Perturbation Analysis for Subspace Decomposition With Applications in Subspace-Based Algorithms

Zhengyuan Xu, *Senior Member, IEEE*

***Abstract*—Subspace decomposition has been exploited in different applications. Due to perturbations from various sources such as finite data samples and measurement noise, perturbations arise in subspaces. Therefore, some loss is introduced to performance of subspace-based algorithms. Although first-order perturbation results have been proposed in the literature and applied to various problems, up to second-order perturbation analysis can provide more accurate analytical results and is studied in this paper. Based on the orthogonality principle, perturbations of subspaces and singular values (or eigenvalues) are derived explicitly as functions of a perturbation in the objective matrix up to the second order, respectively, all in closed forms. It is shown that by keeping only the first-order terms, the derived results reduce to those from existing approaches. Examples to apply the proposed results to both matrix computation and subspace-based channel estimation are provided to verify our analysis.**

***Index Terms*—Channel estimation, second-order perturbation, subspace decomposition.**

I. Introduction

SUBSPACE methods have been employed in solving various statistical problems in array signal processing [13], [15], blind channel estimation [12] and code-division multiple access (CDMA) communications [4], [11], [19]. Due to various perturbation sources arising in different applications, such as finite data effect, modeling errors, and measurement noise, perturbations are introduced to ideal solutions. Analytical results are desired for either performance evaluation or comparison purposes.

Typically, subspace methods are based on subspace decomposition [eigenvalue decomposition (EVD) or singular value decomposition (SVD)] on some matrices. Then, the obtained subspaces will play important roles in subspace processing algorithms. Consider the following matrix

$$\hat{\boldsymbol{X}} = \boldsymbol{X} + \delta\boldsymbol{X} \tag{1}$$

with the number of rows smaller than or equal to the number of columns (fat or square matrix), where $\hat{\boldsymbol{X}}$ is a perturbed version[1] of $\boldsymbol{X}$ with perturbation $\delta\boldsymbol{X}$. If the matrix has more rows than columns (tall matrix), our discussion can be generalized to its Hermitian $\hat{\boldsymbol{X}}^H$ (complex conjugate transpose) first. Then, after Hermitian operation on all corresponding results, results for $\hat{\boldsymbol{X}}$ easily follow. In communications, $\boldsymbol{X}$ can be a data matrix or data covariance matrix depending on applications. Assume $\boldsymbol{X}$ has subspace decomposition

$$\boldsymbol{X} = \boldsymbol{U}_s\boldsymbol{\Lambda}_s\boldsymbol{V}_s^H + \boldsymbol{U}_n\boldsymbol{\Lambda}_n\boldsymbol{V}_n^H \tag{2}$$

where $\boldsymbol{U}_s$ and $\boldsymbol{V}_s$ span the column spaces of $\boldsymbol{X}$ and $\boldsymbol{X}^H$, respectively, whereas $\boldsymbol{U}_n$ and $\boldsymbol{V}_n$ span their orthogonal spaces, and $\boldsymbol{\Lambda}_s$ and $\boldsymbol{\Lambda}_n$ are corresponding singular values (or eigenvalues). Similarly, subspace decomposition on $\hat{\boldsymbol{X}}$ gives

$$\hat{\boldsymbol{X}} = \hat{\boldsymbol{U}}_s\hat{\boldsymbol{\Lambda}}_s\hat{\boldsymbol{V}}_s^H + \hat{\boldsymbol{U}}_n\hat{\boldsymbol{\Lambda}}_n\hat{\boldsymbol{V}}_n^H. \tag{3}$$

Due to $\delta\boldsymbol{X}$, all quantities on the right-hand side of (3) may differ from those on the right-hand side of (2). Detailed study on the first-order approximation has been performed in [16] from different perspectives. It has also appeared in performance analysis for direction-of-arrival (DOA) estimation [10], where the data-plus-noise model is considered. Theoretical approaches to first-order perturbations on eigenvalue problems are proposed in [8] and [14]. The finite data sample effect on statistics of eigenvectors of data covariance matrices is investigated in [7]. When the data covariance matrix is updated by a rank-one matrix, perturbations of its eigenvalues and eigenvectors can be found in [6]. Those results and corresponding methods have been successfully applied to analyze the noise effect on the channel estimators [3], [11], [17] and the multiuser detector [18].

It can be observed that most analyses are based on the first-order perturbation where all perturbed quantities are of the first order of $\delta\boldsymbol{X}$. Meanwhile, results in [10] are bound to a particular model, where $\boldsymbol{\Lambda}_n$ is constrained to be $\mathbf{0}$. Thus, they will encounter problems in serving as references for other applications. However, desires to obtain subspace approximation up to the second order of $\delta\boldsymbol{X}$ arise if the first-order perturbation does not provide sufficient accuracy in data analysis. Since EVD and SVD can both be described by (3), we will not differentiate these cases later.

The objective of this paper is to derive general expressions for approximations to those perturbed terms up to the second order of $\delta\boldsymbol{X}$. The proposed results are expected to be used as references for researchers and engineers in different fields, such as signal processing and communications. It is shown that if our results are truncated at the first order of $\delta\boldsymbol{X}$ and $\boldsymbol{\Lambda}_n$ is set to be $\mathbf{0}$, then they reduce to those in [10]. Besides analysis on a data matrix, in particular, we also apply our results to the case when a data covariance matrix is estimated by its sample average from finite data samples [7]. As expected, our perturbation results exhibit consistency with [7] when only the first-order terms are maintained. In order to verify our analytical results, we first apply them to matrix computation from mathematical point of

Manuscript received January 30, 2001; revised May 23, 2002. This work was supported in part by the National Science Foundation under Grant NSF-CCR 0207931. The associate editor coordinating the review of this paper and approving it for publication was Dr. Alex C. Kot.

The author is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA(e-mail: dxu@ee.ucr.edu).

Digital Object Identifier 10.1109/TSP.2002.804084.

[1]We use $\hat{(\cdot)}$ to represent perturbed quantities and $\delta(\cdot)$ to denote their perturbations.

view. Then, we employ them to derive an analytical solution to subspace-based channel estimation problem [4], [12]. Based on simulation results, it is observed that with the second-order approximation, more accuracy can be achieved compared with the first-order approximation only.

II. Perturbations in Subspace Decomposition

For simplicity, assume $\mathbf{\Lambda}_s$ is a square diagonal matrix and that $\mathbf{\Lambda}_n\mathbf{\Lambda}_n^H$ is a scaled identity matrix

$$\mathbf{\Lambda}_n\mathbf{\Lambda}_n^H = \alpha\boldsymbol{I}. \tag{4}$$

Suppose perturbed quantities on the right-hand side of (3) take the following forms:

$$\begin{aligned}\hat{\boldsymbol{U}}_s &= \boldsymbol{U}_s + \delta\boldsymbol{U}_s = \boldsymbol{U}_s + \boldsymbol{U}_n\boldsymbol{P}_1 + \boldsymbol{U}_s\boldsymbol{P}_2\\ \hat{\boldsymbol{V}}_s &= \boldsymbol{V}_s + \delta\boldsymbol{V}_s = \boldsymbol{V}_s + \boldsymbol{V}_n\bar{\boldsymbol{P}}_1 + \boldsymbol{V}_s\bar{\boldsymbol{P}}_2\end{aligned} \tag{5}$$

$$\begin{aligned}\hat{\boldsymbol{U}}_n &= \boldsymbol{U}_n + \delta\boldsymbol{U}_n = \boldsymbol{U}_n + \boldsymbol{U}_s\boldsymbol{Q}_1 + \boldsymbol{U}_n\boldsymbol{Q}_2\\ \hat{\boldsymbol{V}}_n &= \boldsymbol{V}_n + \delta\boldsymbol{V}_n = \boldsymbol{V}_n + \boldsymbol{V}_s\bar{\boldsymbol{Q}}_1 + \boldsymbol{V}_n\bar{\boldsymbol{Q}}_2\end{aligned} \tag{6}$$

$$\hat{\mathbf{\Lambda}}_s = \mathbf{\Lambda}_s + \delta\mathbf{\Lambda}_s, \quad \hat{\mathbf{\Lambda}}_n = \mathbf{\Lambda}_n + \delta\mathbf{\Lambda}_n \tag{7}$$

where the perturbations in the corresponding orthogonal subspaces are characterized by quantities with subscript $(\cdot)_1$, whereas the "in-space" perturbations in subspaces are explicitly considered by quantities with subscript $(\cdot)_2$. For example, in $\hat{\boldsymbol{U}}_n$, the perturbation $\delta\boldsymbol{U}_n$ consists of two parts: $\boldsymbol{U}_s\boldsymbol{Q}_1$ captures the perturbation in its orthogonal space $\boldsymbol{U}_s$, whereas $\boldsymbol{U}_n\boldsymbol{Q}_2$ represents the "in-space" perturbation. In these equations, there are several unknowns $\boldsymbol{P}_1, \boldsymbol{P}_2, \bar{\boldsymbol{P}}_1, \bar{\boldsymbol{P}}_2, \boldsymbol{Q}_1, \boldsymbol{Q}_2, \bar{\boldsymbol{Q}}_1, \bar{\boldsymbol{Q}}_2, \delta\mathbf{\Lambda}_s$, and $\delta\mathbf{\Lambda}_n$ that are to be determined in order to obtain all perturbations. In general, $\delta\mathbf{\Lambda}_s$ is not necessarily diagonal. As will be seen later, some of them are related to each other. We will start from (2) and (3) and apply the orthogonality principle to $\boldsymbol{X}$ and $\hat{\boldsymbol{X}}$.

With subspace decomposition (2), it is implied that

$$\begin{aligned}\boldsymbol{U}_s^H\boldsymbol{U}_s &= \boldsymbol{I}, \quad \boldsymbol{V}_s^H\boldsymbol{V}_s = \boldsymbol{I}, \quad \boldsymbol{U}_n^H\boldsymbol{U}_n = \boldsymbol{I}\\ \boldsymbol{V}_n^H\boldsymbol{V}_n &= \boldsymbol{I}, \quad \boldsymbol{U}_s^H\boldsymbol{U}_n = \boldsymbol{0}, \quad \boldsymbol{V}_s^H\boldsymbol{V}_n = \boldsymbol{0}\end{aligned} \tag{8}$$

where dimensions of identity matrices and zero matrices have been dropped for notational convenience. Similarly with (3), we have

$$\begin{aligned}\hat{\boldsymbol{U}}_s^H\hat{\boldsymbol{U}}_s &= \boldsymbol{I}, \quad \hat{\boldsymbol{V}}_s^H\hat{\boldsymbol{V}}_s = \boldsymbol{I}, \quad \hat{\boldsymbol{U}}_n^H\hat{\boldsymbol{U}}_n = \boldsymbol{I}\\ \hat{\boldsymbol{V}}_n^H\hat{\boldsymbol{V}}_n &= \boldsymbol{I}, \quad \hat{\boldsymbol{U}}_s^H\hat{\boldsymbol{U}}_n = \boldsymbol{0}, \quad \hat{\boldsymbol{V}}_s^H\hat{\boldsymbol{V}}_n = \boldsymbol{0}\end{aligned} \tag{9}$$

in parallel with (8).

From (2) and (8), a set of projections of $\boldsymbol{X}$ (or $\boldsymbol{X}^H$) onto different subspaces satisfy

$$\begin{aligned}\boldsymbol{X}^H\boldsymbol{U}_s &= \boldsymbol{V}_s\mathbf{\Lambda}_s^H, \quad \boldsymbol{X}^H\boldsymbol{U}_n = \boldsymbol{V}_n\mathbf{\Lambda}_n^H\\ \boldsymbol{X}\boldsymbol{V}_s &= \boldsymbol{U}_s\mathbf{\Lambda}_s, \quad \boldsymbol{X}\boldsymbol{V}_n = \boldsymbol{U}_n\mathbf{\Lambda}_n.\end{aligned} \tag{10}$$

Similarly, from (3) and (9), a set of projections of $\hat{\boldsymbol{X}}$ (or $\hat{\boldsymbol{X}}^H$) onto different perturbed subspaces satisfy

$$\begin{aligned}\hat{\boldsymbol{X}}^H\hat{\boldsymbol{U}}_s &= \hat{\boldsymbol{V}}_s\hat{\mathbf{\Lambda}}_s^H, \quad \hat{\boldsymbol{X}}^H\hat{\boldsymbol{U}}_n = \hat{\boldsymbol{V}}_n\hat{\mathbf{\Lambda}}_n^H\\ \hat{\boldsymbol{X}}\hat{\boldsymbol{V}}_s &= \hat{\boldsymbol{U}}_s\hat{\mathbf{\Lambda}}_s, \quad \hat{\boldsymbol{X}}\hat{\boldsymbol{V}}_n = \hat{\boldsymbol{U}}_n\hat{\mathbf{\Lambda}}_n.\end{aligned} \tag{11}$$

Equations in (10) and (11) will serve as our bases to derive perturbations up to the second order of $\delta\boldsymbol{X}$. Due to considerable length of derivation, the major results are provided in a compact form in the following theorem. The proof of the theorem is detailed in the Appendix, where $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$ are obtained first since other quantities are closely related to these two quantities.

Theorem: Suppose a fat or square matrix $\boldsymbol{X}$ is perturbed to be $\hat{\boldsymbol{X}} = \boldsymbol{X} + \delta\boldsymbol{X}$ with a small perturbation $\delta\boldsymbol{X}$. $\boldsymbol{X}$ and $\hat{\boldsymbol{X}}$ are decomposed as

$$\boldsymbol{X} = \boldsymbol{U}_s\mathbf{\Lambda}_s\boldsymbol{V}_s^H + \boldsymbol{U}_n\mathbf{\Lambda}_n\boldsymbol{V}_n^H, \quad \hat{\boldsymbol{X}} = \hat{\boldsymbol{U}}_s\hat{\mathbf{\Lambda}}_s\hat{\boldsymbol{U}}_s^H + \hat{\boldsymbol{U}}_n\hat{\mathbf{\Lambda}}_n\hat{\boldsymbol{U}}_n^H$$

where $\mathbf{\Lambda}_n\mathbf{\Lambda}_n^H = \alpha\boldsymbol{I}$. Define $\mathbf{\Sigma} = \left(\mathbf{\Lambda}_s\mathbf{\Lambda}_s^H - \alpha\boldsymbol{I}\right)^{-1}$ and different projections of $\delta\boldsymbol{X}$ as

$$\begin{aligned}\boldsymbol{E}_{ss} &= \boldsymbol{U}_s^H\delta\boldsymbol{X}\boldsymbol{V}_s, \quad \boldsymbol{E}_{sn} = \boldsymbol{U}_s^H\delta\boldsymbol{X}\boldsymbol{V}_n\\ \boldsymbol{E}_{ns} &= \boldsymbol{U}_n^H\delta\boldsymbol{X}\boldsymbol{V}_s, \quad \boldsymbol{E}_{nn} = \boldsymbol{U}_n^H\delta\boldsymbol{X}\boldsymbol{V}_n.\end{aligned}$$

If $\hat{\boldsymbol{U}}_s, \hat{\boldsymbol{V}}_s, \hat{\boldsymbol{U}}_n, \hat{\boldsymbol{V}}_n, \hat{\mathbf{\Lambda}}_s, \hat{\mathbf{\Lambda}}_n$ are expressed by (5)–(7), then up to the second order of $\delta\boldsymbol{X}$, these quantities $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\bar{\boldsymbol{P}}_1$, $\bar{\boldsymbol{P}}_2$, $\boldsymbol{Q}_1$, $\boldsymbol{Q}_2$, $\bar{\boldsymbol{Q}}_1$, $\bar{\boldsymbol{Q}}_2$, $\delta\mathbf{\Lambda}_s$ and $\delta\mathbf{\Lambda}_n$ take the following forms:

$$\begin{aligned}\boldsymbol{Q}_1 \approx{}& \mathbf{\Sigma}\left(\boldsymbol{E}_{ss}\mathbf{\Lambda}_s^{-1}\boldsymbol{E}_{sn} - \mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\boldsymbol{E}_{nn} + \alpha\boldsymbol{E}_{ss}\mathbf{\Lambda}_s^{-1}\mathbf{\Sigma}\boldsymbol{E}_{sn}\right.\\ &\left.+\mathbf{\Lambda}_s\boldsymbol{E}_{ss}^H\mathbf{\Sigma}\boldsymbol{E}_{sn} - \mathbf{\Sigma}\mathbf{\Lambda}_s\boldsymbol{E}_{ns}^H\boldsymbol{E}_{nn}\right)\mathbf{\Lambda}_n^H\\ &+\mathbf{\Sigma}\left(\mathbf{\Lambda}_s\boldsymbol{E}_{ss}^H\mathbf{\Sigma}\mathbf{\Lambda}_s\boldsymbol{E}_{ns}^H - \mathbf{\Sigma}\mathbf{\Lambda}_s\boldsymbol{E}_{ns}^H\mathbf{\Lambda}_n\boldsymbol{E}_{nn}^H\right.\\ &\left.+\alpha\boldsymbol{E}_{ss}\mathbf{\Sigma}\boldsymbol{E}_{ns}^H - \mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\mathbf{\Lambda}_n\boldsymbol{E}_{nn}^H - \boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H\right)\\ &+\boldsymbol{F}\end{aligned} \tag{12}$$

$$\begin{aligned}\bar{\boldsymbol{Q}}_1 \approx{}& \mathbf{\Sigma}\left(\boldsymbol{E}_{ss}^H\mathbf{\Sigma}\mathbf{\Lambda}_s\boldsymbol{E}_{ns}^H - \mathbf{\Sigma}\boldsymbol{E}_{ns}^H\mathbf{\Lambda}_n\boldsymbol{E}_{nn}^H\right)\mathbf{\Lambda}_n - \mathbf{\Sigma}\boldsymbol{E}_{ns}^H\boldsymbol{E}_{nn}\\ &+\mathbf{\Sigma}\left(\boldsymbol{E}_{ss}^H\mathbf{\Sigma}\boldsymbol{E}_{sn} - \mathbf{\Sigma}\boldsymbol{E}_{ns}^H\boldsymbol{E}_{nn}\right)\mathbf{\Lambda}_n^H\mathbf{\Lambda}_n\\ &+\mathbf{\Lambda}_s^{-1}\mathbf{\Sigma}\left(\boldsymbol{E}_{ss}\mathbf{\Lambda}_s^{-1}\boldsymbol{E}_{sn} - \mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\boldsymbol{E}_{nn}\right.\\ &\left.+\alpha\boldsymbol{E}_{ss}\mathbf{\Lambda}_s^{-1}\mathbf{\Sigma}\boldsymbol{E}_{sn}\right)\mathbf{\Lambda}_n^H\mathbf{\Lambda}_n\\ &+\mathbf{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\mathbf{\Lambda}_s^{-1}\left(\boldsymbol{E}_{sn} + \mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\mathbf{\Lambda}_n\right)\\ &-\mathbf{\Lambda}_s^{-1}\mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\boldsymbol{E}_{nn} + \mathbf{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\mathbf{\Sigma}\boldsymbol{E}_{ns}^H\mathbf{\Lambda}_n\\ &+\mathbf{\Lambda}_s^{-1}\mathbf{\Sigma}\left(\alpha\boldsymbol{E}_{ss}\mathbf{\Sigma}\boldsymbol{E}_{ns}^H - \mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\mathbf{\Lambda}_n\boldsymbol{E}_{nn}^H\right.\\ &\left.-\boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H\right)\mathbf{\Lambda}_n + \bar{\boldsymbol{F}}\end{aligned} \tag{13}$$

$$\boldsymbol{P}_1 \approx -\boldsymbol{Q}_1^H \tag{14}$$

$$\bar{\boldsymbol{P}}_1 \approx -\bar{\boldsymbol{Q}}_1^H \tag{15}$$

$$\boldsymbol{Q}_2 \approx -\frac{1}{2}\boldsymbol{F}^H\boldsymbol{F} \tag{16}$$

$$\bar{\boldsymbol{Q}}_2 \approx -\frac{1}{2}\bar{\boldsymbol{F}}^H\bar{\boldsymbol{F}} \tag{17}$$

$$\boldsymbol{P}_2 \approx -\frac{1}{2}\boldsymbol{F}\boldsymbol{F}^H \tag{18}$$

$$\bar{\boldsymbol{P}}_2 \approx -\frac{1}{2}\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^H \tag{19}$$

$$\delta\mathbf{\Lambda}_s \approx \boldsymbol{E}_{ss} - \boldsymbol{E}_{sn}\bar{\boldsymbol{F}}^H - \frac{1}{2}\mathbf{\Lambda}_s\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^H + \frac{1}{2}\boldsymbol{F}\boldsymbol{F}^H\mathbf{\Lambda}_s \tag{20}$$

$$\delta\mathbf{\Lambda}_n \approx \boldsymbol{E}_{nn} + \boldsymbol{E}_{ns}\bar{\boldsymbol{F}} - \frac{1}{2}\mathbf{\Lambda}_n\bar{\boldsymbol{F}}^H\bar{\boldsymbol{F}} + \frac{1}{2}\boldsymbol{F}^H\boldsymbol{F}\mathbf{\Lambda}_n \tag{21}$$

where $\boldsymbol{F}$ and $\bar{\boldsymbol{F}}$ represent the first-order terms in $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$, respectively

$$\begin{aligned}\boldsymbol{F} &= -\mathbf{\Sigma}\left(\mathbf{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\right)\\ \bar{\boldsymbol{F}} &= -\mathbf{\Sigma}\boldsymbol{E}_{ns}^H\mathbf{\Lambda}_n - \mathbf{\Lambda}_s^{-1}\mathbf{\Sigma}\boldsymbol{E}_{sn}\mathbf{\Lambda}_n^H\mathbf{\Lambda}_n - \mathbf{\Lambda}_s^{-1}\boldsymbol{E}_{sn}.\end{aligned} \tag{22}$$

Proof: See the Appendix. □

According to this theorem, the procedure to obtain all perturbations up to the second order of $\delta\boldsymbol{X}$ in subspace decomposition on $\boldsymbol{X}$ can be summarized as follows.

1) From subspace decomposition on $\boldsymbol{X}$, obtain $\boldsymbol{U}_s$, $\boldsymbol{V}_s$, $\boldsymbol{U}_n$, $\boldsymbol{V}_n$, $\boldsymbol{\Lambda}_s$, and $\boldsymbol{\Lambda}_n$.
2) Precompute relevant quantities α, $\boldsymbol{\Sigma}$, $\boldsymbol{E}_{ss}$, $\boldsymbol{E}_{sn}$, $\boldsymbol{E}_{ns}$, and $\boldsymbol{E}_{nn}$.
3) Obtain $\boldsymbol{F}$ and $\bar{\boldsymbol{F}}$ and then $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$.
4) Obtain quantities $\boldsymbol{P}_1$, $\bar{\boldsymbol{P}}_1$, $\boldsymbol{Q}_2$, $\bar{\boldsymbol{Q}}_2$, and $\boldsymbol{P}_2$, $\bar{\boldsymbol{P}}_2$ and singular value perturbations $\delta\boldsymbol{\Lambda}_s$, and $\delta\boldsymbol{\Lambda}_n$.
5) Obtain subspace perturbations $\delta\boldsymbol{U}_s$, $\delta\boldsymbol{V}_s$, $\delta\boldsymbol{U}_n$, $\delta\boldsymbol{V}_n$, as

$$\delta\boldsymbol{U}_s = \boldsymbol{U}_n\boldsymbol{P}_1 + \boldsymbol{U}_s\boldsymbol{P}_2, \quad \delta\boldsymbol{V}_s = \boldsymbol{V}_n\bar{\boldsymbol{P}}_1 + \boldsymbol{V}_s\bar{\boldsymbol{P}}_2$$
$$\delta\boldsymbol{U}_n = \boldsymbol{U}_s\boldsymbol{Q}_1 + \boldsymbol{U}_n\boldsymbol{Q}_2, \quad \delta\boldsymbol{V}_n = \boldsymbol{V}_s\bar{\boldsymbol{Q}}_1 + \boldsymbol{V}_n\bar{\boldsymbol{Q}}_2.$$

This theorem provides general results of perturbations for all relevant quantities in subspace decomposition up to the second order. The following remarks can be easily made.

1) If a matrix is tall (with more rows than columns), the theorem can be applied to its Hermitian first to obtain perturbed subspace decomposition. Then, by Hermitian operation, the perturbed subspace decomposition of the desired matrix follows.
2) Perturbations up to the second order give approximations with errors at least in the third order of perturbation $\delta\boldsymbol{X}$. Compared with the first-order perturbation results, approximation errors will be reduced by one order theoretically.
3) Second-order perturbations depend on projections of $\delta\boldsymbol{X}$ onto all subspaces determined by $\boldsymbol{E}_{ss}$, $\boldsymbol{E}_{sn}$, $\boldsymbol{E}_{ns}$, and $\boldsymbol{E}_{nn}$. However, first-order approximations only depend on "cross-space" projections $\boldsymbol{E}_{sn}$ and $\boldsymbol{E}_{ns}$, as seen from (22). Therefore, second-order approximation is able to provide additional information about $\delta\boldsymbol{X}$.
4) It is observed from (16)–(19) that quantities $\boldsymbol{Q}_2$, $\bar{\boldsymbol{Q}}_2$, $\boldsymbol{P}_2$, and $\bar{\boldsymbol{P}}_2$ are square matrices and have a Hermitian (complex conjugate symmetric) property. None of them have first-order terms of $\delta\boldsymbol{X}$. These quantities quantify the "in-space" perturbations according to their definitions. Therefore, "in-space" perturbation is smaller than the corresponding perturbation in its orthogonal subspace as far as a particular subspace is concerned. This is partially explained after (61) in the Appendix.
5) Perturbations $\delta\boldsymbol{\Lambda}_s$ and $\delta\boldsymbol{\Lambda}_n$ up to the second order are not Hermitian in general, even though both $\boldsymbol{\Lambda}_s$ and $\boldsymbol{\Lambda}_n$ are Hermitian in some cases. The first-order perturbation in $\delta\boldsymbol{\Lambda}_s$ is $\boldsymbol{E}_{ss}$. It is thus not affected by either other projections ($\boldsymbol{E}_{sn}$, $\boldsymbol{E}_{ns}$, $\boldsymbol{E}_{nn}$), $\boldsymbol{\Lambda}_s$, or $\boldsymbol{\Lambda}_n$. However, the second-order perturbation in $\delta\boldsymbol{\Lambda}_s$ depends on many other quantities, especially $\boldsymbol{\Lambda}_n$, which exists in $\boldsymbol{F}$ and $\bar{\boldsymbol{F}}$. Similar conclusion can be made for $\delta\boldsymbol{\Lambda}_n$.
6) It is observed that many second-order terms are involved in the associated quantities, which makes analysis of a perturbation problem complicated in many scenarios. However, the number of terms can be significantly reduced in certain applications, leading to simplicity of the analytical results.

More properties of our second-order approximation might be explored when particular applications are considered. Before we provide applications of our theorem in both matrix computation and wireless communication, we will first reveal the connections with existing first-order approaches. It will be shown that they are consistent when our results are truncated at the first order.

III. Connections With Existing Approaches

Currently, there exist first-order methods in noisy data analysis [10] and eigenvalue decomposition of sample covariance matrix [7]. It this section, let us reveal connections of our second-order approximation results with them, respectively, by directly applying our theorem.

A. Perturbation by Noise in the Received Data—Connection With [10]

In this subsection, we show that our perturbation solutions can be significantly simplified when data is perturbed by noise only. We also show that our analytical results are consistent with the first-order based approach [10] when only first-order approximations are adopted in such a scenario.

Consider the following noisy data model [10]

$$\hat{\boldsymbol{Y}} = \boldsymbol{Y} + \delta\boldsymbol{Y} = \boldsymbol{Y} + \boldsymbol{N}$$

where the noise matrix $\boldsymbol{N}$ is a perturbation to the data matrix $\boldsymbol{Y}$. Since $\boldsymbol{\Lambda}_n = \boldsymbol{0}$, $\boldsymbol{\Lambda}_s$ is positive and square diagonal, then $\alpha = 0$ and $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}_s^{-2}$. We will follow our previously described procedure to obtain perturbations and compare them with the results in [10].

According to (22), the first-order terms become

$$\boldsymbol{F} = -\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ns}^H, \quad \bar{\boldsymbol{F}} = -\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}. \tag{23}$$

Then, from (12), we obtain

$$\boldsymbol{Q}_1 \approx -\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ns}^H + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}^H\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ns}^H - \boldsymbol{\Lambda}_s^{-2}\boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H. \tag{24}$$

Compared with long expressions in (12), (24) has a much simpler form. Based on (6) and (24), the perturbation of noise subspace follows:

$$\begin{aligned}\delta\boldsymbol{U}_n \approx &-\boldsymbol{Y}^{\dagger}\delta\boldsymbol{Y}^H\boldsymbol{U}_n + \boldsymbol{Y}^{\dagger}\delta\boldsymbol{Y}^H\boldsymbol{Y}^{\dagger}\delta\boldsymbol{Y}^H\boldsymbol{U}_n\\ &-\boldsymbol{Y}^{\dagger}(\boldsymbol{Y}^{\dagger})^H\delta\boldsymbol{Y}\boldsymbol{V}_n\boldsymbol{V}_n^H\delta\boldsymbol{Y}^H\boldsymbol{U}_n\\ &-\frac{1}{2}\boldsymbol{U}_n\boldsymbol{U}_n^H\delta\boldsymbol{Y}\boldsymbol{V}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{V}_s^H\delta\boldsymbol{Y}^H\boldsymbol{U}_n\end{aligned} \tag{25}$$

where $\boldsymbol{Y}^{\dagger} \triangleq \boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{V}_s^H$, and $\boldsymbol{Q}_2 \approx -(1/2)\boldsymbol{F}^H\boldsymbol{F}$ has been used. Likewise, we can find $\boldsymbol{P}_1$ according to (14) and (24)

$$\boldsymbol{P}_1 \approx \boldsymbol{E}_{ns}\boldsymbol{\Lambda}_s^{-1} - \boldsymbol{E}_{ns}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1} + \boldsymbol{E}_{nn}\boldsymbol{E}_{sn}^H\boldsymbol{\Lambda}_s^{-2}. \tag{26}$$

Therefore, from (5), the perturbation of the signal subspace becomes

$$\begin{aligned}\delta\boldsymbol{U}_s \approx \boldsymbol{U}_n\boldsymbol{U}_n^H\delta\boldsymbol{Y}\Big[&\boldsymbol{V}_s\boldsymbol{\Lambda}_s^{-1} - (\boldsymbol{Y}^{\dagger})^H\delta\boldsymbol{Y}\boldsymbol{V}_s\boldsymbol{\Lambda}_s^{-1}\\ &+\boldsymbol{V}_n\boldsymbol{V}_n^H\delta\boldsymbol{Y}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\Big] - \frac{1}{2}\boldsymbol{Y}^{\dagger}\delta\boldsymbol{Y}^H\boldsymbol{U}_n\boldsymbol{U}_n^H\delta\boldsymbol{Y}\boldsymbol{V}_s\end{aligned} \tag{27}$$

where $\boldsymbol{P}_2 \approx -(1/2)\boldsymbol{F}\boldsymbol{F}^H$ has been used. It can be observed from (25) and (27) that if we focus on the first-order perturbations, then expressions for $\delta \boldsymbol{U}_n$ and $\delta \boldsymbol{U}_s$ reduce to those in [10]. This consistency is not surprising because similar normalization and orthogonality properties are employed in our derivation. However, our results in the theorem are extended to second-order perturbations, where $\boldsymbol{\Lambda}_n = \boldsymbol{0}$ is not required in general. Moreover, other perturbed quantities can be obtained as well from our theorem. Meanwhile, it is worth mentioning that the results we derive in this subsection are not restricted to a data matrix according to our derivation. They are also applicable to matrix $\boldsymbol{Y}$ with $\boldsymbol{\Lambda}_n = \boldsymbol{0}$ and perturbation $\delta \boldsymbol{Y}$.

B. Perturbation by Finite Data in Covariance Matrix—Connection With [7]

Perturbation also arises due to finite data samples in estimation. For example, the data covariance matrix is estimated from N data vectors [7]

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{y}(n)\boldsymbol{y}^H(n). \tag{28}$$

Assume the true covariance is $\boldsymbol{R} = \boldsymbol{Z} + \sigma_v^2 \boldsymbol{I}$, where $\boldsymbol{Z}$ represents the covariance of all signals, and σ_v^2 is the power of the white Gaussian noise. $\boldsymbol{Z}$ may have decomposition

$$\boldsymbol{Z} = \boldsymbol{U}_s \boldsymbol{\Omega} \boldsymbol{U}_s^H. \tag{29}$$

Then, $\boldsymbol{R}$ can be decomposed into

$$\begin{aligned} \boldsymbol{R} =& \boldsymbol{U}_s \boldsymbol{\Lambda}_s \boldsymbol{U}_s^H + \boldsymbol{U}_n \boldsymbol{\Lambda}_n \boldsymbol{U}_n^H \\ \boldsymbol{\Lambda}_s =& \boldsymbol{\Omega} + \sigma_v^2 \boldsymbol{I}, \quad \boldsymbol{\Lambda}_n = \sigma_v^2 \boldsymbol{I}. \end{aligned} \tag{30}$$

It is recognized [5] (under certain ergodicity assumptions) that $\hat{\boldsymbol{R}}$ converges to $\boldsymbol{R}$ as $N \to \infty$. However, for large N, a perturbation $\delta \boldsymbol{R}$ exists in $\hat{\boldsymbol{R}}$ [7]

$$\hat{\boldsymbol{R}} = \boldsymbol{R} + \delta \boldsymbol{R}. \tag{31}$$

Therefore, subspaces are also perturbed. We will apply our theorem to investigate how $\delta \boldsymbol{R}$ affects subspace decomposition on $\hat{\boldsymbol{R}}$. Some results will also be used in Section III-A.

Under sample average estimation method (28), it is first observed from (31) that $\delta \boldsymbol{R}^H = \delta \boldsymbol{R}$ since $\hat{\boldsymbol{R}}^H = \hat{\boldsymbol{R}}$. Therefore, projections of $\delta \boldsymbol{R}$ onto different subspaces have following relations:

$$\boldsymbol{E}_{ss}^H = \boldsymbol{E}_{ss}, \quad \boldsymbol{E}_{nn}^H = \boldsymbol{E}_{nn}, \quad \boldsymbol{E}_{sn}^H = \boldsymbol{E}_{ns}. \tag{32}$$

In addition, $\alpha = \sigma_v^4$, and

$$\boldsymbol{\Sigma} = \left(\boldsymbol{\Lambda}_s^2 - \sigma_v^4 \boldsymbol{I}\right)^{-1} = \boldsymbol{\Omega}^{-1}\left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right)^{-1}. \tag{33}$$

With these simplifications, it is quite straightforward to simplify our perturbation results. We start from (12) for $\boldsymbol{Q}_1$. Since

$$\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right) = \boldsymbol{\Omega}^{-1} \tag{34}$$

according to (33), the first-order term $\boldsymbol{F}$ in (22) becomes

$$\boldsymbol{F} = -\boldsymbol{\Omega}^{-1} \boldsymbol{U}_s^H \delta \boldsymbol{R} \boldsymbol{U}_n. \tag{35}$$

Based on (32), all second-order terms in $\boldsymbol{Q}_1$ are partitioned into two groups for ease of derivation. The first group has five terms with positive signs, each of which has projections $\boldsymbol{E}_{ss}$ and $\boldsymbol{E}_{sn}$

$$\begin{aligned} \boldsymbol{T}_1 = \sigma_v^2 \boldsymbol{\Sigma} \boldsymbol{E}_{ss} \boldsymbol{\Lambda}_s^{-1} \boldsymbol{E}_{sn} + \sigma_v^4 \boldsymbol{\Sigma} \boldsymbol{E}_{ss} \boldsymbol{\Lambda}_s^{-1} \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right) \boldsymbol{E}_{sn} \\ + \boldsymbol{\Sigma} \boldsymbol{\Lambda}_s \boldsymbol{E}_{ss} \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right) \boldsymbol{E}_{sn}. \end{aligned}$$

Applying (34) and combining the first two terms, $\boldsymbol{T}_1$ is simplified to

$$\boldsymbol{T}_1 = \sigma_v^2 \boldsymbol{\Sigma} \boldsymbol{E}_{ss} \boldsymbol{\Omega}^{-1} \boldsymbol{\Lambda}_s^{-1} (\boldsymbol{\Omega} + \sigma_v^2 \boldsymbol{I}) \boldsymbol{E}_{sn} + \boldsymbol{\Sigma} \boldsymbol{\Lambda}_s \boldsymbol{E}_{ss} \boldsymbol{\Omega}^{-1} \boldsymbol{E}_{sn}.$$

Due to (30) and (34), $\boldsymbol{T}_1$ finally becomes

$$\boldsymbol{T}_1 = \boldsymbol{\Omega}^{-1} \boldsymbol{E}_{ss} \boldsymbol{\Omega}^{-1} \boldsymbol{E}_{sn}.$$

Similarly, the second group also has five terms but with negative signs, each of which has projections $\boldsymbol{E}_{sn}$ and $\boldsymbol{E}_{nn}$

$$\begin{aligned} \boldsymbol{T}_2 =& -\left[2\sigma_v^2 \boldsymbol{\Sigma}^2 \left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right) + \boldsymbol{\Sigma}\right] \boldsymbol{E}_{sn} \boldsymbol{E}_{nn} \\ =& -\left(2\sigma_v^2 \boldsymbol{\Sigma} \boldsymbol{\Omega}^{-1} + \boldsymbol{\Sigma}\right) \boldsymbol{E}_{sn} \boldsymbol{E}_{nn}. \end{aligned}$$

Since

$$\begin{aligned} 2\sigma_v^2 \boldsymbol{\Sigma} \boldsymbol{\Omega}^{-1} + \boldsymbol{\Sigma} =& \boldsymbol{\Omega}^{-1} \boldsymbol{\Sigma}\left(2\sigma_v^2 \boldsymbol{I} + \boldsymbol{\Omega}\right) \\ =& \boldsymbol{\Omega}^{-1} \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right) = \boldsymbol{\Omega}^{-2} \end{aligned}$$

$\boldsymbol{T}_2$ is simplified to $\boldsymbol{T}_2 = -\boldsymbol{\Omega}^{-2} \boldsymbol{E}_{sn} \boldsymbol{E}_{nn}$. Combining $\boldsymbol{F}$, $\boldsymbol{T}_1$, and $\boldsymbol{T}_2$, we obtain $\boldsymbol{Q}_1$ as follows:

$$\boldsymbol{Q}_1 \approx -\boldsymbol{\Omega}^{-1} \boldsymbol{E}_{sn} + \boldsymbol{\Omega}^{-1} \boldsymbol{E}_{ss} \boldsymbol{\Omega}^{-1} \boldsymbol{E}_{sn} - \boldsymbol{\Omega}^{-2} \boldsymbol{E}_{sn} \boldsymbol{E}_{nn}. \tag{36}$$

Therefore, the perturbation of the noise subspace is

$$\begin{aligned} \delta \boldsymbol{U}_n \approx & -\boldsymbol{Z}^{\dagger} \delta \boldsymbol{R} \boldsymbol{U}_n + \boldsymbol{Z}^{\dagger} \delta \boldsymbol{R} \boldsymbol{Z}^{\dagger} \delta \boldsymbol{R} \boldsymbol{U}_n \\ & -(\boldsymbol{Z}^{\dagger})^2 \delta \boldsymbol{R} \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} \boldsymbol{U}_n - \frac{1}{2} \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} (\boldsymbol{Z}^{\dagger})^2 \delta \boldsymbol{R} \boldsymbol{U}_n \end{aligned} \tag{37}$$

where $(\cdot)^{\dagger}$ represents the pseudo-inverse, projections $\boldsymbol{E}_{ss}$, $\boldsymbol{E}_{sn}$, and $\boldsymbol{E}_{nn}$ have been replaced by their definitions, and the last term comes from $\boldsymbol{U}_n \boldsymbol{Q}_2$. Since $\delta \boldsymbol{U}_s \approx -\boldsymbol{U}_n \boldsymbol{Q}_1^H - (1/2) \boldsymbol{U}_s \boldsymbol{F} \boldsymbol{F}^H$, we can easily obtain the perturbation of the signal subspace

$$\begin{aligned} \delta \boldsymbol{U}_s \approx & \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} \boldsymbol{U}_s \boldsymbol{\Omega}^{-1} - \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} \boldsymbol{Z}^{\dagger} \delta \boldsymbol{R} \boldsymbol{U}_s \boldsymbol{\Omega}^{-1} \\ & + \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} \boldsymbol{U}_s \boldsymbol{\Omega}^{-2} \\ & - \frac{1}{2} \boldsymbol{Z}^{\dagger} \delta \boldsymbol{R} \boldsymbol{U}_n \boldsymbol{U}_n^H \delta \boldsymbol{R} \boldsymbol{U}_s \boldsymbol{\Omega}^{-1}. \end{aligned} \tag{38}$$

In order to find perturbations in eigenvalues according to (20) and (21), we first need to simplify $\bar{\boldsymbol{F}}$. From (22), we can obtain

$$\begin{aligned} \bar{\boldsymbol{F}} =& -\left(\sigma_v^2 \boldsymbol{\Sigma} + \sigma_v^4 \boldsymbol{\Lambda}_s^{-1} \boldsymbol{\Sigma} + \boldsymbol{\Lambda}_s^{-1}\right) \boldsymbol{E}_{sn} \\ =& -\left[\sigma_v^2 \boldsymbol{\Lambda}_s^{-1} \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s + \sigma_v^2 \boldsymbol{I}\right) + \boldsymbol{\Lambda}_s^{-1}\right] \boldsymbol{E}_{sn}. \end{aligned}$$

Using (30) and (34), $\bar{\boldsymbol{F}}$ can be further simplified to be

$$\begin{aligned} \bar{\boldsymbol{F}} =& -\left(\sigma_v^2 \boldsymbol{\Lambda}_s^{-1} \boldsymbol{\Omega}^{-1} \boldsymbol{\Lambda}_s^{-1}\right) \boldsymbol{E}_{sn} \\ =& -\boldsymbol{\Omega}^{-1} \boldsymbol{\Lambda}_s^{-1}\left(\boldsymbol{\Omega} + \sigma_v^2 \boldsymbol{I}\right) \boldsymbol{E}_{sn} \\ =& -\boldsymbol{\Omega}^{-1} \boldsymbol{U}_s^H \delta \boldsymbol{R} \boldsymbol{U}_n \end{aligned} \tag{39}$$

which is the same as $\boldsymbol{F}$. Therefore, from (20) and (21), we can find

$$\begin{aligned}\delta\boldsymbol{\Lambda}_s \approx &\boldsymbol{U}_s^H\delta\boldsymbol{R}\boldsymbol{U}_s + \boldsymbol{U}_s^H\delta\boldsymbol{R}\boldsymbol{U}_n\boldsymbol{U}_n^H\delta\boldsymbol{R}\boldsymbol{U}_s\boldsymbol{\Omega}^{-1}\\ &- \frac{1}{2}\boldsymbol{\Lambda}_s\boldsymbol{\Omega}^{-1}\boldsymbol{U}_s^H\delta\boldsymbol{R}\boldsymbol{U}_n\boldsymbol{U}_n^H\delta\boldsymbol{R}\boldsymbol{U}_s\boldsymbol{\Omega}^{-1}\\ &+ \frac{1}{2}\boldsymbol{\Omega}^{-1}\boldsymbol{U}_s^H\delta\boldsymbol{R}\boldsymbol{U}_n\boldsymbol{U}_n^H\delta\boldsymbol{R}\boldsymbol{U}_s\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_s \end{aligned} \quad (40)$$

$$\delta\boldsymbol{\Lambda}_n \approx \boldsymbol{U}_n^H\delta\boldsymbol{R}\boldsymbol{U}_n - \boldsymbol{U}_n^H\delta\boldsymbol{R}\boldsymbol{U}_s\boldsymbol{\Omega}^{-1}\boldsymbol{U}_s^H\delta\boldsymbol{R}\boldsymbol{U}_n. \quad (41)$$

It can be easily observed from (37), (38), (40), and (41) that, when we focus on one eigenpair only, the first-order results reduce to those in [7]. However, instead of considering individual eigenpair, we present a compact form of up to the second-order perturbations for all eigenpairs. Although we assume $\boldsymbol{R}$ is a data covariance matrix, the results can be applicable to any Hermitian matrix with Hermitian perturbation $\delta\boldsymbol{R}$.

Up to this point, we have shown connections of our analytical results with typical first-order approaches. We have also simplified expressions for perturbations significantly in those two scenarios. Those results can be served as general references in corresponding applications. However, we have not shown the importance of the second-order perturbations. First-order perturbation results have been widely applied as seen in the literature [3], [17]. Second-order approximation is essential in obtaining more accurate analytical results. The approximation error will be in the third order of perturbation $\delta\boldsymbol{X}$ when second-order terms are included, but the error will be in the second order of perturbation $\delta\boldsymbol{X}$ when only first-order terms are considered. We will present two numerical examples next to illustrate how our second-order approximation can improve the accuracy. Although other applications may also be found, e.g., in high precision data analysis, we will not pursue extensive study on applications in the current paper.

IV. Numerical Examples

In this section, we study applications of our theorem by providing some numerical examples in both matrix computation and wireless communication.

A. Example in Matrix Computation

Suppose a symmetric matrix $\boldsymbol{X}$ is perturbed to be $\hat{\boldsymbol{X}}$ with perturbation $\delta\boldsymbol{X}$ as follows:

$$\boldsymbol{X} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \quad \hat{\boldsymbol{X}} = \begin{bmatrix}1 & \epsilon\\ \epsilon & 0\end{bmatrix}, \quad \delta\boldsymbol{X} = \epsilon\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

where $0 < \epsilon \ll 1$. It can be easily found that the eigenvalue decomposition on $\boldsymbol{X}$ gives

$$\begin{aligned}\boldsymbol{U}_s =&\boldsymbol{V}_s = \begin{bmatrix}1\\0\end{bmatrix}, \quad \boldsymbol{U}_n = \boldsymbol{V}_n = \begin{bmatrix}0\\1\end{bmatrix}\\ \Lambda_s =&1, \quad \Lambda_n = 0.\end{aligned}$$

Here, Λ_s and Λ_n are not boldfaced since they become scalars. Now, consider the eigenvalue decomposition on $\hat{\boldsymbol{X}}$ analytically. It has two eigenvalues $1/2 \pm 1/2\sqrt{1+4\epsilon^2}$. If they are approximated up to ϵ^2, then they become $\hat{\Lambda}_s \approx 1+\epsilon^2$, $\hat{\Lambda}_n \approx -\epsilon^2$.

TABLE I
Matrix Perturbation Results

ϵ	2.0×10^{-1}	1.0×10^{-1}	1.0×10^{-2}	1.0×10^{-3}	1.0×10^{-4}
$e_{u_s,1}$	2.1×10^{-2}	5.1×10^{-3}	5.0×10^{-5}	5.0×10^{-7}	5.0×10^{-9}
$e_{u_s,2}$	1.1×10^{-2}	1.5×10^{-3}	1.5×10^{-6}	1.5×10^{-9}	1.5×10^{-12}
$e_{\Lambda_s,1}$	3.9×10^{-2}	9.9×10^{-3}	1.0×10^{-4}	1.0×10^{-6}	1.0×10^{-8}
$e_{\Lambda_s,2}$	1.5×10^{-3}	9.8×10^{-5}	1.0×10^{-8}	1.0×10^{-12}	0

Therefore, $\delta\Lambda_s \approx \epsilon^2$, $\delta\Lambda_n \approx -\epsilon^2$. Corresponding to $\hat{\Lambda}_s$ and $\hat{\Lambda}_n$, the associated eigenvectors can be found from eigenanalysis

$$\begin{aligned}\hat{\boldsymbol{U}}_s &\approx \begin{bmatrix}1-\frac{1}{2}\epsilon^2\\ \epsilon\end{bmatrix} = \boldsymbol{U}_s + \epsilon\boldsymbol{U}_n - \frac{1}{2}\epsilon^2\boldsymbol{U}_s\\ \hat{\boldsymbol{U}}_n &\approx \begin{bmatrix}-\epsilon\\ 1-\frac{1}{2}\epsilon^2\end{bmatrix} = \boldsymbol{U}_n - \epsilon\boldsymbol{U}_s - \frac{1}{2}\epsilon^2\boldsymbol{U}_n.\end{aligned}$$

If notations in our theorem are used, then the corresponding quantities follow:

$$\begin{aligned}P_1 =&\epsilon, \quad P_2 = -\frac{1}{2}\epsilon^2, \quad Q_1 = -\epsilon, \quad Q_2 = -\frac{1}{2}\epsilon^2\\ \delta\Lambda_s =&\epsilon^2, \quad \delta\Lambda_n = -\epsilon^2.\end{aligned} \quad (42)$$

Next, we obtain these quantities by directly applying our theorem. Following the procedure described before, α, Σ, E_{ss}, E_{sn}, E_{ns}, and E_{nn} can be found first:

$$\begin{aligned}\alpha =&0, \quad \Sigma = 1, \quad E_{ss} = 0, \quad E_{sn} = \epsilon\\ E_{ns} =&\epsilon, \quad E_{nn} = 0.\end{aligned}$$

Then, $F = -\epsilon$, and $\bar{F} = -\epsilon$. With these quantities, we can obtain the same analytical results as (42). This shows that results from our theorem are consistent with analytical results based on conventional techniques.

In order to show the improvement by our second-order approximation, we also test the eigenvalue decomposition results numerically for a large range of ϵ based on a function "eig" in Matlab. Our second-order approximations for $\hat{\boldsymbol{U}}_s$, $\hat{\boldsymbol{U}}_n$, $\hat{\Lambda}_s$, and $\hat{\Lambda}_n$ are compared with their true values obtained in Matlab. The absolute errors, which are denoted as $e_{u_s,2}$, $e_{u_n,2}$, $e_{\Lambda_s,2}$, and $e_{\Lambda_n,2}$ are recorded, respectively. When two vectors are compared, the error is defined as the Euclidean norm of the resulting error vector. Those errors are also compared with the first-order approximation errors [with subscript $(\cdot)_1$]. The results for $\hat{\boldsymbol{U}}_s$ and $\hat{\Lambda}_s$ are presented in Table I, whereas the results for $\hat{\boldsymbol{U}}_n$ and $\hat{\Lambda}_n$ are identical and are thus omitted. The advantage of including second-order terms can be easily observed by comparing the third row with the second row and the last row with the fourth row, respectively.

B. Example in Subspace-Based Channel Estimation

Channel estimation has been extensively studied in wireless communication. One of the most efficient methods is the so-called subspace-based method [4], [12]. If the noise subspace is used, then it can be obtained from subspace decomposition on the sample covariance matrix $\hat{\boldsymbol{R}}$ in (28). As an example, we show in this subsection how our perturbation results can be applied to derive the channel estimation error analytically up to the second order of perturbation $\delta\boldsymbol{R}$. Then, the result is used to analyze such a channel estimator numerically. With perturbation up to the second order, the proposed analysis is more accurate than that based only on the first-order perturbation under a small perturbation assumption.

Consider a synchronous direct sequence CDMA system with K users [11]. For User 1, the desired user has independent and identically distributed (i.i.d.) inputs $w_1(n)$ and is assigned a periodic spreading sequence $c_1(0), \ldots, c_1(P-1)$ with period P. The coefficients of its unknown channel are collected in a normalized vector $\boldsymbol{h}_1$. Similarly, we can define quantities for other users. In the presence of additive white Gaussian noise (AWGN) $\boldsymbol{v}(n)$, the received data is assumed to follow the model [11]

$$\boldsymbol{y}(n) = \sum_{k=1}^{K} \boldsymbol{C}_k \boldsymbol{h}_k w_k(n) + \boldsymbol{v}(n) \tag{43}$$

where $\boldsymbol{C}_k$ is a code filtering Toeplitz matrix constructed from the spreading codes of user k. Assume that the noise subspace from eigenvalue decomposition on $\hat{\boldsymbol{R}}$ is $\hat{\boldsymbol{U}}_n$. Then, the subspace-based method to obtain the desired channel parameters can be formulated as follows:

$$\hat{\boldsymbol{h}}_1 = \arg \min_{\|\hat{\boldsymbol{h}}\|=1} \hat{\boldsymbol{h}}^H \boldsymbol{C}_1^H \hat{\boldsymbol{U}}_n \hat{\boldsymbol{U}}_n^H \boldsymbol{C}_1 \hat{\boldsymbol{h}}. \tag{44}$$

Therefore, $\hat{\boldsymbol{h}}_1$ is an eigenvector of the objective matrix $\hat{\boldsymbol{X}} = \boldsymbol{C}_1^H \hat{\boldsymbol{U}}_n \hat{\boldsymbol{U}}_n^H \boldsymbol{C}_1$ corresponding to its minimum eigenvalue. Due to inaccurate estimate for $\boldsymbol{R}$, perturbation $\delta\boldsymbol{R}$ results in a perturbation $\delta\boldsymbol{U}_n$ in $\hat{\boldsymbol{U}}_n$. Hence, matrix $\hat{\boldsymbol{X}}$ is a perturbed version of $\boldsymbol{X} = \boldsymbol{C}_1^H \boldsymbol{U}_n \boldsymbol{U}_n^H \boldsymbol{C}_1$ with perturbation

$$\begin{aligned}\delta\boldsymbol{X} =& \hat{\boldsymbol{X}} - \boldsymbol{X} \\ =& \boldsymbol{C}_1^H \delta\boldsymbol{U}_n \boldsymbol{U}_n^H \boldsymbol{C}_1 + \boldsymbol{C}_1^H \boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \\ &+ \boldsymbol{C}_1^H \delta\boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1.\end{aligned} \tag{45}$$

Under certain identifiability conditions, the true channel vector $\boldsymbol{h}_1$ is the unique null vector of $\boldsymbol{X}$. Then, $\hat{\boldsymbol{h}}_1$ is a perturbed channel vector with perturbation $\delta\boldsymbol{h}_1$. This perturbation can be obtained analytically in two steps according to our previous results. First, due to $\delta\boldsymbol{X}$, $\delta\boldsymbol{h}_1$ is expressed as a function of $\delta\boldsymbol{X}$ up to its second order. Second, by expressing $\delta\boldsymbol{U}_n$ as a function of $\delta\boldsymbol{R}$, $\delta\boldsymbol{X}$ in (45) can be derived from $\delta\boldsymbol{R}$. Therefore, an analytical result for $\delta\boldsymbol{h}_1$ can be derived from $\delta\boldsymbol{R}$ up to its second order.

In the first step, (25) is directly applicable as explained at the end of Section III-A, which gives a perturbation in the null vector of $\boldsymbol{X}$

$$\begin{aligned}\delta\boldsymbol{h}_1 \approx& -\boldsymbol{X}^\dagger \delta\boldsymbol{X} \boldsymbol{h}_1 + \boldsymbol{X}^\dagger \delta\boldsymbol{X} \boldsymbol{X}^\dagger \delta\boldsymbol{X} \boldsymbol{h}_1 - (\boldsymbol{X}^\dagger)^2 \delta\boldsymbol{X} \boldsymbol{h}_1 \boldsymbol{h}_1^H \delta\boldsymbol{X} \boldsymbol{h}_1 \\ &- \frac{1}{2} \boldsymbol{h}_1 \boldsymbol{h}_1^H \delta\boldsymbol{X} (\boldsymbol{X}^\dagger)^2 \delta\boldsymbol{X} \boldsymbol{h}_1.\end{aligned} \tag{46}$$

Substituting (45) in (46) and using $\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{h}_1 = \boldsymbol{0}$, we obtain $\delta\boldsymbol{h}_1$ up to the second order of $\delta\boldsymbol{U}_n$

$$\begin{aligned}\delta\boldsymbol{h}_1 \approx& -\boldsymbol{X}^\dagger \boldsymbol{C}_1^H \boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{h}_1 - \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \delta\boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &+ \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \delta\boldsymbol{U}_n \boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &+ \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &- \frac{1}{2} \boldsymbol{h}_1 \boldsymbol{h}_1^H \boldsymbol{C}_1^H \delta\boldsymbol{U}_n \boldsymbol{U}_n^H \boldsymbol{C}_1 (\boldsymbol{X}^\dagger)^2 \\ &\times \boldsymbol{C}_1^H \boldsymbol{U}_n \delta\boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{h}_1.\end{aligned} \tag{47}$$

Second, according to (37), $\delta\boldsymbol{U}_n$ is related to $\delta\boldsymbol{R}$. Substituting (37) in (47), we can obtain $\delta\boldsymbol{h}_1$ directly from covariance estimation error $\delta\boldsymbol{R}$

$$\begin{aligned}\delta\boldsymbol{h}_1 \approx& \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &- \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{Z}^\dagger \delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{h}_1 + \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{U}_n \boldsymbol{U}_n^H \delta\boldsymbol{R}(\boldsymbol{Z}^\dagger)^2 \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &- \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \boldsymbol{Z}^\dagger \delta\boldsymbol{R}\boldsymbol{U}_n \boldsymbol{U}_n^H \delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &+ \boldsymbol{X}^\dagger \boldsymbol{C}_1^H \boldsymbol{Z}^\dagger \delta\boldsymbol{R}\boldsymbol{U}_n \boldsymbol{U}_n^H \boldsymbol{C}_1 \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &+ \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{h}_1 \\ &- \frac{1}{2} \boldsymbol{h}_1 \boldsymbol{h}_1^H \boldsymbol{C}_1^H \boldsymbol{Z}^\dagger \delta\boldsymbol{R}\boldsymbol{S}^H \boldsymbol{S}\delta\boldsymbol{R}\boldsymbol{Z}^\dagger \boldsymbol{C}_1 \boldsymbol{h}_1\end{aligned} \tag{48}$$

where $\boldsymbol{Z}$ is defined in (29), and $\boldsymbol{S}$ is defined as

$$\boldsymbol{S} = \left(\boldsymbol{C}_1^H \boldsymbol{U}_n \boldsymbol{U}_n^H \boldsymbol{C}_1\right)^\dagger \boldsymbol{C}_1^H \boldsymbol{U}_n \boldsymbol{U}_n^H.$$

It can be observed that the first term in (48) represents a first-order perturbation due to $\delta\boldsymbol{R}$, which is consistent with the results in [3], [11], and [17], whereas all other terms are in the second order of $\delta\boldsymbol{R}$. With second-order terms included, it is anticipated that the predicted channel estimation error becomes smaller under the small perturbation assumption.

In the current work, we restrict our attention to the perturbation of the channel estimate due to $\delta\boldsymbol{R}$ up to the second order. The statistical property of the subspace-based channel estimator based on (48) will not be further derived analytically, which is beyond the scope of the paper. However, for the first-order based statistical analysis, see [9] for a single-input multiple-output (SIMO) channel or [20] for a CDMA channel. Meanwhile, the asymptotic distribution of the subspace-based channel estimate is derived in [1] and [2]. In those works, it is shown that the second-order statistics of the estimated channel vector depend on up to the fourth-order information of the received data. If our second-order perturbation terms are considered, then those statistics will finally depend on up to the eighth-order statistics of the received data, which turns out to be highly complex. Instead, we turn our attention to simulation study next.

Consider a direct sequence CDMA system with six synchronous users. Gold sequence of length $P = 15$ is used as the spreading sequence for each user. Each user transmits a binary information sequence with equal power. Channel coefficients are arbitrarily selected from a zero-mean Gaussian process with unit variance and put in a matrix shown at the bottom of the next page, where each column represents the normalized channel vector for the corresponding user. Channel $\boldsymbol{h}_1$ is estimated according to (44). The mean square error is defined as [1], [9]

$$\text{MSE} = 10 \log_{10} E\{\|\hat{\boldsymbol{h}}_1 - \boldsymbol{h}_1\|^2\}.$$

In order to gain some insight into the roles of the second-order perturbation terms in (48), this MSE is compared with the ex-

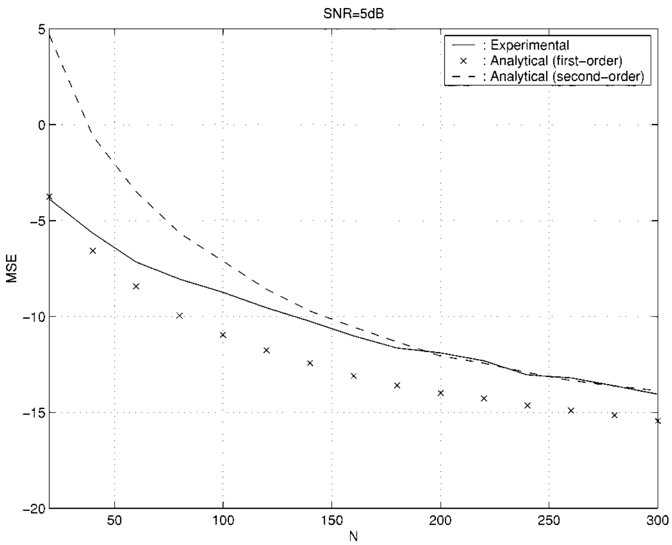

Fig. 1. Channel estimation error versus the sample size when SNR $= 5$ dB.

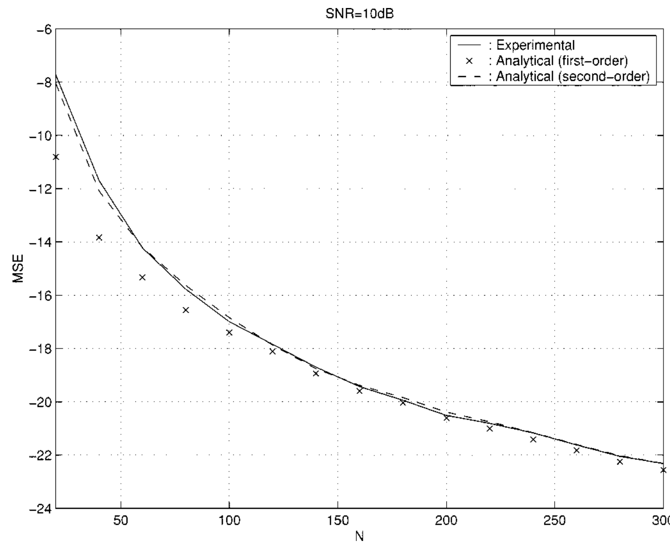

Fig. 2. Channel estimation error versus the sample size when SNR $= 10$ dB.

pected value of the squared norm of (48) and the analytical result by [20], which considers only the first-order perturbation. The average results over 500 realizations of the noise and data processes are obtained. We study the effects of the sample size N and the signal-to-noise ratio (SNR), which are two of the most important factors to cause perturbation in channel estimate. First, the effect of the sample size on the MSE is shown in Fig. 1, starting from $N = 20$, when SNR $= 5$ dB. The solid line is the subspace-based channel estimation result. The dashed line represents the result predicted by our second-order perturbation analysis, and the "x"s denote the result by [20]. It can be observed that the first-order perturbation cannot provide reliable prediction in such a low SNR level, even when N is large compared with P. However, the second-order approximation highly agrees with the experimental result for a large range of N. Certainly, it is inaccurate for small N, where the small perturbation assumption may be violated. Although the first-order approximation appears better for small N, theoretically, it is not reliable either. When SNR increases to 10 dB, the results from experiment and our second-order approximation are almost indistinguishable, as observed from Fig. 2, but the first-order approximation still shows some deviation for small N. However, the gap between the first- and second-order approximation results decreases. It diminishes when SNR further increases to 15 dB, as shown in Fig. 3. This time both approximation methods perform satisfactorily.

The effect of SNR on channel estimation error is also studied. The average results from 2000 independent realizations with $N = 100$ are presented in Fig. 4 for SNR ranging from 0 to 20 dB. It is seen that in the 6– to 16–dB region, the proposed analytical results are in better agreement with the experimental results than the first-order based results [20]. For SNR higher than 16 dB, both first- and second-order analyses show satisfactory performance. In the low SNR region, we should not rely on the perturbation analysis, as expected. In Fig. 5, N increases to 200. Consequently, the gap between two analytical curves decreases. Both analytical results appear more accurate for a large range of SNR when this figure is compared with the previous one. Meanwhile, the experimental results tend to converge to the proposed analytical results even at lower SNRs. For example, at as low as 5 dB SNR, the solid and dashed lines almost overlap. If N increases further to 500, the first- and second-order analyses do not yield much difference, as observed from Fig. 6.

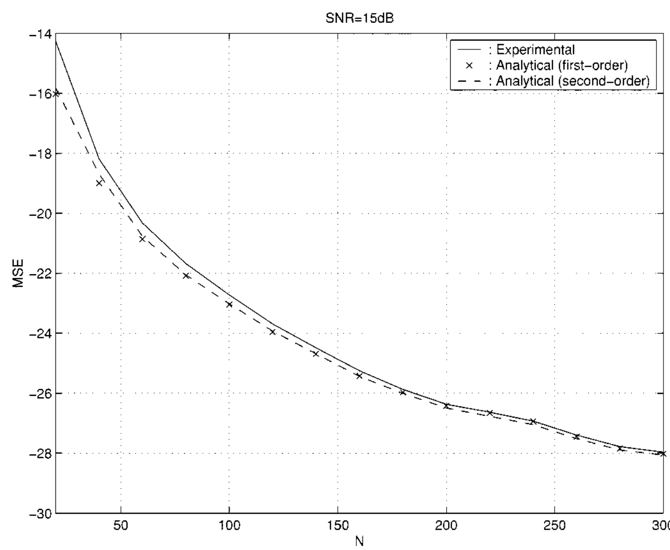

Fig. 3. Channel estimation error versus the sample size when SNR $= 15$ dB.

$$\boldsymbol{H} = \begin{bmatrix} 0.6653 & 0.5771 & 0.5008 & 0.3986 & 0.7373 & 0.3837 \\ -0.4377 & 0.8002 & -0.4519 & -0.7639 & -0.6144 & -0.5279 \\ 0.6043 & -0.1620 & -0.0297 & 0.0336 & 0.2274 & 0.2254 \\ 0.0224 & -0.0192 & -0.7376 & -0.5065 & -0.1647 & -0.7234 \end{bmatrix}$$

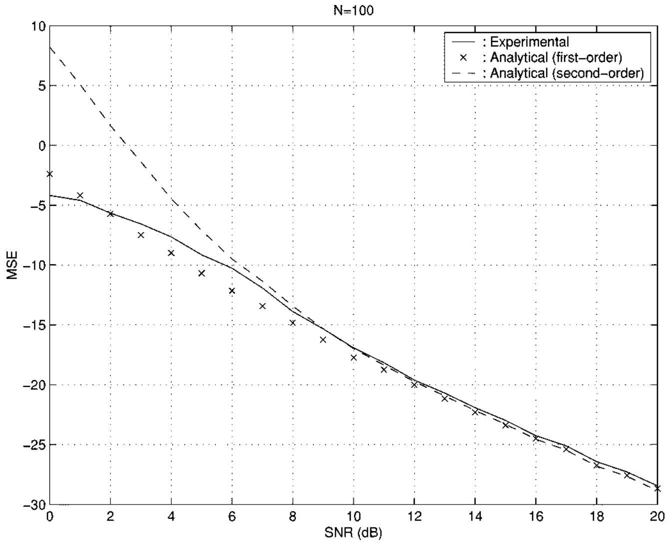

Fig. 4. Channel estimation error versus SNR when $N = 100$.

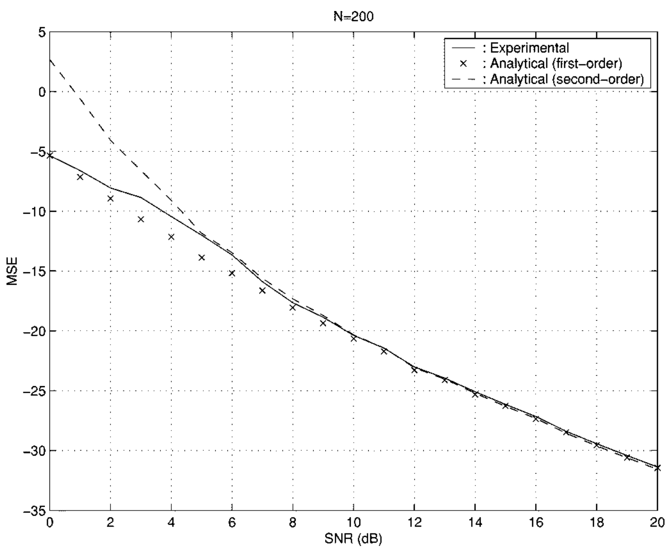

Fig. 5. Channel estimation error versus SNR when $N = 200$.

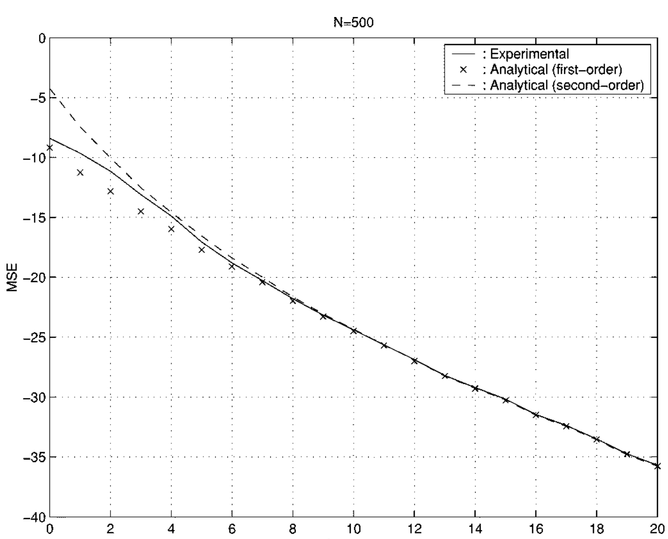

Fig. 6. Channel estimation error versus SNR when $N = 500$.

Based on the previous simulation study, it can be concluded that our second-order perturbation analysis can predict the estimator's performance more accurately for a larger range of number of data samples and SNRs than the first-order perturbation analysis. Thus, the superiority of our second-order perturbation analysis is numerically established.

V. CONCLUSIONS

In this paper, perturbations in subspace decomposition on a perturbed matrix are studied. All relevant quantities such as subspaces and singular values are considered. Their perturbations are derived as functions of the perturbation in the matrix up to the second order, respectively, all in closed forms. It is also shown that if the proposed results are truncated at the first order, they reduce to those of existing approaches, as expected. Demand on the availability of second-order perturbations may arise if the first-order perturbation cannot provide sufficient accuracy in data analysis, depending on applications. Numerical examples in matrix computation and subspace-based blind channel estimation for CDMA systems are provided to illustrate the roles of second-order perturbation terms.

APPENDIX

Proof of the Theorem

We will derive a set of equations for our unknowns. First, substituting (1), (5), and (7) into the first equation of (11) and applying the first and second equations of (10), we obtain

$$\begin{aligned}\delta\boldsymbol{X}^H\boldsymbol{U}_s + \delta\boldsymbol{X}^H\boldsymbol{U}_n\boldsymbol{P}_1 + \boldsymbol{V}_n\boldsymbol{\Lambda}_n^H\boldsymbol{P}_1 &\\ + \boldsymbol{V}_s\boldsymbol{\Lambda}_s^H\boldsymbol{P}_2 + \delta\boldsymbol{X}^H\boldsymbol{U}_s\boldsymbol{P}_2 &\\ = \boldsymbol{V}_s\delta\boldsymbol{\Lambda}_s^H + \boldsymbol{V}_n\bar{\boldsymbol{P}}_1\boldsymbol{\Lambda}_s^H + \boldsymbol{V}_n\bar{\boldsymbol{P}}_1\delta\boldsymbol{\Lambda}_s^H &\\ + \boldsymbol{V}_s\bar{\boldsymbol{P}}_2\boldsymbol{\Lambda}_s^H + \boldsymbol{V}_s\bar{\boldsymbol{P}}_2\delta\boldsymbol{\Lambda}_s^H. &\end{aligned} \tag{49}$$

If we premultiply both sides of (49) by $\boldsymbol{V}_s^H$ and $\boldsymbol{V}_n^H$, respectively, and use (8), then

$$\begin{aligned}\boldsymbol{E}_{ss}^H + \boldsymbol{E}_{ns}^H\boldsymbol{P}_1 + \boldsymbol{\Lambda}_s^H\boldsymbol{P}_2 + \boldsymbol{E}_{ss}^H\boldsymbol{P}_2 =& \delta\boldsymbol{\Lambda}_s^H + \bar{\boldsymbol{P}}_2\boldsymbol{\Lambda}_s^H \\ &+ \bar{\boldsymbol{P}}_2\delta\boldsymbol{\Lambda}_s^H \qquad (50)\\ \boldsymbol{E}_{sn}^H + \boldsymbol{E}_{nn}^H\boldsymbol{P}_1 + \boldsymbol{\Lambda}_n^H\boldsymbol{P}_1 + \boldsymbol{E}_{sn}^H\boldsymbol{P}_2 =& \bar{\boldsymbol{P}}_1\boldsymbol{\Lambda}_s^H \\ &+ \bar{\boldsymbol{P}}_1\delta\boldsymbol{\Lambda}_s^H. \qquad (51)\end{aligned}$$

These are two equations where unknowns are involved based only on the first equation of (11). We will next obtain two equations from each of other three equations of (11).

Second, from the second equation of (11), we can obtain

$$\begin{aligned}\delta\boldsymbol{X}^H\boldsymbol{U}_n + \delta\boldsymbol{X}^H\boldsymbol{U}_s\boldsymbol{Q}_1 + \boldsymbol{V}_s\boldsymbol{\Lambda}_s^H\boldsymbol{Q}_1 &\\ + \boldsymbol{V}_n\boldsymbol{\Lambda}_n^H\boldsymbol{Q}_2 + \delta\boldsymbol{X}^H\boldsymbol{U}_n\boldsymbol{Q}_2 &\\ = \boldsymbol{V}_n\delta\boldsymbol{\Lambda}_n^H + \boldsymbol{V}_s\bar{\boldsymbol{Q}}_1\boldsymbol{\Lambda}_n^H + \boldsymbol{V}_s\bar{\boldsymbol{Q}}_1\delta\boldsymbol{\Lambda}_n^H &\\ + \boldsymbol{V}_n\bar{\boldsymbol{Q}}_2\boldsymbol{\Lambda}_n^H + \boldsymbol{V}_n\bar{\boldsymbol{Q}}_2\delta\boldsymbol{\Lambda}_n^H &\end{aligned} \tag{52}$$

where the first and second equations of (10) have been used once more. If we premultiply both sides of (52) by $\boldsymbol{V}_n^H$ and $\boldsymbol{V}_s^H$, respectively, and use (8), we have

$$\begin{aligned}\boldsymbol{E}_{nn}^H + \boldsymbol{E}_{sn}^H\boldsymbol{Q}_1 + \boldsymbol{\Lambda}_n^H\boldsymbol{Q}_2 + \boldsymbol{E}_{nn}^H\boldsymbol{Q}_2 =& \delta\boldsymbol{\Lambda}_n^H + \bar{\boldsymbol{Q}}_2\boldsymbol{\Lambda}_n^H \\ &+ \bar{\boldsymbol{Q}}_2\delta\boldsymbol{\Lambda}_n^H \qquad (53)\\ \boldsymbol{E}_{ns}^H + \boldsymbol{E}_{ss}^H\boldsymbol{Q}_1 + \boldsymbol{\Lambda}_s^H\boldsymbol{Q}_1 + \boldsymbol{E}_{ns}^H\boldsymbol{Q}_2 =& \bar{\boldsymbol{Q}}_1\boldsymbol{\Lambda}_n^H \\ &+ \bar{\boldsymbol{Q}}_1\delta\boldsymbol{\Lambda}_n^H. \qquad (54)\end{aligned}$$

These are another two equations for our unknowns.

Next, from the third equation of (11), it can be easily verified that the following holds:

$$\begin{aligned}&\delta\boldsymbol{X}\boldsymbol{V}_s + \delta\boldsymbol{X}\boldsymbol{V}_n\bar{\boldsymbol{P}}_1 + \boldsymbol{U}_n\boldsymbol{\Lambda}_n\bar{\boldsymbol{P}}_1 + \boldsymbol{U}_s\boldsymbol{\Lambda}_s\bar{\boldsymbol{P}}_2 + \delta\boldsymbol{X}\boldsymbol{V}_s\bar{\boldsymbol{P}}_2 \\ &\quad = \boldsymbol{U}_s\delta\boldsymbol{\Lambda}_s + \boldsymbol{U}_n\boldsymbol{P}_1\boldsymbol{\Lambda}_s + \boldsymbol{U}_n\boldsymbol{P}_1\delta\boldsymbol{\Lambda}_s \\ &\qquad + \boldsymbol{U}_s\boldsymbol{P}_2\boldsymbol{\Lambda}_s + \boldsymbol{U}_s\boldsymbol{P}_2\delta\boldsymbol{\Lambda}_s. \qquad (55)\end{aligned}$$

Premultiplying both sides of (55) by $\boldsymbol{U}_s^H$ and $\boldsymbol{U}_n^H$, respectively, we arrive at two more equations:

$$\begin{aligned}\boldsymbol{E}_{ss} + \boldsymbol{E}_{sn}\bar{\boldsymbol{P}}_1 + \boldsymbol{\Lambda}_s\bar{\boldsymbol{P}}_2 + \boldsymbol{E}_{ss}\bar{\boldsymbol{P}}_2 =& \delta\boldsymbol{\Lambda}_s + \boldsymbol{P}_2\boldsymbol{\Lambda}_s \\ &+ \boldsymbol{P}_2\delta\boldsymbol{\Lambda}_s \qquad (56)\\ \boldsymbol{E}_{ns} + \boldsymbol{E}_{nn}\bar{\boldsymbol{P}}_1 + \boldsymbol{\Lambda}_n\bar{\boldsymbol{P}}_1 + \boldsymbol{E}_{ns}\bar{\boldsymbol{P}}_2 =& \boldsymbol{P}_1\boldsymbol{\Lambda}_s \\ &+ \boldsymbol{P}_1\delta\boldsymbol{\Lambda}_s. \qquad (57)\end{aligned}$$

Finally, two additional equations are obtained from the fourth equation of (11) by expanding multiplications

$$\begin{aligned}&\delta\boldsymbol{X}\boldsymbol{V}_n + \delta\boldsymbol{X}\boldsymbol{V}_s\bar{\boldsymbol{Q}}_1 + \boldsymbol{U}_s\boldsymbol{\Lambda}_s\bar{\boldsymbol{Q}}_1 + \boldsymbol{U}_n\boldsymbol{\Lambda}_n\bar{\boldsymbol{Q}}_2 + \delta\boldsymbol{X}\boldsymbol{V}_n\bar{\boldsymbol{Q}}_2 \\ &\quad = \boldsymbol{U}_n\delta\boldsymbol{\Lambda}_n + \boldsymbol{U}_s\boldsymbol{Q}_1\boldsymbol{\Lambda}_n + \boldsymbol{U}_s\boldsymbol{Q}_1\delta\boldsymbol{\Lambda}_n \\ &\qquad + \boldsymbol{U}_n\boldsymbol{Q}_2\boldsymbol{\Lambda}_n + \boldsymbol{U}_n\boldsymbol{Q}_2\delta\boldsymbol{\Lambda}_n \qquad (58)\end{aligned}$$

and premultiplying both sides of (58) by $\boldsymbol{U}_n^H$ and $\boldsymbol{U}_s^H$, respectively:

$$\begin{aligned}\boldsymbol{E}_{nn} + \boldsymbol{E}_{ns}\bar{\boldsymbol{Q}}_1 + \boldsymbol{\Lambda}_n\bar{\boldsymbol{Q}}_2 + \boldsymbol{E}_{nn}\bar{\boldsymbol{Q}}_2 =& \delta\boldsymbol{\Lambda}_n + \boldsymbol{Q}_2\boldsymbol{\Lambda}_n \\ &+ \boldsymbol{Q}_2\delta\boldsymbol{\Lambda}_n \qquad (59)\\ \boldsymbol{E}_{sn} + \boldsymbol{E}_{ss}\bar{\boldsymbol{Q}}_1 + \boldsymbol{\Lambda}_s\bar{\boldsymbol{Q}}_1 + \boldsymbol{E}_{sn}\bar{\boldsymbol{Q}}_2 =& \boldsymbol{Q}_1\boldsymbol{\Lambda}_n \\ &+ \boldsymbol{Q}_1\delta\boldsymbol{\Lambda}_n. \qquad (60)\end{aligned}$$

Thus far, we have obtained following eight equations for our unknowns: (50), (51), (53), (54), (56), (57), (59), and (60). We may obtain more equations from (9) when necessary to derive our solutions.

Let us consider $\boldsymbol{Q}_1$ first. It is involved together with $\bar{\boldsymbol{Q}}_1$, $\boldsymbol{Q}_2$, $\bar{\boldsymbol{Q}}_2$, $\delta\boldsymbol{\Lambda}_n$ in (53), (54), (59), and (60). However, $\boldsymbol{Q}_2$ and $\bar{\boldsymbol{Q}}_2$ are related to $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$, respectively. Following the line of [16, ch. V], let us express $\hat{\boldsymbol{U}}_n$ in (6) similarly by another form $(\boldsymbol{U}_n + \boldsymbol{U}_s\boldsymbol{A}_1)\boldsymbol{A}_2$, where $\boldsymbol{A}_2 = \boldsymbol{I} + \boldsymbol{Q}_2$ and $\boldsymbol{Q}_1 = \boldsymbol{A}_1\boldsymbol{A}_2$. Since $\hat{\boldsymbol{U}}_n^H\hat{\boldsymbol{U}}_n = \boldsymbol{I}$, $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ must satisfy normalization requirement $\boldsymbol{A}_2^H\left(\boldsymbol{I} + \boldsymbol{A}_1^H\boldsymbol{A}_1\right)\boldsymbol{A}_2 = \boldsymbol{I}$. We thus may constrain $\boldsymbol{A}_2$ to be a Hermitian matrix $\left(\boldsymbol{I} + \boldsymbol{A}_1^H\boldsymbol{A}_1\right)^{-1/2}$. Therefore, $\boldsymbol{Q}_2$ must be a Hermitian matrix $\boldsymbol{Q}_2^H = \boldsymbol{Q}_2$. Again, applying $\hat{\boldsymbol{U}}_n^H\hat{\boldsymbol{U}}_n = \boldsymbol{I}$ with our current expression of $\hat{\boldsymbol{U}}_n$ in (6), we obtain

$$\boldsymbol{Q}_2 + \boldsymbol{Q}_2^H + \boldsymbol{Q}_2^H\boldsymbol{Q}_2 + \boldsymbol{Q}_1^H\boldsymbol{Q}_1 = \boldsymbol{0}$$

which becomes

$$2\boldsymbol{Q}_2 + \boldsymbol{Q}_2^2 + \boldsymbol{Q}_1^H\boldsymbol{Q}_1 = \boldsymbol{0}$$

after using the Hermitian property of $\boldsymbol{Q}_2$. It can be easily observed that $\boldsymbol{Q}_2$ cannot have first-order terms of $\delta\boldsymbol{X}$, although $\boldsymbol{Q}_1$ may have. Then

$$\boldsymbol{Q}_2 \approx -\frac{1}{2}\boldsymbol{Q}_1^H\boldsymbol{Q}_1. \qquad (61)$$

This result is analogous to a case when a unit-norm vector $\boldsymbol{x}$ is perturbed to be $\hat{\boldsymbol{x}}$. If the angle θ between $\hat{\boldsymbol{x}}$ and $\boldsymbol{x}$ is very small ($\theta \ll 1$), then the perturbation in the direction orthogonal to $\boldsymbol{x}$ is $\sin\theta\boldsymbol{x}^\perp \approx \theta\boldsymbol{x}^\perp$, where vector $\boldsymbol{x}^\perp$ has unit-norm, whereas the "in-direction" perturbation (in parallel with $\boldsymbol{x}$) is $(\cos\theta - 1)\boldsymbol{x} \approx -(1/2)\theta^2\boldsymbol{x}$, which appears as a second-order perturbation. Since $\theta^2 \ll \theta$, the "in-direction" perturbation is much smaller than the perturbation in its orthogonal direction. In a much similar way, we can obtain $\bar{\boldsymbol{Q}}_2$, $\boldsymbol{P}_2$, $\bar{\boldsymbol{P}}_2$ as follows:

$$\bar{\boldsymbol{Q}}_2 \approx -\frac{1}{2}\bar{\boldsymbol{Q}}_1^H\bar{\boldsymbol{Q}}_1, \quad \boldsymbol{P}_2 \approx -\frac{1}{2}\boldsymbol{P}_1^H\boldsymbol{P}_1, \quad \bar{\boldsymbol{P}}_2 \approx -\frac{1}{2}\bar{\boldsymbol{P}}_1^H\bar{\boldsymbol{P}}_1. \qquad (62)$$

Now, we continue our derivation of $\boldsymbol{Q}_1$. Substituting $\delta\boldsymbol{\Lambda}_n$ from (59) into (60), we eliminate $\delta\boldsymbol{\Lambda}_n$ and obtain an equation of only $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$ up to the second order

$$\boldsymbol{E}_{sn} + \boldsymbol{E}_{ss}\bar{\boldsymbol{Q}}_1 + \boldsymbol{\Lambda}_s\bar{\boldsymbol{Q}}_1 \approx \boldsymbol{Q}_1\boldsymbol{\Lambda}_n + \boldsymbol{Q}_1\boldsymbol{E}_{nn} \qquad (63)$$

where higher order terms[2] have been discarded, including those terms with $\boldsymbol{Q}_2$ and $\bar{\boldsymbol{Q}}_2$ multiplied by the first order of $\delta\boldsymbol{X}$. Then, (63) can be written as

$$\bar{\boldsymbol{Q}}_1 \approx \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{\Lambda}_n - \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn} - \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\bar{\boldsymbol{Q}}_1 + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{E}_{nn}. \qquad (64)$$

In order to express $\bar{\boldsymbol{Q}}_1$ by $\boldsymbol{Q}_1$ in a closed form, we adopt an iterative approach. Substituting $\bar{\boldsymbol{Q}}_1$ on the right-hand side of (64) by its expression (64) and neglecting higher order terms, we eliminate $\bar{\boldsymbol{Q}}_1$ on the right-hand side of (64)

$$\begin{aligned}\bar{\boldsymbol{Q}}_1 \approx& \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{\Lambda}_n - \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn} \\ &-\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\left(\boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{\Lambda}_n - \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\right) + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{E}_{nn}. \qquad (65)\end{aligned}$$

Ater rearrangement, (65) becomes

$$\begin{aligned}\bar{\boldsymbol{Q}}_1 \approx& -\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn} + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn} + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{\Lambda}_n \\ &-\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{\Lambda}_n + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{E}_{nn}. \qquad (66)\end{aligned}$$

In a much similar way, if we substitute $\delta\boldsymbol{\Lambda}_n^H$ from (53) into (54) and ignore higher order terms, we obtain another equation for $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$

$$\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{ss}^H\boldsymbol{Q}_1 + \boldsymbol{\Lambda}_s^H\boldsymbol{Q}_1 \approx \bar{\boldsymbol{Q}}_1\boldsymbol{\Lambda}_n^H + \bar{\boldsymbol{Q}}_1\boldsymbol{E}_{nn}^H. \qquad (67)$$

Substituting $\bar{\boldsymbol{Q}}_1$ in (66) into (67) and discarding higher order terms, an equation for $\boldsymbol{Q}_1$ follows:

$$\begin{aligned}\left(\boldsymbol{\Lambda}_s^H - \alpha\boldsymbol{\Lambda}_s^{-1}\right)\boldsymbol{Q}_1 \approx& -\boldsymbol{E}_{ns}^H - \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H \\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H \\ &- \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H - \boldsymbol{E}_{ss}^H\boldsymbol{Q}_1 \\ &- \alpha\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1 + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{E}_{nn}\boldsymbol{\Lambda}_n^H \\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{Q}_1\boldsymbol{\Lambda}_n\boldsymbol{E}_{nn}^H \qquad (68)\end{aligned}$$

[2]In the current context, "higher order" means order greater than two.

where $\boldsymbol{\Lambda}_n\boldsymbol{\Lambda}_n^H$ has been replaced by $\alpha\boldsymbol{I}$. $\boldsymbol{Q}_1$ is not easy to solve in the current matrix-form equation. However, since we are only interested in the expression of $\boldsymbol{Q}_1$ up to the second order of $\delta\boldsymbol{X}$, we may still apply the recursive technique. Multiplying both sides of (68) by $\boldsymbol{\Lambda}_s$, the left-hand side becomes $(\boldsymbol{\Lambda}_s\boldsymbol{\Lambda}_s^H - \alpha\boldsymbol{I})\boldsymbol{Q}_1$. Applying our definition $\boldsymbol{\Sigma} = \left(\boldsymbol{\Lambda}_s\boldsymbol{\Lambda}_s^H - \alpha\boldsymbol{I}\right)^{-1}$, we obtain

$$\begin{aligned}\boldsymbol{Q}_1 \approx &- \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right)\\ &+ \boldsymbol{\Sigma}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H - \boldsymbol{\Sigma}\boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H\\ &- \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ss}^H + \alpha\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\right)\boldsymbol{Q}_1\\ &+ \boldsymbol{\Sigma}\boldsymbol{Q}_1\left(\boldsymbol{E}_{nn}\boldsymbol{\Lambda}_n^H + \boldsymbol{\Lambda}_n\boldsymbol{E}_{nn}^H\right). \end{aligned} \tag{69}$$

Based on our recursive method and keeping terms only up to the second-order perturbations, we can obtain $\boldsymbol{Q}_1$

$$\begin{aligned}\boldsymbol{Q}_1 \approx &- \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right) + \boldsymbol{\Sigma}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\\ &- \boldsymbol{\Sigma}\boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H\\ &+ \boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ss}^H + \alpha\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\right)\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right)\\ &- \boldsymbol{\Sigma}^2\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right)\\ &\times\left(\boldsymbol{E}_{nn}\boldsymbol{\Lambda}_n^H + \boldsymbol{\Lambda}_n\boldsymbol{E}_{nn}^H\right). \end{aligned} \tag{70}$$

After rearranging all terms in (66) and combining corresponding terms, we obtain $\boldsymbol{Q}_1$, which is (12). With $\boldsymbol{Q}_1$, other unknowns can be easily derived, as explained next.

First, $\bar{\boldsymbol{Q}}_1$ can be found from $\boldsymbol{Q}_1$ according to (66). After substituting (12) into (66) and ignoring higher order terms, $\bar{\boldsymbol{Q}}_1$ becomes

$$\begin{aligned}\bar{\boldsymbol{Q}}_1 \approx &- \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn} + \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn}\\ &- \boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right)\boldsymbol{\Lambda}_n\\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{sn} - \boldsymbol{\Sigma}\boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\boldsymbol{E}_{nn}\right.\\ &\qquad\left.+\alpha\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\boldsymbol{E}_{sn}\right)\boldsymbol{\Lambda}_n^H\boldsymbol{\Lambda}_n\\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ss}^H\boldsymbol{\Sigma}\boldsymbol{E}_{sn} - \boldsymbol{\Sigma}\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H\boldsymbol{E}_{nn}\right)\boldsymbol{\Lambda}_n^H\boldsymbol{\Lambda}_n\\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ss}^H\boldsymbol{\Sigma}\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H - \boldsymbol{\Sigma}\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H\boldsymbol{\Lambda}_n\boldsymbol{E}_{nn}^H\right)\boldsymbol{\Lambda}_n\\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\alpha\boldsymbol{E}_{ss}\boldsymbol{\Sigma}\boldsymbol{E}_{ns}^H - \boldsymbol{\Sigma}\boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\boldsymbol{\Lambda}_n\boldsymbol{E}_{nn}^H\right.\\ &\qquad\left.-\boldsymbol{E}_{sn}\boldsymbol{E}_{nn}^H\right)\boldsymbol{\Lambda}_n\\ &+ \boldsymbol{\Lambda}_s^{-1}\boldsymbol{E}_{ss}\boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right)\boldsymbol{\Lambda}_n\\ &- \boldsymbol{\Lambda}_s^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{\Lambda}_s\boldsymbol{E}_{ns}^H + \boldsymbol{E}_{sn}\boldsymbol{\Lambda}_n^H\right)\boldsymbol{E}_{nn}. \end{aligned} \tag{71}$$

With some simplification and reordering, (71) becomes (13), which gives $\bar{\boldsymbol{Q}}_1$.

Once we obtain $\boldsymbol{Q}_1$ and $\bar{\boldsymbol{Q}}_1$, other remaining quantities $\boldsymbol{Q}_2$, $\bar{\boldsymbol{Q}}_2, \boldsymbol{P}_1, \bar{\boldsymbol{P}}_1, \boldsymbol{P}_2, \bar{\boldsymbol{P}}_2, \delta\boldsymbol{\Lambda}_s$, and $\delta\boldsymbol{\Lambda}_n$ can be easily derived as follows. $\boldsymbol{Q}_2$ is related to $\boldsymbol{Q}_1$ by (61). Keeping only the first-order terms in $\boldsymbol{Q}_1$ when using (61), we obtain $\boldsymbol{Q}_2$ as in (16). Similarly, from (62), we obtain $\bar{\boldsymbol{Q}}_2$ as in (17). From orthogonality $\hat{\boldsymbol{U}}_n^H\hat{\boldsymbol{U}}_s = \boldsymbol{0}$, we have

$$\boldsymbol{P}_1 + \boldsymbol{Q}_1^H + \boldsymbol{Q}_1^H\boldsymbol{P}_2 + \boldsymbol{Q}_2^H\boldsymbol{P}_1 = \boldsymbol{0}. \tag{72}$$

Since $\boldsymbol{Q}_2$ and $\boldsymbol{P}_2$ do not contain the first-order perturbations according to our earlier analysis, we conclude from (72) that $\boldsymbol{P}_1 \approx -\boldsymbol{Q}_1^H$, which is (14) and similar to [10]. Similarly, from $\hat{\boldsymbol{V}}_n^H\hat{\boldsymbol{V}}_s = \boldsymbol{0}$, we get $\bar{\boldsymbol{P}}_1 \approx -\bar{\boldsymbol{Q}}_1^H$, which is (15). Therefore, according to (62) and using $\boldsymbol{P}_1 \approx -\boldsymbol{Q}_1^H$, we have $\boldsymbol{P}_2 \approx -(1/2)\boldsymbol{Q}_1\boldsymbol{Q}_1^H$. Keeping only the first-order terms in $\boldsymbol{Q}_1$, we obtain $\boldsymbol{P}_2$, which is given in (18). Similarly, from (62) and using $\bar{\boldsymbol{P}}_1 \approx -\bar{\boldsymbol{Q}}_1^H$, we obtain $\bar{\boldsymbol{P}}_2$, which is given in (19). Other two unknowns $\delta\boldsymbol{\Lambda}_s$ and $\delta\boldsymbol{\Lambda}_n$ can be solved based on (56) and (59), respectively. According to (56), $\delta\boldsymbol{\Lambda}_s$ can be expressed up to the second order as

$$\delta\boldsymbol{\Lambda}_s = \boldsymbol{E}_{ss} + \boldsymbol{E}_{sn}\bar{\boldsymbol{P}}_1 + \boldsymbol{\Lambda}_s\bar{\boldsymbol{P}}_2 - \boldsymbol{P}_2\boldsymbol{\Lambda}_s.$$

After considering (15), (18), and (19) and keeping up to the second order, we obtain $\delta\boldsymbol{\Lambda}_s$ as in (20). Similarly, from (59), we have

$$\delta\boldsymbol{\Lambda}_n = \boldsymbol{E}_{nn} + \boldsymbol{E}_{ns}\bar{\boldsymbol{Q}}_1 + \boldsymbol{\Lambda}_n\bar{\boldsymbol{Q}}_2 - \boldsymbol{Q}_2\boldsymbol{\Lambda}_n.$$

Using (16) and (17), (21) follows, which gives $\delta\boldsymbol{\Lambda}_n$. □

Acknowledgment

The author would like to thank the anonymous referees for the valuable comments that have significantly contributed to the improvement of the quality of the paper.

References

[1] K. Abed-Meraim, J.-F. Cardoso, A. Y. Gorokhov, P. Loubaton, and E. Moulines, “On subspace methods for blind identification of single-input multiple-output FIR systems,” *IEEE Trans. Signal Processing*, vol. 45, pp. 42–55, Jan. 1997.

[2] K. Abed-Meraim, P. Loubaton, and E. Moulines, “A subspace algorithm for certain blind identification problems,” *IEEE Trans. Inform. Theory*, vol. 43, pp. 499–511, Mar. 1997.

[3] E. Aktas and U. Mitra, “Complexity reduction in subspace-based blind channel identification for DS/CDMA systems,” *IEEE Trans. Commun.*, vol. 48, pp. 1392–1404, Aug. 2000.

[4] S. Bensley and B. Aazhang, “Subspace-based channel estimation for code division multiple access communication systems,” *IEEE Trans. Commun.*, vol. 44, pp. 1009–1020, Aug. 1996.

[5] D. Brillinger, *Time Series Data Analysis and Theory*. San Francisco, CA: Holden-Day, 1981.

[6] B. Champagne, “Adaptive eigendecomposition of data covariance matrices based on first-order perturbations,” *IEEE Trans. Signal Processing*, vol. 42, pp. 2758–2770, Oct. 1994.

[7] B. Friedlander and A. Weiss, “On the second-order statistics of the eigenvectors of sample covariance matrices,” *IEEE Trans. Signal Processing*, vol. 46, pp. 3136–3139, Nov. 1998.

[8] G. Golub and C. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

[9] M. Kristensson and B. Ottersten, “A statistical approach to subspace based blind identification,” *IEEE Trans. Signal Processing*, vol. 46, pp. 1612–1623, June 1998.

[10] F. Li, H. Liu, and R. Vaccaro, “Performance analysis for DOA estimation algorithms: Unification, simplification, and observations,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, pp. 1170–1184, Oct. 1993.

[11] H. Liu and G. Xu, “A subspace method for signature waveform estimation in synchronous CDMA systems,” *IEEE Trans. Commun.*, vol. 44, pp. 1346–1354, Oct. 1996.

[12] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, “Subspace methods for the blind identification of multichannel FIR filters,” *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[13] A. Paulraj, R. Roy, and T. Kailath, “Estimation of signal parameters via rotational invariance techniques—ESPRIT,” in *Proc. 19th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1985, pp. 83–89.

[14] F. Rellich, *Perturbation Theory of Eigenvalue Problems*. New York: Gordon and Breach, 1969.

[15] R. Schmidt, "A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1981.
[16] G. Stewart and J. Sun, *Matrix Perturbation Theory*. San Diego, CA: Academic, 1990.
[17] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 137–147, Jan. 1997.
[18] M. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receivers," *IEEE Trans. Signal Processing*, vol. 46, pp. 3014–3022, Nov. 1998.
[19] X. Wang and H. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.
[20] N. Yuen and B. Friedlander, "Asymptotic performance analysis for signature waveform estimation in synchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 46, pp. 1753–1757, June 1998.

Zhengyuan Xu (S'97–M'99–SM'02) received both the B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1989 and 1991, respectively, and the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, in 1999.

From 1991 to 1996, he worked as an engineer and department manager at the Tsinghua Unisplendour Group Corp., Tsinghua University. From 1996 to 1999, he was a research assistant and research associate at Stevens Institute of Technology, working on signal processing for wireless communications, especially for multiuser CDMA systems. In 1999, he joined the Department of Electrical Engineering, University of California, Riverside, as an assistant professor. His current research interests include advanced signal processing, multirate communication, multiuser detection, and system identification.

Dr. Xu received the Outstanding Student Award and the Motorola Scholarship from Tsinghua University and the Peskin Award from Stevens Institute of Technology. He also received the Academic Senate Research Award and the Regents' Faculty Award from University of California, Riverside.