Error analysis and Cramer–Rao bound for convolutional beamspace

Po-Chih Chen, *Student member, IEEE* and P. P. Vaidyanathan, *Life Fellow, IEEE*

***Abstract*—The MSE performance and Cramer–Rao Bound (CRB) for a recently proposed beamspace method, called convolutional beamspace (CBS), are studied. Beamspace processing for DOA estimation offers lower computational complexity and higher DOA resolution. However, traditional beamspace methods do not retain the Vandermonde structure of uniform linear array output, so additional preparation is required to apply root-MUSIC. CBS preserves the Vandermonde structure by using digital filtering. CBS also offers smaller estimation errors for correlated sources. In this paper, we analyze the MSE performance of CBS when MUSIC or root-MUSIC is used. The variance is derived from the asymptotic probability distribution of the eigenvectors of an average finite-snapshot covariance matrix. Meanwhile, the bias due to the filtered stopband sources is given by a first-order perturbation analysis. Known advantages of CBS are confirmed by the MSE analysis. For example, CBS yields smaller MSE for correlated sources than element-space. Moreover, the CRB is also derived. Conventionally, the CRB is a lower bound for unbiased estimators. A lower bound on the variances of the biased CBS estimator is obtained and shown to be well approximated by the classical CRB for unbiased estimators. Two forms of CRB expressions are derived, and they offer different insights as explained in the paper. All the results also apply to element-space since element-space is a special case of CBS. Finally, the theoretical results are verified by simulations.**

***Index Terms*—Convolutional beamspace, DOA estimation, MUSIC, MSE, Cramer–Rao bound.**

I. Introduction

CONVOLUTIONAL beamspace (CBS) is a beamspace method for direction-of-arrival (DOA) estimation recently proposed in [1], [2]. Given an N-sensor uniform linear array (ULA) output $\mathbf{x}$, traditionally one does *beamspace* processing by computing $\mathbf{y} = \mathbf{T}\mathbf{x}$ and estimating DOAs based on $\mathbf{y}$ instead of $\mathbf{x}$. Here $\mathbf{y}$ has length less than N, so there is reduction in computational complexity. Besides, beamspace methods usually enjoy higher DOA resolution and smaller bias compared to *element-space* methods (which estimate DOAs directly using $\mathbf{x}$) [3]–[6]. In traditional beamspace methods [6], [7], the choice of $\mathbf{T}$ makes $\mathbf{y}$ lose the ULA Vandermonde structure, so elaborate steps need to be taken so as to apply standard DOA estimation methods like root-MUSIC or ESPRIT. To avoid this additional preparation, the idea of FIR filtering (convolution) is used in CBS [1]. In this way, the beamspace matrix $\mathbf{T}$ is a banded Toeplitz matrix so that the Vandermonde structure is preserved in $\mathbf{y}$. Therefore, we can readily use root-MUSIC [8] or ESPRIT [9] to estimate DOAs based on $\mathbf{y}$. In [1], uniform decimation (downsampling) on $\mathbf{y}$ is further proposed to achieve significant complexity reduction. More precisely, we compute the average of the covariance matrices of all polyphase components [10] of $\mathbf{y}$, and DOAs can be estimated based on the eigenvalue decomposition of the average covariance. The idea of using FIR filtering in beamspace methods is also proposed in [11], but the detailed development of how we use the filter output to estimate DOAs only appears in [1], [2].

This paper aims to study MSE performance of the CBS method. Besides low computation and compatibility with root-MUSIC and ESPRIT, CBS also offers performance advantage over element-space in some scenarios. While CBS and element-space have similar DOA estimation errors for uncorrelated sources, the estimation error of CBS can be significantly smaller than that of element-space when there are correlated sources. This benefit of CBS is demonstrated in [1] mainly through numerical examples, while theoretical MSE analysis is given only for limited simple cases. Moreover, some details are bypassed in the analysis in [1], so the results are only approximations. One main goal of this paper is thus to develop a rigorous and more accurate analysis for the MSE performance of CBS.

The foundations for the analysis of MSE performance for MUSIC and root-MUSIC were laid many decades ago in the classic papers [8], [12]. We extend the analysis to CBS in this paper. According to the results in [8], [12], MUSIC and root-MUSIC asymptotically have the same MSE in element-space, and we will show that it is also true for CBS. Compared to element-space, we have to tackle two additional complications in CBS. First, the different polyphase components of $\mathbf{y}$ are not independent, so it is more difficult to derive the asymptotic probability distribution of the eigenvectors of the average finite-snapshot covariance matrix. In the traditional asymptotic theory for principal component analysis [13], only independent observations are considered. This is directly applicable to element-space, but some modifications are required to adapt it to CBS. Second, the filter output $\mathbf{y}$ is represented only by the sources in the filter passband since we assume the attenuated stopband sources, if any, can be neglected [1]. This leads to an additional error term that should be analyzed. As we shall see, the effect of these filtered stopband sources is a bias, and it can be analyzed separately from the variance term. A first-order perturbation analysis is adopted. Furthermore, this bias term (squared) is often much smaller than the variance term if the CBS filter has reasonably good stopband attenuation and if the stopband source power is not very large.

In addition, the Cramer–Rao Bound (CRB) [14], [15] for CBS is also derived in this paper. Conventionally, the CRB offers a lower bound on the variances of unbiased estimates of parameters. As explained earlier, CBS yields biased DOA estimates if there are stopband sources. A modified form of CRB for a biased estimator of a scalar parameter is given in [16]. In our case, a biased estimator of a vector of parameters, i.e., DOAs, is considered, and we show that it can be viewed as an unbiased estimator for some transformation of the parameters. Hence, we can use the CRB for transformation of parameters [14] to get a modified lower bound on the

This work was supported by the Office of Naval Research grant N00014-21-1-2521, and the California Institute of Technology.

The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA (email: pochih@caltech.edu; ppvnath@systems.caltech.edu).

variances of CBS DOA estimates. This bound depends on the Jacobian matrix of the transformation and can be numerically computed. Moreover, assuming the stopband sources are reasonably attenuated, the modified bound can be well approximated by the original CRB for unbiased estimates. We study the CRB under the stochastic model [15], [17], where the source amplitudes are assumed random. We derive two forms of CRB expressions. Form 1 is in the same style as the CRB in [17]. Although it is not the focus of this paper, Form 1 can yield a CRB even when the noise is non-white with a singular covariance. On the other hand, Form 2 is in the same style as the CRB in [15] and offers some important insight. For example, the CRB for a DOA is approximately inversely proportional to its SNR and approximately independent of source powers of other DOAs. A necessary and sufficient condition for the CBS CRB to exist is also given. Although the main goal of this paper is to study the MSE and CRB for CBS, all the results also apply to element-space because element-space can be viewed as a special case of CBS.

Paper outline: The basics of convolutional beamspace (CBS) are reviewed in Sec. II, and notions for aiding later analysis are also presented. MSE for CBS is analyzed theoretically in Sec. III. Cramer–Rao Bound (CRB) for CBS is then derived in Sec. IV. In particular, two forms of CRB expressions are proposed in Sec. IV-A and Sec. IV-B. Numerical examples are shown in Sec. V to verify the theory. Finally, Sec. VI concludes the paper. Appendices A, B, and C contain detailed proofs of the results.

Notations: Boldfaced capital letters denote matrices, boldfaced lowercase letters are reserved for column vectors, $[\mathbf{A}]_{i,k}$ and $[\mathbf{A}]_{:,k}$ indicate the (i,k)-entry and the kth column of the matrix $\mathbf{A}$, and $[\mathbf{v}]_i$ is the ith entry of the vector $\mathbf{v}$. For a matrix $\mathbf{A}$, we use $\|\mathbf{A}\|$ to denote its spectral norm, i.e., maximum singular value, and $\mathrm{vec}(\mathbf{A})$ the vectorization of $\mathbf{A}$. We use $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^+$ to denote complex conjugate, transpose, conjugate transpose, and pseudoinverse, respectively. The Kronecker product, Khatri–Rao product, and Hadamard product of two matrices $\mathbf{A}$ and $\mathbf{B}$ are denoted by $\mathbf{A}\otimes\mathbf{B}$, $\mathbf{A}\odot\mathbf{B}$, and $\mathbf{A}\circ\mathbf{B}$, respectively. For any two Hermitian symmetric matrices $\mathbf{A}$ and $\mathbf{B}$, we use $\mathbf{A}\succeq\mathbf{B}$ and $\mathbf{B}\preceq\mathbf{A}$ to denote that $\mathbf{A}-\mathbf{B}$ is positive semidefinite. For square matrices $\mathbf{A}_1,\ldots,\mathbf{A}_n$, we use $\mathrm{diag}(\mathbf{A}_1,\ldots,\mathbf{A}_n)$ to denote the block diagonal matrix having $\mathbf{A}_1,\ldots,\mathbf{A}_n$ in the diagonal. The ith standard basis vector for the k-dimensional space is denoted by $\boldsymbol{\delta}_i^{(k)}$. We use $\mathbf{I}_n$ and $\mathbf{O}_{m,n}$ to denote the $n\times n$ identity matrix and $m\times n$ zero matrix (the subscripts m and n may be dropped if the dimensions are clear from the context), and $\mathrm{E}[\cdot]$ is the expectation operator. Finally, δ_{ik} denotes the Kronecker delta, i.e., $\delta_{ik}=1$ if $i=k$ and $\delta_{ik}=0$ if $i\neq k$.

II. Convolutional Beamspace

We consider an N-sensor ULA with sensor spacing $\lambda/2$, and assume D monochromatic plane waves of wavelength λ impinging on the array with DOAs $\theta_i\in[-\pi/2,\pi/2)$ measured from the normal to the line of array. The array output is thus

$$\mathbf{x}=\mathbf{A}\mathbf{c}+\mathbf{e}, \tag{1}$$

where $\mathbf{c}$ contains source amplitudes c_i, and $\mathbf{e}$ is additive noise. The array manifold $\mathbf{A}=[\mathbf{a}_N(\omega_1)\ \mathbf{a}_N(\omega_2)\cdots\mathbf{a}_N(\omega_D)]$, where $\mathbf{a}_N(\omega)=[1\ e^{j\omega}\ e^{j2\omega}\ \cdots\ e^{j(N-1)\omega}]^T$ and $\omega_i=\pi\sin\theta_i$. The stochastic (also known as unconditional) model [15] is considered. That is, we assume $\mathbf{c}$ is a circularly-symmetric complex Gaussian random vector with covariance $\mathbf{P}$, which can be non-diagonal if sources are correlated. The noise $\mathbf{e}$ is assumed circularly-symmetric complex Gaussian with covariance $p_{\mathrm{e}}\mathbf{I}$ and uncorrelated with the sources. To apply subspace methods like MUSIC [18] or root-MUSIC [8], we compute the array output covariance

$$\mathbf{R}_{\mathbf{xx}}\triangleq\mathrm{E}[\mathbf{x}\mathbf{x}^H]=\mathbf{A}\mathbf{P}\mathbf{A}^H+p_{\mathrm{e}}\mathbf{I}. \tag{2}$$

Then the DOAs ω_i can be estimated by finding the signal and noise subspaces, which are spanned by appropriate subsets of eigenvectors of $\mathbf{R}_{\mathbf{xx}}$.

In [1], a new beamspace method, called convolutional beamspace (CBS) is proposed. This method enjoys the same advantages as classical beamspace methods [4], [19]–[23], including lower computational complexity, increased parallelism of subband processing, and improved DOA resolution. Moreover, it incorporates the idea of digital filtering and preserves the Vandermonde structure of ULA outputs. This allows us to apply subspace methods like root-MUSIC [8] and ESPRIT [9] directly on the CBS output without additional processing. In comparison, classical beamspace methods do not retain the Vandermonde structure and require elaborate preparation to do root-MUSIC [6] or ESPRIT [7].

In CBS [1], the ULA output $x(n), 0\leq n\leq N-1$ is convolved with an FIR filter $H(z)=\sum_{n=0}^{L-1}h(n)z^{-n}$ to obtain the output $y(n)$, where $L<N$. Then the steady-state samples are collected in a vector

$$\begin{aligned}\mathbf{y}&\triangleq[y(L-1)\ y(L)\cdots y(N-1)]^T=\mathbf{H}\mathbf{x}\\&=\mathbf{A}_L\mathbf{d}+\mathbf{H}\mathbf{e},\end{aligned} \tag{3}$$

where

$$\mathbf{H}=\begin{bmatrix}h(L-1) & \cdots & h(0) & 0 & \cdots & 0\\ 0 & h(L-1) & \cdots & h(0) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & h(L-1) & \cdots & h(0)\end{bmatrix}$$

is a $(N-L+1)\times N$ banded Toeplitz matrix,

$$\mathbf{A}_L=[\mathbf{a}_{N-L+1}(\omega_1)\ \cdots\ \mathbf{a}_{N-L+1}(\omega_D)], \tag{4}$$

and $\mathbf{d}=\mathbf{D}_H\mathbf{c}$ with

$$\mathbf{D}_H=\mathrm{diag}(H(e^{j\omega_1})e^{j\omega_1(L-1)},\ldots,H(e^{j\omega_D})e^{j\omega_D(L-1)}). \tag{5}$$

Like the original array output $\mathbf{x}$, the CBS output $\mathbf{y}$ is represented in terms of a Vandermonde matrix, i.e., $\mathbf{A}_L$. Hence, we can compute the covariance

$$\mathbf{R}_{\mathbf{yy}}=\mathbf{A}_L\mathbf{R}_{\mathbf{dd}}\mathbf{A}_L^H+p_{\mathrm{e}}\mathbf{H}\mathbf{H}^H, \tag{6}$$

where

$$\mathbf{R}_{\mathbf{dd}}=\mathrm{E}[\mathbf{d}\mathbf{d}^H]=\mathbf{D}_H\mathbf{P}\mathbf{D}_H^H, \tag{7}$$

and estimate DOAs using root-MUSIC or ESPRIT without any further adjustment or processing to the data. Note from (5) that the source amplitudes c_i are filtered by the frequency response $H(e^{j\omega_i})$. We assume signals in the filter stopband are well attenuated, so $\mathbf{y}$ contains only those DOAs that fall in the passband of $H(e^{j\omega})$. Without loss of generality, assume $\omega_1,\ldots,\omega_{D_0}$ are in the passband. Then, $\mathbf{y}\approx\mathbf{A}_{L,0}\mathbf{d}_0+\mathbf{H}\mathbf{e}$, where $\mathbf{A}_{L,0}$ has the first D_0 columns of $\mathbf{A}_L$, and $\mathbf{d}_0$ has the first D_0 entries of $\mathbf{d}$.

Since $\mathbf{y}$ contains only passband sources, we can decimate $\mathbf{y}$ without causing ambiguity [1]. This gives us complexity reduction, which is an integral part of any beamspace method. In particular, if $H(e^{j\omega})$ has passband width $2\pi/M$ for some integer M, we can decimate $\mathbf{y}$ by M and obtain [1]

$$\mathbf{v}_l = \mathbf{D}_l\mathbf{y} = \mathbf{A}_{\text{dec}}\mathbf{d}_l + \mathbf{D}_l\mathbf{H}\mathbf{e} \approx \mathbf{A}_{\text{dec},0}\mathbf{d}_{l,0} + \mathbf{D}_l\mathbf{H}\mathbf{e} \quad (8)$$

where $\mathbf{D}_l = [\boldsymbol{\delta}_l^{(N-L+1)} \;\; \boldsymbol{\delta}_{l+M}^{(N-L+1)} \;\; \ldots \;\; \boldsymbol{\delta}_{l+(J-1)M}^{(N-L+1)}]^T$ is a decimation matrix, $\mathbf{A}_{\text{dec}} = [\mathbf{a}_J(M\omega_1) \;\cdots\; \mathbf{a}_J(M\omega_D)]$, $J = (N-L+1)/M$ (assumed an integer for simplicity),

$$\mathbf{d}_l = [c_1 e^{j(L-1+l)\omega_1} H(e^{j\omega_1}) \;\cdots\; c_D e^{j(L-1+l)\omega_D} H(e^{j\omega_D})]^T,$$

$\mathbf{A}_{\text{dec},0}$ has the first D_0 columns of $\mathbf{A}_{\text{dec}}$, and $\mathbf{d}_{l,0}$ has the first D_0 entries of $\mathbf{d}_l$. In (8), l can take values $0, \ldots, M-1$, corresponding to the M polyphase components [10] of $\mathbf{y}$. We will estimate only the D_0 passband DOAs based on $\mathbf{v}_l$. We assume $D_0 < J$ for MUSIC to identify DOAs without ambiguity [18]. To avoid wasting data, we compute the average covariance

$$\mathbf{R}_{\text{ave}} = \frac{1}{M}\sum_{l=0}^{M-1} \mathbf{R}_{v_l} = \mathbf{A}_{\text{dec}}\breve{\mathbf{R}}_d\mathbf{A}_{\text{dec}}^H + p_{\text{e}}\mathbf{G}_{\text{dec}}, \quad (9)$$

where $\breve{\mathbf{R}}_d$ is $\mathbf{R}_{d_l} = \mathrm{E}[\mathbf{d}_l\mathbf{d}_l^H]$ averaged over l (polyphase index),

$$\mathbf{R}_{v_l} = \mathrm{E}[\mathbf{v}_l\mathbf{v}_l^H] = \mathbf{A}_{\text{dec}}\mathbf{R}_{d_l}\mathbf{A}_{\text{dec}}^H + p_{\text{e}}\mathbf{G}_{\text{dec}}, \quad (10)$$

and $\mathbf{G}_{\text{dec}} \triangleq \mathbf{D}_l\mathbf{H}\mathbf{H}^H\mathbf{D}_l$ is independent of l. To aid our later analysis, we express

$$\mathbf{R}_{\text{ave}} = \mathbf{A}_{\text{dec},0}\breve{\mathbf{R}}_{d,0}\mathbf{A}_{\text{dec},0}^H + \delta\mathbf{R} + p_{\text{e}}\mathbf{G}_{\text{dec}}, \quad (11)$$

where the covariance perturbation

$$\delta\mathbf{R} = \mathbf{A}_{\text{dec}}\breve{\mathbf{R}}_d\mathbf{A}_{\text{dec}}^H - \mathbf{A}_{\text{dec},0}\breve{\mathbf{R}}_{d,0}\mathbf{A}_{\text{dec},0}^H \quad (12)$$

contains the auto-covariance of stopband sources and the cross-covariance between passband and stopband sources. This covariance perturbation is due to the filtered stopband sources, so we assume $\|\delta\mathbf{R}\|$ is small. In (12), $\breve{\mathbf{R}}_{d,0}$ is $\mathbf{R}_{d_{l,0}} = \mathrm{E}[\mathbf{d}_{l,0}\mathbf{d}_{l,0}^H]$ averaged over l. We assume that the CBS filter is a spectral factor of a Nyquist(M) filter so that $\mathbf{G}_{\text{dec}} = \mathbf{I}$ (see [1]). That is, the noise after filtering and decimation remains white. Then we compute the eigenvalue decomposition

$$\mathbf{R}_{\text{ave}} = \mathbf{E}_{\text{s}}\boldsymbol{\Lambda}_{\text{s}}\mathbf{E}_{\text{s}}^H + \mathbf{E}_{\text{n}}\boldsymbol{\Lambda}_{\text{n}}\mathbf{E}_{\text{n}}^H, \quad (13)$$

where $\mathbf{E}_{\text{s}} = [\mathbf{e}_1 \cdots \mathbf{e}_{D_0}]$ and $\mathbf{E}_{\text{n}} = [\mathbf{e}_{D_0+1} \cdots \mathbf{e}_J]$ contain the signal and noise eigenvectors respectively, and $\boldsymbol{\Lambda}_{\text{s}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_{D_0})$ and $\boldsymbol{\Lambda}_{\text{n}} = \mathrm{diag}(\lambda_{D_0+1}, \ldots, \lambda_J)$ contain the corresponding eigenvalues in descending order. Note that only the first D_0 eigenvalues are assumed dominant and correspond to signals. Then we can estimate the passband DOAs using MUSIC [18] or root-MUSIC [8]. Considering MUSIC as an example, we evaluate the MUSIC spectrum

$$P(\omega) = (\mathbf{a}_J^H(M\omega)\mathbf{E}_{\text{n}}\mathbf{E}_{\text{n}}^H\mathbf{a}_J(M\omega))^{-1} \quad (14)$$

on a dense grid of potential DOAs and identify local maxima as the estimates of $M\omega_i$ mod 2π, or equivalently $\omega_i + 2\pi s_i/M$ for some integers s_i. Since ω_i are known to be in the passband of $H(e^{j\omega})$ which has width $2\pi/M$, the ambiguities s_i can be resolved. In practice, we use a finite number, say K, of independent snapshots to estimate the covariance matrix (9). That is, we compute noise subspace estimate $\widehat{\mathbf{E}}_{\text{n}}$ based on

$$\widehat{\mathbf{R}}_{\text{ave}} = \frac{1}{KM}\sum_{l=0}^{M-1}\sum_{k=1}^{K} \mathbf{v}_l[k]\mathbf{v}_l^H[k] \quad (15)$$

and then evaluate the MUSIC spectrum (14).

III. MSE Analysis for CBS

The goal of this section is to derive the MSE performance when we use MUSIC [18] or root-MUSIC [8] to estimate DOAs based on (15). In [1], comparison of MSE performance between CBS and element-space is given mainly based on numerical examples, and theoretical analysis of CBS MSE is given only for limited simple cases in an approximated way. In the following, we present a rigorous and more accurate MSE analysis for CBS.

MSE of DOA estimates for element-space is analyzed for MUSIC in [12] and for root-MUSIC in [8], and these papers have remained the foundation for such analysis for many years. Here we extend this analysis to CBS. In fact, the probability distributions of DOA estimation errors and thus MSEs in element-space are asymptotically the same for MUSIC and root-MUSIC. We will show that these are also true for CBS. Compared to element-space analysis, there are two complications we should deal with for CBS. First, $\mathbf{v}_l[k]$ is not independent of $\mathbf{v}_m[k]$ for $l \neq m$, which makes the derivation of the distribution of the eigenvectors of (15) more difficult. Independent observations are assumed in the traditional asymptotic theory for principal component analysis [13], which is naturally applicable to element-space. We modify the method so that it can be used for CBS. Second, the presence of $\delta\mathbf{R}$ due to the filtered stopband sources is an additional source of estimation errors. We will show that the effect of the filtered stopband sources is a bias and can be analyzed separately from the variance term as we use a first-order perturbation analysis. To this end, we define $\mathbf{R}_{\mathbf{yy},0} = \mathbf{R}_{\mathbf{yy}}\big|_{\delta\mathbf{R}=\mathbf{O}}$ and

$$\mathbf{R}_{\text{ave},0} = \mathbf{R}_{\text{ave}}\big|_{\delta\mathbf{R}=\mathbf{O}} = \mathbf{E}_{\text{s},0}\boldsymbol{\Lambda}_{\text{s},0}\mathbf{E}_{\text{s},0}^H + \mathbf{E}_{\text{n},0}\boldsymbol{\Lambda}_{\text{n},0}\mathbf{E}_{\text{n},0}^H, \quad (16)$$

to be the covariance matrices when $\delta\mathbf{R}$ is set to be zero (equivalently, when the stopband sources are nulled), where $\mathbf{E}_{\text{s},0} = [\mathbf{e}_{1,0} \cdots \mathbf{e}_{D_0,0}]$ and $\mathbf{E}_{\text{n},0} = [\mathbf{e}_{D_0+1,0} \cdots \mathbf{e}_{J,0}]$ contain the signal and noise eigenvectors respectively, and $\boldsymbol{\Lambda}_{\text{s},0} = \mathrm{diag}(\lambda_{1,0}, \ldots, \lambda_{D_0,0})$ and $\boldsymbol{\Lambda}_{\text{n},0} = \mathrm{diag}(\lambda_{D_0+1,0}, \ldots, \lambda_{J,0}) = p_{\text{e}}\mathbf{I}$ contain the corresponding eigenvalues in the diagonals. Here we only have D_0 signal eigenvectors because the $D - D_0$ stopband sources are null. According to the theory of perturbation of Hermitian matrices [24], [25], the signal eigenvectors of $\mathbf{R}_{\text{ave}}$ are

$$\mathbf{e}_l = \mathbf{e}_{l,0} + \sum_{\substack{r=1\\ r\neq l}}^{J} \frac{\mathbf{e}_{r,0}^H\delta\mathbf{R}\mathbf{e}_{l,0}}{\lambda_{l,0} - \lambda_{r,0}}\mathbf{e}_{r,0}, \quad l = 1, \ldots, D_0 \quad (17)$$

if $\|\delta\mathbf{R}\|$ is small compared to the norm of the first term in (11). Here we assume the signal eigenvalues are distinct so that the denominators $\lambda_{l,0} - \lambda_{r,0} \neq 0$. This is true with probability one if the DOAs are uniformly randomly distributed.

In the following, we first present a lemma for the distribution of the signal eigenvectors $\hat{\mathbf{e}}_{l,0}$ of the K-snapshot estimate $\widehat{\mathbf{R}}_{\text{ave},0}$ when we null the stopband sources. This can be viewed as a generalized version of its element-space counterpart, Lemma 3.1 in [12].

Lemma 1: The signal eigenvectors $\hat{\mathbf{e}}_{l,0}$ of $\widehat{\mathbf{R}}_{\text{ave},0}$ are asymptotically (for large K) jointly complex Gaussian with means $\mathbf{e}_{l,0}$, covariances

$$\begin{aligned}&\mathrm{E}[(\hat{\mathbf{e}}_{l,0} - \mathbf{e}_{l,0})(\hat{\mathbf{e}}_{r,0} - \mathbf{e}_{r,0})^H] \\ &\quad = \sum_{\substack{i=1\\ i\neq l}}^{J}\sum_{\substack{k=1\\ k\neq r}}^{J} \frac{\mathrm{tr}\left(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(r,l)}(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(k,i)})^H\right)}{KM^2(\lambda_{l,0} - \lambda_{i,0})(\lambda_{r,0} - \lambda_{k,0})}\mathbf{e}_{i,0}\mathbf{e}_{k,0}^H \quad (18)\end{aligned}$$

and relation matrices

$$\begin{aligned}&\mathrm{E}[(\hat{\mathbf{e}}_{l,0}-\mathbf{e}_{l,0})(\hat{\mathbf{e}}_{r,0}-\mathbf{e}_{r,0})^T]\\&=\sum_{\substack{i=1\\i\neq l}}^{J}\sum_{\substack{k=1\\k\neq r}}^{J}\frac{\operatorname{tr}\left(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(k,l)}(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(r,i)})^H\right)}{KM^2(\lambda_{l,0}-\lambda_{i,0})(\lambda_{r,0}-\lambda_{k,0})}\mathbf{e}_{i,0}\mathbf{e}_{k,0}^T\end{aligned}\quad (19)$$

for $1\leq l,r\leq D_0$, where $\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(p,q)}\in\mathbb{C}^{M\times M}$ for $1\leq p,q\leq J$ are submatrices of

$$\widetilde{\mathbf{R}}_{\mathbf{yy}}=(\mathbf{E}_0^H\otimes\mathbf{I}_M)\mathbf{R}_{\mathbf{yy},0}(\mathbf{E}_0\otimes\mathbf{I}_M)\quad (20)$$

$$\triangleq\begin{bmatrix}\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(1,1)}&\cdots&\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(1,J)}\\\vdots&\ddots&\vdots\\\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(J,1)}&\cdots&\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(J,J)}\end{bmatrix}.\quad (21)$$

Here $\mathbf{E}_0=[\mathbf{E}_{\mathrm{s},0}\ \mathbf{E}_{\mathrm{n},0}]$ contains the eigenvectors in (16).

Proof: See Appendix A. ∎

Note that although the signals and noise are circularly-symmetric complex Gaussian, the eigenvector estimates $\hat{\mathbf{e}}_{l,0}$ are not in general circularly symmetric since the relation matrices can be nonzero. Our expressions of covariance and relation matrices reduce to those in [12] if we set $L=1$, the CBS filter $H(z)=1$, the decimation ratio $M=1$, and $B_i=0$ (since all sources are in the "passband"). In this case, considering (16), we can simplify (20) as

$$\widetilde{\mathbf{R}}_{\mathbf{yy}}=\mathbf{E}_0^H\mathbf{R}_{\mathbf{yy},0}\mathbf{E}_0=\mathbf{E}_0^H\mathbf{R}_{\mathrm{ave},0}\mathbf{E}_0=\mathbf{\Lambda}_0,\quad (22)$$

where $\mathbf{\Lambda}_0=\operatorname{diag}(\mathbf{\Lambda}_{\mathrm{s},0},\mathbf{\Lambda}_{\mathrm{n},0})$. Hence, we obtain the covariances

$$\begin{aligned}&\mathrm{E}[(\hat{\mathbf{e}}_{l,0}-\mathbf{e}_{l,0})(\hat{\mathbf{e}}_{r,0}-\mathbf{e}_{r,0})^H]\\&=\sum_{\substack{i=1\\i\neq l}}^{J}\frac{\lambda_{l,0}\lambda_{i,0}}{K(\lambda_{l,0}-\lambda_{i,0})^2}\mathbf{e}_{i,0}\mathbf{e}_{i,0}^H\cdot\delta_{lr}\end{aligned}\quad (23)$$

and relation matrices

$$\begin{aligned}&\mathrm{E}[(\hat{\mathbf{e}}_{l,0}-\mathbf{e}_{l,0})(\hat{\mathbf{e}}_{r,0}-\mathbf{e}_{r,0})^T]\\&=\frac{-\lambda_{l,0}\lambda_{r,0}}{K(\lambda_{l,0}-\lambda_{r,0})^2}\mathbf{e}_{r,0}\mathbf{e}_{l,0}^T(1-\delta_{lr}).\end{aligned}\quad (24)$$

These element-space expressions are much simpler than the CBS ones because $\widetilde{\mathbf{R}}_{\mathbf{yy}}$ is diagonal for element-space. Also, as mentioned earlier, CBS eigenvector estimates $\hat{\mathbf{e}}_{l,0}$ are generally not circularly-symmetric complex Gaussian. However, for element-space, they are circularly-symmetric complex Gaussian since $\mathrm{E}[(\hat{\mathbf{e}}_{l,0}-\mathbf{e}_{l,0})(\hat{\mathbf{e}}_{l,0}-\mathbf{e}_{l,0})^T]=0$.

Next, using Lemma 1 and (17), we can derive the distribution of the DOA estimation errors $\hat{\omega}_i-\omega_i$ as follows.

Theorem 1: The CBS DOA estimation errors $\hat{\omega}_i-\omega_i$ for passband sources, when either MUSIC or root-MUSIC is used with (15), are asymptotically (for large K) jointly Gaussian distributed with means B_i and cross-correlations

$$\begin{aligned}&\mathrm{E}[(\hat{\omega}_i-\omega_i)(\hat{\omega}_k-\omega_k)]=B_iB_k+\frac{1}{2KM^2g(\tilde{\omega}_i)g(\tilde{\omega}_k)}\mathrm{Re}\Bigg\{\\&\sum_{l=1}^{D_0}\sum_{r=1}^{D_0}\mathbf{e}_{l,0}^H\mathbf{a}_J(\tilde{\omega}_i)\mathbf{a}_J^H(\tilde{\omega}_k)\mathbf{e}_{r,0}\dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_{\mathrm{n},0}\mathbf{B}_{lr}\mathbf{E}_{\mathrm{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_k)\\&+\sum_{l=1}^{D_0}\sum_{r=1}^{D_0}\mathbf{e}_{l,0}^H\mathbf{a}_J(\tilde{\omega}_i)\mathbf{a}_J^T(\tilde{\omega}_k)\mathbf{e}_{r,0}^*\dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_{\mathrm{n},0}\mathbf{C}_{lr}\mathbf{E}_{\mathrm{n},0}^T\dot{\mathbf{a}}_J^*(\tilde{\omega}_k)\Bigg\}\end{aligned}$$

for $1\leq i,k\leq D_0$, where $\tilde{\omega}_i=M\omega_i$,

$$B_i=\frac{\mathrm{Re}\left\{\mathbf{a}_J^H(\tilde{\omega}_i)\sum_{l=1}^{D_0}\left(\mathbf{e}_{l,0}\delta\mathbf{e}_l^H+\delta\mathbf{e}_l\mathbf{e}_{l,0}^H\right)\dot{\mathbf{a}}_J(\tilde{\omega}_i)\right\}}{Mg(\tilde{\omega}_i)},\quad (25)$$

$$\delta\mathbf{e}_l=\sum_{\substack{r=1\\r\neq l}}^{J}\frac{\mathbf{e}_{r,0}^H\delta\mathbf{R}\mathbf{e}_{l,0}}{\lambda_{l,0}-\lambda_{r,0}}\mathbf{e}_{r,0},\quad l=1,\ldots,D_0,\quad (26)$$

$\dot{\mathbf{a}}_J(\omega)=\frac{\mathrm{d}}{\mathrm{d}\omega}\mathbf{a}_J(\omega)$, $g(\omega)=\dot{\mathbf{a}}_J^H(\omega)\mathbf{E}_{\mathrm{n},0}\mathbf{E}_{\mathrm{n},0}^H\dot{\mathbf{a}}_J(\omega)$, $\mathbf{B}_{lr}$ and $\mathbf{C}_{lr}$ are $(J-D_0)\times(J-D_0)$ matrices with entries

$$[\mathbf{B}_{lr}]_{m,n}=\frac{\operatorname{tr}\left(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(r,l)}(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(D_0+n,D_0+m)})^H\right)}{M^2(\lambda_{l,0}-p_{\mathrm{e}})(\lambda_{r,0}-p_{\mathrm{e}})}\quad (27)$$

$$[\mathbf{C}_{lr}]_{m,n}=\frac{\operatorname{tr}\left(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(D_0+n,l)}(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(r,D_0+m)})^H\right)}{M^2(\lambda_{l,0}-p_{\mathrm{e}})(\lambda_{r,0}-p_{\mathrm{e}})}\quad (28)$$

for $1\leq m,n\leq J-D_0$, $1\leq l,r\leq D_0$, and $\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(p,q)}$, $1\leq p,q\leq J$ are as defined in (21).

Proof: See Appendix B. ∎

In particular, the MSEs $\mathrm{E}[(\hat{\omega}_i-\omega_i)^2]$ of the passband DOAs can be obtained by letting $k=i$ in Theorem 1:

$$\mathrm{E}[(\hat{\omega}_i-\omega_i)^2]=B_i^2+V_i\quad (29)$$

where

$$\begin{aligned}&V_i=\frac{1}{2KM^2g^2(\tilde{\omega}_i)}\mathrm{Re}\Bigg\{\\&\sum_{l=1}^{D_0}\sum_{r=1}^{D_0}\mathbf{e}_{l,0}^H\mathbf{a}_J(\tilde{\omega}_i)\mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{e}_{r,0}\dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_{\mathrm{n},0}\mathbf{B}_{lr}\mathbf{E}_{\mathrm{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)\\&+\sum_{l=1}^{D_0}\sum_{r=1}^{D_0}\mathbf{e}_{l,0}^H\mathbf{a}_J(\tilde{\omega}_i)\mathbf{a}_J^T(\tilde{\omega}_i)\mathbf{e}_{r,0}^*\dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_{\mathrm{n},0}\mathbf{C}_{lr}\mathbf{E}_{\mathrm{n},0}^T\dot{\mathbf{a}}_J^*(\tilde{\omega}_i)\Bigg\}\end{aligned}$$

for $i=1,\ldots,D_0$. As mentioned earlier, since $\mathbf{v}_l[k]$ is not independent of $\mathbf{v}_m[k]$ for $l\neq m$ in (15), it is more difficult to analyze CBS MSE. One can see such complications through the $M\times M$ matrices $\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(p,q)}$ appearing in (27) and (28). On the other hand, the presence of the filtered stopband sources simply results in a bias B_i in the DOA estimates. Thus, the MSE is given in the form of a bias-variance decomposition. As we shall see in the simulations (Fig. 4), the bias term is typically much smaller than the variance term if the CBS filter is properly designed with good stopband attenuation and if the stopband sources are not too powerful. Hence, we may ignore the bias term in many practical cases. Due to the complicated MSE expression, it is not easy to get further insight. However, we will present more insight based on the CRB expressions in Sec. IV.

Our MSE expression can be viewed as a generalized version of the element-space MSE expression in [12]. Our expression reduces to that in [12] if we set $L=1$, the CBS filter $H(z)=1$, the decimation ratio $M=1$, and $B_i=0$. Then again we have $\widetilde{\mathbf{R}}_{\mathbf{yy}}=\mathbf{\Lambda}_0$ as in (22) and thus

$$\mathbf{B}_{lr}=\frac{\lambda_{l,0}p_{\mathrm{e}}\mathbf{I}}{(\lambda_{l,0}-p_{\mathrm{e}})^2}\delta_{lr}\quad (30)$$

and $\mathbf{C}_{lr}=\mathbf{O}$ for all l,r. This is why we can get the much more simplified expression for element-space MSE

$$\mathrm{E}[(\hat{\omega}_i-\omega_i)^2]=\frac{1}{2Kg(\omega_i)}\sum_{l=1}^{D}\frac{\lambda_{l,0}p_{\mathrm{e}}}{(\lambda_{l,0}-p_{\mathrm{e}})^2}|\mathbf{e}_{l,0}^H\mathbf{a}_J(\omega_i)|^2\quad (31)$$

given in [12]. Although it is known that CBS CRB cannot be smaller than element-space CRB [1], [26], it is still unclear whether CBS MSE can be smaller than element-space MSE for uncorrelated sources because of the complicated MSE expression for CBS. For correlated sources, CBS MSE can be much smaller than element-space MSE (see Fig. 5(b)).

IV. Cramer–Rao Bound for CBS

In this section, we derive the Cramer–Rao Bound (CRB) [14] for the DOA estimates based on the CBS output $\mathbf{y}$ in (3), under the stochastic model [15], [17]. The CRB offers a lower bound on the variances of unbiased estimates of parameters. However, as shown in Sec. III, CBS yields biased DOA estimates if there are stopband sources. Hence, some extra care will be taken in order to derive the CRB for biased estimators [14]. We shall derive two forms of CRB expressions. Form 1 is in the same style as the CRB in [17]. It is derived from inverting the Fisher information matrix [14], and a necessary and sufficient condition for the existence of the CRB is naturally obtained in this process. Although it is not the focus of this paper, Form 1 also yields a CRB even when the noise is non-white with a singular covariance. Meanwhile, Form 2 is in the same style as the CRB in [15] and offers some additional insight as shown later.

The probability model for stochastic CRB [15], [17] based on K snapshots of CBS outputs is

$$\begin{bmatrix}\mathbf{y}[1]\\ \mathbf{y}[2]\\ \vdots\\ \mathbf{y}[K]\end{bmatrix} \sim \mathcal{CN}\left(\mathbf{0}, \begin{bmatrix}\mathbf{R_{yy}} & \mathbf{O} & \cdots & \mathbf{O}\\ \mathbf{O} & \mathbf{R_{yy}} & \cdots & \mathbf{O}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{R_{yy}}\end{bmatrix}\right), \quad (32)$$

where $\mathbf{R_{yy}}$ is as in (6). The parameter vector for this probability model is

$$\boldsymbol{\alpha} = \left[[\omega_i]_{i=1}^D \;\; [p_i]_{i=1}^D \;\; [P_{ik}^{(\mathrm{r})}]_{i>k} \;\; [P_{ik}^{(\mathrm{i})}]_{i>k} \;\; p_\mathrm{e}\right]^T, \quad (33)$$

where p_i is the ith diagonal entry of $\mathbf{P}$, and $P_{ik}^{(\mathrm{r})}$ and $P_{ik}^{(\mathrm{i})}$ are the real and imaginary parts of the (i,k)-entry of $\mathbf{P}$, respectively. Thus, there are $D + D^2 + 1$ real parameters, among which only the D parameters $\boldsymbol{\omega} = [\omega_1 \;\cdots\; \omega_D]^T$ are of interest.

For any unbiased estimator $\hat{\boldsymbol{\alpha}}$ with $\mathrm{E}[\hat{\boldsymbol{\alpha}}] = \boldsymbol{\alpha}$, the CRB is given by [14]

$$\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha}) = [\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})]^{-1}, \quad (34)$$

where $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$ is the Fisher information matrix for the model, such that the covariance of the estimator

$$\mathrm{cov}(\hat{\boldsymbol{\alpha}}) \succeq \mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha}). \quad (35)$$

As shown in Sec. III, however, CBS yields biased estimates $\hat{\boldsymbol{\omega}}_0$ for the in-band (passband) DOAs $\boldsymbol{\omega}_0 = [\omega_1 \;\cdots\; \omega_{D_0}]^T$ if there are out-of-band (stopband) sources. A modified form of CRB for a biased estimator of a scalar parameter is given in [16]. To derive a bound for the biased vector estimator $\hat{\boldsymbol{\omega}}_0$, we consider its expectation $\mathrm{E}[\hat{\boldsymbol{\omega}}_0]$. It must be some function of the model parameters $\boldsymbol{\alpha}$, say, $\mathrm{E}[\hat{\boldsymbol{\omega}}_0] = \boldsymbol{\psi}(\boldsymbol{\alpha})$. Then $\hat{\boldsymbol{\omega}}_0$ can be viewed as an unbiased estimator for the transformation $\boldsymbol{\psi}(\cdot)$ of the parameter vector $\boldsymbol{\alpha}$. Hence, assuming $\boldsymbol{\psi}(\boldsymbol{\alpha})$ is differentiable, we can use the CRB for transformations [14] to obtain

$$\mathrm{cov}(\hat{\boldsymbol{\omega}}_0) \succeq \frac{\partial \boldsymbol{\psi}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha}) \left(\frac{\partial \boldsymbol{\psi}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\right)^T, \quad (36)$$

where $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha})$ is defined in (34), and $\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{\alpha}}$ is the Jacobian matrix with entries

$$\left[\frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{\alpha}}\right]_{i,k} = \frac{\partial [\boldsymbol{\psi}]_i}{\partial [\boldsymbol{\alpha}]_k} = \begin{cases} 1 + \frac{\partial B_i}{\partial [\boldsymbol{\alpha}]_i}, & i = k \\ \frac{\partial B_i}{\partial [\boldsymbol{\alpha}]_k}, & i \neq k \end{cases}. \quad (37)$$

Here B_i are the biases defined in (25). Although the derivatives $\frac{\partial B_i}{\partial [\boldsymbol{\alpha}]_k}$ can be analytically computed from (25), the results will be lengthy due to the complicated dependence of $\mathbf{e}_{l,0}$ and $\lambda_{l,0}$ on $[\boldsymbol{\alpha}]_k$. Thus, we omit the tedious derivations, which may not give much insight. In this paper, we only compute the derivatives numerically in the example of Fig. 4. Moreover, these derivatives $\frac{\partial B_i}{\partial [\boldsymbol{\alpha}]_k}$ are typically small since the covariance perturbation $\delta \mathbf{R}$ due to the filtered stopband sources is small. Thus,

$$\frac{\partial \boldsymbol{\psi}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \approx \begin{bmatrix}\mathbf{I} & \mathbf{O}_{D_0, D-D_0+D^2+1}\end{bmatrix} \quad (38)$$

so that we have

$$\mathrm{cov}(\hat{\boldsymbol{\omega}}_0) \succeq \frac{\partial \boldsymbol{\psi}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha}) \left(\frac{\partial \boldsymbol{\psi}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\right)^T \quad (39)$$

$$\approx \mathrm{CRB}_\mathrm{unb}(\boldsymbol{\omega}_0), \quad (40)$$

where $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\omega}_0)$ is the top left $D_0 \times D_0$ block of $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha})$. Note that (40) is precisely the CRB for CBS if there are no stopband sources so that $\hat{\boldsymbol{\omega}}_0$ is unbiased. If stopband sources exist, as we shall see in simulations (Fig. 4), usually (40) also gives a good lower bound. It does not make a big difference to consider (39) unless the out-of-band source power is extremely large. In the following, we focus on the $\boldsymbol{\omega}$-block $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\omega})$, i.e., the top left $D \times D$ block of $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\alpha})$ because it leads to more elegant and insightful expressions. Once we have $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\omega})$, we can immediately obtain $\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\omega}_0)$ as a submatrix. Then in particular, we have that the variance

$$\mathrm{var}(\hat{\omega}_i) \geq [\mathrm{CRB}_\mathrm{unb}(\boldsymbol{\omega}_0)]_{i,i} \quad (41)$$

for each in-band DOA estimate $\hat{\omega}_i$. Now we are ready to derive Form 1 of the CRB by investigating the Fisher information matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$.

A. Form 1

The (i,k)-entry of the Fisher information matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$ for the model (32) can be derived as [17]

$$[\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})]_{i,k} = K \frac{\partial \mathbf{r}_\mathbf{yy}^H}{\partial [\boldsymbol{\alpha}]_i} \left(\mathbf{R}_\mathbf{yy}^T \otimes \mathbf{R_{yy}}\right)^{-1} \frac{\partial \mathbf{r_{yy}}}{\partial [\boldsymbol{\alpha}]_k}, \quad (42)$$

where $\mathbf{r_{yy}} = \mathrm{vec}(\mathbf{R_{yy}})$. Separating the parameters of interest from the other parameters, we have

$$\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha}) = K \begin{bmatrix} \mathbf{G}^H\mathbf{G} & \mathbf{G}^H\boldsymbol{\Delta}_1 \\ \boldsymbol{\Delta}_1^H\mathbf{G} & \boldsymbol{\Delta}_1^H\boldsymbol{\Delta}_1 \end{bmatrix}, \quad (43)$$

where

$$\mathbf{G} = \left(\mathbf{R}_\mathbf{yy}^T \otimes \mathbf{R_{yy}}\right)^{-\frac{1}{2}} \left[\frac{\partial \mathbf{r_{yy}}}{\partial [\boldsymbol{\alpha}]_1} \;\cdots\; \frac{\partial \mathbf{r_{yy}}}{\partial [\boldsymbol{\alpha}]_D}\right], \quad (44)$$

$$\boldsymbol{\Delta}_1 = \left(\mathbf{R}_\mathbf{yy}^T \otimes \mathbf{R_{yy}}\right)^{-\frac{1}{2}} \left[\frac{\partial \mathbf{r_{yy}}}{\partial [\boldsymbol{\alpha}]_{D+1}} \;\cdots\; \frac{\partial \mathbf{r_{yy}}}{\partial [\boldsymbol{\alpha}]_{D+D^2+1}}\right]. \quad (45)$$

Here we assume $\mathbf{R_{yy}}$ is positive definite so that $(\mathbf{R}_\mathbf{yy}^T \otimes \mathbf{R_{yy}})$ is also positive definite. Then $(\mathbf{R}_\mathbf{yy}^T \otimes \mathbf{R_{yy}})^{-1/2}$ denotes the inverse of its positive definite square root. For the CRB (34) to exist, we require $\mathcal{I}(\boldsymbol{\alpha})$ to be invertible, so we obtain the following theorem.

Theorem 2: The CRB (34) exists if and only if

$$\mathbf{M} = \left[\mathbf{M}_1 \quad (\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H) \quad \mathrm{vec}(\mathbf{H}\mathbf{H}^H)\right] \tag{46}$$

has full column rank, where

$$\mathbf{M}_1 = \dot{\mathbf{A}}_L^* \odot (\mathbf{A}_L\mathbf{D}_H\mathbf{P}\mathbf{D}_H^*) + (\mathbf{A}_L\mathbf{D}_H\mathbf{P}\mathbf{D}_H^*)^* \odot \dot{\mathbf{A}}_L. \tag{47}$$

Proof: See Appendix C. ■

Theorem 2 offers a necessary and sufficient condition for the CRB to exist. Since $\mathbf{M}$ has $(N-L+1)^2$ rows and $D+D^2+1$ columns, it can have full column rank only if $D < N - L + 1$. In [17], a similar condition for the existence of the element-space CRB is given assuming sources are known to be uncorrelated. In the following, we assume that $\mathbf{M}$ indeed has full column rank. Now using the ideas of block matrix inversion and Schur complements [17], we can show that (34) and (43) imply that

$$\mathrm{CRB}_{\mathrm{unb}}(\boldsymbol{\omega}) = \frac{1}{K}\left(\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}_1}^{\perp}\mathbf{G}\right)^{-1}, \tag{48}$$

where $\mathbf{\Pi}_{\mathbf{\Delta}_1}^{\perp} = \mathbf{I} - \mathbf{\Delta}_1(\mathbf{\Delta}_1^H\mathbf{\Delta}_1)^{-1}\mathbf{\Delta}_1^H$ denotes the orthogonal projection onto the null space of $\mathbf{\Delta}_1^H$. Computing the above derivatives and simplifying the results, we can obtain the CRB as follows.

Theorem 3: If there are no stopband sources, the CRB for the DOAs $\boldsymbol{\omega}$ for CBS is

$$\mathrm{CRB}_{\mathrm{unb}}(\boldsymbol{\omega}) = \frac{1}{K}\left(\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}}^{\perp}\mathbf{G}\right)^{-1}, \tag{49}$$

where

$$\begin{aligned}\mathbf{G} = \left(\mathbf{R}_{\mathbf{yy}}^T \otimes \mathbf{R}_{\mathbf{yy}}\right)^{-\frac{1}{2}} \big(&\dot{\mathbf{A}}_L^* \odot (\mathbf{A}_L\mathbf{D}_H\mathbf{P}\mathbf{D}_H^*) \\ &+ (\mathbf{A}_L\mathbf{D}_H\mathbf{P}\mathbf{D}_H^*)^* \odot \dot{\mathbf{A}}_L\big), \end{aligned}\tag{50}$$

$$\begin{aligned}\mathbf{\Delta} = &\left(\mathbf{R}_{\mathbf{yy}}^T \otimes \mathbf{R}_{\mathbf{yy}}\right)^{-\frac{1}{2}} \\ &\cdot \left[(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H) \quad \mathrm{vec}(\mathbf{H}\mathbf{H}^H)\right]. \end{aligned}\tag{51}$$

Here $\mathbf{A}_L$ and $\mathbf{D}_H$ are as defined in (4) and (5), and

$$\dot{\mathbf{A}}_L = [\dot{\mathbf{a}}_{N-L+1}(\omega_1) \ \cdots \ \dot{\mathbf{a}}_{N-L+1}(\omega_D)] \tag{52}$$

with $\dot{\mathbf{a}}_{N-L+1}(\omega) = \frac{\mathrm{d}}{\mathrm{d}\omega}\mathbf{a}_{N-L+1}(\omega)$. If stopband sources exist, (49) is an approximate bound.

Remark 1) Suppose there are stopband sources so that the CBS DOA estimator is biased. Then the exact CRB (39) can be computed from (34) and (43) with $\mathbf{G}$ as in (50) and $\mathbf{\Delta}_1$ as in (137). In this paper, we only numerically compute the derivatives in the Jacobian matrix in the example of Fig. 4.

Remark 2) By slightly modifying the proof of Theorem 3, one can verify that if the noise covariance is $p_{\mathrm{e}}\mathbf{R}_{\mathrm{e}}$ instead of $p_{\mathrm{e}}\mathbf{I}$ for any $\mathbf{R}_{\mathrm{e}}$ known a priori, then the CRB is as in (49)–(51) except that $\mathbf{H}\mathbf{H}^H$ is replaced by $\mathbf{H}\mathbf{R}_{\mathrm{e}}\mathbf{H}^H$ in (51). It is valid even when $\mathbf{R}_{\mathrm{e}}$ is singular. This result for a singular noise covariance cannot be obtained using Form 2 in Sec. IV-B.

Proof of Theorem 3: See Appendix C. ■

The CRB in Theorem 3 serves as a good reference for determining how well CBS performs in practice. In Sec. V, we will show that the MSE performance of CBS is close to the CRB in many cases. Besides, since element-space can be viewed as a special case of CBS, we can obtain the following corollary.

Corollary 1: The CRB for the DOAs $\boldsymbol{\omega}$ for element-space is

$$\mathrm{CRB}_{\mathrm{elm}}(\boldsymbol{\omega}) = \frac{1}{K}\left(\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}}^{\perp}\mathbf{G}\right)^{-1}, \tag{53}$$

where

$$\mathbf{G} = \left(\mathbf{R}_{\mathbf{xx}}^T \otimes \mathbf{R}_{\mathbf{xx}}\right)^{-\frac{1}{2}} \left(\dot{\mathbf{A}}^* \odot (\mathbf{A}\mathbf{P}) + (\mathbf{A}\mathbf{P})^* \odot \dot{\mathbf{A}}\right), \tag{54}$$

$$\mathbf{\Delta} = \left(\mathbf{R}_{\mathbf{xx}}^T \otimes \mathbf{R}_{\mathbf{xx}}\right)^{-\frac{1}{2}} \left[\mathbf{A}^* \otimes \mathbf{A} \quad \mathrm{vec}(\mathbf{I})\right]. \tag{55}$$

Here

$$\dot{\mathbf{A}} = [\dot{\mathbf{a}}_N(\omega_1) \ \cdots \ \dot{\mathbf{a}}_N(\omega_D)] \tag{56}$$

with $\dot{\mathbf{a}}_N(\omega) = \frac{\mathrm{d}}{\mathrm{d}\omega}\mathbf{a}_N(\omega)$.

Remark: Similar to Remark 2 for Theorem 3, if the noise covariance is $p_{\mathrm{e}}\mathbf{R}_{\mathrm{e}}$ instead of $p_{\mathrm{e}}\mathbf{I}$ for any $\mathbf{R}_{\mathrm{e}}$ known a priori, then the CRB is as in (53)–(55) except that $\mathbf{I}$ is replaced by $\mathbf{R}_{\mathrm{e}}$ in (55). It is valid even when $\mathbf{R}_{\mathrm{e}}$ is singular.

Proof of Corollary 1: This corollary is obtained by setting the filter length $L = 1$ and CBS filter $H(z) = 1$ in Theorem 3. ■

Note that although Theorem 3 for CBS applies only to ULAs, Corollary 1 applies to any linear arrays. One can verify that our proof under the special case of element-space is valid for sparse arrays. It is valid even when the sources are correlated.

In [17], [27], different expressions of element-space CRBs are also derived, but uncorrelated sources are assumed therein. It is important to consider correlated sources because CBS is especially advantageous over element-space in this case. Yet, to compare with previous works, we also consider uncorrelated sources in the following. When the sources are known to be uncorrelated a priori, i.e., $\mathbf{P} = \mathrm{diag}(p_1, \ldots, p_D)$, the parameter vector becomes

$$\boldsymbol{\alpha} = \left[[\omega_i]_{i=1}^D \ \ [p_i]_{i=1}^D \ \ p_{\mathrm{e}}\right]^T \tag{57}$$

instead of (33), and the CRB for CBS can be derived as follows.

Theorem 4: Suppose the sources are known to be uncorrelated a priori. Then if there are no stopband sources, the CRB for the DOAs $\boldsymbol{\omega}$ for CBS is

$$\mathrm{CRB}_{\mathrm{unb}}^{\mathrm{unc}}(\boldsymbol{\omega}) = \frac{1}{K}\left(\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}_{\mathrm{unc}}}^{\perp}\mathbf{G}\right)^{-1}, \tag{58}$$

where $\mathbf{G}$ is as in (50) and

$$\begin{aligned}\mathbf{\Delta}_{\mathrm{unc}} = &\left(\mathbf{R}_{\mathbf{yy}}^T \otimes \mathbf{R}_{\mathbf{yy}}\right)^{-\frac{1}{2}} \\ &\cdot \left[(\mathbf{A}_L\mathbf{D}_H)^* \odot (\mathbf{A}_L\mathbf{D}_H) \quad \mathrm{vec}(\mathbf{H}\mathbf{H}^H)\right]. \end{aligned}\tag{59}$$

The CRB (58) exists if and only if

$$\mathbf{M}_{\mathrm{unc}} = \left[\mathbf{M}_1 \quad (\mathbf{A}_L\mathbf{D}_H)^* \odot (\mathbf{A}_L\mathbf{D}_H) \quad \mathrm{vec}(\mathbf{H}\mathbf{H}^H)\right] \tag{60}$$

has full column rank, where $\mathbf{M}_1$ is as in (47). If stopband sources exist, (58) is an approximate bound.

Proof: The proof is similar to that of Theorem 2 and Theorem 3. ■

Again, element-space can be viewed as a special case of CBS, so we obtain the following corollary.

Corollary 2: When the sources are known to be uncorrelated a priori, the CRB for the DOAs $\boldsymbol{\omega}$ for element-space is

$$\mathrm{CRB}_{\mathrm{elm}}^{\mathrm{unc}}(\boldsymbol{\omega}) = \frac{1}{K}\left(\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}_{\mathrm{unc}}}^{\perp}\mathbf{G}\right)^{-1}, \tag{61}$$

where $\mathbf{G}$ is as in (54) and

$$\mathbf{\Delta}_{\rm unc} = \left(\mathbf{R}_{\mathbf{xx}}^T \otimes \mathbf{R}_{\mathbf{xx}}\right)^{-\frac{1}{2}} \left[\mathbf{A}^* \odot \mathbf{A} \quad \mathrm{vec}(\mathbf{I})\right]. \tag{62}$$

Proof: This corollary is obtained by setting the filter length $L = 1$ and CBS filter $H(z) = 1$ in Theorem 4. ■

One can check that (61) is equivalent to the CRB expressions in [17], [27]. Besides, comparing Theorem 4 to Theorem 3 and Corollary 2 to Corollary 1, we observe that the only difference in the CRB expressions is the substitution of a Khatri–Rao product for a Kronecker product in the expression for $\mathbf{\Delta}$. Therefore, $\mathbf{\Delta}_{\rm unc}$ is a submatrix of $\mathbf{\Delta}$, obtained by selecting proper columns of $\mathbf{\Delta}$. This allows us to formally establish the following result. This essentially follows the intuition that additional prior knowledge can only decrease the CRB.

Theorem 5: Suppose the sources are uncorrelated. Then the CRB for the DOAs $\boldsymbol{\omega}$ for CBS when the information of the uncorrelatedness of the sources is unknown a priori is not smaller than that when this information is known a priori:

$$\mathrm{CRB}_{\rm unb}(\boldsymbol{\omega})\big|_{\mathbf{P}=\mathrm{diag}(p_1,\ldots,p_D)} \succeq \mathrm{CRB}_{\rm unb}^{\rm unc}(\boldsymbol{\omega}). \tag{63}$$

Similarly, it is true for element-space:

$$\mathrm{CRB}_{\rm elm}(\boldsymbol{\omega})\big|_{\mathbf{P}=\mathrm{diag}(p_1,\ldots,p_D)} \succeq \mathrm{CRB}_{\rm elm}^{\rm unc}(\boldsymbol{\omega}). \tag{64}$$

Proof: Since $\mathbf{\Delta}_{\rm unc}$ is a submatrix of $\mathbf{\Delta}$, obtained by selecting proper columns of $\mathbf{\Delta}$ (for both CBS and element-space), we have $\mathbf{\Pi}_{\mathbf{\Delta}}^{\perp} \preceq \mathbf{\Pi}_{\mathbf{\Delta}_{\rm unc}}^{\perp}$. Hence, $(\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}}^{\perp}\mathbf{G})^{-1} \succeq (\mathbf{G}^H\mathbf{\Pi}_{\mathbf{\Delta}_{\rm unc}}^{\perp}\mathbf{G})^{-1}$, which completes the proof. ■

Theorem 5 shows that the CRB cannot be larger if more prior information is given, which is not surprising. We will verify this theorem by simulations (Fig. 6).

B. Form 2

Now we get back to our correlated model with the parameter vector (33). In [15], an alternative form of element-space stochastic CRB different from Corollary 1 is derived for any linear array under two added assumptions. First, assume the manifold matrix $\mathbf{A}$ has full column rank. Second, assume the number of sources $D < N$, the number of sensors. In the case of a ULA, since $\mathbf{A}$ is Vandermonde, the second assumption implies the first one. Hence, for CBS CRB based on the filter output $\mathbf{y}$, if we assume $D < N - L + 1$, then we can derive a second form of CRB using results in [15]. Note that this same assumption is required for Form 1 to be valid. In the following, we assume this inequality is satisfied. Since the model in [15] only applies to white noise, a noise-whitening transformation is used to obtain the following theorem.

Theorem 6: If there are no stopband sources, the CRB for the DOAs $\boldsymbol{\omega}$ for CBS is

$$\mathrm{CRB}_{\rm unb}(\boldsymbol{\omega}) = \frac{p_{\rm e}}{2K}\left[\mathrm{Re}\{\mathbf{S}_1 \circ \mathbf{S}_2^*\}\right]^{-1}, \tag{65}$$

where

$$\mathbf{S}_1 = \mathbf{P}\mathbf{A}^H\mathbf{H}^H\mathbf{R}_{\mathbf{yy}}^{-1}\mathbf{H}\mathbf{A}\mathbf{P}, \tag{66}$$

$$\mathbf{S}_2 = \dot{\mathbf{A}}^H\mathbf{W}^H\mathbf{\Pi}_{\mathbf{WA}}^{\perp}\mathbf{W}\dot{\mathbf{A}} \tag{67}$$

with $\dot{\mathbf{A}}$ as defined in (56) and

$$\mathbf{W} = (\mathbf{H}\mathbf{H}^H)^{-1/2}\mathbf{H}. \tag{68}$$

Like in Theorem 3, (65) is precisely the CRB for CBS if there are no stopband sources. If stopband sources exist, it is an approximate bound. An exact bound can be computed as in Remark 1 of Theorem 3.

Remark: This noise-whitening method requires a nonsingular noise covariance, i.e., $\mathbf{H}\mathbf{H}^H$ in the case of CBS. By definition we assume $h(0) \neq 0$ and $h(L-1) \neq 0$ so that L is the filter length. This implies that the banded Toeplitz matrix $\mathbf{H}$ has full row rank, so $\mathbf{H}\mathbf{H}^H$ is positive definite, and $\mathbf{W}$ is well defined. This is also why we cannot obtain a Form-2 CRB if the noise covariance is $p_{\rm e}\mathbf{R}_{\rm e}$ instead of $p_{\rm e}\mathbf{I}$ for any singular $\mathbf{R}_{\rm e}$ known a priori (but we can obtain a Form-1 CRB as in Remark 2 for Theorem 3).

Proof of Theorem 6: Consider the noise-whitening transformation

$$\mathbf{z} = (\mathbf{H}\mathbf{H}^H)^{-1/2}\mathbf{y}. \tag{69}$$

Since this transformation is invertible, the CRB based on $\mathbf{z}$ is the same as the CRB based on $\mathbf{y}$. We can show that the covariance of $\mathbf{z}$ is given by

$$\mathbf{R}_{\mathbf{zz}} = \mathbf{A}_{\mathbf{z}}\mathbf{P}\mathbf{A}_{\mathbf{z}}^H + p_{\rm e}\mathbf{I}, \tag{70}$$

where $\mathbf{A}_{\mathbf{z}} = \mathbf{W}\mathbf{A}$ is the equivalent manifold for $\mathbf{z}$, and $\mathbf{W}$ is defined in (68). Since (70) has the same form as the element-space model in [15], we can apply the CRB expression therein and obtain (65), with

$$\mathbf{S}_1 = \mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{R}_{\mathbf{zz}}^{-1}\mathbf{A}_{\mathbf{z}}\mathbf{P} \tag{71}$$

and $\mathbf{S}_2$ as in (67). Using (68) and (69), we can show that (71) is equivalent to (66). ■

Note that if $D < N - L + 1$, one can verify that the RHS of (49) indeed equals the RHS of (65). The detailed derivations are lengthy and omitted here, but the main idea is to start from computing the block matrix inversion $(\mathbf{\Delta}^H\mathbf{\Delta})^{-1}$ in (49), to express quantities in terms of signal and noise subspaces, and to simplify things using properties of the Kronecker product and Khatri–Rao product. Then, one can finally obtain (65). In this process, we need $D < N - L + 1$ to guarantee the noise subspace to have a nonzero dimension.

Theorem 6 offers a second form of CRB for CBS. Some additional insight can be obtained from Form 2. To this end, we first show an approximation as follows.

Theorem 7: Assume

$$\left\| p_{\rm e}(\mathbf{A}^H\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A})^{-1/2}\mathbf{P}^{-1}(\mathbf{A}^H\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A})^{-1/2} \right\| \ll 1. \tag{72}$$

Then the CRB (65) can be approximated by

$$\mathrm{CRB}_{\rm unb}(\boldsymbol{\omega}) \approx \frac{p_{\rm e}}{2K}\left[\mathrm{Re}\{\mathbf{P} \circ \mathbf{S}_2^*\}\right]^{-1}, \tag{73}$$

where $\mathbf{S}_2$ is as in (67).

Remark: The insight to be gained from this theorem will be clear from Corollaries 3 and 4 below.

Proof of Theorem 7: To prove the theorem, we only have to show that $\mathbf{S}_1 \approx \mathbf{P}$ under the assumption (72). Consider (70) and (71). Applying the matrix inversion lemma [28], we have

$$\begin{aligned}\mathbf{R}_{\mathbf{zz}}^{-1} &= p_{\rm e}^{-1}\mathbf{I} - p_{\rm e}^{-2}\mathbf{A}_{\mathbf{z}}(\mathbf{P}^{-1} + p_{\rm e}^{-1}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-1}\mathbf{A}_{\mathbf{z}}^H \\ &= p_{\rm e}^{-1}\mathbf{I} - p_{\rm e}^{-1}\mathbf{A}_{\mathbf{z}}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}(\mathbf{S}+\mathbf{I})^{-1}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}\mathbf{A}_{\mathbf{z}}^H\end{aligned}$$

where

$$\begin{aligned}\mathbf{S} &= p_{\rm e}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}\mathbf{P}^{-1}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}} &(74)\\ &= p_{\rm e}(\mathbf{A}^H\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A})^{-1/2}\mathbf{P}^{-1}(\mathbf{A}^H\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A})^{-1/2}. &(75)\end{aligned}$$

Now since $\|\mathbf{S}\| \ll 1$ due to (72), we have $(\mathbf{S}+\mathbf{I})^{-1} \approx \mathbf{I}-\mathbf{S}$. Thus,

$$\begin{aligned}\mathbf{S}_1 &= \mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{R}_{\mathbf{zz}}^{-1}\mathbf{A}_{\mathbf{z}}\mathbf{P} &(76)\\ &\approx p_{\mathrm{e}}^{-1}\mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}\mathbf{P}\\ &\quad - p_{\mathrm{e}}^{-1}\mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}\mathbf{P}\\ &\quad + p_{\mathrm{e}}^{-1}\mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}\mathbf{S}(\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}})^{-\frac{1}{2}}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}\mathbf{P} &(77)\\ &= p_{\mathrm{e}}^{-1}\mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}\mathbf{P} - p_{\mathrm{e}}^{-1}\mathbf{P}\mathbf{A}_{\mathbf{z}}^H\mathbf{A}_{\mathbf{z}}\mathbf{P} + \mathbf{P} = \mathbf{P}. &(78)\end{aligned}$$

This completes the proof. ∎

The assumption (72) is satisfied in practical CBS and element-space systems with large arrays. To understand why it is the case, note that the LHS of (72) is less than or equal to $\|p_{\mathrm{e}}\mathbf{P}^{-1}\| \cdot \|(\mathbf{A}^H\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A})^{-1}\|$. If the number of sensors N is large and if the DOAs are not very close to one another, then $\mathbf{A}^H\mathbf{A} \approx N\mathbf{I}$. That is, the columns of $\mathbf{A}$ are approximately orthogonal. Meanwhile, according to our numerical experiments, $\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A} \approx \mathbf{A}$ for properly designed CBS filter. Together, we have $\|(\mathbf{A}^H\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A})^{-1}\| \approx N^{-1}$. Moreover, we assume that the SNR is not very small and that the correlation coefficients between the sources are not close to 1, so that $\|p_{\mathrm{e}}\mathbf{P}^{-1}\| \ll N$. Hence, we finally obtain (72). Numerical values for the LHS of (72) will be given in the example of Fig. 3.

The fact that $\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A} \approx \mathbf{A}$ for properly designed CBS filter is expectable to some extent. In [26], it is shown that the CRB based on $\mathbf{y} = \mathbf{T}\mathbf{x}$ for any beamspace matrix $\mathbf{T}$ is larger than or equal to the element-space CRB, and that $\mathbf{\Pi}_{\mathbf{T}^H}\mathbf{A} = \mathbf{A}$ is a necessary condition for the beamspace CRB to equal the element-space CRB. Thus, it is not surprising that the well-performing method CBS leads to $\mathbf{\Pi}_{\mathbf{H}^H}\mathbf{A} \approx \mathbf{A}$. This means that all steering vectors roughly lie in the column space of $\mathbf{H}^H$, so no significant signal information is lost.

Having proved Theorem 7 and discussed the practical side of the assumption (72), we now derive the following corollary relating CRB to source powers.

Corollary 3: Suppose the assumption in Theorem 7 holds so that we have (73). Also assume the correlation coefficient of each pair of sources is fixed. Then the CRB for a DOA is approximately inversely proportional to its own source power and approximately independent of the source powers of the other DOAs. This is true for both CBS and element-space.

Proof: We can express the source covariance $\mathbf{P}$ as

$$\mathbf{P} = \mathbf{D}_p\mathbf{R}_\rho\mathbf{D}_p, \tag{79}$$

where $\mathbf{D}_p = \mathrm{diag}(\sqrt{p_1}, \ldots, \sqrt{p_D})$, and the correlation coefficient matrix $\mathbf{R}_\rho$ has entries $[\mathbf{R}_\rho]_{i,k} = \mathrm{E}[c_i c_k^*]/\sqrt{p_i p_k}$. Hence, (73) implies

$$\begin{aligned}\mathrm{CRB}_{\mathrm{unb}}(\boldsymbol{\omega}) &\approx \frac{p_{\mathrm{e}}}{2K}\left[\mathrm{Re}\{\mathbf{D}_p\mathbf{R}_\rho\mathbf{D}_p \circ \mathbf{S}_2^*\}\right]^{-1}, &(80)\\ &= \frac{p_{\mathrm{e}}}{2K}\mathbf{D}_p^{-1}\left[\mathrm{Re}\{\mathbf{R}_\rho \circ \mathbf{S}_2^*\}\right]^{-1}\mathbf{D}_p^{-1}. &(81)\end{aligned}$$

Thus, the CRB for the ith DOA is

$$\mathrm{CRB}_{\mathrm{unb}}(\omega_i) \approx \frac{p_{\mathrm{e}}}{2Kp_i}\left[\left[\mathrm{Re}\{\mathbf{R}_\rho \circ \mathbf{S}_2^*\}\right]^{-1}\right]_{i,i}, \tag{82}$$

which is inversely proportional to p_i and independent of p_k for all $k \neq i$. This is also true for element-space since CBS reduces to element-space if we set the filter length $L = 1$ and CBS filter $H(z) = 1$. ∎

Corollary 3 shows that the CRB for a DOA almost does not depend on the power of another DOA. In particular, the CBS CRB for in-band DOAs is almost independent of out-of-band source powers. That is, a more powerful out-of-band jammer does not impose a larger lower bound on MSE of in-band DOA estimates. As we shall see in simulations (Fig. 3), CBS can yield an in-band MSE almost independent of the out-of-band source powers as long as the filter stopband attenuation is large enough to sufficiently attenuate the stopband sources.

Another corollary relating CRB to noise power can be obtained from Theorem 7 as stated below.

Corollary 4: Suppose the assumption in Theorem 7 holds so that we have (73). Then the CRB for a DOA is approximately proportional to the noise power p_{e}. This is true for both CBS and element-space.

Proof: It is immediately proved by (73). ∎

Corollary 3 and Corollary 4 together imply that the CRB of a DOA ω_i is approximately inversely proportional to its own SNR p_i/p_{e}. This will be verified by simulations (Fig. 5). Note that the two corollaries may not be easily proved by directly using the CRB expression in Theorem 3 or Theorem 6. Even Theorem 6 does not immediately lead to Corollary 4 because $\mathbf{S}_1$ in (65) depends on p_{e}. In this sense, Theorem 7 is an important result which gives much more insight.

V. Simulations

In the following numerical examples, we assume the number of DOAs is known. To compare with CBS using a filter $H(z)$, for element-space, we just consider DOA estimates in the passband of $H(z)$ and ignore those in the stopband. Whenever we mention mean square errors (MSEs) or root mean square errors (RMSEs) in detected in-band source angles, we refer to averaging square errors measured in ω over all in-band DOAs. Similarly, since the stochastic CRBs [15] differ for different DOAs, we average the variance bounds over all in-band DOAs. The theoretical MSEs of CBS and element-space are computed from (29) and (31), respectively. Unless otherwise stated, the CBS and element-space CRBs are computed from (49) and (53), respectively. Note that (49) is precisely the CRB for CBS only if there are no stopband sources so that the in-band DOA estimates are unbiased. If stopband sources exist, (49) is approximately a lower bound on the variance of DOA estimates of CBS due to (40). The exact lower bound (39) for the biased CBS estimator will be compared to the approximate bound in Fig. 4. The CBS filter $H(z)$ is designed to be a spectral factor of a lowpass Nyquist-equiripple filter [29], with passband edge $\pi/2M$ and stopband edge $3\pi/2M$, where M is the decimation ratio.

In Fig. 1, we compare RMSEs of DOA estimates and CRBs for CBS with various filter length L, and for element-space. The RMSEs based on Monte Carlo simulation and based on our theoretical analysis in Sec. III are both shown for comparison. We consider a ULA with $N = 99$ sensors receiving 6 sources at angles $\theta = -5°, 0°, 5°, 40°$, $60°$, and $80°$. All sources have power 1. Sources n and $n+3$ have a correlation coefficient $\rho = 0.9$ for $n = 1, 2, 3$. For CBS, the decimation ratio is $M = 4$. Hence, the three sources at $-5°, 0°$ and $5°$ are in the passband, and the others are in the stopband. Noise variance is $p_{\mathrm{e}} = 1$. Root-MUSIC is used to estimate DOAs. Covariance estimates are obtained by using 500 snapshots, and 500 Monte Carlo runs are used. Several observations can be made from Fig. 1. First, the theoretical RMSE curves almost coincide with simulated RMSE curves for both CBS and element-space. This verifies our MSE analysis in Sec. III. Second, the CBS CRB is always larger than the element-space CRB. This is consistent with the known

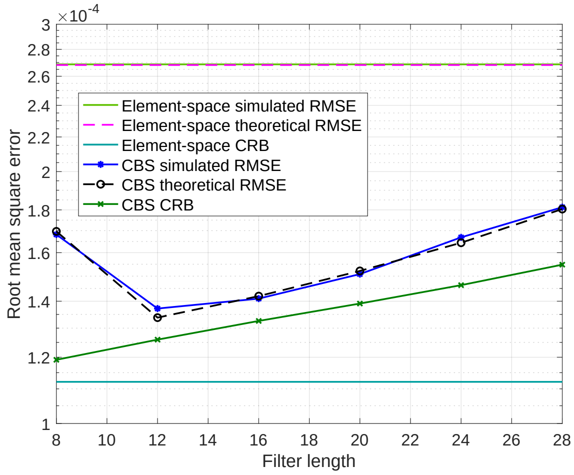

Fig. 1. Simulated RMSE, theoretical RMSE, and CRB for CBS with various filter length L, and for element-space.

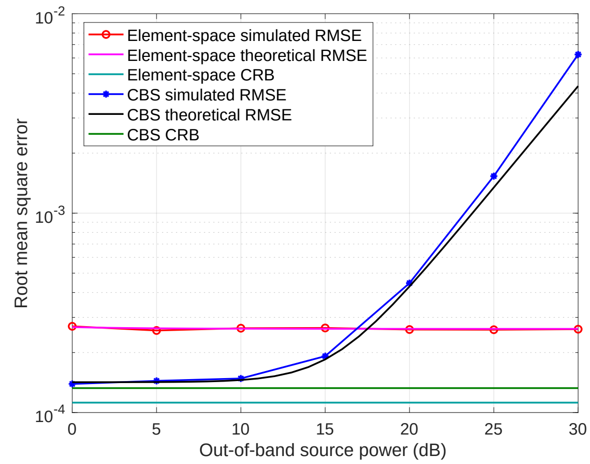

Fig. 3. Simulated RMSE, theoretical RMSE, and CRB for element-space and CBS as the out-of-band source power varies.

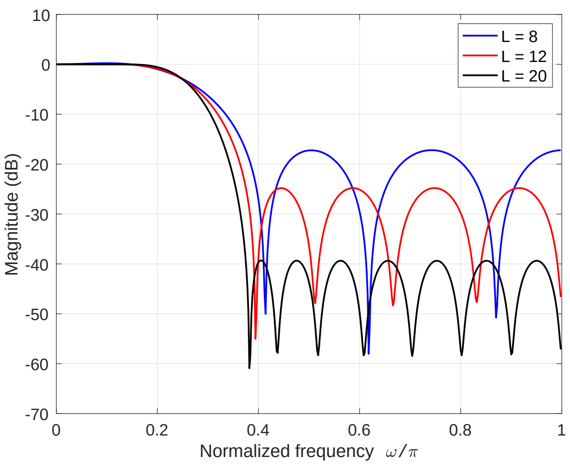

Fig. 2. Magnitude responses of the Nyquist-equiripple filters used for CBS with several typical values of filter length L.

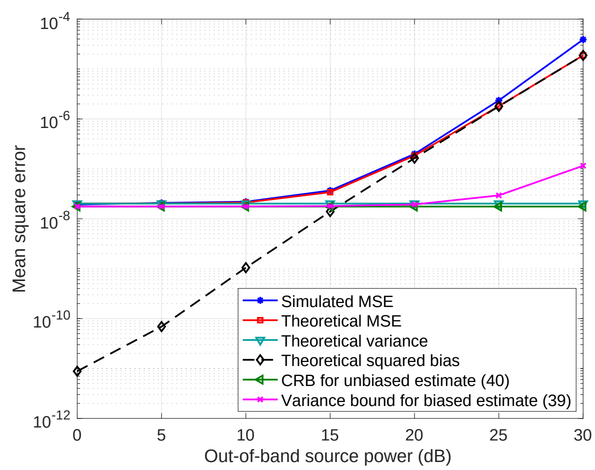

Fig. 4. Simulated MSE, theoretical MSE, theoretical variance, theoretical squared bias, CRB for unbiased estimate (40), and the exact variance bound (39) for biased estimate for CBS as the out-of-band source power varies.

fact that beamspace CRB cannot be smaller than element-space CRB [26]. However, the RMSE of CBS is uniformly smaller than that of element-space in this example. That is, CBS offers a practical algorithm with RMSE approaching the CRB when there are correlated sources, while there is a large gap between RMSE and CRB for element-space. Finally, there is an optimal filter length such that the RMSE of CBS is minimized. If the filter length is too small, the stopband attenuation is not good enough (see Fig. 2), and the filtered stopband sources contribute to a large bias B_i as defined in (25). If the filter length is too large, we need to discard many transient samples in the filter output (since we retain only steady-state samples). This eventually leads to a larger RMSE. By plotting a RMSE curve as a function of the filter length using our analysis in Sec. III, we can determine the optimal filter length. Magnitude responses of the Nyquist-equiripple filters used for CBS with several typical values of filter length L are shown in Fig. 2. These filters indeed have equiripple stopband attenuation. Besides, we obtain better stopband attenuation as the filter length increases.

In Fig. 3, we compare RMSEs of DOA estimates and CRBs for CBS and element-space as we vary the out-of-band source power. The filter length is now fixed at $L = 16$, and all the other simulation parameters are the same as in the example of Fig. 1. According to Fig. 3, both the CBS CRB and element-space CRB almost do not depend on the out-of-band source power, as implied by Corollary 3. To obtain this corollary, we need the assumption (72) in Theorem 7 to be valid. This is indeed the case since the LHS of (72) gradually decreases from 0.112 to 0.0627 as the out-of-band source power increases from 0 dB to 30 dB. Besides, just like in Fig. 1, the CBS CRB is larger than the element-space CRB. However, the RMSE of CBS is smaller than that of element-space if the out-of-band source power is not too large. To better understand how the MSE and CRB of CBS change with the out-of-band source power, in Fig. 4, we show the simulated MSE, theoretical MSE $B_i^2 + V_i$ (as defined in (29)), theoretical variance V_i, theoretical squared bias B_i^2, CRB for unbiased estimate (40), and the exact variance bound (39) for biased estimate for CBS. As long as the out-of-band source power is not extremely large, say, not greater than 20 dB in this example, our theoretical expression gives a good estimate of the true MSE. If the out-of-band source power is extremely large, our assumption that the covariance perturbation $\delta\mathbf{R}$ in (12) due to the filtered stopband sources is small will be

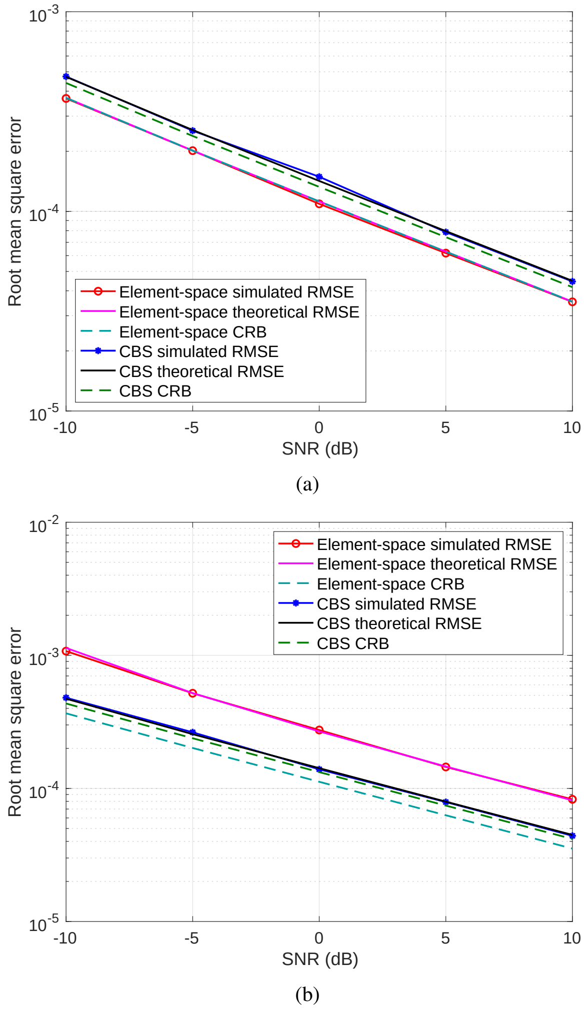

Fig. 5. Simulated RMSE, theoretical RMSE, and CRB for element-space and CBS as the SNR varies. Sources n and $n+3$ have a correlation coefficient ρ for $n = 1, 2, 3$. (a) $\rho = 0$. (b) $\rho = 0.9$.

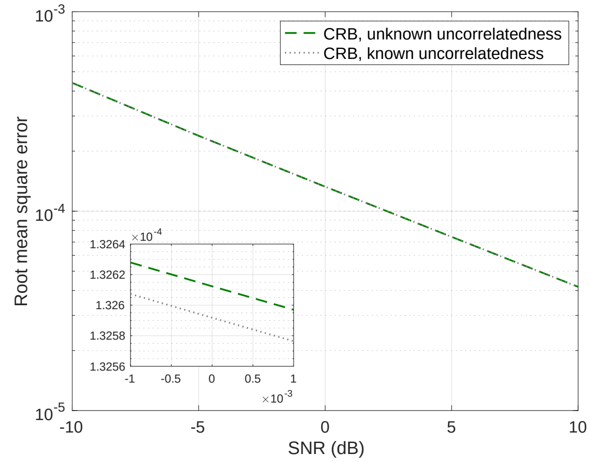

Fig. 6. CRB for CBS as the SNR varies. The uncorrelatedness of the sources are assumed either unknown or known a priori.

invalid. Thus, there is a larger difference between theoretical MSE and simulated MSE. Another observation is that the theoretical squared bias is much smaller than the theoretical variance for out-of-band source power not greater than 10 dB. Hence, we may ignore the bias term in the MSE analysis if the passband and stopband sources have similar powers. Finally, as mentioned in the beginning of this section, the CBS CRB in each previous plot is approximately a lower bound on the variance of DOA estimate due to (40). Here we compare the exact lower bound (39) for the biased CBS estimator with the approximate bound. The exact CRB (39) is computed from (34) and (43) with $\mathbf{G}$ as in (50) and $\boldsymbol{\Delta}_1$ as in (137), and the derivatives in the Jacobian matrix are computed numerically. We see that the approximate bound is similar to the exact bound if the out-of-band source power is not greater than 20 dB. Hence, (40) gives a good lower bound in practice.

In Fig. 5, we compare RMSEs of DOA estimates and CRBs for CBS and element-space as we vary the SNR. We consider the same three passband sources and three stopband sources as in the example of Fig. 1. All sources have power 1. The SNR is thus $1/p_\mathrm{e}$, where the noise power p_e is varied. Sources n and $n+3$ have a correlation coefficient ρ for $n = 1, 2, 3$. All the other simulation parameters are the same as before. In Fig. 5(a), we consider $\rho = 0$. In this uncorrelated case, both CBS and element-space RMSEs approach the CRBs, and the two systems have similar performance. In Fig. 5(b), we consider $\rho = 0.9$. In this correlated case, CBS RMSE still approaches its CRB, but element-space RMSE does not. CBS RMSE is significantly smaller than element-space RMSE. In these two subfigures, the theoretical RMSE curves again almost coincide with simulated RMSE curves for both CBS and element-space. Moreover, both CBS and element-space CRBs look linear in the log-log plots, and the CRBs decrease by a factor of 10 as SNR increases by 20 dB. That is, the CRBs for MSE is inversely proportional to the SNR. This verifies the result of Corollary 4.

As mentioned in the beginning of this section, in all the previous examples, the CBS and element-space CRBs are computed from (49) and (53), respectively. That is, the correlations P_{ik} between sources are assume unknown as in (33). When all the sources are uncorrelated, we can also assume that the uncorrelatedness is known a priori. Under this assumption, we have obtained the CBS and element-space CRBs in (58) and (61), respectively. To compare the two cases, we consider again the example of Fig. 5(a) and compute the CBS CRB assuming either unknown or known uncorrelatedness. The results are shown in Fig. 6. The CRB assuming known uncorrelatedness is smaller than the CRB assuming unknown uncorrelatedness, which verifies Theorem 5. Interestingly, the difference between them is very small, so this prior information is not so influential.

VI. Conclusion

The MSE performance and Cramer–Rao Bound (CRB) for convolutional beamspace (CBS) are analyzed in this paper. Theoretical expressions of MSE are derived assuming that MUSIC or root-MUSIC is used to estimate DOAs. (The performance is the same for both.) The bias of the CBS estimator, though negligible in some cases, is given via a first-order perturbation analysis. To obtain the variance of the CBS estimator, we develop an approach to derive the asymptotic probability distribution of the eigenvectors of the average finite-snapshot covariance matrices of dependent random vectors. This approach can be useful to other applications because previous results are only for independent random vectors. As for CRB, we offer two forms of expressions. Form

1 is derived directly from the Fisher information matrix. Form 2 is derived via a noise-whitening approach and offers more insight. In particular, the CRB for a DOA is approximately inversely proportional to its own source power and approximately independent of the source powers of the other DOAs. Also, the CRB for a DOA is approximately proportional to the noise power. These results are also true for element-space and to the best of our knowledge, have not been theoretically established in previous works. Extensive numerical examples are also given, which verify the theoretical results.

Appendix A
Proof of Lemma 1

Consider the K-snapshot estimate $\widehat{\mathbf{R}}_{\text{ave},0} = \widehat{\mathbf{R}}_{\text{ave}}\big|_{\delta\mathbf{R}=\mathbf{O}}$ with stopband sources nulled, where $\widehat{\mathbf{R}}_{\text{ave}}$ is defined in (15). We have

$$\mathrm{E}[\widehat{\mathbf{R}}_{\text{ave},0}] = \mathbf{R}_{\text{ave},0} = \mathbf{A}_{\text{dec},0}\breve{\mathbf{R}}_{d,0}\mathbf{A}_{\text{dec},0}^H + p_{\text{e}}\mathbf{G}_{\text{dec}} \tag{83}$$

and

$$\begin{aligned}
&K^2M^2\,\mathrm{E}\left[[\widehat{\mathbf{R}}_{\text{ave},0}]_{i,p}[\widehat{\mathbf{R}}_{\text{ave},0}]^*_{g,h}\right]\\
&= \mathrm{E}\left[\sum_{k=1}^{K}\sum_{l=0}^{M-1}[\mathbf{v}_l[k]]_i[\mathbf{v}_l[k]]_p^*\sum_{n=1}^{K}\sum_{m=0}^{M-1}[\mathbf{v}_m[n]]_g^*[\mathbf{v}_m[n]]_h\right]\\
&= \mathrm{E}\left[\sum_{k=1}^{K}\sum_{l=0}^{M-1}\sum_{m=0}^{M-1}[\mathbf{v}_l[k]]_i[\mathbf{v}_l[k]]_p^*[\mathbf{v}_m[k]]_g^*[\mathbf{v}_m[k]]_h\right.\\
&\qquad\left.+\sum_{k=1}^{K}\sum_{n\neq k}\sum_{l=0}^{M-1}\sum_{m=0}^{M-1}[\mathbf{v}_l[k]]_i[\mathbf{v}_l[k]]_p^*[\mathbf{v}_m[n]]_g^*[\mathbf{v}_m[n]]_h\right]\\
&= \mathrm{E}\left[K\sum_{l=0}^{M-1}\sum_{m=0}^{M-1}[\mathbf{v}_l]_i[\mathbf{v}_l]_p^*[\mathbf{v}_m]_g^*[\mathbf{v}_m]_h\right.\\
&\qquad\left.+K(K-1)M^2[\mathbf{R}_{\text{ave},0}]_{i,p}[\mathbf{R}_{\text{ave},0}]^*_{g,h}\right]
\end{aligned}\tag{84}$$

for $0\le i,p,g,h\le J-1$. Since the entries of $\mathbf{v}_l$ for all l are circularly-symmetric complex Gaussian, using Wick's theorem (or Isserlis' theorem) [30], [31], we have

$$\begin{aligned}\mathrm{E}\left[[\mathbf{v}_l]_i[\mathbf{v}_l]_p^*[\mathbf{v}_m]_g^*[\mathbf{v}_m]_h\right] &= \mathrm{E}\left[[\mathbf{v}_l]_i[\mathbf{v}_l]_p^*\right]\mathrm{E}\left[[\mathbf{v}_m]_g^*[\mathbf{v}_m]_h\right]\\
&\quad+\mathrm{E}\left[[\mathbf{v}_l]_i[\mathbf{v}_m]_g^*\right]\mathrm{E}\left[[\mathbf{v}_l]_p^*[\mathbf{v}_m]_h\right].\end{aligned}$$

Hence, we can obtain the covariance

$$\begin{aligned}
&\mathrm{E}\left[([\widehat{\mathbf{R}}_{\text{ave},0}]_{i,p}-[\mathbf{R}_{\text{ave},0}]_{i,p})([\widehat{\mathbf{R}}_{\text{ave},0}]_{g,h}-[\mathbf{R}_{\text{ave},0}]_{g,h})^*\right]\\
&\quad=\frac{1}{KM^2}\sum_{l=0}^{M-1}\sum_{m=0}^{M-1}\mathrm{E}\left[[\mathbf{v}_l]_i[\mathbf{v}_m]_g^*\right]\mathrm{E}\left[[\mathbf{v}_l]_p^*[\mathbf{v}_m]_h\right] \qquad (85)\\
&\quad=\frac{1}{KM^2}\operatorname{tr}\left(\mathbf{R}_{\mathbf{yy}}^{(i,g)}(\mathbf{R}_{\mathbf{yy}}^{(p,h)})^H\right), \qquad (86)
\end{aligned}$$

where $\mathbf{R}_{\mathbf{yy}}^{(i,g)}\in\mathbb{C}^{M\times M}$ are submatrices of

$$\mathbf{R}_{\mathbf{yy},0}=\mathbf{R}_{\mathbf{yy}}\big|_{\delta\mathbf{R}=\mathbf{O}}\triangleq\begin{bmatrix}\mathbf{R}_{\mathbf{yy}}^{(1,1)} & \cdots & \mathbf{R}_{\mathbf{yy}}^{(1,J)}\\ \vdots & \ddots & \vdots\\ \mathbf{R}_{\mathbf{yy}}^{(J,1)} & \cdots & \mathbf{R}_{\mathbf{yy}}^{(J,J)}\end{bmatrix}. \tag{87}$$

Throughout the above derivations, we implicitly assume the stopband sources have been nulled.

Then consider (16) and let $\mathbf{E}_0=[\mathbf{E}_{\text{s},0}\ \mathbf{E}_{\text{n},0}]$ and $\mathbf{\Lambda}_0=\operatorname{diag}(\mathbf{\Lambda}_{\text{s},0},\mathbf{\Lambda}_{\text{n},0})$. Define

$$\mathbf{U}=K^{1/2}(\mathbf{E}_0^H\widehat{\mathbf{R}}_{\text{ave},0}\mathbf{E}_0-\mathbf{\Lambda}_0). \tag{88}$$

Based on the multivariate central limit theorem and (86), we can derive that $\mathbf{U}$ is asymptotically (for large K) complex Gaussian with mean zero and covariance

$$\mathrm{E}\left[[\mathbf{U}]_{i,p}[\mathbf{U}]^*_{g,h}\right]=\frac{1}{M^2}\operatorname{tr}\left(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(i,g)}(\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(p,h)})^H\right), \tag{89}$$

where $\widetilde{\mathbf{R}}_{\mathbf{yy}}^{(i,g)}$ are as defined in (21). Let

$$\mathbf{V}=K^{-1/2}\mathbf{U}+\mathbf{\Lambda}_0=\mathbf{E}_0^H\widehat{\mathbf{R}}_{\text{ave},0}\mathbf{E}_0 \tag{90}$$

and suppose that

$$\mathbf{V}=\mathbf{Q}\mathbf{D}_{\mathbf{V}}\mathbf{Q}^H \tag{91}$$

is the eigenvalue decomposition of $\mathbf{V}$ with eigenvalues in descending order. We partition $\mathbf{U}$ and $\mathbf{Q}$ into

$$\mathbf{U}=\begin{bmatrix}u_{11} & u_{12} & \cdots & u_{1D_0} & \mathbf{u}_{1\text{e}}^T\\ u_{21} & u_{22} & \cdots & u_{2D_0} & \mathbf{u}_{2\text{e}}^T\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ u_{D_01} & u_{D_02} & \cdots & u_{D_0D_0} & \mathbf{u}_{D_0\text{e}}^T\\ \mathbf{u}_{\text{e}1} & \mathbf{u}_{\text{e}2} & \cdots & \mathbf{u}_{\text{e}D_0} & \mathbf{U}_{\text{ee}}\end{bmatrix}, \tag{92}$$

$$\mathbf{Q}=\begin{bmatrix}q_{11} & q_{12} & \cdots & q_{1D_0} & \mathbf{q}_{1\text{e}}^T\\ q_{21} & q_{22} & \cdots & q_{2D_0} & \mathbf{q}_{2\text{e}}^T\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ q_{D_01} & q_{D_02} & \cdots & q_{D_0D_0} & \mathbf{q}_{D_0\text{e}}^T\\ \mathbf{q}_{\text{e}1} & \mathbf{q}_{\text{e}2} & \cdots & \mathbf{q}_{\text{e}D_0} & \mathbf{Q}_{\text{ee}}\end{bmatrix}, \tag{93}$$

where $\mathbf{U}_{\text{ee}},\mathbf{Q}_{\text{ee}}\in\mathbb{C}^{(J-D_0)\times(J-D_0)}$ and $\mathbf{u}_{i\text{e}},\mathbf{u}_{\text{e}i},\mathbf{q}_{i\text{e}},\mathbf{q}_{\text{e}i}\in\mathbb{C}^{J-D_0}$, $i=1,\ldots,D_0$. Following the derivations in [13], we can show that asymptotically (for large K), $K^{1/2}(q_{ll}-1)$ converges stochastically to 0, and the limiting distributions of $K^{1/2}q_{il}$ and $K^{1/2}\mathbf{q}_{\text{e}l}$ are the same as the limiting distributions of $u_{li}^*/(\lambda_{l,0}-\lambda_{i,0})$ and $\mathbf{u}_{l\text{e}}^*/(\lambda_{l,0}-p_{\text{e}})$, $1\le i,l\le D_0$, $i\neq l$. Using (90) and (91), we obtain

$$\widehat{\mathbf{R}}_{\text{ave},0}=(\mathbf{E}_0\mathbf{Q})\mathbf{D}_{\mathbf{V}}(\mathbf{E}_0\mathbf{Q})^H, \tag{94}$$

which is the eigendecomposition of $\widehat{\mathbf{R}}_{\text{ave},0}$. Hence, asymptotically, the signal eigenvector estimates are

$$\begin{aligned}
\hat{\mathbf{e}}_{l,0}&=\mathbf{E}_0[\mathbf{Q}]_{:,l} && (95)\\
&=\mathbf{e}_{l,0}+K^{-1/2}\left[\sum_{\substack{i=1\\i\neq l}}^{D_0}\frac{[\mathbf{U}]^*_{l,i}}{\lambda_{l,0}-\lambda_{i,0}}\mathbf{e}_{i,0}+\frac{\mathbf{E}_{\text{n},0}\mathbf{u}_{l\text{e}}^*}{\lambda_{l,0}-p_{\text{e}}}\right] && (96)\\
&=\mathbf{e}_{l,0}+K^{-1/2}\sum_{\substack{i=1\\i\neq l}}^{J}\frac{[\mathbf{U}]^*_{l,i}}{\lambda_{l,0}-\lambda_{i,0}}\mathbf{e}_{i,0}, && (97)
\end{aligned}$$

$l=1,\ldots,D_0$. Using (89) and (97), we obtain the statement of the lemma. To obtain the relation matrices, we have also used the fact that $\mathbf{U}$ is Hermitian symmetric.

Appendix B
Proof of Theorem 1

Following the idea (Eq. (B.2a)) of [12], asymptotically (for large K), the DOA estimates $\hat{\omega}_i$ of MUSIC based on (15) satisfy

$$\hat{\omega}_i-\omega_i\approx\frac{-\operatorname{Re}\{\mathbf{a}_J^H(\tilde{\omega}_i)\widehat{\mathbf{E}}_{\text{n}}\widehat{\mathbf{E}}_{\text{n}}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)\}}{M\dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{n}}\mathbf{E}_{\text{n}}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)}, \tag{98}$$

where $\tilde{\omega}_i=M\omega_i$, $\dot{\mathbf{a}}_J(\omega)=\frac{\mathrm{d}}{\mathrm{d}\omega}\mathbf{a}_J(\omega)$, and the terms neglected in the approximation are $O(1/K)$. Besides, according to [8], root-MUSIC has the same asymptotic error expression (98). Then in view of (17), for small $\|\delta \mathbf{R}\|$,

$$\dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_\text{n}\mathbf{E}_\text{n}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i) = \dot{\mathbf{a}}_J^H(\tilde{\omega}_i)(\mathbf{I}-\mathbf{E}_\text{s}\mathbf{E}_\text{s}^H)\dot{\mathbf{a}}_J(\tilde{\omega}_i) \tag{99}$$
$$\approx \dot{\mathbf{a}}_J^H(\tilde{\omega}_i)(\mathbf{I}-\mathbf{E}_{\text{s},0}\mathbf{E}_{\text{s},0}^H)\dot{\mathbf{a}}_J(\tilde{\omega}_i) \tag{100}$$
$$= \dot{\mathbf{a}}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i), \tag{101}$$

where the terms neglected in the approximation are $O(\|\delta\mathbf{R}\|)$. Thus,

$$\hat{\omega}_i - \omega_i \approx -\text{Re}\{\mathbf{a}_J^H(\tilde{\omega}_i)\widehat{\mathbf{E}}_\text{n}\widehat{\mathbf{E}}_\text{n}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)\}/(Mg(\tilde{\omega}_i)), \tag{102}$$

where $g(\omega) = \dot{\mathbf{a}}_J^H(\omega)\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\dot{\mathbf{a}}_J(\omega)$. Besides,

$$\widehat{\mathbf{E}}_\text{n}\widehat{\mathbf{E}}_\text{n}^H = \mathbf{I} - \widehat{\mathbf{E}}_\text{s}\widehat{\mathbf{E}}_\text{s}^H \tag{103}$$
$$\approx \mathbf{I} - \widehat{\mathbf{E}}_{\text{s},0}\widehat{\mathbf{E}}_{\text{s},0}^H - \sum_{l=1}^{D_0}\left(\mathbf{e}_{l,0}\delta\mathbf{e}_l^H + \delta\mathbf{e}_l\mathbf{e}_{l,0}^H\right) \tag{104}$$
$$= \widehat{\mathbf{E}}_{\text{n},0}\widehat{\mathbf{E}}_{\text{n},0}^H - \sum_{l=1}^{D_0}\left(\mathbf{e}_{l,0}\delta\mathbf{e}_l^H + \delta\mathbf{e}_l\mathbf{e}_{l,0}^H\right) \tag{105}$$

where the terms neglected in the approximation are $O(\|\delta\mathbf{R}\|^2)$, and $\delta\mathbf{e}_l$ is defined in (26). Hence, for large K and small $\|\delta\mathbf{R}\|$,

$$\hat{\omega}_i - \omega_i = B_i - \frac{\text{Re}\{\mathbf{a}_J^H(\tilde{\omega}_i)\widehat{\mathbf{E}}_{\text{n},0}\widehat{\mathbf{E}}_{\text{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)\}}{Mg(\tilde{\omega}_i)} \tag{106}$$

where B_i is defined in (25). Next, in view of (97), we can write $\widehat{\mathbf{E}}_{\text{s},0} = \mathbf{E}_{\text{s},0} + \widehat{\mathbf{\Xi}}$, where $\|\widehat{\mathbf{\Xi}}\| = O(K^{-1/2})$. Therefore, since $\mathbf{a}_J(\tilde{\omega}_i)$ lies in the signal subspace, we have

$$\mathbf{a}_J^H(\tilde{\omega}_i)\widehat{\mathbf{E}}_{\text{n},0}\widehat{\mathbf{E}}_{\text{n},0}^H = \mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}\mathbf{E}_{\text{s},0}^H\widehat{\mathbf{E}}_{\text{n},0}\widehat{\mathbf{E}}_{\text{n},0}^H \tag{107}$$
$$= \mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}(\mathbf{E}_{\text{s},0} - \widehat{\mathbf{E}}_{\text{s},0})^H\widehat{\mathbf{E}}_{\text{n},0}\widehat{\mathbf{E}}_{\text{n},0}^H \tag{108}$$
$$= -\mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}\widehat{\mathbf{\Xi}}^H(\mathbf{I} - \widehat{\mathbf{E}}_{\text{s},0}\widehat{\mathbf{E}}_{\text{s},0}^H), \tag{109}$$

where we used the fact that $\widehat{\mathbf{E}}_{\text{s},0}^H\widehat{\mathbf{E}}_{\text{n},0} = \mathbf{O}$ to obtain (108). Replacing $\widehat{\mathbf{E}}_{\text{s},0}$ by $\mathbf{E}_{\text{s},0} + \widehat{\mathbf{\Xi}}$ and neglecting $O(1/K)$ terms, we obtain

$$\mathbf{a}_J^H(\tilde{\omega}_i)\widehat{\mathbf{E}}_{\text{n},0}\widehat{\mathbf{E}}_{\text{n},0}^H \approx -\mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}\widehat{\mathbf{\Xi}}^H(\mathbf{I} - \mathbf{E}_{\text{s},0}\mathbf{E}_{\text{s},0}^H) \tag{110}$$
$$= -\mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}\widehat{\mathbf{\Xi}}^H\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H \tag{111}$$
$$= -\mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}\widehat{\mathbf{E}}_{\text{s},0}^H\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H, \tag{112}$$

where we used the fact that $\mathbf{E}_{\text{s},0}^H\mathbf{E}_{\text{n},0} = \mathbf{O}$ to obtain the last equality. Substituting this into (106), we have

$$\hat{\omega}_i - \omega_i = B_i + \widehat{R}_i \tag{113}$$

where

$$\widehat{R}_i = \frac{\text{Re}\{\mathbf{a}_J^H(\tilde{\omega}_i)\mathbf{E}_{\text{s},0}\widehat{\mathbf{E}}_{\text{s},0}^H\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)\}}{Mg(\tilde{\omega}_i)}. \tag{114}$$

Here we see that the DOA estimation errors $\hat{\omega}_i - \omega_i$ are asymptotically (for large K) jointly Gaussian since the signal eigenvector estimates in $\widehat{\mathbf{E}}_{\text{s},0}$ are jointly complex Gaussian due to Lemma 1. Note that $\text{E}\left[\widehat{R}_i\right] = 0$ because $\text{E}\left[\widehat{\mathbf{E}}_{\text{s},0}^H\mathbf{E}_{\text{n},0}\right] = \mathbf{E}_{\text{s},0}^H\mathbf{E}_{\text{n},0} = \mathbf{O}$, so $\text{E}[\hat{\omega}_i - \omega_i] = B_i$. Moreover,

$$\text{E}[(\hat{\omega}_i - \omega_i)(\hat{\omega}_k - \omega_k)]$$
$$= B_iB_k + B_i\,\text{E}\left[\widehat{R}_k\right] + \text{E}\left[\widehat{R}_i\right]B_k + \text{E}\left[\widehat{R}_i\widehat{R}_k\right] \tag{115}$$
$$= B_iB_k + \text{E}\left[\widehat{R}_i\widehat{R}_k\right], \tag{116}$$

where

$$\text{E}\left[\widehat{R}_i\widehat{R}_k\right] = \text{E}\left[\frac{\text{Re}\{\mathbf{a}_J^H(\tilde{\omega}_i)\sum_{l=1}^{D_0}\mathbf{e}_{l,0}\hat{\mathbf{e}}_{l,0}^H\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_i)\}}{Mg(\tilde{\omega}_i)} \cdot \frac{\text{Re}\{\mathbf{a}_J^H(\tilde{\omega}_k)\sum_{r=1}^{D_0}\mathbf{e}_{r,0}\hat{\mathbf{e}}_{r,0}^H\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_k)\}}{Mg(\tilde{\omega}_k)}\right].$$

Using the fact that $\text{Re}\{u\}\text{Re}\{v\} = \text{Re}\{uv + uv^*\}/2$ for any $u, v \in \mathbb{C}$, we obtain

$$\text{E}\left[\widehat{R}_i\widehat{R}_k\right] = \frac{\text{Re}\{\gamma_1 + \gamma_2\}}{2M^2g(\tilde{\omega}_i)g(\tilde{\omega}_k)}, \tag{117}$$

where

$$\gamma_1 = \sum_{l=1}^{D_0}\sum_{r=1}^{D_0}\Big[\mathbf{e}_{l,0}^H\mathbf{a}_J(\tilde{\omega}_i)\mathbf{a}_J^H(\tilde{\omega}_k)\mathbf{e}_{r,0}\dot{\mathbf{a}}_J^H(\tilde{\omega}_i) \cdot \mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\,\text{E}\left[\hat{\mathbf{e}}_{l,0}\hat{\mathbf{e}}_{r,0}^H\right]\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\dot{\mathbf{a}}_J(\tilde{\omega}_k)\Big] \tag{118}$$

and

$$\gamma_2 = \sum_{l=1}^{D_0}\sum_{r=1}^{D_0}\Big[\mathbf{e}_{l,0}^H\mathbf{a}_J(\tilde{\omega}_i)\mathbf{a}_J^T(\tilde{\omega}_k)\mathbf{e}_{r,0}^*\dot{\mathbf{a}}_J^H(\tilde{\omega}_i) \cdot \mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\,\text{E}\left[\hat{\mathbf{e}}_{l,0}\hat{\mathbf{e}}_{r,0}^T\right]\mathbf{E}_{\text{n},0}^*\mathbf{E}_{\text{n},0}^T\dot{\mathbf{a}}_J^*(\tilde{\omega}_k)\Big]. \tag{119}$$

Using (18) and noting that $\mathbf{E}_{\text{n},0}^H\mathbf{e}_{i,0} = 0$ for $i = 1, \ldots, D_0$ and $\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\mathbf{e}_{i,0} = \mathbf{e}_{i,0}$ for $i = D_0+1, \ldots, J$, we obtain

$$\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\,\text{E}\left[\hat{\mathbf{e}}_{l,0}\hat{\mathbf{e}}_{r,0}^H\right]\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H = \mathbf{E}_{\text{n},0}\mathbf{B}_{lr}\mathbf{E}_{\text{n},0}^H/K, \tag{120}$$

where $\mathbf{B}_{lr}$ are as defined in (27). Similarly, using (19), we obtain

$$\mathbf{E}_{\text{n},0}\mathbf{E}_{\text{n},0}^H\,\text{E}\left[\hat{\mathbf{e}}_{l,0}\hat{\mathbf{e}}_{r,0}^T\right]\mathbf{E}_{\text{n},0}^*\mathbf{E}_{\text{n},0}^T = \mathbf{E}_{\text{n},0}\mathbf{C}_{lr}\mathbf{E}_{\text{n},0}^T/K, \tag{121}$$

where $\mathbf{C}_{lr}$ are as defined in (28). Using (116)-(121), we finally derive the expressions for covariance in the theorem statement.

Appendix C
Proof of Theorems 2 and 3

In view of (44)–(48), we need to compute the derivative of $\mathbf{r}_{\mathbf{yy}}$ with respect to each parameter in $\boldsymbol{\alpha}$. According to (5)–(7), we have that

$$\mathbf{r}_{\mathbf{yy}} = \text{vec}(\mathbf{R}_{\mathbf{yy}}) \tag{122}$$
$$= \text{vec}\left(\sum_{i=1}^{D}\sum_{k=1}^{D}P_{i,k}\mathbf{b}(\omega_i)\mathbf{b}^H(\omega_k) + p_\text{e}\mathbf{H}\mathbf{H}^H\right) \tag{123}$$
$$= \sum_{i=1}^{D}\sum_{k=1}^{D}P_{i,k}\mathbf{b}^*(\omega_k)\otimes\mathbf{b}(\omega_i) + p_\text{e}\text{vec}(\mathbf{H}\mathbf{H}^H) \tag{124}$$
$$= \sum_{i=1}^{D}p_i\mathbf{b}^*(\omega_i)\otimes\mathbf{b}(\omega_i) + p_\text{e}\text{vec}(\mathbf{H}\mathbf{H}^H) + \sum_{i>k}P_{ik}^{(\text{r})}\left[\mathbf{b}^*(\omega_k)\otimes\mathbf{b}(\omega_i) + \mathbf{b}^*(\omega_i)\otimes\mathbf{b}(\omega_k)\right] + \sum_{i>k}jP_{ik}^{(\text{i})}\left[\mathbf{b}^*(\omega_k)\otimes\mathbf{b}(\omega_i) - \mathbf{b}^*(\omega_i)\otimes\mathbf{b}(\omega_k)\right] \tag{125}$$

where $\mathbf{b}(\omega_i) = H(e^{j\omega_i})e^{j\omega_i(L-1)}\mathbf{a}_{N-L+1}(\omega_i)$ for each i. We used the fact that $\mathbf{P}$ is Hermitian symmetric to obtain (125). Using (124), we can derive that

$$\left[\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial \omega_1} \cdots \frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial \omega_D}\right] = \dot{\mathbf{A}}_L^* \odot (\mathbf{A}_L\mathbf{D}_H\mathbf{P}\mathbf{D}_H^*) + (\mathbf{A}_L\mathbf{D}_H\mathbf{P}\mathbf{D}_H^*)^* \odot \dot{\mathbf{A}}_L, \quad (126)$$

where $\dot{\mathbf{A}}_L$ is as in (52), and that

$$\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial p_{\mathrm{e}}} = \mathrm{vec}(\mathbf{H}\mathbf{H}^H). \quad (127)$$

Using (125), we can derive that

$$\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial p_i} = \mathbf{b}^*(\omega_i) \otimes \mathbf{b}(\omega_i) \quad (128)$$

$$= [(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]_{:,(i-1)D+i} \quad (129)$$

for $i = 1, \ldots, D$ and

$$\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial P_{ik}^{(\mathrm{r})}} = \mathbf{b}^*(\omega_k) \otimes \mathbf{b}(\omega_i) + \mathbf{b}^*(\omega_i) \otimes \mathbf{b}(\omega_k) \quad (130)$$

$$= [(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]_{:,(k-1)D+i} + [(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]_{:,(i-1)D+k} \quad (131)$$

$$\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial P_{ik}^{(\mathrm{i})}} = j\mathbf{b}^*(\omega_k) \otimes \mathbf{b}(\omega_i) - j\mathbf{b}^*(\omega_i) \otimes \mathbf{b}(\omega_k) \quad (132)$$

$$= j[(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]_{:,(k-1)D+i} - j[(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]_{:,(i-1)D+k} \quad (133)$$

for $i > k$. Hence, there exists a matrix $\mathbf{T} \in \mathbb{C}^{D^2 \times D^2}$ such that

$$\left[\left[\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial p_i}\right]_{i=1}^{D} \left[\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial P_{ik}^{(\mathrm{r})}}\right]_{i>k} \left[\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial P_{ik}^{(\mathrm{i})}}\right]_{i>k}\right] = [(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]\mathbf{T} \quad (134)$$

holds. More precisely, $[\mathbf{T}]_{:,i} = \boldsymbol{\delta}_{(i-1)D+i}^{(D^2)}$ for $i = 1, \ldots, D$, and the column of $\mathbf{T}$ corresponding to $\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial P_{ik}^{(\mathrm{r})}}$, $\frac{\partial \mathbf{r}_{\mathbf{yy}}}{\partial P_{ik}^{(\mathrm{i})}}$ are

$$\boldsymbol{\delta}_{(k-1)D+i}^{(D^2)} + \boldsymbol{\delta}_{(i-1)D+k}^{(D^2)}, j\boldsymbol{\delta}_{(k-1)D+i}^{(D^2)} - j\boldsymbol{\delta}_{(i-1)D+k}^{(D^2)}, \quad (135)$$

respectively, for $i > k$. One can check that $\mathbf{T}$ is invertible. In view of (43), the CRB (34) exists if and only if $[\mathbf{G}\ \boldsymbol{\Delta}_1]$ has full column rank. This proves Theorem 2 when we substitute (126), (127), and (134) into (44) and (45) and note that $(\mathbf{R}_{\mathbf{yy}}^T \otimes \mathbf{R}_{\mathbf{yy}})^{-1/2}$ and $\mathbf{T}$ are invertible. Besides, (48) becomes

$$\mathrm{CRB}_{\mathrm{unb}}(\boldsymbol{\omega}) = \frac{1}{K}\left(\mathbf{G}^H \boldsymbol{\Pi}_{\boldsymbol{\Delta}_1}^{\perp} \mathbf{G}\right)^{-1}, \quad (136)$$

where $\mathbf{G}$ is as in (50), and

$$\boldsymbol{\Delta}_1 = \left(\mathbf{R}_{\mathbf{yy}}^T \otimes \mathbf{R}_{\mathbf{yy}}\right)^{-\frac{1}{2}} \cdot \left[[(\mathbf{A}_L\mathbf{D}_H)^* \otimes (\mathbf{A}_L\mathbf{D}_H)]\mathbf{T} \quad \mathrm{vec}(\mathbf{H}\mathbf{H}^H)\right]. \quad (137)$$

We finally obtain (49) by noting that $\boldsymbol{\Pi}_{\boldsymbol{\Delta}_1}^{\perp} = \boldsymbol{\Pi}_{\boldsymbol{\Delta}}^{\perp}$ since $\mathbf{T}$ is invertible and does not alter the column space.

References

[1] P.-C. Chen and P. P. Vaidyanathan, "Convolutional beamspace for linear arrays," *IEEE Trans. Signal Process.*, vol. 68, pp. 5395–5410, 2020.

[2] P. P. Vaidyanathan and P.-C. Chen, "Convolutional beamspace for array signal processing," in *IEEE Intl. Conf. Acoust., Speech, and Signal Process.*, 2020.

[3] H. B. Lee and M. S. Wengrovitz, "Resolution threshold of beamspace MUSIC for two closely spaced emitters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 9, pp. 1545–1559, Sep. 1990.

[4] H. L. Van Trees, *Optimum array processing*. John Wiley & Sons, Inc., New York, 2002.

[5] X. L. Xu and K. M. Buckley, "Statistical performance comparison of MUSIC in element-space and beam-space," in *Proc. IEEE Intl. Conf. Acoust., Speech, and Signal Process.*, 1989, pp. 2124–2127.

[6] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beamspace root-MUSIC," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 344–364, Jan. 1993.

[7] G. Xu, S. D. Silverstein, R. H. Roy, and T. Kailath, "Beamspace ESPRIT," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 349–356, Feb. 1994.

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

[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. P. Vaidyanathan, *Multirate systems and filter banks*. Prentice Hall, Englewood Cliffs, N.J., 1993.

[11] R. R. Pote and B. D. Rao, "Reduced dimension beamspace design incorporating nested array for mmWave channel estimation," in *2019 Asilomar Conf. on Signal, Syst., Comput.*, 2019, pp. 1212–1216.

[12] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[13] T. W. Anderson, "Asymptotic theory for principal component analysis," *The Annals of Mathematical Statistics*, vol. 34, no. 1, pp. 122–148, 1963.

[14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., 1993.

[15] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 2012.

[17] 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.

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

[19] G. Bienvenu and L. Kopp, "Decreasing high resolution method sensitivity by conventional beamformer preprocessing," in *Proc. IEEE Intl. Conf. Acoust., Speech, and Signal Process.*, Mar. 1984, pp. 714–717.

[20] K. M. Buckley and X. L. Xu, "Reduced-dimension beam-space broadband source localization: preprocessor design," in *Proc. SPIE, Adv. Algorithms and Architectures for Signal Process. III*, Feb. 1988, pp. 368–376.

[21] H. B. Lee and M. S. Wengrovitz, "Improved high-resolution direction-finding through use of homogeneous constraints," in *Fourth Annual ASSP Workshop on Spec. Est. and Modeling*, Aug. 1988, pp. 152–157.

[22] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-D beamspace MUSIC method," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5384–5394, Aug. 2017.

[23] H. Zhao, N. Zhang, and Y. Shen, "Robust beamspace design for direct localization," in *IEEE Intl. Conf. Acoust., Speech and Signal Process.*, May 2019, pp. 4360–4364.

[24] F. Rellich and J. Berkowitz, *Perturbation theory of eigenvalue problems*. CRC Press, 1969.

[25] B. Champagne, "Adaptive eigendecomposition of data covariance matrices based on first-order perturbations," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2758–2770, 1994.

[26] A. J. Weiss and B. Friedlander, "Preprocessing for direction finding with minimal variance degradation," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1478–1485, 1994.

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

[28] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[29] P. P. Vaidyanathan and T. Q. 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.

[30] C. Fassino, G. Pistone, and M. P. Rogantin, "Computing the moments of the complex gaussian: Full and sparse covariance matrix," *Mathematics*, vol. 7, no. 3, p. 263, 2019.

[31] S. Janson, *Gaussian Hilbert Spaces*. Cambridge university press, 1997, no. 129.