IEEE TRANSACTIONS ON

SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

APRIL 2004 VOLUME 52 NUMBER 4 ITPRED (ISSN 1053-587X)

PAPERS

(Contents Continued on Back Cover)

IEEE

(Contents Continued from Front Cover)

Computationally Efficient Subspace-Based Method for Direction-of-Arrival Estimation Without Eigendecomposition

Jingmin Xin, *Member, IEEE,* and Akira Sano, *Member, IEEE*

***Abstract*—A computationally simple direction-of-arrival (DOA) estimation method with good statistical performance is attractive in many practical applications of array processing. In this paper, we propose a new computationally efficient subspace-based method without eigendecomposition (SUMWE) for the coherent narrowband signals impinging on a uniform linear array (ULA) by exploiting the array geometry and its shift invariance property. The coherency of incident signals is decorrelated through subarray averaging, and the null space is obtained through a linear operation of a matrix formed from the cross-correlations between some sensor data, where the effect of additive noise is eliminated. Consequently, the DOAs can be estimated without performing eigendecomposition, and there is no need to evaluate all correlations of the array data. Furthermore, the SUMWE is also suitable for the case of partly coherent or incoherent signals, and it can be extended to the spatially correlated noise by choosing appropriate subarrays. The statistical analysis of the SUMWE is studied, and the asymptotic mean-squared-error (MSE) expression of the estimation error is derived. The performance of the SUMWE is demonstrated, and the theoretical analysis is substantiated through numerical examples. It is shown that the SUMWE is superior in resolving closely spaced coherent signals with a small number of snapshots and at low signal-to-noise ratio (SNR) and offers good estimation performance for both uncorrelated and correlated incident signals.**

***Index Terms*—Direction-of-arrival estimation, eigendecomposition, linear operation, multipath environment, subspace-based method.**

I. Introduction

THE directions-of-arrival (DOAs) estimation of signals impinging on an array of sensors is a fundamental problem in array processing used in many fields such as radar, sonar, communications, seismic data processing, and so on [1]–[6]. For the DOA estimation of narrowband signals, the maximum likelihood (ML) and subspace-based methods have been studied extensively (see [1]–[7] and references therein). The ML methods yield the optimal solutions [10], [13], [34]–[39], [45], but their computational burdens are cumbersome because they typically need a nonlinear and multidimensional optimization procedure [4], [7]. The advantage of most subspace-based methods, e.g., multiple signal classification (MUSIC) [8], over the ML methods is their relatively computational simplicity, where the directions are estimated through the search of a one-dimensional spectrum or the calculation of the roots of a certain polynomial based on either the eigenvalue decomposition (EVD) of an array covariance matrix or the singular value decomposition (SVD) of a matrix of array data. Among subspace-based methods, the method of direction estimation (MODE) [9], [10], [13] and weighted subspace fitting (WSF) method [11]–[13] are derived as a large-sample realization of the ML method, and they can be implemented efficiently, especially for the uniform linear array (ULA) case [4], [9], [40]. Unfortunately, the eigendecomposition process is still computationally intensive and time-consuming, which means that subspace-based methods are unsuitable for some practical situations when the number of sensors is large and/or the directions of impinging signals should be tracked in an on-line manner [14].

For alleviating the difficulty of subspace-based methods, many techniques have been developed to reduce the computational burden involved in eigendecomposition (see [14] and [15] for extensive surveys); however, they are rather complex. The last decade has seen the emergence of computationally simple subspace-based direction estimation methods that do not require EVD or SVD computation [16]–[23]. In linear operation based methods such as the bearing estimation without eigendecomposition (BEWE) [16]–[18], orthonormal propagator method (OPM) [19], [20], and subspace methods without eigendecomposition (SWEDE) [21], the signal or noise (null) subspace is easily obtained from the array data relying on a partition of array response matrix, and then, the directions are estimated in a manner similar to that of the MUSIC [8]. Unfortunately, their accuracy is generally poorer than that of the conventional subspace-based methods (e.g., MUSIC) from the statistical viewpoint [18], [20], [21]. The WSF without eigendecomposition (WSF-E) [22] and alternative ML algorithm [23] achieve the asymptotic efficiency when either the number of snapshots or the signal-to-noise ratio (SNR) is large, but they are computationally much more complicated than linear operation based algorithms [16]–[21]. Furthermore, most of these computationally simple subspace-based methods suffer serious degradation when the incident signals are coherent (i.e., fully correlated) in some practical scenarios due to multipath propagation (e.g., [26], [46], [47]), where the rank of the source signal covariance matrix becomes less than the number of incident signals. Although the WSF-E [22]

Manuscript received November 27, 2002; revised June 6, 2003. The associate editor coordinating the review of this paper and approving it for publication was Prof. Abdelhak M. Zoubir.

J. Xin is with the Mobile Communications Development Laboratories, Fujitsu Laboratories Ltd., Yokosuka 239-0847, Japan (e-mail: jxin@jp.fujitsu.com).

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

Digital Object Identifier 10.1109/TSP.2004.823469

and a variant of BEWE [16] can resolve the coherent signals, their performance degrades severely at low SNR and with a small number of snapshots. We previously presented a simple direction estimation method without eigendecomposition for the coherent cyclostationary signals [41], but it is based on the special temporal property of signals, and its performance is poor when the number of snapshots is small.

Therefore, the purpose of this paper is to develop a computationally efficient direction estimation method that would still be accurate. First, we propose a new computationally efficient subspace-based method without eigendecomposition (SUMWE) for the coherent narrowband signals impinging on a ULA by exploiting the array geometry and its shift invariance property. The coherency of incident signals is decorrelated through subarray averaging, and the null space is obtained by a linear operation of the matrix formed from the cross-correlations between some sensor data, where the effect of additive noise is eliminated. Then, the DOAs are estimated easily without the need to evaluate all correlations of the array data, to perform (forward) spatial smoothing (SS) [24] (or forward-backward SS (FBSS) [33]), which is a fairly common technique for subspace-based methods to combat the problem of signal coherency, and to do eigendecomposition. The proposed SUMWE is suitable for the case of partly coherent or incoherent signals, and it can be extended to the spatially correlated noise by choosing appropriate subarrays (i.e., cross-correlations of array data). As a result, the SUMWE has two notable advantages over the (SS- or FBSS-based) MUSIC [8], [24], [33]: computational simplicity and less restrictive noise model. The statistical analysis of the SUMWE is studied, and the explicit expression of asymptotic (large-sample) mean-squared-error (MSE) (or variance) of the estimation error is derived. Furthermore, an analytical study of the SUMWE error variance is performed for the case of one signal, and the quantitative comparisons show that the SUMWE error variance is bounded from below by the asymptotic error variance of the MUSIC estimator and from above by that of the BEWE algorithm. As shown in [31] and [32], the large number of subarrays and the large working array aperture are expected to decorrelate the signal coherency and to improve the resolution of direction estimation. Thus, the performance advantage of the SUMWE method stems from the exploitation of the maximum possible number of subarrays and working array aperture becuase the subarray size is set as the number of incident signals. The estimation performance of the SUMWE is demonstrated, and the theoretical analysis is confirmed through numerical examples. The simulation results show that the SUMWE is superior in resolving closely spaced coherent signals with a short length of data and at low SNR, and it offers good estimation performance for both uncorrelated and correlated incident signals.

II. Data Model and Basic Assumptions

Consider a ULA of M identical and omnidirectional sensors with spacing d, and suppose that p narrowband signals $\{s_k(n)\}$ with the center frequency f_0 are in the field far from the array and impinge on the array from distinct directions $\{\theta_k\}$. Under the narrowband assumption, the received noisy signal $y_i(n)$ at the ith sensor can be expressed as [1]–[7], [13], [31]

$$y_i(n) = x_i(n) + w_i(n) \tag{1}$$

$$x_i(n) = \sum_{k=1}^{p} s_k(n) e^{j\omega_0(i-1)\tau(\theta_k)} \tag{2}$$

where $x_i(n)$ is the noiseless received signal, $w_i(n)$ is the additive noise, $\omega_0 \triangleq 2\pi f_0, \tau(\theta_k) \triangleq (d/c)\sin\theta_k, c$ is the propagation speed, and $\{\theta_k\}$ are measured relative to the normal of array. The received signals can be reexpressed more compactly as

$$\boldsymbol{y}(n) = \boldsymbol{A}(\theta)\boldsymbol{s}(n) + \boldsymbol{w}(n) \tag{3}$$

where $\boldsymbol{y}(n), \boldsymbol{s}(n)$, and $\boldsymbol{w}(n)$ are the vectors of the received signals, the incident signals, and the additive noise given by $\boldsymbol{y}(n) \triangleq [y_1(n), y_2(n), \ldots, y_M(n)]^T, \boldsymbol{s}(n) \triangleq [s_1(n), s_2(n), \ldots, s_p(n)]^T, \boldsymbol{w}(n) \triangleq [w_1(n), w_2(n), \ldots, w_M(n)]^T, \boldsymbol{A}(\theta)$ is the array response matrix given by $\boldsymbol{A}(\theta) \triangleq [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_p)]$ with $\boldsymbol{a}(\theta_k) \triangleq [1, e^{j\omega_0\tau(\theta_k)}, \ldots, e^{j\omega_0(M-1)\tau(\theta_k)}]^T$, and $(\cdot)^T$ denotes transpose.

In this paper, we make the following basic assumptions on the data model.

A1) The array is calibrated, and the array response matrix $\boldsymbol{A}(\theta)$ is unambiguous, i.e., the array response vectors $\{\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \ldots, \boldsymbol{a}(\theta_p)\}$ are linearly independent for any set of distinct $\{\theta_1, \theta_2, \ldots, \theta_p\}$. Equivalently, the matrix $\boldsymbol{A}(\theta)$ has full rank.

A2) Without loss of generality, the signals $\{s_k(n)\}$ are all coherent so that they are all some complex multiples of a common signal $s_1(n)$; then, under the flat-fading multipath propagation, they can be expressed as [6], [24], [26]–[33], [39]

$$s_k(n) = \beta_k s_1(n), \quad \text{for } k = 1, 2, \ldots, p \tag{4}$$

where β_k is the complex attenuation coefficient with $\beta_k \neq 0$, and $\beta_1 = 1$.

A3) For the simplicity of theoretical performance analysis, the incident signal $s_1(n)$ is a temporally complex white Gaussian random process with zero-mean and the variance given by

$$E\{s_1(n)s_1^*(t)\} = r_s\delta_{n,t}, \quad \forall n, t$$
$$E\{s_1(n)s_1(t)\} = 0, \quad \forall n, t \tag{5}$$

where $E\{\cdot\}, (\cdot)^*$, and $\delta_{n,t}$ denote the expectation, the complex conjugate, and Kronecker delta.

A4) The additive noise $\{w_i(n)\}$ is a temporally and spatially complex white Guassian random process with zero-mean and the following covariance matrix

$$E\{\boldsymbol{w}(n)\boldsymbol{w}^H(t)\} = \sigma^2\boldsymbol{I}_M\delta_{n,t}, \quad \forall n, t$$
$$E\{\boldsymbol{w}(n)\boldsymbol{w}^T(t)\} = \boldsymbol{O}_{M\times M}, \quad \forall n, t \tag{6}$$

where $\boldsymbol{I}_m, \boldsymbol{O}_{m\times q}$, and $(\cdot)^H$ indicate the $m \times m$ identity matrix, the $m \times q$ null matrix, and Hermitian transpose. In addition, the additive noise is uncorrelated with the incident signals.

A5) The number of incident signals p is known, and it satisfies the inequality that $p < M/2$ for an array of M sensors.

Remark A: Although the incident signals are assumed to be fully coherent, the proposed SUMWE algorithm can be extended to the case of partly coherent or incoherent signals (see Remark B). The identifiability condition that guarantees the uniqueness of direction estimation is that $M > 2p$, which is less strict than the sufficient condition $M > 2p-1$ with every single snapshot in the absence of additive noise and the necessary condition $M \geq 3p/2$ with probability one (w.p. 1) [42]. Additionally the number of sensors necessitated by the SUMWE is smaller than that $M \geq 3p$ needed by the SWEDE [21] for the case of incoherent signals. Furthermore, if the number of signals p is unknown, it can be estimated by some proposed techniques (e.g., [31] and references therein). □

In the following, we concentrate on the problem of estimating the DOAs $\{\theta_k\}$ of coherent signals from N snapshots of noisy array data $\{\boldsymbol{y}(n)\}_{n=1}^{N}$ in a computationally simple way with good statistical performance.

III. Subspace-Based Method Without Eigendecomposition—SUMWE

A. Decorrelation With Subarray Averaging

In the presence of coherent signals, the source signal covariance matrix becomes singular, i.e., the rank of the source signal covariance matrix is smaller than the number of signals. In order to estimate the DOAs of incident signals, we must ensure that the dimension of signal subspace used in the direction estimation is equal to the number of signals. Here, we use subarray averaging to tackle the problem of rank deficit.

First, the noiseless received signal $x_i(n)$ in (2) can be reexpressed by

$$x_i(n) = \boldsymbol{b}_i^T(\theta)\boldsymbol{s}(n) = \boldsymbol{s}^T(n)\boldsymbol{b}_i(\theta) \tag{7}$$

where

$$\boldsymbol{b}_i(\theta) \triangleq [e^{j\omega_0(i-1)\tau(\theta_1)}, e^{j\omega_0(i-1)\tau(\theta_2)}, \ldots, e^{j\omega_0(i-1)\tau(\theta_p)}]^T$$

and the correlation r_{ik} of signals $y_i(n)$ and $y_k(n)$ is defined as $r_{ik} \triangleq E\{y_i(n)y_k^*(n)\}$, where $r_{ik} = r_{ki}^*$. Then, the full array can be divided into L overlapping subarrays with p sensors in the forward and backward directions [24], [33], where $L = M-p+1$, and the lth forward or backward subarray comprises $\{l, l+1, \ldots, l+p-1\}$ or $\{M-l+1, M-l, \ldots, L-l+1\}$ sensors, respectively, where $l = 1, 2, \ldots, L$. From (1) and (7), we can express the signal vector $\boldsymbol{y}_{fl}(n)$ of the noisy signals in the lth forward subarray and that $\boldsymbol{y}_{bl}(n)$ of the conjugate noisy signals in the lth backward subarray as [24], [5], [30]–[33]

$$\boldsymbol{y}_{fl}(n) = \boldsymbol{A}_1(\theta)\boldsymbol{D}^{l-1}\boldsymbol{s}(n) + \boldsymbol{w}_{fl}(n) \tag{8}$$

$$\boldsymbol{y}_{bl}(n) = \boldsymbol{A}_1(\theta)\boldsymbol{D}^{-(M-l)}\boldsymbol{s}^*(n) + \boldsymbol{w}_{bl}(n) \tag{9}$$

where $\boldsymbol{y}_{fl}(n) \triangleq [y_l(n), y_{l+1}(n), \ldots, y_{l+p-1}(n)]^T, \boldsymbol{y}_{bl}(n) \triangleq [y_{M-l+1}(n), y_{M-l}(n), \ldots, y_{L-l+1}(n)]^H, \boldsymbol{w}_{fl}(n) \triangleq [w_l(n), w_{l+1}(n), \ldots, w_{l+p-1}(n)]^T, \boldsymbol{w}_{bl}(n) \triangleq [w_{M-l+1}(n), w_{M-l}(n), \ldots, w_{L-l+1}(n)]^H, \boldsymbol{D} \triangleq \text{diag}(e^{j\omega_0\tau(\theta_1)}, e^{j\omega_0\tau(\theta_2)}, \ldots, e^{j\omega_0\tau(\theta_p)})$, and $\boldsymbol{A}_1(\theta)$ is the submatrix of $\boldsymbol{A}(\theta)$ in (3) consisting of the first p rows with the column $\boldsymbol{a}_1(\theta_k) \triangleq [1, e^{j\omega_0\tau(\theta_k)}, \ldots, e^{j\omega_0(p-1)\tau(\theta_k)}]^T$.

Under the basic assumptions, from (4)–(8), we can obtain the correlation $\boldsymbol{\varphi}_{fl}$ between the signal vector $\boldsymbol{y}_{fl}(n)$ of the lth forward subarray and the signal $y_M^*(n)$

$$\begin{aligned}\boldsymbol{\varphi}_{fl} &\triangleq E\{\boldsymbol{y}_{fl}(n)y_M^*(n)\} = [r_{lM}, r_{l+1\,M}, \ldots, r_{l+p-1\,M}]^T \\ &= E\{\boldsymbol{A}_1(\theta)\boldsymbol{D}^{l-1}\boldsymbol{\beta}s_1(n)s_1^*(n)\boldsymbol{\beta}^H\boldsymbol{b}_M^*(\theta)\} \\ &\quad + E\{\boldsymbol{w}_{fl}(n)w_M^*(n)\} \\ &= \rho_M r_s \boldsymbol{A}_1(\theta)\boldsymbol{D}^{l-1}\boldsymbol{\beta}, \quad \text{for } l = 1, 2, \ldots, L-1\end{aligned} \tag{10}$$

where $\rho_i \triangleq \boldsymbol{\beta}^H\boldsymbol{b}_i^*(\theta)$, and $\boldsymbol{\beta} \triangleq [\beta_1, \beta_2, \ldots, \beta_p]^T$. Hence, by concatenating the correlations $\{\boldsymbol{\varphi}_{fl}\}$ in (10) and by performing some algebraic manipulations, we can form a correlation matrix $\boldsymbol{\Phi}_f$ as

$$\begin{aligned}\boldsymbol{\Phi}_f &\triangleq [\boldsymbol{\varphi}_{f1}, \boldsymbol{\varphi}_{f2}, \ldots, \boldsymbol{\varphi}_{f\,L-1}]^T \\ &= \rho_M r_s \begin{bmatrix} \boldsymbol{\beta}^T \\ \boldsymbol{\beta}^T\boldsymbol{D} \\ \vdots \\ \boldsymbol{\beta}^T\boldsymbol{D}^{L-2} \end{bmatrix} \boldsymbol{A}_1^T(\theta) = \rho_M r_s \bar{\boldsymbol{A}}(\theta)\boldsymbol{B}\boldsymbol{A}_1^T(\theta)\end{aligned} \tag{11}$$

where $\boldsymbol{B} \triangleq \text{diag}(\beta_1, \beta_2, \ldots, \beta_p)$, and $\bar{\boldsymbol{A}}(\theta)$ is the submatrix of $\boldsymbol{A}(\theta)$ consisting of the first $L-1$ rows with the column $\bar{\boldsymbol{a}}(\theta_k) \triangleq [1, e^{j\omega_0\tau(\theta_k)}, \ldots, e^{j\omega_0(L-2)\tau(\theta_k)}]^T$. Straightforwardly, we can see that the Hankel matrix $\boldsymbol{\Phi}_f$ is composed of the cross-correlations $\{r_{1M}, r_{2M}, \ldots, r_{M-1M}\}$. Similarly, we can get the correlation $\bar{\boldsymbol{\varphi}}_{fl}$ between the signal vector $\boldsymbol{y}_{fl}(n)$ and the signal $y_1^*(n)$

$$\begin{aligned}\bar{\boldsymbol{\varphi}}_{fl} &\triangleq E\{\boldsymbol{y}_{fl}(n)y_1^*(n)\} = [r_{l1}, r_{l+11}, \ldots, r_{l+p-11}]^T \\ &= \rho_1 r_s \boldsymbol{A}_1(\theta)\boldsymbol{D}^{l-1}\boldsymbol{\beta}, \quad \text{for } l = 2, 3, \ldots, L.\end{aligned} \tag{12}$$

In addition, we can easily obtain a correlation matrix $\bar{\boldsymbol{\Phi}}_f$ formed from the correlations $\{\bar{\boldsymbol{\varphi}}_{fl}\}$ in (12)

$$\begin{aligned}\bar{\boldsymbol{\Phi}}_f &\triangleq [\bar{\boldsymbol{\varphi}}_{f2}, \bar{\boldsymbol{\varphi}}_{f3}, \ldots, \bar{\boldsymbol{\varphi}}_{fL}]^T \\ &= \rho_1 r_s \bar{\boldsymbol{A}}(\theta)\boldsymbol{B}\boldsymbol{D}\boldsymbol{A}_1^T(\theta).\end{aligned} \tag{13}$$

Obviously, the Hankel matrix $\bar{\boldsymbol{\Phi}}_f$ consists of the cross-correlations $\{r_{21}, r_{31}, \ldots, r_{M1}\}$.

By evaluating the correlation $\boldsymbol{\varphi}_{bl} \triangleq E\{y_1(n)\boldsymbol{y}_{bl}(n)\}$ between $y_1(n)$ and $\{\boldsymbol{y}_{bl}(n)\}$ and the correlation $\bar{\boldsymbol{\varphi}}_{bl} \triangleq E\{y_M(n)\boldsymbol{y}_{bl}(n)\}$ between $y_M(n)$ and $\{\boldsymbol{y}_{bl}(n)\}$, we can obtain the correlation matrices $\boldsymbol{\Phi}_b$ and $\bar{\boldsymbol{\Phi}}_b$ for the backward subarrays

$$\begin{aligned}\boldsymbol{\Phi}_b &\triangleq [\boldsymbol{\varphi}_{b1}, \boldsymbol{\varphi}_{b2}, \ldots, \boldsymbol{\varphi}_{bL-1}]^T \\ &= \rho_1^* r_s \bar{\boldsymbol{A}}(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-1)}\boldsymbol{A}_1^T(\theta)\end{aligned} \tag{14}$$

$$\begin{aligned}\bar{\boldsymbol{\Phi}}_b &\triangleq [\bar{\boldsymbol{\varphi}}_{b2}, \bar{\boldsymbol{\varphi}}_{b3}, \ldots, \bar{\boldsymbol{\varphi}}_{bL}]^T \\ &= \rho_M^* r_s \bar{\boldsymbol{A}}(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-2)}\boldsymbol{A}_1^T(\theta).\end{aligned} \tag{15}$$

It is apparent that the Hankel matrix $\boldsymbol{\Phi}_b$ consists of the cross-correlations $\{r_{1M}, r_{1M-1}, \ldots, r_{12}\}$, whereas the Hankel matrix $\bar{\boldsymbol{\Phi}}_b$ is formed by the cross-correlations $\{r_{M\,M-1}, r_{M\,M-2}, \ldots, r_{M1}\}$. Furthermore, we can easily find that $\boldsymbol{\Phi}_b = \boldsymbol{J}_{L-1}\bar{\boldsymbol{\Phi}}_f^*\boldsymbol{J}_p$ and $\bar{\boldsymbol{\Phi}}_b = \boldsymbol{J}_{L-1}\boldsymbol{\Phi}_f^*\boldsymbol{J}_p$, where $\boldsymbol{J}_m$ is

an $m \times m$ counteridentity (or reversal) matrix that has unity elements along the cross-diagonal and zeros elsewhere.

From (11) and (13)–(15), we can straightforwardly see that the correlation matrices $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$ are not affected by the additive noise $\{w_i(n)\}$ (see Remark E for reference). Because $M > 2p$ (i.e., $L-1 > p$), $\beta_k \neq 0$, and $\boldsymbol{A}(\theta)$ is a Vandermonde matrix with full rank, we can find that the ranks of these $(L-1) \times p$ correlation matrices equal p [30]–[32], i.e., the dimension of their signal subspace equals the number of coherent signals. Hence, the directions $\{\theta_k\}$ of coherent signals can be estimated from the subspaces of these matrices $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$.

Remark B: When some of the incident signals are coherent and the others are uncorrelated with these signals and with each other, by assuming that the first q $(1 \leq q \leq p)$ signals are coherent as defined in (4) and performing some algebraic manipulations, we can obtain

$$\boldsymbol{\Phi}_f = \bar{\rho}_M r_{s_1} \bar{\boldsymbol{A}}(\theta)\bar{\boldsymbol{B}}\boldsymbol{A}_1^T(\theta) + \bar{\boldsymbol{A}}(\theta)\boldsymbol{D}^{-(M-1)}\tilde{\boldsymbol{R}}_s\boldsymbol{A}_1^T(\theta) \quad (16)$$

$$\bar{\boldsymbol{\Phi}}_f = \bar{\rho}_1 r_{s_1} \bar{\boldsymbol{A}}(\theta)\bar{\boldsymbol{B}}\boldsymbol{D}\boldsymbol{A}_1^T(\theta) + \bar{\boldsymbol{A}}(\theta)\boldsymbol{D}\tilde{\boldsymbol{R}}_s\boldsymbol{A}_1^T(\theta) \quad (17)$$

$$\boldsymbol{\Phi}_b = \bar{\rho}_1^* r_{s_1} \bar{\boldsymbol{A}}(\theta)\bar{\boldsymbol{B}}^*\boldsymbol{D}^{-(M-1)}\boldsymbol{A}_1^T(\theta) + \bar{\boldsymbol{A}}(\theta)\boldsymbol{D}^{-(M-1)}\tilde{\boldsymbol{R}}_s\boldsymbol{A}_1^T(\theta) \quad (18)$$

$$\bar{\boldsymbol{\Phi}}_b = \bar{\rho}_M^* r_{s_1} \bar{\boldsymbol{A}}(\theta)\bar{\boldsymbol{B}}^*\boldsymbol{D}^{-(M-2)}\boldsymbol{A}_1^T(\theta) + \bar{\boldsymbol{A}}(\theta)\boldsymbol{D}\tilde{\boldsymbol{R}}_s\boldsymbol{A}_1^T(\theta) \quad (19)$$

where $r_{s_k} \triangleq E\{s_k(n)s_k^*(n)\}, \bar{\rho}_i \triangleq \bar{\boldsymbol{\beta}}^H\boldsymbol{b}_i^*(\theta), \bar{\boldsymbol{\beta}} \triangleq [\beta_1, \beta_2, \ldots, \beta_q, 0, \ldots, 0]^T, \bar{\boldsymbol{B}} \triangleq \text{diag}(\beta_1, \beta_2, \ldots, \beta_q, 0, \ldots, 0)$, and $\tilde{\boldsymbol{R}}_s \triangleq \text{diag}(0, \ldots, 0, r_{s_{q+1}}, r_{s_{q+2}}, \ldots, r_{s_p})$. Then, we show that the ranks of the matrices in (16)–(19) still equal the number of incident signals. □

B. Direction Estimation Without Eigendecomposition

Because it is assumed that $M > 2p$ (i.e., $L-1 > p$), we can divide the $(L-1) \times p$ matrix $\bar{\boldsymbol{A}}(\theta)$ in (11) and (13)–(15) into two parts as follows:

$$\bar{\boldsymbol{A}}(\theta) \triangleq \begin{bmatrix} \boldsymbol{A}_1(\theta) \\ \boldsymbol{A}_2(\theta) \end{bmatrix} \begin{matrix} \}p \\ \}L-p-1 \end{matrix} \quad (20)$$

where $\boldsymbol{A}_2(\theta)$ has the column $\boldsymbol{a}_2(\theta_k) \triangleq [e^{j\omega_0 p\tau(\theta_k)}, e^{j\omega_0(p+1)\tau(\theta_k)}, \ldots, e^{j\omega_0(L-2)\tau(\theta_k)}]^T$. Since $\bar{\boldsymbol{A}}(\theta)$ and $\boldsymbol{A}_1(\theta)$ are two submatrices of the Vandemonde matrix $\boldsymbol{A}(\theta)$ with full rank in (3), the rows of $\boldsymbol{A}_2(\theta)$ can be expressed as a linear combination of linearly independent rows of $\boldsymbol{A}_1(\theta)$; equivalently, there is a $p \times (L-p-1)$ linear operator $\boldsymbol{P}$ between $\boldsymbol{A}_1(\theta)$ and $\boldsymbol{A}_2(\theta)$ [19], [20]

$$\boldsymbol{P}^H\boldsymbol{A}_1(\theta) = \boldsymbol{A}_2(\theta). \quad (21)$$

Then, it follows from (21) that

$$\boldsymbol{Q}^H\bar{\boldsymbol{A}}(\theta) = \boldsymbol{O}_{(L-p-1)\times p} \quad (22)$$

where $\boldsymbol{Q} \triangleq [\boldsymbol{P}^T, -\boldsymbol{I}_{L-p-1}]^T$. Because the $(L-1)\times(L-p-1)$ matrix $\boldsymbol{Q}$ has a full rank of $L-p-1$, the columns of $\boldsymbol{Q}$ in fact form the basis for the null space $\mathcal{N}(\bar{\boldsymbol{A}}^H(\theta))$ of $\bar{\boldsymbol{A}}^H(\theta)$, and clearly, the orthogonal projector onto this subspace is given by $\boldsymbol{\Pi}_Q \triangleq \boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H$, which implies that [10]

$$\boldsymbol{\Pi}_Q\bar{\boldsymbol{a}}(\theta) = \boldsymbol{0}_{(L-1)\times 1}, \quad \text{for } \theta = \theta_k \quad (23)$$

where $\bar{\boldsymbol{a}}(\theta) \triangleq [1, e^{j\omega_0\tau(\theta)}, \ldots, e^{j\omega_0(L-2)\tau(\theta)}]^T$, and $\boldsymbol{0}_{m\times 1}$ is an $m \times 1$ null vector. Evidently, the directions can be estimated based on the orthogonal property (23) without any eigendecomposition.

Remark C: By considering the SVD of matrix $\bar{\boldsymbol{A}}(\theta)$, we readily verify that the orthogonal projector $\boldsymbol{\Pi}_Q$ in (23) can be written as $\boldsymbol{\Pi}_Q = \boldsymbol{I}_{L-1} - \bar{\boldsymbol{A}}(\theta)(\bar{\boldsymbol{A}}^H(\theta)\bar{\boldsymbol{A}}(\theta))^{-1}\bar{\boldsymbol{A}}^H(\theta)$ [10], [3]. □

Then, the next problem is how to find the null space of $\bar{\boldsymbol{A}}^H(\theta)$ (i.e., $\boldsymbol{P}$) from the available array data. Based on the partition of $\bar{\boldsymbol{A}}(\theta)$ in (20), the $(L-1) \times p$ correlation matrices $\boldsymbol{\Phi}_f$ in (11), $\bar{\boldsymbol{\Phi}}_f$ in (13), $\boldsymbol{\Phi}_b$ in (14), and $\bar{\boldsymbol{\Phi}}_b$ in (15) can be also divided as

$$\boldsymbol{\Phi}_f = \rho_M r_s \begin{bmatrix} \boldsymbol{A}_1(\theta)\boldsymbol{B}\boldsymbol{A}_1^T(\theta) \\ \boldsymbol{A}_2(\theta)\boldsymbol{B}\boldsymbol{A}_1^T(\theta) \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{\Phi}_{f1} \\ \boldsymbol{\Phi}_{f2} \end{bmatrix} \begin{matrix} \}p \\ \}L-p-1 \end{matrix} \quad (24)$$

$$\bar{\boldsymbol{\Phi}}_f = \rho_1 r_s \begin{bmatrix} \boldsymbol{A}_1(\theta)\boldsymbol{B}\boldsymbol{D}\boldsymbol{A}_1^T(\theta) \\ \boldsymbol{A}_2(\theta)\boldsymbol{B}\boldsymbol{D}\boldsymbol{A}_1^T(\theta) \end{bmatrix} \triangleq \begin{bmatrix} \bar{\boldsymbol{\Phi}}_{f1} \\ \bar{\boldsymbol{\Phi}}_{f2} \end{bmatrix} \begin{matrix} \}p \\ \}L-p-1 \end{matrix} \quad (25)$$

$$\boldsymbol{\Phi}_b = \rho_1^* r_s \begin{bmatrix} \boldsymbol{A}_1(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-1)}\boldsymbol{A}_1^T(\theta) \\ \boldsymbol{A}_2(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-1)}\boldsymbol{A}_1^T(\theta) \end{bmatrix} \triangleq \begin{bmatrix} \boldsymbol{\Phi}_{b1} \\ \boldsymbol{\Phi}_{b2} \end{bmatrix} \begin{matrix} \}p \\ \}L-p-1 \end{matrix} \quad (26)$$

$$\bar{\boldsymbol{\Phi}}_b = \rho_M^* r_s \begin{bmatrix} \boldsymbol{A}_1(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-2)}\boldsymbol{A}_1^T(\theta) \\ \boldsymbol{A}_2(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-2)}\boldsymbol{A}_1^T(\theta) \end{bmatrix} \triangleq \begin{bmatrix} \bar{\boldsymbol{\Phi}}_{b1} \\ \bar{\boldsymbol{\Phi}}_{b2} \end{bmatrix} \begin{matrix} \}p \\ \}L-p-1 \end{matrix}. \quad (27)$$

From (21), we can obtain the following relation between the submatrices of $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$:

$$\boldsymbol{P}^H\boldsymbol{\Phi}_1 = \boldsymbol{\Phi}_2 \quad (28)$$

where $\boldsymbol{\Phi}_1 \triangleq [\boldsymbol{\Phi}_{f1}, \bar{\boldsymbol{\Phi}}_{f1}, \boldsymbol{\Phi}_{b1}, \bar{\boldsymbol{\Phi}}_{b1}]$, and $\boldsymbol{\Phi}_2 \triangleq [\boldsymbol{\Phi}_{f2}, \bar{\boldsymbol{\Phi}}_{f2}, \boldsymbol{\Phi}_{b2}, \bar{\boldsymbol{\Phi}}_{b2}]$. Thus, the matrix $\boldsymbol{P}$ can be found from $\boldsymbol{\Phi}_1$ and $\boldsymbol{\Phi}_2$ as

$$\boldsymbol{P} = \boldsymbol{A}_1^{-H}(\theta)\boldsymbol{A}_2^H(\theta) = (\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H)^{-1}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2^H. \quad (29)$$

Proof: See Appendix A. ■

Remark D: To avoid the effect of additive noise $\{w_i(n)\}$, the auto-correlations $\{r_{ii}\}$ of the sensor data are not used in the computation of the projector onto the null space in the SWEDE and the proposed SUMWE method. In general, the SWEDE needs two $p\times p$ and one $(M-2p)\times p$ subarray cross-covariance matrices of three nonoverlapping subarrays with p, p, and $M-2p$ sensors, and these three subarray cross-covariance matrices share two linear operators $\boldsymbol{P}_1 = -\tilde{\boldsymbol{A}}_1^{-H}(\theta)\tilde{\boldsymbol{A}}_3^H(\theta)$ and $\boldsymbol{P}_2 = -\tilde{\boldsymbol{A}}_2^{-H}(\theta)\tilde{\boldsymbol{A}}_3^H(\theta)$, where $\boldsymbol{A}^H(\theta) \triangleq [\underbrace{\tilde{\boldsymbol{A}}_1^H(\theta)}_{p}, \underbrace{\tilde{\boldsymbol{A}}_2^H(\theta)}_{p}, \underbrace{\tilde{\boldsymbol{A}}_3^H(\theta)}_{M-2p}]$ (see [21] for details). However, in the SUMWE method, the four matrices $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$ share a common linear operator $\boldsymbol{P}$ because of the exploitation of the ULA geometry and its shift invariance property. □

Therefore, when the finite array data are available, the directions $\{\theta_k\}$ of coherent signals can be estimated in a manner similar to that of the MUSIC [8] by minimizing the following cost function:

$$f(\theta) = \bar{\boldsymbol{a}}^H(\theta)\boldsymbol{\Pi}_{\hat{Q}}\,\bar{\boldsymbol{a}}(\theta) \quad (30)$$

where $\boldsymbol{\Pi}_{\hat{Q}} = \hat{\boldsymbol{Q}}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\hat{\boldsymbol{Q}}^H$, and $\hat{\boldsymbol{P}} = (\hat{\boldsymbol{\Phi}}_1\hat{\boldsymbol{\Phi}}_1^H)^{-1}\hat{\boldsymbol{\Phi}}_1\hat{\boldsymbol{\Phi}}_2^H$.

C. Implementation of SUMWE

Based on the above analysis and by using the matrix inversion lemma (e.g., see [3] and [49]), the implementation of the SUMWE for estimating the directions of coherent signals with the finite array data $\{\boldsymbol{y}(n)\}_{n=1}^{N}$ can be summarized as follows.

Step 1) Calculate the estimated correlation vector $\hat{\boldsymbol{\varphi}}$ between $\boldsymbol{y}(n)$ and $y_M^*(n)$ and that $\hat{\bar{\boldsymbol{\varphi}}}$ between $\boldsymbol{y}(n)$ and $y_1^*(n)$ as

$$\hat{\boldsymbol{\varphi}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{y}(n)y_M^*(n)$$
$$\hat{\bar{\boldsymbol{\varphi}}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{y}(n)y_1^*(n) \tag{31}$$

where $\hat{\boldsymbol{\varphi}} = [\hat{r}_{1M}, \hat{r}_{2M}, \ldots, \hat{r}_{MM}]^T$, and $\hat{\bar{\boldsymbol{\varphi}}} = [\hat{r}_{11}, \hat{r}_{21}, \ldots, \hat{r}_{M1}]^T$.

Step 2) Form the estimated correlation matrices $\hat{\boldsymbol{\Phi}}_f$, $\hat{\bar{\boldsymbol{\Phi}}}_f, \hat{\boldsymbol{\Phi}}_b$, and $\hat{\bar{\boldsymbol{\Phi}}}_b$ from $\hat{\boldsymbol{\varphi}}$ and $\hat{\bar{\boldsymbol{\varphi}}}$ by using (11) and (13)–(15).

Step 3) Estimate the linear operator $\boldsymbol{P}$ as

$$\hat{\boldsymbol{P}} = (\hat{\boldsymbol{\Phi}}_1\hat{\boldsymbol{\Phi}}_1^H)^{-1}\hat{\boldsymbol{\Phi}}_1\hat{\boldsymbol{\Phi}}_2^H \tag{32}$$

and calculate the estimated orthogonal projector $\boldsymbol{\Pi}_{\hat{Q}}$ as

$$\boldsymbol{\Pi}_{\hat{Q}} = \hat{\boldsymbol{Q}}(\boldsymbol{I}_{L-p-1} - \hat{\boldsymbol{P}}^H(\hat{\boldsymbol{P}}\hat{\boldsymbol{P}}^H + \boldsymbol{I}_p)^{-1}\hat{\boldsymbol{P}})\hat{\boldsymbol{Q}}^H. \tag{33}$$

Step 4) Estimate the directions $\{\theta_k\}$ by searching the p highest peaks of the spatial spectrum $P(\theta)$ (the spectral approach) or by finding the phases of the p zeros of the polynomial $p(z)$ closest to the unit circle in the z-plane (the root approach), where $P(\theta) \triangleq 1/\bar{\boldsymbol{a}}^H(\theta)\boldsymbol{\Pi}_{\hat{Q}}\bar{\boldsymbol{a}}(\theta), p(z) \triangleq z^{L-2}\boldsymbol{p}^H(z)\boldsymbol{\Pi}_{\hat{Q}}\boldsymbol{p}(z), \boldsymbol{p}(z) \triangleq [1, z, \ldots, z^{L-2}]^T$, and $z \triangleq e^{j\omega_0\tau(\theta)}$.

Remark E: By defining the array covariance matrix $\boldsymbol{R}$ as $\boldsymbol{R} \triangleq E\{\boldsymbol{y}(n)\boldsymbol{y}^H(n)\} = \boldsymbol{A}(\theta)\boldsymbol{R}_s\boldsymbol{A}^H(\theta) + \sigma^2\boldsymbol{I}_M$, where $\boldsymbol{R}_s \triangleq E\{\boldsymbol{s}(n)\boldsymbol{s}^H(n)\}$, we can find that the matrices $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$ are formed from the correlations in the Mth and first columns and those in the first and Mth rows of $\boldsymbol{R}$, except for the auto-correlations r_{11} and r_{MM}, which contain the variance σ^2 of additive noise. Hence, the number of needed correlations $\{r_{ik}\}$ is $4M-6$, which is smaller than $M + (m-1)(2M-m)$, which is needed by the SS-based MUSIC [24], [30], where m is the number of sensors in the subarray with $p < m < M$. Furthermore, because the matrix $\boldsymbol{R}$ is Hermitian, the actual number of needed correlations is $2M-3$. □

Remark F: Like the SWEDE [21] and BEWE [16], the SUMWE can accommodate a more general noise model of the spatially correlated noise if we choose the signal vectors $\boldsymbol{y}_{fl}(n)$ and $\boldsymbol{y}_{bl}(n)$ used to form the matrices $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$ appropriately. If the spatial correlation length of additive noise is $\bar{q}$ (i.e., $E\{w_i(n)w_{i+k}^*(n)\} = 0$ for $|k| > \bar{q}$) [18] and $M > 2p + \bar{q}$, where $0 \le \bar{q} < M-1$, then the noise covariance matrix will become a banded Toeplitz matrix with $2\bar{q}+1$ nondiagonals. To avoid the effect of additive noise in direction estimation, we can define the correlation matrices as $\boldsymbol{\Phi}_f \triangleq [\boldsymbol{\varphi}_{f1}, \boldsymbol{\varphi}_{f2}, \ldots, \boldsymbol{\varphi}_{f\,L-\bar{q}-1}]^T, \bar{\boldsymbol{\Phi}}_f \triangleq [\bar{\boldsymbol{\varphi}}_{f\bar{q}+2}, \bar{\boldsymbol{\varphi}}_{f\bar{q}+3}, \ldots, \bar{\boldsymbol{\varphi}}_{fL}]^T, \boldsymbol{\Phi}_b \triangleq [\boldsymbol{\varphi}_{b1}, \boldsymbol{\varphi}_{b2}, \ldots, \boldsymbol{\varphi}_{bL-\bar{q}-1}]^T$, and $\bar{\boldsymbol{\Phi}}_b \triangleq [\bar{\boldsymbol{\varphi}}_{b\bar{q}+2}, \bar{\boldsymbol{\varphi}}_{b\bar{q}+3}, \ldots, \bar{\boldsymbol{\varphi}}_{bL}]^T$. Then, the expressions of $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$ in (11) and (13)–(15) still hold, except that their dimensions become $(L-\bar{q}-1)\times p$, and their ranks are still p based on the assumption that $M > 2p + \bar{q}$ (i.e., $L - \bar{q} - 1 > p$). Thus, the proposed SUMWE algorithm is still valid. □

Remark G: The implementation of the proposed SUMWE algorithm requires two major steps: i) computation of the correlations $\hat{\boldsymbol{\varphi}}$ and $\hat{\bar{\boldsymbol{\varphi}}}$ to form the matrices $\hat{\boldsymbol{\Phi}}_f, \hat{\bar{\boldsymbol{\Phi}}}_f, \hat{\boldsymbol{\Phi}}_b$, and $\hat{\bar{\boldsymbol{\Phi}}}_b$, and ii) estimation of the orthogonal projector $\boldsymbol{\Pi}_{\hat{Q}}$. The number of flops needed to calculate $\hat{\boldsymbol{\varphi}}$ and $\hat{\bar{\boldsymbol{\varphi}}}$ in (31) is approximately $16NM$, where a flop is defined as a floating-point addition or multiplication operation as adopted by MATLAB software. The computation of $\hat{\boldsymbol{P}}$ in (32) takes about $32p^3 + 40p^2(M-2p) + O(p^3)$ flops, whereas the calculation of $\boldsymbol{\Pi}_{\hat{Q}}$ in (33) requires roughly $16p^2(M-2p) + 8M(M-2p)^2 + 8(M-p)^2(M-2p) + O(p^3)$ flops, where the number of arithmetic operations needed by the inversion of a $p\times p$ Hermitian matrix is in the order of $O(p^3)$, which is about $40p^3$ flops for $p=2$. Thus, the estimated number of MATLAB flops required by the SUMWE algorithm is nearly $16NM + 16M(M-p)^2$ when $N \gg M \gg p$, which occurs often in applications of DOA estimation, where the computations needed in the remaining steps are negligible. □

Remark H: If the orthogonal projector $\boldsymbol{\Pi}_{\hat{Q}}$ is calculated directly by using $\boldsymbol{\Pi}_{\hat{Q}} = \hat{\boldsymbol{Q}}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\hat{\boldsymbol{Q}}^H$ in (30), it will necessitate the inversion of the $(M-2p)\times(M-2p)$ matrix $\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}}$ and take approximately $16(M-p)(M-2p)^2 + 8(M-p)^2(M-2p) + O((M-2p)^3)$ flops. However, the calculation of $\boldsymbol{\Pi}_{\hat{Q}}$ in (33) involves the inversion of the $p\times p$ matrix $\hat{\boldsymbol{P}}\hat{\boldsymbol{P}}^H + \boldsymbol{I}_p$ in $O(p^3)$ operations. From Remark G, we can find that the number of flops needed by the direct computation of $\boldsymbol{\Pi}_{\hat{Q}}$ in (30) is generally larger than that of the calculation in (33) when $M \gg p$. Thus, the proposed SUMWE method can be implemented efficiently by using the matrix inversion lemma [20]. □

IV. Statistical Analysis

A. Asymptotic Properties of SUMWE

As the SUMWE estimator is a complicated nonlinear function of the received array data, its statistical behavior for "small" number of snapshots appears to be difficult to analyze like the other common DOA estimators [35]–[37]. In this section, we study the statistical properties of the proposed SUMWE method for large number of snapshots. First, we can easily obtain the following Lemma on the consistency of the SUMWE estimates.

Lemma: As the number of snapshots N tends to infinity, the estimates $\{\hat{\theta}_k\}$ obtained by minimizing the cost function $f(\theta)$ in (30) approach the true parameters $\{\theta_k\}$ w.p. 1.

Proof: This lemma can be readily established by adopting the proof of Lemma 1 in [18]. Clearly, $f(\theta)$ converges to the true cost function $f_o(\theta) \triangleq \bar{\boldsymbol{a}}^H(\theta)\boldsymbol{\Pi}_Q\bar{\boldsymbol{a}}(\theta)$ w.p. 1 and uniformly in θ when $N\to\infty$, and the minima of $f_o(\theta)$ are achieved if

and only if $\theta = \theta_k$. Thus, the estimates $\{\hat{\theta}_k\}$ approach the true parameters $\{\theta_k\}$ w.p. 1 as $N \to \infty$. ■

From this Lemma, we can find that the estimates $\{\hat{\theta}_k\}$ of the proposed SUMWE method are consistent. To evaluate the estimation accuracy achieved by the SUMWE method, we will investigate the asymptotic (for $N \gg 1$) MSE (or variance) of the estimation error.

As the signal vectors $\{\boldsymbol{y}_{fl}(n)\}$ or $\{\boldsymbol{y}_{bl}(n)\}$ of the first or last $L-1$ forward/backward subarrays are used in forming the correlation matrices $\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$, here, by defining the noisy signals comprised in these forward/backward subarrays as $\boldsymbol{z}_f(n) \triangleq [\boldsymbol{y}_{f1}^T(n), \boldsymbol{y}_{f2}^T(n), \ldots, \boldsymbol{y}_{fL-1}^T(n)]^T, \bar{\boldsymbol{z}}_f(n) \triangleq [\boldsymbol{y}_{f2}^T(n), \boldsymbol{y}_{f3}^T(n), \ldots, \boldsymbol{y}_{fL}^T(n)]^T, z_b(n) \triangleq [\boldsymbol{y}_{b1}^T(n), \boldsymbol{y}_{b2}^T(n), \ldots, \boldsymbol{y}_{b\,L-1}^T(n)]^T, \bar{\boldsymbol{z}}_b(n) \triangleq [\boldsymbol{y}_{b2}^T(n), \boldsymbol{y}_{b3}^T(n), \ldots, \boldsymbol{y}_{bL}^T(n)]^T$, we can introduce the auto/cross-covariance matrices $\boldsymbol{M}_{ff}, \bar{\boldsymbol{M}}_{ff}, \tilde{\boldsymbol{M}}_{ff}, \boldsymbol{M}_{bb}, \bar{\boldsymbol{M}}_{bb}, \tilde{\boldsymbol{M}}_{bb}, \boldsymbol{M}_{fb}, \bar{\boldsymbol{M}}_{fb}, \tilde{\boldsymbol{M}}_{fb}$, and $\breve{\boldsymbol{M}}_{fb}$ of these forward/backward subarrays as

$$\begin{aligned}
&\boldsymbol{M}_{ff} \triangleq E\{\boldsymbol{z}_f(n)\boldsymbol{z}_f^H(n)\}, \quad \bar{\boldsymbol{M}}_{ff} \triangleq E\{\bar{\boldsymbol{z}}_f(n)\bar{\boldsymbol{z}}_f^H(n)\}\\
&\tilde{\boldsymbol{M}}_{ff} \triangleq E\{\boldsymbol{z}_f(n)\bar{\boldsymbol{z}}_f^H(n)\}, \quad \boldsymbol{M}_{bb} \triangleq E\{\boldsymbol{z}_b(n)\boldsymbol{z}_b^H(n)\}\\
&\bar{\boldsymbol{M}}_{bb} \triangleq E\{\bar{\boldsymbol{z}}_b(n)\bar{\boldsymbol{z}}_b^H(n)\}, \quad \tilde{\boldsymbol{M}}_{bb} \triangleq E\{\boldsymbol{z}_b(n)\bar{\boldsymbol{z}}_b^H(n)\}\\
&\boldsymbol{M}_{fb} \triangleq E\{\boldsymbol{z}_f(n)\boldsymbol{z}_b^T(n)\}, \quad \bar{\boldsymbol{M}}_{fb} \triangleq E\{\boldsymbol{z}_f(n)\bar{\boldsymbol{z}}_b^T(n)\}\\
&\tilde{\boldsymbol{M}}_{fb} \triangleq E\{\bar{\boldsymbol{z}}_f(n)\boldsymbol{z}_b^T(n)\}, \quad \breve{\boldsymbol{M}}_{fb} \triangleq E\{\bar{\boldsymbol{z}}_f(n)\bar{\boldsymbol{z}}_b^T(n)\}.
\end{aligned}$$

Then, for a sufficiently larger number of snapshots N, we can obtain the expression of the asymptotic MSE (or variance) of the SUMWE estimates by the following theorem.

Theorem: For the estimates obtained by minimizing the function $f(\theta)$ in (30), the large-sample MSE (or variance) of the estimation errors $\{\hat{\theta}_k - \theta_k\}$ is given by

$$\begin{aligned}
\mathrm{MSE}(\hat{\theta}_k) &= \mathrm{var}(\hat{\theta}_k)\\
&= \frac{1}{2N\bar{\boldsymbol{H}}_{kk}^2}\mathrm{Re}\Big\{2\boldsymbol{g}_k^H\Big(r_{1M}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k1}^T\Big)\\
&\quad\times\boldsymbol{M}_{fb}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k3})\\
&\quad+r_{MM}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k1}^T\Big)\bar{\boldsymbol{M}}_{fb}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k4})\\
&\quad+r_{11}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k2}^T\Big)\tilde{\boldsymbol{M}}_{fb}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k3})\\
&\quad+r_{M1}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k2}^T\Big)\breve{\boldsymbol{M}}_{fb}\,(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k4})\Big)\boldsymbol{g}_k^*\\
&\quad+\boldsymbol{g}_k^H\Big(r_{MM}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k1}^T\Big)\boldsymbol{M}_{ff}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k1}^*)\\
&\quad+r_{11}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k2}^T\Big)\bar{\boldsymbol{M}}_{ff}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k2}^*)\\
&\quad+r_{11}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k3}^T\Big)\boldsymbol{M}_{bb}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k3}^*)\\
&\quad+r_{MM}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k4}^T\Big)\bar{\boldsymbol{M}}_{bb}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k4}^*)\Big)\boldsymbol{g}_k\\
&\quad+2\boldsymbol{g}_k^H\Big(r_{1M}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k1}^T\Big)\tilde{\boldsymbol{M}}_{ff}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k2}^*)\\
&\quad+\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k1}^T\Big)\boldsymbol{M}_w(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k3}^*)\\
&\quad+\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k2}^T\Big)\bar{\boldsymbol{M}}_w(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k4}^*)\\
&\quad+r_{1M}\Big(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k3}^T\Big)\tilde{\boldsymbol{M}}_{bb}(\boldsymbol{I}_{L-1}\otimes\boldsymbol{h}_{k4}^*)\Big)\boldsymbol{g}_k\Big\}
\end{aligned} \tag{34}$$

where $\boldsymbol{g}_k \triangleq \boldsymbol{\Pi}_Q\bar{\boldsymbol{d}}(\theta_k), \boldsymbol{h}_{k1} \triangleq \boldsymbol{\Phi}_{f1}^H\boldsymbol{\Psi}_1^{-1}\boldsymbol{a}_1(\theta_k), \boldsymbol{h}_{k2} \triangleq \bar{\boldsymbol{\Phi}}_{f1}^H\boldsymbol{\Psi}_1^{-1}\boldsymbol{a}_1(\theta_k), \boldsymbol{h}_{k3} \triangleq \boldsymbol{\Phi}_{b1}^H\boldsymbol{\Psi}_1^{-1}\boldsymbol{a}_1(\theta_k), \boldsymbol{h}_{k4} \triangleq \bar{\boldsymbol{\Phi}}_{b1}^H\boldsymbol{\Psi}_1^{-1}\cdot\boldsymbol{a}_1(\theta_k), \boldsymbol{\Psi}_1 \triangleq \boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H, \bar{\boldsymbol{H}} \triangleq \bar{\boldsymbol{D}}^H(\theta)\boldsymbol{\Pi}_Q\bar{\boldsymbol{D}}(\theta), \bar{\boldsymbol{D}}(\theta) \triangleq [\bar{\boldsymbol{d}}(\theta_1), \bar{\boldsymbol{d}}(\theta_2), \ldots, \bar{\boldsymbol{d}}(\theta_p)], \bar{\boldsymbol{d}}(\theta_k) \triangleq (d\bar{\boldsymbol{a}}(\theta))/(d\theta)|_{\theta=\theta_k}$, and the ikth elements of the $(L-1)p \times (L-1)p$ matrices $\boldsymbol{M}_w$ and $\bar{\boldsymbol{M}}_w$ are given, respectively, by

$$(\boldsymbol{M}_w)_{ik} = \begin{cases}\sigma^4, & \text{for } i=k=1\\ 0, & \text{others}\end{cases} \tag{35}$$

$$(\bar{\boldsymbol{M}}_w)_{ik} = \begin{cases}\sigma^4, & \text{for } i=k=(L-1)p\\ 0, & \text{others}\end{cases} \tag{36}$$

whereas $\mathrm{Re}\{\cdot\}$ denotes the real part of the bracketed quantity, $(\cdot)_{ik}$ represents the ikth element of the bracketed matrix, and $\otimes$ denotes the Kronecker product.

Proof: Because the estimate $\hat{\theta}_k$ is obtained by minimizing $f(\theta)$ and it is a consistent estimate, for a sufficiently large number of snapshots N, we can approximate the derivative of $f(\theta)$ using two terms in its Taylor series expansion about the true value θ_k as [10], [13], [18], [35]–[37]

$$0 = f'(\hat{\theta}_k) \approx f'(\theta_k) + f''(\theta_k)(\hat{\theta}_k - \theta_k) \tag{37}$$

where the second- and higher order terms in (37) can be neglected, and the first- and second-order derivatives of $f(\theta)$ with respect to the scalar variable θ are given by

$$f'(\theta) \triangleq \frac{df(\theta)}{d\theta} = 2\mathrm{Re}\{\bar{\boldsymbol{d}}^H(\theta)\boldsymbol{\Pi}_{\hat{Q}}\bar{\boldsymbol{a}}(\theta)\} \tag{38}$$

$$\begin{aligned}
f''(\theta) \triangleq \frac{df'(\theta)}{d\theta} &= 2\mathrm{Re}\{\tilde{\boldsymbol{d}}^H(\theta)\boldsymbol{\Pi}_{\hat{Q}}\bar{\boldsymbol{a}}(\theta)\\
&\quad+\bar{\boldsymbol{d}}^H(\theta)\boldsymbol{\Pi}_{\hat{Q}}\bar{\boldsymbol{d}}(\theta)\}
\end{aligned} \tag{39}$$

in which $\tilde{\boldsymbol{d}}(\theta) \triangleq d\bar{\boldsymbol{d}}(\theta)/d\theta$. From (37)–(39) and Appendix B, the first-order expression for the estimation error $\Delta\theta_k \triangleq \hat{\theta}_k - \theta_k$ can be obtained as

$$\Delta\theta_k \approx -\frac{f'(\theta_k)}{f''(\theta_k)} \approx -\frac{\mathrm{Re}\{\bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{\Pi}_{\hat{Q}}\bar{\boldsymbol{a}}(\theta_k)\}}{\bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{\Pi}_Q\bar{\boldsymbol{d}}(\theta_k)} \tag{40}$$

where the estimated orthogonal projector $\boldsymbol{\Pi}_{\hat{Q}}$ in the denominator of (40), i.e., (39), can be replaced with the true one $\boldsymbol{\Pi}_Q$ without affecting the asymptotic property of estimate $\hat{\theta}_k$ [10], [18], [35]–[37], [43].

Then, by using the fact that $\boldsymbol{Q}^H\bar{\boldsymbol{a}}(\theta_k) = \boldsymbol{0}_{(L-p-1)\times 1}$ and by substituting the approximation of $\boldsymbol{\Pi}_{\hat{Q}}$ presented in Appendix C into (40), the estimation error $\Delta\theta_k$ can be approximately given by

$$\begin{aligned}
\Delta\theta_k &\approx -\frac{\mathrm{Re}\{\bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\hat{\boldsymbol{Q}}^H\bar{\boldsymbol{a}}(\theta_k)\}}{\bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{\Pi}_Q\bar{\boldsymbol{d}}(\theta_k)}\\
&\triangleq -\frac{\mathrm{Re}\{\mu_k\}}{\bar{\boldsymbol{H}}_{kk}}
\end{aligned} \tag{41}$$

where

$$\mu_k \triangleq \bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\hat{\boldsymbol{Q}}^H\bar{\boldsymbol{a}}(\theta_k). \tag{42}$$

Consequently, because the estimate $\hat{\theta}_k$ is consistent, from (41), the MSE (or variance) of the estimation error $\Delta\theta_k$ is given by

$$\begin{aligned}
\mathrm{MSE}(\hat{\theta}_k) &\triangleq E\{(\Delta\theta_k)^2\} = \mathrm{var}(\hat{\theta}_k)\\
&\approx \frac{1}{2\bar{\boldsymbol{H}}_{kk}^2}\mathrm{Re}\left\{E\left\{\mu_k^2\right\} + E\{|\mu_k|^2\}\right\}
\end{aligned} \tag{43}$$

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

By letting $\boldsymbol{\Psi}_2 \triangleq \boldsymbol{\Phi}_2\boldsymbol{\Phi}_1^H$ and $\boldsymbol{\Psi} \triangleq [\boldsymbol{\Psi}_1^T, \boldsymbol{\Psi}_2^T]^T$, we can easily obtain a first-order approximation of $\hat{\boldsymbol{\Psi}}$ [18]

$$\begin{aligned}\hat{\boldsymbol{\Psi}} &= \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\Phi}}_1^H \approx \boldsymbol{\Phi}\boldsymbol{\Phi}_1^H + (\hat{\boldsymbol{\Phi}} - \boldsymbol{\Phi})\boldsymbol{\Phi}_1^H + \boldsymbol{\Phi}\left(\hat{\boldsymbol{\Phi}}_1^H - \boldsymbol{\Phi}_1^H\right) \\ &= \boldsymbol{\Phi}\left(\hat{\boldsymbol{\Phi}}_1^H - \boldsymbol{\Phi}_1^H\right) + \hat{\boldsymbol{\Phi}}\boldsymbol{\Phi}_1^H \end{aligned} \tag{44}$$

where $\boldsymbol{\Phi} \triangleq [\boldsymbol{\Phi}_1^T, \boldsymbol{\Phi}_2^T]^T = [\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b, \bar{\boldsymbol{\Phi}}_b]$. From (32), we get the following approximation as well:

$$\begin{aligned}\hat{\boldsymbol{P}}^H - \boldsymbol{P}^H &= \hat{\boldsymbol{\Psi}}_2\hat{\boldsymbol{\Psi}}_1^{-1} - \boldsymbol{P}^H = (\hat{\boldsymbol{\Psi}}_2 - \boldsymbol{P}^H\hat{\boldsymbol{\Psi}}_1)\hat{\boldsymbol{\Psi}}_1^{-1} \\ &= -\boldsymbol{Q}^H\hat{\boldsymbol{\Psi}}\hat{\boldsymbol{\Psi}}_1^{-1} = -\boldsymbol{Q}^H\hat{\boldsymbol{\Psi}}\boldsymbol{\Psi}_1^{-1} + O(1/N) \end{aligned} \tag{45}$$

where the term of order $O(1/N)$ is neglected [21]. Additionally, from (11) and (13)–(15), the estimated correlation matrices $\hat{\boldsymbol{\Phi}}_f, \hat{\bar{\boldsymbol{\Phi}}}_f, \hat{\boldsymbol{\Phi}}_b$, and $\hat{\bar{\boldsymbol{\Phi}}}_b$ can be expressed by

$$\hat{\boldsymbol{\Phi}}_f = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{Y}_f(n)y_M^*(n) \tag{46}$$

$$\hat{\bar{\boldsymbol{\Phi}}}_f = \frac{1}{N}\sum_{n=1}^{N}\bar{\boldsymbol{Y}}_f(n)y_1^*(n) \tag{47}$$

$$\hat{\boldsymbol{\Phi}}_b = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{Y}_b(n)y_1(n) \tag{48}$$

$$\hat{\bar{\boldsymbol{\Phi}}}_b = \frac{1}{N}\sum_{n=1}^{N}\bar{\boldsymbol{Y}}_b(n)y_M(n) \tag{49}$$

where $\boldsymbol{Y}_f(n) \triangleq [\boldsymbol{y}_{f1}(n), \boldsymbol{y}_{f2}(n), \ldots, \boldsymbol{y}_{f\,L-1}(n)]^T, \bar{\boldsymbol{Y}}_f(n) \triangleq [\boldsymbol{y}_{f2}(n), \boldsymbol{y}_{f3}(n), \ldots, \boldsymbol{y}_{fL}(n)]^T, \boldsymbol{Y}_b(n) \triangleq [\boldsymbol{y}_{b1}(n), \boldsymbol{y}_{b2}(n), \ldots, \boldsymbol{y}_{b\,L-1}(n)]^T$, and $\bar{\boldsymbol{Y}}_b(n) \triangleq [\boldsymbol{y}_{b2}(n), \boldsymbol{y}_{b3}(n), \ldots, \boldsymbol{y}_{bL}(n)]^T$.

It is noted that $\bar{\boldsymbol{a}}(\theta)$ can be partitioned as $\bar{\boldsymbol{a}}(\theta) = [\boldsymbol{a}_1^T(\theta), \boldsymbol{a}_2^T(\theta)]^T$. Then, by using the formula $\boldsymbol{Xc} = (\boldsymbol{I} \otimes \boldsymbol{c}^T)\mathrm{vec}(\boldsymbol{X}^T)$ for matrix $\boldsymbol{X}$ and vector $\boldsymbol{c}$ with compatible dimensions, where $\mathrm{vec}(\boldsymbol{X}^T)$ is a matrix operation stacking the columns of a matrix $\boldsymbol{X}^T$ one under the other to form a single column, from (22) and (44)–(49), we can approximate μ_k in (42) for a sufficiently large N as

$$\begin{aligned}\mu_k &= \bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}(\hat{\boldsymbol{Q}}^H - \boldsymbol{Q}^H)\bar{\boldsymbol{a}}(\theta_k) \\ &= \bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}(\hat{\boldsymbol{P}}^H - \boldsymbol{P}^H)\boldsymbol{a}_1(\theta_k) \\ &\approx -\bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{\Pi}_Q\left(\boldsymbol{\Phi}\left(\hat{\boldsymbol{\Phi}}_1^H - \boldsymbol{\Phi}_1^H\right)\right. \\ &\quad \left. + \hat{\boldsymbol{\Phi}}\boldsymbol{\Phi}_1^H\right)\boldsymbol{\Psi}_1^{-1}\boldsymbol{a}_1(\theta_k) \\ &= -\boldsymbol{g}_k^H\hat{\boldsymbol{\Phi}}\boldsymbol{\Phi}_1^H\boldsymbol{\Psi}_1^{-1}\boldsymbol{a}_1(\theta_k) \\ &\triangleq \mu_{k1} + \mu_{k2} + \mu_{k3} + \mu_{k4} \end{aligned} \tag{50}$$

where

$$\begin{aligned}\mu_{k1} &\triangleq -\boldsymbol{g}_k^H\hat{\boldsymbol{\Phi}}_f\boldsymbol{h}_{k1} = -\frac{1}{N}\boldsymbol{g}_k^H\left(\boldsymbol{I}_{L-1} \otimes \boldsymbol{h}_{k1}^T\right) \\ &\quad \times \left(\sum_{n=1}^{N}\mathrm{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)\right) \\ &= -\frac{1}{N}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^{N}\mathrm{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)\right) \end{aligned} \tag{51}$$

$$\begin{aligned}\mu_{k2} &\triangleq -\boldsymbol{g}_k^H\hat{\bar{\boldsymbol{\Phi}}}_f\boldsymbol{h}_{k2} \\ &= -\frac{1}{N}\boldsymbol{v}_{k2}^H\left(\sum_{n=1}^{N}\mathrm{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right)y_1^*(n)\right) \end{aligned} \tag{52}$$

$$\begin{aligned}\mu_{k3} &\triangleq -\boldsymbol{g}_k^H\hat{\boldsymbol{\Phi}}_b\boldsymbol{h}_{k3} \\ &= -\frac{1}{N}\boldsymbol{v}_{k3}^H\left(\sum_{n=1}^{N}\mathrm{vec}\left(\boldsymbol{Y}_b^T(n)\right)y_1(n)\right) \end{aligned} \tag{53}$$

$$\begin{aligned}\mu_{k4} &\triangleq -\boldsymbol{g}_k^H\hat{\bar{\boldsymbol{\Phi}}}_b\boldsymbol{h}_{k4} \\ &= -\frac{1}{N}\boldsymbol{v}_{k4}^H\left(\sum_{n=1}^{N}\mathrm{vec}\left(\bar{\boldsymbol{Y}}_b^T(n)\right)y_M(n)\right) \end{aligned} \tag{54}$$

and $\boldsymbol{v}_{ki} \triangleq (\boldsymbol{I}_{L-1} \otimes \boldsymbol{h}_{ki}^*)\boldsymbol{g}_k$. Hence, from (50), we obtain the terms $E\{\mu_k^2\}$ and $E\{|\mu_k|^2\}$ in (43) as

$$\begin{aligned}E\{\mu_k^2\} =& E\{\mu_{k1}^2+\mu_{k2}^2+\mu_{k3}^2+\mu_{k4}^2+2(\mu_{k1}\mu_{k2}+\mu_{k1}\mu_{k3} \\ &+ \mu_{k1}\mu_{k4}+\mu_{k2}\mu_{k3}+\mu_{k2}\mu_{k4}+\mu_{k3}\mu_{k4})\} \end{aligned} \tag{55}$$

$$\begin{aligned}E\{|\mu_k|^2\} =& E\{|\mu_{k1}|^2+|\mu_{k2}|^2+|\mu_{k3}|^2+|\mu_{k4}|^2+2(\mu_{k1}\mu_{k2}^* \\ &+ \mu_{k1}\mu_{k3}^*+\mu_{k1}\mu_{k4}^*+\mu_{k2}\mu_{k3}^* \\ &+ \mu_{k2}\mu_{k4}^*+\mu_{k3}\mu_{k4}^*)\}. \end{aligned} \tag{56}$$

By concatenating the results for the expectations $E\{\mu_{ki}\mu_{kq}\}$ and $E\{\mu_{ki}\mu_{kq}^*\}$ in Appendix D with (55) and (56) and by substituting (55) and (56) into (43), we readily get

$$\begin{aligned}\mathrm{MSE}(\hat{\theta}_k) =& \mathrm{var}(\hat{\theta}_k) \\ =& \frac{1}{2N\bar{\boldsymbol{H}}_{kk}^2}\mathrm{Re}\Big\{2\big(r_{1M}\boldsymbol{v}_{k1}^H\boldsymbol{M}_{fb}\boldsymbol{v}_{k3}^* \\ &+ r_{MM}\boldsymbol{v}_{k1}^H\bar{\boldsymbol{M}}_{fb}\boldsymbol{v}_{k4}^* \\ &+ r_{11}\boldsymbol{v}_{k2}^H\tilde{\boldsymbol{M}}_{fb}\boldsymbol{v}_{k3}^* + r_{M1}\boldsymbol{v}_{k2}^H\,\breve{\boldsymbol{M}}_{fb}\,\boldsymbol{v}_{k4}^*\big) \\ &+ r_{MM}\boldsymbol{v}_{k1}^H\boldsymbol{M}_{ff}\boldsymbol{v}_{k1} + r_{11}\boldsymbol{v}_{k2}^H\bar{\boldsymbol{M}}_{ff}\boldsymbol{v}_{k2} \\ &+ r_{11}\boldsymbol{v}_{k3}^H\boldsymbol{M}_{bb}\boldsymbol{v}_{k3} + r_{MM}\boldsymbol{v}_{k4}^H\bar{\boldsymbol{M}}_{bb}\boldsymbol{v}_{k4} \\ &+ 2\Big(r_{1M}\boldsymbol{v}_{k1}^H\tilde{\boldsymbol{M}}_{ff}\boldsymbol{v}_{k2} + \boldsymbol{v}_{k1}^H\boldsymbol{M}_w\boldsymbol{v}_{k3} \\ &+ \boldsymbol{v}_{k2}^H\bar{\boldsymbol{M}}_w\boldsymbol{v}_{k4} + r_{1M}\boldsymbol{v}_{k3}^H\tilde{\boldsymbol{M}}_{bb}\boldsymbol{v}_{k4}\Big)\Big\}. \end{aligned} \tag{57}$$

Therefore, the asymptotic (large-sample) MSE $\mathrm{MSE}(\hat{\theta}_k)$ (or variance $\mathrm{var}(\hat{\theta}_k)$) of the estimation error $\Delta\theta_k$ in (34) can be obtained from (57) immediately by using the definition $\boldsymbol{v}_{ki} \triangleq (\boldsymbol{I}_{L-1} \otimes \boldsymbol{h}_{ki}^*)\boldsymbol{g}_k$. ■

As shown in Remark B, the proposed SUMWE method is suitable for the case of partly coherent or incoherent signals. It is noteworthy that the expression of asymptotic MSE in (34) is valid, regardless of correlation between the incident signals as the proof does not require the signals $\{s_k(n)\}$ to be uncorrelated.

B. Analytic Study of Performance

The general expression of asymptotic MSE (variance) of estimation error in (34) for any (uncorrelated or correlated) incident signals is complicated. In order to gain insights into the proposed SUMWE method, we will specialize in the case of one signal for sake of conciseness and study the asymptotic error variance of the SUMWE estimator in detail. Furthermore, the

statistical performance of the SUMWE is compared quantitatively with that of the MUSIC [8] and BEWE [16] methods and with the stochastic Cramér–Rao lower bound (CRB) [35].

In this case (i.e., $p = 1$ and $L = M$), we can readily have

$$\bar{\boldsymbol{A}}(\theta) = \bar{\boldsymbol{a}}(\theta_1) = [1, e^{j\omega_0\tau(\theta_1)}, \ldots, e^{j\omega_0(M-2)\tau(\theta_1)}]^T \quad (58)$$

$$\begin{aligned}\bar{\boldsymbol{D}}(\theta) &= \bar{\boldsymbol{d}}(\theta_1) \\ &= j\omega[0, e^{j\omega_0\tau(\theta_1)}, 2e^{j2\omega_0\tau(\theta_1)}, \ldots \\ &\quad (M-2)e^{j\omega_0(M-2)\tau(\theta_1)}]^T \end{aligned} \quad (59)$$

$$r_{ik} = r_s e^{j\omega_0(i-k)\tau(\theta_1)} + \sigma^2\delta_{i,k} \quad (60)$$

and $a_1(\theta_1) = 1$, where $\omega \triangleq \omega_0(d/c)\cos\theta_1$. Here, we assume that $M > 4$ and set SNR $\triangleq r_s/\sigma^2$. To avoid a complication of notation, $\bar{\boldsymbol{a}}$ and $\bar{\boldsymbol{d}}$ are used as the brief notation for $\bar{\boldsymbol{a}}(\theta_1)$ and $\bar{\boldsymbol{d}}(\theta_1)$ in the following. From (58)–(60) and Remarks B and C, we can obtain (e.g., see [36])

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{a}} = M - 1 \quad (61)$$

$$\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}} = -j\omega\sum_{k=1}^{M-2} k = -j\omega\frac{(M-1)(M-2)}{2} \quad (62)$$

$$\bar{\boldsymbol{d}}^H\bar{\boldsymbol{d}} = \omega^2\sum_{k=1}^{M-2} k^2 = \omega^2\frac{(M-1)(M-2)(2M-3)}{6} \quad (63)$$

$$\begin{aligned}\bar{H} &= \bar{\boldsymbol{d}}^H\boldsymbol{\Pi}_Q\bar{\boldsymbol{d}} = \bar{\boldsymbol{d}}^H(\boldsymbol{I}_{L-1} - \bar{\boldsymbol{a}}(\bar{\boldsymbol{a}}^H\bar{\boldsymbol{a}})^{-1}\bar{\boldsymbol{a}}^H)\bar{\boldsymbol{d}} \\ &= \omega^2\frac{M(M-1)(M-2)}{12}\end{aligned} \quad (64)$$

$$\boldsymbol{\Phi}_f = \boldsymbol{\Phi}_b = r_s e^{-j\omega_0(M-1)\tau(\theta_1)}\bar{\boldsymbol{a}} \quad (65)$$

$$\bar{\boldsymbol{\Phi}}_f = \bar{\boldsymbol{\Phi}}_b = r_s e^{j\omega_0\tau(\theta_1)}\bar{\boldsymbol{a}}. \quad (66)$$

Furthermore, after some calculations, we get

$$h_{11} = h_{13} = \frac{1}{4r_s}e^{j\omega_0(M-1)\tau(\theta_1)} \quad (67)$$

$$h_{12} = h_{14} = \frac{1}{4r_s}e^{-j\omega_0\tau(\theta_1)} \quad (68)$$

$$\boldsymbol{M}_{ff} = \bar{\boldsymbol{M}}_{ff} = \boldsymbol{M}_{bb} = \bar{\boldsymbol{M}}_{bb} = r_s\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H + \sigma^2\boldsymbol{I}_{M-1} \quad (69)$$

$$\tilde{\boldsymbol{M}}_{ff} = \tilde{\boldsymbol{M}}_{bb} = r_s e^{-j\omega_0\tau(\theta_1)}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H + \sigma^2\bar{\boldsymbol{I}}_{M-1}^{(-1)} \quad (70)$$

$$\bar{\boldsymbol{M}}_{fb} = \tilde{\boldsymbol{M}}_{fb} = r_s e^{-j\omega_0(M-2)\tau(\theta_1)}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^T + \sigma^2\boldsymbol{J}_{M-1} \quad (71)$$

$$\boldsymbol{M}_{fb} = r_s e^{-j\omega_0(M-1)\tau(\theta_1)}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^T + \sigma^2\bar{\boldsymbol{J}}_{M-1}^{(-1)} \quad (72)$$

$$\breve{\boldsymbol{M}}_{fb} = r_s e^{-j\omega_0(M-3)\tau(\theta_1)}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^T + \sigma^2\bar{\boldsymbol{J}}_{M-1}^{(+1)} \quad (73)$$

where $\bar{\boldsymbol{I}}_m^{(-q)}$ denotes an $m\times m$ matrix with unity elements along the qth $(1 \le q \le m-1)$ lower diagonal off the major diagonal and zeros elsewhere, $\bar{\boldsymbol{J}}_m^{(\pm q)}$ represents an $m \times m$ matrix with unity elements along the qth $(1 \le q \le m-1)$ upper (for $+q$) or lower (for $-q$) diagonal off the major cross-diagonal and zeros elsewhere, and their elements are given by [51]

$$\left(\bar{\boldsymbol{I}}_m^{(-q)}\right)_{ik} = \begin{cases} 1, & \text{for } i-k = q \\ 0, & \text{others} \end{cases}$$

$$\left(\bar{\boldsymbol{J}}_m^{(+q)}\right)_{ik} = \begin{cases} 1, & \text{for } i+k = m-q+1 \\ 0, & \text{others} \end{cases}$$

and

$$\left(\bar{\boldsymbol{J}}_m^{(-q)}\right)_{ik} = \begin{cases} 1, & \text{for } i+k = m+q+1 \\ 0, & \text{others.} \end{cases}$$

Thus, by substituting (63)–(73) into (34) and by performing some calculations as shown in Appendix E, the asymptotic (large-sample) error variance $\text{var}_{\text{SUMWE}}(\hat{\theta}_1)$ of the estimate $\hat{\theta}_1$ obtained by the proposed SUMWE method is given by

$$\begin{aligned}\text{var}_{\text{SUMWE}}(\hat{\theta}_1) &= \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{M-2.5}{M(M-1)^2(M-2)} \\ &\quad\times\left(1 + \frac{(2M-3)(M+2)}{2M(2M-5)}\frac{1}{\text{SNR}}\right).\end{aligned} \quad (74)$$

The CRB is a lower bound on the estimation error variance for any unbiased estimator, and the stochastic CRB for the case of one incident signal is readily obtained by [35]

$$\text{CRB}(\hat{\theta}_1) = \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{1}{M(M^2-1)}\left(1 + \frac{1}{M}\frac{1}{\text{SNR}}\right). \quad (75)$$

Following the results presented in [36] and [18], the asymptotic error variances of estimate $\hat{\theta}_1$ obtained by the MUSIC and BEWE estimators are given by

$$\begin{aligned}&\text{var}_{\text{MUSIC}}(\hat{\theta}_1) \\ &\quad= \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{1}{M(M^2-1)}\left(1 + \frac{1}{M}\frac{1}{\text{SNR}}\right)\end{aligned} \quad (76)$$

$$\begin{aligned}&\text{var}_{\text{BEWE}}(\hat{\theta}_1) \\ &\quad= \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{1}{M(M-1)(M-2)}\left(1 + \frac{1}{\text{SNR}}\right).\end{aligned} \quad (77)$$

Note that the MUSIC is a large-sample realization of the ML method in the case considered here, and its error variance asymptotically achieves the stochastic CRB [35], [36] and that we choose $\tilde{y}(n) = y_M(n)$ for getting the BEWE variance $\text{var}_{\text{BEWE}}(\hat{\theta}_1)$ in (77) (see [16] and [18] for more details).

For a large value of the number of sensors M, from (75)–(77), we easily get

$$\text{var}_{\text{SUMWE}}(\hat{\theta}_1) \approx \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{1}{M^3}\left(1 + \frac{1}{2}\frac{1}{\text{SNR}}\right) \quad (78)$$

$$\begin{aligned}\text{CRB}(\hat{\theta}_1) &= \text{var}_{\text{MUSIC}}(\hat{\theta}_1) \\ &\approx \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{1}{M^3}\left(1 + \frac{1}{M}\frac{1}{\text{SNR}}\right)\end{aligned} \quad (79)$$

$$\text{var}_{\text{BEWE}}(\hat{\theta}_1) \approx \frac{1}{\omega^2 N}\frac{6}{\text{SNR}}\frac{1}{M^3}\left(1 + \frac{1}{\text{SNR}}\right). \quad (80)$$

Thus, we can obtain

$$\text{CRB}(\hat{\theta}_1) = \text{var}_{\text{MUSIC}}(\hat{\theta}_1) \le \text{var}_{\text{SUMWE}}(\hat{\theta}_1) \le \text{var}_{\text{BEWE}}(\hat{\theta}_1). \quad (81)$$

Obviously, the SUMWE error variance $\text{var}_{\text{SUMWE}}(\hat{\theta}_1)$ is bounded from below by $\text{var}_{\text{MUSIC}}(\hat{\theta}_1)$ (i.e., $\text{CRB}(\hat{\theta}_1)$) and from above by $\text{var}_{\text{BEWE}}(\hat{\theta}_1)$. In addition, the SUMWE variance $\text{var}_{\text{SUMWE}}(\hat{\theta}_1)$ will near $\text{var}_{\text{MUSIC}}(\hat{\theta}_1)$ when the number of sensors M is sufficiently large and the SNR is reasonably high.

Furthermore, to verify the above observations, we define the relative efficiency ratios between the error variances shown in (74), (76), and (77) as $\text{eff}_{\text{SUMWE}} \triangleq \text{var}_{\text{SUMWE}}(\hat{\theta}_1)/\text{var}_{\text{MUSIC}}(\hat{\theta}_1)$ and $\text{eff}_{\text{BEWE}} \triangleq \text{var}_{\text{BEWE}}(\hat{\theta}_1)/\text{var}_{\text{MUSIC}}(\hat{\theta}_1)$, which are functions of the SNR and number of sensors M. These relative efficiency ratios in terms of the SNR are shown in Fig. 1 for $M = 5, 10, 20$, and 40. Apparently $\text{eff}_{\text{SUMWE}}$ is smaller than eff_{BEWE}, and $\text{eff}_{\text{SUMWE}}$ is close to one for large values of SNR and M, which show that

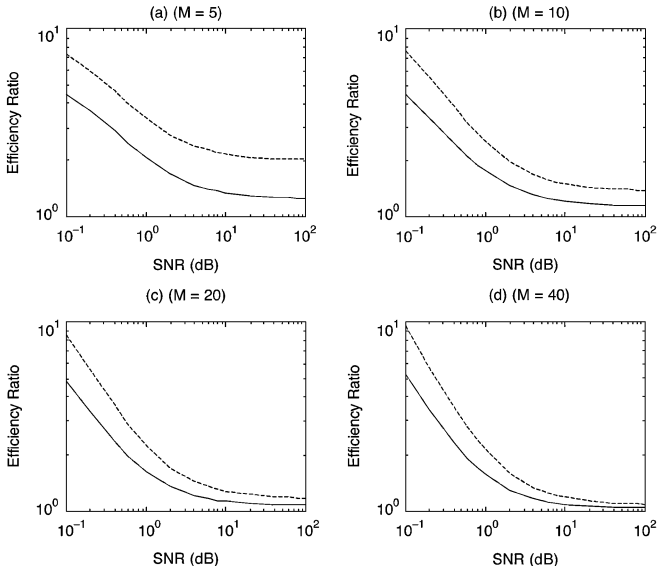

Fig. 1. Relative efficiency ratios of the SUMWE and BEWE estimators versus the SNR for several numbers of sensors. (a) $M = 5$. (b) $M = 10$. (c) $M = 20$. (d) $M = 40$ (dashed line: BEWE; solid line: SUMWE).

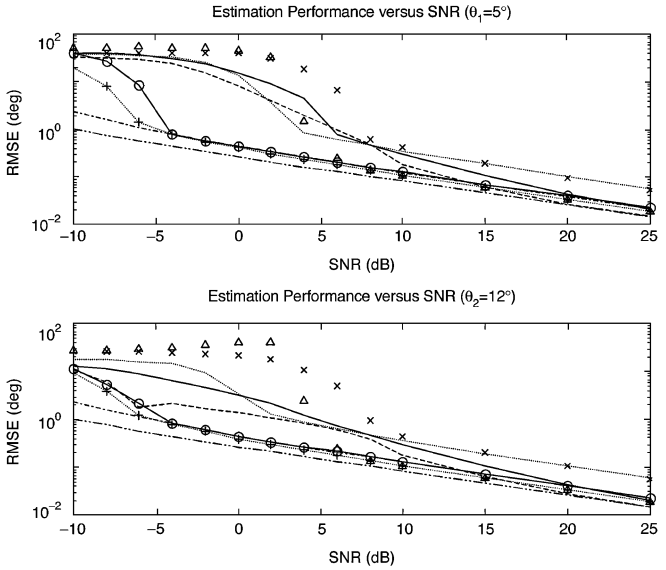

Fig. 2. RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the SNR (dotted line: SS-based root-MUSIC; dotted line with "+": FBSS-based root-MUSIC; "△": FBSS-based OPM; "x": BEWE; solid line: FBSS-based SWEDE; dashed line: WSF-E; solid line with "o": SUMWE; dash-dot line: theoretical RMSE of SUMWE; and dash-dots line: CRB) for Example 1 ($N = 128$ and $M = 10$).

the SUMWE is more accurate than the BEWE and that the SUMWE would be comparable to MUSIC when the SNR and the number of sensors are reasonably large.

Remark I: In the case of uncorrelated signals, the SUMWE may be less accurate than the MUSIC due to its small working array aperture (i.e., $M - p$) but may be more accurate than the BEWE, and it may be equivalent to the MUSIC for reasonably large values of the SNR and M. As shown in Remark B, the SUMWE method is suitable for the case of partly coherent or incoherent signals, and hence, it may be insensitive to the correlation between the signals. Furthermore, as shown in [31] and [32], the large number of subarrays and large working array aperture are expected to decorrelate the signal coherency and to improve the resolution of direction estimation when the incident signals are full correlated (i.e., coherent). In the SUMWE method, the subarray size is set as the number of incident signals (i.e., p) so that the maximum permissible number of subarrays (i.e., $L = M - p + 1$) and the maximum possible array aperture (i.e., $L - 1 = M - p$) are used. Therefore, the SUMWE method usually provides good estimation performance for highly correlated signals. □

V. Numerical Examples

We now evaluate the performance of the proposed SUMWE algorithm in estimating the DOAs of coherent and incoherent signals and confirm the derived statistical analysis through several numerical examples. The ULA with M sensors is separated by a half-wavelength, and two signals with equal power come from angles θ_1 and θ_2. The SNR is defined as the ratio of the power of the source signals to that of the additive noise at each sensor. In the simulations, the root approach version of the SUMWE algorithm is used to measure the performance of the proposed method precisely and appropriately. To compare the estimation performance of the SUMWE, the root-MUSIC [8], [25], SS-based root-MUSIC [24], [25], FBSS-based root-MUSIC [33], [25], BEWE [16], OPM [20], SWEDE (variant G) [21], and WSF-E (equivalently MODE without eigendecomposition) [22] are carried out, and the stochastic CRB [13], [35], [38], [45] is calculated. The FBSS preprocessing [33] is also performed for the OPM and SWEDE to combat the problem of coherency of the coherent signals, and we call these algorithms the FBSS-based OPM and FBSS-based SWEDE. For the BEWE, the variant for coherent signals is used in the first four examples, whereas the general variant for incoherent signals is used in Example 5 (see [16] for more details). Additionally, a two-step procedure of the WSF-E algorithm with the linear constraint is used (see [9] and [40] for more details), and the second step is not iterated to enable comparison of the computational burden. The results shown below are all based on 1000 independent trials.

Example 1—Performance versus SNR: In this example, we examine the performance of the proposed SUMWE algorithm against the SNR. The incident directions of two coherent signals are $\theta_1 = 5°$ and $\theta_2 = 12°$, and their SNR is varied from -10 to 25 dB. The number of sensors is $M = 10$, and the number of snapshots is $N = 128$. Additionally, the subarray size is set as $m = 0.6(M + 1) \approx 7$ [48] for the SS- and FBSS-based algorithms.

The empirical root MSEs (RMSEs) of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are shown in Fig. 2, where the theoretical RMSEs of the SUMWE in (34) and the stochastic CRBs are also plotted for comparison. Because the maximum possible number of subarrays (i.e., $L = M - p + 1$) and working array aperture (i.e., $L - 1 = M - p$) are exploited and the effect of additive noise is eliminated by appropriately choosing the used subarrays, the proposed SUMWE method generally outperforms the SS-based root-MUSIC with EVD and the methods without eigendecomposition such as the BEWE and FBSS-based SWEDE, and it is superior to the FBSS-based OPM without eigendecomposition

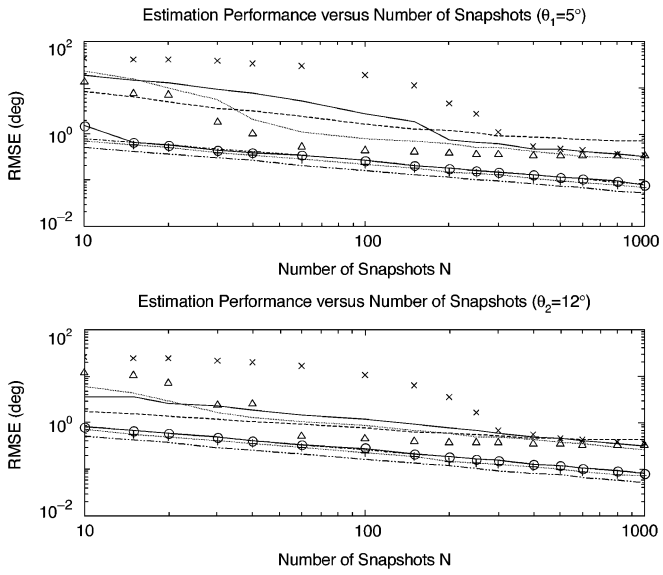

Fig. 3. RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the number of snapshots (dotted line: SS-based root-MUSIC; dotted line with "+": FBSS-based root-MUSIC; "△": FBSS-based OPM; "x": BEWE; solid line: FBSS-based SWEDE; dashed line: WSF-E; solid line with "o": SUMWE; dash-dot line: theoretical RMSE of SUMWE; and dash-dots line: CRB) for Example 2 (SNR $= 5$ dB and $M = 10$).

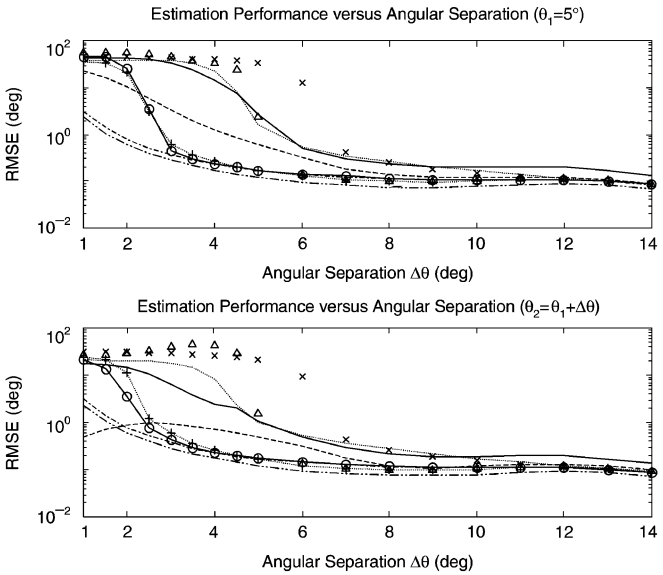

Fig. 4. RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the angular separation (dotted line: SS-based root-MUSIC; dotted line with "+": FBSS-based root-MUSIC; "△": FBSS-based OPM; "x": BEWE; solid line: FBSS-based SWEDE; dashed line: WSF-E; solid line with "o": SUMWE; dash-dot line: theoretical RMSE of SUMWE; and dash-dots line: CRB) for Example 3 (SNR $= 10$ dB, $N = 128$, and $M = 10$).

at low SNRs. Although the WSF-E performs better at high SNR, its performance degrades severely, and its RMSEs are larger than those of the SUMWE at low to moderate SNRs. We also note that the FBSS-based root-MUSIC provides the most accurate estimates than the other methods due to the use of EVD and the forward-backward subarray averaging. From Fig. 2, we can see that the empirical RMSEs of the SUMWE are very close to the theoretical ones (except at low SNR) and that the difference between the theoretical RMSEs and the CRBs is small. It is also shown that the theoretical and empirical RMSEs of the SUMWE decrease monotonically with the increasing SNR. In addition, the proposed SUMWE algorithm is computationally efficient, and the simulation shows that the ratio of the number of MATLAB flops required by the SS-based MUSIC and the WSF-E to that required by the SUMWE is about 7.147 and 18.677, respectively.

Example 2—Performance versus Number of Snapshots: Now, the performance of the SUMWE algorithm versus the number of snapshots is assessed. The simulation conditions are similar to those in Example 1, except that the SNR is set at 5 dB, and the number of snapshots is varied from $N = 10$ to $N = 1000$.

The empirical RMSEs of the estimated directions $\hat{\theta}_1$ and $\hat{\theta}_2$ are plotted in Fig. 3, and they are compared with the theoretical RMSEs of the SUMWE and the CRBs. We can see that the SUMWE is superior to the SS-based root-MUSIC, BEWE, FBSS-based OPM, FBSS-based SWEDE, and WSF-E, even for a small number of snapshots. As described in Example 1, the FBSS-based root-MUSIC performs better than the other methods because of the use of EVD and the forward-backward subarray averaging. Furthermore, we can find that the empirical RMSEs agree very well with the theoretical RMSEs derived in Section IV (their difference is almost indistinguishable), and the empirical and theoretical RMSEs of the SUMWE decrease monotonically with the number of snapshots.

Example 3—Performance versus Angular Separation: Here, the performance of the SUMWE algorithm is studied in terms of the angular separation between two coherent signals. In this example, two coherent signals impinge on the array along $\theta_1 = 5^\circ$ and $\theta_2 = \theta_1 + \Delta\theta$, where $\Delta\theta$ is varied from $\Delta\theta = 1^\circ$ to $\Delta\theta = 14^\circ$, and the other simulation parameters are the same as those in Example 1, except that the SNR is fixed at 10 dB.

Fig. 4 shows the empirical and theoretical RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ of the proposed SUMWE method against the angular separation $\Delta\theta$. The corresponding empirical RMSEs of the estimates obtained by the SS- and FBSS-based root-MUSIC methods, BEWE, FBSS-based OPM, FBSS-based SWEDE, and WSF-E and the CRBs are also plotted. As shown in Fig. 4, the SUMWE generally estimates the directions of closely spaced signals more accurately with a much smaller RMSE than the other methods, and the empirical RMSEs of the SUMWE are much closer to the theoretical ones derived in (34) for large angular separation. Because the two directions θ_1 and $\theta_2 = \theta_1 + \Delta\theta$ are much closer for small angular separation $\Delta\theta$, the WSF-E gives a few better estimates of one of the directions θ_1 and θ_2 for $\Delta\theta = 1^\circ, 1.5^\circ$, and 2°, whereas the other one is estimated inaccurately with a much larger RMSE. In the simulation, the empirical and theoretical RMSEs of the SUMWE do not decrease monotonically with the increasing angular separation as the CRB, which is in contrast to the cases with the increasing SNR and number of snapshots, as shown in Figs. 2 and 3.

Example 4—Performance versus Number of Sensors: We evaluate the performance of the SUMWE algorithm with respect to the number of sensors with the simulation parameters being identical to those in Example 1, except that the SNR is set at 10 dB, and the number of sensors is varied from $M = 5$

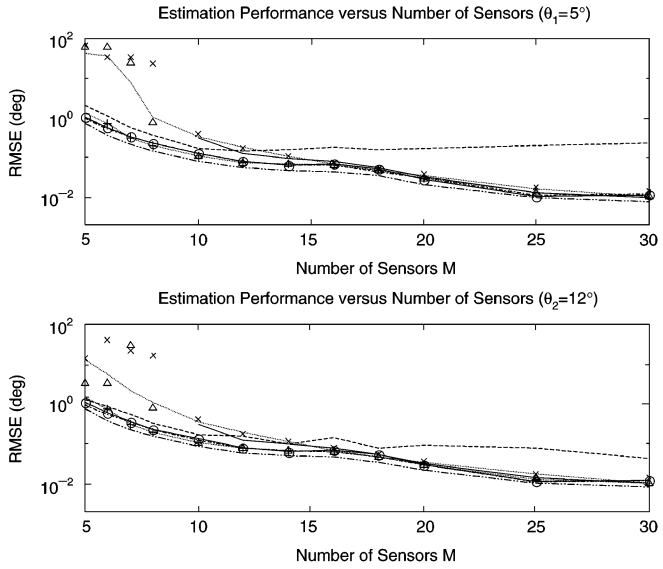

Fig. 5. RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the number of sensors (dotted line: SS-based root-MUSIC; dotted line with "+": FBSS-based root-MUSIC; "△": FBSS-based OPM; "x": BEWE; solid line: FBSS-based SWEDE; dashed line: WSF-E; solid line with "o": SUMWE; dash-dot line: theoretical RMSE of SUMWE; and dash-dots line: CRB) for Example 4 (SNR $=$ 10 dB and $N = 128$).

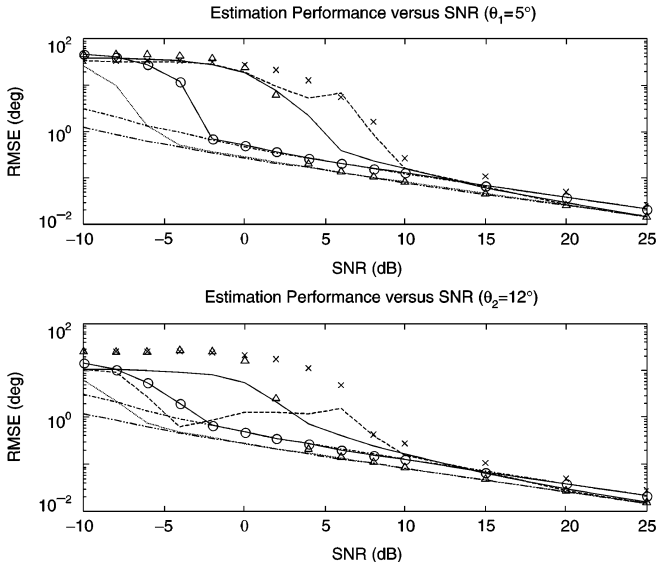

Fig. 6. RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the SNR in the case of uncorrelated signals (dotted line: root-MUSIC; "△": OPM; "x": BEWE; solid line: SWEDE; dashed line: WSF-E; solid line with "o": SUMWE; dash-dot line: theoretical RMSE of SUMWE; and dash-dots line: CRB) for Example 5 ($N = 128$ and $M = 10$).

to $M = 30$, where the identifiability condition $M > 2p$ is satisfied. The subarray size for the SS- and FBSS-based methods is chosen as $m = \text{round}(0.6(M+1))$, where round$(\cdot)$ denotes the round-off operation, and the FBSS-based SWEDE is performed for $M \geq 10$ (i.e., $m > 3p$) to ensure its identifiability.

The empirical RMSEs of arrival angle estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the number of sensors are shown in Fig. 5, where the theoretical RMSEs of the SUMWE and the CRBs are also plotted for reference. In general, the proposed SUMWE algorithm performs better than the SS-based root-MUSIC, BEWE, FBSS-based OPM, FBSS-based SWEDE, and WSF-E, and it has good accuracy like the FBSS-based root-MUSIC. Furthermore, the empirical RMSEs of the SUMWE nearly coincide with the theoretical RMSEs for a large number of sensors, and they do not decrease monotonically with the increasing number of sensors.

Example 5—Performance versus Correlation Between Signals: In the previous examples, the performance of the SUMWE method in estimating the directions of coherent signals is tested. Here, we verify its estimation performance in terms of the correlation between the incident signals, where the correlation factor is denoted by ρ.

First, we consider the estimation performance of the SUMWE algorithm with respect to the SNR when two incident signals are uncorrelated (i.e., $\rho = 0$), where the simulation conditions are similar to that of Example 1. The RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ obtained by the SUMWE method are depicted in Fig. 6, where the results of the root-MUSIC, BEWE, OPM, SWEDE, and WSF-E are plotted as well. Except for the fact that the SUMWE is worse than the root-MUSIC due to the smaller array aperture (i.e., $L - 1 = M - p$) and the avoidance of eigendecomposition, it usually performs better than the BEWE and the OPM, SWEDE, and WSF-E at relatively low SNRs. Furthermore, the difference between the CRBs and the theoretical

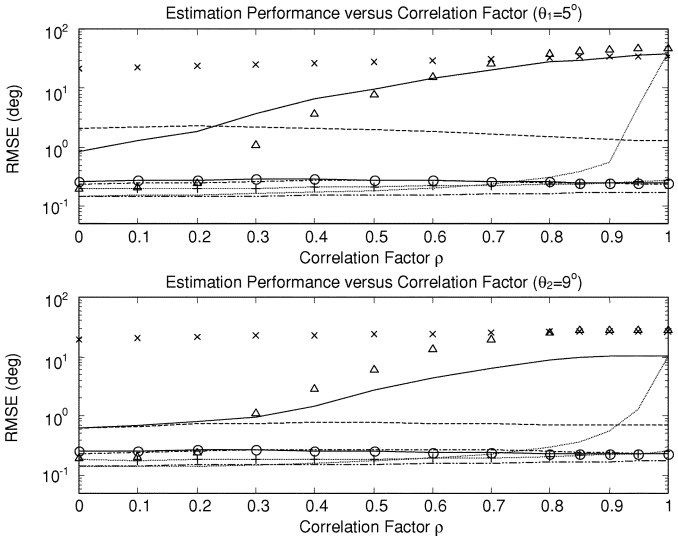

Fig. 7. RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ versus the correlation factor (dotted line: root-MUSIC; dotted line with "+": FBSS-based root-MUSIC; "△": OPM; "x": BEWE; solid line: SWEDE; dashed line: WSF-E; and solid line with "o": SUMWE; dash-dot line: theoretical RMSE of SUMWE; dash-dots line: CRB) for Example 5 (SNR $=$ 10 dB, $N = 128$, and $M = 10$).

RMSEs derived in (34) is small, and there is a very close agreement between the theoretical and empirical RMSEs of the proposed SUMWE method.

Then, we study the performance of the SUMWE in terms of the correlation factor ρ in resolving the closely spaced signals, where two signals arrive from $\theta_1 = 5°$ and $\theta_2 = 9°$, and the signal $s_2(n)$ is a superposition of two uncorrelated signals $s_1(n)$ and $s_{2o}(n)$ with equal power given by [5]

$$s_2(n) = \rho^* s_1(n) + \sqrt{1 - |\rho|^2}\, s_{2o}(n) \tag{82}$$

in which the magnitude of the correlation factor ρ is varied between 0 and 1 (the phase of the correlation factor is assumed to

be zero) for simplicity. The other parameters for simulation are similar to those in Example 1, except that the SNR is fixed at 10 dB.

Fig. 7 shows the empirical RMSEs of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ obtained by the SUMWE, root-MUSIC, FBSS-based root-MUSIC, BEWE, OPM, SWEDE, and WSF-E methods. As shown in Fig. 7, the root-MUSIC has an excellent performance for uncorrelated and weakly correlated signals, but its performance degrades significantly with the increase of correlation between the incident signals. However, the SUMWE is generally superior to the BEWE, OPM, SWEDE, and WSF-E for both uncorrelated and correlated signals and performs better than the root-MUSIC for highly correlated signals. Moreover, we can see that the SUMWE possesses a remarkable insensitivity to the correlation between the incident signals.

VI. Conclusion

In many practical applications of array processing, a computationally simple direction estimation method with good statistical performance is quite attractive. In this paper, we proposed a new computationally efficient subspace-based method called SUMWE that estimates the directions of narrowband signals impinging on a ULA by exploiting the array geometry and its shift invariance property. The proposed SUMWE method does not require the computationally cumbersome eigendecomposition and the evaluation of all correlations of the array data, and the effect of additive noise is eliminated. Furthermore, the SUMWE algorithm can be extended to the spatially correlated noise by appropriately choosing the used subarrays (i.e., cross-correlations of array data). The statistical analysis of the SUMWE was studied, and the explicit expression of asymptotic MSE (or variance) of the estimation error was derived. The performance advantage of the SUMWE method stems from the fact that the maximum permissible number of subarrays and working array aperture are exploited to decorrelate the signal coherency and to improve the resolution of direction estimation by setting the subarray size as the number of incident signals. The estimation performance of the SUMWE was demonstrated, and the theoretical analysis was substantiated through numerical examples. Although the SUMWE method performs slightly worse than the subspace-based methods with eigendecomposition such as the (FBSS-based) root-MUSIC in general, it mostly outperforms the subspace-based methods without eigendecomposition such as the BEWE, OPM, SWEDE, and WSF-E, and the simulation results showed that the SUMWE algorithm has the advantages of reduced computational load and superior estimation performance in resolving closely spaced correlated and uncorrelated signals with a short length of data and at low SNR.

Appendix

A. Proof of Linear Operator in (29)

It follows from Assumption A1) that the Vandermonde matrix $\boldsymbol{A}_1(\theta)$ has full rank and that the inverse matrix $\boldsymbol{A}_1^{-1}(\theta)$ exists; thus, we can readily verify that the first equality $\boldsymbol{P} = \boldsymbol{A}_1^{-H}(\theta)\boldsymbol{A}_2^H(\theta)$ satisfies (28).

Next, we consider the second equality in (29). Under the assumption that $\beta_k \neq 0$, from (24)–(28), some simple manipulations give

$$\boldsymbol{\Phi}_1^H = [\boldsymbol{\Phi}_{f1}, \bar{\boldsymbol{\Phi}}_{f1}, \boldsymbol{\Phi}_{b1}, \bar{\boldsymbol{\Phi}}_{b1}]^H = \rho_M^* r_s \boldsymbol{C}\boldsymbol{B}^* \boldsymbol{A}_1^H(\theta) \tag{A1}$$

$$\left[\boldsymbol{\Phi}_1^H, \boldsymbol{\Phi}_2^H\right] = [\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b, \bar{\boldsymbol{\Phi}}_b]^H = \rho_M^* r_s \boldsymbol{C}\boldsymbol{B}^* \bar{\boldsymbol{A}}^H(\theta) \tag{A2}$$

where

$$\boldsymbol{C} \triangleq \begin{bmatrix} \boldsymbol{A}_1^*(\theta) \\ (\rho_1^*/\rho_M^*)\boldsymbol{A}_1^*(\theta)\boldsymbol{D}^{-1} \\ (\rho_1/\rho_M^*)\boldsymbol{A}_1^*(\theta)\boldsymbol{D}^{M-1}\tilde{\boldsymbol{B}} \\ (\rho_M/\rho_M^*)\boldsymbol{A}_1^*(\theta)\boldsymbol{D}^{M-2}\tilde{\boldsymbol{B}} \end{bmatrix}$$

and $\tilde{\boldsymbol{B}} \triangleq \operatorname{diag}(\tilde{\beta}_1, \tilde{\beta}_2, \ldots, \tilde{\beta}_p)$ with $\tilde{\beta}_i \triangleq \beta_i/\beta_i^*$. Because $\boldsymbol{A}_1(\theta)$ and $\bar{\boldsymbol{A}}(\theta)$ are the submatrices of the Vandermonde matrix $\boldsymbol{A}(\theta)$, we can easily get the ranks of the matrices $\boldsymbol{\Phi}_1^H$ and $[\boldsymbol{\Phi}_1^H, \boldsymbol{\Phi}_2^H]$ as $\operatorname{rank}([\boldsymbol{\Phi}_1^H, \boldsymbol{\Phi}_2^H]) = \operatorname{rank}(\boldsymbol{\Phi}_1^H) = p$ [30]–[32]. Hence, it implies that the matrix $\boldsymbol{\Phi}_2^H$ belongs to the range space $\mathcal{R}(\boldsymbol{\Phi}_1^H)$ of the matrix $\boldsymbol{\Phi}_1^H$ (i.e., $\mathcal{R}(\boldsymbol{\Phi}_2^H) \subset \mathcal{R}(\boldsymbol{\Phi}_1^H)$), and the SVD of the matrix $\boldsymbol{\Phi}_1^H$ is given by [3], [49]

$$\boldsymbol{\Phi}_1^H = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^H = [\underbrace{\boldsymbol{U}_1}_{p}, \underbrace{\boldsymbol{U}_2}_{3p}] \begin{bmatrix} \boldsymbol{\Lambda}_1 \\ \boldsymbol{O}_{3p\times p} \end{bmatrix} \boldsymbol{V}^H = \boldsymbol{U}_1\boldsymbol{\Lambda}_1\boldsymbol{V}^H \tag{A3}$$

where $\boldsymbol{U}\boldsymbol{U}^H = \boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I}_{4p}, \boldsymbol{V}\boldsymbol{V}^H = \boldsymbol{V}^H\boldsymbol{V} = \boldsymbol{I}_p, \boldsymbol{\Lambda}_1 = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$, the $4p \times p$ matrix $\boldsymbol{U}_1$ is an orthonormal basis of $\mathcal{R}(\boldsymbol{\Phi}_1^H)$, and $\boldsymbol{U}_1\boldsymbol{U}_1^H$ is the orthogonal projector onto $\mathcal{R}(\boldsymbol{\Phi}_1^H)$, whereas the dimension of the null space $\mathcal{N}(\boldsymbol{\Phi}_1^H)$ of $\boldsymbol{\Phi}_1^H$ is given by $\dim(\mathcal{N}(\boldsymbol{\Phi}_1^H)) = 0$ (i.e., $\mathcal{N}(\boldsymbol{\Phi}_1^H) = \{0\}$) [3]. By some straightforward manipulations, from (A3), the second equality of (29) can be expressed as

$$\left(\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H\right)^{-1}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2^H = \boldsymbol{V}\boldsymbol{\Lambda}_1^{-1}\boldsymbol{U}_1^H\boldsymbol{\Phi}_2^H \triangleq \boldsymbol{P}_o. \tag{A4}$$

Then, we easily get

$$\boldsymbol{\Phi}_1^H\boldsymbol{P}_o = \boldsymbol{U}_1\boldsymbol{\Lambda}_1\boldsymbol{V}^H\boldsymbol{V}\boldsymbol{\Lambda}_1^{-1}\boldsymbol{U}_1^H\boldsymbol{\Phi}_2^H = \boldsymbol{U}_1\boldsymbol{U}_1^H\boldsymbol{\Phi}_2^H. \tag{A5}$$

Because $\mathcal{R}(\boldsymbol{\Phi}_2^H) \subset \mathcal{R}(\boldsymbol{\Phi}_1^H)$ and $\boldsymbol{U}_1\boldsymbol{U}_1^H$ is the orthogonal projector onto $\mathcal{R}(\boldsymbol{\Phi}_1^H)$, we can conclude that $\boldsymbol{U}_1\boldsymbol{U}_1^H\boldsymbol{\Phi}_2^H = \boldsymbol{\Phi}_2^H$ and, hence, that $\boldsymbol{P}_o$ is a solution of (28), [3]. Furthermore, let it be supposed that there is another solution $\boldsymbol{P}_g$ to (28), except $\boldsymbol{P}_o$, where $\boldsymbol{P}_g$ can be expressed by [3], [50]

$$\boldsymbol{P}_g = \boldsymbol{P}_o + \boldsymbol{\Omega}. \tag{A6}$$

Then, substituting (A6) into (28), we obtain

$$\boldsymbol{\Phi}_1^H\boldsymbol{P}_o + \boldsymbol{\Phi}_1^H\boldsymbol{\Omega} = \boldsymbol{\Phi}_2. \tag{A7}$$

Obviously, this implies that the $4p \times (L-p-1)$ matrix $\boldsymbol{\Omega}$ should satisfy that $\boldsymbol{\Phi}_1^H\boldsymbol{\Omega} = \boldsymbol{O}_{4p\times(L-p-1)}$, i.e., $\mathcal{R}(\boldsymbol{\Omega}) \subset \mathcal{N}(\boldsymbol{\Phi}_1^H)$. However, we can see that the matrix $\boldsymbol{\Omega}$ does not exist by taking into account that the fact $\mathcal{N}(\boldsymbol{\Phi}_1^H) = \{0\}$. Thus, $\boldsymbol{P} = \boldsymbol{P}_o = (\boldsymbol{\Phi}_1\boldsymbol{\Phi}_1^H)^{-1}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2^H$ is the unique solution to (28), and the proof is completed. ■

B. Approximation of Second-Order Derivative of Cost Function $f(\theta)$

As the cost function $f(\theta)$ converges to $f_o(\theta) = \bar{\boldsymbol{a}}^H(\theta)\boldsymbol{\Pi}_Q\bar{\boldsymbol{a}}(\theta)$ w.p. 1 and uniformly in θ when $N \to \infty$, the first-order perturbation expressions of $f'(\theta_k)$ in (38) and $f''(\theta_k)$ in (39) can be obtained as [43]

$$f'(\theta_k) = f'_o(\theta_k) + \Delta f_1 \tag{B1}$$

$$f''(\theta_k) = f''_o(\theta_k) + \Delta f_2 \tag{B2}$$

where $f'_o(\theta_k)$ and $f''_o(\theta_k)$ are the corresponding first- and second-order derivatives of $f_o(\theta)$, and Δf_1, and Δf_2 are the perturbation terms in the first- and second-order derivatives $f'(\theta_k)$ and $f''(\theta_k)$ of $f(\theta)$ with the estimated orthogonal projector $\boldsymbol{\Pi}_{\hat{Q}}$, respectively. Because $\boldsymbol{\Pi}_Q\bar{\boldsymbol{a}}(\theta_k) = \boldsymbol{0}_{(L-1)\times 1}$, we easily get that $f'_o(\theta_k) = 0$ and $f''_o(\theta_k) = 2\bar{\boldsymbol{d}}^H(\theta)\boldsymbol{\Pi}_Q\bar{\boldsymbol{d}}(\theta)$. Hence, from (37), (B1), and (B2), the estimation error $\Delta\theta_k \triangleq \hat{\theta}_k - \theta_k$ can be reexpressed as [43]

$$\begin{aligned}\Delta\theta_k &\approx -\frac{f'_o(\theta_k) + \Delta f_1}{f''_o(\theta_k) + \Delta f_2} \\ &= -\frac{\Delta f_1}{f''_o(\theta_k)}\left(1 + \frac{-\Delta f_2 / f''_o(\theta_k)}{1 - (-\Delta f_2/f''_o(\theta_k))}\right) \\ &= -\frac{\Delta f_1}{f''_o(\theta_k)}\left(1 - \frac{\Delta f_2}{f''_o(\theta_k)} + \left(\frac{\Delta f_2}{f''_o(\theta_k)}\right)^2 \right. \\ &\quad \left. - \left(\frac{\Delta f_2}{f''_o(\theta_k)}\right)^3 + \cdots\right) \\ &\approx -\frac{\Delta f_1}{f''_o(\theta_k)} = -\frac{f'(\theta_k)}{f''_o(\theta_k)}\end{aligned} \tag{B3}$$

where the terms neglected in the approximation tend to zero faster than $\Delta\theta_k$ as $N \to \infty$. It is shown that the asymptotic property of $\Delta\theta_k$ is not affected by replacing $f''(\theta_k)$ with $f''_o(\theta_k)$ in the denominator of $\Delta\theta_k$ in (B3). ■

C. Approximation of Estimated Orthogonal Projector

Here, we derive a first-order approximate of the orthogonal projector $\boldsymbol{\Pi}_{\hat{Q}}$ appearing in the numerator of $\Delta\theta_k$ in (40). Following the idea in [18], we can get the following approximations:

$$\begin{aligned}\boldsymbol{\Pi}_{\hat{Q}} &= \hat{\boldsymbol{Q}}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\hat{\boldsymbol{Q}}^H \approx \boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H + (\hat{\boldsymbol{Q}} - \boldsymbol{Q}) \\ &\quad \times (\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H + \boldsymbol{Q}((\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\hat{\boldsymbol{Q}}^H - (\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H) \\ &= (\hat{\boldsymbol{Q}} - \boldsymbol{Q})(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H + \boldsymbol{Q}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\hat{\boldsymbol{Q}}^H\end{aligned} \tag{C1}$$

where

$$\boldsymbol{Q}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\hat{\boldsymbol{Q}}^H \approx \boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}(\hat{\boldsymbol{Q}} - \boldsymbol{Q})^H + \boldsymbol{Q}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\boldsymbol{Q}^H \tag{C2}$$

in which

$$\begin{aligned}&\boldsymbol{Q}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}\boldsymbol{Q}^H \\ &\quad = \boldsymbol{Q}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}(\boldsymbol{I}_{L-p-1} - \hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1})\boldsymbol{Q}^H + \boldsymbol{\Pi}_Q \\ &\quad = \boldsymbol{Q}(\hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})^{-1}(\boldsymbol{Q}^H\boldsymbol{Q} - \hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H + \boldsymbol{\Pi}_Q \\ &\quad \approx \boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}(\boldsymbol{Q}^H\boldsymbol{Q} - \hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}})(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H + \boldsymbol{\Pi}_Q\end{aligned} \tag{C3}$$

and

$$\begin{aligned}&\boldsymbol{Q}^H\boldsymbol{Q} - \hat{\boldsymbol{Q}}^H\hat{\boldsymbol{Q}} \\ &\quad \approx \boldsymbol{Q}^H\boldsymbol{Q} - \boldsymbol{Q}^H(\hat{\boldsymbol{Q}} - \boldsymbol{Q}) - \hat{\boldsymbol{Q}}^H\boldsymbol{Q} \\ &\quad = 2\boldsymbol{Q}^H\boldsymbol{Q} - \boldsymbol{Q}^H\hat{\boldsymbol{Q}} - \hat{\boldsymbol{Q}}^H\boldsymbol{Q}.\end{aligned} \tag{C4}$$

By substituting (C2)–(C4) into (C1), and after some simple calculations, we can obtain the approximation of $\boldsymbol{\Pi}_{\hat{Q}}$ as

$$\begin{aligned}\boldsymbol{\Pi}_{\hat{Q}} \approx &(\boldsymbol{I}_{L-1} - \boldsymbol{\Pi}_Q)\hat{\boldsymbol{Q}}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\boldsymbol{Q}^H \\ &+ \boldsymbol{Q}(\boldsymbol{Q}^H\boldsymbol{Q})^{-1}\hat{\boldsymbol{Q}}^H(\boldsymbol{I}_{L-1} - \boldsymbol{\Pi}_Q) + \boldsymbol{\Pi}_Q.\end{aligned} \tag{C5}$$

■

D. Calculation of Expectations $E\{\mu_{ki}\mu_{kq}\}$ and $E\{\mu_{ki}\mu^*_{kq}\}$

Under the basic assumptions on the data model, and by using the formula for the expectation of the product of four complex Gaussian random matrices and vectors with zero-mean and compatible dimensions [44]

$$\begin{aligned}E\{\boldsymbol{A}bc^T\boldsymbol{D}\} = &E\{\boldsymbol{A}b\}E\{\boldsymbol{c}^T\boldsymbol{D}\} + E\{\boldsymbol{c}^T \otimes \boldsymbol{A}\} \\ &\cdot E\{\boldsymbol{D} \otimes \boldsymbol{b}\} + E\{\boldsymbol{A}E\{\boldsymbol{b}c^T\}\boldsymbol{D}\}\end{aligned} \tag{D1}$$

from (51) and (11), after some straightforward manipulations, we can obtain

$$\begin{aligned}&E\{\mu^2_{k1}\} \\ &= \frac{1}{N^2}\boldsymbol{v}^H_{k1}E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}^T_f(n)\right)y^*_M(n)y^*_M(t)\right. \\ &\quad \left.\times\ \text{vec}^T\left(\boldsymbol{Y}^T_f(t)\right)\right\}\boldsymbol{v}^*_{k1} \\ &= \frac{1}{N^2}\boldsymbol{v}^H_{k1}\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(E\left\{\text{vec}\left(\boldsymbol{Y}^T_f(n)\right)y^*_M(n)\right\}\right.\right. \\ &\quad \times E\left\{y^*_M(t)\text{vec}^T\left(\boldsymbol{Y}^T_f(t)\right)\right\} \\ &\quad + E\left\{y^*_M(t)\text{vec}\left(\boldsymbol{Y}^T_f(n)\right)\right\} \\ &\quad \times E\left\{\text{vec}^T\left(\boldsymbol{Y}^T_f(t)\right)y^*_M(n)\right\} + E\left\{\text{vec}\left(\boldsymbol{Y}^T_f(n)\right)\right. \\ &\quad \left.\left.\times\ E\{y^*_M(n)y^*_M(t)\}\text{vec}^T\left(\boldsymbol{Y}^T_f(t)\right)\right\}\right)\boldsymbol{v}^*_{k1} \\ &= \frac{1}{N^2}\boldsymbol{v}^H_{k1}\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\boldsymbol{\Phi}^T_f\right)\text{vec}^T\left(\boldsymbol{\Phi}^T_f\right)\right.\right. \\ &\quad \left.\left.+\ \text{vec}\left(\boldsymbol{\Phi}^T_f\right)\text{vec}^T\left(\boldsymbol{\Phi}^T_f\right)\delta_{n,t} + 0\right)\right)\boldsymbol{v}^*_{k1} \\ &= \left(1 + \frac{1}{N}\right)\boldsymbol{g}^H_k\left(\boldsymbol{I}_{L-1} \otimes \boldsymbol{h}^T_{k1}\right)\text{vec}\left(\boldsymbol{\Phi}^T_f\right)\text{vec}^T\left(\boldsymbol{\Phi}^T_f\right)\boldsymbol{v}^*_{k1} \\ &= \left(1 + \frac{1}{N}\right)\bar{\boldsymbol{d}}^H(\theta_k)\boldsymbol{\Pi}_Q\boldsymbol{\Phi}_f\boldsymbol{h}_{k1}\text{vec}^T\left(\boldsymbol{\Phi}^T_f\right)\boldsymbol{v}^*_{k1} = 0\end{aligned} \tag{D2}$$

where the fact that $\boldsymbol{Q}^H\bar{\boldsymbol{A}}(\theta) = \boldsymbol{O}_{(L-p-1)\times p}$ is used. By noting the fact that $\text{vec}(\boldsymbol{Y}^T_f(n)) = \boldsymbol{z}_f(n), \text{vec}(\bar{\boldsymbol{Y}}^T_f(n)) = \bar{\boldsymbol{z}}_f(n), \text{vec}(\boldsymbol{Y}^T_b(n)) = \boldsymbol{z}_b(n)$, and $\text{vec}(\bar{\boldsymbol{Y}}^T_b(n)) = \bar{\boldsymbol{z}}_b(n)$, in a

similar fashion to (D2), from (52)–(54), (13)–(15), and (22), we get

$$E\left\{\mu_{k2}^2\right\} = \frac{1}{N^2}\boldsymbol{v}_{k2}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right)y_1^*(n)y_1^*(t) \times \text{vec}^T\left(\bar{\boldsymbol{Y}}_f^T(t)\right)\right\}\boldsymbol{v}_{k2}^* = \left(1+\frac{1}{N}\right)\boldsymbol{v}_{k2}^H\text{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\text{vec}^T\left(\bar{\boldsymbol{\Phi}}_f^T\right)\boldsymbol{v}_{k2}^* = 0 \quad \text{(D3)}$$

$$E\left\{\mu_{k3}^2\right\} = \frac{1}{N^2}\boldsymbol{v}_{k3}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}_b^T(n)\right)y_1(n)y_1(t) \times \text{vec}^T\left(\boldsymbol{Y}_b^T(t)\right)\right\}\boldsymbol{v}_{k3}^* = \left(1+\frac{1}{N}\right)\boldsymbol{v}_{k3}^H\text{vec}\left(\boldsymbol{\Phi}_b^T\right)\text{vec}^T\left(\boldsymbol{\Phi}_b^T\right)\boldsymbol{v}_{k3}^* = 0 \quad \text{(D4)}$$

$$E\left\{\mu_{k4}^2\right\} = \frac{1}{N^2}\boldsymbol{v}_{k4}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\bar{\boldsymbol{Y}}_b^T(n)\right)y_M(n)y_M(t) \times \text{vec}^T\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}^* = \left(1+\frac{1}{N}\right)\boldsymbol{v}_{k4}^H\text{vec}\left(\bar{\boldsymbol{\Phi}}_b^T\right)\text{vec}^T\left(\bar{\boldsymbol{\Phi}}_b^T\right)\boldsymbol{v}_{k4}^* = 0 \quad \text{(D5)}$$

$$E\left\{\mu_{k1}\mu_{k3}\right\} = \frac{1}{N^2}\boldsymbol{v}_{k1}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)y_1(t) \times \text{vec}^T\left(\boldsymbol{Y}_b^T(t)\right)\right\}\boldsymbol{v}_{k3}^* = \frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\boldsymbol{\Phi}_f^T\right)\text{vec}^T\left(\boldsymbol{\Phi}_b^T\right)+0 + r_{1M}\boldsymbol{M}_{fb}\delta_{n,t}\right)\right)\boldsymbol{v}_{k3}^* = \frac{r_{1M}}{N}\boldsymbol{v}_{k1}^H\boldsymbol{M}_{fb}\boldsymbol{v}_{k3}^* \quad \text{(D6)}$$

$$E\left\{\mu_{k1}\mu_{k4}\right\} = \frac{1}{N^2}\boldsymbol{v}_{k1}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)y_M(t) \times \text{vec}^T\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}^* = \frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\boldsymbol{\Phi}_f^T\right)\text{vec}^T\left(\bar{\boldsymbol{\Phi}}_b^T\right)+0 + r_{MM}\bar{\boldsymbol{M}}_{fb}\delta_{n,t}\right)\right)\boldsymbol{v}_{k4}^* = \frac{r_{MM}}{N}\boldsymbol{v}_{k1}^H\bar{\boldsymbol{M}}_{fb}\boldsymbol{v}_{k4}^* \quad \text{(D7)}$$

$$E\{\mu_{k2}\mu_{k3}\} = \frac{1}{N^2}\boldsymbol{v}_{k2}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right)y_1^*(n)y_1(t) \times \text{vec}^T\left(\boldsymbol{Y}_b^T(t)\right)\right\}\boldsymbol{v}_{k3}^* = \frac{1}{N^2}\boldsymbol{v}_{k2}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\text{vec}^T\left(\boldsymbol{\Phi}_b^T\right)+0 + r_{11}\tilde{\boldsymbol{M}}_{fb}\delta_{n,t}\right)\right)\boldsymbol{v}_{k3}^* = \frac{r_{11}}{N}\boldsymbol{v}_{k2}^H\tilde{\boldsymbol{M}}_{fb}\boldsymbol{v}_{k3}^* \quad \text{(D8)}$$

$$E\{\mu_{k2}\mu_{k4}\} = \frac{1}{N^2}\boldsymbol{v}_{k2}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right)y_1^*(n)y_M(t) \times \text{vec}^T\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}^* = \frac{1}{N^2}\boldsymbol{v}_{k2}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\text{vec}^T\left(\bar{\boldsymbol{\Phi}}_b^T\right)+0 + r_{M1}\breve{\boldsymbol{M}}_{fb}\,\delta_{n,t}\right)\right)\boldsymbol{v}_{k4}^* = \frac{r_{M1}}{N}\boldsymbol{v}_{k2}^H\breve{\boldsymbol{M}}_{fb}\,\boldsymbol{v}_{k4}^*. \quad \text{(D9)}$$

Additionally, by defining $\boldsymbol{M}_f \triangleq E\{\boldsymbol{Y}_f(n)y_1^*(n)\}$ and $\bar{\boldsymbol{M}}_f \triangleq E\{\bar{\boldsymbol{Y}}_f(n)y_M^*(n)\}$, under the basic assumptions in Section II, and after some straightforward manipulations, we get $\boldsymbol{M}_f = \boldsymbol{M}_{xf} + \boldsymbol{M}_{wf}$ and $\bar{\boldsymbol{M}}_f = \bar{\boldsymbol{M}}_{xf} + \bar{\boldsymbol{M}}_{wf}$, where $\boldsymbol{M}_{xf} = \rho_1 r_s\bar{\boldsymbol{A}}(\theta)\boldsymbol{B}\boldsymbol{A}_1^T(\theta), \bar{\boldsymbol{M}}_{xf} = \rho_M r_s\bar{\boldsymbol{A}}(\theta)\boldsymbol{B}\boldsymbol{D}\boldsymbol{A}_1^T(\theta)$, and the ikth elements of the $(L-1)\times p$ matrices $\boldsymbol{M}_{wf}$ and $\bar{\boldsymbol{M}}_{wf}$ are given, respectively, by

$$(\boldsymbol{M}_{wf})_{ik} = \begin{cases}\sigma^2, & \text{for } i=k=1\\ 0, & \text{others}\end{cases} \quad \text{(D10)}$$

and

$$(\bar{\boldsymbol{M}}_{wf})_{ik} = \begin{cases}\sigma^2, & \text{for } i=L-1 \text{ and } k=p\\ 0, & \text{others.}\end{cases} \quad \text{(D11)}$$

Then, from (51), (52), (D10), and (D11), we obtain

$$E\{\mu_{k1}\mu_{k2}\} = \frac{1}{N^2}\boldsymbol{v}_{k1}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)y_1^*(t) \times \text{vec}^T\left(\bar{\boldsymbol{Y}}_f^T(t)\right)\right\}\boldsymbol{v}_{k2}^* = \frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(E\left\{\text{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)\right\}E\left\{y_1^*(t) \times \text{vec}^T\left(\bar{\boldsymbol{Y}}_f^T(t)\right)\right\} + E\left\{y_1^*(t)\text{vec}\left(\boldsymbol{Y}_f^T(n)\right)\right\} \times E\left\{\text{vec}^T\left(\bar{\boldsymbol{Y}}_f^T(t)\right)y_M^*(n)\right\} + E\left\{\text{vec}\left(\boldsymbol{Y}_f^T(n)\right)\right.\right.$$

$$\times\ E\left\{y_M^*(n)y_1^*(t)\right\}\mathrm{vec}^T\left(\bar{\boldsymbol{Y}}_f^T(t)\right)\Big\}\Big)\Big)\boldsymbol{v}_{k2}^*$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\boldsymbol{\Phi}_f^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{\Phi}}_f^T\right)\right.\right.$$
$$\left.+\ \mathrm{vec}\left(\boldsymbol{M}_f^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{M}}_f^T\right)\delta_{n,t}+0\right)\Big)\boldsymbol{v}_{k2}^*$$
$$=\frac{1}{N}\boldsymbol{v}_{k1}^H\mathrm{vec}\left(\boldsymbol{M}_f^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{M}}_f^T\right)\boldsymbol{v}_{k2}^*$$
$$=\frac{1}{N}\boldsymbol{v}_{k1}^H\left(\mathrm{vec}\left(\boldsymbol{M}_{xf}^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{M}}_f^T\right)+\mathrm{vec}\left(\boldsymbol{M}_{wf}^T\right)\right.$$
$$\left.\times\ \mathrm{vec}^T\left(\bar{\boldsymbol{M}}_f^T\right)\right)\boldsymbol{v}_{k2}^*=0 \tag{D12}$$

where the fact that $\boldsymbol{Q}^H\bar{\boldsymbol{A}}(\theta)=\boldsymbol{O}_{(L-p-1)\times p}$ is used implicitly. Further, by defining $\boldsymbol{M}_b\triangleq E\{\boldsymbol{Y}_b(n)y_M(n)\}$ and $\bar{\boldsymbol{M}}_b\triangleq E\{\bar{\boldsymbol{Y}}_b(n)y_1(n)\}$, it follows from some manipulations that $\boldsymbol{M}_b=\boldsymbol{M}_{xb}+\boldsymbol{M}_{wb}$ and $\bar{\boldsymbol{M}}_b=\bar{\boldsymbol{M}}_{xb}+\bar{\boldsymbol{M}}_{wb}$, where $\boldsymbol{M}_{xb}=\rho_M^*r_s\bar{\boldsymbol{A}}(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-1)}\boldsymbol{A}_1^T(\theta)$, $\bar{\boldsymbol{M}}_{xb}=\rho_1^*r_s\bar{\boldsymbol{A}}(\theta)\boldsymbol{B}^*\boldsymbol{D}^{-(M-2)}\boldsymbol{A}_1^T(\theta)$, $\boldsymbol{M}_{wb}=\boldsymbol{M}_{wf}$, and $\bar{\boldsymbol{M}}_{wb}=\bar{\boldsymbol{M}}_{wf}$. Hence, we can get

$$E\{\mu_{k3}\mu_{k4}\}=\frac{1}{N^2}\boldsymbol{v}_{k3}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\boldsymbol{Y}_b^T(n)\right)y_1(n)y_M(t)\right.$$
$$\left.\times\ \mathrm{vec}^T\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}^*$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k3}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\boldsymbol{\Phi}_b^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{\Phi}}_b^T\right)\right.\right.$$
$$\left.\left.+\mathrm{vec}\left(\boldsymbol{M}_b^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{M}}_b^T\right)\delta_{n,t}+0\right)\right)\boldsymbol{v}_{k4}^*$$
$$=\frac{1}{N}\boldsymbol{v}_{k3}^H\mathrm{vec}\left(\boldsymbol{M}_b^T\right)\mathrm{vec}^T\left(\bar{\boldsymbol{M}}_b^T\right)\boldsymbol{v}_{k4}^*=0. \tag{D13}$$

In much the same way as in evaluating the expectation $E\{\mu_{ki}\mu_{kq}\}$ shown in (D2)–(D13), by performing some manipulations, we can obtain the expectation $E\{\mu_{ki}\mu_{kq}^*\}$ in (56) as

$$E\{|\mu_{k1}|^2\}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)y_M(t)\right.$$
$$\left.\times\ \mathrm{vec}^H\left(\boldsymbol{Y}_f^T(t)\right)\right\}\boldsymbol{v}_{k1}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\boldsymbol{\Phi}_f^T\right)\mathrm{vec}^H\left(\boldsymbol{\Phi}_f^T\right)+0+r_{MM}\right.\right.$$
$$\left.\left.\times\ \boldsymbol{M}_{ff}\delta_{n,t}\right)\right)\boldsymbol{v}_{k1}=\frac{r_{MM}}{N}\boldsymbol{v}_{k1}^H\boldsymbol{M}_{ff}\boldsymbol{v}_{k1} \tag{D14}$$

$$E\{|\mu_{k2}|^2\}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k2}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right)y_1^*(n)y_1(t)\right.$$
$$\left.\times\ \mathrm{vec}^H\left(\bar{\boldsymbol{Y}}_f^T(t)\right)\right\}\boldsymbol{v}_{k2}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k2}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\mathrm{vec}^H\left(\bar{\boldsymbol{\Phi}}_f^T\right)+0+r_{11}\right.\right.$$
$$\left.\left.\times\ \bar{\boldsymbol{M}}_{ff}\delta_{n,t}\right)\right)\boldsymbol{v}_{k2}=\frac{r_{11}}{N}\boldsymbol{v}_{k2}^H\bar{\boldsymbol{M}}_{ff}\boldsymbol{v}_{k2} \tag{D15}$$
$$E\{|\mu_{k3}|^2\}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k3}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\boldsymbol{Y}_b^T(n)\right)y_1(n)y_1^*(t)\right.$$
$$\left.\times\ \mathrm{vec}^H\left(\boldsymbol{Y}_b^T(t)\right)\right\}\boldsymbol{v}_{k3}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k3}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\boldsymbol{\Phi}_b^T\right)\mathrm{vec}^H\left(\boldsymbol{\Phi}_b^T\right)+0+r_{11}\right.\right.$$
$$\left.\left.\times\ \boldsymbol{M}_{bb}\delta_{n,t}\right)\right)\boldsymbol{v}_{k3}=\frac{r_{11}}{N}\boldsymbol{v}_{k3}^H\boldsymbol{M}_{bb}\boldsymbol{v}_{k3} \tag{D16}$$
$$E\left\{|\mu_{k4}|^2\right\}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k4}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\bar{\boldsymbol{Y}}_b^T(n)\right)y_M(n)y_M^*(t)\right.$$
$$\left.\times\ \mathrm{vec}^H\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k4}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\bar{\boldsymbol{\Phi}}_b^T\right)\mathrm{vec}^H\left(\bar{\boldsymbol{\Phi}}_b^T\right)+0+r_{MM}\right.\right.$$
$$\left.\left.\times\ \bar{\boldsymbol{M}}_{bb}\delta_{n,t}\right)\right)\boldsymbol{v}_{k4}=\frac{r_{MM}}{N}\boldsymbol{v}_{k4}^H\bar{\boldsymbol{M}}_{bb}\boldsymbol{v}_{k4} \tag{D17}$$
$$E\{\mu_{k1}\mu_{k2}^*\}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)y_1(t)\right.$$
$$\left.\times\ \mathrm{vec}^H\left(\bar{\boldsymbol{Y}}_f^T(t)\right)\right\}\boldsymbol{v}_{k2}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\boldsymbol{\Phi}_f^T\right)\mathrm{vec}^H\left(\bar{\boldsymbol{\Phi}}_f^T\right)+0\right.\right.$$
$$\left.\left.+\ r_{1M}\tilde{\boldsymbol{M}}_{ff}\delta_{n,t}\right)\right)\boldsymbol{v}_{k2}=\frac{r_{1M}}{N}\boldsymbol{v}_{k1}^H\tilde{\boldsymbol{M}}_{ff}\boldsymbol{v}_{k2} \tag{D18}$$
$$E\{\mu_{k1}\mu_{k3}^*\}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^HE\left\{\sum_{n=1}^N\sum_{t=1}^N\mathrm{vec}\left(\boldsymbol{Y}_f^T(n)\right)y_M^*(n)y_1^*(t)\right.$$
$$\left.\times\ \mathrm{vec}^H\left(\boldsymbol{Y}_b^T(t)\right)\right\}\boldsymbol{v}_{k3}$$
$$=\frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^N\sum_{t=1}^N\left(\mathrm{vec}\left(\boldsymbol{\Phi}_f^T\right)\mathrm{vec}^H\left(\bar{\boldsymbol{\Phi}}_b^T\right)\right.\right.$$
$$\left.\left.+\ \mathrm{vec}(\boldsymbol{M}_f^T)\mathrm{vec}^H\left(\boldsymbol{M}_b^T\right)\delta_{n,t}+0\right)\right)\boldsymbol{v}_{k3}$$

$$= \frac{1}{N}\boldsymbol{v}_{k1}^H \text{vec}\left(\boldsymbol{M}_{wf}^T\right) \text{vec}^H\left(\boldsymbol{M}_{wb}^T\right)\boldsymbol{v}_{k3}$$
$$= \frac{1}{N}\boldsymbol{v}_{k1}^H \boldsymbol{M}_w \boldsymbol{v}_{k3} \tag{D19}$$

$$E\{\mu_{k1}\mu_{k4}^*\}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k1}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}_f^T(n)\right) y_M^*(n) y_M^*(t) \times \text{vec}^H\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k1}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\boldsymbol{\Phi}_f^T\right)\text{vec}^H\left(\bar{\boldsymbol{\Phi}}_b^T\right) + \text{vec}\left(\boldsymbol{\Phi}_f^T\right)\text{vec}^H\left(\bar{\boldsymbol{\Phi}}_b^T\right)\delta_{n,t} + 0\right)\right)\boldsymbol{v}_{k4} = 0 \tag{D20}$$

$$E\{\mu_{k2}\mu_{k3}^*\}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k2}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right) y_1^*(n) y_1^*(t) \times \text{vec}^H\left(\boldsymbol{Y}_b^T(t)\right)\right\}\boldsymbol{v}_{k3}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k2}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\text{vec}^H\left(\boldsymbol{\Phi}_b^T\right) + \text{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\text{vec}^H\left(\boldsymbol{\Phi}_b^T\right)\delta_{n,t} + 0\right)\right)\boldsymbol{v}_{k3} = 0 \tag{D21}$$

$$E\{\mu_{k2}\mu_{k4}^*\}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k2}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\bar{\boldsymbol{Y}}_f^T(n)\right) y_1^*(n) y_M^*(t) \times \text{vec}^H\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k2}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\bar{\boldsymbol{\Phi}}_f^T\right)\text{vec}^H\left(\bar{\boldsymbol{\Phi}}_b^T\right) + \text{vec}\left(\bar{\boldsymbol{M}}_f^T\right)\text{vec}^H\left(\bar{\boldsymbol{M}}_b^T\right)\delta_{n,t} + 0\right)\right)\boldsymbol{v}_{k4}$$
$$= \frac{1}{N}\boldsymbol{v}_{k2}^H \text{vec}\left(\bar{\boldsymbol{M}}_{wf}^T\right)\text{vec}^H\left(\bar{\boldsymbol{M}}_{wb}^T\right)\boldsymbol{v}_{k4}$$
$$= \frac{1}{N}\boldsymbol{v}_{k2}^H \bar{\boldsymbol{M}}_w \boldsymbol{v}_{k4} \tag{D22}$$

$$E\{\mu_{k3}\mu_{k4}^*\}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k3}^H E\left\{\sum_{n=1}^{N}\sum_{t=1}^{N}\text{vec}\left(\boldsymbol{Y}_b^T(n)\right) y_1(n) y_M^*(t) \times \text{vec}^H\left(\bar{\boldsymbol{Y}}_b^T(t)\right)\right\}\boldsymbol{v}_{k4}$$
$$= \frac{1}{N^2}\boldsymbol{v}_{k3}^H\left(\sum_{n=1}^{N}\sum_{t=1}^{N}\left(\text{vec}\left(\boldsymbol{\Phi}_b^T\right)\text{vec}^H\left(\bar{\boldsymbol{\Phi}}_b^T\right) + 0 + r_{1M}\tilde{\boldsymbol{M}}_{bb}\delta_{n,t}\right)\right)\boldsymbol{v}_{k4} = \frac{r_{1M}}{N}\boldsymbol{v}_{k3}^H\tilde{\boldsymbol{M}}_{bb}\boldsymbol{v}_{k4} \tag{D23}$$

where $\boldsymbol{M}_w \triangleq \text{vec}(\boldsymbol{M}_{wf}^T)\text{vec}^H(\boldsymbol{M}_{wb}^T)$, and $\bar{\boldsymbol{M}}_w \triangleq \text{vec}(\bar{\boldsymbol{M}}_{wf}^T)\text{vec}^H(\bar{\boldsymbol{M}}_{wb}^T)$, and their ikth elements are given in (35) and (36). ■

E. Calculation of Asymptotic Error Variance for a Single Signal

From (58) and (59), we can obtain, after some manipulations

$$\bar{\boldsymbol{d}}^H\boldsymbol{J}_{M-1}\bar{\boldsymbol{d}}^* = -\omega^2 e^{-j\omega_0(M-2)\tau(\theta_1)}\sum_{k=1}^{M-3}(M-2-k)k = -\omega^2 e^{-j\omega_0(M-2)\tau(\theta_1)}\frac{(M-1)(M-2)(M-3)}{6} \tag{E1}$$

$$\bar{\boldsymbol{a}}^H\boldsymbol{J}_{M-1}\bar{\boldsymbol{d}}^* = -j\omega e^{-j\omega_0(M-2)\tau(\theta_1)}\sum_{k=1}^{M-2}k = -j\omega e^{-j\omega_0(M-2)\tau(\theta_1)}\frac{(M-1)(M-2)}{2} \tag{E2}$$

$$\bar{\boldsymbol{a}}^H\boldsymbol{J}_{M-1}\bar{\boldsymbol{a}}^* = -e^{-j\omega_0(M-2)\tau(\theta_1)}(M-1). \tag{E3}$$

Hence, from (E1)–(E3) and (62), we get

$$\begin{aligned}\boldsymbol{g}_1^H\boldsymbol{J}_{M-1}\boldsymbol{g}_1^* &= \bar{\boldsymbol{d}}^H\boldsymbol{\Pi}_Q\boldsymbol{J}_{M-1}\boldsymbol{\Pi}_Q^*\bar{\boldsymbol{d}}^* \\ &= \bar{\boldsymbol{d}}^H\boldsymbol{J}_{M-1}\bar{\boldsymbol{d}}^* - \frac{2}{M-1}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\boldsymbol{J}_{M-1}\bar{\boldsymbol{d}}^* \\ &\quad + \frac{1}{(M-1)^2}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\boldsymbol{J}_{M-1}\bar{\boldsymbol{a}}^*\bar{\boldsymbol{a}}^T\bar{\boldsymbol{d}}^* \\ &= -\omega^2 e^{-j\omega_0(M-2)\tau(\theta_1)}\frac{M(M-1)(M-2)}{12}.\end{aligned} \tag{E4}$$

Similarly, some calculations give

$$\bar{\boldsymbol{d}}^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\bar{\boldsymbol{d}}^* = -\omega^2 e^{-j\omega_0(M-1)\tau(\theta_1)}\sum_{k=1}^{M-2}(M-1-k)k = -\omega^2 e^{-j\omega_0(M-1)\tau(\theta_1)}\frac{M(M-1)(M-2)}{6} \tag{E5}$$

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\bar{\boldsymbol{d}}^* = -j\omega e^{-j\omega_0(M-1)\tau(\theta_1)}\frac{(M-1)(M-2)}{2} \tag{E6}$$

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\bar{\boldsymbol{a}}^* = -e^{-j\omega_0(M-1)\tau(\theta_1)}(M-2) \tag{E7}$$

$$\bar{\boldsymbol{d}}^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\bar{\boldsymbol{d}}^* = -\omega^2 e^{-j\omega_0(M-3)\tau(\theta_1)}\sum_{k=1}^{M-4}(M-3-k)k = -\omega^2 e^{-j\omega_0(M-3)\tau(\theta_1)} \times \frac{(M-2)(M-3)(M-4)}{6} \tag{E8}$$

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\bar{\boldsymbol{d}}^* = -j\omega e^{-j\omega_0(M-3)\tau(\theta_1)}\frac{(M-2)(M-3)}{2} \tag{E9}$$

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\bar{\boldsymbol{a}}^* = -e^{-j\omega_0(M-3)\tau(\theta_1)}(M-2) \tag{E10}$$

$$\bar{\boldsymbol{d}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{d}} = \omega^2 e^{-j\omega_0\tau(\theta_1)}\sum_{k=1}^{M-3}(k+1)k = \omega^2 e^{-j\omega_0\tau(\theta_1)}\frac{(M-1)(M-2)(M-3)}{3} \tag{E11}$$

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{d}} = j\omega e^{-j\omega_0\tau(\theta_1)}\frac{(M-2)(M-3)}{2} \tag{E12}$$

$$\bar{\boldsymbol{d}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{a}} = -j\omega e^{-j\omega_0\tau(\theta_1)}\frac{(M-1)(M-2)}{2} \tag{E13}$$

$$\bar{\boldsymbol{a}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{a}}^* = e^{-j\omega_0\tau(\theta_1)}(M-2). \tag{E14}$$

Hence, we obtain

$$\begin{aligned}\boldsymbol{g}_1^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\boldsymbol{g}_1^* &= \bar{\boldsymbol{d}}^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\bar{\boldsymbol{d}}^* - \frac{2}{M-1}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\bar{\boldsymbol{d}}^* + \frac{1}{(M-1)^2}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(-1)}\bar{\boldsymbol{a}}^*\bar{\boldsymbol{a}}^T\bar{\boldsymbol{d}}^* \\ &= \omega^2 e^{-j\omega_0(M-1)\tau(\theta_1)}\frac{M(M-2)(M-4)}{12}\end{aligned} \tag{E15}$$

$$\begin{aligned}\boldsymbol{g}_1^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\boldsymbol{g}_1^* &= \bar{\boldsymbol{d}}^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\bar{\boldsymbol{d}}^* - \frac{2}{M-1}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\bar{\boldsymbol{d}}^* + \frac{1}{(M-1)^2}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{J}}_{M-1}^{(+1)}\bar{\boldsymbol{a}}^*\bar{\boldsymbol{a}}^T\bar{\boldsymbol{d}}^* \\ &= \omega^2 e^{-j\omega_0(M-3)\tau(\theta_1)}\frac{M(M-2)(M-4)}{12}\end{aligned} \tag{E16}$$

$$\begin{aligned}\boldsymbol{g}_1^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\boldsymbol{g}_1 &= \bar{\boldsymbol{d}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{d}} - \frac{1}{M-1}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{d}} - \frac{1}{M-1}\bar{\boldsymbol{d}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{d}} + \frac{1}{(M-1)^2} \\ &\quad\times\bar{\boldsymbol{d}}^H\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{I}}_{M-1}^{(-1)}\bar{\boldsymbol{a}}\bar{\boldsymbol{a}}^H\bar{\boldsymbol{d}} \\ &= \omega^2 e^{-j\omega_0\tau(\theta_1)}\frac{M(M-2)(M-4)}{12}.\end{aligned} \tag{E17}$$

Furthermore, it follows from some straightforward manipulations that

$$\boldsymbol{g}_1^H\boldsymbol{M}_w\boldsymbol{g}_1 = \boldsymbol{g}_1^H\bar{\boldsymbol{M}}_w\boldsymbol{g}_1 = \omega^2\sigma^4\frac{(M-2)^2}{4}. \tag{E18}$$

Therefore, by substituting (60) and (64)–(73) into (34) and using the results shown in (E4) and (E15)–(E18), after some manipulations, we can get the asymptotic error variance $\mathrm{var}_{\mathrm{SUMWE}}(\hat{\theta}_1)$ (or MSE $\mathrm{MSE}_{\mathrm{SUMWE}}(\hat{\theta}_1)$) as

$$\begin{aligned}&\mathrm{var}_{\mathrm{SUMWE}}(\hat{\theta}_1) \\ &= \frac{1}{2N\bar{H}^2}\frac{\sigma^2}{16r_s^2}\left\{2\omega^2\left(r_s\frac{M(M-2)(M-4)}{6} + (r_s+\sigma^2)\frac{M(M-1)(M-2)}{6}\right) + 4(r_s+\sigma^2)\bar{H} + 2\omega^2\left(r_s\frac{M(M-2)(M-4)}{6} + \sigma^2\frac{(M-2)^2}{2}\right)\right\} \\ &= \frac{1}{8N\bar{H}}\frac{\sigma^2}{r_s^2}\left\{\left(1+\frac{3(M-3)}{M-1}\right)r_s + \left(1+\frac{M^2+2M-6}{M(M-1)}\right)\sigma^2\right\} \\ &= \frac{1}{\omega^2 N}\frac{6}{\mathrm{SNR}}\frac{M-2.5}{M(M-1)^2(M-2)} \times\left(1+\frac{(2M-3)(M+2)}{2M(2M-5)}\frac{1}{\mathrm{SNR}}\right)\end{aligned} \tag{E19}$$

where the facts that $\boldsymbol{Q}^H\bar{\boldsymbol{a}} = \boldsymbol{0}_{(M-2)\times 1}$ and $\boldsymbol{\Pi}_Q = \boldsymbol{\Pi}_Q^H\boldsymbol{\Pi}_Q$ are used implicitly. ■

ACKNOWLEDGMENT

The authors are gratefully appreciative of the anonymous referees and the associate editor Prof. A. Zoubir for careful review, insightful comments, and valuable suggestions.

REFERENCES

[1] S. Haykin, Ed., *Array Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[2] S. Haykin, Ed., *Advances in Spectrum Analysis and Array Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1991, 1995. Volumes I–III.
[3] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.
[4] H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*. New York: Wiley, 2002.
[5] S. U. Pillai, *Array Signal Processing*. New York: Springer-Verlag, 1989.
[6] L. C. Godara, "Applications of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, pp. 1195–1245, Aug. 1997.
[7] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Mag.*, vol. 13, pp. 67–94, Apr. 1996.
[8] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276–280, Mar. 1986.
[9] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *Proc. Inst. Elect. Eng. F*, vol. 137, no. 1, pp. 19–26, 1990.
[10] ——, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1132–1143, July 1990.
[11] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 1110–1121, May 1991.
[12] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, pp. 2436–2449, Nov. 1991.
[13] 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: Springer-Verlag, 1993, pp. 99–151.
[14] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, pp. 1327–1343, Aug. 1990.
[15] R. D. DeGroat, E. M. Dowling, and D. A. Linebarger, "Subspace tracking," in *The Digital Signal Processing Handbook*, V. K. Madisetti and D. B. Williams, Eds. Boca Raton, FL: CRC, 1998.
[16] C.-C. Yeh, "Simple computation of projection matrix for bearing estimations," *Proc. Inst. Elect. Eng. F*, vol. 134, no. 2, pp. 146–150, 1987.
[17] ——, "Projection approach to bearing estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP–34, pp. 1347–1349, May 1986.
[18] P. Stoica and T. Söderström, "Statistical analysis of a subspace method for bearing estimation without eigendecomposition," *Proc. Inst. Elect. Eng. F*, vol. 139, no. 4, pp. 301–305, 1992.

[19] S. Marcos and M. Benidir, "On a high resolution array processing method nonbased on the eigenanalysis approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, Apr. 1990, pp. 2955–2958.
[20] S. Marcos, A. Marsal, and M. Benidir, "The propagator method for source bearing estimation," *Signal Process.*, vol. 42, no. 2, pp. 121–138, 1995.
[21] A. Eriksson, P. Stoica, and T. Söderström, "On-line subspace algorithms for tracking moving sources," *IEEE Trans. Signal Processing*, vol. 42, pp. 2319–2330, Sept. 1994.
[22] M. Kristensson, M. Jansson, and B. Ottersten, "Modified IQML and weighted subspace fitting without eigendecomposition," *Signal Process.*, vol. 79, no. 1, pp. 29–44, 1999.
[23] A. Swindlehurst, "Alternative algorithm for maximum likelihood DOA estimation and detection," *Proc. Inst. Elect. Eng. Radar, Sonar Navig.*, vol. 141, no. 6, pp. 293–299, 1994.
[24] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 806–811, Apr. 1985.
[25] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, vol. 1, Boston, MA, 1983, pp. 336–339.
[26] J. C. Liberti Jr. and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Upper Saddle River, NJ: Prentice-Hall, 1999.
[27] N. Yuen and B. Friedlander, "DOA estimation in multipath: An approach using fourth-order cumulants," *IEEE Trans. Signal Processing*, vol. 45, pp. 1253–1263, May 1997.
[28] E. Gonen, J. M. Mendel, and M. C. Dogen, "Applications of cumulants to array processing—-Part IV: Direction finding in coherent signals case," *IEEE Trans. Signal Processing*, vol. 45, pp. 2265–2276, Sept. 1997.
[29] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Trans. Signal Processing*, vol. 46, pp. 449–462, Feb. 1998.
[30] J. Xin and A. Sano, "Linear prediction approach to direction estimation of cyclostationary signals in multipath environment," *IEEE Trans. Signal Processing*, vol. 49, pp. 710–720, Apr. 2001.
[31] ——, "MSE-based regularization approach to direction estimation of coherent narrowband signals using linear prediction," *IEEE Trans. Signal Processing*, vol. 49, pp. 2481–2497, Nov. 2001.
[32] ——, "Direction estimation of coherent signals using spatial signature," *IEEE Signal Processing Lett.*, vol. 9, pp. 414–417, Dec. 2002.
[33] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signals identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 8–15, Jan. 1989.
[34] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1553–1560, Oct. 1988.
[35] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1783–1795, Oct. 1990.
[36] ——, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–741, May 1989.
[37] ——, "MUSIC, maximum likelihood, and Cramer-Rao bound: Further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 2140–2150, Dec. 1990.
[38] P. Stoica, B. Ottersten, M. Viberg, and R. L. Moses, "Maximum likelihood array processing for stochastic coherent sources," *IEEE Trans. Signal Processing*, vol. 44, pp. 96–105, Jan. 1996.
[39] J. Sheinvald, M. Wax, and A. J. Weiss, "On maximum-likelihood localization of coherent signals," *IEEE Trans. Signal Processing*, vol. 44, pp. 2475–2482, Oct. 1996.
[40] J. Li, P. Stoica, and Z.-S. Liu, "Comparative study of IQML and MODE direction-of-arrival estimators," *IEEE Trans. Signal Processing*, vol. 46, pp. 149–160, Jan. 1998.
[41] J. Xin and A. Sano, "Direction estimation of coherent signals without eigendecomposition and spatial smoothing," in *Proc. IFAC 12th Symp. Syst. Identification*, vol. 1, Santa Barbara, CA, June 2000, pp. 349–354.
[42] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 996–1000, July 1989.
[43] F. Li and R. J. Vaccaro, "Analysis of min-norm and MUSIC with arbitrary array geometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, pp. 976–985, June 1990.
[44] P. H. M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 867–870, Sept. 1988.
[45] B. Ottersten, M. Viberg, and T. Kailath, "Analysis of subspace fitting and ML techniques for parameter estimation from sensor array data," *IEEE Trans. Signal Processing*, vol. 40, pp. 590–600, Mar. 1992.
[46] N. L. Owsley, "Sonar array processing," in *Array Signal Processing*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[47] S. Haykin, "Radar array processing for angle of arrival estimation," in *Array Signal Processing*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[48] A. B. Gershman and V. T. Ermolaev, "Optimal subarray size for spatial smoothing," *IEEE Signal Processing Lett.*, vol. 2, pp. 28–30, Feb. 1995.
[49] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: John Hopkins Univ. Press, 1989.
[50] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.
[51] J. Xin and A. Sano, "Effect of subarray size on direction estimation of coherent cyclostationary signals based on forward-backward linear prediction," *IEICE Trans. Fundamentals*, vol. E85-A, no. 8, pp. 1807–1821, 2002.

Jingmin Xin (S'92–M'96) received the B.E. degree in information and control engineering from Xi'an Jiaotong University, Xi'an, China, in 1988 and the M.E. and Ph.D. degrees in electrical engineering from Keio University, Yokohama, 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, Xi'an. He was with Communications Research Laboratory, MPT of 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 Co., Ltd., Yokosuka, Japan, from 1999 to 2001. Since 2002, he has been with Fujitsu Laboratories Ltd., Yokosuka. His research interests are in the areas of signal processing, system identification, and their applications to mobile communication systems.

Akira Sano (M'89) received the B.E., M.E., and Ph.D. degrees in mathematical engineering and information physics from the University of Tokyo, Tokyo, Japan, in 1966, 1968, and 1971, respectively.

He joined the Department of Electrical Engineering, Keio University, Yokohama, Japan, in 1971, where he is currently a Professor with the Department of System Design Engineering. He was a Visiting Research Fellow at the University of Salford, Salford, U.K., from 1977 to 1978. Since 1995, he has been a Visiting Research Fellow with the Communication Research Laboratory, Ministry of Posts and Telecommunications of Japan. His current research interests are in 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. He is a coauthor of the textbook *State Variable Methods in Automatic Control* (New York: Wiley, 1988).

Dr. Sano received the Kelvin Premium from the Institute of Electrical Engineering in 1986. He 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 served as 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 has been Chair of the IFAC Technical Committee on Adaptive and Learning Systems since 2002.