

Performance Analysis of Multivariate Complex Amplitude Estimators

Luzhou Xu and Jian Li, *Senior Member, IEEE*

Abstract—We consider multivariate complex amplitude estimation in the presence of unknown interference and noise. Two multivariate approaches [Maximum Likelihood (ML) and Capon] are provided. We derive the closed-form expression of the Cramér–Rao bound (CRB) for the unknown complex amplitudes. We also analyze the bias properties and Mean Squared Errors (MSE) of the two estimators. A comparative study shows that the multivariate ML estimator is unbiased, whereas the multivariate Capon estimator is biased downward for finite snapshots. Both estimators are asymptotically statistically efficient when the number of snapshots is large.

Index Terms—Capon, complex wishart, Cramér–Rao bound, growth curve, ML, multivariate parameter estimation.

I. INTRODUCTION

ESTIMATING unknown signal parameters in the presence of unknown interference and noise via array processing is an important common problem in signal processing. It is well-known that the signal temporal information can be utilized to effectively suppress the interference and noise and hence improve the estimation accuracy (see, e.g., [1]–[10]). Recently, the complex amplitude estimation in the known waveform and steering vector case was studied in [11], where both the temporal and spatial information of a single signal are known and exploited for interference suppression. In [11], both Capon and Maximum Likelihood (ML) estimators are considered. Through theoretical and numerical analyses, it has been shown that both estimators are asymptotically statistically efficient for large number of snapshots, although Capon is biased downward for finite snapshots, while ML is unbiased. The focus of [11] is on the estimation of a scalar parameter, i.e., the univariate amplitude estimation problem. However, in some practical applications, we need to estimate a matrix rather than a scalar parameter, and, hence, multivariate parameter estimators should be considered [12].

In this paper, we consider the following multivariate complex amplitude estimation problem:

$$\mathbf{X} = \mathbf{A}\mathbf{B}\mathbf{S} + \mathbf{Z}. \quad (1)$$

In (1), $\mathbf{X} \in C^{M \times L}$ denotes the observed snapshots with L being the number of snapshots. The columns in $\mathbf{A} \in C^{M \times N}$ are the

known linearly independent spatial vectors, e.g., steering vectors. The rows in $\mathbf{S} \in C^{K \times L}$ are the known temporal vectors, e.g., waveforms, assumed to be linearly independent of each other or not completely correlated with each other. The matrix $\mathbf{B} \in C^{N \times K}$ contains the multivariate unknown complex amplitudes. Throughout this paper, we assume that $M \geq N$ and $L \geq K + M$. The columns of the interference and noise matrix $\mathbf{Z} \in C^{M \times L}$ are statistically independent circularly symmetric complex Gaussian random vectors with zero-mean and unknown covariance matrix \mathbf{Q} . The problem of interest is to estimate the unknown matrix \mathbf{B} .

We note that the data model of (1) has general applications. Its real-valued counterpart, which is called the Growth-Curve Model, has been studied and used widely for investigating growth problems in the statistics field [13]–[15]. This real-valued Growth-Curve model was extended and introduced to the signal processing field in [12]. Using the extended model, the authors in [12] unified many existing algorithms proposed for radar array processing (e.g., [16], [17]), spectral analysis (e.g., [18], [19]), and wireless communication (e.g., [20]–[23]) applications.

The focus of this paper is on the performance analysis of the multivariate ML and Capon estimators for the data model in (1). We derive the closed-form expression of the CRB of the unknown complex amplitude parameters. We also analyze the bias properties and MSE of the two estimators. A comparative study shows that the multivariate ML estimator is unbiased whereas the multivariate Capon estimator is biased downward for finite snapshots. Yet in finite data samples and at low SNR, Capon can provide a smaller MSE than ML. Both estimators are asymptotically statistically efficient when the number of snapshots is large.

The remainder of the paper is organized as follows. Section II provides the multivariate Capon and ML estimators. Section III gives the performance analysis of the two estimators and the CRB of the unknown complex amplitudes. Numerical examples are provided in Section IV. Finally, we present our conclusions in Section V.

II. MULTIVARIATE PARAMETER ESTIMATION

Based on the data model in (1), we describe the multivariate Capon and ML estimators in this section.

A. Multivariate Capon Estimation

The multivariate Capon estimator consists of two main steps. The first is the Capon beamforming [24]–[26]. The other is the Least-Squares (LS) estimation [27], [28], which is basically the matched filtering.

Manuscript received April 22, 2004; revised October 15, 2004. This work was supported in part by the National Science Foundation under Grants CCR-0104887 and CCR-0097114. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jan C. de Munck.

The authors are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: li@dsp.ufl.edu).

Digital Object Identifier 10.1109/TSP.2005.851112

We first consider the Capon beamforming. Let

$$\mathbf{R} = \mathbf{X}\mathbf{X}^H \quad (2)$$

where $(\cdot)^H$ denotes the conjugate transpose of a matrix. Then, the Capon beamformer can be formulated as follows:

$$\hat{\mathbf{W}} = \arg \min_{\mathbf{W}} \text{tr}(\mathbf{W}^H \mathbf{R} \mathbf{W}), \quad \text{subject to } \mathbf{W}^H \mathbf{A} = \mathbf{I} \quad (3)$$

where $\hat{\mathbf{W}}$ is a multivariate weighting matrix for noise and interference suppression while keeping the desired signals undistorted. Solving the above optimization problem yields

$$\hat{\mathbf{W}} = \mathbf{R}^{-1} \mathbf{A} (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})^{-1}. \quad (4)$$

Note that since $M \geq N$ and the columns in \mathbf{A} are linearly independent of each other, $\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A}$ has full rank N with probability one.

The beamforming output, which is denoted by \mathbf{Y} , is

$$\mathbf{Y} = \hat{\mathbf{W}}^H \mathbf{X} = (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{R}^{-1} \mathbf{X}. \quad (5)$$

Now we consider the LS estimation. Substituting (1) into (5) yields

$$\mathbf{Y} = \mathbf{B}\mathbf{S} + (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{R}^{-1} \mathbf{Z}. \quad (6)$$

Estimating \mathbf{B} from \mathbf{Y} based on (6) is a standard Multivariate Analysis of Variance (MANOVA) problem [12]–[14]. Note that after spatial beamforming, the noise vectors remain temporally white, and hence the LS estimator gives the best performance. Using the LS algorithm yields

$$\hat{\mathbf{B}}_{\text{Capon}} = \mathbf{Y}\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}. \quad (7)$$

Substituting (5) into (7), the multivariate Capon estimator has the form

$$\hat{\mathbf{B}}_{\text{Capon}} = (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{R}^{-1} \mathbf{X}\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}. \quad (8)$$

Note that the Capon estimator for the univariate case in [11] is a special case of (8).

B. Multivariate Maximum Likelihood Estimation

A general derivation of the multivariate ML estimator has been given in [12]. In our data model, \mathbf{A} and \mathbf{S} are both full rank, and hence, the multivariate ML estimator can be simplified as follows:

$$\hat{\mathbf{B}}_{\text{ML}} = (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1} \mathbf{X}\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1} \quad (9)$$

where

$$\mathbf{T} = \mathbf{X}[\mathbf{I} - \mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}\mathbf{S}]\mathbf{X}^H. \quad (10)$$

Note again that the ML estimator for the univariate case in [11] is a special case of (9).

To better understand the above ML estimator intuitively, we insert (1) into (10) and get

$$\mathbf{T} = \mathbf{Z}[\mathbf{I} - \mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}\mathbf{S}]\mathbf{Z}^H. \quad (11)$$

It shows that $(1/L)\mathbf{T}$ is an estimate of the unknown noise covariance \mathbf{Q} . We also note that the estimator in (9) can be divided into two steps, including the ML beamforming spatially, corresponding to the left-multiplication matrix $(\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1}$, and the LS estimation temporally, corresponding to the right-multiplication matrix $\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}$.

Note that like the univariate case in [11], the only difference between the Capon and ML estimators is that the matrix \mathbf{R} in (8) is replaced by \mathbf{T} in (9). However, as we will show in the following analysis, this seemingly minor difference in fact leads to significant and interesting performance differences between the two estimators.

III. PERFORMANCE ANALYSIS

A. Performance Analysis of the Multivariate ML Estimator

The statistical performance analysis of the multivariate ML estimator has been considered in [15], where the variables are all restricted to be real-valued. Using the similar technique, we can extend the conclusions in [11] and [15] to the complex-valued and multivariate cases, respectively. Due to space limitations, we present below Theorem 1 without the detailed proof.

Theorem 1: For the data model in (1), the multivariate ML estimate of \mathbf{B} , given by (9), is unbiased and asymptotically statistically efficient for large number of data samples. Its MSE matrix can be expressed as

$$\begin{aligned} \text{MSE}(\hat{\mathbf{B}}_{\text{ML}}) &\triangleq \text{E}[\text{vec}(\hat{\mathbf{B}}_{\text{ML}})\text{vec}(\hat{\mathbf{B}}_{\text{ML}})^H] \\ &= \frac{L-K}{L-K-M+N} \text{CRB}(\mathbf{B}) \end{aligned} \quad (12)$$

where

$$\text{CRB}(\mathbf{B}) = (\mathbf{S}^* \mathbf{S}^T)^{-1} \otimes (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1} \quad (13)$$

$\text{vec}(\cdot)$, $(\cdot)^*$, $(\cdot)^T$ and \otimes denote the direct operator (stacking the columns of a matrix on top of each other), complex conjugate, transpose and Kronecker product of matrices, respectively.

In (12) and (13), $\text{CRB}(\mathbf{B})$ is the CRB of \mathbf{B} , which represents the best possible performance bound for any unbiased estimator. The detailed derivation and discussion on $\text{CRB}(\mathbf{B})$ are given in Appendix A.

It should be noted that, in addition to the complex conjugate, the coefficient in (12) is also different from its real-valued counterpart (see [15, Th. 3.3]). It is due to the fact that the lemma on the mean of real-valued inverted Wishart matrices (see [15, Lemma 3.3] or [13, Th. 2.4.6]) is not directly applicable to the complex-valued case. It should be revised as follows.

Lemma 1: Let \mathbf{Y} be an $m \times m$ random matrix with the complex Wishart distribution with covariance matrix $\Sigma_{m \times m}$ and

l ($l \geq m$) degrees of freedom [29]–[31], denoted by $\Upsilon \sim \text{CW}(l, m; \Sigma)$. Then

$$E(\Upsilon^{-1}) = \frac{\Sigma^{-1}}{(l-m)}. \quad (14)$$

From (12), we note that $\text{MSE}(\hat{\mathbf{B}}_{\text{ML}})$ approaches $\text{CRB}(\mathbf{B})$ for large L , which means that the multivariate ML estimator is asymptotically statistically efficient for large number of snapshots L . Hence, the efficiency condition for the univariate case in [11] can be extended to the multivariate case as well.

Furthermore, the CRB of \mathbf{B} depends on $(\mathbf{S}^* \mathbf{S}^T)^{-1}$ and $(\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1}$. As we show in Appendix A, orthogonalities among the rows of \mathbf{S} and among the columns of $\mathbf{Q}^{-(1/2)} \mathbf{A}$ lead to small diagonal elements for $(\mathbf{S}^* \mathbf{S}^T)^{-1}$ and $(\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1}$, respectively, which in turn reduce the CRB.

We also note that when $M = N$, which implies that \mathbf{A} is a square matrix, the multivariate ML estimator is efficient. However, we should not think of it as a significant advantage to make N as large as possible. As we show in Appendix A, in the case that the columns of $\mathbf{Q}^{-(1/2)} \mathbf{A}$ are not orthogonal to each other, which often happens in practice, large N causes CRB to increase.

B. Performance Analysis of the Multivariate Capon Estimator

We now establish the theoretical properties of the multivariate Capon estimator.

1) *Bias Analysis:* In Section III-A, we know that the multivariate ML estimator is unbiased. We will investigate the bias of the multivariate Capon estimator by studying the relationship between the two estimators.

By (1) and (9), we know that the error of the multivariate ML estimate of \mathbf{B} is

$$\begin{aligned} \Delta \mathbf{B} &\triangleq \hat{\mathbf{B}}_{\text{ML}} - \mathbf{B} \\ &= (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1} \mathbf{Z}_S (\mathbf{S} \mathbf{S}^H)^{-1} \end{aligned} \quad (15)$$

where $\mathbf{Z}_S = \mathbf{Z} \mathbf{S}^H$.

On the other hand, comparing (2) and (10), we note that

$$\mathbf{R} = \mathbf{T} + \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{S}^H)^{-1} \mathbf{S} \mathbf{X}^H. \quad (16)$$

Applying the matrix inversion lemma gives

$$\begin{aligned} \mathbf{A}^H \mathbf{R}^{-1} \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{S}^H)^{-1} \\ = \mathbf{A}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H + \mathbf{S} \mathbf{S}^H)^{-1} \end{aligned} \quad (17)$$

and

$$\begin{aligned} (\mathbf{A}^H \mathbf{R}^{-1} \mathbf{A})^{-1} \\ = [\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A} - \mathbf{A}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H \\ + \mathbf{S} \mathbf{S}^H)^{-1} \mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{A}]^{-1} \\ = (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} - (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H \\ \times [\mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{A} (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H \\ - \mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H - \mathbf{S} \mathbf{S}^H]^{-1} \\ \times \mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{A} (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1}. \end{aligned} \quad (18)$$

Substituting (17) and (18) into (8), and after some straightforward manipulations, we get

$$\hat{\mathbf{B}}_{\text{Capon}} = \hat{\mathbf{B}}_{\text{ML}} \mathbf{A} \quad (19)$$

where

$$\mathbf{A} = [\mathbf{I} + \mathbf{V}]^{-1} \quad (20)$$

with

$$\begin{aligned} \mathbf{V} &= \mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{S}^H)^{-1} - \mathbf{S} \mathbf{X}^H \mathbf{T}^{-1} \\ &\times \mathbf{A} (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1} \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{S}^H)^{-1}. \end{aligned} \quad (21)$$

Then, inserting (1) into (21) gives

$$\mathbf{V} = \mathbf{Z}_S^H [\mathbf{T}^{-1} - \mathbf{T}^{-1} \mathbf{A} (\mathbf{A}^H \mathbf{T}^{-1} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{T}^{-1}] \mathbf{Z}_S (\mathbf{S} \mathbf{S}^H)^{-1}. \quad (22)$$

Note that there are two random matrices, i.e., \mathbf{T} and \mathbf{Z}_S , in (15) and (22). Since the columns of \mathbf{Z} are statistically independent zero-mean Gaussian random vectors while the columns of \mathbf{S}^H are orthogonal to those of $\mathbf{I} - \mathbf{S}^H (\mathbf{S} \mathbf{S}^H) \mathbf{S}$, by the property of joint Gaussian distribution, we know that \mathbf{Z}_S and $\mathbf{Z}_S^\perp \triangleq \mathbf{Z} [\mathbf{I} - \mathbf{S}^H (\mathbf{S} \mathbf{S}^H) \mathbf{S}]$ are two independent Gaussian random matrices. Hence, \mathbf{Z}_S and $\mathbf{T} = \mathbf{Z}_S^\perp (\mathbf{Z}_S^\perp)^H$ are also independent of each other. By [15, Lemmas 1.9 and 1.11], which can be readily extended to the complex-valued case, we have $\mathbf{Z}_S \sim \text{CN}(\mathbf{0}; \mathbf{Q}, \mathbf{S} \mathbf{S}^H)$ and $\mathbf{T} \sim \text{CW}(L-K, M; \mathbf{Q})$ (see [15, (2.38)]).

Since \mathbf{Z}_S and \mathbf{T} are statistically independent of each other and by (15) and (22), we know that $(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}$ is an odd function with respect to \mathbf{Z}_S . Hence, replacing \mathbf{Z}_S with $-\mathbf{Z}_S$ yields

$$E[(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}] |_{\mathbf{Z}_S = -\mathbf{Z}_S} = -E[(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}]. \quad (23)$$

On the other hand, since \mathbf{Z}_S is a zero-mean Gaussian random matrix, $-\mathbf{Z}_S$, as a random matrix transformed from \mathbf{Z}_S , retains all the statistical properties of \mathbf{Z}_S . Hence, replacing \mathbf{Z}_S by $-\mathbf{Z}_S$ will not change the expectation of $(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}$, i.e.,

$$E[(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}] |_{\mathbf{Z}_S = -\mathbf{Z}_S} = E[(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}]. \quad (24)$$

It follows from (23) and (24) that

$$E[(\hat{\mathbf{B}}_{\text{ML}} - \mathbf{B}) \mathbf{A}] = \mathbf{0}. \quad (25)$$

Therefore, by (19) and (25), we have

$$E(\hat{\mathbf{B}}_{\text{Capon}}) = \mathbf{B} E(\mathbf{A}). \quad (26)$$

To calculate the expectation of \mathbf{A} , we use the same technique in [15] to simplify \mathbf{A} and \mathbf{V} via the transformations of random matrices. Let

$$\tilde{\mathbf{T}} \triangleq \mathbf{Q}^{-\frac{1}{2}} \mathbf{T} \mathbf{Q}^{-\frac{1}{2}} \quad (27)$$

and

$$\tilde{\mathbf{Z}}_S = \mathbf{Q}^{-\frac{1}{2}} \mathbf{Z}_S (\mathbf{S} \mathbf{S}^H)^{-\frac{1}{2}} \quad (28)$$

which obviously have the $\text{CW}(L-K, M; \mathbf{I})$ and $\text{CN}(\mathbf{0}; \mathbf{I}_M, \mathbf{I}_K)$ distributions, respectively. Furthermore, denote

$$\tilde{\mathbf{A}} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{A} (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-\frac{1}{2}}. \quad (29)$$

Then, we get

$$\begin{aligned} \mathbf{V} &= (\mathbf{S} \mathbf{S}^H)^{\frac{1}{2}} \tilde{\mathbf{Z}}_S^H [\tilde{\mathbf{T}}^{-1} \\ &- \tilde{\mathbf{T}}^{-1} \tilde{\mathbf{A}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{T}}^{-1} \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H \tilde{\mathbf{T}}^{-1}] \tilde{\mathbf{Z}}_S (\mathbf{S} \mathbf{S}^H)^{-\frac{1}{2}}. \end{aligned} \quad (30)$$

Since $\tilde{\mathbf{A}}^H \tilde{\mathbf{A}} = \mathbf{I}$, we can decompose $\tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{A}} = \mathbf{U}\mathbf{P} \quad (31)$$

where \mathbf{U} is an $M \times M$ unitary matrix with its first N columns being $\tilde{\mathbf{A}}$; $\mathbf{P}_{M \times N} = [\mathbf{I}_N \quad \mathbf{0}]^T$. Since \mathbf{U} is unitary, we have

$$\mathbf{\Gamma} \triangleq \mathbf{U}^H \tilde{\mathbf{T}} \mathbf{U} \sim \text{CW}(L - K, M; \mathbf{I}) \quad (32)$$

and

$$\mathbf{\Xi} \triangleq \mathbf{U}^H \tilde{\mathbf{Z}}_S \sim \text{CN}(0; \mathbf{I}_M, \mathbf{I}_K). \quad (33)$$

We next partition $\mathbf{\Xi}_{M \times K}$, $\mathbf{\Gamma}_{M \times M}$, and $\mathbf{\Gamma}_{M \times M}^{-1}$, respectively, as follows:

$$\begin{aligned} \mathbf{\Xi} &= \begin{bmatrix} \mathbf{\Xi}_1 \\ \mathbf{\Xi}_2 \end{bmatrix}, \quad \mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} \\ \mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} \end{bmatrix}, \quad \text{and} \\ \mathbf{\Gamma}^{-1} &= \begin{bmatrix} \mathbf{\Gamma}^{11} & \mathbf{\Gamma}^{12} \\ \mathbf{\Gamma}^{21} & \mathbf{\Gamma}^{22} \end{bmatrix} \end{aligned} \quad (34)$$

where $\mathbf{\Xi}_1$ and $\mathbf{\Xi}_2$ are $N \times K$ and $(M - N) \times K$, respectively, and both $\mathbf{\Gamma}_{11}$ and $\mathbf{\Gamma}^{11}$ are $N \times N$ matrices.

Inserting (31) to (34) into (30) gives

$$\begin{aligned} \mathbf{V} &= (\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}} \mathbf{\Xi}^H [\mathbf{\Gamma}^{-1} - \mathbf{\Gamma}^{-1} \mathbf{P} (\mathbf{P}^H \mathbf{\Gamma}^{-1} \mathbf{P})^{-1} \mathbf{P}^H \mathbf{\Gamma}^{-1}] \\ &\quad \times \mathbf{\Xi} (\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}} \\ &= (\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}} \mathbf{\Xi}^H \left\{ \begin{bmatrix} \mathbf{\Gamma}^{11} & \mathbf{\Gamma}^{12} \\ \mathbf{\Gamma}^{21} & \mathbf{\Gamma}^{22} \end{bmatrix} \right. \\ &\quad \left. - \begin{bmatrix} \mathbf{\Gamma}^{11} & \mathbf{\Gamma}^{12} \\ \mathbf{\Gamma}^{21} & \mathbf{\Gamma}^{21} (\mathbf{\Gamma}^{11})^{-1} \mathbf{\Gamma}^{12} \end{bmatrix} \right\} \mathbf{\Xi} (\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}} \\ &= (\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}} \mathbf{\Xi}_2^H \mathbf{\Gamma}_{22}^{-1} \mathbf{\Xi}_2 (\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}}. \end{aligned} \quad (35)$$

From (20) and (35) and by the matrix inversion lemma, it follows that

$$\begin{aligned} \mathbf{\Lambda} &= (\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}} [\mathbf{I} + \mathbf{\Xi}_2^H \mathbf{\Gamma}_{22}^{-1} \mathbf{\Xi}_2]^{-1} (\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}} \\ &= \mathbf{I} - (\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}} \mathbf{\Xi}_2^H (\mathbf{\Gamma}_{22} + \mathbf{\Xi}_2 \mathbf{\Xi}_2^H)^{-1} \mathbf{\Xi}_2 (\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}}. \end{aligned} \quad (36)$$

To calculate the expectation of $\mathbf{\Lambda}$, we use the following lemma.

Lemma 2: Let $\mathbf{\Upsilon}_{n \times p}$ and $\mathbf{\Psi}_{n \times n}$ be two independent random matrices, and

$$\mathbf{\Upsilon} \sim \text{CN}(0; \mathbf{I}_n, \mathbf{I}_p), \quad \mathbf{\Psi} \sim \text{CW}(l, n; \mathbf{I}_n). \quad (37)$$

Denote $\mathbf{\Pi} = \mathbf{\Upsilon}^H (\mathbf{\Psi} + \mathbf{\Upsilon} \mathbf{\Upsilon}^H)^{-1} \mathbf{\Upsilon}$. Then the expectation and correlation matrices of the random matrix $\mathbf{\Pi}$ are as follows:

$$E(\mathbf{\Pi}) = \frac{n}{l+p} \mathbf{I}_p \quad (38)$$

$$\begin{aligned} E(\text{vec}(\mathbf{\Pi}) \text{vec}(\mathbf{\Pi})^H) &= \frac{n(n+1)}{(l+p)(l+p+1)} \text{vec}(\mathbf{I}_p) \\ &\quad \times \text{vec}(\mathbf{I}_p)^H + \eta \mathbf{D}_p, \end{aligned} \quad (39)$$

where η is a scalar and approximately equal to $(n(l+p-n))/((l+p)^2(l+p+1))$ for large $l+p$, and \mathbf{D}_p is a $p^2 \times p^2$ matrix with its element at the $[(c_1 - 1)p + r_1]$ th row and the $[(c_2 - 1)p + r_2]$ th column ($r_1, c_1, r_2, c_2 = 1, 2, \dots, p$) being

$$d_{(c_1-1)p+r_1, (c_2-1)p+r_2} = \begin{cases} 1, & \text{when } r_1 = r_2, c_1 = c_2 \text{ but } r_1 \neq c_1 \\ -1, & \text{when } r_1 = c_1, r_2 = c_2 \text{ but } r_1 \neq r_2 \\ 0, & \text{otherwise.} \end{cases} \quad (40)$$

Proof: See Appendix B.

Applying the above lemma to $\mathbf{\Xi}$ and $\mathbf{\Gamma}$, which by construction satisfy the assumptions in the lemma, we have immediately

$$E(\mathbf{\Lambda}) = \left(1 - \frac{M-N}{L}\right) \mathbf{I}. \quad (41)$$

Inserting (41) into (26), we get

$$E(\hat{\mathbf{B}}_{\text{Capon}}) = \left(1 - \frac{M-N}{L}\right) \mathbf{B}. \quad (42)$$

The above equation shows that the multivariate Capon estimator shares the same properties as the univariate Capon in [11]. In other words, it is biased downward for finite snapshot number L . However, for large L , it is asymptotically unbiased. It is also worth noting that the bias of the Capon estimator is not related to K , which means that increasing the number of rows in the temporal information matrix \mathbf{S} will not cause higher bias.

Moreover, we note that when $M = N$, the multivariate Capon estimator becomes unbiased as the multivariate ML estimator. In this case, both ML and Capon reduce to the same estimator $\mathbf{A}^{-1} \mathbf{X} \mathbf{S}^H (\mathbf{S} \mathbf{S}^H)^{-1}$. Hence, for the same reason that we have stated in Section III-A, this unbiasedness of the multivariate Capon estimator should not be seen as a significant advantage. \square

2) *Mean-Squared-Error Analysis:* We investigate the MSE of the multivariate Capon estimator below. Using the same technique to obtain (25), we can prove that $\Delta \mathbf{B} \mathbf{\Lambda}$ and $\mathbf{B}(\mathbf{I} - \mathbf{\Lambda})$ are uncorrelated. Hence, from (15) and (19), we have

$$\begin{aligned} \text{MSE}(\hat{\mathbf{B}}_{\text{Capon}}) &\triangleq E[\text{vec}(\hat{\mathbf{B}}_{\text{Capon}} - \mathbf{B}) \text{vec}(\hat{\mathbf{B}}_{\text{Capon}} - \mathbf{B})^H] \\ &= E\{\text{vec}[\Delta \mathbf{B} \mathbf{\Lambda} - \mathbf{B}(\mathbf{I} - \mathbf{\Lambda})] \text{vec}[\Delta \mathbf{B} \mathbf{\Lambda} - \mathbf{B}(\mathbf{I} - \mathbf{\Lambda})]^H\} \\ &= E[\text{vec}(\Delta \mathbf{B} \mathbf{\Lambda}) \text{vec}(\Delta \mathbf{B} \mathbf{\Lambda})^H] \\ &\quad + E[\text{vec}(\mathbf{B}(\mathbf{I} - \mathbf{\Lambda})) \text{vec}(\mathbf{B}(\mathbf{I} - \mathbf{\Lambda}))^H]. \end{aligned} \quad (43)$$

We first calculate $E[\text{vec}(\Delta \mathbf{B} \mathbf{\Lambda}) \text{vec}(\Delta \mathbf{B} \mathbf{\Lambda})^H]$.

Following the definitions in (27), (28) and (29) and inserting them into (15), we get

$$\begin{aligned} \Delta \mathbf{B} &= (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-\frac{1}{2}} (\tilde{\mathbf{A}}^H \tilde{\mathbf{T}}^{-1} \tilde{\mathbf{A}})^{-1} \\ &\quad \times \tilde{\mathbf{A}}^H \tilde{\mathbf{T}}^{-1} \tilde{\mathbf{Z}}_S (\mathbf{S} \mathbf{S}^H)^{-\frac{1}{2}}. \end{aligned} \quad (44)$$

Then, we adopt the decomposition in (31), the definitions in (32) and (33), and the partitions in (34), and insert them into (30). By the inversion lemma of partitioned matrices

$$\begin{aligned}\Delta\mathbf{B} &= (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-\frac{1}{2}}(\mathbf{P}^H\mathbf{\Gamma}^{-1}\mathbf{P})^{-1}\mathbf{P}^H\mathbf{\Gamma}^{-1}\mathbf{\Xi}(\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}} \\ &= (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-\frac{1}{2}}(\mathbf{\Xi}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)(\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}}.\end{aligned}\quad (45)$$

By (45) and (36), we have

$$\begin{aligned}\Delta\mathbf{B}\mathbf{A} &= (\mathbf{A}^H\mathbf{Q}^{-1}\mathbf{A})^{-\frac{1}{2}}[\mathbf{\Xi}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2] \\ &\quad \times [\mathbf{I} + \mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2]^{-1}(\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}}.\end{aligned}\quad (46)$$

Using the fact that $\text{vec}(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})\text{vec}(\mathbf{Y})$ as well as (13) yields

$$E[\text{vec}(\Delta\mathbf{B}\mathbf{A})\text{vec}(\Delta\mathbf{B}\mathbf{A})^H] = [\text{CRB}(\mathbf{B})]^{-\frac{1}{2}}\mathbf{F}[\text{CRB}(\mathbf{B})]^{-\frac{1}{2}}\quad (47)$$

where

$$\begin{aligned}\mathbf{F} &= E\left\{\text{vec}\left[(\mathbf{\Xi}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)(\mathbf{I} + \mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^{-1}\right] \right. \\ &\quad \left. \times \text{vec}\left[(\mathbf{\Xi}_1 - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)(\mathbf{I} + \mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^{-1}\right]^H\right\}.\end{aligned}\quad (48)$$

Then, using the lemma that $\text{vec}(\mathbf{XY}) = (\mathbf{Y}^T \otimes \mathbf{I})\text{vec}(\mathbf{X})$ and the fact that $\mathbf{\Xi}_1$ and $\mathbf{\Xi}_2$ are independent standard matrix-variate Gaussian distributions, after some manipulations, we get

$$\begin{aligned}\mathbf{F} &= E\left\{\left[(\mathbf{I} + (\mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^*)^{-1} \otimes \mathbf{I}\right] \right. \\ &\quad \times [\mathbf{I} + E_{\mathbf{\Gamma}_{12}|\mathbf{\Gamma}_{22}}(\text{vec}(\mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)\text{vec}(\mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^H)] \\ &\quad \left. \times [(\mathbf{I} + (\mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^*)^{-1} \otimes \mathbf{I}]\right\}.\end{aligned}\quad (49)$$

Note that

$$\begin{aligned}E_{\mathbf{\Gamma}_{12}|\mathbf{\Gamma}_{22}}\left[\text{vec}(\mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)\text{vec}(\mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^H\right] \\ &= [(\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^T \otimes \mathbf{I}]E_{\mathbf{\Gamma}_{12}|\mathbf{\Gamma}_{22}}[\text{vec}(\mathbf{\Gamma}_{12})\text{vec}(\mathbf{\Gamma}_{12})^H] \\ &\quad \times [(\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^* \otimes \mathbf{I}] \\ &= [(\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^T \otimes \mathbf{I}][\mathbf{\Gamma}_{22}^* \otimes \mathbf{I}][(\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^* \otimes \mathbf{I}] \\ &= (\mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^* \otimes \mathbf{I}.\end{aligned}\quad (50)$$

To get the above equation, we have utilized [30, Lemma 2], i.e., $\mathbf{\Gamma}_{12}|\mathbf{\Gamma}_{22} \sim \text{CN}(\mathbf{0}; \mathbf{I}, \mathbf{\Gamma}_{22})$. Hence, by the complex-valued counterpart of [15, Lemma 1.8], we know that the covariance matrix of $\text{vec}(\mathbf{\Gamma}_{12})$ given $\mathbf{\Gamma}_{22}$ is $\mathbf{\Gamma}_{22}^* \otimes \mathbf{I}$.

Inserting (50) into (49) and recalling (36) and (41) yield

$$\begin{aligned}\mathbf{F} &= E\left\{\left[(\mathbf{I} + (\mathbf{\Xi}_2^H\mathbf{\Gamma}_{22}^{-1}\mathbf{\Xi}_2)^*)^{-1}\right] \otimes \mathbf{I} \right. \\ &= [(\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}}E(\mathbf{A}^*)(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}] \otimes \mathbf{I} \\ &= \left(1 - \frac{M-N}{L}\right)\mathbf{I}.\end{aligned}\quad (51)$$

From (47) and (51), the equation below follows directly:

$$E[\text{vec}(\Delta\mathbf{B}\mathbf{A})\text{vec}(\Delta\mathbf{B}\mathbf{A})^H] = \frac{L-M+N}{L}\text{CRB}(\mathbf{B}).\quad (52)$$

Now, we consider the second term in (43). By (36), we know that

$$\mathbf{B}(\mathbf{I} - \mathbf{\Lambda}) = \mathbf{B}(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}\mathbf{\Xi}_2^H(\mathbf{\Gamma}_{22} + \mathbf{\Xi}_2\mathbf{\Xi}_2^H)^{-1}\mathbf{\Xi}_2(\mathbf{S}\mathbf{S}^H)^{-\frac{1}{2}}.\quad (53)$$

We know that $\mathbf{\Xi}_2$ and $\mathbf{\Gamma}_{22}$ are independent of each other with $\text{CN}(\mathbf{0}; \mathbf{I}_{M-N}, \mathbf{I}_K)$ and $\text{CW}(L-K, M-N; \mathbf{I}_{M-N})$ distributions, respectively. Then using the fact that $\text{vec}(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})\text{vec}(\mathbf{Y})$, the following equation is obtained following Lemma 2:

$$\begin{aligned}E\{\text{vec}[\mathbf{B}(\mathbf{I} - \mathbf{\Lambda})]\text{vec}[\mathbf{B}(\mathbf{I} - \mathbf{\Lambda})]^H\} \\ &= \{(\mathbf{S}^*\mathbf{S}^T)^{-\frac{1}{2}} \otimes [\mathbf{B}(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}]\} \\ &\quad \times E\left\{\text{vec}\left[\mathbf{\Xi}_2^H(\mathbf{\Gamma}_{22} + \mathbf{\Xi}_2\mathbf{\Xi}_2^H)^{-1}\mathbf{\Xi}_2\right] \right. \\ &\quad \left. \times \text{vec}\left[\mathbf{\Xi}_2^H(\mathbf{\Gamma}_{22} + \mathbf{\Xi}_2\mathbf{\Xi}_2^H)^{-1}\mathbf{\Xi}_2\right]^H\right\} \\ &\quad \times \{(\mathbf{S}^*\mathbf{S}^T)^{-\frac{1}{2}} \otimes [\mathbf{B}(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}]\}^H \\ &= \frac{(M-N)(M-N+1)}{L(L+1)}\text{vec}(\mathbf{B})\text{vec}(\mathbf{B})^H \\ &\quad + \zeta\{(\mathbf{S}^*\mathbf{S}^T)^{-\frac{1}{2}} \otimes [\mathbf{B}(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}]\} \\ &\quad \times \mathbf{D}_K\{(\mathbf{S}^*\mathbf{S}^T)^{-\frac{1}{2}} \otimes [(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}\mathbf{B}^H]\},\end{aligned}\quad (54)$$

where \mathbf{D}_K is a $K^2 \times K^2$ matrix defined as (40), and ζ is a scalar and approximately equal to $((M-N)(L-M+N)/L^2(L+1))$ for large L .

By (43), (52), and (54), we get the MSE of the multivariate Capon estimator as follows:

$$\begin{aligned}\text{MSE}(\hat{\mathbf{B}}_{\text{Capon}}) &= \frac{L-M+N}{L}\text{CRB}(\mathbf{B}) \\ &\quad + \frac{(M-N)(M-N+1)}{L(L+1)}\text{vec}(\mathbf{B})\text{vec}(\mathbf{B})^H \\ &\quad + \zeta\{(\mathbf{S}^*\mathbf{S}^T)^{-\frac{1}{2}} \otimes [\mathbf{B}(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}]\} \\ &\quad \times \mathbf{D}_K\{(\mathbf{S}^*\mathbf{S}^T)^{-\frac{1}{2}} \otimes [(\mathbf{S}\mathbf{S}^H)^{\frac{1}{2}}\mathbf{B}^H]\}.\end{aligned}\quad (55)$$

Equation (55) gives an approximate closed-form expression of the MSE of the multivariate Capon estimator. In this equation, we note that the MSE consists of three terms. The first term is proportional to $\text{CRB}(\mathbf{B})$. The second term is proportional to the outer-product of $\text{vec}(\mathbf{B})$ and is not related to the parameter K and the temporal information matrix \mathbf{S} . In the third term, although there is no explicit dependence of the parameter K , the number of nonzero elements in \mathbf{D}_K is dependent of K . Hence, the third term will increase as K increases. Moreover, the third term is a function of $(\mathbf{S}^*\mathbf{S}^T)$, which depends on the correlation among the rows of \mathbf{S} . As we will see in the following numerical simulations, for \mathbf{S} with correlated rows, the MSE of an element of \mathbf{B} increases as K increases and/or as the other elements in \mathbf{B} increase. On the contrary, when $K = 1$, the third term is zero because the matrix \mathbf{D} becomes a scalar 0 according

to its definition. If we further set $N = 1$, then (55) reduces to the conclusion in the univariate case in [11].

We also note that when the number of snapshots L is large, the last two terms approach zero while the first term approaches $\text{CRB}(\mathbf{B})$. Hence, the multivariate Capon estimator is also asymptotically statistically efficient for large L .

Furthermore, we note that when $M = N$, the MSE of the multivariate Capon estimator is simplified to $\text{CRB}(\mathbf{B})$ like the multivariate ML estimator. This is consistent with our conclusion in the above subsection that the two multivariate methods reduce to the same estimator when $M = N$. For the same reason that we stated in Section III-A, this efficiency of the multivariate Capon estimator should not be seen as a significant advantage.

Now, we summarize the statistical properties of the multivariate Capon estimator by the following theorem.

Theorem 2: For the data model in (1), the multivariate Capon estimate of \mathbf{B} in (8) is biased downward. However, for large number of data samples, it is asymptotically unbiased and statistically efficient. Its bias and MSE matrices are given by (42) and (55), respectively.

IV. NUMERICAL EXAMPLES

In this section, several numerical examples are presented to verify the performance analysis results of the two multivariate estimators. We consider a uniform linear array with $M = 4$ sensors and half-wavelength spacing. We assume $N = 2$ signals arriving at the sensor array with Direction of Arrival (DOAs) of 0° and 15° relative to the array normal. Unless specified otherwise, we assume that $L = 16$ and $\text{SNR} = 10$ dB and \mathbf{S} is formed by $K = 2$ complex sinusoids with unit amplitudes and frequencies 0.10 Hz and 0.125 Hz, respectively. Except in Fig. 6, the elements in \mathbf{B} are all set to be 1. The interference and noise term in our data model in (1) is temporally white but spatially colored zero-mean circularly symmetric complex Gaussian with the spatial covariance matrix \mathbf{Q} given by

$$[\mathbf{Q}]_{ij} = \rho(0.9)^{|i-j|} \quad (56)$$

where $\rho = 1/\text{SNR}$, and $[\cdot]_{ij}$ denotes the i th row and j th column element of a matrix. The figure below are all for $[\mathbf{B}]_{11}$. The figures for other elements of \mathbf{B} are similar. We obtain the empirical results in Fig. 2, using 10 000 Monte Carlo trials while the others use 1000 trials.

We first investigate the bias performance. Fig. 1 shows the bias properties of the two multivariate estimators (denoted by “MV-ML” and “MV-Capon”) from both theoretical predictions (denoted by “Theo.”) and Monte Carlo trials (denoted by “Empi.”). As expected, the multivariate ML is unbiased whereas Capon is biased downward for finite snapshots. However, when the number of snapshots L is large, the bias of the multivariate Capon approaches zero, as predicted by our theoretical analysis.

Fig. 2 illustrates the relationship between the bias and the number of rows of the temporal information matrix \mathbf{S} , i.e., K , when the frequency difference of the complex sinusoids in \mathbf{S} is 0.04 Hz. As predicted by our theoretical analysis, the bias of the multivariate Capon estimator is independent of K .

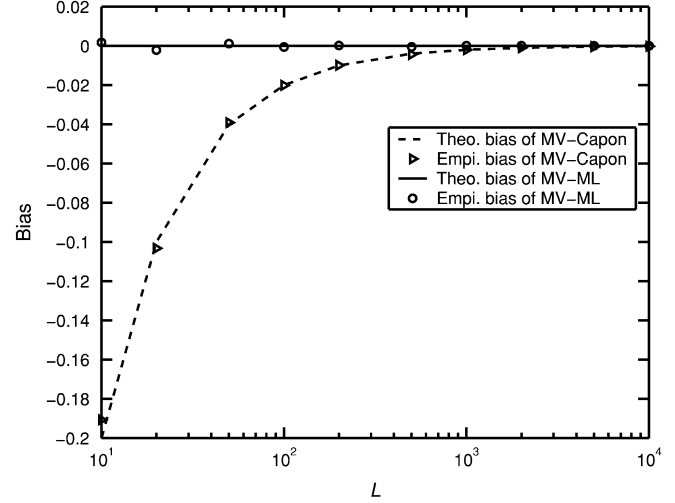


Fig. 1. Bias versus L when $\text{SNR} = 10$ dB, and $K = 2$, $N = 2$.

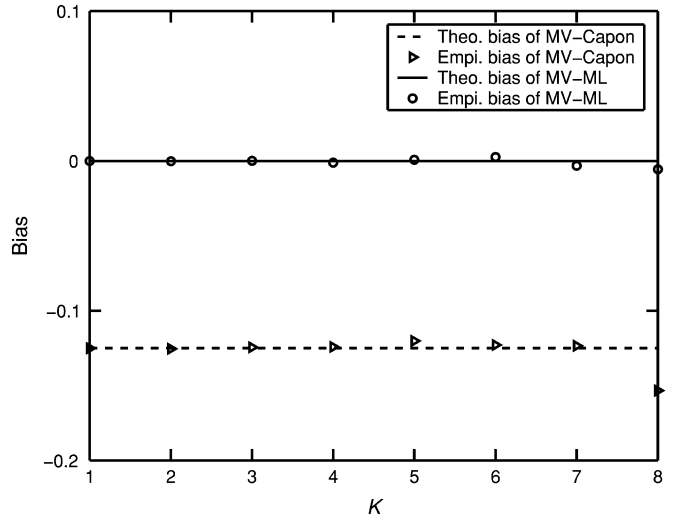


Fig. 2. Bias versus K when $\text{SNR} = 10$ dB, and $L = 16$, $N = 2$.

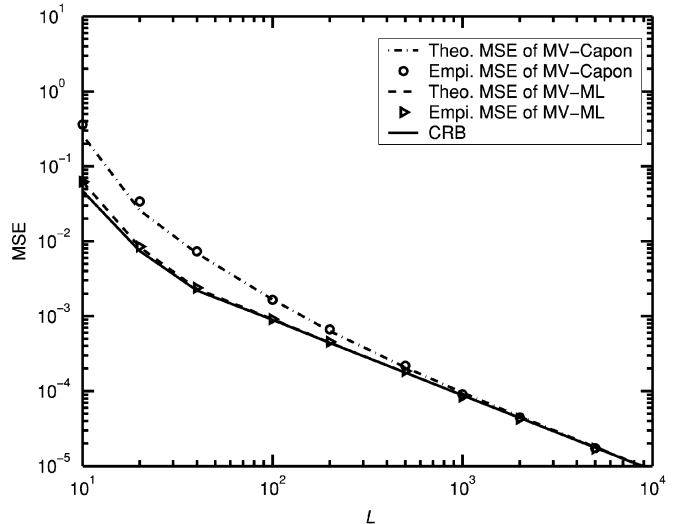


Fig. 3. MSE versus L when $\text{SNR} = 10$ dB, and $K = 2$, $N = 2$.

Fig. 3 illustrates the MSEs of the multivariate estimators as well as the CRB as a function of L . As illustrated, the theoretical and empirical MSE's are consistent. The performance of the multivariate ML estimator is better than the multivariate Capon

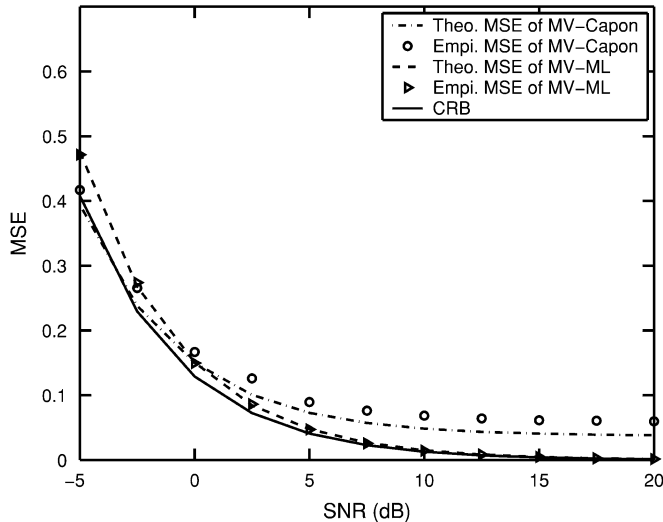


Fig. 4. MSE versus SNR when $L = 16$, $K = 2$, and $N = 2$.

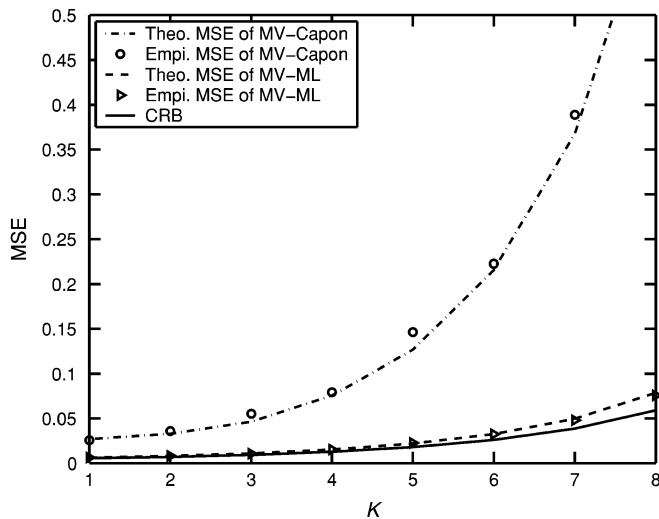


Fig. 5. MSE versus K when SNR = 10 dB, $L = 16$, and $N = 2$.

and very close to the corresponding CRB. As we have predicted in Section III, both multivariate estimators are asymptotically statistically efficient for large number of snapshots, and the performance curves of the two estimators approach the CRB as L increases.

Fig. 4 shows the relationship between the MSE and SNR. Note that the error floor occurs at high SNR for the multivariate Capon estimator due to its bias. As shown in our theoretical analyses, for a fixed M , L , N , and K , the MSE of ML is proportional to $\text{CRB}(\mathbf{B})$, and hence, no “threshold effect” occurs. Note also that like in the univariate case, the Capon estimate can provide a smaller MSE than ML at low SNR. At such a low SNR, however, both ML and Capon perform poorly.

Fig. 5 gives the MSEs of the multivariate Capon and ML estimators as well as the corresponding CRB as a function of K when the frequency difference of the complex sinusoids in \mathbf{S} is 0.04 Hz. As we can see, both the CRB and the MSEs of the two multivariate estimators increase as K increases. However, due to the contribution of the third term in (55), the MSE of Capon increases more quickly than the CRB and the MSE of ML.

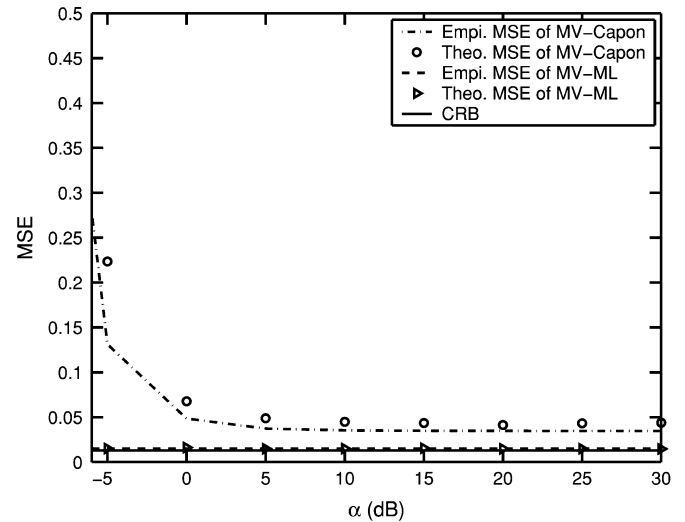


Fig. 6. MSE versus α when SNR = 10 dB, $L = 16$, $K = 2$, and $N = 2$.

In Fig. 6, we consider the case where \mathbf{B} has unequal elements. We set $N = K = 2$, $[\mathbf{B}]_{11} = [\mathbf{B}]_{21} = 1$ and $[\mathbf{B}]_{12} = [\mathbf{B}]_{22} = \alpha^{-1/2}$, where α is the power ratio between the two complex sinusoids in \mathbf{S} . Fig. 6 gives the CRB and MSEs of $[\mathbf{B}]_{11}$ as α varies. As illustrated, the MSE of the multivariate ML estimator is almost constant with respect to α , whereas the MSE of the multivariate Capon estimator increases rapidly when α is decreased to be lower than 0 dB due to its biased nature.

V. CONCLUSION

We have investigated the theoretical performance of two multivariate parameter estimators, namely, the multivariate Capon and ML estimators. Through theoretical analysis and numerical simulations, we conclude that the multivariate ML estimator is unbiased, whereas the multivariate Capon estimator is biased downward for finite snapshots, and both estimators are asymptotically statistically efficient when the number of snapshots is large.

APPENDIX A CRAMÉR–RAO BOUND

In the data model in (1), both \mathbf{Q} and \mathbf{B} are unknown. Let $\boldsymbol{\theta}$ denote the vector containing the real-valued unknowns in \mathbf{Q} and \mathbf{B} . Then, the (i, k) th element of the corresponding Fisher information matrix (FIM) [25], [26] is

$$\begin{aligned} \text{FIM}(\theta_i, \theta_k) = & L \text{tr} \left(\mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial \theta_i} \mathbf{Q}^{-1} \frac{\partial \mathbf{Q}}{\partial \theta_k} \right) \\ & + 2 \text{Re} \left\{ \text{tr} \left[\left(\frac{\partial(\mathbf{ABS})}{\partial \theta_i} \right)^H \mathbf{Q}^{-1} \left(\frac{\partial(\mathbf{ABS})}{\partial \theta_k} \right) \right] \right\} \end{aligned} \quad (57)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix, $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ denote the real and imaginary parts of a complex number (or matrix), respectively, and θ_i denotes the i th element of $\boldsymbol{\theta}$. Because \mathbf{Q} and \mathbf{B} depend on the different variables in $\boldsymbol{\theta}$, FIM will be a block diagonal matrix with respect to the unknowns in \mathbf{Q} and \mathbf{B} . Hence, we can calculate the CRBs of \mathbf{B} and \mathbf{Q} separately. In this paper, we are only interested in the CRB of \mathbf{B} .

Let $b_{ik,R}$ and $b_{ik,I}$ denote the real and imaginary parts of the (i, k) th element in \mathbf{B} , respectively. The corresponding elements in FIM with respect to any two real-valued unknowns in \mathbf{B} are as follows:

$$\text{FIM}(b_{ik,R}, b_{mn,R}) = 2\text{Re} \left[(\mathbf{s}_k^* \mathbf{s}_n^T) (\mathbf{a}_i^H \mathbf{Q}^{-1} \mathbf{a}_m) \right] \quad (58)$$

$$\text{FIM}(b_{ik,R}, b_{mn,I}) = -2\text{Im} \left[(\mathbf{s}_k^* \mathbf{s}_n^T) (\mathbf{a}_i^H \mathbf{Q}^{-1} \mathbf{a}_m) \right] \quad (59)$$

$$\text{FIM}(b_{ik,I}, b_{mn,R}) = 2\text{Im} \left[(\mathbf{s}_k^* \mathbf{s}_n^T) (\mathbf{a}_i^H \mathbf{Q}^{-1} \mathbf{a}_m) \right] \quad (60)$$

$$\text{FIM}(b_{ik,I}, b_{mn,I}) = 2\text{Re} \left[(\mathbf{s}_k^* \mathbf{s}_n^T) (\mathbf{a}_i^H \mathbf{Q}^{-1} \mathbf{a}_m) \right] \quad (61)$$

where $\mathbf{s}_k (k = 1, 2, \dots, K)$ and $\mathbf{a}_i (i = 1, 2, \dots, N)$ are the k th row vector in \mathbf{S} and i th column vector in \mathbf{A} , respectively.

Arranging the complex-valued matrix \mathbf{B} to form a real-valued vector as follows:

$$\mathbf{b} = [\text{Re}(\text{vec}(\mathbf{B}))^T \quad \text{Im}(\text{vec}(\mathbf{B}))^T]^T \quad (62)$$

and arranging (58) to (61) to a matrix according to \mathbf{b} , we get the corresponding FIM

$$\text{FIM}(\mathbf{b}) = 2 \begin{bmatrix} \text{Re}(\Phi) & -\text{Im}(\Phi) \\ \text{Im}(\Phi) & \text{Re}(\Phi) \end{bmatrix} \quad (63)$$

where

$$\Phi = (\mathbf{S}^* \mathbf{S}^T) \otimes (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A}). \quad (64)$$

Using the matrix inversion lemma and the inversion lemma of partitioned matrices [32], we get the CRB in real-valued form:

$$\text{CRB}(\mathbf{b}) = [\text{FIM}(\mathbf{b})]^{-1} = \frac{1}{2} \begin{bmatrix} \text{Re}(\Phi^{-1}) & -\text{Im}(\Phi^{-1}) \\ \text{Im}(\Phi^{-1}) & \text{Re}(\Phi^{-1}) \end{bmatrix}. \quad (65)$$

Transforming the above real-valued CRB into the complex-valued form yields

$$\text{CRB}(\mathbf{B}) \triangleq \text{CRB}(\text{vec}(\mathbf{B})) = (\mathbf{S}^* \mathbf{S}^T)^{-1} \otimes (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1}. \quad (66)$$

Clearly, the diagonal elements of $\text{CRB}(\mathbf{B})$ are determined by the diagonal elements of $(\mathbf{S}^* \mathbf{S}^T)^{-1}$ and $(\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})^{-1}$. To study the influences of the temporal information matrix \mathbf{S} and the spatial information matrix \mathbf{A} on $\text{CRB}(\mathbf{B})$, we denote $\mathbf{R}_S = (\mathbf{S}^* \mathbf{S}^T)$ and $\mathbf{R}_A = (\mathbf{A}^H \mathbf{Q}^{-1} \mathbf{A})$. Without loss of generality, we consider the CRB of b_{11} , i.e., the element on the first column and first row of \mathbf{B} . Partition \mathbf{R}_S , \mathbf{R}_S^{-1} , \mathbf{R}_A and \mathbf{R}_A^{-1} as follows:

$$\mathbf{R}_S = \begin{bmatrix} (r_S)_{11} & (\mathbf{r}_S)_{12} \\ (\mathbf{r}_S)_{21} & (\mathbf{R}_S)_{22} \end{bmatrix}, \quad \mathbf{R}_S^{-1} = \begin{bmatrix} (r_S)_{11}^{-1} & (\mathbf{r}_S)_{12}^{-1} \\ (\mathbf{r}_S)_{21}^{-1} & (\mathbf{R}_S)_{22}^{-1} \end{bmatrix} \quad (67)$$

$$\mathbf{R}_A = \begin{bmatrix} (r_A)_{11} & (\mathbf{r}_A)_{12} \\ (\mathbf{r}_A)_{21} & (\mathbf{R}_A)_{22} \end{bmatrix}, \quad \mathbf{R}_A^{-1} = \begin{bmatrix} (r_A)_{11}^{-1} & (\mathbf{r}_A)_{12}^{-1} \\ (\mathbf{r}_A)_{21}^{-1} & (\mathbf{R}_A)_{22}^{-1} \end{bmatrix}. \quad (68)$$

Obviously, $\text{CRB}(b_{11}) = (r_S)^{11} (r_A)^{11}$. By the inversion lemma of partitioned matrices [32], we have

$$(r_S)^{11} = \frac{1}{(r_S)_{11} - (\mathbf{r}_S)_{12} (\mathbf{R}_S)_{22}^{-1} (\mathbf{r}_S)_{21}}, \quad (69)$$

$$(r_A)^{11} = \frac{1}{(r_A)_{11} - (\mathbf{r}_A)_{12} (\mathbf{R}_A)_{22}^{-1} (\mathbf{r}_A)_{21}}. \quad (70)$$

In (69), $(r_S)_{11} = \|\mathbf{s}_1\|^2$ is the Euclidean norm square of the first row of \mathbf{S} . It is easily verified that $(\mathbf{R}_S)_{22}$ is a positive definite matrix while $(\mathbf{r}_S)_{12} = (\mathbf{r}_S)_{21}^H$. Therefore, $(r_S)^{11}$ is minimized when $(\mathbf{r}_S)_{12} = \mathbf{0}$, i.e., $\mathbf{s}_1^* \mathbf{s}_k^T = 0$ ($k = 2, 3, \dots, K$). Hence, to minimize $\text{CRB}(\mathbf{B})$, the row vectors in \mathbf{S} should be orthogonal to each other.

Similarly, $(r_A)^{11}$ is minimized when $(\mathbf{r}_A)_{12} = \mathbf{0}$, i.e., $(\mathbf{Q}^{-1/2} \mathbf{a}_1)^H (\mathbf{Q}^{-1/2} \mathbf{a}_i) = 0$ ($i = 2, 3, \dots, N$). Note that when this condition is not satisfied, large N causes $(r_A)^{11}$ to increase. Furthermore, We note that $(r_A)_{11} = \mathbf{a}_1^H \mathbf{Q} \mathbf{a}_1$. Therefore, when \mathbf{a}_1 is proportional to the eigenvector of \mathbf{Q} corresponding to its smallest eigenvalue, $(r_A)^{11}$ is minimized. Therefore, to minimize $\text{CRB}(\mathbf{B})$, the columns of $\mathbf{Q}^{-1/2} \mathbf{A}$ should be orthogonal to each other and the columns of \mathbf{A} should correspond to the subspace spanned by the eigenvectors of \mathbf{Q} corresponding to its smallest N eigenvalues. Since \mathbf{Q} is unknown and usually \mathbf{A} is given and cannot be changed in practice, we can only hope for these conditions of \mathbf{A} .

APPENDIX B

PROOF OF LEMMA 2

Let \mathbf{g} be any nonzero $p \times 1$ vector. We first consider the expectation of $\mathbf{g}^H \mathbf{\Pi} \mathbf{g}$.

Decompose \mathbf{g} as follows:

$$\mathbf{g} = \mathbf{U} \tilde{\mathbf{g}} \quad (71)$$

where \mathbf{U} is a $p \times p$ unitary matrix with its first column being $(\mathbf{g} / \|\mathbf{g}\|)$ and $\tilde{\mathbf{g}}_{p \times 1} = \|\mathbf{g}\| [1 \ 0 \ 0 \ \dots \ 0]^T$.

Let $\tilde{\mathbf{Y}} = \mathbf{Y} \mathbf{U}$; obviously $\tilde{\mathbf{Y}} \sim \text{CN}(0; \mathbf{I}_n, \mathbf{I}_p)$. Let $\tilde{\mathbf{v}}_i (i = 1, 2, \dots, p)$ be the i th column vector of $\tilde{\mathbf{Y}}$. Let $\tilde{\Psi} = \Psi + \sum_{i=2}^p \tilde{\mathbf{v}}_i \tilde{\mathbf{v}}_i^H$; obviously $\tilde{\Psi} \sim \text{CW}(l + p - 1, n; \mathbf{I})$ [31]. Then

$$\begin{aligned} \mathbf{g}^H \mathbf{\Pi} \mathbf{g} &= \tilde{\mathbf{g}}^H \tilde{\mathbf{Y}}^H (\Psi + \tilde{\mathbf{Y}} \tilde{\mathbf{Y}}^H)^{-1} \tilde{\mathbf{Y}} \tilde{\mathbf{g}} \\ &= \|\mathbf{g}\|^2 \tilde{\mathbf{v}}_1^H (\tilde{\Psi} + \tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_1^H)^{-1} \tilde{\mathbf{v}}_1 \\ &= \|\mathbf{g}\|^2 \left[1 - \frac{1}{1 + \tilde{\mathbf{v}}_1^H \tilde{\Psi}^{-1} \tilde{\mathbf{v}}_1} \right]. \end{aligned} \quad (72)$$

It has been shown in [16, Appendix] that

$$\frac{1}{1 + \tilde{\mathbf{v}}_1^H \tilde{\Psi}^{-1} \tilde{\mathbf{v}}_1} \sim \beta(l - n + p, n). \quad (73)$$

Hence [33]

$$E \left(\frac{1}{1 + \tilde{\mathbf{v}}_1^H \tilde{\Psi}^{-1} \tilde{\mathbf{v}}_1} \right) = \frac{l - n + p}{l + p}. \quad (74)$$

From (72) and (74), we have

$$\mathbf{g}^H E(\mathbf{\Pi}) \mathbf{g} = \mathbf{g}^H \left(\frac{n}{l+p} \mathbf{I}_p \right) \mathbf{g}. \quad (75)$$

The above equation should be satisfied for any nonzero vector \mathbf{g} , which means $E(\mathbf{\Pi}) = (n/l+p)\mathbf{I}$. To calculate $E(\text{vec}(\mathbf{\Pi})\text{vec}(\mathbf{\Pi})^H)$, we need to calculate

$$E(\pi_{r_1 c_1} \pi_{r_2 c_2}^*) = E \left[\mathbf{v}_{r_1}^H (\mathbf{\Psi} + \mathbf{\Upsilon} \mathbf{\Upsilon}^H)^{-1} \times \mathbf{v}_{c_1} \mathbf{v}_{c_2}^H (\mathbf{\Psi} + \mathbf{\Upsilon} \mathbf{\Upsilon}^H)^{-1} \mathbf{v}_{r_2} \right] \quad (76)$$

where $\pi_{r_1 c_1}$ denotes the (r_1, c_1) -th element in $\mathbf{\Pi}$, \mathbf{v}_k denotes the k th column vector in $\mathbf{\Upsilon}$, and $r_1, c_1, r_2, c_2 = 1, 2, 3, \dots, p$.

Note that (76) is a function of \mathbf{v}_k ($k = 1, 2, \dots, p$). Adopting the same technique used to obtain (25), we can easily show that the expectation is zero when (76) contains odd numbers of \mathbf{v}_k , e.g., $E(\pi_{12} \pi_{34}^*) = E(\pi_{11} \pi_{23}^*) = E(\pi_{12} \pi_{23}^*) = E(\pi_{12} \pi_{32}^*) = 0$.

When $r_1 = c_2, c_1 = r_2$ but $r_1 \neq c_1$, $E(\pi_{r_1 c_1} \pi_{c_1 r_1}^*) = E\{\mathbf{v}_{r_1}^H (\mathbf{\Psi} + \mathbf{\Upsilon} \mathbf{\Upsilon}^H)^{-1} \mathbf{v}_{c_1}\}^2$. Replacing \mathbf{v}_{r_1} by $j \cdot \mathbf{v}_{r_1}$, where j is the unit of the imaginary number, yields

$$E(\pi_{r_1 c_1} \pi_{c_1 r_1}^*) |_{\mathbf{v}_{r_1} = j \cdot \mathbf{v}_{r_1}} = -E(\pi_{r_1 c_1} \pi_{c_1 r_1}^*). \quad (77)$$

On the other hand, since \mathbf{v}_{r_1} is a zero-mean circularly symmetric complex Gaussian random vector, $j \cdot \mathbf{v}_{r_1}$, as a random vector transformed from \mathbf{v}_{r_1} , has the same statistical property as \mathbf{v}_{r_1} . Hence

$$E(\pi_{r_1 c_1} \pi_{c_1 r_1}^*) |_{\mathbf{v}_{r_1} = j \cdot \mathbf{v}_{r_1}} = E(\pi_{r_1 c_1} \pi_{c_1 r_1}^*). \quad (78)$$

Hence, by (77) and (78), we have $E(\pi_{r_1 c_1} \pi_{c_1 r_1}^*) = 0$.

Besides the above cases, there are three cases left in which $E(\mathbf{\Pi}_{r_1 c_1} \mathbf{\Pi}_{r_2 c_2}^*)$ is nonzero, i.e.,

- i) $r_1 = c_1 = r_2 = c_2$;
- ii) $r_1 = c_1, r_2 = c_2$ but $r_1 \neq r_2$;
- iii) $r_1 = r_2, c_1 = c_2$ but $r_1 \neq c_1$.

We know that the expectation in each case is equal. For convenience, we denote the expectations in the three cases as $E(|\pi_{11}|^2)$, $E(\pi_{11} \pi_{22}^*)$, and $E(|\pi_{12}|^2)$, respectively.

First, we calculate $E(|\pi_{11}|^2)$. Let $\check{\mathbf{\Psi}} = \mathbf{\Psi} + \sum_{k=2}^p \mathbf{v}_k \mathbf{v}_k^H$, which obviously has the CW($l+p-1, n; \mathbf{I}$) distribution. Then

$$\begin{aligned} E(|\pi_{11}|^2) &= E \left(\left| \mathbf{v}_1^H (\check{\mathbf{\Psi}}^{-1} + \mathbf{v}_1 \mathbf{v}_1^H)^{-1} \mathbf{v}_1 \right|^2 \right) \\ &= E \left(\left| 1 - \frac{1}{1 + \mathbf{v}_1^H \check{\mathbf{\Psi}}^{-1} \mathbf{v}_1} \right|^2 \right) \\ &= 1 - 2E \left(\frac{1}{1 + \mathbf{v}_1^H \check{\mathbf{\Psi}}^{-1} \mathbf{v}_1} \right) \\ &\quad + E \left[\left(\frac{1}{1 + \mathbf{v}_1^H \check{\mathbf{\Psi}}^{-1} \mathbf{v}_1} \right)^2 \right]. \quad (79) \end{aligned}$$

Again, using the conclusion in [16, Appendix], we know

$$\frac{1}{1 + \mathbf{v}_1^H \check{\mathbf{\Psi}}^{-1} \mathbf{v}_1} \sim \beta(l-n+p, n). \quad (80)$$

Calculating the expectation and the second-order moment of the above β distribution and substituting them into (79) yields

$$E(|\pi_{11}|^2) = \frac{n(n+1)}{(l+p)(l+p+1)}. \quad (81)$$

Second, we consider $E(\pi_{11} \pi_{22}^*)$. It is difficult to calculate this expectation directly. However, note that $\mathbf{\Psi} + \mathbf{\Upsilon} \mathbf{\Upsilon}^H$ can be expressed as the sum of the outer-products of $l+p$ complex Gaussian random vectors with zero mean and covariance \mathbf{I}_n . By the Law of Large Numbers [33], for large $l+p$, it approaches $(l+p)\mathbf{I}$ in probability. Hence

$$E(\pi_{11} \pi_{22}^*) \approx \frac{1}{(l+p)^2} E[(\mathbf{v}_1^H \mathbf{v}_1)(\mathbf{v}_2^H \mathbf{v}_2)] = \frac{n^2}{(l+p)^2}. \quad (82)$$

Third, in order to calculate $E(|\pi_{12}|^2)$, we use the fact that

$$E(|\pi_{11}|^2) = E(|\pi_{12}|^2) + E(\pi_{11} \pi_{22}^*). \quad (83)$$

This equation can be proved as follows. Let us make a transformation from $(\mathbf{v}_1, \mathbf{v}_2)$ to $((\sqrt{2}/2)(\mathbf{v}_1 + \mathbf{v}_2), (\sqrt{2}/2)(\mathbf{v}_1 - \mathbf{v}_2))$. Obviously, $(\sqrt{2}/2)(\mathbf{v}_1 + \mathbf{v}_2)$ and $(\sqrt{2}/2)(\mathbf{v}_1 - \mathbf{v}_2)$ are two independent standard complex Gaussian random vectors and retain the same statistical properties of $(\mathbf{v}_1, \mathbf{v}_2)$. Hence, replacing $(\mathbf{v}_1, \mathbf{v}_2)$ by the new random vectors in $E(|\pi_{11}|^2)$, the expectation will not change:

$$\begin{aligned} E(|\pi_{11}|^2) &= \frac{1}{4} E \left\{ \left| (\mathbf{v}_1^H + \mathbf{v}_2^H) (\mathbf{\Psi} + \mathbf{\Upsilon} \mathbf{\Upsilon}^H)^{-1} (\mathbf{v}_1 + \mathbf{v}_2) \right|^2 \right\} \\ &= \frac{1}{2} E(|\pi_{11}|^2) + \frac{1}{2} E(|\pi_{12}|^2) + \frac{1}{2} E(\pi_{11} \pi_{22}^*). \quad (84) \end{aligned}$$

From (84), (83) is proved.

Using the facts of (81)–(83) and arranging $E(\pi_{r_1 c_1} \pi_{r_2 c_2}^*)$ into $E(\text{vec}(\mathbf{\Pi})\text{vec}(\mathbf{\Pi})^H)$, (39) follows immediately.

REFERENCES

- [1] J. Li and R. T. Compton Jr., "Maximum likelihood angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2850–2862, Sep. 1993.
- [2] J. Li, B. Haider, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2154–2163, Sep. 1995.
- [3] A. Zeira and B. Friedlander, "Direction of arrival estimation using parametric signal models," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 339–350, Feb. 1996.
- [4] M. Cedervall and R. L. Moses, "Efficient maximum likelihood DOA estimation for signals with known waveforms in the presence of multipath," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 808–811, Mar. 1997.
- [5] M. Wax and A. Leshem, "Joint estimation of time delays and directions of arrival of multiple reflections of a known signal," *IEEE Trans. Signal Process.*, vol. 45, no. 10, pp. 2477–2484, Oct. 1997.
- [6] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 449–462, Feb. 1998.

- [7] A. L. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proc. IEEE*, vol. 86, no. 2, pp. 421–441, Feb. 1998.
- [8] A. Jakobsson, A. Swindlehurst, and P. Stoica, "Subspace-based estimation of time delays and doppler shifts," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2472–2483, Sep. 1998.
- [9] H. Li, G. Liu, and J. Li, "Angle estimator for signals with known waveforms," *Electron. Lett.*, vol. 35, pp. 1992–1994, Nov. 1999.
- [10] U. Spagnolini, "An iterative quadratic method for high resolution delay estimation with known waveform," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 1134–1137, Mar. 2000.
- [11] Y. Jiang, P. Stoica, and J. Li, "Array signal processing in the known waveform and steering vector case," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 23–35, Jan. 2004.
- [12] A. Dogandžić and A. Nehorai, "Generalized multivariate analysis of variance," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 39–54, Sep. 2003.
- [13] *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook*, M. Siotani, T. Hayakawa, and Y. Pujikoshi, Eds., Amer. Sciences., Columbus, OH, 1985.
- [14] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.
- [15] J.-X. Pan and K.-T. Fang, *Growth Curve Models and Statistical Diagnostics*. New York: Springer, 2002.
- [16] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 115–127, Mar. 1986.
- [17] H. Wang and L. Cai, "On adaptive multiband signal detection with GLR algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 2, pp. 225–233, Mar. 1991.
- [18] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1469–1484, Jun. 1996.
- [19] P. Stoica and T. Sundin, "Nonparametric NMR spectroscopy," *J. Magn. Reson.*, vol. 152, pp. 57–69, Sep. 2001.
- [20] A. Dogandžić and A. Nehorai, "Space-time fading channel estimation and symbol detection in unknown spatially correlated noise," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 457–474, Mar. 2002.
- [21] E. G. Larsson, J. Liu, and J. Li, "Demodulation of space-time codes in the presence of interference," *Electron. Lett.*, vol. 37, pp. 697–698, May 2001.
- [22] E. G. Larsson, P. Stoica, and J. Li, "On maximum-likelihood detection and decoding for space-time coding systems," *IEEE Trans. Signal Process.*, vol. 50, no. 4, pp. 937–944, Apr. 2002.
- [23] ———, "Orthogonal space-time block codes: Maximum-likelihood detection for unknown channels and unstructured interferences," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 362–372, Feb. 2003.
- [24] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.
- [25] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.
- [26] H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*. New York: Wiley, 2002.
- [27] C. R. Rao, *Linear Statistical Inference and Its Applications*. New York: Wiley, 1973.
- [28] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall, 1989.
- [29] N. R. Goodman, "Statistical analysis based on a certain multi-variate complex Gaussian distribution," *Ann. Math. Stat.*, vol. 34, pp. 152–177, Mar. 1963.
- [30] C. G. Khatri and C. R. Rao, "Effects of estimated noise covariance matrix optimal signal detection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 671–679, May 1987.
- [31] H. H. Andersen, M. Højbjerg, D. Sørensen, and P. S. Eriksen, *Linear and Graphical Models for Multivariate Complex Normal Distribution*. New York: Springer-Verlag, Inc., 1995.
- [32] D. A. Harville, *Matrix Algebra from a Statistician's Perspective*. New York: Springer, Inc., 1997.
- [33] G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Stamford, CT: Thomson Learning, 2002.



Luzhou Xu received the B.Eng. and M.S. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 1996 and 1999, respectively. He is currently pursuing the Ph.D degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville.

From 1999 to 2001, he was with the Zhongxing Telecommunication Company, Shenzhen, China, where he was involved in the system and algorithm design of mobile communication equipments. From 2001 to 2003, he was with the Wireless Communication Group, Philips Research, Shanghai, China. His research interests include biomedical signal processing, wireless communications, and array signal processing.



Jian Li (S'87–M'91–SM'97) received the M.Sc. and Ph.D. degrees in electrical engineering from The Ohio State University (OSU), Columbus, in 1987 and 1991, respectively.

From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, OSU. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. Her current research interests include spectral estimation, array signal processing, and their applications.

Dr. Li is a member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She was an Executive Committee Member of the 2002 International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, May 2002. She has been an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING since 1999. She is presently a member of the Signal Processing Theory and Methods (SPTM) Technical Committee of the IEEE Signal Processing Society.