PUMA: An Improved Realization of MODE for DOA Estimation

CHENG QIAN
Harbin Institute of Technology, Harbin, China

LEI HUANG 🆔 , Senior Member, IEEE
Shenzhen University, Shenzhen, China

MINGYANG CAO, Student Member, IEEE
JUNHAO XIE, Senior Member, IEEE
Harbin Institute of Technology, Harbin, China

HING CHEUNG SO, Fellow, IEEE
City University of Hong Kong, Hong Kong

The method of direction estimation (MODE) offers appealing advantages such as asymptotic efficiency with mild computational complexity and excellent performance in handling coherent signals, which are not shared by conventional subspace-based methods. However, the MODE employs additional assumption and constraints on the symmetry of the root polynomial coefficients, which might cause severe performance degradation in the scenario of low signal-to-noise ratio/small sample size, since any estimation error will be enlarged twice due to the symmetry. Moreover, the standard realization for MODE does not have a closed-form solution for updating its estimates. In this paper, the optimization problem of MODE is proved to be equivalent to that of the *principal-eigenvector utiliztion for modal analysis (PUMA)* algorithm. We show that PUMA which has *closed-form* solution, that does not rely on any additional assumption and constraint on the coefficients, is a better surrogate than MODE for minimizing the same cost function. Extensive simulation results are carried out to support our standpoint.

I. INTRODUCTION

Direction-of-arrival (DOA) estimation is a fundamental array processing problem with numerous applications, e.g., radar [1], sonar [2], and wireless communications [3], [4]. It has been well studied during the past four decades, resulting in many efficient and accurate algorithms [5]–[22].

Among them, the maximum likelihood (ML) [5] approach is able to provide efficient DOA estimation but at the expense of huge complexity which is mainly due to the multidimensional search. Although optimization alternatives such as Newton's method can be applied to avoid the searching step, global convergence is not guaranteed since the ML objective function is nonconvex. Subspace-based DOA estimation algorithms can offer a good tradeoff between the resolution ability and computational complexity, and have been widely studied. In particular, multiple signal classification (MUSIC) [6] and estimation of signal parameters via rotation invariance techniques (ESPRIT) [7] as representatives of the subspace approach have attracted great interests due to their simplicity and high accuracy, and a plenty of their variants [8]–[18] have been developed subsequently. For example, computationally efficient modifications of the MUSIC algorithm have been devised in [14]–[18]. However, these algorithms suffer performance degradation when coherent/high correlated signals appear, thus requiring decoherency technique, e.g., spatial smoothing (SS) [19] or forward–backward SS (FBSS) [20], but at the expense of losing array aperture. In particular, a real-valued version of root-MUSIC, which is named as unitary root-MUSIC, has been proposed [18]. This algorithm exploits a centro-Hermitian property of the uniform linear array (ULA) to transform complex-valued data into the real space. This process is equivalent to performing a one-step forward-backward smoothing to the sample covariance matrix, hence it can handle at most two coherent signals. However, when there are more than two coherent signals, the unitary root-MUSIC algorithm cannot work properly.

Compared to MUSIC and ESPRIT, the method of direction estimation (MODE) [21]–[22] is of great interest, since it has appealing advantages over MUSIC and ESPRIT. These includes the following.

1) It performs like ML but does not require the computationally intensive searching procedure and is known to be approximately efficient in the large sample case.
2) It can handle coherent signals.
3) It is inherently not an iterative method, and hence, does not have the convergence problem.

Moreover, the MODE is solved efficiently through polynomial root finding in the ULA setting. Those properties make MODE a competitive candidate for DOA estimation. However, to minimize the MODE cost function efficiently, Stoica and Sharman [21] assume that the polynomial coefficients are symmetric, which is necessary but not sufficient for solving the root polynomial, and then employ additional constraints, i.e., fixing the real or imaginary part of the first

polynomial coefficient to be one, to avoid nonuniqueness of the coefficients. Note that it is shown in [26] and [27] that such an assumption is not exactly correct and might cause performance loss. In practice, due to the noise corruption, especially in low signal-to-noise ratio (SNR) or small sample size cases, the performance of MODE will be very sensitive to the symmetry assumption, since the coefficients mainly depend on half of its elements. When the estimates are not correct, after compensating the other half using conjugate symmetry, the error will be enlarged twice, resulting in bad performance. According to our experience, the performance degradation happens very frequently for odd source number, even for single source case. Furthermore, the MODE does not have a closed-form solution.

To compensate the performance loss of MODE at the low SNR region, Gershman and Stoica have proposed the MODEX (MODE with extra-roots) algorithm [23], [24]. MODEX runs the MODE twice with assumed number of sources being K and P ($P > K$), respectively, to generate $(P + K)$ DOA candidates, and then employs the deterministic/stochastic ML cost function to help selecting K of them as the final DOA estimates. Provided that K is odd, if we properly choose $P = K + 1$ which is even, MODEX will improve the threshold performance of MODE a lot, where the performance improvement mainly depends on MODE with even P. Recently, we have also employed a similar idea and proposed an enhanced principal-eigenvector utilization for modal analysis (EPUMA) technique to perform DOA estimation. Since we do not have any assumptions on the polynomial coefficient vector, the EPUMA does not suffer performance loss in the case of odd K. Unlike the MODEX, EPUMA performs reliably no matter what K is. More detailed comparison will be provided in the following Sections III and IV.

In this paper, we focus on presenting an efficient realization of MODE, i.e., PUMA, and establishing the analytical variance expression for MODE without introducing any additional assumptions or constraints on the polynomial coefficients. We show that PUMA and MODE, which are derived based on different theories, have the equivalent cost function. Note that Zachariah *et al.,* have also showed the equivalence of the cost functions of these two algorithm [25]. We would like to execute a deep comparison for these two algorithms as well as their respective variants, i.e., EPUMA and MODEX, to show that PUMA has advantages over MODE and it should be a better surrogate than MODE for minimizing the same cost function. Simulation results are provided to showcase that the performance of MODE is sensitive to the source number. Specifically, MODE works well for even source numbers but not for the odd case, and it even does not work when there is only one source. However, PUMA does not have such a problem, it cannot only provide reliable DOA estimates, outperform MODE exceedingly in the odd source number scenarios, but also is easier to be realized and is more computationally attractive than MODE.

NOTATION Throughout the paper, we use boldface lowercase letters for vectors and boldface uppercase letters for matrices. Superscripts $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^{-1}$, and $(\cdot)^\dagger$ represent transpose, complex conjugate, conjugate transpose, matrix inverse, and pseudoinverse, respectively. The $\hat{a}$ denotes an estimate of a, $\odot$ is the element-wise product, $\otimes$ is the Kronecker product, $\mathbb{E}[\cdot]$ is the expectation operator, and vec$(\cdot)$ is the vectorization operator. The tr$(\cdot)$ is the trace operator, Re takes the real part, $\mathbf{0}_{m \times n}$ is the $m \times n$ zero matrix, and $\mathbf{I}_m$ is the $m \times m$ identity matrix. The blkdiag$(\cdot)$ and diag$(\cdot)$ stand for block diagonal and diagonal matrices, respectively. Finally, $\delta_{i,j}$ is the delta function.

II. SIGNAL MODEL AND MODE

A. Signal Model

The problem of estimating the DOAs of K narrowband signals using a M-element ULA can be modeled as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \quad t = 1, \cdots, N \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{C}^M$ is the observation vector, $\mathbf{s}(t) \in \mathbb{C}^K$ is the unknown vector of wave amplitudes, $M \geq K$, $\mathbf{n}(t) \in \mathbb{C}^M$ is the additive noise vector, N is the number of samples, and $\mathbf{A} \in \mathbb{C}^{M \times K}$ is the array manifold matrix which has the form of

$$\mathbf{A} = [\mathbf{a}(\theta_1) \cdots \mathbf{a}(\theta_K)] \tag{2}$$

with its kth steering vector being

$$\mathbf{a}(\theta_k) = \begin{bmatrix} 1 & e^{j2\pi \sin(\theta_k)d/\nu} & \cdots & e^{j2\pi(M-1)\sin(\theta_k)d/\nu} \end{bmatrix}^T. \tag{3}$$

Here, d is the array interelement spacing and ν is the wavelength. The problem dealt with in this paper is to estimate $\boldsymbol{\theta} = [\theta_1 \ldots \theta_K]^T$, under the assumption of ULA and known K, from the received data matrix, i.e.,

$$\mathbf{X} = [\mathbf{x}(1) \cdots \mathbf{x}(N)]. \tag{4}$$

Let us write the eigenvalue decomposition (EVD) of the array covariance matrix as

$$\mathbf{R} = \mathbb{E}\left[\mathbf{X}\mathbf{X}^H\right] = \mathbf{U}_s\boldsymbol{\Lambda}_s\mathbf{U}_s^H + \mathbf{U}_n\boldsymbol{\Lambda}_n\mathbf{U}_n^H \tag{5}$$

where

$$\boldsymbol{\Lambda}_s = \text{diag}(\lambda_1 \ldots \lambda_K) \tag{6}$$
$$\boldsymbol{\Lambda}_n = \text{diag}(\lambda_{K+1} \ldots \lambda_M) \tag{7}$$

contain the K signal and $(M - K)$ noise eigenvalues of $\mathbf{R}$, respectively, whereas $\mathbf{U}_s$ and $\mathbf{U}_n$ are signal and noise subspaces. In the finite sample case, $\mathbf{R}$ is estimated as

$$\hat{\mathbf{R}} = \frac{\mathbf{X}\mathbf{X}^H}{N}. \tag{8}$$

The EVD of $\hat{\mathbf{R}}$ can be written as

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}_s\hat{\boldsymbol{\Lambda}}_s\hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n\hat{\boldsymbol{\Lambda}}_n\hat{\mathbf{U}}_n^H \tag{9}$$

where $\hat{\boldsymbol{\Lambda}}_s$, $\hat{\boldsymbol{\Lambda}}_n$, $\hat{\mathbf{U}}_s$, and $\hat{\mathbf{U}}_n$ are the estimates of $\boldsymbol{\Lambda}_s$, $\boldsymbol{\Lambda}_n$, $\mathbf{U}_s$, and $\mathbf{U}_n$, respectively.

B. MODE

The MODE estimates DOAs by minimizing the following cost function [22]:

$$f(\bar{\mathbf{b}}) = \mathrm{tr}\left(\mathbf{\Pi}\hat{\mathbf{U}}_s\hat{\mathbf{\Gamma}}\hat{\mathbf{U}}_s^H\right) \tag{10}$$

where

$$\bar{\mathbf{b}} = [b_0 \ b_1 \ \ldots \ b_K]^T \tag{11}$$

$$\mathbf{\Pi} = \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H \tag{12}$$

$$\hat{\mathbf{\Gamma}} = (\hat{\mathbf{\Lambda}}_s - \hat{\sigma}_n^2)^2\hat{\mathbf{\Lambda}}_s^{-1}$$

$$= \mathrm{diag}(\hat{\gamma}_1 \ \ldots \ \hat{\gamma}_K) \tag{13}$$

with

$$\mathbf{B} = \begin{bmatrix} b_K & b_{K-1} & \cdots & b_0 & & 0 \\ & \ddots & \ddots & & \ddots & \\ 0 & & b_K & b_{K-1} & \cdots & b_0 \end{bmatrix}^H \tag{14}$$

$$\hat{\gamma}_k = \left(\hat{\lambda}_k - \hat{\sigma}_n^2\right)^2/\hat{\lambda}_k \tag{15}$$

$$\hat{\sigma}_n^2 = \frac{1}{M-K}\mathrm{tr}(\hat{\mathbf{\Lambda}}_n). \tag{16}$$

The connection between $\mathbf{b}$ and $\{\theta_k\}_{k=1}^K$ is built upon the following equation:

$$b_0 z^K + b_1 z^{K-1} + \cdots + b_K = b_0\prod_{k=1}^K\left(1 - e^{j2\pi kd\sin(\theta)/v}\right)$$

$$= 0 \tag{17}$$

where $z = e^{j2\pi d\sin(\theta)/v}$ and the equation holds if and only if θ is the true DOA. Following [21] and [22], DOAs can be estimated via a two-step procedure which are as follows.

1) Initialize $\mathbf{B}^H\mathbf{B}$ using the identity matrix to obtain a guess of $\bar{\mathbf{b}}$, and then refine $\hat{\bar{\mathbf{b}}}$ iteratively by minimizing

$$f(\bar{\mathbf{b}}) = \mathrm{tr}\left(\mathbf{B}(\hat{\mathbf{B}}^H\hat{\mathbf{B}})^{-1}\mathbf{B}\hat{\mathbf{U}}_s\hat{\mathbf{\Gamma}}\hat{\mathbf{U}}_s^H\right).$$

2) Estimate DOAs via finding the roots of (17).

Note that in the first step, when $\hat{\mathbf{b}}$ is obtained, we substitute it into (14) to evaluate $\hat{\mathbf{B}}$.

The detailed steps of minimizing $f(\bar{\mathbf{b}})$ are given as follows [21].

Let

$$[\tilde{\mathbf{s}}_1 \ \cdots \ \tilde{\mathbf{s}}_K] = \begin{bmatrix} \tilde{s}_{1,1} & \cdots & \tilde{s}_{1,K} \\ \vdots & & \vdots \\ \tilde{s}_{M,1} & \cdots & \tilde{s}_{M,K} \end{bmatrix} = \hat{\mathbf{U}}_s\hat{\mathbf{\Gamma}}^{1/2}. \tag{18}$$

Then

$$\mathbf{B}^H\tilde{\mathbf{s}}_k = \begin{bmatrix} \tilde{s}_{K+1,k} & \cdots & \tilde{s}_{1,k} \\ \vdots & & \vdots \\ \tilde{s}_{M,k} & \cdots & \tilde{s}_{M-K,k} \end{bmatrix}\bar{\mathbf{b}} = \tilde{\mathbf{S}}_k\bar{\mathbf{b}}. \tag{19}$$

Define $\mathbf{V} = \mathbf{B}^H\mathbf{B}$, where its Cholesky decomposition is $\mathbf{V}^{1/2}$. This leads to

$$\mathbf{H} = \begin{bmatrix} \mathbf{V}^{1/2}\tilde{\mathbf{S}}_1 \\ \vdots \\ \mathbf{V}^{1/2}\tilde{\mathbf{S}}_K \end{bmatrix} \tag{20}$$

Thus, the MODE cost function becomes

$$f(\bar{\mathbf{b}}) = \|\mathbf{H}\bar{\mathbf{b}}\|_2^2. \tag{21}$$

In order to find the estimate of $\bar{\mathbf{b}}$ efficiently, Stoica and Sharman put an additional constraint on it, i.e.,

$$b_n = b_{K-n}^* \quad \forall n = 0, \ldots, K \tag{22}$$

which means that $\bar{\mathbf{b}}$ is conjugate symmetric. To proceed, it is necessary to introduce the following vector:

$$\boldsymbol{\beta} = [b_0 \ \cdots \ b_d]^T \tag{23}$$

where $d = \lceil(K-1)/2\rceil$ with $\lceil e\rceil$ being the largest integer less than or equal to e. Now the conjugate symmetric constraint can be expressed as

$$\bar{\mathbf{b}} = \begin{cases} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{K}\boldsymbol{\beta}^* \end{bmatrix} & \text{for odd } K \\ \begin{bmatrix} \boldsymbol{\beta} \\ \mu \\ \mathbf{K}\boldsymbol{\beta}^* \end{bmatrix} & \text{for oven } K \end{cases} \tag{24}$$

where $\mathbf{K}$ is an exchanging matrix and μ is a real-valued scalar number.

We further write $\mathbf{H}$ as

$$\mathbf{H} = \begin{cases} [\mathbf{H}_1 \mid \mathbf{H}_2] & \text{for odd } K \\ [\mathbf{H}_1 \mid \mathbf{h} \mid \mathbf{H}_2] & \text{for even } K \end{cases}. \tag{25}$$

For notational simplicity, let subscript "$(\cdot)_R$" and "$(\cdot)_I$" be the real and imaginary parts of a complex number, respectively. When K is odd, (21) becomes

$$f(\bar{\mathbf{b}}) = \left\|\begin{bmatrix} \mathbf{H}_{1R} + (\mathbf{H}_2\mathbf{K})_R & (\mathbf{H}_2\mathbf{K})_I - \mathbf{H}_{1I} \\ \mathbf{H}_{1I} + (\mathbf{H}_2\mathbf{K})_I & \mathbf{H}_{1R} - (\mathbf{H}_2\mathbf{K})_R \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta}_R \\ \boldsymbol{\beta}_I \end{bmatrix}\right\|_2^2$$

$$\triangleq \|\mathbf{G}\boldsymbol{\rho}\|_2^2. \tag{26}$$

When K is even

$$f(\bar{\mathbf{b}})$$

$$= \left\|\begin{bmatrix} \mathbf{H}_{1R} + (\mathbf{H}_2\mathbf{K})_R & (\mathbf{H}_2\mathbf{K})_I - \mathbf{H}_{1I} & \mathbf{h}_R \\ \mathbf{H}_{1I} + (\mathbf{H}_2\mathbf{K})_I & \mathbf{H}_{1R} - (\mathbf{H}_2\mathbf{K})_R & \mathbf{h}_I \end{bmatrix}\begin{bmatrix} \boldsymbol{\beta}_R \\ \boldsymbol{\beta}_I \\ \mu \end{bmatrix}\right\|_2^2$$

$$\triangleq \|\mathbf{G}\boldsymbol{\rho}\|_2^2. \tag{27}$$

In [21], a QR algorithm is proposed to minimize (26) and (27) and the interested reader is referred to [21].

III. NOVEL REALIZATION OF MODE: A CLOSED-FORM SOLUTION

In this section, we show that MODE corresponds to the best linear unbiased estimator (BLUE) and can be solved efficiently using weighted least squares (WLS), which is actually the recently proposed PUMA algorithm.

A. WLS Formulation of MODE

Without loss of generality, let $b_0 = 1$. We should note that the value of b_0 does not affect the cost function, since it will be canceled out in $\mathbf{\Pi}$. Therefore, define

$$\mathbf{b} = [b_1 \ \ldots \ b_K]^T. \tag{28}$$

Then, the MODE cost function can be written as a function of $\mathbf{b}$. It has been demonstrated in [21] that there will be minor performance loss when the symmetric constraint (22) is applied [28], [29]. To avoid the performance loss, in the following, we try to establish a WLS solver for minimizing the MODE cost function without adding any additional assumption on $\mathbf{b}$.

PROPOSITION 1   The MODE estimate of $\mathbf{b}$ is the solution of the following WLS problem:

$$\hat{\mathbf{b}} = \arg\min_{\mathbf{b}} (\hat{\mathbf{F}}\mathbf{b} - \hat{\mathbf{g}})^H \hat{\mathbf{W}} (\hat{\mathbf{F}}\mathbf{b} - \hat{\mathbf{g}}) \tag{29}$$

where $\hat{\mathbf{F}}$ and $\hat{\mathbf{g}}$ are defined in (43) and (44), respectively, and the weighting matrix is

$$\hat{\mathbf{W}} = \hat{\mathbf{\Gamma}} \otimes (\mathbf{B}^H \mathbf{B})^{-1}. \tag{30}$$

PROOF   See Appendix A.   ∎

The above proposition provides us another view of the MODE. Since the cost function in (29) has exactly the same value as (10) (see Appendix A), if $\hat{\mathbf{b}}$ minimizes (29), it also minimizes (10). As a matter of fact, the weighting matrix is very important for WLS problem because it affects the precision of its estimates. However, even the WLS realization of MODE is obtained, we still do not know whether $\hat{\mathbf{W}}$ in (30) is the optimal weighting matrix or not. Therefore, finding the optimal weighting matrix is the most important issue for the WLS technique. Actually, according to Gauss–Markov theorem [30], [31], we have the following proposition.

PROPOSITION 2   The inverse of the weighting matrix $\hat{\mathbf{W}}$ in (30) is a consistent estimate of the covariance of $(\hat{\mathbf{F}}\mathbf{b} - \hat{\mathbf{g}})$.

PROOF   See Appendix B.   ∎

The above proposition shows that $\hat{\mathbf{W}}$ is the inverse of the covariance matrix of $(\hat{\mathbf{F}}\mathbf{b} - \hat{\mathbf{g}})$, and each block in $\hat{\mathbf{W}}$ is equal to the reciprocal of the covariance of residual $(\hat{\mathbf{F}}_i \mathbf{b} - \hat{\mathbf{g}}_i)$, $\forall i = 1, \ldots, K$, which means that $\hat{\mathbf{b}}$ is the BLUE of $\hat{\mathbf{F}}\mathbf{b} \approx \hat{\mathbf{g}}$.

B. WLS Solution for MODE

We now start to provide a new way to minimize the MODE cost function. At first, we introduce the well-known result for WLS, i.e., the solution to (29) is

$$\hat{\mathbf{b}} = \left(\hat{\mathbf{F}}^H \hat{\mathbf{W}} \hat{\mathbf{F}} \right)^{-1} \hat{\mathbf{F}}^H \hat{\mathbf{W}} \hat{\mathbf{g}}. \tag{31}$$

Since $\hat{\mathbf{W}}$ depends on $\mathbf{b}$ inherently, we cannot directly use (31) to compute $\mathbf{b}$. Recall that the WLS is a solution of $\hat{\mathbf{F}}\mathbf{b} \approx \hat{\mathbf{g}}$. We initialize $\hat{\mathbf{b}}$ using its least squares (LS) solution, i.e., $\hat{\mathbf{b}}_{LS} = \hat{\mathbf{F}}^\dagger \mathbf{g}$, and then use it to initialize $\hat{\mathbf{W}}$ [40].

Algorithm 1: WLS Realization of MODE, i.e., PUMA.

1: Calculate $\hat{\mathbf{R}}$ and its EVD via (8) and (9), respectively;
2: Use $\hat{\mathbf{b}}_{LS}$ to initialize $\mathbf{B}$ in (14);
3: **for** $i = 1, 2, \ldots$ **do**
4:    Compute $\hat{\mathbf{W}}$ via (30);
5:    Estimate $\hat{\mathbf{b}}$ using (31);
6:    Compute $\mathbf{B}$ using $\hat{\mathbf{b}}$;
7: **end for**
8: Calculate K roots $\hat{z}_k$ from (17) and estimate $\hat{\theta}_k$ via (32).

After obtaining $\hat{\mathbf{b}}$, similar to MODE, we first compute $\{\hat{z}_k\}_{k=1}^K$ by solving (17), then estimate K DOAs using

$$\hat{\theta}_k = \sin^{-1}\left(\frac{\nu \angle \hat{z}_k}{2\pi d} \right), \ k = 1, \ldots, K. \tag{32}$$

The detailed steps for PUMA are summarized in Algorithm 1.

REMARK 1   Actually, the above WLS realization is an existing algorithm called PUMA, which was first derived for harmonic retrieval [31], and further extended for DOA estimation with a so-called *EPUMA* technique [40]. In the following, we refer to PUMA as the above WLS realization. PUMA and MODE have similar properties but they are not the same. Hence, they have different performance, as we will see later. Their only identicalness is that their respective cost functions are mathematically equivalent. In other words, both of them aim at minimizing the same cost function which is interpreted in two different ways, namely, MODE minimizes (10), while PUMA minimizes (29). Specifically, MODE tries to find $\bar{\mathbf{b}}$ based on adding additional assumption on the conjugate symmetric or unit-norm property of $\bar{\mathbf{b}}$. However, this assumption may be challenged in the case of different number of sources. That is, MODE shows very different performance in even and odd source number cases. More precisely, according to our experience, MODE works very well for even K but not for the odd one. However, PUMA does not have such unstable performance since it does not require any assumption on $\mathbf{b}$.

C. Mean Square Error Analysis

For large enough data size, perturbation analysis can be used for obtaining variance of an estimator [22], [34]–[38]. Variance analysis of MODE has been previously studied in [22], [34], [35]. In this section, using an asymptotic distribution of the sample covariance eigenvectors for complex observations in [37], [38], a new covariance expression that is based on the first- and second-order derivatives of (29) is derived for MODE.

PROPOSITION 3   The asymptotic variance associated with $\{\hat{\theta}_k - \theta_k\}_{k=1}^K$ of MODE is asymptotically jointly Gaussian

TABLE I
Complexity Comparison Between MODE and PUMA

	Main computational steps	Complexity
MODE	Cholesky decomposition of $\mathbf{V}$	$\mathcal{O}((M-K)^3)$
	QR factorization of $\mathbf{G}$	$\mathcal{O}(8K^3(M-K)^3)$
PUMA	Inverse of $(\mathbf{B}^H\mathbf{B})$	$\mathcal{O}((M-K)^3)$
	Calculation of $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{F}}$	$\mathcal{O}(2K^2(M-K)^2)$
	Calculation of $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{g}}$	$\mathcal{O}(K(M-K)^2)$

distributed with mean zero and covariance matrix:

$$\mathbb{E}\left[\Delta\boldsymbol{\theta}\Delta\boldsymbol{\theta}^T\right] \approx \frac{1}{2}\mathrm{Re}\left\{\mathbf{D}\mathbf{Z}^T(\mathbf{F}^H\mathbf{W}_{\mathrm{opt}}\mathbf{F})^{-1}\mathbf{Z}^*\mathbf{D}^H\right\}. \quad (33)$$

PROOF   See Appendix C.   ∎

D.   Complexity Comparison

Table I compares the complexities of MODE and PUMA.

MODE and PUMA require calculating the sample covariance matrix and its EVD, where the computations for these two parts are $\mathcal{O}(M^2N)$ and $\mathcal{O}(M^3)$ flops, respectively. It is seen from Table I that for MODE, the amount of operations involved in each iteration step mainly lies in the Cholesky and QR factorizations of $\mathbf{V}$ and $\mathbf{G}$, which require $\mathcal{O}((M-K)^3)$ and $\mathcal{O}(8K^3(M-K)^3)$ flops, respectively. For PUMA, the complexity is mainly caused by the construction of $\hat{\mathbf{W}}$ and $\hat{\mathbf{b}}$. The computation of $\hat{\mathbf{W}}$ relies on the inverse of $\mathbf{B}^H\mathbf{B}$, which needs $\mathcal{O}((M-K)^3)$. In practice, the inverse of $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{F}} \in \mathbb{C}^{K\times K}$ is almost never explicitly calculated. Instead, the Cholesky decomposition of $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{F}}$ is employed, and for $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{g}}$, forward and backward substitutions are performed to get the corresponding estimate of $\hat{\mathbf{b}}$. Since $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{g}}$ and $\hat{\mathbf{W}}$ are block diagonal, forming $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{F}}$ and $\hat{\mathbf{F}}^H\hat{\mathbf{W}}\hat{\mathbf{g}}$ take $\mathcal{O}(2K^2(M-K)^2)$ and $\mathcal{O}(K(M-K)^2)$ flops, respectively, and finally the back substitution step requires $\mathcal{O}(K^2(M-K))$ flops. Therefore, except for the calculations of $\hat{\mathbf{R}}$ and its EVD, the complexity of the remaining steps for MODE is about $\mathcal{O}(8K^3(M-K)^3)$, while that of the PUMA is about $\mathcal{O}((M-K)^3 + (2K^2+K)(M-K)^2)$ which is smaller than MODE since $(M-K)$ is usually greater than K, which means that PUMA is slightly simpler than MODE.

Generally speaking, no more than three iterations are enough for PUMA to achieve comparable performance. Therefore, we suggest to use a fixed number of iterations (e.g., three iterations) to terminate the algorithm. If M and N are larger than K, the overall complexity is similar to the complexity of one iteration, which is $\mathcal{O}(M^2N + M^3)$.

IV.   NUMERICAL RESULTS

In this section, we examine the stochastic performance of PUMA by comparing it with MODE [22], MODEX [23], [24], EPUMA [40], root-MUSIC, and unitary ESPRIT algorithms. We also include the Cramér–Rao bound (CRB) [5] as a performance benchmark, where the CRB is
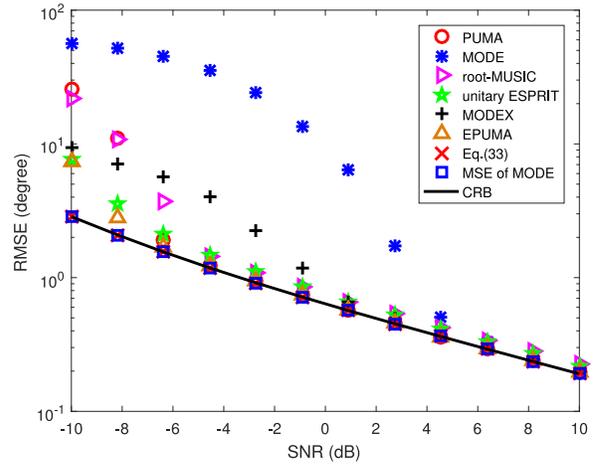
Fig. 1.   RMSE versus SNR when $K = 3$.

computed as

$$\mathrm{CRB} = \frac{\sigma_n^2}{2N}\mathrm{tr}\left(\mathrm{Re}\left((\mathbf{D}^H(\mathbf{I}_M - \mathbf{A}\mathbf{A}^\dagger)\mathbf{D})\odot\mathbf{R}^T\right)^{-1}\right) \quad (34)$$

with $\mathbf{D} = [\partial\mathbf{a}(\theta_1)/\partial\theta_1 \quad \cdots \quad \partial\mathbf{a}(\theta_K)/\partial\theta_K]$ and $\mathbf{R} = \mathbb{E}[\mathbf{x}(t)\mathbf{x}^H(t)] = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I}_M$. Here, $\mathbf{R}_s = \mathbb{E}[\mathbf{s}(t)\mathbf{s}^H(t)]$ is the signal covariance matrix. The number of iterations for our scheme and MODE is set to three. For root-MUSIC and unitary ESPRIT, when coherent signals occur, the FBSS [20] technique is employed to remove the coherency, where the number of forward–backward subarrays is equal to the number of coherent signals. In the following examples, we consider a ULA composed of $M = 10$ sensors receiving $K = 3$ narrowband signals. Furthermore, 2000 Monte-Carlo trials are utilized to compute the root mean square error (RMSE), i.e.,

$$\mathrm{RMSE} = \left(\frac{1}{2000}\sum_{i=1}^{2000}\|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}\|_2^2\right)^{1/2} \quad (35)$$

where $\hat{\boldsymbol{\theta}}_i$ contains K DOA estimates obtained from the ith test.

In the first example, we study the RMSE performance versus SNR. Three signals with DOAs being $[-5°, 2°, 15°]$ are considered, where the first two signals are coherent and uncorrelated with the third one. The number of samples is $N = 80$. The theoretical RMSE which is computed from (33) is included to predict the performance of MODE. To validate the correctness of (33), another RMSE curve which is obtained from [22] is also included for comparison. It is seen from Fig. 1 that PUMA outperforms the original MODE throughout the SNR regime, and its RMSE attains the CRB when SNR ≤ -5 dB. EPUMA achieves the best performance and its performance is better than MODEX. As SNR increases, the RMSE curves of PUMA and original MODE merge together, and their performance attains the theoretical RMSE curve as well as the CRB after SNR ≥ 5 dB. The MODEX has better threshold performance than MODE but not as good as ours. There is a gap between the RMSE curves of root-MUSIC, unitary ESPRIT, and CRB since there is a loss of array aperture after applying the
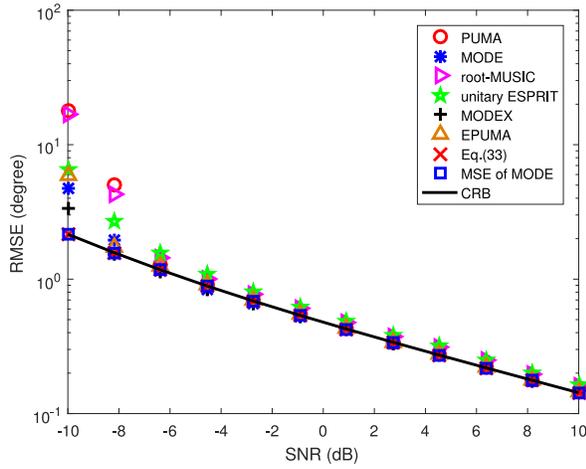
Fig. 2.   RMSE versus SNR when $K = 2$.
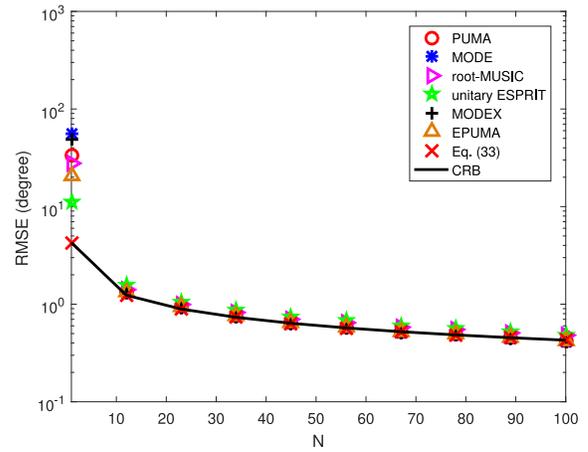
Fig. 4.   RMSE versus N when $K = 2$.
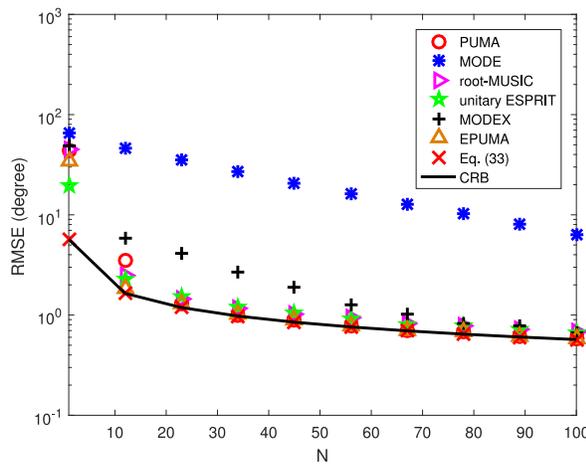
Fig. 3.   RMSE versus N when $K = 3$.

FBSS. Furthermore, the theoretical MSEs computed from (33) and [22] are exactly the same, which validates that our theoretical analysis is correct. Fig. 2 shows the RMSE results when $K = 2$ and DOAs are $\theta_1 = -5°$ and $\theta_2 = 2°$. We see that MODE has a huge performance improvement than the case of $K = 3$. Although for $K = 3$ and $K = 2$, parameter settings are changed and these two examples seem not comparable, we want to demonstrate our analysis in Section III-C that MODE is sensitive to the choice of K. Actually, the sensitivity of MODE shows not only in this example, but also in other parameter settings. It does not perform reliably especially for $K = 3$. One potential reason might be the symmetric assumption used in the MODE solver. The reason behind the performance degradation of MODE requires further theoretical study, which will be our future work.

Fig. 3 shows the RMSE versus N. The SNR is 0 dB and the other parameters are kept the same as Fig. 1. Again, MODE and MODEX do not perform well even when N attains 1000, while PUMA and EPUMA are the best estimators for $N > 20$ and their performance finally attains the CRB and theoretical RMSE curve. At SNR $= 0$ dB, MODE is very sensitive to the symmetry assumption, where the co-

efficients mainly depend on half of its elements, when the estimates are not correct, after compensating the other half using conjugate symmetry, the error will be enlarged twice, resulting in bad performance. We also include an example similar to Fig. 2 to showcase the performance with $K = 2$. It is seen in Fig. 4 that PUMA has very similar performance as the case of $K = 3$. However, MODE and MODEX performs very differently. Specifically, their performance under $K = 2$ is much better than that of $K = 3$. Furthermore, as we can see in Figs. 1 and 3 that MODEX performs much better than MODE. This is mainly due to the fact that in these two examples, MODEX applies the MODE twice by assuming the number of sources being $K = 3$ and $K = 4$, and then employs the ML cost function to judiciously select two final DOAs from the five candidates. Its performance improvement mainly depends on those generated from the case of $K = 4$, where MODE is not sensitive to the symmetry assumption and enables to provide reliable DOA estimates. Such an observation further verifies our analysis in *Remark 1*.

In the third example, we study the resolution ability of PUMA. Three signals are used, where the first two are coherent and uncorrelated with the third one, and the first and third DOAs are fixed at $\theta_1 = 0°$ and $\theta_3 = 20°$. The second DOA θ_2 is varied from $2°$ to $7°$. The SNR is 10 dB and the number of samples is $N = 100$. It is observed from Fig. 5 that for closely spaced DOAs, say, $\Delta\theta = 2°$, unitary ESPRIT performs the best. When $\Delta\theta \geq 2.5°$, PUMA and EPUMA become the best.

In the fourth example, we compare the RMSE performance versus SNR where there is only one signal with its DOA being $0°$. The number of samples is 20. It is seen from Fig. 6 that PUMA works much better than MODE and MODEX, and MODE does not work throughout the SNR region. MODEX performs much better than MODE since it employs MODE twice with assumed number of signals being 1 and 2 and its performance improvement mainly depends on the latter case where the source number is even. This confirms our previous analysis in *Remark 1* again. It should be noted again that PUMA and EPUMA do not have such a problem, i.e., sensitive to the source number.
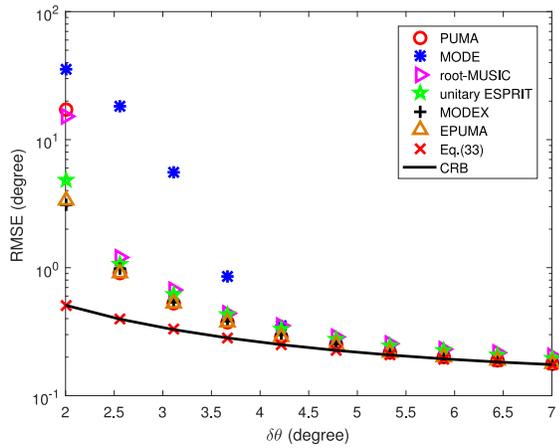
Fig. 5. RMSE versus $\Delta\theta$.
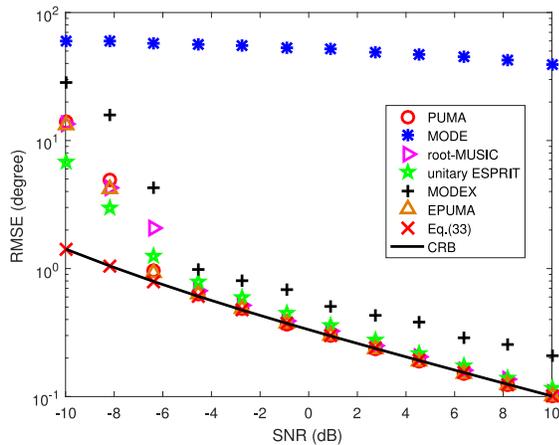
Fig. 6. RMSE versus SNR when $K = 1$.

Finally, we compare the complexity of our algorithm with those of MODE, root-MUSIC, and unitary ESPRIT.[1] We vary M from 10 to 60, and set $K = M/2$ and $N = 5M$. It is seen from Fig. 7 that the unitary ESPRIT and root-MUSIC have almost the same complexity, and both of them are faster than MODE. The standard MODE is the most computationally intensive method among the four competitors. When M becomes larger, our solution is at least ten times faster than the standard MODE. Note that the overall complexity between PUMA and MODE is in the same order, i.e., $\mathcal{O}(MN^2)$. In each iteration, both algorithms should take the inverse of an $K \times K$ matrix. However, the original MODE needs to perform a QR factorization of a $2(M - K)K \times K$ matrix while ours only refers to matrix multiplications. When K approaches M, the complexity of QR factorization becomes burdened. This

---

[1]The MODEX and EPUMA are obviously the most computationally intensive algorithms since both of them require to use the ML cost function at least $(K + P)!/(K!P!)$ times where $P > K$, to help select the final K DOA estimates. When $K \approx M$, their complexity becomes huge, especially for large M. As an example of $M = 20$, if we choose $K = M/2 = 10$ and $P = K + 2 = 12$, we should compute the ML cost function about 6.4×10^5 times, which is very computationally demanding. Due to this reason, we do not include them for complexity comparison.
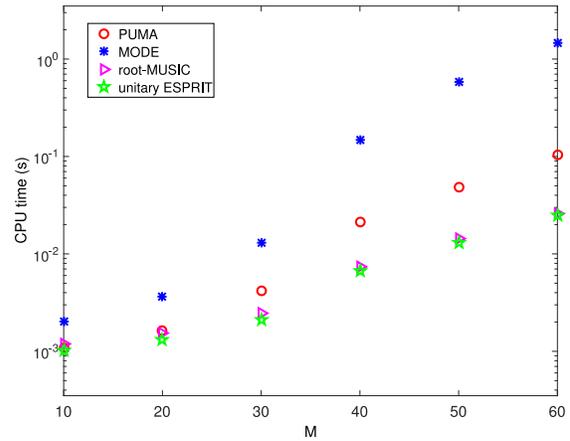
Fig. 7. CPU time versus M.

is the main reason why our method is a little bit faster than MODE.

V. CONCLUSION

In this paper, we have shown that the MODE cost function is equivalent to the recently proposed PUMA method. Unlike the original MODE method, PUMA does not require any assumption and constraints on the polynomial coefficients, and it has a *closed-form* expression for the update of polynomial coefficients. Hence, PUMA is much easier to be implemented than the original one. Numerical results show the advantages of PUMA and its variant.

Some interesting results can be observed from simulations, where MODE performs slightly better than PUMA when the source number is even. However, the former suffers severe performance degradation when the source number is odd. According to our conjecture, the unstable performance of MODE appears not only in our parameter settings but also in some other settings–even (or odd) number of sources helps (or destroys) the MODE. However, PUMA does not have such a problem. Since PUMA and MODE minimize the same unconstrained optimization problem, we recommend employing PUMA instead of MODE to achieve stability. Moreover, MODE still shows advantages over PUMA in certain cases, we will aim at exploiting their advantages and combining them together to further improve the accuracy especially in the low SNR and small sample scenarios.

APPENDIX A
PROOF OF PROPOSITION 1

We first rewrite (10) as

$$f(\mathbf{b}) = \text{tr}\left(\left(\mathbf{B}^H \hat{\mathbf{U}}_s\right)^H \left(\mathbf{B}^H \mathbf{B}\right)^{-1} \mathbf{B}^H \hat{\mathbf{U}}_s \hat{\mathbf{\Gamma}}\right)$$

$$= \sum_{k=1}^{K} \hat{\gamma}_k \left(\mathbf{B}^H \hat{\mathbf{u}}_k\right)^H \left(\mathbf{B}^H \mathbf{B}\right)^{-1} \mathbf{B}^H \hat{\mathbf{u}}_k. \quad (36)$$

Then, it is easily seen that

$$\mathbf{B}^H \hat{\mathbf{u}}_k = \hat{\mathbf{F}}_k \mathbf{b} - \hat{\mathbf{g}}_k, \ k = 1, \ldots, K \quad (37)$$

where

$$\hat{\mathbf{F}}_k = \begin{bmatrix} [\hat{\mathbf{u}}_k]_K & [\hat{\mathbf{u}}_k]_{K-1} & \cdots & [\hat{\mathbf{u}}_1]_1 \\ [\hat{\mathbf{u}}_k]_{K+1} & [\hat{\mathbf{u}}_k]_K & \cdots & [\hat{\mathbf{u}}_k]_2 \\ \vdots & \vdots & & \vdots \\ [\hat{\mathbf{u}}_k]_{M-1} & [\hat{\mathbf{u}}_k]_{M-2} & \cdots & [\hat{\mathbf{u}}_k]_{M-K} \end{bmatrix} \quad (38)$$

$$\hat{\mathbf{g}}_k = -[[\hat{\mathbf{u}}_k]_{K+1} \cdots [\hat{\mathbf{u}}_k]_M]^T. \quad (39)$$

Substituting (37) into (36), we obtain

$$f(\mathbf{b}) = \sum_{k=1}^K \left\| \hat{\mathbf{W}}_k^{1/2} (\hat{\mathbf{F}}_k \mathbf{b} - \hat{\mathbf{g}}_k) \right\|_2^2 \quad (40)$$

where the weighting matrix is

$$\hat{\mathbf{W}}_k = \hat{\gamma}_k \left(\mathbf{B}^H \mathbf{B} \right)^{-1}. \quad (41)$$

Actually, (40) can be written into a more compact form. To this end, we stack (37) for $k = 1, \ldots, K$, into a vector

$$\left[\left(\mathbf{B}^H \hat{\mathbf{u}}_1 \right)^T \cdots \left(\mathbf{B}^H \hat{\mathbf{u}}_K \right)^T \right]^T = \text{vec} \left(\mathbf{B}^H \hat{\mathbf{U}}_s \right)$$
$$= \hat{\mathbf{F}} \mathbf{b} - \hat{\mathbf{g}} \quad (42)$$

where

$$\hat{\mathbf{F}} = \left[\hat{\mathbf{F}}_1^T \ldots \hat{\mathbf{F}}_K^T \right]^T \quad (43)$$

$$\hat{\mathbf{g}} = \left[\hat{\mathbf{g}}_1^T \ldots \hat{\mathbf{g}}_K^T \right]^T. \quad (44)$$

Moreover, define a block diagonal matrix as

$$\hat{\mathbf{W}} = \text{blkdiag} \left(\hat{\mathbf{W}}_1 \ldots \hat{\mathbf{W}}_K \right)$$
$$= \hat{\mathbf{\Gamma}} \otimes \left(\mathbf{B}^H \mathbf{B} \right)^{-1}. \quad (45)$$

Thus, (36) becomes

$$f(\mathbf{b}) = \left(\hat{\mathbf{F}} \mathbf{b} - \hat{\mathbf{g}} \right)^H \hat{\mathbf{W}} \left(\hat{\mathbf{F}} \mathbf{b} - \hat{\mathbf{g}} \right) \quad (46)$$

which is the WLS objective function of the overdetermined system $\hat{\mathbf{F}} \mathbf{b} = \hat{\mathbf{g}}$ with the weighting matrix being $\hat{\mathbf{W}}$.

APPENDIX B
PROOF OF COROLLARY 2

According to the Gauss–Markov theorem [30], the optimal weighting matrix for WLS is

$$\mathbf{W}_{\text{opt}} = \left(\mathbb{E} \left[\left(\hat{\mathbf{F}} \mathbf{b} - \hat{\mathbf{g}} \right) \left(\hat{\mathbf{F}} \mathbf{b} - \hat{\mathbf{g}} \right)^H \right] \right)^{-1}. \quad (47)$$

It follows from (42) that $\hat{\mathbf{F}} \mathbf{b} - \hat{\mathbf{g}} = \text{vec}(\mathbf{B}^H \hat{\mathbf{U}}_s)$, and thus (47) becomes

$$\mathbf{W}_{\text{opt}} = \left(\mathbb{E} \left[\text{vec} \left(\mathbf{B}^H \hat{\mathbf{U}}_s \right) \text{vec} \left(\mathbf{B}^H \hat{\mathbf{U}}_s \right)^H \right] \right)^{-1}. \quad (48)$$

We simplify $\mathbf{B}^H \hat{\mathbf{U}}_s$ as

$$\mathbf{B}^H \hat{\mathbf{U}}_s = \mathbf{B}^H (\mathbf{U}_s + \Delta \mathbf{U}_s) = \mathbf{B}^H \Delta \mathbf{U}_s \quad (49)$$

where $\hat{\mathbf{U}}_s = \mathbf{U}_s + \Delta \mathbf{U}_s$ and the last equation is obtained by considering $\mathbf{B}^H \mathbf{U}_s = \mathbf{0}_{(M-K) \times K}$, which is always true since $\mathbf{B}$ lies in the null space of $\mathbf{A}$ [21].

In the sequel, we will use the following important property of the vectorization operator:

$$\text{vec}(\mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3) = (\mathbf{A}_3^T \otimes \mathbf{A}_1) \text{vec}(\mathbf{A}_2). \quad (50)$$

Using (50) and substituting (49) into (48) yields

$$\mathbf{W}_{\text{opt}} = \left((\mathbf{I}_K \otimes \mathbf{B}) \, \mathbb{E} \left[\Delta \mathbf{u}_s \Delta \mathbf{u}_s^H \right] (\mathbf{I}_K \otimes \mathbf{B})^H \right)^{-1} \quad (51)$$

where $\Delta \mathbf{u}_s = \text{vec}(\Delta \mathbf{U}_s) = [\Delta \mathbf{u}_1^T \ldots \Delta \mathbf{u}_K^T]^T$ with $\Delta \mathbf{u}_K = \hat{\mathbf{u}}_k - \mathbf{u}_k, \forall k$. It is shown in [37] and [38] that the errors between the signal eigenvectors have the following property:

$$\mathbb{E}\left[\Delta \mathbf{u}_i \Delta \mathbf{u}_j^H \right] \approx \frac{\lambda_i}{N} \sum_{\substack{k=1 \\ k \neq i}}^M \frac{\lambda_k}{(\lambda_i - \lambda_k)^2} \mathbf{u}_k \mathbf{u}_k^H \delta_{ij}. \quad (52)$$

It follows from (52) that

$$\mathbb{E}[\Delta \mathbf{u}_s \Delta \mathbf{u}_s^H] \approx \left(\mathbb{E}[\Delta \mathbf{u}_1 \Delta \mathbf{u}_1^H] \cdots \mathbb{E}[\Delta \mathbf{u}_K \Delta \mathbf{u}_K^H] \right). \quad (53)$$

Inserting (53) into (51), we obtain a simplified expression for (51) as

$$\mathbf{W}_{\text{opt}} = \bar{\mathbf{\Gamma}} \otimes (\mathbf{B}^H \mathbf{B})^{-1} \quad (54)$$

where

$$\bar{\mathbf{\Gamma}} = \text{diag} \left(\frac{N(\lambda_1 - \sigma_n^2)^2}{\lambda_1 \sigma_n^2} \cdots \frac{N(\lambda_K - \sigma_n^2)^2}{\lambda_K \sigma_n^2} \right). \quad (55)$$

By replacing $\sigma_n^2, \lambda_1, \ldots, \lambda_K$ by their respective consistent estimates, i.e., $\hat{\sigma}_n^2, \hat{\lambda}_1, \ldots, \hat{\lambda}_K$, we obtain the consistent estimate of $\mathbf{W}_{\text{opt}}$ as

$$\hat{\mathbf{W}}_{\text{opt}} = \hat{\bar{\mathbf{\Gamma}}} \otimes (\mathbf{B}^H \mathbf{B})^{-1} \quad (56)$$

where

$$\hat{\bar{\mathbf{\Gamma}}} = \text{diag} \left(\frac{N}{\hat{\sigma}_n^2} \frac{(\hat{\lambda}_1 - \hat{\sigma}_n^2)^2}{\hat{\lambda}_1} \cdots \frac{N}{\hat{\sigma}_n^2} \frac{(\hat{\lambda}_K - \hat{\sigma}_n^2)^2}{\hat{\lambda}_K} \right). \quad (57)$$

Recall that the solution of (46) is

$$\hat{\mathbf{b}} = \left(\hat{\mathbf{F}}^H \hat{\mathbf{W}} \hat{\mathbf{F}} \right)^{-1} \hat{\mathbf{F}}^H \hat{\mathbf{W}} \hat{\mathbf{g}}. \quad (58)$$

The $\hat{\bar{\mathbf{\Gamma}}}$ can be further simplified since $N/\hat{\sigma}_n^2$ will be canceled out in (58). Now, it is able to replace $\hat{\bar{\mathbf{\Gamma}}}$ by $\hat{\mathbf{\Gamma}}$, such that following (56) will generate the weighting matrix in (30) and the proof is finished.

APPENDIX C
PROOF OF PROPOSITION 3

Our idea to obtain the closed-form formula to the covariance matrix of $\hat{\boldsymbol{\theta}}$ is based on building the connection between θ_k, z_k, and b_k via first- and second-order Taylor series expansion. Similar analysis has been considered in many performance analysis work, e.g., [39]–[41], where a similar analysis is studied for the variance corresponding to a single $\hat{\theta}_k$. However, the covariance matrix of $\hat{\boldsymbol{\theta}}$ is still not available.

It follows from $z_i = e^{j 2\pi d \sin(\theta_i)/v}$ that

$$\Delta \theta_i \approx \frac{v}{2\pi d \cos(\theta_i)} \frac{\Delta z_i}{j z_i} \quad (59)$$

which leads to

$$\Delta \boldsymbol{\theta} \approx [\Delta \theta_1 \ldots \Delta \theta_K]^T$$
$$= -j \mathbf{D} \Delta \mathbf{z} \quad (60)$$

where

$$\mathbf{D} = \text{diag}\left(\frac{\nu}{z_1 2\pi d \cos(\theta_1)} \cdots \frac{\nu}{z_K 2\pi d \cos(\theta_K)}\right). \quad (61)$$

Thus, the covariance matrix of $\hat{\boldsymbol{\theta}}$ is

$$\mathbb{E}\left[\Delta\boldsymbol{\theta}\Delta\boldsymbol{\theta}^T\right] = \frac{1}{4}\mathbb{E}\left[\left(\Delta\boldsymbol{\theta} + \Delta\boldsymbol{\theta}^*\right)\left(\Delta\boldsymbol{\theta} + \Delta\boldsymbol{\theta}^*\right)^H\right]$$

$$\approx \frac{1}{2}\text{Re}\left\{\mathbf{D}\mathbb{E}\left[\Delta\mathbf{z}\Delta\mathbf{z}^T\right]\mathbf{D}\right.$$

$$\left. + \mathbf{D}\mathbb{E}\left[\Delta\mathbf{z}\Delta\mathbf{z}^H\right]\mathbf{D}^H\right\} \quad (62)$$

where from [40], we have

$$\mathbb{E}\left[\Delta\mathbf{z}\Delta\mathbf{z}^H\right] \approx \mathbf{Z}^T(\mathbf{F}^H\mathbf{W}_{\text{opt}}\mathbf{F})^{-1}\mathbf{Z}^* \quad (63)$$

$$\mathbb{E}\left[\Delta\mathbf{z}\Delta\mathbf{z}^T\right] = \mathbf{0}_{K\times K}. \quad (64)$$

Then, (62) can be simplified as

$$\mathbb{E}\left[\Delta\boldsymbol{\theta}\Delta\boldsymbol{\theta}^T\right] \approx \frac{1}{2}\text{Re}\left\{\mathbf{D}\mathbf{Z}^T(\mathbf{F}^H\mathbf{W}_{\text{opt}}\mathbf{F})^{-1}\mathbf{Z}^*\mathbf{D}^H\right\}. \quad (65)$$

Moreover, it has been shown in [40] that the DOA estimate of PUMA is approximately unbiased, i.e.,

$$\mathbb{E}[\Delta\boldsymbol{\theta}] \approx \mathbf{0}_K. \quad (66)$$

This completes the proof.

ACKNOWLEDGMENT

REFERENCES

[1] I. Bekkerman and J. Tabrikian
     Target detection and localization using MIMO radars and sonars
     IEEE Trans. Signal Process., vol. 54, no. 10, pp. 3873–3883, Oct. 2006.

[2] K. T. Wong and M. D. Zoltowski
     Closed-form underwater acoustic direction-finding with arbitrarily spaced vector hydrophones at unknown locations
     IEEE J. Ocean. Eng., vol. 22, no. 3, pp. 566–575, Jul. 1997.

[3] X. Sheng and Y. H. Hu
     Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks
     IEEE Trans. Signal Process., vol. 53, no. 1, pp. 44–53, Jan. 2005.

[4] S. Durrani and M. E. Bialkowski
     Effect of mutual coupling on the interference rejection capabilities of linear and circular arrays in CDMA systems
     IEEE Trans. Antennas Propag., vol. 52, no. 4, pp. 1130–1134, Apr. 2004.

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

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

[7] 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.

[8] N. Hu, Z. Ye, X. Xu, and M. Bao
     DOA estimation for sparse array via sparse signal reconstruction
     IEEE Trans. Aerosp., Electron., Syst., vol. 49, no. 2, pp. 760–773, Apr. 2013.

[9] M. Shaghaghi and S. A. Vorobyov
     Subspace leakage analysis and improved DOA estimation with small sample size
     IEEE Trans. Signal Process., vol. 63, no. 12, pp. 3251–3265, Jun. 2015.

[10] C. Qi, Z. Chen, Y. Wang, and Y. Zhang
     DOA estimation for coherent sources in unknown nonuniform noise fields
     IEEE Trans. Aerosp., Electron., Syst., vol. 43, no. 3, pp. 1195–1204, Jul. 2007.

[11] M. Li and Y. Lu
     Maximum likelihood DOA estimation in unknown colored noise fields
     IEEE Trans. Aerosp., Electron., Syst., vol. 44, no. 3, pp. 1079–1090, Jul. 2008.

[12] C. Qian, L. Huang, and H. C. So
     Computationally efficient ESPRIT algorithm for direction-of-arrival estimation based on Nyström method
     Signal Process., vol. 94, no. 1, pp. 74–80, 2014.

[13] F. J. Chen, S. Kwong, and C. W. Kok
     ESPRIT-like two-dimensional DOA estimation for coherent signals
     IEEE Trans. Aerosp., Electron., Syst., vol. 46, no. 3, pp. 1477–1484, Jul. 2010.

[14] Q. Ren and A. Willis
     Fast root MUSIC algorithm
     Electron. Lett., vol. 33, no. 6, pp. 450–451, 1997.

[15] A. J. Devaney
     Time reversal imaging of obscured targets from multistatic data
     IEEE Trans. Antennas Propag., vol. 53, no. 5, pp. 1600–1610, May 2005.

[16] D. Ciuonzo, G. Romano, and R. Solimene
     Performance analysis of time-reversal MUSIC
     IEEE Trans. Signal Process., vol. 63, no. 10, pp. 2650–2662, May 2015.

[17] E. A. Santiago and M. Saquib
     Noise subspace-based iterative technique for direction finding
     IEEE Trans. Aerosp., Electron., Syst., vol. 49, no. 4, pp. 2281–2295, Oct. 2013.

[18] M. Pesavento, A. B. Gershman, and M. Haardt
     Unitary root-MUSIC with a real-valued eigendecomposition: A theoretical and experimental performance study
     IEEE Trans. Signal Process., vol. 48, no. 5, pp. 1306–1314, May 2000.

[19] 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.

[20] 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.

[21] 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, 1990.

[22] 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.

[23] A. B. Gershman and P. Stoica
MODE with extra-roots (MODEX): A new DOA estimation algorithm with an improved threshold performance
In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1999, vol. 5, pp. 2833–2836.

[24] A. B. Gershman and P. Stoica
New MODE-based techniques for direction finding with an improved threshold performance
Signal Process., vol. 76, no. 3, pp. 221–235, 1999.

[25] D. Zachariah, P. Stoica, and M. Jansson
PUMA criterion = MODE criterion
2017, arXiv:1701.04583.

[26] R. Kumaresan and A. K. Shaw
High resolution bearing estimation without eigen decomposition
In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Tampa, 1985, vol. 10, pp. 576–579.

[27] Y. Bresler and A. Macovski
Exact maximum likelihood parameter estimation of superimposed exponential signals in noise
IEEE Trans. Acoust., Speech, Signal Process., vol. 34, no. 5, pp. 1081–1089, Oct. 1986.

[28] R. Kumaresan, L. L. Scharf, and A. K. Shaw
An algorithm for pole zero-modelling and spectral analysis
IEEE Trans. Acoust., Speech, Signal Process., vol. 34, no. 3, pp. 637–640, Jun. 1986.

[29] Y. Bresler and A. Macovski
Exact maximum likelihood parameter estimation of superimposed exponential signals in noise
IEEE Trans. Acoust., Speech, Signal Process., vol. 34, no. 5, pp. 1081–1089, Oct. 1986.

[30] S. M. Kay
Fundamentals of Statistical Signal Processing: Estimation Theory. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[31] F. K. W. Chan, H. C. So, and W. Sun
Subspace approach for two-dimensional parameter estimation of multiple damped sinusoids
Signal Process., vol. 92, no. 9, pp. 2172–2179, 2012.

[32] M. Viberg, B. Ottersten, and T. Kailath
Detection and estimation in sensor arrays using weighted subspace fitting
IEEE Trans. Signal Process., vol. 39, no. 11, pp. 2436–2449, Nov. 1991.

[33] M. Jansson, A. L. Swindlehurst, and B. Ottersten
Weighted subspace fitting for general array error models
IEEE Trans. Signal Process., vol. 46, no. 9, pp. 2484–2498, Sep. 1998.

[34] 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.

[35] J. Li, P. Stoica, and Z. Liu
Comparative study of IQML and MODE direction-of-arrival estimators
IEEE Trans. Signal Process., vol. 46, no. 1, pp. 149–160, Jan. 1998.

[36] F. Li, H. Liu, and R. J. Vaccaro
Performance analysis for DOA estimation algorithms: Unification, simplication, and observations
IEEE Trans. Aerosp., Electron. Syst., vol. 29, no. 4, pp. 1170–1184, Oct. 1993.

[37] D. J. Jeffries and D. R. Farrier
Asymptotic results for eigenvector methods
IEE Proc. F Commun., Radar Signal Process., vol. 132, no. 7, pp. 589–594, 1985.

[38] M. Kaveh and A. J. Barabell
The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise
IEEE Trans. Acoust., Speech, Signal Process., vol. ASSP-34, no. 2, pp. 331–341, Apr. 1986;

[38] Corrections to: 'The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise'
IEEE Trans. Acoust., Speech. Signal Process vol. ASSP-34, no. 3, p. 633, 1986.

[39] H. C. So and K. W. Chan
A generalized weighted linear predictor frequency estimation approach for a complex sinusoid
IEEE Trans. Signal Process., vol. 54, no. 4, pp. 1304–1315, Apr. 2006.

[40] 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.

[41] H. C. So, Y. T. Chan, K. C. Ho, and Y. Chen
Simple formulas for bias and mean square error computation
IEEE Signal Process. Mag., vol. 30, no. 4, pp. 162–165, Jul. 2013.

Cheng Qian was born in Deqing, Zhejiang, China. He received the B.E. degree in communication engineering from Hangzhou Dianzi University, Hangzhou, China, in 2011, and the M.S. and Ph.D. degrees in information and communication engineering from Harbin Institute of Technology, China, in 2013 and 2017, respectively.

He is currently a Postdoctoral Associate at the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, United States.

His research interests include signal processing and optimization.
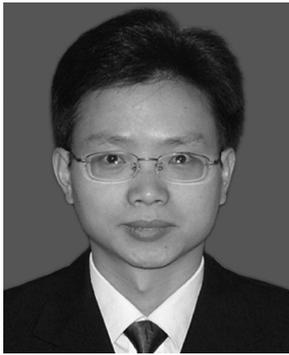
Lei Huang (M'07–SM'14) was born in Guangdong, China. He received the B.Sc., M.Sc., and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2000, 2003, and 2005, respectively.

From 2005 to 2006, he was a Research Associate in the Department of Electrical and Computer Engineering, Duke University, Durham, NC, USA. From 2009 to 2010, he was a Research Fellow in the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, and a Research Associate in the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong. From 2011 to 2014, he was a Professor in the Department of Electronic and Information Engineering, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China. Since November 2014, he has been with the College of Information Engineering, Shenzhen University, Shenzhen, where he is currently a Distinguished Professor. His research interests include spectral estimation, array signal processing, statistical signal processing, and their applications in radar, navigation and wireless communications.

He is currently serving as an Associate Editor for IEEE Transactions on Signal Processing and Digital Signal Processing. Additionally, he is an elected member in Sensor Array and Multichannel (SAM) Technical Committee (TC) of the IEEE Signal Processing Society.

Mingyang Cao (S'15) was born in Daqing, Heilongjiang, China. He received the B.E. degree in electrical and information engineering from Harbin Institute of Technology (HIT), Harbin, China, in 2011, and the M.E. degree in communication and information engineering from HIT Shenzhen Graduate School, Shenzhen, China, in 2013. He is currently working toward the Ph.D. degree in the field of information and communication engineering at HIT. From 2015 to 2016, he was a joint Ph.D. at Aalto University, Espoo, Finland.

His research interests include array signal processing and tensor decomposition.

Hing Cheung So (S'90–M'95–SM'07–F'15) was born in Hong Kong. He received the B.Eng. degree from the City University of Hong Kong, Hong Kong, and the Ph.D. degree from The Chinese University of Hong Kong, Hong Kong, both in electronic engineering, in 1990 and 1995, respectively.

From 1990 to 1991, he was an Electronic Engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. During 1995–1996, he was a Postdoctoral Fellow with The Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor in the Department of Electronic Engineering, City University of Hong Kong, where he is currently a Professor. His research interests include detection and estimation, fast and adaptive algorithms, multidimensional harmonic retrieval, robust signal processing, source localization, and sparse approximation.

Dr. So has been on the editorial boards of IEEE SIGNAL PROCESSING MAGAZINE (2014), IEEE TRANSACTIONS ON SIGNAL PROCESSING(2010–2014), *Signal Processing* (2010), and *Digital Signal Processing* (2011). He is also Lead Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, special issue on "Advances in Time/Frequency Modulated Array Signal Processing."In addtion, he was an elected member in Signal Processing Theory and Methods Technical Committee (2011–2016) of the IEEE Signal Processing Society where he was chair in the awards subcommittee (2015–2016).

Junhao Xie (M'02–SM'15) received the B.S. degree in electronic engineering from Harbin Institute of Shipbuilding Engineering, Harbin, China, the M.S. degree in signal and information processing from Harbin Engineering University, Harbin, and the Ph.D. degree in communication and information system from Harbin Institute of Technology, Harbin, in 1992, 1995, and 2001, respectively.

From 2004 to 2006, he was a Visiting Scholar with the Curtin University of Technology, Perth, W.A., Australia, where he dealt with radar remote sensing signal processing. He is a Professor in the Department of Electronic Engineering, Harbin Institute of Technology. His past and present research interests include radar system analysis and modeling, array signal processing, target detection and estimation, ocean remote sensing for shipborne high-frequency surface wave radar, and inverse synthetic aperture radar.