradient to zero, the optimal weight vector $(w_k^{(p)})_{\theta, opt}$ satisfies

$$a_k^{(p)}(\theta) \left( a_k^{(p)}(\theta) \right)_\theta^H \left( w_k^{(p)} \right)_{\theta, opt} - \lambda^{(p)}_k B_k^{(p)} \left( w_k^{(p)} \right)_{\theta, opt} = 0$$

In other words, $(w_k^{(p)})_{\theta, opt}$ is the generalized eigenvector of $a_k^{(p)}(\theta)(a_k^{(p)}(\theta))^H$ and $B_k^{(p)}$. Here, vector $v_k^{(p)} (\in \mathbb{C}^{M_p \times 1})$ is given as

$$v_k^{(p)} = \left( T_k^{(p)} \right)^{1/2} (w_k^{(p)})_{\theta, opt}$$

where

$$\left( T_k^{(p)} \right)^{1/2}^H \left( T_k^{(p)} \right)^{1/2} = B_k^{(p)}.$$

Therefore,

$$(w_k^{(p)})_{\theta, opt} = \left( T_k^{(p)} \right)^{1/2} v_k^{(p)}.$$

Inserting (64) into (61),

$$a_k^{(p)}(\theta) \left( a_k^{(p)}(\theta) \right)_\theta^H \left( T_k^{(p)} \right)^{1/2} v_k^{(p)} = \lambda^{(p)}_k B_k^{(p)} \left( T_k^{(p)} \right)^{1/2} v_k^{(p)} - \lambda^{(p)}_k \left( T_k^{(p)} \right)^{1/2} v_k^{(p)}$$

and

$$\left( T_k^{(p)} \right)^{1/2} a_k^{(p)}(\theta) \left( a_k^{(p)}(\theta) \right)_\theta^H \left( T_k^{(p)} \right)^{1/2} v_k^{(p)} - \lambda^{(p)}_k \left( T_k^{(p)} \right)^{1/2} v_k^{(p)} = 0.$$

Thus, $v_k^{(p)}$ is one of the eigenvector of the matrix

$$Q_k^{(p)} = \left( T_k^{(p)} \right)^{1/2} - \lambda^{(p)}_k \left( T_k^{(p)} \right)^{1/2}. $$

with

$$b_k^{(p)} = \left[ \left( T_k^{(p)} \right)^{1/2} \right]^H a_k^{(p)}(\theta).$$

Clearly rank($Q_k^{(p)}$) = 1 and the optimal vector of $v_k^{(p)}$ is given as

$$v_k^{(p, opt)} = \beta_k^{(p)} b_k^{(p)}$$

$$- \beta_k^{(p)} \left[ \left( T_k^{(p)} \right)^{1/2} \right]^H a_k^{(p)}(\theta).$$

Therefore, inserting (68) into (64),

$$(w_k^{(p)})_{\theta, opt} = \beta_k^{(p)} \left[ \left( T_k^{(p)} \right)^{1/2} \right]^H a_k^{(p)}(\theta)$$

By inserting (69) into (32) and rearranging (32) with respect to $\beta_k^{(p)}$,

$$\beta_k^{(p)} = \sqrt{c_k^{(p)}} \left\{ (a_k^{(p)}(\theta))^H B_k^{(p)} \right\}^{-1} a_k^{(p)}(\theta).$$

**Appendix B**

Suppose that

$$(w_k^{(p)})_{\theta, opt} = s_k^{(p)}(\theta) u_k^{(p)}(\theta) + n_k^{(p)}(\theta) u_k^{(p, k}(\theta))$$

where $u_k^{(p)}(\theta) (\in \mathbb{R}(U_k^{(p)}))$ and $n_k^{(p)}(\theta) (\in \mathbb{N}(U_k^{(p)}))$ are any unit norm vectors to the look direction $\theta$. Here, $s_k^{(p)}(\theta)$ is the strength of the unit norm vectors, and it should be noted that

$$\left\| (w_k^{(p)})_{\theta, opt} \right\|_2^2 = \left\| s_k^{(p)}(\theta) \right\|_2^2 + \left\| n_k^{(p)}(\theta) \right\|_2^2 - c$$

where $c$ is a constant (in the proposed algorithms, $c = 1$).

**A. The c-2p-KR-MUSIC with estimation error**

Consider the spatial spectrum of the c-2p-KR-MUSIC in (50) and

$$\left\| \left( (w_k^{(p)})_{\theta, opt} \right)^H \tilde{U}_k^{(p)} \right\|_2^2$$

$$- \left\| \left( s_k^{(p)}(\theta) u_k^{(p)}(\theta) + n_k^{(p)}(\theta) u_k^{(p, k}(\theta) \right)^H (U_k^{(p)} + \Delta U_k^{(p)}) \right\|_2^2$$

$$= \left\| s_k^{(p)}(\theta) u_k^{(p)}(\theta) \right\|_2^2 \Delta U_k^{(p)} + \left\| n_k^{(p)}(\theta) u_k^{(p, k}(\theta) \right\|_2^2 \Delta U_k^{(p)}$$

$$+ \left\| n_k^{(p)}(\theta) \right\|_2^2 \left( u_k^{(p)}(\theta) H U_k^{(p)} \right)$$
The fact that \( \{u^{(p)}_{s,k}(\theta)\}^H U_{n,k}^{(p)},\{u^{(p)}_{n,k}(\theta)\}^H \Delta U_{s,k}^{(p)} \) and \( \{u^{(p)}_{n,k}(\theta)\}^H U_{n,k}^{(p)} \) are all null, is used in the derivation (73). It is assumed that the estimation errors are small such that

\[
|\Delta U_{s,k}^{(p)}|_F \ll \|U_{n,k}^{(p)}\|_F
\]

(74)

where \( \| \cdot \|_F \) is the Frobenius norm.

For high-resolution DOA estimation, the following condition must be satisfied:

\[
\begin{align*}
&\left\|s^{(p)}_k(\theta_1)\right\|^2_2 \left((u^{(p)}_{s,k}(\theta_1))^H \Delta U_{s,k}^{(p)}\right)^2 + \left\|n^{(p)}_s(\theta_1)\right\|^2_2 \left((u^{(p)}_{n,k}(\theta_1))^H U_{n,k}^{(p)}\right)^2 \\
\ll &\left\|s^{(p)}_k(\theta \neq \theta_1)\right\|^2_2 \left((u^{(p)}_{s,k}(\theta \neq \theta_1))^H \Delta U_{s,k}^{(p)}\right)^2 + \left\|n^{(p)}_s(\theta \neq \theta_1)\right\|^2_2 \left((u^{(p)}_{n,k}(\theta \neq \theta_1))^H U_{n,k}^{(p)}\right)^2 .
\end{align*}
\]

(75)

1) **Left-hand-side of (75):**

In order for the left-hand-side of (75) to be small,

\[
s^{(p)}_k(\theta_1) \gg n^{(p)}_s(\theta_1) > 0
\]

(76)
due to (72) and (74). In other words, the estimation error of \( (w^{(p)}_{s,k})_{\theta \neq \theta_1} \) in \( \mathcal{N}(U_{s,k}^{(p)}) \) leads to \( (w^{(p)}_{s,k})_{\theta = \theta_1, \theta \neq \theta_1} \in \mathcal{R}(U_{s,k}^{(p)}) \cup \mathcal{N}(U_{s,k}^{(p)}) \), and the error should be small. This happens when \( \alpha^{(p)}_k \gg \max(\Sigma^{(p)}_{s,k}) \) — compare this to the theoretical case when \( \theta \neq \theta_1 \). Here, \( (w^{(p)}_{s,k})_{\theta \neq \theta_1} \in \mathcal{R}(C^{(p)}_{s,k}) \cup \mathcal{N}(C^{(p)}_{s,k}) \) that happens when \( \alpha^{(p)}_k \gg \max(\Sigma^{(p)}_{s,k}) \). Thus, this can lead to a high spatial spectrum value of the c-2p-KR-MUSIC in (50).

2) **Right-hand-side of (75):**

In order for the right-hand-side of (75) to be large,

\[
\begin{align*}
&\left\|s^{(p)}_k(\theta \neq \theta_1)\right\|^2_2 - \left\|s^{(p)}_k(\theta_1)\right\|^2_2 - \zeta > 0, \\
&\left\|n^{(p)}_s(\theta \neq \theta_1)\right\|^2_2 - \left\|n^{(p)}_s(\theta_1)\right\|^2_2 + \zeta > 0
\end{align*}
\]

(77)

where \( \zeta > 0 \). Here, \( \zeta \) must be large so that the right-hand-side of (75) is large as possible. This happens when \( \alpha^{(p)}_k \) is small as possible such that the portion \( n^{(p)}_s(\theta \neq \theta_1) \) is much more than \( s^{(p)}_k(\theta \neq \theta_1) \) — compare this to the theoretical case when \( \theta \neq \theta_1 \) where \( (w^{(p)}_{s,k})_{\theta \neq \theta_1} \in \mathcal{R}(C^{(p)}_{s,k}) \) that happens when \( 0 < \alpha^{(p)}_k \ll \min(\Sigma^{(p)}_{s,k}) \). This, thus, can lead to a low spatial spectrum value of the c-2p-KR-MUSIC in (50).

Therefore, considering \( \alpha^{(p)}_k \) for the conditions (76) and (77), an appropriate \( \alpha^{(p)}_k \) is set to (51) to satisfy (75) for the c-2p-KR-MUSIC.

**B. The c-2p-KR-Capon with estimation error**

Consider the spatial spectrum of the c-2p-KR-Capon in (52) by assuming \( \Sigma^{(p)}_{n,k} \) is null in (47), and

\[
\begin{align*}
&\left\|s^{(p)}_k(\theta)_s^{\text{opt}}\right\|^2_2 \left((w^{(p)}_{s,k})_{\theta} \right)^H U_{s,k}^{(p)} + \left\|n^{(p)}_s(\theta)\right\|^2_2
\end{align*}
\]

(78)

Where \( \{u^{(p)}_{s,k}(\theta)\}^H U_{s,k}^{(p)} + \{u^{(p)}_{n,k}(\theta)\}^H \Delta U_{s,k}^{(p)} \) and \( \{u^{(p)}_{n,k}(\theta)\}^H U_{n,k}^{(p)} \) are all null, is used in the derivation (78). It is assumed that the estimation errors are small such that

\[
\|U_{s,k}^{(p)}\|_F \gg \|\Delta U_{s,k}^{(p)}\|_F
\]

(79)

For high-resolution DOA estimation, the following condition must be satisfied:

\[
\begin{align*}
&\left\|s^{(p)}_k(\theta_1)\right\|^2_2 \left((u^{(p)}_{s,k}(\theta_1))^H U_{s,k}^{(p)}\right)^2 + \left\|n^{(p)}_s(\theta_1)\right\|^2_2 \left((u^{(p)}_{n,k}(\theta_1))^H \Delta U_{s,k}^{(p)}\right)^2 \\
\gg &\left\|s^{(p)}_k(\theta \neq \theta_1)\right\|^2_2 \left((u^{(p)}_{s,k}(\theta \neq \theta_1))^H U_{s,k}^{(p)}\right)^2 + \left\|n^{(p)}_s(\theta \neq \theta_1)\right\|^2_2 \left((u^{(p)}_{n,k}(\theta \neq \theta_1))^H \Delta U_{s,k}^{(p)}\right)^2 .
\end{align*}
\]

(80)

1) **Left-hand-side of (80):**

In order for the left-hand-side of (80) to be large,

\[
s^{(p)}_k(\theta_1) \gg n^{(p)}_s(\theta_1) > 0,
\]

(81)
due to (72) and (79). In other words, the estimation error of \( (w^{(p)}_{s,k})_{\theta \neq \theta_1} \) leads to \( (w^{(p)}_{s,k})_{\theta = \theta_1, \theta \neq \theta_1} \in \mathcal{R}(U_{s,k}^{(p)}) \cup \mathcal{N}(U_{s,k}^{(p)}) \), and the error should be small. This happens when \( \alpha^{(p)}_k \gg \max(\Sigma^{(p)}_{s,k}) \) — compare this to the theoretical case when \( \theta \neq \theta_1 \). Here, \( (w^{(p)}_{s,k})_{\theta \neq \theta_1} \in \mathcal{R}(C^{(p)}_{s,k}) \cup \mathcal{N}(C^{(p)}_{s,k}) \) that happens when \( \alpha^{(p)}_k \gg \max(\Sigma^{(p)}_{s,k}) \). Thus, this can lead to a high spatial spectrum value of the c-2p-KR-Capon in (52).

2) **Right-hand-side of (80):**

In order for the right-hand-side of (80) to be small

\[
\begin{align*}
&\left\|s^{(p)}_k(\theta \neq \theta_1)\right\|^2_2 \left((u^{(p)}_{s,k}(\theta \neq \theta_1))^H U_{s,k}^{(p)}\right)^2 - \left\|s^{(p)}_k(\theta_1)\right\|^2_2 - \zeta > 0, \\
&\left\|n^{(p)}_s(\theta \neq \theta_1)\right\|^2_2 \left((u^{(p)}_{n,k}(\theta \neq \theta_1))^H \Delta U_{s,k}^{(p)}\right)^2 + \zeta > 0
\end{align*}
\]

(82)

where \( \zeta > 0 \). Here, \( \zeta \) must be large so that the right-hand-side of (80) is small as possible. This happens when \( \alpha^{(p)}_k \) is small as
possible such that the portion $r_\ell^{(\theta)}(\theta \neq \theta_1)$ is much more than $s_\ell^{(\theta)}(\theta \neq \theta_1)$-compare this case to the theoretical case when $\theta \neq \theta_1$, where $(w_k^{(\theta)}(\theta_{\text{opt}}) \in \mathcal{N}(\alpha_\ell^{(\theta)}))$ that happens when $0 < \alpha_\ell^{(\theta)} < \min(\Sigma_\ell^{(\theta)})$. Thus, this can lead to a low spatial spectrum value of the c-2-\rho-KR-Capon in (52).

Therefore, considering $s_\ell^{(\theta)}$ for the conditions (81) and (82), an appropriate $s_\ell^{(\theta)}$ is set to (53) to satisfy (80) for the c-2-\rho-KR-Capon.

ACKNOWLEDGMENT

The authors would like to thank Prof. J. P. Delmas and anonymous reviewers for their invaluable comments and suggestions in improving this paper.

REFERENCES


Jin Ho Choi (S'06) received the M.S. degree in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2008. He is currently pursuing the Ph.D degree in the Department of Electrical Engineering, KAIST. His research interest is array signal processing, statistical signal processing and machine learning. From August 2013, he is CEO and CTO of SmartEar Company at KAIST.

Chang D. Yoo (S'92–M'96–SM'11) received the B.S. degree in electrical engineering and applied science from the California Institute of Technology in 1986, the M.S. degree in electrical engineering from Cornell University in 1988, and the Ph. D. degree in electrical engineering from the Massachusetts Institute of Technology (MIT) in 1996. From January 1997 to March 1999, he worked at Korea Telecom as a senior researcher. He joined the Department of Electrical Engineering at the Korea Advanced Institute of Science and Technology in April 1999. From March 2005 to March 2006, he was with the Research Laboratory of Electronics at MIT. His current research interests include the application of machine learning and digital signal processing in multimedia.