

Improved Capon Estimator for High-Resolution DOA Estimation and Its Statistical Analysis

Weiliang Zuo, *Member, IEEE*, Jingmin Xin, *Senior Member, IEEE*, Changnong Liu,
Nanning Zheng, *Fellow, IEEE*, and Akira Sano, *Member, IEEE*

Abstract—Despite some efforts and attempts have been made to improve the direction-of-arrival (DOA) estimation performance of the standard Capon beamformer (SCB) in array processing, rigorous statistical performance analyses of these modified Capon estimators are still lacking. This paper studies an improved Capon estimator (ICE) for estimating the DOAs of multiple uncorrelated narrowband signals, where the higher-order inverse (sample) array covariance matrix is used in the Capon-like cost function. By establishing the relationship between this nonparametric estimator and the parametric and classic subspace-based MUSIC (multiple signal classification), it is clarified that as long as the power order of the inverse covariance matrix is increased to reduce the influence of signal subspace components in the ICE, the estimation performance of the ICE becomes equivalent to that of the MUSIC regardless of the signal-to-noise ratio (SNR). Furthermore the statistical performance of the ICE is analyzed, and the large-sample mean-squared-error (MSE) expression of the estimated DOA is derived. Finally the effectiveness and the theoretical analysis of the ICE are substantiated through numerical examples, where the Cramer-Rao lower bound (CRB) is used to evaluate the validity of the derived asymptotic MSE expression.

Index Terms—Capon beamformer, direction-of-arrival (DOA) estimation, large-sample mean-squared-error (MSE), subspace-based methods, uniform linear array.

I. INTRODUCTION

IN various applications such as radar, sonar, astronomy, seismology, biomedicine, and communications, the direction-of-arrival (DOA) estimation of multiple narrowband sig-

nals impinging on an array of sensors is very important [1]–[17]. Extensive research has been conducted on this estimation problem for decades, and many estimation methods have been proposed in the literature (see, e.g., [1]–[3], [7], [12], [15]–[17] and references therein), where the nonparametric methods (such as beamforming techniques) and the parametric methods (for example, the maximum likelihood (ML), the subspace-based methods) are of great significance and are most widely used. The nonparametric methods are based on the concept of data-adaptive finite-impulse response (FIR) filtering (see, e.g., [5], [12], [15]–[18]), while the parametric methods assume a model for the array data (see, e.g., [19]–[29]), where the former possess better robustness than the latter, but the latter has high resolution and good accuracy [30].

The beamforming is one of the oldest methods for nonparametric DOA estimation [15], and perhaps the most representative data-independent one is the Capon beamformer (i.e., the minimum variance distortionless response (MVDR)) [31], which was proposed in the late 1960s as a better alternative to the traditional Fourier analysis based Bartlett beamformer, where a non-linear and non-quadratic cost function with the inverse array covariance matrix is used. Compared to parametric and classic subspace-based MUSIC [20], the Capon beamformer is in general more practical in terms of implementation as it requires less a priori knowledge on the statistical properties of array data and the number of incident signals and is applicable for any array geometry [32], where these properties are not available (see, [15]–[18]), and the geometries of the arrays are usually limited by physical factors, but the standard Capon beamformer (SCB) generally suffers from low resolution and poor accuracy at low signal-to-noise ratio (SNR) (see, e.g., [1]–[5], [16], [33]–[43]), while the MUSIC is the large-sample realization of the ML method in the presence of uncorrelated incident signals [44] and it involves the computationally intensive eigendecomposition, which may become a tremendous computational burden in the practical applications with a large number of array sensors [16], [38], [45].

In fact, the beamforming including Capon beamformer is also a classic yet continuously developing field, which has practical and sensible applications in array processing (see, e.g., [13], [39], [46]–[64] and references therein), and many efforts and attempts were made to develop some variants of the Capon beamformer to improve the estimation performance from different perspectives in the past few decades

Manuscript received December 30, 2022; accepted March 9, 2023. This work was supported in part by the National Natural Science Foundation of China (62201447) and the Project Supported by Natural Science Basic Research Plan in Shaanxi Province of China (2022JQ-640). This paper was presented in part at 2014 Int. Symposium Nonlinear Theory and Its Applications (NOLTA'2014), Luzern, Switzerland, Sept. 14–18, 2014. Recommended by Associate Editor Qinglai Wei. (*Corresponding author: Jingmin Xin.*)

Citation: W. L. Zuo, J. M. Xin, C. N. Liu, N. N. Zheng, and A. Sano, "Improved Capon estimator for high-resolution DOA estimation and its statistical analysis," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 8, pp. 1716–1729, Aug. 2023.

W. L. Zuo, J. M. Xin, and N. N. Zheng are with the National Key Laboratory of Human-Machine Hybrid Augmented Intelligence, the National Engineering Research Center for Visual Information and Applications, and the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: weiliang.zuo@xjtu.edu.cn; jxin@mail.xjtu.edu.cn; nnzheng@mail.xjtu.edu.cn).

C. N. Liu is with the Data Center, China Construction Bank, Beijing 100033, China (e-mail: liu.ch.n@stu.xjtu.edu.cn).

A. Sano is with the Department of System Design Engineering, Keio University, Yokohama 223-8522, Japan (e-mail: sano@sd.keio.ac.jp).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JAS.2023.123549

(see, e.g., [10], [32], [37], [38], [65]–[85]). Since the array covariance matrix should be estimated from finite array data, and the (high-order) inverse sample covariance matrix is usually involved in the non-linear and non-quadratic cost function, it is rather complicated and difficult to analyze the statistical performance of the standard and modified Capon estimators. In particular, a class of non-quadratic spectral estimators (i.e., Pisarenko framework [86]) was proposed from the analogy of Capon beamformer [67], where the higher-order array covariance matrix is adopted in the non-quadratic cost function. Although the statistical performances for the Pisarenko framework [67], the SCB [31] and the weighted Capon beamformer (WCB) [38] were analyzed in [37]–[39] only simulation examples were used to evaluate the estimation performance of other variants of the SCB (see, e.g., [38]). Further the derived analytical expression [40] is not strict and accurate when the number of snapshots is small and/or the SNR is relatively low, while the statistical analysis of the WCB derived in [38] is only applicable to the modified Capon beamformer (MCB) with the second-order inverse sample covariance matrix [9] and it is still difficult to determine the optimal weighting matrix of the WCB [38]. Although the similar Capon estimator was considered in our previous work [87], its statistical performance was not completed therein.

Therefore in this paper, we study an improved Capon estimator (ICE) for estimating the DOAs of multiple uncorrelated narrowband signals, where the higher-order inverse array covariance matrix is used in the cost function, and we focus on the performance analyses of such DOA estimator. First the relationship between the ICE and the classic subspace-based MUSIC is established. It is clarified that as long as the power order of inverse covariance matrix is larger enough, the ICE can greatly outperform the SCB and is equivalent to the MUSIC regardless of the SNR. Then the statistical properties of the ICE are analyzed, and the asymptotic MSE expressions of DOA estimates are derived for a sufficiently large number of snapshots. In addition, an analytical and numerical study of the performance is performed for the case of one signal impinging on the uniform linear array (ULA) to get insights into the ICE. The simulation results show that the DOA estimation performance of the ICE is significantly improved compared with the SCB, where the Cramer-Rao lower bound (CRB) is also used to evaluate the validity of the derived MSE analytical expression.

The main contribution of this paper: 1) A new improved Capon estimator for DOA estimation is proposed by using the higher-order inverse array covariance matrix in the cost function. 2) The relationship between the ICE and the MUSIC is established that as long as the power order of inverse covariance matrix is larger enough, the ICE can greatly outperform the SCB and is equivalent to the MUSIC regardless of the SNR. 3) The statistical properties of the ICE are derived explicitly, and the asymptotic MSE expressions of DOA estimates are derived for a sufficiently large number of snapshots.

Notations: Throughout the paper, $\mathbf{O}_{m \times n}$, \mathbf{I}_m , $\mathbf{0}_{m \times 1}$, \mathbf{e}_l , and $\delta_{n,t}$ stand for an $m \times n$ null matrix, $m \times m$ identity matrix, $m \times 1$ null vector, an $l \times 1$ unit vector with one as the first element

whereas zeros elsewhere, and Kronecker delta, while $E\{\cdot\}$, $\{\cdot\}^*$, and $(\cdot)^H$ represent the statistical expectation, complex conjugate, and Hermitian transposition, respectively. Additionally, $\text{diag}\{\cdot\}$ and $\text{blkdiag}\{\cdot\}$ indicate the diagonal matrix and block diagonal matrix operators, respectively, and $\text{tr}\{\cdot\}$ signifies the trace operator, while $\text{Re}\{\cdot\}$ denotes the real part of the bracketed quantity, and \hat{x} means the estimate of x .

II. PROBLEM FORMULATION

We consider an array composed of L identical and omnidirectional sensors indexed by $l = 1, 2, \dots, L$, where the coordinate of the l th sensor is (\bar{x}_l, \bar{y}_l) , and we assume that K uncorrelated narrowband signals $\{s_k(n)\}_{k=1}^K$ with the wavelength λ are in the far-field and impinge from distinct directions $\{\theta_k\}_{k=1}^K$. Herein the first sensor of the array (i.e., sensor 1) is assumed to be the phase reference point (i.e., $\bar{x}_1 = 0$, $\bar{y}_1 = 0$), and the DOA θ_k of $s_k(n)$ is measured at the reference sensor relative to the normal of the array. Then the received noisy signal $x_l(n)$ at the l th sensor is expressed as

$$x_l(n) = \sum_{k=1}^K s_k(n) e^{j\tau_l(\theta_k)} + w_l(n) \quad (1)$$

where $w_l(n)$ is the additive noise, and $\tau_l(\theta_k)$ indicates the phase delay of the k th signal $s_k(n)$ due to the propagation time between the reference sensor and the l th sensor, which is given by

$$\tau_l(\theta_k) = \frac{2\pi}{\lambda} (\bar{x}_l \sin \theta_k + \bar{y}_l \cos \theta_k). \quad (2)$$

Then the compact model of array data is given by

$$\mathbf{x}(n) = \mathbf{A}(\theta) \mathbf{s}(n) + \mathbf{w}(n) \quad (3)$$

where $\mathbf{x}(n)$, $\mathbf{s}(n)$ and $\mathbf{w}(n)$ are the vectors of the received noisy data, the incident signals and the additive noises given by $\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_L(n)]^T$, $\mathbf{s}(n) = [s_1(n), s_2(n), \dots, s_K(n)]^T$ and $\mathbf{w}(n) = [w_1(n), w_2(n), \dots, w_L(n)]^T$, respectively, $\mathbf{A}(\theta)$ is the array response matrix given by $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)]$, while the array response vector $\mathbf{a}(\theta_k)$ is given by

$$\mathbf{a}(\theta_k) = [1, e^{j\tau_2(\theta_k)}, \dots, e^{j\tau_L(\theta_k)}]^T. \quad (4)$$

Now we make the basic assumptions on the data model as follows.

Assumption 1: The array is calibrated and the array response matrix \mathbf{A} has full-rank and unambiguous.

Assumption 2: The incident signals $\{s_k(n)\}$ are temporally complex white Gaussian random processes with zero-mean and the variance given by $E\{s_k(n)s_k^*(t)\} = r_{s_k} \delta_{n,t}$ and $E\{s_k(n)s_k(t)\} = 0, \forall n, t$.

Assumption 3: The additive noises $\{w_l(n)\}$ are temporally and spatially complex white Gaussian random processes with zero-mean and the covariance matrices $E\{\mathbf{w}(n)\mathbf{w}^H(t)\} = \sigma^2 \mathbf{I}_L \delta_{n,t}$, and $E\{\mathbf{w}(n)\mathbf{w}^T(t)\} = \mathbf{O}_{L \times L}, \forall n, t$. In addition, the additive noise is independent of the incident signals.

Assumption 4: The number of incident signals K is known or estimated by the existing number detection techniques in advance (cf., [88] and references therein), and it satisfies the relation that $K < L$.

Under the basic assumptions, from (3), we easily obtain the

covariance matrix \mathbf{R} of the received array data

$$\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^H(n)\} = \mathbf{A}(\theta)\mathbf{R}_s\mathbf{A}^H(\theta) + \sigma^2\mathbf{I}_L \quad (5)$$

where \mathbf{R}_s is the signal covariance matrix with full-rank and defined as $\mathbf{R}_s = E\{s(n)s^H(n)\} = \text{diag}\{r_{s_1}, r_{s_2}, \dots, r_{s_K}\}$. In practice, when the finite snapshots of array data are available, the true covariance matrix \mathbf{R} should be replaced by its sample estimate $\hat{\mathbf{R}}$ given by

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}(n)\mathbf{x}^H(n) \quad (6)$$

where N is the number of snapshots. In this paper, we consider a new modified Capon beamformer (i.e., the ICE herein) for estimating the DOAs $\{\theta_k\}_{k=1}^K$ of multiple signals with high-resolution from the sample covariance matrix $\hat{\mathbf{R}}$ and study its statistical performance analysis.

III. IMPROVED CAPON ESTIMATOR FOR DOA ESTIMATION—ICE

A. Preliminary—SCB

The beamforming is considered as a spatial filter that multiplies the received signals from each sensor with complex weights and then sums them to form the array output (see [5], [17], [18] for details). The SCB uses the array weights to maintain a unity constraint in the specific “look direction” while maximally suppressing the signals from other directions and additive noises to minimize the mean-squared power of the array output [15], [31], [36], [66], [67]. The optimal weight vector $\bar{\mathbf{w}}$ of the SCB is obtained as the solution of the following constrained quadratic problem [15]

$$\min_{\bar{\mathbf{w}}} \bar{\mathbf{w}}^H \mathbf{R} \bar{\mathbf{w}} \quad \text{s.t.} \quad \bar{\mathbf{w}}^H \mathbf{a}(\theta) = 1. \quad (7)$$

By using techniques such as the Lagrange optimization, we can obtain the weight vector $\bar{\mathbf{w}}_{\text{SCB}}$ as

$$\bar{\mathbf{w}}_{\text{SCB}} = \frac{\mathbf{R}^{-1} \mathbf{a}(\theta)}{\mathbf{a}^H(\theta) \mathbf{R}^{-1} \mathbf{a}(\theta)}. \quad (8)$$

Then the Capon estimates of the DOAs $\{\theta_k\}_{k=1}^K$ are obtained from the positions of the K smallest minima of the following function $f_{\text{SCB}}(\theta)$ defined as:

$$f_{\text{SCB}}(\theta) = \mathbf{a}^H(\theta) \mathbf{R}^{-1} \mathbf{a}(\theta) \quad (9)$$

i.e., the DOAs $\{\theta_k\}_{k=1}^K$ can be estimated from the locations of the K highest peaks of the Capon spatial spectrum (i.e., the array output power) $P_{\text{SCB}}(\theta)$ given by

$$P_{\text{SCB}}(\theta) = \frac{1}{f_{\text{SCB}}(\theta)} = \frac{1}{\mathbf{a}^H(\theta) \mathbf{R}^{-1} \mathbf{a}(\theta)}. \quad (10)$$

From (5), the eigenvalue decomposition (EVD) of the covariance matrix \mathbf{R} is given by [89]

$$\mathbf{R} = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^H = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{U}_s^H + \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{U}_n^H \quad (11)$$

where $\mathbf{U} = [\mathbf{U}_s, \mathbf{U}_n]$, $\mathbf{U}_s = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_K]$, $\mathbf{U}_n = [\mathbf{u}_{K+1}, \mathbf{u}_{K+2}, \dots, \mathbf{u}_L]$, $\mathbf{\Sigma} = \text{blkdiag}\{\mathbf{\Sigma}_s, \mathbf{\Sigma}_n\}$, $\mathbf{\Sigma}_s = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_K\}$, and $\mathbf{\Sigma}_n = \text{diag}\{\lambda_{K+1}, \lambda_{K+2}, \dots, \lambda_L\}$, in which $(\mathbf{u}_l, \lambda_l)$ is the l th eigenvector and eigenvalue pair with $\lambda_1 \geq \dots \geq \lambda_K \geq \lambda_{K+1} = \dots = \lambda_L = \sigma^2 \neq 0$, while \mathbf{U}_s and \mathbf{U}_n correspond to the signal and noise

subspaces, respectively, and $\mathbf{U} \mathbf{U}^H = \mathbf{U}^H \mathbf{U} = \mathbf{I}_L$. Then by substituting (11) into (9), we easily have

$$f_{\text{SCB}}(\theta) = \sum_{l=1}^L \frac{1}{\lambda_l} |\mathbf{a}^H(\theta) \mathbf{u}_l|^2 = \sum_{l=1}^L \frac{1}{\lambda_l} f_l(\theta) \quad (12)$$

where $f_l(\theta) = |\mathbf{a}^H(\theta) \mathbf{u}_l|^2$, while $\mathbf{a}^H(\theta_k) \mathbf{u}_l \neq 0$ for $l = 1, 2, \dots, K$, and $\mathbf{a}^H(\theta_k) \mathbf{u}_l = 0$ for $l = K+1, K+2, \dots, L$. As a result, we can obtain

$$f_{\text{SCB}}(\theta_k) = \sum_{l=1}^K \frac{1}{\lambda_l} f_l(\theta_k) \neq 0 \quad (13)$$

i.e., the estimated DOAs of the SCB are asymptotically biased and also not consistent. Obviously, we can easily find that the signals subspace is involved by the cost function $f_{\text{SCB}}(\theta)$, which is not orthogonal to the array response vector, and it may dominate the estimation performance in some cases. Hence the SCB performs worse than the MUSIC.

B. Review of Various Modifications of SCB

In the past few decades, various modifications were proposed in an attempt to improve the performance of the SCB, and they are briefly summarized as follows.

Firstly, by using a truncated array covariance matrix to reduce the influence of signal subspace components in (13), a modified Capon beamformer with noise eigenvectors (MCB-NEV) was presented [15], [32]

$$f_{\text{MCB-NEV}}(\theta) = \sum_{l=K+1}^L \frac{1}{\lambda_l} |\mathbf{a}^H(\theta) \mathbf{u}_l|^2. \quad (14)$$

Clearly the noise eigenvector is weighted by the corresponding inverse eigenvalue (i.e., $1/\lambda_l$) in the MCB-NEV, which will makes the effect of the noise eigenvector on the DOA estimation different for a finite number of snapshots, while the unity weight is used in the MUSIC.

Secondly, for the ULA [66], it was clarified that the “harmonic averaging” [45] (or “parallel resistor network averaging” [10]) effect of combining the LP spectra with all different models from the lowest to highest resolution results in the reduced resolution of the SCB, i.e.,

$$\frac{1}{P_{\text{SCB}}(\theta)} = \frac{1}{L} \sum_{l=1}^L \frac{1}{P_{\text{LP}}^{(l)}(\theta)} \quad (15)$$

where the LP spatial spectrum $P_{\text{LP}}^{(l)}(\theta)$ with the l th LP order is defined as [45]

$$P_{\text{LP}}^{(l)}(\theta) = \frac{\mathbf{e}_l^H \mathbf{R}_l^{-1} \mathbf{e}_l}{|\mathbf{e}_l^H \mathbf{R}_l^{-1} \mathbf{a}_l(\theta)|^2} \quad (16)$$

in which \mathbf{R}_l is the $l \times l$ subarray covariance matrix with the first l sensors, and $\mathbf{a}_l(\theta)$ is the corresponding response vector. By deleting the lower-order LP spectra to enhance the resolution, a pseudospectrum estimation method (PEM) was suggested [73], where its spatial spectrum is defined as

$$\frac{1}{P_{\text{PEM}}(\theta)} = \sum_{l=K}^L \frac{1}{P_{\text{LP}}^{(l)}(\theta)}. \quad (17)$$

Evidently the computational burden of the PEM is heavier than that of the SCB, and the number of incident signals is required.

On the other hand, by direct analogy with the SCB in (9), a class of non-quadratic spectral estimators (i.e., Pisarenko framework [86]) was proposed in [67], where the higher-order array covariance matrix is adopted in its non-quadratic cost function

$$f_{\text{PISARENKO}}(\theta) = (\mathbf{a}^H(\theta)\mathbf{R}^r\mathbf{a}(\theta))^{-\frac{1}{r}} \quad (18)$$

where r is an integer with $r \neq 0$. Obviously the SCB in (9) is obtained with $r = -1$. In [70], a stable nonlinear method (SNLM) was suggested, and its cost function $f_{\text{SNLM}}(\theta)$ is given by

$$f_{\text{SNLM}}(\theta) = \mathbf{a}^H(\theta)\mathbf{R}^{-\frac{1}{2}}(\mathbf{R} + \alpha^2\mathbf{R}^{-1})^{-p}\mathbf{R}^{-\frac{1}{2}}\mathbf{a}(\theta) \quad (19)$$

where $p = 0, 1, 2, \dots$, and α are the parameters introduced to emphasize the noise eigenvectors while deemphasize the signal eigenvectors with different weights, but it is difficult to determine them without proper a priori information.

Additionally a generalized Capon beamformer (GCB) was considered as follows [69], [71]:

$$f_{\text{GCB}}(\theta) = \frac{\mathbf{a}^H(\theta)\mathbf{R}^{-q}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{R}^{-q+1}\mathbf{a}(\theta)}. \quad (20)$$

When $q = 0, 1$, or 2 , the GCB will reduce to the Bartlett beamformer [17], [35], the SCB in (9), or the simple modified Capon beamformer (SMCB) [69], respectively. By using the squared array covariance matrix, another modified Capon beamformer (MCB) was recommended [10], [67], where its cost function $f_{\text{MCB}}(\theta)$ is given by

$$f_{\text{MCB}}(\theta) = \mathbf{a}^H(\theta)\mathbf{R}^{-2}\mathbf{a}(\theta) \quad (21)$$

but its performance was not shown therein. Further a weighted Capon beamformer (WCB) was proposed [38], where its cost function $f_{\text{WCB}}(\theta)$ is defined as

$$f_{\text{WCB}}(\theta) = \mathbf{a}^H(\theta)\mathbf{R}^{-1}\bar{\mathbf{W}}\mathbf{R}^{-1}\mathbf{a}(\theta) \quad (22)$$

where $\bar{\mathbf{W}}$ is the weighting matrix. Clearly the WCB is equal to the MCB in (21) when $\bar{\mathbf{W}} = \mathbf{I}_L$. Unfortunately, there is no existing prior technique to determine the optimal weight matrix.

In short, some variants were proposed from different perspectives to enhance the DOA estimation performance of the SCB as mentioned above. However, they usually need to know the number of incident signals in the MCB-NEV [15], [32] and the PEM [73], or alternatively, they need to select the appropriate parameters in the SNLM [70], the GCB [69], [71] and the WCB [38]. More importantly, except that the statistical analyses of a special case of the Pisarenko framework in (18), the SCB in (9) and the WCB in (22) were well studied in [37], [38], these modified methods were only verified by numerical simulations.

Remark 1: By exploiting the statistical results on the inverted Wishart distribution of the inverse sample covariance matrix (i.e., $\hat{\mathbf{R}}^{-1}$) [90] and by approximating the high-order inverse sample covariance matrix (i.e., $\hat{\mathbf{R}}^r$) with the contour integrals, the estimation performance of a special case of

Pisarenko framework (i.e., $\bar{f}_{\text{PISARENKO}}(\theta) = \mathbf{a}^H(\theta)\hat{\mathbf{R}}^r\mathbf{a}(\theta)$, for $r \leq -1$) and the SCB were analyzed with the first-order Taylor series expansion around the asymptotic DOA estimate and the theoretical inverse covariance matrix \mathbf{R}^{-1} in [39], [40], respectively. Additionally, by using the first-order Taylor series expansion around the asymptotic DOA estimate and the inverse sample covariance matrix $\hat{\mathbf{R}}^{-1}$, the asymptotic MSEs of the WCB and the SCB were studied [37], [38]. However, the derived analytical expression [40] is not strict and accurate for a small number of snapshots and/or at SNR (see Section V for details), while the statistical analysis of the WCB derived in [38] is only applicable to the MCB in (21) and it is still difficult to determine the optimal weighting matrix of the WCB [38].

C. Improved Capon Without Eigendecomposition

As shown in (12) and (13), the low resolution of the SCB is due to the combination of the signal eigenvectors $\{\mathbf{u}_l\}_{l=1}^K$ in the cost function $f_{\text{SCB}}(\theta)$, that are not orthogonal to the array response vector $\mathbf{a}(\theta)$. Differently from the classic subspace-based MUSIC, we consider eliminating the signal subspace components to improve the SCB estimation performance without knowing the number of incident signals in advance and performing eigendecomposition.

By defining a “pseudo-power” of the ICE as $P_{\text{ICE}} = \bar{\mathbf{w}}^H\mathbf{R}^m\bar{\mathbf{w}}$, where m is a positive integer constant, and $m \geq 2$, we can determine the weight $\bar{\mathbf{w}}$ by solving the following constrained quadratic problem:

$$\min_{\bar{\mathbf{w}}} \bar{\mathbf{w}}^H\mathbf{R}^m\bar{\mathbf{w}} \quad \text{s.t.} \quad \bar{\mathbf{w}}^H\mathbf{a}(\theta) = 1. \quad (23)$$

Consequently we easily get the weight vector $\bar{\mathbf{w}}_{\text{ICE}}$ as

$$\bar{\mathbf{w}}_{\text{ICE}} = \frac{\mathbf{R}^{-m}\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{R}^{-m}\mathbf{a}(\theta)}. \quad (24)$$

As a result, we can estimate the DOAs $\{\theta_k\}_{k=1}^K$ from the locations of K highest peaks of the ICE spatial spectrum $P_{\text{ICE}}(\theta)$ defined as

$$P_{\text{ICE}}(\theta) = \frac{1}{f_{\text{ICE}}(\theta)} = \frac{1}{\mathbf{a}^H(\theta)\mathbf{R}^{-m}\mathbf{a}(\theta)} \quad (25)$$

where the ICE cost function $f_{\text{ICE}}(\theta)$ is given by

$$f_{\text{ICE}}(\theta) = \mathbf{a}^H(\theta)\mathbf{R}^{-m}\mathbf{a}(\theta). \quad (26)$$

Obviously, the ICE turns into the SCB in (9) and the MCB in (21) with $m = 1$ or $m = 2$, respectively.

Then by substituting (11) into (26), the ICE cost function $f_{\text{ICE}}(\theta)$ can be expressed as [87]

$$\begin{aligned} f_{\text{ICE}}(\theta) &= \sum_{l=1}^L \frac{1}{\lambda_l^m} |\mathbf{a}^H(\theta)\mathbf{u}_l|^2 \\ &= \frac{1}{(\sigma^2)^m} \left(\sum_{l=1}^K \left(\frac{\sigma^2}{\lambda_l} \right)^m |\mathbf{a}^H(\theta)\mathbf{u}_l|^2 + f_{\text{MUSIC}}(\theta) \right). \end{aligned} \quad (27)$$

Clearly when the SNR is sufficiently large or the power order m becomes larger, we can obtain $(\sigma^2/\lambda_l)^m \rightarrow 0$ and $(\sigma^2)^m f_{\text{ICE}}(\theta) = f_{\text{MUSIC}}(\theta)$, i.e., the serious influence of the signal subspace components $\{\mathbf{u}_l\}_{l=1}^K$ (i.e., \mathbf{U}_s) on the DOA estima-

tion performance can be effectively eliminated, and hence the ICE cost function can be close to that of the MUSIC method, and the DOA estimation performance of the ICE will be improved with a resolution comparable to the MUSIC.

As mentioned above, the array covariance matrix \mathbf{R} should be estimated from the finite available array data $\{\mathbf{x}(n)\}_{n=1}^N$ as (6), then the implementation of the ICE DOA estimation can be summarized as follows.

1) Calculate the sample array covariance matrix $\hat{\mathbf{R}}$ with (6).

2) Calculate the matrix inversion $\hat{\mathbf{R}}^{-1}$ from $\hat{\mathbf{R}}$ by using some technique (i.e., the Gaussian elimination method), and calculate its m th power (i.e., $(\hat{\mathbf{R}}^{-1})^m$).

3) Estimate the directions $\{\theta_k\}_{k=1}^K$ by searching the K highest peaks of the spatial spectrum $\hat{P}_{\text{ICE}}(\theta)$ (the spectral approach), where $\hat{P}_{\text{ICE}}(\theta)$ is given by

$$\hat{P}_{\text{ICE}}(\theta) = \frac{1}{\hat{f}_{\text{ICE}}(\theta)} \quad (28)$$

in which

$$\hat{f}_{\text{ICE}}(\theta) = \mathbf{a}^H(\theta)(\hat{\mathbf{R}}^{-m})\mathbf{a}(\theta) \quad (29)$$

or by finding the phases of the K zeros of the polynomial $\hat{f}_{\text{ICE}}(z)$ closest to the unit circle in the z -plane for the ULA (the root approach), where $\hat{f}_{\text{ICE}}(z)$ is given by

$$\hat{f}_{\text{ICE}}(z) = z^{L-1} \mathbf{p}^H(z)(\hat{\mathbf{R}}^{-m})\mathbf{p}(z) \quad (30)$$

in which $\mathbf{p}(z) = [1, z, \dots, z^{L-1}]^T$, $z = e^{j2\pi d \sin \theta / \lambda}$, and d is the sensor spacing of the ULA.

Remark 2: In practice, there is a trade-off between good estimation performance and computational complexity. Since the computationally intensive eigendecomposition is not required in the SCB and some its variants, the computational complexity of these Capon-like methods is relatively small compared with the classic subspace-based MUSIC, and many effective algorithms were suggested to calculate the inverse sample covariance matrix (i.e., $\hat{\mathbf{R}}^{-1}$) in the literature (see, e.g., [91]–[97]). In fact, the computational complexity of calculating the m th power of the inverse sample covariance matrix (i.e., $(\hat{\mathbf{R}}^{-1})^m$) will be increased with the increase of the power order m . Hence, the trade-off value of the power order m should be determined by a balance between performance and computational complexity, and numerous experiments show that the power order $m = 4$ is generally sufficient for the ICE to achieve comparable performance (see Section V for details).

Remark 3: As shown in above, the implementation of the ICE involves the calculation of $\hat{\mathbf{R}}$, the inversion of $\hat{\mathbf{R}}^{-1}$ and the calculation of $(\hat{\mathbf{R}}^{-1})^m$, which approximately require $10L^2N + 4L^2$ flops, $25L^3 + 7L^2 + 24L$ flops, and $8(m-1)L^3$ flops, respectively, where a flop is defined as a floating-point addition or multiplication operation as adopted by MATLAB software, and the LU decomposition and Gaussian elimination can be used for inversion [98]. In addition, the rooting of $\hat{f}_{\text{ICE}}(z)$ also needs about $8L^3 + 8L$ flops. Hence the estimated number of MATLAB flops required by the ICE is nearly $10L^2N + 8(m-1)L^3 + 33L^3 + 8L^2$. The quantitative comparisons of computational complexities between the ICE and the

root-MUSIC in MATLAB flops are shown in Fig. 1, where the number of flops required by the root-MUSIC is rough $10L^2N + O(L^3) + 8(L-K)L^2 + 20L^2$ [99]. Obviously, the ICE is generally more efficient than the MUSIC for a moderate power order m (i.e., $m < 24$), even if the number of flops required by the ICE becomes very large in case of the large power order m .

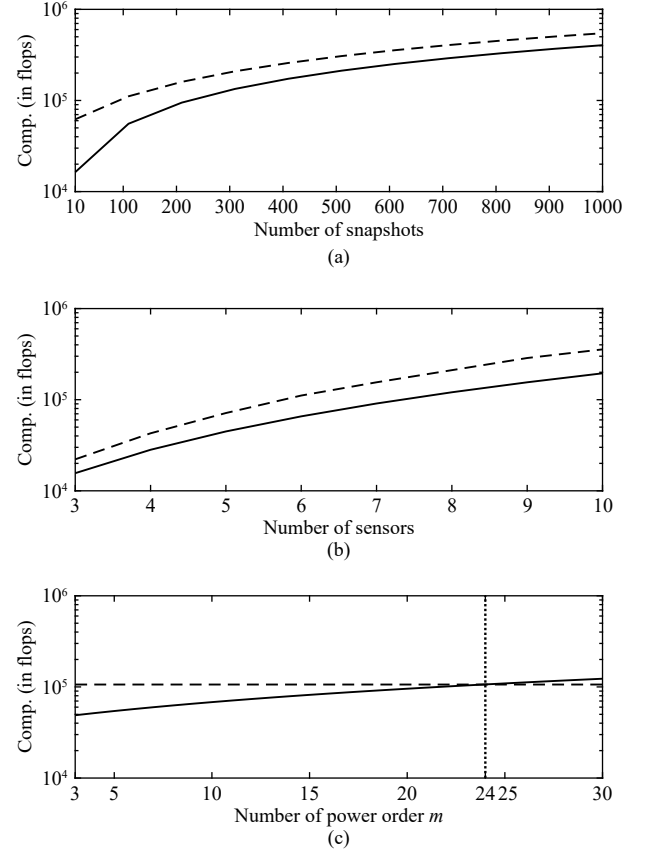


Fig. 1. Comparison of computational complexities between the MUSIC and the ICE in MATLAB flops versus (a) the number of snapshots ($L = 7$); (b) the number of sensors ($N = 200$); and (c) the number of power order ($N = 200$, $L = 7$) (dashed line: MUSIC, and solid line: ICE).

Remark 4: The ICE can be extended to the coherent (i.e., fully correlated) incident signals by using the forward or forward-backward spatial averaging (see, e.g., [100]–[102]), though the incident signals are considered to be uncorrelated to simplify the statistical analysis in this paper.

IV. STATISTICAL ANALYSIS

Because the ICE cost function $\hat{f}_{\text{ICE}}(\theta)$ in (29) is a complicated nonlinear function of the finite received array data, and its statistical behavior for “small” number of snapshots is difficult to analyze like the other DOA estimators [37], [38], [44], [103], we study the asymptotic statistical properties of the ICE for a large number of snapshots in this section.

A. Consistency of ICE Estimate

The consistency of the ICE estimate obtained by minimizing the cost function $\hat{f}_{\text{ICE}}(\theta)$ in (29) is given below.

Lemma 1: As the number of snapshots N tends to infinity, the ICE estimates $\{\hat{\theta}_k\}$ approach the true parameters $\{\theta_k\}$ with the probability one (w.p.1) under any of the following conditions i) the power order m is sufficiently large, or ii) the SNR is sufficiently large, or iii) the power order m and the SNR are both sufficiently large.

Proof: Firstly, from (6), when $N \rightarrow \infty$, we easily obtain the consistencies of the sample covariance matrix $\hat{\mathbf{R}}$ and its inverse $\hat{\mathbf{R}}^{-1}$ as $E\{\hat{\mathbf{R}}\} = \mathbf{R}$, $E\{\hat{\mathbf{R}}^{-1}\} = \mathbf{R}^{-1}$ w.p.1 [39], [42], [90], and consequently we have $E\{\hat{\mathbf{u}}_l\} = \mathbf{u}_l$ and $E\{\hat{\lambda}_l\} = \lambda_l$ for $l = 1, 2, \dots, L$. By using the eigenpairs of the sample covariance matrix $\hat{\mathbf{R}}$ in (6), from (11) and (27), the cost function $\hat{f}_{\text{ICE}}(\theta)$ in (29) can be rewritten as

$$\hat{f}_{\text{ICE}}(\theta) = \sum_{l=1}^K \frac{1}{\hat{\lambda}_l^m} |\mathbf{a}^H(\theta) \hat{\mathbf{u}}_l|^2 + \sum_{l=K+1}^L \frac{1}{\hat{\lambda}_l^m} |\mathbf{a}^H(\theta) \hat{\mathbf{u}}_l|^2 \quad (31)$$

where $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_K \geq \hat{\lambda}_{K+1} \geq \dots \geq \hat{\lambda}_L > 0$.

i) If $N \rightarrow \infty$ and $m \rightarrow \infty$, then $\sigma^2/\lambda_l < 1$ and $(\sigma^2/\lambda_l)^m \rightarrow 0$ for $l = 1, 2, \dots, K$, hence we get

$$\begin{aligned} \lim_{N \rightarrow \infty} \hat{f}_{\text{ICE}}(\theta) &= \frac{1}{\sigma^{2m}} \sum_{l=1}^K \left(\frac{\sigma^2}{\lambda_l} \right)^m |\mathbf{a}^H(\theta) \mathbf{u}_l|^2 + \frac{1}{\sigma^{2m}} f_{\text{MUSIC}}(\theta) \\ &\rightarrow \frac{1}{\sigma^{2m}} f_{\text{MUSIC}}(\theta). \end{aligned} \quad (32)$$

ii) If $N \rightarrow \infty$ and $\text{SNR} \rightarrow \infty$, then $\lambda_l \gg \sigma^2$ (i.e., $\sigma^2/\lambda_l \rightarrow 0$) for $l = 1, 2, \dots, K$, and similarly we obtain

$$\lim_{N \rightarrow \infty} \hat{f}_{\text{ICE}}(\theta) \rightarrow \frac{1}{\sigma^{2m}} f_{\text{MUSIC}}(\theta). \quad (33)$$

iii) If $N \rightarrow \infty$, $m \rightarrow \infty$ and $\text{SNR} \rightarrow \infty$, then $(\sigma^2/\lambda_l)^m \rightarrow 0$ for $l = 1, 2, \dots, K$, hence we also have $\lim_{N \rightarrow \infty} \hat{f}_{\text{ICE}}(\theta) \rightarrow (1/\sigma^{2m}) f_{\text{MUSIC}}(\theta)$.

As a result, under any of the above three conditions, we can get $\hat{f}_{\text{ICE}}(\theta) \rightarrow 1/(\sigma^2)^m f_{\text{MUSIC}}(\theta) = \hat{f}_{\text{ICE}}(\theta)$, where the MUSIC is a consistent estimator (see, e.g., [44]), and consequently the minima of $\hat{f}_{\text{ICE}}(\theta)$ are achieved if and only if $\theta = \theta_k$, i.e., $\hat{f}_{\text{ICE}}(\theta_k) = 0$ w.p.1. ■

B. Asymptotic MSE Expression

Now we study the asymptotic (for $N \gg 1$) MSE of the estimation error in order to evaluate the estimation accuracy of the ICE presented in Section III-C.

As discussed above, since the estimation performance of the ICE is dominated by the power order m , we represent the minimizers of the asymptotic ICE cost function $\hat{f}_{\text{ICE}}(\theta)$ in (26) with $\{\bar{\theta}_k\}_{k=1}^K$, where these asymptotic estimates $\{\bar{\theta}_k\}_{k=1}^K$ may be significantly different from the true DOAs $\{\theta_k\}_{k=1}^K$ under adverse conditions such as low SNR or/and small power order m (see, e.g., [37], [38], [42], [83]). Herein by defining the asymptotic bias $\Delta\theta_k$ and the additional bias $\Delta\bar{\theta}_k$ as $\Delta\theta_k = \bar{\theta}_k - \theta_k$ and $\Delta\bar{\theta}_k = \hat{\theta}_k - \bar{\theta}_k$, when the number of snapshots N is large, the ICE estimates $\{\hat{\theta}_k\}$ obtained with $\hat{f}_{\text{ICE}}(\theta)$ in (29) will fluctuate around $\{\bar{\theta}_k\}$ with a variance $\text{var}\{\hat{\theta}_k\}$, and hence the asymptotic MSE of the estimates $\{\hat{\theta}_k\}$ can be expressed by [37]–[39], [42]

$$\begin{aligned} \text{MSE}(\hat{\theta}_k) &= E\{(\hat{\theta}_k - \theta_k)^2\} \\ &= E\{(\Delta\bar{\theta}_k)^2\} + 2E\{\Delta\bar{\theta}_k\Delta\theta_k + (\Delta\theta_k)^2\} \\ &= \text{var}\{\hat{\theta}_k\} + (\Delta\theta_k)^2 \end{aligned} \quad (34)$$

where the variance $\text{var}\{\hat{\theta}_k\}$ is defined as

$$\text{var}\{\hat{\theta}_k\} = E\{(\Delta\bar{\theta}_k)^2\} \quad (35)$$

and $E\{\Delta\bar{\theta}_k\} = 0$. Obviously the actual error of the ICE estimate $\hat{\theta}_k$ can be obtained by the asymptotic bias $\Delta\theta_k$ and the additional bias $\Delta\bar{\theta}_k$. Then we can obtain the expression of the variance $\text{var}\{\hat{\theta}_k\}$ of the ICE by the following theorem.

Theorem 1: For the estimates $\hat{\theta}_k$ obtained by minimizing the ICE cost function in (29), its large-sample variance $\text{var}\{\hat{\theta}_k\}$ defined in (35) is given by

$$\text{var}\{\hat{\theta}_k\} = \frac{1}{2NH_k^2} \sum_{r=1}^m \sum_{p=1}^m (\Gamma_{rp} + \Upsilon_{rp}) \quad (36)$$

where

$$H_k = \text{Re}\{\mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{d}(\bar{\theta}_k) + \tilde{\mathbf{d}}^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{a}(\bar{\theta}_k)\} \quad (37)$$

$$\Gamma_{rp} = \text{Re}\{(\kappa_k^H \mathbf{R}^{-(m-r+p-2)} \mathbf{v}_k) \times (\kappa_k^H \mathbf{R}^{-(m+r-p-2)} \mathbf{v}_k)\} \quad (38)$$

$$\Upsilon_{rp} = (\kappa_k^H \mathbf{R}^{-(2m-r-p-1)} \kappa_k) \times (\mathbf{v}_k^H \mathbf{R}^{-(r+p-3)} \mathbf{v}_k) \quad (39)$$

while $\kappa_k = \mathbf{R}^{-1} \mathbf{d}(\bar{\theta}_k)$, $\mathbf{v}_k = \mathbf{R}^{-1} \mathbf{a}(\bar{\theta}_k)$, $\mathbf{d}(\bar{\theta}_k) = d\mathbf{a}(\theta)/d\theta|_{\theta=\bar{\theta}_k}$, and $\tilde{\mathbf{d}}(\bar{\theta}_k) = d\mathbf{d}(\theta)/d\theta|_{\theta=\bar{\theta}_k}$.

Proof: See Appendix. ■

More especially, from Theorem 1, after some straightforward manipulations, when the power order m is set to $m = 1$ or $m = 2$, we can easily get the corresponding asymptotic MSE expression of the ICE estimate.

Remark 5: From Theorem 1, by setting the power order as $m = 1$, we can easily obtain the asymptotic MSE of the SCB [31], [35] as

$$\begin{aligned} \text{MSE}_{\text{SCB}}\{\hat{\theta}_k\} &= \frac{1}{2NH_k^2} \left(\mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-1} \mathbf{d}(\bar{\theta}_k) \mathbf{a}^H(\bar{\theta}_k) \mathbf{R}^{-1} \mathbf{a}(\bar{\theta}_k) \right. \\ &\quad \left. - |\mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-1} \mathbf{a}(\bar{\theta}_k)|^2 \right) + (\bar{\theta}_k - \theta_k)^2 \end{aligned} \quad (40)$$

which equals to the expression derived in [37], [38], [42].

Remark 6: From Theorem 1, when the power order is set to $m = 2$, the asymptotic MSE of the MCB [10] is given by

$$\begin{aligned} \text{MSE}_{\text{MCB}}\{\hat{\theta}_k\} &= \frac{1}{2NH_k^2} \left(2(\kappa_k^H \kappa_k \mathbf{v}_k^H \mathbf{v}_k - |\kappa_k^H \mathbf{v}_k|^2) \right. \\ &\quad \left. + \kappa_k^H \mathbf{R} \kappa_k \mathbf{v}_k^H \mathbf{R}^{-1} \mathbf{v}_k + \kappa_k^H \mathbf{R}^{-1} \kappa_k \mathbf{v}_k^H \mathbf{R} \mathbf{v}_k \right. \\ &\quad \left. + 2\text{Re}\{\kappa_k^H \mathbf{R} \mathbf{v}_k \kappa_k^H \mathbf{R}^{-1} \mathbf{v}_k\} \right) + (\bar{\theta}_k - \theta_k)^2 \end{aligned} \quad (41)$$

which equals to the expression of the WCB with the weight matrix $\tilde{\mathbf{W}} = \mathbf{I}_L$ derived in [38].

Remark 7: The asymptotic bias $\Delta\bar{\theta}_k$ was approximated by exploiting the first- and second-order Taylor series expansion of $f_{\text{SCB}}(\theta)$ around $\bar{\theta}_k$ in [39], [42]. However, as discussed in Section III-C, since this bias may be larger for small power order m , the asymptotic estimate $\bar{\theta}_k$ obtained from the cost function $\hat{f}_{\text{ICE}}(\theta)$ in (26) can be used to directly calculate the

asymptotic bias $\Delta\bar{\theta}_k$ in (34), similarly to the statistical analyses of the SCB and the WCB studied in [37], [38].

C. Analytic and Numerical Study of Performance

In order to gain insights into the ICE, here we specialize in the case of one signal impinging on a ULA with sensor spacing d and study the asymptotic $\text{var}(\hat{\theta})$ (i.e., $\text{MSE}(\hat{\theta})$) of the ICE estimator in detail. Further we quantitatively compare the asymptotic MSE of the ICE with that of the MUSIC and the stochastic CRB, which is the lower bound on the estimation error for any unbiased estimator [44].

In this case (i.e., $K = 1$), we readily have [29]

$$\mathbf{A}(\theta) = \mathbf{a}(\theta_1) = [1, e^{j\omega(\theta_1)}, e^{j2\omega(\theta_1)}, \dots, e^{j(L-1)\omega(\theta_1)}]^T \quad (42)$$

$$\mathbf{d}(\theta_1) = j\bar{\omega}(\theta_1)[0, e^{j\omega(\theta_1)}, \dots, (L-1)e^{j(L-1)\omega(\theta_1)}]^T \quad (43)$$

$$\begin{aligned} \tilde{\mathbf{d}}(\theta_1) = & -j\bar{\omega}(\theta_1)[0, e^{j\omega(\theta_1)}, \dots, (L-1)e^{j(L-1)\omega(\theta_1)}]^T \\ & - \bar{\omega}^2(\theta_1)[0, e^{j\omega(\theta_1)}, \dots, (L-1)^2 e^{j(L-1)\omega(\theta_1)}]^T \end{aligned} \quad (44)$$

$$\mathbf{R}_s = r_s \mathbf{I}_L \quad (45)$$

$$\mathbf{R} = r_s \mathbf{a}(\theta_1) \mathbf{a}^H(\theta_1) + \sigma^2 \mathbf{I}_L \quad (46)$$

where $\omega(\theta) = 2\pi(d/\lambda)\sin\theta$, and $\bar{\omega}(\theta) = 2\pi(d/\lambda)\cos\theta$. To avoid a complication of notation, \mathbf{a} , \mathbf{d} , $\tilde{\mathbf{d}}$, and $\bar{\omega}$ are used as the brief notation for $\mathbf{a}(\theta_1)$, $\mathbf{d}(\theta_1)$, $\tilde{\mathbf{d}}(\theta_1)$, and $\bar{\omega}(\theta_1)$, respectively, in the following. From (44)–(46), we can obtain [29]

$$\mathbf{a}^H \mathbf{a} = L \quad (47)$$

$$\mathbf{d}^H \mathbf{a} = -j\bar{\omega} \sum_{l=1}^{L-1} l = -j\bar{\omega} \frac{L(L-1)}{2} \quad (48)$$

$$\mathbf{d}^H \mathbf{d} = \bar{\omega}^2 \sum_{l=1}^{L-1} l^2 = \bar{\omega}^2 \frac{L(L-1)(2L-1)}{6} \quad (49)$$

$$\begin{aligned} \tilde{\mathbf{d}}^H \mathbf{a} = & -j\bar{\omega} \sum_{l=1}^{L-1} l - \bar{\omega}^2 \sum_{l=1}^{L-1} l^2 \\ = & -j\bar{\omega} \frac{L(L-1)}{2} - \bar{\omega}^2 \frac{L(L-1)(2L-1)}{6}. \end{aligned} \quad (50)$$

By setting $\text{SNR} = r_s/\sigma^2$ and using matrix inversion lemma, from (48), we have

$$\mathbf{R}^{-1} = \sigma^{-2} (\mathbf{I}_L + \xi \mathbf{a} \mathbf{a}^H) \quad (51)$$

$$\mathbf{R}^{-m} = \sigma^{-2m} (\mathbf{I}_L + \xi \mathbf{a} \mathbf{a}^H)^m \quad (52)$$

where $\xi = -(L + \text{SNR}^{-1})^{-1}$. Then by using the fact that $\mathbf{a}^H \mathbf{a} = L$ and the binomial theorem, from (54), we can obtain

$$\begin{aligned} \mathbf{R}^{-m} = & \sigma^{-2m} (C(m, 0) \mathbf{I}_L + C(m, 1) \xi \mathbf{a} \mathbf{a}^H + C(m, 2) \xi^2 \\ & \times L \mathbf{a} \mathbf{a}^H + \dots + C(m, m) \xi^m L^{m-1} \mathbf{a} \mathbf{a}^H) \\ = & \sigma^{-2m} (\mathbf{I}_L + (C(m, 1) \xi + C(m, 2) \xi^2 L + \dots \\ & + C(m, m) \xi^m L^{m-1}) \mathbf{a} \mathbf{a}^H) \\ = & \sigma^{-2m} (\mathbf{I}_L + \frac{1}{L} ((1 + \xi L)^m - 1) \mathbf{a} \mathbf{a}^H) \end{aligned} \quad (53)$$

where $C(m, \bar{m})$ is the binomial coefficient given by $C(m, \bar{m}) = m!/\bar{m}!(m-\bar{m})!$ for $1 \leq \bar{m} \leq m$.

Finally, from (37)–(39), (49)–(52) and (55), after some straightforward calculations, we get

$$\begin{aligned} H_k = & \text{Re}\{\mathbf{d}^H \mathbf{R}^{-m} \mathbf{d} + \tilde{\mathbf{d}}^H \mathbf{R}^{-m} \mathbf{a}\} \\ = & -\bar{\omega}^2 \sigma^{-4m} \frac{L(L^2-1)}{12} ((L \times \text{SNR} + 1)^{-m} - 1) \end{aligned} \quad (54)$$

$$\begin{aligned} \Gamma_{rp} = & \text{Re}\{(\mathbf{d}^H \mathbf{R}^{-(m-r+p)} \mathbf{a})(\mathbf{d}^H \mathbf{R}^{-(m+r-p)} \mathbf{a})\} \\ = & -\bar{\omega}^2 \sigma^{-4m} \frac{L^2(L-1)}{12} ((3L-3)(L \times \text{SNR} + 1)^{-2m} \\ & + (L+1)(L \times \text{SNR} + 1)^{-(r+p-1)}) \end{aligned} \quad (55)$$

$$\begin{aligned} \Upsilon_{rp} = & (\mathbf{d}^H \mathbf{R}^{-(2m-r-p+1)} \mathbf{d})(\mathbf{a}^H \mathbf{R}^{-(r+p-1)} \mathbf{a}) \\ = & \bar{\omega}^2 \sigma^{-4m} \frac{L^2(L-1)^2}{4} (L \times \text{SNR} + 1)^{-2m}. \end{aligned} \quad (56)$$

Then by substituting (56)–(58) into (36), we can obtain the asymptotic variance $\text{var}_{\text{ICE}}(\hat{\theta})$ as

$$\begin{aligned} \text{var}_{\text{ICE}}(\hat{\theta}) \approx & \frac{1}{\bar{\omega}^2 N} \frac{6}{(L^2-1)} \frac{1}{((L \times \text{SNR} + 1)^{-m} - 1)^2} \\ & \cdot \sum_{r=1}^m \sum_{p=1}^m (L \times \text{SNR} + 1)^{-(r+p-1)}. \end{aligned} \quad (57)$$

Further when the number of sensors L and the power order m are reasonably larger, from (34) and (59), we can get the asymptotic MSE expression $\text{MSE}_{\text{ICE}}(\hat{\theta})$ as [29], [44]

$$\begin{aligned} \text{MSE}_{\text{ICE}}(\hat{\theta}) \approx & \frac{1}{\bar{\omega}^2 N} \frac{6}{\text{SNR}} \frac{1}{L(L^2-1)} \left(1 + \frac{1}{L} \frac{1}{\text{SNR}}\right) \\ = & \text{MSE}_{\text{MUSIC}}(\hat{\theta}) = \text{CRB}(\hat{\theta}). \end{aligned} \quad (58)$$

i.e., the asymptotic MSE of the ICE will near that of the MUSIC and the asymptotic MSE of the ICE asymptotically achieves the stochastic CRB.

V. NUMERICAL EXAMPLES

Now we evaluate the effectiveness of the ICE for estimating DOAs of uncorrelated narrowband signals impinging on the ULA and confirm the derived theoretical analysis through numerical examples, where the sensors are separated by a half-wavelength (i.e., $d = \lambda/2$), and two signals with equal power come from far-field with the angles θ_1 and θ_2 . In the simulations, some existing DOA estimators such as the SCB [31], the MCB/WCB with the weight matrix $\bar{\mathbf{W}} = \mathbf{I}_L$ [38], the MUSIC [20], [104], the GCB [71] (with $q = 4$), the sparse and parametric approach (SPA) [105], the enhanced principal-singular-vector utilization for modal analysis (EPUMA) [106], and the multi-snapshot Newtonized orthogonal matching pursuit (MNOMP) [107] are carried out for comparison, and the stochastic CRB [44], [103] is also calculated as a reference. Additionally we define an empirical root-MSE (ERMSE) by averaging the estimated directions $\{\hat{\theta}_k\}$ as

$$\text{ERMSE}\{\hat{\theta}_k\} = \sqrt{\frac{1}{K\bar{N}} \sum_{k=1}^K \sum_{i=1}^{\bar{N}} (\hat{\theta}_k^{(i)} - \theta_k)^2} \quad (59)$$

where $\hat{\theta}_k^{(i)}$ denotes the estimates at the i th trial, and \bar{N} is the

number of trials. The SNR is defined as the ratio of the power of the incident signals to that of the additive noise at each sensor, and the results shown below are all based on 1000 independent trials (i.e., $\bar{N} = 1000$).

Example 1—Estimation Performance Versus SNR: Two distinct DOAs are $\theta_1 = -4^\circ$ and $\theta_2 = 7^\circ$, and the SNR varies from -5 dB to 20 dB, while the number of sensors is $L = 5$, and the number of snapshots is $N = 100$.

When $\text{SNR} = 5$ dB, the averaged spatial spectra of the ICE with different power orders m such as $m = 2, 3, 4$ are plotted and compared with the SCB and the spectral MUSIC in Fig. 2. Clearly the SCB (i.e., the ICE with $m = 1$) cannot provide clear peaks at the incoming directions and fails to distinguish two DOAs due to the severe influence of signal subspace components at low SNR. However, by increasing the power order m , the influence of signal subspace components is efficiently weakened, so that the resolution of the ICE is gradually improved, and the DOAs can be successfully estimated with $m = 3, 4$ in this empirical scenario. Further when the power order m becomes larger, the spatial spectrum of the ICE can accurately approximate that of the MUSIC with a scale of $1/(\sigma^2)^m$.

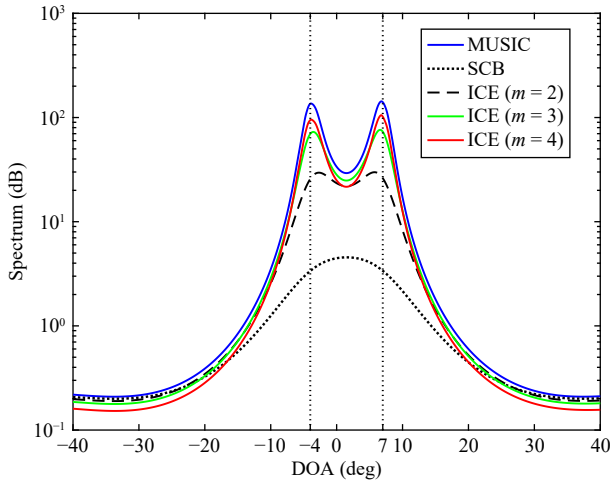


Fig. 2. The averaged spatial spectra of the SCB, the MUSIC and the ICE with different power order m for Example 1 ($\text{SNR} = 5$ dB, $N = 100$, $L = 5$, $\theta_1 = -4^\circ$, and $\theta_2 = 7^\circ$).

The empirical and theoretical RMSEs of the estimated DOAs against the SNR are shown in Fig. 3, where the theoretical results presented in [40] (denoted as VB-MSE), and the stochastic CRB are also plotted for comparison. Obviously and the ICE estimation performance becomes better than the SCB as the power order m increases especially in the case of low SNRs, and the ICE and the MUSIC provide similar small estimation errors for high SNRs. Compared with the theoretical derivations of [40], the empirical RMSEs of the ICE are very close to the theoretical ones derived in Section IV-C (except for low SNR and small m) and the difference between the theoretical RMSEs and the stochastic CRBs is small when the power order m is larger. Although the GCB performs better than the SCB, the ICE with higher power order generally

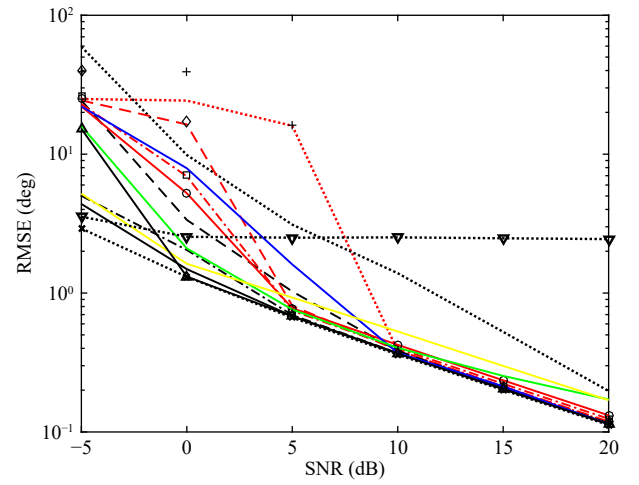


Fig. 3. The empirical and theoretical RMSEs of the ICE estimates versus the SNR for Example 1 (red dotted line: SCB; red dashed line: MCB; red dash-dotted line: ICE ($m = 3$); red solid line: ICE ($m = 4$); “+”: T.-MSE ($m = 1$); “ \diamond ”: T.-MSE ($m = 2$); “ \square ”: T.-MSE ($m = 3$); “ \circ ”: T.-MSE ($m = 4$); black dotted line: VB-MSE ($m = 1$); black dashed line: VB-MSE ($m = 2$); black dash-dotted line: VB-MSE ($m = 3$); black solid line: VB-MSE ($m = 4$); blue line: GCB; green line: SPA; yellow line: EPUMA; dotted line with “ ∇ ”: MNOMP; purple line: MUSIC; and dash-dotted line with “ \times ”: CRB ($N = 100$, $L = 5$, $\theta_1 = -4^\circ$, and $\theta_2 = 7^\circ$)).

outperforms the GCB. Although the SPA, the MNOMP and the EPUMA perform better only in the case of low SNRs and/or for few snapshots, but the SPA and the MNOMP are based on the sparse signal representation and require quite complicated calculations, while the EPUMA estimation criterion is equivalent to that of the direction estimation (MODE) [24], [25] (see [108] for details), and the EPUMA requires *a priori* knowledge about the number of incident signals and several computationally complex iterations.

Example 2—Estimation Performance Versus Number of Snapshots: The simulation conditions are similar to those in Example 1, except that the SNR is set at 0 dB, and the number of snapshots varies from 10 to 1000 .

As illustrated in Fig. 4, the estimation performance of the ICE can be effectively improved by increasing the number of snapshots N or the power order m , and the ICE with larger power order performs better than the SCB and the MCB. Obviously as the number of snapshots increases, both the empirical and theoretical RMSEs of the ICE decrease significantly, and the empirical RMSEs agree with the theoretical RMSEs very well for $m > 1$, while the RMSEs of the SCB remain large for all numbers of snapshots due to the lower SNR. In addition, we can find that the empirical and theoretical RMSEs of the ICE become close to the RMSE of the MUSIC and the CRB by increasing the power order m , although the number of snapshots is small. Although the ICE performs worse than the computationally complicated sparsity-based MNOMP and SPA and the eigentrusture-based EPUMA for small number of snapshots, the ICE generally performs better than the EPUMA, the MNOMP, and the SPA by increasing the number of snapshots.

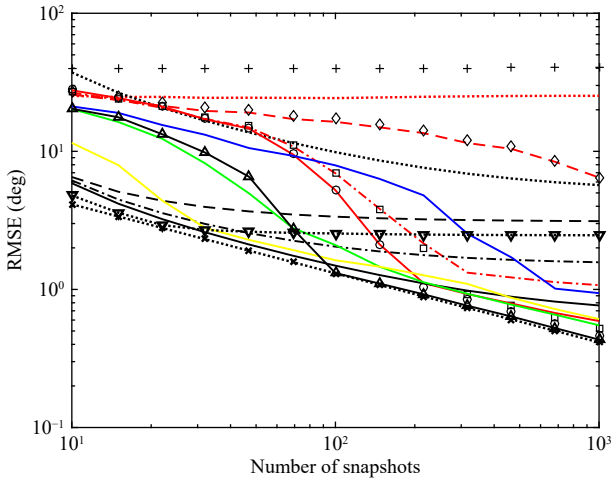


Fig. 4. The empirical and theoretical RMSEs of the ICE estimation versus the number of snapshots for Example 2 (red dotted line: SCB; red dashed line: MCB; red dash-dotted line: ICE ($m=3$); red solid line: ICE ($m=4$); “+”: T-MSE ($m=1$); “ \diamond ”: T-MSE ($m=2$); “ \square ”: T-MSE ($m=3$); “ \circ ”: T-MSE ($m=4$); black dotted line: VB-MSE ($m=1$); black dashed line: VB-MSE ($m=2$); black dash-dotted line: VB-MSE ($m=3$); black solid line: VB-MSE ($m=4$); blue line: GCB; green line: SPA; yellow line: EPUMA; dotted line with “ ∇ ”: MNOMP; purple line: MUSIC; and dash-dotted line with “ \times ”: CRB) (SNR = 0dB, $L=5$, $\theta_1 = -4^\circ$, and $\theta_2 = 7^\circ$).

Example 3—Estimation Performance With Versus Angular Separation: The simulation conditions are similar to those in Example 1, except that the SNR is fixed at 0 dB, and the DOAs of two incident signals are -4° and $-4^\circ + \Delta\theta$, respectively, where $\Delta\theta$ is varied from 5° to 37° with $\Delta\theta = 4^\circ$. As shown in Fig. 5, the ICE with the larger power order m generally estimates the directions of closely spaced signals more accurately with a much smaller RMSE than the SCB, the empirical RMSEs of the estimated DOAs are close to the theoretical RMSEs derived in Section IV. Additionally the empirical and theoretical RMSEs decrease with with increasing angular separation rather than monotonously.

VI. CONCLUSIONS

In this paper, a modified Capon estimator called the ICE was investigated for DOA estimation of uncorrelated narrow-band signals, where the higher-order inverse array covariance matrix is used in the Capon-like cost function. The relationship between the ICE and the MUSIC was studied, where the serious influence of the signal subspace components on the DOA estimation can be eliminated by increasing the power order of the covariance matrix, so that the performance of ICE is better than the SCB. Further the statistical properties of the ICE were analyzed, and the closed-form asymptotic MSE expressions of the DOA estimates were derived. Finally the effectiveness of the ICE and the validity of the theoretical analysis are verified through numerical examples.

APPENDIX

PROOF OF THEOREM 1

Proof: Since the estimate $\hat{\theta}_k$ is obtained by minimizing the

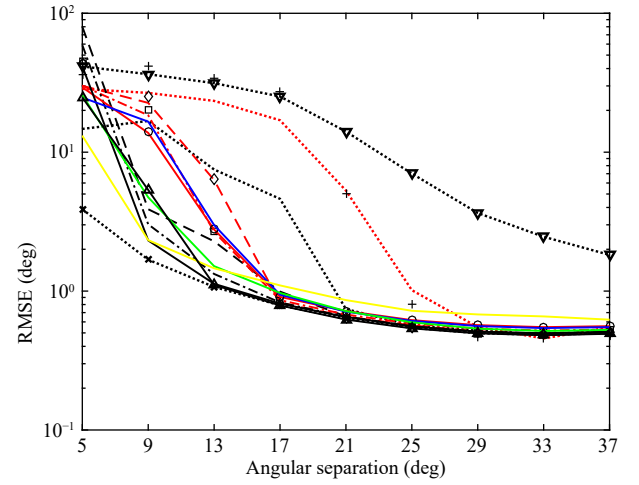


Fig. 5. The empirical and theoretical RMSEs of the ICE estimation versus the angular separation for Example 3 (red dotted line: SCB; red dashed line: MCB; red dash-dotted line: ICE ($m=3$); red solid line: ICE ($m=4$); “+”: T-MSE ($m=1$); “ \diamond ”: T-MSE ($m=2$); “ \square ”: T-MSE ($m=3$); “ \circ ”: T-MSE ($m=4$); black dotted line: VB-MSE ($m=1$); black dashed line: VB-MSE ($m=2$); black dash-dotted line: VB-MSE ($m=3$); black solid line: VB-MSE ($m=4$); blue line: GCB; green line: SPA; yellow line: EPUMA; dotted line with “ ∇ ”: MNOMP; purple line: MUSIC; and dash-dotted line with “ \times ”: CRB) (SNR = 0dB, $N=100$, $L=5$, $\theta_1 = -4^\circ$, and $\theta_2 = -4^\circ + \Delta\theta$).

cost function $\hat{f}_{ICE}(\theta)$ in (29), for a sufficiently large number of snapshots N , we can obtain the second-order approximation of the derivative of $\hat{f}_{ICE}(\theta)$ about the parameter $\bar{\theta}_k$ as (cf., [27], [29], [38], [44], [103], [109], and references therein)

$$0 = f'(\hat{\theta}_k) \approx f'(\bar{\theta}_k) + f''(\bar{\theta}_k)(\hat{\theta}_k - \bar{\theta}_k) \quad (60)$$

where $\hat{f}_{ICE}(\theta)$ is represented by $f(\theta)$ for simplicity of notation, the second- and higher order terms in (60) can be neglected, and the first- and second-order derivatives of $f(\theta)$ with respect to the scalar variable θ are given by

$$f'(\theta) = \frac{df(\theta)}{d\theta} = 2\text{Re}\{d^H(\theta)\hat{\mathbf{R}}^{-m}\mathbf{a}(\theta)\} \quad (61)$$

$$f''(\theta) = \frac{df'(\theta)}{d\theta} = 2\text{Re}\{d^H(\theta)\hat{\mathbf{R}}^{-m}d(\theta) + \tilde{d}^H(\theta)\hat{\mathbf{R}}^{-m}\mathbf{a}(\theta)\}. \quad (62)$$

From (60), the estimation error (i.e., the additional bias) $\Delta\bar{\theta}_k$ can be asymptotically obtained as

$$\begin{aligned} \Delta\bar{\theta}_k &\approx -\frac{f'(\bar{\theta}_k)}{f''(\bar{\theta}_k)} \approx -\frac{\text{Re}\{d^H(\bar{\theta}_k)\hat{\mathbf{R}}^{-m}\mathbf{a}(\bar{\theta}_k)\}}{\text{Re}\{d^H(\bar{\theta}_k)\mathbf{R}^{-m}d(\bar{\theta}_k) + \tilde{d}^H(\bar{\theta}_k)\mathbf{R}^{-m}\mathbf{a}(\bar{\theta}_k)\}} \\ &= -\frac{\text{Re}\{\mu_k\}}{H_k} \end{aligned} \quad (63)$$

where the sample array covariance matrix $\hat{\mathbf{R}}$ in the denominator of (63) (i.e., (62)) is replaced by the true one \mathbf{R} without affecting the asymptotic property of estimate $\hat{\theta}_k$ [27], [29], [38], [44], [103], [109], [110], while

$$\mu_k = d^H(\bar{\theta}_k)\mathbf{R}^{-m}\mathbf{a}(\bar{\theta}_k). \quad (64)$$

Consequently, from (63), the variance of the estimation error $\Delta\theta_k$ is given by

$$\text{var}\{\hat{\theta}_k\} = E\{(\Delta\theta_k)^2\} \approx \frac{1}{2H_k^2} \text{Re}\{E\{\mu_k^2\} + E\{|\mu_k|^2\}\} \quad (65)$$

where the identity that $\text{Re}\{\mu_i\}\text{Re}\{\mu_k\} = 0.5(\text{Re}\{\mu_i\mu_k\} + \text{Re}\{\mu_i\mu_k^*\})$ is used implicitly.

Now by defining an auxiliary variable $\bar{\mu}_k$ with the true parameters $\mathbf{d}^H(\bar{\theta}_k)$, \mathbf{R}^{-1} and $\mathbf{a}(\bar{\theta}_k)$ being

$$\bar{\mu}_k = \mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{a}(\bar{\theta}_k) \quad (66)$$

and by using the following formula for the derivative of matrix power term with respect to the matrix [111]

$$\frac{d\mathbf{a}^T \mathbf{X}^n \mathbf{b}}{d\mathbf{X}} = \sum_{r=1}^n (\mathbf{X}^{n-r})^T \mathbf{a} \mathbf{b}^T (\mathbf{X}^{r-1})^T \quad (67)$$

the first-order Taylor series expansion of μ_k in (64) around the inverse covariance matrix \mathbf{R}^{-1} is given by

$$\begin{aligned} \mu_k &\approx \bar{\mu}_k + \text{tr}\left\{\left(\frac{d\bar{\mu}_k}{d(\mathbf{R}^{-1})}\right)^T (\hat{\mathbf{R}}^{-1} - \mathbf{R}^{-1})\right\} \\ &= \bar{\mu}_k + \sum_{r=1}^m \text{tr}\{(\mathbf{R}^{-(m-r)})^T \mathbf{d}^*(\bar{\theta}_k) \mathbf{a}^T(\bar{\theta}_k) \\ &\quad \times (\mathbf{R}^{-(r-1)})^T \Delta_{\mathbf{R}^{-1}}\} \\ &= \bar{\mu}_k + \sum_{r=1}^m \mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-(m-r)} \Delta_{\mathbf{R}^{-1}} \mathbf{R}^{-(r-1)} \mathbf{a}(\bar{\theta}_k) \end{aligned} \quad (68)$$

where the second- and higher-order terms $O(\Delta_{\mathbf{R}^{-1}})$ in (68) can be neglected since $\hat{\mathbf{R}}^{-1}$ converges to \mathbf{R}^{-1} as $N \rightarrow \infty$ [44], [103], and the bias $\Delta_{\mathbf{R}^{-1}}$ of the sample inverse array covariance matrix $\hat{\mathbf{R}}^{-1}$ is given by [37], [38]

$$\begin{aligned} \Delta_{\mathbf{R}^{-1}} &= \hat{\mathbf{R}}^{-1} - \mathbf{R}^{-1} = \hat{\mathbf{R}}^{-1} (\mathbf{R} - \hat{\mathbf{R}}) \mathbf{R}^{-1} \\ &= (\mathbf{R}^{-1} + \Delta \mathbf{Y}) \Delta \mathbf{X} \mathbf{R}^{-1} \\ &= \mathbf{R}^{-1} \Delta \mathbf{X} \mathbf{R}^{-1} + \Delta \mathbf{Y} \Delta \mathbf{X} \mathbf{R}^{-1} \\ &\approx -\mathbf{R}^{-1} (\hat{\mathbf{R}} - \mathbf{R}) \mathbf{R}^{-1} \end{aligned} \quad (69)$$

where $\Delta \mathbf{X} = \hat{\mathbf{R}} - \mathbf{R}$ and $\Delta \mathbf{Y} = \hat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}$, while the second-order small quantity $\Delta \mathbf{Y} \Delta \mathbf{X} \mathbf{R}^{-1}$ has been neglected. By substituting (69) into (68), we obtain

$$\begin{aligned} \mu_k &\approx \bar{\mu}_k - \sum_{r=1}^m \kappa_k^H \mathbf{R}^{-(m-r)} (\hat{\mathbf{R}} - \mathbf{R}) \mathbf{R}^{-(r-1)} \mathbf{v}_k \\ &= (m+1) \mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{a}(\bar{\theta}_k) + \sum_{r=1}^m \bar{\mu}_{kr} \end{aligned} \quad (70)$$

where

$$\bar{\mu}_{kr} = -\kappa_k^H \mathbf{R}^{-(m-r)} \hat{\mathbf{R}} \mathbf{R}^{-(r-1)} \mathbf{v}_k \quad (71)$$

for $r = 1, 2, \dots, m$. In addition, since the asymptotic estimates $\{\bar{\theta}_k\}_{k=1}^K$ are the minimizers of the cost function $f_{\text{ICE}}(\theta)$ in (26), we have

$$0 = f'_{\text{ICE}}(\bar{\theta}_k) = \left. \frac{df_{\text{ICE}}(\theta)}{d\theta} \right|_{\theta=\bar{\theta}_k} = 2\text{Re}\{\mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{a}(\bar{\theta}_k)\}. \quad (72)$$

Hence, from (70)–(72), we obtain the terms $E\{\mu_k^2\}$ and $E\{|\mu_k|^2\}$ in (65) as

$$\begin{aligned} E\{\mu_k^2\} &= E\{\bar{\mu}_{k1}^2 + \bar{\mu}_{k1}\bar{\mu}_{k2} + \dots + \bar{\mu}_{k1}\bar{\mu}_{km} \\ &\quad + \bar{\mu}_{k2}\bar{\mu}_{k1} + \bar{\mu}_{k2}^2 + \dots + \bar{\mu}_{k2}\bar{\mu}_{km} + \dots \\ &\quad + \bar{\mu}_{km}\bar{\mu}_{k1} + \bar{\mu}_{km}\bar{\mu}_{k2} + \dots + \bar{\mu}_{km}^2\} \end{aligned} \quad (73)$$

$$\begin{aligned} E\{|\mu_k|^2\} &= E\{|\bar{\mu}_{k1}|^2 + \bar{\mu}_{k1}\bar{\mu}_{k2}^* + \dots + \bar{\mu}_{k1}\bar{\mu}_{km}^* \\ &\quad + \bar{\mu}_{k2}\bar{\mu}_{k1}^* + |\bar{\mu}_{k2}|^2 + \dots + \bar{\mu}_{k2}\bar{\mu}_{km}^* + \dots \\ &\quad + \bar{\mu}_{km}\bar{\mu}_{k1}^* + \bar{\mu}_{km}\bar{\mu}_{k2}^* + \dots + |\bar{\mu}_{km}|^2\} \end{aligned} \quad (74)$$

where the fact $\text{Re}\{\mathbf{d}^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{a}(\bar{\theta}_k)\} = 0$ in (72) is implicitly used.

Under the basic assumptions on the data model, and by using the well-known formula for the expectation of four Gaussian random variables with zero-mean (e.g., [112])

$$\begin{aligned} E\{x_1 x_2 x_3 x_4\} &= E\{x_1 x_2\} E\{x_3 x_4\} + E\{x_1 x_3\} E\{x_2 x_4\} \\ &\quad + E\{x_1 x_4\} E\{x_2 x_3\} \end{aligned} \quad (75)$$

and by performing some straightforward manipulations, we can obtain the terms in (73) and (74) as

$$\begin{aligned} E\{\bar{\mu}_{kr}\bar{\mu}_{kp}\} &= \frac{1}{N^2} \sum_{n=1}^N \sum_{t=1}^N E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{R}^{-(r-1)} \mathbf{v}_k \\ &\quad \times \kappa_k^H \mathbf{R}^{-(m-p)} \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{R}^{-(p-1)} \mathbf{v}_k\} \\ &= \frac{1}{N^2} \sum_{n=1}^N \sum_{t=1}^N \{E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{R}^{-(r-1)} \mathbf{v}_k\} \\ &\quad \times E\{\kappa_k^H \mathbf{R}^{-(m-p)} \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{R}^{-(p-1)} \mathbf{v}_k\} \\ &\quad + E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^T(t) (\mathbf{R}^{-(m-p)})^* \kappa_k^*\} \\ &\quad \times E\{\mathbf{v}_k^T (\mathbf{R}^{-(r-1)})^* \mathbf{x}^*(n) \mathbf{x}^H(t) \mathbf{R}^{-(p-1)} \mathbf{v}_k\} \\ &\quad + E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^H(t) \mathbf{R}^{-(p-1)} \mathbf{v}_k\} \\ &\quad \times E\{\mathbf{v}_k^T (\mathbf{R}^{-(r-1)})^* \mathbf{x}^*(n) \mathbf{x}^T(t) (\mathbf{R}^{-(m-p)})^* \kappa_k^*\}\} \\ &= (\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k)^2 + 0 \\ &\quad + \frac{1}{N} \kappa_k^H \mathbf{R}^{-(m-r+p-2)} \mathbf{v}_k \mathbf{v}_k^T (\mathbf{R}^{-(m+r-p-2)})^* \kappa_k^* \\ &= (\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k)^2 + \frac{1}{N} (\kappa_k^H \mathbf{R}^{-(m-r+p-2)} \mathbf{v}_k) \\ &\quad \times (\kappa_k^H \mathbf{R}^{-(m+r-p-2)} \mathbf{v}_k) \end{aligned} \quad (76)$$

and

$$\begin{aligned} E\{\bar{\mu}_{kr}\bar{\mu}_{kp}^*\} &= \frac{1}{N^2} \sum_{n=1}^N \sum_{t=1}^N E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{R}^{-(r-1)} \mathbf{v}_k \\ &\quad \times \kappa_k^T (\mathbf{R}^{-(m-p)})^* \mathbf{x}^*(t) \mathbf{x}^T(t) (\mathbf{R}^{-(p-1)})^* \mathbf{v}_{mk}^*\} \\ &= \frac{1}{N^2} \sum_{n=1}^N \sum_{t=1}^N \{E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{R}^{-(r-1)} \mathbf{v}_k\} \\ &\quad \times E\{\kappa_k^T (\mathbf{R}^{-(m-p)})^* \mathbf{x}^*(t) \mathbf{x}^T(t) (\mathbf{R}^{-(p-1)})^* \mathbf{v}_{mk}^*\} \\ &\quad + E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^H(t) \mathbf{R}^{-(m-p)} \kappa_k\} \\ &\quad \times E\{\mathbf{v}_k^H \mathbf{R}^{-(r-1)} \mathbf{x}(n) \mathbf{x}^H(t) \mathbf{R}^{-(p-1)} \mathbf{v}_k\} \end{aligned}$$

$$\begin{aligned}
& + E\{\kappa_k^H \mathbf{R}^{-(m-r)} \mathbf{x}(n) \mathbf{x}^T(t) (\mathbf{R}^{-(p-1)})^* \mathbf{v}_{mk}^*\} \\
& \times E\{\mathbf{v}_k^T (\mathbf{R}^{-(r-1)})^* \mathbf{x}^*(n) \mathbf{x}^H(t) \mathbf{R}^{-(m-p)} \kappa_k\} \\
& = |\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k|^2 + \frac{1}{N} \kappa_k^H \mathbf{R}^{-(2m-r-p-1)} \kappa_k \\
& \quad \times \mathbf{v}_k^T (\mathbf{R}^{-(r+p-3)})^* \mathbf{v}_k^* + 0 \\
& = |\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k|^2 + \frac{1}{N} (\kappa_k^H \mathbf{R}^{-(2m-r-p-1)} \kappa_k) \\
& \quad \times (\mathbf{v}_k^H \mathbf{R}^{-(r+p-3)} \mathbf{v}_k). \tag{77}
\end{aligned}$$

Additionally by using the relation in (72), we can get

$$\begin{aligned}
& \text{Re}\{(\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k)^2 + |\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k|^2\} \\
& = \left(\text{Re}\{(\kappa_k^H \mathbf{R}^{-m+2} \mathbf{v}_k)\}\right)^2 \\
& = \left(\text{Re}\{d^H(\bar{\theta}_k) \mathbf{R}^{-m} \mathbf{a}(\bar{\theta}_k)\}\right)^2 = 0. \tag{78}
\end{aligned}$$

Therefore by substituting (73)–(78) into (65) and performing some straightforward manipulations, the large-sample variance $\text{var}\{\hat{\theta}_k\}$ in (36) can be established immediately. ■

REFERENCES

- [1] S. Haykin, *Array Signal Processing*. Englewood Cliffs, USA: Prentice-Hall, 1985.
- [2] S. U. Pillai and C. S. Burrus, *Array Signal Processing*. New York, USA: Springer-Verlag, 1989.
- [3] S. Haykin, *Advances in Spectrum Analysis and Array Processing, Volumes I–III*. Englewood Cliffs, USA: Prentice-Hall, 1991–1995.
- [4] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, USA: Prentice Hall, 1993.
- [5] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, USA: Prentice-Hall, 1997.
- [6] P. R. P. Hoole, *Smart Antennas and Signal Processing: For Communications, Biomedical and Radar Systems*. Ashurst, UK: WIT Press, 2001.
- [7] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York: John Wiley & Sons, 2002.
- [8] S. M. Kay and S. L. Marple, “Spectrum analysis — A modern perspective,” *Proc. IEEE*, vol. 69, no. 11, pp. 1380–1419, Nov. 1981.
- [9] S. L. Marple Jr, *Digital Spectral Analysis*. Englewood Cliffs, USA: Prentice-Hall, 1987.
- [10] S. M. Kay, *Modern Spectral Estimation: Theory and Applications*. Englewood Cliffs, USA: Prentice Hall, 1988.
- [11] S. Li, R. X. He, B. Lin, and F. Sun, “DOA estimation based on sparse representation of the fractional lower order statistics in impulsive noise,” *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 4, pp. 860–868, Jun. 2018.
- [12] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, USA: Prentice-Hall, 2005.
- [13] A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, *Academic Press Library in Signal Processing, Volume 3: Array and Statistical Signal Processing*. New York, USA: Academic Press, 2014.
- [14] Y. B. Gao, “Adaptive generalized eigenvector estimating algorithm for hermitian matrix pencil,” *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 11, pp. 1967–1979, Nov. 2022.
- [15] D. H. Johnson, “The application of spectral estimation methods to bearing estimation problems,” *Proc. IEEE*, vol. 70, no. 9, pp. 1018–1028, Sept. 1982.
- [16] H. Krim and M. Viberg, “Two decades of array signal processing research: The parametric approach,” *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
- [17] L. C. Godara, “Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations,” *Proc. IEEE*, vol. 85, no. 8, pp. 1195–1245, Aug. 1997.
- [18] B. D. Van Veen and K. M. Buckley, “Beamforming: A versatile approach to spatial filtering,” *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
- [19] V. F. Pisarenko, “The retrieval of harmonics from a covariance function,” *Geophys. J. Roy. Astron. Soc.*, vol. 33, no. 3, pp. 347–366, Sept. 1973.
- [20] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [21] D. W. Tufts and R. Kumaresan, “Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood,” *Proc. IEEE*, vol. 70, no. 9, pp. 975–989, Sept. 1982.
- [22] R. Kumaresan and D. W. Tufts, “Estimating the angles of arrival of multiple plane waves,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 1, pp. 134–139, Jan. 1983.
- [23] 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.
- [24] P. Stoica and K. C. Sharman, “Novel eigenanalysis method for direction estimation,” *IEE Proc. F (Radar Signal Process.)*, vol. 137, no. 1, pp. 19–26, Feb. 1990.
- [25] P. Stoica and K. C. Sharman, “Maximum likelihood methods for direction-of-arrival estimation,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1132–1143, Jul. 1990.
- [26] M. Viberg and B. Ottersten, “Sensor array processing based on subspace fitting,” *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1110–1121, May 1991.
- [27] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, “Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing,” in *Radar Array Processing*, S. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin, Germany: Springer-Verlag, 1993, pp. 99–151.
- [28] L. C. Liu, Y. Li, and S. M. Kuo, “Feed-forward active noise control system using microphone array,” *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 5, pp. 946–952, Sept. 2018.
- [29] J. M. Xin and A. Sano, “Computationally efficient subspace-based method for direction-of-arrival estimation without eigendecomposition,” *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 876–893, Apr. 2004.
- [30] Y. W. Wang, J. Li, and P. Stoica, “Rank-deficient robust Capon filter bank approach to complex spectral estimation,” *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2713–2726, Aug. 2005.
- [31] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.
- [32] D. H. Johnson and S. DeGraaf, “Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, no. 4, pp. 638–647, Aug. 1982.
- [33] C. D. Seligson, “Comments on ‘High-resolution frequency-wavenumber spectrum analysis’,” *Proc. IEEE*, vol. 58, no. 6, pp. 947–949, Jun. 1970.
- [34] J. Capon and N. R. Goodman, “Probability distributions for estimators of the frequency-wavenumber spectrum,” *Proc. IEEE*, vol. 58, no. 10, pp. 1785–1786, Oct. 1970.
- [35] R. T. Lacoss, “Data adaptive spectral analysis methods,” *Geophysics*, vol. 36, no. 4, pp. 661–675, Aug. 1971.
- [36] T. Marzetta, “A new interpretation of Capon’s maximum likelihood method of frequency-wavenumber spectral estimation,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 2, pp. 445–449, Apr. 1983.

- [37] P. Händel, P. Stoica, and T. Söderström, "Capon method for DOA estimation: Accuracy and robustness aspects," in *Proc. IEEE Winter Workshop on Nonlinear Digital Signal Processing*, Tampere, Finland, 1993, pp. P_7.1–P_7.5.
- [38] P. Stoica, P. Händel, and T. Söderström, "Study of Capon method for array signal processing," *Circ. Syst. Signal Process.*, vol. 14, no. 6, pp. 749–770, Apr. 1995.
- [39] C. Vaidyanathan and K. M. Buckley, "Performance analysis of the MVDR spatial spectrum estimator," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1427–1437, Jun. 1995.
- [40] C. Vaidyanathan and K. M. Buckley, "Performance analysis of DOA estimation based on nonlinear functions of covariance matrix," *Signal Process.*, vol. 50, no. 1–2, pp. 5–16, Apr. 1996.
- [41] C. Vaidyanathan and K. M. Buckley, "Performance analysis of the enhanced minimum variance spatial spectrum estimator," *IEEE Trans. Signal Process.*, vol. 46, no. 8, pp. 2202–2206, Aug. 1998.
- [42] M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and Capon direction estimation," *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2291–2304, Sept. 1998.
- [43] J. Benesty, J. D. Chen, and Y. T. Huang, "A generalized MVDR spectrum," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 827–830, Dec. 2005.
- [44] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 720–741, May 1989.
- [45] U. Nickel, "Algebraic formulation of Kumaresan-Tufts superresolution method, showing relation to ME and MUSIC methods," *IEEE Proc. F (Commun., Radar Signal Process.)*, vol. 135, no. 1, pp. 7–10, Feb. 1988.
- [46] A. Jakobsson, S. L. Marple, and P. Stoica, "Computationally efficient two-dimensional capon spectrum analysis," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2651–2661, Sept. 2000.
- [47] P. Stoica, Z. S. Wang, and J. Li, "Robust capon beamforming," *IEEE Signal Process. Lett.*, vol. 10, no. 6, pp. 172–175, Jun. 2003.
- [48] J. Li, P. Stoica, and Z. S. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
- [49] R. Abrahamsson, A. Jakobsson, and P. Stoica, "A Capon-like spatial spectrum estimator for correlated sources," in *Proc. 12th European Signal Processing Conf.*, Vienna, Austria, 2004, pp. 1265–1268.
- [50] L. Du, T. Yardibi, J. Li, and P. Stoica, "Review of user parameter-free robust adaptive beamforming algorithms," *Digital Signal Process.*, vol. 19, no. 4, pp. 567–582, Jul. 2009.
- [51] J. Yang, X. C. Ma, C. H. Hou, and Y. C. Liu, "Shrinkage-based Capon and APES for spectral estimation," *IEEE Signal Process. Lett.*, vol. 16, no. 10, pp. 869–872, Oct. 2009.
- [52] P. Lopez-Dekker and J. J. Mallorqui, "Capon and APES-based SAR processing: Performance and practical considerations," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2388–2402, May 2010.
- [53] A. Aubry, V. Carotenuto, and A. De Maio, "A new optimality property of the Capon estimator," *IEEE Signal Process. Lett.*, vol. 24, no. 11, pp. 1706–1708, Jul. 2017.
- [54] R. Chellappa and S. Theodoridis, *Academic Press Library in Signal Processing, Volume 7: Array, Radar and Communications Engineering*. New York, USA: Academic Press, 2017.
- [55] J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 30–39, Mar. 2002.
- [56] Y.-S. Yoon, M. G. Amin, and F. Ahmad, "MVDR beamforming for through-the-wall radar imaging," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 347–366, Jan. 2011.
- [57] S. D. Somasundaram, "Wideband robust Capon beamforming for passive sonar," *IEEE J. Ocean. Eng.*, vol. 38, no. 2, pp. 308–322, Apr. 2013.
- [58] K. Luo and A. Manikas, "Superresolution multitarget parameter estimation in MIMO Radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3683–3693, Jun. 2013.
- [59] T. Tirer and A. J. Weiss, "High resolution direct position determination of radio frequency sources," *IEEE Signal Process. Lett.*, vol. 23, no. 2, pp. 192–196, Feb. 2016.
- [60] M. D. Hossain and A. S. Mohan, "Eigenspace time-reversal robust Capon beamforming for target localization in continuous random media," *IEEE Antennas Wirel. Propag. Lett.*, vol. 16, pp. 1605–1608, Jan. 2017.
- [61] L. S. Yang, M. R. McKay, and R. Couillet, "High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1933–1947, Apr. 2018.
- [62] P. Chevalier, J.-P. Delmas, and M. Sadok, "Third-order Volterra MVDR beamforming for non-Gaussian and potentially non-circular interference cancellation," *IEEE Trans. Signal Process.*, vol. 66, no. 18, pp. 4766–4781, Sept. 2018.
- [63] L. Zhang, L. Huang, B. Li, M. Huang, J. M. Yin, and W. M. Bao, "Fast-moving jamming suppression for UAV navigation: A minimum dispersion distortionless response beamforming approach," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 7815–7827, Aug. 2019.
- [64] S. R. Tuladhar and J. R. Buck, "Unit circle rectification of the minimum variance distortionless response beamformer," *IEEE J. Ocean. Eng.*, vol. 45, no. 2, pp. 500–510, Apr. 2020.
- [65] N. L. Owsley, "Signal subspace based minimum-variance spatial array processing," in *Proc. IEEE 19th Asilomar Conf. Circuit, Systems and Computers*, Pacific Grove, USA, 1985, pp. 94–97.
- [66] J. P. Burg, "The relationship between maximum entropy spectra and maximum likelihood spectra," *Geophysics*, vol. 37, no. 2, pp. 375–376, Apr. 1972.
- [67] V. F. Pisarenko, "On the estimation of spectra by means of non-linear functions of the covariance matrix," *Geophys. J. Roy. Astron. Soc.*, vol. 28, no. 5, pp. 511–531, Jun. 1972.
- [68] G. Borgiotti and L. Kaplan, "Superresolution of uncorrelated interference sources by using adaptive array techniques," *IEEE Trans. Antennas Propag.*, vol. 27, no. 6, pp. 842–845, Nov. 1979.
- [69] M. Lagunas-Hernandez and A. Gasull-Llampallas, "An improved maximum likelihood method for power spectral density estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 1, pp. 170–173, Feb. 1984.
- [70] C. L. Byrne and A. K. Steele, "Stable nonlinear methods for sensor array processing," *IEEE J. Ocean. Eng.*, vol. 10, no. 3, pp. 255–259, Jul. 1985.
- [71] M. A. Lagunas, M. E. Santamaria, A. Gasull, and A. Moreno, "Maximum likelihood filters in spectral estimation problems," *Signal Process.*, vol. 10, no. 1, pp. 19–34, Jan. 1986.
- [72] J. Munier and G. Y. Delisle, "Spatial analysis in passive listening using adaptive techniques," *Proc. IEEE*, vol. 75, no. 11, pp. 1458–1471, Nov. 1987.
- [73] M. K. Ibrahim, "New high-resolution pseudospectrum estimation method," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 7, pp. 1071–1072, Jul. 1987.
- [74] M. D. Zoltowski and F. Haber, "A multiply constrained minimum variance approach to multiple source parameter estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 9, pp. 1358–1360, Sept. 1987.
- [75] P. S. Naidu and V. V. Krishna, "Improved maximum likelihood spectrum for direction of arrival (DOA) estimation," in *Proc. 1988 IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, New York, USA, 1988, pp. 2901–2904.
- [76] M. A. Lagunas and F. Vallverdu, "Rayleigh estimates for high resolution direction finding," in *Underwater Acoustic Data Processing*, Y. T. Chan, Ed. Dordrecht, the Netherlands: Springer, 1989, pp. 267–271.

- [77] J. Choi, I. Song, S. I. Park, and J. S. Yun, "Direction of arrival estimation with unknown number of signal sources," in *Proc. IEEE Int Conf Communication Systems*, Singapore, Singapore, 1992, pp. 30–34.
- [78] U. Nickel, "Radar target parameter estimation with array antennas," in *Radar Array Processing*, S. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin, Germany: Springer, 1993, pp. 47–98.
- [79] A. Hassaniien, S. Shahbazpanahi, and A. B. Gershman, "A generalized Capon estimator for localization of multiple spread sources," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 280–283, Jan. 2004.
- [80] B. D. Van Veen, W. Van Drongelen, M. Yuchtman, and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 867–880, Sept. 1997.
- [81] M.-X. Huang, J. J. Shih, R. R. Lee, D. L. Harrington, R. J. Thoma, M. P. Weisend, F. Hanlon, K. M. Paulson, T. Li, K. Martin, G. A. Miller, and J. M. Canive, "Commonalities and differences among vectorized beamformers in electromagnetic source imaging," *Brain Topogr.*, vol. 16, no. 3, pp. 139–158, Mar. 2004.
- [82] M. Costa and V. Koivunen, "Application of manifold separation to polarimetric Capon beamformer and source tracking," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 813–827, Feb. 2014.
- [83] C. D. Richmond, "Capon algorithm mean-squared error threshold SNR prediction and probability of resolution," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2748–2764, Aug. 2005.
- [84] A. K. Steele, C. L. Byrne, J. L. Riley, and M. Swift, "Performance comparison of high resolution bearing estimation algorithms using simulated and sea test data," *IEEE J. Ocean. Eng.*, vol. 18, no. 4, pp. 438–446, Oct. 1993.
- [85] R. Klemm, "High-resolution analysis of non-stationary data ensembles," in *Signal Processing: Theories and Applications*, M. Kunt and F. de Coulon, Eds. Amsterdam, the Netherlands: North-Holland, 1980, pp. 711–714.
- [86] P. Stoica, J. Li, and X. Tan, "On spatial power spectrum and signal estimation using the Pisarenko framework," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5109–5119, Oct. 2008.
- [87] C. N. Liu, J. M. Xin, Y. Nishio, N. N. Zheng, and A. Sano, "Modified Capon beamformer for high-resolution direction-of-arrival estimation," in *Proc. Int. Symp. Nonlinear Theory and Its Applications*, Luzern, Switzerland, 2014, pp. 116–119.
- [88] J. M. Xin, N. N. Zheng, and A. Sano, "Simple and efficient nonparametric method for estimating the number of signals without eigendecomposition," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1405–1420, Apr. 2007.
- [89] G. H. Golub and C. F. Van Loan, *Matrix Computations*. 2nd ed. Baltimore, USA: Johns Hopkins University Press, 1989.
- [90] P. Shaman, "The inverted complex Wishart distribution and its application to spectral estimation," *J. Multivar. Anal.*, vol. 10, no. 1, pp. 51–59, Mar. 1980.
- [91] B. Musicus, "Fast MLM power spectrum estimation from uniformly spaced correlations," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 5, pp. 1333–1335, Oct. 1985.
- [92] T. Ekman, A. Jakobsson, and P. Stoica, "On efficient implementation of the Capon algorithm," in *Proc. 8th Europ. Signal Process. Conf.*, Tampere, Finland, 2000.
- [93] L. Wei and S. L. Marple, "Fast algorithms for least-squares-based minimum variance spectral estimation," *Signal Process.*, vol. 88, no. 9, pp. 2181–2192, Sept. 2008.
- [94] Z.-S. Liu, H. Li, and J. Li, "Efficient implementation of Capon and APES for spectral estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 4, pp. 1314–1319, Oct. 1998.
- [95] G.-O. Glentis, "A fast algorithm for APES and Capon spectral estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4207–4220, Sept. 2008.
- [96] C. S. Withers and S. Nadarajah, "The n th power of a matrix and approximation for large n ," *N. Z. J. Math.*, vol. 38, pp. 171–178, 2008.
- [97] G.-O. Glentis, "Efficient algorithms for adaptive Capon and APES spectral estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 84–96, Jan. 2010.
- [98] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, USA: Prentice-Hall, 2000.
- [99] J. M. Xin and A. Sano, "Efficient subspace-based algorithm for adaptive bearing estimation and tracking," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4485–4505, Dec. 2005.
- [100] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, Aug. 1985.
- [101] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 8–15, Jan. 1989.
- [102] J. M. Xin and A. Sano, "MSE-based regularization approach to direction estimation of coherent narrowband signals using linear prediction," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2481–2497, 2001.
- [103] 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.
- [104] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Boston, USA, 1983, pp. 336–339.
- [105] Z. Yang, L. H. Xie, and C. S. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4959–4973, Oct. 2014.
- [106] C. Qian, L. Huang, N. D. Sidiropoulos, and H. C. So, "Enhanced PUMA for direction-of-arrival estimation and its performance analysis," *IEEE Trans. Signal Process.*, vol. 64, no. 16, pp. 4127–4137, Aug. 2016.
- [107] J. Zhu, L. Han, R. S. Blum, and Z. W. Xu, "Multi-snapshot Newtonized orthogonal matching pursuit for line spectrum estimation with multiple measurement vectors," *Signal Process.*, vol. 165, pp. 175–185, Dec. 2019.
- [108] D. Zachariah, P. Stoica, and M. Jansson, "Comments on 'Enhanced PUMA for direction-of-arrival estimation and its performance analysis'," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 6113–6114, Nov. 2017.
- [109] P. Stoica and T. Söderström, "Statistical analysis of a subspace method for bearing estimation without eigendecomposition," *IEEE Proc. F (Radar Signal Process.)*, vol. 139, no. 4, pp. 301–305, Aug. 1992.
- [110] G. M. Wang, J. M. Xin, N. N. Zheng, and A. Sano, "Computationally efficient subspace-based method for two-dimensional direction estimation with L-shaped array," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3197–3212, Jul. 2011.
- [111] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov. 14, 2008. [Online]. Available: <http://faculty.bicmr.pku.edu.cn/~wenzw/bigdata/matrix-cook-book.pdf>, 2012.
- [112] P. H. M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Autom. Control*, vol. 33, no. 9, pp. 867–870, Sept. 1988.



Weiliang Zuo (Member, IEEE) received the B.E. degree in electrical engineering and the Ph.D. degree in control science and engineering from Xi'an Jiaotong University in 2010 and 2018, respectively.

During 2016 to 2017, he was a visiting Ph.D. candidate at the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech). He is currently an Associate Professor at the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University. His current research interests include array and statistical signal processing, pattern recognition, and medical imaging processing.



Jingmin Xin (Senior Member, IEEE) received the B.E. degree in information and control engineering from Xi'an Jiaotong University, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from Keio University, Japan, in 1993 and 1996, respectively.

From 1988 to 1990, he was with the Tenth Institute of Ministry of Posts and Telecommunications (MPT) of China. He was with the Communications Research Laboratory, Japan, as an Invited Research Fellow of the Telecommunications Advancement Organization of Japan (TAO) from 1996 to 1997 and as a Postdoctoral Fellow of the Japan Science and Technology Corporation (JST) from 1997 to 1999. He was also a Guest (Senior) Researcher with YRP Mobile Telecommunications Key Technology Research Laboratories Company, Limited, Japan, from 1999 to 2001. From 2002 to 2007, he was with Fujitsu Laboratories Limited, Japan. Since 2007, he has been a Professor at Xi'an Jiaotong University. His research interests include adaptive filtering, statistical and array signal processing, system identification, and pattern recognition.



Changnong Liu received the B.S. degree in automation in 2012 and the M.S. degree in control science and engineering (pattern recognition and intelligence systems) in 2015 from Xi'an Jiaotong University. Currently she is an Application System Administrator at the Data Center of China Construction Bank.



Nanning Zheng (Fellow, IEEE) graduated from the Department of Electrical Engineering, Xi'an Jiaotong University in 1975, and received the M.S. degree in information and control engineering from Xi'an Jiaotong University in 1981 and the Ph.D. degree in electrical engineering from Keio University, Japan, in 1985.

He joined Xi'an Jiaotong University in 1975, and is currently a Professor and the Director of the Institute of Artificial Intelligence and Robotics, Xi'an

Jiaotong University. His research interests include computer vision, pattern recognition and image processing, and hardware implementation of intelligent systems.

Dr. Zheng became a Member of the Chinese Academy of Engineering in 1999, and he is the Chinese Representative on the Governing Board of the International Association for Pattern Recognition. He also serves as an Executive Deputy Editor of the Chinese Science Bulletin.



Akira Sano (Member, IEEE) received the B.E., M.S., and Ph.D. degrees in mathematical engineering and information physics from the University of Tokyo, Japan, in 1966, 1968, and 1971, respectively.

In 1971, he joined the Department of Electrical Engineering, Keio University, Japan, where he was a Professor with the Department of System Design Engineering till 2009. He is currently Professor Emeritus of Keio University. He has been a Member of Science Council of Japan since 2005. He was a Visiting Research Fellow at the University of Salford, U.K., from 1977 to 1978. He is the coauthor of the textbook *State Variable Methods in Automatic Control* (Wiley, 1988). His current research interests include adaptive modeling and design theory in control, signal processing and communication, and applications to control of sounds and vibrations, mechanical systems, and mobile communication systems.

Dr. Sano is a Fellow of the Society of Instrument and Control Engineers and is a Member of the Institute of Electrical Engineering of Japan and the Institute of Electronics, Information and Communications Engineers of Japan. He was General Co-Chair of 1999 IEEE Conference of Control Applications and an IPC Chair of 2004 IFAC Workshop on Adaptation and Learning in Control and Signal Processing. He was the Chair of IFAC Technical Committee on Modeling and Control of Environmental Systems from 1996 to 2001. He has also been Vice Chair of IFAC Technical Committee on Adaptive Control and Learning since 1999 and the Chair of IFAC Technical Committee on Adaptive and Learning Systems since 2002. He was also on the Editorial Board of Signal Processing. He was the recipient of the Kelvin Premium from the Institute of Electrical Engineering in 1986.